Investigating the structure-function relationships of *Plasmodium* Haem Detoxification Protein and Phosphatidylinositol 4-kinase

Lynn Wambua

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

In the Department of Chemistry

UNIVERSITY OF CAPE TOWN

July 2022

Supervisors: Professor Kelly Chibale and Professor Timothy J. Egan

Co-supervisor: and Dr Lauren B. Arendse

Declaration

I, Lynn Wambua declare that:

1. The above thesis is my own unaided work, both in concept and execution, and that apart from the normal guidance from my supervisor, I have received no assistance except as stated below
2. That in cases where others' work has been cited, this has been acknowledged and referenced
3. Neither the substance nor any part of the above thesis has been in the past, or is being, or is to be submitted for a degree at this University, or any other university, except as stated below
4. I grant the University of Cape Town free licence to reproduce the above thesis, in whole or in part, for the purpose of research

I am now presenting the thesis for examination for the degree of DOCTOR OF PHILOSOPHY in the Department of Chemistry, University of Cape Town.

Signature of Candidate:_______________________ **Date:** 03/07/2022

Abstract

The human malaria parasite *Plasmodium falciparum* relies on proteins and protein-mediated processes to survive, replicate in the host and evade the host's immune response. This study focuses on two *Plasmodium* proteins: (i) Haem Detoxification Protein (HDP) which has been reported to catalyse haemozoin formation in *Plasmodium* and (ii) Phosphatidylinositol-4-Kinase β (PI4Kβ) which plays an essential role in multiple stages of the parasite lifecycle and is a clinically validated drug target for malaria chemotherapy.

Previous efforts to characterise HDP have been limited to experiments using the refolded form of recombinant protein and the role of HDP in haemozoin formation is still unclear. The aim of this research project was to optimise expression and purification of soluble *Plasmodium falciparum* HDP and characterise the protein's role in haemozoin formation

Efforts to optimise soluble HDP expression were effective albeit exhaustive efforts to purify HDP from the soluble fraction were unsuccessful. Purification of HDP under denaturing conditions was achieved using previously reported methods. Refolded HDP was evaluated for β-hematin formation activity, but the results indicated β-hematin formation was mediated by sodium dodecyl sulfate used in the assay rather than by HDP. HDP was co-crystalised in the presence of haem but crystals diffracted poorly. Evaluation of the predicted HDP structure based on homology modelling showed that the four histidine residues predicted to facilitate haemozoin formation are not aligned in a way that would facilitate haemozoin crystal growth. This, coupled with the data from the biochemical assays, suggests HDP is unlikely to be involved in haemozoin formation.

The second aim of this project was to use *Plasmodium vivax* PI4Kβ inhibition assays coupled with site-directed mutagenesis and mass spectrometry to support target-based malaria drug discovery programs focused on the development of both ATP-competitive and covalent *Plasmodium* PI4Kβ inhibitors.

Two residues of interest in *Plasmodium vivax* PI4Kβ unique to *Plasmodium*, F832 and C1327, were mutated to alanine. F832 is thought to form key Pi-Pi interactions with inhibitors and C1327, found on the periphery of the catalytic site, is a potential target for covalent inhibitors. Site-directed mutagenesis was used to introduce mutations at F832 and C1327 in wild-type *Plasmodium vivax* PI4Kβ and wild-type PI4Kβ and the two mutants were expressed and purified. Kinetic characterisation revealed the three enzymes had similar kinetic parameters.

Inhibition data indicated the F832 mutation to alanine had a minimal effect on inhibition of PI4Kβ. A time-dependent inhibition assay was established to evaluate targeted covalent inhibitors (TCIs) using wild-type PI4Kβ and the PI4Kβ C1327A mutant. A distinct decrease in wild-type PI4Kβ IC_{50} was observed with increasing enzyme-inhibitor pre-incubation time for inhibitors containing chloroacetamide and acrylamide warheads. In contrast, PI4Kβ C1327A IC_{50} values were independent of enzyme-inhibitor pre-incubation time supporting C1327-mediated covalent inhibition. Mass spectrometry was used to confirm covalent modification of the targeted cysteine residue and to assess the rate of covalent bond formation. These assays provide valuable insights, which can be used to guide the optimisation of PI4Kβ inhibitors and the choice of warheads for TCIs.

Dedication

To my dearest grandmother and namesake, Lois Mutio

Acknowledgements

I am beyond grateful to my family, my parents Gregory and Jacinta for your endless support and always having faith in me. To my brother Lloyd, thank you for listening to my complaining, walking me to my car, making me dinner and providing comedic relief when I needed it.

To my supervisor Professor Kelly Chibale, thank you for providing me with the opportunity to pursue my PhD and for the incredible opportunities I have had while under your supervision. Your humility, tireless work ethic and dedication continue to inspire me. To Professor Timothy Egan, thank you for your kindness and patience and for teaching me to follow the science. You were always supportive and provided guidance when I needed it and for that I am sincerely grateful. You will be dearly missed, may you rest in peace. To my co-supervisor Dr Lauren Arendse, thank you for going above and beyond for me and being great mentor. I have learned so much from you and your support has been invaluable through this journey.

Thank you to the Haem Team, particularly Dr Roxanne Mohunlal, for creating such an encouraging work environment and for the strength and love shown especially in this last year. Thank you Kyllen Dilsook for helping with the structure predictions.Thank you to the K.C academic group for the camaraderie particularly Stephanie Kamunya for teaching me the computational basics and Ferdinand Ndubi for providing the TCI compounds used in the project. Special thanks to Elaine Rutherfoord-Jones for always helping me with admin when I needed it. I would also like to thank the H3D biology group for being a source of joy, support and a sounding board when needed.

Thank you to Brandon Weber for teaching me more about protein purification than I thought was possible and for your endless optimism. Thank you to Dr Stephen Fienberg for your help and guidance with the docking. Thank you to Dr Andrew Nel for all your help with the insect cell work and to Sylva Schwager and Professor Edward Sturrock for allowing me to use your equipment. Thank you to Professor Trevor Sewell for allowing me to use your facility and for guiding me on the structural work. Thank you to Phillip Venter for teaching me crystallography and helping me with data collection even at ridiculous hours of night. Thank you to Mohamed Jaffer for your help with the electron microscopy work. A big thank you to Tariq Ganief for your guidance and collaborating with me on the mass spectrometry work.

I am grateful to my friends who have encouraged and supported me through this journey, Glyn Oree, Kaylene Baron, Winnie and Rosa and the Muthama family (my home away from home).

To my sisters of fate, Farrah Khan and Leigh-anne Derry, I am so lucky to have you as my best friends. Thank you for being my biggest cheerleaders and creating our weird little family. Leigh-anne, to be honest girl you didn’t help me write this thesis but thank you for being there when I wanted to discuss murder mysteries and drama. I love you and cannot wait to be in the same country as you.

Finally, I would like to thank the National Research Foundation, the South African Medical Research Council and the Future Leaders-African Independent Research Fellowship Programme for the financial support provided during this project.

Table of Contents

Abbreviations

5-ALA- 5-aminolevulinic acid

ACBD$_3$- Acyl-CoA binding domain 3

ACT- Artemisinin-based combination therapy

ADP- Adenosine diphosphate

AGE- Agarose gel electrophoresis

ART- Ataxia telangiectasia and Rad3-related protein

AT- Adenine-thymine

aTc- Anhydrotetracycline

ATP- Adenosine triphosphate

BSA- Bovine serum albumin

CASP- Critical assessment of protein structure prediction

CRT- Chloroquine resistance transporter

CryoEM- Cryogenic electron microscopy

CV- Coefficient of variation

DHFR-TS- Dihydrofolate reductase-thymidylate synthase

DHPS- Dihydropteroate synthase

DMSO- Dimethyl sulfoxide

DNA- Deoxyribonucleic acid

dNTP- Deoxyribonucleoside triphosphate

DTT- Dithiothreitol

E.coli- *Escherichia coli*

EDTA- Ethylenediaminetetraacetic acid

GC- Guanine-cytosine

GST- Glutathione S-transferase

HbD- Haemoglobin binding

HDP- Haem detoxification protein

HeD- Haem-binding

His-tag- polyHistidine tag
HRP- Horseradish peroxidase
HRPII- Histidine Rich Protein II
IPTG- Isopropyl β-d-1-thiogalactopyranoside
IR- Infrared
kDa- Kilodalton
LB- Luria broth
mdr1- Multidrug resistance protein 1
MMV- Medicines for Malaria Venture
MS- Mass spectrometry
MW- Molecular weight
NTA- Nitriloacetic acid
PBS-Phosphate buffered saline
PCR- Polymerase chain reaction
PDB- Protein Data Bank
Pf- *Plasmodium falciparum*
PI- Phosphatidylinositol
$PI(3,4)P_2$- phosphatidylinositol 3,4-bisphosphate
$PI(3,5)P_2$- Phosphatidylinositol 3,5-bisphosphate
$PI(4,5)P_2$- phosphatidylinositol 4,5-bisphosphate
PI3K- Phosphatidylinositol 3-kinase
PI3P- Phosphatidylinositol 3-phosphate
PI4K- Phosphatidylinositol 4-kinase
PI4P- Phosphatidylinositol 4-phosphate
PI5P- Phosphatidylinositol 5-phosphate
PIK- Phosphatidylinositol kinase
PIP- Phosphoinositide
PIP_3- Phosphatidylinositol 3,4,5-triphosphate
PIPK- Phosphatidylinositol phosphate kinases
Pv- *Plasmodium vivax*

rHDP- Refolded HDP

RMSD- Root mean square deviation

rpm- Revolutions per minute

SAR- Structure activity relationship

SDS- PAGE- Sodium dodecyl-sulfate

SDS- PAGE- Sodium dodecyl-sulfate polyacrylamide gel electrophoresis

SEC- Size exclusion chromatography

Sf9- Spodoptera frugiperda

TAE- Tris-acetate-EDTA

TB- Terrific broth

TCPs- Target Candidate Profiles

TEM- Transmission electron microscopy

TEMED- Tetramethylethylenediamine

Tris- Tris(hydroxymethyl)aminomethane

TTC7- Tetratricopeptide repeat domain 7A

TTPs- Target Product Profiles

UCT- University of Cape Town

WHO- World Health Organisation

WT-Wild-type

X-Gal- 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1

Literature Review

1.1 Malaria

Malaria is one of the most prevalent diseases in the world and occurs in 85 countries. According to the World Health Organisation (WHO) 2021 Malaria Report, there was an estimated 241 million cases of malaria in 2020 and 627,000 deaths worldwide (WHO Malaria Report, 2021). Of the estimated cases, 95 % occurred in the WHO African region and 96 % of all malaria deaths recorded in 2020 also occurred in the WHO African region (WHO Malaria Report, 2021). Furthermore, 80% of the deaths in the WHO African region occurred in children under 5 years (WHO Malaria Report, 2021). Since 2000, there has been a steady decrease in malaria incidence and malaria deaths worldwide until 2019 which saw an increase in malaria cases and deaths between 2019 and 2020 (WHO Malaria Report, 2021). Malaria remains a major threat to global health particularly in Asia, sub-Saharan Africa and South America (Forte *et al*., 2021).

1.1.1 Malaria Lifecycle

Malaria is caused by protozoan parasites of the *Plasmodium* genus Phillips *et al*., 2017). There are five *Plasmodium* species that cause malaria in humans: *Plasmodium malariae* (*P. malariae*), *Plasmodium falciparum* (*P. falciparum*), *Plasmodium ovale* (*P. ovale*), *Plasmodium vivax* (*P. vivax*) and simian *Plasmodium knowlesi* (*P. knowlesi*) (Talapko *et al*., 2019). *Plasmodium falciparum* is the most prevalent malaria parasite and is predominantly found in sub-Saharan Africa whereas *Plasmodium vivax* (*P. vivax*) is predominantly found in the WHO Region of the Americas where it is responsible for 75 % of malaria cases (Ashley *et al*., 2018; WHO, 2021). *Plasmodium falciparum* causes the majority of malaria deaths (WHO, 2021).

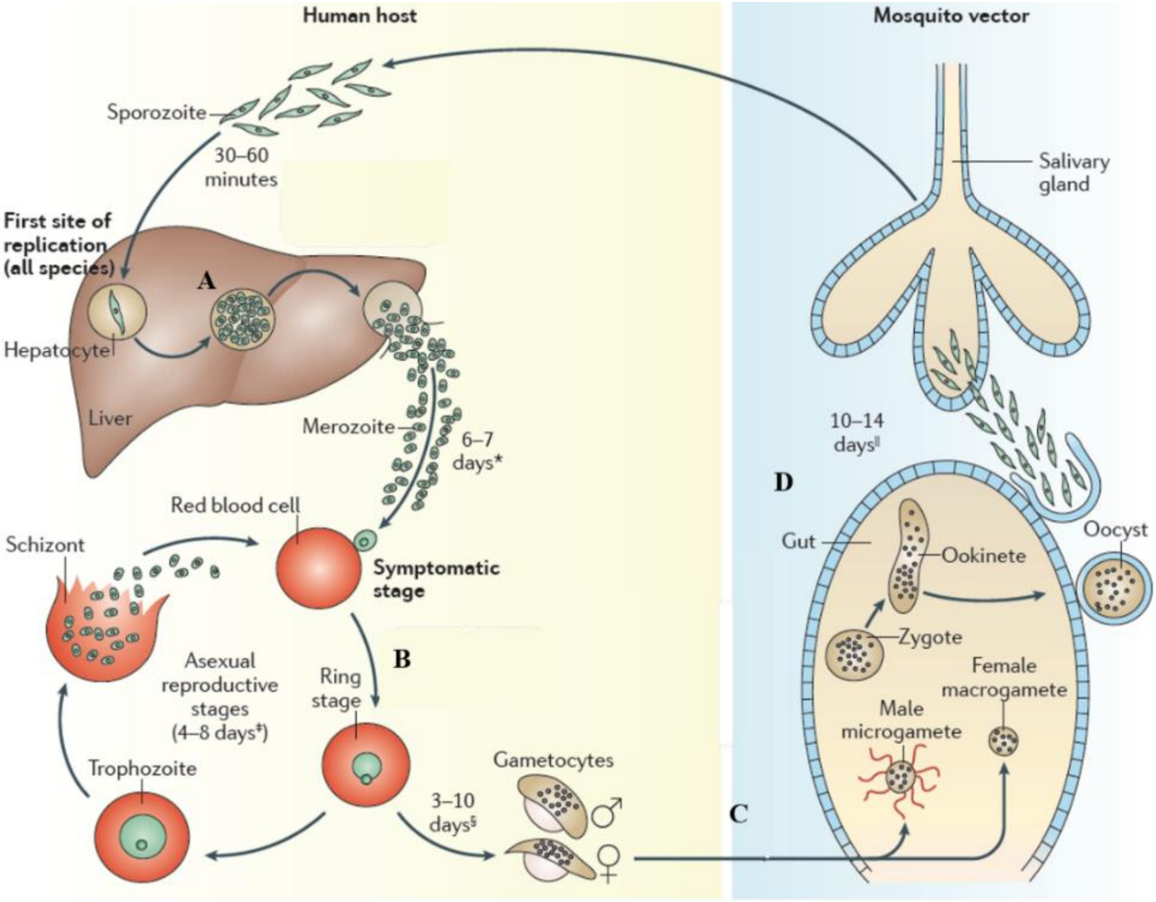


Figure 1.1: The lifecycle of malaria causing *Plasmodium* (adapted from Phillips *et al*., 2017).

There are four stages in the *Plasmodium* lifecycle as discussed in the reviews by Biamonte *et al*. (2013) and Phillips *et al*. (2017): **A**: liver stage, **B**: erythrocyte stage, **C**: transmission stage and **D**: mosquito stage as shown in Figure 1.1. Human infection begins once an infected female *Anopheles* mosquito inoculates sporozoites into the blood stream during a blood feed (Biamonte *et al*., 2013). The sporozoites are transported to the liver where they invade hepatocytes and replicate, producing schizonts – known as the liver stage – Figure 1.1A (Phillips *et al*., 2017). After 5 – 10 days, the parasites enter the erythrocyte stage (Figure 1.1B) where the liver cells rupture releasing merozoites which invade red blood cells and proliferate rapidly by asexual schizogony (Phillips *et al*., 2017). In the red blood cell, parasites go through various forms i.e. rings to trophozoites to schizonts which form approximately 20 daughter merozoites that are released into the blood stream, re-initiating the erythrocyte stage cycle (Biamonte *et al.*, 2013). The erythrocyte stage is the symptomatic stage of malaria as a result of high parasitaemia in the host (Ashley *et al*., 2018). Symptoms of uncomplicated malaria include fever, headache, chills, cough, body-ache and diarrhoea while in more severe cases, multi-organ system failure, coma, pulmonary complications, severe anaemia and hypoglycaemia can occur (Phillips *et al*., 2017). Drugs that target the erythrocyte stage control malaria symptoms (Biamonte *et al*., 2013).

In the transmission stage, (Figure 1.1C), some parasites undergo sexual differentiation and develop into male and female gametocytes (Biamonte *et al*., 2013). Parasites enter the mosquito stage (Figure 1.1D) when the male and female gametocytes are ingested by a mosquito via a blood feed and fuse in the midgut forming a diploid zygote (Phillips *et al*., 2017). The zygote develops into a mobile ookinete which passes through the gut wall as an oocyst (Ashley *et al*., 2018). The oocyst undergoes several replication cycles, forming sporozoites which migrate to the mosquito's salivary glands thus completing the lifecycle (Phillips *et al*., 2017). Drugs that target the transmission and mosquito stages benefit the malaria eradication agenda and would prevent the infection of other humans (Biamonte *et al*., 2013).

1.1.2 Malaria prevention strategies and treatment

There are three malaria prevention strategies: chemoprophylaxis, vaccination and vector control as reviewed by Ashley *et al*. (2018). Chemoprophylaxis is often targeted at young children, pregnant women and travellers (Phillips *et al*., 2017). Vector control mainly focuses on the use of long-lasting insecticide (pyrethroid) treated bed nets and indoor residual spraying with insecticides and has been relatively successful in preventing malaria cases (Ashley *et al*., 2018). However, the success of these measures is threatened by several factors: behavioural changes in the mosquito (biting during the day), use of treated bed nets for fishing and the widespread resistance to insecticides, especially pyrethroids, necessitating the urgent need for new insecticides (Phillips *et al*., 2017).

Currently, there is only one malaria vaccine that has been approved for widespread use – RTS,S/AS01 (RTS,S) (The Lancet, 2021). The vaccine targets the sporozoite stage of the parasite lifecycle – blocking infection of the liver (The Lancet, 2021). There are several other vaccines in development targeting the erythrocyte stage, the mosquito stage, the placenta malaria erythrocyte stage and the pre-erythrocytic stage (Duffy and Gorres, 2020).

Antimalarial drugs have been the main method of treatment, control and chemoprevention against malaria (Cowman *et al*., 2016). There are three major classes of antimalarial drugs currently used – quinoline derivatives, antifolates and artemisinins – administered alone or in combination (Shahinas *et al*., 2013). Most antimalarial drugs target the parasite's erythrocyte stage while some also target the gametocyte stage of the parasite lifecycle (Santos and Torres, 2013). Quinoline derivatives such as quinine, mefloquine, amodiaquine, lumefantrine and chloroquine (CQ) to varying degrees share a similar mode of action, interfering with hemozoin formation – a process where toxic free haem is crystalised into a non-toxic biomineral known as the haemozoin - in the parasite's digestive vacuole thus inhibiting the parasite's ability to remove the toxic haem (Shahinas *et al*., 2013). Quinine, lumefantrine and mefloquine have also been proposed to inhibit haemoglobin import (Ross and Fidock, 2019). Antifolates such as proguanil, pyrimethamine and sulfadoxine target the folate biosynthesis pathway by blocking dihydrofolate reductase and dihydropteroate synthetase enzymes in the parasite thus inhibiting folic acid synthesis (Daskum *et al*., 2020). The mode of action of artemisinin and its derivatives is widely debated but most commonly it is suggested they function by generating free radicals in the presence of free haem which subsequently damages proteins required for parasite survival (Tse et al *et al*., 2019).

Artemisnin based combination therapy (ACT) is the first line of treatment for uncomplicated *P. falciparum* malaria (Naing *et al*., 2019). The five ACTs recommended by WHO for *P. falciparum* malaria treatment are: artemether-lumefantrine, artesunate-mefloquine, artesunate-amodiaquine, dihydroartemisinin-piperaquine and artesunate in combination with sulfadoxine-pyrimethamine (Naing *et al*., 2019). Artemisinin and its derivatives are highly efficacious and are fast acting but have a short half-life and thus are rapidly eliminated (Ashley and Phyo, 2018). Combining artemisinin and its derivatives with more slowly eliminated drugs allows for an initial fast decrease in parasitaemia and elimination of remaining parasites over an extended period of time thus providing highly effective treatment aimed at improving cure rates and reducing the risk of resistance developing (Ashley and Phyo, 2018; Talapko *et al*., 2019).

1.1.3 Antimalarial drug resistance

Resistance has emerged to nearly all existing antimalarial drugs (Cui *et al*., 2015). Antimalarial resistance is often caused by poor adherence, incorrect dosing, poor absorption, misdiagnosis, overuse of antimalarials and poor drug quality (Watsierah *et al.* 2010; Shahinas *et al*., 2013). Antimalarial drug resistance often begins in low transmission areas such as South America or Southeast Asia before spreading to high transmission areas such as sub-Saharan Africa (Menard and Dondorp, 2017). Genetic and biochemical approaches have shown several molecular markers are associated with antimalarial drug resistance as reviewed by Cui *et al*. (2015) and Menard and Dondorp, (2017).

Mutations in several *Plasmodium* proteins have been linked to antimalarial drug resistance (Cui *et al*., 2015). Point mutations in the *P. falciparum* chloroquine resistance transporter (*Pf*CRT) gene - a food vacuole transporter protein - mediates resistance to chloroquine and other 4-aminoquinolines by increasing drug export from the food vacuole, away from their primary site of action (Fidock *et al.* 2000; Cui *et al*., 2015). 4-aminoquinolines are weak bases in neutral pH (Cowell and Winzeler, 2019). Once the drugs enter the acidic food vacuole, they become protonated and accumulate. Additionally, drug accumulation has been shown to occur due to the drug i.e. chloroquine binding to heme (Bray *et al.* 1999). *Pf*CRT facilitates the efflux of protonated drug molecules out of the food vacuole (Cowell and Winzeler, 2019; Wicht, Mok and Fidock, 2020).

Plasmodium falciparum multidrug resistance protein 1 (*Pf*mdr1) gene encodes an ABC transporter (ATP-binding cassette) which is a P-glycoprotein homolog (Menard and Dondorp, 2017). The protein is localised in the food vacuole membrane and is thought to transport antimalarial drugs into the food vacuole (Rohrbach *et al*., 2006). Data shows that changes in the copy number or sequence of *Pf*mdr1 alters the transport of multiple drugs in to or out of the parasite food vacuole, however some changes decrease the parasite’s susceptibility for one drug while increasing sensitivity for another drug. For example, increased *Pf*mdr1copy number i.e. amplification of the *Pf*mdr1 gene, decreases *P. falciparum* drug susceptibility to mefloquine, lumefantrine, halofantrine and artemisinins by transporting them away from their site of action but also increases the parasites drug sensitivity to chloroquine and piperaquine by increasing their concentration in the food vacuole (Menard and Dondorp, 2017; Wicht, Mok and Fidock, 2020). Furthermore, mutations in *Pf*mdr1 are thought to inhibit transport of antimalarial drugs into the food vacuole thus decreasing the concentration of drugs such as chloroquine, amodiaquine and piperaquine in the food vacuole resulting in decreased *P. falciparum* dug susceptibility (Cowell and Winzeler, 2019).

Resistance to antifolate drugs such pyrimethamine, proguanil and sulfa drugs such as sulfadoxine occurs as a result of point mutations in the *P. falciparum* bifunctional dihydrofolate reductase-thymidylate synthase (*Pf*DHFR-TS) gene and the dihydropteroate synthase (*Pf*DHPS) gene (Menard and Dondorp, 2017). *Pf*DHFR-TS encodes an enzyme involved in the pathway of folate synthesis and is the target of antifolate drugs (Menard and Dondorp, 2017). Point mutations alter binding of drugs to the enzyme, resulting in resistance (Cowell and Winzeler, 2019). In *Pf*DHFR-TS, point mutations occur in a particular order starting with codon 108 (S→N), followed by N51I, C59R and lastly I164L thus susceptibility to pyrimethamine after the initial mutation can be decreased further by mutations at codons 51,59 and 164 (Sharma *et al*., 2015; Cowell and Winzeler, 2019). The *Pf*DHPS gene encodes an enzyme involved in the de novo synthesis of essential folate co-enzymes and resistance to sulfa drugs such as sulfadoxine often occurs as a result of mutations in the following codons: I1431V, S436A/F, A437G, K540E, A581G and A613S/T (Xu *et al*., 2019). A combination of mutations in the *Pf*DHFR-TS and *Pf*DHPS genes is associated with sulfadoxine – pyrimethamine combination therapy failure (Menard and Dondorp, 2017).

Artemisinin resistance was first identified in 2007 in Southeast Asia and is a major threat to malaria control as ACTs are the main method of treatment for uncomplicated malaria (Noedl *et al.*, 2008). Artemisinin resistance is characterised by delayed parasite clearance (Conrad and Rosenthal, 2019). Ariey *et al*. (2014) found mutations in the *Pf* kelch13 propeller domain were responsible for reduced artemisinin susceptibility. Today, WHO considers 10 *Pf*kelch13 mutations -that have been identified and validated - as molecular markers for artemisinin resistance; F446I, N458Y, M476I, Y493H, R539T, I543T, P553L, R561H, P574L, and C580Y with a further 11 mutations as candidate or associated markers of artemisinin resistance (Rosenthal, 2021). A study by Uwimana *et al*. (2021) found the R561H mutation was present in 12.8 % of the *P. falciparum* malaria cases in Rwanda in 2018 and delayed parasite clearance was observed. Since 2017, the emergence and spread of artemisinin resistance has also been seen in Northern Uganda which has been associated with candidate (A675V) and associated (C469Y) markers of artemisinin resistance (Balikagala *et al*., 2021; Rosenthal, 2021).

Widespread resistance to most antimalarials and emerging resistance to artemisinin drives drug development research for alternative drugs with novel mechanisms of action and new therapeutic approaches (Shahinas *et al.*, 2013).

1.1.4 Antimalarial drug discovery

Exploring new drug combinations and formulations of existing antimalarial drugs is one way to overcome drug resistance or improve drug delivery. Tse *et al*. (2019) reviews these aspects of antimalarial drug. Alternatively, drugs that are used for other purposes but are efficacious against malaria can be repurposed as new antimalarial compounds as discussed by Tse *et al*. (2019). As previously mentioned, drug resistance drives the need for new antimalarials with novel modes of action.

Ideally, new antimalarials should be safe, cheap, kill the parasite quickly and address the issue of resistance (Shahinas *et al*., 2013). Compounds with activity against all stages of the parasite will contribute to the goal of elimination (Flannery *et al*., 2013). Fast treatment would ensure compliance and minimise the risk of resistance emerging while drugs with new mechanisms of action would allow for the possibility of synergy with existing antimalarials and reduce the likelihood of cross-resistance developing (Shahinas *et al*., 2013). The process of identifying new antimalarials has evolved over the last decade with a major shift towards phenotypic screening and more recently towards target-based approaches as more potential *Plasmodium* drug targets are being validated (Hovlid and Winzeler, 2016; Forte *et al*., 2021). Large compound libraries have been screened for potential antimalarial activity and identified compounds act as a starting point for drug development (Cowell and Winzeler, 2019).

To guide antimalarial drug discovery and development, the Medicines for Malaria Venture (MMV) developed Target Product Profiles (TTPs) and Target Candidate Profiles (TCPs) (Burrows *et al*., 2013). TPPs describe goals for antimalarial medications while TCPs describe goals for antimalarial compounds (Burrows *et al*., 2013). There are two TPP classifications TPP-1 and TPP-2 with the goals of treating active disease and chemoprotection respectively (Burrows *et al*., 2017). Table 1.1 shows the recently defined TCP guidelines for antimalarial compounds.

Table 1.1: Target Candidate Profile Definitions (adapted from Forte *et al*., 2021).

TCP	Definition	Goal
TCP-1	Active against the erythrocytic stage. Active against all resistant strains.	Treatment of severe and uncomplicated malaria. Chemoprophylaxis.
TCP-3	Active against liver stage hypnozoites	Anti-relapse
TCP-4	Active against liver stage	Prophylaxis
TCP-5	Active against *Plasmodium* gametocytes	Transmission blocking
TCP-6	Active against insect vector- blocking transmission	Transmission blocking

New antimalarial compounds are evaluated based on a number of requirements: potential single-dose cures, novel modes of action with no cross-resistance to existing drugs, ability to prevent infection, activity against both the asexual blood stage and the gametocyte stage and the ability to clear *P. vivax* hypnozoites from the liver (Tse *et al.*, 2019). Table 1.2 shows some of the newer drug targets currently pursued in the effort to develop drugs with novel modes of action.

Table 1.2: Summary of antimalarials with novel targets currently under development (adapted from Tse *et al.*, 2019 and Belete, 2020).

