1-11-2022

The role of endoplasmic reticulum modifications in protein export in Plasmodium falciparum

Alexander Polino
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Molecular Microbiology and Microbial Pathogenesis

Dissertation Examination Committee:
Daniel Goldberg, Chair
Sergej Djuranovic
Tamara Doering
Stuart Kornfeld
David Sibley

The Role of Endoplasmic Reticulum Modifications in Protein Export in *Plasmodium falciparum*

by
Alexander Polino

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2021
St. Louis, Missouri

Table of Contents

List of Figures

Chapter 1: Introduction

Chapter 2: Assessment of biological role and insight into druggability of plasmepsin V

Chapter 3: Functional analysis of unique structural features of the malaria parasite plasmepsin V

Chapter 4: An essential ER-resident N-acetyltransferase in *P. falciparum*

List of Tables

Acknowledgments

The several years of dedicated labor described in this thesis were enabled by the support, friendship, and occasional timely intervention of countless people, of whom I have space to thank just a few here. First, I'd like to thank the faculty and staff who made the MMMP program a fantastic place to train, particularly program directors David Sibley and Christina Stallings who served as guard rails to keep me from academic trouble, as well as program coordinators Jeanne Silvestrini and Sara Holmes who kept me from an equal measure of administrative trouble. At Washington University I've had the privilege to meet excellent scientists of all stripes, but I am most indebted to my MMMP classmates: Lisa McLellan, Lizzy Mueller, Jennifer Elliott, Amy Ly, Robert Potter, Michael McAllaster, and add-ons Justin Miller and Jimmy Weagley – each among the cleverest, kindest, and most giving people I've ever had the good fortune to meet.

Second, I owe a deep debt of gratitude to the members of the Goldberg Lab, who provided constant counsel throughout my stay, and who lost innumerable otherwise-productive hours helping to answer and re-answer questions on my behalf. A particular thank you to Josh Beck for taking me under his wing and getting me started in the lab, and then for Eva Istvan and Sumit Mukherjee who gave generously of their time and insight over years of my pestering. Of course, the lab would run aground without the steady hand of Dan Goldberg. I was cautioned before I picked a lab that we all take on a bit of our mentors. As I've wandered through my PhD training, Dan has always been prepared to inject optimism, an idea, or a well-timed silence each in its own turn. His wisdom and experience are, and will continue to be, invaluable to me. I hope I'm fortunate enough to take on a bit of Dan as I go.

Lastly, and no doubt most importantly, I'm deeply grateful to the people who support me outside of the lab. My batteries were repeatedly recharged by interaction with friends – most often those same classmates mentioned above, plus Brad Hiller, Philip Frasse, Greg Harrison, Jerome Prusa, and many others. For three years, I was fortunate to share an apartment with Justin Miller and eventual add-on Elizabeth Kennedy, both inestimable friends, guides, and scientists to whom I'll forever be indebted. As Covid-19 came and locked us down, that burden fell to my girlfriend, roommate, and co-dogparent Celia McKee who often served as support-system-of-one through the most challenging parts of graduate school. To Celia (and to Benji, who says little, but listens well), thank you, I'm not sure I could've made it through without you. Finally, to my parents, who made me, formed me, support me unconditionally, and are always available to hear me out or distract from my frustrations: thank you most of all.

Alex Polino

Washington University in St. Louis

December 2021

ABSTRACT OF THE DISSERTATION

The Role of Endoplasmic Reticulum Modifications in Protein Export in *Plasmodium falciparum*
by
Alexander Polino

Doctor of Philosophy in Biology and Biomedical Sciences

Molecular Microbiology and Microbial Pathogenesis

Washington University in St. Louis, 2021

Professor Daniel Goldberg, Chair

Following invasion of a host red blood cell (RBC), the malaria parasite *Plasmodium falciparum* undertakes a program of dramatic host cell renovation. The parasite exports hundreds of effectors into the RBC, altering its shape and rigidity, increasing its permeability to nutrients, and constituting an intracellular transport network that traffics adhesins to the RBC surface. These adhesins mediate binding of the parasitized RBC to the host vascular endothelium, allowing the parasite to avoid destruction in the spleen, but contributing to the pathogenesis of cerebral malaria. Most of these exported effectors bear a pentameric amino acid sequence that marks them for protein export, called the *Plasmodium* export element (PEXEL), with consensus sequence RxLxE/Q/D. PEXEL is cleaved in the parasite endoplasmic reticulum (ER) by the aspartic protease plasmepsin V (PM V), a step necessary for export. Following cleavage, the new N-terminus is acetylated by some unknown N-acetyltransferase (NAT), the role of which remains unclear. For my Ph. D. thesis work, I used parasite genetics to interrogate the roles of these early players in the protein export pathway. First, I used a recently described post-transcriptional depletion method to lethally deplete PM V, finding this enzyme to be maintained by the parasite in substantial excess, and involved in an export-independent process that remains

to be elucidated. Second, I used this depletion line to dissect PM V's structure, finding that the enzyme requires an unusual nepenthesin insert to carry out its function. Finally, I searched for the PEXEL NAT, lethally depleting the likeliest candidate, and finding that it is required for parasite growth, but uninvolved in protein export. Taken together, this data provides further insight into how the parasite uses these ER-resident enzymes to enable its complex growth and protein sorting needs. Further, this serves as a foundation for future research, opening questions about the role of PM V, the nepenthesin insert's contribution to PM V function (and protease function more generally), and the contribution of N-terminal acetylation to PEXEL function.

Chapter 1: Introduction

1.1 Preface

Parts of this chapter were previously published in:

Nasamu AS*, Polino AJ*, Istvan ES, Goldberg DE (2020) Malaria parasite plasmepsins: More than just plain old degradative pepsins. *J Biol. Chem.* 295(25):8425-8441

(* indicates co-first authors)

1.2 Malaria

Malaria remains a scourge of the equatorial world, with over 200 million cases resulting in nearly 500,000 deaths annually (WHO, 2019). Disease is characterized by periodic fever and chills sometimes accompanied by headaches, joint pain, diarrhea, vomiting, and/or cough (Despommier, et al. 2019; Milner 2018). In severe cases, disease can progress to life-threatening anemia (in young children), acidosis, and/or coma (Milner 2018). Severe disease disproportionately affects the very young, with two thirds of malaria deaths in children under five years of age (WHO, 2019).

1.3 Plasmodium

Malaria is caused by infection with haemosporidian parasites in the genus *Plasmodium*. These parasites undertake a complex lifecycle, requiring alternate passage between mosquito and human hosts (Figure 1.1). Parasites reside in the salivary glands of infected female mosquitos. When the mosquito bites a human, it also injects parasites into the dermis. Sporozoites rapidly enter the bloodstream and follow the blood circulation to the liver. Here they take up residence in hepatocytes, undergoing a massive round of replication to produce thousands of daughter cells.

As an infected liver cell begins to break down, fragments of its plasma membrane enveloping *Plasmodium* cells bud off into the blood stream. These parasite-filled structures eventually burst, releasing the invasive merozoites into the blood (Cowman, et al. 2016). Merozoites invade red blood cells (RBCs), progressing through visually distinguishable stages, termed the ring, trophozoite, and schizont. The schizont, filled with 16-32 daughter cells, eventually bursts, releasing new merozoites into the blood. These merozoites then invade new RBCs, resulting in an intraerythrocytic replicative cycle that amplifies infection. Occasionally, parasites instead initiate a sexual replicative cycle, forming either of two elongated forms called gametocytes (a "male" and a "female" form). These are eventually taken up by a mosquito as it feeds from an infected person. In the mosquito midgut, the gametocytes merge, forming a zygote that develops into a motile form called the ookinete. The ookinete burrows through the mosquito's midgut wall forming an oocyst on the midgut exterior. The oocyst fills with sporozoite daughter cells, which eventually migrate to the mosquito salivary glands and await injection during its next blood meal.

Several species of *Plasmodium* can cause human disease, however most cases and nearly all deaths are caused by *Plasmodium falciparum*. For this reason, and the relative ease of maintaining *P. falciparum* in laboratory culture, *P. falciparum* will be the focus of this thesis, with other species mentioned where comparison is relevant.

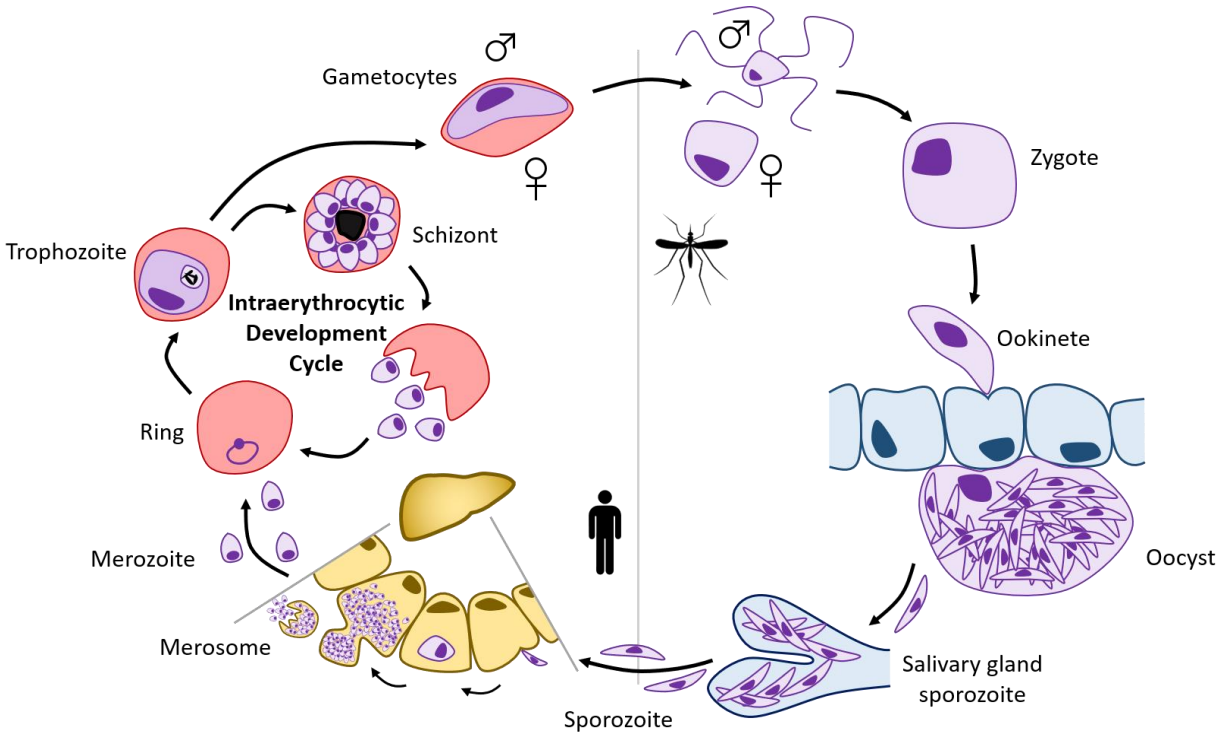


Figure 1.1 – Diagram of the *Plasmodium falciparum* life cycle illustrating the stages that infect mosquitos (right) and humans (left).

1.4 Antimalarials

For nearly two decades, the mainstay of antimalarial treatment has been the administration of combinations of the fast-acting phytocompound artemisinin (or one of its derivatives) and a longer-lived partner drug (Bosman and Mendis, 2007; WHO, 2015). However, as with other microbial infections, resistance to frontline antimalarials has spread, threatening our ability to control this widespread disease (Blasco, et al. 2017). The epicenter of resistance has been Southeast Asia, where regular treatment failures have been reported for four of the five WHO-recommended artemisinin combination therapies (ACTs) (WHO, 2018). This growing threat has spurred calls for new antimalarial compounds to feed the pharmaceutical pipeline, as well as the characterization of essential parasite pathways that could serve as targets for

antimalarial development, and guide the optimization and eventual combination of antimalarial compounds for treatment.

1.5 Protein export as an essential pathway

After invading an RBC, the parasite traffics a menagerie of exported proteins into the host cell. These effectors comprise a dramatic host-modification program that establishes a complex sorting system in the RBC cytosol, rigidifies the host cell, enhances its nutrient permeability, and modulates its endothelial binding properties (Spillman, et al. 2015; Matthews, et al. 2019). Altogether, at least 400 *P. falciparum* gene products are predicted to be exported in a genome of just 5,000 genes (Boddey, et al. 2013; Heiber, et al. 2013). Of these, nearly half belong to expanded families of membrane adhesins: ~120 rifins, 50 PfEMP1s, and 30 stevors (Sargeant, et al. 2006). The remaining proteins defy easy categorization, as many lack homology to known genes or to each other. No doubt, some contribute to known pathways of host modification. However, as currently understood these functions require just a fraction of proteins predicted to be exported. Other functions of protein export likely remain to be uncovered, both in RBC infection, and in other stages of the *P. falciparum* life cycle.

Protein export is coordinated through an elaborate sorting process that is beginning to be clarified (Figure 1.2). Most known exported proteins bear both a canonical signal sequence for ER translocation, as well as a second sorting signal termed the *Plasmodium* export element (PEXEL) (Hiller, et al. 2004; Marti, et al. 2004). PEXEL is recognized and cleaved in the parasite ER by the aspartic protease plasmepsin V (PM V) (Chang, et al. 2008; Boddey, et al. 2009; Osborne, et al. 2010; Russo, et al. 2010; Boddey, et al. 2010). Cleaved PEXEL proteins are acetylated at their new N-terminus, then secreted into the surrounding "parasitophorous vacuole" (PV), where they are somehow recognized and translocated across the PV membrane

by the Parasite translocon for exported proteins (PTEX) (de Koning-Ward, et al. 2009; Beck, et al. 2014; Elsworth, et al. 2014; Ho, et al. 2018). In the RBC cytosol, these proteins are refolded and ushered to any of several final destinations. Several exported proteins lack a PEXEL, collectively referred to as PEXEL-negative exported proteins (PNEPs) (Spielmann, et al. 2006; Haase, et al. 2009). Most of these have a canonical signal sequence for ER-translocation, which is likely cleaved in the ER by signal peptidase, revealing an N-terminus that resembles that of the cleaved PEXEL (Grüring, et al. 2012).

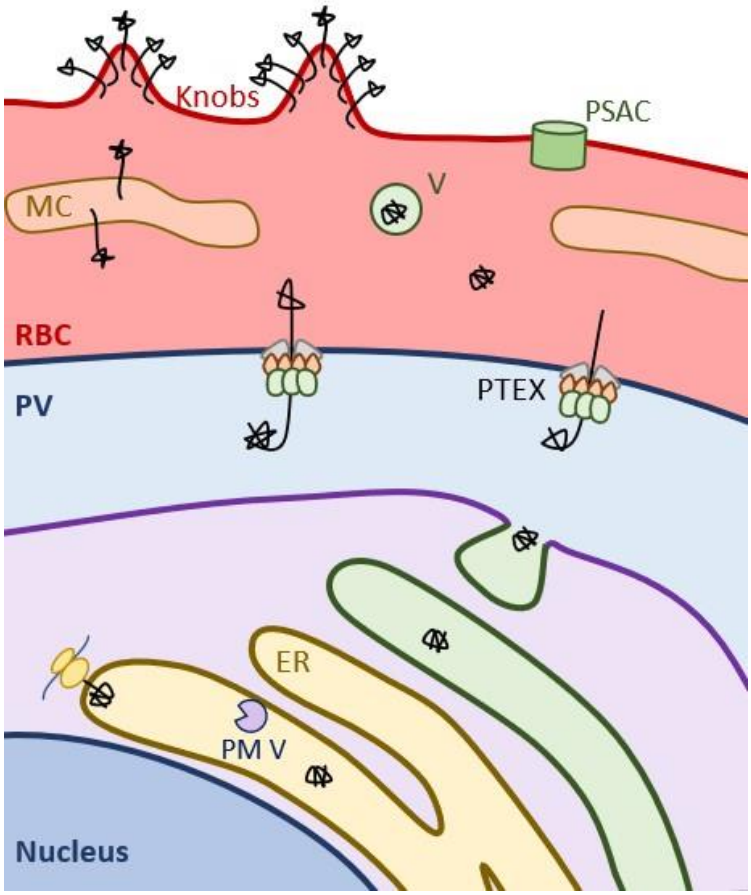


Figure 1.2 – Diagram of *P. falciparum* protein export pathway. Exported proteins are translated into the ER (yellow), then trafficked via the secretory system (green) to the parasitophorous vacuole (PV; blue). Exported proteins are threaded through PTEX, across the PV and into the host cytosol (RBC; red). Some exported proteins traffic via vesicles (V; green) and/or Maurer's clefts (MC; orange) to reach the RBC surface.

Given the central role of protein export in parasite survival and virulence, there has been substantial interest in identifying and characterizing the essential components of the export pathway. The core components of PTEX: the AAA+ ATPase Hsp101, the channel Exp2, and the adapter PTEX150 are each required for protein export and parasite survival (Elsworth, et al. 2014; Beck and Muralidharan, et al. 2014; Garten, et al. 2018; Ho, et al. 2018). Besides PTEX, the export pathway remains murky. Several parasite proteins in the RBC cytosol have been implicated in trafficking adhesins to the RBC surface, but so far none have been shown to be required for parasite growth in culture. Upstream of PTEX, we have evidence that PM V is likely essential for growth, but the extent of its role in export remains unclear, and any additional nodes in the export pathway remain undiscovered.

1.6 Plasmepsin V

Study of PM V began nearly two decades ago, when the sequencing of the P. falciparum genome revealed genes for eight aspartic proteases (in addition to the known plasmepsins I and II) whose function was completely unknown. Antibodies raised against these newly annotated aspartic proteases revealed one, termed plasmepsin V, to be an ER-resident integral membrane protein constitutively expressed across the intraerythrocytic development cycle (Banerjee et al., 2002; Klemba and Goldberg, 2005). The 2004 discovery of PEXEL, and the subsequent realization that the PEXEL motif is cleaved in the ER, sparked a search for the PEXEL protease. The advent of tools for reverse genetics in P. falciparum enabled the assignment of the PEXEL-processing function to PM V (Boddey et al., 2010; Russo et al., 2010). Development of peptidomimetic inhibitors of PM V has driven further discovery, validating PM V as essential for parasite survival in RBCs during the asexual cycle and gametocytogenesis (Jennison et al., 2019; Sleebs et al., 2014a, 2014b) In addition to uncovering novel parasite biology, these inhibitors

have enabled the determination of a high-resolution PM V structure, opening the door to biochemical and pharmacological investigation of this essential enzyme (Hodder et al., 2015). A number of tools have now been turned towards the study of PM V including inducible Di-Cre excision (Boonyalai et al., 2018) and post-transcriptional depletion (Gambini et al., 2015; Polino et al., 2020; Sleebs et al., 2014a), as well as expression, purification and activity assessment in PM V expressed from E. coli (Boddey et al., 2010; Boonyalai et al., 2015; Loymunkong et al., 2019; Sittikul et al., 2018; Xiao et al., 2014), insect cells (Hodder et al., 2015), and in parasite culture (Boddey et al., 2010; Russo et al., 2010), enabling rapid progress in our knowledge of this enzyme, and its development as a potential antimalarial target.

1.6.1 Role of PM V in Protein Export

We now know PM V to be the PEXEL protease. PEXEL-containing proteins are translated into the ER, immediately processed by PM V after the PEXEL Leu, then acetylated at the new N-terminus by an unknown N-acetyltransferase (NAT) (Boddey et al., 2009, 2010; Chang et al., 2008; Osborne et al., 2010; Russo et al., 2010). Preventing PEXEL processing by mutating PEXEL or depleting/inhibiting PM V blocks protein export, indicating that PM V serves as a gatekeeper for the export pathway (Hiller et al., 2004; Marti et al., 2004; Polino et al., 2020; Sleebs et al., 2014a).

PM V is not the only gatekeeper to protein export; several mutated PEXEL reporters are correctly cleaved by PM V but still retained in the parasitophorous vacuole (Boddey et al., 2013; Osborne et al., 2010). Two non-mutually-exclusive models have been put forward to explain this. First, export-destined proteins may contain trafficking information in addition to PEXEL that targets them for export. This information would have to be C-terminal to PEXEL, as N-terminal sequence is removed by PM V. Exported reporter studies have greatly limited the

possible locations of such information. The N-terminus of PEXEL proteins, containing the PEXEL motif followed by just 11 amino acids, supports export of a GFP fusion (Hiller et al., 2004; Marti et al., 2004). Following PEXEL cleavage, the only detectable sequence conservation is the P2' residue (largely restricted to Gln, Glu, or Asp). Residues beyond P2' are important for export, as an inserted PEXEL was not sufficient to re-target the normally vacuole-resident SERA5 to the RBC, but if the subsequent 18 amino acids from the PEXEL protein PfEMP3 were added, export was restored (Boddey et al., 2013). However, the properties of these sequences that support export have remained elusive; they have no obvious conserved sequence, structure, or biochemical properties; and Boddey, et al. (Boddey et al., 2013) reported the surprising finding that even replacing the post-P2' amino acids with all Ala supports protein export, leaving us to wonder what signaling information this sequence could contain.

A variant of this model would be that secreted proteins bind HSP101 and get exported by default, unless they are too tightly folded (Matthews et al., 2019) or contain an N-terminal sequence that prevents chaperone recognition.

The second model posits that following PEXEL cleavage, PM V hands off export-destined cargo to an ER chaperone that ushers it through the export pathway (Russo et al., 2010). In this model, PM V is required not just for cleavage, but also for some downstream step. Tests of this model were attempted with constructs that would create a post-PEXEL N-terminus independent of PM V using either an engineered signal peptidase or a fused viral protease; these gave discordant results, with the signal peptidase-generated N-terminus unable to support export, while the viral protease-generated N-terminus did support export (Boddey et al., 2010; Tarr et al., 2013).

Another avenue for elucidating the post-PM V path is to determine its interacting partners, a task that was carefully undertaken by Danushka, et al. (Marapana et al., 2018). They found that PM V interacts with a non-catalytic component of the signal peptidase complex SPC25 and, if cross-linked, with the ER translocon Sec61, its accessory proteins for post-translational import Sec62/63, as well as signal peptidase. They posited that secretory proteins follow one of two distinct paths: either they are cleaved by the canonical signal peptidase complex, or they are recognized and cleaved by a distinct signal processing complex consisting of PM V and SPC25 (Marapana et al., 2018) (Figure 1.3). Perhaps there are chaperones that recognize this complex selectively, though their identity is not apparent at this time. A variant of the chaperone hand-off model is that PM V could be in a subregion of the ER with a direct route to the PTEX translocon at the PVM (Boddey et al., 2010; Goldberg and Cowman, 2010). Subregions of the PVM have recently been described, with each region hosting distinct secretory proteins (Garten et al., 2020; Nessel et al., 2020). However, there is no information yet about whether these PVM subregions are fed by distinct regions of the ER.

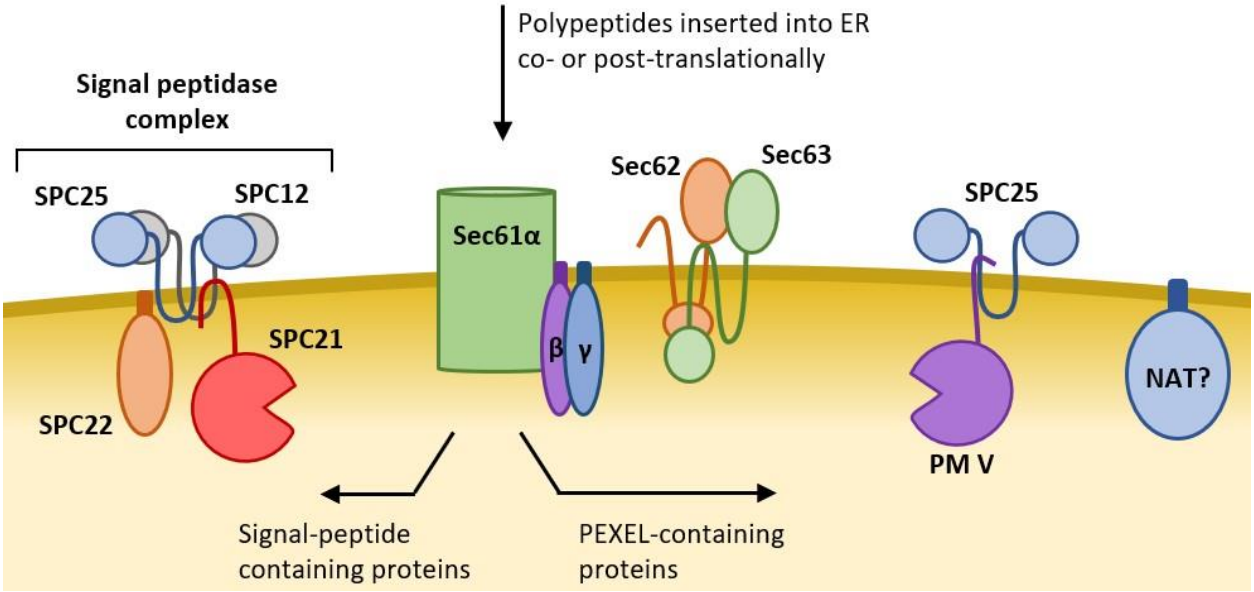


Figure 1.3 – Plasmepsin V in the endoplasmic reticulum. Marapana, et al. (2018) propose a model where proteins translated into the ER are processed either by a signal peptidase-containing complex (left) or a PM V-containing complex (right).

1.6.2 Specificity of PM V

Unlike most other pepsin-like proteases, PM V has a highly specific recognition sequence, cleaving only after the conserved Leu of the consensus PEXEL sequence RxLxE/Q/D (Boddey et al., 2009; Chang et al., 2008; Osborne et al., 2010). Mutation of the highly conserved PEXEL residues P1 or P3 ablate cleavage (with the exception of P1 Ile, which inconsistently supports some PM V cleavage) (Boddey et al., 2013; Hiller et al., 2004; Marti et al., 2004; Schulze et al., 2015). PM V does not cleave the PNEPs SBP1, PfEMP1, and REX2 (Boddey et al., 2013), or a fluorogenic globin-derived peptide cleaved by PM I-IV (Klemba and Goldberg, 2005). Functional PEXELs are found 15 to 30 amino acids downstream of the signal sequence (Hiss et al., 2008); however, whether this location is required for PM V cleavage has not been reported.

1.6.3 PM V Structure

The structure of the *P. vivax* PM V bound to a peptidomimetic inhibitor has been determined and reveals that PM V largely resembles the digestive vacuole plasmepsins, with a few notable differences (Hodder et al., 2015) (Figure 1.4). Like PM I-IV, PM V is produced with a poorly conserved N-terminal extension speculated to be a self-inhibiting prodomain. However, processing of PM V has not been detected (Klemba and Goldberg, 2005) and the full-length recombinant protein is reportedly active (Boonyalai et al., 2015; Xiao et al., 2014). As in other plasmepsins, the PM V active site is covered by a flap that may serve to regulate substrate access. The PM V flap is several amino acids longer than the analogous region of the digestive vacuole plasmepsins, a difference that could perhaps underlie PM V's unique substrate specificity. Along the flap (but not predicted to contact the substrate) is an unpaired Cys, which has been implicated in PM V's sensitivity to Hg^{2+} (Loymunkong et al., 2019; Sittikul et al., 2018; Xiao et al., 2014). Connected directly to the flap is an insert made up of two cysteine pairs folded into a cloverleaf-like structure called a "nepenthesin insert", which it shares with Nep1 from the pitcher plant *Nepenthesia*. The nepenthesin insert is present in PM V across the genus, but absent from other plasmepsins. Its location adjoined to the flap over the active site tempts us to speculate that the nepenthesin insert may respond to some environmental condition or binding partner and regulate access to the active site. Alternatively, the nepenthesin insert could be a set distance from the active site and be involved in measuring distance from the N-end of potential substrates (Goldberg, 2015). No clear hint comes from its ortholog Nep1, a broad-substrate digestive protease present in the low pH of the pitcher plant lumen (Athauda et al., 2004). Another unusual feature is a helix-turn-helix motif, again not present in other plasmepsins. While this is often a nucleic acid-interacting motif, it is not clear what purpose such a motif might have

in the ER lumen, where all but the non-essential C-terminal tail of PM V is found (Russo et al., 2010; Tarr and Osborne, 2015). Projecting away from the active site of the enzyme is a poorly structured insert whose length and sequence varies among different *Plasmodium* species and isolates from different regions (Rawat et al., 2013; Sappakhaw et al., 2015); mutagenesis studies on recombinant enzyme have not yet revealed a function for this structure (Sappakhaw et al., 2015). Further work has the potential to both uncover novel parasite biology and perhaps to unearth secrets of biochemistry broadly applicable to enzymes outside of *Plasmodium*.

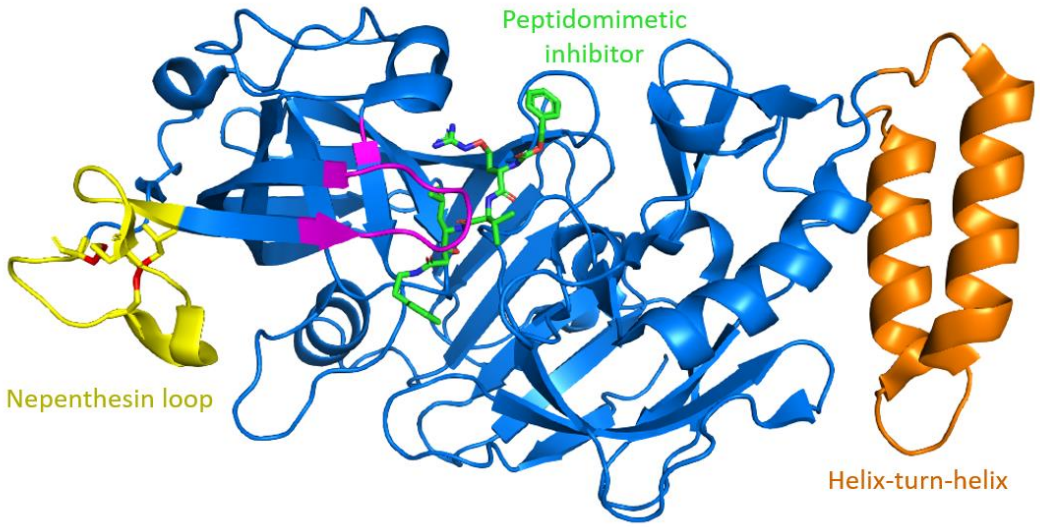


Figure 1.4 – Model of the P. vivax plasmepsin V structure based on Hodder, et al. (2015). The enzyme was crystallized with a peptidomimetic inhibitor (light green). Colored structures are described in the text: nepenthesin loop (yellow), flap over the active site (magenta), and helix-turn-helix motif (orange).

1.7 The PEXEL N-terminus is Acetylated

As mentioned above, the mass spectrometry work that revealed PEXEL to be a proteolytic cleavage site also revealed a surprise: following PEXEL cleavage, the new protein N-terminus is acetylated (Chang, et al. 2008). Brefeldin A treatment (blocking anterograde secretory traffic from the ER) did not block acetylation, suggesting that acetylation – as with PEXEL cleavage – was happening in the parasite ER. This finding was quickly replicated by other groups (Boddey, et al. 2009; Osborne, et al. 2010) leaving us with a series of results that suggest some correlation between N-terminal acetylation and export competence (Table 1.1). Acetylation is not sufficient to drive export, as some mutated PEXEL reporters are properly cleaved and acetylated, yet not exported into the host cell (Chang, et al. 2008; Boddey, et al. 2009). Nor is acetylation completely unique to the post-cleavage PEXEL N-terminus as both the PNEP PfHsp70x and a PEXEL mutant KAHRP-GFP reporter have been shown to be acetylated at the N-terminus following signal peptidase cleavage (Boddey, et al. 2009; Rhiel, et al. 2016). The only protein reported to be exported but not acetylated is REX2, a PNEP that lacks a signal peptide yet is cleaved near its N-terminus by an unknown protease – an unusual case complicated by the fact that we do not yet know where in the secretory pathway REX2 is cleaved (Haase, et al. 2009). Otherwise, the limited data available leave unclear whether acetylation is typically required for protein export. The identity of the responsible NAT(s) remains unknown.

Table 1.1 – Data for acetylation of exported proteins from the literature.

Target	PEXEL	Export?	Acetyl.?	Reference
HRPII	RLLHE	Yes	Yes	Chang, et al. 2008
$HRPII^{N\text{-}term}$-myc	RLLHE	Yes	Yes	Chang, et al. 2008
$PfEMP2^{1\text{-}87}$-GFP	RILSE	Yes	Yes	Chang, et al. 2008
$KAHRP^{1\text{-}60}$-GFP	RILAQ	No	Yes	Chang, et al. 2008
$KAHRP^{1\text{-}69}$-GFP (PEXEL R>A)	AILAQ	No	Yes (after sig. pept. cleavage)	Boddey, et al. 2009
$GBP130^{1\text{-}99}$-YFP	RILAE	Yes	Yes	Boddey, et al. 2009
$GBP130^{1\text{-}99}$-YFP (PEXEL E>A)	RILAA	No	Yes	Boddey, et al. 2009
StAR-related lipid transfer protein (PF3D7_0104200)	Yes		No	Boddey, et al. 2009
Exported protein, unknown function (PF3D7_0402400)	Yes		Yes	Boddey, et al. 2009
Exported protein, unknown function (PF3D7_0501000)	Yes		Yes	Boddey, et al. 2009
PHISTc (PF3D7_0936800)	No (KSLAE)		Yes	Boddey, et al. 2009
Exported lipase 2 (PF3D7_1001600)	Yes		Yes	Boddey, et al. 2009
Exported protein, unknown function (PF3D7_1353100)	Yes		Yes	Boddey, et al. 2009
Exported protein, unknown function (PF3D7_0301700)	No (KSLAE)		Yes	Boddey, et al. 2009
$REX3^{1\text{-}61}$-GFP	RQLSE	Yes	Yes	Osborne, et al. 2010; Tarr, et al. 2013
$REX3^{1\text{-}61}$-GFP-SDEL	RQLSE	No	Yes	Osborne, et al. 2010
$REX3^{1\text{-}61}$-GFP (P1' S>D)	RQLDE	No	No	Tarr, et al. 2013
$REX3^{1\text{-}61}$-GFP (P1'P2' SE>YG}	RQLYG	No	No	Tarr, et al. 2013
REX2-GFP	No (xx|LAE)	Yes	No	Haase, et al. 2009
PfHsp70x-GFP	No	Yes	Yes (after sig. pept. cleavage)	Rhiel, et al. 2016

1.7.1 N-terminal Acetylation

N-terminal acetylation (NTA) is a co- or post-translational modification common to all of life, but particularly widespread in eukaryotes (Deng and Marmorstein, 2021). NTA is a modification of the polypeptide backbone, whereby an acetyl group from acetyl-coenzyme A is irreversibly transferred to the amino group at the polypeptide N-terminus. This acetylation slightly increases the N-terminus' size and blunts its charge. The effects of NTA can vary widely depending on the substrate, with acetylation sometimes required for protein-protein interactions or proper localization, and sometimes affecting protein half-life and stability (Aksnes et al., 2016; Behnia et al., 2004; Monda et al., 2013; Scott et al., 2011). In model eukaryotes, broad swaths of the proteome are N-terminally acetylated: approximately 50-70% of yeast proteins (Arnesen et al., 2009), 70% of Drosophila melanogaster proteins (Goetze et al., 2009), and 80-90% of human proteins (Arnesen et al., 2009). Outside of eukaryote lineages that contain model organisms, the preponderance of NTA remains unexplored.

N-terminal acetylation is undertaken by a family of enzymes called N-terminal acetyltransferases (NATs). Eight eukaryotic NAT complexes have been described, termed NatA through NatH, with the active subunits of each complex designated Naa10, Naa20… Naa80 (Deng and Marmorstein, 2020). NatA, NatB, NatC, NatD, and NatE are typically associated with the ribosome, acetylating peptides as they emerge from the ribosome exit channel (Aksnes et al., 2016). NatF is associated with the cytosolic face of the Golgi and acetylates the cytosolic terminus of transmembrane secretory proteins (Aksnes et al., 2015). NatG is unique to plants and resides in the chloroplast (Dinh et al., 2015). NatH is unique to animals and acetylates β- and γ-actin post-translationally in the cytosol (Wiame et al., 2018). Thus, with the exception of NATG,

all described NAT activity occurs in the cytosol, making the reported NAT activity in the Plasmodium ER all the more intriguing.

