Summer 8-15-2021

The Regulation of Plasmodium falciparum Metabolism by Haloacid Dehalogenase Proteins

Philip Frasse
Washington University in St. Louis

WASHINGTON UNIVERSITY IN SAINT LOUIS

Division of Biology and Biomedical Sciences
Molecular Microbiology and Microbial Pathogenesis

Dissertation Examination Committee:
Daniel Goldberg, Chair
Audrey R. Odom John, Co-Chair
Tamara Doering
Jeffrey Henderson
Christina Stallings

The Regulation of *Plasmodium falciparum* Metabolism by Haloacid Dehalogenase Proteins

By
Philip Michael Frasse

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

August 2021
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2dRib5P	2-deoxyribose 5-phosphate
3PGA	3-phosphoglyceric acid
8oxodGMP	8-oxo-deoxyguanosine monophosphate
Å	angstrom
A, Ala	Alanine
Abs	absorbance
ACT	Artemisinin combination therapy
ADP	adenosine diphosphate
AMP	adenosine monophosphate
AMR1	Apicoplast Minus Rescue 1 protein
aTc	anhydrotetracycline
ATP	adenosine triphosphate
BCE	before common era
BSD	blasticidin
BSTFA	N, O-bistrifluoroacetamide
CDC	Centers for Disease Control and Prevention
CE	common era
CMP	cytosine monophosphate
COVID-19	coronavirus disease
CQ	chloroquine
CRISPR	clustered regularly interspaced short palindromic repeats

CRT	chloroquine resistance transporter
CSP	Circumsporozoite protein
CTP	cytidine triphosphate
D, Asp	Aspartate
dADP	deoxyadenosine diphosphate
DAG	diacyl glycerol
dAMP	deoxyadenosine monophosphate
dATP	deoxyadenosine triphosphate
dCMP	deoxycytosine monophosphate
dCTP	deoxycytosine triphosphate
DDT	dichlorophenyltrichloroethane
dGDP	deoxyguanosine diphosphate
dGMP	deoxyguanosine monophosphate
dGTP	deoxyguanosine triphosphate
DHAP	dihydroxyacetone phosphate
DHFR	dihydrofolate reductase
DHODH	dihydroorotate dehydrogenase
DHPS	dihydropteroate synthase
dIMP	deoxyinosine monophosphate
DMP	Dess-Martin periodinane
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DOXP	deoxy-xylulose-5-phosphate

DOZI	Development of zygote inhibited, RNA helicase
dTMP	deoxythymidine monophosphate
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
dUTP	deoxyruridine triphosphate
DXR	deoxy-xylulose-5-phosphate reductase
E4P	erythrose-4-phosphate
E64	epoxysuccinyl-L-leucylamido(4-guanidino)butane
EC_{50}	half-maximal effective concentration
EDTA	ethylenediaminetetraacetic acid
F, Phe	Phenylalanine
F1P	fructose 1-phosphate
F6P	fructose 6-phosphate
FAD	flavin adenine dinucleotide
FBP	fructose-1,6-bisphosphate
FL	full length
FMN	flavin mononucleotide
FSM	fosmidomycin
G-1,6-P	glucose-1,6-bisphosphate
G6PDH	glucose 6-phosphate dehydrogenase
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GC-MS	gas chromatography mass spectrometry
gDNA	genomic DNA

GDP	guanosine diphosphate
GFP	green fluorescent protein
Glc1P	glycerol 1-phosphate
Glc2P	glycerol 2-phosphate
GlcN	glucosamine
glmS	glucosamine 6-phosphate activated ribozyme
Glu1P, G1P	glucose 1-phosphate
Glu6P, G6P	glucose 6-phosphate
Gly3P, GAP	glyceraldehyde 3-phosphate
GMP	guanosine monophosphate
GPI	glycosyl phosphatidylinositol
GPI-AP	GPI-anchored protein
GTP	guanosine triphosphate
H, His	histidine
HA	hemagglutinin tag
HAD	Haloacid Dehalogenase
HR	homologous region
Hs, h	*Homo sapiens*
HSP110	heat shock protein 110
HSP70	heat shock protein 70
I, Ile	isoleucine
IC_{50}	half-maximal inhibitory concentration
IDC	intraerythrocytic developmental cycle

IMP	inosine monophosphate
Ins	inositol
IPP	isopentyl pyrophosphate
IPTG	isopropyl- β-D-thiogalactoside
iRBC	infected RBC
K, Lys	Lysine
K13	Kelch-13
k_{cat}	turnover number
k_{cat}/K_m	catalytic efficiency
kDa	kilodaltons
K_m	Michaelis-Menten constant
LC-MS	Liquid chromatography – mass spectrometry
LDH1	lactate dehydrogenase 1
LIC	ligation independent cloning
LMIC	lower- and middle-income countries
LPA	lyso-phosphatidic acid
m/z	mass-to-charge ratio
M1P	mannose 1-phosphate
M6P	mannose 6-phosphate
ManC	mannose-1-phosphate guanylyltransferase
ManN	mannosamine
MDR1	multidrug resistance protein 1
MEP	Methyl-erythritol phosphate

MESG	methyl purine riboside
MetHb	methemoglobin
mL	milliliter
mM	millimolar
MMV	Medicines for Malaria Venture
MSP	merozoite surface protein
MW	molecular weight
N, Asn	Asparagine
nM	nanomolar
NMR	nuclear magnetic resonance
ns	not significant
OBP	octulose-1,8-bisphosphate
PBS	phosphate buffered saline
Pc	*Plasmodium cynomolgi*
PCR	polymerase chain reaction
PEG	polyethylene glycol
PEP	phosphoenolpyruvate
Pf	*Plasmodium falciparum*
PFK	phosphofructokinase
PFK9	phosphofructokinase 9
PGM	phosphoglucomutase
PGP	phosphoglycolate phosphatase
P_i	inorganic phosphate

PK	pyruvate kinase
PLP	pyridoxal-5’-phosphate, vitamin B6
PM5	plasmepsin 5
PMM	phosphomannomutase
PMSF	phenylmethylsulfonyl fluoride
PNP	purine nucleoside phosphorylase
*p*NPP	*para*-nitrophenyl phosphate
PPase	pyrophosphatase
PP_i	inorganic pyrophosphate
PV	parasitophorous vacuole
Pv	*Plasmodium vivax*
PVM	parasitophorous vacuole membrane
R, Arg	Arginine
R5P	ribose 5-phosphate
RAMA	rhoptry-associated membrane antigen
RBC	red blood cell
RFU	relative fluorescent units
RNA	ribonucleic acid
S, Ser	Serine
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SBP	sedoheptulose-1,7-bisphosphate
SCD	sickle cell disease
SD	standard deviation

SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sedo7P	Sedoheptulose 7-phosphate
SEM	Standard error of the mean
SNP	single nucleotide polymorphism
SP	sulphadoxine-pyrimethamine
Suc6P	sucrose 6-phosphate
T, Thr	Threonine
TAG	triacyl glycerol
TCA	Tricarboxylic acid
TetR	Tet repressor protein
TLC	thin layer chromatography
TMSI	trimethylsilyl iodine (iodotrimetylsilane)
tRNA	transfer RNA
TTP	thymidine triphosphate
UMP	uridine monophosphate
uRBC	uninfected RBC
US, USA	United States of America
UTR	untranscribed region
V_{max}	maximal enzyme velocity
WGS	whole genome sequencing
WHO	World Health Organization
WT	Wild type
WWI	World War I

X	premature stop codon (in context of “R157X”)
X	unspecified amino acid (in context of “DXD”)
XMP	xanthosine monophosphate
YbjI	5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase
YigB	5-amino-6-(5-phospho-D-ribitylamino) uracil phosphatase
YrfG	GMP/IMP nucleotidase
µg	microgram
µL	microliter
µM	micromolar
µmol	micromole

Acknowledgments

I am incredibly grateful for the support and guidance from my many mentors, colleagues, friends, and loved ones, without whom this work would never have been possible. Thank you to everyone who had a part in helping me through this process.

I would like to especially thank Dr. Audrey Odom John for her mentorship and relentless encouragement. Audrey has been a persistent force of optimism and guidance, regardless of whether experiments were working or grant applications were successful. She has given me confidence in my own work by acknowledging my strengths while pushing me to improve as a scientist, and she always inspires me to remain curious. Thank you, Audrey, for all your guidance during my thesis work and in planning for my future career.

Thank you also to Dr. Dan Goldberg for generously accepting me into the Goldberg lab for the past two years. I have felt incredibly welcomed and included as a bona fide Goldberger, and have benefited greatly from Dan's wisdom and experience as another mentor to turn to. Thank you so much, Dan.

Many thanks, also, to my thesis committee for sharing their advice and wisdom, teaching me when to drop a project and when to add more experiments to round out another one. Thank you also to the entire DBBS community and the microbiology (MMMP) program directors and coordinators who have guided me through the entire graduate school process.

Thank you to the peers and mentors I have gained in both the Odom John and Goldberg labs, who have taught me parasite culturing and biochemistry techniques, helped brainstorm experiments, and kept me sane with casual conversations, float trips, and puns. Thanks especially to Dana Hodge and Eva Istvan, who have together taught me all I know about culturing malaria

parasites. Thank you to Andrew Jezewski, Justin Miller, and Ann Guggisberg, with whom I spent many days and evenings learning about protein purification, crystallization, enzymology, and whom I could always rely on for help. A big thank you also to Alex Polino, my bay-mate in the Goldberg lab who was consistently the first person to field my countless questions as I became acquainted with a new lab. Thank you also to everyone else in both labs, who taught me protocols, helped with experimental design, and made my workdays enjoyable.

Thank you to our collaborators, Simon Cobbold, Malcolm McConville, and David Jakeman, who provided reagents, performed experiments, and brainstormed ideas. A special thank you to to Joe Jez for helping with crystallography and structure solving.

I am immensely appreciative of the wonderful friends that I have met during my time in graduate school. Many of them have doubled as colleagues and mentors, and I have valued all their input on my projects and experiments. I am also endlessly grateful for our non-scientific endeavors together, be they boardgames, rock-climbing, marathons, or musical theater.

Lastly, thank you to my family for always supporting me, for the phone calls and vacations, for hearing me present posters and for watching me dance on stage. Most of all, thank you to my all-time favorite person, my partner Leah. You are my biggest fan, and I am yours. I would not be who I am nor where I am without your love and encouragement.

Philip Frasse

Washington University in St. Louis

August 2021

Dedicated to Leah.

May we always continue to learn and

grow as scientists and people together.

ABSTRACT OF THE DISSERTATION

The Regulation of *Plasmodium falciparum* Metabolism by Haloacid Dehalogenase Proteins

Philip M Frasse

Doctor of Philosophy in Biology and Biomedical Sciences

Molecular Microbiology and Microbial Pathogenesis

Washington University in St. Louis, 2021

Professor Daniel E Goldberg MD, PhD, Chair

Professor Audrey R Odom John MD, PhD, Co-Chair

Malaria is an enormous financial and public health burden for much of the world, infecting over 200 million and killing over 400,000 people every year. While much progress has been made combating malaria in the past few decades, those advances have slowed in recent years, partially due to the emergence of resistance to all known antimalarials used to date. To achieve the goal of eliminating malaria as a major global health problem, new therapeutics need to be developed, targeting novel categories of parasite biology. One poorly understood area of parasite biology is the regulation of various metabolic pathways. We have recently identified a superfamily of proteins, named haloacid dehalogenase (HAD) proteins, that are implicated in resistance to metabolic inhibitors and regulation of essential metabolic pathways in *Plasmodium falciparum* malaria parasites. Here, we investigate how HAD2 (PF3D7_1226300) regulates metabolism of the isoprenoid biosynthesis pathway, using biochemical, metabolomic, and genetic tools. We find that HAD2 is a phosphatase with a preference for triose phosphates. We then investigate the related HAD proteins—HAD4 (PF3D7_1118400), Lipin (PF3D7_0303200), and HAD5 (PF53D7_1017400)—for their roles in regulating parasite metabolism and the implications for

future drug design. We find that HAD4 and Lipin are dispensable for growth in asexual malaria parasites. Lipin disruption causes significant growth reduction and accumulation of lipid species, while HAD4 is a dispensable nucleotide phosphatase. We also find that HAD5 is a phosphomannomutase that is essential for parasite egress and invasion. We solve the three-dimensional crystal structure of HAD5 and demonstrate our ability to selectively inhibit it compared to human phosphomannomutases. All of these findings add to our understanding of metabolic regulation in malaria parasites, illuminating key ways that targeting different metabolic pathways could work synergistically in development of novel antimalarial therapeutic strategies.

Chapter 1: Introduction

1.1 Malaria, and its primary causative agent, *Plasmodium falciparum*

1.1.1 Perspectives on infectious disease research in the COVID-19 era

Over the past year and a half, the world has been ravaged by a respiratory disease, COVID-19, caused by the virus SARS-CoV-2, which, in the year of 2020 killed approximately 1.8 million people globally[1]. In the midst of this pandemic, with evolving virus variants, changes to public health policy, and a complete uprooting of everyday life, it is difficult to view this pandemic through the historical lens of infectious diseases as a whole. Communicable diseases have been an ever-present and ever-evolving reality that humanity has had to face. Globally, communicable diseases were responsible for roughly one quarter of all deaths in 2019, and this number is far higher for lower-income countries[2]. One of the oldest and most deadly infectious diseases is malaria, which has plagued humanity since at least 4700 years ago, with references to its quintessential cyclical fevers in the ancient Chinese medical text, *Nei Ching* and the *Ebers Papyrus* from 1534 BCE, and *Plasmodium falciparum* DNA has been isolated from 4000 year old mummies[3–5]. Malaria parasites even predate humanity, with *Plasmodium spp.* DNA isolated from 30 million-year old *Culex* mosquitos preserved in amber[6]. So how does malaria stack up to a modern-day emergent pathogen like SARS-CoV-2? To provide some perspective, in the 20th Century alone, malaria killed approximately 150-300 million people[3], roughly equal to a COVID-19 pandemic every year of that century. However, it should be noted that recent decades have seen a great decline in malaria deaths, and we in affluent Western nations rarely hear or think about malaria as a result of the successful malaria elimination campaigns of the mid-20th

Century. Therefore, virtually all of the annual malaria cases and deaths occur in Africa, Southeast Asia, and South America[7].

Of course, this perspective should not be used to minimize the tragedy caused by COVID-19, nor to support notions that COVID-19 is an inconsequential "mild cold". Indeed, lower- and middle-income countries (LMICs) have also been hit hard by COVID-19 and have been given far fewer resources and vaccine doses than affluent countries. 85% of globally administered vaccine doses have been given to citizens in high- and upper-middle-income countries as of June 14, 2021, despite accounting for just 53% of the global population. Meanwhile, low-income countries that account for 8.7% of the global population have administered just 0.3% of global vaccines as of June 14 (vaccination data compiled from national ministries of health by OurWorldInData.org[8], population data from WorldBank.org[9]). Therefore, rather than trivialize the effects of COVID-19, let this subchapter serve as a reminder that although devastating infectious diseases like malaria have been largely eliminated from privileged nations like the USA, LMICs struggle with diseases on-par with COVID-19 every year, and those diseases have not taken a hiatus during the COVID-19 pandemic. Perhaps this outlook may rally some to harness their passion for funding and discovering cures and vaccines for COVID-19 and channel it toward diseases that continue to ravage communities abroad.

1.1.2 Current and historical global health impact of malaria

Malaria has shaped the lives, geopolitics, and even biology of humanity for millennia. References to malaria-like illnesses date back thousands of years to ancient Mesopotamia, China, and Egypt. Some believe that it was malaria that killed Alexander the Great (though others contend it was typhoid fever)[10], and Alaric I, King of the Visigoths was also likely killed by the disease after he sacked the city of Rome in 410 CE[11]. DNA evidence has even been found of

malaria parasites in the graves of Roman children circa 450 CE, lending credence to the idea that malaria epidemics were a key factor in the ultimate fall of the Roman Empire in 476 CE[12,13].

While treatments for this disease were occasionally rooted in effective methodologies, such as herbal remedies containing active antimalarial compounds, other treatments were less...robust. One of the most famous methods of treating malaria came from the roman physician Quintus Serenus Sammonicus, who instructed patients to write "ABRACADABRA" on a talisman in eleven rows, removing one letter per row until only an "A" remained, and wear it around their necks[14]. Although this was probably not a particularly effective treatment, we therefore have malaria to thank for one of our most well-known magical words.

Skipping ahead by a millennium or two, malaria continued to shape geopolitics into the 20th Century, as the advent of World Wars and a more global society led to an influx of immunologically naïve soldiers into malaria-endemic regions, resulting in massive numbers of infections and greater spread of different strains and species of parasites, even introducing the parasites to non-endemic regions[15]. At least 1.5 million soldiers were believed to be infected in World War I, with up to 5% case fatality[15]. In addition to the inevitable non-combat deaths of these malaria patients, large swaths of military personnel were put out of commission by the cumbersome illness, thwarting military plans.[15]

Although WWI provided a wake-up call to governments that malaria was more prevalent than they had appreciated, mitigation efforts did not prevent malaria from playing a role in the next World War as well, especially in the South Pacific Theater. Of note, the American campaign to prevent the Japanese invasion of the Philippines in 1942 was sidelined by malaria. Roughly 1/3 of the American and Filipino troops suffered from the disease, dramatically drawing out the war effort on that front. All-told, 60-65% of American soldiers in the South Pacific

reported contracting malaria, and there are some estimates that up to 90% of deployed Japanese troops at a time were rendered ineffective from the disease[16].

Besides military difficulties, malaria has been shaping the lives and even the genetics of humanity for thousands of years. Several genetic polymorphisms are known to confer degrees of resistance to malaria and have been selected for in populations where malaria is endemic. These polymorphisms include mutations that give rise to sickle cell trait, alpha- and beta-thalassemia, glucose-6-phosphate dehydrogenase deficiency, and Duffy antigen negativity, among others[17]. In the case of sickle cell trait, while people who are homozygous for sickle cell allele of the beta-globin gene are at risk of life-threatening sickle-cell anemia, termed Sickle Cell Disease (SCD), heterozygous carriers (who are said to have "sickle cell trait") have a 50-90% reduction in malaria parasite density[18], leading to a selection for that allele in malaria-endemic regions of sub-Saharan Africa. Additionally, the Duffy antigen, Fy glycoprotein, which is required for *Plasmodium vivax* infection, is largely missing in sub-Saharan African populations, leading to near elimination of *P. vivax* malaria from that region. These genetic selection factors have even further-reaching effects than current populations of Africa, as they are also common among people of African descent, such as African Americans, and can affect healthcare of these patients. Duffy antigen negativity, which is found in 65%-68% of African Americans, is associated with lower allograft procedure survival[19], and SCD occurs in 1 out of every 365 African Americans[20]. The presence of the sickle cell trait has even been controversially cited as an underlying cause of death in Black American men who died in police custody, relieving police of culpability in those deaths[21] and highlighting the pervasive effects that malaria has had on modern society.

By the mid 20th Century, decades of malaria-ridden wars led the United States to establish the Office of Malaria Control in War Areas, the predecessor to the Centers for Disease Control and Prevention (CDC). When WWII ended, the CDC (at the time, the Communicable Diseases Center) was established to commence the effort to eliminate malaria from the US: The National Malaria Eradication Program, which primarily consisted of vector control through the application of the insecticide dichlorodiphenyltrichloroethane (DDT). By 1949, the USA was declared free of malaria as a significant public health problem[22]. Similar efforts were successfully implemented throughout Europe, leading to the elimination of malaria in Europe by 1978[23].

However, despite effective elimination efforts in the USA and Europe, malaria remains a part of daily life in much of the world. In 2019, there were an estimated 229 million cases of malaria throughout the world, leading to 409,000 deaths, two-thirds of which were children under the age of five[7]. Approximately 94% of the cases and deaths from malaria occurred in the WHO Africa region, with nearly ¼ of deaths and 1/3 of cases in Nigeria alone[7]. While all of these numbers have dropped significantly in the last few decades, there remains a great burden caused by malaria on these countries. The financial costs, as well as opportunity costs of constant illness, are dramatic in malaria endemic countries. A 2001 study came to the conclusion that between 1965 and 1995, malaria endemic countries exhibited a 70% reduction in income per capita compared to non-endemic countries after correcting for other factors[24] . More recently, a similar analysis found a correlation between reduction of malaria incidence and an increase in gross domestic product per capita[25]. It is therefore quite clear that centuries of malaria endemicity have shaped the world as we know it, and there remains a huge burden of this disease that must be overcome to achieve true financial and healthcare equity across the globe.

1.1.3 Life cycle of malaria parasites

Malaria is caused by protozoan, apicomplexan parasites of the genus *Plasmodium*. While there are dozens of species of these parasites, which are known to infect diverse animals from birds and reptiles to rodents and deer[26–29], there are five *Plasmodium* species known to infect humans: *P. falciparum, P. vivax, P. malariae, P. ovale,* and *P. knowlesi*. Each species has its idiosyncrasies in disease and life cycle, but the deadliest and most common in modern times is *P. falciparum*, which is responsible for roughly 90% of the annual malaria deaths[30]. There also exists a robust and reliable culturing system to grow *P. falciparum* in the laboratory[31,32], allowing for much of malaria research to be conducted on this species. *P. falciparum* parasites are therefore the focus of the research in this dissertation, and it can thus be assumed that mentions of "parasites" or "malaria" are specifically referring to *P. falciparum* parasites or *P. falciparum*-caused malaria disease, unless otherwise specified.

P. falciparum parasites have a complex life cycle, composed of several transcriptionally, metabolically, and morphologically distinct stages. At the beginning of a human infection, the sporozoite form of the parasite is injected into the bloodstream from an infected female *Anopheles* mosquito while taking a bloodmeal. These sporozoites migrate to the human liver, where they then invade and take-up residence within the hepatocytes. The parasites grow and replicate within hepatocytes, multiplying their numbers to develop into upwards of 40,000 merozoites, capable of infecting red blood cells (RBCs). After 2-10 days, the merozoites are released out of the hepatocytes and into the blood stream, where they invade RBCs and commence their asexual replicative cycle.[33]

The RBC invasion process involves protein-protein interactions between the merozoite and the RBC, and several of these interacting proteins have been identified and well-documented.[34]

As the merozoite invades the erythrocyte, a process that takes mere minutes, the RBC membrane invaginates and encloses around the parasite, forming a parasite-containing vacuole within the host cell, termed the parasitophorous vacuole (PV) surrounded by the parasitophorous vacuolar membrane (PVM). Within the RBC, the parasite grows for 48 hours, with morphologically distinguishable phases. Soon after invasion, stained parasites resemble a diamond ring under the microscope, and have been termed "rings". As they grow and digest the RBC hemoglobin and other nutrients, the parasites start to encompass nearly the entire volume of the host cell, and form a distinctive digestive vacuole filled with crystallized heme from digested hemoglobin, termed hemozoin. These larger parasites are termed trophozoites and can further be informally subdivided by size into early- or late-trophozoites. Near the end of the 48-hour intraerythrocytic cycle, trophozoites undergo schizogony, a form of cell division where organelles and membranes divide to form roughly 16 to 32 daughter merozoite cells, still enclosed by the PVM. Finally, the parasites release proteins to degrade the PVM and subsequently the RBC membrane, releasing the daughter merozoites in a process called egress. These merozoites can then reinvade fresh RBCs and start the cycle anew. It is this 48-hour cycle, culminating with a large burst of egressing parasites, that causes the 48-hour cyclical fevers that are emblematic of *P. falciparum* malaria.

Occasionally, intraerythrocytic parasites follow an alternate route of development, which ends in their maturation into the sexual stages of parasites: male or female gametocytes.[35] The causative factors leading a parasite to develop into gametocytes rather than remaining asexual continues to be an active area of research. It is these gametocytes that are drawn up by mosquitoes during another bloodmeal, allowing them to develop further, undergo sexual reproduction, and grow in the midgut of the mosquito. Eventually, the progeny of that

reproduction develop once again into sporozoites and migrate to the mosquito salivary glands to be deposited into a new human host.

1.1.4 Unique biology and metabolism of malaria parasites

As a member of the apicomplexan phylum, *Plasmodium falciparum* is part of a unique branch of evolution, which is believed to be the result of a secondary endosymbiosis event[36]. Long ago, a eukaryotic algae cell engulfed a cyanobacterium in a primary endosymbiosis event, leading to a symbiotic relationship between the two. When that algal cell was itself engulfed by another eukaryote in a second endosymbiosis event, the resulting cell contained a unique organelle that formed from the membranes of the original cyanobacteria and algae and included a diverse set of genes. This is the evolutionary origin of apicomplexans like *P. falciparum*, and the unique endosymbiotic organelle is called the apicoplast[37]. This non-photosynthetic, four-membraned, plastid-like organelle has its own self-replicating circular genome, and is the site of several metabolic pathways in the parasites, including the methyl-erythritol-phosphate (MEP) pathway to produce isoprenoids, fatty acid biosynthesis, and heme biosynthesis[36,38]. Not only are the evolutionary origins of apicomplexans intellectually intriguing, however, they also provide unique opportunities for therapeutically targeting these pathogens. Because of the algal or bacterial origins of many parasite genes and metabolic pathways, several parasitic proteins are more similar to those of plants and bacteria than they are to metazoans[39]. This can be seen specifically with the MEP pathway.

The MEP pathway is a multiple-step enzymatic process of converting the glycolytic intermediates pyruvate and glyceraldehyde 3-phosphate (Gly3P) into isopentyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP)[40]. These products serve as the building blocks for all isoprenoids, including farnesyl and geranylgeranyl groups, which are post-translational

protein modifications, as well as dolichol, ubiquinone and more essential products[41]. In addition to malaria parasites, many bacteria and the chloroplasts of plants follow the MEP pathway to produce their IPP and DMAPP. However, other eukaryotes, including fungi and humans, follow an entirely distinct biosynthetic route involving mevalonate[40,42]. This provides an opportunity to develop antimalarial therapies that have a low risk of inadvertently targeting human enzymes.

For many years, it was believed that the sole essential function of the apicoplast in intraerythrocytic parasites was isoprenoid production, as disruption of apicoplast biogenesis entirely can be rescued with the MEP pathway product, IPP[43]. However, it was recently demonstrated that at least one apicoplast-targeted gene, dephospho-CoA kinase, remains active and essential even after the apicoplast is disrupted[44]. Nonetheless, the MEP pathway remains one of the key essential functions of the apicoplast, and has great potential for parasite-specific therapeutic development.

Even parasite genes in pathways are shared with humans, such as glycolysis, are sometimes sufficiently distinct from human orthologs due to our distant evolutionary paths, allowing for specific inhibitors to be made against parasite proteins without the risk of inhibiting human enzymes. For example, both humans and parasites encode the dihydrofolate reductase (DHFR) gene, an enzyme that reduces dihydrofolate to tetrahydrofolate, a necessary step in producing purines and some amino acids. Despite both species encoding DHFR, these enzymes are sufficiently unique to develop potent and specific inhibitors to parasite DHFR (*Pf*DHFR) without inhibiting human DHFR (*Hs*DHFR). Indeed, the previously used antimalarial pyrimethamine was widely used as a specific inhibitor of *Pf*DHFR, until resistance emerged[45]. Similarly, compound WR99210 is a potent inhibitor of *Pf*DHFR, but not *Hs*DHFR. This allows *Hs*DHFR to be introduced to parasites in the laboratory as a drug selection marker with WR99210

treatment when researchers are genetically editing parasites[46]. Thus, the unique metabolism and evolutionary distinctiveness of malaria parasites provides many opportunities for specific targeting of parasitic pathways.

1.2 Antimalarial therapeutics

1.2.1 History and targets of antimalarials

For as long as malaria has plagued humankind, people across the world have searched for ways to combat this deadly febrile illness. In South America, generations of native peoples learned that the bark of the chinchona tree offered abatement from fevers. When Jesuit missionaries arrived in the Andes in the 17th Century CE., they learned of these healing properties of chinchona bark and brought that knowledge back to Europe. It wasn't until 1820 that the bitter-tasting active compound from the bark was isolated and named: quinine.[47] Quinine, and other alkaloid compounds from chinchona bark, as well as their synthetic derivatives, have long been some of the most widely used antimalarial agents to this day. They are believed to interfere with the parasite's ability to safely digest hemoglobin, preventing the formation of hemozoin crystals and causing the accumulation of toxic free heme in the parasite.[48–50]

A similar story describes the discovery of another antimalarial game-changer. For thousands of years, ancient Chinese herbalists knew of the healing properties of sweet wormwood (*Artemisia annua)*, which they called Qinghao. In the 1960s, the Chinese government established a project to screen plants for antimalarial compounds, and by 1971, Dr. Youyou Tu and colleagues extracted the active component from the *Artemisia* plant, named artemisinin, and showed its efficacy in treating malaria, for which Dr. Tu was awarded the Nobel

Prize in 2015. Artemisinin and its derivatives then began use in the mid 1990s as an antimalarial therapy, and they continue to be used as a primary component of therapies, particularly in Southeast Asia, where they are used with other partner drugs in artemisinin combination therapies (ACTs). While the mechanism of action and mechanisms of resistance of artemisinin and its derivatives is an area of active research, it appears that these compounds kill parasites by causing non-specific damage through alkylating several proteins and cellular structures in the parasite.[51–55]

Many other antimalarial compounds have also been developed and administered, including drugs that target folate metabolism (proguanil, pyrimethamine/sulfadoxine), mitochondrial electron transport (atovaquone), and other inhibitors of heme detoxification (lumefantrine). Several metabolic pathways in the parasite are under active investigation for drug development, some of which have even gone through initial clinical trial stages, including drugs targeting pyrimidine synthesis enzyme dihydroorotate dehydrogenase (DHODH)[56] and the MEP pathway enzyme deoxy-xylulose-5-phosphate reductase (DXR)[57]. There also remains much active research into antimalarial compounds with unknown functions and unknown cellular targets. One frequently used repository of such compounds is the Malaria Box, a set of 400 compounds with antimalarial activity, funded through the Medicines for Malaria Venture (MMV).[58–61]

1.2.2 Rise of antimalarial resistance

The recurring story of antimalarial therapies is the resistance that has invariably emerged to all antimalarials known to date. One of the most successful antimalarials, chloroquine (CQ), which was first used in 1934, first faced resistance in the field by the late 1950s, and continued to spread throughout Africa, Asia, and South America through the 1960s and 1970s[62,63]. Parasites

had developed mutations in the *Pf*CRT gene or *Pf*MDR1, which allowed these proteins to shuttle chloroquine out of the parasite's digestive vacuole, its locus of action.[64–68]

Following the replacement of chloroquine with sulphadoxine-pyrimethamine (SP), resistance quickly emerged to that treatment as well. This time, the resistance was obtained primarily through target site mutation, with multiple mutations in the *DHFR* and *DHPS* (dihydropteroate synthase) genes emerging soon after SP's widespread use.[69–73]

With the introduction of artemisinin and ACTs in the 1990s, there was renewed hope as these drugs quickly became the first-line of defense[74–76]. However, as before, treatment failures were observed as early as the mid 2000s[77]. This time, parasites developed a unique form of resistance, termed "delayed clearance", wherein parasites would eventually succumb to drug treatment, but it took longer to achieve that clearance.[78] Thus, this form of resistance would theoretically be acceptable if patients were treated for longer, but difficulties with accessibility and patient compliance make this resistance concerning. The delayed clearance phenotype is a result of mutations to the Kelch13 (K13) gene in *P. falciparum*.[79–83] Several studies have investigated the natural role for K13 in the parasites, and how mutation confers resistance[84–86], but the most recent prevailing evidence points toward K13 involvement in endocytosis. It is the iron released from hemoglobin endocytosis and digestion that activates artemisinin, and thus K13 mutation slows down endocytosis and limits artemisinin activity[51,87]. While these mutations, and thus the treatment failures, are widespread in Southeast Asia[80,88], other regions of the world have thus-far been mostly spared. However, there have been some potentially concerning reports of K13 mutations in African parasite isolates recently[89,90]. If these became widespread, it could be devastating for the region, which already bears the brunt of *P. falciparum* cases and deaths[7].

1.2.3 Strategies to overcome antimalarial resistance

One of the most commonly implemented methods to prevent antimalarial resistance is the use of combination therapies. As previously mentioned, sulfadoxine-pyrimethamine served as a one-two hit against the folate biosynthesis pathway, but unfortunately these targets were easily able to evolve resistance. More recently, ACTs have taken the field by storm, by combining the potent, but short-acting artemisinin derivatives with a longer-lasting partner drug such as piperaquine or lumefantrine.[91]

Another strategy that has been put forward is to cycle through different drugs to take advantage of waning resistance.[92] This idea has been propelled by the finding that chloroquine resistance, and particularly the *Pf*CRT K76T mutations associated with CQ resistance, have been diluted out of the parasite population following the discontinuation of CQ in certain regions. In the absence of selective pressure from widespread CQ, sensitivity has re-emerged.[93,94] This finding could be to our advantage, cycling through different drug regimens as parasites develop and lose resistance.

Of course, one of the most commonly discussed strategies in antimalarial efforts is to develop new therapeutics that the parasites have not been exposed to. This consists of new derivatives of older drugs, which was seen with the advent of quinine derivatives such as chloroquine and primaquine[95], as well as the artemisinin derivatives dihydroartemisinin and artemether[96]. While this strategy has certainly proven useful, providing a wealth of new antimalarials to our repertoire with varying degrees of drug-like properties, parasites continue to develop resistance easily.

Therefore, the final method of combatting antimalarial resistance is, of course, the development of new drugs that target novel pathways. This is an enticing strategy against

malaria parasites, as they harbor unique, essential metabolic pathways[97]. Some of the metabolic pathways that have shown potential include the shikimate pathway[98,99], the MEP pathway[40,57,100], and pyrimidine synthesis[56,101]. However, target-based approaches to drug discovery can prove difficult; while it is relatively simple to identify essential and unique potential targets, developing a compound *de novo* to inhibit that target is not trivial. Therefore, much of the work being done on the novel drug target front is through identification of the targets of Malaria Box compounds, followed by compound optimization[60,61,102].

1.2.4 Vaccines against malaria

The Holy Grail of tools to combat infectious diseases is, of course, the vaccine. Unfortunately, an effective malaria vaccine has remained elusive, despite decades of attempts. Part of the reason for this lack of success is the complex parasite life cycle, with a variety of antigens that change throughout the parasite's life cycle, and even some surface antigens that can be mixed and matched within the intraerythrocytic stages.[103] Furthermore, the parasites live within host cells and sequester those infected RBCs in small blood vessels, largely shielding them from the immune system[104,105]. This may be why the most successful vaccine candidate to date, the RTS,S subunit vaccine, targets the circumsporozoite protein (CSP) on an extracellular stage of the parasites, the sporozoite[106,107]. However, this vaccine remains suboptimal, with ≤ 55% efficacy in children up to 17 months old, and this efficacy wanes over time[107]. By contrast, the World Health Organization (WHO) has called for efficacy of at least 75% for a vaccine to be sufficient for widespread use[108,109].

Recently, a new vaccine candidate, R21/MM completed its phase II clinical trial with very promising results, demonstrating up to 77% efficacy when administered with a high dose of adjuvant[110,111], surpassing the WHO guidelines. The R21 vaccine improved upon RTS,S by

increasing the relative concentration of CSP antigen compared to other antigens present on the vaccine particle, which appeared to be effective. While we will still await further phase II trials in a second malaria season, as well as phase III trials, these results show much promise for the field.

To hedge bets in case the CSP-targeting vaccines face future challenges, many labs continue to search for other potential vaccine targets that may help with future efforts to combat escape mutants and non-*falciparum* malaria species. Clinical trials are already underway that target antigens from sporozoites[112–114], blood stages[115–117], and sexual stages of the parasites[118–120]. And we are constantly searching for new antigens that may serve as vaccine candidates.

While the field continues to make progress on the development of new therapeutics and vaccines, and the cases and deaths from malaria have greatly declined in recent decades, the road to global malaria eradication remains long. It is therefore prudent that we advance our understanding of the cellular and metabolic processes within malaria parasites, which will allow further development of novel vaccines and therapies.

1.3 The Haloacid dehalogenase-like hydrolase (HAD) Superfamily of proteins

1.3.1 Introduction to HAD proteins

Haloacid Dehalogenase (HAD) and HAD-like proteins (IPR023214 and IPR036412) are a vast and varied superfamily of proteins found ubiquitously throughout life[121,122]. Although the superfamily is named for founding members that catalyze haloacid dehalogenase activity[123,124] the majority of these proteins actually catalyze phosphatase and phosphoryl-transferase

reactions[121]. These enzymes canonically consist of a two-domain structure, the first being a catalytic core domain containing a Rossmannoid fold. Distinct from other Rossmannoid fold-comprised proteins, HAD proteins also contain two structural motifs, termed the "squiggle" and "flap", and four conserved sequence motifs (I-IV).[125] The first of these conserved sequence motifs is the DXD di-aspartate catalytic site which coordinates a Mg^{2+} ion to facilitate catalysis. In addition to the catalytic core domain, the other notable feature of HAD proteins is the Cap domain, which is linked to the core domain through a hinge-like mechanism that allows the cap to close upon a substrate in the active site[126]. HAD enzyme cap domains can be categorized into C1 or C2, depending on the location of their insertion between other structural motifs of the enzyme, or C0 if the cap is a simple *beta* hairpin loop.[121] An analogous way of categorizing the HAD superfamily is to divide them subfamilies (I, II, and III) based their cap domain, with subfamilies I and II corresponding to C1 and C2 caps, respectively, and subfamily III corresponding to C0 caps[127].

HAD proteins have been observed catalyzing a diverse array of reactions across life. The majority of documented HAD proteins are small-molecule phosphatases, and several efforts have been made to characterize the substrate profiles of HAD proteins from various organisms, including the model organisms of *Escherichia coli* and *Saccharomyces cerevisiae*[128,129]. Other documented HAD proteins catalyze phosphomutase and phospho-transferase reactions upon small metabolites[130,131]. Yet other HAD proteins have also been characterized as protein-phosphatases[132,133]. While there have not been many conclusive links between subfamily categorization and substrate profile, Subfamily III proteins, marked by small C0 cap domains, appear capable of accommodating larger substrates like tRNA[134], or tyrosine-phosphates on polypeptides[135]. However, this is not cut-and-dry, as there are Subfamily III enzymes that

dephosphorylate smaller substrates[136], as well as Subfamily II members that catalyze protein phosphatase activity[132]. Some HAD protein have even been demonstrated to catalyze dephosphorylation both of proteins and small metabolites, prompting the suggestion that small-molecule phosphatase HAD enzymes may also conduct "moonlighting" activity as protein phosphatases.[133]

1.3.2 HAD Proteins in *P. falciparum*

Genome-wide searching of *P. falciparum* for proteins that contain HAD or HAD-like domains reveals approximately 23 protein-coding sequences (Fig. 1). Some of these, such as P-type ATPases and guanylyl cyclases, have been previously studied for their cellular functions[137–140], while others remain uncharacterized. In recent years, a handful of these uncharacterized HADs have begun catching our interest for the roles they play in parasite metabolism.

In 2014, Guggisberg et al. published a study[141] documenting the discovery and characterization of *Pf*HAD1, the first of the uncharacterized metabolic HADs to be described in *P. falciparum*. That study found that parasites lacking *Pf*HAD1 were resistant to fosmidomycin (FSM), a competitive inhibitor of the methyl-erythritol phosphate (MEP) pathway of isoprenoid biosynthesis. Mutation of HAD1 led to an increase in metabolism through the MEP pathway, marked by elevated levels of the MEP pathway starting substrate, deoxyxylulose 5-phosphate (DOXP), thus outcompeting FSM. In 2018, Dumont et al.[142] found another *P. falciparum* HAD protein, the phosphoglycolate phosphatase (PGP; PF3D7_0715000), that affected MEP pathway metabolism. When PGP was knocked out, dramatic changes were seen in parasite metabolism, particularly in the pentose phosphate pathway and flavin mononucleotide metabolism. Furthermore, PGP-knockout parasites were hypersensitized to FSM by decreasing metabolic flux toward the MEP pathway intermediates, the opposite effect seen with HAD1. Finally, another

HAD protein, termed *Pf*HAD2, was also found in resistance screens described by Guggisberg et al., and will be examined in the next chapter. All three of these enzymes are members of HAD Subfamily II. The implication of three Subfamily II members in FSM resistance begged the question: is there something inherent about this subfamily of proteins that links them to MEP pathway metabolism? Alternatively, is the MEP pathway uniquely sensitive to metabolic perturbations, such that many metabolic disruptions might confer changes to FSM sensitivity? If so, this would support further investigation of this pathway as an antimalarial target, pairing FSM with other antimalarial drugs that target metabolism for FSM hypersensitization.

1.3.3 Hypotheses on the mechanism(s) of HAD protein function

Although increasing evidence indicates that HAD protein family members, and specifically HAD subfamily II members, modulate metabolic homeostasis in *P. falciparum*, the cellular mechanism by which HAD enzymes control metabolism remains yet undefined. In a 2019 *Microbiology Insights* commentary[143] on the 2018 *mBio* paper from Guggisberg et al.[144], we proposed 2 hypothetical models of how HAD protein family members may regulate cellular metabolite levels (Figure 2). In the first model, HAD proteins regulate metabolism directly by enzymatically converting substrates that would otherwise directly feed into biosynthetic pathways (Fig. 2A). Thus, loss of HAD activity causes a build-up of phosphosubstrate levels, leading to increased availability of these compounds for downstream metabolic reactions. To date, however, most of the genetic changes described in HADs not only reduce catalytic activity but also lead to substantial reduction of overall HAD protein levels. Therefore, the metabolic changes in *had* mutants may be due to the loss of HAD activity or, alternatively, due in part to some non-catalytic "moonlighting" function of HADs. In this alternative model of HAD-mediated metabolic regulation, enzymatic activity serves as a molecular switch between "active"

and “inactive” conformational states (Fig. 2B). Following the paradigm established by small GTPases (such as Ras and Rab homologs),[145] the conformational change in a given HAD might modulate localization or protein-protein interactions to mediate a regulatory effect. In support of this alternative model, structural studies of substrate-bound HAD1 by Park et al.[126] found that HAD1 undergoes a large conformational change (from an “open” to “closed” conformation) during cap domain closure on substrate binding. It remains to be seen if the intriguing *P. falciparum* HAD proteins act through one or both of these mechanisms to exert their effect on parasite metabolism and cellular function.

1.4 Aim and Scope of the Dissertation

With the growing appreciation for the diverse roles that HAD proteins play in cell biology and metabolism, and recent findings of their roles in *P. falciparum* parasites, we sought to further understand a subset of HAD proteins in these parasites. Due to its emergence in FSM-resistance screens and strong homology to HAD1, HAD2 was chosen as the first protein to explore for interesting and important parasite biology and is covered in Chapter 2. Based on phylogeny and membership in related HAD subfamilies to HAD1 and PGP, PF3D7_1118400 (HAD4) and PF3D7_1017400 (HAD5) were next chosen as interesting HAD proteins to investigate. Investigation of HAD4 is described in Chapter 3 alongside another HAD protein termed Lipin, while HAD5 is characterized in Chapter 4. Other related HAD enzymes that are outside the scope of this dissertation include HAD6 (ISN1; PF3D7_1206100), which has recently been characterized as an IMP-specific nucleotidase[146], many P-type ATPases, and several more (Fig. 1).

This dissertation therefore aims to further characterize the growing list of Subfamily II and related HAD phosphatases in *P. falciparum* by investigating the above genes through genetic, biochemical, and metabolomic approaches. It also aims to further understand the link between the MEP pathway of isoprenoid biosynthesis (and specifically FSM sensitivity) and HAD proteins. Understanding these metabolic regulatory processes, particularly as they relate to essential metabolic functions of the parasite, may give insight into how these processes could be exploited for the future development of antimalarial therapeutics.

1.5 Figures

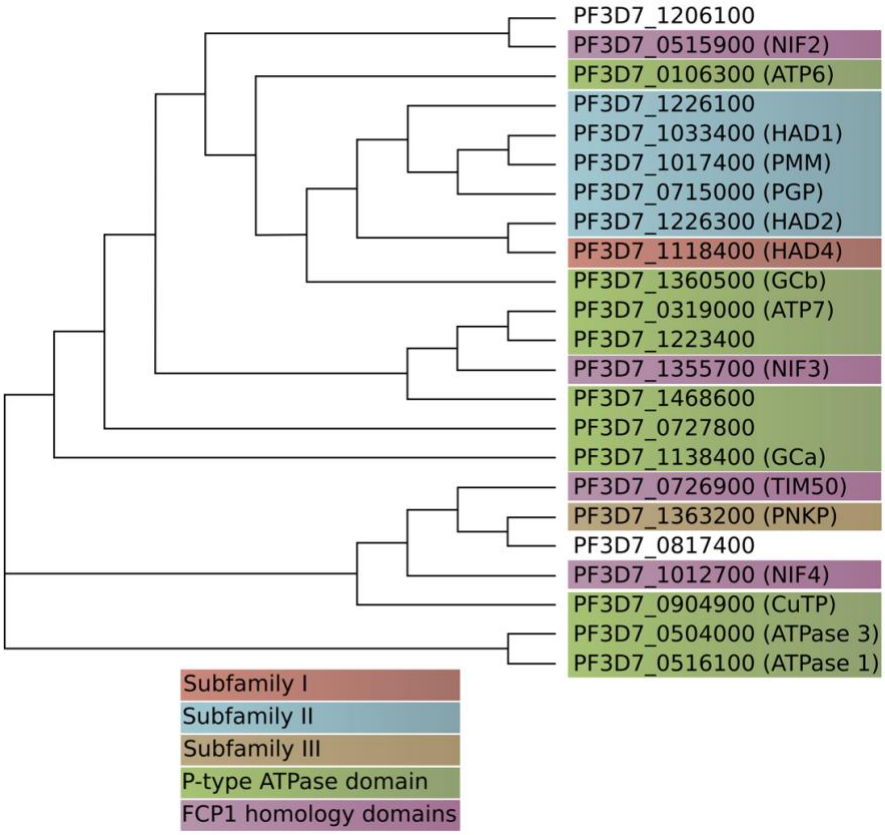


Figure 1. Phylogenetic tree of HAD superfamily (IPR023214) proteins from *P. falciparum*.

Protein names are given in parentheses, if applicable. HAD1, HAD2, and PGP all fall within Subfamily II of HAD proteins, and cluster near each other phylogenetically. Proteins are color-coordinated by the presence of InterPro domains: HAD Subfamily I (IPR006439), HAD Subfamily II (IPR006379 and IPR006357), HAD Subfamily III (IPR006549), P-type ATPase (IPR006539 and IPR008250), and FCP1 homology domain (IPR004274). Phylogenetic analysis was performed by ClustalW2 from EMBL-EBI, using the Neighbor-joining method with distance correction.[143]

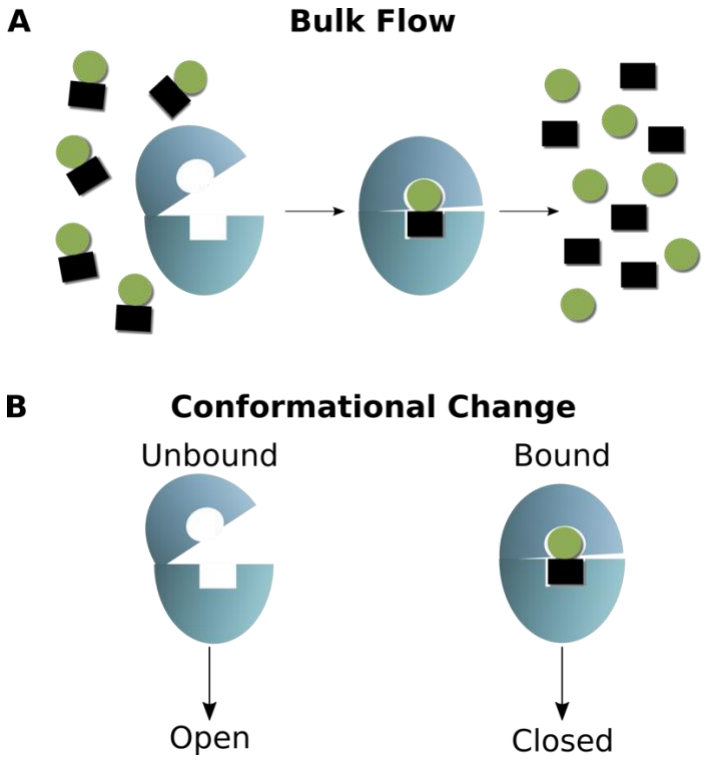


Figure 2. Model of hypothetical mechanisms of HAD-mediated metabolic regulation.

(**A**) HAD proteins catalyze the bulk flow of metabolites from one state to another. (**B**) Binding and hydrolyzing substrates induces conformational changes that dictate the regulatory functions of HAD proteins.[143]

1.6 References

1. Team, W. E. R. *COVID-19 Weekly Epidemiological Update - 5 January 2021*. (2021).
2. The top 10 causes of death. *World Health Organization* (2020). Available at: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. (Accessed: 14th May 2021)
3. Carter, R. & Mendis, K. N. Evolutionary and historical aspects of the burden of malaria. *Clinical Microbiology Reviews* **15**, 564–594 (2002).
4. Azab, M. An interview with Dr. Magda Azab. *Trop. Parasitol.* **3**, 170 (2013).
5. Taylor, G., Rutland, P. & Molleson, T. A sensitive polymerase chain reaction method for the detection of Plasmodium species DNA in ancient human remains – ScienceOpen. *Anc. Biomol.* **1**, 193–203 (1997).
6. Poinar, G. Plasmodium dominicana n. sp. (Plasmodiidae: Haemospororida) from Tertiary Dominican amber. doi:10.1007/s11230-004-6354-6
7. World Health Organization. *World Malaria Report 2020: 20 years of global progress and challenges*. (2020).
8. Coronavirus (COVID-19) Vaccinations - Statistics and Research - Our World in Data. Available at: https://ourworldindata.org/covid-vaccinations. (Accessed: 15th June 2021)
9. Data for High income, Upper middle income, Lower middle income, Low income | Data. Available at: https://data.worldbank.org/?locations=XD-XT-XN-XM. (Accessed: 15th June 2021)
10. Cunha, B. A. The death of Alexander the Great: Malaria or typhoid fever? *Infectious Disease Clinics of North America* **18**, 53–63 (2004).

11. Galassi, F. M. *et al.* The sudden death of Alaric i (c. 370-410 AD), the vanquisher of Rome: A tale of malaria and lacking immunity. *Eur. J. Intern. Med.* **31**, 84–87 (2016).

12. Thompson, A. BBC - History - Ancient History in depth: Malaria and the Fall of Rome. *BBC* (2011). Available at: https://www.bbc.co.uk/history/ancient/romans/malaria_01.shtml. (Accessed: 15th June 2021)

13. Sallares, R. & Gomzi, S. Biomolecular Archaeology of Malaria. *Anc. Biomol.* **3**, 195–213 (2001).

14. Hempelmann, E. & Krafts, K. Bad air, amulets and mosquitoes: 2,000 years of changing perspectives on malaria. *Malar. J.* **12**, 232 (2013).

15. Brabin, B. J. Malaria's contribution to World War One - The unexpected adversary. *Malaria Journal* **13**, 1–22 (2014).

16. Malaria in World War II - Army Heritage Center Foundation. Available at: https://www.armyheritage.org/soldier-stories-information/malaria-in-world-war-ii/. (Accessed: 14th May 2021)

17. Kariuki, S. N. & Williams, T. N. Human genetics and malaria resistance. *Human Genetics* **139**, 801–811 (2020).

18. Williams, T. N. *et al.* Sickle cell trait and the risk of Plasmodium falciparum malaria and other childhood diseases. *J. Infect. Dis.* **192**, 178–186 (2005).

19. Akalin, E. & Neylan, J. F. The influence of Duffy blood group on renal allograft outcome in African Americans. *Transplantation* **75**, 1496–1500 (2003).

20. Data & Statistics on Sickle Cell Disease | CDC. Available at: https://www.cdc.gov/ncbddd/sicklecell/data.html. (Accessed: 14th May 2021)

21. LaForgia, M. & Valentino-DeVries, J. How a Genetic Trait in Black People Can Give the Police Cover. *The New York Times* (2021). Available at: https://www.nytimes.com/2021/05/15/us/african-americans-sickle-cell-police.html?action=click&module=Top Stories&pgtype=Homepage. (Accessed: 15th May 2021)

22. Prevention, C.-C. for D. C. and. CDC - Malaria - About Malaria - History - Elimination of Malaria in the United States (1947-1951). (2019).

23. Piperaki, E. T. & Daikos, G. L. Malaria in Europe: emerging threat or minor nuisance? *Clinical Microbiology and Infection* **22**, 487–493 (2016).

24. Gallup, J. L. & Sachs, J. D. The Economic Burden of Malaria. *Am. J. Trop. Med. Hyg.* (2001).

25. Sarma, N., Patouillard, E., Cibulskis, R. E. & Arcand, J. L. The economic burden of Malaria: Revisiting the evidence. *Am. J. Trop. Med. Hyg.* **101**, 1405–1415 (2019).

26. Valkiunas, G. & Iezhova, T. A. Keys to the avian malaria parasites. *Malaria Journal* **17**, 212 (2018).

27. Guggisberg, A. M., Sayler, K. A., Wisely, S. M. & Odom John, A. R. Natural History of Plasmodium odocoilei Malaria Infection in Farmed White-Tailed Deer. *mSphere* **3**, e00067-18 (2018).

28. Silva, J. C. *et al.* Genome sequences reveal divergence times of malaria parasite lineages. *Parasitology* **138**, 1737–1749 (2011).

29. Boundenga, L. *et al.* Rodent malaria in Gabon: Diversity and host range. *Int. J. Parasitol. Parasites Wildl.* **10**, 117–124 (2019).

30. Snow, R. W. Global malaria eradication and the importance of Plasmodium falciparum epidemiology in Africa. *BMC Medicine* **13**, EE (2015).

31. Haynes, J. D., Diggs, C. L., Hines, F. A. & Desjardins, R. E. Culture of human malaria parasites Plasmodium falciparum. *Nature* **263**, 767–769 (1976).

32. Trager, W. & Jensen, J. B. Human malaria parasites in continuous culture. *Science (80-.).* **193**, 673–675 (1976).

33. Cowman, A. F., Healer, J., Marapana, D. & Marsh, K. Malaria: Biology and Disease. *Cell* **167**, 610–624 (2016).

34. Weiss, G. E. *et al.* Revealing the Sequence and Resulting Cellular Morphology of Receptor-Ligand Interactions during Plasmodium falciparum Invasion of Erythrocytes. *PLoS Pathog.* **11**, 1004670 (2015).

35. Talman, A. M., Domarle, O., McKenzie, F. E., Ariey, F. & Robert, V. Gametocytogenesis: The puberty of Plasmodium falciparum. *Malaria Journal* **3**, 24 (2004).

36. Vaishnava, S. & Striepen, B. The cell biology of secondary endosymbiosis - How parasites build, divide and segregate the apicoplast. *Molecular Microbiology* **61**, 1380–1387 (2006).

37. Janouškovec, J., Horák, A., Oborník, M., Lukeš, J. & Keeling, P. J. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 10949–10954 (2010).

38. Ralph, S. A. *et al.* Metabolic maps and functions of the Plasmodium falciparum apicoplast. *Nature Reviews Microbiology* **2**, 203–216 (2004).

39. Nkrumah, L. J. *et al.* Efficient site-specific integration in Plasmodium falciparum

chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat. Methods* **3**, 615–621 (2006).

40. Guggisberg, A. M., Amthor, R. E. & Odom, A. R. Isoprenoid Biosynthesis in Plasmodium falciparum. *Eukaryot. Cell* **13**, 1348–1359 (2014).

41. Imlay, L. & Odom, A. R. Isoprenoid Metabolism in Apicomplexan Parasites. *Current Clinical Microbiology Reports* **1**, 37–50 (2014).

42. Kuzuyama, T. Mevalonate and nonmevalonate pathways for the biosynthesis of isoprene units. *Biosci. Biotechnol. Biochem.* **66**, 1619–1627 (2002).

43. Yeh, E. & DeRisi, J. L. Chemical rescue of malaria parasites lacking an apicoplast defines organelle function in blood-stage plasmodium falciparum. *PLoS Biol.* **9**, (2011).

44. Swift, R. P., Rajaram, K., Liu, H. B. & Prigge, S. T. Dephospho-CoA kinase, a nuclear-encoded apicoplast protein, remains active and essential after Plasmodium falciparum apicoplast disruption . *EMBO J.* (2021). doi:10.15252/embj.2020107247

45. Chulay, J. D., atkins, W. M. & Sixsmith, D. G. Synergistic antimalarial activity of pyrimethamine and sulfadoxine against Plasmodium falciparum in vitro. *Am. J. Trop. Med. Hyg.* **33**, 325–330 (1984).

46. Rieckmann, K. H., Yeo, A. E. & Edstein, M. D. Activity of PS-15 and its metabolite, WR99210, against Plasmodium falciparum in an in vivo-in vitro model. *Trans. R. Soc. Trop. Med. Hyg.* **90**, 568–571 (1996).

47. Achan, J. *et al.* Quinine, an old anti-malarial drug in a modern world: Role in the treatment of malaria. *Malaria Journal* **10**, 144 (2011).

48. Egan, T. J. *et al.* Structure-function relationships in aminoquinolines: Effect of amino and

chloro groups on quinoline-hematin complex formation, inhibition of β- hematin formation, and antiplasmodial activity. *J. Med. Chem.* **43**, 283–291 (2000).

49. Kaschula, C. H. *et al.* Structure - Activity relationships in 4-aminoquinoline antiplasmodials. The role of the group at the 7-position. *J. Med. Chem.* **45**, 3531–3539 (2002).

50. Corrêa Soares, J. B. R. *et al.* Interference with hemozoin formation represents an important mechanism of schistosomicidal action of antimalarial quinoline methanols. *PLoS Negl. Trop. Dis.* **3**, (2009).

51. Birnbaum, J. *et al.* A Kelch13-defined endocytosis pathway mediates artemisinin resistance in malaria parasites. *Science (80-.).* **367**, 51–59 (2020).

52. Talman, A. M., Clain, J., Duval, R., Ménard, R. & Ariey, F. Artemisinin Bioactivity and Resistance in Malaria Parasites. *Trends in Parasitology* **35**, 953–963 (2019).

53. O'Neill, P. M., Barton, V. E. & Ward, S. A. The molecular mechanism of action of artemisinin - The debate continues. *Molecules* **15**, 1705–1721 (2010).

54. Hou, H. ping *et al.* Effects and Mechanism of Action of Artemisinin on Mitochondria of Plasmodium berghei. *Chin. J. Integr. Med.* **26**, 277–282 (2020).

55. Meshnick, S. R. Artemisinin: Mechanisms of action, resistance and toxicity. in *International Journal for Parasitology* **32**, 1655–1660 (Pergamon, 2002).

56. Kokkonda, S. *et al.* Lead optimization of a pyrrole-based dihydroorotate dehydrogenase inhibitor series for the treatment of malaria. *J. Med. Chem.* **63**, 4929–4956 (2020).

57. Wang, X. *et al.* MEPicides: α,β-Unsaturated Fosmidomycin Analogues as DXR Inhibitors against Malaria. *J. Med. Chem.* **61**, 8847–8858 (2018).

58. Spangenberg, T. *et al.* The Open Access Malaria Box: A Drug Discovery Catalyst for Neglected Diseases. *PLoS One* **8**, (2013).

59. Lucantoni, L., Duffy, S., Adjalley, S. H., Fidock, D. A. & Avery, V. M. Identification of MMV malaria box inhibitors of Plasmodium falciparum early-stage gametocytes using a luciferase-based high-throughput assay. *Antimicrob. Agents Chemother.* **57**, 6050–6062 (2013).

60. Malebo, H. M. *et al.* In vitro multistage malaria transmission blocking activity of selected malaria box compounds. *Drug Des. Devel. Ther.* **14**, 1593–1607 (2020).

61. Chirawurah, J. D. *et al.* Antimalarial activity of Malaria Box Compounds against Plasmodium falciparum clinical isolates. *Int. J. Parasitol. Drugs Drug Resist.* **7**, 399–406 (2017).

62. Payne, D. Spread of chloroquine resistance in Plasmodium falciparum. *Parasitology Today* **3**, 241–246 (1987).

63. Wellems, T. E. & Plowe, C. V. Chloroquine-resistant malaria. *Journal of Infectious Diseases* **184**, 770–776 (2001).

64. Foote, S. J. *et al.* Several alleles of the multidrug-resistance gene are closely linked to chloroquine resistance in Plasmodium falciparum. *Nature* **345**, 255–258 (1990).

65. Wellems, T. E., Walker-Jonah, A. & Panton, L. J. Genetic mapping of the chloroquine-resistance locus on Plasmodium falciparum chromosome 7. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 3382–3386 (1991).

66. Ecker, A., Lehane, A. M., Clain, J. & Fidock, D. A. PfCRT and its role in antimalarial drug resistance. *Trends in Parasitology* **28**, 504–514 (2012).

67. Reiling, S. J. & Rohrbach, P. Monitoring PfMDR1 transport in Plasmodium falciparum. *Malar. J.* **14**, 270 (2015).

68. Duraisingh, M. T. & Cowman, A. F. Contribution of the pfmdr1 gene to antimalarial drug-resistance. *Acta Trop.* **94**, 181–190 (2005).

69. Eskandarian, A. A., Keshavarz, H., Basco, L. K. & Mahboudi, F. Do mutations in Plasmodium falciparum dihydropteroate synthase and dihydrofolate reductase confer resistance to sulfadoxine-pyrimethamine in Iran? *Trans. R. Soc. Trop. Med. Hyg.* **96**, 96–98 (2002).

70. Cowman, A. F., Morry, M. J., Biggs, B. A., Cross, G. A. M. & Foote, S. J. Amino acid changes linked to pyrimethamine resistance in the dihydrofolate reductase-thymidylate synthase gene of Plasmodium falciparum. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 9109–9113 (1988).

71. Vasconcelos, K. F. *et al.* Mutations in Plasmodium falciparum Dihydrofolate Reductase and Dihydropteroate Synthase of Isolates from the Amazon Region of Brazil. *Mem. Inst. Oswaldo Cruz* **95**, 721–728 (2000).

72. Triglia, T., Menting, J. G. T., Wilson, C. & Cowman, A. F. Mutations in dihydropteroate synthase are responsible for sulfone and sulfonamide resistance in Plasmodium falciparum. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 13944–13949 (1997).

73. Gatton, M. L., Martin, L. B. & Cheng, Q. Evolution of resistance to sulfadoxine-pyrimethamine in Plasmodium falciparum. *Antimicrob. Agents Chemother.* **48**, 2116–2123 (2004).

74. White, N. J. Artemisinin: Current status. *Trans. R. Soc. Trop. Med. Hyg.* **88**, 3–4 (1994).

75. White, N. J. The Treatment of Malaria. *N. Engl. J. Med.* **335**, 800–806 (1996).

76. Klayman, D. L. Qinghaosu (artemisinin): An antimalarial drug from China. *Science (80-.).* **228**, 1049–1055 (1985).

77. Noedl, H. *et al.* Evidence of Artemisinin-Resistant Malaria in Western Cambodia. *N. Engl. J. Med.* **359**, 2619–2620 (2008).

78. Dondorp, A. M. *et al.* Artemisinin Resistance in Plasmodium falciparum Malaria . *N. Engl. J. Med.* **361**, 455–467 (2009).

79. Mbengue, A. *et al.* A molecular mechanism of artemisinin resistance in Plasmodium falciparum malaria. *Nature* **520**, 683–687 (2015).

80. Ashley, E. A. *et al.* Spread of Artemisinin Resistance in *Plasmodium falciparum* Malaria. *N. Engl. J. Med.* **371**, 411–423 (2014).

81. Ariey, F. *et al.* A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature* **505**, 50–55 (2014).

82. Ménard, D. *et al.* A Worldwide Map of Plasmodium falciparum K13-Propeller Polymorphisms . *N. Engl. J. Med.* **374**, 2453–2464 (2016).

83. Straimer, J. *et al.* Drug resistance. K13-propeller mutations confer artemisinin resistance in Plasmodium falciparum clinical isolates. *Science* **347**, 428–31 (2015).

84. Tilley, L., Straimer, J., Gnädig, N. F., Ralph, S. A. & Fidock, D. A. Artemisinin Action and Resistance in Plasmodium falciparum. *Trends in Parasitology* **32**, 682–696 (2016).

85. Wicht, K. J., Mok, S. & Fidock, D. A. Molecular Mechanisms of Drug Resistance in Plasmodium falciparum Malaria. *Annual Review of Microbiology* **74**, 431–454 (2020).

86. Rosenthal, M. R. & Ng, C. L. Plasmodium falciparum Artemisinin Resistance: The Effect of Heme, Protein Damage, and Parasite Cell Stress Response. *ACS Infectious Diseases* **6**,

1599–1614 (2020).

87. Yang, T. *et al.* Decreased K13 Abundance Reduces Hemoglobin Catabolism and Proteotoxic Stress, Underpinning Artemisinin Resistance. *Cell Rep.* **29**, 2917-2928.e5 (2019).

88. Woodrow, C. J. & White, N. J. The clinical impact of artemisinin resistance in Southeast Asia and the potential for future spread. *FEMS Microbiology Reviews* **41**, 34–48 (2017).

89. Association of mutations in the Plasmodium falciparum Kelch13 gene (Pf3D7_1343700) with parasite clearance rates after artemisinin-based treatments-a WWARN individual patient data meta-analysis. *BMC Med.* **17**, 1 (2019).

90. Uwimana, A. *et al.* Emergence and clonal expansion of in vitro artemisinin-resistant Plasmodium falciparum kelch13 R561H mutant parasites in Rwanda. *Nat. Med.* **26**, 1602–1608 (2020).

91. Cui, L. & Su, X. Z. Discovery, mechanisms of action and combination therapy of artemisinin. *Expert Review of Anti-Infective Therapy* **7**, 999–1013 (2009).

92. Nguyen, T. D. *et al.* Optimum population-level use of artemisinin combination therapies: A modelling study. *Lancet Glob. Heal.* **3**, e758–e766 (2015).

93. Dagnogo, O. *et al.* Towards a re-emergence of chloroquine sensitivity in Côte d'Ivoire? *Malar. J.* **17**, 413 (2018).

94. Mita, T. *et al.* Recovery of chloroquine sensitivity and low prevalence of the Plasmodium falciparum chloroquine resistance transporter gene mutation K76T following the discontinuance of chloroquine use in Malawi. *Am. J. Trop. Med. Hyg.* **68**, 413–415 (2003).

95. Al-Bari, A. A. Chloroquine analogues in drug discovery: New directions of uses,

mechanisms of actions and toxic manifestations from malaria to multifarious diseases. *J. Antimicrob. Chemother.* **70**, 1608–1621 (2014).

96. Krishna, S., Bustamante, L., Haynes, R. K. & Staines, H. M. Artemisinins: their growing importance in medicine. *Trends in Pharmacological Sciences* **29**, 520–527 (2008).

97. Nigussie, D., Beyene, T., Shah, N. A. & Belew, S. New Targets in Malaria Parasite Chemotherapy: A Review. (2015). doi:10.4172/2470-6965/1000S1-007

98. Mcconkey, G. A. Targeting the shikimate pathway in the malaria parasite Plasmodium falciparum. *Antimicrob. Agents Chemother.* **43**, 175–177 (1999).

99. Derrer, B., Macheroux, P. & Kappes, B. The shikimate pathway in apicomplexan parasites: implications for drug development. *Front. Biosci. (Landmark Ed.* **18**, 944–69 (2013).

100. Zhang, B. *et al.* A second target of the antimalarial and antibacterial agent fosmidomycin revealed by cellular metabolic profiling. *Biochemistry* **50**, 3570–3577 (2011).

101. Kokkonda, S. *et al.* Isoxazolopyrimidine-Based Inhibitors of Plasmodium falciparum Dihydroorotate Dehydrogenase with Antimalarial Activity. *ACS Omega* **3**, 9227–9240 (2018).