Target	Function	Inhibitor	Phase***
Translational elongation factor 2 (*Pf*EF2)	Catalyses the translocation of tRNA and mRNA*	**M5717**	1
V-type H+- ATPase	Efflux of H+ to maintain intracellular pH of 7.3	**MMV674253**	1
Phosphatidylinositol 4 -kinase (*Pf*PI4K)	Regulates intracellular signalling and trafficking	**MMV390048**	2 – development discontinued
P-type Na+ - ATPase transporter (*Pf*ATP4)	Shuttles sodium ions out of the cell	**SJ733**	1
Dihydroorotate dehydrogenase (*Pf*DHODH)	Catalyses the oxidation of dihydroorotate to produce orotate **	**DSM265**	2
***Pf* Duffy-binding like domain 1α (DBL1α)**	Involved in sequestration of infected red blood cells	**Sevuparin**	2

*Essential factor for eukaryotic protein synthesis /**A key step in pyrimidine biosynthesis /*** Clinical trial phase

The haemoglobin degradation pathway has been an attractive drug target for antimalarials for many years. Identifying drugs that act on new targets and additionally, drug therapies that target various stages of the parasite lifecycle in order to reduce the risk of resistance developing is important in light of increasing antimalarial drug resistance to currently used treatments (Hovlid and Winzeler, 2016). Kinases, both protein and phosphatidylinositol kinases, have emerged as promising novel targets for malaria drug discovery as demonstrated by KDU691 target elucidation (McNamara *et al.*, 2013).

1.2 Haemoglobin degradation

Haemoglobin degradation is an essential process for the development of malaria parasites during the erythrocytic stage of the parasite lifecycle and thus is an attractive target in malaria drug discovery (Goldberg, 2013; de Villiers and Egan, 2021). Haemoglobin is internalised by the parasite, transported to and digested in the food vacuole (Elliot *et al.*, 2008). Internalisation of haemoglobin from the red blood cell cytosol is thought to occur via a localised invagination of the parasites' parasitophorous vacuolar membrane and plasma membrane known as cytostomes (Milani *et al*., 2015). Cytostomes are transported to the parasite's food vacuole and fuse with the food vacuole membrane, releasing haemoglobin into the food vacuole (Milani *et al*., 2015). Haemoglobin is degraded inside the parasite's food vacuole resulting in the generation of amino acids, which *Plasmodium* parasites use for growth and development, and free haem, which is toxic to the parasite as it can lead to the formation of reactive oxygen species that can consequently induce oxidative stress (Jani *et al*., 2008; Gupta *et al*., 2017). In *Plasmodium* parasites, free haem is detoxified by conversion into an inert, insoluble crystalline dimer called haemozoin (Egan, 2008). The haemozoin crystal structure is identical to that of β-haematin and consists of a haem dimer in which the ferric iron of one haem molecule binds to the oxygen of the propionate carboxylate group of an adjacent haem molecule via an iron-oxygen coordinate bond (Pagola *et al*., 2000).

The molecular mechanism of haemozoin formation is highly disputed and several parasite factors have been proposed to catalyse haem detoxification (Nakatani *et al*., 2014). Some research groups have suggested that lipids, proteins, or a combination of the two, catalyse the formation of haemozoin while others have hypothesised that lipids could be the primary factor, as the lipid fraction from the *P. falciparum* (*Pf*) food vacuole promotes haemozoin formation *in vitro* (Jani *et al*., 2008; Nakatani *et al*., 2014). The first protein shown to promote β-haematin formation was histidine-rich protein II (HRPII), although *Pf*HRPII knockouts retain the ability to grow and survive normally, indicating that other factors are involved in the formation of haemozoin (Gupta *et al*., 2017). Recently, haem detoxification protein (HDP) has been reported as a potent β-haematin producing protein (Chugh *et al*., 2013). According to Jani *et al*. (2008), HDP is several times more efficient than lipids and HRPII in converting haem into haemozoin. HDP has no homology with known haem proteins but the C-terminal region is homologous to fasciclin-1 (Nakatani *et al*., 2014).

1.2.1 Haem Detoxification Protein

Haem Detoxification Protein is composed of 205 amino acids and is found in the parasite food vacuole, reportedly in a complex known as the degradosequestrome of approximately 200 kDa consisting of HDP, falcipain-2/falcipain-2', plasmepsin II, plasmepsin IV, and histo-aspartic protease (Gupta *et al*., 2017; Jani *et al*., 2008). The detailed molecular mechanism of haemozoin formation is yet to be determined although four histidine residues have been shown to be critical for haem binding: His122, His172, His175, and His197 (Nakatani *et al*., 2014). A proposed mechanism (Figure 1.2) suggests that two haem molecules bind to His122 and His175 in HDP and His172 and His197 properly align haem for the reaction, while reciprocal iron-oxygen bonds are subsequently formed between the two molecules (Nakatani *et al*., 2014). Thereafter, HDP releases the haem dimer which then serves as a seed crystal for haemozoin growth, possibly by interaction with lipids (Nakatani *et al*., 2014). Haem binds to HDP with high affinity (K_d = 80 nM) and HDP orthologs have been identified in seven other *Plasmodium* species, suggesting its functionality is conserved across *Plasmodium* genes (Jani *et al*., 2008).

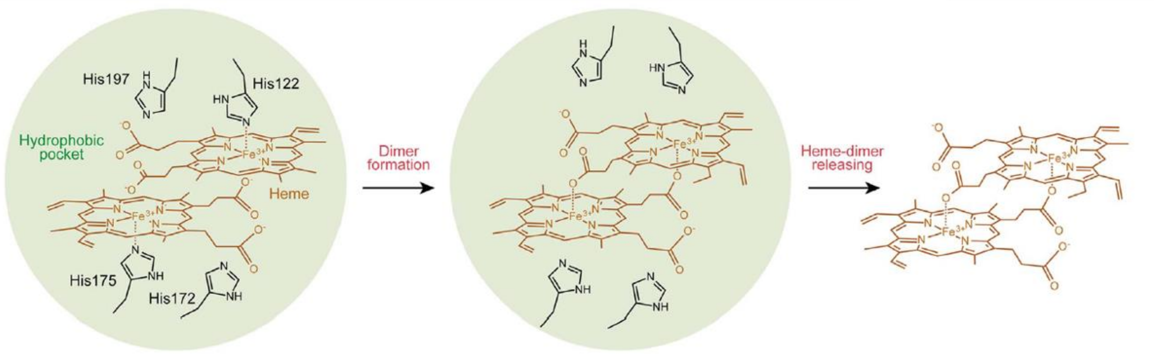


Figure 1.2: Suggested mechanism of haemozoin formation (Nakatani *et al*., 2014).

Limited research has been conducted on HDP and its role in haem detoxification. Jani *et al*. (2008) were the first to conduct extensive characterisation studies on *Pf*HDP and reported that HDP has a high affinity for haem (Kd = 80 nM) with 2.7:1 binding stoichiometry, and is up to 2000 times more efficient than neutral lipids in converting haem to haemozoin. A haemozoin formation assay was used to determine whether the protein converts haem to haemozoin and the optimal pH for HDP-mediated haemozoin production was reported to be 5.2 (similar to the *in vivo* conditions of the food vacuole, pH 4.5-5.2; Jani *et al*., 2008). Their studies also revealed the high thermostability of HDP, which shows activity at temperatures as high as 94 °C and attempts to mutate the HDP gene locus were unsuccessful suggesting the protein is an essential *Plasmodium* protein (Jani *et al*., 2008). Additionally, Jani *et al* (2008) observed in the ring stage of the parasite lifecycle, HDP was secreted into the host cell cytosol and accumulated in the cytosol. As the parasite developed, HDP was trafficked together with the host haemoglobin to the food vacuole via the cytostome-mediated pathway.

Further characterisation studies conducted by Nakatani *et al*. (2013 and 2014) suggested that the binding stoichiometry of untagged HDP to haem was 2:1 and thus differed from the binding stoichiometry of polyhistidine (His)-tagged HDP to haem (2.7:1), which suggested that the histidine tag may interact with haem. Although the presence of the His-tag did not change the haemozoin formation capability of HDP, the C-terminal histidine tag on His-tagged HDP appeared to influence the electronic absorption spectrum of HDP-bound haem when compared to the absorption spectrum of haem bound to untagged HDP (Nakatani *et al*., 2013). Single-point mutations suggested four of the nine histidine residues found in HDP are involved in haemozoin formation (H122, H172, H175, and H197) as their mutations (H→A) reduced haemozoin formation by approximately 50% (Nakatani *et al*., 2014).

Chugh *et al*. (2013) observed that the addition of lipids extracted from lysed parasites or mono-oleoyl glycerol had an additive effect on HDP-mediated haemozoin formation and also reported that falcipain-2 co-exists with HDP to form a haemozoin formation complex in the parasite's food vacuole, suggesting falcipain-2 may be a major haemoglobinase involved in haemoglobin degradation. An alternative haemozoin formation mechanism suggested by Chugh *et al*. (2013) involves HDP in complex with falcipain-2, promoting the nucleation of haemozoin crystals from haem released by the protease action of falcipain-2 on haemoglobin, which may then interact with lipid nanospheres resulting in further growth and correct alignment of the haemozoin crystals.

Gupta *et al*. (2017) used *in silico* and biochemical approaches to identify two possible haem binding sites and a possible haemoglobin binding site in *Pf*HDP (Figure 1.3), revealing that HDP may have a role in trafficking haemoglobin to the food vacuole. The haem binding sites (aa 171-181 and aa 191-200) were identified by aligning the *Pf*HDP sequence to *Pf*HRPII, which contains two histidine-rich binding sequences and the haemoglobin-binding domain (aa 154-172) was identified by aligning the *Pf*HDP sequence to the 14-amino acid haemoglobin-binding loop of falcipain-2 (Gupta *et al*., 2017). As previously mentioned, the C-terminal domain of HDP (aa 60-204) which contains both haem binding sites and the haemoglobin binding domain -is homologous to fasciclin-1. Fasciclin-1 domains often occur in tandem arrangement and have multiple surfaces which enable different ligands to bind (Seifert, 2018). Studies on fasciclin-1 proteins in fungi, archaea and eubacteria suggest they play a role in pathogenicity and symbiosis.

The identified haem binding sites contained 3 of the 4 histidine residues reported as essential for haem to haemozoin activity in *Pf*HDP (Nakatani *et al*., 2014). Deletion mutants generated by Gupta *et* al. (2017) and their haemozoin-formation capabilities were compared to the activity of wild-type (WT) *Pf*HDP as shown in Tables 1.3 and 1.4.

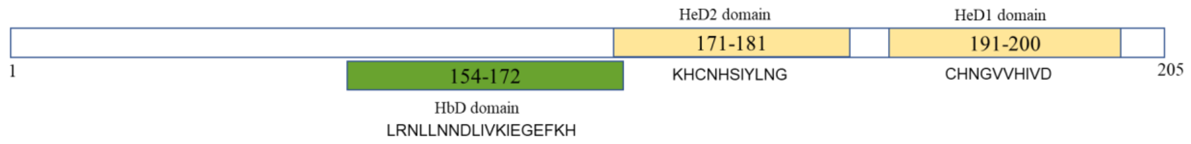


Figure 1.3: Possible *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) haem-binding sites (HeD2 and HeD1) and haemoglobin-binding site (HbD) (Adapted from Gupta *et al.*, 2017).

Table 1.3: Generated deletion mutants (adapted from Gupta *et al.*, 2017).

Mutant	Deletion
*Pf*HDP$_{\text{HeD1}}$	aa 191-205 (lacking HeD$_1$ domain)
*Pf*HDP$_{\text{HeD2}}$	aa 171-205 (lacking HeD$_1$ and HeD$_2$ domains)
*Pf*HDP$_{\text{(HbD)}}$	aa 154-170 (lacking HbD domain)
*Pf*HDP-N	aa 120-205 (C-terminus truncated)
*Pf*HDP-C	aa 1-87 (N-terminus truncated

*Pf*HDP *Plasmodium falciparum* haem detoxification protein; (HeDx) Haem binding domain; HbD, haemoglobin binding domain

Table 1.4: Haemozoin-formation activity of wild-type *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) and generated deletion mutants (adapted from Gupta *et al.*, 2017).

Mutant	Approx. % haem converted to haemozoin	Approx. % haem converted to haemozoin in the presence of falcipain-2
$_{\text{WT}}$*Pf*HDP	50	70
*Pf*HDP $_{\text{HeD1}}$	35	63
*Pf*HDP $_{\text{HeD2}}$	20	50
*Pf*HDP $_{\text{(HbD)}}$	48	42
*Pf*HDP-N	12.5	-
*Pf*HDP-C	14.5	-

The study by Gupta *et al*. (2017) showed increased haemozoin formation in the presence of falcipain-2, which is similar to the findings reported by Chugh *et al*. (2013). Haemozoin formation was still observed upon deletion of the haem-binding domains, suggesting that HDP may have other haem-binding regions or calls into question the actual role of this protein (Gupta *et al*., 2017). Since the *Pf*HDP-C construct contained all domains reportedly involved in haemozoin formation, the study hypothesised that the decrease in haemozoin formation observed in the *Pf*HDP-C mutant may be due to the hydrophobic core required by haem for binding to HDP. The *Pf*HDP-C protein may not be able to provide this on its own, thus indicating the protein must be folded properly for activity (Gupta *et al*., 2017).

Extensive studies by Roxanne Mohunlal (PhD thesis 2019, University of Cape Town [UCT]) optimised expression of soluble and refolded recombinant HDP in *E. coli* C41 cells using the pColdI expression vector. Expression was carried out at low temperatures in nutrient rich media supplemented with 1 % glucose. Soluble and refolded HDP were observed to have a 1:1 HDP to haem binding stoichiometry and relatively weak binding affinity (K_d) for haem was observed: 1.2 ± 0.5 µM and 0.35 ± 0.04 µM for soluble and refolded HDP respectively. Soluble and refolded HDP were also shown to differ in their secondary and tertiary structure. Following the published β-haematin formation assay conditions (Sullivan *et al*., 1996;), the study found that the addition of SDS (sodium dodecyl sulfate) was responsible for promoting β-haematin formation and soluble and refolded HDP was incapable of catalysing β-haematin formation in the absence of SDS. Knockout studies have shown HDP is essential for parasite survival thus elucidating that its role in *Plasmodium* could lead to a potential new drug target.

1.3 Kinases

Kinases are enzymes that catalyse the transfer of the γ-phosphoryl group of ATP onto a target protein in a process known as phosphorylation (Shchenelinin *et al*., 2006). Phosphorylation mediates most signal transduction pathways and is responsible for the regulation of various cellular activities such as transcription, proliferation, and differentiation (Wu *et al*., 2015). Kinases regulate various cellular activities and as a result, mutations and dysregulation of kinases has been shown to play a role in human diseases. These enzymes have thus been studied as possible drug targets over the last 20 years (Roskoski, 2015).

According to literature, the *Plasmodium* kinome consists of 85 protein kinases and protein kinase-related proteins (or 99 depending on the inclusion of borderline sequences) (Ward *et al*., 2004; Srinivasan and Krupa, 2004). Many kinases in the *Plasmodium* kinome can be grouped into the typical eukaryotic protein kinase (ePK) classes, AGC (cyclic-nucleotide- and calcium/phospholipid-dependent kinases), CK1 (casein kinases), CMGC (cyclin-dependent- (CDK), mitogen-activated- (MAPK), glycogen-synthase- (GSK) and CDK-like kinases), CaMK (calmodulin-dependent kinases) and TKL (tyrosine kinase-like kinases) (Talevich *et al*., 2012). Further study of the *Plasmodium* kinome found several kinases did not fall within any of the known ePK groups and a novel family of 20 ePK-related kinases known as FIKKs named after a conserved phenylalanine-isoleucine-lysine-lysine motif which is found only in

Apicomplexa (Ward *et al*., 2004). The kinases in the *Plasmodium* kinome are highly conserved among the different *Plasmodium* species, a significant number do not have a human homologue and 36 have been identified as potentially essential for the asexual blood stage (Mustière, Vanelle and Primas, 2020). There are key differences between the *Plasmodium* kinome and the human kinome: the human kinome consists of 538 kinases while the *Plasmodium* kinome consists of 99 at most, two classes of ePKs found in the human kinome – tyrosine kinases and the STE group (includes enzymes functioning in MAPK pathways) – are absent in the *Plasmodium* kinome and the presence of the FIKK family in the *Plasmodium* kinome (Ward *et al*., 2004; Zhang *et al*., 2021). Enzymes with kinase activity but are unrelated to ePKs at a primary structure level are known as 'atypical protein kinases' which includes RIO (right open reading frame) kinases, phosphoinositide kinases and DNA-dependent protein kinase (Ward *et al*., 2004).

1.3.1 Phosphoinositides

Phosphoinositides (PIPs) are phosphorylated derivatives of phosphatidylinositol (PI) which are generated by several kinases – known as PI kinases (PIKs) (Boura and Nencka, 2015). PI contains a *myo*-inositol head group and consists of an inositol ring (Figure 1.4; which can be phosphorylated at positions 3, 4, and 5) linked to a diacylglycerol backbone (Balla, 2013). These three phosphorylation positions give rise to seven possible PIPs: phosphatidylinositol 3-phosphate (PI3P), phosphatidylinositol 4-phosphate (PI4P), phosphatidylinositol 5-phosphate (PI5P), phosphatidylinositol 3,4-bisphosphate (PI(3,4)P_2), phosphatidylinositol 3,5-bisphosphate (PI(3,5)P_2), phosphatidylinositol 4,5-bisphosphate (PI(4,5)P_2) and phosphatidylinositol 3,4,5-triphosphate (PIP$_3$) (Balla, 2013; McPhail and Burke, 2020). PIPs serve as membrane markers usually in conjunction with organelle-specific proteins involved in signal transduction and are considered essential molecules in cellular signalling (Boura and Nencka, 2015).

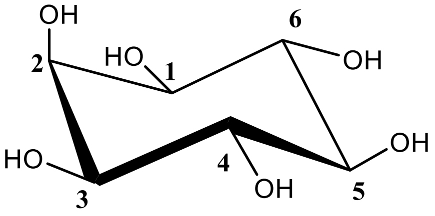


Figure 1.4: Structure of *myo*-inositol.

1.3.2 PIKs

There are 3 types of PIKs: phosphatidylinositol 3-kinases (PI3Ks), phosphatidylinositol 4-kinases (PI4Ks) and phosphatidylinositol phosphate kinases (PIPKs) which can be further classed by their substrate specificity, regulatory binding partners and their distinct phosphorylated PI products (Hassett and Roepe, 2018; Burke, 2018). Figure 1.5 summarises the role PI4K and different classes of PI3Ks and PIPKs play in the PIP synthesis pathway.

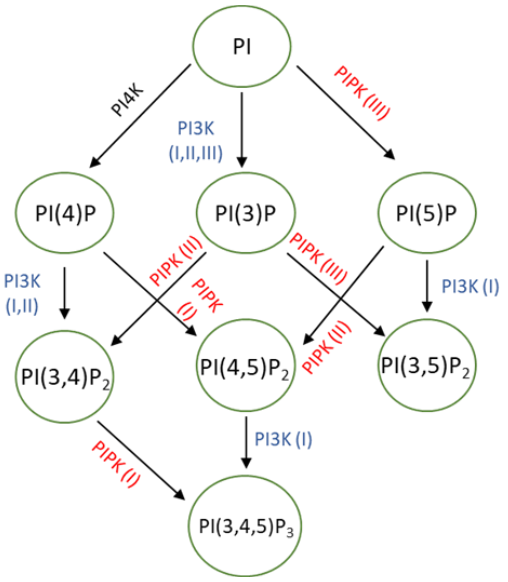


Figure 1.5: Known eukaryotic phosphatidylinositol phosphate metabolic pathways (adapted from Hassett and Roepe, 2018).

Kinase classes in parentheses.

In eukaryotic PIKs, Class I, II and III PI3Ks all phosphorylate PI to produce PI3P (Hassett and Roepe, 2018). Additionally, Class I and II PI3Ks phosphorylate PI4P to produce PI(3,4)P_2 and only class I PI3Ks phosphorylate PI(4,5)P_2 to produce PIP_3 (Hassett and Roepe, 2018). PIPKs are less studied. Class I PIPKs produce PIP_3 and PI(4,5)P_2 by phosphorylating PI(3,4)P_2 and PI4P respectively, class II PIPKs produce PI(3,4)P_2 and PI(4,5)P_2 by phosphorylating PI3P and PI5P respectively and class III PIPKs phosphorylate the 5' position of PI and PI3P resulting in PI5P and PI(3,5)P_2 respectively (Hassett and Roepe, 2018). PI4K has been extensively studied. All eukaryotes possess only two classes of PI4Ks: type II and type III, with type III kinases further divided into type IIIα and type IIIβ isoforms (Balla, 2013). All PI4Ks phosphorylate PI to produce phosphatidylinositol 4-phosphate (PI4P) (Boura and Nencka, 2015).

There are 19 human PIKs which have been extensively studied (review by Burke 2018) and for which 13 structures have been solved either for the human PIK or a closely related orthologue (Burke, 2018). The *P. falciparum* genome only has 7 known PIKs of which only 3 have been studied in any detail (Hasset and Roepe, 2018). Human PIK orthologues can be found for all the eukaryotic PIK classes while *P,falciparum* PIK orthologues only exist for 5 PIK classes, class III PI3K, class I and III PIPK and class II and III PI4K (Arendse *et al*., 2021). *Plasmodium falciparum* PI4K is a clinically validated drug target but the other PIKs have yet to be extensively studied as potential drug targets (Arendse *et al*., 2021). *Plasmodium falciparum* PI3K has been identified as an essential kinase at the blood stage of the parasite lifecycle, is reported to play a role in trafficking to the apicoplast and food vacuole as well as possible involvement in the regulation of autophagy (Tawk *et al*., 2010; Hain and Bosch, 2013). The essentiality of *Pf*PI3K taken together with the strong synergy observed for artemisinin-PI3K drug pairs (Sternberg and Roepe, 2020) indicates *Plasmodium* PI3K is a good potential target for further studies in kinase-based antimalarial drug discovery. However, phenotypic validation of this target has not been reported. Studies have identified compounds that inhibit PI3K in vitro but the whole cell activity of these compounds has not been conclusively linked to *Pf*PI3K inhibition in the parasite and other targets may be implicated.

1.3.3 Phosphatidylinositol 4-Kinase

In eukaryotic cells, phosphorylation of PI to PI4P by PI4Ks primarily occurs in endomembranes such as the Golgi, endosomes, and the *trans*-Golgi network (Balla, 2013). Because of their similarity to phosphatidylinositol 3-kinases (PI3Ks), type III kinases are known as typical PI4Ks whereas type II kinases are known as atypical, as they are distinct from any other lipid kinases (Boura and Nencka, 2015). Type IIIα PI4K is an essential enzyme primarily found in the plasma membrane whereas type IIIβ PI4K is found primarily in the Golgi and *trans*-Golgi network and plays a role in several cell processes: cytokinesis, lipid transport, preserves lysosomal identity and membrane trafficking (Burke, 2018).

Type II PI4Ks have a novel lipid kinase fold and a conserved stretch of cysteines in their catalytic domains which, upon palmitoylation, tightly associate with the membrane and become active (Balla, 2013). The enzyme is mostly associated with the *trans*-Golgi network and endosomes and reportedly plays an additional role in epidermal growth factor receptor (EGFR) trafficking as well as the lysosomal trafficking of glucocerebrosidase (Balla, 2013).

Eukaryotic protein kinases, aPKs and PIKs share a 3D kinase fold and several key structural features (Kanev *et al.,* 2019). Kinase domains have an N-terminal lobe containing β-sheets, a C-terminal lobe containing α-helices, and a connecting hinge region that contains a conserved valine residue (Wu *et al*., 2015). The ATP binding site (Figure 1.6) is located in between the N-terminal lobe and the C-terminal lobe and consists of a hinge region, a ribose binding (front) pocket, a gatekeeper residue and a phosphate binding loop (P-loop) (Fienberg *et al*., 2020). A catalytic lysine/acid pair is found at the end of the P-loop where the terminal phosphate is hydrolysed (Fienberg *et al*., 2020).

Access to the active site is controlled by the DFG (aspartic acid- phenylalanine-glycine) motif as the aspartic acid binds to magnesium ions that coordinate the β- and γ-phosphates of ATP in the ATP-binding pocket (Fabbro, 2014). When the kinase is in a catalytically inactive conformation, the DFG motif is flipped 'out', known as the 'DFGout' conformation (Treiber and Shah, 2013). In the DFGout conformation, the phenylalanine residue occupies the ATP binding pocket and the aspartic acid residue is out of the active site (Modi and Dunbrack, 2019). In the kinase active conformation, the DFG motif is flipped 'in', known as the DFGin conformation (Treiber and Shah, 2013).

Key differences in the ePK and aPK kinase domains include, the glycine-rich sequence motif in the P-loop found in ePKs which is not conserved in the PIK P-loop region, the catalytic lysine which is found in a conserved AxK signature sequence is ePKs sometimes has a larger hydrophobic amino acid in the alanine position of the signature sequence in PIKs and the activation segment – a region found in the C-terminal lobe that contains the DFG motif, phosphorylation sites and a highly conserved APE motif in ePKs- contains a conserved PFxLT motif in PIKs (Arendse *et al*., 2021).

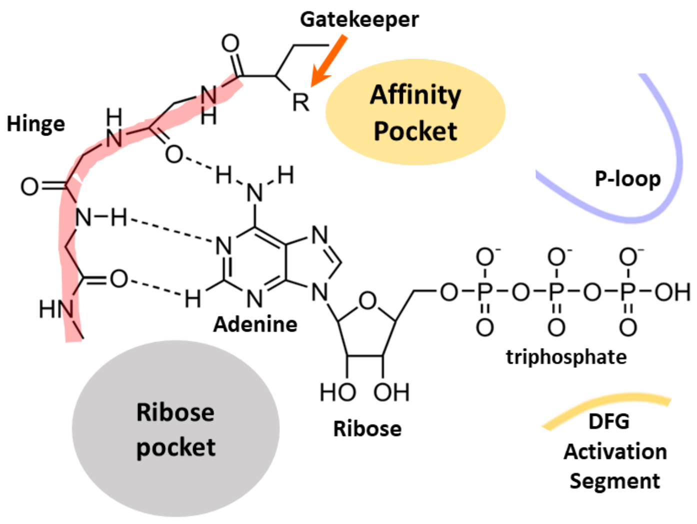


Figure 1.6: 2D schematic of the ATP binding site (adapted from Arendse *et al*., 2021).

Domain organisation of human PI4K type IIIα and type IIIβ (Figure 1.7A) shows both types of the protein consist of a proline rich (Pro-rich) domain, a helical domain, and a kinase domain which consists of the N-terminal (N) lobe and the C- terminal (C) lobe. Studies conducted on yeast orthologues of type IIIα and type IIIβ PI4Ks show type IIIα PI4K supplies the plasma membrane with PI4P and is possibly involved in the synthesis of additional PI pools in the endoplasmic reticular membranes (Nakagawa *et al.*, 1996; Wong *et al.*, 1997). The structure of human type IIIα PI4K determined by cryogenic electronic microscopy (cryoEM) shows a complex form composed of PI4KIIIα bound to TTC7 (tetratricopeptide repeat domain 7) and FAM126 (hyccin protein) regulatory proteins in a dimer (Baskin *et al.*, 2016).

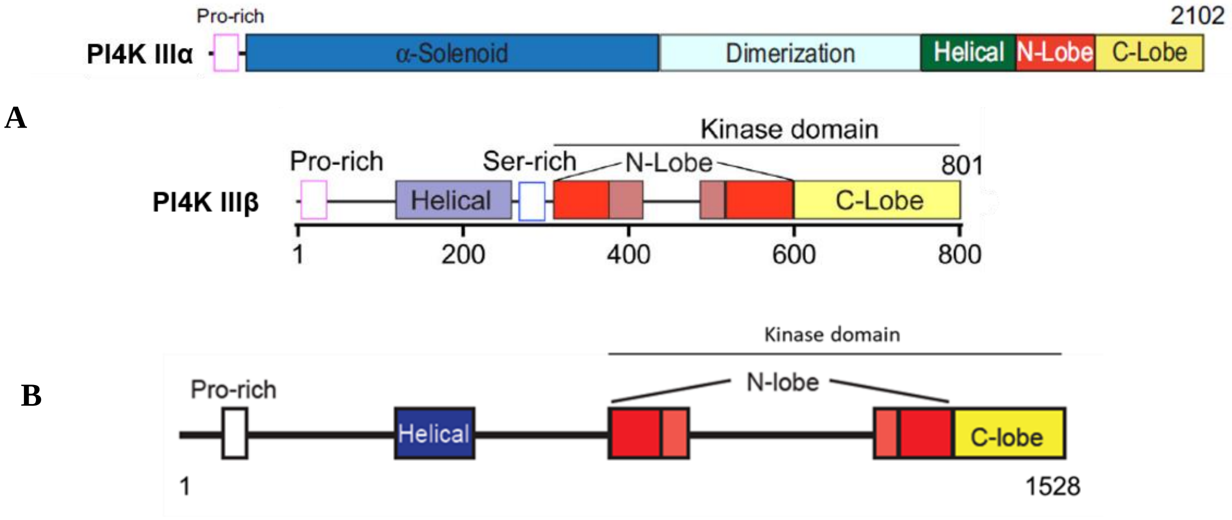


Figure 1.7: (A) Domain organisation of human PI4K type IIIα and type IIIβ (Burke, 2014; Burke, 2018). (B) Domain organisation of Plasmodium falciparum PI4K IIIβ (adapted from McPhail and Burke, 2020).