All NATs belong to the GCN5-related N-acetyltransferases (GNAT) superfamily of enzymes, which contains all enzymes that transfer an acetyl group from acetyl-coenzyme A to a primary amine. In addition to NATs, this includes the lysine acetyltransferases known for their role in epigenetic modification, as well as acetyltransferases involved in various metabolic processes. Despite overall low sequence homology, GNAT members share four conserved motifs that typically span around 100 amino acids (Vetting et al., 2005). These shared sequences allow the computational prediction of GNAT members, though assignment to subfamily or assessment of biological function is largely impossible without targeted biological assays.

Eight genes in the *Plasmodium falciparum* genome are predicted to contain the GNAT-defining domain: Pf3D7_0109500, Pf3D7_0629000, Pf3D7_0805400, Pf3D7_0823300, Pf3D7_1003300, Pf3D7_1020700, Pf3D7_1323300, and Pf3D7_1437000. Of these, Pf3D7_0629000 has been shown to act as a glucosamine-phosphate N-acetyltransferase (Cove, et al. 2018), while Pf3D7_0823300 has been shown to acetylate histones (Cui, et al. 2007). Five of the remaining six have apparent orthologs in metazoans. The Pf3D7_1020700 ortholog in humans is NAT10, an RNA cytidine acetyltransferase (Lv, et al. 2003). Pf3D7_1003300, Pf3D7_0109500, Pf3D7_0805400, and Pf3D7_1323300 are the closest orthologs of the human Naa10, Naa20, Naa30, and Naa40 respectively (Chen, et al. 2006). The remaining *Plasmodium* GNAT protein, Pf3D7_1437000, is unique to Apicomplexa. It lacks a classical signal sequence for ER entry, but is predicted to contain two transmembrane domains that could grant it access to the secretory compartment. For this reason, we focus on Pf3D7_1437000 as the most likely candidate for the PEXEL NAT.

1.8 Aims and Scope of Thesis

Over the last fifteen years, we have uncovered key elements of the system by which parasites export effectors into the host cell – most recently the function and structure of PTEX. This thesis aims to revisit the contribution of ER-resident proteins to this process in three aims.

The first aim is to characterize a lethal depletion of PM V. We found that depletion of PM V leads to parasite death early in the life cycle, a result distinct from those obtained by disrupting PTEX function. This suggests PM V substrates could have some role(s) within the parasite or parasitophorous vacuole. The second aim probes the role of unusual structural motifs of PM V revealed by a recent structural model. We find that several of these motifs are not required for PM V function; however, a tetracysteine nepenthesin loop is. The third aim seeks to identify the ER-resident NAT responsible for acetylating the PEXEL N-terminus. We have identified an ER-localized NAT and shown it to be essential for parasite growth, but not required for protein export.

Together, the completion of these aims has furthered our understanding of the function of PM V, and the parasite's process for sorting proteins for export into the host cell. Additionally, the aims leave important questions for future investigation: Why is PM V required early in the parasite life cycle? What role does the nepenthesin loop play in PM V activity? Why must the parasite acetylate secretory proteins to complete its life cycle? Addressing these questions promises to further evolve our understanding of *P. falciparum* replication and virulence, as well as the biochemical roles and regulation of proteases.

1.9 References

Aksnes H, Drazic A, Marie M, Arnesen T. 2016. First things first: vital protein marks by N-terminal acetyltransferases. *Trends Biochem Sci* **41**:746–760.

Aksnes H, Van Damme P, Goris M, Starheim KK, et al. 2015. An organellar nα-acetyltransferase, naa60, acetylates cytosolic N-termini of transmembrane proteins and maintains golgi integrity. *Cell Rep* **10**:1362–1374.

Arnesen T, Van Damme P, Polevoda B, Helsens K, et al. 2009. Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proc Natl Acad Sci U S A* **106**:8157–62.

Athauda SBP, Matsumoto K, Rajapakshe S, Kuribayashi M, Ket al. 2004. Enzymic and structural characterization of nepenthesin, a unique member of a novel subfamily of aspartic proteinases. *Biochem J* **381**:295–306.

Banerjee R, Liu J, Beatty W, Pelosof L, et al. 2002. Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. *Proc Natl Acad Sci U S A* **99**:990–5.

Beck JR, Muralidharan V, Oksman A, Goldberg DE. 2014. PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes. *Nature* **511**:592–5.

Behnia R, Panic B, Whyte JRC, Munro S. 2004. Targeting of the Arf-like GTPase Arl3p to the Golgi requires N-terminal acetylation and the membrane protein Sys1p. *Nat Cell Biol* **6**:405–413.

Blasco B, Leroy D, Fidock DA. 2017. Antimalarial drug resistance: linking *Plasmodium falciparum* parasite biology to the clinic. *Nat Med* **23**:917–928.

Boddey JA, Carvalho TG, Hodder AN, Sargeant TJ, et al. 2013. Role of plasmepsin V in export of diverse protein families from the *Plasmodium falciparum* exportome. *Traffic* **14**:532–50.

Boddey JA, Hodder AN, Günther S, Gilson PR, et al. 2010. An aspartyl protease directs malaria effector proteins to the host cell. *Nature* **463**:627–631.

Boddey JA, Moritz RL, Simpson RJ, Cowman AF. 2009. Role of the *Plasmodium* export element in trafficking parasite proteins to the infected erythrocyte. *Traffic* **10**:285–99.

Boonyalai N, Collins CR, Hackett F, Withers-Martinez C, Blackman MJ. 2018. Essentiality of *Plasmodium falciparum* plasmepsin V. *PLoS One* **13**:e0207621.

Boonyalai N, Sittikul P, Yuvaniyama J. 2015. *Plasmodium falciparum* Plasmepsin V (PfPMV): Insights into recombinant expression, substrate specificity and active site structure. *Mol Biochem Parasitol* **201**:5–15.

Bosman A, Mendis KN. 2007. A major transition in malaria treatment: the adoption and deployment of artemisinin-based combination therapies. *Am J Trop Med Hyg* **77**:193–7.

Chang HH, Falick AM, Carlton PM, Sedat JW, et al. 2008. N-terminal processing of proteins exported by malaria parasites. *Mol Biochem Parasitol* **160**:107–15.

Chen F, Mackey AJ, Stoeckert CJ, Roos DS. 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res* **34**:D363-8.

Cova M, López-Gutiérrez B, Artigas-Jerónimo S, González-Díaz A, et al. 2018. The Apicomplexa-specific glucosamine-6-phosphate N-acetyltransferase gene family encodes a key enzyme for glycoconjugate synthesis with potential as therapeutic target. *Sci Rep* **8**:4005.

Cowman AF, Healer J, Marapana D, Marsh K. 2016. Malaria: Biology and Disease. *Cell* **167**:610–624.

de Koning-Ward TF, Gilson PR, Boddey JA, Rug M, et al. 2009. A newly discovered protein export machine in malaria parasites. *Nature* **459**:945–9.

Deng S, Marmorstein R. 2021. Protein N-terminal acetylation: structural basis, mechanism, versatility, and regulation. *Trends Biochem Sci* **46**:15–27.

Despommier DD, Griffin DO, Gwadz RW, Hotez PJ, Knirsch CA. 2019. "The Malarias" in *Parasitic Diseases*, 7th ed. New York: Parasites Without Borders. pp. 93–122.

Dinh TV, Bienvenut WV, Linster E, Feldman-Salit A, et al. 2015. Molecular identification and functional characterization of the first Nα-acetyltransferase in plastids by global acetylome profiling. *Proteomics* **15**:2426–2435.

Elsworth B, Matthews K, Nie CQ, Kalanon M, et al. 2014. PTEX is an essential nexus for protein export in malaria parasites. *Nature* **511**:587–91.

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

Gambini L, Rizzi L, Pedretti A, Taglialatela-Scafati O, et al. 2015. Picomolar inhibition of plasmepsin V, an essential malaria protease, achieved exploiting the prime region. *PLoS One* **10**:e0142509.

Garten M, Beck JR, Roth R, Tenkova-Heuser T, et al. 2020. Contacting domains segregate a lipid transporter from a solute transporter in the malarial host–parasite interface. *Nature Comm* **11**:3825.

Garten M, Nasamu AS, Niles JC, Zimmerberg J, et al. 2018. EXP2 is a nutrient-permeable channel in the vacuolar membrane of *Plasmodium* and is essential for protein export via PTEX. *Nat Microbiol* **3**:1090–1098.

Goetze S, Qeli E, Mosimann C, Staes A, et al. 2009. Identification and functional characterization of N-terminally acetylated proteins in *Drosophila melanogaster*. *PLoS Biol* **7**.

Goldberg DE. 2015. Plasmepsin V shows its carnivorous side. *Nat Struct Mol Biol* **22**:647–648.

Goldberg DE, Cowman AF. 2010. Moving in and renovating: exporting proteins from *Plasmodium* into host erythrocytes. *Nat Rev Microbiol* **8**:617–621.

Grüring C, Heiber A, Kruse F, Flemming S, et al. 2012. Uncovering common principles in protein export of malaria parasites. *Cell Host Microbe* **12**:717–729.

Haase S, Herrmann S, Grüring C, Heiber A, et al. 2009. Sequence requirements for the export of the *Plasmodium falciparum* Maurer's clefts protein REX2. *Mol Microbiol* **71**:1003–1017.

Heiber A, Kruse F, Pick C, Grüring C, et al. 2013. Identification of new PNEPs indicates a substantial non-PEXEL exportome and underpins common features in *Plasmodium falciparum* protein export. *PLoS Pathog* **9**:e1003546.

Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, et al. 2004. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**:1934–7.

Hiss JA, Przyborski JM, Schwarte F, Lingelbach K, Schneider G. 2008. The *Plasmodium* export element revisited. *PLoS One* **3**:e1560.

Ho C-M, Beck JR, Lai M, Cui Y, et al. 2018. Malaria parasite translocon structure and mechanism of effector export. *Nature* **561**:70–75.

Hodder AN, Sleebs BE, Czabotar PE, Gazdik M, et al. 2015. Structural basis for plasmepsin V inhibition that blocks export of malaria proteins to human erythrocytes. *Nat Struct Mol Biol* **22**:590–6.

Ito S, Horikawa S, Tateki S, Kawauchi H, et al. 2014. Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). *J Biol Chem* **289**:35724–30.

Jennison C, Lucantoni L, O'Neill MT, McConville R, et al. 2019. Inhibition of plasmepsin V activity blocks *Plasmodium falciparum* gametocytogenesis and transmission to mosquitoes. *Cell Rep* **29**:3796-3806.e4.

Klemba M, Goldberg DE. 2005. Characterization of plasmepsin V, a membrane-bound aspartic protease homolog in the endoplasmic reticulum of *Plasmodium falciparum*. *Mol Biochem Parasitol* **143**:183–191.

Loymunkong C, Sittikul P, Songtawee N, Wongpanya R, Boonyalai N. 2019. Yield improvement and enzymatic dissection of *Plasmodium falciparum* plasmepsin V. *Mol Biochem Parasitol* **231**:111188.

Marapana DS, Dagley LF, Sandow JJ, Nebl T, et al. 2018. Plasmepsin V cleaves malaria effector proteins in a distinct endoplasmic reticulum translocation interactome for export to the erythrocyte. *Nat Microbiol* **3**:1010–1022.

Marti M, Good RT, Rug M, Knuepfer E, Cowman AF. 2004. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**:1930–3.

Matthews KM, Kalanon M, de Koning-Ward TF. 2019a. Uncoupling the threading and unfoldase actions of P*lasmodium* HSP101 reveals differences in export between soluble and insoluble proteins. *mBio* **10**.

Matthews KM, Pitman EL, de Koning-Ward TF. 2019b. Illuminating how malaria parasites export proteins into host erythrocytes. *Cell Microbiol* **21**:e13009.

Milner DA. 2018. Malaria Pathogenesis. *Cold Spring Harb Perspect Med* **8**:a025569.

Monda JK, Scott DC, Miller DJ, Lydeard J, et al. 2013. Structural conservation of distinctive N-terminal acetylation-dependent interactions across a family of mammalian NEDD8 ligation enzymes. *Structure* **21**:42–53.

Nessel T, Beck JM, Rayatpisheh S, Jami-Alahmadi Y, et al. 2020. EXP1 is required for organisation of EXP2 in the intraerythrocytic malaria parasite vacuole. *Cell Microbiol* **22**:e13168.

Osborne AR, Speicher KD, Tamez PA, Bhattacharjee S, et al. 2010. The host targeting motif in exported *Plasmodium* proteins is cleaved in the parasite endoplasmic reticulum. *Mol Biochem Parasitol* **171**:25–31.

Polino AJ, Armiyaw SN, Niles JC, Goldberg DE. 2020. Assessment of biological role and insight into druggability of the *Plasmodium falciparum* protease plasmepsin V. *ACS Infect Dis* **6**:738–746.

Rawat M, Vijay S, Gupta Y, Tiwari PK, Sharma A. 2013. Imperfect duplicate insertions type of mutations in plasmepsin V modulates binding properties of PEXEL motifs of export proteins in Indian *Plasmodium vivax*. *PLoS One* **8**:e60077.

Rhiel M, Bittl V, Tribensky A, Charnaud SC, et al. 2016. Trafficking of the exported *P. falciparum* chaperone PfHsp70x. *Sci Rep* **6**:36174.

Russo I, Babbitt S, Muralidharan V, Butler T, et al. 2010. Plasmepsin V licenses *Plasmodium* proteins for export into the host erythrocyte. *Nature* **463**:632–6.

Sappakhaw K, Takasila R, Sittikul P, Wattana-Amorn P, et al. 2015. Biochemical characterization of plasmepsin V from *Plasmodium vivax* Thailand isolates: Substrate specificity and enzyme inhibition. *Mol Biochem Parasitol* **204**:51–63.

Sargeant TJ, Marti M, Caler E, Carlton JM, et al. 2006. Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol* **7**:R12.

Schulze J, Kwiatkowski M, Borner J, Schlüter H, et al. 2015. The *Plasmodium falciparum* exportome contains non-canonical PEXEL/HT proteins. *Mol Microbiol* **97**:301–314.

Scott DC, Monda JK, Bennett EJ, Harper JW, Schulman BA. 2009. N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. Science 334:674–678.

Sittikul P, Songtawee N, Kongkathip N, Boonyalai N. 2018. *In vitro* and *in silico* studies of naphthoquinones and peptidomimetics toward *Plasmodium falciparum* plasmepsin V. *Biochimie* **152**:159–173.

Sleebs BE, Gazdik M, O'Neill MT, Rajasekaran P, et al. 2014a. Transition state mimetics of the *Plasmodium* export element are potent inhibitors of plasmepsin V from *P. falciparum* and *P. vivax*. *J Med Chem* **57**:7644–7662.

Sleebs BE, Lopaticki S, Marapana DS, O'Neill MT, et al. 2014b. Inhibition of plasmepsin V activity demonstrates its essential role in protein export, PfEMP1 display, and survival of malaria parasites. *PLoS Biol* **12**:e1001897.

Spielmann T, Hawthorne PL, Dixon MWA, Hannemann M, et al. 2006. A cluster of ring stage–specific genes linked to a locus implicated in cytoadherence in *Plasmodium falciparum* codes for PEXEL-negative and PEXEL-positive proteins exported into the host cell. *Mol Biol Cell* **17**:3613–3624.

Spillman NJ, Beck JR, Goldberg DE. 2015. Protein export into malaria parasite–infected erythrocytes: mechanisms and functional consequences. *Annu Rev Biochem* **84**:813–841.

Tarr SJ, Cryar A, Thalassinos K, Haldar K, Osborne AR. 2013. The C-terminal portion of the cleaved HT motif is necessary and sufficient to mediate export of proteins from the malaria parasite into its host cell. *Mol Microbiol* **87**:835–850.

Tarr SJ, Osborne AR. 2015. Experimental determination of the membrane topology of the *Plasmodium* protease Plasmepsin V. *PLoS One* **10**:e0121786.

Vetting MW, Luiz LP, Yu M, Hegde SS, et al. 2005. Structure and functions of the GNAT superfamily of acetyltransferases. *Arch Biochem Biophys* **433**:212–226.

Wiame E, Tahay G, Tyteca D, Vertommen D, et al. 2018. NAT6 acetylates the N-terminus of different forms of actin. *FEBS J* **285**:3299–3316.

World Health Organization. 2019. 2019 World Malaria Report.

World Health Organization. 2018. 2018 World Malaria Report.

World Health Organization. 2015. *Guidelines for the Treatment of Malaria*, 3rd ed.

Xiao H, Bryksa BC, Bhaumik P, Gustchina A, et al. 2014. The zymogen of plasmepsin v from *Plasmodium falciparum* is enzymatically active. *Mol Biochem Parasitol* **197**:56–63.

Chapter 2: Assessment of biological role and insight into druggability of plasmepsin V

2.1 Preface

This chapter was previously published:

> Polino, A.J., Nasamu, A.S., Niles, J.C., and Goldberg, D.E. (2020). Assessment of biological role and insight into druggability of the Plasmodium falciparum protease plasmepsin V. ACS Infect. Dis. 6, 738-746.

In original publication, this manuscript was accompanied by four supplemental figures:

Figure S1 – Cloning schematic for using pSN054.

Figure S2 – Uncut gels used to generate Figure 2.

Figure S3 – Uncut gels used to generate Figure 3B and C.

Figure S4 – PMV is required for processing and export of RESA.

These are referenced in this chapter as Fig. S1, Fig. S2, Fig. S3, and Fig. S4. They are reproduced at the end of the chapter.

2.2 Abstract

Upon infecting a red blood cell (RBC), the malaria parasite Plasmodium falciparum drastically remodels its host by exporting hundreds of proteins into the RBC cytosol. This protein export program is essential for parasite survival, hence export-related proteins could be potential drug targets. One essential enzyme in this pathway is plasmepsin V (PMV), an aspartic protease that processes export-destined proteins in the parasite endoplasmic reticulum (ER) at the Plasmodium export element (PEXEL) motif. Despite long-standing interest in PMV, functional studies have been hindered by the inability of previous technologies to produce a regulatable lethal depletion of PMV. To overcome this technical barrier, we designed a system for stringent

post-transcriptional regulation allowing a tightly controlled, tunable knockdown of PMV. Using this system, we found that PMV must be dramatically depleted to affect parasite growth, suggesting the parasite maintains this enzyme in substantial excess. Surprisingly, depletion of PMV arrested parasite growth immediately after RBC invasion, significantly before the death from exported protein deficit that has previously been described. The data suggest that PMV inhibitors can halt parasite growth at two distinct points in the parasite life cycle. However, overcoming the functional excess of PMV in the parasite may require inhibitor concentrations far beyond the enzyme's IC50.

2.3 Introduction

Malaria remains a scourge of the developing world, causing nearly 500,000 deaths per year, with the overwhelming majority due to *Plasmodium falciparum* infection (WHO, 2018). While the life cycle of *P. falciparum* includes replication in both the liver and blood, symptomatic human disease is caused by infection of red blood cells (RBCs) (Cowman et al. 2016). Upon infection of a host RBC, the parasite executes a dramatic program of protein export, sending hundreds of proteins through the secretory system, across the surrounding vacuole (parasitophorous vacuole, PV) through a parasite-encoded translocation complex called PTEX, and into the host cytosol (de Koning-Ward et al. 2016; Spillman, Beck, and Goldberg 2015). These exported effectors drastically remodel the host cell, setting up new solute permeability pathways, modifying the RBC shape and rigidity, and reconstituting trafficking machinery in the RBC cytosol to send parasite-encoded adhesins to the RBC surface (de Koning-Ward et al. 2016; Spillman et al. 2015). These adhesins mediate binding of infected RBCs to vascular endothelia allowing parasites to avoid splenic clearance. Adherent parasites in the brain can cause vascular blockage leading to death in severe cases (Cowman et al. 2016).

Due to the central role of protein export in the survival and virulence of *P. falciparum*, there has been interest in this pathway as a source of potential drug targets. One essential enzyme in the pathway is the parasite aspartic protease plasmepsin V (PMV) (Boonyalai et al. 2018; Meyers and Goldberg 2012). PMV processes exported proteins in the parasite ER by cleaving them co-translationally in a variant signal recognition particle complex (Marapana et al. 2018). Cleavage occurs at the conserved amino acid motif RxLxE/Q/D, termed the Plasmodium export element (PEXEL) (Boddey et al. 2010, 2013; Hiller et al. 2004; Marti et al. 2004; Russo et al. 2010). PMV is highly specific for RxL in the PEXEL and cleaves after the leucine (Boddey et al.

2009, 2013; Chang et al. 2008). PEXEL processing is a critical step in protein export, as mutations in PEXEL that block PMV processing also block protein export (Hiller et al. 2004; Marti et al. 2004). Furthermore, processing of PEXEL proteins is likely an essential function in the parasite, as PMV is essential for survival in both *P. falciparum* and the related rodent parasite *P. berghei*, and treatment with a PEXEL peptidomimetic is lethal to intraerythrocytic parasites (Boonyalai et al. 2018; Bushell et al. 2017; Hodder et al. 2015; Sleebs, Lopaticki, et al. 2014).

A number of tools have been used to study PMV, including peptidomimetic inhibitors that block its function in vitro and are lethal to parasites in culture, a DiCre-mediated inducible excision of the gene, and crystallographic studies of *P. vivax* PMV (Boonyalai et al. 2018; Hodder et al. 2015; Sleebs, Gazdik, et al. 2014; Sleebs, Lopaticki, et al. 2014). However, study of PMV function has been hindered by an inability of previous depletions to yield a phenotype in RBC culture. The most robust knockdown described used the glmS ribozyme system, reducing PMV levels 10-fold with no measurable effect on parasite growth or PEXEL processing (Marapana et al. 2018; Sleebs, Lopaticki, et al. 2014). Here, we sought to apply the recently described TetR-DOZI aptamer system for stringent and tunable regulation of PMV (Ganesan et al. 2016). This system can deplete a reporter gene 45 to 70-fold when aptamers are installed in both the 5' and 3' untranslated regions (UTRs) of the target gene (Ganesan et al. 2016). However, cloning such a construct in traditional plasmid systems requires the assembly and maintenance of large circular plasmids that are prone to deletions and vector rearrangements during propagation in *E. coli* (Godiska et al. 2010). To overcome this technical challenge, we assembled a number of tools for genetic manipulation onto a previously described linear vector. Using this new vector system, we achieved substantially greater depletion of PMV than had been reported. By tuning the degree of knockdown, we confirmed that PMV is maintained in

substantial excess during RBC infection and must be suppressed to nearly undetectable levels to affect parasite growth. Finally, we found that PMV-depleted parasites die immediately after invasion in a manner distinct from that of disruptions of other protein export machinery, suggesting PMV may have additional roles beyond those previously described.

2.4 Results

Construction of a linear vector for aptamer knockdowns

To overcome the challenges associated with maintaining *P. falciparum* genomic material in circular plasmids, we utilized the pJAZZ linear vector system (Godiska et al. 2010) as a chassis for DNA assembly. This system has previously been used to manipulate large [A+T]-rich genomic fragments, including those derived from the rodent malaria parasite, *P. berghei* (Godiska et al. 2010; Pfander et al. 2011). We constructed a plasmid ("pSN054") to allow facile cloning, endogenous tagging, robust regulation of expression and inducible knockout of *P. falciparum* genes (Fig. 2.1A, Fig. S1). pSN054 has the following features: a single 5' aptamer, 10x array of 3' aptamers, regulatory protein (TetR-DOZI complex) (Ganesan et al. 2016), parasitemia-tracking component (Renilla luciferase), drug selection marker (Blasticidin S deaminase) (Mamoun et al. 1999), cassette for generation of sgRNA for CRISPR/Cas9 genome editing, cloning sites for inserting homologous sequences for genome repair, modularized affinity tags for tagging genes of interest at the N- or C-terminus, and loxP sites for gene excision (Fig. 2.1A).

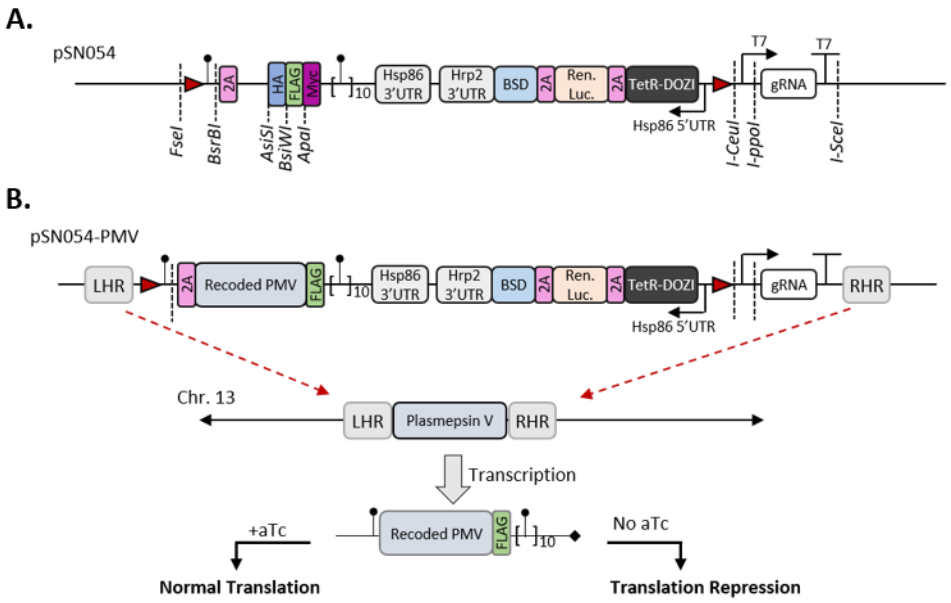


Figure 2.1 – Architecture of pSN054 and application to editing PMV. (A) Schematic of pSN054 showing restriction sites (dashed lines), loxP (red triangles), aptamers (black lollipops), and 2A skip peptide (pink rectangle). Restriction sites allow the choice of 3x-HA (blue), FLAG (green), or Myc (purple). An expression cassette drives production of the Tet repressor-DOZI helicase fusion (TetR-DOZI, black), Renilla luciferase (Ren. Luc.), and blasticidin-S deaminase selectable marker (BSD). The T7 expression cassette drives transcription of CRISPR guide RNAs (gRNA). (B) Cloning strategy for editing of the PMV locus. Left and right homologous regions (LHR and RHR) were inserted at FseI and I-SceI respectively, while the recoded gene sequence was inserted into plasmid cut with AsiSI and BsiWI. The endogenous PMV sequence was disrupted by CRISPR/Cas9 gene editing. When transcribed, aptamers are bound by TetR-DOZI in the absence of aTc, repressing translation. In the presence of aTc, TetR-DOZI does not bind the aptamers and translation occurs as normal.

To utilize pSN054 for genome editing, a gene of interest must be recodonized to replace the current gene, preventing aberrant homologous repair from truncating the construct. To facilitate repair, 400-600bp of homologous sequence corresponding to the 5' and 3' UTR of a gene, are cloned into the FseI and I-SceI restriction sites respectively (Fig. S1). This intermediate vector can be used to create gene knockouts since it does not possess the coding sequence of the gene of interest. A subsequent Gibson assembly inserts the coding sequence of the gene. pSN054 contains the 2A skip peptide (Wagner et al. 2013) such that cloning into the AsiSI site produces a protein with no tag on the N-terminus. If an N- or C-terminal tag is desired, the relevant

restriction site is used for gene insertion (Fig. S1). This donor plasmid can be adapted for use with the T7-RNAP CRISPR/Cas9 system (Wagner et al. 2014) by cloning an sgRNA into the I-ppoI site. Alternatively, a U6 promoter-containing gRNA plasmid (Spillman et al. 2017) can be used and co-transfected with a finished pSN054 donor plasmid.

TetR-DOZI aptamer system for tagging and regulation of PMV

To apply pSN054 to PMV, we cloned pieces of the 5' UTR and 3' UTR into FseI and I-SceI respectively, as well as a recodonized PMV coding sequence into plasmid cut with AsiSI and BsiWI (Fig. 2.1B). This plasmid enabled us to replace the endogenous PMV gene with a FLAG-tagged recodonized PMV flanked by aptamers. The construct was co-transfected into the *P. falciparum* strain NF54^{attB} (Nkrumah et al. 2006) (referred to as "NF54" throughout) along with a separate gRNA-containing plasmid pAIO (Spillman et al. 2017), and parasites selected and cloned. Incorporation of the construct into the genome was verified by Southern blot (Fig. 2.2A). Expression of a FLAG-tagged protein of the expected size was verified by western blot (Fig. 2.2B). Similarly, a western blot with anti-PMV verified that modification of this locus did not substantially change PMV expression levels of the edited line relative to the parent (Fig. 2.2C).

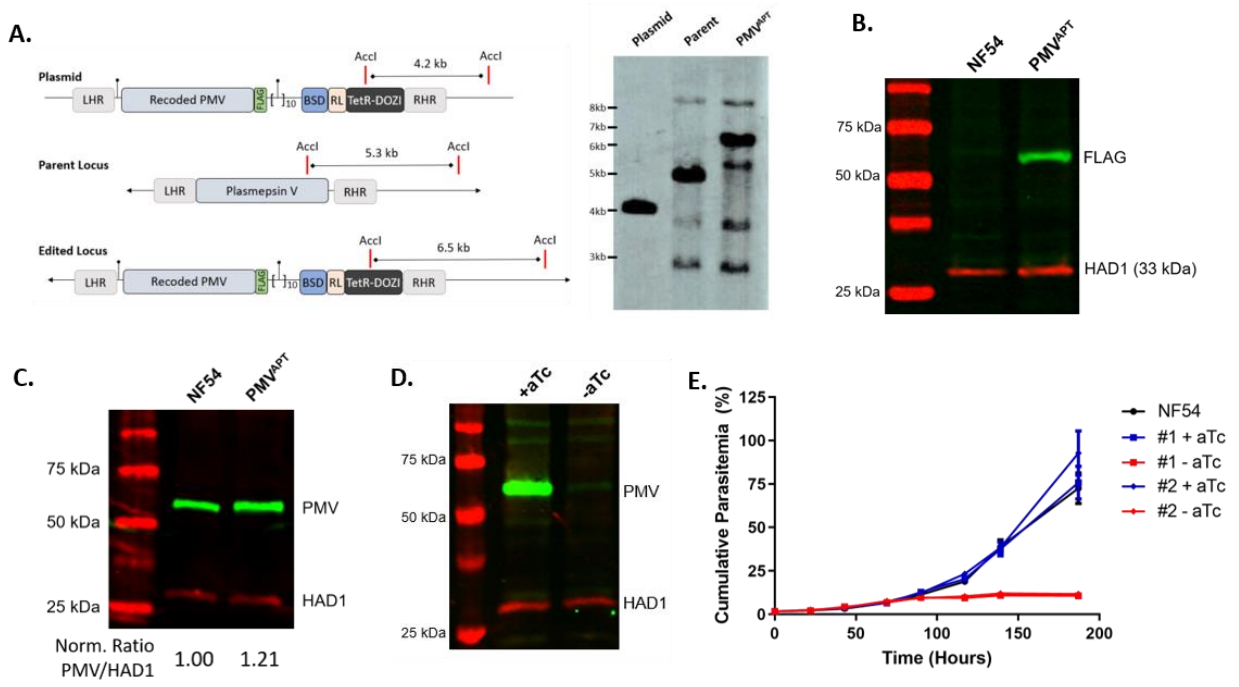


Figure 2.2 – Depletion of PMV demonstrates its essentiality in parasite culture. (A) Tagging was verified by Southern blot with right homologous region (RHR) used as probe. Digest schematic shows expected size of bands. The top-most and bottom three bands in the "PMVAPT" lane are detected in the parent as well and are likely non-specific. (B) Protein tagging was verified by western blot using an anti-FLAG antibody as well as an antibody to a cytosolic sugar phosphatase HAD1 as loading control. (C) Expression levels were compared to the parent strain by western blot using anti-PMV and anti-HAD1 antibodies. An average ratio of PMV to HAD1 signal from two experiments is shown. (D) Knockdown of PMV was initiated by washing aTc from ring-stage cultures and adding either 500nM aTc (+aTc) or an equal volume of DMSO (-aTc). After 72 hours, knockdown was assessed by western blot, probing with anti-PMV and anti-HAD1. Uncut gels used to generate panels (C) and (D) are shown in Fig. S2. (E) Two separate transfected PMV-regulatable lines ("#1" and "#2") were split +/- aTc as above and growth monitored daily by flow cytometry. Experiment was done twice in technical triplicate. A representative experiment is shown, with data points representing mean and error bars standard deviation.

PMV is maintained in substantial excess during infection

We then utilized the TetR-DOZI aptamer system to post-transcriptionally deplete PMV. Knockdown was initiated by washing out anhydrotetracycline (aTc) from young ring-stage parasites in RBC culture. In the absence of aTc, PMV levels were depleted about 50-fold (Fig. 2.2D) and parasite growth was arrested after 96 hours (Fig. 2.2E). This confirms previous reports that PMV is essential for intraerythrocytic growth (Boonyalai et al. 2018; Sleebs, Lopaticki, et al.

2014; Zhang et al. 2018) and showcases the ability of the TetR-DOZI aptamer system to drive more substantial depletion of proteins than previously possible.

We next utilized the tunability of the TetR system to determine the amount of PMV required for parasite survival. To this end, we titrated aTc levels and followed parasite growth by flow cytometry. Parasites maintained in 3 nM aTc or above grew normally. Parasites maintained in DMSO or 1 nM aTc arrested by 96h, while parasites maintained in 2 nM aTc survived an additional cycle before arresting at 120h (Fig. 2.3A). Given this, we sought to quantify the PMV depletion necessary to affect parasite growth. We synchronized parasites and washed out aTc from cultures containing predominantly young ring-stage parasites, then harvested samples for western blot at 72h following washout to compare the PMV levels in parasites that would go on to survive another cycle (2 nM aTc) to those that would arrest in the next 24h (1 nM aTc). Over three independent experiments, we estimate that approximately 8% of wildtype PMV (2 nM aTc) supports another cycle of growth, while 3% (1 nM aTc) was insufficient and led to death upon reinvasion (Fig. 2.3B, C). Since these parasites presumably divide their PMV, diluting it among their daughter cells before the next cycle begins, our data suggest that PMV is maintained in enormous functional excess in parasite culture, with perhaps less than 1% of wildtype PMV likely sufficient to support growth in parasite culture.

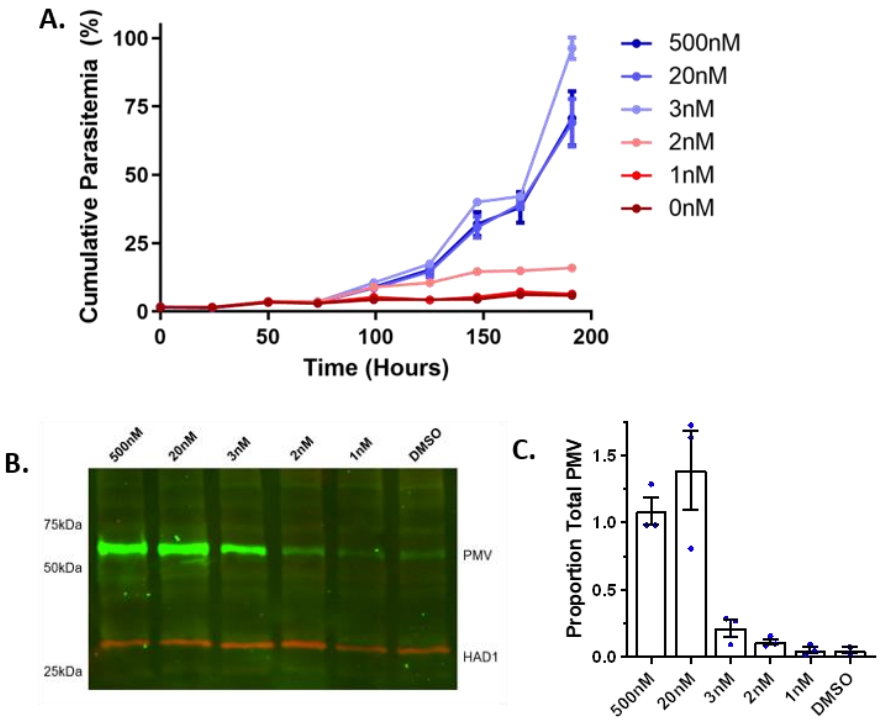


Figure 2.3 – PMV is maintained in substantial excess. (A) Knockdown of PMV was initiated by washing aTc from ring-stage cultures, and parasites maintained in 500nM, 20nM, 3nM, 2nM, 1nM, or 0nM (DMSO) aTc. Growth was monitored daily by flow cytometry. Three experiments were performed with each sample done in technical triplicate. A representative experiment is shown with points representing mean and error bars the standard deviation. (B) Samples were prepared as in (A) but harvested at 72h for western blot. Three experiments were performed. A representative blot is shown. Uncut gels are shown in Fig. S3. (C) Quantification of the three western blots. Blue dots represent mean PMV signal relative to the 500nM signal, as quantified on a diluted standard curve for each experiment (see Methods); bar height represents the mean of the three experiments; error bars represent the standard error of the mean.