102. Imlay, L. S. *et al.* Plasmodium IspD (2-C-Methyl-D-erythritol 4-Phosphate Cytidyltransferase), an Essential and Druggable Antimalarial Target. *ACS Infect. Dis.* **1**, 157–167 (2015).

103. Hviid, L. & Jensen, A. T. R. PfEMP1 - a parasite protein family of key importance in plasmodium falciparum malaria immunity and pathogenesis. *Adv. Parasitol.* **88**, 51–84 (2015).

104. Belachew, E. B. Immune Response and Evasion Mechanisms of Plasmodium falciparum

Parasites. *Journal of Immunology Research* **2018**, (2018).

105. Dinko, B. & Pradel, G. Immune evasion by <i>Plasmodium falciparum</i> parasites: converting a host protection mechanism for the parasite's benefit. *Adv. Infect. Dis.* **06**, 82–95 (2016).

106. Casares, S., Brumeanu, T. D. & Richie, T. L. The RTS,S malaria vaccine. *Vaccine* **28**, 4880–4894 (2010).

107. Laurens, M. B. RTS,S/AS01 vaccine (Mosquirix^{TM}): an overview. *Hum. Vaccines Immunother.* **16**, 480–489 (2019).

108. *Malaria Vaccine Technology Roadmap*. (2006).

109. *2013 Update to the Malaria Vaccine Technology Roadmap*. (2013).

110. Moorthy, V. & Binka, F. R21/Matrix-M: a second malaria vaccine? *The Lancet* **397**, 1782–1783 (2021).

111. Datoo, M. S. *et al.* Efficacy of a low-dose candidate malaria vaccine, R21 in adjuvant Matrix-M, with seasonal administration to children in Burkina Faso: a randomised controlled trial. *Lancet* **397**, 1809–1818 (2021).

112. Roestenberg, M. *et al.* A double-blind, placebo-controlled phase 1/2a trial of the genetically attenuated malaria vaccine PfSPZ-GA1. *Sci. Transl. Med.* **12**, (2020).

113. Mordmüller, B. *et al.* Sterile protection against human malaria by chemoattenuated PfSPZ vaccine. *Nature* **542**, 445–449 (2017).

114. Molina-Franky, J. *et al.* Plasmodium falciparum pre-erythrocytic stage vaccine development. *Malaria Journal* **19**, (2020).

115. Ellis, R. D. *et al.* Phase 1 trial of the Plasmodium falciparum blood stage vaccine MSP1 42-

C1/alhydrogel with and without CPG 7909 in malaria naïve adults. *PLoS One* **5**, (2010).

116. Blank, A. *et al.* Immunization with full-length Plasmodium falciparum merozoite surface protein 1 is safe and elicits functional cytophilic antibodies in a randomized first-in-human trial. *NPJ vaccines* **5**, 10 (2020).

117. Chitnis, C. E. *et al.* Phase I clinical trial of a recombinant blood stage vaccine candidate for Plasmodium falciparum malaria based on MSP1 and EBA175. *PLoS One* **10**, (2015).

118. Chan, J. A. *et al.* Malaria vaccine candidates displayed on novel virus-like particles are immunogenic and induce transmission-blocking activity. *PLoS One* **14**, (2019).

119. Lee, S. M. *et al.* The Pfs230 N-Terminal fragment, Pfs230D1+: Expression and characterization of a potential malaria transmission-blocking vaccine candidate. *Malar. J.* **18**, 356 (2019).

120. Theisen, M., Jore, M. M. & Sauerwein, R. Towards clinical development of a Pfs48/45-based transmission blocking malaria vaccine. *Expert Review of Vaccines* **16**, 329–336 (2017).

121. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D. & Aravind, L. Evolutionary Genomics of the HAD Superfamily: Understanding the Structural Adaptations and Catalytic Diversity in a Superfamily of Phosphoesterases and Allied Enzymes. *J. Mol. Biol.* **361**, 1003–1034 (2006).

122. Koonin, E. V. & Tatusov, R. L. Computer Analysis of Bacterial Haloacid Dehalogenases Defines a Large Superfamily of Hydrolases with Diverse Specificity. *J. Mol. Biol.* **244**, 125–132 (1994).

123. Van der Ploeg, J., van Hall, G. & Janssen, D. B. Characterization of the haloacid

dehalogenase from Xanthobacter autotrophicus GJ10 and sequencing of the dhlB gene. *J. Bacteriol.* **173**, 7925–7933 (1991).

124. Weightman, A. J., Weightman, A. L. & Slater, J. H. Stereospecificity of 2-monochloropropionate dehalogenation by two dehalogenases of Pseudomonas putida PP3: evidence for two different dehalogenation mechanisms. *J. Gen. Microbiol.* **128**, 1755–1762 (1982).

125. Aravind, L., Galperin, M. Y. & Koonin, E. V. The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold. *Trends Biochem. Sci.* **23**, 127–129 (1998).

126. Park, J., Guggisberg, A. M., Odom, A. R. & Tolia, N. H. Cap-domain closure enables diverse substrate recognition by the C2-type haloacid dehalogenase-like sugar phosphatase Plasmodium falciparum HAD1. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1824–34 (2015).

127. Allen, K. N. & Dunaway-Mariano, D. Phosphoryl group transfer: Evolution of a catalytic scaffold. *Trends in Biochemical Sciences* **29**, 495–503 (2004).

128. Kuznetsova, E. *et al.* Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family. *J. Biol. Chem.* **281**, 36149–61 (2006).

129. Kuznetsova, E. *et al.* Functional Diversity of Haloacid Dehalogenase Superfamily Phosphatases from Saccharomyces cerevisiae: BIOCHEMICAL, STRUCTURAL, AND EVOLUTIONARY INSIGHTS. *J. Biol. Chem.* **290**, 18678–98 (2015).

130. Singh, S. K. *et al.* The thrH Gene Product of Pseudomonas aeruginosa Is a Dual Activity Enzyme with a Novel Phosphoserine:Homoserine Phosphotransferase Activity. *J. Biol. Chem.* **279**, (2004).

131. Kedzierski, L. *et al.* Structure of Leishmania mexicana Phosphomannomutase Highlights

Similarities with Human Isoforms. *J. Mol. Biol.* **363**, (2006).

132. Gohla, A., Birkenfeld, J. & Bokoch, G. M. Chronophin, a novel HAD-type serine protein phosphatase, regulates cofilin-dependent actin dynamics. *Nat. Cell Biol.* **7**, 21–29 (2005).

133. Gohla, A. Do metabolic HAD phosphatases moonlight as protein phosphatases? *Biochim. Biophys. Acta - Mol. Cell Res.* **1866**, 153–166 (2019).

134. Galburt, E. A., Pelletier, J., Wilson, G. & Stoddard, B. L. Structure of a tRNA repair enzyme and molecular biology workhorse: T4 polynucleotide kinase. *Structure* **10**, 1249–1260 (2002).

135. Selengut, J. D. MDP-1 is a new and distinct member of the haloacid dehalogenase family of aspartate-dependent phosphohydrolases. *Biochemistry* **40**, 12704–12711 (2001).

136. Wu, J. & Woodard, R. W. Escherichia coli YrbI is 3-deoxy-D-manno-octulosonate 8-phosphate phosphatase. *J. Biol. Chem.* **278**, 18117–18123 (2003).

137. Kenthirapalan, S., Waters, A. P., Matuschewski, K. & Kooij, T. W. A. Copper-transporting ATPase is important for malaria parasite fertility. *Mol. Microbiol.* **91**, 315–325 (2014).

138. Taylor, C. J., McRobert, L. & Baker, D. A. Disruption of a Plasmodium falciparum cyclic nucleotide phosphodiesterase gene causes aberrant gametogenesis. *Mol. Microbiol.* **69**, 110–118 (2008).

139. Carucci, D. J. *et al.* Guanylyl cyclase activity associated with putative bifunctional integral membrane proteins in Plasmodium falciparum. *J. Biol. Chem.* **275**, 22147–22156 (2000).

140. Krishna, S. *et al.* A family of cation ATPase-like molecules from Plasmodium falciparum. *J. Cell Biol.* **120**, 385–398 (1993).

141. Guggisberg, A. M. *et al.* A sugar phosphatase regulates the methylerythritol phosphate

(MEP) pathway in malaria parasites. *Nat. Commun.* **5**, 4467 (2014).

142. Dumont, L. *et al.* The metabolite repair enzyme phosphoglycolate phosphatase regulates central carbon metabolism and fosmidomycin sensitivity in plasmodium falciparum. *MBio* **10**, e02060-19 (2019).

143. Frasse, P. M. & Odom John, A. R. Haloacid Dehalogenase Proteins: Novel Mediators of Metabolic Plasticity in Plasmodium falciparum. *Microbiol. Insights* (2019). doi:DOI: 10.1177/1178636119848435

144. Guggisberg, A. M. *et al.* Suppression of Drug Resistance Reveals a Genetic Mechanism of Metabolic Plasticity in Malaria Parasites. *MBio* **9**, e01193-18 (2018).

145. Mishra, A. K. & Lambright, D. G. Invited review: Small GTPases and their GAPs. *Biopolymers* **105**, 431–448 (2016).

146. Carrique, L. *et al.* Structure and catalytic regulation of Plasmodium falciparum IMP specific nucleotidase. *Nat. Commun.* **11**, 1–11 (2020).

Chapter 2: HAD2, an inorganic phosphate-sensitive phosphatase with specificity toward triose phosphates

Preface

Below is a compilation of data from both the publication Guggisberg et al., 2018 in *mBio*[1], as well as subsequent and parallel findings that have not been formulated into a complete publication yet. Much of the data presented in the *mBio* publication was performed before my time and is considered background information, and it is therefore omitted from this thesis chapter. This chapter therefore focuses less on the original discovery of HAD2 as a mediator of FSM-resistance and more on attempts to understand the mechanism of that phenomenon and the cellular function of HAD2. All data presented here was produced by myself except for metabolomics and parasite growth in phosphate-depleted media, which were performed by Dr. Simon Cobbold.

2.1 Abstract

HAD2 is a *Plasmodium falciparum* enzyme of the haloacid dehalogenase (HAD) superfamily of proteins. Premature stop codon mutations in HAD2 confer resistance to the inhibitor of parasite isoprenoid biosynthesis, fosmidomycin (FSM), at the expense of diminished growth. To understand how loss of HAD2 confers FSM resistance, we characterize the biochemical and metabolic role of HAD2 in parasites. Using whole-parasite metabolic profiling and biochemical assays of recombinant HAD2, we provide evidence that HAD2 is a small metabolite phosphatase, with some specificity for triose phosphates and fructose 1,6-bisphosphate. We also explore the role that serine phosphorylation plays in regulation of HAD2 activity and substrate specificity. Finally, we demonstrate that HAD2 is highly sensitive to free

phosphate inhibition, which may indicate a role for HAD2 in regulating intracellular phosphate levels for the parasite.

2.2 Introduction

2.2.1 The identification of HAD2, and its complex relationship with glycolysis[1]

Malaria is caused by infection with unicellular eukaryotic parasites of the genus *Plasmodium*. The species *Plasmodium falciparum* is responsible for most life-threatening malarial disease. As an obligate intracellular parasite of human erythrocytes, *Plasmodium falciparum* has unique metabolic features that may be exploited to discover new drug targets and develop new therapies. In the red blood cell niche, *Plasmodium* parasites are highly dependent on glucose metabolism. Infection with *Plasmodium* spp. results in a nearly 100-fold increase in glucose import in red blood cells[2–4]. Despite these energy requirements, the parasite demonstrates little aerobic respiration via the tricarboxylic acid (TCA) cycle. Instead, it relies on anaerobic glycolysis to produce ATP[5–8].

Besides ATP production, glucose also has a number of anabolic fates in *P. falciparum*. One such fate is the synthesis of isoprenoids. Isoprenoids are a large class of hydrocarbons with extensive structural and functional diversity[9]. In the malaria parasite, isoprenoids perform several important functions, including protein prenylation, dolichylation, and synthesis of GPI anchors[10–12]. Despite this diversity, all isoprenoids are synthesized from a common five-carbon building block, isopentyl pyrophosphate (IPP). Evolution has produced two distinct routes for IPP synthesis: the mevalonate pathway, found in archaea, fungi, animals, and the cytoplasm of plants; and the methylerythritol phosphate (MEP) pathway, found in most eubacteria, plant chloroplasts, and apicomplexan parasites such as *P. falciparum*[13]. Because it is both essential for

the parasite and absent from the human host, the MEP pathway is a compelling target for antimalarial development. The antibiotic and antimalarial fosmidomycin (FSM) is a competitive inhibitor of the first committed enzymatic step of the MEP pathway, catalyzed by 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR; EC 1.1.1.267)[14–16]. FSM has been validated as a specific inhibitor of the MEP pathway in *P. falciparum*[17]and is a valuable chemical tool to study MEP pathway biology and essential metabolism in the parasite. In this study, we found that FSM is also a useful tool for probing glycolytic metabolism upstream of the essential MEP pathway.

Parasites are likely to control the proportion of glucose used for energy production versus production of secondary metabolites, such as isoprenoids. We previously used a screen for FSM resistance to identify HAD1, a metabolic regulator whose loss results in increased levels of MEP pathway intermediates and resistance to MEP pathway inhibition. HAD1 is a cytoplasmic sugar phosphatase that dephosphorylates a number of sugar phosphate intermediates upstream of the MEP pathway[18,19]. Hypomorphic mutation of HAD1 allows buildup of those upstream intermediates, culminating in increased availability of 1-deoxy-D-xylulose-5-phosphate (DOXP). This increased substrate availability outcompetes the inhibitor FSM to confer resistance.

HAD1 belongs to the haloacid dehalogenase-like hydrolase (HAD) enzyme superfamily and, more specifically, to the IIB and Cof-like hydrolase subfamilies[20]. While HADs are found in all kingdoms of life, HAD1 is most closely related to bacterial members of this superfamily[18,21], which have been implicated in metabolic regulation, stress response, and phosphate homeostasis[22–26]. However, most members of this superfamily remain uncharacterized.

We next described the discovery of HAD2 (PF3D7_1226300), a second HAD family member in *P. falciparum*[1]. We found that HAD2 is a cytosolic phosphatase required for metabolic homeostasis. Loss of HAD2 by acquiring a premature stop codon, R157X,

dysregulates glycolysis and misroutes metabolites toward the MEP pathway, conferring drug resistance. In our study, we harnessed a fitness defect in *HAD2*^{R157X} parasite strains to employ an innovative screen for suppression of drug resistance in the parasite. Selection for suppression of drug resistance identified mutations in *PFK9*, which encodes the canonical glycolytic regulatory enzyme phosphofructokinase (PFK). Reduction in PFK9 activity rescued the metabolic dysregulation in our resistant mutants and restored FSM sensitivity. Our unique approach thus revealed PFK9 as a site of exceptional metabolic plasticity in the parasite and uncovers a novel genetic mechanism by which *P. falciparum* malaria parasites may adapt to metabolic stress and drug selective pressure.

2.2.2 Hypotheses on HAD2 function and regulatory mechanisms

Having identified HAD2 as a small molecule phosphatase with some promiscuity in its substrate use[1], we sought to further characterize how HAD2 functions in a cellular environment, what its *in vivo* substrate(s) is/are, and how its activity is regulated. HAD2 is highly homologous to HAD1, as it is also a member of the IIB and Cof-like hydrolase subfamilies of HAD proteins[20]. Because HAD1 is a sugar phosphatase, and unpublished metabolomic work of HAD1-overexpressing parasites indicates that its primary natural substrate is glycerol 1-phosphate, we hypothesized that HAD2 would also have sugar phosphatase activity. In this work, we demonstrate that HAD2 has promiscuous activity against several small phosphometabolites, and observe a preference toward triose phosphates.

One distinction between HAD2 and HAD1 is the proteomic description of two phosphorylated serine residues (Ser4 and Ser9) near the N-terminus of HAD2[27]. The phosphorylation by kinases and dephosphorylation by phosphatases of proteins have long been known as mechanisms of enzymatic regulation[28,29]. These phosphorylated serines therefore may

regulate the enzymatic function of HAD2. Understanding this regulation, and which kinase(s) or phosphatase(s) carry it out, may be a useful tool in future efforts to therapeutically target the MEP pathway. We therefore investigate these two serine residues in HAD2 by generating recombinant enzyme that ablates phosphorylation with serine-to-alanine mutation, or enzyme that mimics serine phosphorylation with serine-to-aspartate mutation. We demonstrate that phosphorylation of Ser4 and Ser9 may confer subtle increases to HAD2 phosphatase activity and modest changes in substrate specificity.

Finally, we demonstrate that HAD2 is highly sensitive to inhibition by physiologically relevant levels of inorganic phosphate (P_i). While this has been seen in other phosphatases, including HAD1 and HAD4[30], the high sensitivity of HAD2 to low concentrations of free phosphate suggests that this may be a pivotal aspect of the cellular function of HAD2 in parasites.

2.3 Results

2.3.1 HAD2 is a functional phosphometabolite phosphatase.

We have previously established that *P. falciparum* HAD1 is a promiscuous sugar phosphatase, with activity against a wide range of phosphometabolites. Similarly, *P. vivax* HAD2 (PVX_123945) has been enzymatically characterized and found to possess phosphatase activity against various monophosphorylated substrates, including glycerol 2-phosphate (Glc2P) and pyridoxal-5'-phosphate (PLP)[31]. Recombinant *Pv*HAD2 also utilizes additional monophosphorylated substrates, such as adenosine monophosphate (AMP) and glycerol 1-phosphate (Glc1P), with moderate activity.

On the basis of the previous characterization of a close *Plasmodium* homolog, as well as sequence homology to HAD1 and other HAD proteins, we predicted that *Pf*HAD2 would also function enzymatically as a phosphatase. We successfully purified recombinant *Pf*HAD2 in *Escherichia coli* and confirmed the phosphatase activity of recombinant *Pf*HAD2 using *para*-nitrophenyl phosphate (*p*NPP), a promiscuous, chromogenic phosphosubstrate (Fig. 1A)[21,32,33]. Because *E. coli* expresses a number of HAD-like phosphatases[21], we confirmed that the phosphatase activity was specific to purified *Pf*HAD2 by expression and purification of a catalytically inactive mutant (HAD2^{D26A}). The Asp26 residue was chosen for mutagenesis because the corresponding residue in *Pf*HAD1 (Asp27) has been previously shown to be required for catalysis[19].

We also established the activity of *Pf*HAD2 against a panel of phosphorylated substrates and determined that its substrate profile closely mirrors that of *Pv*HAD2 (Fig. 1B). Overall, we found that *Pf*HAD2 is a phosphatase with activity against small phosphosubstrates, such as Glc2P. These data suggest that, like HAD1 and related HADs in microbes and plants[21,33,34], HAD2 is a phosphatase with the potential to utilize a variety of monophosphorylated phosphometabolites.

2.3.2 HAD2 enzymatic activity is strongly inhibited by free phosphate

We have previously seen that HAD1 (as-yet unpublished) and HAD4[30] are both inhibited by high concentrations of inorganic phosphate (P_i) in the reaction buffer, acting as a feedback-inhibitory mechanism to regulate enzyme function. To test whether HAD2 was also subject to P_i inhibition, we tested HAD2 activity in our *p*NPP assay with varying concentrations of P_i. Changing the concentration of the *p*NPP substrate along with the concentration of P_i, we found that HAD2 is competitively inhibited by free phosphate (Fig. 2A). We observed that HAD1 was

inhibited by free phosphate with a half-maximal inhibitory concentration (IC_{50}) of ~5 mM. By contrast, we found that HAD2 was 5- to 10-fold more sensitive to inorganic phosphate inhibition than HAD1 (Fig. 2B), with a mean ± SEM IC_{50} of 0.46 ± 0.15 mM. Notably, parasites' intracellular phosphate concentration is linked to that of the extracellular media, equilibrating in a matter of hours (Dr. Simon Cobbold, personal communication). Therefore, under typical culturing conditions, and in plasma, in which [P_i] equals 1-5 mM, we would expect the majority of HAD2 activity to be inhibited, as it is 2-10x the enzyme's IC_{50}. Furthermore, parasites exhibit diminished growth when grown in low concentrations of P_i, and HAD2 mutant parasites are hypersensitive to this (Fig. 2C). This suggests that the metabolic role of HAD2 could be linked to maintaining the balance of phosphate in the parasites, so that loss of HAD2 dysregulated phosphate homeostasis.

2.3.3 Phosphorylation of HAD2 may modestly enhance enzymatic function

To probe the question of how phosphorylation affects enzymatic function of HAD2, we generated recombinant HAD2 mutant enzymes that could mimic either an unphosphorylated or phosphorylated state, in addition to the wild-type (WT) sequence. We predict that the WT enzyme is likely in its unphosphorylated state, as the necessary protein kinase is likely absent from the *E. coli* cells used to generate the enzyme. However, to ensure an unphosphorylated state, we generated a mutant in which both Ser4 and Ser9 are mutated to alanine, termed "S4AS9A". In contrast, due to its structural resemblance to phosphorylated serine, aspartic acid has previously been used to mimic the phosphorylated state of enzymes[35,36]. We therefore also generated a double-phospho-mimetic mutant in which both Ser4 and Ser9 are replaced with aspartic acid residues, termed "S4DS9D". These two enzymes were then compared to each other and to WT for differences in enzymatic activity.

When testing these enzymes for phosphatase activity against the robust substrate, Glc2P, we find that WT and S4AS9A enzymes had similar activities, while the S4DS9D enzyme exhibited a modest but significant increase in activity (Fig. 3). One caveat to this approach is that comparisons between different enzyme preparations can yield differing results, depending on the relative purification of the enzymes. Other metrics that are less dependent on exact concentration of active enzyme, such as Michaelis-Menten kinetics and relative activity against other substrates, may be more informative than this approach.

To further evaluate the potential regulatory role of serine phosphorylation on HAD2, we reasoned that perhaps inorganic phosphate inhibition of HAD2 was dependent upon serine phosphorylation. We therefore evaluated our recombinant S4AS9A and S4DS9D enzymes for inorganic phosphate inhibition. As shown in Fig. 2D, however, there was no statistically significant difference in phosphate inhibition between the WT enzyme (presumed to be unphosphorylated), the phospho-ablated, and the phospho-mimetic. We conclude that inorganic phosphate inhibition of HAD2 is independent of phosphorylation status of the enzyme.

2.3.4 Metabolomic Analysis of HAD2 mutant parasites

To investigate the mechanism by which HAD2 mutation confers FSM resistance, as well as the natural function of HAD2, we collaborated with Drs. Malcolm McConville and Simon Cobbold at the University of Melbourne to perform metabolic profiling on our parasites. This profiling was performed with the FSM-resistant parasites with HAD2^{R157X} mutations from our resistance screen, termed “E2-D2”. Additionally, we profiled parasites in which FSM had been removed and suppressor mutations arose in *PFK9*, termed “E2-D2S”, signifying that they are descended from E2-D2, but are again FSM-sensitive. Finally, the E2-D2S parasites were complemented with a copy of HAD2-GFP, under a constitutive HSP110 promoter. These

parasites are termed “E2-A8”. Parasites were fed ^{13}C-glucose over a period of 0-30 minutes, with time points taken to measure metabolic uptake of ^{13}C in a variety of metabolites using LC-MS, comparing the pool size of each metabolite to that of wild-type 3D7 parasites. As shown in Figure 4, dramatic changes were seen in E2-D2 parasites compared to wild-type. Some of the most striking changes occurred in nucleotide phosphates, which were decreased in E2-D2 parasites, and increases in certain glycolytic and pentose phosphate pathway metabolites, such as fructose 1,6-bisphosphate and the bisphosphates of both octulose and sedoheptulose. Based on these data from HAD2^{R157X} parasites, we hypothesized that perhaps HAD2 dephosphorylates sugar phosphates, as has been shown for HAD1, and somehow promotes the synthesis of nucleotide phosphates. One caveat to these parasites, however, is that HAD2 mutation is not the sole polymorphism difference between E2-D2 parasites and wild-type, so other polymorphisms in the genome might account for some of these metabolic changes.

We next investigated E2-D2S parasites, which had developed suppressor mutations in *PFK9* compared to E2-D2. These parasites had mostly inverted metabolic signatures compared to E2-D2 parasites, in regard to nucleotide phosphates and the pentose phosphate pathway. Unfortunately, as PFK9 is such an important regulator of carbon metabolism as a whole, it was difficult to fit a model to the effects we observed.

Finally, E2-A8 parasites were metabolically profiled. Because these parasites were transgenically generated from E2-D2S parasites, rather than evolved in a screen, the comparison between E2-A8 and E2-D2S is the cleanest comparison we currently have available to assess HAD2 metabolic effects. While there are not many differences between these strains, one notable exception is again seen with octulose 1,8-bisphosphate (OBP) and sedoheptulose 1,7-bisphosphate (SBP). Both of these metabolites were elevated in the original E2-D2 strain, and

are decreased upon complementation with HAD2-GFP in the E2-A8 strain. This suggests that these metabolites could be the natural substrates for HAD2.

2.3.5 HAD2 has a preference for triose phosphates *in vitro*

Because the substrate profile in Figure 1 was limited and guided mostly by the findings from the *P. vivax* homolog, we sought to expand our substrate profile analysis of HAD2 to more phospho-metabolites. Furthermore, while Glc2P is the substrate with highest activity, it is not a naturally occurring metabolite, so the natural substrate(s) of HAD2 in parasites remained undetermined. Moreover, our metabolic data revealed specific substrates of interest that could potentially serve as HAD2 substrates, which needed to be validated *in vitro*. We therefore tested the enzymatic activity of WT HAD2 against a wide panel of potential substrates (Fig. 5). Again, it is clear that the substrate for which HAD2 has highest activity is Glc2P. While no single natural substrate stands out above the rest, and indeed HAD2 may act upon multiple substrates within parasites, the trend that emerges is the preference for triose phosphates. After Glc2P, the next substrates with highest activity are fructose bisphosphate (FBP) racemic (D/L) glyceraldehyde-3-phosphate (Gly3P) and D-Gly3P. Glycerol-1-phosphate (Glc1P) is also among the substrates with relatively higher activity, as is dihydroxyacetone phosphate (DHAP). Although FBP is an interesting outlier, the trend we observe is that the binding pocket for HAD2 is appears tuned to accommodate triose phosphates. Unlike the *P. vivax* homolog, which has specific activity toward pyridoxal-5'-phosphate (PLP; Vitamin B6) that is comparable to Glc2P [31], *Pf*HAD2 had far less activity toward PLP, at roughly 20% the activity of Glc2P. This discrepancy highlights the difficulty in predicting substrate specificity of HAD proteins through homology alone.

Examining the phosphatase activity of our phosphomutant enzymes against the same panel of substrates, we identified some substrates that were differentially utilized between WT, S4AS9A, and S4DS9D enzymes (Fig. 5). For instance, the S4DS9D enzyme could more robustly dephosphorylate PLP and deoxyuridine triphosphate (dUTP), whereas the S4AS9A mutant was much less active against fructose 6-phosphate (F6P). Although this data is preliminary, and requires replication, it raises the possibility that phosphorylation of HAD2 alters its substrate specificity. Based on the crystal structure of *Pv*HAD2, we would expect Ser4 and Ser9 to be relatively distal from the HAD2 active site, suggesting that serine phosphorylation does not result in direct contact with substrate, but rather may cause some form of conformational change that changes the binding pocket allosterically. It is also interesting to note that while we hypothesized the S4AS9A enzyme would mimic WT, this is not universally true, indicating that the hydroxyl group of the serine residue is important for HAD2 function, even before it becomes phosphorylated.

Interestingly, many of the triose phosphates we observed *in vitro*, namely Gly3P and Glc1P, showed very little change upon HAD2 loss in the metabolomic data. One exception to this is DHAP which was elevated in E2-D2 strains, and decreased upon HAD2-GFP complementation. DHAP is rapidly interchanged with Gly3P by triose phosphate isomerase, so perhaps this is the source of the discrepancy[37]. Furthermore, this method of metabolic profiling measures the steady state pool levels of metabolites and is not suited to detect flux through a pathway. Thus, if these metabolites are quickly funneled to other fates, their buildup may not be detected.

We also tested HAD2 activity toward nucleotide phosphates. Despite the broad effect of HAD2 mutation on nucleotide phosphates (Fig. 4), HAD2 did not have robust phosphatase

activity upon these, indicating that the nucleotide effect of HAD2 is indirect. Although they are omitted from Figure 5, guanosine monophosphate (GMP) and cytidine triphosphate (CTP), were also tested with WT HAD2, yielding activity of <0.015 μmol/min/mg. Finally, our most promising metabolomic candidates, OBP and SBP, are not readily accessible for purchase from any chemical vendors, so they could not be directly tested in our assays. To test the phosphatase activity of HAD2 against OBP and SBP, we instead developed a crude assay to synthesize these substrates *in vitro*, incubating aldolase with DHAP and ribose 5-phosphate (R5P) to generate OBP, or DHAP and erythrose 4-phosphate (E4P) to generate SBP. Shown in Figure 6, we then measured the phosphatase activity of HAD2 upon the incubated mixture. When compared to phosphate release of the reactants without aldolase, there was no significant increase in HAD2 phosphatase activity when the full reaction was in place (Fig. 6). This suggests that HAD2 activity against OBP and SBP is no higher than its activity against DHAP, E4P, and R5P. However, it should be noted that we were not able to control the concentration of OBP or SBP present, nor even confirm that these compounds were successfully generated. To conclusively test the hypothesis of HAD2 dephosphorylating SBP and OBP, confirmation of those compounds' presence will be necessary.

2.3.6 Genetic manipulation of HAD2

Because of the inherent complexity of our metabolomic work due to the multitude of genetic differences between *HAD2*-null mutants and their parental strains, we sought a cleaner genetic approach to evaluating HAD2 function in parasites. By generating parasites whose only difference from parental strains was the presence or absence of HAD2, particularly in an inducible manner, we can be more confident that metabolic differences we observe are due to HAD2, either directly or indirectly, rather than polymorphisms elsewhere in the genome.

Furthermore, while we already have had success rescuing *had2, pfk9* mutant parasites with a *HAD2-GFP* construct (i.e. the E2-A8 strain), we were concerned that the underlying *pfk9* mutations might mask or further complicate any metabolic changes observed by the addition of *HAD2-GFP* due to the key role PFK9 plays in regulation of all of carbon metabolism[38].

Therefore, the first genetic method we used was to transfect an extra copy of pTEOE110: HAD2-GFP into a wild-type background, using the same transposase-mediated method as in the complementation parasites E2-A8. This resulted in successfully acquired parasites with exogenous HAD2-GFP expressed cytosolically (Fig. 7). These parasites have yet to be characterized metabolically.

There are several disadvantages to using the pTEOE/piggyBac system. First, as a transposase-mediated system, there is no control over the locus of insertion, meaning that in the absence of multiple transfections, the metabolic or phenotypic observations cannot be confidently attributed to the extra copy of HAD2 without the consideration that an endogenous gene's disruption is the source of the phenotype. Secondly, depending on the relationship between endogenous HAD2 activity and its substrate availability, addition of more HAD2 may not have much effect if the system is already saturated with HAD2 activity. Finally, because this method of transfection is static rather than inducible, the long culturing period before attaining transgenic parasites allows compensatory mutations to arise before parasites are even acquired.

We therefore decided to generate an inducible knockdown of HAD2 at the endogenous locus using the previously described anhydrotetracycline (aTc)-repressible system[39,40]. Using CRISPR/Cas9, we inserted a 3xHA tag and a 10x aptamer array to the C-terminus of the endogenous HAD2 gene, so that translation of HAD2 could be repressed in the absence of aTc and permitted in the presence of aTc (Fig. 8A). We successfully acquired transfectants of this

system, calling this strain “PF4”, and immunoblotting for HA demonstrates that removal of aTc results in modest reduction of HAD2 expression (Fig. 8B).

To confirm whether this modest reduction of HAD2 was sufficient to phenocopy the *HAD2*^{R157X} parasites originally obtained in FSM-resistance screens, we evaluated these parasites for their FSM resistance and their growth rate. As seen in Figure 8C and 8D, respectively, the PF4 parasites had minimal difference in growth and FSM resistance when comparing +aTc conditions to -aTc conditions (note that while it may appear in Fig. 8C that a difference in growth is emerging over time, carrying this out for longer did not reveal a true difference in growth; not shown). The FSM half-maximal effective concentration (EC_{50}) did increase by 1.2-fold, but this is significantly less than the 5-fold increase observed in *HAD2*^{R157X} parasites from our original screen[1].

2.4 Discussion

We have reported that, similar to the homologous enzyme *Pf*HAD1, mutation to *Pf*HAD2 confers fosmidomycin resistance and dramatic changes to parasite metabolism. To better understand how parasites regulate metabolism and evolve FSM resistance, we sought to uncover the mechanism by which HAD2 mutation confers FSM resistance by elucidating its natural role in the cell and how its function might be regulated by the parasite and its environment.

While several hypotheses arose from metabolomic profiling of E2-D2, E2-D2S, and E2-A8 parasites, perhaps most promising of which was the dephosphorylation of SBP and OBP, these hypotheses could not be validated by *in vitro* biochemical assays, and the metabolomic data is complicated by the complex genetic background of these parasites. By contrast, enzymatic data suggests that triose phosphates are among the most readily dephosphorylated by

HAD2. The dephosphorylation of triose phosphates by HAD2 does align with our model for how its homolog, HAD1 functions. HAD1 mutants have elevated glycerol 1-phosphate (Glc1P) which may be easily oxidized to Gly3P, the direct precursor to DOXP, to allow elevated DOXP levels and FSM resistance. As shown in Figure 5, *Pf*HAD2 also dephosphorylates Glc1P, like HAD1, so this may be the mechanism for HAD2 as well. Alternatively, or in addition, one of the highest naturally-occurring HAD2 substrates is Gly3P itself. The dephosphorylation of Gly3P would likely directly impair the synthesis of DOXP, explaining how loss of HAD2 promotes elevated levels of DOXP and subsequent FSM resistance. Additionally, FBP is another substrate upon which HAD2 has high activity, and loss of HAD2 results in elevated FBP. The fact that suppressor mutations in the E2-D2S parasites arose in PFK9, which produces FBP, indicates that this may be a true function of HAD2 as well. As HAD proteins are notorious for their substrate promiscuity, any or all of these might be true substrates of HAD2 within the parasite.

These models are complicated by our findings that the phosphatase activity of HAD2 is strongly inhibited by free phosphate. Because this activity is so strongly inhibited by cellularly relevant concentrations of phosphate, it may be that phosphatase activity is not the cellular role of HAD2. Perhaps another reaction, such as phosphotransferase activity, is less inhibited by free phosphate. Unfortunately, this is a difficult reaction to assay *in vitro*, as the correct phosphate-donor and acceptor would need to be chosen to assay. Alternatively, some promiscuous phosphatases are known to catalyze sulfate transfer activity as well[41]. Metabolic profiling may be able to assist in teasing out these possible alternative reactions for HAD2, highlighting the usefulness of our HAD2-GFP-overexpressing parasites.

This also highlights the need for an inducible knockdown system. Unfortunately, our inducible knockdown was unable to phenocopy HAD2^{R157X} parasites obtained in FSM resistance

screens. We hypothesize that this is due to the incomplete knockdown, as there is still detectable HAD2 present in the parasites, and previous work by others has demonstrated the precedent for certain *P. falciparum* proteins being expressed in great excess, requiring robust knockdown to observe a phenotype[42]. We therefore propose that a more substantial inducible knockout/knockdown will be an invaluable tool moving forward in our understanding of HAD2, and suggest future work that implements a Di-Cre system for inducible knockout or a more complete knockdown with a TetR-DOZI system that introduces RNA aptamers both upstream and downstream of the coding sequence. With a robust knockdown, in addition to the HAD2-GFP overexpressing parasites, metabolomic studies may be able to conclusively ascertain the metabolic role of HAD2 in parasites.

If these metabolic studies confirmed a role for HAD2 dephosphorylating triose phosphates or FBP, or identified other substrates, an open question would remain as to the biological purpose of this activity. Our observation that inorganic phosphate inhibits HAD2 activity offers a possible answer. Wild-type parasites require ~5 mM inorganic phosphate to grow optimally, and are sensitive to decreases in the phosphate concentration, while *HAD2*^{R157X} parasites are hypersensitized to this drop in phosphate (Fig 3C). Perhaps HAD2 serves as a regulator of the intracellular phosphate concentration, and utilizes triose phosphates and FBP as the source of phosphate release. This may explain why HAD2 would have such robust capacity for dephosphorylation despite existing in a prohibitively high concentration of phosphate. Intracellular phosphate homeostasis is critical for countless purposes, including the energy needs of the cell and biosynthesis of macromolecules[43–47], which might explain why dysregulation of this system by loss of HAD2 leads to growth defects in parasites.

Finally, the effect that serine phosphorylation of HAD2 has on its function remains unclear. We showed by our phospho-mimetic mutants that phosphorylation of both Ser4 and Ser9 may confer a slight (~20%) increase in phosphatase activity, which is relatively subtle compared to other enzymes in which phosphorylation canonically confers a more dramatic shift in on/off status[48,49]. Additionally, phosphomutant enzymes appeared to have a slightly altered substrate specificity, notably increasing the activity of HAD2 toward PLP, to better align with the *P. vivax* HAD2. Perhaps phosphorylation of HAD2 is a response to availability of phosphate sources in the cell, allowing the enzyme to utilize readily available phosphometabolites to maintain phosphate homeostasis. Alternatively, or perhaps in conjunction, perhaps phosphorylation controls the localization of HAD2, another common function of enzyme phosphorylation[50,51], to provide access to different substrates. Future studies investigating the role of phosphorylation on HAD2 may benefit from rescuing HAD2-knockdown parasites with our phospho-ablated or phospho-mimetic constructs to detect phosphorylation-dependent differences in localization and metabolomics in parasites. Protein crystallization and analysis of the three-dimensional structure of HAD2's binding pocket, and how it might change upon phosphorylation, may also be valuable in understanding how phosphorylation controls HAD2 function.

2.5 Methods

2.5.1 Parasite strains and culture.

All parasites used were of the 3D7 parental strain with the exception of PF4, which was in an NF54^{attB} background. Unless otherwise indicated, parasites were maintained at 37°C in 5% O_2–5% CO_2–90% N_2 in a 2% suspension of human erythrocytes in RPMI medium (Sigma-

Aldrich) modified with 27 mM $NaHCO_3$, 11 mM glucose, 5 mM HEPES, 0.01 mM thymidine, 1 mM sodium pyruvate, 0.37 mM hypoxanthine, 10 µg/ml gentamicin, and 5 g/liter Albumax (Thermo Fisher Scientific).

2.5.2 Plasmodium growth measurement

Parasitemia in daily growth assays was measured via flow cytometry by incubating 10 µL of parasite culture with 190 µL of 0.4 µg/mL acridine orange (Invitrogen) in PBS for 1 minute. Stained parasites were analyzed on a BD FACS Canto flow cytometer gating on DNA and RNA-bound dye signal using FITC and PerCP-Cy5.5 filters, respectively. 50,000 events were recorded for each sample.

2.5.3 Generation of recombinant HAD2.

The predicted coding sequence of *HAD2* was amplified using the HAD2_LIC_F and HAD2_LIC_R primers (Table 1). A catalytic mutant (D26A) and phospho-mutants (S4AS9A and S4DS9D) were also generated. The *had2*^{D26A} allele was created using the HAD2_D26A_F and HAD2_D26A_R site-directed mutagenesis primers (Table 1). The S4AS9A allele was generated with PMF003 and PMF007, and the S4DS9D allele was generated with PMF006 and PMF007.

Ligation-independent cloning was used to clone *HAD2* (WT and mutants) into vector BG1861[52], which introduces an N-terminal 6×His fusion into the expressed protein. BG1861:6×His-HAD2 was transformed into One Shot BL21(DE3)pLysS *Escherichia coli* cells (Thermo Fisher Scientific). Protein expression was induced for 2.5-3 h with 1 mM isopropyl-β-D-thiogalactoside at mid-log phase (optical density at 600 nm [OD_{600}] of 0.4 to 0.5). Cells were collected by centrifugation and stored at −20°C.

Cells were lysed in buffer containing 1 mg/ml lysozyme, 20 mM imidazole, 1 mM dithiothreitol, 1 mM $MgCl_2$ 10 mM Tris HCl (pH 7.5), 30 U benzonase (EMD Millipore), and cOmplete EDTA-free protease inhibitor tablets (Roche). 6×His-HAD2 was bound to nickel agarose beads (Gold Biotechnology), washed with a mixture containing 20 mM imidazole, 20 mM Tris HCl (pH 7.5), and 150 mM NaCl, and eluted in a mixture containing 300 mM imidazole, 20 mM Tris HCl (pH 7.5), and 150 mM NaCl. This eluate was further purified by size exclusion gel chromatography using a HiLoad 16/600 Superdex 200-pg column (GE Healthcare) equilibrated in a mixture containing 25 mM Tris HCl (pH 7.5), 250 mM NaCl, and 1 mM $MgCl_2$. The elution fractions containing HAD2 were pooled and concentrated, and glycerol was added to reach a concentration of 10% (wt/vol). Protein solutions were immediately flash frozen and stored at −80°C.

2.5.4 HAD2 activity assays.

The rate of *para*-nitrophenyl phosphate (*p*NPP; Sigma-Aldrich S0942) hydrolysis by HAD2 was determined by continuous measurement of absorbance at 405 nm. Assays were performed at 37°C in a 50-µl volume consisting of 50 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 10 mM *p*NPP, and 1.2 µM enzyme. Hydrolysis of other phosphorylated substrates by HAD2 was measured using an EnzChek phosphate assay kit (Life Technologies). The reaction buffer was modified to contain 50 mM Tris-HCl (pH 8.0), 20 mM $MgCl_2$, 0.2 mM 2-amino-6-mercapto-7-methylpurine riboside (MESG), and 1 U/ml purine nucleoside phosphorylase (PNP). Reaction mixtures contained 5 mM substrate and 730 nM enzyme. The activity of catalytically inactive 6×His-$HAD2^{D26A}$ was measured for all substrates, and the data were used to normalize the activity found for the WT HAD2 enzyme. Activity was normalized to that obtained from

catalytically inactive 6×His-HAD2^{D26A}. All data represent means of results from ≥3 independent experiments performed with technical replicates.

2.5.4 pTEOE110:HAD2 and pSN054:HAD2 plasmid construction and transfection

The pTEOE110 construct contains the heat shock protein 110 gene (PF3D7_0708800) 5′ untranscribed region (UTR) and a C-terminal GFP tag[18]. Human dihydrofolate reductase (hDHFR) is present as a selectable marker. Inverted terminal repeats are included for genome integration by a cotransfected piggyBac transposase (pHTH, MRA-912 from MR4; ATCC, Manassas, VA). *HAD2* was amplified with the HAD2_XhoI_F and HAD2_AvrII_R primers (Table 1) and cloned into AvrII and XhoI sites in the pTEOE110 plasmid.

The HAD2 conditional knockdown strain, "PF4", was generated by transfecting NF54^{attB} parasites[53] using previously described methods[39,40]. CRISPR/Cas9 was used to edit the endogenous HAD2 locus, adding a 3xHA tag to the C-terminus of the gene, followed by a 10x aptamer array, a blasticidin (BSD) resistance marker, and the TetR-DOZI fusion protein. Primers PMF140 and PMF141 were used to amplify the 5' homologous region (HR) from parasite gDNA, and primers PMF142 and PMF143 were used to amplify the 3' HR. Resultant parasites were maintained in the presence of anhydrotetracycline (aTc; Cayman Chemicals) in DMSO at 500nM unless otherwise specified.

Transfections were performed as previously described[18]. Briefly, 50 to 100 μg of plasmid DNA was precipitated and resuspended in Cytomix (25 mM HEPES [pH 7.6], 120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 10 mM K_2HPO_4). A ring-stage *P. falciparum* culture was washed with Cytomix and resuspended in the DNA/Cytomix solution. Cells were electroporated using a Bio-Rad Gene Pulser II electroporator at 950 μF and 0.31 kV.

Electroporated cells were washed with media and returned to normal culture conditions. Parasites expressing the construct were selected by continuous treatment with 5 nM WR92210 for the pTEOE plasmid (Jacobus Pharmaceuticals) or with 2.5 μg/mL BSD (Fisher) for the pSN054 plasmid. Transfectants were cloned by limiting dilution, and the presence of the HAD2-GFP construct was verified by PCR using gene- and GFP-specific primers (HAD2_R157X_F and GFP_R; Table 1). Maintenance of the endogenous *HAD2* and *PFK9* genotypes was verified by Sanger sequencing.

2.5.5 Microscopy

Fluorescent microscopy images were taken using a Zeiss Axio Observer D1 (Carl Zeiss Inc.) at the Washington University Molecular Microbiology Imaging Facility. Images were acquired with a Plan-Apochromat 100x (NA1.4) objective using the Zen 2.3 pro (blue edition) software.

2.5.6 Quantification of FSM resistance

Asynchronous *P. falciparum* cultures were diluted to 1% parasitemia and were treated with FSM (Life Technologies) in water at concentrations up to 500 μM. Growth inhibition assays were performed in opaque 96-well plates at 100 μL culture volume. After 3 days, parasite growth was quantified by measuring DNA content using Quant-iT PicoGreen dsDNA reagent (Thermofisher), as previously described[54]. Picogreen fluorescence was measured on a Perkin Elmer multimode Envision microplate reader. EC_{50} values were calculated by nonlinear regression analysis using GraphPad Prism software.

2.6 Figures

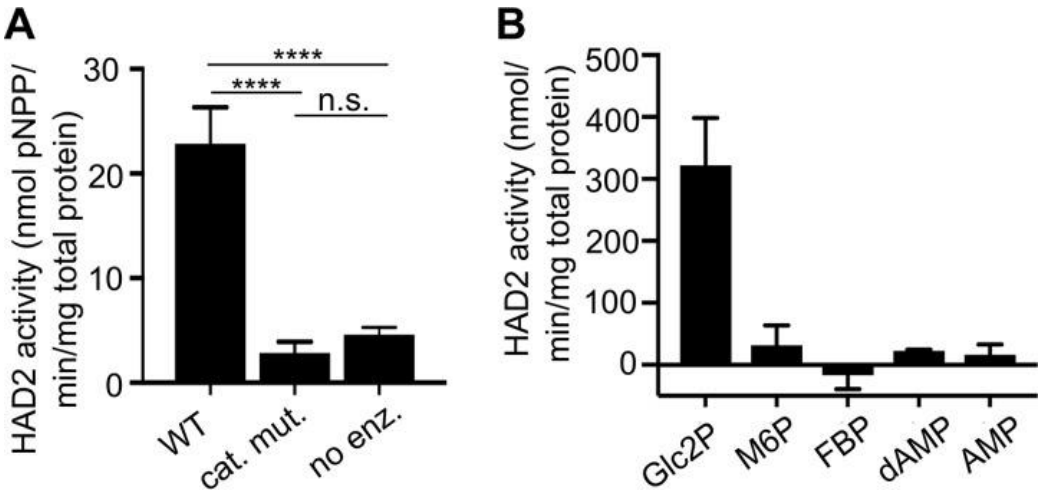


Figure 1. *Pf*HAD2 is a small molecule phosphatase.

(A) HAD2 is an active phosphatase, and HAD2^{D26A} is a catalytic mutant (cat. mut.) that can be used as a negative control for HAD2-specific activity. "no enz." represents a no-enzyme control. Data shown represent the enzyme activities seen using the synthetic phosphatase substrate *p*NPP. Error bars represent standard errors of the means (SEM) (****, $P \leq 0.0001$ [unpaired t test]; n.s., not significant). (B) Activity of HAD2, normalized to the activity of the catalytic mutant (HAD2^{D26A}), for a variety of substrates (Glc2P, glycerol 2-phosphate; M6P, mannose 6-phosphate; FBP, fructose 2,6-bisphosphate; dAMP, deoxy-AMP).

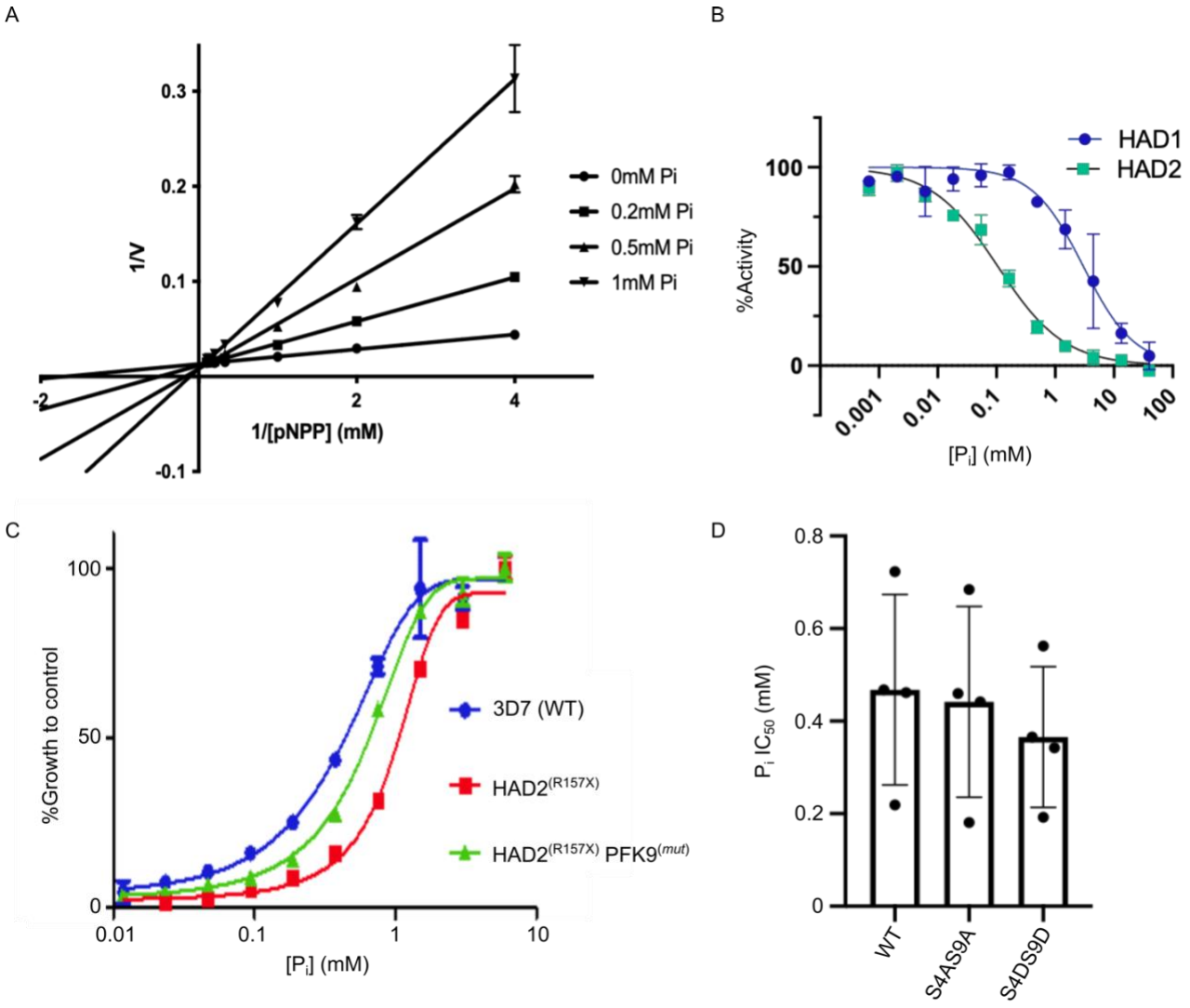


Figure 2. HAD2 is inhibited by inorganic phosphate.

(A) Lineweaver-Burke plot of HAD2 inhibition by free inorganic phosphate (P_i). The pattern of a fixed V_{max} with variable x-intercept (K_m) is consistent with competitive inhibition. **(B)** Inorganic phosphate IC_{50} of HAD2 enzymatic activity compared to HAD1. **(C)** Parasite sensitivity to low media phosphate. **(D)** Comparison of wild-type (WT), phospho-ablated (S4AS9A), and phospho-mimetic (S4DS9D) HAD2 IC_{50} for inorganic phosphate. No statistically significant differences, using ordinary one-way ANOVA and Tukey's test for multiple comparisons. All data represent 2-4 independent experiments, each with technical replicates.

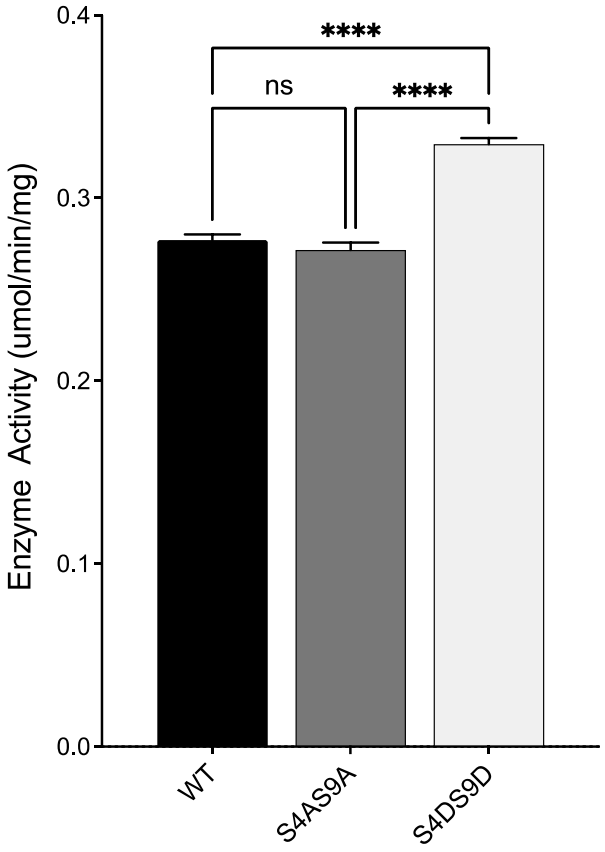


Figure 3: Comparison of S4AS9A and S4DS9D.

Wild-type, S4AS9A, and S4DS9D recombinant enzyme was tested for activity against glycerol-2-phosphate (Glc2P) in the EnzChek phosphate release assay. Statistics were performed using an ordinary one-way ANOVA test with Tukey's test for multiple comparisons. Data represent the means ± SEM of 3 technical replicates of the same enzyme preparations.

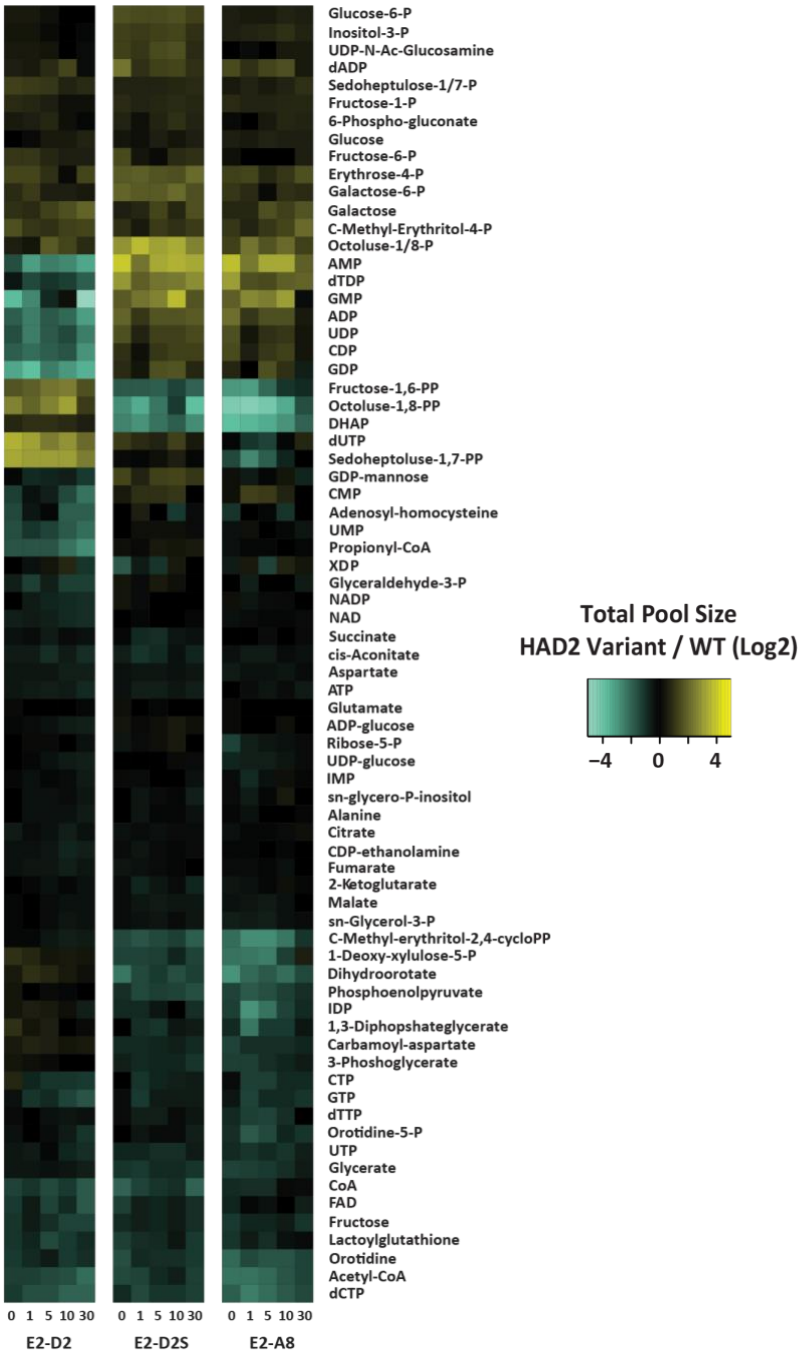


Figure 4: Metabolic profiling of HAD2 mutant parasites.

HAD2-null, PFK9 suppressor mutant, and HAD2-GFP-complemented parasites were fed ^{13}C-labeled glucose, and metabolites were extracted and analyzed by LC-MS at time points of 0, 1, 5, 10, and 30 minutes. All colors are relative to a WT control.

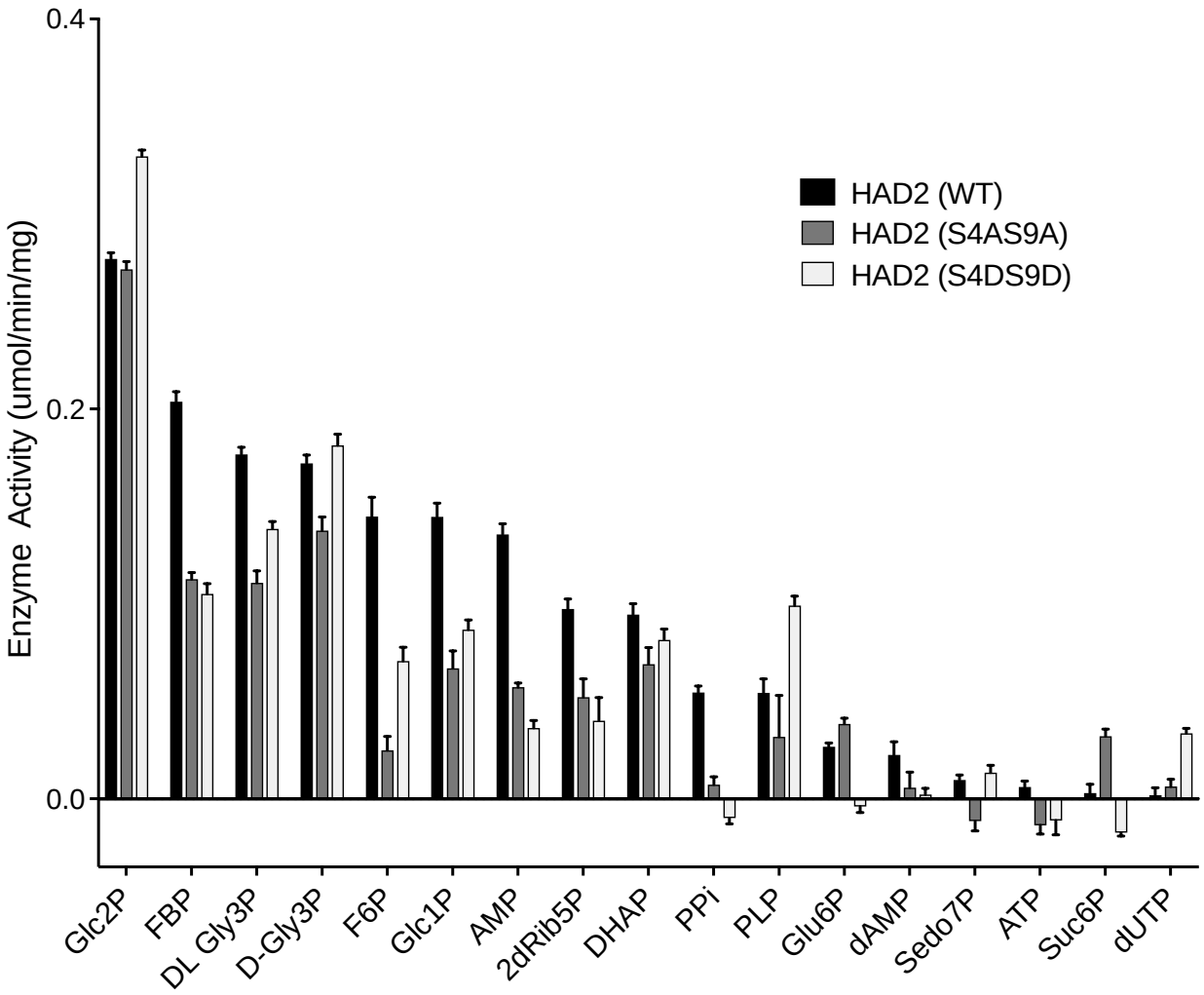


Figure 5: Phosphatase activity of HAD2 on a wide substrate panel.

Phosphatase activity on substrates at 1 mM concentration, was performed using the 50 µL EnzChek phosphate release assay. Data represent mean ± SEM of three replicate experiments. Abbreviations: Glc2P, glycerol 2-phosphate; FBP, fructose-1,6-bisphosphate; DL Gly3P, (D/L)-glyceraldehyde 3-phosphate; D-Gly3P, (D)-glyceraldehyde 3-phosphate; F6P, fructose 6-phosphate; Glc1P, glycerol 1-phosphate; AMP, adenosine monophosphate; 2dRib5P, 2-deoxyribose 5-phosphate; DHAP, dihydroxyacetone phosphate; PPi, inorganic pyrophosphate; PLP, pyridoxal-5'-phosphate; Glu6P, glucose 6-phosphate; dAMP, deoxyadenosine

monophosphate; Sedo7P, sedoheptulose 7-phosphate; ATP, adenosine triphosphate; Suc6P, sucrose 6-phosphate; dUTP, deoxyuridine triphosphate.

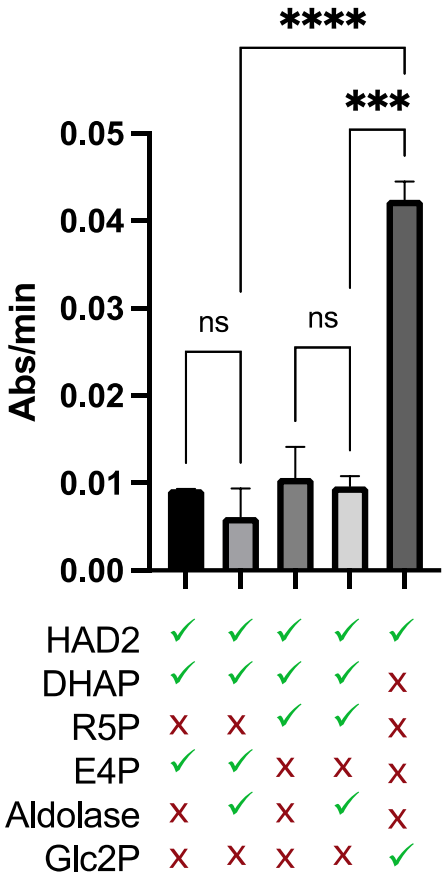


Figure 6: Phosphatase activity of HAD2 upon SBP and OBP.

Using a the EnzChek phosphate release assay, HAD2 was incubated with either DHAP and E4P or DHAP and R5P, either in the presence or absence of pre-incubation with aldolase to generate SBP or OBP, respectively. Glc2P was used as a positive control for HAD2 activity. Data represents two independent experiments, each performed in triplicate. Statistics were performed with a one-way ANOVA. ***p=0.0001, ****p<00001, ns = not significant.

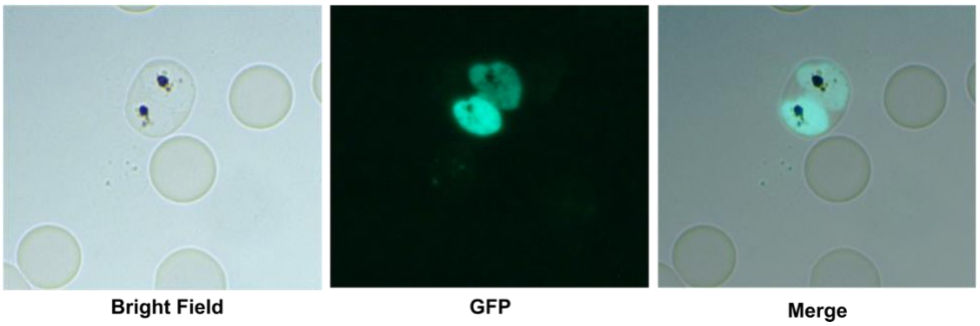


Figure 7: Successful generation of HAD2-GFP overexpressing parasites in 3D7.

Bright field, GFP channel, and merged images of transgenic parasites that were transfected with HAD2-GFP under an HSP110 promoter using the piggyBac transposase.

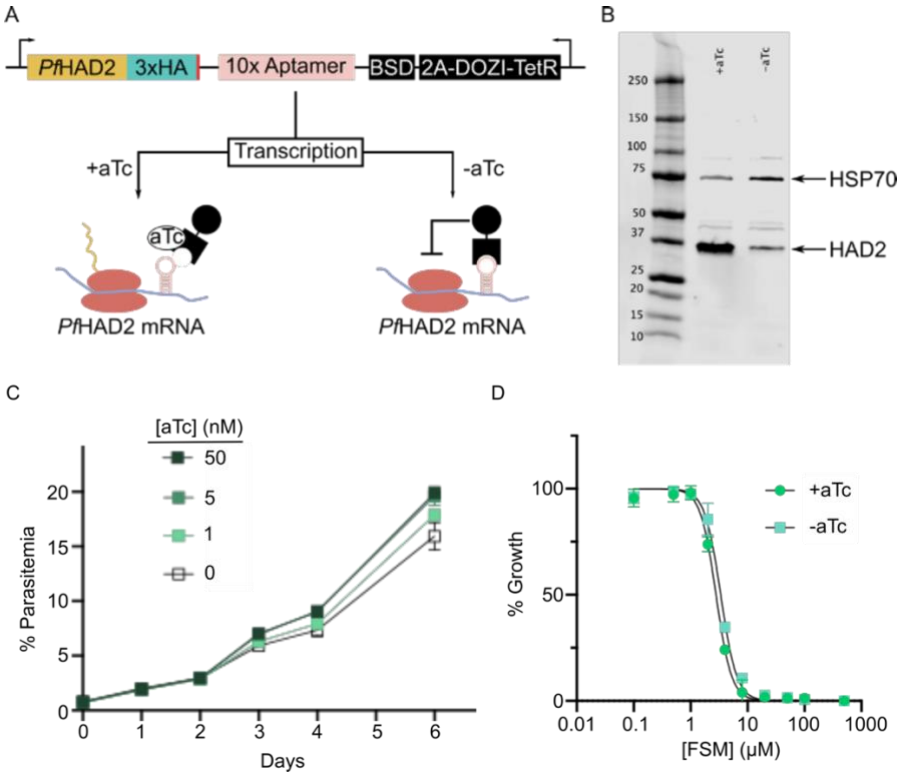


Figure 8: Partial knockdown of HAD2 in WT background.