Type IIIβ PI4K localises to the nucleus and Golgi (controlled by ADP-ribosylation factor 1) and in addition to the catalysis of PI4P formation, regulates trafficking in the late secretory pathway and recruits Rab11 (Boura and Nencka, 2015). In addition to Rab11, PI4KIIIβ also interacts with acyl CoA-binding protein 3 (ACBD3) at the N-terminus and 14-3-3 proteins between the helical and kinase domains (Burke, 2018). The helical and kinase domains of PI4KIIIβ are similar to type IIIα with an extension in the N-lobe of the kinase domain (Burke, 2018). Human PI4KIIIβ has a number of disordered regions - that play an essential role in interacting with protein binding partners - at the N-terminus where ACBD3 binds, between the helical and kinase domains where 14-3-3 proteins bind, within the N-lobe of the kinase domain and at the C-terminus of the C-lobe in the kinase domain (Boura and Nenck, 2015; Burke, 2018).

The domain organisation of *Pv*PI4KIIIβ (Figure 1.7B) showed the protein is significantly larger than human PI4KIIIβ but still contains similar domain organisation, a Pro-rich domain, a helical domain, a N-lobe, a C-lobe and an insertion within the N-lobe of the kinase domain (McPhail and Burke, 2020). As the protein is larger, the disordered regions – at the N-terminus, between the helical and kinase domains and within the N-lobe- in *Pv*PI4KIIIβ are also larger. The C-terminal region of the N-lobe and the entire C-lobe of *Pv*PI4KIIIβ is well conserved with human PI4KIIIβ (62 % similarity) but the N-terminal region of the N-lobe has poor sequence similarity to human PI4KIIIβ and also has a significantly different P-loop (McPhail and Burke, 2020).

1.3.4 Kinase inhibitors

Kinase inhibitors can be classed into two groups: reversible and irreversible inhibitors (Figure 1.8) (Wu *et al*., 2015). Reversible kinase inhibitors can be further classified into five types of inhibitors: type I to V (Wu *et al*., 2015). Irreversible inhibitors are typically composed of a ligand - that binds in the ATP-binding site- and a reactive electrophilic group which covalently binds to a cysteine, lysine, tyrosine or aspartic acid residue proximal to the ATP site, blocking the ATP-binding site of the kinase (Martinez, Defnet and Shapiro, 2020).

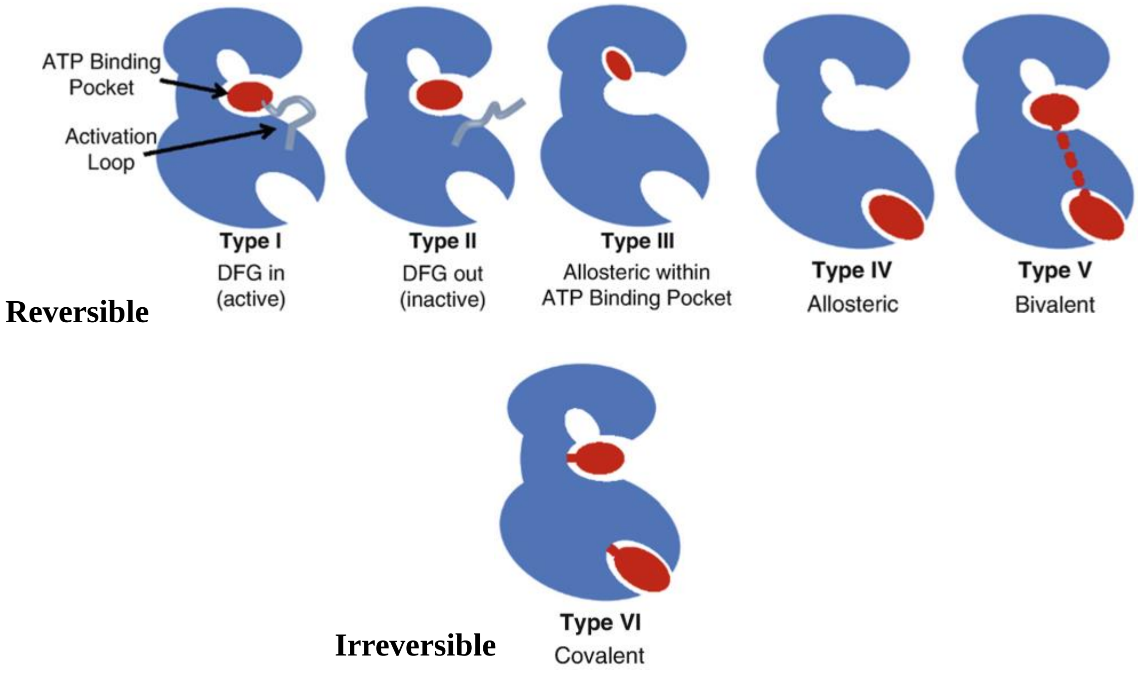


Figure 1.8: Types of kinase inhibitors (Martinez, Defnet and Shapiro, 2020).

Type I inhibitors bind to the ATP binding pocket of the kinase in the active 'DFG-in' conformation (Roskoski, 2016). These inhibitors often bind where ATP binds in the kinase (thus are ATP competitors) and extend into regions surrounding this site such as the P-loop, hydrophobic back pocket, the DFG motif and the front pocket (Zhao and Bourne, 2020). Type I inhibitors often have low selectivity as the ATP binding site is often conserved through the kinome, increasing the possibility of off-target effects (Bhullar *et al*., 2018).

Type II inhibitors are also ATP competitors but bind to the inactive 'DFG-out' conformation of the kinase (Roskoski, 2016). These inhibitors form hydrogen bonds with the DFG motif residues in the DFGout kinase conformation and exploit regions adjacent to the ATP binding site exposed due to the DFGout conformation of the kinase (Wu *et al*., 2015). Type II inhibitors generally display higher selectivity compared to type I inhibitors as they do not affect activated kinases (Bhullar *et al*., 2018).

Type III and type IV inhibitors are allosteric inhibitors and bind to kinases at an allosteric site adjacent to the ATP-binding site, and at an allosteric site isolated from the ATP-binding site, respectively (Wu *et al*., 2015). As type III inhibitors bind to a site adjacent to the ATP binding site, they have reduced promiscuity and a decreased possibility of gatekeeper mutations arising in the ATP binding site (which often results in inhibitor resistance) – two common challenges observed for type I and type II inhibitors (Martinez, Defnet and Shapiro, 2020). Type IV inhibitors alter enzyme activity by preventing the phosphorylation of select downstream substrates by the kinase or by disrupting access to the upstream activators (Martinez, Defnet and Shapiro, 2020). Although type IV inhibitors may potentially block specific kinase functions associated with a particular disease while still allowing other kinase functions to occur, targeting sites outside the ATP binding region to determine which are important for biological functions is difficult (Martinez, Defnet and Shapiro, 2020). Type III and IV kinase inhibitors are non-competitive and uncompetitive ATP inhibitors and have a high degree of kinase selectivity (Bhullar *et al*., 2018).

Inhibitors that target both the ATP-binding site and a different site on the kinase are known as type V inhibitors (Lee, Yeoh and Low, 2022). These inhibitors often consist of a small ATP-binding site targeting molecule coupled to another ligand that binds to a region outside the ATP-binding site (Gower, Chang and Maly, 2014). Type V inhibitors are generally potent and highly selective (Lee, Yeoh and Low, 2022).

Type VI (irreversible) inhibitors bind covalently to the kinase (Roskoski, 2016). Commonly, these inhibitors bind to a catalytic nucleophile cysteine residue in the kinase active site (Lee, Yeoh and Low, 2022). Covalent inhibition is a two-step process. Initially, the compounds form a reversible complex followed by the formation of an irreversible enzyme-inhibitor complex as shown in the equation below where K_I is the reversible binding constant i.e. the affinity of the ligand for the enzyme and K_{inact} is a rate constant describing the maximum rate of covalent bond formation (Baillie, 2016). The efficiency of covalent bond formation is described by the second-order rate constant K_{inact}/K_I (Strelow, 2017).

$$E + I \overset{K_I}{\leftrightarrows} \underset{\text{Non-covalent complex}}{E \bullet I} \xrightarrow{K_{inact}} \underset{\text{Enzyme-Inhibitor complex}}{E - I}$$

Covalent inhibitors are often called ‘targeted covalent inhibitors’ (TCIs) and typically consist of an ATP-binding site targeting ligand with a weakly electrophilic Michael acceptor – also known as a ‘warhead’ (Baillie, 2016; Bhullar *et al.*, 2018). The benefits of targeting kinases with TCIs include high potency and selectivity, prolonged action in patients and low dosage (Baillie, 2016). A major concern of using TCIs is toxicity and the potential formation of toxic covalent drug metabolites e.g. the hepatotoxic cellular metabolites formed during metabolism of acetaminophen (Ghosh *et al.*, 2019). Using TCIs to target kinases requires careful design/selection of the ATP-binding site ligand structure and careful selection of the warhead in order to maximise inhibitor potency, selectivity, duration of action while mitigating toxicity risks (Baillie, 2016).

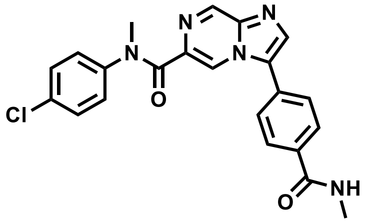
According to the Blue Ridge Institute for Medical Research, there are currently 71 clinically approved protein kinase inhibitors and many more in development (Roskoski, 2022). There are 5 FDA approved PI3K inhibitors, all of which are for cancer treatment (Ayala-Aguilera *et al.*, 2021). The majority of PI3K inhibitors are highly selective which occurs due to inhibitor interactions with the hinge region, affinity pocket and catalytic lysine of HuPI3K (Ayala-Aguilera *et al.*, 2021).

As previously mentioned, the main challenge faced in kinase inhibitor drug discovery is minimising off-target interactions. As most kinase inhibitors are type I inhibitors, high sequence similarity around the ATP-binding pockets of kinases increases the likelihood of activity against other kinases as well as the desired target (Wu *et al*., 2015). Inhibitors developed that target *Plasmodium* PI4K should ideally inhibit PI4K with high potency and selectivity (Fienberg *et al*., 2020).

1.3.5 *Plasmodium* PI4K

As already mentioned, the *P. falciparum* genome has 7 putative PIKs and only one is a clinically validated drug target for malaria – PI4KIIIβ (Arendse *et al*., 2021). *Pf*PI4KIIIβ (which will be referred to as *Pf*PI4K in this thesis) is an essential protein required for membrane ingression when daughter merozoites are generated in the erythrocyte stage of the parasite lifecycle and also regulates intracellular signalling and trafficking (McNamara *et al*., 2013; Phillips *et al*., 2017). *Pf*PI4K's role in the erythrocyte, liver stage and mosquito stages of the parasite lifecycle makes it an attractive malaria drug target (McNamara *et al*., 2013).

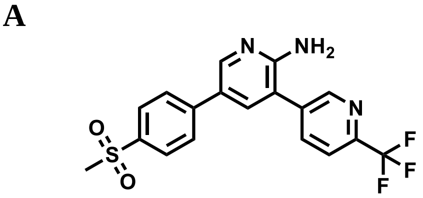
Antiplasmodium imidazopyridazines have been found to target *Plasmodium* PI4K (Boura and Nencka, 2015). A study conducted by McNamara *et al*. (2013) determined that imidazopyridazines target several stages of the parasite lifecycle. Compound KDU691 (Figure 1.9) is active against the liver stage of *P. yoelii*, *P. cynomolgi*, and *P. berghei* and this compound was shown to be potent against blood-stage *P. vivax* and *P. falciparum* field isolates. KDU691 was also shown to reduce gamete formation by 60% as well as completely inhibit transmission to the mosquito at 1 µM (McNamara *et al*., 2013). The study also showed imidazopyridazines bind to the ATP-binding pocket of *Plasmodium* PI4K (McNamara *et al*., 2013).



***Pv*PI4K IC$_{50}$**	**1.5 nM**
***Pf* EC$_{50}$**	118 nM
HuPI4Kβ IC$_{50}$	7.9 µM

Figure 1.9: (A) KDU691 structure. (B) KDU691 inhibition data.

Drug candidate MMV390048 (Figure 1.10) is an aminopyridine compound that targets *Plasmodium* PI4K - by binding to the ATP- binding pocket - and shows activity against early schizonts in the blood stage of the parasite lifecycle as well as potential transmission inhibition (Paquet *et al*., 2017). Additionally, the compound displayed potential as a prophylactic and chemoprotective agent for human malaria infections as it prevented *P. cynomolgi* hypnozoite and schizont development in the liver both *in vitro* and *in vivo* (Paquet *et al*., 2017). *Pf*PI4K was identified as the target of MMV390048 by chemoproteomic pull-down experiments in which the compound competitively inhibited binding of *Pf*PI4K to beads, as well as by single-nucleotide mutations in the *Pf*PI4K gene of MMV390048-resistant parasite strains (Paquet *et al*., 2017).



B

***Pv*PI4K IC$_{50}$**	**3.4 nM**
***Pf*NF54 EC$_{50}$**	28 nM
HuPI4Kβ IC$_{50}$	9.7 µM

Figure 1.10: (A) MMV390048 structure. (B) MMV390048 inhibition data.

MMV390048 had strong potency and favourable pharmacokinetic behaviour in mouse models and maintained a low clearance rate and long half-life upon investigation in various infected species (Paquet *et al*., 2017). The compound is selective for PI4K and does not bind to any other *P. falciparum* kinases, although it does bind to human PIP4K2C, the consequences of which are unknown. Further investigation is thus required to determine the significance of blocking human PIP4K2C and its effect on the host (Paquet *et al*., 2017).

To date, no *Plasmodium* PIK structures have been elucidated. At the amino acid level, orthologues of PI4K (IIIβ) found in *Plasmodium* species are conserved, with 97% similarity observed in the catalytic domain of *P. falciparum* and *P. vivax* orthologues, whereas human PI4KIIIβ (which will be referred to as HuPI4K for the remainder of this thesis) only shows 43% similarity to the catalytic region (McNamara *et al*., 2013). *Plasmodium* PI4K is much larger than HuPI4K, has a larger regulatory domain and has extended loop regions within the kinase domain which are not found in HuPI4K (Fienberg *et al*., 2020). These differences make *Plasmodium* PI4K difficult to express recombinantly. Expression is made more difficult for *Pf*PI4K due to the AT-rich sequence (79.3% in *Pf*PI4K, 57.3 % in *Pv*PI4K), and long asparagine repeats typical in *P.falciparum* which form low-complexity regions that tend to form loops (Muralidharan and Goldberg, 2013). In *P.falciparum*, these repeats are in approximately 30 % of the proteome but are rare in other *Plasmodium* species (Muralidharan and Goldberg, 2013). Full length PvPI4K has been successfully expressed recombinantly (McNamara *et al*., 2013).

A *Pf*PI4K homology model (Figure 1.11) was built using the HuPI4K crystal structure (PDB ID: 4D0L) as a template (Burke *et al*., 2014; Fienberg *et al*., 2020). Human PI4K and *Pf*PI4K share 48 % sequence identity. The homology model can be used to evaluate the binding modes of *Pf*PI4K inhibitors and rationalise interactions between the inhibitor and the enzyme hinge region, catalytic site and front and back pockets in relation to the observed antiplasmodium activity (Fienberg *et al*., 2020).

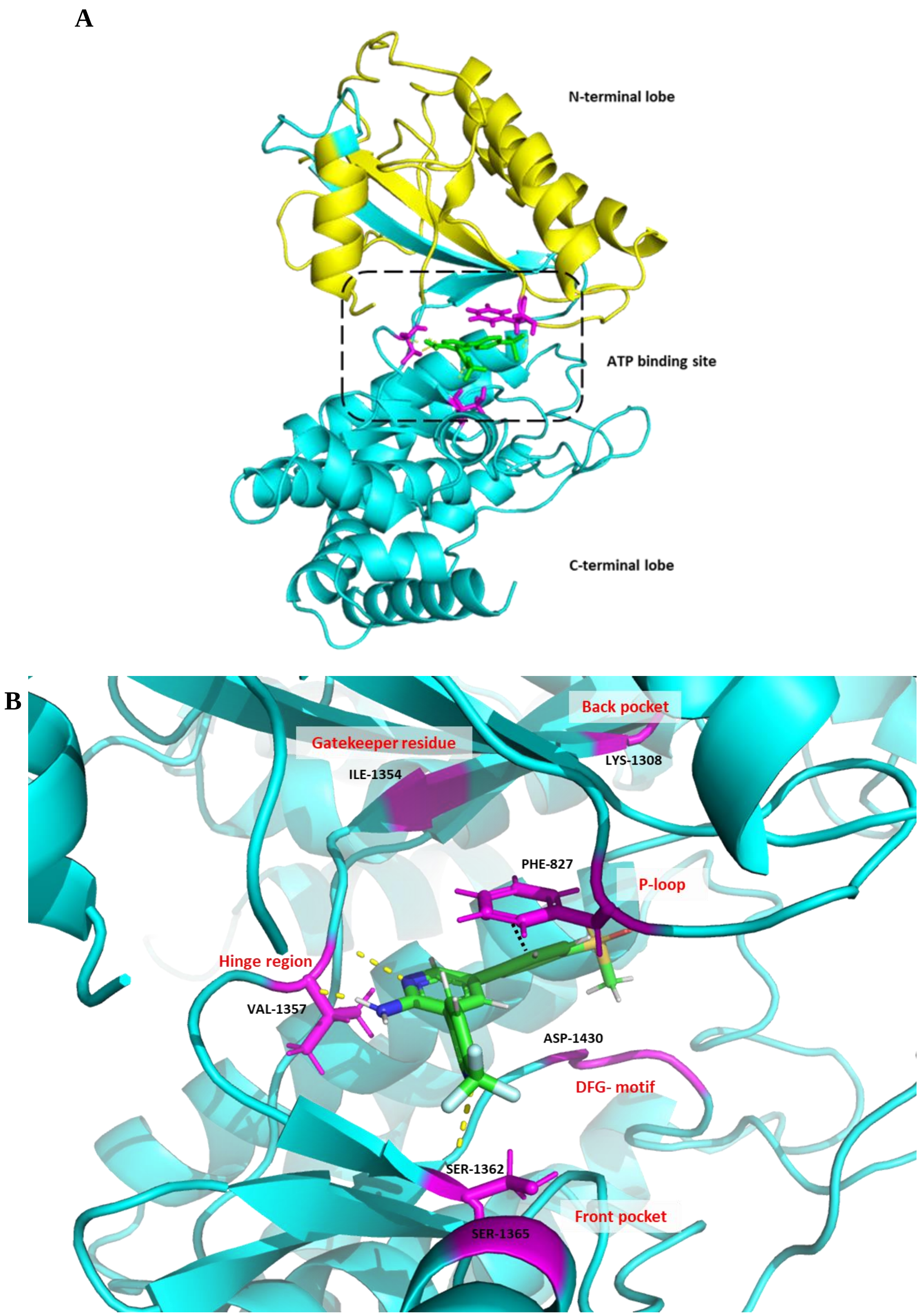


Figure 1.11: *Plasmodium falciparum* phosphatidylinositol 4-kinase homology model. *Plasmodium falciparum* phosphatidylinositol 4-kinase homology model generated by Stephen Fienberg (UCT). (**A**) *Plasmodium* PI4K kinase domain. N-terminal lobe highlighted in yellow, C-terminal lobe highlighted in cyan, ATP-binding site shown in the box. (**B**) *Plasmodium* PI4K ATP-binding site. Key features shown in magenta and labelled in red. Yellow lines indicate hydrogen bonds, black line indicates Pi-Pi stacking.

MMV390048 forms 2 hydrogen bonds with the conserved hinge region valineV1357), the phenyl sulfone forms a hydrogen bond with K1308 in the affinity (back) pocket of the enzyme and the trifluoromethyl-pyridyl group sits in the ribose (front) pocket (Fienberg *et al*., 2020).

As previously mentioned, *Plasmodium* PI4K is a clinically validated drug target which is involved in multiple stages of the parasite lifecycle. The *Pf*PI4K homology model and previous *Plasmodium* PI4K inhibitor studies indicate various possible inhibitor-kinase interactions which can be exploited to improve inhibitor selectivity and potency and also potential new interactions that can be probed in the search for new PI4K targeting inhibitors.

1.4 Aims

This thesis will focus on two essential *Plasmodium* proteins: HDP (Section A) and PI4K (Section B).

Section A

Previous research conducted by Roxanne Mohunlal (PhD Thesis 2019, UCT) has shown HDP binds to haem in a 1:1 stoichiometric ratio, which differs from results reported in previous studies (Jani *et al*., 2008; Nakatani *et al*., 2014). Furthermore, both soluble and insoluble HDP used in the haemozoin-formation assay have been found to mediate β-haematin formation. However, further investigation has suggested β-haematin formation may be mediated by sodium dodecyl sulphate (SDS) used in the haemozoin-formation assay, thus raising questions regarding the role of HDP if it does not in catalyse haemozoin formation, as label-free semi-quantitative proteomic analysis (Roxanne Mohunlal, UCT) has shown its entrapment by haemozoin. The aim of this research project was to optimise expression and purification of soluble *Pf*HDP and characterise the protein’s role in haemozoin formation.

Section B

Increasing resistance to artemisinin drives the need for new antimalarials which ideally have novel modes of action, could form part of a single dose combination regimen and act on multiple stages of the parasite lifecycle. *Pf*PI4K is a clinically validated drug target which was found to be an essential *Plasmodium* kinase involved in multiple stages of the parasite lifecycle. Docking of known PI4K inhibitors in a *Pf*PI4K homology model has shown specific kinase residues form hydrogen bonds and Pi-Pi stacking interactions with inhibitors which may affect inhibitor potency and selectivity.

There is no known structure for *Pf*PI4K as the protein is difficult to express and crystallise, however, *Pv*PI4K has been expressed in a baculovirus-insect cell expression system thus *Pv*PI4K was used in this study. The aim of this project was to use *Pv*PI4K inhibition assays coupled with site-directed mutagenesis and mass spectrometry to probe key inhibitor- protein interactions within the ATP binding site in order to support target-based malaria drug discovery programs focused on the development of both ATP-competitive and covalent *Plasmodium* PI4Kβ inhibitors.

Section A

Haem Detoxification Protein

Chapter 2

Expression and Purification of Haem Detoxification Protein

2.1 Introduction

Protein production for biochemical analysis and structural studies is dependent on the success of three factors: protein expression, protein solubility and protein purification (Esposito and Chatterjee, 2006). A number of expression hosts are available for recombinant protein expression including mammalian cells, insect cells and yeast, but the *Escherichia coli* (*E. coli*) expression system is the most widely used due to its ease of use, speed and low cost (Braun and LaBaer, 2003). Some proteins do not express in soluble form possibly due to protein misfolding, lack of post-translational modifications or precipitation through the formation of inclusion bodies (Gräslund *et al*., 2008). To address this, changes of the expression conditions such as temperature, inducer concentration, expression media and addition of molecular chaperones and folding modulators can be used to improve protein solubility (Esposito and Chatterjee, 2006).

Another method used to enhance protein solubility is the addition of affinity tags which also significantly increase the efficiency of protein purification (Stevens, 2000). Peptide tags (such as the cost-effective poly-histidine tag) allow for one step purification, improved solubility and a method of tracking the recombinant protein through the expression and purification process (Rosano and Ceccarelli, 2004). An additional advantage of using the poly-histidine affinity tag is that its small size and charge ensure protein activity and structure is rarely affected, thus the tag typically does not need to be removed after purification (Gräslund *et al*., 2008).

Immobilized Metal Affinity Chromatography (IMAC) is a type of affinity chromatography that separates proteins from a sample based on their affinity for metal ions (Bornhorst and Falke, 2000). His-tagged proteins are purified based on the histidine residue’s affinity for metal ions such as Ni^{2+} and Cu^{2+} immobilized on a chromatographic matrix by a chelating ligand, commonly nitriloacetic acid (NTA) Figure 2.1 (Spriestersbach *et al*., 2015).

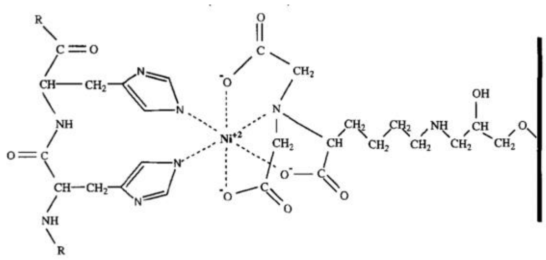


Figure 2.1: Interaction between a poly-histidine tag and nickel-NTA (Bornhorst and Falke, 2000).

The nickel-NTA matrix has a high affinity for the poly-histidine tag and the protein can be eluted by the use of mild buffer conditions and imidazole (Bornhorst and Falke, 2000).

Major improvements in the last few decades have ensured protein expression and purification are no longer major limiting steps in protein production, however, soluble protein production for purification can be a considerable hindrance to the process (Esposito and Chatterjee, 2006).

Recombinant expression of *P. falciparum* proteins in *E. coli* is often challenging because *P. falciparum* has an A-T rich genome (approximately 80%) and the codons preferentially used by the parasite for protein expression are rarely used by *E. coli* (Baca and Hol, 2000). The *P. falciparum* haem detoxification protein (HDP) expression construct has an A-T content of 58 %. Previous recombinant expression of *Pf*HDP has used a number of different expression vectors under varying conditions (Table 2.1), yet all protein produced has been localised to inclusion bodies (Jani *et al*., 2008; Nakatani *et al*., 2013; Nakatani *et al*., 2014 and Gupta *et al*., 2017).

Table 2.1: Previous HDP expression vectors and conditions (adapted from Jani *et al*., 2008; Nakatani *et al*., 2013; Nakatani *et al*., 2014 and Gupta *et al*., 2017).

Vector	Tag	Expression cells	Growth conditions
pET101	C-terminal his-tag	BL21	N/A*
pCOLD IV**	None	BL21 (DE3)	15 °C - overnight
pCOLD I	N-terminal his-tag	BL21 (DE3)	15 °C - overnight
pET-28a	N-terminal his-tag	BL21	37 °C - 4 hours
pCOLD I***	N-terminal his-tag	C41	15 °C - 48 hours

*conditions not reported; ** constructs confirmed to be codon optimised; *** Roxanne Mohunlal PhD Thesis, UCT 2019

Soluble recombinant protein is considered an important requirement for proper analysis of biological function and is less time consuming than purifying and refolding from inclusion bodies (Bhatwa *et al.*, 2021; Sorensen and Mortensen, 2005).

This chapter focuses on optimisation of the expression and purification of soluble *P. falciparum* HDP. The following *E. coli* expression systems were used: ArcticExpress cells, C41 cells and BL21 cells. Challenges in purifying soluble HDP and low protein yields obtained resulted in a switch to purification under denaturing conditions.

2.2 Methods

Plasmodium falciparum HDP (Appendix Figure A1) cloned into pColdI (Appendix Figure A2) at NdeI and BamHI in the multiple cloning site was the vector used for all *Pf*HDP expression (Roxanne Mohunlal PhD Thesis, UCT 2019). The construct was codon optimised for expression in *E.coli*.

2.2.1 Transformation of ArcticExpress cells

pColdI-*Pf*HDP was transformed into ArcticExpress competent *Escherichia coli* cells (Agilent Technologies). The competent cells were thawed on ice and 1 µL plasmid DNA was added to 100 µL competent cells. The cells were mixed gently and incubated on ice for 30 min. The cells were then heat-pulsed at 42 ºC for 23 s and incubated on ice for a further 2 min. Pre-heated Luria-Bertani broth (10 g/L tryptone, 10 g/L NaCl, and 5 g/L yeast extract) without antibiotic was added to the transformation reaction and the cells were incubated at 37 ºC for 1 h with shaking at 180-200 rpm. The transformation culture was plated onto LB-agar plates containing 100 µg/mL ampicillin. The plates were incubated overnight at 37 ºC in an inverted position.

A single transformation colony was inoculated into LB broth containing 100 µg/mL ampicillin and incubated with shaking overnight at 37 ºC. Sterile glycerol was added to 850 µL of the overnight culture to a final concentration of 17% (v/v) and the culture was stored at -80 °C.

2.2.2 Small-scale protein expression

A scraping of the expression construct glycerol stock was inoculated into 10 mL LB broth containing 100 µg/mL ampicillin and 20 µg/mL gentamycin and 100 µg/mL ampicillin for ArcticExpress cells and BL21 cells respectively. The culture was incubated with shaking at 37 ºC overnight. The overnight culture was then inoculated into 100 mL LB broth containing 1%

(w/v) glucose and no antibiotic and 1% (w/v) glucose and 100 µg/mL ampicillin for ArcticExpress cells and BL21 cells respectively. Cultures containing no glucose were also set up. The cultures were incubated at 37 ºC with shaking until the optical density measured at 600 nm (OD_{600}) using a Helios γ instrument (Thermo Fisher Scientific) was 0.6-0.8. A 1-mL aliquot was collected from the culture, pelleted by centrifugation at 4000 × *g* for 3 min in a Centrifuge 5415D (Eppendorf). The supernatant was discarded and the pellet was resuspended in (50 x OD_{600}) µL 4× SDS sample buffer (10 mL stacking buffer, 8 mL glycerol, 0.8 g SDS, 0.8 mL 2-mercapto-ethanol, 0.2 mg bromophenol blue, 1.2 mL water) in a 3:1 sample:sample buffer ratio. The SDS sample was heated at 100 ℃ for 5 min and stored at -20 ℃ until analysis via SDS-polyacrylamide gel electrophoresis (SDS-PAGE; uninduced sample). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM and the culture was incubated with shaking at 11 ºC for a further 24 h. After expression, 1 mL was collected from the culture and pelleted by centrifugation at 4000 × *g* for 3 min in a Centrifuge 5415D (Eppendorf). The supernatant was discarded and the pellet was resuspended in (50 x OD_{600}) µL 4× SDS sample buffer. The SDS sample was heated at 100 ℃ for 5 min and stored at -20 ℃ until analysis via SDS-PAGE (induced sample). The cells were harvested by centrifugation with a F0630 rotor at 4 ºC and centrifuged at 4000 × *g* for 20 min using an Allegra X-30R Centrifuge (Beckman Coulter). The bacterial pellet was stored at -20 ºC overnight. The pellet was thawed on ice, resuspended in 5 mL 50 mM Tris-500mM NaCl (pH 8.0), and sonicated for 8 cycles with a microtip at power 4 (15 s on, 15 s off) using a Misonix Sonicator 3000 (Cole-Parmer). The expression samples were centrifuged at 20,000 × *g* for 20 min and the soluble (supernatant) and insoluble (pellet) fractions were separated. A 60-µL aliquot was collected and mixed with 4×SDS sample buffer in a 3:1 sample:sample buffer ratio. The samples were analysed via SDS-PAGE (section 2.2.3).