PMV-depleted parasites arrest early in life cycle

Given the canonical role of PMV in protein export, we expected PMV depletion to phenocopy disruption of other critical export machinery such as components of the *Plasmodium* translocon for exported proteins (PTEX) which mediates translocation of effectors across the PV membrane (Ho et al. 2018; de Koning-Ward et al. 2009, 2016). PTEX components Hsp101, PTEX150 and EXP2 are all essential in parasite culture, and previously described depletion of

any of these caused parasite arrest during the early trophozoite stage (Beck et al. 2014; Elsworth et al. 2014; Garten et al. 2018). Therefore, we were surprised to find that PMV depletion caused growth arrest very early in the intraerythrocytic development cycle, arresting as “dots” shortly after invasion (Fig. 2.4A, B). Arrested parasites were further investigated by transmission electron microscopy and were found to have gross structural abnormalities (Fig. 2.4C). PMV-depleted parasites generally showed more electron density throughout the parasite cytosol, failure to expand much beyond the size of merozoites, and large unidentified vacuolar structures within the parasite. In contrast, when parasites were fixed as schizonts from the preceding cycle, there was no obvious morphological defect (Fig. 2.4D). This suggests that PMV plays some critical role(s) in successfully initiating the intraerythrocytic development cycle, distinct from its canonical protein export role.

To assess whether PEXEL processing was defective at the time of parasite death, we purified PMV-depleted schizonts and allowed them to invade fresh RBCs for four hours, then probed the status of the early exported protein RESA by western blot and immunofluorescence. PMV depletion led to an accumulation of unprocessed RESA as well as a corresponding defect in RESA export to the surface of the infected RBC (Fig. S4).

Figure 2.4 – PMV-depleted parasites arrest early in life cycle. (A) aTc was washed from ring-stage parasites, which were then maintained in 500nM aTc (+aTc) or an equal volume of DMSO (-aTc). Parasite growth was monitored by Hemacolor-stained thin smear. Experiment was performed three times (B) Phenotype quantification for 0- to 4-hour rings (N = 208 parasites +aTc; 105 parasites –aTc). (C) Parasites maintained as above. At 96h, parasites were fixed, and early ring-stage parasites visualized by transmission electron microscopy. (D) Parasites were maintained as in (C) except schizonts were harvested at 90h following aTc washout. Scale bars represent 1 μm.

PMV inhibitors kill parasites at multiple points in the intraerythrocytic development cycle

Previous work indicated that PMV inhibitors were lethal at the ring-trophozoite transition, consistent with depletion of PTEX components (Hodder et al. 2015). Since this differs from our finding that PMV depletion arrests parasite growth early after invasion, we sought to recapitulate the previously described early trophozoite death as well as our early post-invasion death with the peptidomimetic PMV inhibitor WEHI-842 (Hodder et al. 2015; Sleebs, Gazdik, et

al. 2014). We treated synchronized ring-stage parasites or schizonts with 5 μM WEHI-842 for an 8-hour window, then monitored parasites by thin smear. We found that ring-stage parasites treated with WEHI-842 arrested as early-trophozoites as previously described (Fig. 2.5A). However, parasites treated with WEHI-842 beginning in schizogony arrested immediately after invasion, as was seen in our genetic PMV depletion line (Fig. 2.5B). Taken together, our data suggest that parasites are sensitive to PMV inhibition at two points during asexual growth in RBCs. The first is immediately after invasion. The second is in early trophozoites and is phenocopied by PTEX disruptions.

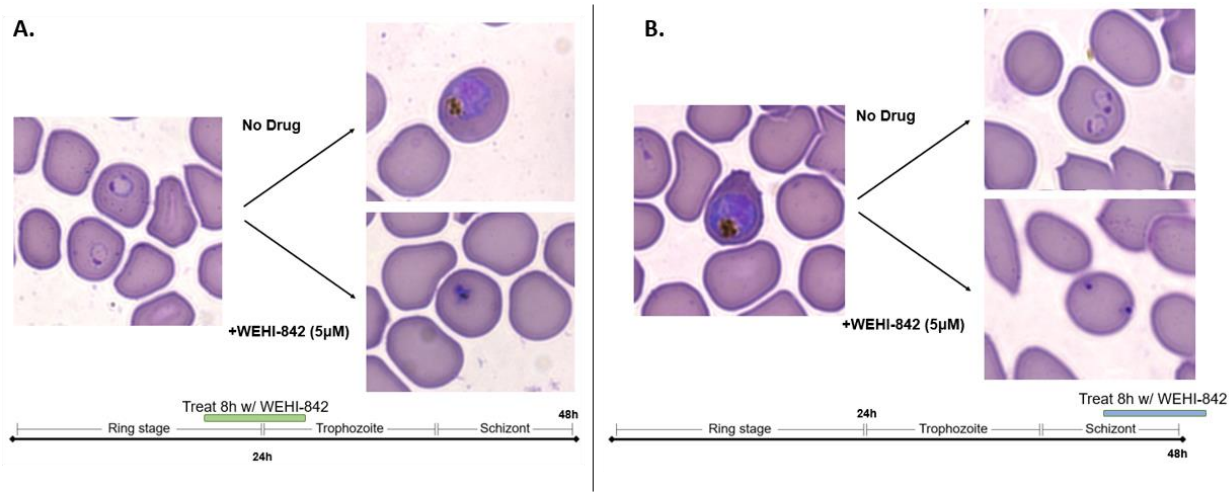


Figure 2.5 – PMV inhibition arrests parasite growth at two distinct points in the life cycle. Parasites were synchronized to within 3 hours, then treated with 5 μM WEHI-842 for an 8-hour window beginning either in (A) ring-stage (12 to 15 hours after invasion) or (B) early schizont (41 to 44 hours after invasion). At the beginning and end of each window, parasites were monitored by Hemacolor-stained thin smear. Experiment was performed twice; representative images from one experiment are shown.

2.5 Discussion

We report the first regulatable knockdown that lowers PMV levels enough to reveal a lethal phenotype. This work overcomes the technical limitations of past knockdown systems by utilizing the TetR system with aptamers on both the 5' and 3' end of the gene of interest. This manipulation was facilitated by the plasmid described here, pSN054, which enables a suite of

previously described genetic tools to be utilized with relative ease for gene editing to achieve protein tagging, regulation of gene expression, parasite growth monitoring and inducible knockout as required.

An unexpected result is that PMV depletion does not phenocopy the disruption of PTEX components (Beck et al. 2014; Elsworth et al. 2014; Garten et al. 2018). The fact that the early death phenotype described here is not seen in disruption of other export machinery suggests that PMV has a role independent of protein export that is essential for parasite survival in RBCs. A reasonable hypothesis is that this role could be the cleavage of a critical substrate early in the life cycle to allow it to perform an essential activity. In this case, the substrate would likely be acting within the parasite or parasitophorous vacuole, since disruption of Hsp101 function with a destabilization domain blocked nearly all exported effectors within the vacuole but only arrested growth in early trophozoites (Beck et al. 2014). One possible source of essential PMV substrates is parasite secretory organelles called dense granules that are involved in establishing the PV (Sheiner and Soldati-Favre 2008). In the Apicomplexan parasite *Toxoplasma gondii*, the PMV ortholog Asp5 primarily cleaves dense granule effectors at a PEXEL-like sequence near their N-termini (Coffey et al. 2018; Hammoudi et al. 2015). Similarly, in *Plasmodium*, the dense granule protein RESA is cleaved at a "relaxed" PEXEL sequence of RxLxxE (Boddey et al. 2013) by PMV (Fig. S4). It is then rapidly secreted into the PV during parasite invasion (Boddey et al. 2013). While RESA is dispensable for intraerythrocytic growth (Silva et al. 2005), other PMV substrates may follow a similar trafficking route and may be required early in the intraerythrocytic development cycle. Alternatively, early death could be a non-specific result of PMV deficiency, such as a buildup of uncleaved PEXEL proteins in the ER. Consistent with this, treatment with the canonical ER-stress inducer DTT arrested growth in *P. falciparum* with

similar morphology by Giemsa stain to that caused by PMV depletion described here (Chaubey, Grover, and Tatu 2014).

One encouraging note for the development of PMV inhibitors as antimalarials is our finding that PMV inhibition can lead to parasite death at two distinct points within intraerythrocytic development. Knockdown of PTEX components seems to cause growth arrest only at the early trophozoite stage in blood-stage parasites (Beck et al. 2014; Elsworth et al. 2014; Garten et al. 2018). Due to this, drugs inhibiting the function of PTEX components may take up to a full intraerythrocytic cycle (48 hours) to reach the point in the cycle where growth arrest occurs. In contrast, PMV inhibitors may arrest growth more quickly by acting upon intraerythrocytic parasites at more than one point in the life cycle. Recent work has also shown that PMV inhibitors can block development of early-stage gametocytes, and have a transmission-blocking effect (Jennison et al. 2019). Together, these findings bolster the case that PMV inhibitors can have properties in line with the target antimalarial profiles put forward by the Medicines for Malaria Venture (Burrows et al. 2017). However, these beneficial characteristics are counterbalanced by our finding that PMV must be suppressed to barely detectable levels to affect parasite growth. Peptidomimetic inhibitors of PMV that have been developed are generally greater than 100-fold less potent in culture than on isolated enzyme (Gambini et al. 2015; Hodder et al. 2015; Sleebs, Lopaticki, et al. 2014). It has been presumed that potency against parasites is limited by cellular permeability. Our functional genetics data would suggest that an additional, and possibly major, component of the potency drop-off is the need to inhibit nearly all the cellular enzyme to kill parasites. It is possible that a potent compound with good biochemical and pharmacokinetic properties could overcome this barrier to sufficiently inhibit PMV during human infection to achieve high antimalarial efficacy.

Taken together, our study provides further data on the proposed antimalarial drug target plasmepsin V. Future work is needed to determine if PMV is maintained at excessive levels *in vivo* as it is *in vitro*, and to elucidate the cause of growth arrest after invasion in PMV-depleted parasites.

2.6 Materials and Methods

Parasite lines and culture. *P. falciparum* strain NF54^{attB} (referred to as NF54 throughout) was used as a parent strain for transfections (Nkrumah et al. 2006). Asexual parasites were cultured in RPMI 1640 (Gibco) supplemented with 0.25% (w/v) Albumax, 15mg/L hypoxanthine, 110mg/L sodium pyruvate, 1.19g/L HEPES, 2.52g/L sodium bicarbonate, 2g/L glucose, and 10mg/L gentamycin. Deidentified RBCs were obtained from the Barnes-Jewish Hospital blood bank (St. Louis, MO), St. Louis Children's Hospital blood bank (St. Louis, MO), and from American Red Cross Blood Services (St. Louis, MO).

Generation of knockdown line. The construct for aptamer regulation of PMV was constructed using pSN054, described above. The right homologous region (3' UTR) was amplified from NF54 genomic DNA using primers AGTGGTGTACGGTACAAACCCGGAATTCGAGCTCGGGGAATCAACATAGAAACGTTAAAG and GATTGGGTATTAGACCTAGGGATAACAGGGTAATGTACTAGGTCATTTTCTTTATTTTAC, and cloned into the I-SceI site using Gibson Assembly (NEB). The left homologous region (5' UTR) was amplified from NF54 genomic DNA using primers TTGGTTTTCAAACTTCATTGACTGTGCCGACATTAATTTGTGTAACATATAAATATGTAG and AAGTTATGAGCTCCGGCAAATGACAAGGGCCGGCCCTTTCCTTAAAAAATAATTATTGAT, and cloned into the FseI site. PMV was codon-optimized for expression in *Saccharomyces cerevisiae* and synthesized as gene blocks by Integrated DNA Technologies (Coralville, IA) then cloned into the

vector at the AsiSI and BsiWI sites. The plasmid was grown in BigEasy Electrocompetent Cells (Lucigen) with 12.5 μg/mL chloramphenicol and 0.01% (w/v) arabinose.

CRISPR/Cas9 editing was performed as previously described (Spillman, et al. 2017). Guide RNA sequences were inserted into the pAIO vector by annealing oligonucleotides of the sequences

ATTAAGTATATAATATT**TGTAATGGTTGTAAAGATTG**GTTTTAGAGCTAGA and

TCTAGCTCTAAAAC**CAATCTTTACAACCATTACA**AATATTATATACTTAAT and inserting them into BtgZI-cut pAIO by In-Fusion HD Cloning (Clontech). pAIO was maintained in XL10 Gold cells (Agilent Technologies). Bold sequences represent the gRNA site.

For each transfection, 100 μg of donor vector and 50 μg of pAIO were transfected into early ring-stage parasites in 2mm gap electroporation cuvettes (Fisher) using a BioRad Gene Pulser II. Transfectants were maintained in 0.5 μM anhydrotetracycline (aTc; Cayman Chemical) and were selected beginning 24-hours post-transfection with Blasticidin S (2.5 μg/mL; Fisher). Parasites were obtained from several independent transfections and clones obtained by limiting dilution.

Validation of PMV^{APT} line. Proper integration of our construct was verified by Southern Blot as in (Spillman, et al. 2017). For a probe, the right homologous region was amplified from NF54 genomic DNA using primers described above.

To verify tagging of protein, schizonts of NF54 and PMV^{APT} were first synchronized by purifying on magnetic columns (Miltenyi Biotech) then allowed to invade fresh uninfected RBCs for 3 hours before remaining schizonts were cleared with 5% sorbitol. Parasites were then allowed to progress for 40h, then RBCs lysed with cold PBS + 0.035% saponin. Samples were centrifuged to pellet parasites and remove excess hemoglobin, then parasites lysed in RIPA

(50mM Tris, pH 7.4; 150mM NaCl; 0.1% SDS; 1% Triton X-100; 0.5% DOC) plus HALT-Protease Inhibitor Cocktail, EDTA-free (Thermo Fisher). Lysates were centrifuged at high speed to pellet and remove hemozoin. Cleared lysates were then diluted in SDS sample buffer (10% SDS, 0.5M DTT, 2.5mg/mL bromophenol blue, 30% 1M Tris pH 6.8, 50% glycerol) and boiled for 5 minutes. Lysates were separated by SDS-PAGE, then transferred to 0.45 μm nitrocellulose membrane (BioRad). Membranes were blocked in PBS + 3% bovine serum albumin, then probed with primary antibodies mouse anti-PMV 1:25 (Banerjee et al. 2002) or anti-FLAG 1:500 (M2, Sigma), and rabbit anti-HAD1 1:1000 (Guggisberg et al. 2014). Membranes were washed in PBS + 0.1% Tween 20, then incubated with secondary antibodies goat anti-mouse IRDye 800CW 1:10,000 (Licor) and donkey anti-rabbit IRDye 680RD 1:10,000 (Licor). Membranes were then washed in PBS + 0.1% Tween 20 and imaged on a Licor Odyssey platform.

Assessment of knockdown. To assess the effect of PMV knockdown on parasite growth, aTc was removed from cultures by washing 3 times for 5 minutes each in media without aTc, then either 500nM aTc ("+ aTc") or DMSO ("- aTc") was added and parasites split into 1mL triplicate cultures for each condition. Parasite growth was monitored daily by flow cytometry (BD FACSCanto) using acridine orange (1.5 μg/mL in PBS). Parasites were sub-cultured every 48 hours to prevent overgrowth. "Cumulative parasitemia" was back-calculated based on the sub-culture schedule. Flow cytometry data is plotted with each point representing the mean of three technical replicates with error bars showing the standard deviation. Experiments were done three times unless otherwise noted.

For titrations of aTc, parasite cultures were washed as above to remove aTc, then cultures were maintained in media containing the shown concentrations of aTc or DMSO (aTc stocks were diluted to ensure each culture received equal volumes of the solvent DMSO in order to

attain the desired final aTc concentration). To determine the effect of aTc concentration on PMV levels, cultures were prepared as above, and samples taken for western blot 72 hours after aTc was removed. Sample preparation and western blotting was done as above (see **Validation of PMV^{APT} line**). To quantitate PMV titration blots, the 500nM sample was diluted out by factors of two to draw a standard curve correlating PMV signal to relative amount of the 500nM sample (see Figure S3). Blots were quantitated using Licor Image Studio. The experiment was performed three times and the mean for those three experiments is plotted with error bars representing the standard error of the mean.

WEHI-842 treatment. NF54^{attB} parasites were synchronized to within 3 hours as above (see **Validation of PMV^{APT} line** above, paragraph 2). Then either 5 μM WEHI-842, or an equal volume of DMSO was added to late rings (15 hours after invasion initiated; parasites 12-15h old), or late trophozoites (44 hours after invasion initiated; parasites 41-44h old). Parasites were incubated with drug or DMSO for 8 hours, then assessed by thin smear. The experiment was performed twice, with similar results in each. Representative images from one experiment are shown.

Microscopy. Parasites monitored by thin smear were dyed using Harleco Hemacolor stains (MilliporeSigma). Images were taken using a Zeiss Axio Observer.D1 at the Washington University Molecular Microbiology Imaging Facility. For transmission electron microscopy, infected RBCs were fixed in 2% paraformaldehyde/2.5% glutaraldehyde (Polysciences Inc., Warrington, PA) in 100 mM sodium cacodylate buffer, pH 7.2 for 1 hour at room temperature. Samples were washed in sodium cacodylate buffer at room temperature and post-fixed in 1% osmium tetroxide (Polysciences Inc.) for 1 hour. Samples were rinsed extensively in deionized water prior to en bloc staining with 1% aqueous uranyl acetate (Ted Pella Inc., Redding, CA) for

1 hour. Following several rinses in deionized water, samples were dehydrated in a graded series of ethanol and embedded in Eponate 12 resin (Ted Pella Inc.). Sections of 95 nm were cut with a Leica Ultracut UCT ultramicrotome (Leica Microsystems Inc., Bannockburn, IL), stained with uranyl acetate and lead citrate, and viewed on a JEOL 1200 EX transmission electron microscope (JEOL USA Inc., Peabody, MA) equipped with an AMT 8-megapixel digital camera and AMT Image Capture Engine V602 software (Advanced Microscopy Techniques, Woburn, MA).

2.7 Closing Material

Author Contributions:

A.S.N and J.C.N conceived the plasmid platform. A.S.N designed and built pSN054 and the pSN054 construct for PMV with supervision from J.C.N. A.J.P designed and performed PMV experiments under supervision of D.E.G.

Database links:

Sequences for genes used in this study were obtained from PlasmoDB (Release 38) using the following gene IDs: plasmepsin V (PF3D7_1323500), HAD1 (PF3D7_1033400) and IspD (Pf3D7_0106900). Also discussed were Hsp101 (PF3D7_1116800), Exp2 (PF3D7_1471100), PTEX150 (PF3D7_1436300), plasmepsin IX (Pf3D7_1430200), and plasmepsin X (Pf3D7_0808200).

Acknowledgements:

We would like to thank Anna Oksman for technical assistance, Eva Istvan and Josh Beck for invaluable comments and suggestions, Dr. Audrey Odom John (Dept. of Pediatrics, Washington University School of Medicine) for anti-HAD1 and anti-IspD, Dr. Wandy Beatty (Dept. of Molecular Microbiology, Washington University School of Medicine) for electron

microscopy, and Dr. Justin Boddey (Walter and Eliza Hall Institute of Medical Research) for WEHI-842.

2.8 References

Banerjee R, Liu J, Beatty W, Pelosof L, et al. 2002. Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. *Proc Natl Acad Sci U S A* **99**:990–5.

Beck JR, Muralidharan V, Oksman A, Goldberg DE. 2014. PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes. *Nature* **511**:592–5.

Boddey JA, Carvalho TG, Hodder AN, Sargeant TJ, et al. 2013. Role of plasmepsin V in export of diverse protein families from the *Plasmodium falciparum* exportome. *Traffic* **14**:532–50.

Boddey JA, Hodder AN, Günther S, Gilson PR, et al. 2010. An aspartyl protease directs malaria effector proteins to the host cell. *Nature* **463**:627–631.

Boddey JA, Moritz RL, Simpson RJ, Cowman AF. 2009. Role of the *Plasmodium* export element in trafficking parasite proteins to the infected erythrocyte. *Traffic* **10**:285–99.

Boonyalai N, Collins CR, Hackett F, Withers-Martinez C, Blackman MJ. 2018. Essentiality of *Plasmodium falciparum* plasmepsin V. *PLoS One* **13**:e0207621.

Burrows JN, Duparc S, Gutteridge WE, Hooft van Huijsduijnen R, et al. 2017. New developments in anti-malarial target candidate and product profiles. *Malar J* **16**:26.

Bushell E, Gomes AR, Sanderson T, Anar B, et al. 2017. Functional profiling of a *Plasmodium* genome reveals an abundance of essential genes. *Cell* **170**:260-272.e8.

Chang HH, Falick AM, Carlton PM, Sedat JW, et al. 2008. N-terminal processing of proteins exported by malaria parasites. *Mol Biochem Parasitol* **160**:107–15.

Chaubey S, Grover M, Tatu U. 2014. Endoplasmic reticulum stress triggers gametocytogenesis in the malaria parasite. *J Biol Chem* **289**:16662–74.

Coffey MJ, Dagley LF, Seizova S, Kapp EA, et al. 2018. Aspartyl Protease 5 matures dense granule proteins that reside at the host-parasite interface in *Toxoplasma gondii*. *MBio* **9**: e01796–18

Cowman AF, Healer J, Marapana D, Marsh K. 2016. Malaria: Biology and Disease. *Cell* **167**:610–624.

de Koning-Ward TF, Dixon MWA, Tilley L, Gilson PR. 2016. *Plasmodium* species: master renovators of their host cells. *Nat Rev Microbiol* **14**:494–507.

de Koning-Ward TF, Gilson PR, Boddey JA, Rug M, et al. 2009. A newly discovered protein export machine in malaria parasites. *Nature* **459**:945–9.

Elsworth B, Matthews K, Nie CQ, Kalanon M, et al. 2014. PTEX is an essential nexus for protein export in malaria parasites. *Nature* **511**:587–91.

Gambini L, Rizzi L, Pedretti A, Taglialatela-Scafati O, et al. 2015. Picomolar inhibition of Plasmepsin V, an essential malaria protease, achieved exploiting the prime region. *PLoS One* **10**:e0142509.

Ganesan SM, Falla A, Goldfless SJ, Nasamu AS, Niles JC. 2016. Synthetic RNA-protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. *Nat Commun* **7**:10727.

Garten M, Nasamu AS, Niles JC, Zimmerberg J, et al. 2018. EXP2 is a nutrient-permeable channel in the vacuolar membrane of *Plasmodium* and is essential for protein export via PTEX. *Nat Microbiol* **3**:1090–1098.

Godiska R, Mead D, Dhodda V, Wu C, et al. 2010. Linear plasmid vector for cloning of repetitive or unstable sequences in *Escherichia coli*. *Nucleic Acids Res* **38**:e88–e88.

Guggisberg AM, Park J, Edwards RL, Kelly ML, et al. 2014. A sugar phosphatase regulates the methylerythritol phosphate (MEP) pathway in malaria parasites. *Nat Commun* **5**:4467.

Hammoudi P-M, Jacot D, Mueller C, Di Cristina M, et al. 2015. Fundamental roles of the golgi-associated *Toxoplasma* aspartyl protease, ASP5, at the host-parasite interface. *PLoS Pathog* **11**:e1005211.

Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, et al. 2004. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**:1934–7.

Ho C-M, Beck JR, Lai M, Cui Y, et al. 2018. Malaria parasite translocon structure and mechanism of effector export. *Nature* **561**:70–75.

Hodder AN, Sleebs BE, Czabotar PE, Gazdik M, et al. 2015. Structural basis for Plasmepsin V inhibition that blocks export of malaria proteins to human erythrocytes. *Nat Struct Mol Biol* **22**:590–6.

Jennison C, Lucantoni L, O'Neill MT, McConville R, et al. 2019. Inhibition of Plasmepsin V activity blocks *Plasmodium falciparum* gametocytogenesis and transmission to mosquitoes. *Cell Rep* **29**:3796-3806.e4.

Mamoun CB, Gluzman IY, Goyard S, Beverley SM, Goldberg DE. 1999. A set of independent selectable markers for transfection of the human malaria parasite Plasmodium falciparum. *Proc Natl Acad Sci U S A* **96**:8716–20.

Marapana DS, Dagley LF, Sandow JJ, Nebl T, et al. 2018. Plasmepsin V cleaves malaria effector proteins in a distinct endoplasmic reticulum translocation interactome for export to the erythrocyte. *Nat Microbiol* **3**:1010–1022.

Marti M, Good RT, Rug M, Knuepfer E, Cowman AF. 2004. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**:1930–3.

Meyers MJ, Goldberg DE. 2012. Recent advances in plasmepsin medicinal chemistry and implications for future antimalarial drug discovery efforts. *Curr Top Med Chem* **12**:445–55.

Nkrumah LJ, Muhle RA, Moura PA, Ghosh P, et al. 2006. Efficient site-specific integration in *Plasmodium falciparum* chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat Methods* **3**:615–621.

Pfander C, Anar B, Schwach F, Otto TD, et al. 2011. A scalable pipeline for highly effective genetic modification of a malaria parasite. *Nat Methods* **8**:1078–1082.

Russo I, Babbitt S, Muralidharan V, Butler T, et al. 2010. Plasmepsin V licenses *Plasmodium* proteins for export into the host erythrocyte. *Nature* **463**:632–6.

Sheiner L, Soldati-Favre D. 2008. Protein Trafficking inside *Toxoplasma gondii*. *Traffic* **9**:636–646.

Silva MD, Cooke BM, Guillotte M, Buckingham DW, et al. 2005. A role for the *Plasmodium falciparum* RESA protein in resistance against heat shock demonstrated using gene disruption. *Mol Microbiol* **56**:990–1003.

Sleebs BE, Gazdik M, O'Neill MT, Rajasekaran P, et al. 2014a. Transition state mimetics of the *Plasmodium* export element are potent inhibitors of Plasmepsin V from *P. falciparum* and *P. vivax*. *J Med Chem* **57**:7644–7662.

Sleebs BE, Lopaticki S, Marapana DS, O'Neill MT, et al. 2014b. Inhibition of Plasmepsin V activity demonstrates its essential role in protein export, PfEMP1 display, and survival of malaria parasites. *PLoS Biol* **12**:e1001897.

Spillman NJ, Beck JR, Ganesan SM, Niles JC, Goldberg DE. 2017. The chaperonin TRiC forms an oligomeric complex in the malaria parasite cytosol. *Cell Microbiol* **19**:e12719.

Spillman NJ, Beck JR, Goldberg DE. 2015. Protein export into malaria parasite–infected erythrocytes: mechanisms and functional consequences. *Annu Rev Biochem* **84**:813–841.

Wagner JC, Goldfless SJ, Ganesan SM, Lee MCS, et al. 2013. An integrated strategy for efficient vector construction and multi-gene expression in *Plasmodium falciparum*. *Malar J* **12**:373.

Wagner JC, Platt RJ, Goldfless SJ, Zhang F, Niles JC. 2014. Efficient CRISPR-Cas9-mediated genome editing in *Plasmodium falciparum*. *Nat Methods* **11**:915–8.

World Health Organization. 2018. 2018 World Malaria Report.

Zhang M, Wang C, Otto TD, Oberstaller J, et al. 2018. Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360**:eaap7847.

2.9 Supplemental Figures

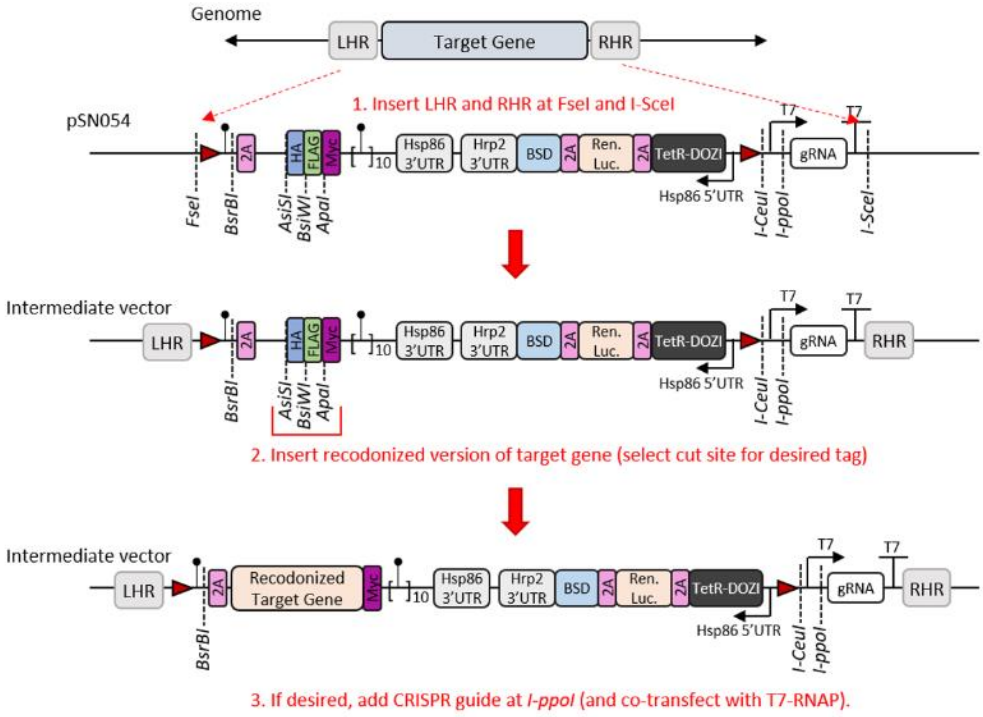


Figure S1 – Cloning schematic for using pSN054. LHR, left homologous region; RHR, right homologous region.

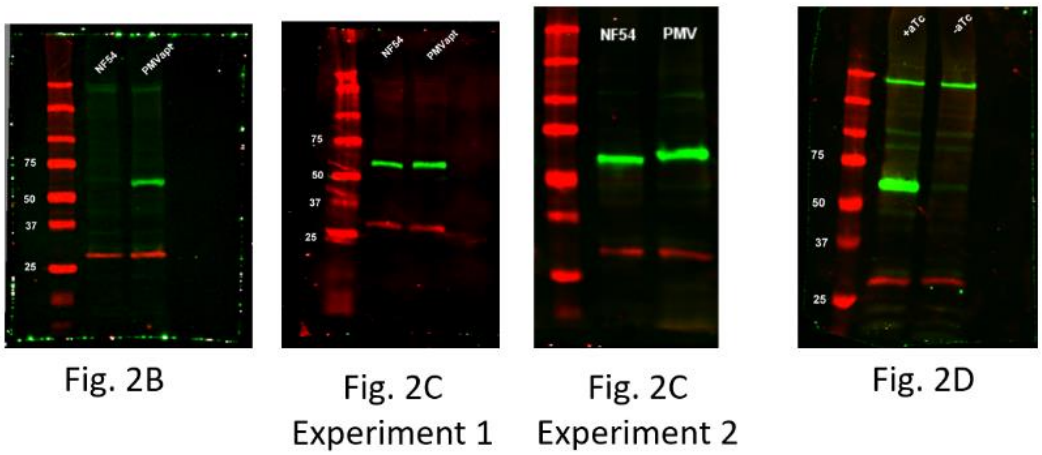
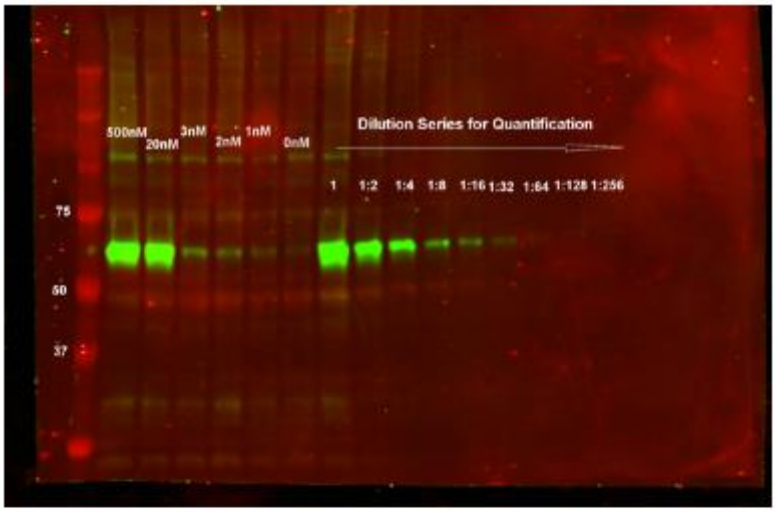
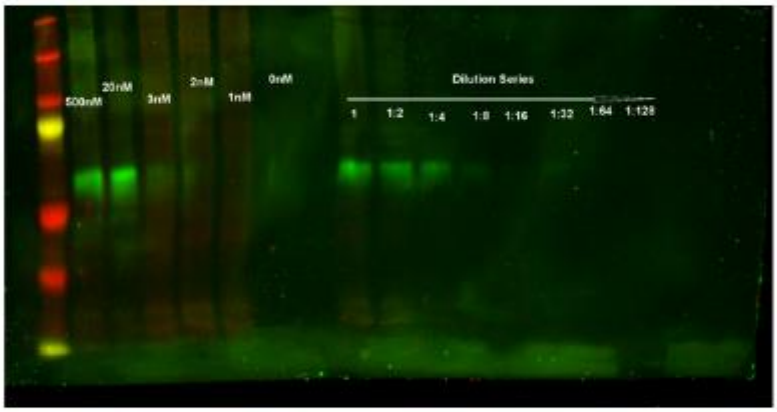


Figure S2 – Uncut gels used to generate Figure 2.





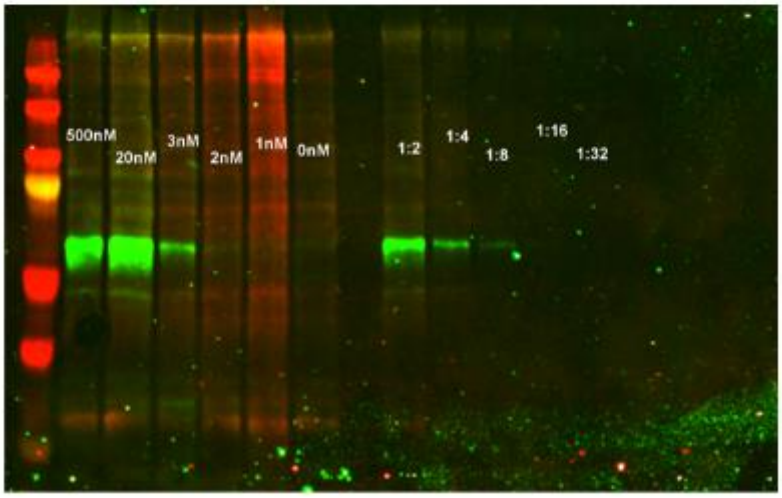


Figure S3 – Uncut gels used to generate Figure 3B and C. Gels for each of the experiments described in Figure 3B and 3C. For each experiment, a dilution series was used to generate a standard curve relating western blot signal to PMV amounts. Experimental samples are lanes 1-6. Dilution series used to generate the curve is labelled.