(A) Scheme of TetR-DOZI regulatable system **(B)** Western blot of HAD2 partial knockdown, with HSP70 loading control. **(C)** Growth of PF4 parasites upon aTc removal, measured by

acridine orange staining and flow cytometry. Data represent 3 replicates ± SEM. **(D)** 72-hr dose-response of FSM on PF4 parasites ±aTc. +atc EC_{50} = 2.82 ± 0.06 µM; -aTc EC_{50} = 3.37 ± 0.06 µM. Data represent means of 3 experiments ± SEM.

2.7 Tables

Table 1. Primers used in this study.

Primer name	Purpose	Sequence (5'->3')
HAD2_R157X_F	Genotyping *had2*^{R157X} allele, detection of HAD2-GFP construct	AGGATATACTTTATTAGATGAGAC
HAD2_R157X_R	Genotyping *had2*^{R157X} allele	ATTATATGTTGTAAATATGGTCAATT
PFK9_F	Amplifying *PFK9*	CTCACCACCACCACCACCATATGGATACC AAGAGTGGAGATAAAA
PFK9_R	Amplifying *PFK9*	ATCCTATCTTACTCACTTAGTTCATTCTTT TTCTCTGGTTTTC
PFK9_seq_1	Sequencing *PFK9*	CACAGGTAATGAATTCCCAGC
PFK9_seq_2	Sequencing *PFK9*	ATCATCGGCATTCTGACATAAC
PFK9_seq_3	Sequencing *PFK9*	GATCTATGCGTTTTGAACAATTAG
PFK9_seq_4	Sequencing *PFK9*	CCTGTTAAGCCACTTTCAATAAC
PFK9_seq_5	Sequencing *PFK9*	CCCAGTTGATTCTCTATCCATTAA
PFK9_seq_6	Sequencing *PFK9*	GGTCCATTTGATGCTTCGAAAC
PFK9_seq_7	Sequencing *PFK9*	CGTTGTTATACTGTAAAGACTCCA
PFK9_seq_8	Sequencing *PFK9*	CTAGTCCCATTTCTTTTGAAATAAGAA
HAD2_LIC_F	Cloning *HAD2* into vector BG1861	CTCACCACCACCACCACCATATGGCTTCT AGTAACGATGTACA

HAD2_LIC_R	Cloning *HAD2* into vector BG1861	ATCCTATCTTACTCACTTATTTTTTTTTCAAGTCAAATACTTTTTTTAA
HAD2_D26A_F	Site-directed mutagenesis of HAD2	GTTATTATTAATTGCTTTTGATGGTACAT
HAD2_D26A_R	Site-directed mutagenesis of HAD2	ATGTACCATCAAAAGCAATTAATAATAAC
DXS_F	qPCR of *DXS* (Pf3D7_1337200) mRNA	AACGTGGATAAAGTACACATTGC
DXS_R	qPCR of *DXS* (Pf3D7_1337200) mRNA	TGATATACTTACGGCATTTGTTGG
DXR_F	qPCR of *DXR* (PF3D7_1467300) mRNA	ACATGGCCTGATAGAATAAAAACA
DXR_R	qPCR of *DXR* (PF3D7_1467300) mRNA	TTCATTTGACGCATTTAGTACAGTT
betatub_F	qPCR of *beta tubulin* (PF3D7_1008700) mRNA	ATCCCATTCCCACGTTTACATT
betatub_R	qPCR of *beta tubulin* (PF3D7_1008700) mRNA	TCCTTTGTGGACATTCTTCCTC
18S_F	qPCR of *18S rRNA* (PF3D7_0112300)	GAACGAGGAATGCCTAGTAAGCA
18S_R	qPCR of *18S rRNA* (PF3D7_0112300)	TTCATCATATCTTTCAATCGGTAGGA
HAD2_XhoI_F	Cloning *HAD2* into vector pTEOE110	GATCCTCGAGATGGCTTCTAGTAACGATGTACATT

HAD2_AvrII_R	Cloning *HAD2* into vector pTEOE110	GATCCCTAGGTTTTTTTTTCAAGTCAAATACTTTTTTTAATAAG
GFP_R	Detection of HAD2-GFP construct	CCGTATGTTGCATCACCTTC
PMF140	Cloning 5' HR of HAD2 into pSN054	TTGAAGAAAATCCAGGTCCAGCGATAGAGGTGAATCGAATTATGTAACTGAA
PMF141	Cloning 5' HR of HAD2 into pSN054	AGTCTGGAACGTCATAAGGGTAGGCTTTTTTTTTCAAGTCAAATACTTTTTTTAATAAGTATGCGACAGCTCTATCTTTATGT
PMF142	Cloning 3' HR of HAD2 into pSN054	ATTCGAGCTCGGATTACCCTGTTATCGTGTCTTGTATTACTTCCATATGAT
PMF143	Cloning 3' HR of HAD2 into pSN054	GAGAGATTGGGTATTAGACCTAGGGGTAGAATTGTGAGAAATGGGTTACACT
PMF003	Cloning HAD2 (S4AS9A) into BG1861	CTCACCACCACCACCACCATATGGCTTCTGCTAACGATGTACATGCTGA
PMF006	Cloning HAD2 (S4DS9D) into BG1861	CTCACCACCACCACCACCATATGGCTTCTGATAACGATGTACATGATGA
PMF007	Cloning HAD2 (S4AS9A/S4DS9D) into BG1861	ATCCTATCTTACTCACTTATTTTTTTTCAAGTCAAATACTTTTTTTAA

2.8 References

1. Guggisberg, A. M. *et al.* Suppression of Drug Resistance Reveals a Genetic Mechanism of Metabolic Plasticity in Malaria Parasites. *MBio* **9**, e01193-18 (2018).
2. Mehta, M., Sonawat, H. M. & Shobhona, S. Glycolysis in Plasmodium falciparum results in modulation of host enzyme activities. *J Vector Borne Dis* **43**, 95–103 (2006).
3. Roth, E. Plasmodium falciparum carbohydrate metabolism: a connection between host cell and parasite. *Blood Cells* **16**, 453–60; discussion 461-6 (1990).
4. Roth, E. F. Malarial parasite hexokinase and hexokinase-dependent glutathione reduction in the Plasmodium falciparum-infected human erythrocyte. *J Biol Chem* **262**, 15678–15682 (1987).
5. Bowman, I. B. R., Grant, P. T., Kermack, W. O. & Ogston, D. The metabolism of Plasmodium berghei, the malaria parasite of rodents. 2. An effect of mepacrine on the metabolism of glucose by the parasite separated from its host cell*. *Biochem. J.* **78**, 472–478 (1961).
6. Krungkrai, J., Burat, D., Kudan, S., Krungkrai, S. & Prapunwattana, P. Mitochondrial oxygen consumption in asexual and sexual blood stages of the human malarial parasite, Plasmodium falciparum - PubMed. *Southeast Asian J Trop Med Public Heal.* **30**, 636–642 (1999).
7. MacRae, J. I. *et al.* Mitochondrial metabolism of sexual and asexual blood stages of the malaria parasite Plasmodium falciparum. *BMC Biol.* **11**, (2013).
8. Scheibel, L. & Miller, J. Glycolytic and cytochrome oxidase activity in Plasmodia - PubMed. *Mil Med* **134**, 1074–1080 (1969).
9. Gershenzon, J. & Dudareva, N. The function of terpene natural products in the natural

world. *Nat. Chem. Biol.* **3**, 408–414 (2007).

10. Guggisberg, A. M., Amthor, R. E. & Odom, A. R. Isoprenoid Biosynthesis in Plasmodium falciparum. *Eukaryot. Cell* **13**, 1348–1359 (2014).
11. Imlay, L. & Odom, A. R. Isoprenoid Metabolism in Apicomplexan Parasites. *Current Clinical Microbiology Reports* **1**, 37–50 (2014).
12. Jordão, F. M., Kimura, E. A. & Katzin, A. M. Isoprenoid biosynthesis in the erythrocytic stages of Plasmodium falciparum. *Memorias do Instituto Oswaldo Cruz* **106**, 134–141 (2011).
13. Lange, B. M., Rujan, T., Martin, W. & Croteau, R. Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13172–13177 (2000).
14. Koppisch, A. T., Fox, D. T., Blagg, B. S. J. & Poulter, C. D. E. coli MEP synthase: Steady-state kinetic analysis and substrate binding. *Biochemistry* **41**, 236–243 (2002).
15. Kuzuyama, T., Shimizu, T., Takahashi, S. & Seto, H. Fosmidomycin, a specific inhibitor of 1-deoxy-D-xylulose 5-phosphate reductoisomerase in the nonmevalonate pathway for terpenoid biosynthesis. *Tetrahedron Lett.* **39**, 7913–7916 (1998).
16. Steinbacher, S. *et al.* Structural basis of fosmidomycin action revealed by the complex with 2-C-methyl-D-erythritol 4-phosphate synthase (IspC). Implications for the catalytic mechanism and anti-malaria drug development. *J. Biol. Chem.* **278**, 18401–18407 (2003).
17. Zhang, B. *et al.* A second target of the antimalarial and antibacterial agent fosmidomycin revealed by cellular metabolic profiling. *Biochemistry* **50**, 3570–3577 (2011).
18. Guggisberg, A. M. *et al.* A sugar phosphatase regulates the methylerythritol phosphate (MEP) pathway in malaria parasites. *Nat. Commun.* **5**, 4467 (2014).

19. Park, J., Guggisberg, A. M., Odom, A. R. & Tolia, N. H. Cap-domain closure enables diverse substrate recognition by the C2-type haloacid dehalogenase-like sugar phosphatase Plasmodium falciparum HAD1. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1824–34 (2015).

20. Hunter, S. *et al.* InterPro in 2011: New developments in the family and domain prediction database. *Nucleic Acids Res.* **40**, (2012).

21. Kuznetsova, E. *et al.* Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family. *J. Biol. Chem.* **281**, 36149–61 (2006).

22. Kang, M. J. *et al.* Identification of genes affecting lycopene accumulation in Escherichia coli using a shot-gun method. *Biotechnol. Bioeng.* **91**, 636–642 (2005).

23. Kang, Y., Weber, K. D., Qiu, Y., Kiley, P. J. & Blattner, F. R. Genome-wide expression analysis indicates that FNR of Escherichia coli K-12 regulates a large number of genes of unknown function. *J. Bacteriol.* **187**, 1135–1160 (2005).

24. Gawand, P. & Mahadevan, R. Engineering Escherichia coli for D-ribose production from glucose-xylose mixtures. *Ind. Biotechnol.* **10**, 106–114 (2014).

25. Roberts, A., Lee, S. Y., McCullagh, E., Silversmith, R. E. & Wemmer, D. E. YbiV from Escherichia coli K12 is a HAD phosphatase. *Proteins Struct. Funct. Genet.* **58**, 790–801 (2005).

26. Sun, Y. & Vanderpool, C. K. Physiological consequences of multiple-target regulation by the small RNA SgrS in escherichia coli. *J. Bacteriol.* **195**, 4804–4815 (2013).

27. Lasonder, E. *et al.* The *Plasmodium falciparum* Schizont Phosphoproteome Reveals Extensive Phosphatidylinositol and cAMP-Protein Kinase A Signaling. *J. Proteome Res.* **11**, 5323–5337 (2012).

28. Hunter, T. Signaling-2000 and Beyond. *Cell* **100**, 113–127 (2000).

29. Krebs, E. G. & Beavo, J. A. Phosphorylation-dephosphorylation of enzymes. *Annual review of biochemistry* **48**, 923–959 (1979).

30. Cobbold, S. A. *et al.* Non-canonical metabolic pathways in the malaria parasite detected by isotope-tracing metabolomics. *Mol. Syst. Biol.* **17**, (2021).

31. Srinivasan, B., Kempaiah Nagappa, L., Shukla, A. & Balaram, H. Prediction of substrate specificity and preliminary kinetic characterization of the hypothetical protein PVX_123945 from Plasmodium vivax. *Exp. Parasitol.* **151**–**152**, 56–63 (2015).

32. Kuznetsova, E. *et al.* Enzyme genomics: Application of general enzymatic screens to discover new enzymes. *FEMS Microbiology Reviews* **29**, 263–279 (2005).

33. Kuznetsova, E. *et al.* Functional Diversity of Haloacid Dehalogenase Superfamily Phosphatases from Saccharomyces cerevisiae: BIOCHEMICAL, STRUCTURAL, AND EVOLUTIONARY INSIGHTS. *J. Biol. Chem.* **290**, 18678–98 (2015).

34. Caparrós-Martín, J. A., McCarthy-Suárez, I. & Culiáñez-Macià, F. A. HAD hydrolase function unveiled by substrate screening: Enzymatic characterization of Arabidopsis thaliana subclass I phosphosugar phosphatase AtSgpp. *Planta* **237**, 943–954 (2013).

35. Ueda, E. K. M. *et al.* A molecular mimic of phosphorylated prolactin (S179D PRL) secreted by eukaryotic cells has a conformation with an increased positive surface charge compared to that of unmodified prolactin. *Biochemistry* **48**, 6887–6897 (2009).

36. Salman, E. D., He, D., Runge-Morris, M., Kocarek, T. A. & Falany, C. N. Site-directed mutagenesis of human cytosolic sulfotransferase (SULT) 2B1b to phospho-mimetic Ser348Asp results in an isoform with increased catalytic activity. *J. Steroid Biochem. Mol. Biol.* **127**, 315–323 (2011).

37. Bandyopadhyay, D., Murthy, M. R. N., Balaram, H. & Balaram, P. Probing the role of highly conserved residues in triosephosphate isomerase – analysis of site specific mutants at positions 64 and 75 in the Plasmodial enzyme. *FEBS J.* **282**, 3863–3882 (2015).

38. UNDERWOOD, A. H. & NEWSHOLME, E. A. PROPERTIES OF PHOSPHOFRUCTOKINASE FROM RAT LIVER AND THEIR RELATION TO THE CONTROL OF GLYCOLYSIS AND GLUCONEOGENESIS. *Biochem. J.* **95**, 868–875 (1965).

39. Ganesan, S. M., Falla, A., Goldfless, S. J., Nasamu, A. S. & Niles, J. C. Synthetic RNA-protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. *Nat. Commun.* **7**, 10727 (2016).

40. Nasamu, A. S. *et al.* Plasmepsins IX and X are essential and druggable mediators of malaria parasite egress and invasion. *Science (80-.).* **358**, 518–522 (2017).

41. Pabis, A., Duarte, F. & Kamerlin, S. C. L. Promiscuity in the Enzymatic Catalysis of Phosphate and Sulfate Transfer. *Biochemistry* **55**, 3061–3081 (2016).

42. Polino, A. J., Nasamu, A. S., Niles, J. C. & Goldberg, D. E. Assessment of Biological Role and Insight into Druggability of the Plasmodium falciparum Protease Plasmepsin V. *ACS Infect. Dis.* **6**, 738–746 (2020).

43. Eskes, E., Deprez, M. A., Wilms, T. & Winderickx, J. pH homeostasis in yeast; the phosphate perspective. *Current Genetics* **64**, 155–161 (2018).

44. Fabiańska, I., Bucher, M. & Häusler, R. E. Intracellular phosphate homeostasis – A short way from metabolism to signaling. *Plant Science* **286**, 57–67 (2019).

45. Dick, C. F., Dos-Santos, A. L. A. & Meyer-Fernandes, J. R. Inorganic phosphate uptake in unicellular eukaryotes. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 2123–2127 (2014).

46. Saliba, K. J. *et al.* Sodium-dependent uptake of inorganic phosphate by the intracellular malaria parasite. *Nat. 2006 4437111* **443**, 582–585 (2006).

47. Lee, P., Ye, Z., Dyke, K. Van & Kirk, R. G. X-Ray Microanalysis of Plasmodium Falciparum and Infected Red Blood Cells: Effects of Qinghaosu and Chloroquine on Potassium, Sodium, and Phosphorus Composition. *Am. J. Trop. Med. Hyg.* **39**, 157–165 (1988).

48. Goyal, R., Rathert, P., Laser, H., Gowher, H. & Jeltsch, A. Phosphorylation of Serine-515 Activates the Mammalian Maintenance Methyltransferase Dnmt1. *http://dx.doi.org/10.4161/epi.2.3.4768* **2**, 155–160 (2007).

49. Agius, L. Role of glycogen phosphorylase in liver glycogen metabolism. *Mol. Aspects Med.* **46**, 34–45 (2015).

50. Ptacek, J. & Snyder, M. Charging it up: global analysis of protein phosphorylation. *Trends Genet.* **22**, 545–554 (2006).

51. Nishi, H., Hashimoto, K. & Panchenko, A. R. Phosphorylation in protein-protein binding: effect on stability and function. *Structure* **19**, 1807 (2011).

52. Alexandrov, A. *et al.* A Facile Method for High-throughput Co-expression of Protein Pairs. *Mol. Cell. Proteomics* **3**, 934–938 (2004).

53. Nkrumah, L. J. *et al.* Efficient site-specific integration in Plasmodium falciparum chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat. Methods* **3**, 615–21 (2006).

54. COLEY, P. D. *et al.* A NOVEL DNA-BASED MICROFLUORIMETRIC METHOD TO EVALUATE ANTIMALARIAL DRUG ACTIVITY. *Am. J. Trop. Med. Hyg.* **70**, 119–124 (2004).

Chapter 3: Metabolic signatures of the dispensable HAD proteins, HAD4 and Lipin

Preface

The following work was incorporated into a larger collaborative study published in *Molecular Systems Biology*, authored by Simon A Cobbold, Madel V Tutor*, Philip M Frasse*, Emma McHugh, Markus Karnthaler, Darren J Creek, Audrey Odom John, Leann Tilley, Stuart A Ralph, and Malcolm J McConville (*authors contributed equally)[1]. The purposes of the study were to understand metabolic fates of different carbon sources in *P. falciparum* and to compile a full draft metabolome of the parasites. The work utilized HAD4 as a proof-of-principle that metabolic changes seen in cultured parasites could be validated by *in vitro* enzymatic assays. Another HAD protein, Lipin, was also investigated, demonstrating further utility of the metabolic labeling approach. While the manuscript examines several conditions and enzymes, this dissertation chapter will focus on the methods, results, and discussion of the sections specific to HAD proteins. These sections have been altered slightly to focus on HAD proteins and provide additional discussion of how this HAD data fits into the framework of this thesis. PMF conducted the recombinant protein purification and biochemical assays for HAD4, gene knockout was performed by SAC and MVT. Mass Spectrometry by SAC and DJC, mutant phenotype analysis by SAC, PMF, AOJ, EM, and MK.

3.1 Abstract

The malaria parasite, *Plasmodium falciparum*, proliferates rapidly in human erythrocytes by actively scavenging multiple carbon sources and essential nutrients from its host cell. However, a global overview of the metabolic capacity of intraerythrocytic stages is missing. Using multiplex ^{13}C-labelling coupled with untargeted mass spectrometry and unsupervised isotopologue grouping, we have generated a draft metabolome of *P. falciparum* and its host erythrocyte consisting of 911 and 577 metabolites, respectively, corresponding to 41% of metabolites and over 70% of the metabolic reaction predicted from the parasite genome. An additional 89 metabolites and 92 reactions were identified that were not predicted from genomic reconstructions, with the largest group being associated with metabolite damage-repair systems. Validation of the draft metabolome revealed four previously uncharacterized enzymes which impact isoprenoid biosynthesis, lipid homeostasis and mitochondrial metabolism and are necessary for parasite development and proliferation, two of which are haloacid dehalogenase (HAD) proteins. This study defines the metabolic fate of multiple carbon sources in *P. falciparum*, and highlights the activity of metabolite repair pathways in these rapidly growing parasite stages, opening new avenues for drug discovery. The metabolic roles of two HAD proteins are demonstrated and discussed.

3.2 Introduction

3.2.1 Introduction to metabolic labeling

Considerable progress has been made in reducing the incidence of malaria over the last decade, although the decline in malaria cases has stalled in recent years and resistance to frontline antimalarials is on the rise[2]. Identifying new antimalarials with novel targets therefore remains a priority, and significant investment has been made in expanding the drug development pipeline with novel classes of antimalarials[3,4]. Metabolic enzymes and metabolite transporters are direct or indirect targets of most of the existing antimalarials and current lead compounds[5,6]. However, the total number of enzymes/transporters that have been rigorously validated as drug targets remains small. A detailed understanding of the metabolism of the different developmental stages of the malaria parasite, *Plasmodium falciparum*, and the host cells within which they live is therefore necessary for informing the development of new antimalarial therapies.

Plasmodium falciparum progresses through a number of different developmental stages during its life cycle in the *Anopheles* mosquito and its human host. Infection in humans is initiated by infective sporozoites that develop asymptomatically in the liver, resulting in the release of thousands of merozoites that initiate repeated cycles of infection and asexual replication in erythrocytes [i.e., red blood cells (RBCs)] that cause the clinical symptoms associated with malaria. The *P. falciparum* intraerythrocytic developmental cycle (IDC) takes approximately 48 hr and involves the development of the metabolically active trophozoite and schizont stages, followed by cell division of individual parasites into 16–32 new merozoites. The massive expansion of parasite biomass during development is fueled by the uptake and catabolism of glucose, as well as a number of other essential nutrients (e.g. amino acids, purines and vitamins) that are either directly scavenged from the RBC or derived from breakdown of

RBC hemoglobin and other proteins[7–10]. Considerable progress has been made in delineating key salvage and metabolic pathways involved in *P. falciparum* asexual development, which has formed the basis for genome-scale models of parasite metabolism[11–14]. Despite these advances, > 40% of the protein-encoding genome remains unannotated and a significant fraction of annotated metabolic genes have yet to be assigned to specific metabolic pathways or reactions. In the phylum Apicomplexa, many genes have also been repurposed to fulfil non-canonical functions, impeding genomic reconstructions and a systematic understanding of the total metabolic capacity of the pathogen[15,16]. Finally, enzyme promiscuity and side reactions can result in the production of unanticipated and novel metabolites that can have important roles in regulating cellular metabolism[17–19], further complicating predictions of enzyme function based on gene homology.

Defining the observable metabolic capacity of key developmental stages of *P. falciparum* and its host cell is required to verify the accuracy of genomic reconstructions and to identify unexpected metabolic pathways and gene functions. A number of approaches have been used to undertake a global analysis of the metabolic capacity of other organisms. For example, a system-wide reverse genetics approach was used to identify the metabolic function or indirect metabolic impact of each gene within *E. coli*[20]. The emergence of genome-wide disruption libraries in *P. falciparum* and *P. berghei* makes this approach theoretically possible[21,22]. However, these studies have highlighted the essentiality of many metabolic enzymes in *Plasmodium* spp., limiting the effectiveness of this approach. Even viable but slow-growing mutants generated through such approaches are likely to be difficult to compare directly to parental parasites at a metabolic level. The converse approach – acquiring untargeted mass spectrometry data and verifying the "observed" metabolome – has not yet been fully exploited

because of the lack of well-established pipelines for data filtering and metabolite identification. Current liquid chromatography–mass spectrometry platforms allow detection of > 10,000 mass-to-charge (*m/z*) features, yet a significant majority (> 90%) of these features correspond to background noise or degeneracy[23–25]. The absence of autonomous methods for controlling the false discovery rate has hampered the compilation of an accurate metabolome for most organisms to date.

Here we use stable-isotope resolved metabolomics to prioritize *m/z* features corresponding to metabolites actively synthesized by *P. falciparum* or the host RBC[26,27]. Previous work has demonstrated the ability of this approach to define the extent of active ^{13}C-glucose metabolism in RBCs[28], and here we expand this approach to ten biologically relevant ^{13}C-substrates in *P. falciparum*-infected RBCs. Filtering for actively-labelled metabolites enabled > 95% of *m/z* features to be removed, and the remaining *m/z* features were then identified and the active metabolome defined. This approach led to the identification of 577 metabolites in uninfected human RBCs and 911 metabolites in *P. falciparum*-infected RBCs corresponding to 41% coverage across the predicted metabolome of *P. falciparum* (the summation of all expected metabolites from all known pathways inferred from a genomic reconstruction irrespective of enzyme gaps). The pattern of stable isotope labelling for each metabolite allowed us to further infer metabolic reactions corresponding to 70.5% coverage across predicted reactions in *P. falciparum*, with the mismatch between metabolite and reaction coverage largely due to a subset of metabolites participating in many reactions. Defining the "observed" metabolome without constraining the results to the expected composition inferred from genomic reconstructions revealed 89 metabolites and 92 reactions not predicted from genomic reconstructions. These studies have highlighted unanticipated complexity

in *P. falciparum* metabolism, including the presence of active metabolite damage and repair systems in rapidly dividing parasite stages.

3.2.2 The HAD proteins examined with this approach

We wanted to explore the role of two uncharacterized haloacid dehalogenase (HAD) phosphatases, given the important role this protein family plays in regulating metabolic flux and metabolite repair. We targeted PF3D7_1118400 (annotated as HAD4) and PF3D7_0303200 (annotated as a member of the HAD superfamily, hereon referred to as Lipin) for inducible disruption. These enzymes were chosen to validate the metabolic labeling approach with a parallel *in vitro* biochemical approach, as well as demonstrate the utility of this method of isotope-tracing metabolomics. It should be noted that Lipin contains a HAD-*like* Superfamily domain (IPR036412) and therefore is not included in the phylogenetic tree shown in Chapter 1, Figure 1 of this dissertation, which is strictly limited to proteins with HAD Superfamily domains (IPR023214).

3.3 Results

Endogenous loci of both HAD proteins, HAD4 and Lipin, were targeted via Cas9-mediated double-stranded break towards the center of each gene and the 3'-half replaced with a recodonized sequence flanked by two loxP sites[29]. The 3'-end was also HA-tagged and a glmS ribozyme introduced in the 3'-UTR. Transfection of this construct into a 3D7 parasite line expressing an integrated copy of the dimerizable cre-recombinase then enabled inducible excision of the gene upon rapamycin addition, or transcript degradation when glucosamine was added. Transfections were performed with linearized rescue DNA template and integration confirmed by PCR (Fig. S1). HA-tagged proteins of the expected size were detected and were

efficiently depleted following the addition of 100 nM rapamycin and 2.5 mM glucosamine (Fig 1A). A transgenic line with lactate dehydrogenase 1 (LDH1) under inducible disruption was used as a positive control. Disruption of Lipin significantly impaired parasite growth, whereas loss of HAD4 had no effect (Fig 1B). Loss of LDH1 led to specific metabolic changes, with increased abundance of pyruvate (the enzyme's substrate) and reduced abundance of lactate (Fig 1C), demonstrating the validity of the approach for inferring enzyme function.

3.3.1 Lipin is a lipid phosphatase important for optimal parasite growth

Loss of Lipin was associated with elevated levels of ceramides and lyso-phosphatidic acid (LPA) lipid species in iRBC (Fig 1D and Table 1). Sphingosine 1-P and sphinganine 1-P were also elevated, suggesting that this enzyme dephosphorylates diverse lipid substrates *in vivo*. This phenotype is consistent with the function of Lipin domain-containing proteins in other eukaryotes, in regulating intracellular pools of phosphatidic acid and diacylgycerols, and flux of these lipids into other pathways of bulk lipid synthesis (phospholipid, triacylglycerol). Interestingly, infection with wild-type *P. falciparum* leads to a depletion of phosphatidic acid and accumulation of diacylglycerol in the iRBC[30].

3.3.2 HAD4 is a dispensable nucleotide phosphatase

HAD4 disruption led to an accumulation of several nucleotides and glycolytic intermediates which feed into lipid metabolism (Fig 1E, Table 2). In order to gain further insights into the potential function of HAD4, we recombinantly expressed and purified HAD4 for analysis of *in vitro* activity (Fig 2). Phosphatase activity was confirmed (Fig 1F) and inhibited by free phosphate with an IC_{50} of 2.1 ± 0.21 mM (Fig 1G). HAD4 appeared to dephosphorylate a wide range of nucleotide substrates with a preference for nucleotide monophosphates (Fig 1H). HAD4 may therefore function as a nucleotide monophosphate

phosphatase, consistent with accumulation of monophosphates following depletion, including the elevation of orotidine-P and reduced levels of the dephosphorylated orotidine following HAD4 depletion (Fig 1E). The preference of HAD4 towards the monophosphates versus diphosphates (up to sixfold) is similar to the HAD nucleotide monophosphatase YrfG present in *E. coli*[32]. This preference is likely an underestimate as diphosphate nucleotide substrates will produce monophosphate nucleotides in the assay leading to further hydrolysis and higher activity measurements.

Interestingly, BLASTp searches for *E. coli* genes most homologous to HAD4 reveal that the nearest *E. coli* homolog is not YrfG, but rather YigB, which is predominantly a flavin mononucleotide (FMN) phosphatase[32]. To assess whether HAD4 also catalyzes dephosphorylation of FMN, we implemented a BIOMOL green assay to detect inorganic phosphate release by HAD4, as our original assay was incompatible with the intrinsic absorption of FMN. As compared to deoxyguanosine monophosphate (dGMP), which was the strongest substrate in our larger assay, FMN was far less-robustly dephosphorylated (Figure 3). Although FMN was not detected in metabolomics, the related metabolite, flavin adenine dinucleotide (FAD) had minimal changes upon loss of HAD4, in line with the negative enzymatic findings (Table 2).

3.4 Discussion

We performed a preliminary characterization on two putative haloacid dehalogenases. Lipin was shown to dephosphorylate a subset of lipid species and to be necessary for normal parasite growth. The *P. berghei* ortholog is also predicted to be essential[21] but no data are available at other lifecycle stages. It will be of interest to determine the precise lipid substrate

specificity of this phosphatase and determine whether the enzyme is necessary for bulk lipid regulation or is necessary for specialized lipid signalling[30,31]. In contrast, HAD4 appears to dephosphorylate a range of nucleotides, with a preference for nucleotide monophosphates. Despite the lack of a growth defect in asexual RBC stages and the dispensability of this protein in *P. berghei* liver stage development[33], HAD4 may contribute to optimal metabolic flux and nucleotide homeostasis by repressing monophosphate accumulation. In particular, loss of HAD4 led to an accumulation of orotidine phosphate and a reciprocal decrease in the non-canonical orotidine (Fig 1E), suggesting HAD4 may regulate pyrimidine biosynthetic flux via dephosphorylation of a key intermediate. Furthermore, the dysregulation of downstream glycolytic intermediates, despite low or absent phosphatase activity upon those substrates *in vitro*, suggests that HAD4 exerts glycolytic effects through the regulatory capacity of nucleotides[34,35].

As described in Chapter 2, related HAD proteins (HAD1, HAD2, and PGP) have all been linked to parasite sensitivity to the isoprenoid biosynthesis inhibitor fosmidomycin (FSM)[19,36,37]. Although FSM resistance was not tested in Lipin- or HAD4-knockout parasites, we hypothesize that the glycolytic effects of HAD4 could also confer resistance to FSM, as was seen with HAD1 and HAD2.

Interestingly, homology to documented HAD proteins from *E. coli* was not predictive of the substrate profile of HAD4, as its activity more closely resembled YrfG than the more homologous YigB. This discrepancy speaks to the nature of HAD proteins. It is well-documented that HAD proteins can be both specific and promiscuous in their substrate profiles[32,38], mediated by the Cap domain that closes upon substrates in the binding pocket[39]. Perhaps the tendency of these enzymes to be promiscuous, as well as their divergent

evolutionary origins[40], may explain why it is difficult to predict function from polypeptide sequence homology alone. This supports the notion that biochemical and/or metabolic studies like this one will be necessary for future characterization of HAD proteins, as homology alone is insufficient and can lead to erroneous functional assignment.

3.5 Methods

3.5.1 Recombinant Protein Expression

*Pf*HAD4 (PlasmoDB ID 3D7_1118400) was amplified from *P. falciparum* genomic DNA using the following primers:

5'-CTCACCACCACCACCACCATATGAAAGATGAACAAATATCATGTTATTATC −3'

5'- ATCCTATCTTACTCACTTATGCAAGTATACTATCTAGATCTCG −3'

The PCR product was cloned by ligation-independent cloning into vector BG1861[41], which introduces an N-terminal 6X-His tag. The catalytic mutant HAD4^{D29A} was generated by PCR amplification with the above primers in addition to the following primers:

5'- AATTGATAACGTTCGCCCTTGACCATACGATATG −3'

5'- ATCGTATGGTCAAGGGCGAACGTTATCAATTTTA −3'

These primers were used to introduce an A→C point mutation in the coding sequence. The HAD4^{D29A} construct was cloned by ligation-independent cloning into a modified BG1861 vector that also includes a KFS motif in front of the 6xHis tag to increase protein expression[42]. Constructs were verified by Sanger sequencing.

Constructs were transformed into BL21 (DE3) pLysS *Escherichia coli* (Life Technologies). Cells were grown at 37˚C and induced with 1 mM isopropyl-β-D-thiogalactoside (IPTG). Cells were harvested by centrifugation and resuspended in lysis buffer containing 10

mM Tris–HCL (pH 7.5), 20 mM imidazole, 1 mM $MgCl_2$, 200 mM NaCl, 1 mg/ml lysozyme, 1 mM dithiothreitol (DTT) and cOmplete Mini EDTA-free Protease Inhibitor tablets (Roche), and sonicated. Soluble protein was bound to nickel agarose beads (Gold Biotechnology), eluted in 20 mM Tris–HCl pH 7.5, 150 mM NaCl and 300 mM imidazole, and further purified by size exclusion chromatography using a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) equilibrated with 25 mM Tris–HCl (pH 7.5), 250 mM NaCl and 1 mM $MgCl_2$ buffer. Fractions containing HAD4 were pooled, concentrated with a centrifugal filter, flash-frozen and stored at −80°C.

3.5.2 HAD4 enzymatic assays

All assays were performed in clear 96-well half-area plates using a FLUOstar Omega microplate reader (BMG Labtech) at 37°C. Reaction rates were determined using GraphPad Prism software. All reaction rates represent the mean and standard error of at least three experiments, each with technical replicates. Substrates were purchased from Sigma-Aldrich, except for GTP, dGTP, dATP, dCTP, dTTP, dUTP (Roche), dTMP (BioBasic Canada), DOXP (Echelon Biosciences) and fructose 1-phosphate (Santa Cruz Biotechnology).

Other substrates tested that had activity < 0.1 µmol/min/mg were as follows: cytosine triphosphate; adenosine triphosphate; uridine diphosphate; deoxyuridine triphosphate; deoxycytosine triphosphate; deoxythymidine triphosphate; inosine triphosphate; deoxyadenosine triphosphate; D/L-glyceraldehyde-3-phosphate; sorbitol-6-phosphate; sodium pyrophosphate; deoxyxylulose-5-phosphate; galactose-1-phosphate, sucrose-6-phosphate; 2-deoxy-ribose-5-phosphate; glucose-6-phosphate; glycerol-2-phosphate; mannose-6-phosphate; thiamine monophosphate; fructose-6-phosphate; thiamine pyrophosphate; sedoheptulose-7-phosphate; racemic glycerol-1-phosphate; 2-phosphoglyceric acid; 2,3-diphosphoglycerate; myo-inositol-2-

phosphate; mannose-1-phosphate; trehalose-6-phosphate; glucose-1-phosphate; phosphoenol pyruvate; and dihydroxyacetone phosphate.

General phosphatase activity and phosphate inhibition were measured using the substrate *para*-nitrophenyl phosphate (*p*NPP) (New England Biolabs). Reactions were performed in 50 µL volumes with 1 mM *p*NPP, 50 mM Tris–HCl (pH 7.5), 5 mM $MgCl_2$ and inorganic phosphate (0 mM–42 mM). Reactions contained 1 µg purified recombinant enzyme. *Para*-nitrophenyl production was quantified by absorbance at 405 nm.

Enzyme activity against phosphorylated substrates was measured using the EnzChek Phosphate Assay Kit (Invitrogen) according to supplier instructions. Each 50µl assay contained 200 ng recombinant purified enzyme and 1 mM substrate. The reaction was quantified by absorbance at 360nm, and reactions were linear with respect to time and enzyme concentration. Enzyme activity against flavin mononucleotide (FMN) was measured using the BIOMOL Green kit (Enzo Life Sciences) to account for the overlap in absorbance between FMN and the EnzChek kit. Each 50 µl assay contained 200 ng recombinant purified enzyme, 25 mM Tris–HCl (pH 7.5), 250 mM NaCl, 1 mM $MgCl_2$ and 1 mM substrate. A time course was taken to obtain kinetic data, and the reaction was quantified by absorbance at 620 nm.

Substrate abbreviations are as follows: dGMP, deoxyguanosine monophosphate; dIMP, deoxyinosine monophosphate; dCMP, deoxycytosine monophosphate; GMP, guanosine monophosphate; IMP, inosine monophosphate; XMP, xanthosine monophosphate; AMP, adenosine monophosphate; 8oxodGMP, 8-oxo-deoxyguanosine monophosphate; dAMP, deoxyadenosine monophosphate; GDP, guanosine diphosphate; UMP, uridine monophosphate; CMP, cytosine monophosphate; dTMP, deoxythymidine monophosphate; dGDP, deoxyguanosine diphosphate; GTP, guanosine triphosphate; dADP, deoxyadenosine

diphosphate; ADP, adenosine diphosphate; TTP, thymidine triphosphate; dGTP, deoxyguanosine triphosphate; PLP, pyridoxal-5-phosphate; R5P, ribose-5-phosphate; FBP, fructose-1,6-bisphosphate; E4P, erythrose-4-phosphate; and 3PGA, 3-phosphoglyceric acid.

3.5.3 Cultivation and stable-isotope labelling

Plasmodium falciparum-infected RBCs were cultivated in RPMI 1640 GlutaMAX supplemented with 500 µM hypoxanthine, 22 µg/ml gentamicin, 0.2% (w/v) D-glucose, 25 mM HEPES and albumax II (0.5% w/v). Cultures were routinely synchronized with 5% sorbitol (w/v) to maintain a 12-hour developmental range. Transfectant parasite cultures were maintained with 5% human serum and 0.25% (w/v) albumax II. Uninfected RBCs from the same donor were maintained under identical conditions for 48 h prior to experimentation. Trophozoite-stage *P. falciparum*-infected RBCs were magnetically enriched to > 95% parasitemia (Colebrook Bioscience) and cell density determined using a Neubauer Haemocytometer. Following a one-hour recovery in fresh RPMI 1640 at 37°C, 1×10^8 cells were centrifuged and the media replaced with 5 ml of RPMI containing unlabeled substrates or 11 mM $^{13}C_6$-glucose, 2 mM $^{13}C_5$ glutamine, 1.1 mM $^{13}C_6$ arginine, 1 mM $^{13}C_3$ serine, 1 mM $^{13}C_5$ proline, 5 mM $^{13}C_2$ acetate, 23 mM $^{13}C_1$ bicarbonate, 1 mM $^{13}C_6$ lysine, 10 mM $^{13}C_3$ glycerol or ^{13}C-amino acid mix at 1 mg/ml (and a ^{12}C-amino acid mix 1mg/ml used as a control). iRBC and matched uRBC samples were incubated for 5 h under standard culturing conditions.

3.5.4 Metabolite extraction

An aliquot of culture (1×10^8 cells/sample) was transferred to a microcentrifuge tube and centrifuged at 14,000 *g* for 30 s. The medium was aspirated, and the cell pellet was resuspended in ice-cold PBS (1 ml) to quench cell metabolism and transferred equally into two microcentrifuge tubes. Following centrifugation (14,000 *g*, 30 s), the PBS was aspirated and

200 µl of 80% acetonitrile (containing 1 µM $^{13}C_5$ $^{15}N_1$ aspartate or 5 µM $^{13}C_5$ valine as the internal standard) was added and rapidly mixed for polar LC-MS analysis. Samples were centrifuged to remove precipitated protein (14,000 *g*, 5 min) and the supernatant transferred to a glass mass spectrometry vial (containing an insert) and stored at −80°C until LC-MS analysis. For GC-MS and LC-MS lipid analysis, the quenched cell pellet was suspended in 100 µl chloroform and vortex-mixed, prior to addition of 400 µl of methanol:H_2O (3:1 v/v) with further vigorous mixing. Samples were centrifuged for 5 min at 14,000 *g* and the supernatant transferred to a fresh microcentrifuge tube. H_2O (200 µl) was added to generate a biphasic mixture which was vortex-mixed and then centrifuged for 1 min at 14,000 *g*, and the top aqueous layer was collected for derivatization and GC-MS analysis. The lower organic phase was collected for lipid analysis. For GC-MS analysis, samples were dried by vacuum centrifugation, then resuspended in 90% methanol (100 µl) and transferred to a glass mass spectrometry insert. Samples were dried by vacuum centrifugation, washed with 100 µl methanol and dried. Samples were resuspended in 20 µl of methoxyamine (20 mg/ml) prepared in pyridine, sealed and incubated overnight with shaking at ambient temperature. The next day, 20 µl of BSTFA was added to each sample and incubated for 1 h prior to GC-MS analysis.

Lipid analysis was carried out on the remaining lower organic phase. The organic layer was transferred to a fresh microfuge tube and dried down under nitrogen flow. Dried samples were stored at −80°C until ready for lipid LC-MS analysis.

3.5.5 LC-MS acquisition

Polar metabolite detection was performed on an Agilent 6550 Q-TOF mass spectrometer operating in negative mode. Metabolites were separated on a SeQuant ZIC-pHILIC column (5 µM, 150 × 4.6 mm, Millipore) using a binary gradient with a 1200 series HPLC system across

a 45-min method using 20 mM ammonium carbonate (pH 9) and acetonitrile as outlined in Cobbold *et al*, (2016)[43]. Two independent replicates of the metabolite profiling following AMR1 and Lipin depletion were performed using the same ZIC-pHILIC chromatography on a Thermo Q-Exactive operating in both positive and negative mode (rapid switching) as described previously[44]. Lipid extracts were analyzed on an Agilent 6550 Q-TOF using the reverse phase chromatography outlined by Bird *et al* (2011)[45].

GC-MS analysis was performed using methods previously described[46]. Metabolites were separated using a BD5 capillary column (J&W Scientific, 30 m × 250 µM × 0.25 µM) on a Hewlett Packard 6890 system (5973 EI-quadrupole MS detector). The oven temperature gradient was 70 °C (1 min); 70°C to 295°C at 12.5°C/min, 295°C to 320°C at 25°C/min; 320°C for 2 min. MS data were acquired using scan mode with a *m/z* range of 50–550, threshold 150 and scan rate of 2.91 scans/second. GC retention time and mass spectra were compared with authentic standards analyzed in the same batch for metabolite identification.

3.5.6 Draft metabolome compilation

Raw Agilent.d files were converted to mzXML with MSconvert and analyzed using the X^{13}CMS R package[26]. XCMS centWave peak detection was performed with a 10 ppm mass tolerance, following obiwarp retention time correction. The getIsoLabelReport function was performed with $P < 0.01$ (Welch's *t*-test), to identify features that significantly labelled under each ^{13}C-labelling condition in both iRBC and uRBC cell types (isotope mass tolerance = 15 ppm). The getIsoDiffReport function then compared the significantly labelled features between iRBC and uRBC using $P < 0.01$ (Welch's *t*-test). IsoDiffReport *m/z* features were removed if > 1,200 *m/z*, and the putative M_0 species queried to the METLIN database with a 10 ppm mass tolerance, excluding all toxicants and adducts, except M-H, from the search.

Only *m/z* features with a putative METLIN match were retained for further analysis. *m/z* features were further curated to remove mis-annotated in-source fragments and isotopologues, and where ^{13}C-enrichment was < 1%. *m/z* features with a retention time of < 250 s were removed from the polar metabolite analysis (assuming these lipid-like species were also detected in the lipid C18 LC-MS analysis). This processing was performed for all ^{13}C-labelling conditions across both C18 lipid and pHILIC polar LC-MS data and a single metabolite list complied. Metabolite identification was performed either with retention time matching to authentic standards (approximately 150 in-house metabolite standards or the HMDB compound library), or MS/MS matching. Polar MS/MS data were collected with an auto preferred MS/MS method at a collision energy of 10V, 20V and 40V at a threshold of 10,000 on an Agilent 6545 Q-TOF using the same chromatography outlined above. MS/MS spectra were searched against the METLIN PCDL database via the Agilent Qualitative software package and manually curated. MS/MS spectra without a match were reacquired in positive and negative mode and manually searched against the online METLIN database. Lipid MS/MS data were collected as above and converted to (ABFconverter) and analyzed with the MSDial software package[47].

Polar GC-MS data were analyzed using the DeXsi software package[48] and metabolite identifications confirmed using an in-house metabolite standards library and merged with the LC-MS datasets.

A first iteration of the observable metabolome of *P. falciparum* was compiled, and unique KEGG IDs (acquired from MetaboAnalyst 3.0[49]) were used to compare to the predicted metabolome of *P. falciparum* reported by Huthmacher *et al* (2010)[50]. Under the conditions tested, we could not confidently assign stereoisomers or precise structures (e.g. D-glucose) and assigned KEGG IDs to be consistent with the level of evidence present in the literature and with

the format used by Huthmacher *et al* (2010)[50]. uRBC and iRBC datasets were then re-extracted for expected *m/z* of metabolites predicted from Huthmacher *et al* (2010)[50], to assess the rate of false negatives but also because certain classes of metabolites were not expected to label under any of the conditions tested (e.g. vitamins and some cofactors). Putative matches were then verified using authentic standards and MS/MS matching. This iteration of the observable metabolome of iRBC and uRBC cell types was then grouped according to the level of identification: (i) match to authentic standard and MS/MS match, (ii) match to authentic standard or MS/MS match, (iii) ^{13}C-labelling data consistent with known metabolic pathway and (iv) exact mass match (≤ 10 ppm). To capture additional metabolites that were unlabeled across all ^{13}C-substrates, all *m/z* features from unlabeled uRBC and iRBC extracts were compared with the expected exact mass of all metabolites in the predicted *P. falciparum* and RBC metabolomes. Putative hits were confirmed via pure standards or MS/MS matching where possible.

3.5.7 Pathway reconstruction

Metabolite enrichment analysis was performed with Metaboanalyst 3.0, and pathway reconstruction was performed for each stable-isotope condition. Observed metabolites that were labelled under a given condition were compared to the theoretical reactions for both iRBC and uRBC[50]. A matched reaction was defined as having a labelled metabolite corresponding to either an expected product or substrate. Unpredicted reactions were reported when an identified metabolite was labelled under a given ^{13}C-substrate but did not match to a predicted reaction. A putative reaction was proposed based on the nearest proximity to an identified metabolite with consistency in the labelling pattern.

3.5.8 Generation of transgenic parasites for inducible protein disruption

For each gene of interest, a protospacer adjacent motif (PAM) was selected at the center or towards the start of the gene which ranked the highest via CHOP-CHOP v2[51]. Guide oligos were synthesized for cloning into the BtgZI site of pUF-Cas9g using the In-Fusion cloning kit (Takara). The gene was modified using a rescue template containing a 5'-homology arm immediately upstream of the PAM site, an artificial intron containing a loxP site, a recodonized version of the down-stream gene sequence, a 3xHA tag, which was synthesized by GeneArt. The rescue template was cloned into the pGlmS plasmid using the BglII and SpeI[52]. The 3'-homology arm was generated via PCR of *P. falciparum* genomic DNA and cloned into the pGlmS plasmid with EcoRI and KasI. Following verification of the correct DNA sequence, both pUF-Cas9g and pGlmS plasmids were purified with a midi-prep kit (Macherey-Nagel). pGlmS plasmid was linearized overnight using BglI, BglII and PvuI. 75–150 µg of pUF-Cas9g and linearized pGlmS were transformed into 3D7 ring-stage parasites with dimerizable Cre-recombinase integrated into its genome[29]. Cultures were maintained on WR99210 and blasticidin and complete integration was confirmed on recovered parasites via PCR.

Inducible disruption of each targeted protein was performed cultivating transfected parasites under standard conditions in the presence or absence of 100 nM rapamycin and 2.5 mM glucosamine. Initial tests were performed to determine length of treatment required to deplete the target protein and effectiveness of the DiCre-loxP and GlmS ribozyme system independently.

3.6 Figures

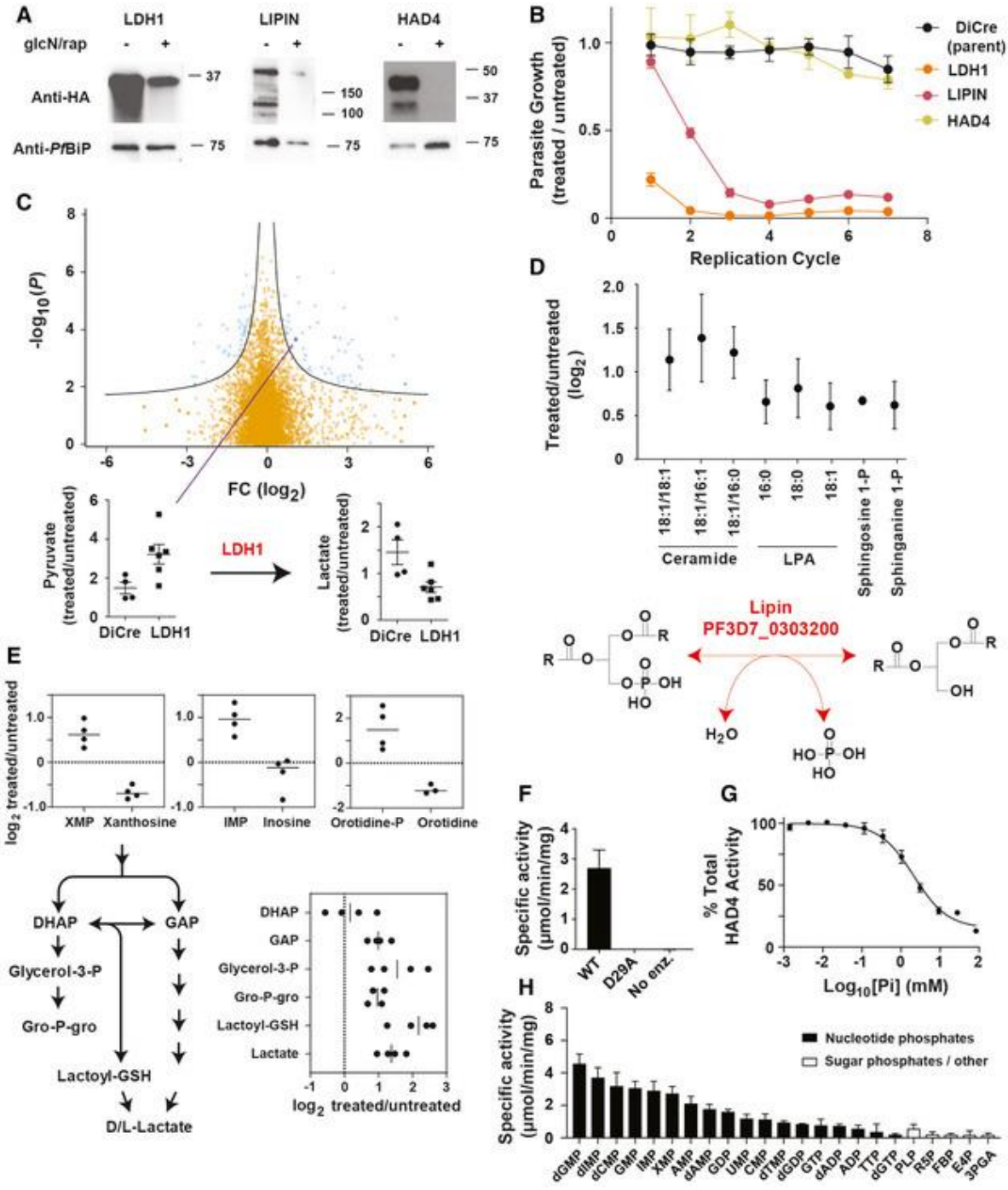


Figure 1. Inducible disruption of haloacid dehalogenase 4 and a putative Lipin.

(A) Anti-HA Western blots for each protein targeted for disruption with a detectable band corresponding to the expected product size. 100 nM rapamycin and 2.5 mM glucosamine were added to induce protein depletion for either one cycle (LDH1) or three cycles (Lipin and HAD4). Anti-*Pf*BiP was used as the loading control. **(B)** Growth of DiCre-3D7 transgenic lines following LDH1, HAD4 and Lipin depletion with 100 nM rapamycin and 2.5 mM glucosamine was assessed relative to uninduced parasite cultures. The untransfected DiCre-3D7 parasite line was used as a negative control. Data are presented as the mean ± standard error of the mean (SEM) from three independent replicates. **(C)** Untargeted LC-MS analysis of LDH1 depletion (one cycle of rapamycin and glucosamine treatment) indicated minimal but selective metabolic changes. *m/z* feature intensities are plotted as the $\log_2$ ratio of treated/untreated and the Benjamini–Hochberg-corrected *P* values across six biological replicates plotted as –log10(*P*). Below, LDH1 substrate and product (pyruvate and lactate, respectively) abundance plotted as the ratio of treated/untreated from biological replicates (mean ± SEM) and the parental DiCre line presented as the negative control. **(D)** Loss of Lipin (PF3D7_0303200) leads to accumulation of various lipid species (three cycles of rapamycin and glucosamine treatment). Data are presented as the mean $\log_2$ fold change Lipin-depleted (treated) enriched trophozoite-stage iRBCs to untreated controls from three to six biological replicates (± SD). The schematic depicts the proposed lipid phosphatase activity of Lipin. **(E)** Loss of HAD4 leads to increases in intracellular levels of several nucleotides and intermediates in lower glycolysis. Data are represented as the mean $\log_2$ fold change of enriched trophozoite-stage iRBCs treated for three cycles compared to untreated controls from three/four biological replicates. Three nucleotide–nucleoside pairs are depicted in the top inset panels, and HAD4 disruption leads to accumulation of the nucleotide monophosphate and depletion of the nucleoside. The schematic depicts lower glycolysis and

triose-phosphate interconversion, with the corresponding metabolite levels following HAD4 disruption depicted in the lower right inset. **(F)** Phosphatase activity against the generic substrate pNPP of wild-type HAD4, catalytically inactive HAD4^{D29A} and a no-enzyme (No enz.) control, presented as the mean ± SEM (three biological replicates). **(G)** Inhibition of HAD4 phosphatase activity by inorganic phosphate. Phosphatase activity against the generic substrate pNPP as the mean ± SEM from three independent experiments. The inorganic phosphate IC_{50} of HAD4 is 2.1 ± 0.21 mM. **(H)** Substrate specificity of HAD4. Substrates are divided into nucleotide phosphates (black) and other metabolites such as sugar phosphates and vitamins (white). Presented is the enzyme activity from three independent experiments mean ± SEM. Substrate concentrations were 1 mM, and abbreviations are listed in methods.

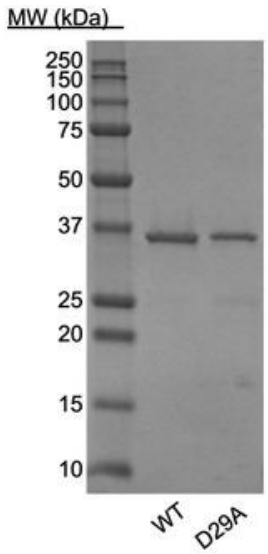


Figure 2. Coomassie-stained SDS–PAGE gel of purified HAD4.

Wild type (WT) and catalytically inactive (D29A) HAD4 were expressed and purified from *E. coli*.

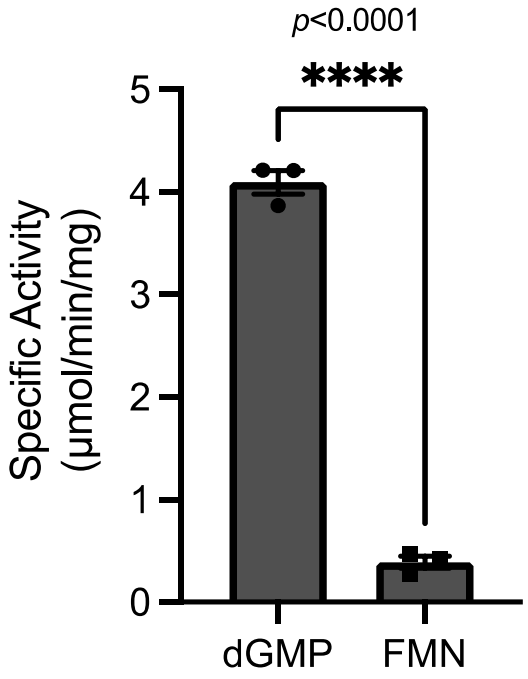


Figure 3. FMN phosphatase activity of HAD4.

Phosphate release by HAD4 incubated with the given substrate was measured by BIOMOL Green detection reagent over time. Shown are the mean and SEM of three independent experiments, each with technical replicates. Statistics were performed with unpaired t-test. Abbreviations: dGMP, deoxyguanosine monophosphate; FMN, flavin mononucleotide.

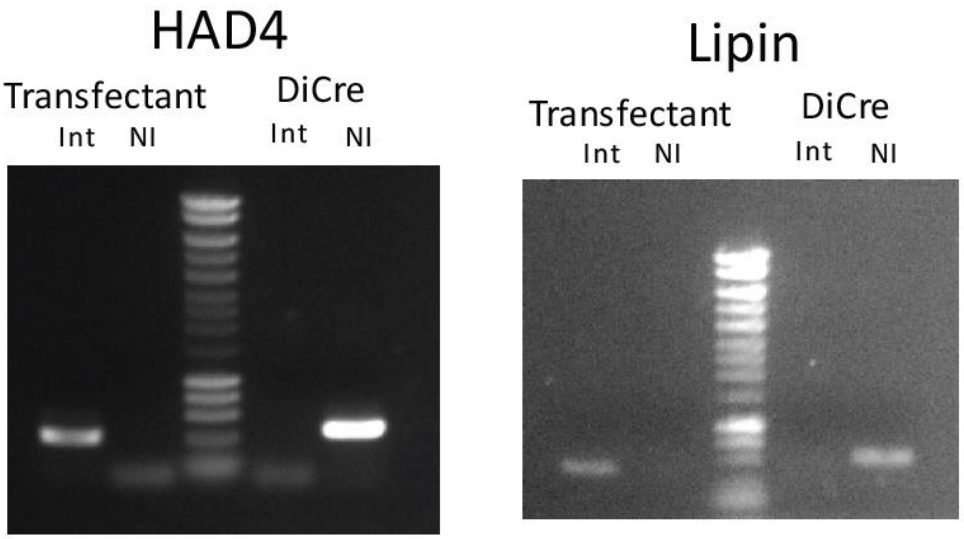


Figure S1. PCR confirmation of integration.

'Int' indicates primers specific for the integrated locus and 'NI' indicates primers specific for the endogenous (nonintegrated)locus. DiCre represents the untransfected parental line stably expressing the DiCre recombinase.