2.2.3 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein expression samples were analysed via 10% SDS-PAGE according to the protocol described by Laemmli (1970). The gel comprised 10% (w/v) resolving gel (3.33 mL 40% (w/v) acrylamide, 2.5 mL lower gel buffer (1.5 M Tris-HCl, 0.4% (w/v) SDS, pH 8.8), 4.16 mL water, 35 µL 10% (w/v) ammonium persulfate, and 7 µL TEMED, tetramethylethylenediamine) or 12% (w/v) resolving gel (4 mL 40% (w/v) acrylamide and 3.5 mL water substitution) and 4% stacking gel (0.7 mL 40% (w/v) acrylamide, 1.25 mL stacking gel buffer (0.5 M Tris-HCl, 0.4% (w/v) SDS, pH 6.8), 3 mL water, 25 µL 10% (w/v) ammonium persulfate, and 6 µL TEMED). The soluble and insoluble fractions were mixed with 4× SDS sample buffer in a 3:1 sample:sample buffer ratio. The samples were heated at 100 ºC for 5 min and loaded onto the SDS-PAGE gel alongside a Color Prestained Protein Standard (NEB). Electrophoresis was conducted using a Mini-PROTEAN® 3 Cell (Bio-Rad) in SDS running buffer (25 mM Tris-HCl, 0.2 M glycine, and 3.5 mM SDS) at 120 V constant voltage for 1.5 h. The gel was stained with Coomassie stain (45% (v/v) water, 45% (v/v) methanol, 10% (v/v) acetic acid, and 0.25% (w/v) Coomassie Brilliant Blue R-250) overnight at room temperature with gentle shaking (Orbital shaker SH30). The gel was destained with destain solution (44% (v/v) water, 44% (v/v) methanol, and 11% (v/v) acetic acid) at room temperature with gentle shaking. The destain solution was changed every 15 min until the protein bands were visible. Alternatively, the gel was stained with PageBlue™ (Thermo Fisher Scientific) using a microwave procedure in which the gel was microwaved in ultrapure water for 60 s, followed by gentle shaking for 4 min for 3 cycles. The gel was microwaved in PageBlue™ stain for 25 s and stained for 20 min with gentle shaking before rinsing with ultrapure water and washing twice with ultrapure water for 5 min. The gel was then photographed using a camera.

2.2.4 Large-scale protein expression

A large-scale expression culture of pColdI-*Pf*HDP was produced for protein purification. A scraping of the expression construct glycerol stock was inoculated into 10 mL LB broth containing 100 µg/mL ampicillin and 20 µg/mL gentamycin. The culture was incubated with shaking overnight at 37 ºC. The overnight culture was inoculated into 1 L LB broth containing 1% glucose and no antibiotic. The culture was incubated with shaking at 37 ºC until the OD_{600} measured using a Helios γ instrument was 0.6-0.8. A 1-mL aliquot was collected from the culture and pelleted by centrifugation at 4000 × *g* for 3 min in a Centrifuge 5415D. The supernatant was discarded and the pellet was resuspended in 50 x OD_{600} µL 4× SDS sample buffer. The SDS sample was heated at 100 °C for 5 min and stored at -20 °C until analysis via SDS-PAGE (uninduced sample). IPTG was added to a final concentration of 0.5 mM and the culture was incubated with shaking at 11 ºC for 24 h (30 h after optimisation).

In further optimisation experiments, a large-scale expression culture of pColdI-*Pf*HDP ArcticExpress cells was produced for protein purification. A scraping of the expression construct glycerol stock was inoculated into 10 mL LB broth containing) 100 µg/mL ampicillin and 20 µg/mL gentamycin. The culture was incubated with shaking overnight at 37 ºC. The overnight culture was inoculated into 1 L Terrific Broth (TB) media containing 1% glucose, 2.5 mg 5-aminolevulinic acid hydrochloride and no antibiotic. The culture was incubated with shaking at 30 °C until the OD_{600} measured using a Helios γ instrument reached 0.6–0.8. A 1-mL aliquot was collected from the culture and pelleted by centrifugation at 4000 × *g* for 3 min in a Centrifuge 5415D. The supernatant was discarded and the pellet was resuspended in 50 × OD_{600} µL 4× SDS sample buffer. The SDS sample was heated at 100 °C for 5 min and stored at -20 °C until analysis via SDS-PAGE (uninduced sample). IPTG was added to a final concentration of 0.5 mM and the culture was incubated with shaking at 11 ºC for 30 h.

In later experiments, a large-scale expression culture of pColdI-*Pf*HDP in (A) C41 cells and (B) BL21 cells was produced for protein purification. A scraping of the expression construct glycerol stock was inoculated into 10 mL LB broth containing 100 µg/mL ampicillin. The culture was incubated with shaking overnight at 37 ºC. The overnight culture was inoculated into (A) 1 L TB medium containing 1% glucose and 100 µg/mL ampicillin and (B) 250 mL LB broth containing 1% glucose and 100 µg/mL ampicillin. The culture was incubated with shaking at 37 ºC until the OD_{600} measured using a Helios γ instrument reached 0.6–0.8. A 1-mL aliquot was collected from the culture and pelleted via centrifugation at 4000 × *g* for 3 min in a Centrifuge 5415D. The supernatant was discarded and the pellet was resuspended in 50 × OD_{600} µL 4× SDS sample buffer. The SDS sample was heated at 100 °C for 5 min and stored at -20 °C until analysis via SDS-PAGE (uninduced sample). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM and the culture was incubated with shaking at (A) 11 ºC for 48 h or (B) 15 °C for 24 h.

After expression, 1 mL was collected from the culture, pelleted by centrifugation at 4000 × *g* for 3 min in a Centrifuge 5415D. The supernatant was discarded and the pellet was resuspended in 50 × OD_{600} µL 4× SDS sample buffer. The SDS sample was heated at 100 °C for 5 min and stored at -20 °C until analysis via SDS-PAGE (induced sample). The cells were harvested by centrifugation in a SX4400 rotor at 4 ºC and 4000 × *g* for 20 min using an Allegra X-30R centrifuge. The bacterial pellet was stored at -20 ºC overnight.

2.2.5 His-tag affinity chromatography

The bacterial pellet was thawed on ice and resuspended in 40 mL 50 mM Tris-500mM NaCl (pH 8.0) and sonicated for 8 cycles with a flat tip at power 6 (15 s on, 15 s off) using a Misonix Sonicator 3000 (Cole-Parmer). The expression samples were centrifuged at 20,000 × *g* for 20 min and the soluble (supernatant) and insoluble (pellet) fractions were separated. A 60-µL aliquot was collected and mixed with 4× SDS sample buffer in a 3:1 sample:sample buffer ratio. The samples were then analysed via SDS-PAGE (section 2.2.3). HDP was purified from the soluble fraction (filtered through a 0.45-µm filter) using a HiTrap™ 5 mL HP column (GE Healthcare) and the AKTA explorer purification system (GE Healthcare) with fraction collection. The HiTrap™ HP column was equilibrated with five column volumes of equilibration buffer (20 mM $NaHPO_4$, 0.5 M NaCl, 10% glycerol, pH 7.4) and the soluble fraction was loaded onto the column. The column was washed with equilibration buffer and a low pH wash buffer (20 mM $NaHPO_4$, 0.5 M NaCl, 10% glycerol, pH 6.5). The bound protein was eluted from the column using elution buffer (20 mM $NaHPO_4$, 0.5 M NaCl, 500 mM imidazole, 10% glycerol, pH 7.4) via step-wise elution with 2.5, 7.5, 50, and 100% imidazole. Aliquots (60 µL) of the sample flow-through, column wash fractions, and elution fractions were collected and 4× SDS sample buffer was added at a 3:1 sample:sample buffer ratio. The SDS samples were heated at 100 ℃ for 5 min and analysed via SDS-PAGE (section 2.2.3). The HiTrap™ HP column was rinsed with water and stored in 20% (v/v) ethanol.

2.2.6 Protein Concentration

Purified HDP was concentrated using a 3-kDa Amicon® Ultra centrifugal filter (Merck). The protein was loaded onto the centrifugal filter and the concentrator was centrifuged at 4000 × g at 10-min intervals in an Allegra X-30R centrifuge until the volume was approximately 5 mL (prior to size-exclusion chromatography, SEC) or until the protein concentration was ≥1 mg/mL (after size exclusion chromatography).

2.2.7 Protein Concentration Determination

The concentration of purified HDP was determined using a Multiskan GO (Thermo Fisher Scientific). Absorbance of the protein sample and buffer solution (blank) was measured at 280 nm. The sample values were corrected using blank absorbance values and protein concentration was determined using the following equation based on Beer’s Law:

$$c = \frac{A280 \text{ x } 19.608}{Ab\ 0.1\ \%}$$

c = concentration in mg/mL

A280 = absorbance measured at 280 nm

Ab 0.1% = Extinction coefficient in 0.1% solution (1.479 for HDP*)

*Determined by Expasy ProtParam tool (Swiss Institute of Bioinformatics)

2.2.8 Size-exclusion chromatography (SEC)

The concentrated protein (section 2.2.6) was loaded onto a HiLoad 16/600 Superdex 200 pg column for SEC. The column was equilibrated overnight with 1× phosphate buffered saline (PBS, Sigma) at pH 7.4 containing 10% glycerol. The protein was loaded onto the column and SEC was conducted using a Gilson chromatography system controlled by Uniprot software with 3 mL fraction collection. A 60-µl aliquot of the SEC fractions was collected and mixed with 4×SDS sample buffer in a 3:1 sample:sample buffer ratio for SDS-PAGE analysis and the remaining fractions were stored at 4 ℃. The SDS samples were heated at 100 ℃ for 5 min and analysed via SDS-PAGE (section 2.2.3). The column was rinsed with water and stored in 20% (v/v) ethanol.

2.2.9 Ammonium sulphate precipitation

Ammonium sulphate precipitation of the soluble fraction (section 2.2.4) was performed using solid ammonium sulphate with 10% increments in salt saturation using the following formula:

$$G = \frac{533\,(S2 - S1)}{100 - (0.3 \text{ x } S1)}$$

G = ammonium sulphate (g) added per L
S1 = % salt saturation in starting solution
S2 = % salt saturation in final solution

Crushed ammonium sulphate was added to a known volume of the soluble fraction to obtain 20% salt saturation and stirred. Once dissolved, the lysate was kept on ice for 20 min followed by centrifugation at 10,000 × g for 10 min in an Allegra X-30R centrifuge. The supernatant was collected and the pellet was resuspended in lysis buffer equal to the volume of supernatant collected. A 60-µL aliquot of the pellet fraction was collected and 4× SDS sample buffer was added in a 3:1 sample:sample buffer ratio for SDS-PAGE analysis. The remaining pellet fraction was stored on ice and this was repeated with 30, 40, 50, and 60% salt saturation using the resulting supernatant as the starting solution. The SDS samples were heated at 100 ℃ for 5 min and analysed via SDS-PAGE (section 2.2.3).

2.2.10 Anion-exchange chromatography

HDP was purified from the soluble fraction using a 10-mL Q Sepharose High Performance packed column (GE Healthcare) and AKTA explorer purification system (GE Healthcare) with fraction collection. The column was equilibrated with four column volumes of binding buffer (50 mM Tris, pH 7) and the soluble fraction was loaded onto the column. The column was washed with 20 column volumes of binding buffer and the bound protein was eluted via linear gradient elution using 50 mM Tris-1 M NaCl (pH 7) from 0 to 100% NaCl. Aliquots (60 µL) of the sample flow-through, column wash fractions, and elution fractions were collected and 4× SDS sample buffer was added in a 3:1 sample:sample buffer ratio. The SDS samples were heated at 100 ℃ for 5 min and analysed via SDS-PAGE (section 2.2.3). The column was rinsed with water and stored in 20% (v/v) ethanol.

2.2.11 Purification under denaturing conditions

The HDP-BL21 pellet was thawed on ice and resuspended in equilibration buffer (50 mM sodium phosphate, 8 M urea, 300 mM NaCl, 20 mM imidazole (pH 7.4)) containing 0.1% DNase 1 and the lysate was incubated at 20.5 ℃ with gentle agitation for 1 h. The lysate was pelleted via centrifugation at 10,000 x *g* for 30 min and the supernatant and pellet fractions were separated. A 60-µL aliquot was collected and mixed with 4X SDS sample buffer in a 3:1 sample:sample buffer ratio. The samples were then analysed via SDS-PAGE (section 2.2.3). HDP was purified under denaturing conditions using a 1-mL His60 Ni gravity column (Takara Bio). The His60 Ni column was equilibrated with 10 column volumes of equilibration buffer. The supernatant fraction was loaded onto the column and the flow-through was collected. The column was washed with 10 column volumes of wash buffer (50 mM sodium phosphate, 8 M urea, 300 mM NaCl, 40 mM imidazole (pH 7.4)). The bound protein was eluted from the column using elution buffer (50 mM sodium phosphate, 8 M urea, 300 mM NaCl, 300 mM imidazole (pH 7.4)). Aliquots (60 µL) of the sample flow-through, column wash fractions, and elution fractions were collected and 4× SDS sample buffer was added at a 3:1 sample:sample buffer ratio. The SDS samples were heated at 100 ℃ for 5 min and analysed via SDS-PAGE (section 2.2.3). The His60 Ni column washed with 20 column volumes of equilibration buffer, 5 column volumes of water and stored in 20 % ethanol.

2.2.12 Desalting

After purification under denaturing conditions, purified HDP was refolded via buffer exchange using an 8.3-mL Sephadex™ G-25 PD-10 Desalting Column (GE Healthcare). The column was equilibrated with 1× PBS (Sigma) at pH 7.4 containing 10% glycerol, and 2.5 mL eluate containing purified protein was loaded onto the column. The protein was eluted with 3.5 mL PBS under gravity. A 120-µL aliquot of the eluted protein was set aside for SDS-PAGE analysis (section 2.2.3) and protein concentration determination (section 2.2.7). The desalted protein eluate was stored at -80 ºC. The column was washed with two bed volumes of water and stored in 50 % (v/v) ethanol.

2.2.13 Western blot

Soluble and refolded HDP purification samples were analysed via SDS-PAGE and the gel was placed in a transblot sandwich (Whatman filter paper, nitrocellulose membrane, SDS-PAGE gel, Whatman filter paper). The transblot sandwich was placed in a Transblot® Turbo™ Transfer system (BioRad) and the transfer was conducted for 15 min at 25 V. The blot was incubated in blocking buffer (Tris-buffered saline, TBS (25 mM Tris, 150 mM NaCl, pH 7.2)–0.1% Tween 20 containing 1% bovine serum albumin (BSA) and 5% non-fat milk powder) for 1 h. The blocking buffer was discarded and primary antibody (5 mL blocking buffer containing 2.5 µL monoclonal anti-polyHistidine-peroxidase clone HIS-1 [Sigma-Aldrich A7058]) was added to the membrane and incubated at room temperature for 1 h. The blot was washed with TBS–0.1% Tween 20 four times and 3 mL SuperSignal® working solution (Thermo Fisher Scientific) was added to the blot for 5 min. The blot was visualised using Syngene G:Box Chemi and photographed.

2.2.14 Mass spectrometry (MS)

Mass spectrometry analysis of purified soluble and refolded HDP was performed by Tariq Ganief (UCT) using a Q Exactive™ mass spectrometer (Thermo Scientific). Suspected HDP protein bands and co-purified protein bands were excised from an SDS-PAGE gel, cut into 1 x 1 mm pieces, and placed in a clean microfuge tube. The gel pieces were washed three times in washing buffer (5mM NH_4HCO_3 in 50 % v/v acetonitrile) for 20 min under agitation. The supernatant was discarded and the gel pieces were incubated in acetonitrile for 10 min under agitation. 10 mM dithiothreitol was added to the gel pieces and the samples were incubated for 45 min at 56 °C. The supernatant was discarded and 20 mM iodoacetamide was added to the gel pieces. The samples were incubated in the dark for 45 min. The supernatant was discarded and the gel pieces were washed twice in washing buffer and incubated in acetonitrile for 15 min × 2 and left in a fume hood to air dry for approximately 20 min. The gel pieces were incubated in 12.5 ng/µL trypsin for 10 min. 20 mM ammonium bicarbonate buffer was added to the gel pieces and the samples were incubated for 30 min at room temperature. The samples were transferred to 37 °C and incubated overnight under agitation.

The supernatant was transferred to a new microfuge tube (Tube A). Extraction A (50 % v/v acetonitrile, 0.1 % formic acid) was added to the gel pieces and incubated, shaking at room temperature for 30 min. The samples were centrifuged at 2000 rpm for 30 s and the supernatant was transferred to Tube A. Extraction buffer B (80 % v/v acetonitrile, 0.1 % formic acid) was added to the gel pieces and incubated, shaking at room temperature for 30 min. The samples were centrifuged at 2000 rpm for 30 s and the supernatant was transferred to Tube A. Acetonitrile was added to the gel pieces and the samples were incubated, shaking at room temperature for 30 min. The samples were centrifuged at 2000 rpm for 30 s and the supernatant was transferred to Tube A. Tube A was sent to Tariq Ganief (UCT) for vacuum centrifugation and mass spectrometry analysis.

2.3 Results and Discussion

Previously, Roxanne Mohunlal (PhD Thesis, UCT 2019) expressed soluble *Pf*HDP in C41 cells - using *Pf*HDP codon optimised for expression in *E.coli* - in terrific broth growth medium at 15 °C expressed for 48 h after 0.5 mM IPTG induction. Heating of the soluble fraction was used to isolate the protein from contaminating proteins present in the fraction. Purhification by His-tag affinity chromatography and subsequent size exclusion chromatography yielded 1.8 mg protein at the expected size of HDP (25 kDa) from a 4 L culture.

2.3.1 Soluble *Pf*HDP Expression Optimisation

Small-scale expression studies were performed to determine HDP expression levels using the pColdI-*Pf*HDP vector under various conditions, such as varying expression times, presence/absence of glucose, type of cell line, and IPTG concentration. Initial small scale expression trials were carried out in BL21 cells. The culture was incubated with shaking at 37 °C until the OD_{600} was 0.6-1. Expression was induced by the addition of IPTG. The cultures were incubated with shaking at 37 °C for 2 h and at 15 °C for 24 h. The cells were harvested by centrifugation and stored at -20 °C. The cells were lysed by sonication and the soluble (supernatant) and insoluble (pellet) fractions were separated by centrifugation and analysed by SDS-PAGE (Figure 2.2).

Large distinct bands were observed at 25 kDa – the expected size of *Pf*HDP- in the induced whole-cell and insoluble fractions at 15 °C but not in the soluble fraction. In comparison, no distinct protein bands were observed at the expected size in the uninduced samples and in the 37 °C expression culture. This suggests that HDP is forming inclusion bodies, possibly as a result of overexpression or incorrect protein folding.

The addition of supplements such as glucose to the expression media can increase the yield of soluble protein (Rosano and Ceccarelli, 2014). Glucose represses induction of the *lac* promoter resulting in tightly regulated protein expression (Rosano and Ceccarelli, 2014). To improve expression of soluble *Pf*HDP, 1% glucose was added to the expression media, expression was conducted as previously described and samples were analysed by SDS-PAGE (Figure 2.3).

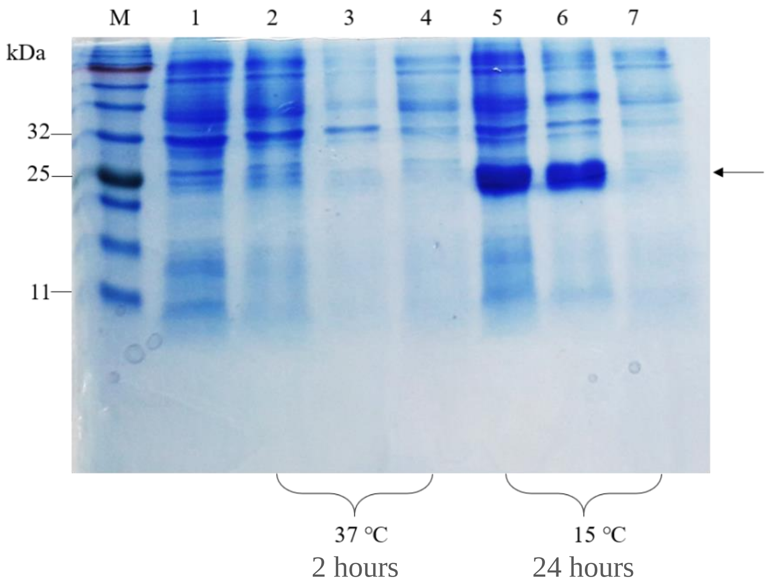


Figure 2.2: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) small-scale expression in BL21 cells.

Small-scale expression samples analysed via SDS-PAGE on a 12% gel. M: Molecular mass marker; Lane 1: uninduced whole-cell fraction; Lane 2: induced whole-cell fraction at 37 °C; Lane 3: induced insoluble fraction at 37 °C; Lane 4: induced soluble fraction at 37 °C; Lane 5: induced whole-cell fraction at 15 °C; Lane 6: induced insoluble fraction at 15 °C; Lane 7: induced soluble fraction at 15 °C.

In Figure 2.3, faint bands at the expected size of HDP, which were absent in the uninduced fraction, were observed in the whole-cell and insoluble fractions at 37 °C. Distinct bands were seen in the whole-cell and insoluble fractions at 15 °C as previously observed. However, the level of soluble *Pf*HDP expression in both the 37 °C and 15 °C samples remained too low to proceed to large-scale expression and protein purification.

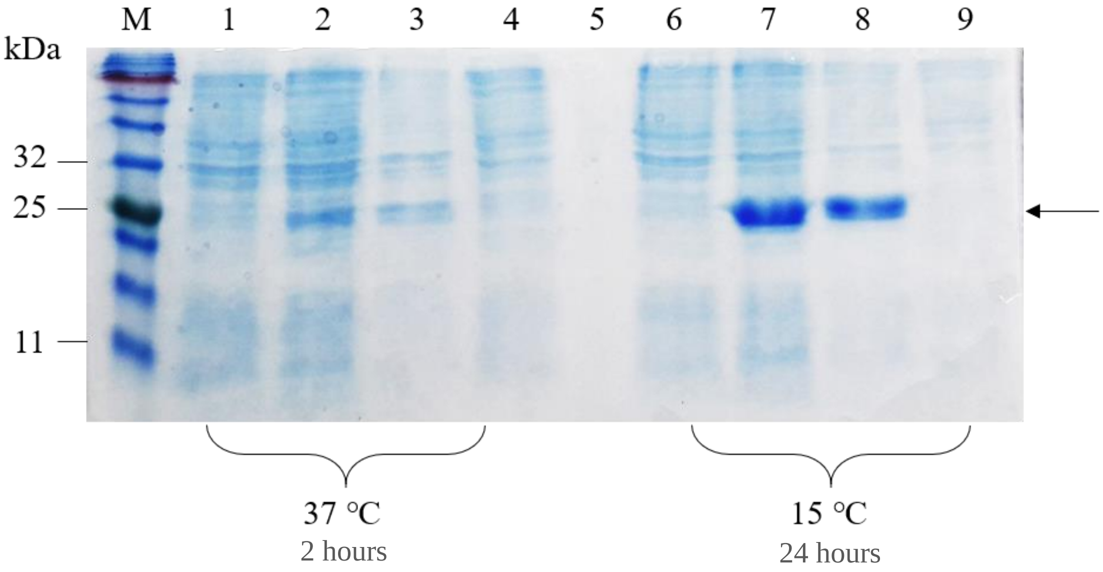


Figure 2.3: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) small-scale expression in glucose-supplemented medium.

Small-scale expression samples analysed via SDS-PAGE on a 12% gel. M: Molecular mass marker; Lane 1: uninduced whole-cell fraction; Lane 2: induced whole-cell fraction at 37 °C; Lane 3: insoluble fraction at 37 °C; Lane 4: induced soluble fraction at 37 °C; Lane 5: empty; Lane 6: uninduced whole-cell fraction; Lane 7: induced whole-cell fraction at 15 °C; Lane 8: induced insoluble fraction; Lane 9: induced soluble fraction.

Expression at lower temperatures is often used to avoid the formation of inclusion bodies during recombinant protein expression (Miyake *et al*., 2007). At lower temperatures, the rate of protein expression is reduced and thus the formation of inclusion bodies resulting from overexpression is reduced (Miyake *et al*., 2007). ArcticExpress cells are *E. coli* cells engineered to increase the yield of soluble protein produced by facilitating expression at low temperatures. The cells co-express chaperonin Cpn60 and co-chaperonin Cpn10, which are cold-adapted chaperonins isolated from the psychrophilic bacterium *Oleispira antarctica* (Hartinger *et al*., 2010). Cpn60 and Cpn10 show high protein refolding activities at 4-12 °C, resulting in improved protein processing at lower temperatures and possibly increased yield of soluble protein.

pColdI-*Pf*HDP was transformed into ArcticExpress cells. Overnight cultures containing 100 µg/mL ampicillin and 20 µg/mL gentamycin were inoculated into 100 mL expression cultures containing no antibiotic (with and without glucose). The cultures were incubated with shaking at 30 °C until the OD_{600} was 0.6-1. Expression was induced by the addition of IPTG. The cultures were incubated with shaking at 11 °C for 24 h. The cells were harvested by centrifugation and stored at -20 °C. The cells were lysed by sonication and the soluble and insoluble fractions were separated via centrifugation. Samples were analysed by SDS-PAGE.

In the expression culture lacking glucose, faint bands were observed at 25 kDa in the insoluble and soluble fractions, whereas relatively distinct bands were observed at 25 kDa in the soluble and insoluble fractions of the glucose-containing expression culture (Figure 2.4). When comparing Figures 2.3 and 2.4, an overall decrease in *Pf*HDP expression was observed, however, the band at 25 kDa observed in the soluble fraction in Figure 2.4 indicates an increase in soluble *Pf*HDP expression. Another distinct band (A) suspected to be Cpn60 was observed at approximately 60 kDa in all the fractions of the glucose-containing expression culture but most of the protein was observed in the soluble fraction. This suggests Cpn60 may help correct protein folding and thus improve soluble protein expression.

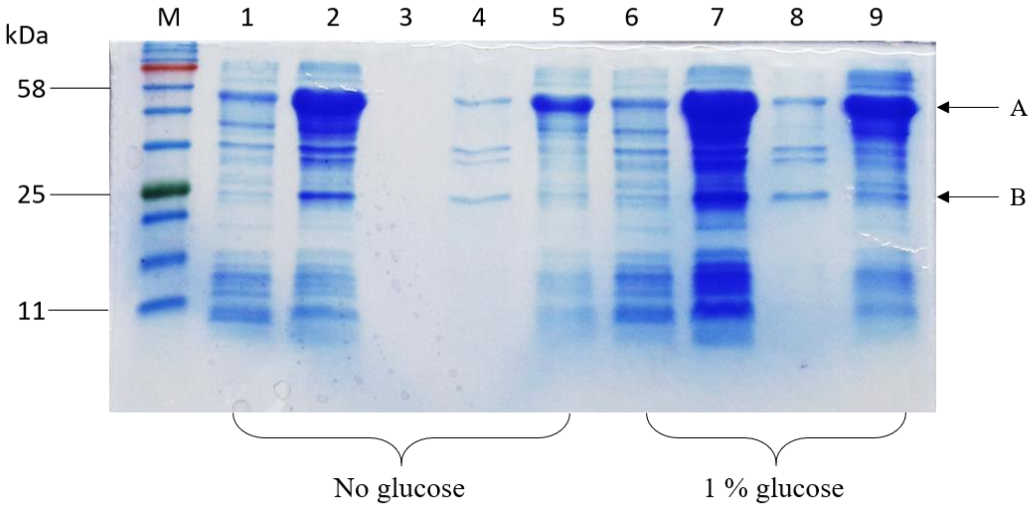


Figure 2.4: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) small-scale expression in ArcticExpress cells.

Small-scale expression samples analysed via SDS-PAGE on a 12% gel. M: Molecular mass marker; Lane 1: no glucose uninduced whole-cell fraction; Lane 2: no glucose induced whole-cell fraction; Lane 3: empty; Lane 4: no glucose insoluble fraction; Lane 5: no glucose soluble fraction; Lane 6: 1 % glucose uninduced whole-cell fraction; Lane 7: 1 % glucose induced whole-cell fraction; Lane 8: 1 % glucose insoluble fraction; Lane 9: 1 % glucose soluble fraction. A - Cpn60; B - *Pf*HDP.