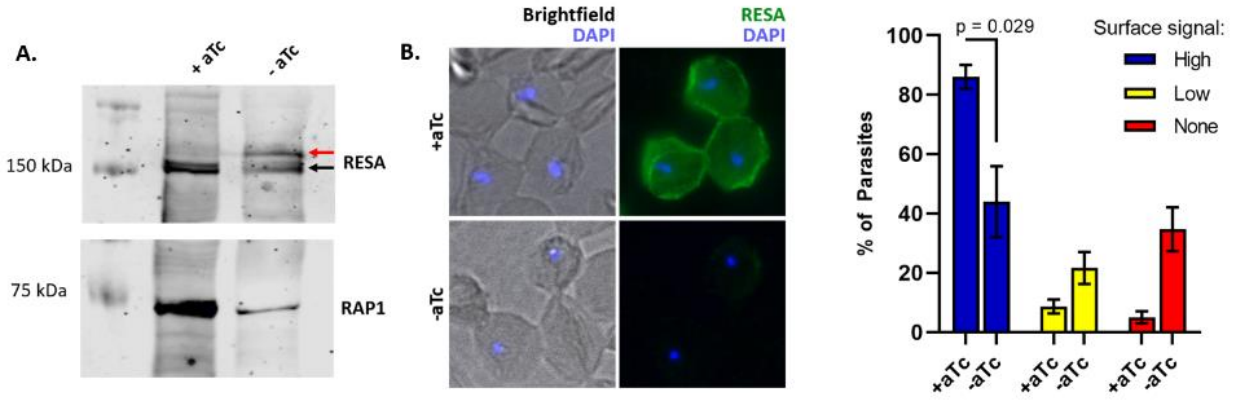


Figure S4 – PMV is required for processing and export of RESA. (A) We assessed RESA processing in synchronized, newly invaded parasites (0 to 4 hours after invasion) by western blot. (B) Newly invaded PMV-depleted parasites were fixed and stained with anti-RESA (green) and DAPI (blue), then scored as high, low, or no surface signal. This experiment was performed three times with 10 images at 63x were taken per condition; illustrative images are shown at left; quantification of the three experiments shown at right (Total parasites scored = 544 +aTc; 194 – aTc). Bar height represents the mean of the three experiments, with error bars representing the SEM; two bars are compared with a Student's t-test and the p-value shown. Intracellular RESA was not detectable by IFA. Antibodies used were mouse anti-RESA (monoclonal 28/2) and anti-RAP1 (monoclonal 2.29).

Chapter 3: Functional analysis of unique structural features of the malaria parasite plasmepsin V

3.1 Preface

This chapter is being prepared for deposition on bioRxiv as:

> Polino AJ, Bobba S, Goldberg DE (2021). Functional analysis of unique structural features of the malaria parasite plasmepsin V. *bioRxiv*.

The text refers to five supplementary figures and one supplementary table, titled:

- Fig. S1: Alignment of the experimentally determined P. vivax PM V and the P. falciparum PM V model
- Fig. S2: Sequence alignment of the reference sequences for P. falciparum, P. vivax, and P. berghei PM V
- Fig. S3: Western blot showing detection of mutant PM V enzymes
- Fig. S4: Full growth curves for Figures 3.2D and 3.3D
- Fig. S5: Inconsistent expression of PM V mutants in HEK293 cells
- Table S1: Primers used in this study

These are not displayed inline, but are instead reproduced at the end of this chapter.

3.2 Abstract

Plasmepsin V (PM V) is a pepsin-like aspartic protease essential for growth of the malaria parasite Plasmodium falciparum. Previous work has shown PM V to be an ER-resident protease that processes parasite proteins destined for export into the host cell. Depletion or inhibition of the enzyme is lethal during asexual replication within red blood cells, as well as during the formation of sexual stage gametocytes. The structure of the P. vivax PM V has been characterized by x-ray crystallography, revealing a canonical pepsin fold punctuated by

structural features uncommon to secretory aspartic proteases. Here we use parasite genetics to probe these structural features by attempting to rescue lethal PM V depletion with various mutant enzymes. We find an unusual nepenthesin 1-type insert to be essential for parasite growth and PM V activity. Mutagenesis of the nepenthesin insert suggests that both its amino acid sequence and one of the two disulfide bonds that undergird its structure are required for the nepenthesin insert's role in PM V function.

3.3 Introduction

Malaria remains a substantial burden on the developing world, with over 200 million cases each year resulting in over 400,000 deaths (World Health Organization, 2020). While the parasite life cycle includes several host tissues, symptomatic human disease results entirely from infection of red blood cells (RBCs). After invasion of an RBC, the parasite dramatically modifies the host cell by exporting hundreds of effector proteins into the host cytosol. This program of protein export results in altered host cell rigidity, increased permeability to nutrients, and the constitution of a secretory apparatus that facilitates the display of parasite-encoded adhesins on the RBC surface (Matthews et al., 2019; Spillman et al., 2015). These adhesins mediate binding of parasitized RBCs to host vascular endothelia, resulting in vascular disruption and, in extreme cases, death of the host.

Most known exported *P. falciparum* effectors bear an export motif termed the *Plasmodium* export element (PEXEL) near the protein N-terminus (Hiller et al., 2004; Marti et al., 2004). PEXEL – with consensus sequence RxLxE/Q/D – is cleaved in the parasite ER by the aspartic protease plasmepsin V (PM V) (Boddey et al., 2010; Russo et al., 2010). Consistent with its role in this essential parasite process, inhibition of PM V activity disrupts protein export and kills parasites (Sleebs et al., 2014). PM V is essential at two distinct points of intraerythrocytic development (Polino et al., 2020), as well as during gametocytogenesis (Jennison et al., 2019), raising the profile of this enzyme as a potential target for chemotherapeutic intervention.

A recent x-ray crystallographic structure of the *P. vivax* PM V bound to the peptidomimetic inhibitor WEHI-842 revealed features unusual for a secretory aspartic protease (Hodder et al., 2015). A nepenthesin 1-type insert and a helix-turn-helix domain feature prominently at opposite surfaces of the protein. The nepenthesin insert – so-called as it was

identified in the protease nepenthesin 1 from the digestive fluid of the pitcher plant *Nepenthesia* – is a tri-lobed structure held together by two cysteine pairs (Athauda et al., 2004). Its functional role is unknown. The helix-turn-helix domain is enigmatic in its own right as the motif canonically mediates protein-nucleic acid binding, which is not expected for PM V in the ER lumen. The structure additionally revealed an unpaired cysteine at the flap over the active site, raising the possibility of redox regulation. Lastly, not visible in the *P. vivax* PM V structure is a predicted flexible loop projecting away from the active site. In *P. falciparum* PM V the corresponding sequence represents a substantially larger and highly charged insert, the role of which we were curious to determine.

Here we utilized our previously described lethal depletion of PM V (Polino et al., 2020) to probe the above features. Using a genetic rescue method, we find the helix-turn-helix domain, unpaired cysteine, and *P. falciparum*-specific charged insert to be dispensable for parasite growth. In contrast, PM V without a nepenthesin insert was unable to rescue PM V depletion. Further mutagenesis of the nepenthesin insert revealed that the sequence of the insert's lobes as well as some elements of its structure are required for its function. Taken together, these data provide further insight into how PM V enacts its essential function(s), and the role of the nepenthesin insert in protease biochemistry.

3.4 Results

PfPM V has several unusual features.

We used Phyre2 to model PM V from the genetically tractable *P. falciparum*, using the recently published *P. vivax* PM V structure (Fig. 3.1) (Hodder et al., 2015; Kelley et al., 2015). The model closely resembles the *P. vivax* structure (Fig. S1), suggesting that the *P. falciparum* enzyme bears the same unusual motifs found in the *P. vivax* enzyme. The nepenthesin insert is

present in PM V across the *Plasmodium* genus, yet is absent in other plasmepsins (Fig. 3.1A). The spacing between the four key cysteines that comprise the insert's structure is conserved among PM V sequences, while the intervening lobe sequences contain some variation (Fig. 3.1A, Fig. S2). The unpaired cysteine over the active site (Fig. 3.1B) is completely conserved in PM V sequences across the genus, but absent in Asp5, its ortholog in the related apicomplexan *Toxoplasma gondii*. The charged insertion (Fig. 3.1C) projecting away from the active site is substantially expanded in *P. falciparum* relative to *P. vivax* and includes an unusually dense run of charged amino acids. Lastly, the helix-turn-helix domain is conserved throughout the genus, but absent in other *P. falciparum* plasmepsins (Fig. 3.1D, Fig. S2).

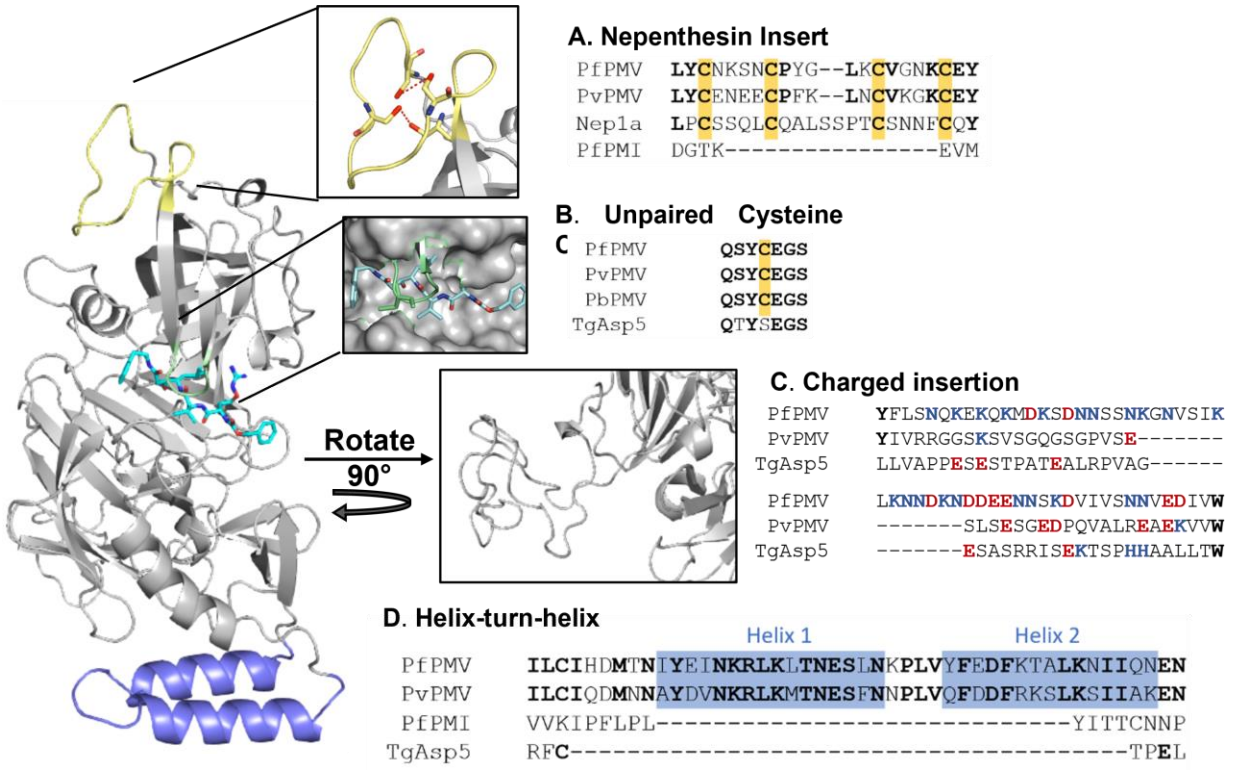


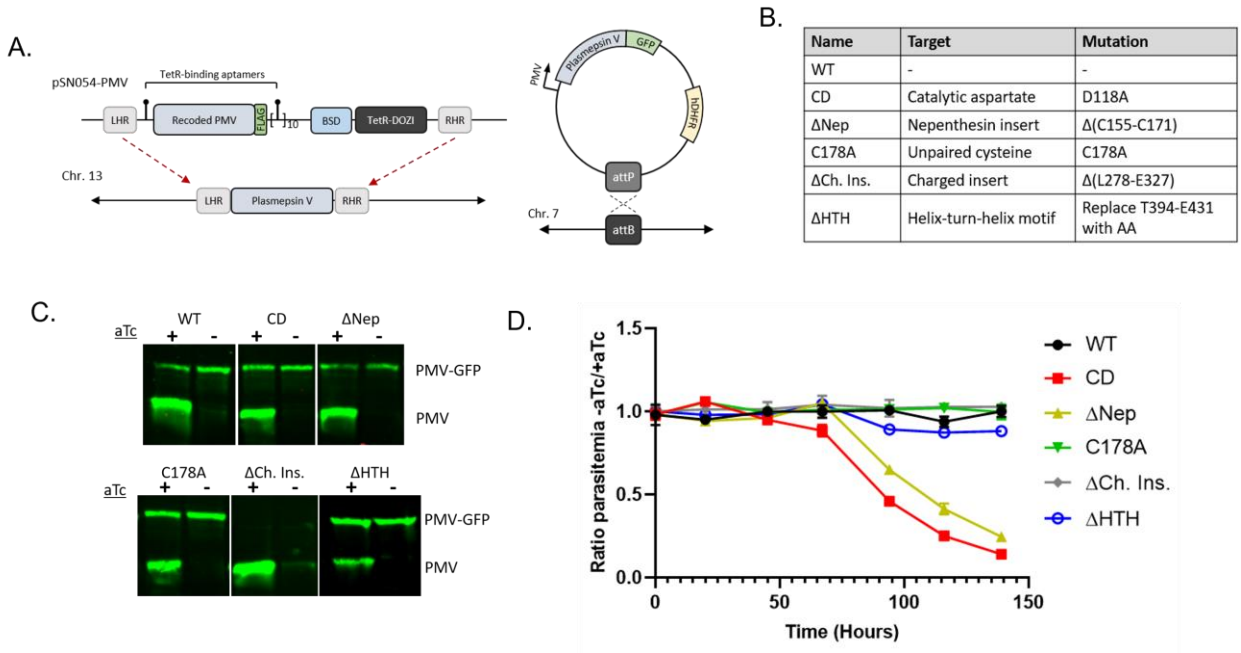
Figure 3.1 – PfPM V structure bears unusual motifs: The structure of PfPM V was modeled using Phyre2 and the published structure from Hodder, Sleebs, Czabotar, et al (2015). The peptidomimetic inhibitor WEHI-842 is shown in bright blue. In this study we investigated (A) the nepenthesin insert, shown in yellow (inset, showing the two Cys-Cys bonds in red), (B) an unpaired cysteine (green) near the active site, (C) a charged insert projecting away from the active site, and (D) a helix-turn-helix domain in blue. At right, the PfPM V sequence aligned next to that of the human parasite *P. vivax*, the rodent parasite *P. berghei*, the related Apicomplexan *Toxoplasma gondii*, and the paralog plasmepsin I (PfPMI). Shared amino acids are bolded. In (A) and (B) cysteines are highlighted in yellow. In (C) positively and negatively charged amino acids are shown in red and blue respectively.

Regulatable system for PM V depletion and rescue

To probe the function of these structural motifs we utilized a previously described post-transcriptional depletion of PM V (Polino et al., 2020). Briefly, we used CRISPR/Cas9 editing to replace the endogenous PM V sequence with a recodonized copy flanked by an N-terminal aptamer and a 10x array of C-terminal aptamers (Fig. 3.2A). In the mRNA, the aptamers fold into hairpins evolved to bind the Tet Repressor (TetR). TetR is fused to the RNA helicase DOZI, and the binding of the TetR-DOZI fusion protein to the mRNA results in sequestration of the

mRNA and depletion of protein levels. The system is controlled by the presence of anhydrotetracycline (aTc). In the presence of aTc, TetR-DOZI preferentially binds aTc and translation continues unabated (Ganesan et al., 2016). In the absence of aTc, TetR-DOZI binds the aptamers and translation is suppressed. Knockdown depletes PM V levels greater than 50-fold and is lethal to parasites (Polino et al., 2020). We complemented the depletion with a second copy of PM V, tagged with C-terminal GFP, and integrated into the neutral *cg6* locus of NF54^{attB} parasites using the previously described attP/attB system (Fig. 3.2A) (Nkrumah et al., 2006).

To probe the role of PM V structural motifs, we made six PM V-GFP constructs: a full-length second copy (WT), a D118A mutation that alters one of the catalytic aspartates ablating catalytic activity ("catalytic dead", or CD), a complete deletion of the nepenthesin insert from the first to fourth cysteine (ΔNep), replacement of the unpaired flap cysteine with alanine (C178A), deletion of the charged insert (ΔCh. Ins.), and replacement of the helix-turn-helix motif with two alanines (ΔHTH) (Fig. 3.2B). Expression of each is driven by the PM V promoter, with the exception of ΔHTH, for which cloning issues forced us to use an alternative plasmid driven by the Hsp86 promoter. Parasite lines were selected using blasticidin S (for the knockdown construct) and WR-99210 (for the second copies) and cloned by limiting dilution. The resulting lines each expressed the endogenous PM V as well as the second copy at approximately 2:1 molar ratio (Fig. S3, Fig. 3.2C); the ΔHTH line, being driven by a stronger promoter, expressed the second copy at an approximately equimolar ratio to the endogenous enzyme (Fig. S3, Fig. 3.2C). Deletion of the charged insert resulted in an enzyme that was no longer recognized by our previously described PM V monoclonal antibody (Banerjee et al., 2002; Klemba and Goldberg, 2005), but was still recognized by anti-GFP (Fig. S3), suggesting that this region contains the binding site for the PM V monoclonal antibody.



Name	Target	Mutation
WT	-	-
CD	Catalytic aspartate	D118A
ΔNep	Nepenthesin insert	Δ(C155-C171)
C178A	Unpaired cysteine	C178A
ΔCh. Ins.	Charged insert	Δ(L278-E327)
ΔHTH	Helix-turn-helix motif	Replace T394-E431 with AA

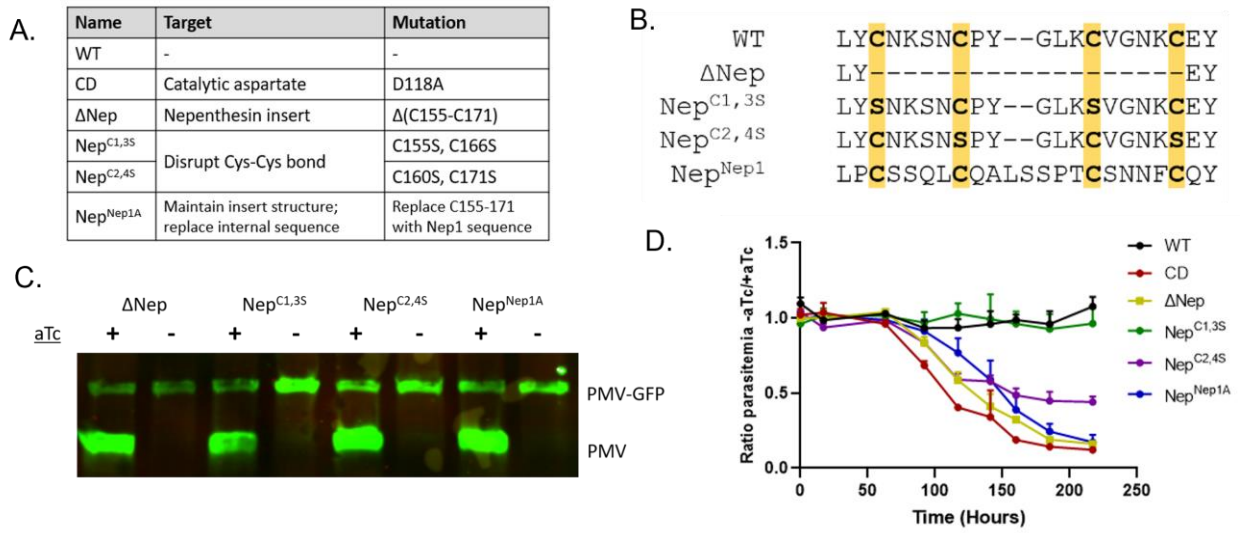
Figure 3.2 – Regulatable system for PM V depletion and rescue: (A) Diagrams of the plasmid system combining pSN054-PMV for tetracyline-regulatable PM V expression, and a GFP-tagged second copy integrated at a genomic attB site and driven by the endogenous promoter. (B) A table of the mutant names in panel C, with the mutations they represent at right. (C) Western blot using anti-PMV antibody. The endogenous copy of PM V is regulatable with aTc; the second copy of mutant PMV-GFP is not. (D) aTc was washed out from asynchronous parasites and growth followed daily by flow cytometry. The experiment was performed three times, each time in technical triplicate. Points represent the mean of technical triplicates, error bars the standard deviation. Growth curves adjusted for subculturing are shown in Supp. Fig. 4.

The nepenthesin insert is required for parasite growth

To probe the role of each of these motifs in PM V function, we depleted endogenous PM V by washing out aTc from parasite culture and followed parasite growth daily by flow cytometry. In each case, aTc washout caused substantial depletion of endogenous PM V levels (Fig. 3.2C). The second copy of WT PM V rescued parasite growth, while CD PMV did not (Fig. 3.2D). To our surprise, parasite lines expressing PM V lacking the unpaired cysteine, charged insert, or helix-turn-helix domain all grew indistinguishably +/- aTc, suggesting each of these motifs is unnecessary for PM V's essential function in parasite growth. Only the ΔNep mutant

failed to rescue growth in the absence of aTc, demonstrating that the nepenthesin insert is required for proper PM V function.

To further dissect the functional contribution of the nepenthesin insert, we mutagenized the insert, attempting to separate its structure from the sequence of its lobes. To this end we made three additional mutants: replacing either the first or second cysteine pairs with serines ($Nep^{C1,3S}$ and $Nep^{C2,4S}$ respectively), or replacing the entire nepenthesin insert with the orthologous sequence from the pitcher plant protease Nep1A (Nep^{Nep1A}) (Fig. 3.3A, B), to preserve the insert's fold while largely altering the lobe sequences. As above, we transfected these as second copies into parasite lines with regulatable endogenous PM V. Washing out aTc from the cultures depleted endogenous PM V levels but had no effect on levels of the second copy (Fig. 3.3C). After washout, we followed growth daily by flow cytometry (Fig. 3.3D). PM V with the Nep1A sequence failed to rescue growth, suggesting that the general structure of the insert is not sufficient for its role in PM V function. Of the two cysteine pair mutants, replacement of the first pair had no effect on growth, while replacement of the second pair resulted in a partial growth phenotype – these parasites grew at approximately half the rate without aTc as with, suggesting that the insert's structure has some contribution to its function.



Name	Target	Mutation
WT	-	-
CD	Catalytic aspartate	D118A
ΔNep	Nepenthesin insert	Δ(C155-C171)
Nep^{C1,3S}	Disrupt Cys-Cys bond	C155S, C166S
Nep^{C2,4S}		C160S, C171S
Nep^{Nep1A}	Maintain insert structure; replace internal sequence	Replace C155-171 with Nep1 sequence

Figure 3.3 – The nepenthesin insert scaffold is not sufficient to rescue nepenthesin loop function (A) Table of mutants for Figure 3 (B) Sequence of the nepenthesin insert for the mutants described in (A). (C) Western blot with anti-PMV showing regulation of the endogenous gene but not the mutant second copy. (D) aTc was washed out from asynchronous parasites and growth followed daily by flow cytometry. The experiment was performed three times, each time in technical triplicate. Points represent the mean of technical triplicates, error bars the standard deviation. Growth curves adjusted for subculturing are shown in Supp. Fig. 4.

The nepenthesin insert is required for PM V catalysis

We next queried whether the nepenthesin mutants failed to rescue growth due to an effect on PM V catalytic activity, or whether they interfered with some other essential PM V role (e.g. a protein-protein interaction). Recombinant expression in mammalian cells afforded production of some but not all of the enzyme mutants (Fig. S5). To employ a system where we could compare the full range of mutants, we immunoprecipitated PM V enzyme mutants using anti-GFP from the parasites described above, and used a previously described FRET peptide assay to assess their activity (Russo et al., 2010). Briefly, we incubated each enzyme prep with either a "PEXEL" peptide derived from the recognition sequence of the PM V substrate HRP2, or a "mutant" peptide that has a single amino acid change, abrogating PM V recognition. Peptide cleavage liberates the fluorophore EDANS from its quencher DABCYL, resulting in fluorescence. The eluate contains each mutant PM V without any detectable signal from the

endogenous enzyme (Fig. 3.4A). Incubation with the FRET peptides indicated that the enzymes that could rescue growth (WT, Nep^{C1,3S}) were active against PEXEL peptide, but not the mutant peptide (Fig. 3.4B). The enzymes that could not rescue growth (CD, ΔNep, Nep^{Nep1A}) could cleave neither peptide. Strangely enough, the Nep^{C2,4S} enzyme which partially rescues growth appears inactive in our enzyme assay, perhaps due to the *in vitro* activity buffer not fully recapitulating the environment of the parasite ER.

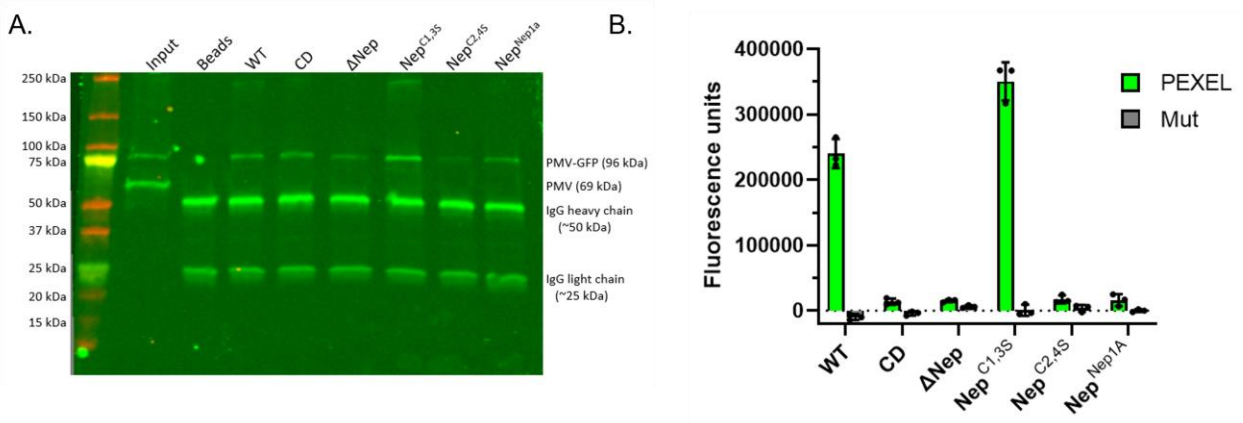


Figure 3.4 – Nepenthesin mutants interfere with PM V catalysis (A) Western blot with anti-PM V showing immunoprecipitations of mutant PM V-GFPs. (B) Mutant enzymes were incubated with a PEXEL FRET peptide or a non-recognized mutant peptide for 12 hours, then fluorescence intensity read as a measure of catalytic activity. Each sample was assayed in technical triplicate, mean fluorescence is shown, with error bars representing standard deviation of the measurements. The experiment was performed three times; a representative experiment is shown.

Given the canonical role of pH in activating pepsin-like proteases, and the example of the Nep1A protease being active at such a low pH, we sought to determine whether nepenthesin mutants would have an altered pH profile. To this end, we ran the same activity assay in buffers ranging from pH 3 to pH 7. We found that the two enzymes active at pH 6.5 (WT and Nep^{C1,3S}) were the only ones to display activity at any pH tested, and each displayed a similar pH to profile

to that previously described (Fig. 3.5) (Russo et al., 2010). Thus, perturbations to the nepenthesin insert do not obviously seem to influence the enzyme's pH/activity relationship.

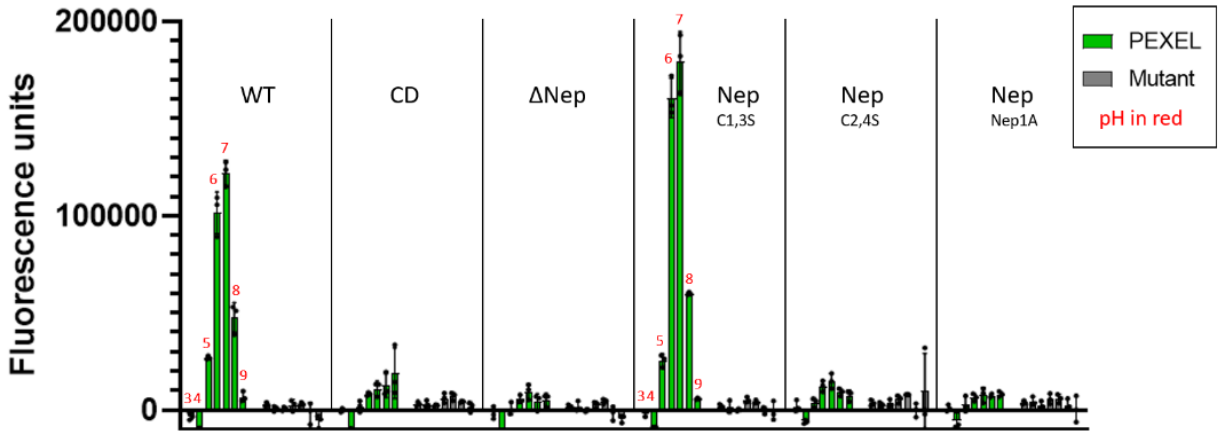


Figure 3.5 – Nepenthesin mutants do not impact pH profile. Activity of each enzyme was assayed as above (Fig. 3.5B) at buffers that differed only in pH, with pH ranging from 3 to 9 (shown above bars with red numbers). Each sample was incubated with either the PEXEL peptide (green) or the mutant (grey). Each sample was run in technical triplicate. The bar height represents the mean of those three measurements; error bars represent the standard deviation. This experiment was performed twice; a representative experiment is shown.

3.5 Discussion

Here we used parasite genetics to investigate the relationship between PM V's unusual structure and its function in parasite biology. We found that the PMV nepenthesin insert is required for parasite survival and for catalysis of a substrate peptide *in vitro*, whereas a helix-turn-helix motif, unpaired cysteine near the active site, and large charged insert were dispensable for parasite survival (and presumably for PM V catalysis).

The most intriguing motif in our story is the nepenthesin insert. This motif was originally described by Athauda, et al. (Athauda et al., 2004) as an approximately 22-residue insertion unique to nepenthesin-like aspartic proteases (and absent from their closest plant orthologues).

The finding that PM V bears a very similar insert was a surprise (Klemba and Goldberg, 2005), given that it has little in common with the nepenthesin-type aspartic proteases. The nepenthesin proteases are digestive enzymes, secreted into the pitcher plant lumen, and activated at very acidic pH (optimal activity at pH ~2.5) to participate in the digestion of insect prey material. Both nepenthesin I and nepenthesin II have broad substrate specificity, and can cleave before or after many residues (Yang et al., 2015). In contrast PM V is retained in the ER at much higher pH (optimal activity at pH ~6.5) and has only been described to cleave C-terminal to RxL (Boddey et al., 2013; Russo et al., 2010). The same motif in both sets of aspartic proteases, but not in others, defies obvious explanation.

Athauda, et al. speculated that the nepenthesin insert may influence localization or regulation of the enzyme; the subsequent finding of a nepenthesin insert in PM V makes at least the former seem less likely. Hodder, Sleebs, Czabotar, et al. suggest that the insert may be involved in interaction with other proteins, perhaps the PEXEL substrates themselves (Hodder et al., 2015). We have not tested this hypothesis here, but can comment that the *Toxoplasma gondii* ortholog Asp5 cleaves an RRL motif that is very similar to the RxL recognized by PM V (18–20) yet appears to have a large insert between the third and fourth cysteine of its orthologous nepenthesin insert sequence, giving spacing of $CX_5CX_3CX_{19}C$ compared to the $CX_4CX_5CX_4C$ spacing in *Plasmodium*. One could imagine that this extra 15-amino acid run would disrupt the nepenthesin insert, however without structural information on Asp5, this remains unclear. We have speculated that the insert may serve as a docking point for chaperones (Goldberg, 2015), yet to our surprise all nepenthesin insert mutations that disrupted parasite growth also blocked catalysis *in vitro*, suggesting that the insert somehow contributes to substrate recognition, access, or cleavage. Thus we are left with an enduring mystery. To our knowledge, this work is the first

to use mutations to probe the nepenthesin insert in any protein. We hope this provides a basis for future hypotheses and experimentation into this poorly understood motif.

As for the other structural motifs studied here, the helix-turn-helix, unpaired cysteine, and large charged insertion each appear dispensable for parasite growth in culture. Helix-turn-helix domains canonically mediate protein-DNA interactions, but have also been described to have roles in protein-RNA interactions, protein-protein interactions, and structural support for enzymatic domains (Aravind et al., 2005). In this case, the conservation of the domain in PM V sequences throughout the genus suggests a significant role for this motif. Our data are limited to blood-stage parasites in culture; perhaps studies in different *Plasmodium* life stages or animal models of infection would reveal a yet unappreciated role.

Similarly, the unpaired cysteine appears dispensable for parasite growth. The enzyme domain of PM V is rich in cysteines, with 15 cysteines arranged into seven pairs – compared to plasmepsin II which has five cysteines arranged into two pairs (Hodder et al., 2015; Silva et al., 1996). This disulfide-rich architecture has earned PM V classification into MEROPS peptidase subfamily A1B, with type member Nep1, rather than the canonical pepsin subfamily A1A (which includes all other plasmepsins) (Rawlings et al., 2014). In A1B subfamily members, the richer network of disulfide bonds has been hypothesized to bolster the enzyme structure against harsh conditions (Athauda et al., 2004). This makes sense for the nepenthesins, which must act in acidic digestive conditions; perhaps for PM V this enhanced structure protects against the oxidizing environment of the ER. The C178A mutation we describe here was also made in *E. coli*-produced PM V by Loymunkong, et al., who observed it to have slightly reduced catalytic efficiency and substantial resistance to inhibition by Hg^{2+} (Loymunkong et al., 2019). Our data suggest these changes to the mutant enzyme do not impact parasite growth in culture. C178 and

its surrounding amino acids are perfectly conserved across the *Plasmodium* genus. Whether that conservation presages a functional contribution at another part of the parasite life cycle, we cannot yet say.

Lastly, the *P. falciparum* genome is inexplicably full of insertions of charged and presumably poorly structured sequences in proteins that lack these insertions in the rest of the genus (with the possible exception of *P. reichenowi*) (Gardner et al., 2002; Singh et al., 2004). These insertions are thought to form surface loops with no known function (Singh et al., 2004). Muralidharan et al. showed that deletion of an asparagine-rich repeat from the proteasome subunit rpn6 had no discernible effect on the protein's essential function (Muralidharan et al., 2011), and suggested that the *P. falciparum* Hsp110 may be so effective at preventing aggregation of repetitive charged sequence as to nearly eliminate the fitness cost of maintaining these sequences (Muralidharan et al., 2012). Here, we describe a similar example: a surface-exposed, low complexity loop, rich in charged amino acids, that has no apparent role in PM V function. Whether these sequences play roles in certain setting or are merely an evolutionary quirk merits further exploration.

In sum, our data provide insight into the essential function of PM V in the *P. falciparum* intraerythrocytic cycle. Given the unusual nature of this enzyme's structure, this also provides initial insight into the contribution of the nepenthesin insert, which we hope will support further investigation into this motif and A1B subfamily proteases more broadly.

3.6 Closing material

Acknowledgements:

We thank members of the Goldberg lab for invaluable suggestions over the course of this project, particularly Eva Istvan and Sumit Mukherjee for sharing their expertise in protein

biochemistry, and Barbara Vaupel for assistance making plasmids. Additionally, we thank Nichole Salinas for her technical assistance with mammalian cell protein expression.

Funding:

This work was supported by the National Institute of Allergy and Infection Diseases (RO1 AI047798; awarded to D.E.G.). A.J.P. was supported by an American Heart Association Predoctoral Fellowship (#18PRE33960417).

3.7 Material & Methods

Modeling and Alignments:

The *P. falciparum* PM V structure was modelled using Phyre2. Mutant sites were selected by aligning the *P. falciparum* PM V structural model with the published *P. vivax* PM V (Hodder et al., 2015) one using PyMOL (v 2.3). Sequence alignments were performed using reference genome sequences retrieved from PlasmoDB (Aurrecoechea et al., 2009) and aligned with Clustal Omega (Madeira et al., 2019).

DNA constructs.