3.7 Tables/Datasets

Table 1: Mass Spectrometry data for treated/untreated parasites of Lipin KO

m/z #	compound	m/z	Ratio1_log	Ratio2_log	Ratio3_log	Ratio4_log	Ratio5_log	Ratio6_log	Log mean	SD
1	CE(18:4)	665.5324	1.218509	2.691747	-0.28798	0.427853	1.160949	1.170339	1.06357	0.992639
2	CE(20:7)	687.5134	0.984557		0.962456	0.640828	1.095088	0.772094	0.920732	0.181821
3	Ceramide(d18:1/16:0)	536.5038	0.82066				1.076939	1.513982	1.137194	0.350566
4	Ceramide(d18:1/16:1)	534.4887	0.807051				1.691366	1.656766	1.385061	0.50087
5	Ceramide(d18:1/18:1)	562.5189	0.881667				1.371109	1.406989	1.219922	0.293486
6	Ceramide(d18:1/14:0)	508.473	-0.14188				1.608448	1.849055	1.105208	1.086689
7	LPA(16:0)	409.2353		0.546128	0.86132		0.861857	0.359568	0.657218	0.247973
8	LPA(18:0)	437.269		0.711578	1.283715		0.764426	0.487799	0.811879	0.336635
9	LPA(18:1)	435.2505		0.717914	0.466401		0.91902	0.320506	0.60596	0.265519
10	LPA(20:0)	465.3042	0.384655	1.076788	-0.59449		0.703461	0.391278	0.392339	0.620313
11	LPE(18:0)	480.3084	0.076995	0.565154	-0.37462		0.599038	0.625034	0.29832	0.438808

12	LPE(18:1)	478.293	0.994185	0.309762		0.761719	0.374189	-0.07305	0.47336	0.415204
13	LPE(18:2)	476.2777	0.937968	0.037158		0.865188	0.359487	-0.14355	0.411249	0.483198
14	LPE(20:4)	500.2785	0.801744	0.511443		0.962743	0.354774	0.003374	0.526816	0.377308
15	Sphingosine 1-phosphate	378.2405		0.710905	0.604823	0.695733	0.64786	0.693794	0.670623	0.043697
16	Sphinganine 1-phosphate	380.2559	0.150293	0.650656	0.798919	0.810598		0.687429	0.619579	0.271316
17	(-)-jasmonic acid methyl ester	223.1331	-0.39032	NA	NA	-0.15792			-0.27412	0.164333
18	(cis)-aconitate	173.0909	0.963053	-0.2088	0.243187	0.307415			0.326215	0.482717
19	(R)-S-Lactoylglutathione	378.0977	1.20872	5.332734	-0.86133	0.350279			1.5076	2.687764
20	1-3 Bisphosphoglycerate / 2-3-Bisphosphoglycerate	264.9524	-1.30275	-0.67161	1.027085	-0.27651			-0.30594	0.984066
21	1-Pyrroline-5-carboxylic acid	112.0403	-0.76104	-1.79024	0.067741	0.176405			-0.57678	0.910882

22	2-C-Methyl-D-erythritol 2-4-cyclodiphosphate	276.9883	1.975813	-1.92657	0.170254	-0.3242			-0.02618	1.60695
23	2-C-Methyl-D-erythritol 4-phosphate	215.0327	-0.32747	-0.27279	0.321332	-0.24876			-0.13192	0.303959
24	2-deoxyribose-5-P	213.0168	-0.11184	-0.98789	0.822998	0.494608			0.05447	0.795514
25	2-Hydroxyglutarate	147.0297	-0.9265	0.111833	-0.0651	0.351402			-0.13209	0.556428
26	2-Oxobutanoate	101.0244	-0.91693	-0.34994	0.033594	0.2447			-0.24715	0.509873
27	2-Oxoglutaramate	144.03	0.79373	-1.00281	0.586567	0.014278			0.09794	0.804475
28	3-Phosho-D-glycerate	184.9857	-0.22184	-0.04642	0.026716	-0.03649			-0.06951	0.106595
29	4-Aminobutyric acid	102.0561	-0.09138	-0.43338	0.029057	-0.12078			-0.15412	0.197138
30	4-Methylene-L-glutamate	158.0459	0.376534	-0.8654	-0.06769	0.125342			-0.1078	0.536814

31	6-Phospho-D-gluconate	275.0176	0.358472	-1.21434	0.897791	-0.1477			-0.02644	0.899661
32	a-ketoglutarate	145.0143	-1.05281	-0.21074	-0.02204	0.243535			-0.26051	0.560105
33	Acetyl-CoA	808.1183	0.642466	-0.99213	-0.00111	0.542104			0.047832	0.748736
34	ADP	426.0221	-0.5252	-1.10293	-0.23971	-0.54243			-0.60256	0.361306
35	AMP	346.0558	-0.92696	-1.56851	-0.80805	-0.44237			-0.93647	0.469107
36	Arginine	173.1043	0.230491	-0.1374	-0.47092	0.33949			-0.00958	0.369077
37	ATP	505.9893	0.449412	-0.66531	0.19529	-0.13284			-0.03836	0.481144
38	CDP-choline	487.1003	-0.55184	-0.80466	-0.20256	-0.18708			-0.43653	0.297649
39	CDP-ethanolamine	445.054	-0.88889	-1.28712	-0.17415	-0.31713			-0.66682	0.516105

40	Citrate	191.0199	0.513383	-0.41764	0.72276	0.407522			0.306507	0.500217
41	CMP	322.0447	-0.225	-1.31455	-1.09962	0.203973			-0.6088	0.718071
42	CoA	766.1071	-0.7192	-0.79635	0.076327	0.052267			-0.34674	0.475768
43	CTP	481.9775	-0.21617	-0.91599	-0.33126	-0.12086			-0.39607	0.357127
44	D-4--Phosphopantothenate	298.0698	-0.0894	-0.50597	-0.02832	0.402604			-0.05527	0.371784
45	D-Glyceraldehyde 3-phosphate	168.9909	1.05251	-1.03082	0.311542	0.080229			0.103365	0.862386
46	D-Histidine	154.0624	0.21682	-0.70381	-0.09656	0.119872			-0.11592	0.413241
47	Dihydroorotate	157.0255	0.962911	0.256836	-0.47556	-1.05872			-0.07863	0.878535
48	dihydroxy-acetone-phosphate	168.9909	-0.97895	-0.47874	-0.16379	0.215681			-0.35145	0.505573

49	DL-Glutamate	146.0458	-0.19129	-0.38712	-0.06349	-0.13543			-0.19433	0.138763
50	dTTP	480.9817	-1.26127	-0.68546	-0.47651	0.184703			-0.55963	0.596958
51	Ethanolamine phosphate	140.0119	0.321737	-0.88385	-0.47383	0.210237			-0.20643	0.572415
52	FAD	784.1483	0.120986	-1.356	-0.47552	-0.16652			-0.46926	0.639371
53	fructose-1-6-bisphosphate	338.9883	0.879374	-0.58682	0.457815	-0.05198			0.174599	0.634562
54	Fructose-1-P	259.0226	-0.29176	-0.16197	0.864512	-0.51665			-0.02647	0.611791
55	Fructose-6-phosphate	259.0229	-0.09615	0.140539	0.151465	0.169703			0.091389	0.125604
56	Fumarate	115.0036	0.510249	-0.0315	0.039111	-0.38102			0.034209	0.366684
57	GDP	442.0176	0.079948	-2.0613	-0.56295	-0.34392			-0.72206	0.931861
58	GDP-mannose	604.0706	0.202525	-1.79715	-0.52385	-0.52709			-0.66139	0.831316

59	Gluconate	195.051	0.302399	-0.01756	0.089506	0.113881			0.122057	0.133095
60	Glucose-6-phosphate	259.0226	-0.35003	0.049538	-0.10932	0.075115			-0.08367	0.195416
61	glutathione	306.0766	0.403798	-0.38028	-0.34373	0.034808			-0.07135	0.368177
62	Glutathione disulfide	611.1453	1.226363	-2.76934	-0.21087	0.575537			-0.29458	1.751365
63	GMP	362.051	0.806146	-0.88523	-1.23498	0.245913			-0.26704	0.954623
64	GTP	521.9836	-0.14555	-1.03529	0.022485	-0.08388			-0.31056	0.488113
65	HLD	382.1727	0.798341	-0.40454	0.219072	0.446332			0.264801	0.505875
66	L-Aspartate	132.0304	-0.18288	-0.60052	-0.37418	-0.91586			-0.51836	0.31522
67	L-Citrulline	174.0882	-0.777	0.199999	0.787933	-0.46201			-0.06277	0.698156
68	L-Lactic acid	89.02435	0.825037	-0.33393	0.013419	0.691055			0.298895	0.551538

69	L-Ornithine	131.0825	-0.27455	0.16167	0.294691	0.063067			0.061219	0.243136
70	L-Pipecolate	128.0718	0.985193	-0.48072	0.212919	0.573272			0.322667	0.621614
71	Malate	133.0144	0.832778	-0.18015	0.10056	0.126312			0.219875	0.431532
72	myo-inositol-3-P	259.0227	-0.28286	-0.79551	-0.0223	-0.07433			-0.29375	0.352948
73	N-Carbamoyl-L-aspartate	175.0361	0.964764	0.032014	0.170318	-0.47632			0.172694	0.596767
74	N-Succinyl-L-L-2-6-diaminopimelate	289.1046	1.037043	-1.03398	-0.14991	0.082626			-0.01606	0.851048
75	N2-Succinyl-L-ornithine	231.0987	0.640428	-0.54235	0.301076	0.607908			0.251765	0.551044
76	NAD+	NA	-0.26234	-0.9392	-0.14276	-0.08304			-0.35684	0.395337
77	NADH	664.1166	0.27344	-0.84666	-0.29249	-0.46915			-0.33372	0.466117
78	NADP+	NA	-0.61505	-0.69973	0.15424	-0.25115			-0.35292	0.390111

79	NADPH	744.0842	0.174137	-6.24507	1.833144	0.527039			-0.92769	3.616027
80	O-Succinyl-L-homoserine	218.0669	0.311765	-0.8659	-0.56136	0.737723			-0.09444	0.746249
81	Octoluse 8/1P	319.0435	0.162403	-0.04357	0.56766	0.380901			0.266849	0.265055
82	Octulose-1-8-bisphosphate	399.0103	0.022249	-2.70135	0.640363	-0.28546			-0.58105	1.465023
83	Orotate	155.0096	0.344978	-0.44316	-0.18793	-0.70784			-0.24849	0.448988
84	Orotidine	287.0517	1.020474	-0.70235	-1.10622	0.485071			-0.07576	0.995109
85	PEE	372.1417	0.757823	-0.53035	0.128532	0.426346			0.195587	0.547978
86	PEEK	500.2365	-0.11885	-1.20902	0.085348	0.342568			-0.22499	0.682643
87	Phosphoenolpyruvate	166.9751	-0.39695	-0.14434	-0.02176	-0.15626			-0.17983	0.156996

88	Proline	114.0561	-0.04769	-0.45205	0.182756	0.095656			-0.05533	0.281025
89	Pyridoxine	168.0664	0.059274	0.243304	0.346306	0.122055			0.192735	0.127735
90	Ribose-5-Phosphate	229.0119	0.157274	-0.25578	0.133644	0.18925			0.056097	0.209163
91	S-Adenosyl-L-homocysteine	383.1138	1.026715	-0.61926	-0.52311	0.256559			0.035227	0.768576
92	S-Adenosyl-L-methionine	397.1303	0.457451	-0.61886	0.125654	0.135433			0.024918	0.456034
93	S-Inosyl-L-homocysteine	384.0981	-0.33056	-2.19842	-3.03651	-0.12637			-1.42296	1.423537
94	S-Methyl GSH	320.0924	-0.07763	-0.01997	1.176013	-0.42281			0.163899	0.697795
95	Sedoheptulose-1-7-bisphosphate	368.9996	0.063344	-0.86941	0.610933	0.171442			-0.00592	0.622461
96	Sedoheptulose 7-phosphate	289.0334	-0.13041	0.373168	0.974678	-0.01179			0.301412	0.497661
97	sn-glycero-3-Phosphoethanolamine	214.0482	0.579944	-0.02244	0.068756	0.099684			0.181485	0.270651
98	sn-Glycerol 3-phosphate	171.0063	-0.22065	-1.02638	0.393603	0.301972			-0.13787	0.651213

99	sphinganine 1-phosphate	380.2562	-0.15135	0.826025	0.64615	0.295281			0.404028	0.43088
100	Succinate	117.0192	0.489903	-0.87516	-0.08761	-0.09599			-0.14221	0.560326
101	Succinoadenosine	462.0672	0.060305	-0.99537	-1.07374	0.267197			-0.4354	0.697714
102	TPDAVM	631.2766	0.521332	-1.73744	0.300412	0.627714			-0.072	1.118634
103	UDP	402.953	-0.36463	-1.11653	-0.65652	-0.30541			-0.61078	0.370457
104	UMP	323.029	0.170821	-1.15574	-1.12172	-0.35641			-0.61576	0.641237
105	UTP	482.9627	0.054281	-0.95937	-0.17194	0.050316			-0.25668	0.480244
106	VDEVG	516.2309	-0.26438	-1.13121	-0.25421	0.519922			-0.28247	0.674614

Table 2: Mass Spectrometry data for treated/untreated parasites of HAD4 KO

m/z #	compound	mz_m	RT_m	LogRatio1	LogRatio2	LogRatio3	LogRatio4	Log mean	SD
1	(-)-jasmonic acid methyl ester	223.1332	3.405483	2.202308	-0.15503	0.038674	-0.11402	0.492984	1.142593
2	(cis)-aconitate	173.0089	21.02408	0.742625	1.138888	0.579695	1.234507	0.923929	0.313068
3	(R)-S-Lactoylglutathione	378.0972	14.81235	1.952281	2.595178	1.256222	2.409859	2.053385	0.596184
4	1-3 Bisphosphoglycerate / 2-3-Bisphosphoglycerate	264.9523	22.202	2.505718	2.806278	0.739404	4.773988	2.706347	1.652699
5	1-5-Anhydrosorbitol	163.0611	13.47387	-0.54308	-0.29524	-0.59348	0.089008	-0.3357	0.3117
6	1-Pyrroline-5-carboxylic acid	112.0403	18.03903	0.050178	0.842594	0.431556	1.349005	0.668333	0.557333
7	13S-HODE	295.2265	3.388041	-0.34124	-0.27018	1.686668	1.063205	0.534613	1.003569
8	2-3-6-Trihydroxypyridine	126.0196	12.4667	0.887678	1.47143	0.988734	1.030451	1.094573	0.25829
9	2-3-Bisphosphoglycerate	264.9515	20.59875	0.392509	0.751475	0.333324	-1.04484	0.108118	0.790529
10	2-C-Methyl-D-erythritol 2-4-cyclodiphosphate	276.9883	18.39288	0.308355	0.614085	0.280203	0.491675	0.42358	0.157864

11	2-C-Methyl-D-erythritol 4-phosphate	215.033	16.52829	-0.06133	0.075647	-0.09663	0.277093	0.048703	0.16942
12	2-deoxyribose-5-P	213.017	16.33859	0.558174	1.534558	0.35752	2.681129	1.282845	1.064571
13	2-Hydroxyethanesulfonate	124.9912	12.81872	-0.47097	0.399192	-0.08139	0.131727	-0.00536	0.367438
14	2-Hydroxyglutarate	147.0298	18.0183	-0.02614	0.452004	-0.18043	-0.04441	0.050257	0.276535
15	2-Keto-3-deoxy-D-gluconic acid	177.0401	15.33347	-0.26054	0.725185	-0.60944	-0.10242	-0.0618	0.565803
16	2-keto-3-deoxy-D-glycero-D-galacto-nononic acid	267.0723	17.50992	-0.23912	-0.89307	-0.40594	-1.37391	-0.72801	0.51224
17	2-keto-D-Gluconic acid	193.0352	17.36146	1.870132	2.366606	2.772885	3.192241	2.550465	0.565098
18	2-octaprenylphenol	637.5335	3.218483	-0.06097	-0.34912	-1.0043	-2.33034	-0.93618	1.009777
19	2-Oxoglutaramate	144.0301	12.44133	1.053995	1.423596	0.925055	0.952719	1.088841	0.229949
20	2-Phosphinomethylmalate	211.0013	17.35812	0.720171	0.890229	0.32799	0.893665	0.708014	0.265979
21	3-7-11-Trimethyl-6E-10-dodecadien-1-ol	223.2062	3.619578	0.174898	0.130344	-0.27157	2.109408	0.535771	1.068134

22	3-octaprenyl-4-hydroxybenzoate	681.523	3.08555	-1.31131	0.861338	-0.41941	-1.267	-0.53409	1.016791
23	3-Phosho-D-glycerate	184.9859	19.75729	0.098785	0.232833	0.078406	0.551085	0.240277	0.218235
24	4-(glutamylamino) butanoate	231.0984	16.08309	1.912612	2.469261	0.841198	1.247483	1.617638	0.719292
25	4-Aminobutyric acid	102.0561	17.24997	0.00802	0.251974	-0.05875	0.214383	0.103907	0.152511
26	4-hydroxybutyrate	103.04	18.008	-0.16792	0.158436	0.006491	0.226169	0.055793	0.175159
27	4-Hydroxycinnamyl aldehyde	147.0443	10.77349	0.273315	1.810675	0.081894	0.524696	0.672645	0.780054
28	4-hydroxyphenylpyruvate	179.0349	13.60848	0.756413	-0.79548	0.048442	0.030172	0.009887	0.634511
29	4-hydroxytoluene	107.05	4.129083	0.454398	-0.13476	0.183594	0.151653	0.16372	0.24092
30	4-Methylene-L-glutamate	158.0458	16.88183	0.989414	1.743342	0.849949	1.23773	1.205109	0.393032
31	4-Oxoglutaramate	144.0301	9.0471	1.106961	1.147414	0.767629	1.016875	1.00972	0.170367
32	4-Oxoproline	128.0352	11.77852	0.583591	0.912465	0.290564	0.560902	0.586881	0.25462
33	5-amino-6-(5-phospho-D-ribosylamino)uracil	353.0487	5.123167	0.64937	0.623516	0.390893	0.204393	0.467043	0.210167

34	5-deoxy-D-glucuronate	177.0402	14.3451	0.082381	1.53937	0.147817	0.216739	0.496577	0.697356
35	5-Formiminotetrahydrofolate	471.1757	13.1381	1.290518	2.05279	0.951249	2.110675	1.601308	0.572264
36	5-methyl-tetrahydrofolate	458.1806	13.17111	4.621828	5.672918	5.514713	6.926607	5.684017	0.94887
37	5-methyltetrahydropteroyltri-L-glutamate	716.2624	20.20983	-0.45846	0.136315	-0.16594	0.244299	-0.06095	0.316823
38	5S-HPETE	335.2225	3.521283	1.186015	0.620909	-0.19293	-0.68969	0.231075	0.83496
39	6-7-dimethyl-8-(1-D-ribityl)lumazine	325.1153	15.5504	0.94856	1.915534	0.56916	1.376869	1.202531	0.578627
40	6-hydroxymethyl-7-8-dihydropterin	194.0686	10.76133	3.108681	3.441123	0.662271	2.345339	2.389353	1.239409
41	6-Phospho-D-gluconate	275.017	20.68252	0.083245	0.233045	0.226498	0.510641	0.263357	0.178762
42	9Z-Hexadecenoic acid	253.217	3.237067	-0.25445	-0.7014	0.078443	-0.9974	-0.4687	0.475724
43	a-ketoglutarate	145.0143	18.61592	0.589292	1.277141	0.508531	0.78443	0.789849	0.344889
44	Acetoacetic acid	101.0244	13.80461	0.586479	-4.64432	-0.25698	0.157916	-1.03923	2.427941
45	Acetyl-alanine	130.0509	9.554242	1.476433	2.240361	0.766986	1.114146	1.399481	0.630996

46	Acetyl-CoA	808.1165	14.29273	0.913903	1.454242	0.463493	1.361696	1.048333	0.455729
47	Acetyllysine	187.1082	16.07188	1.399095	2.088967	0.487416	1.536074	1.377888	0.664341
48	adenosine-3--5--bisphosphate	426.0223	16.19158	0.780951	0.910806	0.372514	1.043525	0.776949	0.290152
49	ADP	426.0212	17.51417	0.931819	0.139756	0.071395	0.582951	0.43148	0.403339
50	Ala-Trp-OH	382.1035	16.5027	3.635336	4.61803	0.313911	5.839581	3.601714	2.370072
51	alpha-linolenate	277.2156	3.226757	0.055853	-0.37201	0.310501	-1.88678	-0.47311	0.983619
52	AMP	346.0558	16.04092	1.367785	0.470876	0.159614	0.660646	0.66473	0.512193
53	Arginine	173.1042	26.10383	0.309647	0.694492	0.203393	0.955377	0.540727	0.347741
54	Ascorbic acid	175.0247	18.17703	0.575869	0.90688	0.530995	0.912804	0.731637	0.206601
55	ATP	505.9872	19.065	0.023927	0.203489	0.15819	0.62141	0.251754	0.257963
56	CDP	402.0103	19.76295	0.995869	-0.39587	0.242767	0.509737	0.338126	0.580207
57	CDP-choline	487.0991	17.48767	0.201066	0.45527	0.173722	0.481895	0.327988	0.16309
58	CDP-ethanolamine	445.0522	18.82722	0.292997	0.828938	0.250447	0.901836	0.568555	0.344481

59	Citrate	191.0197	21.09378	0.231084	0.45518	0.217231	0.568405	0.367975	0.172472
60	CMP	322.0443	18.23792	0.426004	1.406998	-0.146	-0.11469	0.393078	0.725155
61	CMP-2-aminoethylphosphonate	429.0562	17.58739	0.76941	1.304754	0.287598	0.819667	0.795357	0.41576
62	CoA	766.1053	15.82273	1.086245	2.05091	0.628939	1.460456	1.306638	0.601512
63	CTP	481.9761	21.26975	0.194197	0.47351	-0.13233	0.832945	0.34208	0.410349
64	cyclic-AMP	328.0457	17.59967	-0.18058	0.252257	-0.80978	-0.50281	-0.31023	0.454548
65	Cyclic ADP-ribose	540.0534	16.17321	0.587392	0.81698	0.372332	0.722862	0.624892	0.19295
66	Cys Phe Phe Gln	542.2121	12.30385	0.429983	0.811129	0.223018	0.265518	0.432412	0.267788
67	Cysteinylglycine	177.0345	16.71567	0.428724	0.849723	0.285234	0.346622	0.477576	0.254967
68	Cytidine	242.0795	16.1937	0.556251	0.542349	0.054504	0.930478	0.520896	0.359161
69	D-4--Phosphopantothenate	298.0695	17.6836	0.262808	0.660326	0.055182	0.442002	0.35508	0.257672
70	D-Fructose	179.0562	15.30254	0.491629	1.458742	1.39904	0.849026	1.049609	0.462265

71	D-Glucose	179.05 59	17.046 31	-0.4593	-0.6045	-0.24358	-0.20937	-0.37919	0.186555
72	D-Glyceraldehyde 3-phosphate	168.99 06	18.837 36	1.042834	1.381513	0.67222	0.942209	1.009694	0.293139
73	D-Histidine	154.06 19	16.994 58	-0.94375	-0.41059	0.13553	-0.13816	-0.33924	0.460565
74	D-Phenylalanine	164.07 15	10.809 66	0.739921	0.910228	0.539248	0.842444	0.75796	0.161744
75	D-Ribonate	165.04	15.586 7	0.472625	1.350942	0.251441	0.484418	0.639856	0.486017
76	dATP	489.99 29	18.156 11	0.63844	1.055833	0.02804	0.951807	0.66853	0.462379
77	dCDP	386.01 5	16.557 09	-0.23984	-0.76101	-0.54905	-0.31291	-0.4657	0.237005
78	dCTP	465.98 14	20.249 64	1.120825	1.374497	0.389361	1.412871	1.074389	0.474712
79	Deoxycytidine	226.08 29	11.133 05	0.91664	1.871524	0.869693	1.244041	1.225475	0.461766
80	Deoxyuridine	227.06 69	15.783 51	-0.31681	0.321253	-0.37315	0.871672	0.125741	0.58861
81	Dihydroorotate	157.02 55	13.249 75	0.578435	1.154922	-0.07406	-0.1707	0.372149	0.618891
82	dihydroxy-acetone-phosphate	168.99 06	17.963 89	-0.56619	-0.0794	0.416509	0.957864	0.182195	0.654491

83	DL-Glutamate	146.0461	17.21433	-0.00291	0.153841	-0.10899	0.143393	0.046333	0.12587
84	DL-Tyrosine	180.0665	14.86191	0.305994	0.77337	0.242023	0.249621	0.392752	0.255344
85	Dodecanoic acid	199.1702	3.521086	-0.05068	-0.14721	0.028153	-0.17296	-0.08568	0.092352
86	dTTP	480.9806	18.68096	0.303209	0.750911	0.382494	1.27383	0.677611	0.442764
87	Dyspropterin	236.0777	14.85107	0.623168	0.933368	-0.24801	0.717627	0.506537	0.519517
88	Ethanolamine phosphate	140.0117	18.25181	0.271988	0.794447	0.212833	0.872928	0.538049	0.343724
89	FAD	784.1474	13.40414	0.092578	0.628085	0.371642	0.776594	0.467225	0.3006
90	Farnesyl pyrophosphate	381.1241	13.19374	3.537935	5.264932	1.972559	1.824041	3.149867	1.609137
91	Folic acid	440.1311	19.42083	1.600153	1.761737	0.365597	0.269796	0.999321	0.790799
92	fructose-1-6-bisphosphate	338.9873	20.92094	-0.43221	-0.2426	0.303454	2.465447	0.523522	1.33165
93	Fructose-1-P	259.0223	17.77271	1.051941	1.451599	0.553101	0.464431	0.880268	0.46038
94	Fructose-6-phosphate	259.0222	18.62672	0.08141	0.543762	0.174343	0.82271	0.405556	0.342369

95	gamma-Glutamyl-S-(1-propenyl)cysteine sulfoxide	305.0808	12.88609	0.612222	1.22445	0.4759	1.256403	0.892244	0.40609
96	Gamma-Glutamylcysteine	249.055	16.64154	0.740073	1.193943	0.535606	1.515498	0.99628	0.442162
97	GDP	442.0163	20.7763	1.786747	0.839522	0.599877	0.610688	0.959209	0.562652
98	GDP-4-Dehydro-6-deoxy-D-mannose	586.0606	16.11578	1.022331	1.009748	0.061422	0.571895	0.666349	0.454424
99	GDP-D-fucose	588.0738	20.41408	0.327247	0.877676	-0.05761	0.218738	0.341514	0.392447
100	GDP-mannose	604.0685	21.01942	0.401778	0.464258	0.08822	0.277686	0.307985	0.165764
101	Gln-Al	216.099	14.55132	-0.12995	-0.00905	-0.15551	0.28488	-0.00241	0.201897
102	Gluconate	195.051	16.61179	1.08164	1.042407	0.526305	1.866359	1.129178	0.552775
103	Glucosamine-1P	258.0385	17.9545	0.096407	0.223192	0.313555	0.616193	0.312337	0.221286
104	Glucose-6-phosphate	259.0222	19.55747	-0.15673	0.085374	0.012326	0.169919	0.027721	0.138811
105	Glutamyl-Glutamate	273.0728	17.37159	1.406832	1.775135	0.885845	1.19694	1.316188	0.373392
106	glutathione	306.0766	16.737	0.716486	1.039453	0.242668	0.601703	0.650077	0.328809

107	Glutathione disulfide	611.1428	19.93672	0.804333	0.394413	0.119427	0.993625	0.577949	0.394953
108	Gly-Gly	131.0461	17.58911	0.202083	0.404127	0.060401	0.672699	0.334828	0.265764
109	Glycerate	105.0192	14.22076	-0.30232	0.914073	-0.01577	0.263806	0.214948	0.520242
110	Glycerophosphoglycerol	245.0431	15.12984	0.836984	1.167556	0.683133	1.096269	0.945986	0.225575
111	GMP	362.059	19.3449	-0.04906	0.652794	-0.72194	-0.9393	-0.264368	0.719365
112	GTP	521.9821	22.27695	0.558159	1.028408	0.432797	0.73758	0.689236	0.258403
113	Heme	559.233	16.801	2.619124	1.465102	0.448059	0.598524	1.282702	0.997333
114	Hydroxyacetone phosphate	152.9957	13.69577	0.438764	1.010689	0.634686	0.412482	0.624155	0.276101
115	Hydroxypropionic acid	89.02451	18.97983	1.245407	1.609612	0.385015	0.831635	1.017917	0.528242
116	Hydroxypyruvic acid	103.0038	17.35525	1.849708	3.206034	2.091402	2.868768	2.503978	0.638812
117	Inosine	267.0724	12.16215	-0.03155	0.029655	-0.20901	-0.83348	-0.261099	0.394796
118	IMP	347.0392	17.92923	0.858988	1.331441	0.569614	1.058462	0.954626	0.321534

119	Isoleucine	130.0874	12.34015	0.533725	0.6649	0.36312	0.809786	0.592883	0.190197
120	Itaconic acid	129.0192	14.89908	-0.30448	0.41437	-0.52098	0.676661	0.066392	0.570401
121	L-2-3-Dihydrodipicolinate	168.0294	17.20502	1.095318	2.551536	3.059259	2.377491	2.270901	0.835388
122	L-2-Amino-3-oxobutanoic acid	116.0353	12.59931	0.944947	1.147214	0.273645	0.846816	0.803156	0.374508
123	L-Alanine	88.04044	16.84886	0.713416	1.179822	0.426826	0.683365	0.750857	0.313562
124	L-Aspartate	132.0303	17.58292	-1.38646	-1.21972	-1.11676	-1.24624	-1.2423	0.111157
125	L-galactono-1-4-lactone	177.04	16.57455	-0.36158	1.607776	1.614601	1.325959	1.046689	0.94843
126	L-Glutamic gamma-semialdehyde	130.0514	16.42995	0.122954	2.140258	-0.73424	0.071616	0.400148	1.224688
127	L-Glutamine	145.0618	17.18333	0.385714	0.289774	0.154503	0.647955	0.369486	0.208471
128	L-Histidinol phosphate	220.0493	15.9413	-0.07165	0.321032	-0.16717	0.378253	0.115118	0.274595
129	L-Lactic acid	89.02448	10.86634	1.279188	1.816764	0.954823	1.48097	1.382936	0.361414
130	L-Ornithine	131.0826	24.27133	0.374867	0.848643	-0.26345	0.506235	0.366587	0.465054
131	L-Pipecolate	128.0716	13.73792	1.299214	1.958111	0.843876	1.338114	1.359829	0.457637

132	Leucine	130.0874	11.6794	0.747734	1.009123	0.522504	0.91669	0.799013	0.213761
133	Linoleate	279.2325	3.23105	-0.24817	-0.69639	-0.1161	-1.38113	-0.61045	0.570661
134	Linoleoyl-CoA	1028.333	4.574313	0.987887	0.768342	0.50712	1.245185	0.877133	0.314364
135	lysine	145.0982	25.07483	1.681581	2.641688	0.073379	1.474445	1.467773	1.059587
136	LysoPG(18:0/0:0)	511.3026	3.336171	1.21126	1.018987	0.48695	0.138485	0.71392	0.490936
137	Malate	133.0144	18.97519	1.244105	1.727678	0.42552	0.724798	1.030526	0.574786
138	Maleamate	114.0197	17.5654	-1.18717	-0.77545	-0.6437	-0.72826	-0.83365	0.241905
139	malonate semialdehyde	87.00832	8.560785	1.497225	2.236796	1.280185	2.086407	1.775153	0.459068
140	Mesaconic acid	129.0192	17.95233	-0.16659	0.602641	4.798846	0.382185	1.404271	2.286045
141	methionine	148.0439	12.82232	0.202276	0.28302	0.173493	0.372694	0.257871	0.089493
142	Methionine sulfoxide	164.0378	15.35264	0.211737	0.535361	0.336127	0.243773	0.331749	0.145625
143	Myo-Inositol	179.0555	19.6805	0.333724	0.510152	0.270328	0.666146	0.445087	0.178926

144	myo-inositol-3-P	259.0222	20.01667	0.175018	0.491472	0.114377	0.253276	0.258536	0.165372
145	N-Acetyl-aspartate	174.0405	17.37467	0.506939	0.953513	-0.4071	0.500788	0.388535	0.571213
146	N-Acetyl-D-fucosamine	204.0878	13.78685	0.585345	0.669136	0.153859	-0.8088	0.149884	0.677827
147	N-Acetyl-D-glucosamine 6-phosphate	300.0481	17.40883	6.62468	6.995202	4.646308	3.114649	5.34521	1.809547
148	N-Acetyl-L-glutamate	188.0566	16.79618	-0.53869	-0.49828	-0.22338	-0.16671	-0.35677	0.188886
149	N-Acetyl-L-methionine	190.054	5.587245	0.950165	1.66211	0.649936	0.928239	1.047613	0.431857
150	N-Acetylaspartylglutamic acid	303.0832	15.5504	1.5525	2.561577	1.017555	1.861205	1.748209	0.644591
151	N-Acetylmannosamine	220.0828	14.95685	4.221896	3.678384	3.286565	4.52687	3.928429	0.553404
152	N-Acetylornithine	173.0931	10.18914	1.859562	2.171926	0.818143	1.039489	1.47228	0.646708
153	N-Carbamoyl-L-aspartate	175.0361	19.77583	0.69358	0.679133	-0.10288	-0.68499	0.146211	0.667472
154	N-Formimino-L-glutamate	173.056	12.59457	1.062943	1.341025	0.682612	0.429419	0.879	0.403317
155	N-Methylhydantoin	113.0354	17.4585	0.745903	0.571896	-0.41335	0.439628	0.336019	0.515084

156	N-N--Diacetylhydrazine	115.0513	16.90614	0.001979	0.36053	-0.1461	-0.02673	0.047421	0.218365
157	N-Succinyl-L-glutamate	246.0619	16.86098	1.33238	1.706196	1.051342	1.376024	1.366485	0.268311
158	N-Succinyl-L-L-2-6-diaminopimelate	289.1038	15.29767	0.575924	1.352664	0.573953	1.346056	0.962149	0.447121
159	Nα-Acetyl-L-glutamine	187.0714	11.03774	1.049182	1.107389	1.271916	0.90137	1.082464	0.153206
160	N2-Succinyl-L-ornithine	231.0978	9.358067	1.343151	1.853718	0.921214	0.932579	1.262666	0.440215
161	NAD+	663.1042	16.04662	0.338788	0.178974	0.461565	1.085238	0.516141	0.396646
162	NADH	664.1155	15.68322	-0.48374	0.176623	-0.2087	0.53913	0.005828	0.446945
163	NADP+	743.0692	19.21533	0.59082	1.108302	2.843418	0.957282	1.374955	1.002797
164	NADPH	744.0818	19.75983	1.304695	0.60888	0.536263	0.746994	0.799208	0.348142
165	Nicotinic acid mononucleotide	335.0406	17.55305	-0.52173	-0.23501	-0.9238	0.693943	-0.246665	0.687761
166	O-Acetyl-L-serine	146.0459	12.56727	0.783835	1.23687	0.319052	0.657113	0.749218	0.379705
167	O-Succinyl-L-homoserine	218.0668	16.52067	1.060916	2.540204	0.782957	1.531763	1.47896	0.772055
168	octanoate	286.2025	3.767033	-0.00788	0.20729	-0.48833	-0.18309	-0.118	0.294006

169	Octanoic acid	143.1076	3.857234	-1.61405	0.634036	0.341202	0.489178	-0.03741	1.05787
170	Octoluse 8/1P	319.0431	19.34933	0.968609	0.993811	0.116053	0.448121	0.631649	0.425921
171	Oleate	281.2483	3.22395	-0.27693	-0.62566	-0.26122	-1.49589	-0.66492	0.578956
172	Ophthalmic acid	288.1198	15.71033	1.029524	1.352568	0.815876	1.437973	1.158985	0.288555
173	Orotate	155.0099	11.80484	0.730978	1.717808	0.140103	0.186342	0.693808	0.7335
174	Orotidine	287.0516	14.88205	-0.93659		-1.33266	-1.22947	-1.16624	0.205467
175	Orotidine 5-phosphate	367.0177	20.2313	2.067332	2.559124	0.619012	0.899714	1.536295	0.926431
176	Palmitic acid	255.2327	3.277046	-0.54468	-1.33778	-0.20826	-1.17138	-0.81552	0.529628
177	Palmityl-CoA	1004.334	4.588634	0.496662	0.479393	0.124494	0.829095	0.482411	0.287812
178	Pantetheine 4--phosphate	357.0885	11.52454	1.46013	2.083835	0.772586	1.851839	1.542098	0.573959
179	Pantothenic acid	218.1036	8.7806	0.513674	0.660553	0.561747	0.88208	0.654514	0.163569

180	PE(16:0/0:0)	452.2773	3.8991	-0.22837	-0.04442	-0.13073	0.254474	-0.03726	0.208503
181	phenylpyruvate	163.0403	14.80093	1.448616	0.462263	0.264405	0.149999	0.581321	0.59241
182	Phosphate	96.96987	18.55239	-0.01151	-0.04869	-0.02464	0.450979	0.091534	0.240124
183	Phosphodimethylethanolamine	168.0429	16.98917	0.170149	0.206497	0.30755	0.848165	0.38309	0.315452
184	Phosphoenolpyruvate	166.9752	20.47893	0.014679	0.154513	-0.01433	0.165346	0.080053	0.093096
185	Phosphonoacetaldehyde	122.9854	17.03626	0.097881	0.579046	0.143209	0.928035	0.437043	0.392686
186	Phosphoribosyl-AMP	558.0638	16.15455	0.60249	1.228076	0.467601	1.734455	1.008156	0.586697
187	Pro Glu Phe	390.1696	4.163484	0.887376	2.116165	1.55218	0.68708	1.3107	0.651969
188	Proline	114.0561	14.30817	0.454494	1.046584	0.278053	0.626242	0.601343	0.32911
189	PS-Lipid	836.5389	3.111483	-0.41468	-0.21272	-2.32596	-0.16825	-0.7804	1.035935
190	PS(18:0/0:0)	524.2966	3.428083	0.151804	0.29618	-0.31034	0.169268	0.076729	0.265944
191	Pseudouridine 5--phosphate	323.0282	19.17512	0.279648	0.519911	-0.04422	0.347662	0.27575	0.236066
192	Pyridoxine	168.0664	7.428357	0.487236	0.750562	0.251386	0.54236	0.507886	0.205182

193	Pyridoxine 5--phosphate	248.0331	16.15506	0.536841	1.015385	0.364803	0.746288	0.665829	0.280427
194	Pyroglutamic acid	128.0353	17.22755	0.098423	0.331841	-0.0277	0.277991	0.170139	0.165391
195	Pyrophosphate	176.9363	18.5524	0.006491	-0.16913	-0.1032	0.361289	0.023862	0.236326
196	Ribose	149.0454	14.2761	-0.05554	1.336115	-0.1556	-0.12285	0.25053	0.724921
197	Ribose-5-Phosphate	229.0122	18.36192	0.081426	0.246604	0.143404	0.270556	0.185498	0.088641
198	S-(2-Hydroxyethyl)glutathione	350.1028	16.53906	1.138082	1.057076	-0.0643	0.71356	0.711104	0.548715
199	S-(2-Succinyl)glutathione	422.0868	21.37333	0.926052	1.314687	0.010598	0.334013	0.646338	0.584991
200	S-Adenosyl-L-homocysteine	383.114	15.21284	0.8545	1.319438	0.372124	0.744837	0.822725	0.390231
201	S-Inosyl-L-homocysteine	384.0992	17.288	0.495084	0.899046	0.993793	1.578958	0.991721	0.447247
202	S-Methyl GSH	320.092	15.09923	0.387686	1.193807	0.391369	1.270618	0.81087	0.487536
203	Sarcosine	88.0404	15.78723	0.281189	0.599688	0.561388	0.795668	0.559481	0.212015
204	Sedoheptulose 7-phosphate	289.0325	18.9432	-0.35057	0.664362	0.76229	0.632603	0.427172	0.521422
205	Ser Asp Tyr	382.1279	12.36332	0.984116	1.466728	0.974413	1.122579	1.136959	0.230027

206	sn-glycero-3-Phosphoethanolamine	214.0481	17.85393	0.572872	0.493651	0.06575	0.546682	0.419739	0.238282
207	sn-Glycerol 3-phosphate	171.0059	17.51014	1.942317	2.459661	0.78722	1.164709	1.588479	0.754043
208	Sorbitol	181.0716	15.93577	-0.16948	0.471862	-0.29257	0.176153	0.046492	0.346101
209	Sorbitol 6-phosphate	261.038	17.80467	1.864118	2.315153	0.57977	0.826974	1.396504	0.827445
210	sphinganine 1-phosphate	380.2557	4.3853	0.163879	0.321673	0.201478	1.022156	0.427296	0.402243
211	Succinate	117.0194	18.02894	1.067147	1.495579	0.571994	1.104818	1.059884	0.378567
212	Succinic acid semialdehyde	101.0245	18.62831	0.590458	1.281023	0.545805	0.821912	0.8098	0.336651
213	Succinoadenosine	462.0658	20.97133	-0.49573	-0.07679	-0.09384	-0.02063	-0.17175	0.218243
214	Tetradecanoic acid	227.2012	3.292817	-0.57964	-0.31402	0.235332	-0.49448	-0.2882	0.36617
215	Threonate/Erythronate	135.0299	15.02476	-0.12934	0.836085	-0.02695	0.062755	0.185636	0.440677
216	threonine	118.0506	16.61755	0.978586	1.282426	0.626818	1.182313	1.017536	0.289538
217	UDP	402.9941	19.24154	1.025286	-0.04222	0.007134	0.45106	0.360314	0.495713
218	UDP-D-glucose	565.0464	19.10383	0.295154	0.627118	0.02435	0.579626	0.381562	0.279639

219	UDP-N-acetyl-D-glucosamine	606.0728	17.72981	6.471388	6.90763	6.254382	6.503308	6.534177	0.272426
220	UMP	323.0284	17.69289	-0.75505	-0.33699	-0.96614	-0.84368	-0.72546	0.273062
221	Uridine	243.0619	16.79033	-0.14923	0.310004	-0.44071	-0.42602	-0.17649	0.35095
222	Uridine diphosphate glucuronic acid	579.0251	22.15642	1.826655	2.616627	1.11272	3.362599	2.22965	0.973527
223	UTP	482.961	20.7221	0.862504	1.482242	0.436681	1.846189	1.156904	0.628848
224	Valine	116.0717	13.94702	0.763006	1.329502	0.519071	0.957749	0.892332	0.34227
225	Xanthine	151.0264	13.20648	-0.81748	-0.65258	-0.4864	-0.73886	-0.67383	0.141948
226	XMP	363.0342	20.83739	0.516989	0.98162	0.319211	0.712355	0.632544	0.282698

3.8 References

1. Cobbold, S. A. *et al.* Non-canonical metabolic pathways in the malaria parasite detected by isotope-tracing metabolomics. *Mol. Syst. Biol.* **17**, (2021).
2. World Health Organization. *World Malaria Report 2019*. (2019).
3. Antonova-Koch, Y. *et al.* Open-source discovery of chemical leads for next-generation chemoprotective antimalarials. *Science (80-.).* **362**, (2018).
4. Hooft van Huijsduijnen, R. & Wells, T. N. The antimalarial pipeline. *Current Opinion in Pharmacology* **42**, 1–6 (2018).
5. Cowell, A. N. *et al.* Mapping the malaria parasite druggable genome by using in vitro evolution and chemogenomics. *Science (80-.).* **359**, 191–199 (2018).
6. Ross, L. S. & Fidock, D. A. Elucidating Mechanisms of Drug-Resistant Plasmodium falciparum. *Cell Host and Microbe* **26**, 35–47 (2019).
7. Roth, E. J. Plasmodium falciparum carbohydrate metabolism: a connection between host cell and parasite - PubMed. *Blood Cells* **16**, 453–466 (1990).
8. Atamna, H., Pascarmona, G. & Ginsburg, H. Hexose-monophosphate shunt activity in intact Plasmodium falciparum-infected erythrocytes and in free parasites. *Mol. Biochem. Parasitol.* **67**, 79–89 (1994).
9. Liu, J., Istvan, E. S., Gluzman, I. Y., Gross, J. & Goldberg, D. E. Plasmodium falciparum ensures its amino acid supply with multiple acquisition pathways and redundant proteolytic enzyme systems. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 8840–8845 (2006).
10. Olszewski, K. L. *et al.* Host-Parasite Interactions Revealed by Plasmodium falciparum Metabolomics. *Cell Host Microbe* **5**, 191–199 (2009).
11. Fatumo, S. *et al.* Estimating novel potential drug targets of Plasmodium falciparum by

analysing the metabolic network of knock-out strains in silico. *Infect. Genet. Evol.* **9**, 351–358 (2009).

12. Plata, G., Hsiao, T. L., Olszewski, K. L., Llinás, M. & Vitkup, D. Reconstruction and flux-balance analysis of the Plasmodium falciparum metabolic network. *Mol. Syst. Biol.* **6**, (2010).
13. Bazzani, S., Hoppe, A. & Holzhütter, H. G. Network-based assessment of the selectivity of metabolic drug targets in Plasmodium falciparum with respect to human liver metabolism. *BMC Syst. Biol.* **6**, (2012).
14. Tymoshenko, S., Oppenheim, R. D., Soldati-Favre, D. & Hatzimanikatis, V. Functional genomics of plasmodium falciparum using metabolic modelling and analysis. *Brief. Funct. Genomics* **12**, 316–327 (2013).
15. Oppenheim, R. D. *et al.* BCKDH: The Missing Link in Apicomplexan Mitochondrial Metabolism Is Required for Full Virulence of Toxoplasma gondii and Plasmodium berghei. *PLoS Pathog.* **10**, (2014).
16. Ke, H. *et al.* Genetic investigation of tricarboxylic acid metabolism during the plasmodium falciparum life cycle. *Cell Rep.* **11**, 164–174 (2015).
17. Linster, C. L., Van Schaftingen, E. & Hanson, A. D. Metabolite damage and its repair or pre-emption. *Nature Chemical Biology* **9**, 72–80 (2013).
18. Bommer, G. T., Van Schaftingen, E. & Veiga-da-Cunha, M. Metabolite Repair Enzymes Control Metabolic Damage in Glycolysis. *Trends in Biochemical Sciences* **45**, 228–243 (2020).
19. Dumont, L. *et al.* The metabolite repair enzyme phosphoglycolate phosphatase regulates central carbon metabolism and fosmidomycin sensitivity in plasmodium falciparum. *MBio*

10, e02060-19 (2019).

20. Fuhrer, T., Zampieri, M., Sévin, D. C., Sauer, U. & Zamboni, N. Genomewide landscape of gene–metabolome associations in Escherichia coli . *Mol. Syst. Biol.* **13**, 907 (2017).
21. Bushell, E. *et al.* Functional Profiling of a Plasmodium Genome Reveals an Abundance of Essential Genes. *Cell* **170**, 260-272.e8 (2017).
22. Zhang, M. *et al.* Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360**, DOI: 10.1126/science.aap7847 (2018).
23. Creek, D. J. *et al.* Toward global metabolomics analysis with hydrophilic interaction liquid chromatography-mass spectrometry: Improved metabolite identification by retention time prediction. *Anal. Chem.* **83**, 8703–8710 (2011).
24. Mahieu, N. G. & Patti, G. J. Systems-Level Annotation of a Metabolomics Data Set Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. *Anal. Chem.* **89**, 10397–10406 (2017).
25. Wang, L. *et al.* Peak Annotation and Verification Engine for Untargeted LC-MS Metabolomics. *Anal. Chem.* **91**, 1838–1846 (2019).
26. Huang, X. *et al.* X13CMS: Global tracking of isotopic labels in untargeted metabolomics. *Anal. Chem.* **86**, 1632–1639 (2014).
27. Sévin, D. C., Fuhrer, T., Zamboni, N. & Sauer, U. Nontargeted in vitro metabolomics for high-throughput identification of novel enzymes in Escherichia coli. *Nat. Methods* **14**, 187–194 (2017).
28. Srivastava, A., Evans, K. J., Sexton, A. E., Schofield, L. & Creek, D. J. Metabolomics-Based Elucidation of Active Metabolic Pathways in Erythrocytes and HSC-Derived

Reticulocytes. *J. Proteome Res.* **16**, 1492–1505 (2017).

29. Wilde, M. L. *et al.* Protein kinase A is essential for invasion of Plasmodium falciparum into human erythrocytes. *MBio* **10**, (2019).

30. Gulati, S. *et al.* Profiling the Essential Nature of Lipid Metabolism in Asexual Blood and Gametocyte Stages of Plasmodium falciparum. *Cell Host Microbe* **18**, 371–381 (2015).

31. Bullen, H. E. *et al.* Phosphatidic Acid-Mediated Signaling Regulates Microneme Secretion in Toxoplasma. *Cell Host Microbe* **19**, 349–360 (2016).

32. Kuznetsova, E. *et al.* Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family. *J. Biol. Chem.* **281**, 36149–61 (2006).

33. Stanway, R. R. *et al.* Genome-Scale Identification of Essential Metabolic Processes for Targeting the Plasmodium Liver Stage. *Cell* **179**, 1112-1128.e26 (2019).

34. Locasale, J. W. New concepts in feedback regulation of glucose metabolism. *Current Opinion in Systems Biology* **8**, 32–38 (2018).

35. Sauro, H. M. Control and regulation of pathways via negative feedback. *Journal of the Royal Society Interface* **14**, (2017).

36. Guggisberg, A. M. *et al.* A sugar phosphatase regulates the methylerythritol phosphate (MEP) pathway in malaria parasites. *Nat. Commun.* **5**, 4467 (2014).

37. Guggisberg, A. M. *et al.* Suppression of Drug Resistance Reveals a Genetic Mechanism of Metabolic Plasticity in Malaria Parasites. *MBio* **9**, e01193-18 (2018).

38. Kuznetsova, E. *et al.* Functional Diversity of Haloacid Dehalogenase Superfamily Phosphatases from Saccharomyces cerevisiae: BIOCHEMICAL, STRUCTURAL, AND EVOLUTIONARY INSIGHTS. *J. Biol. Chem.* **290**, 18678–98 (2015).

39. Park, J., Guggisberg, A. M., Odom, A. R. & Tolia, N. H. Cap-domain closure enables

diverse substrate recognition by the C2-type haloacid dehalogenase-like sugar phosphatase Plasmodium falciparum HAD1. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1824–34 (2015).

40. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D. & Aravind, L. Evolutionary Genomics of the HAD Superfamily: Understanding the Structural Adaptations and Catalytic Diversity in a Superfamily of Phosphoesterases and Allied Enzymes. *J. Mol. Biol.* **361**, 1003–1034 (2006).

41. Alexandrov, A. *et al.* A facile method for high-throughput co-expression of protein pairs. *Mol. Cell. Proteomics* **3**, 934–938 (2004).

42. Verma, M. *et al.* A short translational ramp determines the efficiency of protein synthesis. *Nat. Commun.* **10**, 5774 (2019).

43. Cobbold, S. A. *et al.* Metabolic Dysregulation Induced in *Plasmodium falciparum* by Dihydroartemisinin and Other Front-Line Antimalarial Drugs. *J. Infect. Dis.* **213**, 276–286 (2016).

44. Creek, D. J. *et al.* Metabolomics-Based Screening of the Malaria Box Reveals both Novel and Established Mechanisms of Action. *Antimicrob. Agents Chemother.* **60**, 6650–6663 (2016).

45. Bird, S. S., Marur, V. R., Sniatynski, M. J., Greenberg, H. K. & Kristal, B. S. Lipidomics profiling by high-resolution LC-MS and high-energy collisional dissociation fragmentation: Focus on characterization of mitochondrial cardiolipins and monolysocardiolipins. *Anal. Chem.* **83**, 940–949 (2011).

46. Saunders, E. C. *et al.* Isotopomer profiling of Leishmania mexicana promastigotes reveals important roles for succinate fermentation and aspartate uptake in Tricarboxylic Acid

Cycle (TCA) anaplerosis,,glutamate synthesis, and growth. *J. Biol. Chem.* **286**, 27706–27717 (2011).

47. Tsugawa, H. *et al.* MS-DIAL: Data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* **12**, 523–526 (2015).

48. Dagley, M. J. & McConville, M. J. DExSI: A new tool for the rapid quantitation of 13 C-labelled metabolites detected by GC-MS. *Bioinformatics* **34**, 1957–1958 (2018).

49. Xia, J. & Wishart, D. S. Using metaboanalyst 3.0 for comprehensive metabolomics data analysis. *Curr. Protoc. Bioinforma.* **2016**, 14.10.1-14.10.91 (2016).

50. Huthmacher, C., Hoppe, A., Bulik, S. & Holzhütter, H.-G. Antimalarial drug targets in Plasmodium falciparum predicted by stage-specific metabolic network analysis. *BMC Syst. Biol.* **4**, 120 (2010).

51. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **44**, W272–W276 (2016).

52. Prommana, P. *et al.* Inducible knockdown of Plasmodium gene expression using the glmS ribozyme. *PLoS One* **8**, e73783 (2013).

Chapter 4: Enzymatic and structural characterization of HAD5, an essential phosphomannomutase of malaria parasites

Preface

The following work has been submitted for publication at the *Journal of Biological Chemistry* and is currently under review. The authors are Philip M. Frasse, Justin J. Miller, Ebrahim Soleimani, Jian-She Zhu, David L. Jakeman, Joseph M. Jez, Daniel E. Goldberg, and Audrey R. Odom John. PMF designed and executed the majority of experiments with the following exceptions: ES, JSZ, and DLJ synthesized and provided phosphomannomutase inhibitors. JJM and JMJ assisted with X-ray crystallography and structure deposition, and JJM conducted *in silico* computational docking.

4.1 Abstract

The malaria parasite *Plasmodium falciparum* is responsible for over 200 million infections and 400,000 deaths per year. At multiple stages during its complex life cycle, *P. falciparum* expresses several essential proteins tethered to its surface by glycosylphosphatidylinositol (GPI) anchors, which are critical for biological processes such as parasite egress and reinvasion of host red blood cells. Targeting this pathway therapeutically has the potential to broadly impact parasite development across several life stages. Here, we characterize an upstream component of GPI anchor biosynthesis, the putative phosphomannomutase (EC 5.4.2.8) of the parasites, HAD5 (PF3D7_1017400). We confirm the phosphomannomutase and phosphoglucomutase activity of purified recombinant HAD5. By regulating expression of HAD5 in transgenic parasites, we demonstrate that HAD5 is required for malaria parasite egress and erythrocyte reinvasion. Finally, we determine the three-dimensional crystal structure of HAD5 and identify a substrate analog that specifically inhibits

HAD5, compared to orthologous human phosphomannomutases. These findings demonstrate that the GPI anchor biosynthesis pathway is exceptionally sensitive to inhibition, and that HAD5 has potential as a multi-stage antimalarial target.

4.2 Introduction

Malaria remains an enormous global health burden for much of the world, resulting in over 200 million infections and 400,000 deaths every year, the majority of which are in children under the age of five[1]. One of the primary barriers to effective malaria treatment and control is the emergence of resistance to all approved antimalarial chemotherapeutics[2,3], prompting an urgent call for the development of new therapies and the identification of novel drug targets. Malaria is caused by apicomplexan parasites of the genus *Plasmodium*, primarily the species *Plasmodium falciparum*. *P. falciparum* has a complex life cycle, in which parasites develop in a mosquito host, are deposited into a human during a mosquito blood meal, and migrate to the liver where they infect hepatocytes. Parasites are released from the liver into the bloodstream and begin an asexual cycle of replication within red blood cells, occasionally branching off into sexual stage gametocytes that can be taken up by mosquitoes to start the cycle anew[4].

This complex, multi-host life cycle has stymied efforts to develop therapeutics and vaccines to eradicate this disease, as it has been difficult to identify effective vaccine and therapeutic targets that span multiple life stages. In recent years, the push for new therapeutics and vaccines has focused on a strategy of developing transmission-blocking vaccines and therapies, which impair the development or viability of gametocytes or target the mosquito vector itself, thus preventing vector-borne transmission[5–7]. Our goal is therefore to identify novel targets for antimalarial therapeutics that are not only essential for intraerythrocytic growth of the

parasite, but are also essential for sexual- and/or mosquito-stage parasites, indicating their potential in transmission blocking strategies.

Malaria parasites are highly metabolically active[8,9], and several known antimalarials target unique and essential metabolic processes in the parasites[10–14]. Metabolic enzymes have great potential for chemical inhibition, as substrate analogs can be rationally designed and developed as potential inhibitors, making these enzymes well-suited as "druggable" targets[15]. We therefore sought a metabolic enzyme that is expressed and essential during multiple life stages of *Plasmodium falciparum*. These constraints narrowed our search to the upstream steps of glycosyl phosphatidylinositol (GPI) anchor synthesis.

GPI anchors are an essential component of all life stages of *Plasmodium falciparum*. In intraerythrocytic parasites, these glycolipid anchors tether several essential proteins to the parasite plasma membrane prior to red blood cell egress and reinvasion[16]. Most abundant amongst these GPI-anchored proteins (GPI-APs) is merozoite surface protein 1 (MSP1)[16–18]. Proper MSP1 localization and processing is necessary for parasites to egress from the erythrocyte, and mature MSP1 anchored to the surface of free merozoites facilitates the binding and invasion of new red blood cells. In the absence of the GPI-anchoring C-terminus of MSP1, parasites are defective in their ability to egress[19], and antibodies developed against MSP1 prevent merozoite reinvasion[20]. Several other GPI-APs are also involved in parasite egress and invasion, including other merozoite surface proteins (MSPs) and rhoptry-associated membrane antigen (RAMA)[16,18,21]. GPI anchored proteins are also expressed in other life stages of *P. falciparum*. Gamete- and ookinete-stage GPI anchored proteins include Pfs25 and Pfs230, which are considered as possible vaccine candidates[22–25], while circumsporozoite protein (CSP), the critical antigen of the RTS,S and R21 vaccines, is an essential GPI-anchored protein of the sporozoite

stages[26–30]. Thus, it is clear that successful targeting of GPI anchor biosynthesis would not only effectively treat symptomatic blood-stage infection, but may also block transmission of the parasites at multiple stages.

One enzyme of the GPI anchor biosynthesis pathway that has yet to be characterized in *P. falciparum* is the putative phosphomannomutase, HAD5 (PMM; PF3D7_1017400). Phosphomannomutases (PMMs) are responsible for the conversion of mannose 6-phosphate (M6P) to mannose 1-phosphate (M1P), the precursor to GDP-mannose. GDP-mannose is then converted to dolichol-p-mannose, which is the building block for incorporating mannose into glycolipids. In asexual *P. falciparum*, the dominant mannosylated glycolipids are GPI anchors[31], as *N*-glycans in these parasites only contain *N*-acetyl glucosamine[32–34], so targeting of mannose metabolism is predicted to specifically inhibit GPI anchor synthesis. HAD5 is also predicted to be essential[35], and transcriptomic studies show its expression during the blood stage[36,37] and sexual stages[38], making it a potential multi-stage antimalarial drug target. In this study, we characterize the putative phosphomannomutase of *P. falciparum*, HAD5, demonstrating its essentiality for parasite growth and its potential for specific targeting by future antimalarial therapies.

4.3 Results

4.3.1 HAD5 is a phosphomannomutase

HAD5 (PF3D7_1017400) has been annotated as a putative phosphomannomutase through homology to other known phosphomannomutases. Phosphomannomutases are responsible for interconverting M6P and M1P to generate M1P for downstream glycolipid production, most notably GPI anchors in *P. falciparum* (Fig. 1A)[39–41]. To determine the

biochemical function of HAD5, we purified recombinant HAD5 (Fig. 1B) and examined its hexose phosphate mutase activities (Fig. 1C, S1). We found that HAD5 was active in the phosphomannomutase assay (0.56 ± 0.48 µmol/min/mg), with even more robust activity (7.67 ± 0.06 µmol/min/mg) upon addition of the known co-factor glucose-1,6-bisphosphate (G-1,6-P)[41], which is comparable to the specific activity of other phosphomannomutases[42,43]. Furthermore, as has been seen for other phosphomannomutases[43,44], HAD5 exhibits some promiscuity in its substrate preference, as it also displays phosphoglucomutase activity at 0.31 ± 0.03 µmol/min/mg. Comparing the catalytic efficiencies of HAD5 toward phosphomannose and phosphoglucose suggests that phosphomannomutase activity is the dominant enzymatic function of HAD5, with a roughly 4-fold higher catalytic efficiency (k_{cat}/K_m) compared to its phosphoglucomutase activity (Table 1).

4.3.2 HAD5 is essential for intraerythrocytic parasite growth

To assess the essentiality of HAD5 in intraerythrocytic parasite stages, we used a previously described conditional knockdown system in cultured asexual *P. falciparum*[45,46] (Fig. 2A). We placed a Tet repressor-binding aptamer array at the 3'- and 5'-end of the endogenous HAD5 locus by CRISPR/Cas9-mediated integration. In parasite integrants, the presence of anhydrotetracycline (aTc) promotes HAD5 translation, while washing out aTc leads to inhibition of translation, and we term this conditional knockdown strain "HAD5^{KD}". Immunoblotting confirms substantial reduction in cellular abundance of HAD5 in the absence of aTc (Fig. 2B). In HAD5^{KD} parasites, 0 nM aTc conditions led to an absence of growth, whereas addition of aTc promoted growth in a dose-dependent manner (Fig. 2C), indicating that HAD5 is essential for asexual growth of *P. falciparum*.

Because phosphosugar mutases often utilize more than one substrate, we examined whether the phosphomannomutase activity of HAD5 was responsible for its essential function in asexual parasites. Attempts to chemically rescue parasite growth with hexose phosphates such as M6P and M1P were unsuccessful (Fig. S2A), as was expected due to the impermeability of the erythrocyte and parasite membranes to such highly charged compounds. However, we found that simple chemical supplementation of the media with D-mannose was sufficient to rescue growth when HAD5 expression is reduced. This indicates that the primary mechanism of death in these parasites is due to defects in mannose metabolism (Fig. 2D, S2B). Notably, while artificially elevated concentrations of D-mannose completely rescued parasite growth, a physiologically relevant concentration of 50 µM [equivalent to that of human serum[47]] did not provide a statistically significant rescue of parasite growth.

4.3.3 HAD5 is required for parasite egress and invasion

GPI anchors, which contain mannose[16,18,31], are required for egress of the malaria parasite from the infected host cell, as well as reinvasion of merozoites[46,48–55]. Because of its essential role in mannose metabolism, we hypothesized that HAD5 may be required for efficient egress and invasion. To evaluate this possibility, we washed out aTc at the beginning of the life cycle in synchronized parasites. Parasites grown in -aTc conditions over the course of one life cycle developed morphologically normally through the majority of life cycle stages, including the development of multi-nucleated schizonts. Whereas +aTc conditions allowed parasites to continue into the next life cycle and form newly reinvaded “ring”-stage parasites, -aTc parasites were arrested in late schizogony (Fig. 2E), which suggests a defect in parasite egress when HAD5 expression is knocked down. Notably, when HAD5 expression is reduced, parasites retain

normal mature schizont architecture by transmission electron microscopy, indicating that neither gross developmental defects nor structural aberrations prevent parasites from egressing (Fig. S3).

To examine whether HAD5 is required for invasion, as well as egress, we examined the reinvasion capacity of HAD5^{KD} parasites. Segmented schizont-enriched cultures of HAD5^{KD} ±aTc were mechanically lysed and the freed merozoites were allowed to reinvade fresh RBCs. Under these conditions, knockdown parasites were unable to reinvade new host cells (Fig. S4). Thus, the knockdown of HAD5 confers defects to parasite biology that prevent both egress and reinvasion of the parasites.

4.3.4 HAD5 knockdown disrupts GPI anchor synthesis

Mannose metabolism is linked to parasite egress through biosynthesis of GPI anchors[56]. In *P. falciparum*, GPI anchors are synthesized through addition of one glucosamine (GlcN) and three to four mannose residues to a phosphatidylinositol backbone. These mannose residues are derived from the product of phosphomannomutase, M1P, which is converted to GDP-mannose and subsequently to dolichol-phosphate mannose, the direct mannose donor to GPI anchors (Fig. 3A)[57]. Several GPI-APs contribute to egress and invasion of parasites[16,19,21]. We therefore hypothesized that reduced HAD5 expression leads to loss of phosphomannomutase activity and causes parasite death by disruption of GPI anchor biosynthesis. To directly evaluate the effect of HAD5 knockdown on GPI anchor biosynthesis, we labeled mid- to late-trophozoite parasites with [^{3}H]-GlcN and extracted GPI precursors as previously described[18,56,58,59]. HAD5^{KD} parasites grown in +aTc conditions had the expected repertoire of GPI anchor precursors (Fig. 3B, 3C). A variety of precursors are observed, with earlier, less polar species (with fewer mannose groups) migrating further than more polar, highly-mannosylated species. When HAD5 expression is reduced, there is a relative accumulation of the earlier precursors, as well as a reduced

production of fully mature, highly mannosylated precursors (Fig. 3B), indicating a defect in GPI anchor biosynthesis. In particular, there was a significant reduction of the highly polar band 9 and significant buildup of less polar band 4 when HAD5 expression was reduced (Fig. 3C). Intriguingly, despite the substantial knockdown of HAD5 and the complete loss of growth in these parasites, the abundance of many mannosylated precursors is unchanged and highly mannosylated GPI precursors are still observed, suggesting that this biosynthetic pathway is not completely ablated.

To confirm the role of HAD5 in GPI biosynthesis, we deployed two established chemical tools, mannosamine (ManN) and GlcN, which impair growth of *P. falciparum* through inhibition of GPI anchor biosynthesis[60,61]. We expected that if HAD5 knockdown disrupts GPI anchor biosynthesis, parasites should be hypersensitized to ManN and GlcN treatment. Indeed, when expression of HAD5 is reduced with an intermediate concentration of aTc (3 nM) that still permits modest asexual growth, knockdown parasites yielded a marked shift in half-maximal effective concentration (EC_{50}) (Fig. 3D, 3E). These results further implicate HAD5 in the production of GPI anchors in *P. falciparum*.

4.3.5 HAD5-dependent GPI anchor synthesis enables proper anchoring of MSP1

Reduced GPI anchor biosynthesis in malaria parasites is expected to impact the localization and function of a number of essential GPI-anchored parasite proteins. While several GPI-anchored proteins have been characterized in *P. falciparum* intraerythrocytic stages, the most abundant is MSP1[16,17]. MSP1 must be targeted and anchored through GPIs and proteolytically processed in order for schizont-stage parasites to egress from the erythrocyte, and the MSP1 complex is also critical for binding and reinvading new red blood cells[19,62]. For this

reason, we investigated whether HAD5-dependent GPI anchor synthesis is required for localization and anchoring of MSP1. We expected that, when the pathway is intact, MSP1 is successfully anchored to the parasite plasma membrane. In contrast, when HAD5 expression is knocked down, GPI anchors will fail to fully incorporate mannose and GPI-anchored proteins, including MSP1, will remain untethered to the membrane (Fig. 3A). To evaluate this effect, we used immunofluorescence to detect the localization of MSP1. When schizonts grown in ±aTc conditions were mechanically lysed and the resultant merozoites were imaged, there was a modest but significant decrease in MSP1 signal surrounding the daughter merozoites when HAD5 expression is reduced (Fig. 4A, 4B), indicating that MSP1 membrane attachment is diminished, causing it to diffuse away from the cell.

To independently confirm this finding, we partitioned lysate from early schizont-stage parasites into membranous and soluble fractions[63]. Whole lysate and fractions were assessed by immunoblotting for MSP1, and the relative proportion of MSP1 in the membrane fraction compared to lysate was calculated. In +aTc conditions, a large proportion of MSP1 separates with the detergent-enriched phase, with some full-length MSP1 in the soluble phase. Upon HAD5 knockdown, there was a subtle reduction in the proportion of cellular MSP1 bound to the membrane (Fig. S5). This difference did not reach significance; however, substantial biological variability in MSP1 expression is present within merozoite populations, as indicated in Figure 4B, and post-hoc power analysis of this data indicates that we were underpowered to observe a significant difference (Power = 0.290)[64]. Together with our analysis of GPI anchor biosynthesis in these parasites (Fig. 3), these data confirm some functional preservation of GPI biosynthesis when HAD5 expression is reduced.

4.3.6 Disruption of mannose metabolism by HAD5 knockdown does not affect fosmidomycin sensitivity in parasites

HAD5 and other phosphomannomutases are part of the haloacid dehalogenase (HAD) superfamily of proteins (Interpro: IPR023214), a ubiquitous family of enzymes that primarily conduct phosphatase and phosphotransferase reactions[65]. HAD5 is a member of subfamily IIB (IPR006379) of the HAD superfamily and has substantial sequence homology to other *P. falciparum* HAD proteins in this subfamily (HAD1 and HAD2) and the related subfamily IIA (IPR006357) protein, phosphoglycolate phosphatase (PGP)[66]. One notable commonality between these subfamily II *P. falciparum* HAD proteins is their effect on the parasite's sensitivity to the antimalarial fosmidomycin (FSM), which is a well-validated inhibitor of the apicoplast methylerythritol phosphate pathway of isoprenoid biosynthesis[10]. Mutations in either HAD1[67] or HAD2[68] render parasites resistant to FSM, and PGP-knockout parasites are hypersensitive to FSM[69]. To examine whether HAD5 also plays a role in FSM susceptibility, we assessed the FSM dose-response of parasites grown in saturating or intermediate aTc conditions. Unlike its close homologs, HAD5 knockdown had no effect on the FSM EC_{50} of parasites (Fig. S6).

4.3.7 HAD5 is distinct from human PMMs and can be specifically inhibited

Previous studies have investigated egress and invasion as a promising target for antimalarial drug discovery, suggesting that HAD5 may likewise be of interest for antimalarial drug discovery[46,48–51]; however, phosphomannomutases are found widely throughout nature, including two genes in the human genome, *Hs*PMM1 and *Hs*PMM2[70,71]. We therefore evaluated the potential for selective inhibition of *P. falciparum* HAD5 over human PMM1 and PMM2 (Fig. S7A). Previous work has successfully demonstrated the use of substituted ketoheptoses and other phosphosugar analogues as inhibitors of microbial phosphomannomutases[72,73], which we

sought to replicate with HAD5. Using a panel of 11 phosphosugar analogues (Fig. S8A), we screened each compound for its ability to inhibit recombinant HAD5, PMM1, and PMM2. The majority of the compounds had negligible inhibition against all three enzymes (Fig. S8). Compound D9, however, inhibits purified recombinant HAD5 with a half-maximal inhibitory concentration of 79 ± 2.6 μM, several-fold more potently than the inhibition of either PMM1 or PMM2 (Fig. 5A). Moreover, we find dramatic time-dependent effects on the ability of D9 to inhibit HAD5, as pre-incubating HAD5 with D9 prior to assaying activity substantially increased D9 potency, such that a 60-minute pre-incubation yielded HAD5 activity of only 4.5% of a vehicle-treated control (Fig. 5B). This effect was not seen for *Hs*PMM1, demonstrating that the potential to specifically inhibit HAD5 may be even greater under ideal binding conditions. As expected given its poor drug-like characteristics (and likely inadequate cellular permeability), compound D9 did not impair the growth of asexual *P.* falciparum at concentrations up to 100 μM (Fig. S9A). The selective inhibition of HAD5 by D9 is an important proof-of-concept that distinct structural features of HAD5 may be harnessed for parasite-specific inhibitor development.

To uncover the structural basis for HAD5-specific inhibition, we solved the crystal structure of HAD5 to 3.5 Å resolution. Like other known phosphomannomutases and members of the HAD superfamily[65,74], HAD5 comprises two domains with an overall fold similar to that of *Hs*PMM1 (Fig. 5C, Table S1). There were no large structural differences between the human and parasite enzymes, so we next evaluated the substrate-binding pocket by computationally docking D9 onto the HAD5 crystal structure (Fig. S9B). Although the overall binding pockets were similar and the amino acid sequence within the pocket are highly conserved, subtle differences in side-chain positioning may account for the parasite specificity of D9. Further

analysis of the binding pocket is restricted by the current resolution of 3.5Å. Furthermore, comparing D9 to the inactive compounds D7 and D11, the singular difference is the presence of a hydroxyaminomethyl group on carbon 6 of compound D9, whereas an aminomethyl group or hydroxymethyl group are present on D7 and D11, respectively (Figure S8A), a difference that may explain the disparity in potency and specificity.

4.4 Discussion

We report here the lethal knockdown and biochemical characterization of the phosphomannomutase in *P. falciparum*, HAD5. Loss of HAD5 leads to growth arrest in asexually replicating parasites, marked by defects in egress and reinvasion. This growth defect can be rescued by media supplementation with D-mannose, indicating that disruption of mannose metabolism is the primary mechanism of death in HAD5^{KD} parasites; however, a physiologically relevant concentration of 50 µM D-mannose is unable to significantly rescue growth, bolstering the case for this pathway as a therapeutic target. We further report the specific inhibition of HAD5 enzymatic activity compared to orthologous human phosphomannomutases by the hexose-phosphate analogue, compound D9, highlighting the potential for specific therapeutics to be developed to HAD5.

Unexpectedly, despite the dramatic decrease in HAD5 protein levels and the compelling loss of growth upon HAD5 knockdown, GPI anchor synthesis and MSP1 anchoring to merozoite surfaces are only modestly impacted. Residual HAD5 activity may be present after knockdown and sufficient for generating detectable GPI precursors and subsequent tethering of some GPI-APs, including MSP1. Alternatively, *P. falciparum* expresses two annotated phosphoglucomutases (PGMs) (PF3D7_1012500 and PF3D7_0413500)[75] that could potentially

catalyze phosphomannomutase activity as well, indicating some functional redundancy to HAD5. Residual HAD5 or functional redundancy by PGMs could also explain our observation that D-mannose rescues HAD5 knockdown. Intracellular mannose is likely converted to M6P by hexokinase, and thus some phosphomannomutase activity would still be required for successful rescue of the GPI-anchor pathway. In either case, while residual phosphomannomutase activity present in HAD5^{KD} may not eliminate GPI synthesis entirely, it is insufficient to sustain parasite growth. This speaks to an exquisite sensitivity of malaria parasite cells to disruption in this pathway, whereby minor perturbations in GPI anchor synthesis nonetheless completely interrupt parasite growth, highlighting the promise of this pathway as a therapeutic target.

In addition, we find that HAD5^{KD} parasites have a complete cell cycle arrest, although others have found that untethering MSP1 from the membrane by removing its GPI-anchoring C-terminus still allows for minimal parasite growth[19]. We expect that this discrepancy is due to the role of HAD5 in function of all GPI-anchored parasite proteins, not solely MSP1. These other GPI-APs include related MSPs, RAMA, and 6-cysteine proteins, many of which are refractory to deletion and likely essential[16,21,35,76]. With the incomplete loss of GPI synthesis, it may be that each one of these GPI-APs, including MSP1, are only relatively de-anchored, but the modest reduction in this post-translational modification across multiple cellular proteins works in concert to cause parasite growth arrest.

We therefore propose that HAD5, as an upstream member of the GPI biosynthesis, has great potential as an antimalarial target. We expect that HAD5 inhibition will have broad downstream effects on parasite biology across several life cycle stages. That compound D9 has markedly improved potency against malaria HAD5 compared to orthologous human enzymes provides key proof-of-concept for ongoing development of specific HAD5-directed antimalarial

therapeutics. While compound D9 has limited antimalarial efficacy, this is likely due to its charged phosphonate, expected to have poor cellular penetration. This liability may be improved through a variety of medicinal chemistry strategies, including the addition of pro-drug moieties to mask this charge, a strategy that has been highly effective for other phosphonate antimalarials in development [77–79]. Finally, the crystal structure of HAD5 is likely to be valuable to ongoing efforts to develop more potent and specific HAD5 inhibitors as antimalarials.

This study also adds to the growing literature on HAD-like proteins in *P. falciparum*. Three related HAD proteins each independently modulate parasite sensitivity to the isoprenoid biosynthesis inhibitor FSM, which prompts the question of whether this effect would be similarly seen with other HAD proteins[67–69]. Alternatively, the non-mevalonate isoprenoid biosynthesis pathway may be particularly sensitive to cellular metabolic perturbations. We find that HAD5 serves as an interesting counterexample. HAD5 knockdown yields no changes to FSM sensitivity, providing evidence of a HAD protein and a metabolic perturbation that does not impact the sensitivity of parasites to inhibition of isoprenoid metabolism.

Finally, we note that HAD5 and the GPI anchor biosynthesis pathway are expressed throughout the parasite's life cycle. Several GPI-anchored proteins are expressed in gamete- and oocyst-stage parasites[80,81], and CSP, the target of the RTS,S malaria vaccine and the more recent R21 vaccine candidate[26–29], is a GPI-anchored protein expressed on the surface of sporozoites that facilitates sporozoite development[30] and targets sporozoites to the liver[82]. Furthermore, sexual-stage parasites harbor additional fates of mannose 1-phosphate, including *C*-mannosylation[83,84] and *O*-fucosylation[85,86], suggesting that HAD5 will play a critical role in parasite biology across several life stages. Hence, HAD5 serves an essential role in parasite metabolism and shows promise as a specific therapeutic target. The inhibition of HAD5 not only

has potential to treat intraerythrocytic *P. falciparum* infection, but may also serve to block transmission.