Further optimisation of soluble HDP expression was conducted by varying the IPTG concentration used for induction as changing the concentration of IPTG is often used to increase the yield of soluble protein (Prasad *et al*., 2011). The soluble *Pf*HDP expression observed using ArcticExpress cells was significant enough to proceed to large-scale expression. Overnight cultures of transformed ArcticExpress cells were inoculated into two 5-L flasks containing 1 L LB broth. The cultures were incubated with shaking at 30 ℃ until the OD_{600} was 0.6-0.8. Expression was induced by the addition of 0.1 mM and 0.5 mM IPTG and the cultures were incubated with shaking at 11 ℃ for 24 h. The cells were harvested by centrifugation and stored at -20 ℃. The cells were lysed by sonication and the soluble and insoluble fractions were separated via centrifugation and analysed by SDS-PAGE (Figure 2.5).

The relatively significant difference in *Pf*HDP expression under 0.1 mM and 0.5 mM IPTG induction is shown in Figure 2.5. A larger more distinct protein band was observed in all the induced fractions of the culture induced with 0.5 mM IPTG, indicating higher protein expression. Therefore, 0.5 mM IPTG was used as the optimised IPTG induction concentration for *Pf*HDP expression.

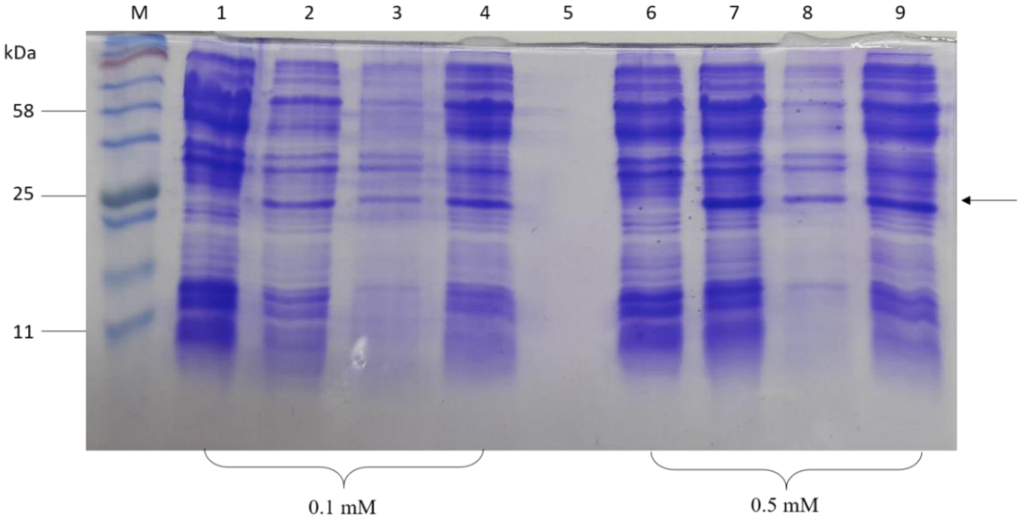


Figure 2.5: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of large-scale *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) expression after induction with 0.1 mM and 0.5 mM IPTG.

Large-scale expression samples analysed via SDS-PAGE on a 12% gel. M: Molecular mass marker; Lane 1: uninduced whole-cell fraction; Lane 2: 0.1 mM induced whole-cell fraction; Lane 3: 0.1 mM insoluble fraction; Lane 4: 0.1 mM induced soluble fraction; Lane 5: empty; Lane 6: uninduced whole-cell fraction; Lane 7: 0.5 mM induced whole-cell fraction; Lane 8: 0.5 mM induced insoluble fraction; Lane 9: 0.5 mM induced soluble fraction.

Another protein expression optimisation technique often used is change in incubation time (Terpe, 2006). Large-scale expression of two 1-L cultures was performed as previously explained using 0.5 mM IPTG to induce expression and incubation with shaking at 11 ℃ for 30 and 48 h. The cells were harvested and stored at -20 ℃. The cells were lysed by sonication and the soluble and insoluble fractions were separated via centrifugation and analysed by SDS-PAGE (Figure 2.6)

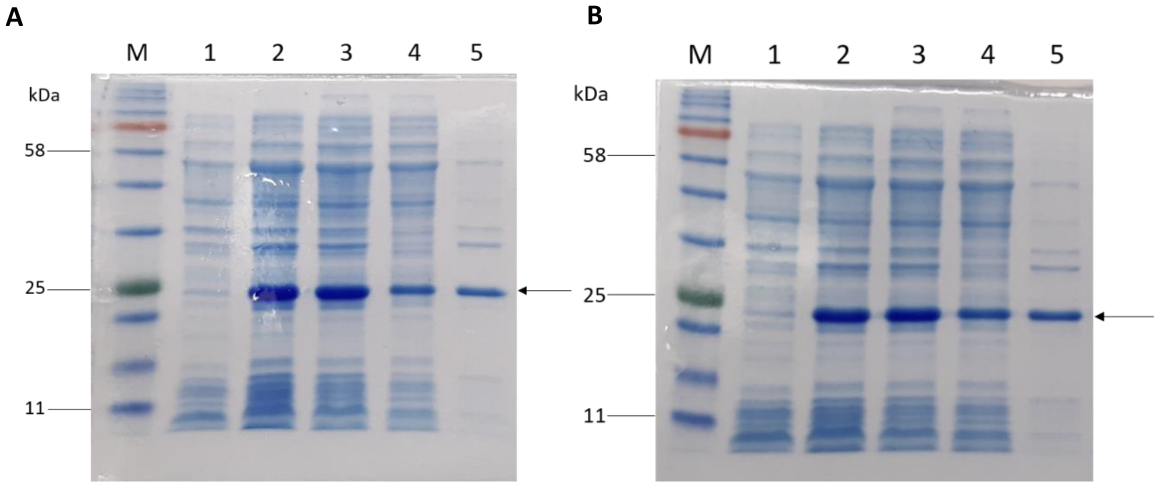


Figure 2.6: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of 30-h (A) and 48-h (B) *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) large-scale expression in ArcticExpress cells.

Large-scale expression samples analysed via SDS-PAGE on a 10% gel. **(A)** M: Molecular mass marker; Lane 1: 30-h uninduced whole-cell fraction; Lane 2: 30-h induced whole-cell fraction; Lane 3: 30-h induced sonicated fraction; Lane 4: 30-h induced soluble fraction; Lane 5: 30-h induced insoluble fraction. **(B)** M: Molecular mass marker; Lane 1: 48-h uninduced whole-cell fraction; Lane 2: 48-h induced whole-cell fraction; Lane 3: 48-h induced sonicated fraction; Lane 4: 48-h induced soluble fraction; Lane 5: 48-h induced insoluble fraction.

After 30- and 48-h expression, relatively similar levels of *Pf*HDP expression were observed in the soluble and insoluble fractions (Figure 2.6). The small difference in *Pf*HDP levels at the different time-points resulted in the selection of 30 h as the optimised expression time in order to save time and resources and minimise the risk of cells losing the expression plasmid as a result of long expression periods.

Optimised expression conditions for soluble *Pf*HDP were determined as *Pf*HDP in ArcticExpress cells cultured in 1L LB broth supplemented with 1% glucose with induction by addition of 0.5 mM IPTG, expressed for 30 h. The optimised conditions produced a sufficient amount of *Pf*HDP to proceed to protein purification.

2.3.2 *Pf*HDP Purification Strategies

Large-scale expression was performed as previously described (section 2.2.4). *Pf*HDP was purified from the soluble fraction via His-tag affinity chromatography followed by SEC. Samples collected during affinity chromatography and SEC were analysed via SDS-PAGE.

The gel shown in Figure 2.7A indicates that a significant amount of soluble *Pf*HDP was expressed. Analysis of the purification fractions showed a very faint band at 25 kDa in elution fraction 54 and contaminating proteins were observed at higher molecular weights (>58 kDa) in all the elution fractions analysed (52, 54, 56, and 58). A distinct band at 25 kDa in the flow-through (fraction 4) suggests that most of the protein expressed did not bind to the HisTrap HP column. This may possibly be due to the presence of chaperones which may mask the his-tag or perhaps due to the protein folding in a way that buries the his-tag.

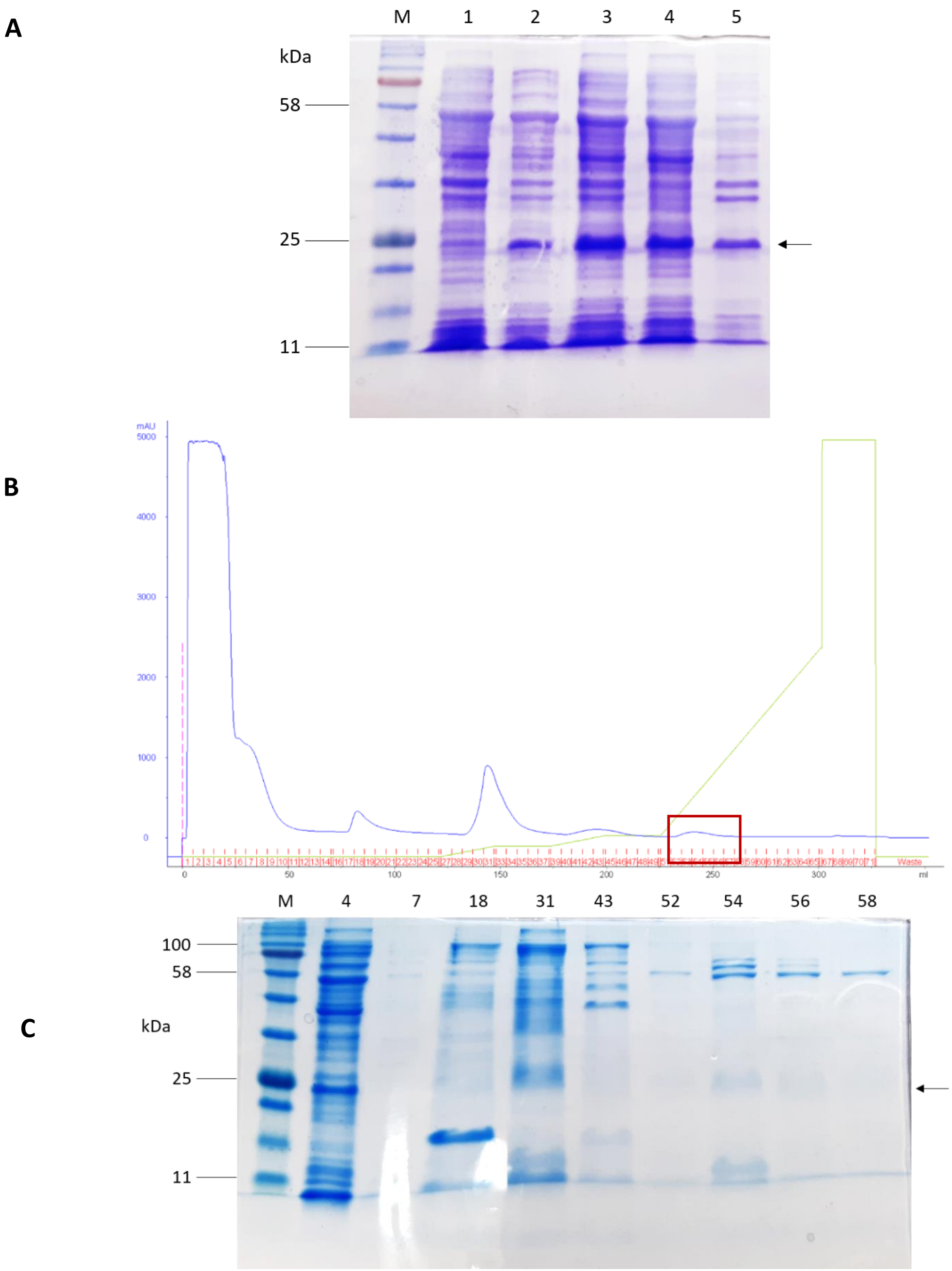


Figure 2.7: Affinity chromatography of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) expressed in LB broth.

(A) Large-scale expression samples analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; Lane 1: uninduced whole-cell fraction; Lane 2: induced whole-cell fraction; Lane 3: induced sonicated fraction; Lane 4: induced soluble fraction; Lane 5: induced insoluble fraction. **(B)** *Pf*HDP HisTrap HP purification profile, red box indicates the elution peak. **(C)** HisTrap HP purification fractions analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; Fractions analysed: 4; 7; 18; 31; 43; 52; 54; 56; 58.

An unusually high baseline was observed at the beginning of the chromatography profile likely due to the presence of contaminating buffers in the exclusion column (Figure 2.8). The *Pf*HDP peak observed was small indicating low protein concentration and was confirmed via SDS-PAGE analysis of the chromatography fractions which showed very faint bands at 25 kDa. The protein peaks were pooled, concentrated, and glycerol was added before storing at -80 °C. The final concentration of the purified *Pf*HDP was 0.2 mg/mL in 500 µL.

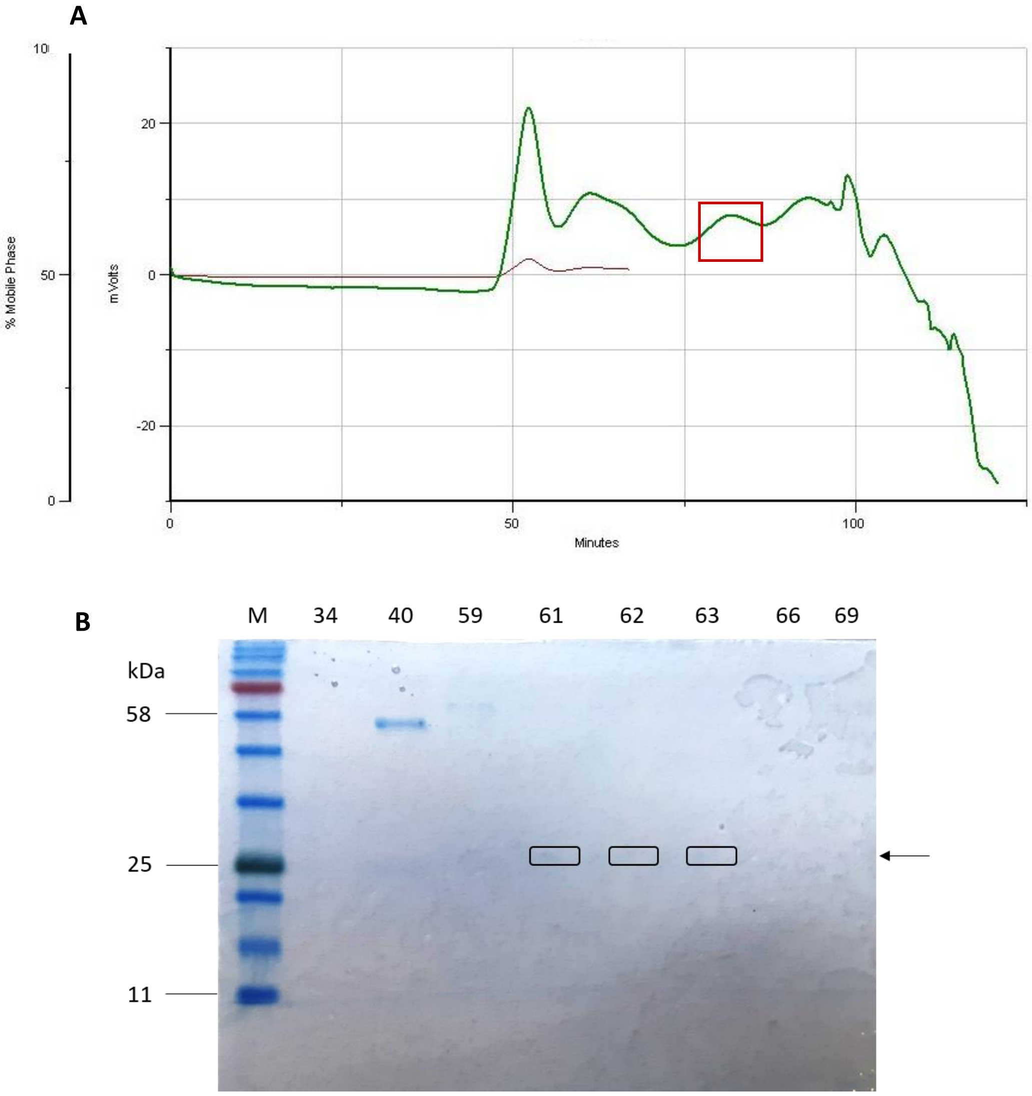


Figure 2.8: Size-exclusion chromatography of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP).
(A) HiLoad 16/600 Superdex 200 pg size -chromatography profile; red box indicates the *Pf*HDP peak. **(B)** Size-exclusion chromatography fractions analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; Fractions analysed: 34; 40; 59; 61; 62 ;63; 66; 69.

*Pf*HDP was expressed in ArcticExpress cells in 1 L terrific broth, a nutrient-rich growth medium that allows cultures to reach high cell densities by extending the exponential phase of *E. coli* in order to increase the amount of soluble protein in the lysate (Kram and Finkel, 2015). Furthermore, ammonium sulphate precipitation was used to isolate *Pf*HDP from contaminating proteins in the cell lysate. The principle of ammonium sulphate precipitation is based on the decrease in protein solubility at higher salt concentrations, leading to protein precipitation (Wingfield, 2001). Large-scale expression was performed as previously described. The soluble fraction isolated from the lysed cell suspension, was filtered and divided in half. Half of the obtained soluble fraction was used in ammonium sulphate precipitation.

A slight increase in soluble *Pf*HDP expression in terrific broth growth medium is observed in Figure 2.9 compared to that in LB medium, indicating that terrific broth should be used instead of LB medium in the optimised *Pf*HDP expression protocol. Ammonium sulphate precipitation of *Pf*HDP showed the protein started to precipitate out at relatively low ammonium sulphate concentrations (20%) up to 50% ammonium sulphate saturation.

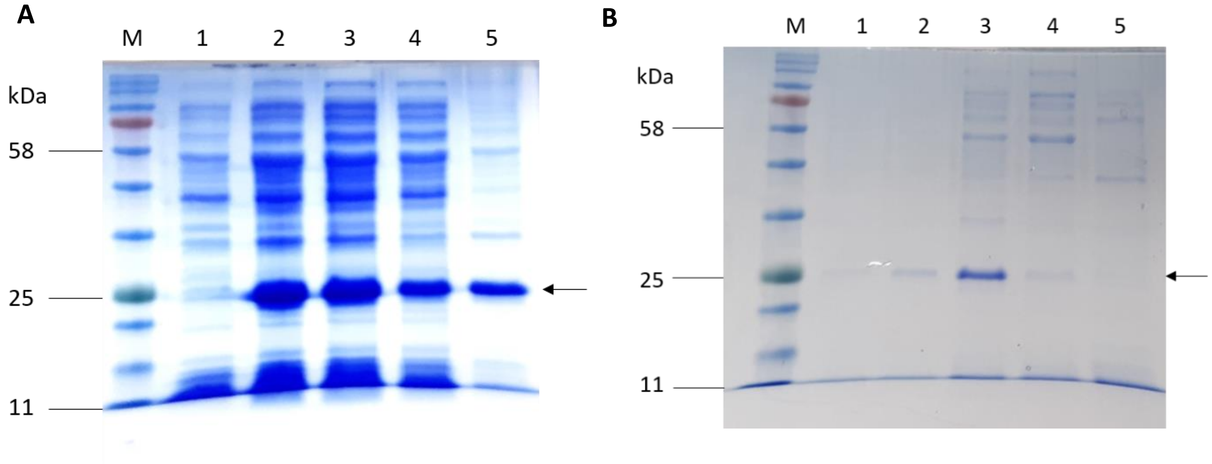


Figure 2.9: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of large-scale *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) expression in terrific broth medium and ammonium sulphate precipitation.
(**A**) PageBlue™ stained large-scale expression samples analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; Lane 1: uninduced whole-cell fraction; Lane 2: induced whole-cell fraction; Lane 3: induced sonicated fraction; Lane 4: induced soluble fraction; Lane 5: induced insoluble fraction. **(B)** Protein precipitate samples analysed via SDS-PAGE and Coomassie-stained. M: Molecular mass marker; Lane 1: 20% ammonium sulphate saturation; Lane 2: 30% ammonium sulphate saturation; Lane 3: 40% ammonium sulphate saturation; Lane 4: 50% ammonium sulphate saturation; Lane 5: 60% ammonium sulphate saturation.

As the *Pf*HDP levels observed at 20% and 50% ammonium sulphate saturation were relatively low, the 30% and 40% ammonium sulphate saturation fractions were pooled and concentrated using an Amicon® Ultra 10-kDa centrifugal filter. The protein was desalted and loaded onto a

HisTrap HP purification column to isolate *Pf*HDP from contaminating proteins in the sample and subsequently loaded onto the HiLoad 16/600 Superdex 200 pg column for SEC. The samples collected during these purification steps were analysed via SDS-PAGE.

As observed in the previous purification, the bulk of *Pf*HDP in the starting sample eluted out with the flow-through, indicating that the inclusion of the ammonium sulphate precipitation step did not improve binding to the nickel column. Although no protein bands were observed in the elution fractions analysed via SDS-PAGE, the fractions corresponding to the peak observed in Figure 2.10A were pooled, concentrated, and loaded onto the SEC column.

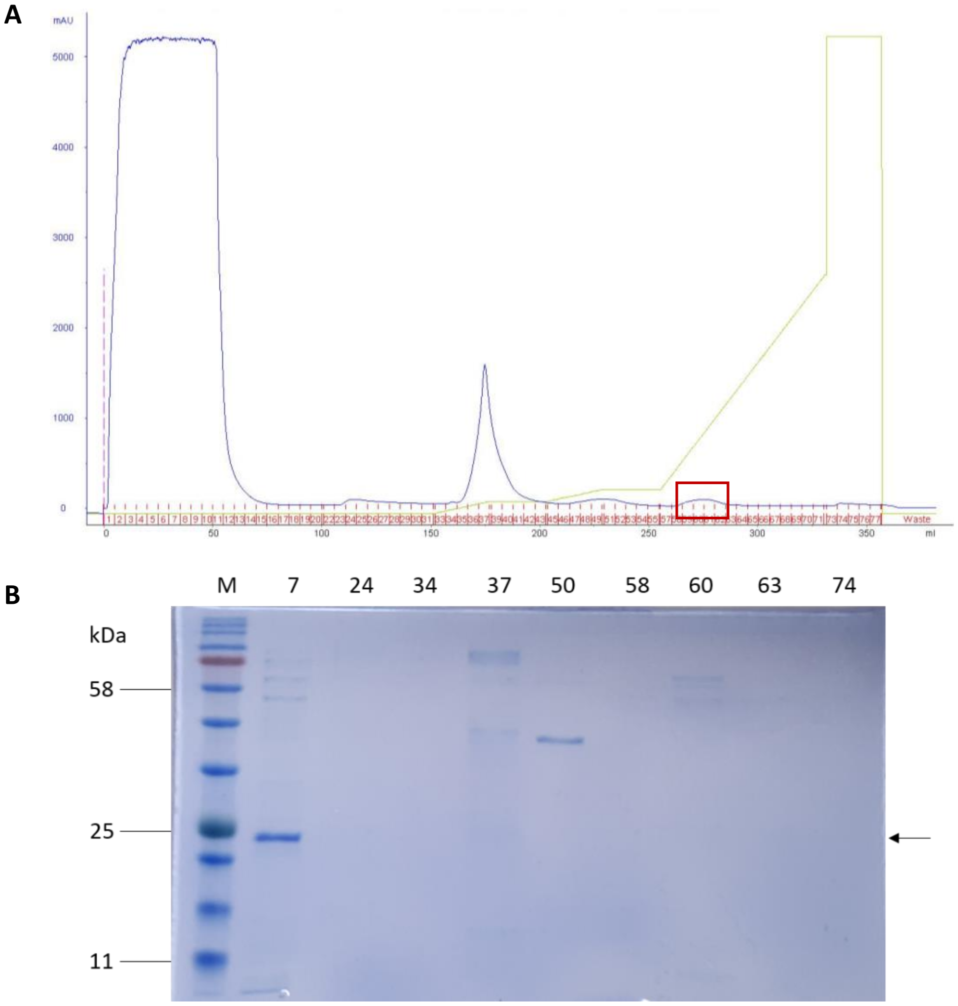


Figure 2.10: Affinity chromatography of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) expressed in terrific broth.

(A) *Pf*HDP HisTrap HP purification profile; red box indicates the elution peak. **(B)** HisTrap HP purification fractions analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; Fractions analysed: 7; 24; 34; 37; 50; 58; 60; 63; 74.

The size-exclusion profile and SDS-PAGE analysis in Figure 2.11 shows successful purification of a protein corresponding to the size of *Pf*HDP, although a very low concentration of the protein was obtained (approximately 0.03 mg) and some contaminating proteins at very low concentrations were observed to co-elute with *Pf*HDP.

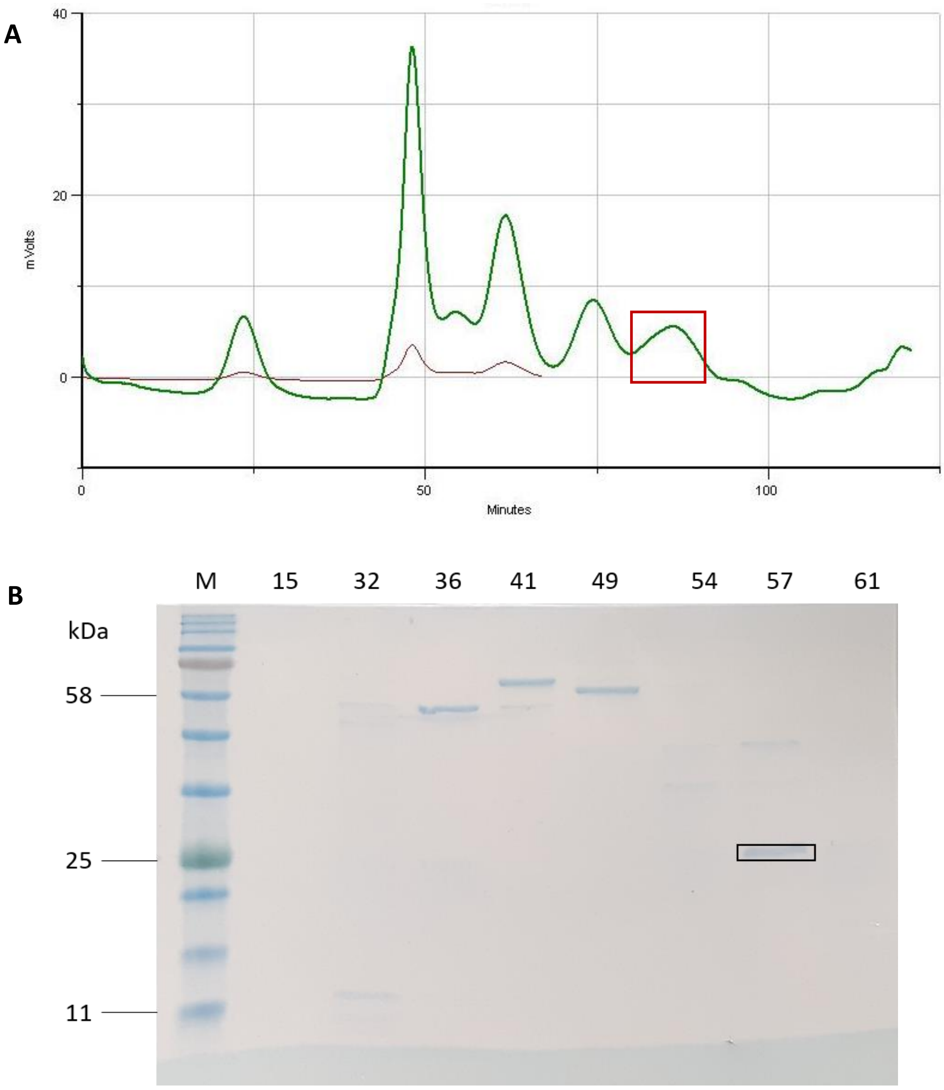


Figure 2.11: Size-exclusion chromatography of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP).

(A) HiLoad 16/600 Superdex 200 pg size-exclusion chromatography profile; red box indicates the *Pf*HDP peak. **(B)** Size-exclusion chromatography fractions analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; Fractions analysed: 15; 32; 36; 41; 49; 54; 57; 61.

The remaining half of the cell suspension was loaded onto a HisTrap HP purification column and subsequently a 16/600 Superdex 200 pg column for SEC. The samples collected during affinity chromatography and SEC were analysed via SDS-PAGE.

The purification profile observed was similar to the profiles of previous purification experiments. The bulk of the protein eluted out in the flow-through fraction and was stored, although low concentrations of a protein corresponding to the size of *Pf*HDP were observed in the elution fractions (46, 57, and 62), which corresponds to the small peak (red box) seen in the purification profile (Figure 2.12).