To modify the parasite genome, we used the previously described AttP/AttB recombinase system which utilizes a recombinase-expressing plasmid pINT (Nkrumah et al., 2006) along with a donor plasmid that gets integrated into the genome at the *cg6* locus. Here, our donor plasmid was a derivative of pEOE (Muralidharan et al., 2012; Spillman et al., 2017) modified to include the attP sequence using the QuikChange Lightning Multi Site Directed Mutagenesis kit and the primer **1** (see Sup. Table 1 for list of primers used in this study) as in (Garten et al., 2018) to create the plasmid pEOE-AttP. To make PM V complementation vectors, the full PM V sequence and its 5' UTR were PCR amplified from the NF54 genome using primers **2** and **3**, gel

purified, and inserted by Infusion (Clontech) into pEOE-AttP cut with AatII and AvrII. The resulting plasmid, pEOE-AttP-PMV$_{prom}$-PMV-GFP, was then mutagenized using the QuikChange Lightning Multi Site Directed Mutagenesis kit and primers **5**, **6**, **7**, **8**, **10**, **11**, or **12** to make rescue plasmids with each PM V mutant in Fig. 2B and Fig. 4A. The exception, as noted above, was the ΔHTH mutant. This plasmid was instead made by amplifying the plasmepsin V sequence with primers **2** and **4**, and inserting into pEOE-AttP cut with XhoI and AvrII (maintaining the existing *hsp86* promoter), then mutagenizing as above but with primer **9**. Plasmids sequences were verified by Sanger sequencing (Genewiz; South Plainfield, NJ) using primers **13**, **14**, **15**, and **16**.

For mammalian cell expression, we ordered a synthesized sequence for the soluble domain of PM V (V36-S520) codon optimized for expression in human cells (Integrated DNA Technologies). This sequence was PCR amplified with primers **17** and **18**, then inserted into KpnI-digested pHLsec (Aricescu et al., 2006) by Infusion. Plasmids for expressing mutant enzyme were made as above, using primers **19**, **20**, **21**, **22**, and **23**. Completed plasmids were verified by Sanger sequencing using primers **24** and **25**.

Parasite lines and culture.

P. falciparum strain NF54^{attB} (Nkrumah et al., 2006) was cultured in RPMI 1640 (Gibco) supplemented with 0.25% (w/v) Albumax (Gibco), 15 mg/L hypoxanthine, 110 mg/L sodium pyruvate, 1.19 g/L HEPES, 2.52 g/L sodium bicarbonate, 2 g/L glucose, and 10 mg/L gentamicin as well as human red blood cells at 2% hematocrit. We obtain deidentified human red blood cells from St. Louis Children's Hospital blood bank as well as American Red Cross Blood Services (St. Louis, MO). Cultures were grown in the presence of 100 nM anhydrotetracycline (aTc; Cayman Chemical) unless noted otherwise.

To generate the lines used in this study, we combined 50 µg of each pEOE-AttP-PMV plasmid mentioned above with 50 µg pINT in cytomix (120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 10 mM K_2HPO_4, 25 mM HEPES) and used that to resuspend a pellet from 5 mL parasite culture with ~5% young ring-stage parasites. We then electroporated in a 2mm gap cuvette with a BioRad Gene Pulser II (0.31 kV, 0.950 µF, "High Cap", "Infinite" resistance on Pulse Controller II). After 24 hours we added 10 nM WR99210 as a selection agent, then monitored cultures thrice weekly by thin smear until parasites emerged. Clonal lines were isolated by limiting dilution and verified by PCR and western blot.

Western blotting

Western blots were performed as in (Polino et al., 2020). Briefly, RBCs were lysed in PBS + 0.035% saponin, washed to remove hemoglobin, and the resulting parasite pellets solubilized with RIPA (50 mM Tris pH 7.4, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 0.5% DOC) plus HALT-Protease Inhibitor Cocktail (Thermo Fisher), cleared by centrifugation to remove hemozoin crystals, and then diluted into 4X Protein Loading Buffer (Licor), and boiled at 99°C. Prepared lysates were separated on Mini-PROTEAN TGX Gels (4-15%, BioRad), transferred to 0.45 µm nitrocellulose, blocked in StartingBlock blocking buffer (ThermoFisher Scientific), and probed with mouse anti-PMV 1:50 (Banerjee et al., 2002) and/or rabbit anti-GFP (1:100; Abcam, ab6556), followed by secondary antibodies goat anti-mouse IRDye 800CW 1:10,000 (Licor) and donkey anti-rabbit IRDye 680RD 1:10,000 (Licor). Stained membranes were imaged on a Licor Odyssey platform, and brightness/contrast adjusted for display using ImageStudio Lite 5.2.

Growth assays

To assess the ability of each mutant enzyme to rescue parasite growth, cultures were washed to remove aTc (3 washes, 5 minutes each) and resuspended in medium containing either 100 nM aTc or an equal volume of DMSO (to deplete endogenous PM V). Growth was monitored daily by flow cytometry using a BD FACS Canto and measuring fluorescence from acridine orange (1.5 μg/mL in PBS). Parasites were subcultured as needed to prevent overgrowth. Parasitemias were quantified using BD FACSDiva, and are displayed as a ratio of the parasitemias of the DMSO culture ("-aTc") to the aTc culture (+aTc). Graphs were made using Prism (Version 9).

Peptide assays:

PM V activity assays were performed as in (Russo et al., 2010). Briefly, parasite cultures were pelleted and then RBCs lysed in PBS + 0.035% saponin (10 mL PBS + saponin per 50 mL culture) for 15 min at 4°C. Parasites were pelleted at 15,500 g for 3 min, resuspended in PBS, pelleted again at 15,500 g for 3 min, then frozen at -80°C for days to weeks until the day of the assay. Eventually pellets were thawed and solubilized in PBS + 0.5% Triton X-100, sonicated to disrupt the pellet, then incubated at 4°C for one hour. Lysates were then incubated for one hour at 4°C with Dynabeads-Protein A or Dynabeads-Protein G (ThermoFisher) that had been bound to anti-GFP (clone 3E6, ThermoFisher). Beads were collected on a magnet stand, washed five times in cold PBS + 0.02% Tween-20, and then incubated directly with either the fluorigenic PEXEL peptide DABCYL-LNKRLLHETQ-EDANS, or the mutant peptide DABCYL-LNKRLAHETQ-EDANS at 1 μM in PM V activity buffer (50 mM Tris-Malate, 50 mM NaCl, 0.05% Triton X-100, pH 6.5 unless otherwise noted) for 12 hours at 37°C. Total fluorescence was read on a multimodal plate reader (PerkinElmer Envision). The pH profiling assay was

performed as above, but with PM V activity buffer brought to pH 3 with hydrochloric acid, or brought up to pH 4, 5, 6, 7, 8, or 9 with sodium hydroxide.

Mammalian cell expression:

To express PM V in mammalian cells, we transfected HEK293-F cells (ThermoFisher) using a modified version of the protocol in (Aricescu et al., 2006). Briefly, HEK293-F cells were cultured in FreeStyle 293 Expression Medium (ThermoFisher). A days before transfection, cells were seeded at 6 x 10^5 cells/mL, 100 mL/transfection. The next day, with cells ~10^6 cells/mL, transfection was performed by combining 300 μg polyethylenimine (PEI, 25 kDa branched) with 15 mL FreeStyle 293 Expression Medium, slowly adding 100 μg of the plasmid of interest, allowing to sit for 15 min at room temperature, then adding all to the flask of cells. Cells were incubated shaking at 37°C for 48 hours, then supernatants collected and frozen at -80 °C for up to a week.

On the day of purification, supernatants were thawed and bound to a column of nickel agarose beads (GoldBio; St. Louis, MO), washed five times in wash buffer (20 mM imidazole, 20 mM $H_2NaO_4P*2H_2O$, 500 mM NaCl, pH 7.4), and eluted three times in elution buffer (500 mM imidazole, 20 mM H2NaO4P*2H2O, 500 mM NaCl, pH 7.4). Elutions were then prepared for western blot as above (see “Western blotting”) or used directly for activity assay as above (see “Peptide assays”).

3.8 References

Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM. 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* **29**:231–62.

Aricescu AR, Lu W, Jones EY. 2006. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr Sect D Biol Crystallogr* **62**:1243–1250.

Athauda SBP, Matsumoto K, Rajapakshe S, Kuribayashi M, et al. 2004. Enzymic and structural characterization of nepenthesin, a unique member of a novel subfamily of aspartic proteinases. *Biochem J* **381**:295–306.

Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, et al. 2009. PlasmoDB: A functional genomic database for malaria parasites. *Nucleic Acids Res*.

Banerjee R, Liu J, Beatty W, Pelosof L, et al. 2002. Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. *Proc Natl Acad Sci U S A* **99**:990–5.

Boddey JA, Carvalho TG, Hodder AN, Sargeant TJ, et al. 2013. Role of plasmepsin V in export of diverse protein families from the *Plasmodium falciparum* exportome. *Traffic* **14**:532–50.

Boddey JA, Hodder AN, Günther S, Gilson PR, et al. 2010. An aspartyl protease directs malaria effector proteins to the host cell. *Nature* **463**:627–631.

Ganesan SM, Falla A, Goldfless SJ, Nasamu AS, Niles JC. 2016. Synthetic RNA-protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. *Nat Commun* **7**:10727.

Gardner MJ, Hall N, Fung E, White O, et al. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**:498–511.

Garten M, Nasamu AS, Niles JC, Zimmerberg J, et al. 2018. EXP2 is a nutrient-permeable channel in the vacuolar membrane of *Plasmodium* and is essential for protein export via PTEX. *Nat Microbiol* **3**:1090–1098.

Goldberg DE. 2015. Plasmepsin V shows its carnivorous side. *Nat Struct Mol Biol* **22**:647–648.

Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, et al. 2004. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**:1934–7.

Hodder AN, Sleebs BE, Czabotar PE, Gazdik M, et al. 2015. Structural basis for plasmepsin V inhibition that blocks export of malaria proteins to human erythrocytes. *Nat Struct Mol Biol* **22**:590–6.

Jennison C, Lucantoni L, O'Neill MT, McConville R, et al. 2019. Inhibition of plasmepsin V activity blocks *Plasmodium falciparum* gametocytogenesis and transmission to mosquitoes. *Cell Rep* **29**:3796-3806.e4.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**:845–858.

Klemba M, Goldberg DE. 2005. Characterization of plasmepsin V, a membrane-bound aspartic protease homolog in the endoplasmic reticulum of *Plasmodium falciparum*. *Mol Biochem Parasitol* **143**:183–191.

Loymunkong C, Sittikul P, Songtawee N, Wongpanya R, Boonyalai N. 2019. Yield improvement and enzymatic dissection of *Plasmodium falciparum* plasmepsin V. *Mol Biochem Parasitol* **231**:111188.

Madeira F, Park YM, Lee J, Buso N, et al. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**:W636–W641.

Marti M, Good RT, Rug M, Knuepfer E, Cowman AF. 2004. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**:1930–3.

Matthews KM, Pitman EL, de Koning-Ward TF. 2019. Illuminating how malaria parasites export proteins into host erythrocytes. *Cell Microbiol* **21**:e13009.

Muralidharan V, Oksman A, Iwamoto M, Wandless TJ, Goldberg DE. 2011. Asparagine repeat function in a *Plasmodium falciparum* protein assessed via a regulatable fluorescent affinity tag. *Proc Natl Acad Sci U S A* **108**:4411–6.

Muralidharan V, Oksman A, Pal P, Lindquist S, Goldberg DE. 2012. *Plasmodium falciparum* heat shock protein 110 stabilizes the asparagine repeat-rich parasite proteome during malarial fevers. *Nat Commun* **3**:1310.

Nkrumah LJ, Muhle RA, Moura PA, Ghosh P, et al. 2006. Efficient site-specific integration in *Plasmodium falciparum* chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat Methods* **3**:615–621.

Polino AJ, Nasamu AS, Niles JC, Goldberg DE. 2020. Assessment of biological role and insight into druggability of the *Plasmodium falciparum* protease plasmepsin V. *ACS Infect Dis* **6**:738–746.

Rawlings ND, Waller M, Barrett AJ, Bateman A. 2014. *MEROPS* : the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* **42**:D503–D509.

Russo I, Babbitt S, Muralidharan V, Butler T, et al. 2010. Plasmepsin V licenses *Plasmodium* proteins for export into the host erythrocyte. *Nature* **463**:632–6.

Silva AM, Lee AY, Gulnik S V, Maier P, et al. 1996. Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* **93**:10034–9.

Singh GP, Chandra BR, Bhattacharya A, Akhouri RR, et al. 2004. Hyper-expansion of asparagines correlates with an abundance of proteins with prion-like domains in *Plasmodium falciparum*. *Mol Biochem Parasitol* **137**:307–319.

Sleebs BE, Lopaticki S, Marapana DS, O'Neill MT, et al. 2014. Inhibition of Plasmepsin V activity demonstrates its essential role in protein export, PfEMP1 display, and survival of malaria parasites. *PLoS Biol* **12**:e1001897.

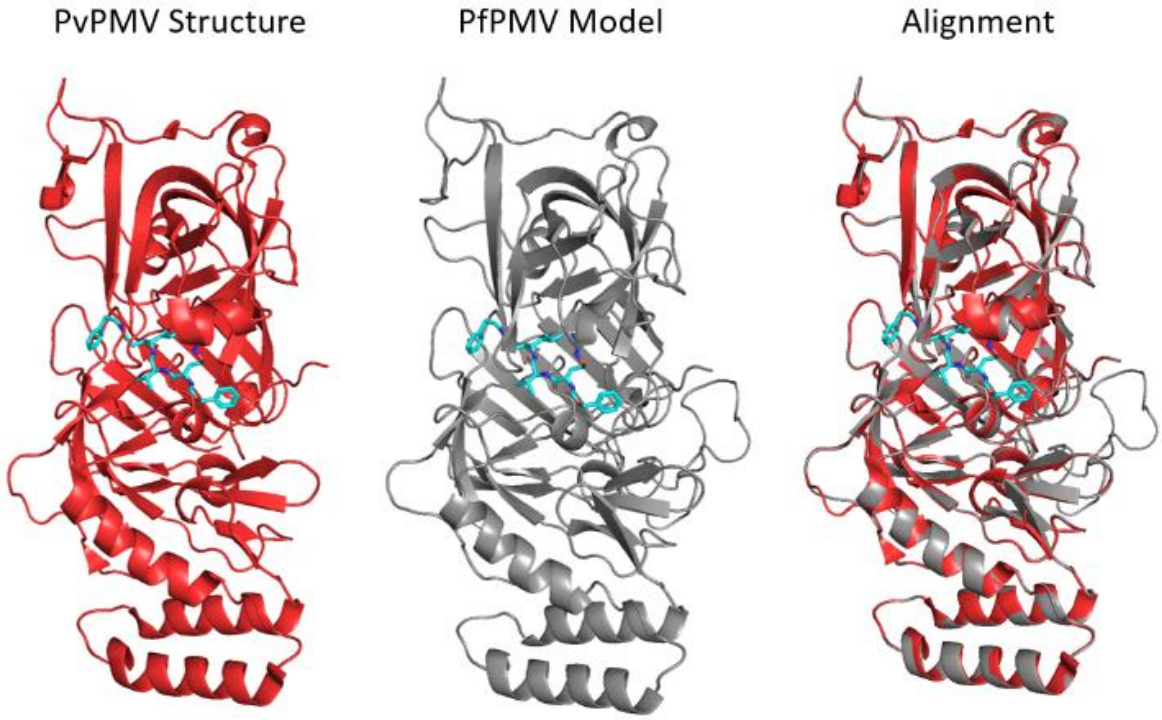
Spillman NJ, Beck JR, Ganesan SM, Niles JC, Goldberg DE. 2017. The chaperonin TRiC forms an oligomeric complex in the malaria parasite cytosol. *Cell Microbiol* **19**:e12719.

Spillman NJ, Beck JR, Goldberg DE. 2015. Protein export into malaria parasite–infected erythrocytes: mechanisms and functional consequences. *Annu Rev Biochem* **84**:813–841.

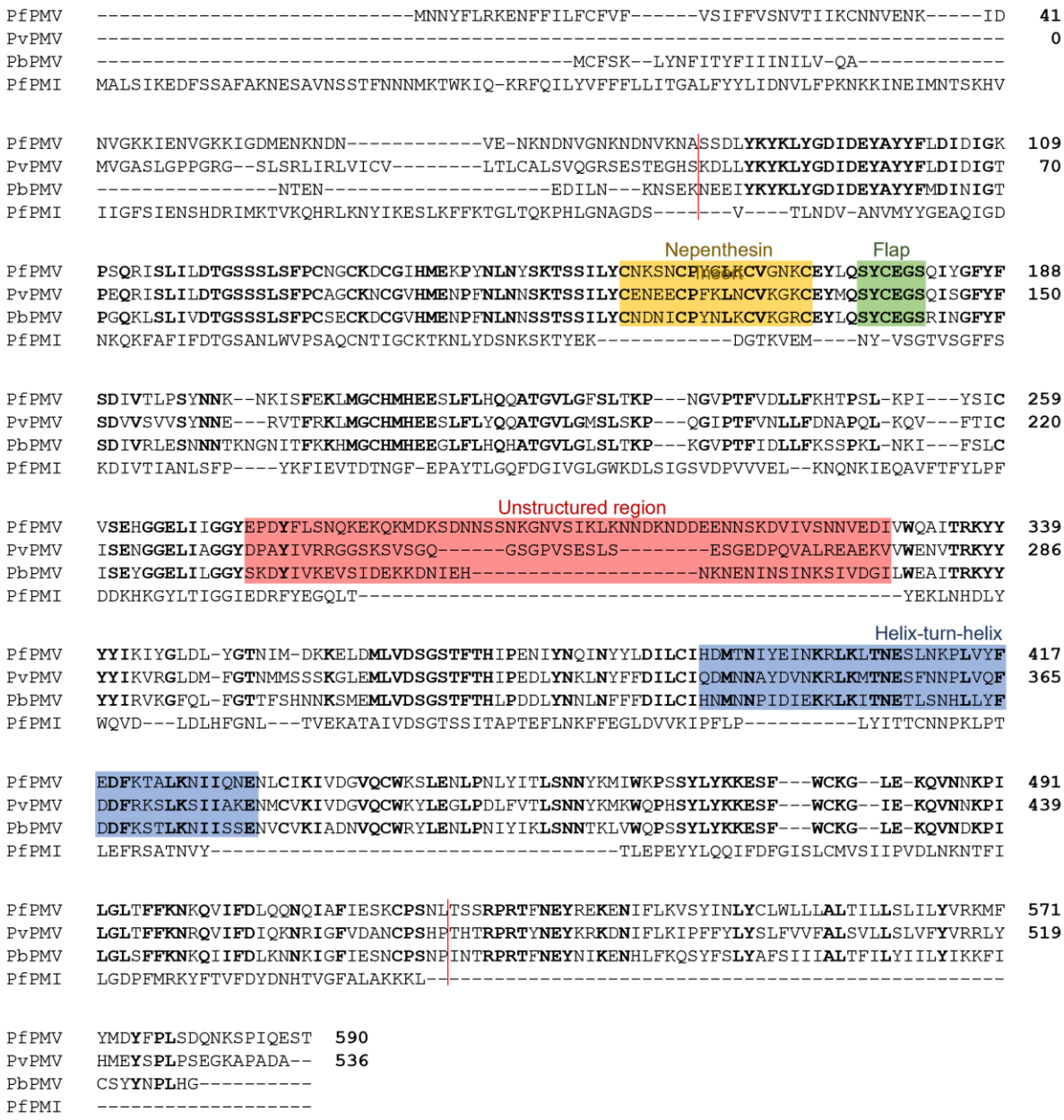
World Health Organization. 2020. World Malaria Report 2020.

Yang M, Hoeppner M, Rey M, Kadek A, et al. 2015. Recombinant nepenthesin II for hydrogen/deuterium exchange mass spectrometry. *Anal Chem* **87**:6681–6687.

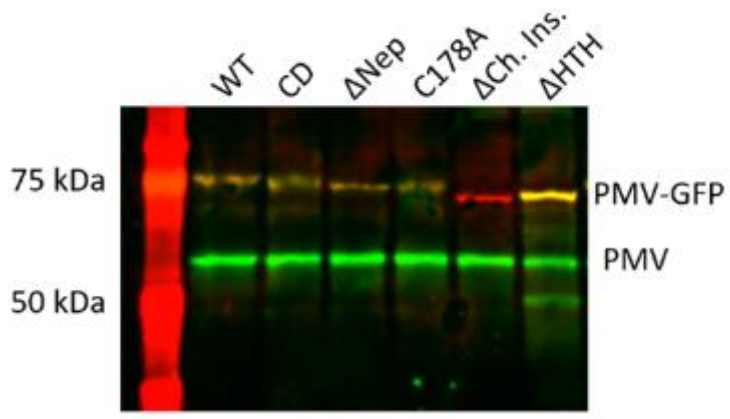
3.9 Supplemental Material



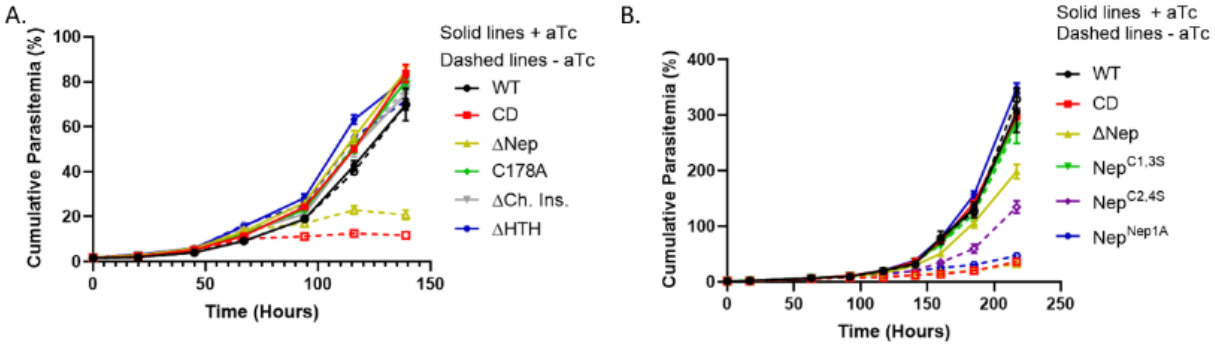
Supplemental Figure 1 – Alignment of the experimentally determined P. vivax PM V (PvPMV) structure (left, red) and the P. falciparum PM V (PfPMV) model predicted by Phyre2 (grey, center). At right, the two structures are aligned, showing their gross structural similarity. Peptidomimetic inhibitor WEHI-842 is shown in blue.



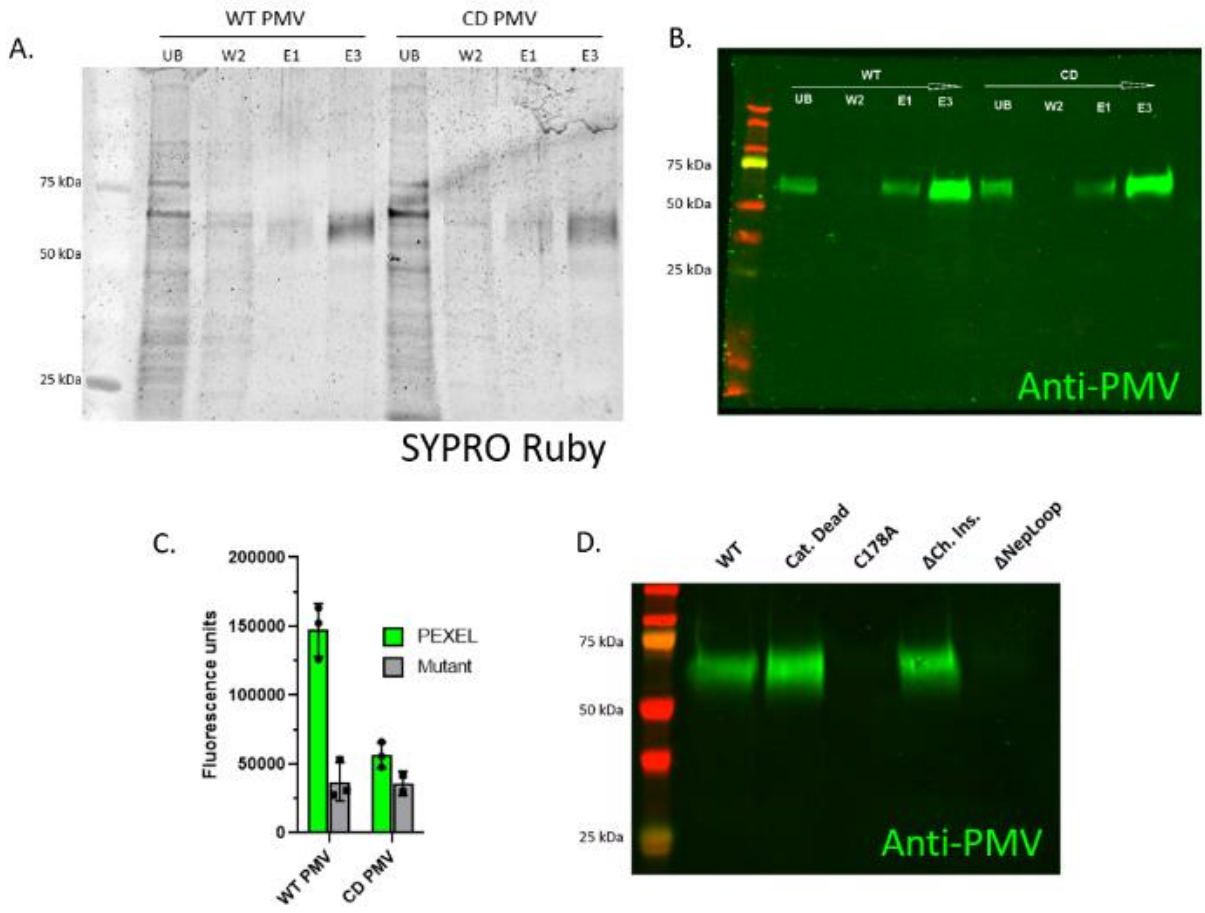
Supplemental Figure 2 – Sequence alignment of the reference sequences for P. falciparum, P. vivax, and P. berghei PM V, as well as P. falciparum plasmepsin I (PfPMI). Alignments were performed using Clustal Omega. Residues shared among the PM V sequences are bolded. Structures of interest are highlighted.



Supplemental Figure 3 – Western blot with anti-PMV (green) and anti-GFP (red) showing that mutants enzymes can be detected by monoclonal anti-PMV or anti-GFP, with the exception of the ΔCh. Ins. mutant, which is no longer recognized by anti-PMV.



Supplemental Figure 4 – Full growth curves for Fig. 2D and Fig. 3D. Knockdown was performed as described above and parasite growth measured daily by flow cytometry. Parasites were subcultured 1:2 any time parasitemias grew above 5%. Reads were then multiplied appropriately to calculate "cumulative parasitemias".



Supplemental Figure 5 – Inconsistent expression of PM V mutants in HEK293 cells (A) SYPRO Ruby-stained gel showing purification scheme for secreted WT and CD PM V from HEK293 conditioned media. Fractions shown are unbound (UB), a wash fraction (W2), and two elution fractions (E1 and E3). (B) Western blot of the same samples probed with anti-PM V. (C) Activity assay for WT and CD PM V showing recombinant PM V to be active against a PEXEL peptide but not the mutant peptide. Assay was performed as in Fig. 5B. (D) Western blot probed with anti-PM V.

Name (use)	**Sequence**
1 (attB into EOE)	GGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCTGGTTTGTCTGGTCAACCACCGCGGTCTCAGTGGTGTAC GGTACAAACCCGAATTCTGGTTTGTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCCGGAATTCTAGAT TTAATAAATATGTTCTTATATATAATG
2 (PM V FWD)	GTGCCACCTGACGTCGAGGTGGGTGGAGGGGTTTTATCCATGAGAGG
3 (PM V REV)	CTGCACCTGGcctaggTGTTGATTCCTGTATGGGAG
4 (PM V FWD HTH)	ACGATTTTTTctcgagATGAATAATTATTTTTTAAGGAAAGAAAATTTTTTTATATTG
5 (CD)	CCATCGCAAAGAATTTCTTTAATTCTAGcTACAGGTTCATCTTCGTTAAGTTTCCCGTG
6 (ΔNep)	AAACATCATCTATTTTATATGAATATCTTCAATCGTATTG
7 (C178A)	GTGAATATCTTCAATCGTATgctGAAGGGTCTCAAATATATGG
8 (ΔCh. Ins.)	GTTATGAACCGGATTACTTCGATATTGTGTGGCAAGCTAT
9 (ΔHTH)	CCAAATAAATTATTATTTAGATATTTTATGTATACATGATATGgccgccAATTTATGTATTAAAATAGTTGATGGAGTACAA TGTTGG
10 ($Nep^{C1,3S}$)	CAAAAACATCATCTATTTTATATaGTAATAAATCCAATTGTCCTTATGGTTTAAAAaGTGTAGGAAATAAATGTGAATATC
11 ($Nep^{C2,4S}$)	CATCTATTTTATATTGTAATAAATCCAATaGTCCTTATGGTTTAAAATGTGTAGGAAATAAAaGTGAATATCTTCAATCGT ATTGTG
12 (Nep^{Nep1A})	GAATTATTCAAAAACATCATCTATTTTATATTGCTCAAGCCAACTCTGTCAAGCCCTTTCAAGCCCGACATGCTCTAATA ATTTCTGCGAATATCTTCAATCGTATTGTGAAGGG
13 (Seq 1)	GGTATTCATATGGAAAAACCATATAACTTG
14 (Seq 2)	CACATATTCCAGAAAATATTTATAACC
15 (Seq 3)	GTGGAAAATAAAAATGACAATGTGGGAAATAAAAATGACAATG
16 (Seq 4)	CGTTAAGTTTCCCGTGTAATGGTTG
17 (PMV into pHL)	GCGTAGCTGAAACCGGTgtaGAGAATAAAATCGATAATGTTGG
18 (PMV into pHL)	GATGGTGGTGCTTGGTACCTGACGGGCACTTGC
19 (pHL CD)	CATCCCAGCGTATATCTTTGATCTTAGcTACCGGTAGCAGCTCTTTGTCTTTTCC
20 (pHL ΔNep)	CTACTCAAAAACAAGTTCTATTCTTTAtgAATACCTACAAAGTTATTGTGAAGGAAG
21 (pHL C178A)	GGAAATAAATGTGAATACCTACAAAGTTATgctGAAGGAAGTCAAATATATGGTTTC
22 (pHL ΔCh. Ins.)	CATCGGCGGTTATGAACCCGATTACTTcgATATTGTATGGCAAGCAATAACAAGG
23 (pHL ΔHTH)	GGATATTTTGTGCATACATGACATGgccgccAACTTGTGTATAAAAATAGTTGATGG
24 (pHL seq 1)	GACATAGGAAAACCATCCCAGCG
25 (pHL seq 2)	GGTAGCACTTTCACGCATATACCTG

Supplemental Table 1 – Primers used in this study.

Chapter 4: An essential ER-resident N-acetyltransferase in *Plasmodium falciparum*

4.1 Preface

This chapter has been deposited on BioRxiv as:

> Polino AJ, Floyd K, Avila-Cruz Y, Yang Y, Goldberg DE (2021). An essential ER-resident N-acetyltransferase in *Plasmodium falciparum*. *BioRxiv*.

The text makes reference to five supplemental figures and one supplemental table titled:

> Supp. Fig. 1: Southern blot
>
> Supp. Fig. 2: Uncut gel for Figure 1
>
> Supp. Fig. 3: Supplement to Figure 3B
>
> Supp. Fig. 4: DiCre excision of Pf3D7_1473000 locus also doesn't affect HRP2/HRP3 acetylation
>
> Supp. Table 1: Primers used in this study

These are not displayed inline with the main figures, but are reproduced at the end of this chapter.

4.2 Abstract

N-terminal acetylation is a common eukaryotic protein modification that involves the addition of an acetyl group to the N-terminus of a polypeptide. This modification is largely performed by cytosolic N-terminal acetyltransferases (NATs). Most associate with the ribosome, acetylating nascent polypeptides co-translationally. In the malaria parasite Plasmodium falciparum, exported effectors are translated into the ER, processed by the aspartic protease plasmepsin V and then N-acetylated, despite having no clear access to cytosolic NATs. Here, we used post-transcriptional knockdown to investigate the most obvious candidate, Pf3D7_1437000.

We found that it co-localizes with the ER-resident plasmepsin V and is required for parasite growth. However, depletion of Pf3D7_1437000 had no effect on protein export or acetylation of the exported proteins HRP2 and HRP3. Pf3D7_1437000-depleted parasites arrested later in their development cycle than export-blocked parasites, suggesting the protein's essential role is distinct from protein export.

4.3 Introduction

N-terminal acetylation is among the most common modifications to eukaryotic proteins (Deng and Marmorstein, 2020). Typically during or soon-after translation, an acetyl group is transferred from acetyl-CoA to the N-terminus of a polypeptide. This alteration blunts the N-terminal charge, changing its chemical properties, and altering the way the polypeptide interacts with various biological systems. In some cases, N-terminal acetylation changes a protein's half-life (Hwang et al., 2010; Park et al., 2015). In others, proper interaction with binding partners relies on acetylation (Scott et al., 2011). Acetylation is performed by N-acetyltransferases (NATs), nearly all of which are cytosolic enzymes that typically associate with the ribosome. In eukaryotes, eight currently known NAT complexes combine to acetylate most cytosolic N-termini, as well as some proteins in the chloroplast lumen.

The malaria parasite *Plasmodium falciparum* marks effectors for export into the host cell with a pentameric amino acid sequence called the *Pl*asmodium export element (PEXEL) (Hiller et al., 2004; Marti et al., 2004). PEXEL-containing proteins are translated into the parasite endoplasmic reticulum (ER), where PEXEL is cleaved by the aspartic protease plasmepsin V (PM V) (Boddey et al., 2010; Russo et al., 2010) after a conserved leucine. Previous work with exported reporters revealed that following PEXEL cleavage in the ER, the new N-terminus is somehow acetylated (Boddey et al., 2009; Chang et al., 2008; Osborne et al., 2010). This acetylation occurred even if exported proteins were sequestered in the ER by brefeldin A treatment (blocking anterograde traffic from the ER) or addition of an ER retention signal (Chang et al., 2008; Osborne et al., 2010). This suggests that a yet unidentified NAT exists in the *P. falciparum* ER. Subsequently Tarr, et al. (2013) mutagenized the PEXEL motif of the exported protein REX3 and found that point mutants that were not acetylated also were not

exported (Tarr et al., 2013), demonstrating a coincidence between these two processes. The identity of the PEXEL NAT, and its role in export, if any, remains to be determined.

Here, we searched the genome for putative NATs in *P. falciparum* that might access the secretory system. We identified one, Pf3D7_1437000, as a likely candidate for follow-up investigation. Depletion of Pf3D7_1437000 arrested parasite growth in culture, but had no detectable effect on protein export or on exported protein N-acetylation. The phenotype manifested by Pf3D7_1437000 depletion is distinct from that seen after disruption of an essential component of the export pathway, suggesting that the essential role of Pf3D7_1437000 is not in facilitating protein export.

4.4 Results

To search for an ER-resident NAT, we used PlasmoDB to identify *P. falciparum* genes whose sequence was annotated to contain a motif assigning them to the GNAT enzyme superfamily (Aurrecoechea et al., 2009). This search yielded eight genes. Two – Pf3D7_ 0823300 and Pf3D7_0629000 – have been previously studied in *P. falciparum* and are proposed to act as a GCN5 histone acetyltransferase and a glucosamine-phosphate N-acetyltransferase respectively (Cova et al., 2018; Fan et al., 2004). Five of the remaining six have orthologs in metazoans. Pf3D7_1020700 appears similar to the human NAT10, an RNA cytidine acetyltransferase (Ito et al., 2014). Pf3D7_1003300, Pf3D7_0109500, Pf3D7_0805400, and Pf3D7_1323300 are the closest orthologs of the cytosolic human N-acetyltransferases Naa10, Naa20, Naa30, and Naa40 respectively (Fig. 4.1A) (Chen et al., 2006). The remaining candidate NAT, Pf3D7_1437000, appears to have orthologs among the Apicomplexa, but no obvious representative outside the phylum. None of the candidate NATs has an apparent signal or retention sequence to drive ER localization. However, Pf3D7_1437000 has two predicted

transmembrane domains (Fig. 4.1B), making it the likeliest candidate to access the secretory system. With that in mind, we focused our efforts on characterizing Pf3D7_1437000.

A.

GeneID	Conserved Among	Ortholog	ER Targeting?
Pf3D7_1003300	Life	Naa10	-
Pf3D7_0109500	Eukaryotes	Naa20	-
Pf3D7_0805400	Eukaryotes	Naa30	-
Pf3D7_1323300	Eukaryotes	Naa40	-
Pf3D7_1437000	Apicomplexa	-	TM

B.