4.5 Methods

4.5.1 Parasite Strains and Culturing

Unless otherwise indicated, parasites were maintained at 37°C in 5% O_2, 5% CO_2, 90% N_2 in a 2% suspension of human red blood cells in RPMI medium (Gibco) modified with the addition of 30 mM $NaHCO_3$, 11 mM glucose, 5 mM HEPES, 1 mM sodium pyruvate, 110 µM hypoxanthine, 10 µg/mL gentamicin, (Sigma) and 2.5 g/L AlbuMAX I (Gibco). 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 American Red Cross Blood Services (St. Louis, MO).

The HAD5 conditional knockdown strain, "HAD5^{KD}", was generated by transfecting NF54^{attB} parasites[87] using previously described methods[45,46]. Resultant parasites were maintained in the presence of anhydrotetracycline (aTc; Cayman Chemicals) in DMSO at 500 nM unless otherwise specified. Parasites were synchronized with a combination of 5% sorbitol (Fisher Bioreagents) and 1.5 µM Compound 1 (MedChemExpress)[88,89]

4.5.2 Plasmodium growth measurement

Parasitemia in daily growth assays was measured via flow cytometry by incubating 10 µL of parasite culture with 190 µL of 0.4 µg/mL acridine orange (Invitrogen) in PBS for 1 minute. Stained parasites were analyzed on a BD FACS Canto flow cytometer gating on DNA and RNA-

bound dye signal using FITC and PerCP-Cy5.5 filters, respectively. 50,000 events were recorded for each sample.

4.5.3 Merozoite reinvasion

Merozoite reinvasion was assessed as previously described[90]. Parasites were synchronized, grown in ±aTc, and treated with 10 µM epoxysuccinyl-L-leucylamido(4-guanidino)butane (E64) (Sigma) for 8 hours to stall them in late schizont stages. E64 was washed out, and cultures were passed through a 1.2 µm filter to lyse the schizonts. Lysates were incubated with fresh red blood cells for 1 hour to allow reinvasion before cultures were washed again to remove debris. Parasitemia was assessed 24 hours post-reinvasion by acridine orange staining and flow cytometry. Reinvasion efficiency was assessed by normalizing resultant parasitemia to the measured parasitemia prior to lysis.

4.5.4 Light and Fluorescent Microscopy

Parasite development was monitored by thin smear of synchronized parasites that were dyed with modified Giemsa Stain (Sigma). For fluorescent microscopy, E64-treated schizont stage parasites were mechanically lysed by passing them through a 1.2 µm filter[90]. Lysates were added to poly-lysine-coated glass slides, fixed in 4% paraformaldehyde, 0.0075%glutaraldehyde in PBS, and permeabilized in 0.1% TritonX-100. Immunofluorescence was performed using mouse anti-MSP1 monoclonal antibody (Novus Biologicals) and goat-anti-mouse Alexafluor 488 secondary antibody (Life Technologies). Samples were then preserved with ProLong Gold antifade reagent with DAPI (Life Technologies).

All immunofluorescence and bright field images were taken using a Zeiss AxioObserver D1 inverted microscope (Carl Zeiss Inc.), equipped with a Axiocam 503 color camera, at the Washington University Molecular Microbiology Imaging facility. Images were acquired with a

Plan-Apochromat 100x (NA 1.4) objective using the ZEN 2.3 pro (blue edition) software. Fluorescent signal was quantified using the ImageStudioLite software from LI-COR.

4.5.5 Transmission Electron Microscopy

For ultrastructural analyses, highly synchronized 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 hr at room temperature. Samples were washed in sodium cacodylate buffer and postfixed in 1% osmium tetroxide (Polysciences Inc.) for 1 hr at room temperature. Samples were then rinsed in dH_20, dehydrated in a graded series of ethanol, and embedded in Eponate 12 resin (Ted Pella Inc., Redding, CA). 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).

4.5.6 Cloning

The coding sequence of HAD5 was cloned from cDNA of 3D7 parasites using primers P1 and P2 (Table S2) and cloned into the BG1861 vector[91], which introduces an N-terminal 6xHis-tag, by ligation independent cloning (LIC). This coding sequence was subsequently cut-and-pasted into a pET28a vector with NdeI and BamHI-HF (NEB), followed by ligation with NEB Quick Ligase using manufacturer's protocols. The coding sequence did not match published reference sequence of PF3D7_1017400, as an adenosine-to-guanosine mutation yielded an Asn-to-Ser substitution at residue 100. To revert this sequence to the reference sequence, primer P3 (Table S2) was used in the QuikChange Multisite-directed mutagenesis kit

(Agilent), and the resulting plasmid was transformed into XL10 Gold ultracompetent *Escherichia coli* cells. The HAD5^{D11A} allele was generated from the WT plasmid, again using the QuikChange multisite-directed mutagenesis kit and primer P4 (Table S2). *Homo sapiens* PMM1 and PMM2 coding sequences were identified from UniProt, codon-optimized for *E. coli*, and synthesized by Integrated DNA Technologies (Table S3). These gene blocks were amplified and extended using primers P5+P6 (PMM1) and P7+P6 (PMM2) and PrimeSTAR GXL DNA polymerase (Takara)(Table S2). These sequences were then cloned into the pET28a plasmid by Gibson assembly at NdeI and BamHI-HF cut sites, and plasmids were transformed into XL10 Gold ultracompetent *E. coli* cells.

E. coli mannose-1-phosphate guanylyl transferase (ManC) was cloned for use in phosphomannomutase assays. The coding sequence of *Ec*ManC was identified from UniProt and a gene block of the sequence was ordered from Integrated DNA Technologies (Table S3). This sequence was cloned by LIC into a BG1861 vector that had been modified with a starting KFS sequence downstream of the 6xHis-tag to enhance protein expression[92], and the resulting plasmid was transformed into Stellar competent cells (Takara).

4.5.7 Recombinant Protein Expression & Purification

To generate recombinant protein for enzyme assays and crystallography, the pET28a plasmids were transformed into BL21 Gold (DE3) competent *E. coli* cells (Agilent). Cells were grown to optical density at 600nm (OD_{600}) of 0.6-0.8 in Terrific Broth medium (24 g/L yeast extract, 20 g/L tryptone, 4 mL/L glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4) shaking at 37˚C, and then induced for approximately 18 hours with 1mM isopropyl-beta-D-thiogalactoside (IPTG) at 16˚C. Cells were collected by centrifugation and resuspended in lysis buffer containing 25 mM Tris HCl (pH 7.5), 250 mM NaCl, 1 mM $MgCl_2$, 10% glycerol, 20 mM imidazole, and

100 µM phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication. Lysate was centrifuged at 18,000xg for 1 hr and the supernatant containing soluble protein was bound to nickel agarose beads (Gold Biotechnology), washed with buffer containing 25mM Tris HCl (pH 7.5), 250 mM NaCl, 1 mM $MgCl_2$, 10% glycerol, 20 mM imidazole, and eluted with 10 mL of buffer containing 25mM Tris HCl (pH 7.5), 250mM NaCl, 1mM $MgCl_2$, 10% glycerol, and 320mM imidazole. Eluate was dialyzed overnight at 4˚C in the presence of 30 units of thrombin protease (Sigma) into buffer containing 25 mM Tris HCl (pH 7.5), 250mM NaCl, 1 mM $MgCl_2$, 10% glycerol. The dialyzed, thrombin-cleaved mixture was then run over nickel agarose beads and benzamidine-sepharose beads (GE Healthsciences) to remove the cleaved His-tag and thrombin. The flow through protein was collected and further purified by size-exclusion chromatography using a HiLoad 16/60 Superdex-200 column (GE Healthsciences), equilibrated in dialysis buffer. Elution fractions containing the protein of interest were identified by UV absorbance and SDS-PAGE, pooled, and concentrated to 6 mg/mL. Protein solutions were flash frozen in liquid nitrogen and stored at -80˚C (Figure 1B, S7A).

*Ec*ManC for enzyme activity assays was expressed by cloning the BG1861:ManC vector into BL21(DE3) pLysS *E. coli* cells (Life Technologies). Cells were grown in LB broth to an OD_{600} of 0.7-0.8 and induced with 1mM IPTG for 2hr at 37˚C. Cells were harvested and protein was purified as described for other proteins above. However, as this construct lacked the thrombin-cleavable site from the pET28a vector, the elution from nickel beads was directly run over the size exclusion column, then pooled and concentrated (Figure S7B).

4.5.8 HAD5 activity assays

Phosphomannomutase activity of HAD5 was measured using a linked enzyme scheme (Fig. S1A), modified from the EnzChek phosphate release kit (ThermoFisher). 200 µM of 2-

amino-6-mercapto-7-methyl-purine riboside (MESG) and 1x EnzChek reaction buffer (50 mM Tris HCl, 1 mM $MgCl_2$, pH 7.5, containing 100 µM sodium azide) were incubated with 1 U/mL Purine Nucleoside Phosphorylase (PNP), 1 U/mL pyrophosphatase, 125 µM GTP, 52 µg/mL recombinant purified *E. coli* ManC, and 1mM mannose-6-phosphate. 12.5 µM glucose-1,6-bisphosphate was also added to demonstrate its activating properties on HAD5. All reagents were obtained from Sigma-Aldrich. The reaction was started by adding 50 ng recombinant HAD5 (WT or D11A), which is within the linear range of the assay (Figure S1B). Reactions took place in 40 µL, and the production of nucleotide by PNP was measured at 360nM.

Phosphoglucomutase assays (Fig. S1C) were developed by incubating 1mM glucose-1-phosphate, 0.75 mM $NADP^+$, 2.5 U/mL glucose-6-phosphate dehydrogenase, and 10 µM glucose-1,6-bisphosphate in 50 µL reactions containing 25 mM NaCl, 25 mM Tris HCl pH 7.5, and 1 mM $MgCl_2$. All reagents were obtained from Sigma-Aldrich. Again, 50 ng of HAD5 was added to start the reaction (Fig. S1D), and the production of reduced NADPH was measured by absorbance at 340 nm.

All reactions took place at room temperature in clear CoStar 96-well half-area plates and absorbances were measured by a Perkin Elmer multimode Envision plate reader. For Michaelis-Menten kinetics of each assay, two-fold serial dilutions of substrate concentrations (either M6P or G1P, respectively) were made and all other components of the assays were kept constant.

4.5.9 Inhibition assays of recombinant HAD5

Synthesis of compounds D1-D11 (Fig. S8A) have been described previously[72,73], with the exception of compound D9, whose synthesis is described, and characterization data included, in the attached supporting information. To assess inhibition of HAD5, compounds D1-D11 were suspended in water and added to the phosphoglucomutase assay of HAD5 activity (the

phosphomannomutase assay was not used, as cross inhibition was seen with downstream components of that assay, but not with the phosphoglucomutase assay; Fig S8D). 5 µL of water volume in the assay was replaced with 5 µL serial dilutions of analogs, with final concentrations ranging from 0 µM – 1 mM. The rate of product formation in each condition was used to determine the half maximal inhibitory concentration (IC_{50}) for each inhibitor. In addition to HAD5, these assays were performed with recombinantly purified human PMM1 and PMM2. For these inhibition assays, 200 ng of HAD5 (139 nM), PMM1 (133 nM), or PMM2 (65 nM) was used to start the reaction.

Time-dependent inhibition was assessed by pre-incubating HAD5 or PMM1 in 416 µM D9 for time points up to 60 minutes, before adding the enzyme+inhibitor mixture to the full reaction mixture, achieving a final [D9] of 50 µM.

4.5.10 Parasite growth inhibition assays

Dose-response inhibition experiments were performed on asynchronous parasites. Three experimental replicates were performed, with technical duplicates, for each experiment. GlcN and ManN inhibition experiments were performed by adding 2-fold dilutions of each compound in water, from 0 mM to 2 mM, to a 100 µL culture of parasites. 1 µM of chloroquine was used as a positive control. Inhibition of cultured parasites by compound D9 was assessed using 100 µM D9 or equivalent volume of water. Parasite growth in these experiments was assessed by acridine-orange staining and flow cytometry measurement of parasitemia 72 hours after treatment.

Dose-response of fosmidomycin (FSM) was assessed using concentrations up to 50 µM FSM in water, and growth was measured after 72 hour treatment on a PerkinElmer multimode Envision plate reader by Quant-iT PicoGreen dsDNA reagent (Thermofisher) staining.

4.5.11 Western blotting

To verify tagging and knockdown of protein, cultures of HAD5^{KD} were grown for 24 hours in ±aTc conditions, then RBCs were lysed with cold PBS + 0.1% saponin. Samples were centrifuged to pellet parasites and remove excess hemoglobin, then parasites were lysed in 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, 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, 0.5 M DTT, 2.5 mg/mL bromophenol blue, 30% 1 M Tris pH 6.8, 50% glycerol) and boiled for 5 min. Lysates were separated by SDS-PAGE, then transferred to 0.45 µm nitrocellulose membrane (Thermo Scientific). Membranes were blocked in PBS + 3% bovine serum albumin, then probed with primary antibodies: mouse anti-FLAG (M2, Sigma), rabbit anti-HSP70 (Agrisera). Membranes were washed in PBS + 0.5% Tween 20, then probed with secondary antibodies goat anti rabbit IRDye 800CW 1:20,000 (LI-COR) and donkey anti-mouse IRDye 680LT 1:20,000 (LI-COR). Membranes were then washed in PBS + 0.1% Tween 20 and imaged on a LI-COR Odyssey platform.

4.5.12 Autoradiography and GPI anchor quantification

GPI anchor autoradiography was performed as previously described[18,56,58]. Briefly, 33-39 hr old parasites that had been grown in ±aTc conditions were washed and resuspended in glucose-free media (RPMI R1383 + 20 mM D-fructose, 25 mM HEPES, 21 mM $NaHCO_3$, 0.37 mM hypoxanthine, 11 mM glutathione, 5 g/L Albumax I, and 10 µg/mL gentamicin). 150 µCi of 40 Ci/mmol [^{3}H]GlcN (American Radiochemicals) was added to each 10 mL culture and incubated for 3 hours. Parasites were saponin lysed and glycolipids were extracted with chloroform:methanol:water (10:10:3), nitrogen evaporation, and n-butanol partitioning[59].

Resultant glycolipids were run on TLC Silica gel 60 F_{254} plates (Millipore Sigma) using chloroform:methanol:water (10:10:3) as a solvent. TLC plates were exposed to autoradiography films (MidSci), which were developed after 1 week of exposure. Films were imaged on a BIO-RAD ChemiDoc MP imaging system, and signal was quantified by the ImageLab software from Bio-Rad.

4.5.13 TritonX-114 Membrane Partitioning

Membrane partitioning was adapted from Doering et al.[63]. Briefly, at approximately 40-44 hours post invasion, synchronized parasite cultures were magnetically purified (Miltenyi Biotec), and elution was centrifuged and resuspended in 1 mL ice-cold Tris-buffered saline (TBS) + 1x HALT protease inhibitor (PI) cocktail, EDTA-free (Thermo Fisher). 200 µL precondensed Triton X-114 was added to lyse parasites, and lysates were incubated on ice for 15 min. Lysates were centrifuged to remove debris. Then followed a series of 5 extractions with cold TBS+PI and precondensed TritonX-114, followed by warming to 37˚C and centrifugation to separate phases. The lysate, detergent-enriched phase, and the soluble phase were analyzed by SDS-PAGE and Western blotted as described above using mouse anti-plasmepsin V 1:20[93] and rabbit anti-HAD1 1:10[67], or rabbit anti-MSP1 1:1000 MSP1[94]. Detergent samples were diluted 5-fold in TBS prior to SDS-PAGE to avoid TritonX-114 interference with the gel.

4.5.14 Protein Crystallography and Ligand Docking

Crystals of *P. falciparum* HAD5 were grown at 4˚C using vapor diffusion in hanging drops comprised of 2 µL protein (6 mg/mL) and 2 µL crystallization buffer (0.1M bis -tris-propane pH 6.5, 0.1 M calcium acetate, 16% (w/v) PEG 8000, and 2% (v/v) benzamidine hydrochloride). Prior to data collection, crystals were flash frozen in mother liquor supplemented with 25% glycerol. All diffraction images were collected at 100K at beamline 19-ID of the

Structural Biology Center at Argonne National Laboratory Advanced Photon Source. HKL3000 was used to index, integrate, and scale the data sets[95]. Prior to phasing, a homology model of PfHAD5 was created using SWISS-MODEL[96] based on human alpha phosphomannomutase 1 (PDB: 6CFR[97], sequence identity 51%). Molecular replacement was performed using PHASER[98] and the homology model of PfHAD5 as a search model. COOT and PHENIX were used for iterative rounds of model building and refinement[99,100]. Data collection and refinement statistics are summarized in Table S1. Atomic coordinates and structure factors of PfHAD5 have been deposited in the RCSB Protein Data Bank (PDB: 7MYV).

The PfHAD5 structure was prepared for docking using AutoDock Tools 1.5.7[101]. The three-dimensional structure of compound D9 was prepared using Avogadro[102], and was prepared for docking using AutoDock Tools 1.5.7. Docking was performed with AutoDock Vina using default search settings[103].

4.6 Figures

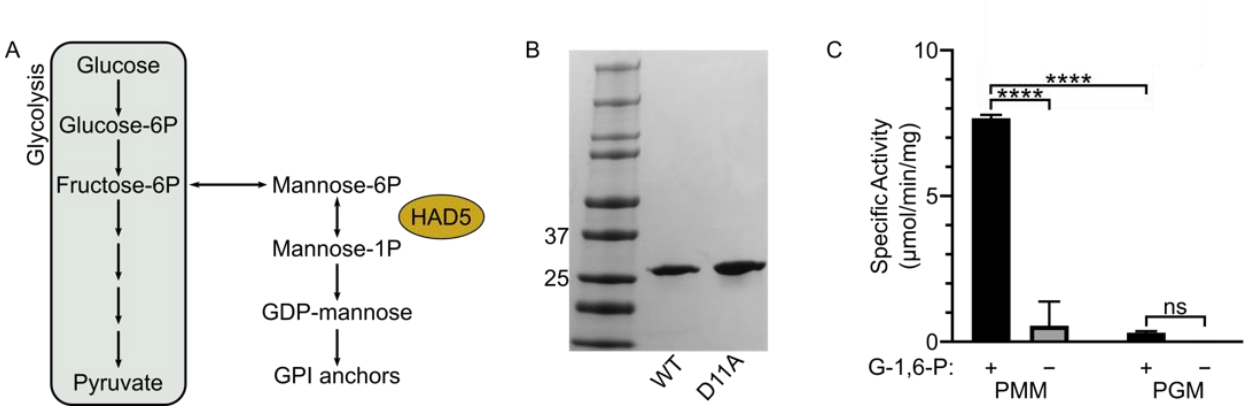


Figure 1. HAD5 is a bifunctional phosphomannomutase/phosphoglucomutase.

(A) Schematic of phosphomannomutase's role in metabolism. Phosphomannomutases like HAD5 interconvert mannose 6-phosphate and mannose 1-phosphate, providing the latter for downstream glycolipid production and synthesis of GPI anchors in *P. falciparum*. **(B)** SDS-PAGE gel of recombinant wild-type (WT) HAD5 and a catalytically inactive mutant (D11A). **(C)** Displayed are the mean ± standard error of the mean (SEM) of HAD5 activity across three independent trials, with D11A activity subtracted as background. Abbreviations: G-1,6-P, glucose-1,6-bisphosphate; PMM, phosphomannomutase assay; PGM, phosphoglucomutase assay. *p*-values were determined using an ordinary two-way ANOVA (Tukey's test for multiple comparisons, $\alpha = 0.05$). ****$p<0.0001$; ns = not significant.

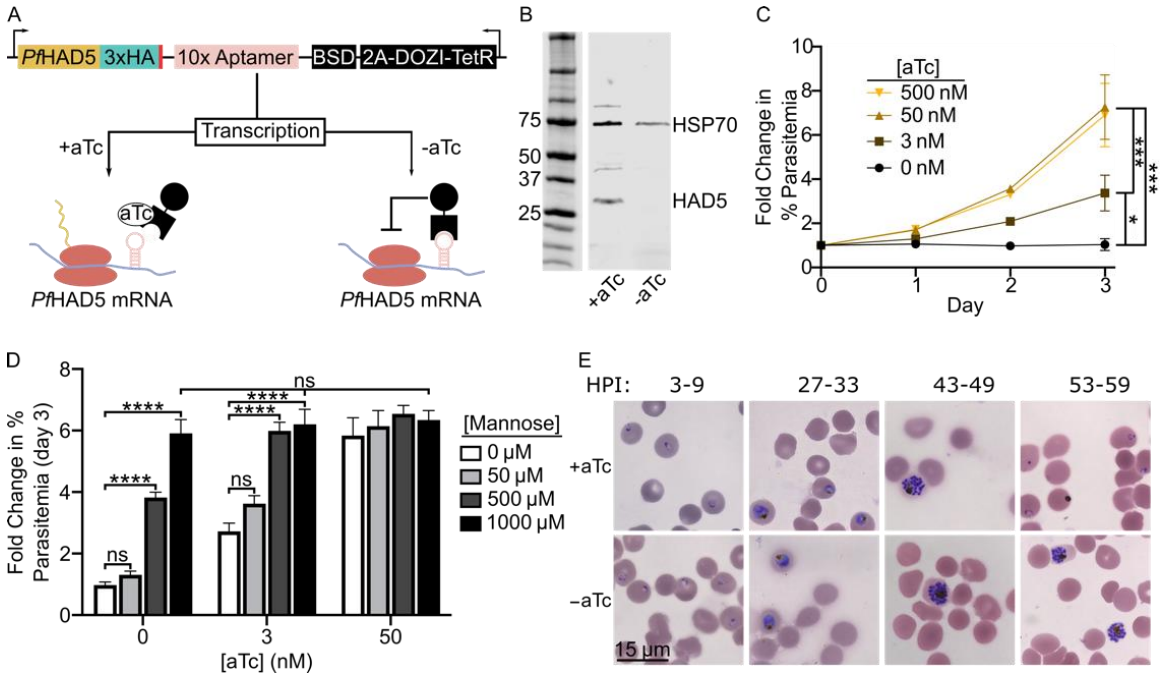


Figure 2. HAD5 is essential for intraerythrocytic parasite growth.

(A) Schematic of the regulatable knockdown system introduced at the native locus of *Pf*HAD5[45,46]. **(B)** Western blot of transgenic HAD5^{KD} parasite lysate ±aTc, using α-HA to detect HAD5 and α-HSP70 as a loading control. Removal of aTc results in successful knockdown of HAD5. Approximate expected protein masses: *Pf*HSP70, 74 kDa; *Pf*HAD5-3xHA, 32 kDa. **(C)** Fold-change in parasitemia over time of HAD5^{KD} parasites grown in varying concentrations of aTc. Data represent mean ± SEM of three independent experiments with technical duplicates. Significance was determined by one-way ANOVA with Fisher's LSD; *p=0.03; ***p<0.001. **(D)** Fold change in parasitemia of HAD5^{KD} parasites grown in varying aTc concentrations with D-mannose rescue. Significance was determined by ordinary two-way ANOVA with Tukey's correction for multiple comparisons. ****p<0.0001, ns = not significant. Data represent mean ± SEM of three independent experiments with technical duplicates. **(E)** Bright-field images of Giemsa-stained thin-smear synchronized parasites over time. HPI = hours post-invasion.

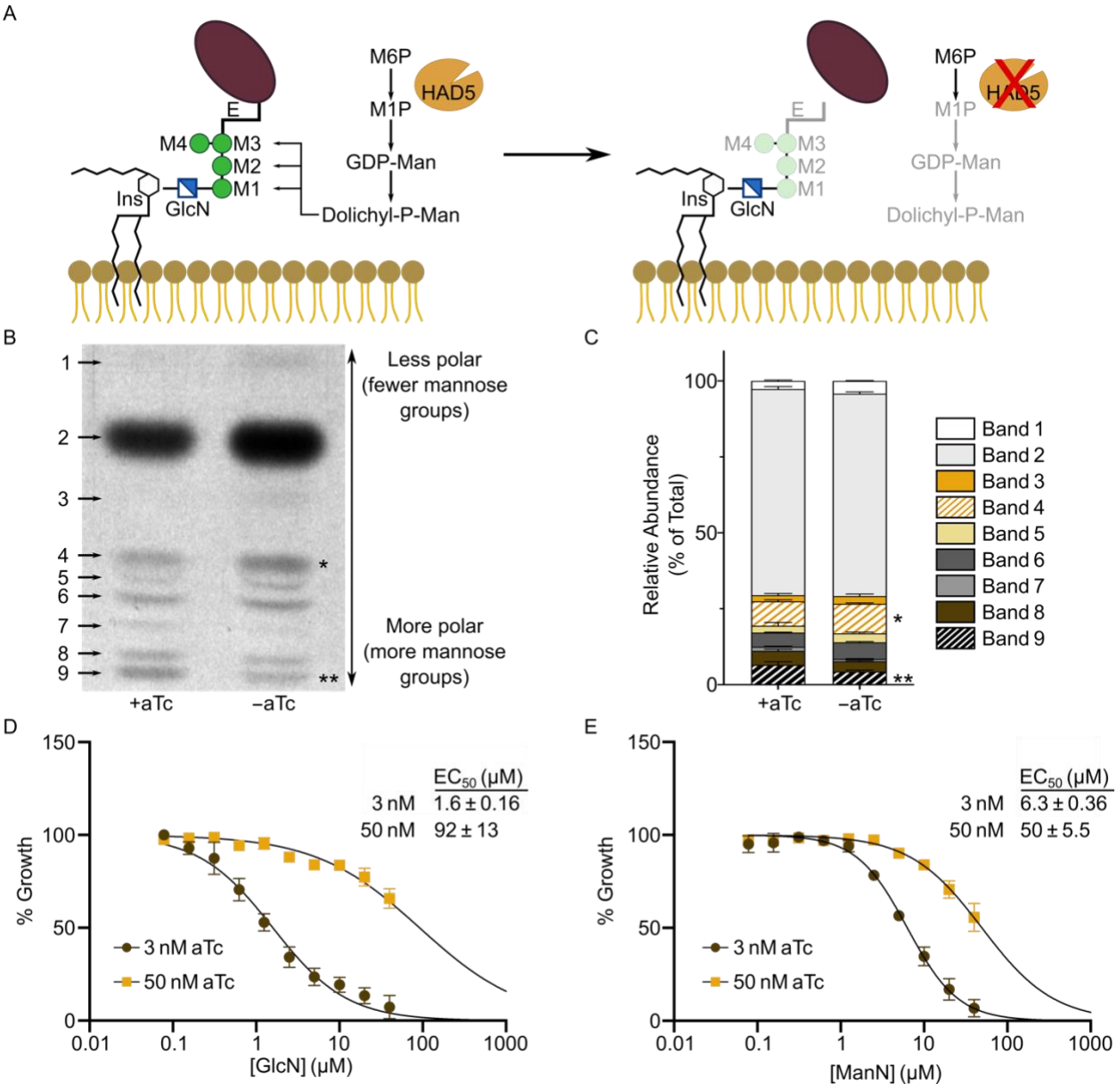


Figure 3. Knockdown of HAD5 disrupts GPI anchor biosynthesis.

(A) Model of the predicted effect that knocking down HAD5 will have on GPI anchor precursor synthesis and subsequent anchoring of GPI-APs. Abbreviations: Ins, inositol; GlcN, glucosamine; M, mannose; E, ethanolamine; M6P, mannose 6-phosphate; M1P, mannose 1-phosphate. **(B)** Representative autoradiography film of GPI-anchor-precursors from [^{3}H]GlcN-labeled parasites. Bands toward the top migrated farthest on a silica TLC plate, indicating they

are less polar and less mannosylated. **(C)** The radiographic signal was quantified and is represented as proportions of total signal. Band numbers indicate the corresponding band from **A**. Shown are the mean and SEM of three independent experiments, analyzed by ordinary two-way ANOVA with Fisher's LSD test. *p=0.038, **p=0.008. **(D,E)** Dose response curve of parasite growth in the presence of glucosamine (GlcN; **D**) or mannosamine (ManN; **E**). Data represent the means and SEM of three independent experiments with technical replicates.

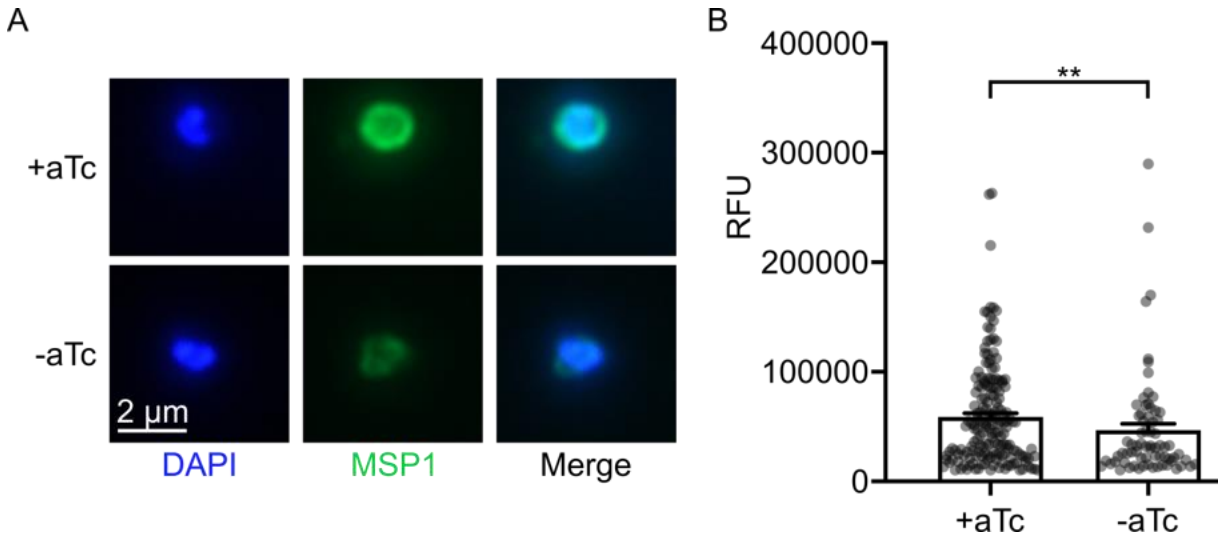


Figure 4. Knockdown of HAD5 diminishes membrane anchoring of the egress and invasion protein MSP1.

(A) Representative immunofluorescent images of mechanically freed merozoites that were grown in ±aTc conditions and schizont enriched by E64 treatment. **(B)** Quantification of MSP1 signal from **A**. Data points represent three independent experiments, each with >25 observed merozoites, removing those under a threshold of 10,000 RFU. Bar graphs represent the mean + SEM of all data points. Statistics were performed by Mann-Whitney test. **p=0.008.

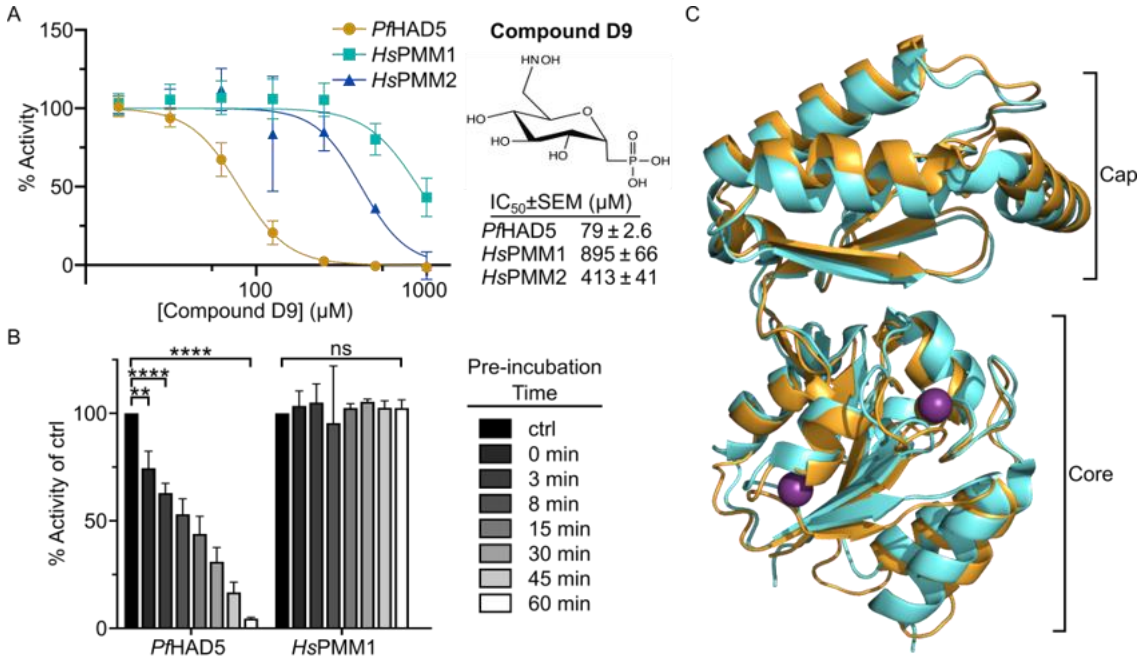


Figure 5. HAD5 is sufficiently distinct from human PMMs to be specifically inhibited.

(A) Dose response curve of compound D9 against recombinant *Pf*HAD5, *Hs*PMM1, and *Hs*PMM2. Data represents the mean ± SEM of three independent experiments, each with technical replicates. **(B)** Activity of *Pf*HAD5 or *Hs*PMM1 was assayed after preincubating enzymes for the given time with 416 μM of D9 prior to adding the preincubation to the reaction mix (final [D9] = 50μM). As a control (ctrl), enzymes were incubated with an equal volume of water for 60 minutes. Statistics were performed with an ordinary Two-way ANOVA, using Dunnett's test for multiple comparisons. ** p=0.0046 **** $p < 0.0001$, ns = not significant. **(C)** 3.5 Å resolution crystal structure of *Pf*HAD5 (orange) aligned to *Hs*PMM1 (cyan; PDB 2FUC) with Mg^{2+} ions (purple). Indicated are the Cap and Core domains typical of HAD enzymes.

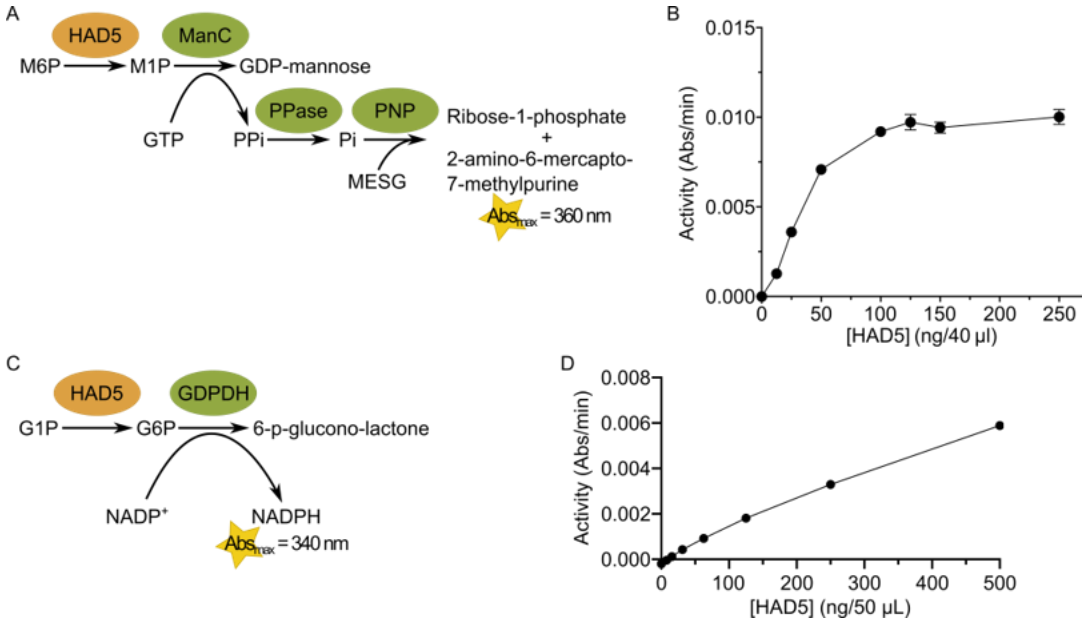


Figure S1. Enzymatic activity assays.

(A) Schematic of the phosphomannomutase (PMM) activity assay, in which HAD5 converts mannose 6-phosphate (M6P) to mannose 1-phosphate (M1P). A series of linked-enzyme steps translates that activity into a spectrophotometric signal at 360 nm. (B) Graph depicting enzymatic activity in the PMM assay with varying HAD5 concentration, demonstrating that our chosen value of 50 ng / 40 µL reaction is within the linear range of the assay with respect to enzyme concentration. (C) Schematic of the phosphoglucomutase (PGM) activity assay, in which HAD5 converts glucose 1-phosphate (G1P) to glucose 6-phosphate (G6P), which is then used by glucose 6-phosphate dehydrogenase (G6PDH) to generate NADPH, which can be measured at 340 nm. (D) Graph depicting enzymatic activity in the PGM assay with varying HAD5 concentration, demonstrating that our chosen value of 50 ng / 50 µL reaction is within the linear range of the assay with respect to enzyme concentration.

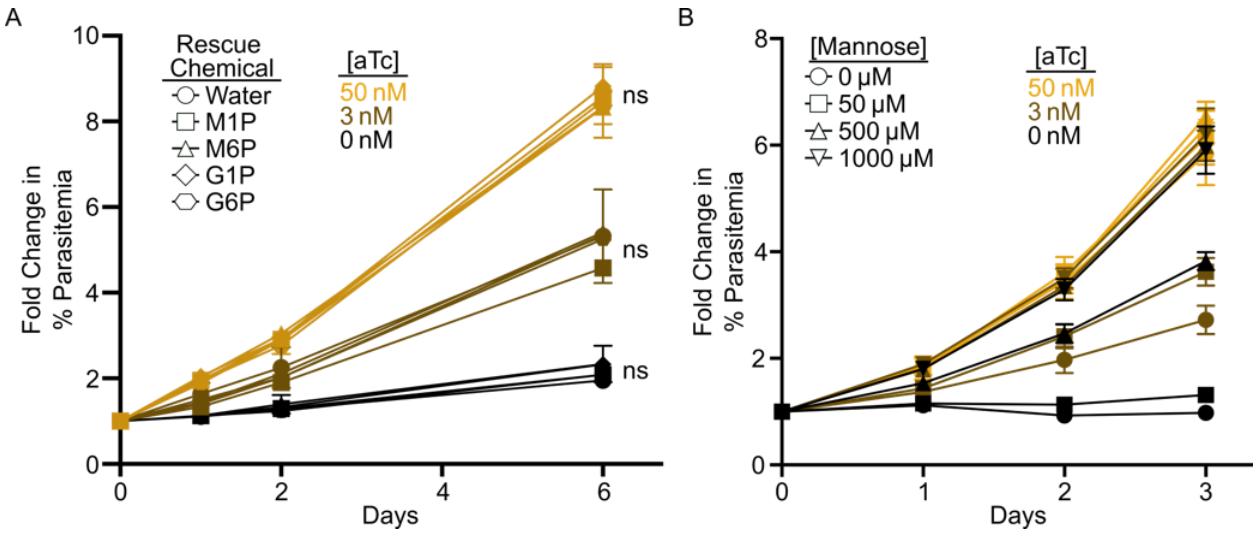


Figure S2. Chemical rescue of parasite growth.

(A) HAD5^{KD} parasite growth was measured in the presence of varying anhydrotetracycline (aTc) concentrations with chemical rescue by four different chemicals: mannose 1-phosphate (M1P), mannose 6-phosphate (M6P), glucose 1-phosphate (G1P), and glucose 6-phosphate (G6P) at 20 µM concentration. Data depicts the mean ± standard error of the mean (SEM) of duplicate experiments and represents fold change in parasitemia over time. Statistics were performed on the Day 6 parasitemia, using a two-way ANOVA with Dunnett's multiple comparisons test, with individual variances computed for each comparison. In all cases, chemical rescues were compared to the vehicle control within a given aTc concentration, with no significant rescue of growth observed (ns = not significant). (B) Shown are the complete data of the fold change in parasitemia over time when HAD5^{KD} parasites are grown in varying aTc and D-mannose concentrations. Data represent mean ± SEM of three independent experiments with technical duplicates. The data from the Day 3 time point was used to generate Figure 2B.

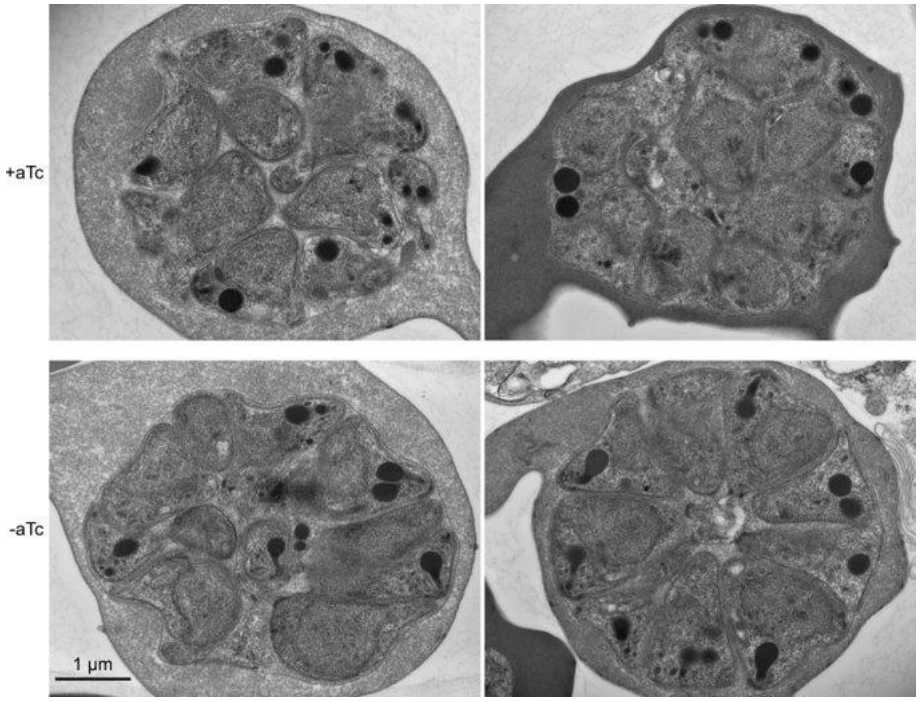


Figure S3. Ultrastructures of HAD5^{KD} schizonts by transmission electron micrograph.

Transmission electron microscopy of highly synchronized schizont parasites showing successful schizogony in parasites grown under -aTc conditions compared to +aTc.

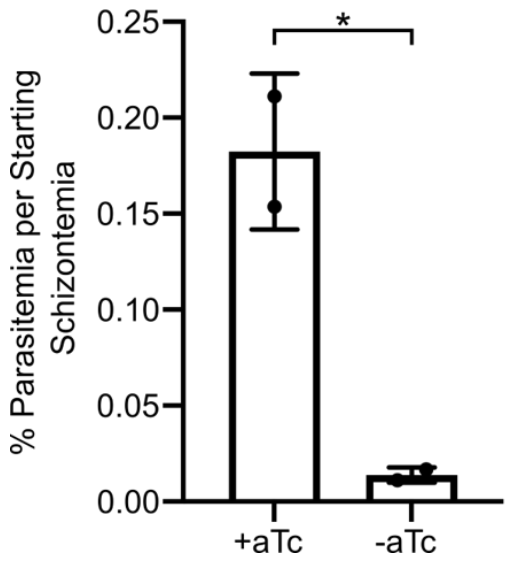


Figure S4. HAD5 KD Parasites are deficient in reinvasion.

Schizont-stage parasites 40-44 hours post invasion, grown in ±aTc conditions, were E64-arrested as late schizonts for 8 hours, mechanically lysed, and allowed to reinvade over fresh red blood cells. 24 hours later, parasitemia was assessed by flow cytometry and normalized to pre-lysis schizontemia, demonstrating a deficiency in reinvasion by parasites grown in –aTc conditions. Statistics were performed using an unpaired two-tailed t-test. *p=0.028

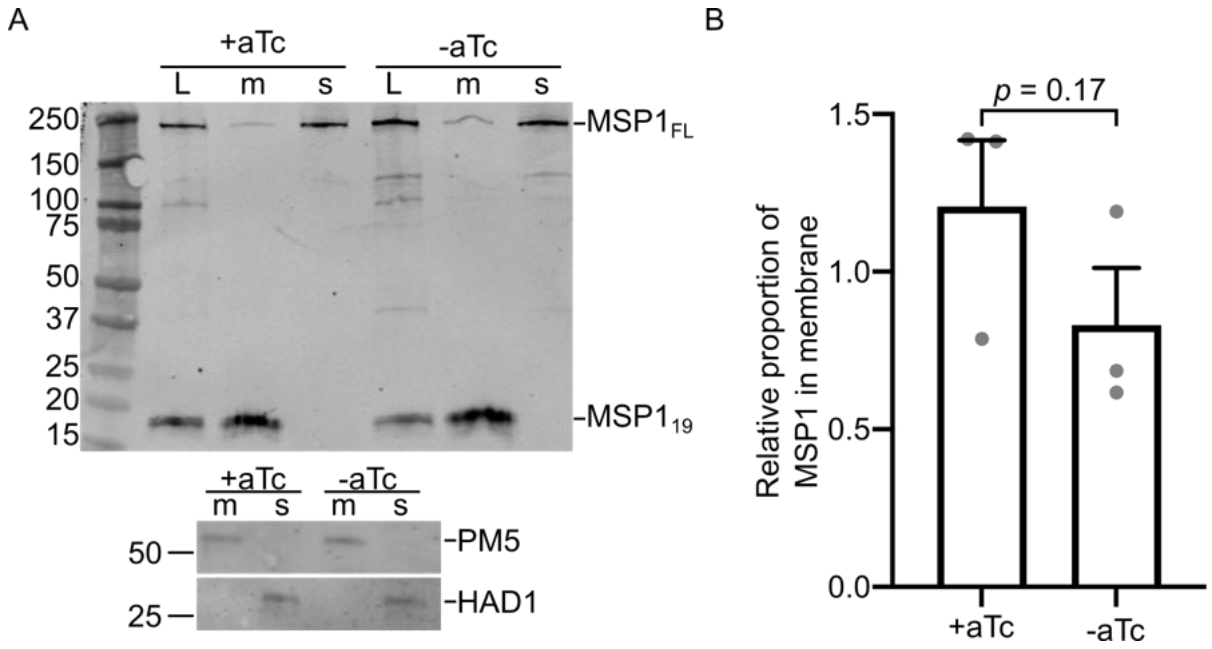


Figure S5. TritonX-114 partition of MSP1.

(A) Representative Western blot of whole lysate (L), membranous fraction (m), and soluble fraction (s) of 40-44hr parasites grown in ±aTc, and blotted for MSP1. Full Length (FL) and 19 kDa fragments of MSP1 are indicated. Plasmepsin 5 (PM5) and HAD1 were used as membrane and soluble controls, respectively. Molecular Weight ladder sizes are indicated in kDa on the left. (B) Quantification of A. Full Length and $MSP1_{19}$ were summed in each lane, and the relative signal in membranous lanes were compared to that of whole lysate. Data represent mean + SEM of three independent experiments. Statistics were performed using a paired, two-tailed t-test.

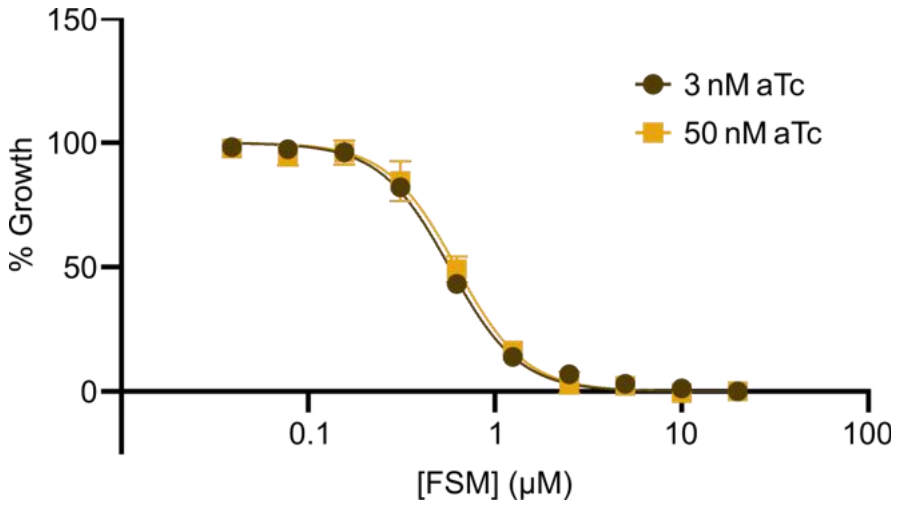


Figure S6. Knockdown of HAD5 has no effect on FSM sensitivity.

72-hour dose-response curves of asynchronous parasites treated with fosmidomycin in either intermediate or saturating aTc concentrations. % Growth was relative to a vehicle control. The EC_{50} ± SEM (in μM) for each condition were: 3 nM, 0.57 ± 0.02; 50 nM, 0.62 ± 0.04. Not significant (p = 0.39) by unpaired t-test.

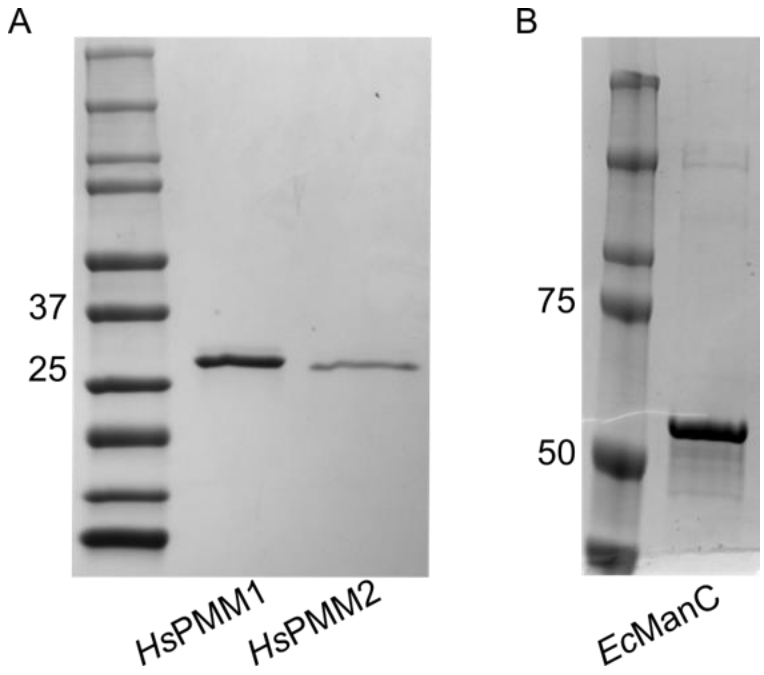
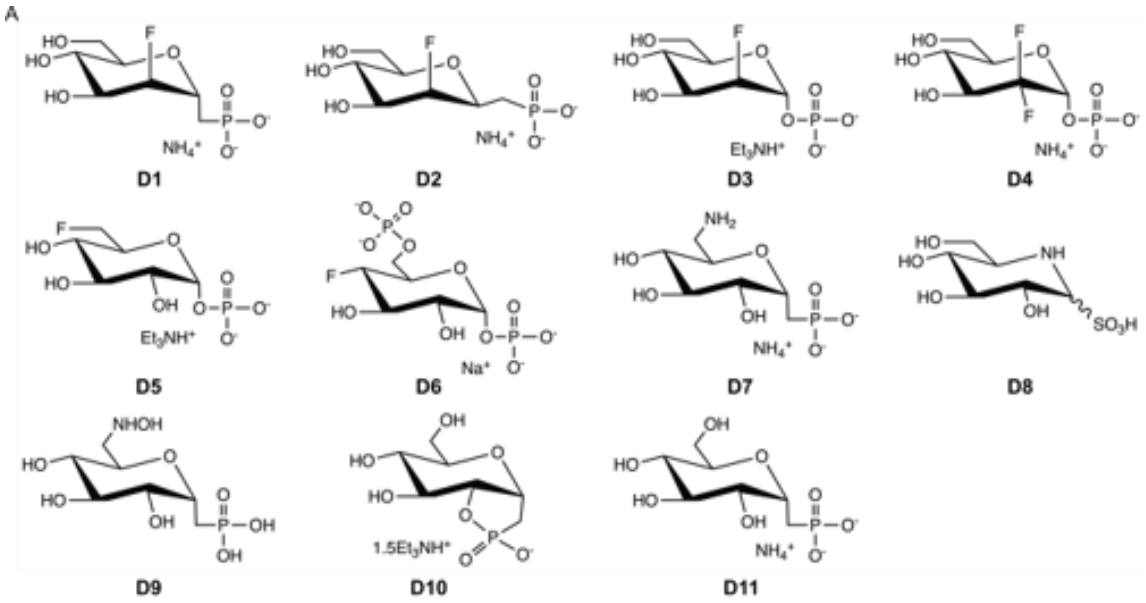
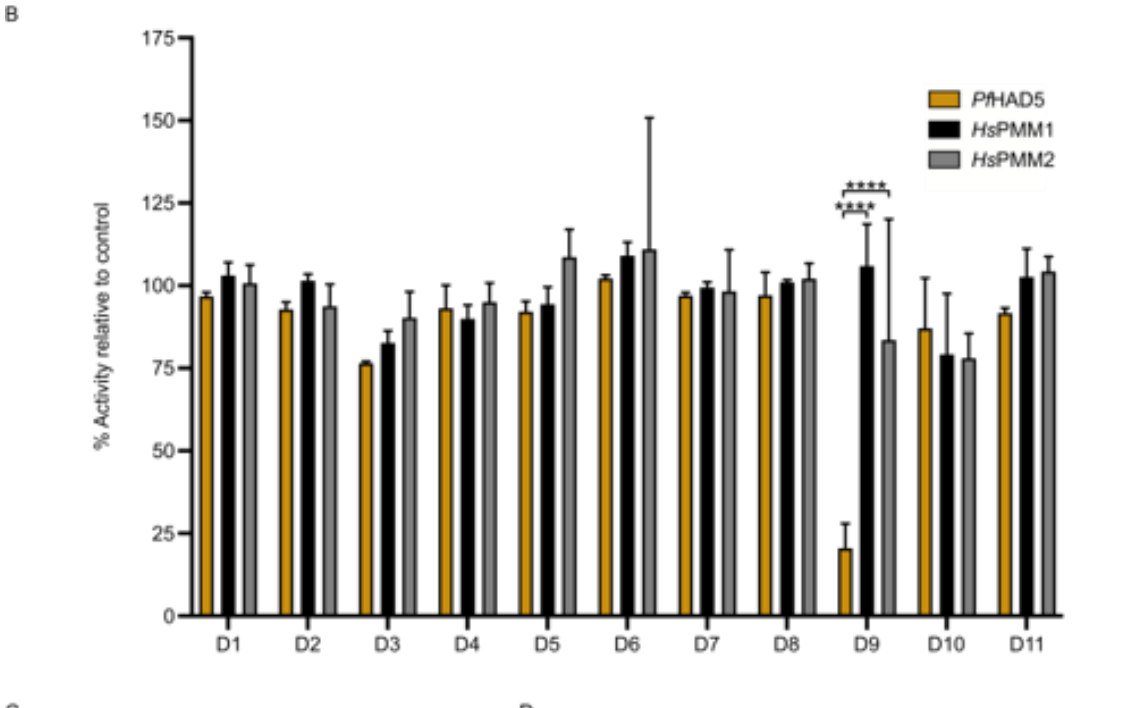
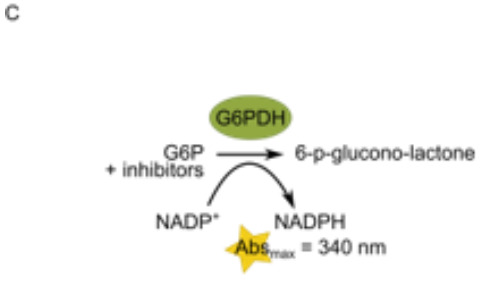


Figure S7. Purified recombinant proteins.

Coomassie-stained SDS-PAGE gels of final purified forms of recombinant *Hs*PMM1, *Hs*PMM2 (A) and *Ec*ManC (B).


A
D1
D2
D3
D4
D5
D6
D7
D8
D9
D10
D11


B
% Activity relative to control
PfHAD5
HsPMM1
HsPMM2
****
****
D1
D2
D3
D4
D5
D6
D7
D8
D9
D10
D11


C
G6PDH
G6P
+ inhibitors
6-p-glucono-lactone
NADP+
NADPH
Abs max = 340 nm

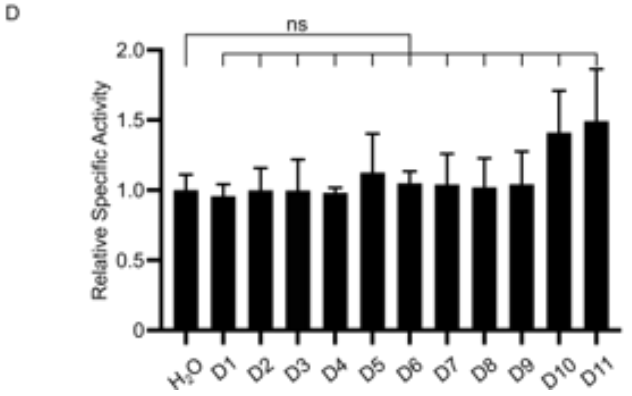

D
ns
Relative Specific Activity
H2O
D1
D2
D3
D4
D5
D6
D7
D8
D9
D10
D11

Figure S8. Structures and evaluation of compounds D1-D11.

(A) Structures of the 11 compounds tested for inhibition against *Pf*HAD5, *Hs*PMM1, and *Hs*PMM2. (B) Activity of recombinant *Pf*HAD5, *Hs*PMM1, and *Hs*PMM2 when treated with 125 µM of the indicated compounds as a percentage of treatment with a vehicle control. Data represent the mean ± SEM of 3 independent experiments with technical replicates. Statistics were performed with a two-way ANOVA using Tukey's test for multiple comparisons. ****$p<0.0001$. All other comparisons between enzymes for a given compound were not significant. (C) Diagram of the amended assay to assess compound inhibition of downstream components of the assay. Assays were commenced with the addition of substrate (G6P). (D) Quantification of assay shown in C. Assay activity is depicted relative to treatment with a vehicle control. Data represent the mean ± SEM of 3 independent experiments with technical replicates. Statistics were performed with an ordinary one-way ANOVA using Dunnett's test for multiple comparisons with single pooled variance. ns = not significant.

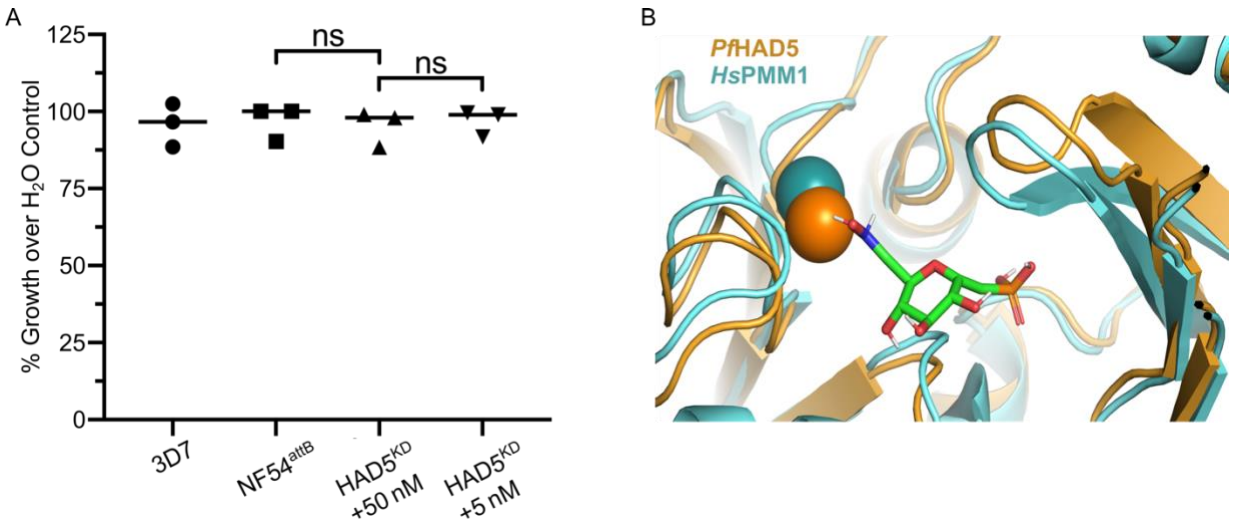
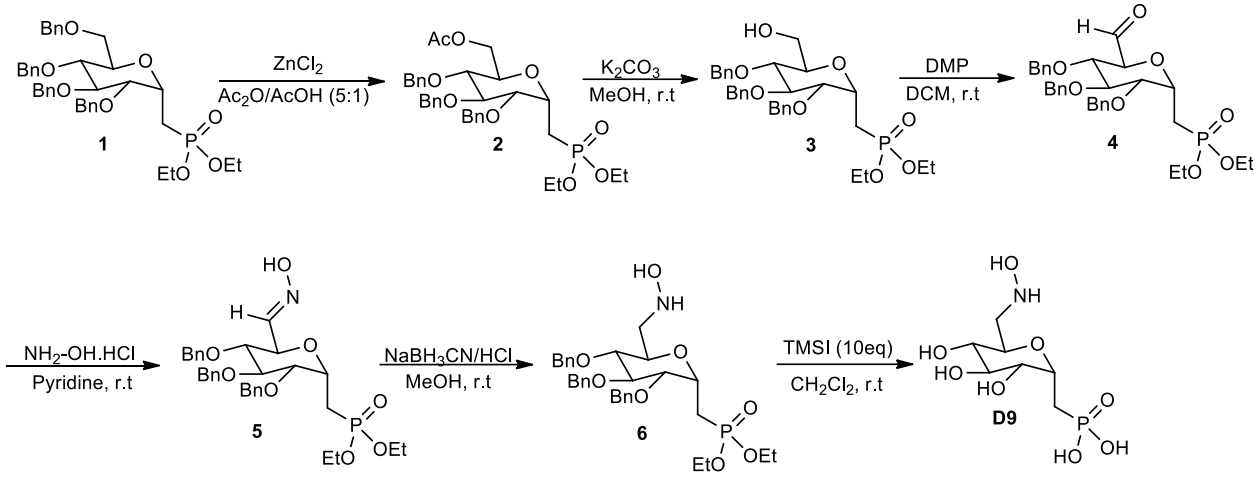


Figure S9. Compound D9 does not inhibit parasites in culture.

(A) Displayed is the percent growth after 72 hours of parasites (starting parasitemia ~1.2%) when treated with 100 µM compound D9, compared to vehicle control. The experiment was performed on wild type parasites of two different strains (3D7 and NF54), as well as HAD5^{KD} parasites ±aTc. Three independent experiments, each with technical replicates, were performed. Statistics were performed with an ordinary one-way ANOVA using Tukey's test for multiple comparisons, ns = not significant. (B) Shown is the D9 compound (green), computationally docked to *Pf*HAD5 crystal structure (orange), highlighting its location in the binding pocket, compared to subtle differences in the binding pocket of *Hs*PMM1 (cyan). Mg2+ ions are depicted from *Pf*HAD5 (dark orange) and *Hs*PMM1 (dark cyan) as spheres.

Synthesis of phosphonate analogue of 6NHOH-G1CP (D9)

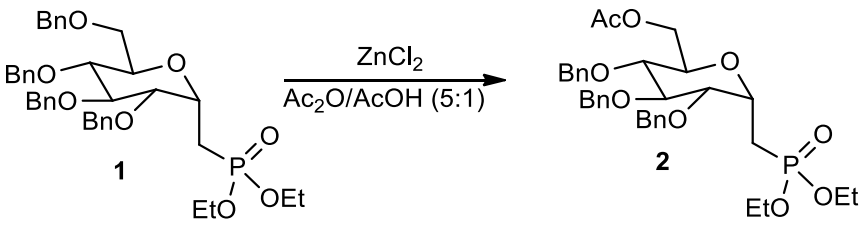
Scheme 1. Synthesis of phosphonate analogue of 6NHOH-G1CP (D9)



The synthesis of α-D-glucose-1-phosphonate **D9** analogue commenced with the selective debenzylation-acetolysis (104) of the benzyl C-6 substituent of diethyl-C-(1-Deoxy 2,3,5,6-tetra-O-benzyl-α-D-gluctopyranosyl) methanephosphonate (**1**) using a solution of freshly fused $ZnCl_2$ (5.2 equiv) in 1:5 HOAc–Ac2O in 98% yield. (Scheme 1). We observed that it was important to keep the temperature at 0 °C during the course of the reaction (30 min). The precursor phosphonate of **1** was synthesized according to our previously reported in three steps starting from 2,3,4,6-tetra-*O*-benzyl-D-glucopyranose [105]. Conventional de-O-acetylation of **2** with potassium carbonate in methanol gave the alcohol **3** in high yield. Although deacetylation of the acetyl protecting group at C-6 in **2** was first attempted using a mixture 1:3:7 TEA-H_2O-MeOH, but only poor yields of alcohol **2** was obtained mainly due to low solubility of **2**. Synthetic alcohol **3** was converted to the corresponding aldehyde **4** by using the Dess-Martin periodinane (DMP) in 98% yield (106). First we examined reaction with using 1 equivalent of DMP, but the yield was low, then we increased amounts of DMP to 1.5 and 2 equivalents. Our optimal result was obtained when we used 1.5 equivalents of DMP. Without further purification, condensation

of aldehyde **4** with hydroxylamine hydrochloride in pyridine afforded the desired oxime **5** in 94% yield [106]. Further reduction of **5** with sodium cyanoborohydride ($NaBH_3CN$) under acidic condition provided the hydroxylamine **6** in 78% yield [106]. Deprotection of both benzyl ether and ethyl ester groups was readily accomplished with an excess of iodotrimethylsilane (TMSI) [105,107]. For preparation of target phosphonate **D9**, different quantities of TMSI were examined. With 10 equivalents of TMSI compound **6,** was converted to the final product **D9** in 95% yield. It should be mentioned, that when more than 10 equivalents of TMSI was used, the reaction led to decomposition.

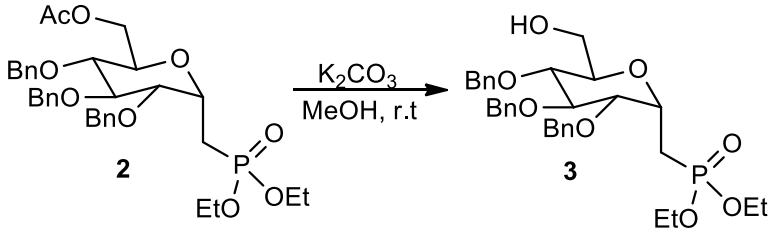
((2R,3R,4S,5R,6S)-3,4,5-tris(benzyloxy)-6-((diethoxyphosphoryl)methyl)tetrahydro-2H-pyran-2-yl)methyl acetate (2).