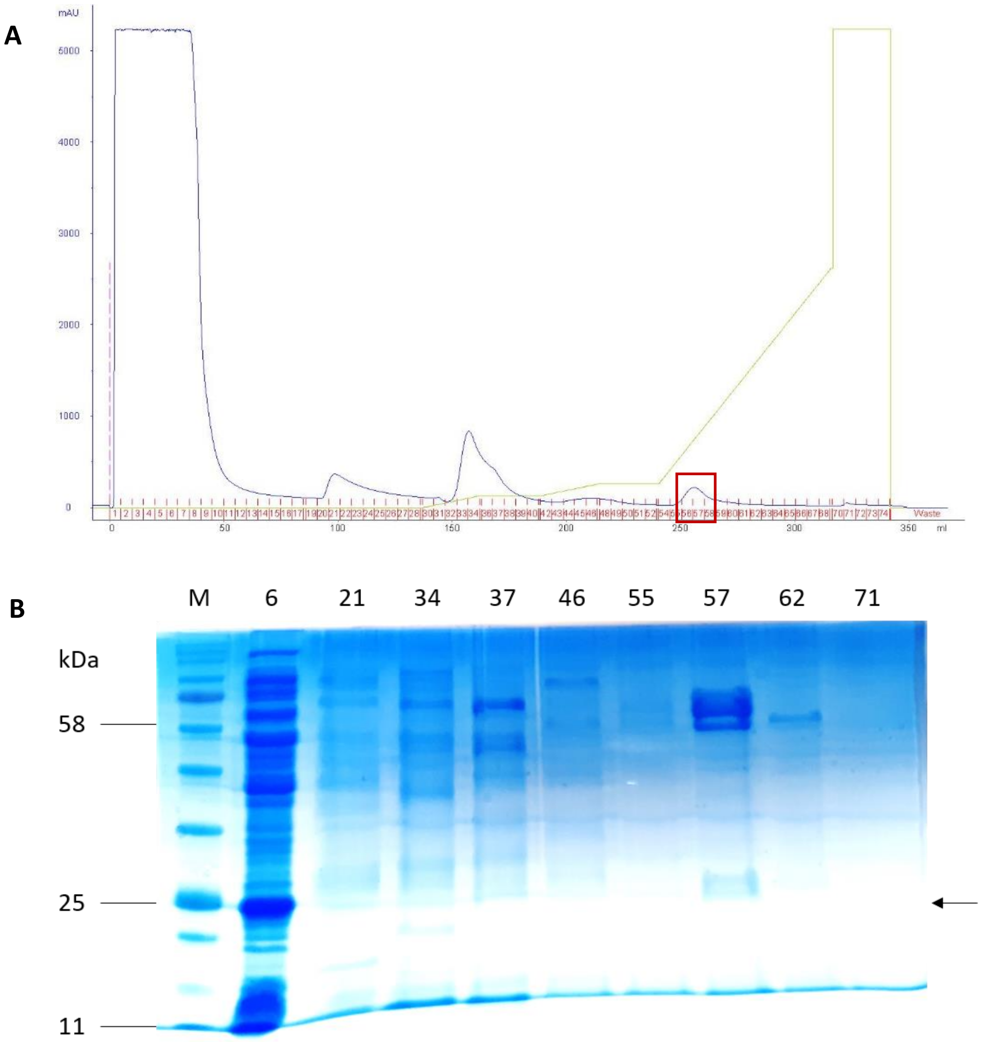


Figure 2.12: Affinity chromatography of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) expressed in terrific broth.
(A) *Pf*HDP HisTrap HP purification profile; red box indicates the elution peak. **(B)** HisTrap HP purification fractions analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; Fractions analysed: 6; 21; 34; 37; 46; 55; 57; 62; 71.

Suspected *Pf*HDP was observed in various fractions of the SEC profile, even in fractions containing high molecular-weight proteins as is evident in Figure 2.13B, in which high molecular-weight proteins at approximately 58 kDa and 80 kDa were observed to co-elute with *Pf*HDP in fractions 37 and 41 respectively. This may be due to HDP binding to other proteins in the sample.

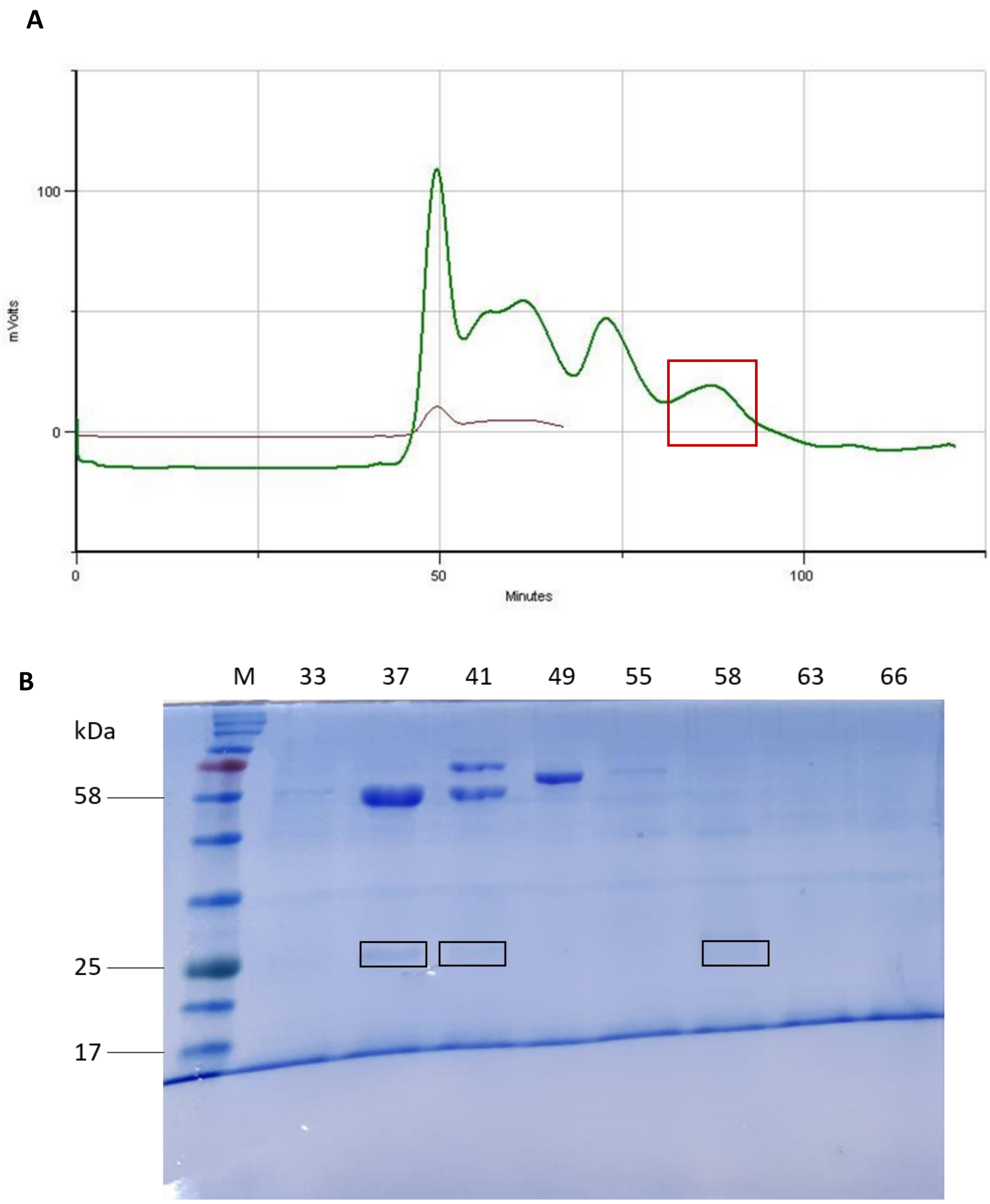


Figure 2.13: Size-exclusion chromatography of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP).
(A) HiLoad 16/600 Superdex 200 pg size-exclusion chromatography profile; red box indicates the *Pf*HDP peak. **(B)** Size-exclusion chromatography fractions analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; Fractions analysed: 33; 37; 41; 49; 55; 58; 63; 66.

Anion-exchange chromatography uses a positively charged stationary phase to isolate negatively charged anions (Wilson and Walker, 2010). The theoretical isoelectric point of *Pf*HDP was determined as 9.29 (EXPASY), although experiments showed that the protein was negatively charged at pH 7, and thus anion exchange was conducted using Tris-buffers at pH 7. The flow-through from the previous purification contained the bulk of *Pf*HDP expressed, possibly because of the presence of other proteins that may prevent *Pf*HDP from binding to the Ni-NTA column by masking the His-tag. The flow-through fractions were pooled and loaded onto an anion-exchange column composed of Q Sepharose High Performance Resin.

Anion-exchange chromatography removed the bulk of the contaminating proteins present in the sample (Figure 2.14). HDP was found to co-elute in fractions 17 – 21 with other proteins. A band thought to be Cpn60 was observed at approximately 60 kDa in fraction 19, which also showed a significant band at approximately 25 kDa. This suggests that Cpn60 may interact with HDP and help solubilise the protein although it may also mask the His-tag by doing so, which could explain why HDP was not binding to the nickel column. As the peaks in the chromatography profile were of low resolution and a significant amount of HDP was observed in the flow-through with very few contaminants, the flow-through fraction was loaded onto the HisTrap HP purification column and subsequently the 16/600 Superdex 200 pg column for further purification. The samples collected during affinity chromatography and SEC were analysed via SDS-PAGE (Figure 2.15).

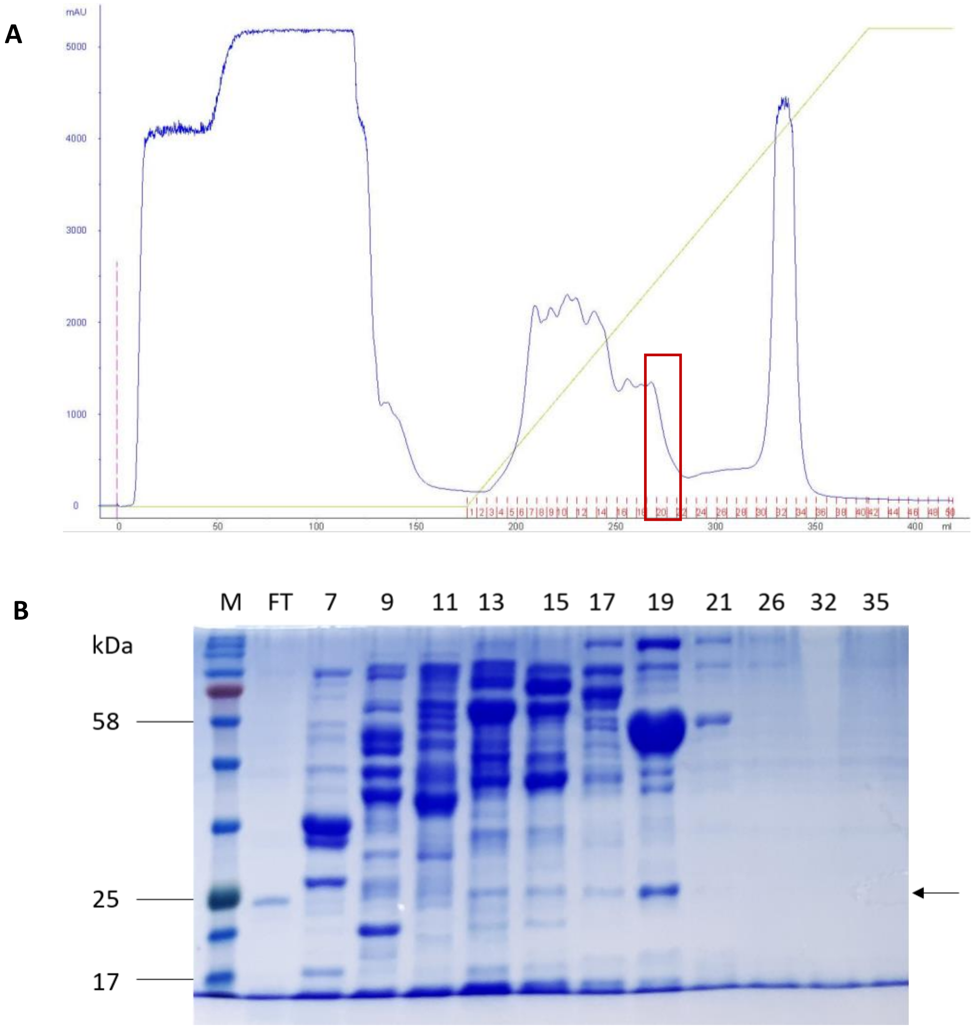


Figure 2.14: Anion-exchange chromatography of affinity chromatography flow-through fraction.

(A) *Pf*HDP anion exchange purification profile; red box indicates the elution peak. **(B)** Anion-exchange chromatography fractions analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; FT: Flow-through; Fractions analysed: 7; 9; 11; 13; 15; 17; 19; 21; 26; 32; 35.

Both the purification and gel analysis in Figure 2.15 show very few contaminating proteins were present in the starting sample. No HDP was observed in the flow-through fraction indicating successful binding of the His-tag to the Ni^{2+}-NTA column and suggesting that the presence of other proteins in the sample was masking the His-tag. Low concentrations of HDP were observed in the earlier fractions (prior to the elution step) but significant bands were seen at the expected size corresponding to the elution peak of the purification profile despite its low resolution. A high molecular-weight protein at approximately 100 kDa also co-eluted with HDP, as has been observed in previous experiments (Roxanne Mohunlal PhD Thesis, UCT 2019).

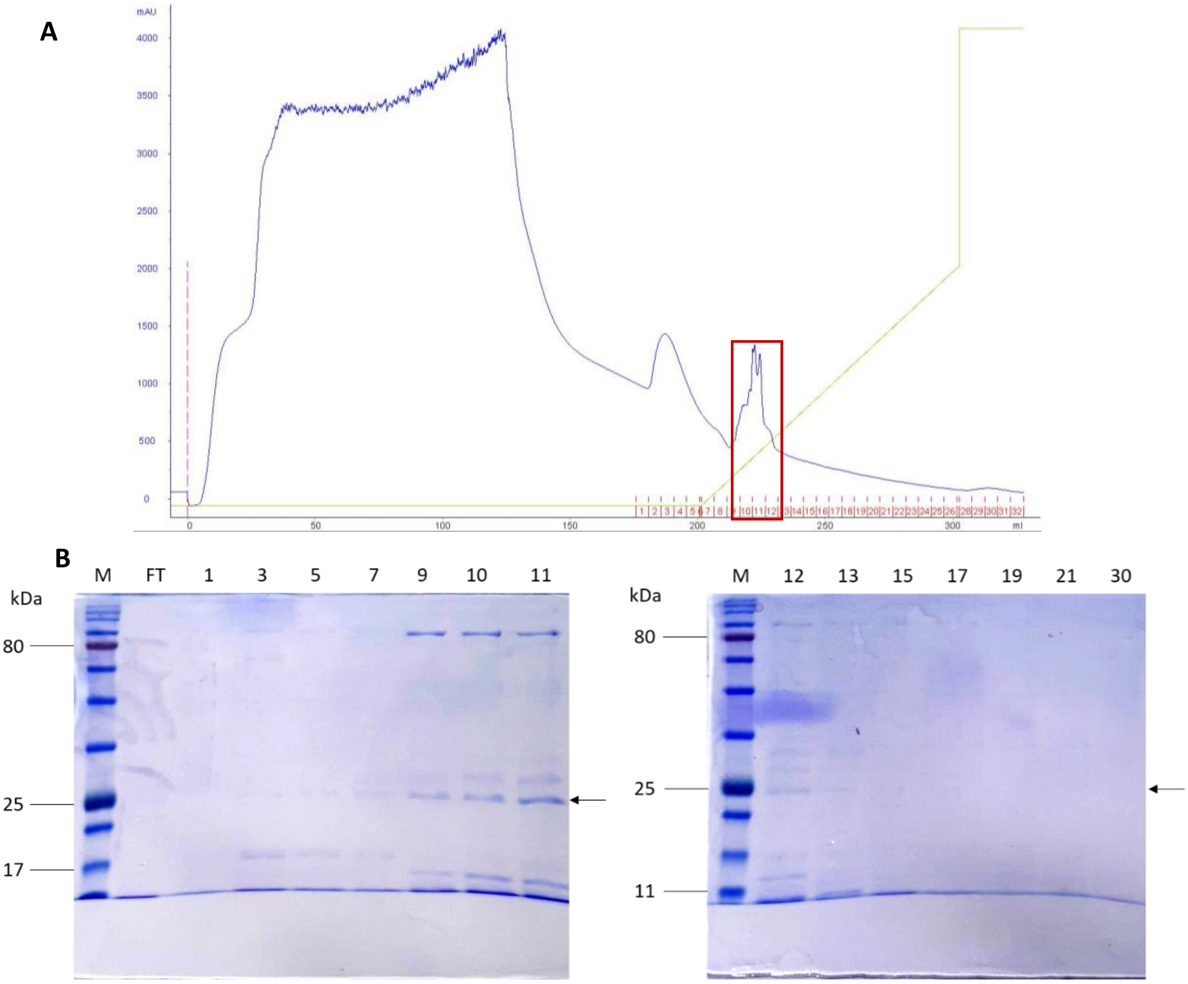


Figure 2.15: Affinity chromatography of anion exchange chromatography flow-though. (**A**) HisTrap HP purification profile; red box indicates the elution peak. **(B)** Purification fractions analysed via SDS-PAGE on 10-% gels. M: Molecular mass marker; FT: Flow-through; Fractions analysed: 1; 3; 5; 7; 9; 10; 11; 12; 13; 15; 17; 19; 21; 30.

SEC was conducted to separate the two proteins (Figure 2.16). In Figure 2.16A, one large peak at position 3 was expected as distinct bands were observed at 25 kDa and 27 kDa after affinity chromatography. However, only one large peak was observed (at position 1 – the void volume) and smaller peaks were also observed later in the analysis at 75 minutes and 100 minutes (the expected elution time for *Pf*HDP). SDS-PAGE analysis (Figure 2.16B) revealed distinct protein bands at approximately 100, 27, and 25 kDa. The protein at 100 kDa was found in peak 1 and peak 2 while in peak 3, the two smaller protein bands were observed.

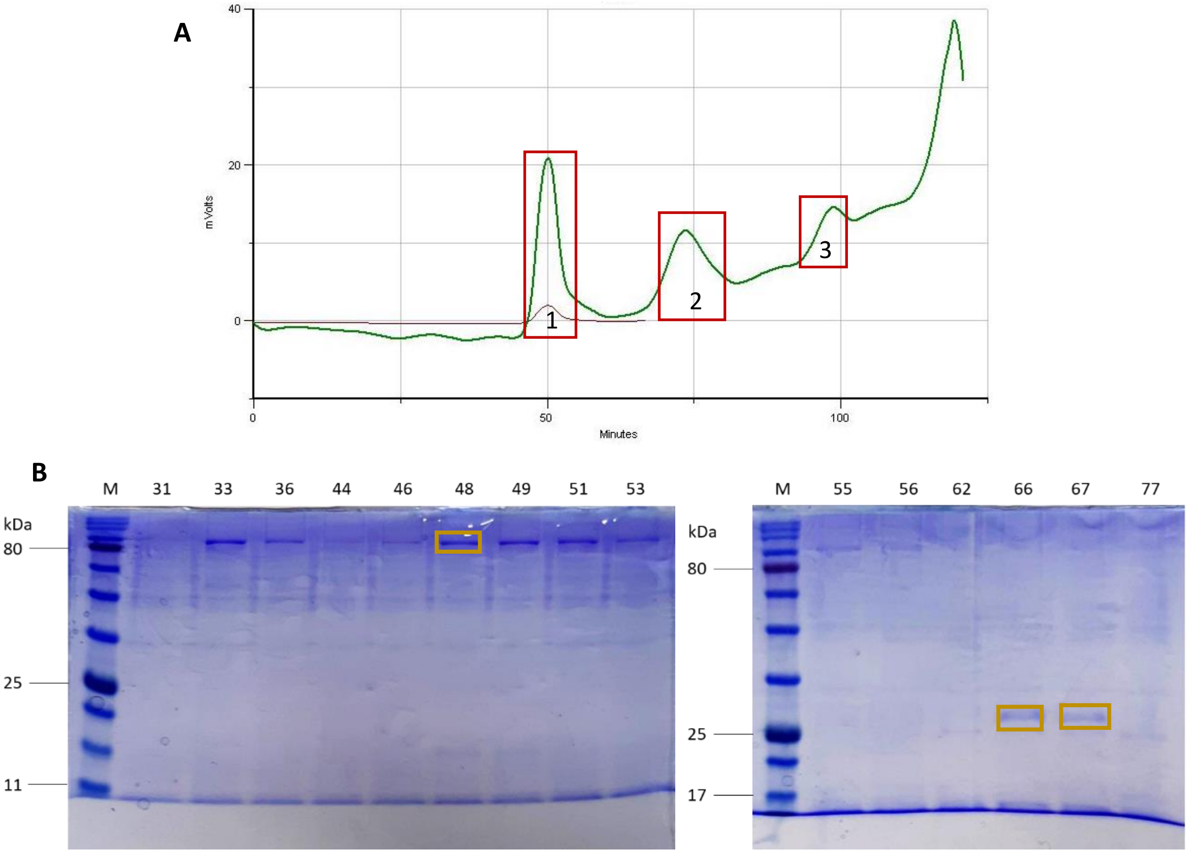


Figure 2.16: Size-exclusion chromatography of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP).

(A) HiLoad 16/600 Superdex 200 pg size-exclusion chromatography profile; red boxes show Peak 1, Peak 2, and Peak 3. **(B)** Size-exclusion chromatography fractions analysed via SDS-PAGE on 10% gels. M: Molecular mass marker; Fractions analysed: 31; 33; 36; 44; 46; 48; 49; 51; 53; 55; 56; 62; 66; 67; 77. Yellow boxes show distinct protein bands.

Although soluble *Pf*HDP expression was previously optimised, insufficient His-tag binding to the Ni^{2+}-NTA column hindered protein purification and protein precipitation during the final stages of the purification process further reduced the protein yield. Recombinant production of active haem proteins is often limited by insufficient incorporation of haem (Krainer *et al*., 2015). In *Escherichia coli*, the formation of 5-aminolevulinic acid hydrochloride (5-ALA) is the rate-limiting step in the haem biosynthesis pathway and supplementation of the medium with 5-ALA is thus a possible strategy used to improve haem biosynthesis. As *Pf*HDP has been shown to bind to haem (Jani *et al*., 2008), the expression of *Pf*HDP was supplemented with 5-ALA in an attempt to enhance protein folding during expression and potentially improve protein binding to the HiTrap™ HP column.

Previously determined HDP expression conditions (*Pf*HDP in ArcticExpress cells cultured in 1 L TB medium supplemented with 1% glucose with induction by addition of 0.5 mM IPTG, expressed for 30 h) were further optimised by adding 5-ALA. Overnight cultures of ArcticExpress cells were inoculated into 100 mL TB media supplemented with 1% glucose containing no antibiotic, with and without 5-ALA. The cultures were incubated with shaking at 30 °C until the OD_{600} reached 0.6–1. Expression was induced via addition of 0.5 mM IPTG and the cultures were incubated with shaking at 11 °C for 30 h. The cells were harvested via centrifugation and stored at -20 °C. The cells were then lysed via sonication and the soluble and insoluble fractions were separated via centrifugation and analysed by SDS-PAGE.

In the expression culture containing 5-ALA (Figure 2.17B), large bands were observed at 25 kDa in the induced and soluble fractions, indicating a significant increase in *Pf*HDP expression compared to that observed in the culture lacking 5-ALA (Figure 2.17A). When comparing expression in the soluble and insoluble fractions, the culture lacking 5-ALA showed relatively equal amounts of soluble and insoluble HDP whereas more soluble HDP was observed in the culture containing ALA. 5-ALA was thus included in the optimised expression conditions for soluble HDP expression in ArcticExpress cells, and large scale expression was carried out.

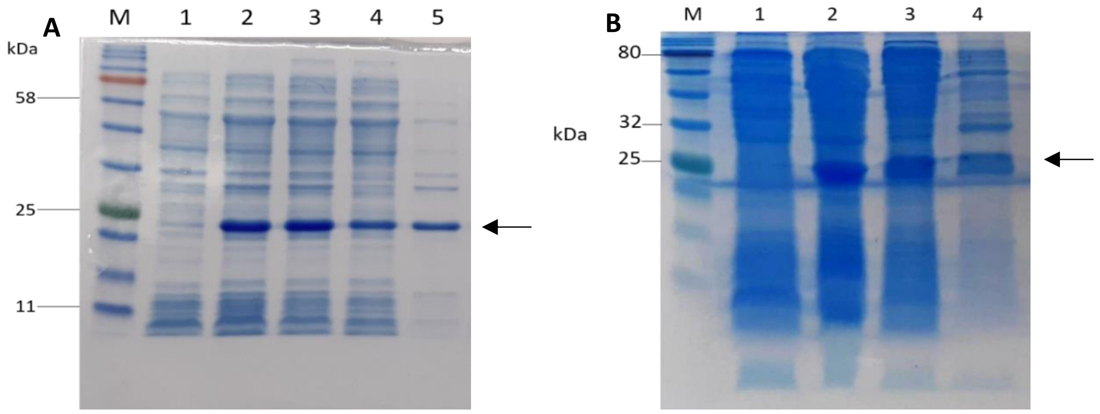


Figure 2.17: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) small-scale expression in ArcticExpress cells cultured without 5-ALA (A) and with 5-ALA (B).
Large-scale expression samples analysed via SDS-PAGE on a 10% gel. **(A)** M: Molecular mass marker; Lane 1: no 5-ALA, uninduced whole-cell fraction; Lane 2: no 5-ALA, induced whole-cell fraction; Lane 3: no 5-ALA, induced sonicated fraction; Lane 4: no 5-ALA, induced soluble fraction; Lane 5: no 5-ALA, induced insoluble fraction. **(B)** M: Molecular mass marker; Lane 1: + 5-ALA, uninduced whole-cell fraction; Lane 2: + 5-ALA, induced whole-cell fraction; Lane 3: + 5-ALA, induced soluble fraction; Lane 4: + 5-ALA, induced insoluble fraction.

Large-scale expression of *Pf*HDP in previously optimised expression conditions - ArcticExpress cells cultured in 1 L TB medium supplemented with 1% glucose with induction by addition of 0.5 mM IPTG, expressed for 30 h – and addition of 5-ALA was performed as previously described (section 2.2.4). The soluble fraction isolated from the lysed cell suspension was filtered and loaded onto a HiTrap™ 5 mL HP column. SDS-PAGE analysis of the purification fractions showed most of the protein did not bind to the HiTrap™ HP column thus the flow-through fractions were pooled and loaded onto an anion-exchange column composed of Q Sepharose High-Performance Resin. The anion exchange flow-through fractions were pooled and concentrated using a 3-kDa Amicon® Ultra centrifugal filter prior to SEC using a HiLoad 16/600 Superdex 200 pg column. Samples collected during affinity chromatography, anion exchange, and SEC were analysed via SDS-PAGE.

Analysis of purification fractions obtained via SDS-PAGE (Figure 2.18) revealed a very faint band in the affinity chromatography elution fraction (fraction 41) and a contaminating band at approximately 58 kDa. A distinct band at 25 kDa was observed in the affinity chromatography flow-through fraction (fraction 4), indicating that most of the protein in the lysate did not bind to the HiTrap™ HP column thus addition of 5-ALA in the expression media did not improve *Pf*HDP binding to the column. Similarly, a faint band at 25 kDa was observed in the anion exchange elution fraction (fraction 20) and contaminating proteins were present in the sample, although a distinct band at 25 kDa was also observed in the anion exchange flow-through fraction (fraction 8) with a single band at approximately 58 kDa.

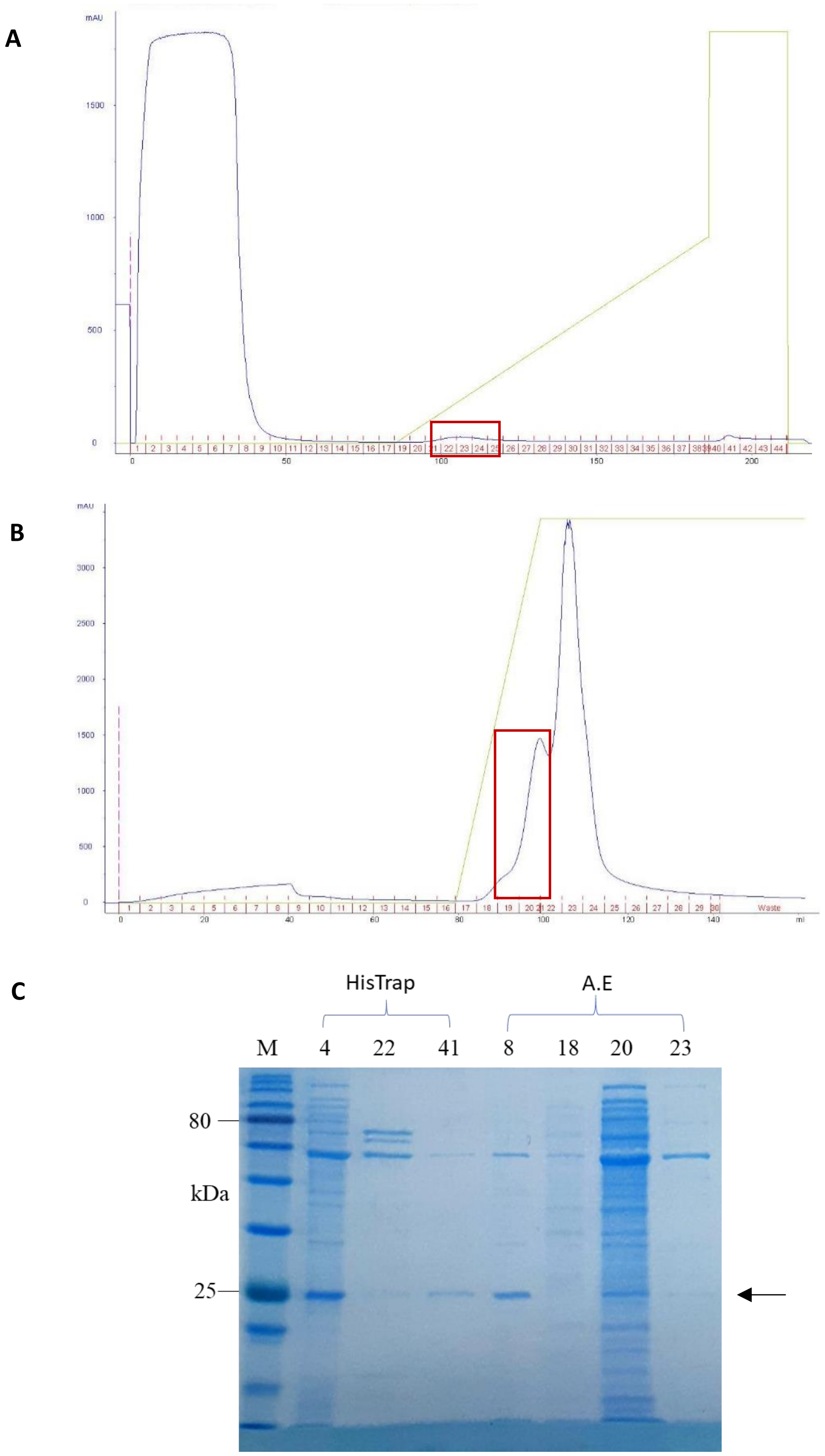


Figure 2.18: Purification of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) expressed ArcticExpress cells.