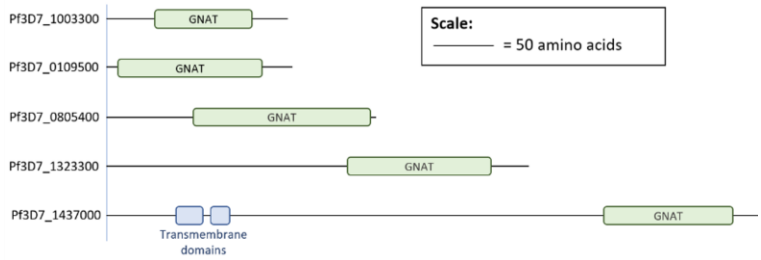


Figure 4.1 - Putative protein N-acetyltransferases encoded by P. falciparum: (A) Table of putative N-acetyltransferases (NATs) in the *P. falciparum* genome. Genes were included based on assignment to the GNAT enzyme superfamily in PlasmoDB, and subsequently excluded if they had a different predicted function (see text). Orthology was based on grouping in OrthoMCL, and supported by reciprocal BLAST searches. ER targeting sequences were sought with SignalP 4.1, and transmembrane domains (TM) annotated based on TMHMM predictions. (B) Diagram of the identifiable domains in the NAT candidates in (A). Diagram is to scale as shown.

We targeted Pf3D7_1437000 using CRISPR/Cas9 editing and the previously described pSN054 vector to replace the gene with a recodonized version that is C-terminally HA-tagged and flanked with loxP sites for gene excision, as well as TetR-binding aptamers for post-transcriptional depletion (Fig. 4.2A) (Polino et al., 2020). Proper genome editing was confirmed by Southern blot (Supp. Fig. 1). Western blotting for HA revealed a single band consistent with tagged full-length Pf3D7_1437000 (expected size 64 kDa) (Fig. 4.2B). Washing out aTc depleted Pf3D7_1437000 levels to approximately 40% of the +aTc levels within twelve hours (Fig. 4.2C). Washing out aTc in ring-stage parasites resulted in parasite death within a single intraerythrocytic development cycle, showing Pf3D7_1437000 to be essential for growth in parasite culture (Fig. 4.2D).

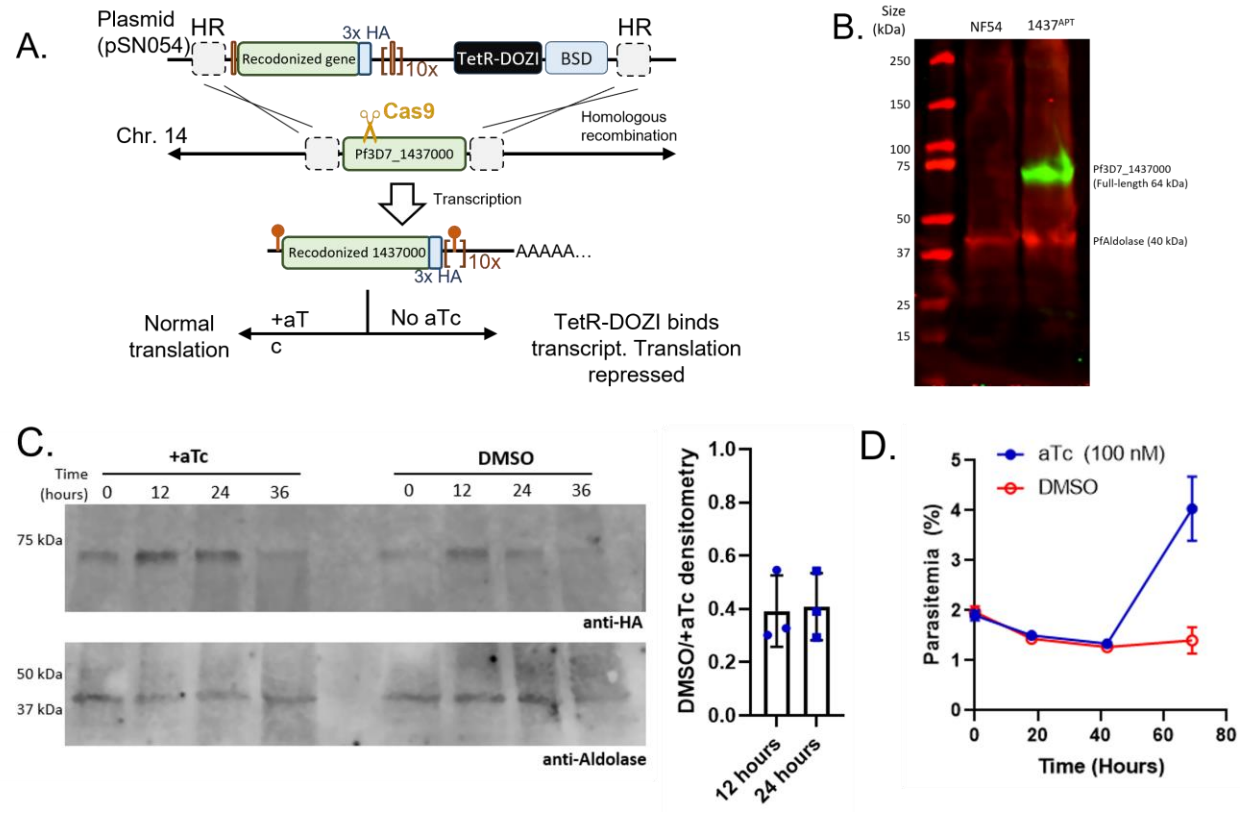


Figure 4.2 – TetR-Aptamer system for inducible Pf3D7_1437000 depletion: (A) Diagram of genome editing and inducible depletion of Pf3D7_1437000. The endogenous gene was disrupted by Cas9 and replaced by a recodonized version flanked with aptamers. When transcribed, the aptamers fold to bind TetR. In the absence of anhydrotetracycline (aTc), TetR binds the aptamers; its fusion protein DOZI sequesters the bound mRNA, repressing translation. (B) Western blot probing lysate from the parent (NF54) or tagged line (1437APT) with anti-HA (green) and anti-*P. falciparum* fructose-bisphosphate aldolase (PfAldolase, red). (C) Western blot demonstrating Pf3D7_1437000-HA levels over part of the intraerythrocytic development cycle, and the effect of washing out aTc ("DMSO" as aTc was replaced with an equal volume of DMSO) on Pf3D7_1437000-HA levels. At right, knockdown was quantified for three separate experiments. Bar height represents mean; error bars standard deviation. (D) aTc was washed from ring-stage parasites and their growth monitored daily by flow cytometry. The experiment was performed three times, with each culture in technical triplicate. A representative experiment is shown: points represent the mean of technical triplicates, error bars the standard deviation of those measurements. Uncut gel for panel C is shown in Supp. Fig. 2.

Pf3D7_1437000 localizes to the ER

We assessed the localization of Pf3D7_1437000-HA in fixed parasites by immunofluorescence. Staining with anti-HA antibodies suggests that Pf3D7_1437000-HA is found in a perinuclear ring, consistent with ER localization (Fig. 4.3A) (Klemba et al., 2004).

This is supported by co-staining with antibodies against organellar markers: anti-plasmepsin V for the ER (Banerjee et al., 2002; Klemba and Goldberg, 2005), anti-ACP for the apicoplast (Waller et al., 1998), and anti-aldolase for the cytosol. The staining pattern of anti-PM V largely resembles that of anti-HA, while anti-ACP and anti-aldolase clearly highlight patterns distinct from anti-HA. To test whether this localization is an artifact of cell fixation, we constructed a parasite line with Pf3D7_1437000 fused to the fluorescent protein mNeonGreen and 3xHA, and PM V fused to mRuby3 and 3xFLAG. Western blot showed most tagged protein to be at the expected size for full-length Pf3D7_1437000-mNeonGreen-3xHA and PM V-mRuby3-3xFLAG (Supp. Fig. 3). Live microscopy on trophozoites showed colocalization between Pf3D7_1437000-mNeonGreen and PM V-mRuby3 (Fig. 4.3B, C), again suggesting that Pf3D7_1437000 localizes to the parasite ER. Interestingly, we note that both by immunofluorescence and live fluorescence Pf3D7_1437000 and PM V appear to occupy the same space, but their intensities are not perfectly correlated throughout that space: PM V is present in a ring around the nucleus with a single protruding bleb. Pf3D7_1437000 is present over the same space, but disproportionately concentrated in the bleb. The significance of this distinction is not obvious to us.

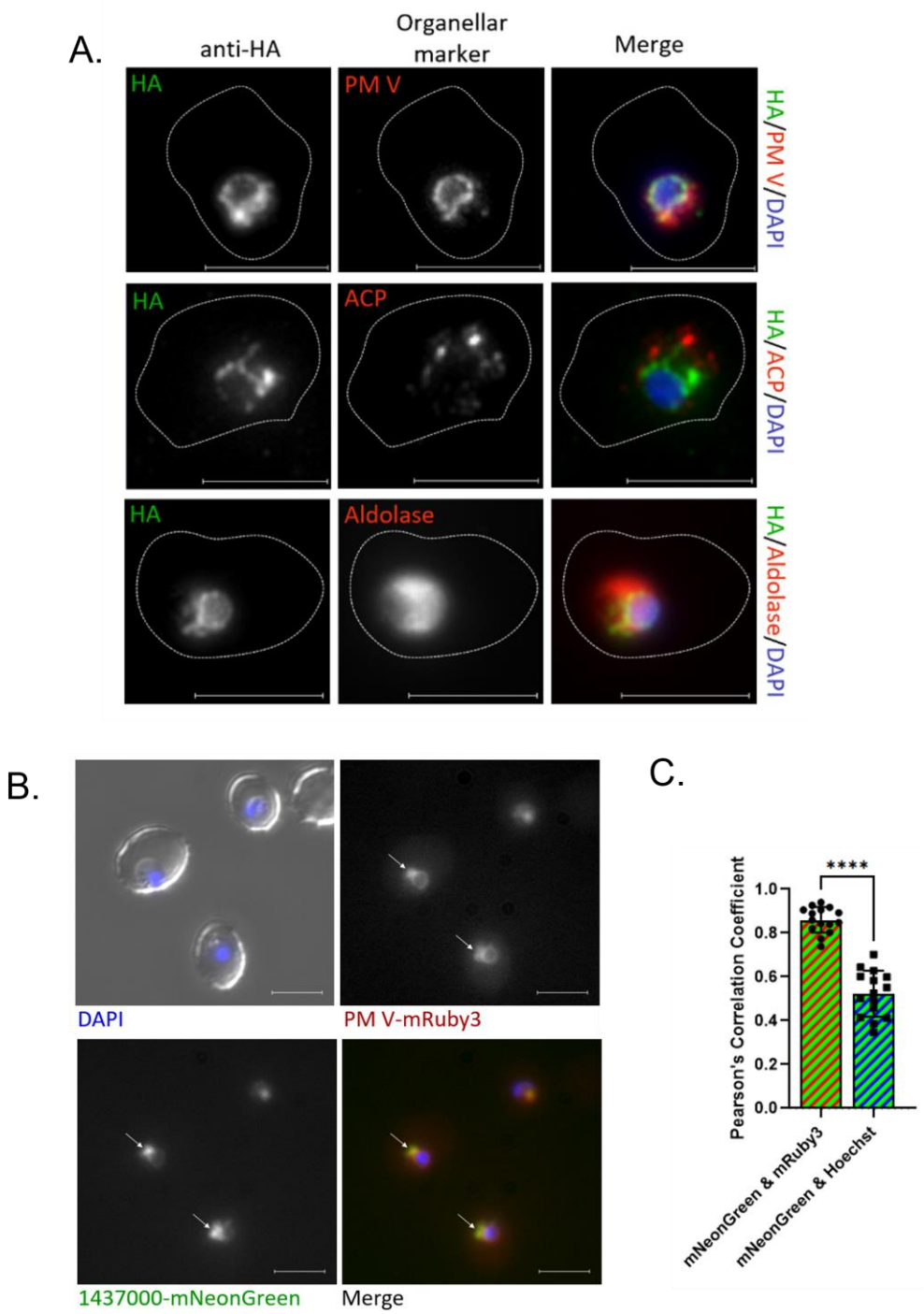


Figure 4.3 – Pf3D7_1437000 resides in the ER: (A) Micrographs of immunofluorescence assay to localize Pf3D7_1437000-HA (green in merge) alongside PM V, ACP, or Aldolase (red in merge). Experiment was performed three times; representative images are shown. (B) Live epifluorescence microscopy with Pf3D7_1437000-mNeonGreen (green) and PM V-mRuby3 (red). Scale bar represents 5 μm. White arrows are to give the reader spatial references between pictures. A representative image is displayed here. Additional zoomed out images are in Supp. Fig. 3. (C) Quantification of signal overlap from (B) by Pearson's Correlation Coefficient. N = 15 cells. Groups were compared by an unpaired t-test, with $p < 0.0001$, designated by "****".

Pf3D7_1437000 is not required for protein export

Since our original interest was in the acetylation of exported proteins, we next sought to assess whether Pf3D7_1437000 had a role in exporting proteins into the host red blood cell. We used a previously described line where the PTEX component Hsp101 is fused to a DHFR destabilization domain (abbreviated Hsp101-DD), and its function requires the stabilizing ligand trimethoprim (TMP). When TMP is removed, Hsp101 is destabilized and exported proteins accumulate in the parasitophorous vacuole (Beck et al., 2014). Alongside this line, we depleted Pf3D7_1437000 by washing out aTc from young ring-stage parasites (0-4 hours old), then fixed trophozoites and stained for exported proteins by immunofluorescence assay (Fig. 4.4A). We found that while Hsp101 destabilization blocked protein export as previously described, depletion of Pf3D7_1437000 had no discernible effect on export of the PEXEL proteins HRP II or FIKK4.2 (Fig. 4.4A, B).

To get an immunofluorescence-independent view of export competence we assessed establishment of the exported protein-dependent nutrient import channel on the infected erythrocyte surface (Beck et al., 2014; Ginsburg et al., 1985). To do so, we depleted Pf3D7_1437000 or disrupted Hsp101, pelleted trophozoites, and resuspended them in 5% sorbitol. Export-competent parasites have increased solute uptake into the host RBC, leaving them susceptible to osmotic lysis in 5% sorbitol. Parasites that cannot export proteins are not able to increase solute uptake, and are protected from sorbitol lysis (Beck et al., 2014; Ginsburg et al., 1985; Kirk et al., 1994). We monitored lysis by measuring the release of hemoglobin into the supernatant over 25 minutes. Disrupting Hsp101 protected parasites from lysis as previously described (Beck et al., 2014) but depleting Pf3D7_1437000 had no effect on sorbitol sensitivity (Fig. 4.4C), again suggesting Pf3D7_1437000 is not involved in protein export.

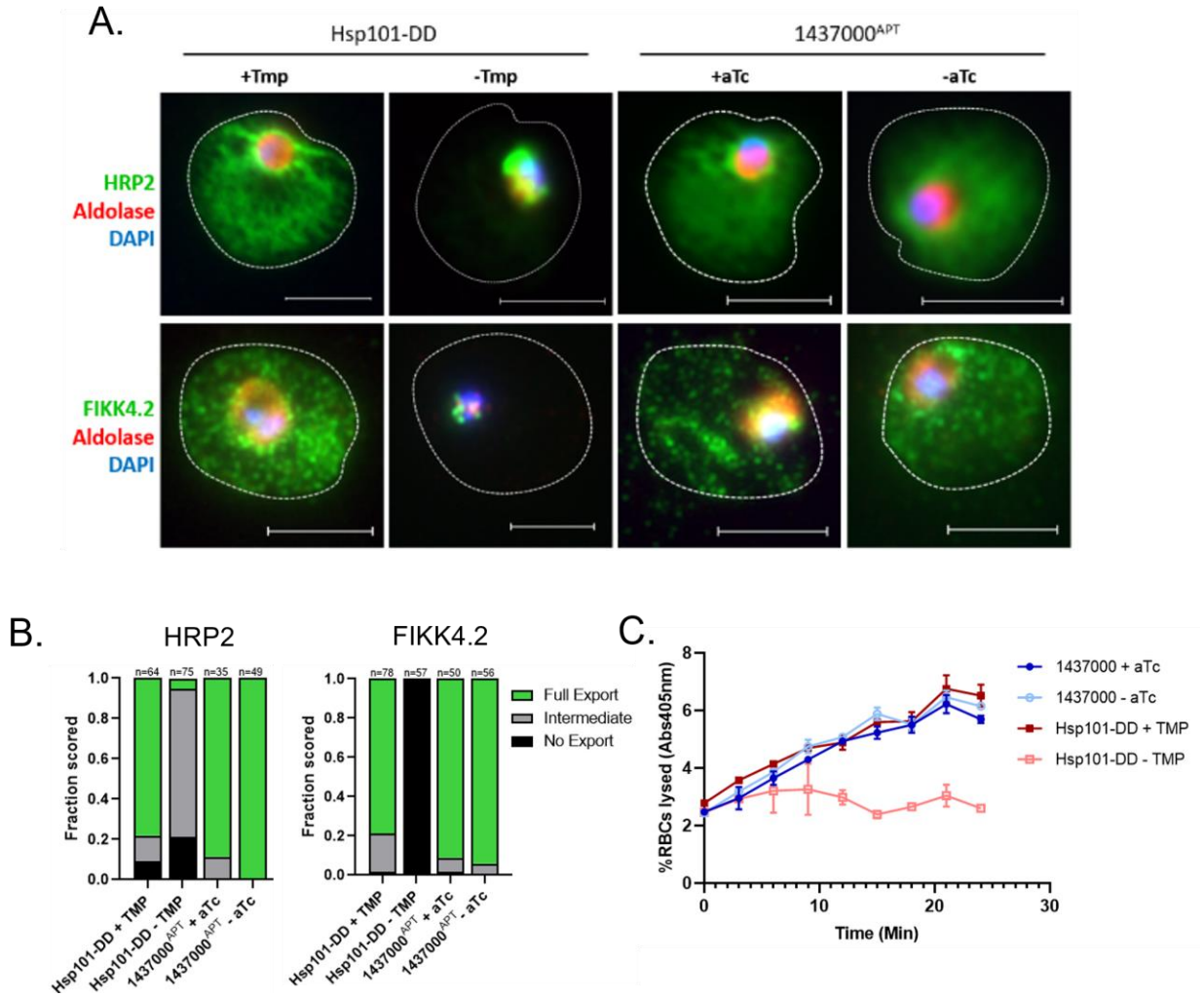


Figure 4.4 – Pf3D7-1437000 depletion does not affect protein export: (A) Trophozoites were fixed and stained to compare protein export in Hsp101-DD +/- TMP and 1437APT +/- aTc. Exported proteins are marked by anti-HRP2 or anti-FIKK4.2, each shown in green; the parasite cytosol is marked by anti-aldolase shown in red; parasite nucleus marked by DAPI in blue. Experiment was performed three times; representative images shown. Scale bars represent 5 μm. (B) Quantification of ten fields by a blinded scorer for the experiment shown in (A). Number of cells counted is shown above each bar. (C) Sorbitol sensitivity was monitored by measuring hemoglobin release into the supernatant (via absorbance at 405 nm) over time. Experiment was performed three times. A representative experiment is shown, with each sample in technical triplicate. Data points represent the measured mean; error bards the standard deviation.

To determine whether Pf3D7_1437000 is responsible for acetylating PEXEL proteins, we isolated two abundant exported parasite proteins, histidine-rich proteins II and III (HRP2 and HRP3) (Fig. 4.5A) and measured the mass of each intact protein by liquid chromatography/ mass

spectrometry. In the presence of aTc, we detected substantial peaks consistent with acetylated HRP2 and acetylated HRP3, as expected (Fig. 4.5B). When we depleted Pf3D7_1437000, we were surprised to find no change in the mass of the peaks (Fig. 4.5B). In fact, the chromatogram in both cases revealed no peak at the mass expected for unacetylated HRP2 or HRP3.

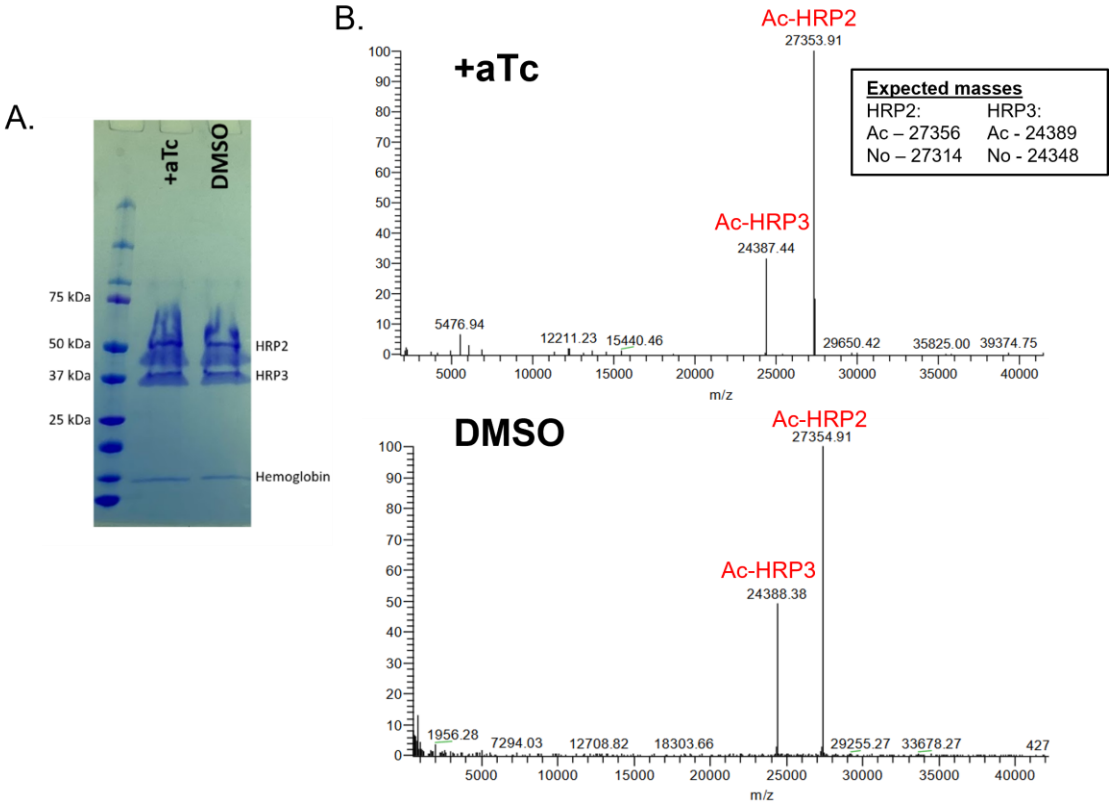


Figure 4.5 – Pf3D7_1437000 depletion does not affect HRP2/HRP3 acetylation: (A) Coomassie-stained SDS-PAGE gel showing HRP2 and HRP3 isolated from the supernatants of saponin-released parasites on nickel resin. Both migrate through the gel less than one would anticipate from their linear size due to their extreme positive charge. (B) Deconvoluted mass spectra from analysis of intact HRP2 and HRP3. Inset shows anticipated sizes for acetylated and un-acetylated HRP2 and HRP3 after PEXEL cleavage. This experiment was performed twice; a representative experiment is shown.

Concerned that this could be due to the slow nature of the aptamer-driven knockdown, we remade the aptamer construct in a parasite line that constitutively expresses a dimerizable Cre recombinase (DiCre) that is activated by the ligand rapamycin (Beck, Manuscript in Preparation;

Jullien et al., 2003). Since our aptamer construct already resulted in a gene flanked by loxP sites (Fig. 4.2A) this allowed us to inducibly excise the Pf3D7_1437000 gene from the genome. Addition of 50 nM rapamycin to the culture media of synchronized ring-stage parasites resulted in complete excision of Pf3D7_1437000 within 24 hours (Supp. Fig. 4A) and parasite death within a single replicative cycle (Supp. Fig. 4B). As before, we purified HRP2 and HRP3 from parasite culture and again found the mass of each unchanged by the excision of Pf3D7_1437000 from the genome (Supp. Fig. 4C). This suggests that Pf3D7_1437000 is not the NAT that acetylates PEXEL proteins.

Pf3D7_1437000 is involved in parasite growth and entry into schizogony

We next turned our attention to the role of Pf3D7_1437000 in intraerythrocytic development. Normally, parasites are classified as ring forms for the first 20-24 hours after invasion, followed by maturation to fuller trophozoite forms and finally to schizonts that have replicated their DNA and formed daughter merozoites, the invasive forms that will invade new RBCs. We synchronized parasites within a three-hour window, washed out aTc in young rings, and monitored parasite development by thin smear. By 24 hours after invasion, Pf3D7_1437000-depleted parasites appeared smaller than their non-depleted counterparts (Fig. 4.6A, mean area +aTc = 6.95 μm2 vs. DMSO 5.45 μm2, unpaired t-test gives $p < 0.0001$). The disparity in size persisted throughout the intraerythrocytic development cycle, with Pf3D7_1437000-depleted parasites approximately half the size of non-depleted parasites by 36 hours (mean area = 22 μm2 vs. 12 μm2) (Fig. 4.6A). By 48 hours after invasion, non-depleted parasites had formed schizonts filled with clearly distinguishable daughter cells. Pf3D7_1437000-depleted parasites apparently remained trophozoites: some mature but without distinguishable daughter cells, others shrunken (Fig. 4.6B). Correspondingly Pf3D7_1437000-depleted parasites remained much smaller than

non-depleted parasites (mean = 18 μm2 vs. 30 μm2) (Fig. 4.6A). Thus, we conclude that Pf3D7_1437000 is important for parasite growth, and for completion of proper schizogony.

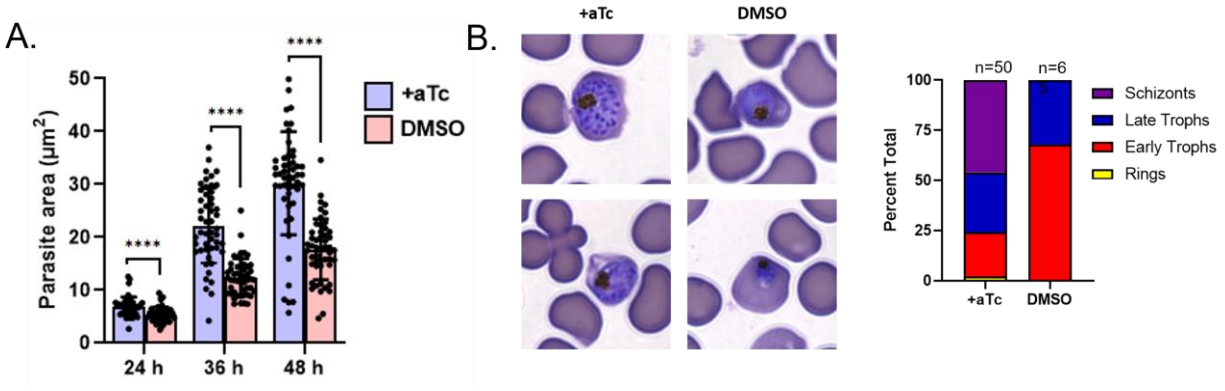


Figure 4.6 – Pf3D7_1437000 depletion results in growth arrest before schizogony: (A) Area of parasites +/- aTc examined by thin smear at indicated times post-infection. Bar height indicates mean of 50 measured parasites, error bars indicate standard deviation. For each time point the measured areas were compared by an unpaired t-test, which yielded p-values < 0.0001 (marked with "****"). (B) At left, demonstrative images showing phenotypes described in the text. +aTc parasites were mostly schizonts and late trophozoites; parasites without aTc arrested before schizogony. At right, quantification of parasite life stages of 10 microscopic fields (n = 50 or 65 parasites, shown atop bar). Experiment was performed twice. A representative experiment is shown.

The Pf3D7_1437000 depletion phenotype appeared to us quite distinct from death due to protein export block, which arrests growth at the transition from rings to trophozoites (Beck et al., 2014). To distinguish these phenotypes more clearly, we fixed Pf3D7_1437000-regulated parasites alongside Hsp101-DD parasites at several time points during and assessed DNA content with the dye SYBR Green I by flow cytometry (Fig. 4.7). By 28 hours after invasion, Hsp101-disrupted parasites lag behind non-disrupted parasites, while Pf3D7_1437000-depleted parasites are indistinguishable from their non-depleted partners. By 40 hours after invasion Hsp101-disrupted parasites have the same DNA content that they did 20 hours earlier. Pf3D7_1437000-depleted parasites too lag behind their non-depleted counterparts, but with a substantially larger DNA content. Thus, it appears that growth arrest caused by Pf3D7_1437000

depletion is distinct from and substantially later than arrest caused by export disruption. Disrupting export prevents most of S-phase, while Pf3D7_1437000 depleted parasites continue until almost the final rounds of DNA replication.

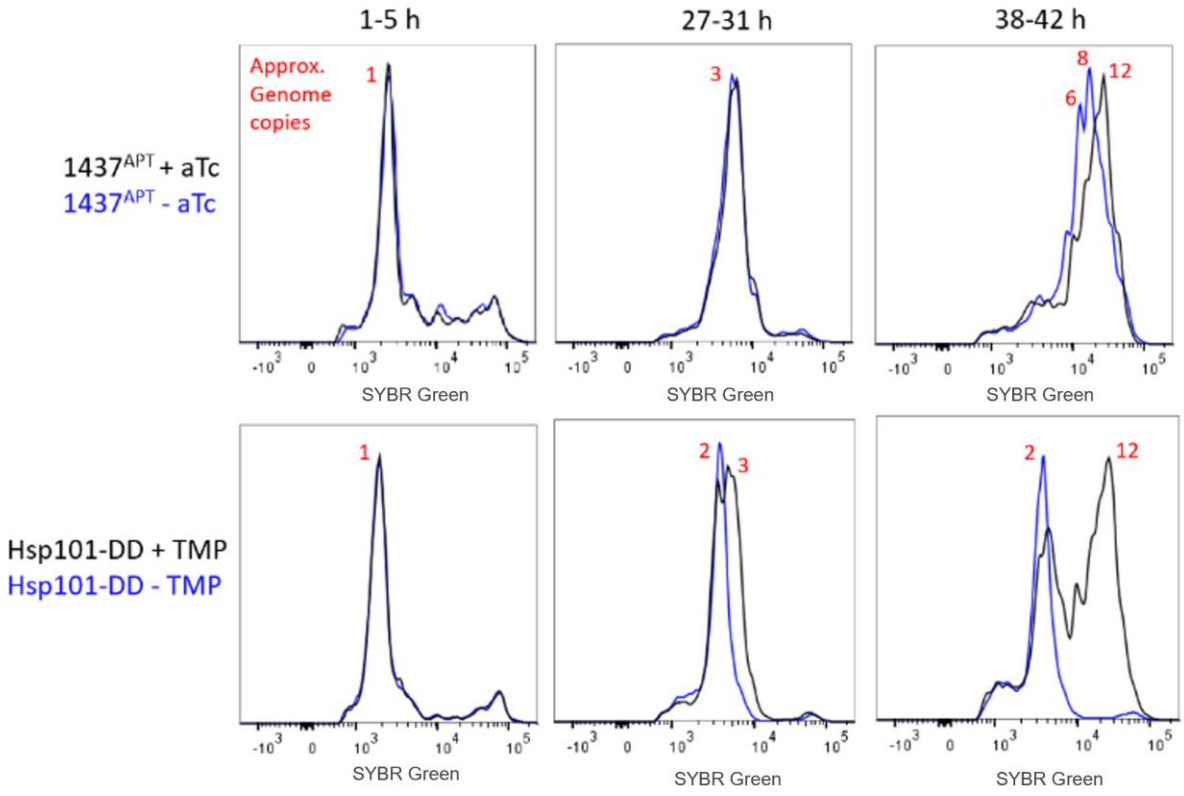


Figure 4.7 – Pf3D7_1437000 depleted parasites do not complete DNA replication: Synchronized parasites were fixed at the indicated times, and their DNA content measured with SYBR Green I and flow cytometry. Samples were run in technical triplicate (N=100,000 cells per sample). Uninfected RBCs were gated out by their low SYBR Green I fluorescence. Histograms are shown, with SYBR Green I fluorescence along the x-axis, and frequency on the y-axis. Approximate genome copy numbers were estimated from each peak's fluorescence intensity, assuming most parasites at 1 to 5 hours after invasion have a single copy of the genome. This experiment was performed twice; a representative experiment is shown.

4.5 Discussion

Here we sought to investigate the acetylation of PEXEL proteins, instead finding a putative NAT that is essential for parasite growth but appears to be uninvolved in protein export

and likely not the PEXEL NAT. Pf3D7_1437000 seemed a likely candidate (and really the only plausible candidate) to accomplish this post-PM V-cleavage acetylation. Indeed there is explicit speculation in the literature that it may serve such a role (Boddey and Cowman, 2013; Osborne et al., 2010). Supporting that is its putative ER localization: it appears in Marapana, et al.'s extensive immunoprecipitation/mass spectrometry results when pulling down SPC21 (signal peptidase) and SPC25 (Marapana et al., 2018). It has a clear *T. gondii* ortholog predicted to reside in the *Toxoplasma* ER (Barylyuk et al., 2020). Our findings are consistent with this ER localization, yet argue against a model where Pf3D7_1437000 acetylates PEXEL N-termini following PM V cleavage. The *Toxoplasma* PMV ortholog, ASP5, resides in and processes proteins in the Golgi, so N-acetylation after protein processing in this organism is not likely to involve the same NAT ortholog.

Our data most parsimoniously support a model where Pf3D7_1437000 is not the PEXEL NAT; however, our data cannot yet exclude alternative models. First, PF3D7_1437000 could be so efficient at acetylating its substrates that the small amount of it remaining after knockdown could be capable of acetylating all HRP2 and HRP3 N-termini that we could detect. We tried to minimize this possibility by harvesting schizonts for our mass spectrometry experiments, giving any acetylation phenotype as much time as possible to manifest. At the time of harvest, Pf3D7_1437000 knockdown parasites have been physically affected by the knockdown (as measured by parasite size) for at least 20 hours, suggesting that the remaining enzyme is not able to accomplish its full function for much of the life cycle. Still, the possibility remains that Pf3D7_1437000 is acetylating HRP2 and HRP3 in our assay, and that parasite growth problems are caused by a distinct second function of this protein that is more sensitive to knockdown. Second, it remains possible that Pf3D7_1437000 depletion affects HRP2/HRP3 acetylation, but

we failed to detect the unacetylated forms due to their poor stability in the parasites, solubility in our purification system, or ionization for mass spectrometry. The only data we have to consider in this light is that Tarr, et al. has described mutants of the reporter fusion $REX3^{1-61}$-GFP that are unacetylated (and not exported) and appear in similar abundance to the wild-type protein (Tarr et al., 2013). That is, acetylation of the PEXEL N-terminus is not universally required for protein stability and detection; however, we cannot exclude the possibility that such problems have frustrated our analysis here.

Regardless of whether Pf3D7_1437000 is the PEXEL NAT, we report here that its depletion does not affect protein export, and that – while Pf3D7_1437000 is essential for parasite development – the phenotype of its depletion differs from described disruptions of export machinery. Depletions/disruptions of the major PTEX components Hsp101 and PTEX150 have been described, each of which arrested parasite growth as early trophozoites (Beck et al., 2014; Elsworth et al., 2014). Chemical inhibition or depletion of PM V can arrest parasite growth immediately after invasion, or as early trophozoites (Boonyalai et al., 2018; Polino et al., 2020; Sleebs et al., 2014). Here we show that Pf3D7_1437000 depletion arrests parasite growth substantially later than we observe when blocking export via Hsp101 disruption. Our finding that Pf3D7_1437000 depletion affects parasite size through much of intraerythrocytic development cycle is curious, but its causes could be manifold, and further study of Pf3D7_1437000-depleted parasites would be needed to elucidate how this putative enzyme could be influencing cell size.

Among our field's directives is to identify and characterize essential parasite proteins and pathways sufficiently diverged from their hosts' orthologues as to be specifically targeted with chemotherapeutics. In some ways, Pf3D7_1437000 fits that bill: an essential putative NAT with orthologues throughout Apicomplexa, yet without a clear orthologue in mammals. However,

hurdles remain before its druggability can be assessed. First, the degree to which this enzyme is essential in other pathogenic apicomplexans is not yet known. The *Plasmodium berghei* knockout screen scored the *P. berghei* orthologue PbANKA_0611800 as "essential" albeit with low confidence (Bushell et al., 2017), while the genome-wide CRISPR screen in *Toxoplasma gondii* assigned the likely orthologue TgME49_305450 to have an intermediate phenotype score, higher than most validated essential genes, and lower than most validated dispensable genes (Sidik et al., 2016). Delineating how well-conserved the function and essentiality of Pf3D7_1437000 is remains for future work. Perhaps more importantly – assuming Pf3D7_1437000's essential function is as an NAT – chemical inhibition of NATs has only been described with acetyl-CoA-peptide conjugates (Foyn et al., 2013). The development of more druglike and cell permeant NAT inhibitors might raise the profile of this enzyme class for additional chemotherapeutic development.