A solution of $ZnCl_2$ (708 mg, 5.2 eq, 5.2 mmol) in Ac_2O/AcOH (5:1, 12 mL) was cooled to 0 °C then a solution of **1** (674 mg, 1mmol) in Ac_2O/AcOH (5:1, 12 mL) was added dropwise. The solution was stirred for overnight at room temperature under N_2. After evaporation, the residue was suspended in water (20 mL) and extracted with DCM (40 mL x 2). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated. Purification by flash chromatography (Hexanes-EtOAc, 1:1) afforded compound **2** (660 mg, 98%) as a colorless liquid; Rf 0.50 (Hexanes-EtOAc, 1:3). ^{1}H NMR (500 MHz, $CDCl_3$), δ_H (ppm) 1.35 (6H, td, $^3J_{HH}$ = 7.1 Hz, $^4J_{HP}$ = 2.4 Hz, $2OCH_2CH_3$), 2.06 (3H, s, CH_3CO), 2.20-2.26 (2H, m, CH_2-P), 3.57 (1H, t, $^3J_{HH}$ = 9.1

Hz, H-4), 3.72-3.75 (2H, m, H-2, H-3), 3.80 (1H, dt, $^{3}J_{HH}$ = 9.7 Hz, $^{3}J_{HH}$ = 2.8 Hz, H-5), 4.10-4.16 (4H, m, 2OC*H*$_2$CH$_3$), 4.24 (1H, dd, $^{2}J_{HH}$ = 12.0 Hz, $^{3}J_{HH}$ = 2.1 Hz, H-6a or H-6b), 4.45 (1H, dd, $^{2}J_{HH}$ = 12.0 Hz, $^{3}J_{HH}$ = 3.6 Hz, H-6a or H-6b), 4.35-4.56 (1H, m, H-1), 4.58 (1H, d, $^{2}J_{HH}$ = 10.9 Hz, CH$_2$ of Bn), 4.68 (1H, d, $^{2}J_{HH}$ = 11.5 Hz, CH$_2$ of Bn), 4.71 (1H, d, $^{2}J_{HH}$ = 11.5 Hz, CH$_2$ of Bn), 4.81 (1H, d, $^{2}J_{HH}$ = 11.5 Hz, CH$_2$ of Bn), 4.87 (1H, d, $^{2}J_{HH}$ = 10.9 Hz, CH$_2$ of Bn), 4.94 (1H, d, $^{2}J_{HH}$ = 10.9 Hz, CH$_2$ of Bn), 7-28-7.36 (15H, m, 3Ph). ^{13}C NMR (125 MHz, CDCl$_3$), δ_C (ppm) 16.69 (d, $^{3}J_{CP}$ = 3.7 Hz, 2OCH$_2$*C*H$_3$), 16.73 (d, $^{3}J_{CP}$ = 3.7 Hz, 2OCH$_2$*C*H$_3$), 21.08 (*C*H$_3$CO), 22.87 (d, $^{1}J_{CP}$ = 144.0 Hz, CH$_2$-P), 61.81 (d, $^{2}J_{CP}$ = 6.4 Hz, O*C*H$_2$CH$_3$), 62.03 (d, $^{2}J_{CP}$ = 6.1 Hz, O*C*H$_2$CH$_3$), 63.31 (CH$_2$-6), 69.88 (d, $^{2}J_{CP}$ = 5.1 Hz, C-1), 70.49 (C-5), 73.34, 75.14, 75.58 (3CH$_2$ of Bn), 77.40 (C-4), 79.31 (d, $^{3}J_{CP}$ = 12.7 Hz, C-2), 77.40 (C-3), 127.98, 128.14, 128.29, 128.67, 128.71, 137.95, 137.99, 138.53 (3Ph), 170.92 (C=O).^{31}P NMR (121 MHz, CDCl$_3$), δ_P (ppm) 28.88. HRMS (ESI, Positive Mode) calcd for $C_{34}H_{43}NaO_9P$ [M + Na] $^{+}$ 649.2537, found 649.2534.

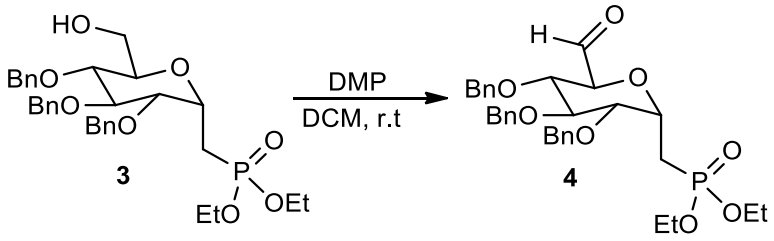
((2R,3R,4S,5R,6S)-3,4,5-tris(benzyloxy)-6-((diethoxyphosphoryl)methyl)tetrahydro-2H-pyran-2-yl)methyl acetate (3).



To a solution of **2** (626 mg, 1 mmol) in MeOH (30 mL), K_2CO_3 (500mg, 3.6 mmol) was added at room temperature. The resulting solution was then stirred overnight. After evaporation, the residue was suspended in water (30 mL) and extracted with DCM (40 mL x 2). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated to provide the crude product (543

mg, 93%) as a colorless liquid which was used in subsequent reactions without further purification (Rf 0.22 (Hexanes-EtOAc, 1:3). ^{1}H NMR (500 MHz, $CDCl_3$), δ_H (ppm) 1.35 (6H, t, $^{3}J_{HH}$ = 7.1 Hz, $2OCH_2CH_3$), 1.78 (1H, brs, OH), 2.04-2.12 (1H, m, CH_2-P), 2.32-2.41 (1H, m, CH_2-P), 3.40 (1H, t, $^{3}J_{HH}$ = 8.1 Hz, H-4), 3.60-3.67 (1H, m, H-2), 3.73-3.81 (3H, m, H-3, H-6a, H-6b), 3.85 (1H, td, $^{3}J_{HH}$ = 8.3 Hz, $^{3}J_{HH}$ = 2.6 Hz, H-5), 4.12-4.17 (4H, m, $2OCH_2CH_3$), 4.43-4.45 (1H, m, H-1), 4.61-4.83 (6H, m, $3CH_2$ of Bn), 7-29-7.40 (15H, m, 3Ph). ^{13}C NMR (125 MHz, $CDCl_3$), δ_C (ppm) 16.62 (d, $^{3}J_{CP}$ = 5.8 Hz, $2OCH_2CH_3$), 24.38 (d, $^{1}J_{CP}$ = 144.1 Hz, CH_2-P), 61.97 (d, $^{2}J_{CP}$ = 5.9 Hz, OCH_2CH_3), 62.00 (CH_2-6), 62.25 (d, $^{2}J_{CP}$ = 6.4 Hz, OCH_2CH_3), 68.66 (d, $^{2}J_{CP}$ = 5.6 Hz, C-1), 73.42 (CH_2 of Bn), 73.89 (C-5), 74.76, 75.14 ($2CH_2$ of Bn), 77.42 (C-4), 78.88 (d, $^{3}J_{CP}$ = 13.7 Hz, C-2), 80.79 (C-3), 127.95, 128.02, 128.09, 128.17, 128.19, 128.64, 128.67, 128.70, 137.98, 138.11, 138.47 (3Ph).^{31}P NMR (121 MHz, $CDCl_3$), δ_P (ppm) 29.79. HRMS (ESI, Positive Mode) calcd for $C_{32}H_{41}NaO_8P$ $[M + Na]^+$ 607.2431, found 607.2420.

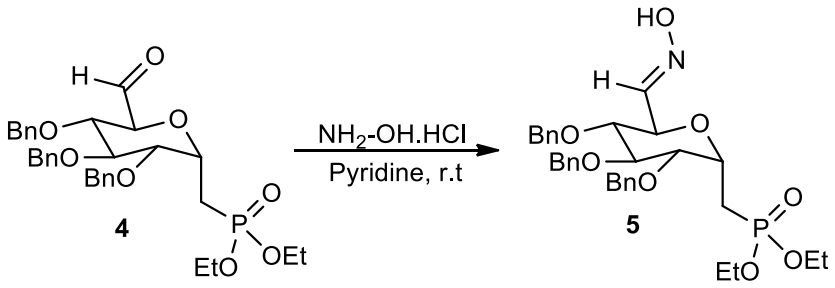
Diethyl(((2S,3R,4S,5S,6S)-3,4,5-tris(benzyloxy)-6-formyltetrahydro-2H-pyran-2-yl) methyl) phosphonate (4).



To a solution of **3** (584 mg, 1 mmol) in anhydrous CH_2Cl_2 (30 mL), Dess-Martin periodinane (636 mg, 1.5 mmol, 1.5 equiv) was added at room temperature. The resulting solution was then stirred at same temperature for 4h under nitrogen. The solution was diluted CH_2Cl_2 (30 mL), and

then saturated $NaHCO_3$(aq) and saturated $Na_2S_2O_3$(aq) were added to the reaction sequentially. The resulting mixture was stirred for another 30 min at room temperature. The organic layer was separated, and the aqueous layer was extracted with EtOAc. The combined organic extracts were dried ($MgSO_4$), filtered and concentrated to provide the crude product (547 mg, 94%) as a colorless liquid which was used in subsequent reactions without further purification (Rf 0.30 (Hexanes-EtOAc, 1:3). ^{1}H NMR (500 MHz, $CDCl_3$), δ_H (ppm) 1.28-1.35 (6H, m, $2OCH_2CH_3$), 2.11-2.19 (1H, m, CH_2-P), 2.31-2.39 (1H, m, CH_2-P), 3.52 (1H, dd, $^3J_{HH}$ = 5.3 Hz, $^4J_{HP}$ = 3.3 Hz, H-2), 3.76 (1H, t, $^3J_{HH}$ = 5.3 Hz, H-3), 3.81 (1H, t, $^3J_{HH}$ = 4.8 Hz, H-4), 4.09-4.18 (4H, m, $2OCH_2CH_3$), 4.28 (1H, d, $^3J_{HH}$ = 4.7 Hz, H-5), 4.46 (1H, d , $^2J_{HH}$ = 11.6 Hz, CH_2 of Bn), 4.53-4.57 (3H, m, H-1 and CH_2 of Bn), 4.62-4.70 (3H, m, CH_2 of Bn), 7-20-7.36 (15H, m, 3Ph), 9.83 (1H, s, CHO). ^{31}P NMR (121 MHz, $CDCl_3$), δ_P (ppm) 28.67.

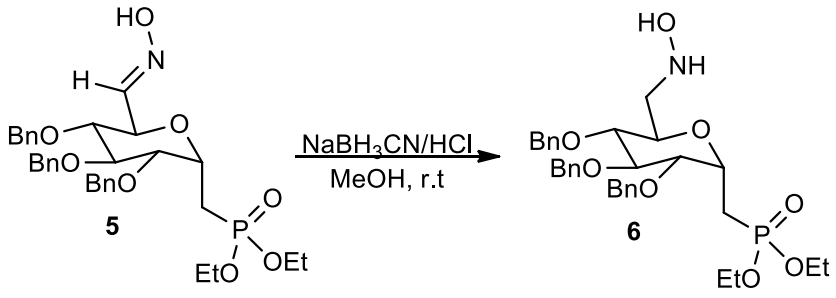
Diethyl (((2S,3R,4S,5R,6R)-3,4,5-tris(benzyloxy)-6-((E)-(hydroxyimino)methyl)tetrahydro-2H-pyran-2-yl) methyl)phosphonate (5).



To a solution of **4** (582 mg, 1 mmol) in anhydrous pyridine (15 mL), hydroxylamine hydrochloride (104 mg, 1.5 mmol, 1.5 equiv) was added at room temperature. The resulting solution was then stirred at same temperature for 5h and then pyridine was removed at 50 °C under high vacuum. Water (30 mL) was added to the residue and the mixture was extracted with

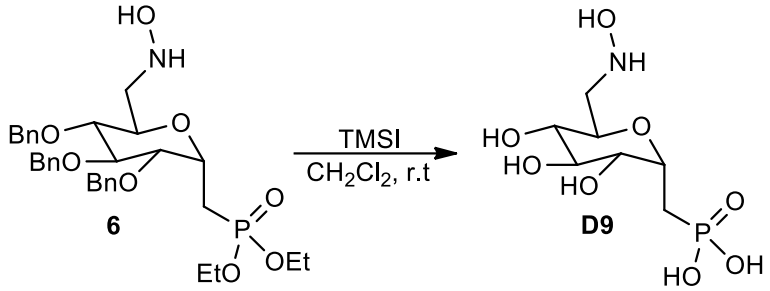
CH_2Cl_2 (30 mL x 2). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated. Purification by flash column chromatography (Hexanes-EtOAc, 1:2) afforded compound **5** as a mixture of E/Z isomers (561 mg, 94%) as a colorless liquid; Rf 0.41 (Hexanes-EtOAc, 1:3). ^{1}H NMR (500 MHz, $CDCl_3$), δ_H (ppm) 1.29-1.37 (6H, m, $2OCH_2CH_3$), 2.07-2.50 (2H, m, CH_2-P), 3.43-3.80 (3H, m, H-2, H-3, H-4), 4.06-4.19 (4H, m, $2OCH_2CH_3$), 4.23-4.36 (1H, m, H-5), 3.48-4.55 (1H, m, H-1), 4.57-4.95 (6H, m, $3CH_2$ of Bn), 7-15-7.37 (16H, m, 3Ph, CH=N), 9.38 (1H, brs, OH). ^{13}C NMR (125 MHz, $CDCl_3$), δ_C (ppm) 16.52 (d, $^3J_{CP}$ = 6.5 Hz, OCH_2CH_3), 16.60 (d, $^3J_{CP}$ = 6.4 Hz, OCH_2CH_3), 22.36 (d, $^1J_{CP}$ = 143.2 Hz, CH_2-P), 62.05 (d, $^2J_{CP}$ = 6.4 Hz, OCH_2CH_3), 62.20 (d, $^2J_{CP}$ = 5.7 Hz, OCH_2CH_3), 70.10 (d, $^2J_{CP}$ = 5.4 Hz, C-1), 71.31 (C-5), 73.39, 75.24, 75.43 ($3CH_2$ of Bn), 78.81 (d, $^3J_{CP}$ = 12.9 Hz, C-2), 80.13, 80.71 (C-3, C-4), 127.74, 127.82, 128.05, 128.12, 128.21, 128.51, 128.63, 128.70, 137.91, 138.01, 138.71 (3Ph), 148.14 (CH=N).^{31}P NMR (121 MHz, $CDCl_3$), δ_P (ppm) 28.96 (minor isomer), 30.46 (major isomer). HRMS (ESI, Positive Mode) calcd for $C_{32}H_{40}NNaO_8P$ $[M + Na]^+$ 620.2384, found 620.2391.

Diethyl (((2S,3R,4S,5R,6R)-3,4,5-tris(benzyloxy)-6-((hydroxyamino)methyl)tetrahydro-2H-pyran-2-yl)methyl)phosphonate (6).



To a solution of **5** (597 mg, 1 mmol) and $NaBH_3CN$ (125 mg, 2 mmol, 2 equiv) in MeOH (30 mL), a solution of HCl (6N in MeOH) was added dropwise at 0 °C until pH 1-3. The resulting solution was then stirred at room temperature for 10h. The reaction was quenched by the addition of saturated $NaHCO_3$ (aq) and the mixture was extracted with EtOAc (40 mL x 2). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated. Purification by flash column chromatography (CH_2Cl_2-MeOH, 40:1) afforded compound **5** (467 mg, 78%) as a colorless liquid; Rf 0.50 (CH_2Cl_2-MeOH, 40:1). ^{1}H NMR (500 MHz, $CDCl_3$), δ_H (ppm) 1.35 (6H, td, $^3J_{HH}$ = 7.1 Hz, $^4J_{HP}$ = 1.7 Hz, 2OCH$_2$C*H*$_3$), 2.13-2.21 (1H, m, CH$_2$-P), 2.33-2.42 (1H, m, CH$_2$-P), 2.82 (1H, dd, $^2J_{HH}$ = 13.6 Hz, $^3J_{HH}$ = 9.5 Hz, H-6a or H-6b), 3.27 (1H, t, $^3J_{HH}$ = 9.3 Hz, H-4), 3.40 (1H, dd, $^2J_{HH}$ = 13.6 Hz, $^3J_{HH}$ = 1.7 Hz, H-6a or H-6b), 3.70-3.79 (2H, m, H-2, H-3), 4.00 (1H, dd, $^3J_{HH}$ = 9.5 Hz, $^3J_{HH}$ = 1.7 Hz, H-5), 4.09-4.22 (4H, m, 2OC*H*$_2$CH$_3$), 4.35-4.41 (1H, m, H-1), 4.62 (1H, d, $^2J_{HH}$ = 11.0 Hz, CH$_2$ of Bn), 4.64 (1H, d, $^2J_{HH}$ = 11.7 Hz, CH$_2$ of Bn), 4.76 (1H, d, $^2J_{HH}$ = 11.7 Hz, CH$_2$ of Bn), 4.83 (1H, d, $^2J_{HH}$ = 10.9 Hz, CH$_2$ of Bn), 4.87 (1H, d, $^2J_{HH}$ = 11.0 Hz, CH$_2$ of Bn), 4.92 (1H, d, $^2J_{HH}$ = 10.9 Hz, CH$_2$ of Bn), 7-29-7.36 (15H, m, 3Ph). ^{13}C NMR (125 MHz, $CDCl_3$), δ_C (ppm) 16.61 (d, $^3J_{CP}$ = 4.8 Hz, 2OCH$_2$*C*H$_3$), 16.64 (d, $^3J_{CP}$ = 4.6 Hz, 2OCH$_2$*C*H$_3$), 24.16 (d, $^1J_{CP}$ = 145.7 Hz, CH$_2$-P), 55.59 (CH$_2$-6), 61.96 (d, $^2J_{CP}$ = 6.3 Hz, O*C*H$_2$CH$_3$), 62.68 (d, $^2J_{CP}$ = 6.5 Hz, O*C*H$_2$CH$_3$),68.97 (C-5), 68.93 (d, $^2J_{CP}$ = 5.6 Hz, C-1), 73.70, 75.20, 75.71 (3CH$_2$ of Bn), 79.63 (d, $^3J_{CP}$ = 13.4 Hz, C-2), 80.80 (C-4), 82.13 (C-3) 172.89, 128.00, 128.04, 128.07, 128.09, 128.17, 128.60, 128.71, 138.00, 138.07, 138.58 (3Ph).^{31}P NMR (121 MHz, $CDCl_3$), δ_P (ppm) 28.83. HRMS (ESI, Positive Mode) calcd for $C_{32}H_{43}NO_8P$ $[M]^+$ 600.2721, found 600.2700.

(((2S,3R,4S,5S,6R)-3,4,5-Trihydroxy-6-((hydroxyamino)methyl)tetrahydro-2H-pyran-2-yl) methyl)phosphonic acid (D9).



To a solution of **6** (180 mg, 0.3 mmol) in anhydrous CH_2Cl_2 (3 mL), iodotrimethylsilane (427 µL, 3.0 mmol, 10 equiv) was added dropwise at 0 °C under nitrogen. The resulting solution was then stirred at room temperature for 3h. The reaction was quenched by the addition of methanol. The mixture was concentrated, dissolved in H_2O (10 mL) and washed with diethyl ether (20mL x 5). The aqueous layer was lyophilized to afford the target **D9** as a colorless foam (78 mg, 95%). ^{1}H NMR (500 MHz, D_2O), δ_H (ppm) 2.01-2.06 (1H, m, CH_2-P), 2.21-2.27 (1H, m, CH_2-P), 3.31 (1H, t, $^{3}J_{HH}$ = 9.5 Hz, H-4), 3.57 (1H, dd, $^{2}J_{HH}$ = 13.6 Hz, $^{3}J_{HH}$ = 10.0 Hz, H-6a or H-6b), 3.57-3.62 (2H, m, H-3, H-6a or H-6b), 3.70-3.75 (1H, m, H-2), 3.92 (1H, dd, $^{3}J_{HH}$ = 9.8 Hz, $^{3}J_{HH}$ = 2.3 Hz, H-5), 4.39-4.41 (1H, m, H-1). ^{13}C NMR (125 MHz, D_2O), δ_C (ppm) 23.29 (d, $^{1}J_{CP}$ = 136.9 Hz, CH_2-P), 52.50 (CH_2-6), 66.28 (C-5), 70.55 (d, $^{3}J_{CP}$ = 12.9 Hz, C-2), 71.73 (C-4), 72.26 (d, $^{2}J_{CP}$ = 5.8 Hz, C-1), 72.60 (C-3). ^{31}P NMR (121 MHz, D_2O), δ_P (ppm) 24.50. HRMS (ESI, Positive Mode) calcd for $C_7H_{17}NO_8P$ $[M+1]^+$ 274.0686, found 274.0694.