(A) *Pf*HDP HiTrap™ HP purification profile, red box indicates the elution peak. **(B)** *Pf*HDP anion exchange purification profile, red box indicates the elution peak. **(C)** *Pf*HDP purification samples analysed via SDS-PAGE on a 10% gel. M: Molecular marker; HiTrap™ HP fractions analysed: 4, 22, 41; Anion Exchange (A.E) fractions analysed: 8, 18, 20, 23.

As most of the contaminants present in the lysate were removed by affinity chromatography and anion exchange, the anion exchange flow-through fraction was concentrated and analysed via SEC. The *Pf*HDP peak observed was small, indicating low protein concentration and this was confirmed via SDS-PAGE analysis of the SEC fractions in which no bands were observed at 25 kDa, possibly because of protein precipitation occurring prior to SEC (Figure 2.19).

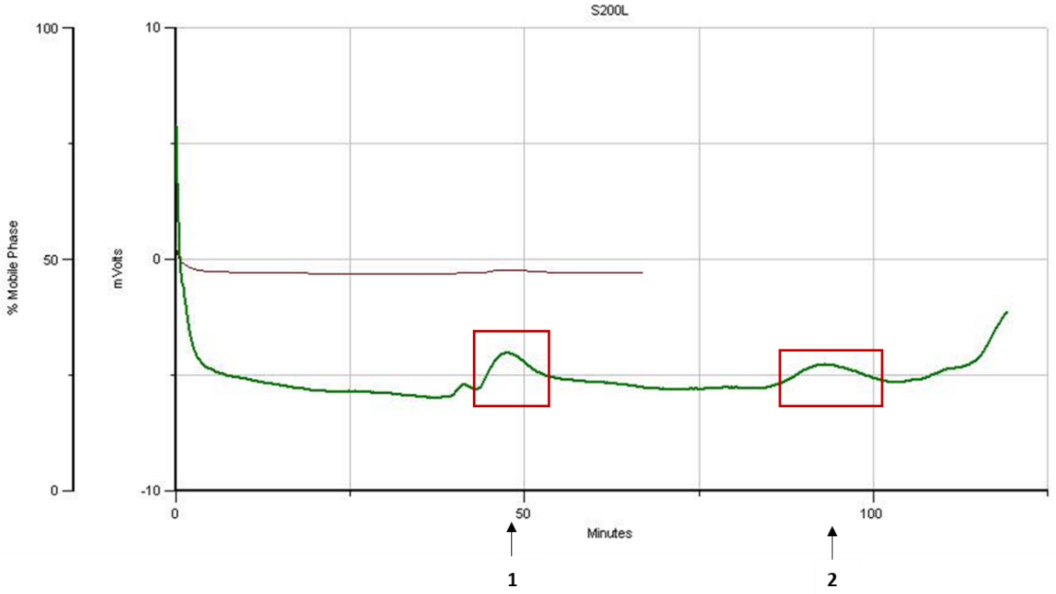


Figure 2.19: Size-exclusion chromatography of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP).
HiLoad 16/600 Superdex 200 pg size-exclusion chromatography profile; red box indicates the isolated peaks.

Despite improved soluble expression of *Pf*HDP as seen in the small-scale expression tests (Figure 2.17), poor binding of the expressed protein to the Ni-NTA column via the his-tag made purification difficult. There was a significant loss of protein during the purification process and very low concentrations of *Pf*HDP were obtained using ArcticExpress cells thus expression of *Pf*HDP in C41 cells and subsequent purification was performed as this had been relatively successful in previous studies by Roxanne Mohunlal (PhD Thesis, UCT 2019).

2.3.3 *Pf*HDP expression in C41 cells and purification

C4 cells – originally described by Miroux and Walker (1996) – are a mutant strain of BL21 cells used to overcome the toxicity associated with overexpression of recombinant proteins that may be toxic to the *E.coli* expression system. (Dumon-Seignovert *et al*., 2004). *Plasmodium falciparum* HDP was expressed in C41 cells in 3 L TB supplemented with 1% glucose as previously described (section 2.2.4). The soluble fraction isolated from the lysed cell suspension was filtered and loaded onto an equilibrated HiTrap™ 5 mL HP column. The elution fractions were pooled and loaded onto a hydrophobic interaction chromatography column before concentrating prior to SEC using a HiLoad 16/600 Superdex 200 pg column. Samples collected during affinity chromatography, HIC, and SEC were analysed via SDS-PAGE.

Analysis of the purification fractions (Figure 2.20B) revealed a large distinct band at 25 kDa in later elution fractions (mostly fraction 77), although a large number of contaminating proteins were also present. Further purification via hydrophobic interaction chromatography removed a number of contaminating proteins (Figure 2.20C), and protein bands were observed at approximately 72, 55, 34, 25, 17, and 10 kDa in the elution fractions.

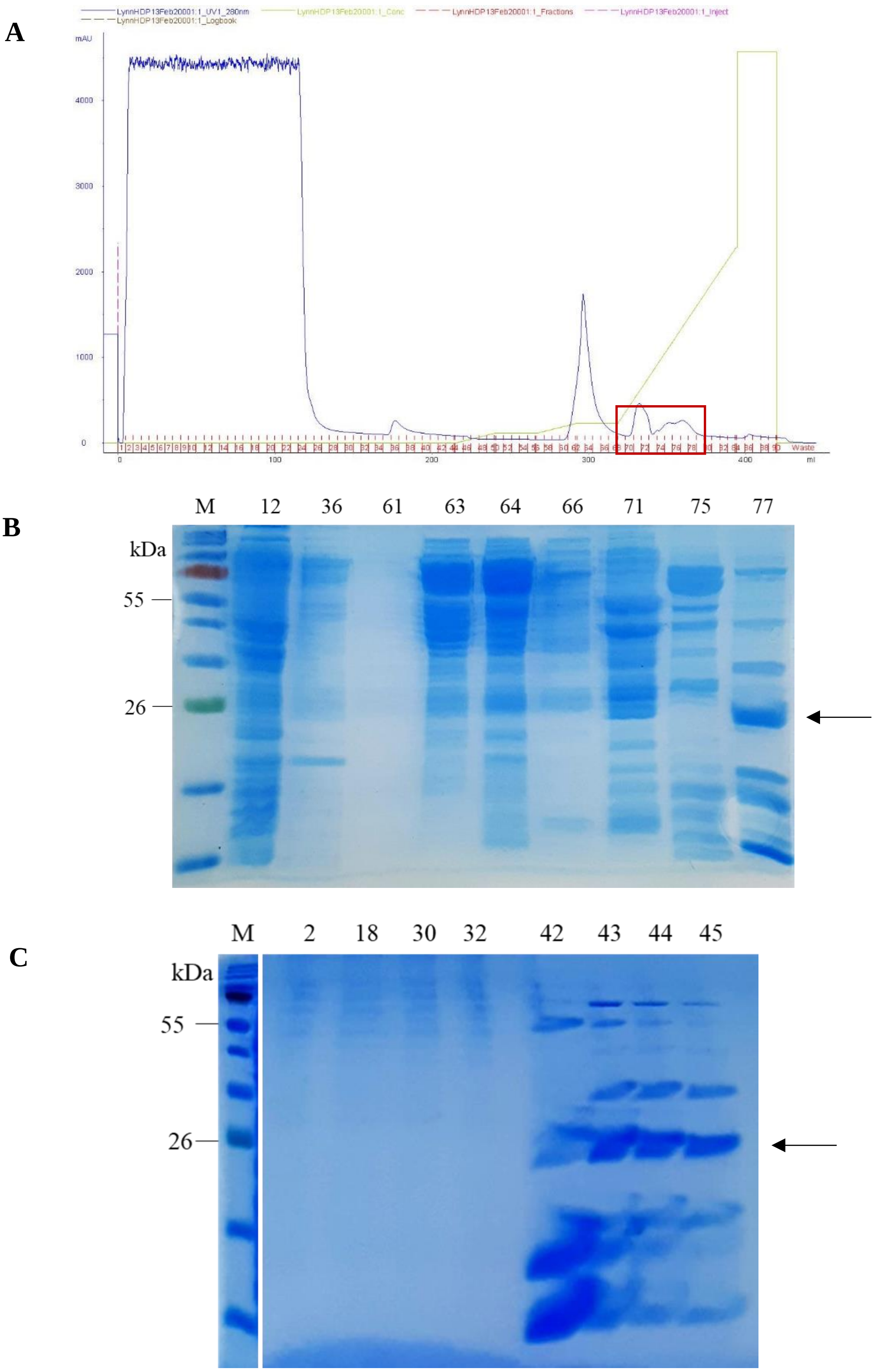


Figure 2.20: Purification of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) expressed in C41 cells.

(A) *Pf*HDP HiTrap™ HP purification profile. Red box indicates the elution peak. **(B)** *Pf*HDP HiTrap™ HP purification samples analysed via SDS-PAGE on a 10% gel. M: Molecular marker, HiTrap™ HP fractions analysed: 12, 36, 61, 63, 64, 66, 71, 75, 77. **(C)** *Pf*HDP HIC purification samples analysed via SDS-PAGE on a 10% gel. M: Molecular marker, HIC fractions analysed: 2, 18, 30, 32, 42, 43, 44, 45.

The elution fractions were analysed via SEC on a HiLoad 16/600 Superdex 200 pg column to separate the remaining contaminating proteins. A significant peak corresponding to *Pf*HDP was observed and analysis of the SEC fraction via SDS-PAGE revealed two distinct bands at approximately 70 kDa in earlier fractions and 25 kDa in later elution fractions (Figure 2.21).

The 25 kDa protein peak was pooled, concentrated to 300 µL and stored at -80 °C in glycerol at 0.48 mg/mL. A relatively low yield of protein was obtained compared to band intensity on the SDS-PAGE gel (Figure 2.21B) due to protein precipitation during concentration. Despite the addition of glycerol to all purification buffers and working at low temperatures, protein precipitation remains a common issue faced during the purification process. High salt conditions and octyl-glucoside have also been used to reduce protein precipitation with little success.

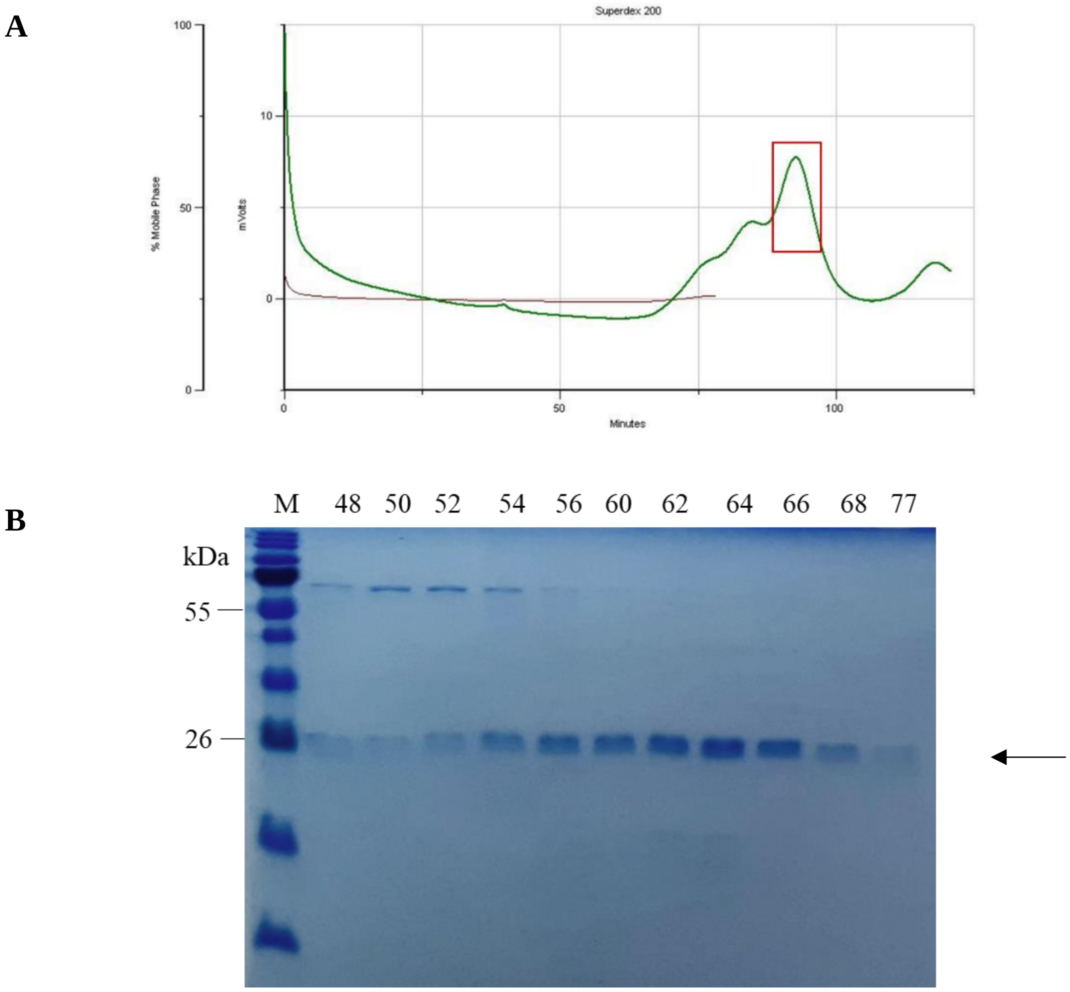


Figure 2.21: Size-exclusion chromatography of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP).

(A) HiLoad 16/600 Superdex 200 pg size-exclusion chromatography profile. The red box indicates the *Pf*HDP peak. **(B)** Size-exclusion chromatography fractions analysed via SDS-PAGE on a 10% gel. M: Molecular mass marker; fractions analysed: 48, 50, 52, 54, 56, 60, 62, 64, 66, 68, 77.

2.3.4 *Pf*HDP purification under denaturing conditions

*Pf*HDP has previously been purified successfully under denaturing conditions (Nakatani *et al*., 2013). Here, *Pf*HDP was expressed in BL21 cells in LB broth supplemented with 1% glucose as previously described (section 2.2.4). The protein was localised in inclusion bodies and was purified using His60 Ni gravity flow columns. *Pf*HDP was refolded via buffer exchange using a Sephadex PD-10 desalting column and the purification fractions were analysed via SDS-PAGE (Figure 2.22).

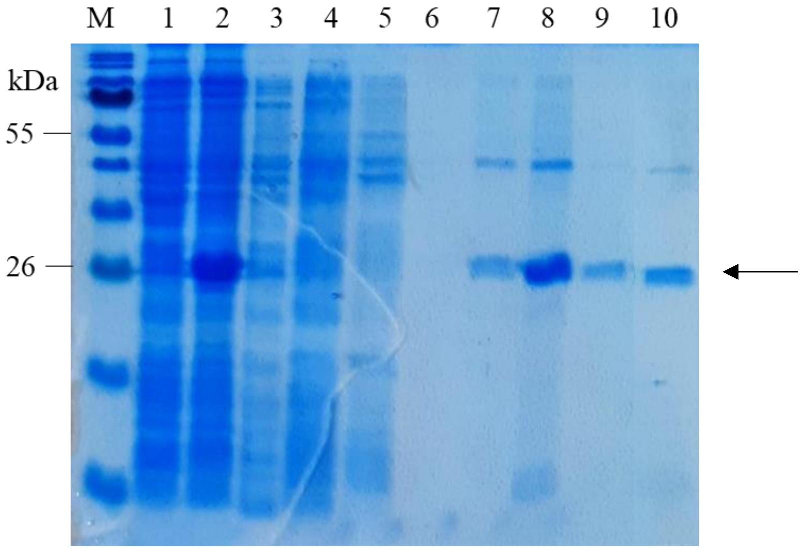
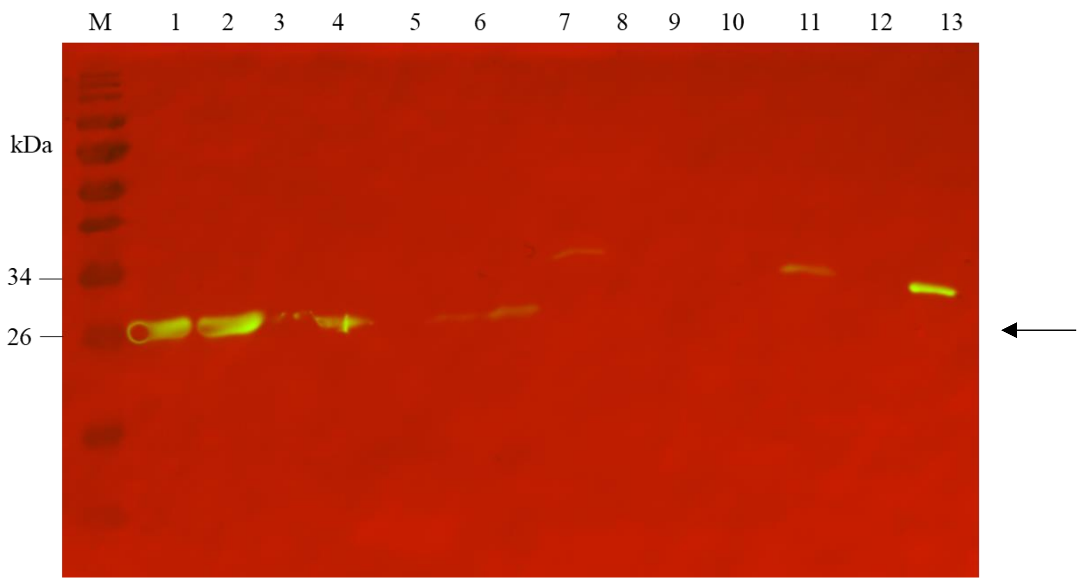


Figure 2.22: Purification of *Plasmodium falciparum* haem detoxification protein (*Pf*HDP) expressed in BL21 cells under denaturing conditions.

*Pf*HDP His60 Ni purification samples analysed by SDS-PAGE on a 10% gel. M: Molecular marker; 1: Uninduced whole-cell fraction, 2: Induced whole-cell fraction; 3: Cleared lysate; 4: Flow-through; 5: Wash fraction 1; 6: Wash fraction 2; 7: Elution fraction 1; 8: Elution fraction 2; 9: Elution fraction 3; 10: Refolded protein.

A distinct band was observed at 25 kDa after refolding and a very faint band at approximately 40 kDa, which is similar to previous purification profiles observed by Nakatani *et al*. (2013). *Pf*HDP purification under denaturing conditions yielded approximately 0.46 mg protein from 250 mL expression culture. As previously described, purification of *Pf*HDP under native conditions is a three-step, laborious process with low yield, which may explain why all previous work on *Pf*HDP using recombinant protein has involved protein purified under denaturing conditions, a two-step process with relatively high yield.

The *Pf*HDP purification samples (soluble and refolded) were analysed via western blot. In Figure 2.23, prominent luminescence was observed in the refolded *Pf*HDP purification samples, while a faint band was seen in the cleared lysate and flow-through fractions of the soluble *Pf*HDP purification samples. The band observed in the soluble *Pf*HDP HiTrap™ HP elution fraction and final soluble *Pf*HDP sample appeared slightly higher on the gel at ~34 kDA, raising questions as to the identity of the proteins in these fractions.



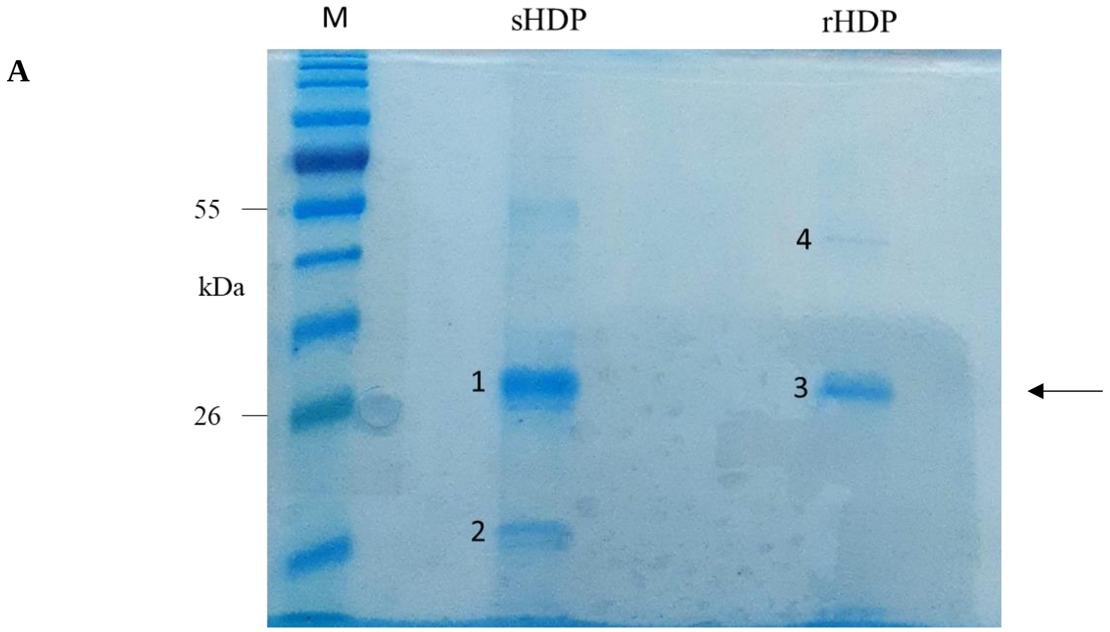
rHDP- refolded HDP; sHDP-soluble HDP

Figure 2.23: Western blot analysis of *Plasmodium falciparum* haem detoxification protein purification samples.

*Pf*HDP purification samples analysed by western blot using monoclonal anti-polyHistidine-peroxidase clone HIS-1 antibody. M: Molecular marker; 1: rHDP elution fraction; 2: rHDP pooled elution fractions; 3: rHDP desalted fraction; 4: final rHDP protein; 5: sHDP cleared lysate; 6: sHDP flow-through; 7: sHDP HiTrap HP elution fraction; 8: sHDP HIC elution fraction; 9: sHDP SEC fraction; 10: empty; 11: sHDP final protein; 12: empty; 13: his-tag positive control.

2.3.5 Mass Spectrometry

Mass spectrometry was used to confirm the identity of the proteins in the soluble *Pf*HDP and refolded *Pf*HDP samples purified from C41 cells as described above. Final purified soluble and refolded *Pf*HDP samples were analysed by SDS-PAGE (Figure 2.24A) and the protein bands were cut out and prepared for mass spectrometry as previously described (section 2.2.14).



B

Species	Protein ID (uniprot)	Molecular Weight (kDa)	Gene names	Signal (%)			
				Band 1	Band 2	Band 3	Band 4
P. falciparum	A0A144A6G1	24.3	*Pf*3D7_1446800	0	0	89	0
E. coli	P0A9B1	16.7	fur	0	98	0	0
E. coli	P0A9L1	20.1	slyD	95	0	10	0
E. coli	P0A6N3	43.3	tufA	0	0	0	99

Figure 2.24: *Pf*HDP samples analysed by mass spectrometry.
(**A**) Purified *Pf*HDP samples analysed by SDS-PAGE on a 10% gel. M: Molecular marker; sHDP: soluble HDP; rHDP: refolded HDP. 1-4 correspond to bands excised for mass spectrometry. **(B)**Table showing the identified proteins in each sample.

The mass spectrometry results (Figure 2.24B) show that the protein thought to be *Pf*HDP purified under native conditions (Band 1) is FKBP-type peptidyl-prolyl cis-trans isomerase (SlyD) – a 20.1 kDa *E. coli* protein with chaperone activity that helps prevent aggregation of unfolded/partially folded proteins and also promotes correct folding (Bernhardt *et al*., 2002). The protein appears to run at a slightly higher molecular weight when analysed by SDS-PAGE. Additionally, it also has peptidyl-prolyl cis-trans isomerase (PPIase) activity which catalyses the cis-trans isomerization of Xaa-Pro (unspecified amino acid – proline) bonds in peptides, speeding up the protein folding process (Bernhardt *et al*., 2002). It is a common contaminant found in proteins purified by IMAC as the C-domain contains 15 histidine residues. Band 3 was identified as *Pf*HDP (uniport ID: A0A144A6G1) with some SlyD contamination in the sample which suggests the protein may help correctly fold *Pf*HDP during the purification process.

Band 2, which co-purified with band 1 is the ferric uptake regulator (FUR) protein which uses iron as a co-factor and controls the expression of enzymes that protect against damage by reactive oxygen species thus maintaining iron homeostasis in *E. coli* (Troxell and Hassan, 2013). Band 4 was identified as elongation factor thermo unstable (EF-Tu), which is a G-protein that transports aminoacylated tRNA's (transfer ribonucleic acid) to the ribosome during protein biosynthesis (Harvey *et al*., 2019). Elongation factor Tu makes up almost 6% of the total protein expressed in *E. coli* (Furano, 1975). The abundance and function of EF-Tu indicate its importance to *E. coli* and may explain why some of the protein co-purifies with *Pf*HDP.

2.4 Conclusion

Efforts to increase the expression levels of soluble *Pf*HDP were successful using ArcticExpress cells in TB media containing 5-ALA but isolating the protein from other contaminants proved challenging. Initially this was thought to be due to chaperones, coeluting with *Pf*HDP, masking the his-tag and preventing binding to the Ni^{2+}-NTA column. As a result, expression was carried out in C41 cells (lacking the Cpn60 and Cpn10 chaperones) using a previously optimised method. However, protein purified from C41 cells that was initially thought to be soluble *Pf*HDP was later shown to be SlyD, a common contaminant found in proteins purified by IMAC. The results showed that soluble *Pf*HDP expressed in both C41 and ArcticExpress cells, did not bind to the Ni^{2+}-NTA chromatography columns. Since *Pf*HDP can be purified under denaturing conditions, the inability of soluble *Pf*HDP to bind to the Ni^{2+}-NTA column suggests that in its native form, the protein is folded in a way that masks the N-terminal his-tag – discussed further in Chapter 3. A possible alternative may be to move the his-tag to the C-terminus but since such a construct has been used previously (Jani *et al*., 2008) and the protein was also purified under denaturing conditions, it may not be effective. This is most likely why all published data on recombinant *Pf*HDP has been on HDP purified from inclusion bodies. An alternative expression system was considered – insect cell expression system – however, the data indicates that the expression levels were not the issue. Thus it was likely that similar challenges would be encountered at the purification stage. Exhaustive efforts to purify recombinant *Pf*HDP from the soluble fraction expressed in the *E. coli* system have been unsuccessful. On the other hand, purification of *Pf*HDP under denaturing conditions was achieved using the guidance of previously published methods and due to the ease and relatively high yield purification under denaturing conditions, further studies on *Pf*HDP were conducted on refolded HDP.

Chapter 3

Plasmodium falciparum Haem Detoxification Protein function evaluation, protein crystallization attempts and *in silico* structure prediction

3.1 Introduction

The β-hematin formation assay is an important tool for evaluating antimalarials that potentially inhibit haem metabolism (Sinha *et al*., 2017). The procedure involves evaluating β-hematin formation from hemin or hematin. This can be done using various methods such as spectrophotometry, Fourier-transform infrared spectroscopy, fluorometry, high-performance liquid chromatography or radioisotope evaluation (Sinha *et al*., 2017). In vitro β-hematin formation is performed at the physiological pH and temperature of the parasite digestive vacuole and can only occur in the presence of biological factors such as proteins, lipids or pre-formed β-hematin (Kumar *et al*., 2007).

X-ray crystallography is the most widely used method in 3D protein structure determination (Srivastava *et al*., 2018). Currently, there are approximately 183,980 crystal structures available in the Protein Data Bank (PDB) – 160,690 of which are protein structures and of these, 88 % were solved by x-ray diffraction (rcsb.org/stats/summary, 2021). Protein structures often provide numerous insights into protein function and its role in the biological system (Papageorgiou and Mattson, 2014).