Our data leave us with the question: if Pf3D7_1437000 is not the PEXEL NAT, what is? The simplest possibility is that another NAT resides in the parasite ER, with the most obvious candidates being the additional members of the GNAT superfamily listed in Figure 4.1. Each has an ortholog with an alternative described function in yeast and mammals, but perhaps they serve a distinct role in Apicomplexa. An alternative is that the genome continues to hide an as-yet cryptic NAT whose sequence defies our attempts to computationally predict its function. Unfortunately no obvious candidate jumps out from the Marapana, et al. immunoprecipitations of various ER proteins involved in PEXEL processing (Marapana et al., 2018). A third possibility is that no protein NAT truly resides in the parasite ER, but instead that following PEXEL cleavage by PM V, the new N-terminus somehow accesses the cytosol and is acetylated by the regular cadre of ribosome-associated NATs. Tarr, et al. used a split-GFP setup to show

that the active site of PM V is oriented into the ER lumen (Tarr and Osborne, 2015), and we have since presumed that the post-cleavage PEXEL N-terminus is limited to the ER lumen; however, our understanding of the dynamics of secretory import and traffic in *Plasmodium* is itself limited, and perhaps peptides can somehow transiently access cytosolic NATs.

Lastly, the determinants of exported protein trafficking in *P. falciparum* remain unclear. Here, we were unable to assess the role of PEXEL acetylation on export competence. However, our mass spectrometry approach measured the mass of intact HRP2 and HRP3 and found each to be the exact mass expected from the polypeptide backbone and N-terminal acetylation alone, suggesting that export competent proteins need no further post-translational modification beyond what has already been described.

Taken together, our data provide new insight into the processing of exported proteins – if largely by excluding the primary candidate for involvement in the process. Additionally we provide initial characterization of the putative NAT Pf3D7_1437000, whose essential function in schizogony remains unclear. We hope that future work will uncover the identity and role of the PEXEL NAT, and also clarify the role of Pf3D7_1437000 in the *Plasmodium* life cycle.

4.6 Closing material

Acknowledgement:

Our mass spectrometry approach was formed in consultation with and performed by the Donald Danforth Plant Science Center Proteomics & Mass Spectrometry Facility, particularly Brad Evans and Shin-Cheng Tzeng. Additionally we thank Geoffrey McFadden for the anti-ACP antibody, Dianne Taylor for anti-HRP2, John Adams for anti-ERD2 (through the Malaria Research and Reference Reagent Resource Center), Odile Mercereau-Puijalon for anti-FIKK4.2 (through the European Malaria Reagent Repository), and Joshua Beck for sharing the NF54^{attB}-

DiCre parasite line, pM2GT-mNeonGreen-3xHA plasmid, and pM2GT-mRuby3-3xFLAG plasmids pre-publication.

Funding:

This work was supported by an American Heart Association Predoctoral Fellowship (#18PRE33960417, awarded to A.P.) and by the National Institute of Allergy and Infectious Diseases (RO1 AI047798, awarded to D.G.). The Donald Danforth Plant Science Center Proteomics and Mass Spectrometry Facility acknowledges the support of the National Science Foundation (#DBI-0922879) for acquisition of the LTQ-Velos Pro Orbitrap LC-MS/MS.

Accession Numbers:

For all Plasmodium proteins referenced in this study, the PlasmoDB accession numbers are below:

Putative NATs**:** Pf3D7_0109500, Pf3D7_0629000, Pf3D7_0805400, Pf3D7_0823300, Pf3D7_1003300, Pf3D7_1020700, Pf3D7_1323300, and Pf3D7_1437000. Organellar markers**:** plasmepsin II (PF3D7_1408000), plasmepsin V (PF3D7_1323500), ERD2 (PF3D7_1353600), ACP (PF3D7_0208500), and aldolase (PF3D7_1444800). Export-related proteins**:** Hsp101 (PF3D7_1116800) HRP2 (Pf3D7_0831800), FIKk4.2 (PF3D7_0424700), HRP3 (PF3D7_1372200). Discussion**:** SPC21 (PF3D7_1331300), SPC25 (PF3D7_0320700), REX3 (PF3D7_0936300), PTEX150 (PF3D7_1436300).

Author contributions:

AJP conceptualized this work, performed experiments, and wrote the manuscript under the supervision of DEG. YAC, KF, and YY performed experiments and influenced the path of this project by sharing their ideas.

4.7 Materials & Methods:

Generation of plasmids:

The construct for regulating Pf3D7_1437000 levels was made using the previously described pSN054 vector (Nasamu et al., 2021; Polino et al., 2020). Primer sequences are listed in Supplementary Table 1; shorthand names will be used here. Homologous sequences for genome repair were amplified from NF54^{attB} (Nkrumah et al., 2006) with primers 14APT-1/14APT-2 for the left/upstream homologous region, and 14APT-3/14APT-4 for the right/downstream homologous region. The resulting PCR products were inserted into pSN054 cut with FseI and I-SceI respectively via Gibson Assembly (New England Biolabs). The recodonized Pf3D7_1437000 gene was synthesized by GENEWIZ (South Plainfield, NJ) and inserted into pSN054 containing the above homologous regions. Plasmid was cut with AsiSI and the recodonized gene inserted via Gibson Assembly, to make the final vector called pSN054-1437000-3xHA. The vector was transformed into BigEasy-TSA Electrocompetent Cells (Lucigen) for propagation. When amplifying vector for harvest, 0.01% weight/volume arabinose was added to stimulate plasmid replication.

The two endogenous tagging vectors used for live microscopy: pM2GT-1437000-mNeonGreen-3xHA (yDHOD) and pM2GT-PMV-mRuby3-3xFLAG (hDHFR) are derived from pM2GT-EXP2-mNeonGreen (yDHOD) (Glushakova et al., 2017) and pM2GT-Hsp101-3xFLAG (Garten et al., 2018; Ho et al., 2018). For the former, mNeonGreen-3xHA was amplified (and the HA added) by primers NG-HA-F/NG-HA-R. The amplicon was inserted into pM2GT-EXP2-mNeonGreen cut with AvrII/EagI via In-Fusion Cloning. For tagging Pf3D7_1437000, homologous sequences were amplified from the NF54^{attB} genome with primers 14NG-1/14NG-2 for the right homologous region (in the 3’ UTR), and 14NG-3/14NG-4 for the left homologous region. The two PCR products were combined in an overlap-extension PCR with the right homologous region forward and left homologous region reverse primers, and the resulting PCR

product gel extracted and inserted into the XhoI/NheI-cut plasmid via In-Fusion Cloning (Takara) to make the donor vector pM2GT-1437000-mNeonGreen-3xHA, which was then transformed into XL10-Gold Ultracompetent Cells (Stratagene) for propagation.

Synthesis of the PM V tagging vector was analogous. First mRuby3 was amplified with primers Rub-H-F/Rub-H-R and added to pM2GT-Hsp101-3xFLAG (hDHFR) (Garten et al., 2018; Ho et al., 2018) cut with AvrII using In-Fusion Cloning. To adapt the plasmid to PM V tagging we used primers PMVR-1/PMVR-2 and PMVR-3/PMVR-4 to amplify the right and left homologous regions respectively. These PCR products were inserted into the Xho/NheI-cut plasmid in a single pot reaction with In-Fusion Cloning, and the resulting vector transformed into XL10-Gold cells for propagation.

CRIPSR/Cas9 targeting plasmids for each were made in the previously described pAIO3 plasmid (Nessel et al., 2020). Primers 14G-1, 14G-2, 14G-3, 14G-4, PMVG-1, and PMVG-2 were each ordered along with their reverse complement sequences, annealed in the thermal cycler, and then inserted into AvrII-cut pAIO3 by In-Fusion Cloning. Completed vectors were transformed into XL10 Gold cells for propagation.

Parasite culture:

For all experiments described here we cultured *P. falciparum* in PRMI 1640 (Gibco) supplemented with 0.25% (w/v) Albumax I, 15 mg/L hypoxanthine, 110 mg/L sodium pyruvate, 1.19 g/L HEPES, 2.52 g/L sodium bicarbonate, 2 g/L glucose and 10 mg/L gentamycin with human RBCs added to 2% hematocrit. Parasite cultures were maintained in sealed chambers under a gas mixture consisting of 5% O_2, 5% CO_2, and 90% N_2 at 37°C. Deidentified RBCs were obtained from Barnes-Jewish Hospital blood bank (St. Louis, MO), St. Louis Children's Hospital blood bank (St. Louis, MO), and the American National Red Cross.

Generation of parasite lines:

All genetic modifications described here were done in the parasite line $NF54^{AttB}$ (Nkrumah et al., 2006) referred to as "NF54" throughout. For each transfection, donor vectors were linearized if necessary (pSN054 is already linear) and 50 μg each of donor vector and pAIO3 with relevant guide were combined, ethanol precipitated to ensure sterility, and dissolved in 100 μL sterile water. At time of transfection, the dissolved DNA was brought up to 400 μL in cytomix (120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 10 mM K_2HPO_4, and 25 mM HEPES adjusted to pH 7.6 with KOH; plasmid is solubilized more effectively in water than cytomix, so we typically allow DNA to dissolve in 100 μL water, then when dissolved add 100 μL 2x cytomix and 200 μL 1x cytomix to bring it up to transfection volume) and transfected into ~5% young ring-stage parasites with a BioRad Gene Pulser II (Settings: 0.31 kV, 0.950 μF, capacitance set to "High Cap", resistance on the Pulse Controller II set to "Infinite"). Successful transfectants were selected with the relevant drug beginning 24 hours after transfection: Blasticidin S (2.5 μg/mL; Fisher) for the aptamer line, DSM-1 (2 μM; Asinex) and WR-99210 (5 nM; gift from D. Jacobus of Jacobus Pharmaceutical Co.) for the mNeonGreen- and mRuby3-tagged line. Medium was changed daily for the week following transfection, then thrice weekly until parasites could be visualized by thin smear, typically two to four weeks after transfection.

Validation of lines:

Proper integration of the pSN054-1437000-3xHA vector was confirmed by Southern blotting. Genomic DNA from the parent and edited lines was isolated (QIAamp DNA Blood Miniprep Kit) and 10 μg each digested with HinDIII, separated overnight on a 0.7% agarose gel, and transferred to nylon (Nytran SuPerCharge TurboBlotter, 0.45 μm, GE) overnight. The blot was then probed with the left homologous region (PCR product of 14APT-1/14APT-2) labelled

with alkaline phosphatase (Amersham AlkPhos Direct Labelling Kit; GE) in hybridization buffer (Amerhsam) at 55°C overnight, washed twice each in primary wash buffer (120 g/L urea, 1 g/L SDS, 100 mL/L 0.5 M sodium phosphate pH 7, 8.7 g/L NaCl, 2 g/L Amersham blocking reagent) and secondary wash buffer (6.05 g/L Tris base, 5.6 g/L NaCl, 2 mL/L 1M $MgCl_2$, pH 10), then detected with Amersham CDP-Star Detection Reagent (GE) and exposed to blue autoradiography film (MidSci, BX810) overnight.

Additional validation was by western blot done exactly as described in (Polino et al., 2020) (section "Validation of PMV^{APT} line").

Assessment of Pf3D7_1437000 depletion:

In all experiments shown, infection with the aTc-regulatable line (1437^{APT}) was synchronized by purifying schizonts on LD columns (Miltenyi Biotech), eluting into fresh blood and media, lacking aTc and allowing parasites to invade for 3-4 hours. Invasion was halted by replacing media with 5% sorbitol, lysing any un-egressed schizonts. We then washed in fresh media one additional time for five minutes, to ensure aTc was removed from the culture. These synchronized parasites were cultured in either the presence of 100 nM aTc ("+aTc" throughout) or an equal volume of DMSO ("DMSO" throughout). DiCre excision experiments were performed as above, but parasites were cultured in the presence of 100 nM aTc and either 50 nM rapamycin or an equal volume of DMSO.

Western blotting:

We performed western blotting as in (Polino et al., 2020) with primary antibodies mouse anti-HA diluted 1:1000 (clone 16B12; Biolegend #901501), rabbit anti-PfAldolase diluted 1:2000 (abcam #ab207494, targets the protein with PlasmoDB accession PF3D7_1444800), and mouse anti-FLAG diluted 1:500 (Sigma #F1804), followed by secondary antibodies goat anti-

mouse IRDye 800CW (Licor) and donkey anti-rabbit IRDye 680RD (Licor) both diluted 1:10,000. For Figure 2, panel C parasites were harvested at the indicated times and Pf3D7_1437000-3xHA levels quantified using ImageStudio Lite v. 5.2 (Licor). The sizes of bands were approximated using the Precision Plus Protein Dual Color Standards (Bio-Rad; #1610374).

Flow cytometery:

To assess the effect of Pf3D7_1437000 depletion on parasite growth parasites were maintained in technical triplicate (3 x 1mL culture) and their growth monitored daily by flow cytometry (BD FACSCanto with attached High Throughput Sampler) by diluting culture 1:20 into PBS with 0.8 μg/mL acridine orange (Molecular Probes).

We assessed parasite progress through the DNA replication cycle as in (Perrin et al., 2021). Briefly, parasites were synchronized as above, then fixed at the indicated times (see Figure 7) by doubling their volume in 2x PBS + 0.4% glutaraldehyde (final concentration, 1x PBS + 0.2% glutaraldehyde) and stored at 4°C until all time points collected. Then DNA was stained using SYBR Green I (ThermoFisher), measured on a BD FACSCanto, and analyzed with FloJo v. 10.7.1.

Microscopy:

For localization of Pf3D7_1437000, parasites were synchronized as above to within a 4-hour window, then harvested 24 hours after invasion ended (i.e. parasites 24-28 hours old), washed once in PBS, and prepared for immunofluorescence imaging as recommended in (Tonkin et al., 2004) with the modification that cells were settled onto concanavalin A-coated coverslips (0.5 mg/mL) for 10 minutes prior to fixation, and that the wash following primary and secondary antibody incubation consisted of five PBS washes for three minutes each. Primary antibodies:

mouse anti-HA (close 16B12, Biolegend #901501) diluted 1:100, rabbit anti-HA (Sigma-Aldrich, H6908) diluted 1:100, mouse anti-PMV (Banerjee et al., 2002) diluted 1:50, rabbit anti-ACP (Waller et al., 1998) diluted 1:100, and rabbit anti-aldolase (abcam #ab207494) diluted 1:500. Secondary antibodies: goat anti-mouse IgG-AlexaFluor488, goat anti-rabbit IgG-AlexaFluor555 (both from Invitrogen) diluted 1:2000. Coverslips were mounted in ProLong Gold Antifade with DAPI (ThermoFisher), allowed to cure for 24 hours, then imaged on a Zeiss AxioImager.M1 epifluorescence microscope with a Hamamatsu ORCA-ER CCD camera and AxioVision v. 4.8.1. Images were cropped, scale bars added, and brightness/contrast adjusted for presentation with Zen Lite v. 2.5 (Zeiss).

Parasites for live microscopy were synchronized as in the preceding paragraph, harvested 24 hours after invasion, washed once in PBS, incubated with 1 μg/μL Hoechst 33342 (Invitrogen; H3570) for 30 minutes, then imaged on the same AxioImager.M1 described above. Images were analyzed in ImageJ: regions-of-interest were drawn around the mNeonGreen signal by hand, applied to all channels, and Pearson's Correlation Coefficients calculated using ImageJ's coloc.pearsons() function.

To monitor parasite size, 1437^{APT} was synchronized as above and monitored by thin smear at the indicated times (see Figure 6). Thin smears were fixed and stained with Harleco Hemacolor (MilliporeSigma), then imaged using a Zeiss Axio Observer.D1 at the Washington University Molecular Microbiology Imaging Facility. Parasite size was assessed using ImageJ, by manually drawing parasite borders and calculating their area.

To assess the effect of Pf3D7_1437000 depletion on protein export, 1437^{APT} was synchronized as above alongside Hsp101-DD (Beck et al., 2014), with Hsp101-DD maintained in 10 μM trimethoprim (TMP) or an equal volume of DMSO. Parasites were harvested 28 hours

after the invasion ended (i.e. parasites were 28-32 hours old), fixed, and processed as above. Primary antibodies are mouse anti-HRP2 clone 2G12 (Rock et al., 1987) diluted 1:500, mouse anti-FIKK4.2 (Kats et al., 2014) diluted 1:500, and mouse anti-KARHP clone 18.2 diluted 1:500.

Sorbitol lysis assay:

To measure sensitivity to sorbitol lysis, parasites were maintained as in the preceding paragraph. 28 hours after invasion, parasites were moved to a 96-well plate (100 μL/well), pelleted, and the media supernatant aspirated off. Every 5 minutes, we resuspended another row of the plate (4 conditions, 3 technical replicates per condition) in 5% sorbitol. One row was instead resuspended in deionized water to fully lyse the RBCs. After 24 minutes, the infected RBCs were again pelleted and the supernatant transferred to a new 96-well plate. Lysis of infected RBCs was measured by absorbance at 405 nm, a measure of hemoglobin abundance, using an Envision Multimodal Plate Reader (PerkinElmer). Values are expressed relative to the deionized water control (representing 100% lysis). The cultures were at 6% parasitemia at the outset of the experiment, so a maximum expected value is 6% lysis.

Assessment of HRP2/3 acetylation:

To investigate the acetylation status of HRP2 and HRP3, 1437^{APT} parasites were synchronized and maintained as above. 44 hours after invasion, aTc was washed from cultures (3 washes; 5 min. each), and cultures resuspended in 500 nM aTc or an equal volume of DMSO. For the DiCre experiments, either 50 nM rapamycin or an equivalent volume of DMSO were added. At the end of the following cycle (again, 44 hours after inivasion) parasites were pelleted, media was aspirated off, and then the RBC and PV lysed in 0.035% saponin in PBS at 4°C for 15 minutes. Parasite material was then pelleted at 17,000 g and the supernatant poured onto nickel HTC agarose beads (Gold Biotechnology) that had been equilibrated in PBS + 100 mM

imidazole. Columns were then washed in 20 column volumes of PBS + 100 mM imidazole to remove the abundant hemoglobin, then eluted in 15 mL PBS + 1 M imidazole. Eluate was concentrated in 10K cutoff Amicon Ultra-15 Centrifugal Filters (Millipore), flash frozen in liquid nitrogen, then sent to the Danforth Plant Science Cetnter's Proteomics and Mass Spectrometry Facility.

There, samples were acidified by formic acid to 1% then cleaned up with C4 ZipTip (Millipore). The captured samples were eluted with 50% acetonitrile in 0.1% formic acid, dried down and then resuspended in 10 μL 3% acetonitrile in 1% formic acid. 5 μL of sample was analyzed by LCMS with a Dionex RSLCnano HPLC coupled to an OrbiTrap Fusion Lumos (Thermo Fisher Scientific) mass spectrometer using a 60 min gradient (2-90% acetonitrile). Sample was resolved using a 75 μm x 150 cm PepMap C4 column (Thermo Scientific). MS spectra of protein ions of different charge-states were acquired in positive ion mode with a resolution setting of 120,000 (at m/z 200) and accurate mass was deconvoluted using Xcalibur (Thermo Scientific).

4.8 References

Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, et al. 2009. PlasmoDB: A functional genomic database for malaria parasites. Nucleic Acids Res 37:D539–543

Banerjee R, Liu J, Beatty W, Pelosof L, et al. 2002. Four plasmepsins are active in the Plasmodium falciparum food vacuole, including a protease with an active-site histidine. Proc Natl Acad Sci U S A 99:990–995.

Barylyuk K, Koreny L, Ke H, Butterworth S, et al. 2020. A comprehensive subcellular atlas of the Toxoplasma proteome via hyperLOPIT provides spatial context for protein functions. Cell Host Microbe 28: 752–766.e9

Beck JR, Muralidharan V, Oksman A, Goldberg DE. 2014. PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes. Nature 511:592–5.

Boddey JA, Cowman AF. 2013. Plasmodium nesting: remaking the erythrocyte from the inside out. Annu Rev Microbiol 67:243–269.

Boddey JA, Hodder AN, Günther S, Gilson PR, et al. 2010. An aspartyl protease directs malaria effector proteins to the host cell. Nature 463:627–631.

Boddey JA, Moritz RL, Simpson RJ, Cowman AF. 2009. Role of the Plasmodium export element in trafficking parasite proteins to the infected erythrocyte. Traffic 10:285–99.

Boonyalai N, Collins CR, Hackett F, Withers-Martinez C, Blackman MJ. 2018. Essentiality of Plasmodium falciparum plasmepsin V. PLoS One 13:e0207621.

Bushell E, Gomes AR, Sanderson T, Anar B, et al 2017. Functional profiling of a Plasmodium genome reveals an abundance of essential genes. Cell 170:260-272.e8.

Chang HH, Falick AM, Carlton PM, Sedat JW, et al. 2008. N-terminal processing of proteins exported by malaria parasites. Mol Biochem Parasitol 160:107–15.

Chen F, Mackey AJ, Stoeckert CJ, Roos DS. 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Res 34:D363-8.

Cova M, López-Gutiérrez B, Artigas-Jerónimo S, González-Díaz A, et al. 2018. The Apicomplexa-specific glucosamine-6-phosphate N-acetyltransferase gene family encodes a key enzyme for glycoconjugate synthesis with potential as therapeutic target. Sci Rep 8:4005.

Deng S, Marmorstein R. 2020. Protein N-terminal acetylation: structural basis, mechanism, versatility, and regulation. Trends Biochem Sci 46:15–27.

Elsworth B, Matthews K, Nie CQ, Kalanon M, et al. 2014. PTEX is an essential nexus for protein export in malaria parasites. Nature 511:587–91.

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

Foyn H, Jones JE, Lewallen D, Narawane R, et al. 2013. Design, synthesis, and kinetic characterization of protein N-terminal acetyltransferase inhibitors. ACS Chem Biol 8:1121–1127.

Garten M, Nasamu AS, Niles JC, Zimmerberg J, et al. 2018. EXP2 is a nutrient-permeable channel in the vacuolar membrane of Plasmodium and is essential for protein export via PTEX. Nat Microbiol 3:1090–1098.

Ginsburg H, Kutner S, Krugliak M, Ioav Cabantchik Z. 1985. Characterization of permeation pathways appearing in the host membrane of Plasmodium falciparum infected red blood cells. Mol Biochem Parasitol 14:313–322.

Glushakova S, Busse BL, Garten M, Beck JR, et al. 2017. Exploitation of a newly-identified entry pathway into the malaria parasite-infected erythrocyte to inhibit parasite egress. Sci Rep 7:12250.

Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, et al. 2004. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science 306:1934–7.

Ho C-M, Beck JR, Lai M, Cui Y, et al. 2018. Malaria parasite translocon structure and mechanism of effector export. Nature 561:70–75.

Hwang C-S, Shemorry A, Varshavsky A. 2010. N-terminal acetylation of cellular proteins creates specific degradation signals. Science 327:973–7.

Ito S, Horikawa S, Suzuki Tateki, Kawauchi H, et al. 2014. Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). J Biol Chem 289:35724–30.

Jullien N, Sampieri F, Enjalbert A, Herman J-P. 2003. Regulation of Cre recombinase by ligand-induced complementation of inactive fragments. Nucleic Acids Res 31:e131.

Kats LM, Fernandez KM, Glenister FK, Herrmann S, et al. 2014. An exported kinase (FIKK4.2) that mediates virulence-associated changes in Plasmodium falciparum-infected red blood cells. Int J Parasitol 44:319–328.

Kirk K, Horner HA, Elford BC, Ellory JC, Newbold CI. 1994. Transport of diverse substrates into malaria-infected erythrocytes via a pathway showing functional characteristics of a chloride channel. J Biol Chem 269:3339–3347.

Klemba M, Beatty W, Gluzman I, Goldberg DE. 2004. Trafficking of plasmepsin II to the food vacuole of the malaria parasite Plasmodium falciparum. J Cell Biol 164.

Klemba M, Goldberg DE. 2005. Characterization of plasmepsin V, a membrane-bound aspartic protease homolog in the endoplasmic reticulum of Plasmodium falciparum. Mol Biochem Parasitol 143:183–191.

Marapana DS, Dagley LF, Sandow JJ, Nebl T, et al. 2018. Plasmepsin V cleaves malaria effector proteins in a distinct endoplasmic reticulum translocation interactome for export to the erythrocyte. Nat Microbiol 3:1010–1022.

Marti M, Good RT, Rug M, Knuepfer E, Cowman AF. 2004. Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science 306:1930–3.

Nasamu AS, Falla A, Pasaje CFA, Wall BA, et al. 2021. An integrated platform for genome engineering and gene expression perturbation in Plasmodium falciparum. Sci Rep 11:342.

Nessel T, Beck JM, Rayatpisheh S, Jami-Alahmadi Y, et al. 2020. EXP1 is required for organization of EXP2 in the intraerythrocytic malaria parasite vacuole. Cell Microbiol cmi.13168.

Nkrumah LJ, Muhle RA, Moura PA, Ghosh P, et al. 2006. Efficient site-specific integration in Plasmodium falciparum chromosomes mediated by mycobacteriophage Bxb1 integrase. Nat Methods 3:615–621.

Osborne AR, Speicher KD, Tamez PA, Bhattacharjee S, et al. 2010. The host targeting motif in exported Plasmodium proteins is cleaved in the parasite endoplasmic reticulum. Mol Biochem Parasitol 171:25–31.

Park S-E, Kim J-M, Seok O-H, Cho H, et al. 2015. Control of mammalian G protein signaling by N-terminal acetylation and the N-end rule pathway. Science (80-) 347:1249–1252.

Perrin AJ, Bisson C, Faull PA, Renshaw MJ, et al. 2021. Malaria parasite schizont egress antigen-1 plays an essential role in nuclear segregation during schizogony. mBio 12.

Polino AJ, Nasamu AS, Niles JC, Goldberg DE. 2020. Assessment of biological role and insight into druggability of the Plasmodium falciparum protease plasmepsin V. ACS Infect Dis 6:738–746.

Rock EP, Marsh K, Saul AJ, Wellems TE, et al. 1987. Comparative analysis of the Plasmodium falciparum histidine-rich proteins HRP-I, HRP-II and HRP-III in malaria parasites of diverse origin. Parasitology 95 (Pt 2):209–27.

Russo I, Babbitt S, Muralidharan V, Butler T, et al. 2010. Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte. Nature 463:632–6.

Scott DC, Monda JK, Bennett EJ, Harper JW, Schulman BA. 2011. N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. Science 334:674–678.

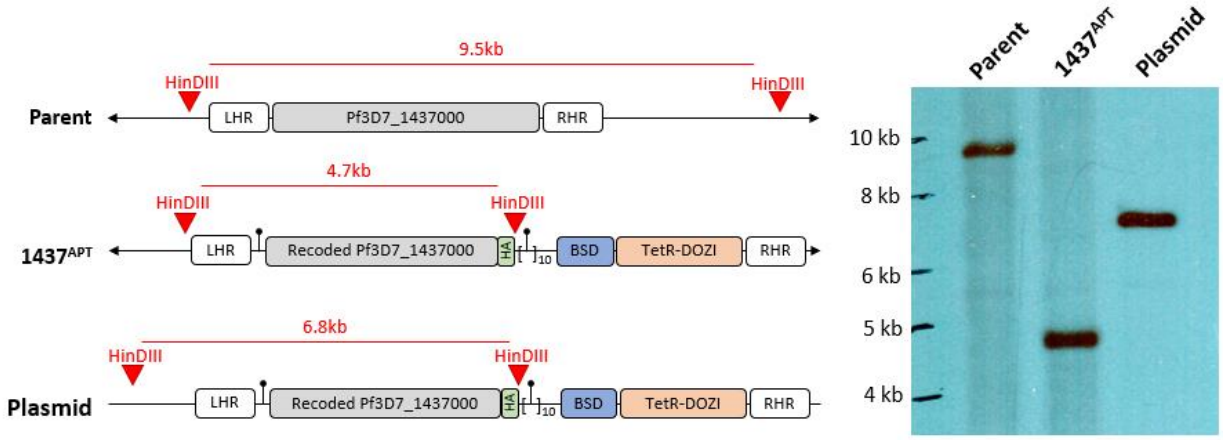
Sleebs BE, Lopaticki S, Marapana DS, O'Neill MT, et al. 2014. Inhibition of plasmepsin V activity demonstrates its essential role in protein export, PfEMP1 display, and survival of malaria parasites. PLoS Biol 12:e1001897.

Tarr SJ, Cryar A, Thalassinos K, Haldar K, Osborne AR. 2013. The C-terminal portion of the cleaved HT motif is necessary and sufficient to mediate export of proteins from the malaria parasite into its host cell. Mol Microbiol 87:835–850.

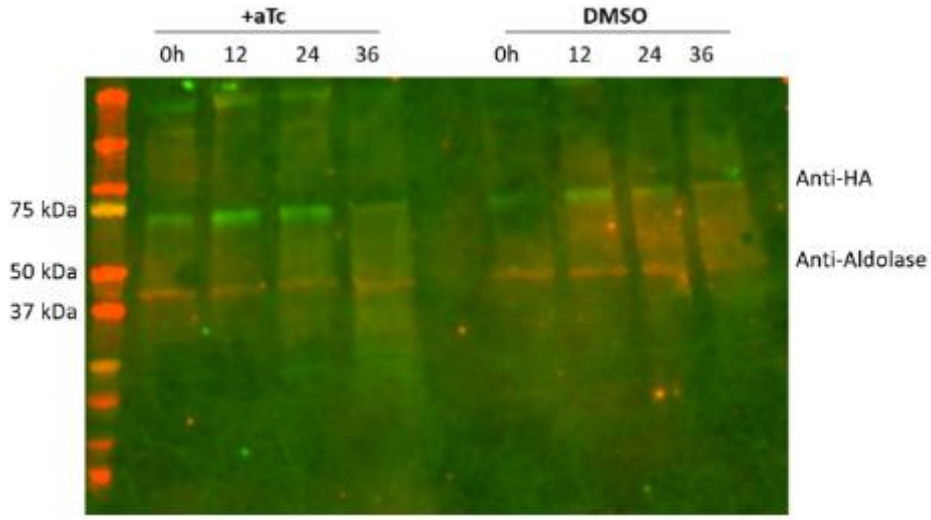
Tarr SJ, Osborne AR. 2015. Experimental determination of the membrane topology of the Plasmodium protease plasmepsin V. PLoS One 10:e0121786.

Waller RF, Keeling PJ, Donald RG, Striepen B, et al. 1998. Nuclear-encoded proteins target to the plastid in Toxoplasma gondii and Plasmodium falciparum. Proc Natl Acad Sci U S A 95:12352–7.

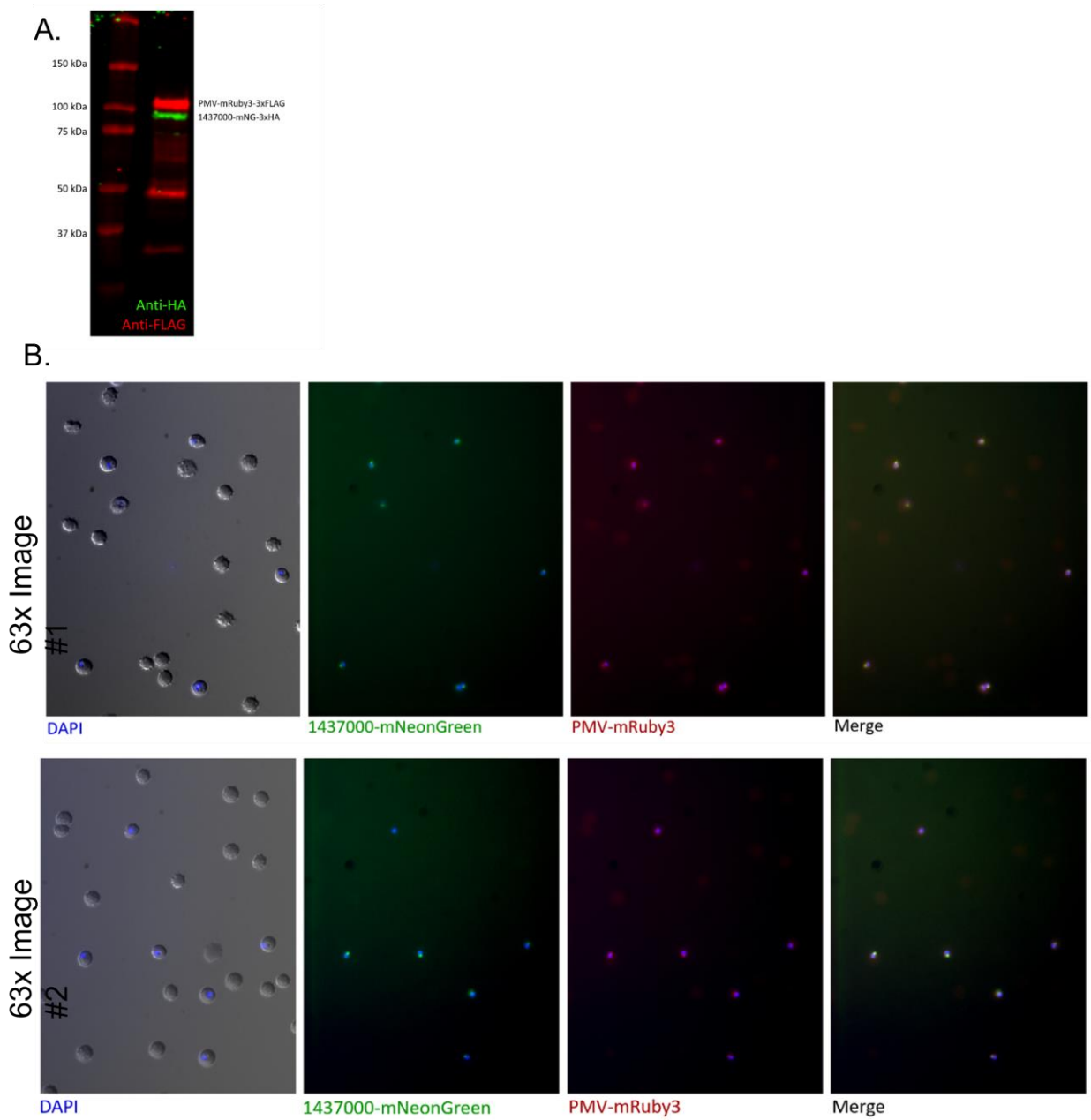
4.9 Supplemental Figures



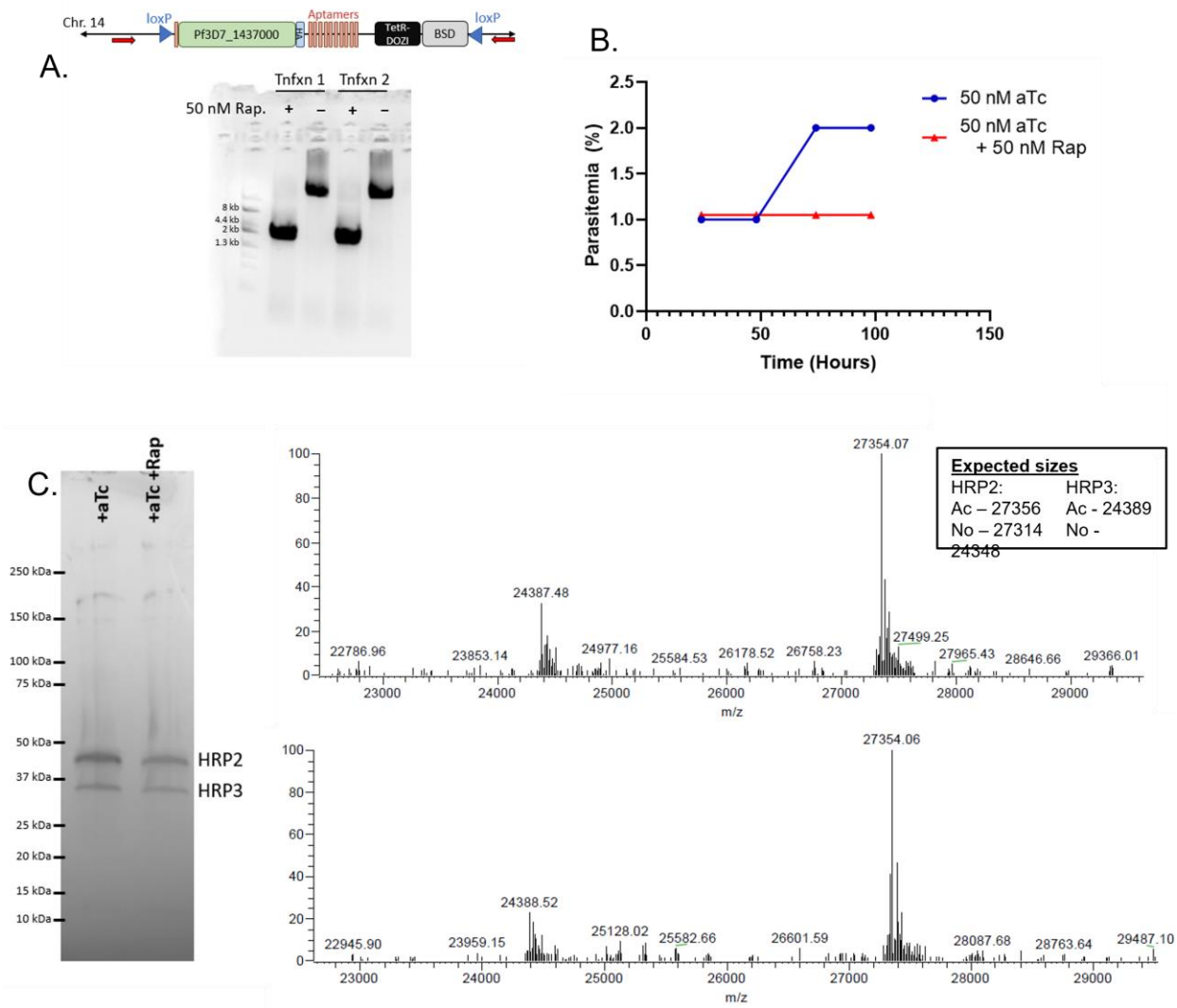
Supp. Fig. 1 – Southern blot: Correct genome editing was verified by Southern blot with left homologous region (LHR) used to probe the HinDIII-fragmented genome. Digest schematic shows expected size of bands for the parent, edited line (1437APT) and donor plasmid.