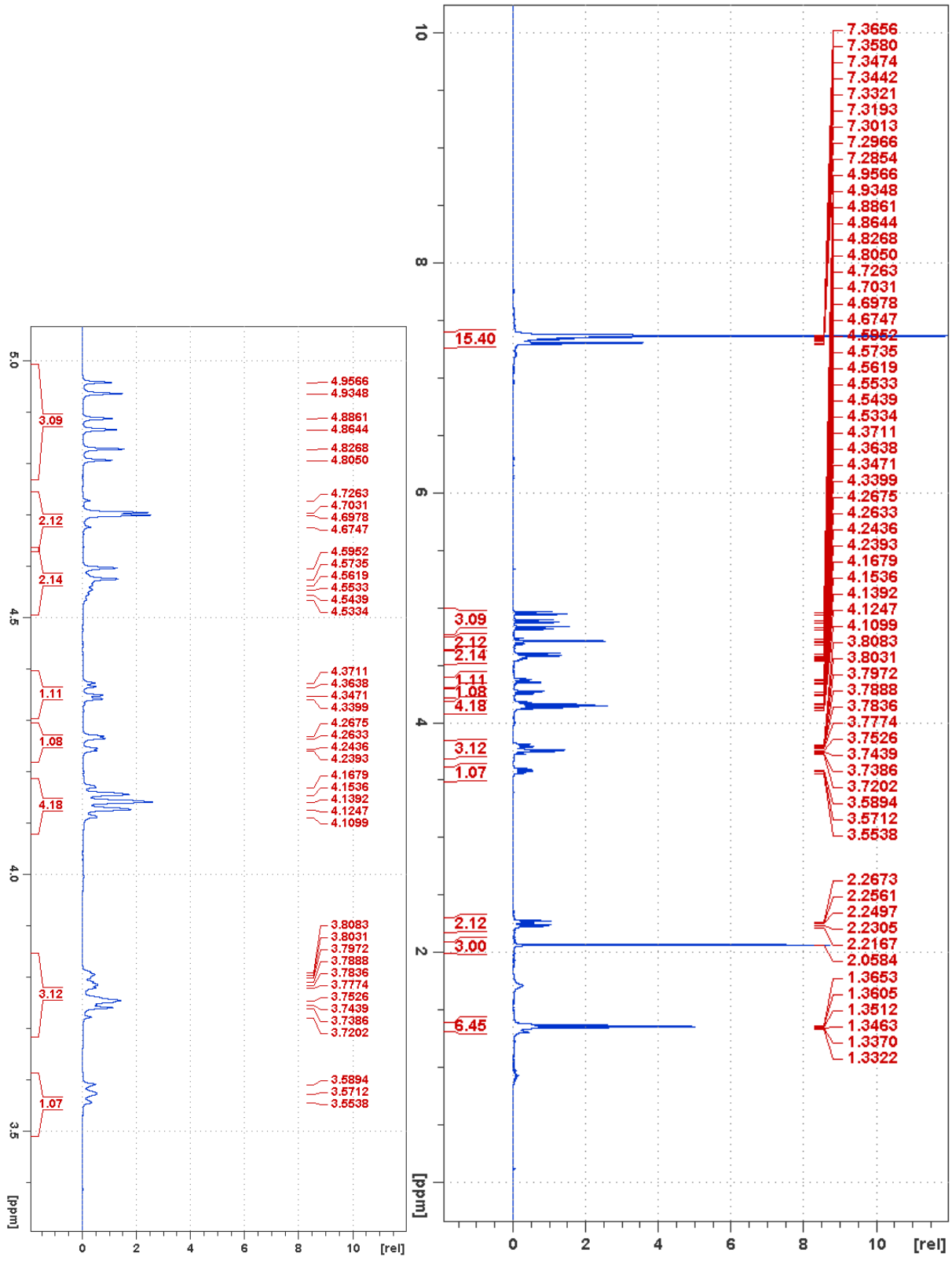


Figure S10. ^{1}H NMR of compound 2

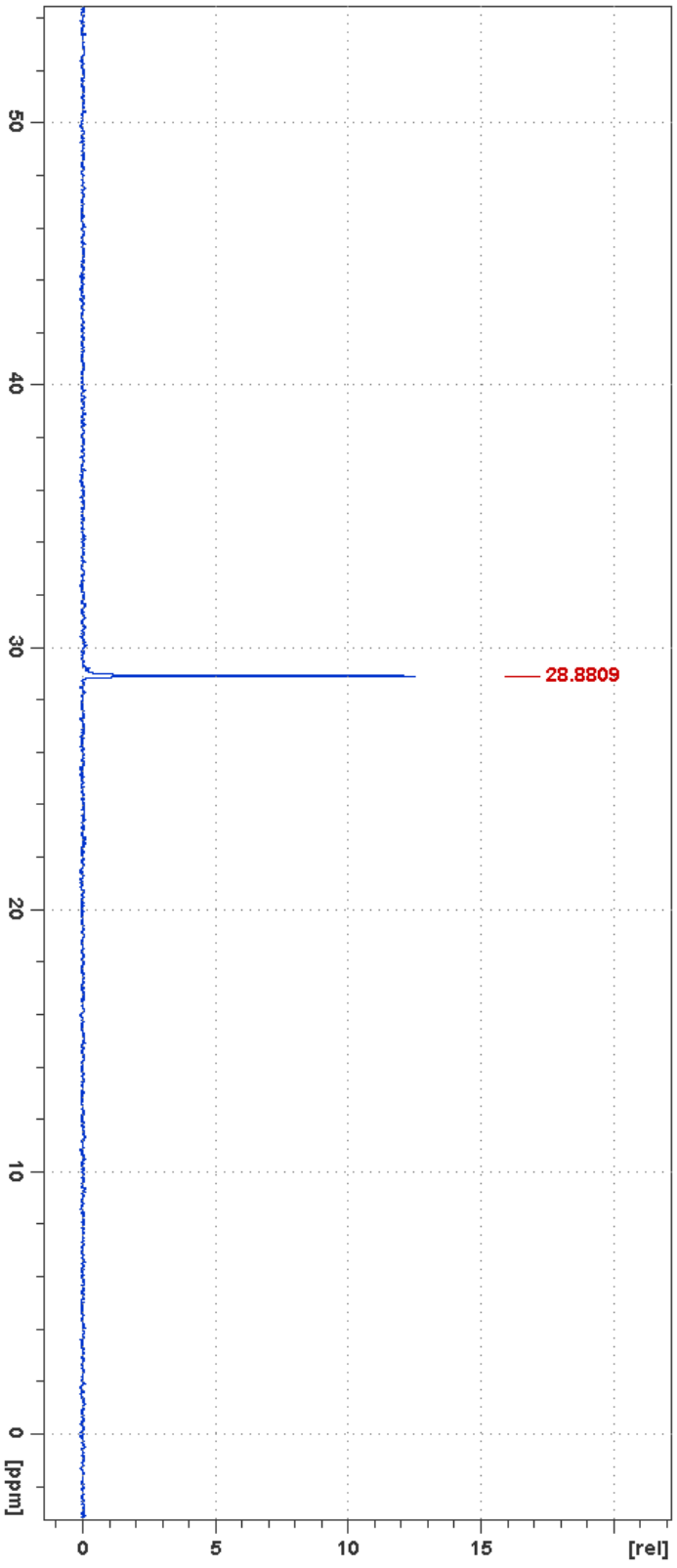


Figure S11. ^{31}P NMR of compound 2

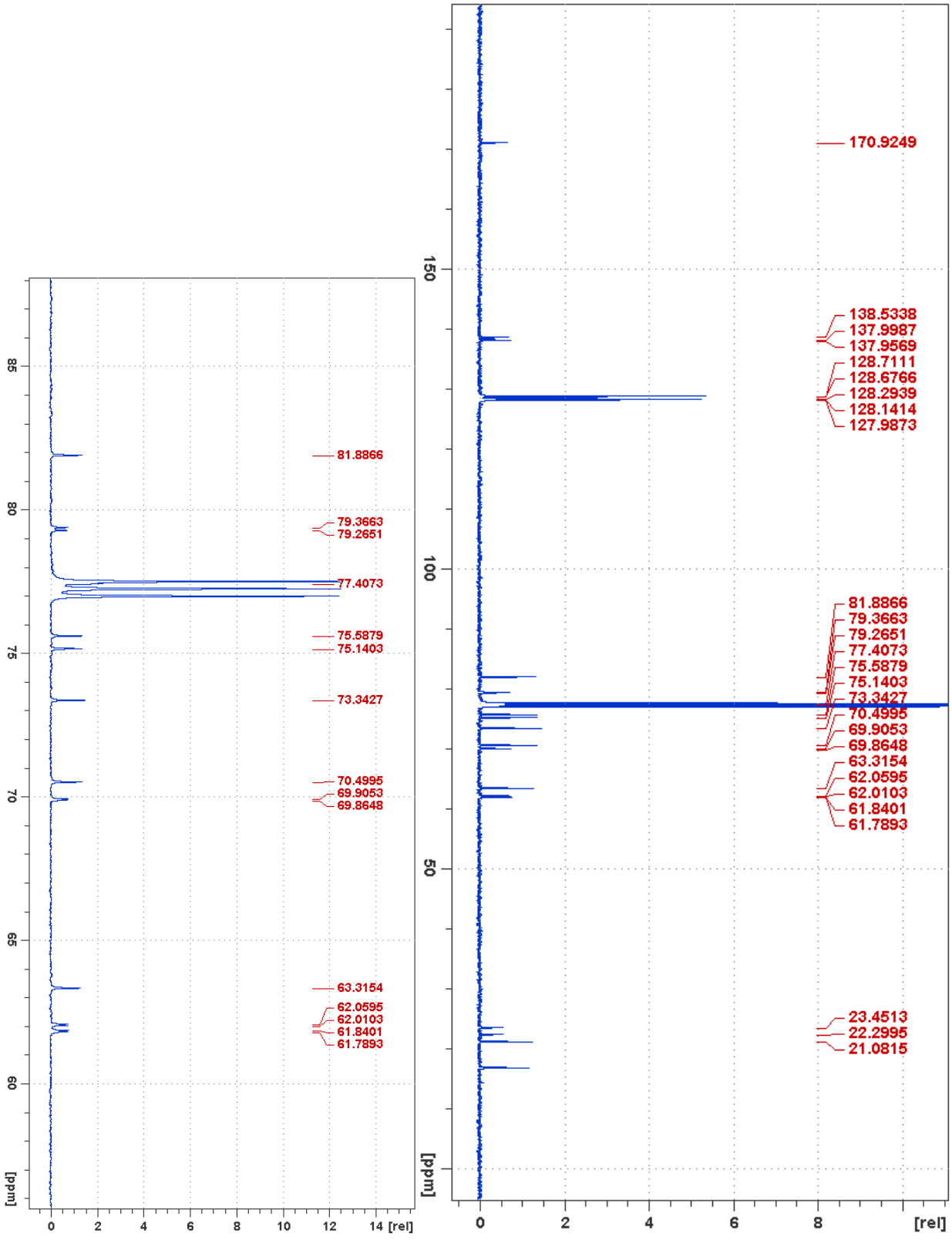


Figure S12. ^{13}C NMR of compound 2

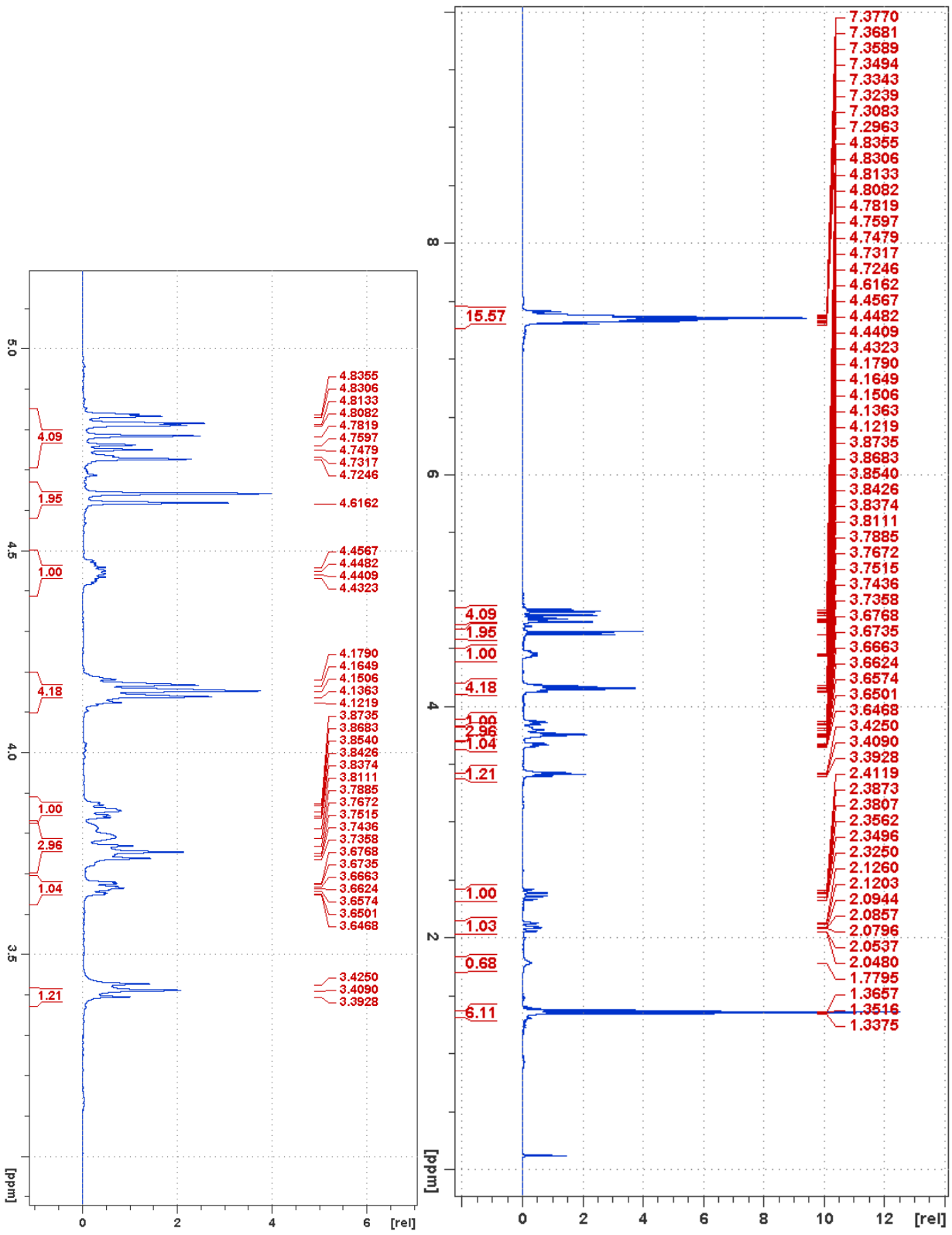


Figure S13. ^{1}H NMR of compound 3

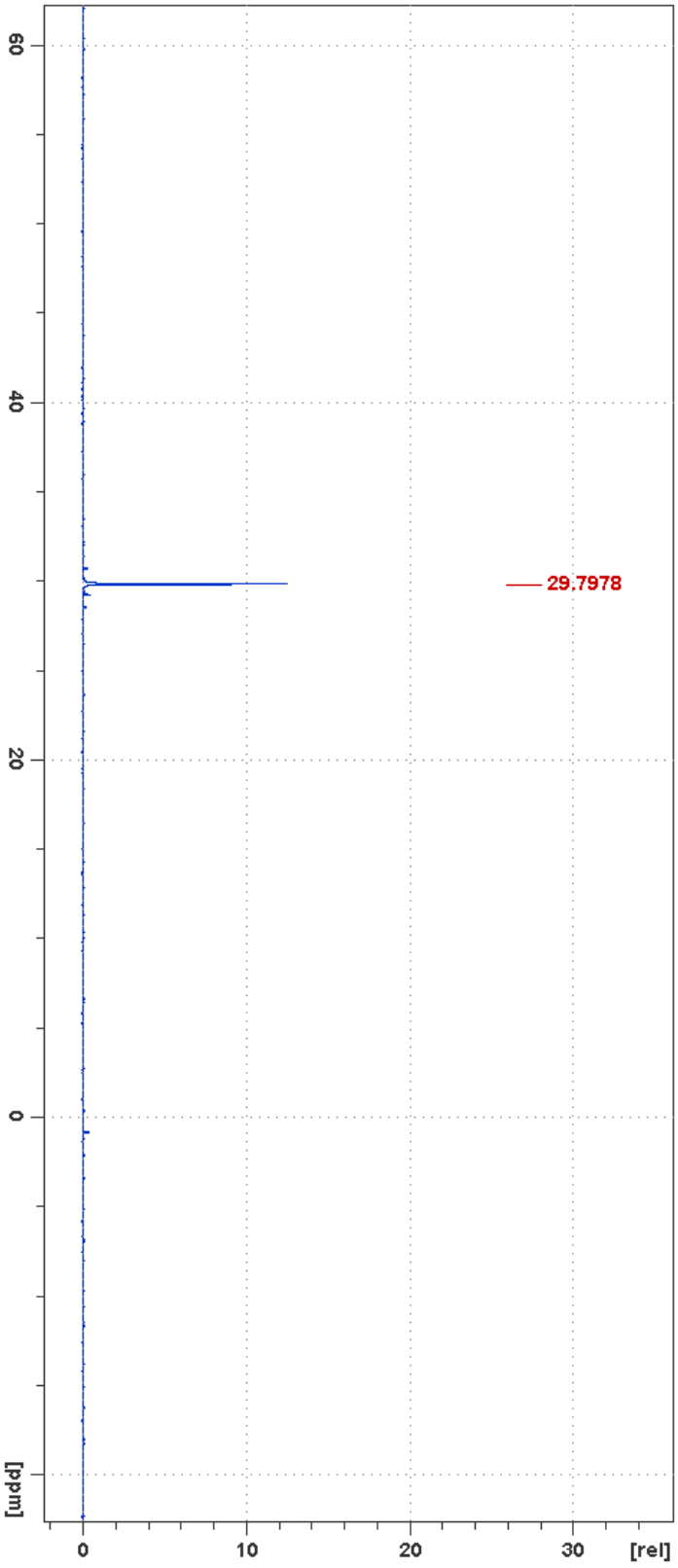


Figure S14. ^{31}P NMR of compound 3

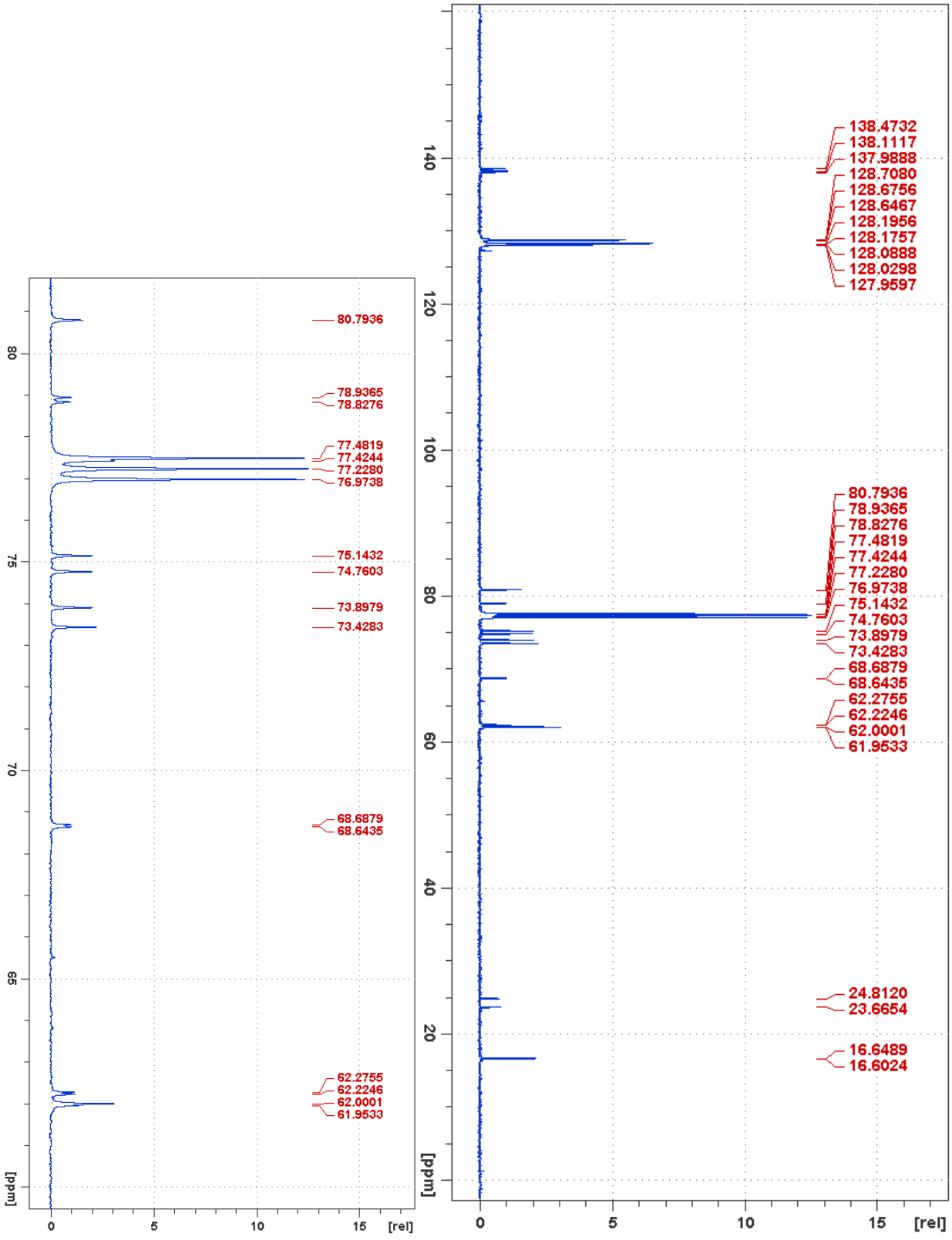


Figure S15. ^{13}C NMR of compound 3

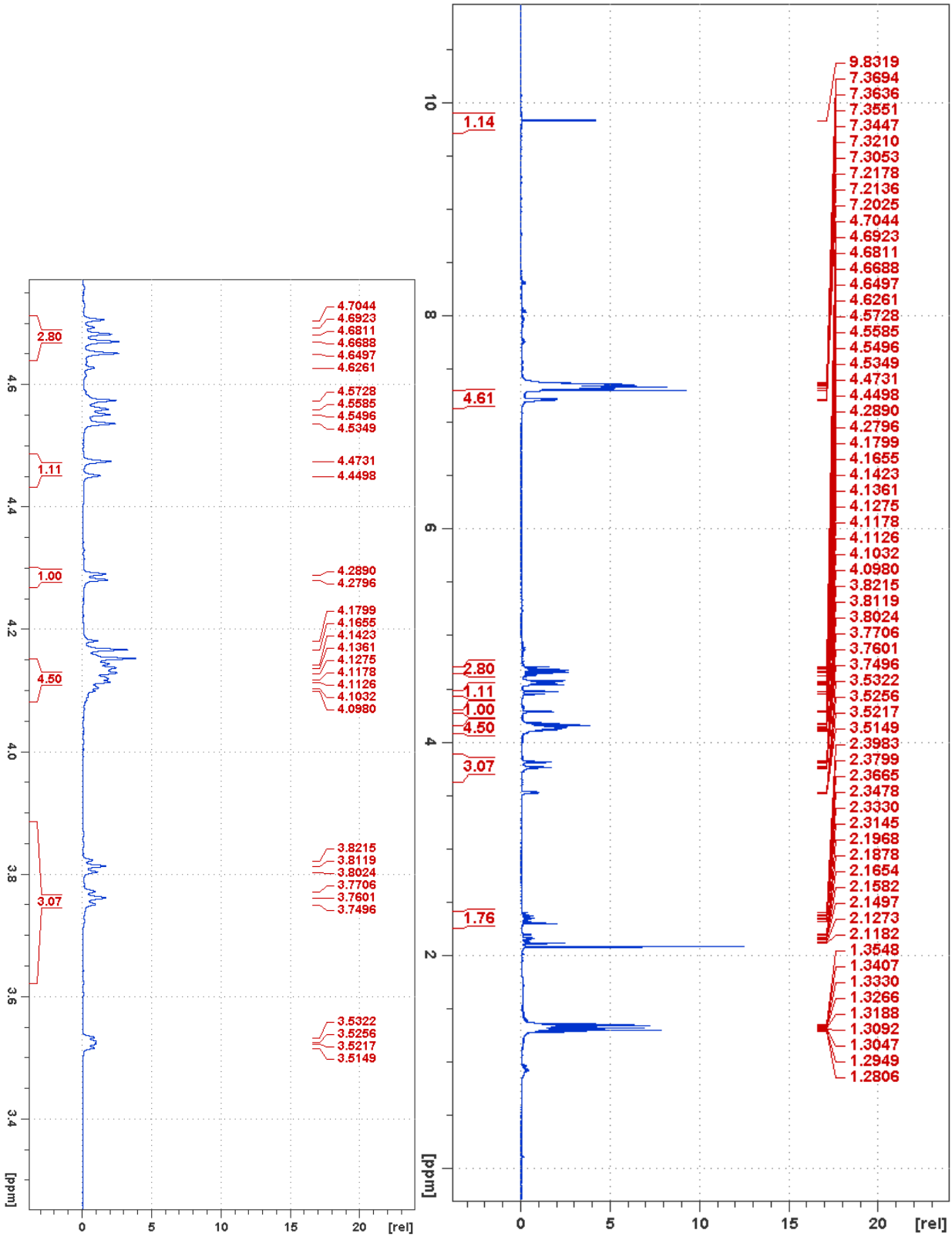


Figure S16. ^{1}H NMR of compound 4

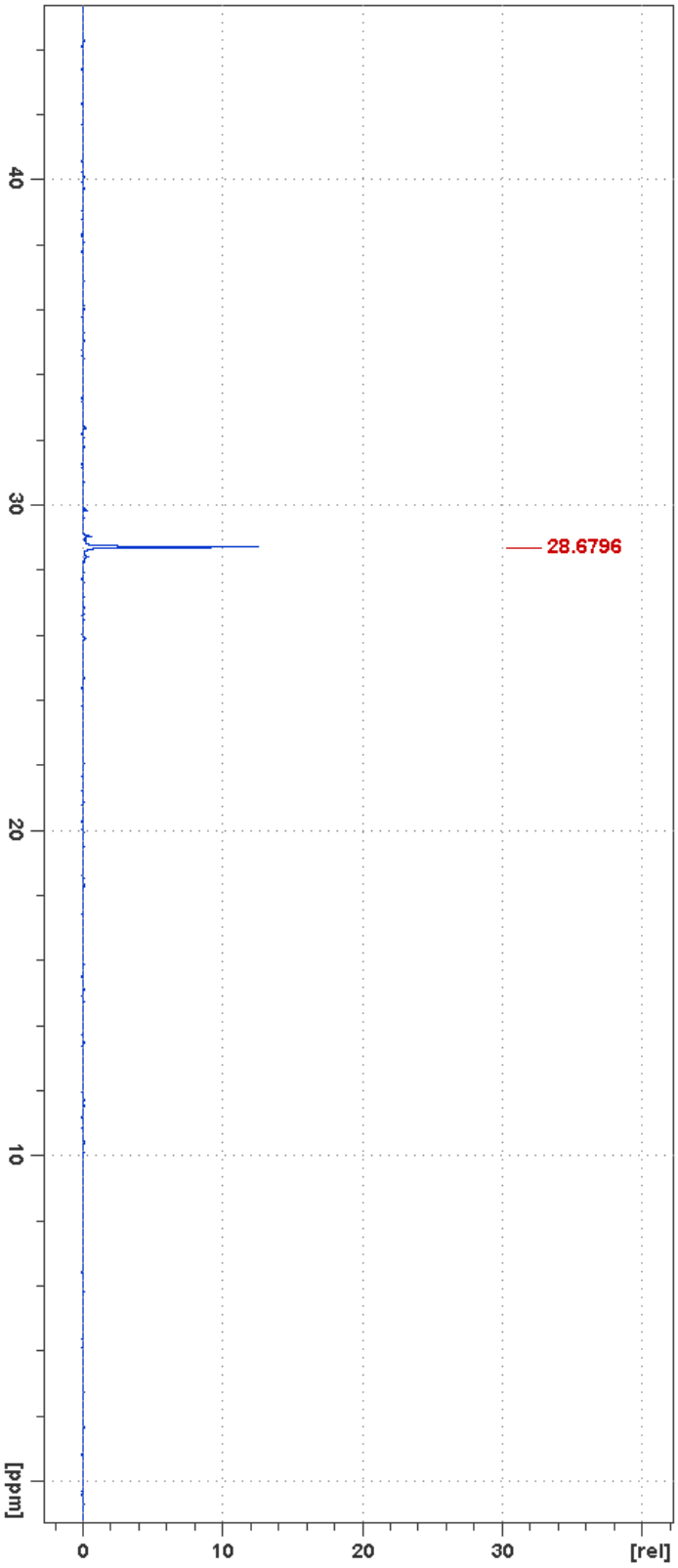


Figure S17. ^{31}P NMR of compound 4

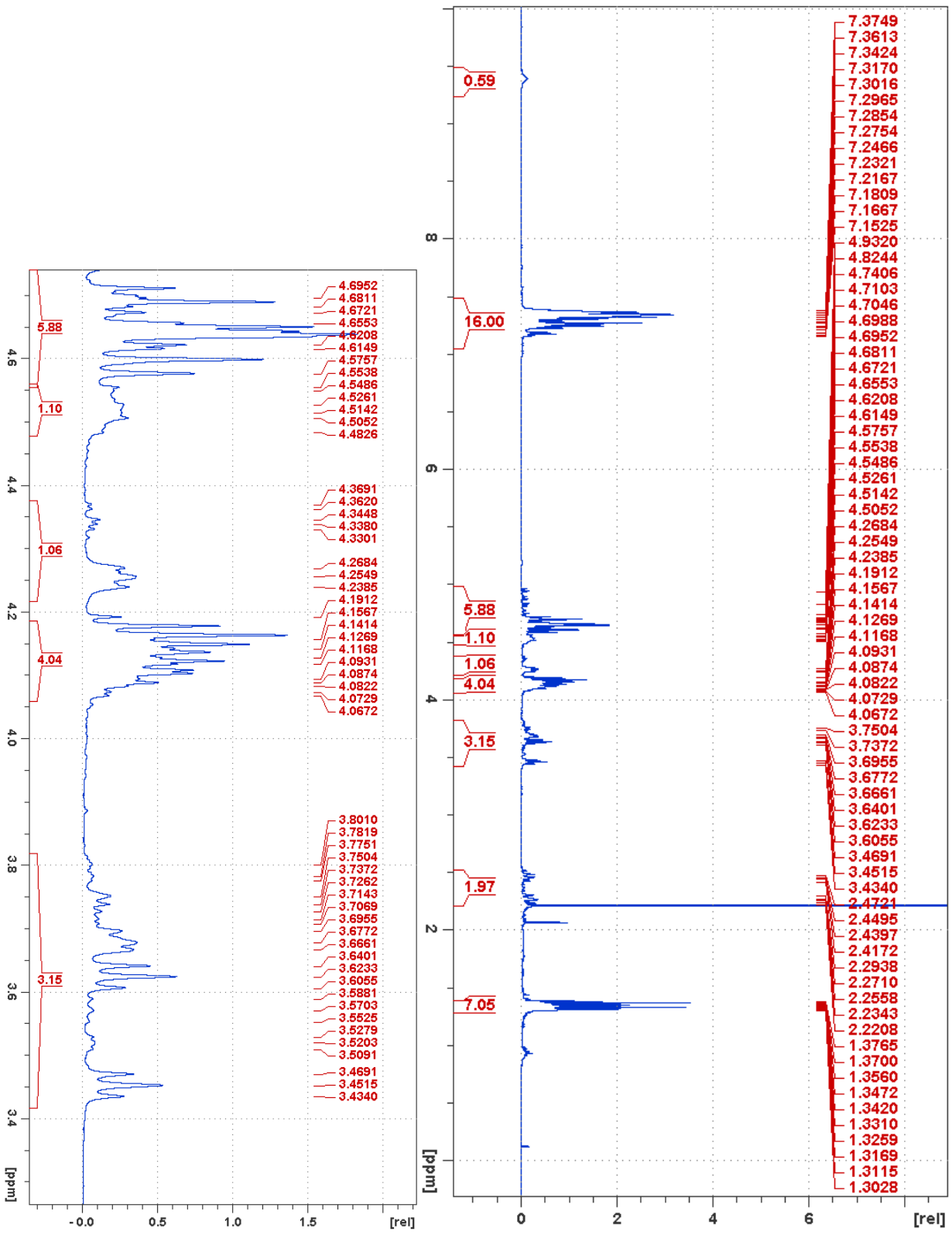


Figure S18. ^{1}H NMR of compound 5

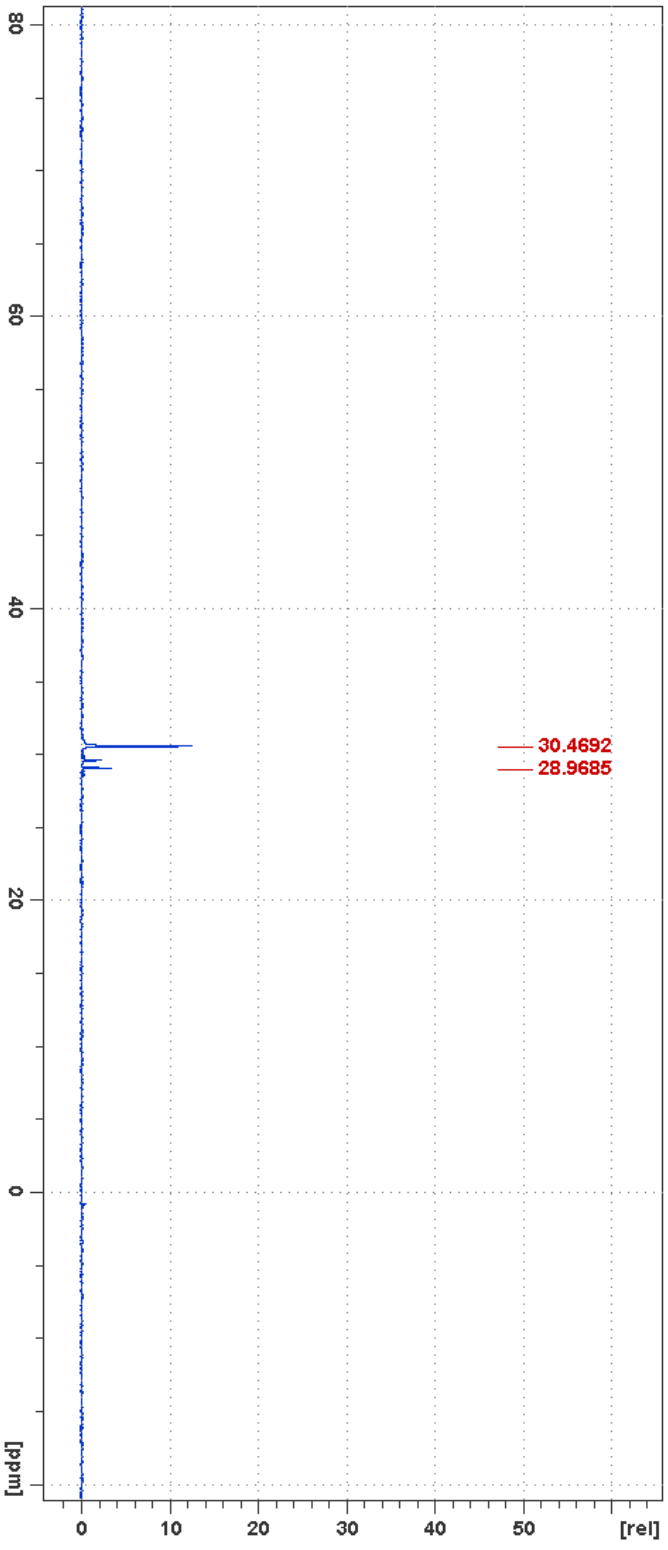


Figure S19. ^{31}P NMR of compound 5

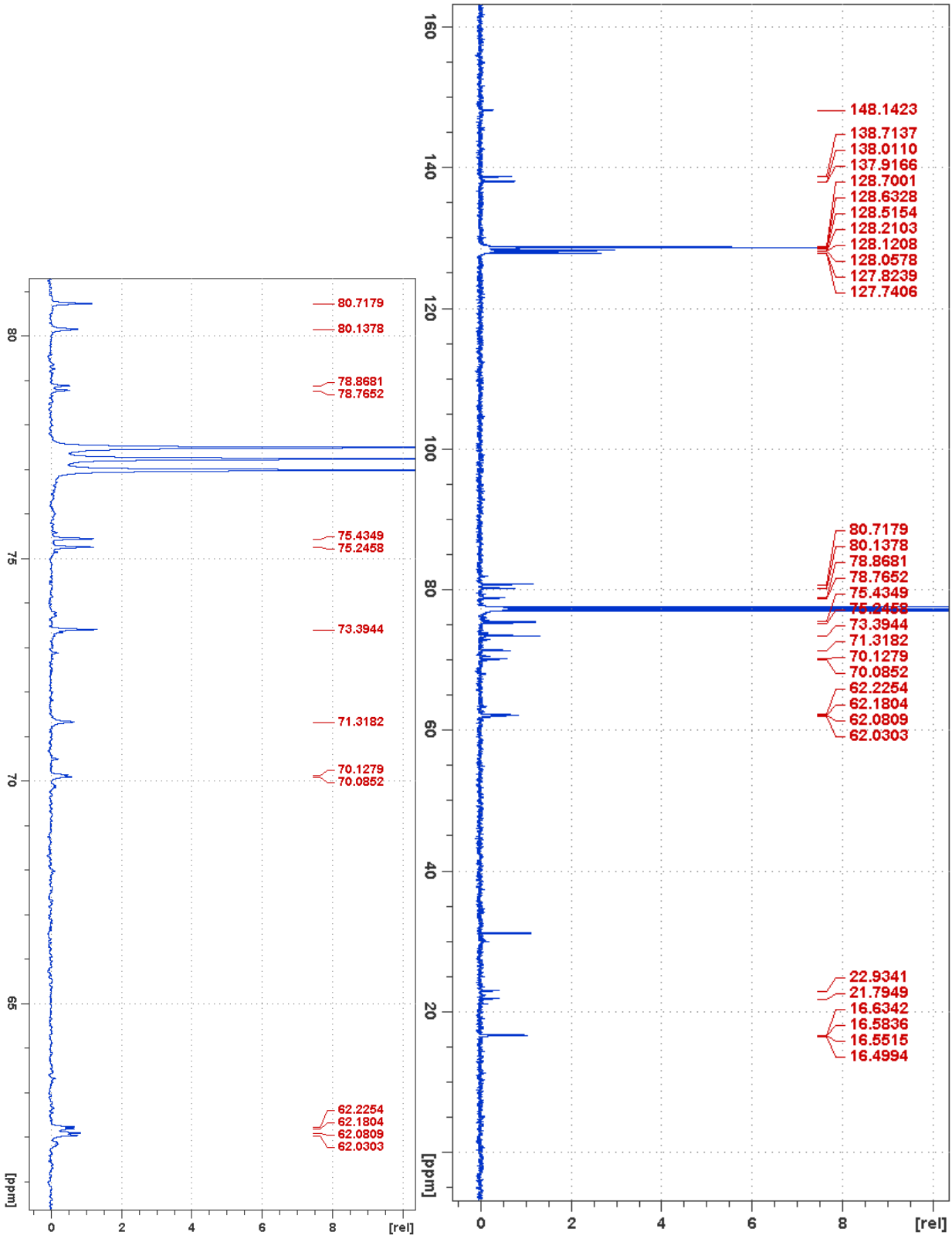


Figure S20. ^{13}C NMR of compound 5

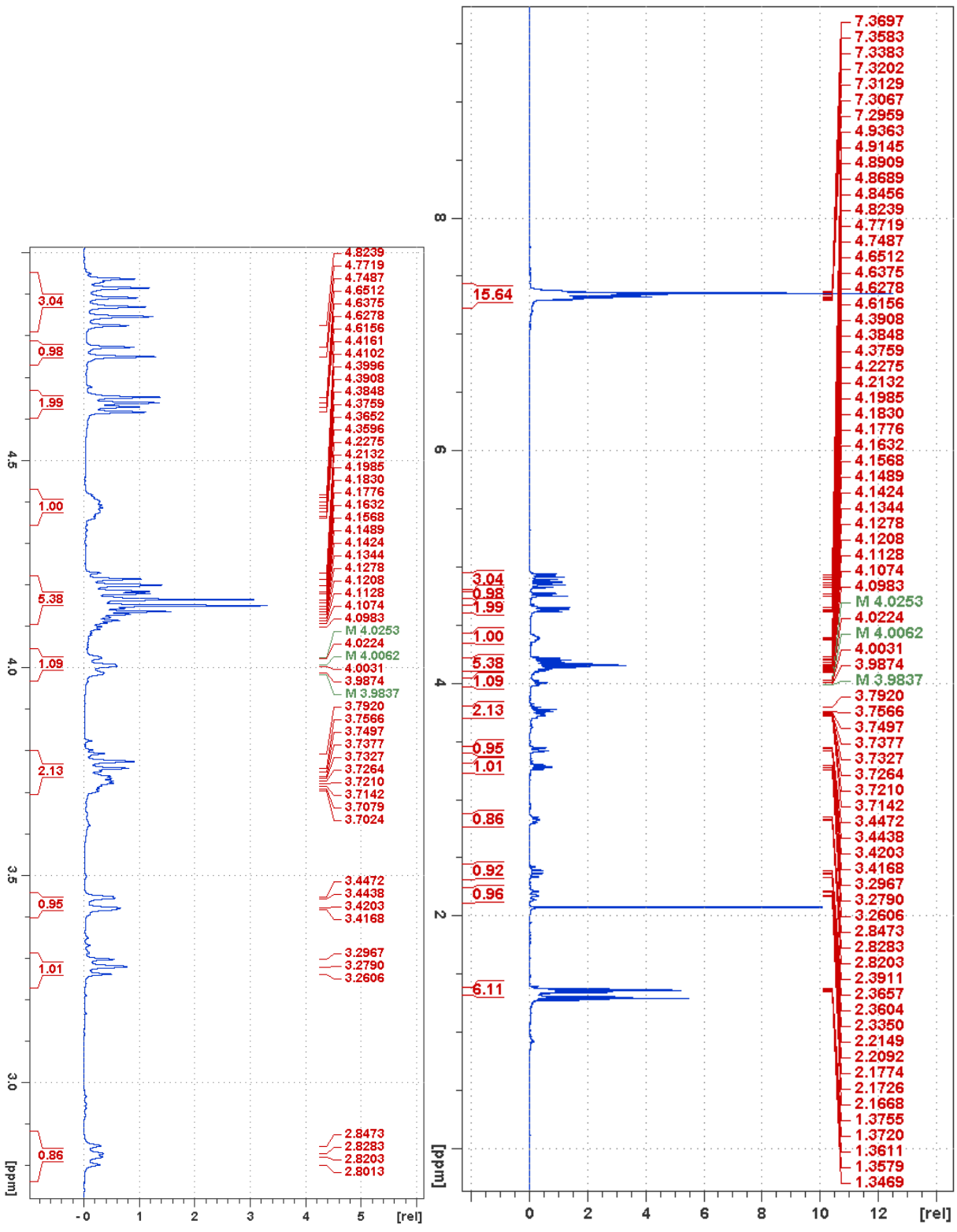


Figure S21. ^{1}H NMR of compound 6

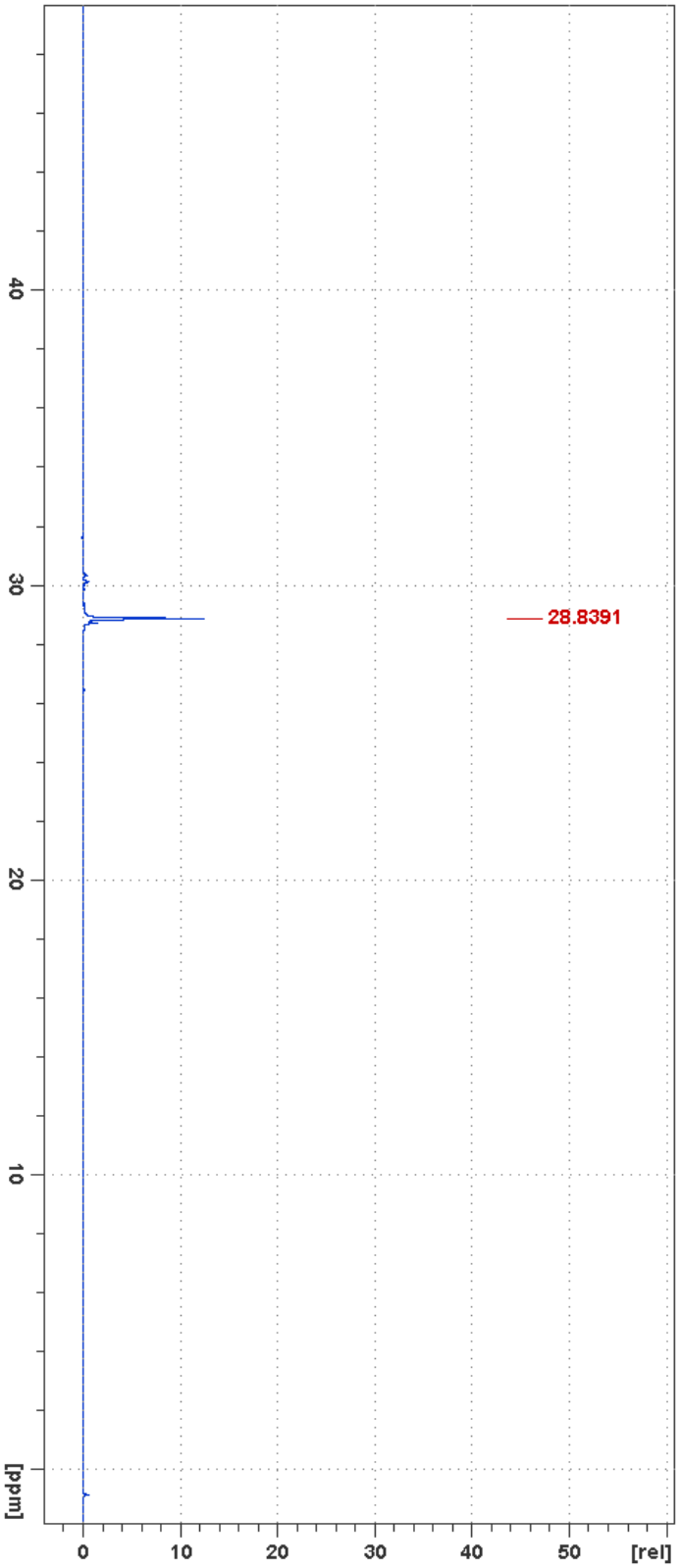


Figure S22. ^{31}P NMR of compound 6

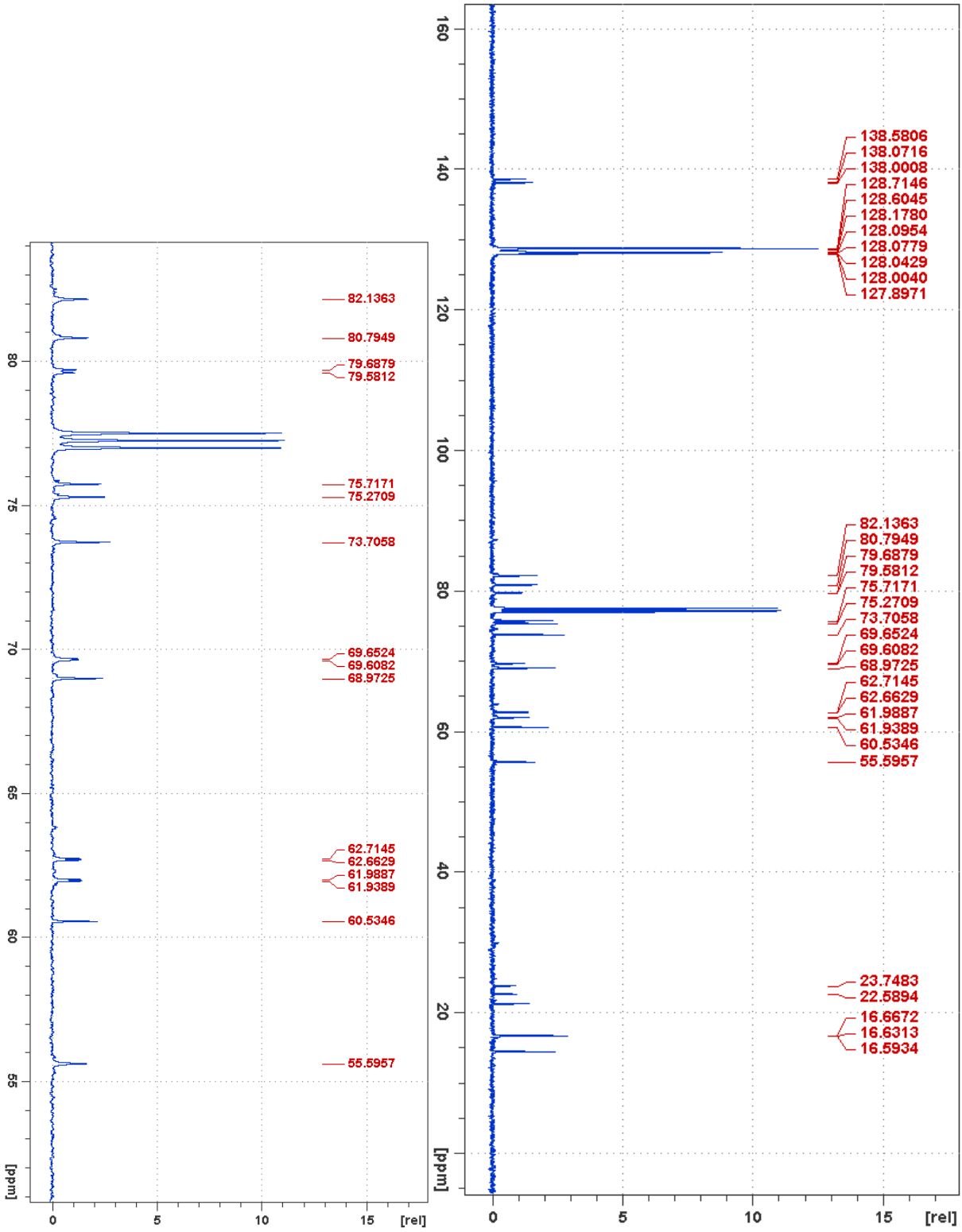


Figure S23. ^{13}C NMR of compound 6

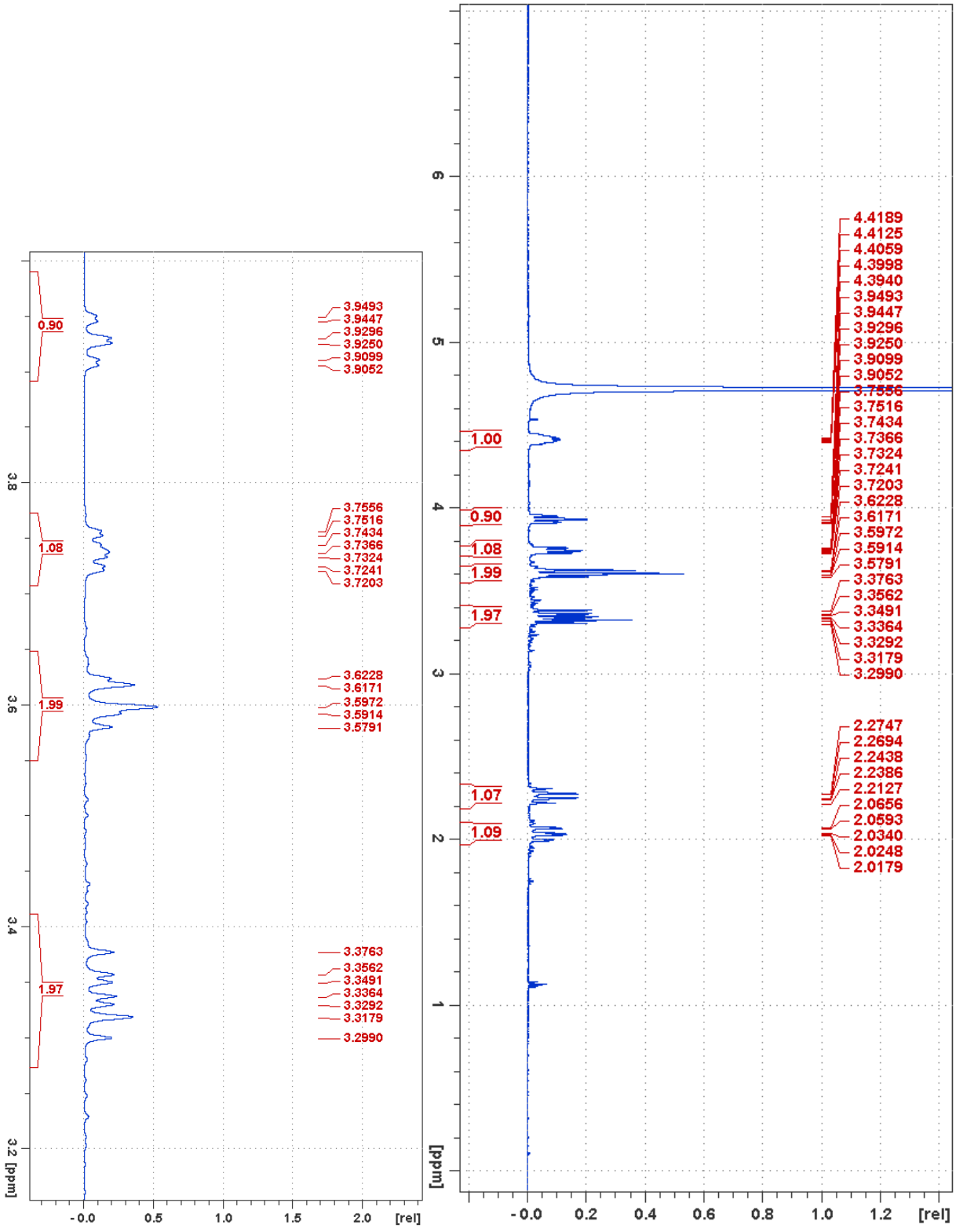


Figure S24. ^{1}H NMR of compound D9

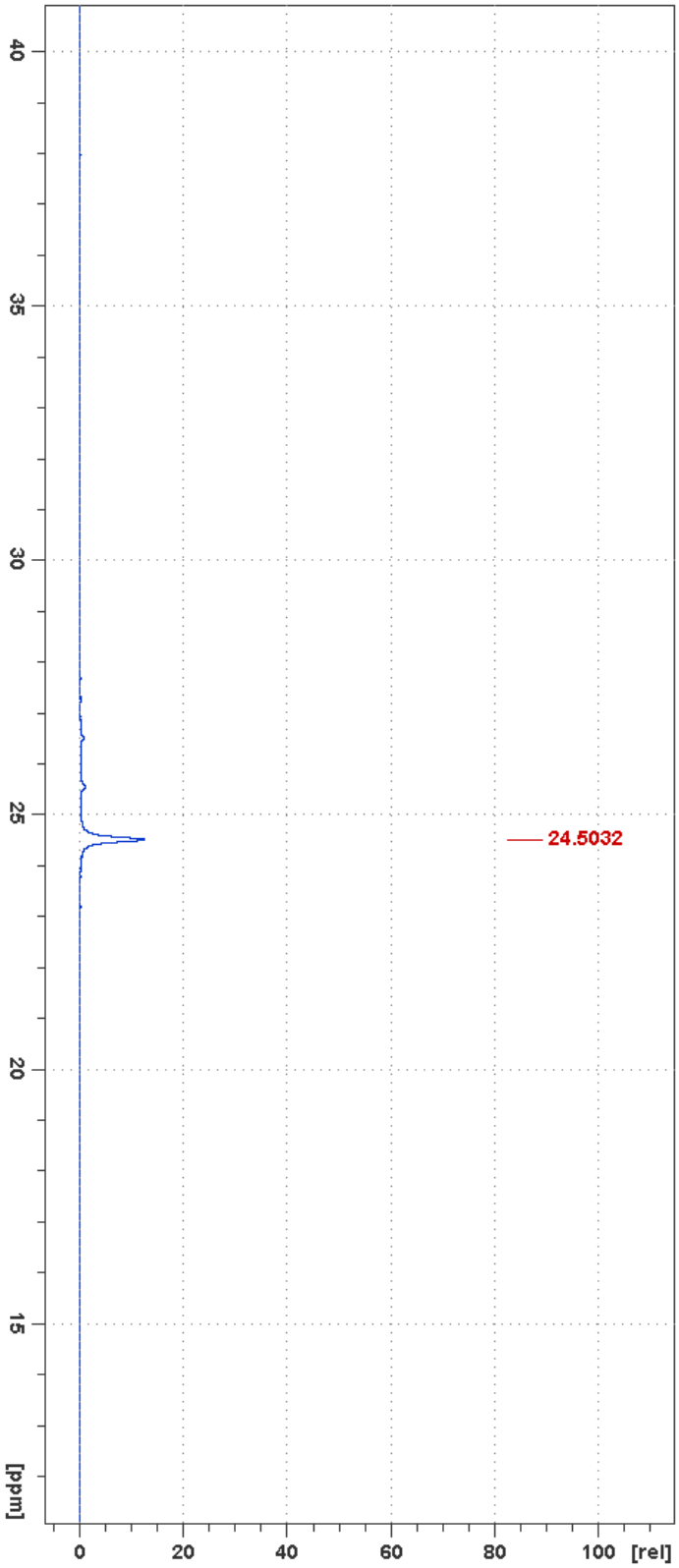


Figure S25. ^{31}P NMR of compound D9

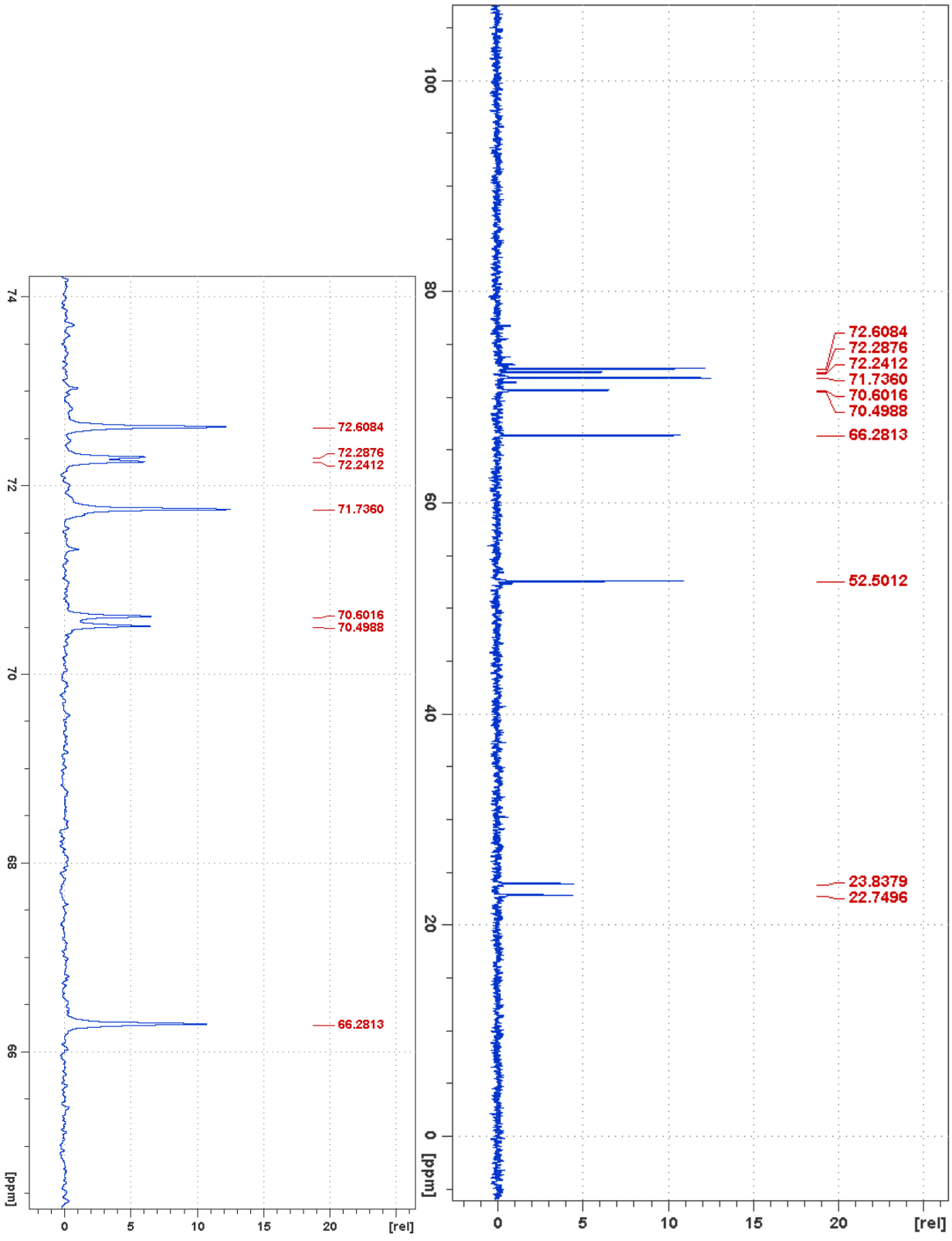


Figure S26. ^{13}C NMR of compound D9

4.7 Tables

Table 1. Phosphomannomutase and phosphoglucomutase activity of HAD5.

Displayed are the mean ± SEM of three independent trials for the kinetic parameters of HAD5 converting M6P to M1P (PMM) or G1P to G6P (PGM).

Enzyme Assay	k_{cat} (s^{-1})	K_m (µM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
PMM	$1.7 \times 10^{-4} \pm 1.9 \times 10^{-5}$	32 ± 3.6	5.4
PGM	$7.8 \times 10^{-6} \pm 1.2 \times 10^{-6}$	5.2 ± 1.4	1.5

Table S1. Summary of crystallographic statistics.

Crystal	PfHAD5
Space group	$P6_522$
Cell dimensions	*a*=*b*=161.3. Å, *c*=109.7 Å
Data Collection	
Wavelength	0.979 Å
Resolution range (highest shell)	40.3-3.50 Å (3.56-3.520 Å)
Reflections (total/unique)	354,596 / 11,030
Completeness (highest shell)	99.9% (100%)
<I/□> (highest shell)	18.5 (2.7)
R_{sym} (highest shell)	33.4% (143%)
Refinement	
R_{crys} / R_{free}	25.4% / 31.0%
No. of protein atoms	3,941
No. of ligand atoms	2
R.m.s. deviation, bond lengths	0.003 Å
R.m.s. deviation, bond angles	0.74°
Avg. B-factor: protein, ligand	92.7, 59.1 $Å^2$
Stereochemistry: favored, allowed, outliers	94.7, 4.5, 0.8 %

Table S2: Primers used for cloning

Primer Name	**Sequence**	**Notes**
P1	CTCACCACCACCACCACCATATGAATAAGAAAAAAGGCAATATTTCTGT	For cloning HAD5 sequence from *P. falciparum* cDNA into BG1861 vector
P2	ATCCTATCTTACTCACTTACAAGAAATTCTCTCTTAAAATTTTAAC	For cloning HAD5 sequence from *P. falciparum* cDNA into BG1861 vector
P3	GAGAAGACCGATTAAAAAAATTGATAAATTATAGTTTAAAATATATTGCC	For reverting the HAD5 gene sequence to the reference sequence.
P4	GGCAATATTTCTGTTTGCTGTAGATGGGACCC	For generating $HAD5^{D11A}$ mutant for recombinant protein.
P5	GCCGCGCGGCAGCCATATGGCAGTTACA	Forward primer for cloning HsPMM1 gblock into pET28a vector.
P6	CGGAGCTCGAATTCGGATCCTATCTTACTC	Reverse primer for cloning HsPMM1 and Hs PMM2 gblocks into pET28a vector.
P7	CGGAGCTCGAATTCGGATCCTATCTTACTCACTTA	Forward primer for cloning HsPMM2 gblock into pET28a vector.

Table S3: gBlock sequences used in cloning

Gene Name	Sequence
*Hs*PMM1	ATGGCAGTTACAGCCCAGGCAGCCCGTCGTAAGGAGCGTGTCTTATGTC TGTTCGATGTAGACGGAACTCTGACCCCCGCACGTCAAAAAATCGACCC GGAAGTTGCAGCTTTTTTGCAGAAGCTGCGTTCGCGCGTCCAGATCGGT GTAGTCGGCGGATCAGATTACTGCAAAATCGCCGAGCAACTTGGAGAT GGCGACGAAGTGATCGAGAAGTTTGACTACGTCTTCGCCGAGAATGGG ACAGTTCAATACAAGCATGGGCGCTTATTGAGTAAGCAGACTATTCAGA ACCATCTGGGGGAGGAGTTGCTTCAAGATCTTATTAATTTTTGTTTATCC TATATGGCCTTACTTCGCCTGCCCAAAAAGCGCGGTACTTTCATTGAGT TCCGTAACGGGATGCTGAACATCAGTCCAATCGGTCGCTCATGCACTCT GGAGGAGCGTATCGAGTTTTCTGAACTTGACAAGAAAGAGAAAATTCG TGAGAAATTCGTCGAGGCGTTAAAAACGGAGTTTGCAGGGAAGGGATT ACGCTTTTCTCGCGGAGGCATGATTTCATTCGACGTGTTTCCAGAAGGT TGGGACAAGCGCTACTGCTTGGACTCATTAGATCAAGATAGCTTTGATA CCATTCACTTTTTCGGGAACGAAACCTCGCCTGGGGGTAACGACTTCGA GATCTTTGCGGACCCTCGTACGGTCGGGCACTCGGTAGTGAGCCCTCAG GACACCGTGCAACGTTGTCGTGAGATTTTTTTCCCAGAGACGGCGCATG AAGCGTAAGTGAGTAAGATAGGATCCGAATTCGAGCTCCG
*Hs*PMM2	ATGGCGGCTCCGGGCCCAGCATTATGTTTATTTGACGTTGACGGAACCC TTACCGCACCGCGTCAAAAGATCACGAAGGAAATGGATGATTTTTTGCA GAAGTTACGTCAGAAGATCAAAATCGGGGTGGTCGGTGGTTCCGATTTT GAGAAAGTTCAGGAGCAGCTTGGAAACGACGTGGTTGAGAAGTACGAT

	TACGTCTTTCCGGAAAATGGGTTGGTCGCGTATAAGGACGGTAAACTGC TTTGTCGTCAAAATATTCAGTCCCATCTGGGCGAAGCCTTGATTCAAGA TTTAATCAATTATTGCTTATCCTATATCGCTAAGATCAAATTGCCCAAGA AACGCGGCACCTTTATTGAGTTTCGTAATGGCATGTTGAACGTGTCCCC GATCGGACGTTCGTGTTCCCAGGAGGAACGCATTGAGTTCTATGAACTG GATAAAAAAGAAAATATCCGTCAAAAGTTCGTTGCCGATCTTCGCAAG GAGTTCGCAGGCAAAGGTTTAACGTTCTCAATCGGCGGTCAAATCTCTT TCGATGTGTTCCCAGACGGATGGGACAAACGTTACTGTCTTCGCCATGT AGAAAATGATGGATATAAAACCATCTACTTTTTTGGGGACAAAACAAT GCCAGGAGGGAATGACCATGAAATTTTCACGGACCCCCGTACAATGGG CTACTCAGTAACCGCACCGGAAGATACCCGTCGTATTTGCGAGCTGTTG TTCTCTTAAGTGAGTAAG
*Ec*ManC	CTCACCACCACCACCACCATATGGCTCAAAGTAAACTTTACCCAGTGGT AATGGCGGGGGGAAGTGGCTCTCGTCTTTGGCCTTTATCTCGCGTTCTTT ATCCAAAACAGTTCTTGTGCCTTAAGGGTGATTTGACAATGTTGCAAAC AACGATCTGCCGCCTGAATGGTGTCGAATGTGAGTCACCTGTGGTAATT TGCAATGAACAGCACCGTTTCATCGTAGCAGAACAACTGCGTCAGCTGA ATAAATTAACGGAAAATATTATCCTGGAGCCTGCTGGTCGCAACACGGC ACCTGCAATCGCCTTGGCTGCACTTGCGGCCAAGCGTCACTCACCAGAA AGTGACCCGCTTATGCTTGTCTTGGCCGCCGATCATGTGATCGCAGATG AAGATGCTTTTCGTGCTGCTGTCCGTAATGCGATGCCATATGCCGAAGC TGGAAAGTTAGTTACGTTTGGTATCGTGCCGGATTTGCCGGAGACTGGA TATGGTTATATCCGCCGTGGGGAGGTCAGCGCTGGAGAACAAGACATG

	GTAGCCTTCGAGGTGGCACAATTTGTTGAAAAGCCAAATCTTGAGACAG CACAAGCGTACGTGGCTAGTGGGGAGTATTACTGGAACTCTGGGATGTT TTTATTCCGTGCGGGGCGTTACCTTGAAGAACTGAAAAAATATCGTCCT GACATTTTAGACGCGTGTGAAAAGGCAATGAGTGCAGTGGACCCAGAC TTAAACTTTATTCGTGTAGACGAAGAGGCTTTCTTGGCATGTCCAGAAG AATCCGTTGACTATGCCGTGATGGAGCGTACGGCTGATGCGGTAGTTGT CCCAATGGACGCTGGATGGTCCGATGTTGGCAGCTGGTCATCGCTTTGG GAAATTAGCGCCCACACCGCGGAAGGAAATGTATGTCACGGCGACGTG ATTAACCATAAAACAGAAAATTCATACGTTTATGCGGAATCCGGCTTGG TCACTACTGTCGGAGTGAAGGATTTGGTTGTGGTGCAAACGAAAGATGC AGTATTAATTGCGGACCGCAATGCTGTCCAAGATGTGAAGAAAGTAGTT GAACAGATTAAGGCTGATGGTCGTCACGAGCATCGCGTCCACCGCGAA GTATATCGCCCATGGGGAAAGTATGACTCTATTGATGCGGGTGACCGCT ATCAAGTTAAACGTATCACAGTCAAACCCGGTGAGGGGCTTTCGGTGCA AATGCATCATCATCGCGCAGAGCATTGGGTAGTGGTTGCGGGTACTGCG AAAGTAACAATTGATGGCGATATCAAGTTGCTTGGCGAAAATGAGTCA ATTTACATCCCGCTGGGCGCGACACACTGTCTTGAGAATCCGGGGAAAA TCCCATTGGATTTAATCGAAGTGCGTTCTGGATCTTACTTAGAGGAGGA TGACGTTGTTCGTTTTGCTGATCGTTACGGACGCGTTtAAGTGAGTAAGA TAGGAT

4.8 References

1. World Health Organization. *World Malaria Report 2020: 20 years of global progress and challenges*. (2020).
2. Haldar, K., Bhattacharjee, S. & Safeukui, I. Drug resistance in Plasmodium. *Nature Reviews Microbiology* **16**, 156–170 (2018).
3. Yeung, S. Malaria - Update on Antimalarial Resistance and Treatment Approaches. *Pediatr. Infect. Dis. J.* **37**, 367–369 (2018).
4. Meibalan, E. & Marti, M. Biology of malaria transmission. *Cold Spring Harb. Perspect. Med.* **7**, DOI: 10.1101/cshperspect.a025452 (2017).
5. Alonso, P. L. *et al.* A research Agenda to underpin Malaria Eradication. *PLoS Medicine* **8**, DOI: 10.1371/journal.pmed.1000406 (2011).
6. Burrows, J. N. *et al.* New developments in anti-malarial target candidate and product profiles. *Malaria Journal* **16**, 26 (2017).
7. Moyo, P. *et al.* Natural products: A potential source of malaria transmission blocking drugs? *Pharmaceuticals* **13**, 251 (2020).
8. Mehta, M., Sonawat, H. M. & Sharma, S. Malaria parasite-infected erythrocytes inhibit glucose utilization in uninfected red cells. *FEBS Lett.* **579**, 6151–6158 (2005).
9. Roth, E. F., Raventos-Suarez, C., Perkins, M. & Nagel, R. L. Glutathione stability and oxidative stress in P. falciparum infection in vitro: Responses of normal and G6PD deficient cells. *Biochem. Biophys. Res. Commun.* **109**, 355–362 (1982).
10. Zhang, B. *et al.* A second target of the antimalarial and antibacterial agent fosmidomycin revealed by cellular metabolic profiling. *Biochemistry* **50**, 3570–3577 (2011).
11. Ferone, R. Folate metabolism in malaria. *Bulletin of the World Health Organization* **55**,

291–298 (1977).

12. Anderson, A. C. Targeting DHFR in parasitic protozoa. *Drug Discovery Today* **10**, 121–128 (2005).

13. Rieckmann, K. H., Yeo, A. E. & Edstein, M. D. Activity of PS-15 and its metabolite, WR99210, against Plasmodium falciparum in an in vivo-in vitro model. *Trans. R. Soc. Trop. Med. Hyg.* **90**, 568–571 (1996).

14. Kokkonda, S. *et al.* Isoxazolopyrimidine-Based Inhibitors of Plasmodium falciparum Dihydroorotate Dehydrogenase with Antimalarial Activity. *ACS Omega* **3**, 9227–9240 (2018).

15. Azema, L., Baron, R. & Ladame, S. Targeting Enzymes with Phosphonate-Based Inhibitors: Mimics of Tetrahedral Transition States and Stable Isosteric Analogues of Phosphates. *Curr. Enzym. Inhib.* **2**, 61–72 (2006).

16. Gilson, P. R. *et al.* Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite Plasmodium falciparum. *Mol. Cell. Proteomics* **5**, 1286–1299 (2006).

17. von Itzstein, M., Plebanski, M., Cooke, B. M. & Coppel, R. L. Hot, sweet and sticky: the glycobiology of Plasmodium falciparum. *Trends in Parasitology* **24**, 210–218 (2008).

18. Gerold, P., Schofield, L., Blackman, M. J., Holder, A. A. & Schwarz, R. T. Structural analysis of the glycosyl-phosphatidylinositol membrane anchor of the merozoite surface proteins-1 and -2 of Plasmodium falciparum. *Mol. Biochem. Parasitol.* **75**, 131–143 (1996).

19. Das, S. *et al.* Processing of Plasmodium falciparum Merozoite Surface Protein MSP1 Activates a Spectrin-Binding Function Enabling Parasite Egress from RBCs. *Cell Host*

Microbe **18**, 433–444 (2015).

20. Woehlbier, U. *et al.* Analysis of antibodies directed against merozoite surface protein 1 of the human malaria parasite Plasmodium falciparum. *Infect. Immun.* **74**, 1313–1322 (2006).

21. Sherling, E. S. *et al.* The Plasmodium falciparum rhoptry bulb protein RAMA plays an essential role in rhoptry neck morphogenesis and host red blood cell invasion. *PLoS Pathog.* **15**, DOI: 10.1371/journal.ppat.1008049 (2019).

22. Kapoor, N. *et al.* Malaria Derived Glycosylphosphatidylinositol Anchor Enhances Anti-Pfs25 Functional Antibodies That Block Malaria Transmission. *Biochemistry* **57**, 516–519 (2018).

23. Williamson, K. C. Pfs230: From malaria transmission-blocking vaccine candidate toward function. *Parasite Immunology* **25**, 351–359 (2003).

24. Lee, S. M. *et al.* The Pfs230 N-Terminal fragment, Pfs230D1+: Expression and characterization of a potential malaria transmission-blocking vaccine candidate. *Malar. J.* **18**, 356 (2019).

25. Scaria, P. V. *et al.* Outer membrane protein complex as a carrier for malaria transmission blocking antigen Pfs230. *NPJ Vaccines* **4**, 24 (2019).

26. Casares, S., Brumeanu, T. D. & Richie, T. L. The RTS,S malaria vaccine. *Vaccine* **28**, 4880–4894 (2010).

27. Laurens, M. B. RTS,S/AS01 vaccine (Mosquirix^{TM}): an overview. *Hum. Vaccines Immunother.* **16**, 480–489 (2019).

28. Moorthy, V. & Binka, F. R21/Matrix-M: a second malaria vaccine? *The Lancet* **397**, 1782–1783 (2021).

29. Datoo, M. S. *et al.* Efficacy of a low-dose candidate malaria vaccine, R21 in adjuvant Matrix-M, with seasonal administration to children in Burkina Faso: a randomised controlled trial. *Lancet* **397**, 1809–1818 (2021).

30. Ménard, R. *et al.* Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes. *Nature* **385**, 336–340 (1997).

31. Gowda, D. C., Gupta, P. & Davidson, E. A. Glycosylphosphatidylinositol anchors represent the major carbohydrate modification in proteins of intraerythrocytic stage Plasmodium falciparum. *J. Biol. Chem.* **272**, 6428–6439 (1997).

32. Bushkin, G. G. *et al.* Suggestive evidence for Darwinian selection against asparagine-linked glycans of Plasmodium falciparum and Toxoplasma gondii. *Eukaryot. Cell* **9**, 228–241 (2010).

33. Samuelson, J. *et al.* The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets glycosyltranferases. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 1548–1553 (2005).

34. Cova, M., Rodrigues, J. A., Smith, T. K. & Izquierdo, L. Sugar activation and glycosylation in Plasmodium. *Malar. J.* **14**, 427 (2015).

35. Zhang, M. *et al.* Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360**, DOI: 10.1126/science.aap7847 (2018).

36. Otto, T. D. *et al.* New insights into the blood-stage transcriptome of Plasmodium falciparum using RNA-Seq. *Mol. Microbiol.* **76**, 12–24 (2010).

37. Toenhake, C. G. *et al.* Chromatin Accessibility-Based Characterization of the Gene Regulatory Network Underlying Plasmodium falciparum Blood-Stage Development. *Cell*

Host Microbe **23**, 557-569.e9 (2018).

38. López-Barragán, M. J. *et al.* Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum. *BMC Genomics* **12**, 587 (2011).

39. Kepes, F. & Schekman, R. The Yeast SEC53 Gene Encodes Phosphomannomutase. *J. Biol. Chem.* **263**, 9155–9161 (1988).

40. Oesterhelt, C., Schnarrenberger, C. & Gross, W. The reaction mechanism of phosphomannomutase in plants. *FEBS Lett.* **401**, 35–37 (1997).

41. Rose, Z. B. The glucose biphosphate family of enzymes. *Trends in Biochemical Sciences* **11**, 253–255 (1986).

42. Li, L., Kim, S. A., Fang, R. & Han, N. S. Expression of manB gene from escherichia coli in Lactococcus lactis and characterization of its bifunctional enzyme, phosphomannomutase. *J. Microbiol. Biotechnol.* **28**, 1293–1298 (2018).

43. Veiga-Da-Cunha, M., Vleugels, W., Maliekal, P., Matthijs, G. & Van Schaftingen, E. Mammalian phosphomannomutase PMM1 is the brain IMP-sensitive glucose-1,6-bisphosphatase. *J. Biol. Chem.* **283**, 33988–33993 (2008).

44. Shackelford, G. S., Regni, C. A. & Beamer, L. J. Evolutionary trace analysis of the alpha-D-phosphohexomutase superfamily. *Protein Sci.* **13**, 2130–2138 (2004).

45. Ganesan, S. M., Falla, A., Goldfless, S. J., Nasamu, A. S. & Niles, J. C. Synthetic RNA-protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. *Nat. Commun.* **7**, 10727 (2016).

46. Nasamu, A. S. *et al.* Plasmepsins IX and X are essential and druggable mediators of malaria parasite egress and invasion. *Science (80-.).* **358**, 518–522 (2017).

47. Wishart, D. S. *et al.* HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids*

Res. **46**, D608–D617 (2018).

48. Dans, M. G. *et al.* Screening the Medicines for Malaria Venture Pathogen Box for invasion and egress inhibitors of the blood stage of Plasmodium falciparum reveals several inhibitory compounds. *Int. J. Parasitol.* **50**, 235–252 (2020).

49. Lidumniece, E. *et al.* Peptidic boronic acids are potent cell-permeable inhibitors of the malaria parasite egress serine protease SUB1. *Proc. Natl. Acad. Sci. U. S. A.* **118**, DOI: 10.1073/pnas.2022696118 (2021).

50. Tan, M. S. Y. *et al.* Autocatalytic activation of a malarial egress protease is druggable and requires a protein cofactor. *EMBO J.* **40**, DOI: 10.15252/embj.2020107226 (2021).

51. Vanaerschot, M. *et al.* Inhibition of Resistance-Refractory P. falciparum Kinase PKG Delivers Prophylactic, Blood Stage, and Transmission-Blocking Antiplasmodial Activity. *Cell Chem. Biol.* **27**, 806-816.e8 (2020).

52. Favuzza, P. *et al.* Dual Plasmepsin-Targeting Antimalarial Agents Disrupt Multiple Stages of the Malaria Parasite Life Cycle. *Cell Host Microbe* **27**, 642-658.e12 (2020).

53. Wickham, M. E., Culvenor, J. G. & Cowman, A. F. Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte. *J. Biol. Chem.* **278**, 37658–37663 (2003).

54. Soldati, D., Foth, B. J. & Cowman, A. F. Molecular and functional aspects of parasite invasion. *Trends in Parasitology* **20**, 567–574 (2004).

55. Dowse, T. J., Koussis, K., Blackman, M. J. & Soldati-Favre, D. Roles of proteases during invasion and egress by plasmodium and toxoplasma. *Subcell. Biochem.* **47**, 121–139 (2008).

56. Santos de Macedo, C. *et al.* Inhibition of glycosyl-phosphatidylinositol biosynthesis in

Plasmodium falciparum by C-2 substituted mannose analogues. *Eur. J. Biochem.* **268**, 6221–6228 (2001).

57. Kurucz, R., Seeberger, P. H. & Varón Silva, D. Glycosylphosphatidylinositols in Malaria: GPI Biosynthesis and GPI-Derived Proteins. in *Encyclopedia of Malaria* (eds. Hommel, M. & Kremsner, P.) (Springer, New York, NY, 2013). doi:DOI: 10.1007/978-1-4614-8757-9_22-1

58. Gerold, P., Dieckmann-Schuppert, A. & Schwarz, R. Glycosylphosphatidylinositols synthesized by asexual erythrocytic stages of the malarial parasite, Plasmodium falciparum. Candidates for plasmodial glycosylphosphatidylinositol membrane anchor precursors and pathogenicity factors. *J Biol Chem* **269**, 2597–2606 (1994).

59. Doering, T. L., Raper, J., Buxbaum, L. U., Hart, G. W. & Englund, P. T. Biosynthesis of glycosyl phosphatidylinositol protein anchors. *Methods* **1**, 288–296 (1990).

60. Naik, R. S., Davidson, E. A. & Gowda, D. C. Developmental Stage-specific Biosynthesis of Glycosylphosphatidylinositol Anchors in Intraerythrocytic Plasmodium falciparum and Its Inhibition in a Novel Manner by Mannosamine. *J. Biol. Chem.* **275**, 24506–24511 (2000).

61. Naik, R. S., Krishnegowda, G. & Gowda, D. C. Glucosamine inhibits inositol acylation of the glycosylphosphatidylinositol anchors in intraerythrocytic Plasmodium falciparum. *J. Biol. Chem.* **278**, 2036–2042 (2003).

62. Chandramohanadas, R. *et al.* Small molecule targeting malaria Merozoite surface protein-1 (MSP-1) prevents host invasion of divergent Plasmodial species. *J. Infect. Dis.* **210**, 1616–1626 (2014).

63. Doering, T. L., Englund, P. T. & Hart, G. W. Detection of Glycophospholipid Anchors on

Proteins. *Curr. Protoc. Protein Sci.* DOI: 10.1002/0471140864.ps1205s02 (1993). doi:DOI: 10.1002/0471142727.mb1708s22

64. Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* **41**, 1149–1160 (2009).

65. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D. & Aravind, L. Evolutionary Genomics of the HAD Superfamily: Understanding the Structural Adaptations and Catalytic Diversity in a Superfamily of Phosphoesterases and Allied Enzymes. *J. Mol. Biol.* **361**, 1003–1034 (2006).

66. Frasse, P. M. & Odom John, A. R. Haloacid Dehalogenase Proteins: Novel Mediators of Metabolic Plasticity in Plasmodium falciparum. *Microbiol. Insights* (2019). doi:DOI: 10.1177/1178636119848435

67. Guggisberg, A. M. *et al.* A sugar phosphatase regulates the methylerythritol phosphate (MEP) pathway in malaria parasites. *Nat. Commun.* **5**, 4467 (2014).

68. Guggisberg, A. M. *et al.* Suppression of Drug Resistance Reveals a Genetic Mechanism of Metabolic Plasticity in Malaria Parasites. *MBio* **9**, e01193-18 (2018).

69. Dumont, L. *et al.* The metabolite repair enzyme phosphoglycolate phosphatase regulates central carbon metabolism and fosmidomycin sensitivity in plasmodium falciparum. *MBio* **10**, e02060-19 (2019).

70. Matthijs, G. *et al.* PMM (PMM1), the human homologue of SEC53 or yeast phosphomannomutase, is localized on chromosome 22q13. *Genomics* **40**, 41–47 (1997).

71. Matthijs, G. *et al.* Mutations in PMM2, a phosphomannomutase gene on chromosome 16p13, in carbohydrate-deficient glycoprotein type 1 syndrome (Jaeken syndrome). *Nat.*

Genet. **16**, 88–92 (1997).

72. Zhu, J. S. *et al.* Synthesis, Derivatization, and Structural Analysis of Phosphorylated Mono-, Di-, and Trifluorinated d -Gluco-heptuloses by Glucokinase: Tunable Phosphoglucomutase Inhibition. *ACS Omega* **4**, 7029–7037 (2019).

73. Zhu, J. S. *et al.* Inhibitory Evaluation of αPMM/PGM from Pseudomonas aeruginosa: Chemical Synthesis, Enzyme Kinetics, and Protein Crystallographic Study. *J. Org. Chem.* **84**, 9627–9636 (2019).

74. Park, J., Guggisberg, A. M., Odom, A. R. & Tolia, N. H. Cap-domain closure enables diverse substrate recognition by the C2-type haloacid dehalogenase-like sugar phosphatase Plasmodium falciparum HAD1. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1824–34 (2015).

75. Gardner, M. J. *et al.* Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419**, 498–511 (2002).

76. Reddy, K. S. *et al.* Multiprotein complex between the GPI-anchored CyRPA with PfRH5 and PfRipr is crucial for Plasmodium falciparum erythrocyte invasion. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 1179–1184 (2015).

77. Schultz, C. Prodrugs of biologically active phosphate esters. *Bioorganic Med. Chem.* **11**, 885–898 (2003).

78. Edwards, R. L. *et al.* MEPicides: Potent antimalarial prodrugs targeting isoprenoid biosynthesis. *Sci. Rep.* **7**, 8400 (2017).

79. Istvan, E. S. *et al.* Esterase mutation is a mechanism of resistance to antimalarial compounds. *Nat. Commun.* **8**, 14240 (2017).

80. Sidén-Kiamos, I. *et al.* Distinct roles for Pbs21 and Pbs25 in the in vitro ookinete to

oocyst transformation of Plasmodium berghei. *J. Cell Sci.* **113**, 3419–3426 (2000).

81. Van Dijk, M. R. *et al.* A central role for P48/45 in malaria parasite male gamete fertility. *Cell* **104**, 153–164 (2001).

82. Pradel, G., Garapaty, S. & Frevert, U. Proteoglycans mediate malaria sporozoite targeting to the liver. *Mol. Microbiol.* **45**, 637–651 (2002).

83. Lopez-Gutierrez, B., Cova, M. & Izquierdo, L. A Plasmodium falciparum C-mannosyltransferase is dispensable for parasite asexual blood stage development. *Parasitology* **146**, 1767–1772 (2019).

84. Hoppe, C. M. *et al.* Apicomplexan C-Mannosyltransferases Modify Thrombospondin Type I-containing Adhesins of the TRAP Family. *Glycobiology* **28**, 333–343 (2018).

85. Lopaticki, S. *et al.* Protein O-fucosylation in Plasmodium falciparum ensures efficient infection of mosquito and vertebrate hosts. *Nat. Commun.* **8**, 561 (2017).

86. Sanz, S. *et al.* Protein O-Fucosyltransferase 2 Is Not Essential for Plasmodium berghei Development. *Front. Cell. Infect. Microbiol.* **9**, 238 (2019).

87. Nkrumah, L. J. *et al.* Efficient site-specific integration in Plasmodium falciparum chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat. Methods* **3**, 615–21 (2006).

88. Collins, C. R. *et al.* Malaria Parasite cGMP-dependent Protein Kinase Regulates Blood Stage Merozoite Secretory Organelle Discharge and Egress. *PLoS Pathog.* **9**, e1003344 (2013).

89. Donald, R. G. K. *et al.* Toxoplasma gondii Cyclic GMP-Dependent Kinase: Chemotherapeutic Targeting of an Essential Parasite Protein Kinase. *Eukaryot. Cell* **1**, 317–328 (2002).

90. Boyle, M. J. *et al.* Isolation of viable Plasmodium falciparum merozoites to define erythrocyte invasion events and advance vaccine and drug development. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 14378–14383 (2010).

91. Alexandrov, A. *et al.* A Facile Method for High-throughput Co-expression of Protein Pairs. *Mol. Cell. Proteomics* **3**, 934–938 (2004).

92. Verma, M. *et al.* A short translational ramp determines the efficiency of protein synthesis. *Nat. Commun.* **10**, 5774 (2019).

93. Banerjee, R. *et al.* 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 (2002).

94. Das, S. *et al.* Na+ Influx Induced by New Antimalarials Causes Rapid Alterations in the Cholesterol Content and Morphology of Plasmodium falciparum. *PLoS Pathog.* **12**, e1005647 (2016).

95. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: The integration of data reduction and structure solution - From diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 859–866 (2006).

96. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).

97. Ji, T., Zhang, C., Zheng, L., Dunaway-Mariano, D. & Allen, K. N. Structural Basis of the Molecular Switch between Phosphatase and Mutase Functions of Human Phosphomannomutase 1 under Ischemic Conditions. *Biochemistry* **57**, 3480–3492 (2018).

98. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

99. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

100. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

101. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).

102. Hanwell, M. D. *et al.* Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **4**, 17 (2012).

103. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2009).

104. Lu, W., Navidpour, L. & Taylor, S. D. An expedient synthesis of benzyl 2,3,4-tri-O-benzyl-β-D- glucopyranoside and benzyl 2,3,4-tri-O-benzyl-β-D-mannopyranoside. *Carbohydr. Res.* **340**, 1213–1217 (2005).

105. Beaton, S. A., Huestis, M. P., Sadeghi-Khomami, A., Thomas, N. R. & Jakeman, D. L. Enzyme-catalyzed synthesis of isosteric phosphono-analogues of sugar nucleotides. *Chem. Commun.* 238–240 (2009). doi:10.1039/b808078j

106. Yuan, M. C. *et al.* Identification of an oxime-containing C-glucosylarene as a potential inhibitor of sodium-dependent glucose co-transporter 2. *Eur. J. Med. Chem.* **143**, 611–620 (2018).

107. Nicotra, F., Ronchetti, F. & Russo, G. Stereospecific Synthesis of the Phosphono Analogues of α- and β-D-Glucose 1-Phosphate. *J. Org. Chem.* **47**, 4459–4462 (1982).

Chapter 5: Conclusions and future directions

5.1 Summary

The human malaria parasite *Plasmodium falciparum* has a complex metabolism, with unique pathways and regulatory mechanisms that can be exploited for antimalarial therapeutic design. To further understand the metabolic networks in these parasites, we sought to characterize a subset of the known metabolic regulators, haloacid dehalogenase (HAD) proteins. We described the substrate specificity of the enzyme HAD2, explored the function of serine phosphorylation in regulating its activity, and characterized its sensitivity to inorganic phosphate. We thus developed a model explaining how loss of HAD2 confers resistance to fosmidomycin and hypersensitivity to low phosphate media. We also described two related HAD proteins, HAD4 and Lipin, demonstrating their dispensability and showing the effects they have on metabolism. HAD4 is a nucleotide phosphatase *in vitro*, and its *in vivo* metabolic effects are also linked to downstream glycolysis, possibly through the regulatory effect that nucleotides exert on glycolytic enzymes. Lipin is a phospholipid phosphatase, loss of which causes growth defects and dysregulated phospholipid metabolism. Finally, we described HAD5, an indispensable, druggable phosphomannomutase of malaria parasites. We demonstrated the essentiality of HAD5, and the effect that its knockdown has on parasite egress. We also demonstrated the druggability of HAD5 and showed that it can be specifically inhibited compared to human phosphomannomutases. Finally, we solved the crystal structure of HAD5, which may aid in future drug development against this enzyme.

5.2 Diverse functions of HAD proteins in malaria parasites

HAD enzymes catalyze a diverse assortment of biochemical reactions in cells, primarily involving phosphate transfer, which includes phosphatase, phosphotransferase, and phosphosugar mutase reactions[1]. Herein, we characterized the different reactions and substrate specificities of four HAD proteins, demonstrating the breadth of roles that HAD proteins play in parasite metabolism.

5.2.1 The hypothetical role of HAD2 in parasite metabolism

Truncating R157X mutations in *Pf*HAD2 have been found to confer resistance to fosmidomycin (FSM) at the expense of diminished growth[2]. We not only found that HAD2 is a small molecule phosphatase, but further characterized its substrate specificity. Recombinant HAD2 most actively dephosphorylates fructose phosphates and triose phosphates *in vitro*. The dephosphorylation of triose phosphates may help to explain why metabolomic data also shows elevated deoxyxylulose 5-phosphate (DOXP) in these parasites, as DOXP is a downstream product of triose phosphates. This DOXP elevation if the proximal cause of FSM resistance[2].

We further concluded that fructose 1,6-bisphosphate (FBP) may also be a natural substrate of HAD2, explaining the buildup of this metabolite in *HAD2*^{R157X} parasites as well. This also helps explain why these parasites acquired hypomorphic suppressor mutations in phosphofructokinase (PFK9), and could suggest that a buildup of FBP might be toxic to the parasites. FBP toxicity has been observed when delivered intraperitoneally or intravenously in rodents[3,4], but we do not have a model for how this might be toxic toward the highly glycolytically active parasites, and it merits further investigation. The mechanism of rodent toxicity of FBP remains unclear, but could be linked to Ca^{2+} balance or changes in pentose

phosphate pathway functionality[4,5], so these may be worth testing in parasites that have elevated FBP as well.

As its natural function in parasites, it remains unclear why it would be advantageous for HAD2 to dephosphorylate these triose phosphates or fructose phosphates. Based on our *in vitro* demonstration that HAD2 is potently inhibited by inorganic phosphate (P_i), and the observation that *HAD2*^{R157X} parasites are hypersensitized to diminished growth in low phosphate, we concluded that HAD2 may be involved in regulating phosphate homeostasis in the parasites. In this model, the triose phosphates and fructose phosphates are simply the sources of phosphate for HAD2 to maintain intracellular phosphate levels.

Finally, we investigated the possible functions of phosphorylating Ser4 and Ser9 of HAD2. We produced preliminary evidence that phosphorylation of these residues may lead to increased overall HAD2 activity, and possibly to changes in substrate specificity. We reason that the kinase(s) responsible for HAD2 phosphorylation may be sensitive to phosphate levels or to concentrations of certain phosphometabolites, which has been seen in other systems[6]. Thus, the regulation of HAD2 activity could serve as a feedback mechanism to maintain a balance of free phosphate and the phosphometabolites used as phosphate sources.

Future investigations of HAD2 should focus on clean genetic mutants, including a more substantial knockdown or knockout, as well as HAD2-overexpressing parasites. This will provide unambiguous genetic differences that can be used for metabolic validation of our model. It will also be informative to investigate our FBP toxicity hypothesis, as well as the regulation of phosphate homeostasis in parasites. We are particularly interested to know the direct cause of defective growth when HAD2 is lost. We hypothesize that reactions requiring free phosphate may be dysregulated in these parasites, and propose investigating the important glycolytic

enzyme glyceraldehyde 3-phosphate dehydrogenase (GAPDH), which uses P_i to produce 1,3-bisphosphoglycerate[7]. Another enzyme that could potentially be affected by dysregulated P_i is purine nucleoside phosphorylase (PNP), a validated target of quinine and mefloquine[8] that is involved in the purine salvage pathway and is required for robust parasite growth[9]. Based on our model of phosphate dysregulation, we hypothesize that *HAD2*^{R157X} parasites will be hypersensitized to P_i-dependent enzyme inhibitors like quinine, mefloquine, other PNP inhibitors[10,11], or GAPDH inhibitors[12–14].

5.2.2 Lessons from HAD5 and the GPI-anchor pathway for antimalarial development

Glycosyl phosphatidylinositol (GPI) anchors have long been of interest in protozoan parasites, particularly in trypanosomes, which are densely covered by GPI-anchored proteins (GPI-APs)[15,16]. Even in *Plasmodium falciparum, in silico* prediction of antimalarial drug targets has identified phosphomannose isomerase (PMI), which is directly upstream of phosphomannomutase in GPI biosynthesis, as a potentially promising antimalarial target[17]. The allure of inhibiting GPI-anchors in parasites stems from their universal expression, and likely essentiality, throughout all stages of the parasite's life cycle, as well as evidence that protein-free GPI anchors contribute to toxicity during infection[18]. Despite this, few studies[19] have intentionally sought to design or repurpose existing drugs like those that target PMI in other organisms[20–22] as antimalarials, nor has PMI in *P. falciparum* been published on at all.

With our characterization of HAD5, a phosphomannomutase in parasites that is essential for development, we bolster the case for targeting the GPI-anchor biosynthesis pathway in malaria parasites. At present, our HAD5 inhibitor D9 is not effective in cultured parasites, likely due to limited cell permeability, but we propose that simple strategies could be implemented to

optimize potency and increase permeability of D9. In parallel, repurposing known inhibitors of PMI and PMM[23] from other systems could also be tested for use in intraerythrocytic parasites as well as sexual-, mosquito-, and liver-stage parasites. With our crystal structure of HAD5 in hand, more rational design and optimization of these compounds can also be implemented. Although our 3.5Å resolution crystal structure may not be perfected for rational drug design yet, as side chain orientations are not fully resolved, this structure does lay the foundation for future structure determination, and our experience demonstrates the ease with which HAD5 can be purified and crystallized.

In addition to developing drugs to target intraerythrocytic parasites, we also hope to develop a genetic strategy for validating mosquito-stage essentiality of *P. falciparum* genes like HAD5. Using our HAD5^{KD} parasites, we plan to induce gametocytogenesis and feed gametocytes to mosquitos, which will then feed on either +aTc or -aTc nectar. At present, it remains unclear if ingested aTc will be able to penetrate the mosquito midgut epithelium to access developing parasites, so we hope to validate this as a method to inducibly test gene essentiality throughout multiple life stages. If successful, this method of inducible knockdown within mosquitos could then be applied to many other potential drug targets believed to be essential in mosquito-stage parasites, a strategy that is currently lacking in the field of malaria parasite genetic editing[24].

Finally, the story of inhibiting HAD5 adds to an age-old debate of target-based approaches versus whole-cell phenotypic screening in antimalarial discovery[25,26]. The potential disadvantages of starting with a target first, as we did with HAD5, are: there may not be any known compounds to inhibit the target of interest, those that do exist may not have good drug-like qualities, and there may be a disconnect between *in vitro* potency and *in vivo* efficacy or

essentiality. However, in the modern era of parasite genetics, with CRISPR/Cas9 and inducible knockdown or knockout systems, our ability to effectively predict *in vivo* efficacy and gene essentiality has been greatly improved. While we did encounter issues with drug-like qualities of our D9 compound, great strides have also been made in medicinal chemistry to improve permeability[27–29], making this strategy more feasible. Lastly, the critique that target-based approaches risk the absence of existing inhibitors remains valid. However, we argue that this risk can be minimized by focusing on metabolic targets. As we observed with HAD5 in collaboration with Dr. David Jakeman, inhibitors to metabolic enzymes can easily be conceived and synthesized using the enzyme's natural substrate as a backbone. This has successfully been used for other phosphomannomutases[30,31], and we demonstrated its effectiveness with HAD5. Therefore, with modern advances in genetic editing of parasites and optimization of drug-like characteristics, we argue that target-based approaches remain a valid strategy to compete with high-throughput phenotypic screening, and the GPI-anchor biosynthesis pathway is a promising candidate for antimalarial drug development.

5.2.3 Other HAD proteins of interest

We have demonstrated that PF3D7_0303200 is a phospholipid phosphatase, termed Lipin, and have shown that it is necessary for robust growth of the parasites[32]. Future studies using *in vitro* biochemical assays to validate the substrate specificity of *Pf*Lipin may be interesting, as well as uncovering how Lipin activity is regulated, as it is expected to have many phosphorylation sites[33–38]. In other systems, Lipins are dephosphorylated by the Dullard/Nem1 protein phosphatase, which activates Lipin[39,40]. Notably, Dullard proteins contain a Nuclear LIM Interactor-interacting factor (NLI-IF) domain[41], which is also present in at least three *P. falciparum* HAD genes: NIF2 (PF3D7_0515900), NIF3 (PF3D7_1355700), and NIF4

(PF3D7_0726900)[42]. Of particular interest to us is NIF4, because both *Pf*Lipin and *Pf*NIF4 have been found to be mutated in a resistance screen for the compound MMV665917 (Appendix A). Determining the target of this antimalarial compound, as well as elucidating how Lipin and NIF4 might be involved in resistance, is worth investigating further.

We also use both *in vitro* biochemical assays and cultured parasite metabolic labeling to demonstrate that HAD4, which clusters closely with previously described *P. falciparum* HAD members[42], is a nucleotide phosphatase. Although loss of HAD4 did not affect parasite viability or growth, the disruption of nucleotide metabolism may be exploitable with inhibitors of nucleotide salvage and synthesis pathways. We also find that downstream glycolytic intermediates are elevated in HAD4 knockouts. Based on trends we have seen with HAD1, HAD2, and phosphoglycolate phosphatase (PGP)[2,43,44], we hypothesize that this may also lead to elevated MEP pathway intermediates, and thus resistance to FSM. Lastly, further investigation of how HAD4 is indirectly linked to glycolysis may be useful in our greater understanding of parasite metabolic networks.

The HAD superfamily is a vast and varied assortment of proteins, and the list of *P. falciparum* HAD proteins is representative of this[42]. For that reason, future endeavors to study HAD proteins must prioritize those with higher likelihood of importance, interest, and expected function. Because the subfamily II members have thus far been among the most well characterized, we recommend continuing to investigate those and their nearest homologs. One particularly intriguing enzyme has been termed HAD3 (PF3D7_1226100), the last of the subfamily II HAD proteins in the genome. HAD3 is predicted to be both essential[45] and widely expressed across many parasite life stages[46,47]. Although we have experienced that homology cannot predict substrate specificity, we note that homologous proteins to HAD3 are involved in

flavin metabolism[48,49], which would be worth exploring with HAD3 as well. Our attempts to purify soluble, recombinant HAD3 are described in Appendix B, as are recommendations for how to move forward with optimizing its purification and using genetic tools to test its function.

5.3 Reexamining the link between HAD proteins and the MEP pathway

We have previously hypothesized that the emergence of HAD proteins as mediators of FSM resistance in parasites[2,43,44] may be evidence for one of two possible scenarios: 1) HAD proteins, or at least subfamily II HAD proteins, are somehow inherently linked to MEP pathway metabolism or 2) the MEP pathway is finely metabolically tuned, and is sensitive to slight perturbations in parasite metabolism. Our observations that HAD5 knockdown yields no changes to FSM sensitivity in parasites serve as a counterexample to both of these hypotheses. While large-scale metabolomics was not performed with HAD5^{KD} parasites, we did demonstrate HAD5's function as a metabolic enzyme, and therefore conclude that perturbation of at least one branch of metabolism – mannose and GPI metabolism – has no effect on FSM sensitivity.

Resistance to FSM was not investigated in Lipin and HAD4 knockout parasites, but these would be interesting additions to the dataset. As a lipid phosphatase, we expect that Lipin may be too far removed from the MEP pathway to confer changes to FSM resistance, similar to HAD5. Conversely, HAD4 knockout led to changes in glycolytic intermediates, particularly triose phosphates like glycerol-3-phosphate and glyceraldehyde 3-phosphate (Gly3P)[32]. While we believe this is only an indirect effect of HAD4's activity upon regulatory nucleotides, the increased triose concentrations, especially Gly3P, could easily feed into the MEP pathway and

confer FSM resistance. Interestingly, HAD4 was never found to be mutated in screens for FSM resistance, despite its amenability to knockout. This could suggest that HAD4 mutation might not, in fact, alter FSM sensitivity, or it could simply be the result of an unsaturated screen.

In light of our HAD5 findings and further investigation of HAD2, we amend our model of HAD proteins and FSM resistance. It seems the critical changes to metabolism required to alter FSM sensitivity may involve the triose phosphates. Rather than being uniquely sensitive to all metabolic perturbations, we hypothesize that the MEP pathway in parasites is particularly sensitive to disruptions in triose phosphate metabolism, as these are the precursors that feed into the pathway[50]. In this model, HAD enzymes are therefore linked to FSM resistance only insofar as they perturb triose phosphate levels, an effect that may be enriched in subfamily II HAD proteins. In the case of HAD1 and HAD2, these two highly homologous enzymes likely act directly upon triose phosphates, and are thus the most closely linked to FSM resistance. However, because of its sensitivity to inorganic phosphate inhibition and its possible role in intracellular phosphate homeostasis, the loss of HAD2 has less effect on triose phosphate levels than HAD1, explaining why it emerged less frequently than HAD1 in FSM resistance screens. As for phosphoglycolate phosphatase (PGP), knockout of this enzyme leads to lower levels of dihydoxyacetone phosphate (DHAP)[44], which enters the apicoplast to feed into the MEP pathway[50], and thus causes FSM hypersensitivity.

Because HAD enzymes are typically believed to be promiscuous in substrate utilization[51], and are an enormous superfamily of primarily phosphatase enzymes, it is reasonable that several HAD enzymes have been implicated in triose phosphate metabolism and subsequently MEP metabolism. While not all HAD enzymes alter triose phosphate metabolism, we hypothesize that disrupted triose phosphate levels – by HAD enzymes or by other means –

will reliably yield altered FSM sensitivity in organisms that harbor the MEP pathway. It may be interesting to test this hypothesis in other organisms and with other enzymes, such as with *E. coli*, whose HAD enzymes and other phosphatases have been more extensively characterized[49,52].

5.4 Final Thoughts

The urgent need for new antimalarial therapeutics and vaccines is clear, and a deeper understanding of parasite biology will lead to the discovery of new potential drug targets. Given the many prior successes of targeting metabolic pathways with antimalarials[53,54], further elucidating key metabolic enzymes in parasites is a promising avenue of drug target discovery. As established regulators of metabolism, HAD proteins represent an untapped resource for our understanding of parasite metabolism. Because of their wide variety in function, from P-type ATPases to sugar phosphatases to phosphomannomutases [42,49,51,55], as well as the incongruence between sequence homology and substrate specificity, it is difficult to predict the function of novel HAD proteins *a priori*. However, at the very least, the type of enzymatic activity (e.g., phosphatase vs. phosphomutase) appears predictable, and the mechanism of catalysis, with Mg^{2+} ions coordinated by catalytic aspartate residues, is generally conserved[1,51]. Therefore, HAD proteins represent a reservoir of metabolic regulators with relatively predictable function and ease of investigation. Some, such as HAD5, are essential and possible drug targets themselves, while dispensable HAD proteins like HAD2 or HAD4 may still play important and exploitable roles in regulating metabolic flux, carbon utilization, phosphate homeostasis, and drug susceptibility.

5.7 References

1. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D. & Aravind, L. Evolutionary Genomics of the HAD Superfamily: Understanding the Structural Adaptations and Catalytic Diversity in a Superfamily of Phosphoesterases and Allied Enzymes. *J. Mol. Biol.* **361**, 1003–1034 (2006).
2. Guggisberg, A. M. *et al.* Suppression of Drug Resistance Reveals a Genetic Mechanism of Metabolic Plasticity in Malaria Parasites. *MBio* **9**, e01193-18 (2018).
3. Nunes, F. B. *et al.* Intravenous toxicity of fructose-1,6-bisphosphate in rats. *Toxicol. Lett.* **143**, 73–81 (2003).
4. Vexler, Z. *et al.* Toxicity of fructose-1,6-bisphosphate in developing normoxic rats. *Pharmacol. Toxicol.* **84**, 115–121 (1999).
5. Shao, L.-R., Wang, G. & Stafstrom, C. E. The Glycolytic Metabolite, Fructose-1,6-bisphosphate, Blocks Epileptiform Bursts by Attenuating Voltage-Activated Calcium Currents in Hippocampal Slices. *Front. Cell. Neurosci.* **12**, 168 (2018).
6. Wösten, M. M. S. M. *et al.* The Campylobacter jejuni PhosS/PhosR operon represents a non-classical phosphate-sensitive two-component system. *Mol. Microbiol.* **62**, 278–291 (2006).
7. Daubenberger, C. A. *et al.* Identification and recombinant expression of glyceraldehyde-3-phosphate dehydrogenase of Plasmodium falciparum. *Gene* **246**, 255–264 (2000).
8. Dziekan, J. *et al.* Identifying purine nucleoside phosphorylase as the target of quinine using cellular thermal shift assay. *Sci. Transl. Med.* **11**, (2019).
9. Madrid, D. C., Ting, L.-M., Waller, K. L., Schramm, V. L. & Kim, K. Plasmodium falciparum

Purine Nucleoside Phosphorylase Is Critical for Viability of Malaria Parasites *. *J. Biol. Chem.* **283**, 35899–35907 (2008).

10. Kicska, G. A. *et al.* Purine-less Death in Plasmodium falciparumInduced by Immucillin-H, a Transition State Analogue of Purine Nucleoside Phosphorylase *. *J. Biol. Chem.* **277**, 3226–3231 (2002).
11. Kicska, G. A. *et al.* Transition State Analogue Inhibitors of Purine Nucleoside Phosphorylase from Plasmodium falciparum*. *J. Biol. Chem.* **277**, 3219–3225 (2002).
12. Liao, S. *et al.* 4-Octyl itaconate inhibits aerobic glycolysis by targeting GAPDH to exert anti-inflammatory effects. *Nat. Commun.* **10**, (2019).
13. Liberti, M. *et al.* A Predictive Model for Selective Targeting of the Warburg Effect through GAPDH Inhibition with a Natural Product. *Cell Metab.* **26**, 648-659.e8 (2017).
14. Galbiati, A., Zana, A. & Conti, P. Covalent inhibitors of GAPDH: From unspecific warheads to selective compounds. *Eur. J. Med. Chem.* **207**, (2020).
15. Yadav, U. & Khan, M. A. Targeting the GPI biosynthetic pathway. *Pathog. Glob. Health* **112**, 115 (2018).
16. Hong, Y. & Kinoshita, T. Trypanosome glycosylphosphatidylinositol biosynthesis. *Korean J. Parasitol.* **47**, 197–204 (2009).
17. Ludin, P., Woodcroft, B., Ralph, S. A. & Mäser, P. In silico prediction of antimalarial drug target candidates. *Int. J. Parasitol. Drugs Drug Resist.* **2**, 191–199 (2012).
18. Arrighi, R. & Faye, I. Plasmodium falciparum GPI toxin: a common foe for man and mosquito. *Acta Trop.* **114**, 162–165 (2010).
19. Smith, T. *et al.* Substrate specificity of the Plasmodium falciparum

glycosylphosphatidylinositol biosynthetic pathway and inhibition by species-specific suicide substrates. *Biochemistry* **41**, 12395–12406 (2002).

20. Foret, J., de Courcy, B., Gresh, N., Piquemal, J. P. & Salmon, L. Synthesis and evaluation of non-hydrolyzable d-mannose 6-phosphate surrogates reveal 6-deoxy-6-dicarboxymethyl-d-mannose as a new strong inhibitor of phosphomannose isomerases. *Bioorg. Med. Chem.* **17**, 7100–7107 (2009).

21. Mastrolorenzo, A., Scozzafava, A. & Supuran, C. T. Antifungal Activity of Ag(I) and Zn(Ii) Complexes of Aminobenzolamide (5-Sulfanilylamido-1,3,4-Thiadiazole-2-Sulfonamide) Derivatives. *http://dx.doi.org/10.3109/14756360009040707* **15**, 517–531 (2008).

22. Sharma, V. *et al.* Phosphomannose isomerase inhibitors improve N-glycosylation in selected phosphomannomutase-deficient fibroblasts. *J. Biol. Chem.* **286**, 39431–39438 (2011).

23. Liu, H.-Y., Wang, Z., Regni, C., Zou, X. & Tipton, P. A. Detailed Kinetic Studies of an Aggregating Inhibitor; Inhibition of Phosphomannomutase/Phosphoglucomutase by Disperse Blue 56 †. (2004). doi:10.1021/bi0491907

24. Kudyba, H. M., Cobb, D. W., Vega-Rodríguez, J. & Muralidharan, V. Some conditions apply: Systems for studying Plasmodium falciparum protein function. *PLoS Pathogens* **17**, (2021).

25. Chatterjee, A. K. & Yeung, B. K. S. Back to the Future: Lessons Learned in Modern Target-based and Whole-Cell Lead Optimization of Antimalarials. *Curr. Top. Med. Chem.* **12**, 473–483 (2012).

26. Hughes, J., Rees, S., Kalindjian, S. & Philpott, K. Principles of early drug discovery. *Br. J.*

Pharmacol. **162**, 1239 (2011).

27. Matsson, P., Doak, B. C., Over, B. & Kihlberg, J. Cell permeability beyond the rule of 5. *Adv. Drug Deliv. Rev.* **101**, 42–61 (2016).

28. Miller, J. J. *et al.* Structure-guided microbial targeting of antistaphylococcal prodrugs. *bioRxiv* 2020.12.15.408237 (2020). doi:DOI: 10.1101/2020.12.15.408237

29. Schultz, C. Prodrugs of biologically active phosphate esters. *Bioorganic Med. Chem.* **11**, 885–898 (2003).

30. Zhu, J. S. *et al.* Synthesis, Derivatization, and Structural Analysis of Phosphorylated Mono-, Di-, and Trifluorinated d -Gluco-heptuloses by Glucokinase: Tunable Phosphoglucomutase Inhibition. *ACS Omega* **4**, 7029–7037 (2019).

31. Zhu, J. S. *et al.* Inhibitory Evaluation of αPMM/PGM from Pseudomonas aeruginosa: Chemical Synthesis, Enzyme Kinetics, and Protein Crystallographic Study. *J. Org. Chem.* **84**, 9627–9636 (2019).

32. Cobbold, S. A. *et al.* Non-canonical metabolic pathways in the malaria parasite detected by isotope-tracing metabolomics. *Mol. Syst. Biol.* **17**, (2021).

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

34. Pease, B. *et al.* Global analysis of protein expression and phosphorylation of three stages of Plasmodium falciparum intraerythrocytic development. *J. Proteome Res.* **12**, 4028–4045 (2013).

35. Lasonder, E., Green, J., Grainger, M., Langsley, G. & Holder, A. Extensive differential

protein phosphorylation as intraerythrocytic Plasmodium falciparum schizonts develop into extracellular invasive merozoites. *Proteomics* **15**, 2716–2729 (2015).

36. Solyakov, L. *et al.* Global kinomic and phospho-proteomic analyses of the human malaria parasite Plasmodium falciparum. *Nat. Commun.* **2**, (2011).

37. Lasonder, E. *et al.* The *Plasmodium falciparum* Schizont Phosphoproteome Reveals Extensive Phosphatidylinositol and cAMP-Protein Kinase A Signaling. *J. Proteome Res.* **11**, 5323–5337 (2012).

38. Ganter, M. *et al.* Plasmodium falciparum CRK4 directs continuous rounds of DNA replication during schizogony. *Nat. Microbiol.* **2**, (2017).

39. Kim, Y. *et al.* A conserved phosphatase cascade that regulates nuclear membrane biogenesis. *Proc. Natl. Acad. Sci.* **104**, 6596–6601 (2007).

40. Wu, R., Garland, M., Dunaway-Mariano, D. & Allen, K. N. Homo sapiens Dullard Protein Phosphatase Shows a Preference for the Insulin-Dependent Phosphorylation Site of Lipin1. *Biochemistry* **50**, 3045–3047 (2011).

41. Satow, R., Chan, T. & Asashima, M. Molecular cloning and characterization of dullard: a novel gene required for neural development. *Biochem. Biophys. Res. Commun.* **295**, 85–91 (2002).

42. Frasse, P. M. & Odom John, A. R. Haloacid Dehalogenase Proteins: Novel Mediators of Metabolic Plasticity in Plasmodium falciparum. *Microbiol. Insights* (2019). doi:DOI: 10.1177/1178636119848435

43. Guggisberg, A. M. *et al.* A sugar phosphatase regulates the methylerythritol phosphate (MEP) pathway in malaria parasites. *Nat. Commun.* **5**, 4467 (2014).

44. Dumont, L. *et al.* The metabolite repair enzyme phosphoglycolate phosphatase regulates central carbon metabolism and fosmidomycin sensitivity in plasmodium falciparum. *MBio* **10**, e02060-19 (2019).

45. Zhang, M. *et al.* Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360**, DOI: 10.1126/science.aap7847 (2018).

46. López-Barragán, M. J. *et al.* Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum. *BMC Genomics* **12**, (2011).

47. Gómez-Díaz, E. *et al.* Epigenetic regulation of Plasmodium falciparum clonally variant gene expression during development in Anopheles gambiae. *Sci. Rep.* **7**, (2017).

48. Haase, I. *et al.* Enzymes from the haloacid dehalogenase (HAD) superfamily catalyse the elusive dephosphorylation step of riboflavin biosynthesis. *ChemBioChem* **14**, 2272–2275 (2013).

49. Kuznetsova, E. *et al.* Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family. *J. Biol. Chem.* **281**, 36149–61 (2006).

50. Guggisberg, A. M., Amthor, R. E. & Odom, A. R. Isoprenoid Biosynthesis in Plasmodium falciparum. *Eukaryot. Cell* **13**, 1348–1359 (2014).

51. Huang, H. *et al.* Panoramic view of a superfamily of phosphatases through substrate profiling. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E1974-83 (2015).

52. Wang, T. *et al.* Genome-wide screening identifies promiscuous phosphatases impairing terpenoid biosynthesis in Escherichia coli. *Appl. Microbiol. Biotechnol.* **102**, 9771–9780 (2018).

53. Ross, L. S. & Fidock, D. A. Elucidating Mechanisms of Drug-Resistant Plasmodium falciparum. *Cell Host and Microbe* **26**, 35–47 (2019).

54. Cowell, A. N. *et al.* Mapping the malaria parasite druggable genome by using in vitro evolution and chemogenomics. *Science (80-.).* **359**, 191–199 (2018).

55. Kuznetsova, E. *et al.* Functional Diversity of Haloacid Dehalogenase Superfamily Phosphatases from Saccharomyces cerevisiae: BIOCHEMICAL, STRUCTURAL, AND EVOLUTIONARY INSIGHTS. *J. Biol. Chem.* **290**, 18678–98 (2015).

Appendix A: Searching for the target of MMV665917 in *Plasmodium falciparum*

PREFACE

This work was conducted by Dana Hodge and myself, under the guidance of Dr. Audrey Odom John and in collaboration with Dr. Chris Huston and his lab. DH performed the resistance screening and quantified parasite resistance, and I entered the project to help analyze whole genome sequencing and validate hits through Sanger sequencing and recombinant enzyme assays. Because I was not involved in the resistance screening and quantification, the methods from that part of the project are omitted here. The presented data, as well as commentary on other genes of interest, was handed off to Dr. Huston's lab to pursue further with a targeted approach in *Cryptosporidium.*

A.1 INTRODUCTION

One valuable toolbox at researchers' disposal for the development of new antimalarial therapeutics is the Malaria Box, a selection of ~400 compounds known to exhibit antimalarial activity that were collected by the Medicines for Malaria Venture (MMV)[1]. To further optimize these compounds, many researchers have sought to uncover these compounds' targets so that future medical chemists can optimize the target specificity and potency. One frequently implemented strategy to identify these targets is through drug resistance screens. In this method, parasites are cultured in low concentrations of the compound, gradually increasing the concentration in a stepwise fashion over long periods of time—several weeks in the case of *Plasmodium falciparum*. This process pressures parasites to evolve resistance, and the resultant growing parasites' whole genomes are sequenced to identify any single nucleotide polymorphisms (SNPs) or other genetic aberrations that could feasibly be causing the resistance

phenotype. This was the method our lab implemented to evolve resistance to fosmidomycin, uncovering resistant parasites with hypomorphic SNPs in HAD1 and HAD2[2,3].

Malaria Box compounds often not only exhibit antimalarial activity, but are also active against other parasites. The Malaria Box compound MMV665917 has shown promise as an anti-cryptosporidial agent, with *in vitro* and *in vivo* efficacy against the apicomplexan parasites *Cryptosporidium parvum* and *Cryptosporidium hominis*[4–7]. To further optimize this compound, which is a derivative of piperazine, Dr. Christopher Huston's lab at the University of Vermont was interested in determining the *Cryptosporidium* target of MMV665917. However, culturing methods of *Cryptosporidium spp.* are in their nascent stages[8–10], and time-intensive techniques like a long resistance screen was not a practical method of identifying the target of MMV665917 in those parasites. Therefore, Dr. Huston's lab reached out to us to evolve resistance to the compound in *P. falciparum*. We reasoned that, because both *Plasmodium* and *Cryptosporidum* share an apicomplexan parasite lineage, identifying the target in the former would strongly suggest the target in the latter. With that in mind, Dana Hodge and I set about trying to identify the *P. falciparum* target of MMV665917.

A.2 RESULTS

A.2.1 Evolving resistance to MMV665917 in *P. falciparum*

Wild-type *P. falciparum* parasites ("MR4E", 3D7 background) exhibit a roughly 4.9 μM half-maximal effective concentration (EC_{50}) to MMV665917. To evolve resistance, parasites were grown in three different concentrations higher than the EC_{50} : 10 μM (2x EC_{50}), 20 μM (4x EC_{50}) and 50 μM (10x EC_{50}). However, these concentrations appeared to be too high, and no parasites were capable of definitive growth.

We next set up a step-wise selection of parasites, working up to 5 μM, which ultimately led to successful continuous growth of parasites in 5 μM MMV665917. The EC_{50}'s were only modestly increased however, and clean curves were difficult to obtain due to the insolubility of MMC665917 at high concentrations. To demonstrate resistance through another method, we chose a single concentration of the compound, 25 μM, and compared growth of parasites as a percentage of maximum (vehicle control) growth. Six parasite strains were obtained that demonstrated resistance compared to the MR4 control (Figure 1). Each of the resistant strains were cloned to obtain clonal populations.

A.2.2 Sequencing of MMV665917-resistant clones

Two clones from each resistant strain were subjected to whole genome sequencing (WGS). Across all clones, 2693 SNPs were identified from the WGS, which were sorted and prioritized based on the following criteria on the basis of likelihood to be real and likelihood to be causative: presence in both clones, not in known highly-variable genes such as *var* genes, and not found in our previous screens. Several genes were eliminated from these criteria, and the genes that remained were: TFIIH complex member XPB (PF3D7_1037600), the HAD superfamily member Lipin (PF3D7_0303200), the HAD superfamily member NLI-interacting phosphatase (PF3D7_1012700), HAD5 (PF3D7_1017400), aldo-keto reductase (PF3D7_1409100), serine/threonine protein kinase (PF3D7_0214600), structural maintenance of chromosomes protein 1 (PF3D7_1130700), histone deacetylase 2 (PF3D7_1008000), PRP2 (PF3D7_1231600), deoxyribodipyrimidine photo-lyase (PF3D7_0513600), and pyruvate kinase (PF3D7_0626800) (Table 2). Some of these were further ruled out because although the specific SNPs had not been observed before, other SNPs within those genes had been seen in other screens.

Due to our interest in HAD proteins, we were especially interested in the three HAD proteins that emerged from these screens, particularly HAD5, as we had already begun independently investigating HAD5 when we received these WGS results. However, subsequent Sanger sequencing of the HAD5 locus revealed that this was not a real SNP, and the connection to HAD5 was lost. Nonetheless, the other two HAD proteins remained of interest. Furthermore, while the mutation in pyruvate kinase (PK) was only found in the clones of one resistant strain, it was in strain R5.9, which had the highest resistance. Taken together with the high likelihood of it being an impactful mutation (Ile408Asn), Sanger sequencing confirmation that this SNP was real, and the pivotal role that PK plays in central carbon metabolism, we were interested in further pursuing PK as an important gene to investigate.

A.2.3 MMV665917 does not target RBC glycolysis

We first aimed to investigate the possibility that MMV665917 might be targeting (either primarily or via off-target effects) a host enzyme within the RBC. In particular, if pyruvate kinase is the parasite target of the compound, then it could feasibly also inhibit the host pyruvate kinase as well. To evaluate whether RBC pyruvate kinase is inhibited by MMC665917, we quantified methemoglobin (MetHb) in MMV665917-treated, uninfected RBCs (uRBCs). MetHb is hemoglobin in which the heme-bound iron is in its oxidized (Fe^{3+}) form, rather than reduced (Fe^{2+}). Under normal conditions, when the cell can maintain proper redox metabolism, MetHb is quickly reduced to the Fe^{2+} form, which permits the binding of oxygen and therefore proper functioning. This reductive capacity is in part fueled by glutathione metabolism and NADH as an electron transporter[11,12]. Previous work from our lab demonstrated that inhibiting downstream components of RBC glycolysis results in a loss of reductive capacity, and an increase in relative MetHb levels[13]. Our previous work demonstrated this effect when subjecting RBCs to the

compound POMHEX, an inhibitor of RBC enolase, which is directly upstream of pyruvate kinase in glycolysis. We reasoned that a similar effect would be observed if RBC pyruvate kinase was inhibited, and tested the ability for MMV665917 to induce a buildup of MetHb in uRBCs. We found that MMV665917 treatment does not result in increased MetHb signal, even at concentrations much higher than that of POMHEX (Figure 2). This demonstrates that redox metabolism of the RBC is not disrupted by MMV665917, and suggests that RBC glycolytic enzymes like pyruvate kinase are not targeted by the drug.

A.2.4 Testing parasite Pyruvate Kinase as the target of MMV665917

To test the hypothesis that MMV665917 inhibits *Pf*PK directly, we expressed and purified recombinant wild-type *Pf*PK from *E. coli*, as well as the I408N mutant (Figure 3). Using a KinaseGlo kit that detects the production of ATP when PK is incubated with ADP and phosphoenolpyruvate (PEP), we found that the I408N mutation did not have any observable effect on PK catalysis of the ADP+PEP to ATP+pyruvate reaction (Figure 4). Moreover, using concentrations up to 1200 µM, there was little detectable inhibition of either WT or I408N PK activity by MMV665917 (Figure 4). While statistically significant differences are observed between DMSO treatment and several of the treated conditions, the effect was not dose-dependent, and we hypothesize that more experimental replicates would erase those differences. These data demonstrate that *Pf*PK is not the direct target of MMV665917, but raise the possibility that the I408N mutation somehow helps *Pf*PK compensate for other mutations in the genome, similar to the suppressor mutations that emerge in *PFK9* after HAD2 is mutated. Notably, Ile408 is located near the interfaces between monomers within the homotetrameric structure of *Pf*PK (Figure 5), suggesting that the mutation might affect oligomerization states. However, size exclusion chromatography of both the WT and I408N mutant reveal that both

proteins eluted after roughly the same volume, 66-70 mL, which corresponds to the expected size of the tetramer in both cases (Figure 3).

A.3 Discussion

The cellular target of MMV665917 in *P. falciparum*, and in *Cryptosporidium spp.*, remains unresolved. While we demonstrate here that pyruvate kinase activity is not inhibited by MMV665917, it is unclear what might have driven the I408N mutation to arise. This point mutation, and indeed any mutations to *Pf*PK, have not been observed in any other resistance screens, suggesting that some form of selective pressure drove it to arise in the R5.9 parasites. This mutation does not appear to affect PK activity, nor does it alter oligomerization state of the enzyme.

There do remain some unexplored metrics of PK function that could be assayed to identify the function of the I408N mutation. *Pf*PK is unlike many PKs from other species, as it is not activated by fructose 1,6-bisphosphate nor fructose 2,6-bisphosphate. However, *Pf*PK is activated by glucose 6-phosphate (G6P) and inactivated by pyridoxal-5'-phosphate (PLP), ATP, oxalate, and citrate, so perhaps the I408N mutation sensitizes or ablates regulation by those means[14,15]. Notably, this residue is directly adjacent to the identified "effector site" of *Pf*PK where G6P potentially binds (Fig. 5)[15], suggesting that *Pf*PK^{I408N} may lose its allosteric regulation. Furthermore, some glycolytic enzymes are known to serve multiple "moonlighting" functions apart from their glycolytic activity. This is particularly notable in glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which has been implicated in membrane fusion, endocytosis, iron transport, and RNA binding[16–20]. As for pyruvate kinase, are documented cases of its ability to moonlight as a protein kinase of histones[21] and as a mediator of plant

mitochondria-chloroplast colocalization[22]. Perhaps *Pf*PK shares some of these moonlighting roles (the *P. falciparum* mitochondrion and apicoplast are also closely associated[23]), or harbors another as-yet undescribed function, and the I408N mutation affects a non-glycolytic role of the enzyme. More rigorous phenotypic characterization of these parasites, as well as further assessment of allosteric regulation with *in vitro* assays, will help illuminate the function of the I408N mutation.

Apart from pyruvate kinase, several other genes of interest arose from our WGS that have yet to be fully examined. Of particularly interest to us are the two HAD proteins, Lipin and NLI-interacting protein. We have previously seen that knockout of Lipin causes dramatic growth defects in parasites and accumulation of various lipid species like ceramides and phosphatidic acid[24]. Furthermore, NLI-interacting protein shares homology to the mammalian Dullard protein, a known phosphatase of Lipin that activates Lipin activity by dephosphorylating it. That two genes within the same pathway (one being a possible activator of the other) would be mutated in this screen at random is unlikely, and indicates that the function of Lipin as a lipid phosphatase is linked to the mechanism of MMV665917 inhibition.

While it remains unclear if or how pyruvate kinase mutation is also linked to Lipin/NLI-interacting protein, a tempting parallel can be drawn between this mutation pattern and that of HAD2 and PFK9: a HAD protein is mutated to confer resistance to a given inhibitor, causing a growth defect. Subsequently, mutation to a glycolytic enzyme compensates for that growth defect. In this case, we know that loss of Lipin diminishes growth, so if the Lipin mutation (Asn337Lys) is hypomorphic, that may have been the selective pressure driving PK mutation.

A.4 Methods

A.4.1 Cloning of Recombinant Enzyme Plasmids

The pET15-MHL plasmid containing wild-type *Pf*PK was obtained from Addgene (Catalog # 25278), which was deposited by Wernimont et al. after its use in the solving of the PfPyrK crystal structure (PDB: 3KHD). To prepare for robust protein expression, this PyrK sequence was cloned out of the pET15-MHL vector using primers PMF115 and PMF117, and subsequently with PMF116+PMF117, which add a ligation-independent cloning (LIC) extension onto each end (Table 1). This sequence was then inserted into our pKFS plasmid, a derivative of BG1861[25] that has a KFS sequence at the N-terminus to promote expression[26].

To clone the I408N mutant, a TAT sequence encoding Ile was replaced with an ATT sequence that encodes Asn in the PMF116/PMF117 amplicon. This was achieved by amplifying the sequence in two segments, with primers PMF116+PMF119 and PMF117+PMF118 (Table 1), representing the N-terminal half and the C-terminal half of the sequence, respectively. PMF118 and PMF119 primers are long enough to anneal with high fidelity to most of the sequence, but introduce the mutated codon in the middle. These two amplicons were then amplified together with PMF116+PMF117 to obtain a final mutated coding sequence, which was LIC-cloned into the pKFS vector as well.

A.4.2 Purification of Recombinant Enzymes

pKFS plasmids containing either WT or I408N variants of *Pf*PK were transformed into chemically competent BL21(DE3) pLysS *E. coli* cells (ThermoFisher). These cells were grown in 4 L of LB media at 37˚C, shaking, until mid-log phase, then induced with 1mM IPTG at 37˚C for 2 hours. Cells were spun down, resuspended in Lysis Buffer (10 mM Tris HCl (pH 7.5), 200mM NaCl, 1 mM $MgCl_2$, 4 µL/L benzonase, 20 mM imidazole, 1 mM DTT, 1 mg/mL

lysozyme, and cOmplete EDTA-free protease inhibitor cocktail), and sonicated. Lysates were spun down, and soluble supernatants were purified by affinity to nickel agarose beads (Gold Biotechnology). Nickel-eluted protein was loaded onto a HiLoad 16/60 Superdex 200 column (GE HealthSciences), run with protein dilution buffer (25 mM Tris pH 7.5, 250 mM NaCl, 1 mM $MgCl_2$), and fractions containing protein were pooled, concentrated, and flash-frozen for storage.

A.4.3 Pyruvate Kinase enzyme assays

Pyruvate kinase activity was assayed using the KinaseGlo Luminescent Kinase Assay (Promega), following manufacturers protocols. Briefly, 100µl assays were performed, containing 50 µL KinaseGlo Reagent, 4 µL of serial dilutions of MMV665917, 5 µL of 10 ng/µL pyruvate kinase, and 500 µM each of ADP and PEP, with protein dilution buffer filling the balance to 100µL. Assay was allowed to proceed for at least 5 minutes, measuring the production of luminescence, and was performed in duplicate.

A.4.4 Detection of MetHemoglobin

To assess whether MMV665917 was inhibiting the RBC glycolytic enzymes, such as RBC pyruvate kinase, we spectrophotometrically measured the production of MetHb in RBCs treated with the drug, as has been described previously[13]. 25 µl of serial drug dilutions was added to 75 µL of 3% hematocrit human blood in RPMI media, both with MMV665917 diluted in DMSO, as well as POMHEX diluted in DMSO, which is an inhibitor of RBC enolase[13]. This was allowed to incubate 72 hours, after which the production of MetHb was detected by Absorbance at 630 nm, and compared to oxyhemoglobin absorption at 575 nm.

A.5 Figures

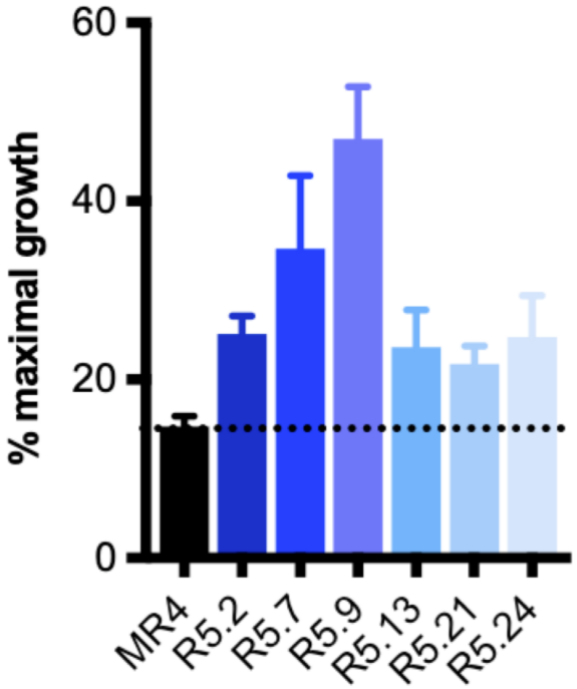


Figure 1. Relative growth inhibition in MMV-resistant parasites.

Parasite strains that emerged from stepwise resistance screens up to 5 µM were assessed for % maximal growth when grown at 25 µM MMV665917. Dashed line represents % growth exhibited by WT 3D7 parasites (MR4).

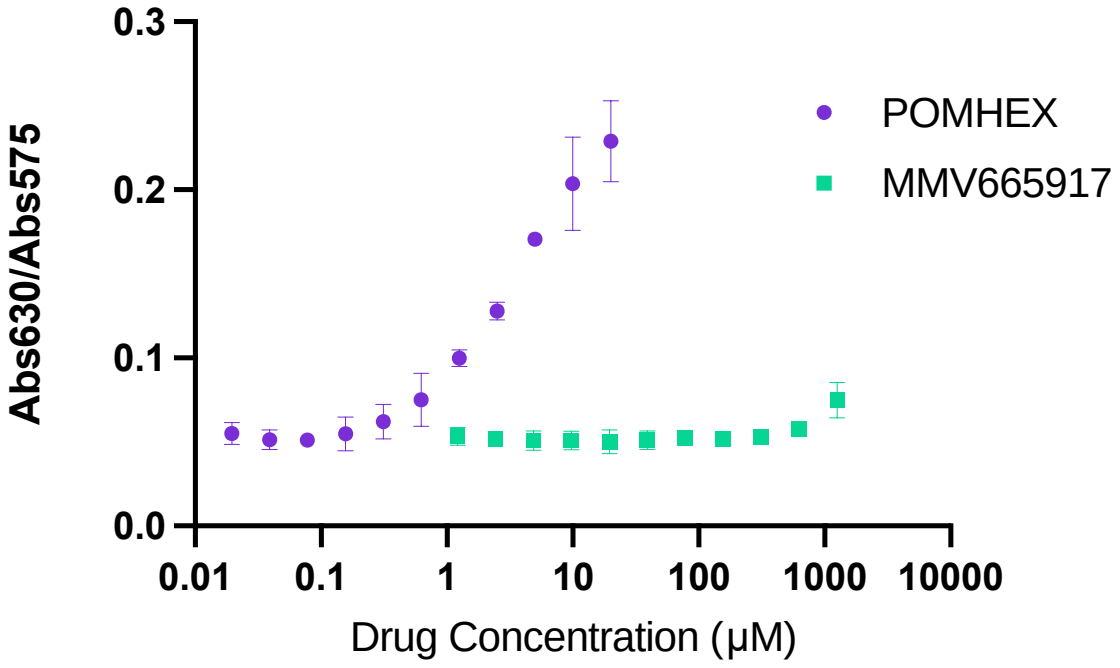


Figure 2. MetHemoglobin production in RBCs treated with MMV665917.

Ratio of MetHemoglobin (Abs_{630}) to Oxyhemoglobin (Abs_{575}) when uRBCs are treated with MMV665917. POMHEX is an inhibitor of RBC enolase, and serves as a positive control of MetHemoglobin formation. Depicted are the means ± SEM of three independent experiments with technical replicates.

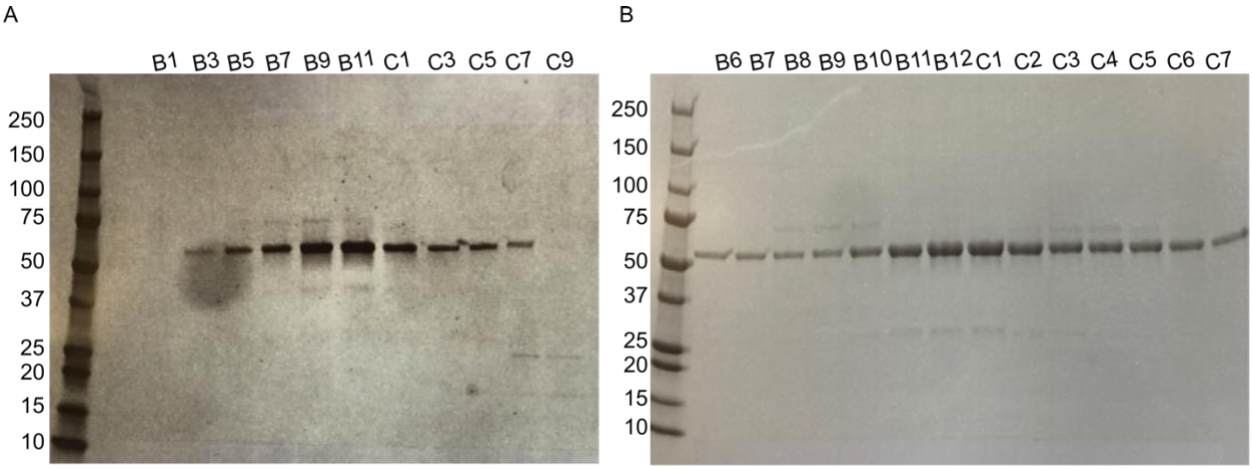


Figure 3. Size Exclusion Chromatography gels of recombinant PK.

(A) Wild-type and **(B)** I408N *Pf*PK were expressed and purified from *E. coli*, nickel purified, and fractionated after size exclusion chromatography. Shown are Coomassie-stained gels of the size exclusion fractions, demonstrating purity of the recombinant protein. Lanes are labeled with fraction number.

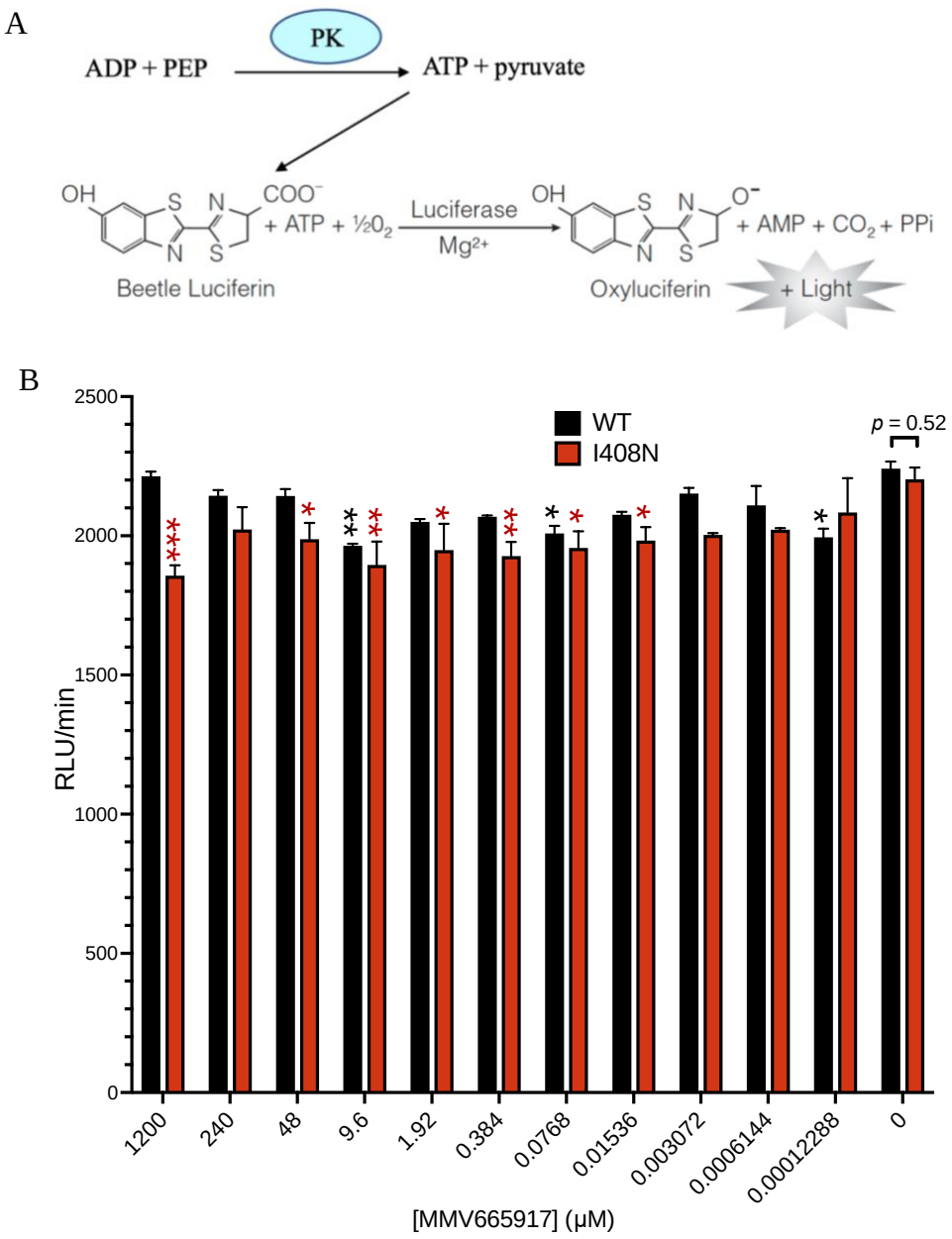


Figure 4. Activity of PK.

(A) Diagram of the pyruvate kinase assay, which generates ATP, to be used with a luciferin-based KinaseGlo kit**. (B)** Activity of wild-type and I408N PK in the pyruvate kinase assay when MMV665917 is added in increasing dosage. Data represent mean ± SEM of two replicates. Statistics were performed by 2-way ANOVA with Dunnett's test for multiple comparisons. For a given enzyme, each MMV665917 concentration was compared to the 0 µM control. *p<0.05,

$^{**}p<0.01$,$^{***}p<0.001$. Columns without significance indicated are not significant. Statistics within the 0 µM control between enzymes were performed using a two-tailed unpaired t-test.

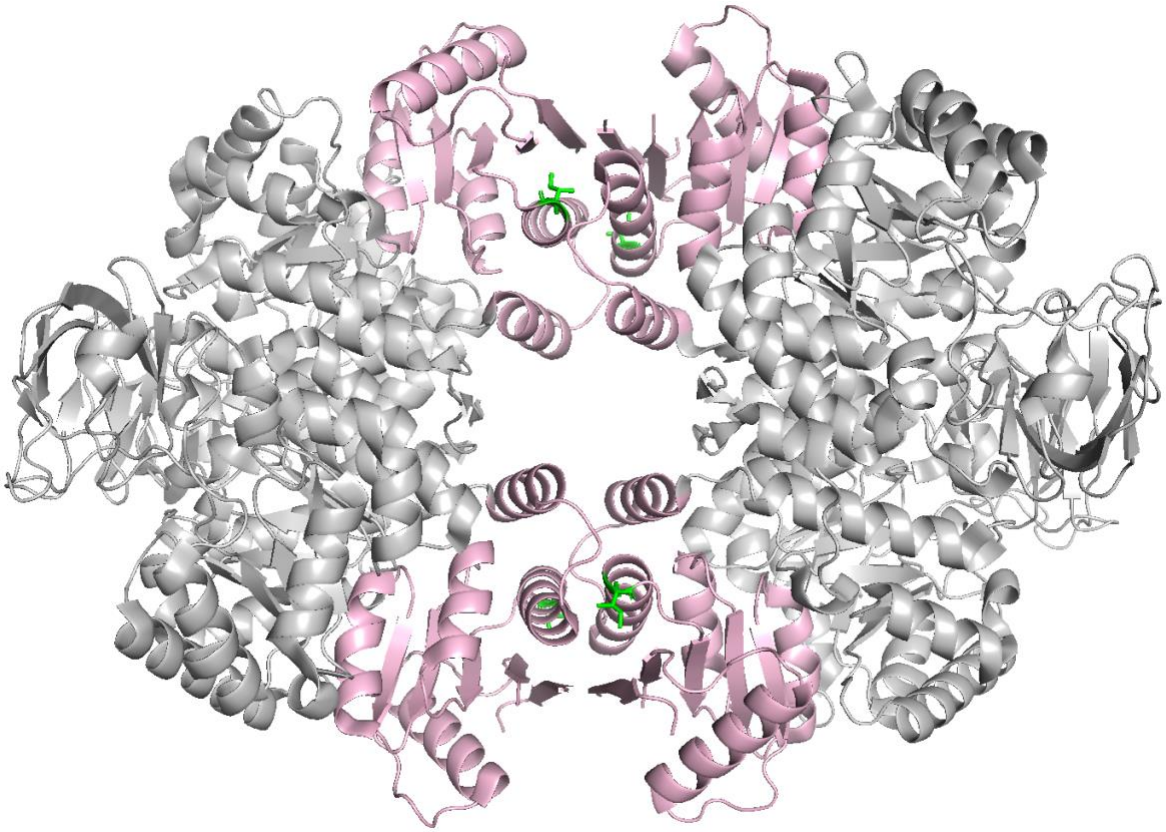


Figure 5. *Pf*PK structure (PDB: 6KSH).

Highlighted in green is Ile408, demonstrating its position near the interfaces between chains of the homotetramer and location within "Effector Site"-containing C-domain (pink). Structure and C-domain identification from Zhong et al., (2020)[15].

A.6 Tables

Table 1. Primers used for cloning

NAME	SEQUENCE
PMF115	5’ CATCCTTCAAGTACAAGAATTCGGCGGCCGGCGCAAGTAT 3’
PMF116	5’ CTCACCACCACCACCACCATTCATCCTTCAAGTAC 3’
PMF117	5’ ATCCTATCTTACTCACTCACTCAATTTGTACCACC 3’
PMF118	5’ TGCCGTAGAAACCGCAGAGTCAAATCAGGCGAGCC 3’
PMF119	5’ AGGGCAATGATCAGGCTCGCCTGATTTGACTCTGC 3’

Table 2. SNP list of specific genes of interest.

Single Nucleotide Polymorphisms, and their expected amino acid changes, are given for specific genes of interest out of the 2693 identified SNPs across all strains. Numbers indicate the number of WGS reads at that locus that found the given SNP.

Gene	Gene Name	cDNA change	Protein Change	Clone Name											
				2a	2b	7a	7b	9a	9b	13a	13b	21 a	21 b	24 a	24 b
PF3D7_ 0303200	HAD superfamily protein, putative	1011T>A	N337K	64	87	32		98				22	73	40	32
PF3D7_ 0303200	HAD superfamily protein, putative	4T>G	W2G					26	45						
PF3D7_ 0303200	HAD superfamily protein, putative	3-3C>T		20											
PF3D7_ 0303200	HAD superfamily protein, putative	3387T>A	D1129E					6							
PF3D7_ 1012700	NLI interacting factor-like phosphatase, putative	3136A>G	R1046G	43	41										

PF3D7_1012700	NLI interacting factor-like phosphatase, putative	2650G>T	E884*						6						
PF3D7_1037600	TFIIH basal transcription factor complex helicase XPB subunit, putative	248C>A	T83K							43	10		31	17	12
PF3D7_1037600	TFIIH basal transcription factor complex helicase XPB subunit, putative	235T>A	L79I			11	18								
PF3D7_1037600	TFIIH basal transcription factor complex helicase XPB subunit, putative	703T>C	Y235H								5				
PF3D7_1037600	TFIIH basal transcription factor complex helicase XPB subunit, putative	2617G>T	D873Y									6			
PF3D7_0214600	serine/threonine protein kinase, putative	3153C>A	N1051K		31		11	24	23	33			15		
PF3D7_0214600	serine/threonine protein kinase, putative	391G>T	V131L		6										
PF3D7_0214600	serine/threonine protein kinase, putative	548C>A	P183H			7									
PF3D7_0214600	serine/threonine protein kinase, putative	3150T>A	N1050K		38										
PF3D7_1409100	aldo-keto reductase, putative	423T>A	N141K					76	47						
PF3D7_1409100	aldo-keto reductase, putative	435T>A	N145K					87							
PF3D7_1409100	aldo-keto reductase, putative	3589T>C	Y1197H				20								
PF3D7_0626800	pyruvate kinase	1223T>A	I408N					123	119						
PF3D7_1017400	phosphomannomutase, putative	175+5C>A						11		25	5				

PF3D7_1017400	phosphomannomutase, putative	586G>T	E196*		7										
PF3D7_1130700	structural maintenance of chromosomes protein 1, putative	3053T>G	V1018G							171	54	30	98	68	43
PF3D7_1130700	structural maintenance of chromosomes protein 1, putative	4741G>T	D1581Y											5	
PF3D7_1130700	structural maintenance of chromosomes protein 1, putative	2218G>T	D740Y											5	
PF3D7_1130700	structural maintenance of chromosomes protein 1, putative	2278G>T	E760*						7						
PF3D7_1130700	structural maintenance of chromosomes protein 1, putative	833C>A	S278*									7			
PF3D7_1008000	histone deacetylase 2	2363C>A	T788K											7	
PF3D7_1008000	histone deacetylase 2	601C>A	Q201K	8	13	6									
PF3D7_1008000	histone deacetylase 2	1178C>A	S393*						8						
PF3D7_1008000	histone deacetylase 2	941C>T	S314F			5									
PF3D7_1231600	pre-mRNA-splicing factor ATP-dependent RNA helicase PRP2, putative	1115T>C	V372A			23	32								
PF3D7_1231600	pre-mRNA-splicing factor ATP-dependent RNA helicase PRP2, putative	1042T>A	Y348N			16									
PF3D7_0513600	deoxyribodipyrimidine photo-lyase, putative	643C>A	Q215K											16	
PF3D7_0513600	deoxyribodipyrimidine photo-lyase, putative	633T>A	N211K					39	28	53				15	

PF3D7_0513600	deoxyribodipyrimidine photo-lyase, putative	657T>A	N219K						31	51			34	15	
PF3D7_0513600	deoxyribodipyrimidine photo-lyase, putative	647C>A	T216K						32				36	17	
PF3D7_0513600	deoxyribodipyrimidine photo-lyase, putative	638T>A	M213K	34						53				16	
PF3D7_0513600	deoxyribodipyrimidine photo-lyase, putative	627T>A	D209E	32					29	51					
PF3D7_0513600	deoxyribodipyrimidine photo-lyase, putative	3341G>T	*1114L			6									

A.7 References

1. Spangenberg, T. *et al.* The Open Access Malaria Box: A Drug Discovery Catalyst for Neglected Diseases. *PLoS One* **8**, (2013).
2. Guggisberg, A. M. *et al.* A sugar phosphatase regulates the methylerythritol phosphate (MEP) pathway in malaria parasites. *Nat. Commun.* **5**, 4467 (2014).
3. Guggisberg, A. M. *et al.* Suppression of Drug Resistance Reveals a Genetic Mechanism of Metabolic Plasticity in Malaria Parasites. *MBio* **9**, e01193-18 (2018).
4. Jumani, R. S. *et al.* A novel piperazine-based drug lead for cryptosporidiosis from the medicines for malaria venture open-access malaria box. *Antimicrob. Agents Chemother.* **62**, (2018).
5. Chavez, M. A. & White, A. C. Novel treatment strategies and drugs in development for cryptosporidiosis. *Expert Review of Anti-Infective Therapy* **16**, 655–661 (2018).
6. Stebbins, E. *et al.* Clinical and microbiologic efficacy of the piperazine-based drug lead MMV665917 in the dairy calf cryptosporidiosis model. *PLoS Negl. Trop. Dis.* **12**, (2018).
7. Lee, S. *et al.* Piperazine-Derivative MMV665917: An Effective Drug in the Diarrheic Piglet Model of Cryptosporidium hominis. *J. Infect. Dis.* **220**, 285–293 (2019).
8. Karanis, P. The truth about in vitro culture of Cryptosporidium species. *Parasitology* **145**, 855–864 (2018).
9. Jossé, L., Bones, A. J., Purton, T., Michaelis, M. & Tsaousis, A. D. A Cell Culture Platform for the Cultivation of Cryptosporidium parvum. *Curr. Protoc. Microbiol.* **53**, (2019).
10. Wilke, G. *et al.* In Vitro Culture of Cryptosporidium parvum Using Stem Cell-Derived Intestinal Epithelial Monolayers. in *Methods in Molecular Biology* **2052**, 351–372

(Humana Press Inc., 2020).

11. Wright, R. O., Magnani, B., Shannon, M. W. & Woolf, A. D. N-acetylcysteine reduces methemoglobin in vitro. *Ann. Emerg. Med.* **28**, 499–503 (1996).

12. Wright, R. O., Lewander, W. J. & Woolf, A. D. Methemoglobinemia: Etiology, pharmacology, and clinical management. *Annals of Emergency Medicine* **34**, 646–656 (1999).

13. Jezewski, A. J. *et al.* [preprint] Targeting host glycolysis as a strategy for antimalarial development. SHORT TITLE: Host glycolysis as an antimalarial target. *BioRxiv* (2020). doi:10.1101/2020.10.09.331728

14. Chan, M. & Sim, T. S. Functional analysis, overexpression, and kinetic characterization of pyruvate kinase from Plasmodium falciparum. *Biochem. Biophys. Res. Commun.* **326**, (2004).

15. Zhong, W. *et al.* Pyruvate kinase from Plasmodium falciparum: Structural and kinetic insights into the allosteric mechanism. *Biochem. Biophys. Res. Commun.* **532**, (2020).

16. Sirover, M. A. Pleiotropic effects of moonlighting glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in cancer progression, invasiveness, and metastases. *Cancer and Metastasis Reviews* **37**, 665–676 (2018).

17. Zaffagnini, M., Fermani, S., Costa, A., Lemaire, S. D. & Trost, P. Plant cytoplasmic GAPDH: Redox post-translational modifications and moonlighting properties. *Frontiers in Plant Science* **4**, (2013).

18. Garcin, E. D. GAPDH as a model non-canonical AU-rich RNA binding protein. *Seminars in Cell and Developmental Biology* **86**, 162–173 (2019).

19. Boradia, V. M., Raje, M. & Raje, C. I. Protein moonlighting in iron metabolism:

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH). *Biochem. Soc. Trans.* **42**, 1796–1801 (2014).

20. Henderson, B. & Martin, A. C. R. Protein moonlighting: a new factor in biology and medicine. *Biochem. Soc. Trans.* **42**, 1671–1678 (2014).
21. Yang, W. *et al.* PKM2 phosphorylates histone H3 and promotes gene transcription and tumorigenesis. *Cell* **150**, 685–696 (2012).
22. Zhang, Y. *et al.* A moonlighting role for enzymes of glycolysis in the co-localization of mitochondria and chloroplasts. *Nat. Commun.* **11**, (2020).
23. Van Dooren, G. G. *et al.* Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of Plasmodium falciparum. *Mol. Microbiol.* **57**, (2005).
24. Cobbold, S. A. *et al.* Non-canonical metabolic pathways in the malaria parasite detected by isotope-tracing metabolomics. *Mol. Syst. Biol.* **17**, (2021).
25. Alexandrov, A. *et al.* A facile method for high-throughput co-expression of protein pairs. *Mol. Cell. Proteomics* **3**, 934–938 (2004).
26. Verma, M. *et al.* A short translational ramp determines the efficiency of protein synthesis. *Nat. Commun.* **10**, 5774 (2019).

Appendix B: HAD3 from *Plasmodium spp.* is highly insoluble as a recombinant enzyme

B.1 INTRODUCTION

Between our findings with HAD1, HAD2, HAD4, HAD5, and Lipin[1–3] as well as Dumont et al.'s published work on phosphoglycolate phosphatase[4], there is a growing appreciation for the regulatory role that HAD proteins serve in parasite metabolism and drug resistance. The HAD proteins that have been described thus far generally cluster together, and one phylogenetically related gene that remains unstudied is HAD3 (PF3D7_1226100) (Chapter 1, Figure 1).

In addition to its phylogenetic similarity to other HADs of interest, HAD3 is intriguing because published transcriptomics demonstrate its expression in asexually replicating parasites, as well as even higher expression in gametocytes, oocysts and sporozoites[5,6]. As these stages of the parasite are important for transmission of malaria, deeper understanding of their metabolic regulation could be invaluable. Furthermore, transposon mutagenesis screens show that HAD3 is non-mutable, suggesting its essentiality in blood-stage parasites[7]. However, the *Plasmodium berghei* ortholog of HAD3 (PBANKA_1441000) is dispensable in a similar screen for essentiality, the PlasmoGem screen[8]. This data also claimed the HAD5 ortholog to be dispensable, which is not true of *P. falciparum*, highlighting the differences between *P. falciparum* and *P. berghei* and leaving open the question of HAD3 essentiality.

Many of the HAD proteins from *Escherichia coli* have been described and characterized[9], and that data serves to generate hypotheses of functions to orthologous proteins in other organisms. The *E. coli* gene most closely related to HAD3 is the 5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase, YbjI. This enzyme in *E. coli* catalyzes an essential step in riboflavin biosynthesis[10], as well as dephosphorylation of flavin mononucleotide (FMN)[9]. Thus, we hypothesize that HAD3, too, may be involved in flavin metabolism in *P. falciparum*

parasites. It should be noted however, that *Pf*HAD4 also demonstrated highest homology to a FMN phosphatase, YigB, though its activity more closely mimicked that of YrfG, as shown in Chapter 3. Thus, homology alone is not a strong predictor of the substrate specificity of HAD proteins, and direct biochemical and metabolic investigation is necessary.

In *P. falciparum*, riboflavin is taken up from the host cells and media and is essential for or utilized by >30 different parasite enzymes, including dihydroorotate dehydrogenase (DHODH) and the shikimate pathway enzyme, chorismate synthase, both of which are under investigation as drug targets[11–15]. If HAD3 is involved in regulating flavin metabolism in parasites, it may represent a critical point of metabolic control that could be exploited in conjunction with therapeutics that target these flavin-dependent pathways.

Historically in our lab, several others have attempted to purify HAD3, to no avail. However, because I set my sights on HAD proteins as a whole, I believed that I could spend more time and effort optimizing its purification. Therefore, to investigate the function of HAD3, I attempted to express and purify recombinant HAD3 to assay its enzymatic activity.

B.2 RESULTS

B.2.1 Expression and purification of HAD3 from *P. falciparum*

Earlier work from our lab had cloned the coding sequence of *Pf*HAD3 from the genome into the BG1861 vector. Although this was successful, IPTG-induction of this strain exhibited minimal increased protein expression (Figure 1A). We reasoned that the gene sequence was simply not optimal for protein synthesis in *E. coli*, so we obtained a recodonized form of the gene, optimized for expression in *E. coli*. This was cloned into our pKFS expression vector. Expression of the codon-optimized *Pf*HAD3 continued to be only modestly induced (Figure 1B)

but we decided to continue with it in the hopes of obtaining enough protein with large volumes of culture.

After lysis and separation of these cells, however, it became clear that the majority of *Pf*HAD3 was insoluble, as it mostly appears to be present in the insoluble pellet rather than the soluble supernatant. A small portion of nickel-bound protein appears after the soluble fraction is bound, washed, and eluted from nickel beads, but the low signal suggests that a huge volume of culture would be necessary to eventually get enough HAD3 for enzyme assays (Figure 2), and it appears lower than expected so likely is not even HAD3.

Other strategies to help obtain more soluble HAD3 were attempted, including the use of ArcticExpress competent cells, which induce at 4˚C to allow slower synthesis and less potential for aggregation, as well as the addition of a maltose-binding-protein tag to confer solubility. In both cases, the resultant protein continued to be primarily or exclusively found in the insoluble fraction (not shown).

B.2.2 Expression of orthologous HAD3 enzymes

Because of our lack of success purifying recombinant *Plasmodium falciparum* HAD3 from *E. coli*, we reasoned that perhaps an orthologue from a related *Plasmodium* species may be more amenable to recombinant expression and purification. Using the protein solubility predictor ProsoII[16], we found that the *Plasmodium* orthologs with the highest predicted solubility were that of *P. vivax* (PVX_123935) and *Plasmodium cynomolgi* HAD3 (PCYB_145180). We therefore obtained gene blocks for these enzymes, codon-optimized for *E. coli* expression, and cloned them into the pKFS expression vector. Again, however, despite fairly robust expression, virtually all of the protein of interest was observed in the insoluble pellet (Figure 3).

B.3 DISCUSSION

It is not uncommon for recombinantly expressed heterologous proteins to be insoluble. The added effort of extra protein production can stress the protein folding machinery and cause excess misfolded proteins and aggregation[17]. This can be compounded by mismatched codon availability between the heterologous gene and the expression organism, as well as unavailability of eukaryote-specific post-translational modifications when expressing in prokaryotic systems like *E. coli*[18]. In *Pf*HAD3, there are no predicted post-translational modifications, and codon-optimization for *E. coli* expression should have ameliorated most of the problems with mismatched codon availability, increasing the %GC content of native HAD3 from 28% to 41%, much closer to the ~50% GC content of the *E. coli* genome.

While the exact reason for the insolubility of *Pf*HAD3 remains unclear, several strategies can be employed to overcome insolubility. One common method of maximizing protein solubility is low growth temperature. At a lower temperature, protein synthesis slows down, allowing protein folding machinery more time to correctly fold newly synthesized polypeptides[18,19]. Unfortunately, our attempts to express HAD3 at lower temperatures were not successful in preventing insolubility.

Another strategy that has been described is to isolate the aggregated insoluble proteins, termed inclusion bodies in bacteria, and subject the misfolded proteins to denaturing conditions. That denatured protein can then be applied to a wide screen of buffer conditions to allow refolding to occur[20–22]. Additionally, expression in eukaryotic systems, such as mammalian, yeast, or insect cells[23–25] may be more amenable to maintaining HAD3 as a soluble product.

Lastly, we can turn to endogenous HAD3 within parasites. In the last several years, genetic modification of *P. falciparum* parasites has improved immensely, allowing for the

tagging, knockout, knockdown, and knock-sideways of native parasite genes, often in an inducible manner[26]. Targeting the endogenous locus of HAD3 would not only help provide a phenotypic understanding of HAD3's role in parasite metabolism and biology, but would also allow for immunoprecipitation of tagged HAD3, which could then be used for *in vitro* biochemical assays to validate activity and substrate specificity.

We propose that future efforts to investigate HAD3 should follow the above strategies of endogenous genetic modification, and attempts at recombinant protein expression would be wise to utilize a eukaryotic expression system, as mentioned. Deeper understanding of the potential role of HAD3 in parasite biology, possibly through flavin metabolism, could prove valuable for related efforts to therapeutically target flavin-dependent enzymes in the parasites. Furthermore, understanding this sexual-stage expressed enzyme will deepen our understanding of extraerythrocytic parasite metabolism.

B.4 METHODS

B.4.1 Expression and Purification of *Pf*HAD3

The native sequence of PfHAD3 was originally cloned from genomic *P. falciparum* DNA using primers SE03 and SE04 (Table 1) and ligation-independent cloning (LIC) was used to insert it into the BG1861 vector as previously described. This plasmid was kept stored in a frozen glycerol stock of Top10 chemically competent cells for long-term storage, and was transformed into BL21(DE3) pLysS cells (ThermoFisher) for protein expression.

Codon optimized *Pf*HAD3, pre-cloned into the BG1861 vector, was ordered directly from Genewiz. To assist with protein induction, this sequence was cloned out of BG1861 and

into the BG1861-derived pKFS vector using primers PMF49 and PMF50 (Table 1). Cloning was again performed by LIC.

For both the native and codon-optimized *Pf*HAD3 sequences, cultures of BL21(DE3) pLysS cells containing the plasmid of interest were grown shaking at 37˚C in LB media until they reached mid-log phase. Protein expression was then induced with 1mM IPTG, and cultures were grown overnight at 18˚C. Cells were then pelleted, resuspended in Lysis buffer (25 mM Tris HCl (pH 7.5), 250 mM NaCl, 1 mM $MgCl_2$, 20 mM imidazole, and cOmplete EDTA-free protease inhibitor cocktail), and sonicated. Lysates were then spun to pellet the insoluble fraction, and the soluble supernatant was bound to nickel agarose beads (Gold Biotechnology) for 2 hours. Beads were washed 3 times in Wash buffer (Lysis buffer without imidazole or protease inhibitors) and eluted with 10ml of elution buffer (lysis buffer + 300 mM imidazole). After failing to see appreciable quantities of HAD3 in this elution by Coomassie gel, no further purification was attempted.

B.4.2 Expression and Purification of *Pv*HAD3 and *Pc*HAD3

Gene blocks of PVX_123935 and PCYB_145180, including 5' and 3' extensions for LIC cloning, were ordered from IdtDNA. These gene blocks were LIC-cloned into the pKFS vector and kept in Stellar competent *E. coli* cells.

These plasmids were then transformed into BL21(DE3) pLysS cells for protein expression. Cells were grown in just 4 mL of LB until mid-log phase, and induced with 1 mM IPTG at 16˚C overnight. Cells were pelleted, and lysed with Lysis Buffer + 1x BugBuster Protein Extraction Reagent (Novagen) for 20 minutes at room temperature. Lysates were spun down to separate soluble and insoluble fractions.

B.5 FIGURES

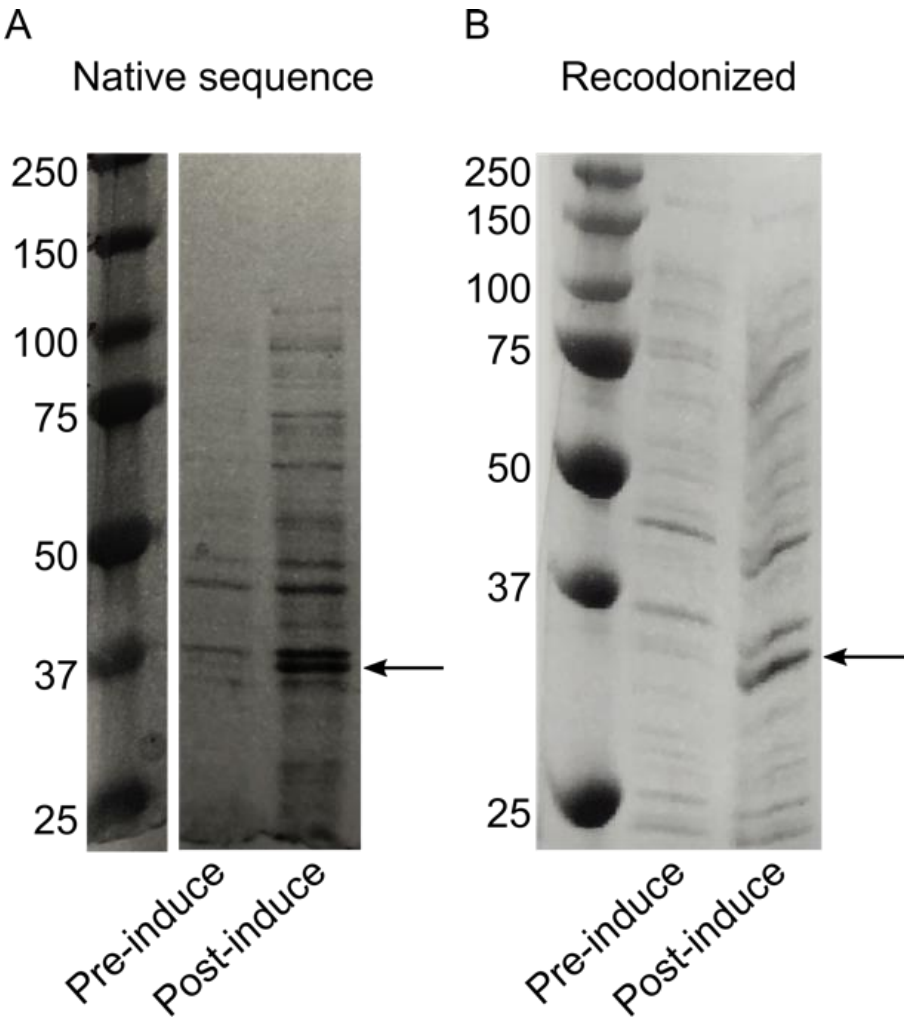


Figure 1. Induction of *Pf*HAD3 protein expression.

*Pf*HAD3 was expressed using either the native sequence cloned from genomic DNA (A) or a recodonized sequence that was optimized for *E. coli* expression (B). In both cases, induction appears modest, and it is likely the lower band of the observed doublet, indicated by arrows, that represents HAD3. Expected protein size for HAD3 is 37.2 kDa for the native sequence in the BG1861 vector and 37.4 kDa for the recodonized sequence in the pKFS vector.

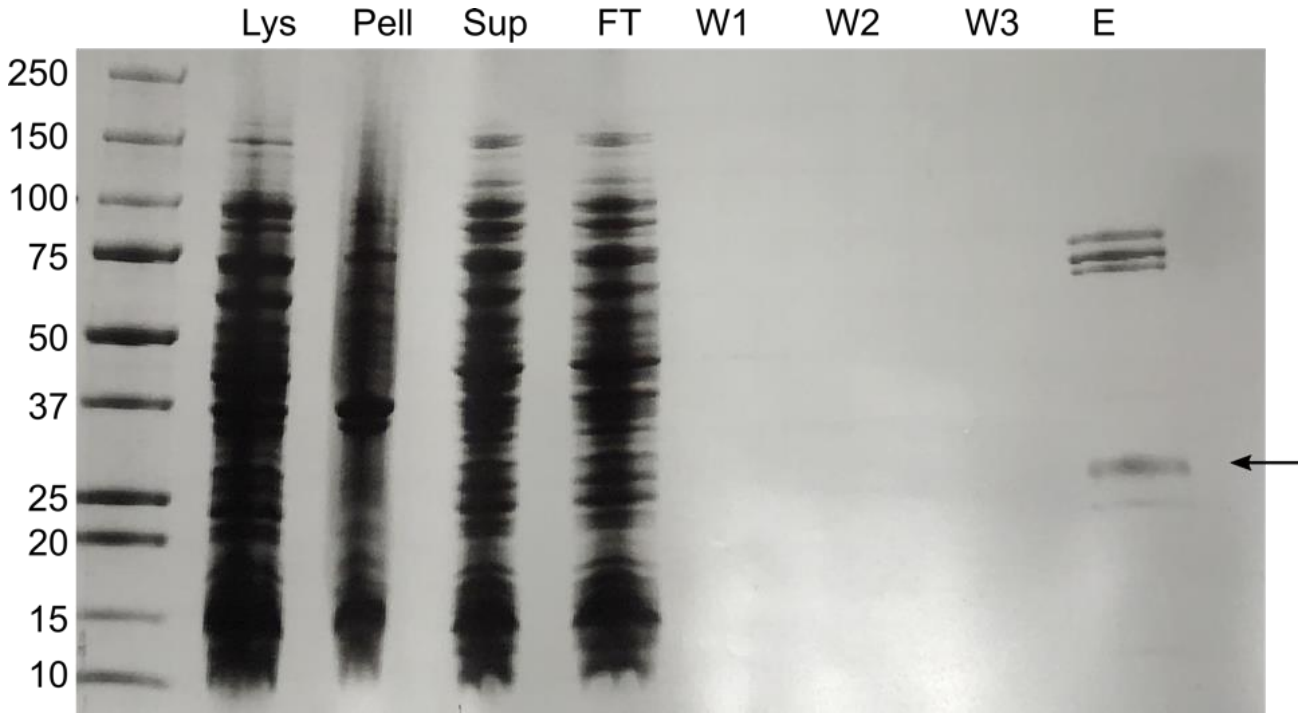


Figure 2. Lysis and Nickel affinity of *Pf*HAD3.

Shown is a Coomassie-stained gel of post-sonication crude lysate (Lys), the insoluble pellet (Pell), and soluble supernatant (Sup) samples of post-induction pKFS-HAD3 (codon optimized). Then, the protein was poured over nickel beads, and the flowthrough (FT) and three washes (W1-W3) are shown, as well as the elution off nickel beads (E). The arrow indicates possible HAD3 protein after elution, though it appears lower than the expected 37.4 kDa.

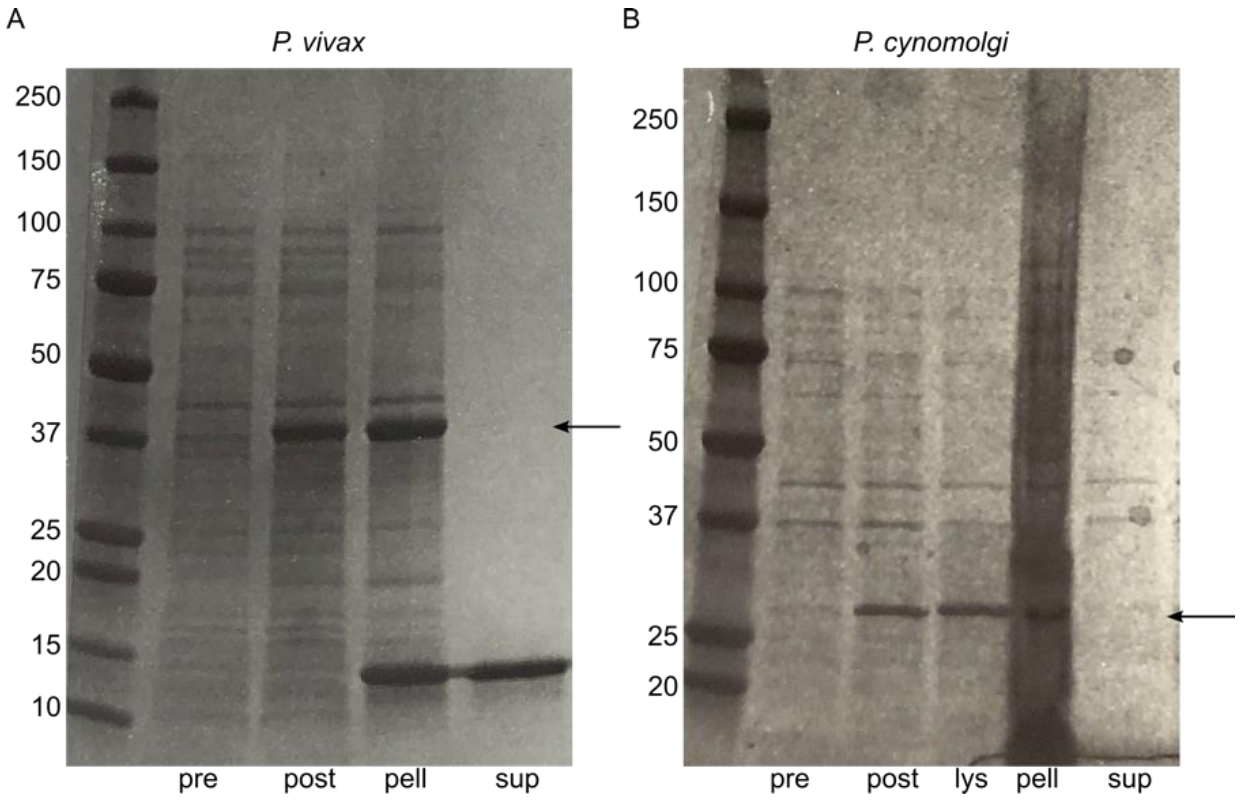


Figure 3. Expression of HAD3 from *P. vivax* and *P. cynomolgi*.

Coomassie-stained gels of the induction and partial purification of HAD3 from *P. vivax* (A) and *P. cynomolgi* (B). Expected sizes of *Pv*HAD3 and *Pc*HAD3 are 39.5 kDa and 28.2 kDa, respectively, indicated by the arrows. Abbreviations: pre, pre-induction sample; post, post-induction sample; lys, crude lysate sample; pell, insoluble pellet sample; sup, soluble supernatant sample.

B.6 Tables

Table 1. Primers Used for Cloning

Name	Sequence
SE03	5’ CTCACCACCACCACCACCATATGTTAAAACACAAATATAACGATTCATGT 3’
SE04	5’ ATCCTATCTTACTCACTTAAGAGCGCCTTCCAGATAAAAC 3’
PMF49	5’ CTCACCACCACCACCACCATCTGAAACAT 3’
PMF50	5’ ATCCTATCTTACTCACTTAACTGCGGCGGCC 3’

B.7 REFERENCES

1. Guggisberg, A. M. *et al.* A sugar phosphatase regulates the methylerythritol phosphate (MEP) pathway in malaria parasites. *Nat. Commun.* **5**, 4467 (2014).
2. Guggisberg, A. M. *et al.* Suppression of Drug Resistance Reveals a Genetic Mechanism of Metabolic Plasticity in Malaria Parasites. *MBio* **9**, e01193-18 (2018).
3. Cobbold, S. A. *et al.* Non-canonical metabolic pathways in the malaria parasite detected by isotope-tracing metabolomics. *Mol. Syst. Biol.* **17**, (2021).
4. Dumont, L. *et al.* The metabolite repair enzyme phosphoglycolate phosphatase regulates central carbon metabolism and fosmidomycin sensitivity in plasmodium falciparum. *MBio* **10**, e02060-19 (2019).
5. López-Barragán, M. J. *et al.* Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum. *BMC Genomics* **12**, (2011).
6. Gómez-Díaz, E. *et al.* Epigenetic regulation of Plasmodium falciparum clonally variant gene expression during development in Anopheles gambiae. *Sci. Rep.* **7**, (2017).
7. Zhang, M. *et al.* Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360**, DOI: 10.1126/science.aap7847 (2018).
8. Schwach, F. *et al.* PlasmoGEM, a database supporting a community resource for large-scale experimental genetics in malaria parasites. *Nucleic Acids Res.* **43**, D1176–D1182 (2015).
9. Kuznetsova, E. *et al.* Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family. *J. Biol. Chem.* **281**, 36149–61 (2006).
10. Haase, I. *et al.* Enzymes from the haloacid dehalogenase (HAD) superfamily catalyse the

elusive dephosphorylation step of riboflavin biosynthesis. *ChemBioChem* **14**, 2272–2275 (2013).

11. Hoelz, L. V., Calil, F. A., Nonato, M. C., Pinheiro, L. C. & Boechat, N. Plasmodium falciparum dihydroorotate dehydrogenase: A drug target against malaria. *Future Medicinal Chemistry* **10**, 1853–1874 (2018).
12. Fitzpatrick, T. *et al.* Subcellular localization and characterization of chorismate synthase in the apicomplexan Plasmodium falciparum. *Mol. Microbiol.* **40**, 65–75 (2001).
13. Hossain, M. M. *et al.* Structural analysis and molecular docking of potential ligands with chorismate synthase of Listeria monocytogenes: a novel antibacterial drug target. *Indian J. Biochem. Biophys.* **52**, 45–59 (2015).
14. B. Dias, M. *et al.* Chorismate Synthase: An Attractive Target For Drug Development Against Orphan Diseases. *Curr. Drug Targets* **8**, 437–444 (2007).
15. Kokkonda, S. *et al.* Lead optimization of a pyrrole-based dihydroorotate dehydrogenase inhibitor series for the treatment of malaria. *J. Med. Chem.* **63**, 4929–4956 (2020).
16. Smialowski, P., Doose, G., Torkler, P., Kaufmann, S. & Frishman, D. PROSO II - A new method for protein solubility prediction. *FEBS J.* **279**, (2012).
17. Ferrer-Miralles, N., Saccardo, P., Corchero, J. L., Xu, Z. & Garcfa-Fruitós, E. General introduction: Recombinant protein production and purification of insoluble proteins. *Methods Mol. Biol.* **1258**, 1–24 (2015).
18. Francis, D. M. & Page, R. Strategies to optimize protein expression in E. coli. *Current Protocols in Protein Science* **61**, 5.24.1-5.24.29 (2010).
19. Song, J. M., An, Y. J., Kang, M. H., Lee, Y. H. & Cha, S. S. Cultivation at 6-10°C is an effective strategy to overcome the insolubility of recombinant proteins in Escherichia coli.

Protein Expr. Purif. **82**, 297–301 (2012).

20. Armstrong, N., Lencastre, A. De & Gouaux, E. A new protein folding screen: Application to the ligand binding domains of a glutamate and kainate receptor and to lysozyme and carbonic anhydrase. *Protein Sci.* **8**, 1475–1483 (1999).
21. Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M. & Panda, A. K. Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microbial Cell Factories* **14**, (2015).
22. Ramakrishnan, B. & Qasba, P. K. In vitro folding of β-1,4Galactosyltransferase and Polypeptide-α-N-Acetylgalactosaminyltransferase from the inclusion bodies. *Methods Mol. Biol.* **1022**, 321–333 (2013).
23. Wurm, F. M. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnology* **22**, 1393–1398 (2004).
24. Radner, S. *et al.* Transient transfection coupled to baculovirus infection for rapid protein expression screening in insect cells. *J. Struct. Biol.* **179**, 46–55 (2012).
25. Böer, E., Steinborn, G., Kunze, G. & Gellissen, G. Yeast expression platforms. *Applied Microbiology and Biotechnology* **77**, 513–523 (2007).
26. Kudyba, H. M., Cobb, D. W., Vega-Rodríguez, J. & Muralidharan, V. Some conditions apply: Systems for studying Plasmodium falciparum protein function. *PLoS Pathogens* **17**, (2021).