Supp. Fig. 2 – Uncut gel for Figure 2



Supp. Fig. 3 – Supplement to Figure 3B: (A) Western blot probing Pf3D7_1437000-mNeonGreen-3xHA and PM V-mRuby3-3xFLAG double-tagged line with antibodies for anti-HA (green) and anti-FLAG (red). (B) Additional images from Fig. 3B.



Supp. Fig. 4 – DiCre excision of Pf3D7_1437000 locus also doesn't affect HRP2/HRP3 acetylation. (A) PCR to check excision of Pf3D7_1437000 locus upon addition of rapamycin (Rap) to cultures. At top, red arrows represent PCR primer locations. Expected sizes are 11 kb for the parent locus, 4 kb for the excised locus. Below, an ethidium bromide-stained agarose gel showing PCR products 24 hours after excision from two different transfections ("Tnfxn1" and "Tnfxn2"). (B) Rapamycin was added to ring stage parasites and growth followed daily by flow cytometry. Experiment was performed twice; a representative experiment shown. (C) Deconvoluted mass spectra from analysis of intact HRP2 and HRP3. Inset shows anticipated sizes for acetylated and un-acetylated HRP2 and HRP3 after PEXEL cleavage. This experiment was performed twice; a representative experiment is shown.

Name	Sequence
14APT-1	caaacttcattgactgtgccggccggccGAAGATTTTTCAATTGTTGATGGATGTGG
14APT-2	atgagctccggcaaatgacaagCCTTTTTGAAATCCATTTATACATCCATAAAAAAAATG
14APT-3	aacccggaattcgagctcggCCGTATTAATATTTATTTAATTTTAATGAAAAAAAATGTC
14APT-4	cgagagattgggtattagaccCATAATTTAAAAAAGATAGAAAAGAAGCACG
14NG-1	tgacactatagaactcgagGTATATATAATGATAAAAGGAATcGTCaC
14NG-2	cacatcctacttcgaaGATATACATGAAGAAACACACCCTCA
14NG-3	CTTCATGTATATCttcgaaGTAGGATGTGTAGGTATAGTACCG
14NG-4	ctccacttcccctaggATCtAgGATATTATTTGTATAGCTCTGTATTTCTG
PMVR-1	ggtgacactatagaactcgagCCCAATGCGCACATTTTATACTTACATC
PMVR-2	gtactacctgaatctacttcgaaCACATTTTGGCTAGTTTGTCCATGAATAATAAAAG
PMVR-3	ctagccaaaatgtgttcgaaGTAGATTCAGGTAGTACATTTACACATATTCCAG
PMVR-4	CtcctccacttccgctagcTGTTGATTCCTGTATGGGAGATTTATTTTGATCAgaCAgAGGAAAATAATCC
14G-1	taagtatataatatt**GAGCTATACAAATAATATCT**gttttagagctagaa
14G-2	taagtatataatatt**TTCATACATTTTTAATATTT**gttttagagctagaa
14G-3	taagtatataatatt**AAATTAAATAAATATTAATA**gttttagagctagaa
14G-4	taagtatataatatt**AAAGTCTTTAAAATATAAAT**gttttagagctagaa
PMVG-1	taagtatataatatt**AAATGTGCGCATTGGGAATG**gttttagagctagaa
PMVG-2	taagtatataatatt**GATTTATTTTGATCACTCAA**gttttagagctagaa
NG-HA-F	gatgaaaataaagaagctagcGGAAGTGGAGGAGTGAGCAAGGGCGAGGAGGATAA
NG-HA-R	TtatataactcgacgcggccgtcaGGCATAATCTGGAACATCGTAAGGATACGCATAATCGGGCACATCATAGGG ATAGCCAGCGTAGTCCGGGACGTCGTACGGGTAcctaggcttgtacagctcgtccatgcc
Rub-H-F	acttatctaagaccgctagcggaagtggaggaGTGTCTAAGGGCG
Rub-H-R	cgtccttgtagtccctaggCTTGTACAGCTCGTCCATGCCACC

Supp. Table 1: Primers used in this study. Note (1) Lowercase letters are overhanging sequence homologous to the vector backbone for In-Fusion or Gibson cloning, (2) Letters highlighted in yellow are shield mutations to avoid targeting the donor vector with CRISPR guides., (3) Bolded sequences represent the CRISPR guide sequences inserted into pAIO3, and (4) Sequence highlighted in red is 3xHA (or rather its reverse complement)

Chapter 5: Major findings and future directions

I set out with the goal of examining the contribution of ER-resident proteins in the process of protein export. Here I summarize our findings and the relevant questions going forward:

5.1 Major findings

5.1.1 PM V has an export-independent role

At the outset of the project, the canonical view of protein export had PM V as the PEXEL maturase, required only for protein export. In chapter 2 of this thesis, we describe our use of a post-transcriptional depletion technique to probe PM V function. We find that PM V is required for parasite survival, not just at the ring-trophozoite transition when other export machinery is known to be required (Beck et al., 2014; Elsworth et al., 2014; Garten et al., 2018), but also immediately after invasion (Fig. 2.4, 2.5). Parasites lacking PM V fail to grow much beyond the size of merozoites, and have dramatic morphological defects (Fig. 2.4). This post-invasion phenotype has not been seen in rapid destabilization of Hsp101 (Beck et al., 2014), or post-transcriptional depletion of Hsp101, PTEX150, or Exp2 (Elsworth et al., 2014; Garten et al., 2018), suggesting that PM V has some role independent of these other export-related proteins. The fact that this phenotype is recapitulated by inhibition with a peptidomimetic PM V inhibitor suggests that early death is caused by a failure of PM V catalysis, rather than some other function of the protein (Fig. 2.5). Together, our data suggest a PEXEL-containing protein (or set of proteins) that requires PM V cleavage to perform its essential role soon after invasion of a new red blood cell.

5.1.2 PM V is maintained in substantial excess

Our work overcomes a previous challenge for the field demonstrated by (Sleebs et al., 2014), where *glmS*-mediated ten-fold depletion of PM V levels resulted in no inhibition of parasite growth. The TetR-DOZI 5' and 3' aptamer system we describe here enabled more stringent regulation than was previously possible, allowing us to validate PM V's essentiality (Fig. 2.2). This system allowed us to titrate PM V levels, probing the relationship between PM V protein levels and parasite growth. We found that parasites with ~8% of wildtype PM V levels could yield viable daughter cells, while those with ~3% could not (Fig. 2.3). This suggests that PM V is maintained in substantial excess by intraerythrocytic parasites, at least in RBC culture.

5.1.3 PM V requires its nepenthesin insert for proper function

Combining our lethal depletion of PM V with the attP/attB recombination system (Nkrumah et al., 2006) allowed us to probe the role of PM V's structural anomalies by rescuing knockdown with mutant enzymes lacking structures of interest. Using this method, we found that a helix-turn-helix domain, an unpaired cysteine near the active site, and a highly charged surface loop were dispensable for parasite survival in culture (Fig. 3.2). In contrast, enzyme lacking a nepenthesin 1-type insert was unable to rescue parasite growth when endogenous PM V was depleted (Fig. 3.2). Replacing the insert with the presumedly similar structured orthologous sequence from the pitcher plant enzyme Nep1a also yielded an enzyme unable to rescue parasite growth from PM V depletion (Fig. 3.3), suggesting that the nepenthesin insert structure is not sufficient to recapitulate its role in PM V function.

5.1.4 Pf3D7_1437000 is uninvolved in protein export

We used the same TetR-DOZI system to probe the role of the putative N-acetyltransferase Pf3D7_1437000, which had been speculated to be the PEXEL N-

acetyltransferase. We used immunofluorescence and live microscopy to show that tagged Pf3D7_1437000 colocalizes with PM V, suggesting it resides in the ER (Fig. 4.3). Depletion of Pf3D7_1437000 was lethal to parasites in RBC culture, resulting in smaller parasites that failed to undergo schizogony at the appropriate time (Fig. 4.6, 4.7). This phenotype is distinct both from previous disruptions of protein export, and from the early death phenotype from PM V depletion, suggesting Pf3D7_1437000's essential role is distinct from known export machinery. Consistent with this, we find that depleted Pf3D7_1437000 has no effect on export of the PEXEL proteins HRP2 or FIKK4.2, again suggesting a role unrelated to protein export (Fig. 4.4). When we use nickel resin to pull down the naturally histidine-rich PEXEL proteins HRP2 and HRP3, we find that both are PEXEL-cleaved and acetylated despite Pf3D7_1437000 depletion (Fig. 4.5), suggesting that Pf3D7_1437000 may not be the PEXEL N-acetyltransferase.

5.2 Future Directions

5.2.1 Defining the role(s) of PM V

Our work in chapter 2 suggests that PM V's role in the parasite's intraerythrocytic development cycle is not yet completely defined. The simplest explanation for our findings is that all points where PM V is essential for parasite survival are due to its role as the PEXEL maturase, and that there exists some PEXEL protein(s) that play an essential role within the parasite or parasitophorous vacuole and require PM V processing in order to play this role. Candidates are many, and data to narrow them few. The *Plasmodium falciparum* genome contains an estimated 400 genes bearing the PEXEL sequence (Boddey et al., 2013), nearly 150 of which belong to expanded gene families that encode surface immunogens and adhesion-related proteins unlikely to be required for growth. The only genome-scale phenotype data

available to us is a recent *piggyBac* transposon screen, which predicts around 85 PEXEL-containing genes to be essential for parasite growth (Zhang et al., 2018). A recent review synthesized published data on PEXEL proteins with the *piggyBac* screen, arriving at a predicted 75 essential PEXEL proteins (Jonsdottir et al., 2021). Nearly all of these proteins are yet to be studied, and only a few contain predicted domains that would suggest a function.

In Chapter 2 we raise one possibility: that an essential PEXEL protein is processed and packaged into the dense granules of daughter cells. Upon reinvasion of a new host cell, this dense granule effector (or effectors) would be secreted into the nascent parasitophorous vacuole. Perhaps without PM V cleavage, some PEXEL-containing dense granule effector fails to fulfill its essential function, resulting in parasite death soon after invasion. Two tangential pieces of evidence could support such a theory: first, the dense granule protein RESA has a "relaxed PEXEL" – RxLxxE – yet is still cleaved by PM V (Chapter 2, Supplemental Figure 4), packaged into the dense granules, then exported via PTEX immediately upon invasion in the following intraerythrocytic cycle (Beck et al., 2014). Thus, we know it is possible for a PEXEL-containing protein to be diverted from the export pathway and packaged into the dense granules. Second, the PM V ortholog in *Toxoplasma gondii*, Asp5, cleaves dense granule effectors (Coffey et al., 2018, Coffey, et al. 2015; Hammoudi et al., 2015). Parasites lacking Asp5 still package these effectors into the dense granules, but upon the next invasion, the effectors are dysfunctional and the parasitophorous vacuole is grossly malformed (Hammoudi et al., 2015). Perhaps an analogous effector in *P. falciparum* could explain the grossly deformed parasitophorous vacuoles we see when PM V is depleted (Fig. 2.4). RESA is dispensable for intraerythrocytic parasite growth (Silva et al., 2005), but perhaps the dense granules hold other PEXEL proteins. Efforts to identify such an effector are currently hindered by the fact that we know neither what signal(s)

drives traffic to the dense granules, nor what the dense granules' contents are at any point in the life cycle. A dense granule proteome is sorely needed.

As for the second point of essentiality, the requirement for PM V early in the trophozoite stage (Boonyalai et al., 2018; Sleebs et al., 2014; this thesis Fig. 2.5), we assume that PM V is required for some essential protein(s) to be exported into the host RBC. Depletion of the PTEX components Hsp101 or PTEX150 arrested parasite growth at the same time (Beck et al., 2014; Elsworth et al., 2014), suggesting some key exported effector is required to progress into the trophozoite stage. Beck and Muralidharan, et al. showed that PTEX disruption protected parasites from sorbitol lysis, suggesting that protein export is required for the implantation or activation of channels on the host RBC surface (Beck et al., 2014), a function presumably essential for parasite survival. The known parasite proteins that contribute to this function: RhopH2, RhopH3, and CLAG3.1 (Counihan et al., 2017; Ito et al., 2017; Sherling et al., 2017) all lack a PEXEL motif, suggesting they are not the effectors impacted by PM V depletion (though CLAG3.1 has a few unexplored RxL sequences). Whether this is the only essential function of PTEX in culture remains unknown, but perhaps our work can help narrow the search. PTEX disruption blocked export of all studied PEXEL proteins and PEXEL-negative exported proteins. The fact that PM V depletion phenocopied the timing of growth arrest caused by PTEX disruption suggests that at least some essential effector is a PEXEL protein that requires PM V cleavage to function (though of course, several PEXEL proteins, or a mix of PEXEL and PNEP proteins may all be essential). As mentioned above, some efforts have been undertaken to predict essential PEXEL proteins; however, our current efforts have been insufficient to suggest an effector or set of effectors that may be playing this crucial role. Future work characterizing these

PEXEL proteins that are predicted to be essential will likely clarify why parasites require protein export in culture.

5.2.2 PM V as an antimalarial target

Part of the rationale for our group's longstanding interest in aspartic proteases is that they could serve as targets for antimalaria development (Meyers and Goldberg, 2012). What are the current prospects for PM V as a drug target? Some indicators point in favor. PM V is essential for parasite survival at several points of the intraerythrocytic cycle as well as in transmission stage gametocytes (Boonyalai et al., 2018; Jennison et al., 2019; Polino et al., 2020; Sleebs et al., 2014). Its structure has been determined to high resolution in complex with a PEXEL peptidomimetic inhibitor (Hodder et al., 2015). However, we sound two cautionary notes:

First, we show that PM V must be suppressed to very low levels to impact parasite growth. Consistent with this, a *glmS* ribozyme depletion of PM V depleted protein levels 10-fold with no affect on parasite growth (Sleebs et al., 2014) or processing of PEXEL proteins (Marapana et al., 2018). While it remains possible that parasites are more sensitive to PM V inhibition during an actual infection, current data suggest the enzyme is maintained in substantial excess. As we note in Chapter 2 those who screen compounds for activity against PM V should be aware that a compound concentration that serves as an IC_{50} *in vitro* may be insufficient to perturb parasite growth. Instead, an IC_{99} or higher may be more appropriate. This is consistent with described PM V inhibitors, each of which were potent *in vitro*, but at least 100-fold less so in tissue culture (Gambini et al., 2015; Hodder et al., 2015; Sleebs et al., 2014).

Second, despite some effort, no potent druglike PM V inhibitor has yet been described. The only compounds that have been validated to kill via PM V inhibition in parasites are the peptidic compounds WEHI-842, WEHI-916, and compound 29 (Gambini et al., 2015; Hodder et

al., 2015; Sleebs et al., 2014). Some other compounds have been shown to inhibit PM V activity *in vitro* but await validation in parasites. Phenotypic screens have revealed several compounds that kill parasites with promising drug-like properties, and large efforts are underway to evolve resistance to these compounds and thus identify the proteins and/or pathways they target (Yang et al., 2021). However, to our knowledge resistance-conferring mutations in PM V have yet to be discovered. Why no PM V inhibitors? Membrane permeability has been suggested as a possible impediment, and indeed to reach PM V a compound must cross the RBC membrane, parasitophorous vacuole, parasite plasma membrane, and enter the ER. However, inhibitors of secretory system-resident proteins are many in the literature; a similar feat is accomplished when chloroquine, piperaquine, and artemisinin enter the digestive vacuole, or atovaquone finds the mitochondrial inner membrane. A second possibility is that the PM V active site is sufficiently similar to the active sites of PM IX or PM X that most compounds that inhibit PM V would inhibit those proteases as well, and parasites may be more sensitive to PM IX and X inhibition than PM V. Alternatively, perhaps PM V is targetable, but we have yet to find appropriate compounds. Recently several studies have used *in silico* and *in vitro* approaches to identify potential PM V inhibitors (Meissner et al., 2019; Sharma et al., 2021; Sittikul et al., 2018); whether PM V is the killing target of any of these inhibitors remains to be seen.

5.2.3 The role of the nepenthesin insert in PM V function

In chapter 3, we show that the nepenthesin insert in PM V is required for its essential function and catalytic activity. What role does the nepenthesin insert play in this and other proteases? Besides a high cysteine content, PM V has little in common with the nepenthesin proteases, which are active at very low pH with extremely broad substrate specificities. Hodder, Sleebs, Czabotar, et al. hypothesized that the insert may be involved in interactions with the

PEXEL substrates (Hodder et al., 2015); however, the *Toxoplasma gondii* ortholog Asp5 has a nepenthesin insert interrupted by a run of 15 amino acids between cysteines three and four, yet still recognizes a PEXEL-like sequence. Structural information clarifying the orientation of the Asp5 nepenthesin insert, and/or domain swaps to determine if the Asp5 insert can replace the PM V insert would help to clarify whether these two enzymes may be using a nepenthesin insert similarly, or whether they are a dissimilar pair that can help narrow down the possible roles for this motif. (Goldberg, 2015) raised the possibility that the insert may be involved in measuring the N-terminal end of substrate proteins, allowing only PEXELs that are at a designated length from the N-terminus. Known PM V substrates have PEXEL motifs 15 to 30 amino acids downstream of the signal peptide; however, whether this location is required for PM V recognition remains completely unexplored.

Another possibility raised in (Goldberg, 2015) is that the nepenthesin loop is a structural support, perhaps used in nepenthesins to bolster their structure against the harsh digestive environment in the pitcher plant lumen, and in PM V to rigidify the flap over the active site to enhance specificity. Consistent with this idea, while the cysteines and the spacing between them are completely conserved in PM V sequences across the genus, the actual sequence between the cysteines is quite variable. Additionally the spacing differs slightly in nepenthesins compared to PM V, perhaps to support a slightly different structural role. A closer study of these intervening sequences, as well as molecular modeling to infer the affect of this loop on PM V dynamics may help generate further hypotheses. Alternatively, further study of this motif in nepenthesin enzymes may more rapidly clarify its role. Does removal of the nepenthesin insert disrupt nepenthesin I catalytic activity as it does for PM V? Does it influence stability in pitcher plant fluid? Perhaps it holds open the catalytic site to allow for its very broad substrate specificity at

low pH? To our knowledge no study has mutagenized any part of a nepenthesin protease to probe the relationship between its structural components and its function. Perhaps such a study is now due.

5.2.4 What acetylates PEXEL proteins?

We initiated the work described Chapter four with the hope of identifying the N-acetyltransferase (NAT) that acetylates PEXEL proteins after cleavage by PM V. While our data don't formally exclude the possibility that Pf3D7_1437000 is that NAT, we have no data to suggest that it is. The likeliest alternative would be one of the other annotated NATs in the *P. falciparum* genome, listed in Figure 4.1. We have assumed that the PEXEL NAT would be in the parasite ER since (A) it occurs after cleavage by PM V, an enzyme oriented into the ER lumen (Tarr and Osborne, 2015), (B) it occurs even if traffic from the ER is blocked by the fungal toxin brefeldin A, or the PEXEL substrate is retained in the ER by addition of an -SDEL retention sequence (Boddey et al., 2009; Chang et al., 2008; Osborne et al., 2010). However, none of the annotated NAT candidates have a hydrophobic stretch to suggest a mode of access to the secretory system, and each has an ortholog in other eukaryotes known to be a cytosolic protein. Despite this, it remains possible that one these NAT candidates, or some yet-undiscovered NAT accesses the ER and performs the PEXEL acetylation as originally conceived.

In Chapter 4 we raise the alternative possibility that following PEXEL cleavage, the new N-terminus is somehow accessed by cytosolic NATs before it continues on its secretory path. Chang, et al. found that when they pulled down various PEXEL proteins and submitted them to tandem mass spectrometry, all N-terminal peptides detected were consistent with a cleaved, acetylated N-terminus, suggesting that all of this processing happens very rapidly upon translation (Chang et al., 2008). The dynamics of ER import and PEXEL processing in *P.*

falciparum are little-studied, and we have some suggestion from Marapana, et al. that they may differ substantially from model eukaryotes (Marapana et al., 2018). Perhaps further work to delineate how and when PEXEL proteins are processed will help guide the search for the NAT(s) that processes these proteins. An obvious first step may be inducible depletions of the NAT candidates in Figure 4.1; however, if these proteins are anything like their counterparts in model eukaryotes, we can expect these NATs to modify many target proteins, and thus their phenotypes may be complex and challenging to interpret. *In vitro* work to recapitulate the activity of each and test the substrate specificity could clarify which enzymes *can* acetylate the post-cleavage PEXEL N-terminus, however determining which enzyme *does* perform this acetylation in parasites may be less straightforward.

5.2.5 How are PEXEL proteins marked for export?

A now long-standing mystery in *P. falciparum* protein export is this: how do parasites mark certain proteins for export into the host cell? We know that the PEXEL motif (with appropriate downstream sequence, see section 1.6.1) is one way to drive export of a protein. However many proteins are exported despite lacking a PEXEL. The only known node common to all exported proteins is PTEX, yet how this complex can distinguish between export-ready proteins and other parasitophorous vacuole residents remains unclear. In initiating the project described in Chapter 4 we had hoped to assess whether N-terminal acetylation contributed to export competence; a more thorough assessment may need to wait for identification of the PEXEL NAT. Additionally it remains unclear whether N-terminal acetylation is unique to the PEXEL N-terminus, or is widespread in *Plasmodium* secretory proteins. Deep proteomics may help to clarify this point. As for other modifications besides acetylation, we show in Figure 4.5 that when we isolate parasite HRP2 and HRP3 and submit them intact to mass spectrometry, we

find masses consistent with PEXEL-cleaved acetylated HRP2 and HRP3, with no evidence of further modification. This suggests that at least for HRP2 and HRP3, no permanent modification besides cleavage at PEXEL and acetylation is required for export competence.

So what does control export competence? Two broad possibilities merit further exploration. First, what are the sequence determinants of export competence? A handful of mutants in fluorescence reporter proteins have been described (summarized here in pages 9-10) but a comprehensive analysis of sequences has yet to be performed. Unfortunately, given our current limitations of transfection efficiency, such an undertaking would require a heterologous system for testing (e.g. binding of a substrate peptide library to Hsp101), improved transfection efficiency, or a heroic undertaking with at least hundreds of individual transfections. The other connected possibility is that some chaperone(s) control recognition of export-competent proteins and transmit them to PTEX. Marapana, et al. performed extensive immunoprecipitations of PM V and other ER machinery +/- crosslinking (Marapana et al., 2018); no obvious candidate(s) is revealed by their results, but perhaps their peptide lists could serve as a starting point for such a study.

5.2.6 What is the role of Pf3D7_1437000

Lastly, what is the role of Pf3D7_1437000, the putative NAT that is the focus of Chapter 4? We know that it is essential for survival, and that depletion of Pf3D7_1437000 reduces parasite size and interferes with schizogony, but such a broad phenotype is challenging to connect to a particular parasite process. A likely first step for future investigation is to validate that Pf3D7_1437000 actually acts as an N-acetyltransferase. This could be done with recombinant enzyme against a PEXEL peptide, or a library of peptides, followed by mass spectrometry to detect acetylation. If recombinant enzyme production fails, these same

experiments could be performed by pulling down enzyme for parasite culture, as we have done with PM V in Chapter 3. If Pf3D7_1437000 is indeed an NAT, tracking down the substrates that contribute to the knockdown phenotype may still be a challenge. Each NAT tends to have many substrates; teasing out which substrates contribute to various aspects of the Pf3D7_1437000 depletion phenotype with reverse genetics may be beyond our current capacity.

An alternative path to gaining insight into Pf3D7_1437000's role in parasite biology may be to target its ortholog in *Toxoplasma gondii*, TgME49_305450. The genome-wide CRISPR screen in *T. gondii* assigned TgME49_305450 an intermediate growth score, higher than most validated essential proteins; lower than most validated dispensable proteins, leaving its essentiality challenging to guess at (Sidik et al., 2016). If the protein is essential in *T. gondii* perhaps the phenotype would shed some light on what processes are involved. If not, a clean knockout in *T. gondii* may prove a fruitful tool to explore the TgME49_305450 substrate profile using proteomics, as was done for the PM V ortholog Asp5 (Coffey et al., 2018).

5.3 Closing Remarks

The six years that I've spent in the malaria field have seen exciting progress on the mechanisms of protein export: the PTEX structure (Ho et al., 2018), the realization of Exp2's dual role as nutrient importer/protein exporter (Garten et al., 2018), extensive pulldowns of PM V and other ER machinery to define complexes involved in PEXEL processing (Marapana et al., 2018), and more. Alongside these advances came the advent of powerful tools to further our studies: the *piggyBac* transposon screen in *P. falciparum* (Zhang et al., 2018), the CRISPR indel screen in *T. gondii* (Sidik et al., 2016), hyperLOPIT for localizations of *T. gondii* proteins (Barylyuk et al., 2020), and the rise of proximity biotinylation (Khosh-Naucke et al., 2017; Nessel et al., 2020; Schnider et al., 2018), DiCre gene excision (Collins et al., 2013), and the

improved aptamer system for more convenient post-transcriptional depletion (Rajaram et al., 2020). As the culmination of my thesis work, I present three stories that each leave substantial questions – PM V is required immediately after invasion, though we don't know why; a nepenthesin insert is required for PM V activity, though its role remains unclear; Pf3D7_1437000 is essential for parasite survival, though we know neither what it does, nor the identity of the PEXEL NAT. It has been an exciting time to work and train in this collegial field. I hope my contributions, however small, help to guide the thinking of others as we attempt to puzzle out the workings of this interesting parasite. Many questions remain: I've listed a few that connect most directly to my thesis work above, two additional questions stand out as stubborn areas that have eluded us in my time here. (1) The export pathway remains shrouded in fog between PM V cleavage (or for PNEPs between ER import) and PTEX, and (2) the essential functions of exported proteins remain completely unknown. I am hopeful that we'll continue to see rapid progress in this field, both in terms of understanding gained and tools developed. As I prepare to leave the field behind, I look forward to reading answers to these and other questions in the near future. To the field, thank you for six years of challenging questions and interesting data. To the reader, thank you for your patience and consideration of this thesis.

5.4 References

Barylyuk K, Koreny L, Ke H, Butterworth S, et al. 2020. A comprehensive subcellular atlas of the *Toxoplasma* proteome via hyperLOPIT provides spatial context for protein functions. *Cell Host Microbe* **28**:752-766.e9.

Beck JR, Muralidharan V, Oksman A, Goldberg DE. 2014. PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes. *Nature* **511**:592–5.

Boddey JA, Carvalho TG, Hodder AN, Sargeant TJ, et al. 2013. Role of plasmepsin V in export of diverse protein families from the *Plasmodium falciparum* exportome. *Traffic* **14**:532–50.

Boddey JA, Moritz RL, Simpson RJ, Cowman AF. 2009. Role of the *Plasmodium* export element in trafficking parasite proteins to the infected erythrocyte. *Traffic* **10**:285–99.

Boonyalai N, Collins CR, Hackett F, Withers-Martinez C, Blackman MJ. 2018. Essentiality of *Plasmodium falciparum* plasmepsin V. *PLoS One* **13**:e0207621.

Chang HH, Falick AM, Carlton PM, Sedat JW, et al. 2008. N-terminal processing of proteins exported by malaria parasites. *Mol Biochem Parasitol* **160**:107–15.

Coffey MJ, Dagley LF, Seizova S, Kapp EA, et al. 2018. Aspartyl protease 5 matures dense granule proteins that reside at the host-parasite interface in *Toxoplasma gondii*. *mBio* **9**:e01796–18

Coffey MJ, Sleebs BE, Uboldi AD, Garnham A, et al. 2015. An aspartyl protease defines a novel pathway for export of *Toxoplasma* proteins into the host cell. *eLife* **4:**e10809.

Collins CR, Das S, Wong EH, Andenmatten N, et al. 2013. Robust inducible Cre recombinase activity in the human malaria parasite *Plasmodium falciparum* enables efficient gene deletion within a single asexual erythrocytic growth cycle. *Mol Microbiol* **88**:687–701.

Counihan N, Chisholm SA, Bullen HE, Srivastava A, et al. 2017. *Plasmodium falciparum* parasites deploy RhopH2 into the host erythrocyte to obtain nutrients, grow and replicate. *eLife* **6**:e23217.

Elsworth B, Matthews K, Nie CQ, Kalanon M, et al. 2014. PTEX is an essential nexus for protein export in malaria parasites. *Nature* **511**:587–91.

Gambini L, Rizzi L, Pedretti A, Taglialatela-Scafati O, et al. 2015. Picomolar inhibition of plasmepsin V, an essential malaria protease, achieved exploiting the prime region. *PLoS One* **10**:e0142509.

Garten M, Nasamu AS, Niles JC, Zimmerberg J, et al. 2018. EXP2 is a nutrient-permeable channel in the vacuolar membrane of *Plasmodium* and is essential for protein export via PTEX. *Nat Microbiol* **3**:1090–1098.

Goldberg DE. 2015. Plasmepsin V shows its carnivorous side. *Nat Struct Mol Biol* **22**:647–648.

Hammoudi P-M, Jacot D, Mueller C, Di Cristina M, et al. 2015. Fundamental roles of the golgi-associated *Toxoplasma* aspartyl protease, ASP5, at the host-parasite interface. *PLoS Pathog* **11**:e1005211.

Ho C-M, Beck JR, Lai M, Cui Y, et al. 2018. Malaria parasite translocon structure and mechanism of effector export. *Nature* **561**:70–75.

Hodder AN, Sleebs BE, Czabotar PE, Gazdik M, et al. 2015. Structural basis for plasmepsin V inhibition that blocks export of malaria proteins to human erythrocytes. *Nat Struct Mol Biol* **22**:590–6.

Ito D, Schureck MA, Desai SA. 2017. An essential dual-function complex mediates erythrocyte invasion and channel-mediated nutrient uptake in malaria parasites. *Elife* **6**:e23485.

Jennison C, Lucantoni L, O'Neill MT, McConville R, et al. 2019. Inhibition of plasmepsin V activity blocks *Plasmodium falciparum* gametocytogenesis and transmission to mosquitoes. *Cell Rep* **29**:3796-3806.e4.

Jonsdottir TK, Gabriela M, Crabb BS, F de Koning-Ward T, Gilson PR. 2021. Defining the essential exportome of the malaria parasite. *Trends Parasitol* **37**:664–675.

Khosh-Naucke M, Becker J, Mesén-Ramírez P, Kiani P, et al. 2017. Identification of novel parasitophorous vacuole proteins in *P. falciparum* parasites using BioID. *Int J Med Microbiol* 1–12.

Marapana DS, Dagley LF, Sandow JJ, Nebl T, et al. 2018. Plasmepsin V cleaves malaria effector proteins in a distinct endoplasmic reticulum translocation interactome for export to the erythrocyte. *Nat Microbiol* **3**:1010–1022.

Meissner KA, Kronenberger T, Maltarollo VG, Trossini GHG, Wrenger C. 2019. Targeting the *Plasmodium falciparum* plasmepsin V by ligand-based virtual screening. *Chem Biol Drug Des* **93**:300–312.

Meyers MJ, Goldberg DE. 2012. Recent advances in plasmepsin medicinal chemistry and implications for future antimalarial drug discovery efforts. *Curr Top Med Chem* **12**:445–55.

Nessel T, Beck JM, Rayatpisheh S, Jami-Alahmadi Y, et al. 2020. EXP1 is required for organisation of EXP2 in the intraerythrocytic malaria parasite vacuole. *Cell Microbiol* **22**:e13168.

Nkrumah LJ, Muhle RA, Moura PA, Ghosh P, et al. 2006. Efficient site-specific integration in *Plasmodium falciparum* chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat Methods* **3**:615–621.

Osborne AR, Speicher KD, Tamez PA, Bhattacharjee S, et al. 2010. The host targeting motif in exported *Plasmodium* proteins is cleaved in the parasite endoplasmic reticulum. *Mol Biochem Parasitol* **171**:25–31.

Polino AJ, Nasamu AS, Niles JC, Goldberg DE. 2020. Assessment of biological role and insight into druggability of the *Plasmodium falciparum* protease plasmepsin V. *ACS Infect Dis* **6**:738–746.

Rajaram K, Liu HB, Prigge ST. 2020. Redesigned TetR-aptamer system to control gene expression in *Plasmodium falciparum*. *mSphere* **5**:2020.05.19.105411.

Schnider CB, Bausch-Fluck D, Brühlmann F, Heussler VT, Burda P-C. 2018. BioID reveals novel proteins of the *Plasmodium* parasitophorous vacuole membrane. *mSphere* **3**:e00522-17.

Sharma PP, Kumar S, Kaushik K, Singh A, et al. 2021. *In silico* validation of novel inhibitors of malarial aspartyl protease, plasmepsin V and antimalarial efficacy prediction. *J Biomol Struct Dyn* 1–13.

Sherling ES, Knuepfer E, Brzostowski JA, Miller LH, et al. 2017. The *Plasmodium falciparum* rhoptry protein RhopH3 plays essential roles in host cell invasion and nutrient uptake. *eLife* **6**:e23239.

Sidik SM, Huet D, Ganesan SM, Huynh M-H, et al. 2016. A genome-wide CRISPR screen in *Toxoplasma* identifies essential apicomplexan genes. *Cell* **166**:1423-1435.e12.

Sittikul P, Songtawee N, Kongkathip N, Boonyalai N. 2018. *In vitro* and *in silico* studies of naphthoquinones and peptidomimetics toward *Plasmodium falciparum* plasmepsin V. *Biochimie* **152**:159–173.

Sleebs BE, Lopaticki S, Marapana DS, O'Neill MT, et al. 2014. Inhibition of plasmepsin V activity demonstrates its essential role in protein export, PfEMP1 display, and survival of malaria parasites. *PLoS Biol* **12**:e1001897.

Tarr SJ, Osborne AR. 2015. Experimental determination of the membrane topology of the *Plasmodium* protease Plasmepsin V. *PLoS One* **10**:e0121786.

Yang T, Ottilie S, Istvan ES, Godinez-Macias KP, et al. 2021. MalDA, Accelerating Malaria Drug Discovery. *Trends Parasitol*.

Zhang M, Wang C, Otto TD, Oberstaller J, et al. 2018. Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science* **360**:eaap7847.