Obtaining protein crystals is predominantly a ‘trial and error’ process in which proteins in a precipitant solution are precipitated and sometimes form crystals by transfer of water from the protein to the precipitant solution (Rondeau and Schreuder, 2015). Vapour diffusion is the most commonly used technique in protein crystallisation either by the hanging drop method or the sitting drop method, illustrated in Figure 3.1A, where water diffuses from the protein drop to the precipitant solution resulting in the formation of crystals (Papageorgiou and Mattson, 2014).

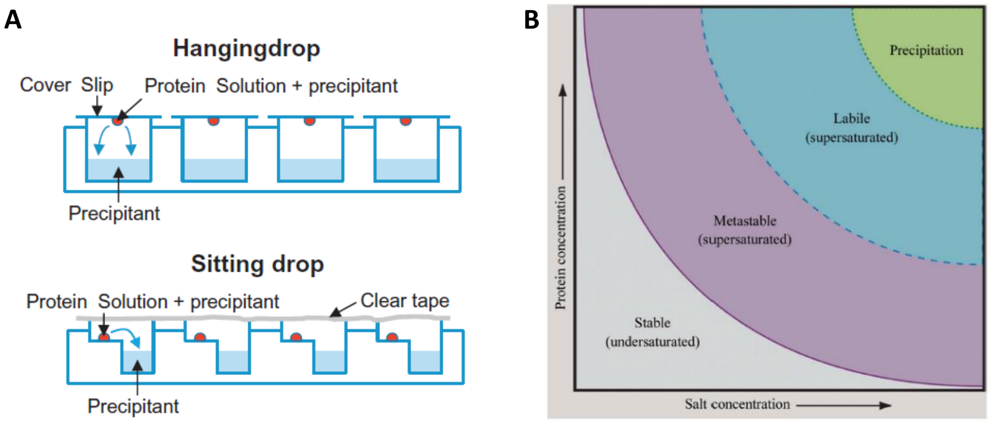


Figure 3.1: Illustration of the protein crystallisation process.
A) Vapour diffusion crystallization techniques (Rondeau and Schreuder, 2015). **B**) Crystallization phase diagram (McPherson and Gavira, 2014).

The phase diagram (Figure 3.1B) is a representation of the various states possible in a protein crystallization environment. In the stable state, the protein is undersaturated and thus soluble (McPherson and Gavira, 2014). For crystallization to occur, the protein must be in a supersaturated state i.e., the protein (in solution) at a certain quantity reaches the solubility limit and equilibrium is restored by the formation of a solid state such as crystals (McPherson and Gavira, 2014). In the metastable and labile state, the protein is supersaturated, and crystals can form (McPherson and Gavira, 2014). In the metastable state, crystal growth occurs with no nucleation while in the labile state, nucleation followed by fast crystal growth often occurs (Krauss *et al*., 2013). Protein crystals contain an average of 50 % (v/v) solvent thus the protein structure determined by crystallography are highly similar to the structure of the protein in solution (McPherson and Gavira, 2014). However, the high solvent content also results in the crystals being soft and brittle, sensitive to dehydration and limited in size (Rondeau and Schreuder, 2015).

Despite the success of protein structure determination by x-ray crystallography over the last six decades, there are still numerous challenges faced when trying to obtain protein crystals that can make it a laborious, time-consuming process. Challenges often faced in the protein crystallization process include: protein solubility, unstable crystals as a result of protein disorder, degradation and product limitation (Srivastava *et al*., 2018). Even with advances such as commercial screening kits and crystallization robots, protein crystallization is still a labour-intensive, error-prone process. If protein structure determination by x-ray diffraction is unsuccessful, *in silico* structure prediction is an alternative that can also provide insights into protein function (Sleator and Walsh, 2010).

Protein structure prediction is based on the theory that proteins from the same evolutionary family have similar sequences and will often also have similar 3D structures (Deng *et al.*, 2018). There are three main methods of protein structure prediction – comparative modelling, threading and free modelling - which vary in accuracy as illustrated in Figure 3.2 (Zhang, 2009).

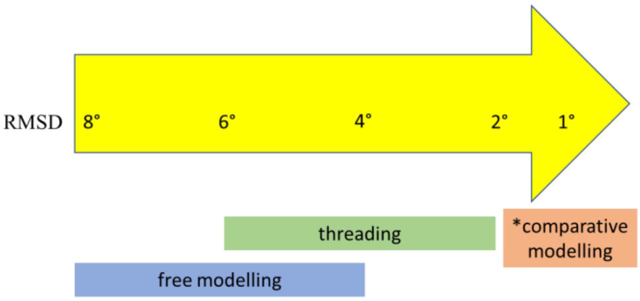


Figure 3.2: Accuracy of structure prediction methods (adapted from Baker and Sali, 2001; Zhang, 2009).
* For highly homologous templates
RMSD – root mean square deviation from experimental structure in Angstrom (°)

Comparative modelling and threading are template-based prediction strategies which are based on comparing a 'related' template protein by sequence and protein fold respectively (Deng *et al.*, 2018). Free modelling builds structures based on physics and is often limited to small proteins for successful prediction (Zhang, 2009). Protein structure prediction methods are compared and assessed by the CASP (critical assessment of protein structure prediction) experiments wherein predicted protein structures are compared to their experimentally solved, yet to be published, counterparts (Jumper *et al.*, 2021).

This chapter focuses on evaluation of recombinant *Pf*HDP using the β-hematin formation assay, protein crystallization attempts and *in silico* structure prediction of *Pf*HDP.

3.2 Methods

3.2.1 β-hematin formation assay

The β-hematin formation assay was performed as described by Sullivan *et al.* (1996) with modifications.

A 10 mM hemin (Fe^{3+} protoporphyrin IX with a coordinating chlorine) stock solution was prepared in *N*, *N*-dimethylformamide (Sigma). Each reaction was set up in a 1 mL final volume containing 600 µM hemin and 2 µM rHDP in 0.5 M sodium acetate buffer (pH 5.2). A no-protein control was set up using volume of buffer equal to the rHDP in assay reaction. The reactions were incubated at 37 °C for 1 hour. SDS was added to a final concentration of 0.1 % (v/v). The samples were centrifuged at 15,000 *xg* for 10 minutes at room temperature. The supernatant was discarded and the pellet was washed twice with HEPES buffer (0.02 M, pH 7.4) containing 5 % (v/v) pyridine. The resulting pellet was washed three times in 1 mL Milli-Q® water. The precipitate was washed with 500 µL acetone/methanol solution at a 1:9 (v/v) ratio, dried overnight and analysed by infrared (IR) spectroscopy.

3.2.2 Imaging of rHDP by electron microscopy

Enriched rHDP was visualised using a transmission electron microscope (TEM) with help from Mohamed Jaffer (UCT). Carbon-coated copper grids were placed in a glow discharger (Electron Microscopy Sciences) at 25 mA for 25 seconds. A 3 µl aliquot of the purified protein was placed on the grid for 30 seconds. Excess sample was wicked off. The grid was washed twice with water and stained with 2 % (v/v) aqueous uranyl acetate. The grid was left to airdry and kept at room temperature. The grid viewed under a Tecnai F20 Transmission Electron Microscope operated at 200 kV. Images were rendered by a 4k x 4k CCD camera (Gatan US400).

3.2.3 Crystal screening and optimisation

Initial crystal screening was performed using a Morpheus MD1-46 screening kit (Molecular Dimensions) via sitting drop method using the mosquito® (SPT Labtech). Screening was done in 96-well plates in which each well contained a 1:1 protein:precipitant solution. Conditions that produced crystals were used in fine screens. Two crystallisation conditions were used in the fine screens: buffer system 1 (0.1 M imidazole, 0.1 M MES monohydrate acid (pH 6.5), 0.06 M magnesium chloride hexahydrate, 0.06 M calcium chloride dihydrate) + precipitant mix 2 (40 % v/v ethylene glycol, 20 % w/v PEG 8,000) and buffer system 2 (0.1 M sodium HEPES, 0.1 M MOPS acid (pH 7.5), 0.06 M magnesium chloride hexahydrate, 0.06 M calcium chloride dihydrate) + precipitant mix 1 (40 % v/v PEG 550 MME, 20 % w/v PEG 20,000). Conditions were optimised by varying conditions of precipitant mix in each buffer system from 38% to 59% in 3% increments and varying the ratio of protein to precipitant mix (2:1; 1.8:1.2; 1.6:1.4; 1.4:1.6; 1.2:1.8, and 1:2).

The crystal screen was repeated using the Morpheus MD1-46 screening kit and 100 µM hemin by the sitting drop method. The drops in subwell 1 and 2 contained a 1:1:0.25 ratio of protein:precipitant solution:100 µM hemin. The drops in subwell 3 contained a 1:1:0.25 ratio of buffer:precipitant solution:100 µM hemin. The crystals observed were mounted on a cryoloop and analysed at Diamond Light Source by x-ray diffraction on the I04 beamline with assistance from Phillip Venter (UCT).

3.2.4 *In silico* protein structure prediction

The haem detoxification protein sequence was evaluated by PSIPRED.

Plasmodium falciparum HDP models were generated using I-TASSER and RoseTTAFold structure prediction programs by Kyllen Dilsook (UCT). A predicted HDP protein structure was also found of the AlphaFold Protein Structure Database. The three predicted structures were aligned in Maestro and the RMSD was determined.

3.3 Results and Discussion

3.3.1 β-hematin formation assay

Enriched rHDP was evaluated for β-hematin formation activity based on an assay developed by Sullivan *et al*. (1996) with modifications. Three separate reactions were set up: 1 and 2 containing rHDP and 600 µM hemin and a control reaction (3) containing buffer and 600 µM hemin. The reactions were incubated at 37 °C for 1 hour. To quench the reaction, 0.1 % (v/v) SDS was added to reactions 1 and 3, no SDS was added to reaction 2. The reactions were centrifuged, washed in HEPES buffer (0.02 M, pH 7.4) containing 5 % (v/v) pyridine, MilliQ® water and acetone/methanol at a 1:9 (v/v) ratio. The samples were dried overnight and analysed by IR spectroscopy. Infrared spectroscopy of β-hematin and haemozoin shows peaks with characteristic peaks at approximately 1210 cm^{-1} and 1660 cm^{-1} (Sullivan *et al*., 1996).

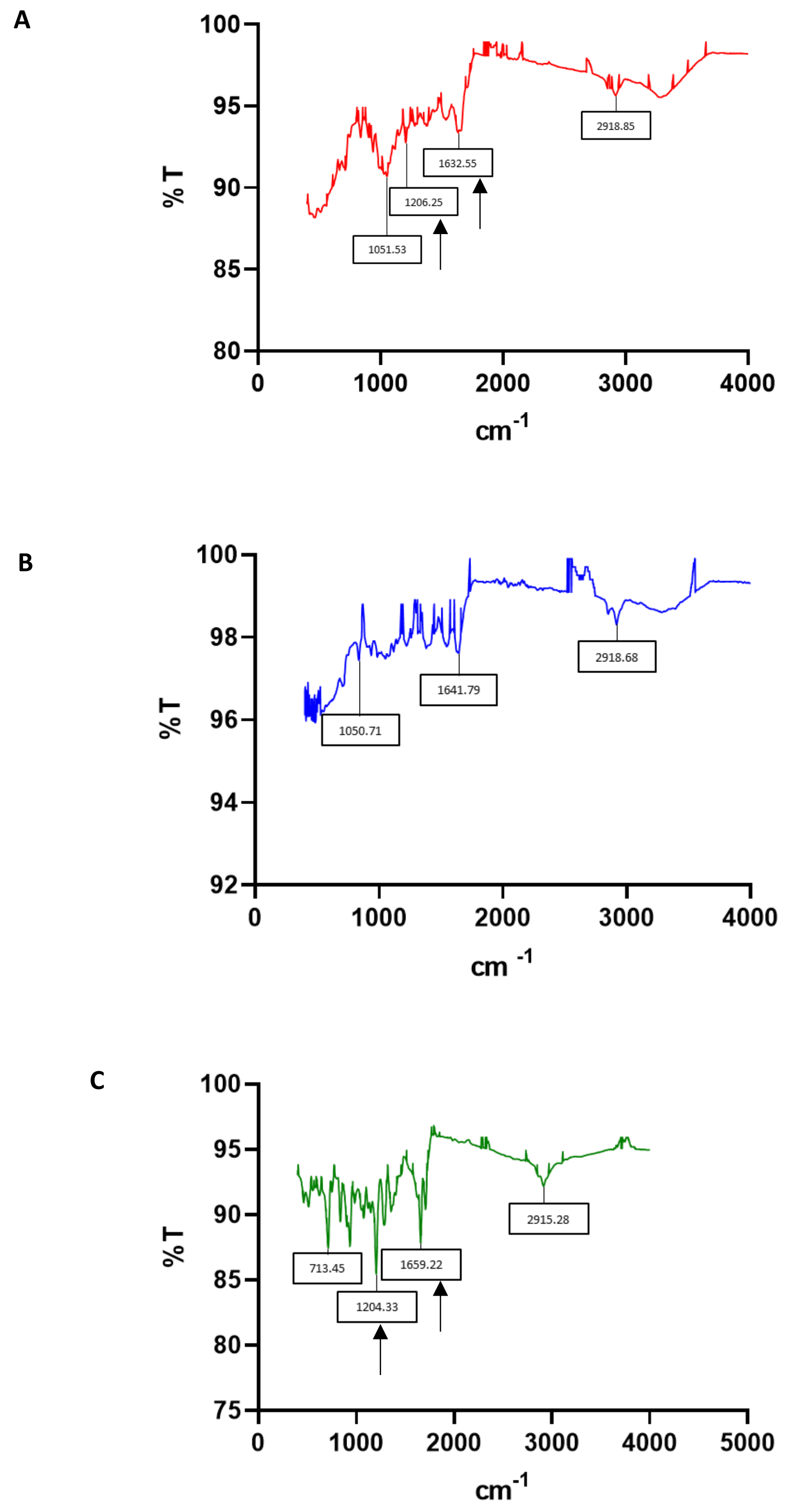


Figure 3.3: Infrared spectra of β-hematin formation.
A) rHDP + 600 µM hemin reaction quenched with 0.1 % SDS; **B)** rHDP + 600 µM hemin reaction not quenched with 0.1 % SDS and **C)** Buffer + 600 µM hemin reaction quenched with 0.1 % SDS.

The spectra in Figure 3.3 shows the presence of β-hematin in the HDP + SDS and buffer + SDS samples (Figure 3.3A and 3.3C) characterised by peaks at 1632 cm^{-1} and 1206 and 1659 and 1204 cm^{-1} respectively. The lower-than-expected 1600 peak in Figure 3.3A may be due to the presence of water in the sample prior to evaluation which has previously been observed by Egan *et al.* (2001). The absence of a peak at approximately 1200-1211 cm^{-1} and the peak observed at 1641 cm^{-1} instead of at approximately 1655-1664 cm^{-1} in Figure 3.3B indicates the absence of β-hematin in the sample. This data suggests that β-hematin formation is mediated by SDS in the assay reaction rather than by HDP as previously reported.

3.3.2 Secondary and Tertiary structure prediction of *Pf*HDP

Plasmodium falciparum HDP was expressed, purified under denaturing conditions, refolded, stained using uranyl acetate and visualised by transmission electron microscopy (Figure 3.4).

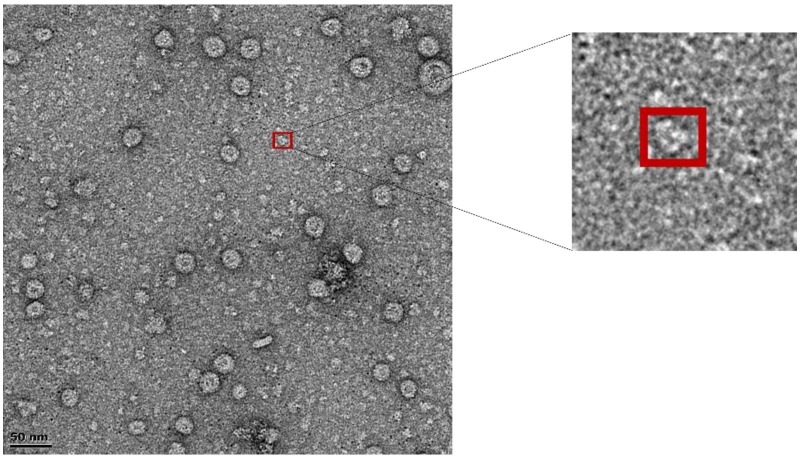


Figure 3.4: Enriched rHDP viewed by transmission electron microscopy (red box showing single protein particle approximately 15 nm in diameter).

Limitations in imaging due to the small size of the protein restricted the structural information obtained by TEM. Large particles were observed on the grid possibly due to aggregation of *E.coli* tufA (Elongation factor Tu 1) which co-purifies with rHDP as shown in Figure 2.24. However, the grid also indicates rHDP does not form a tertiary complex and is likely a monomer as the protein structures observed were singular with defined edges.

The HDP amino acid sequence was analysed by PSIPRED (a computational secondary structure prediction server) to provide insights into HDP’s secondary and tertiary structure.

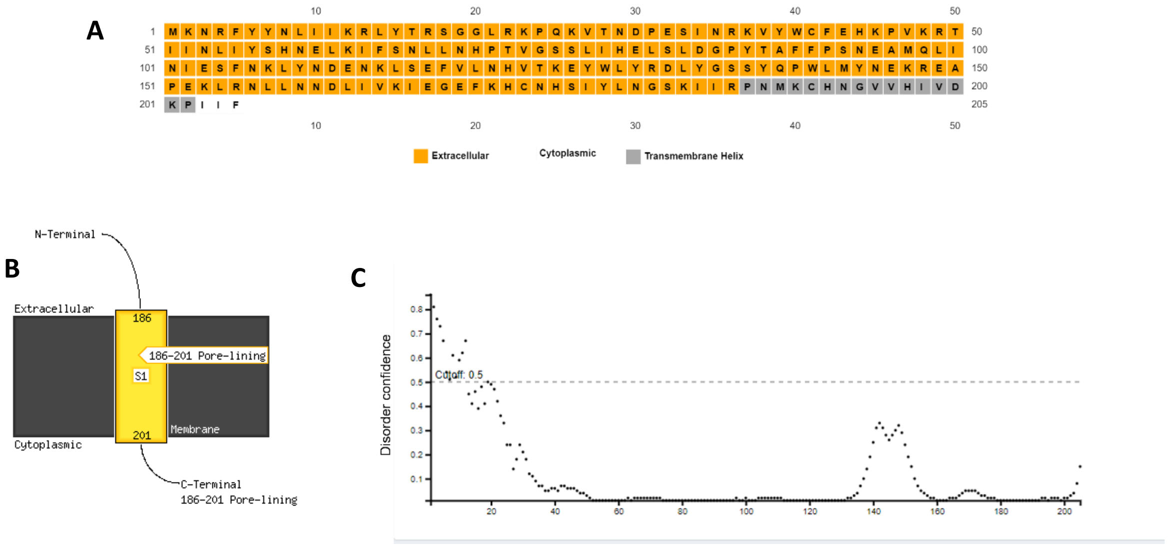


Figure 3.5: PSIPRED analysis of the HDP amino acid sequence.
A) MEMSAT sequence plot, **B)** MEMSAT schematic diagram, **C)** DISOPRED plot.

MEMSAT is a membrane protein prediction method that predicts the potential length, topological orientation and location of transmembrane proteins (Nugent and Jones, 2009). MEMSAT prediction of the HDP amino sequence (Figure 3.5A and 3.5B) suggests most of the protein is extracellular with 16 pore-lining amino acids and 3 cytoplasmic amino acids which previously has not been reported. As previously discussed in Chapter 1, Jani *et al*. (2008) reported HDP was secreted into the red blood cell cytosol and trafficked to the food vacuole via the cytostome-mediated pathway. The difficulty in obtaining soluble HDP supports the membrane protein prediction as membrane proteins often express at low levels in recombinant systems, aggregate/form inclusion bodies, misfold and have low stability (Hering *et al*., 2020). Many of the listed issues were faced when trying to express and purify *Pf*HDP.

Interestingly, TMHMM analysis of *Pf*HDP sequence suggested the entire protein was extracellular (Appendix H) and predicted to be a globular protein. Further studies into this prediction are required however, globular proteins often play a role in cell signaling, regulation, molecule transport and binding (Shen, 2019).

DISOPRED predicts the probability of ordered and disordered amino acids within a protein sequence (Jones and Cozzetto, 2015). The DISOPRED plot (Figure 3.5C) shows the first approximately 35 amino acids are highly disordered as well as residues 137-147 in the sequence as well as some disorder at residues 165-178 and 203-205. The highly disordered regions on the N-terminus may be the reason for the low his-tag to Ni^{2+}-NTA binding observed between recombinant HDP and the HiTrap™ purification column in Chapter 2. This region may cause the protein to fold in a way that results in masking of the N-terminus or alternatively masking may also occur due to other small proteins in the lysate binding to the disordered regions.

Following this, determination of rHDP tertiary structure by crystallization was attempted. An initial crystallization screen was set up using the Morpheus MD1-46 kit in a 96-well plate. Only three conditions resulted in crystal-like solids. Using refolded HDP, small crystal-like deposits were observed after 5 days Figure 3.6).

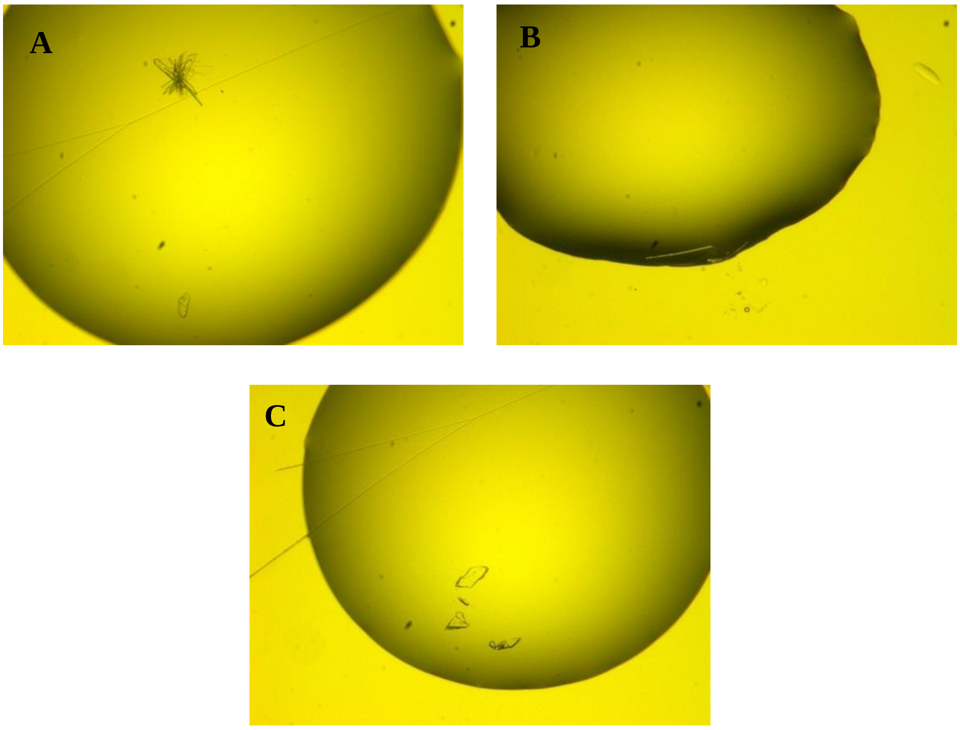


Figure 3.6: Refolded *Plasmodium falciparum* haem detoxification protein crystals.
HDP crystals formed in **(A)** 0.06 M magnesium chloride hexahydrate, 0.06 M calcium chloride dihydrate, 0.1 M imidazole, 0.1 M MES monohydrate acid (pH 6.5), 50 % (v/v) precipitant mix 4 (25% v/v MPD, 25% PEG 1000, 25% w/v PEG 3350); **(B)** 0.06 M magnesium chloride hexahydrate, 0.06 M calcium chloride dihydrate, 0.1 M sodium HEPES, 0.1 M MOPS acid (pH 7.5), 50 % v/v precipitant mix 1 (40 % v/v PEG 550 MME, 20 % w/v PEG 20,000); **(C)** 0.06 M magnesium chloride hexahydrate, 0.06 M calcium chloride dihydrate, 0.1 M imidazole, 0.1 M MES monohydrate acid (pH 6.5), 50 % (v/v) precipitant mix 2 (40 % v/v ethylene glycol, 20 % w/v PEG 8,000).

Two fine screens were carried out using the available reagents: Buffer system 1 (0.1 M imidazole, 0.1 M MES monohydrate acid (pH 6.5), 0.06 M magnesium chloride hexahydrate, 0.06 M calcium chloride dihydrate) + Precipitant Mix 2 (40 % v/v ethylene glycol, 20 % w/v

PEG 8,000) and Buffer system 2 (0.1 M sodium HEPES, 0.1 M MOPS acid (pH 7.5), 0.06 M magnesium chloride hexahydrate, 0.06 M calcium chloride dihydrate) + Precipitant Mix 1 (40 % v/v PEG 550 MME, 20 % w/v PEG 20,000), using freshly purified rHDP at 0.5 mg/ml at a varying range of precipitant concentration.

Under Buffer system 1+ Precipitant mix 2 conditions, shard-like crystals were observed after 4 days (Figure 3.7). The crystals formed in clumps at higher protein:precipitant ratios. Although more crystals were observed, the size of the crystals did not increase.

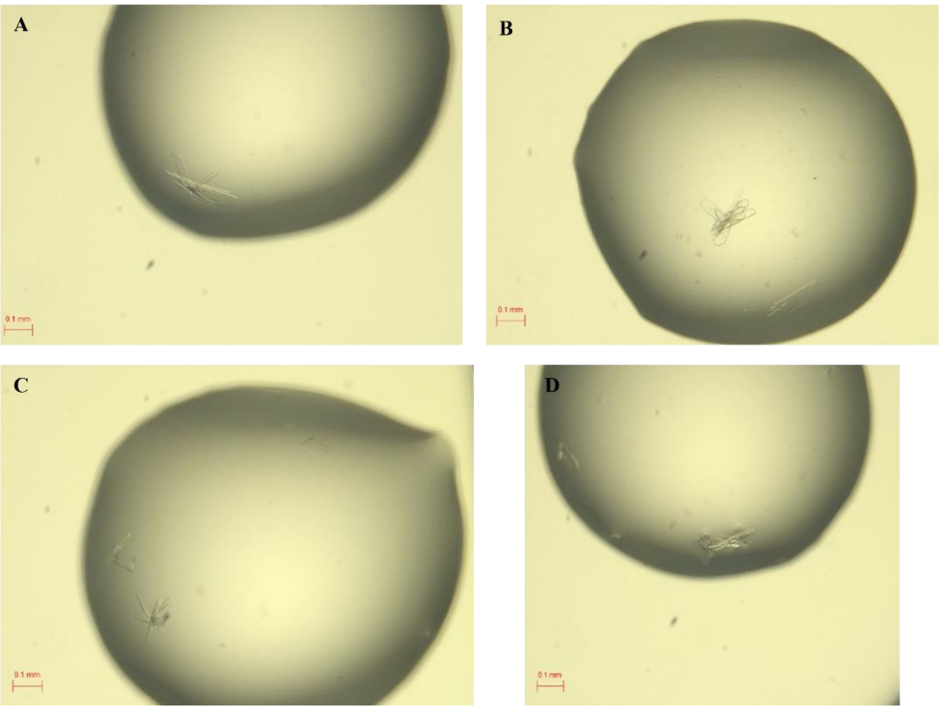


Figure 3.7: HDP crystals observed in Buffer system 1+ Precipitant mix 2.

Under Buffer system 2 + Precipitant mix 1 conditions, small cube-shaped crystals were observed (Figure 3.8). These crystals were viewed under a polarizer they were weakly birefringent – characteristic of protein crystals.

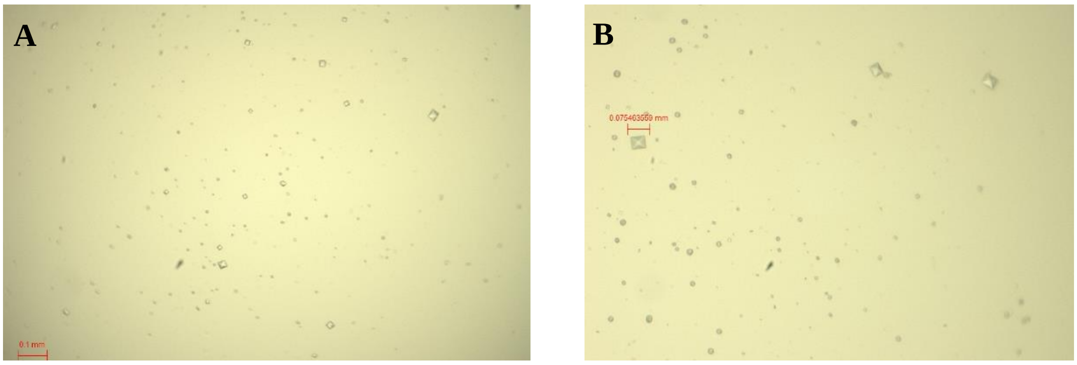


Figure 3.8: Crystals observed in Buffer system 2 + Precipitant mix 1.

The crystals observed in both conditions were analysed by X-ray crystallography however, the crystals were very small, thus mounting them on a cryoloop was difficult and x-ray crystallography was difficult as the beam could not centre on the crystal.

A new crystallization screen was set up in the presence of (11 µM) hemin as an optimisation strategy as the protein is reported to have a haem binding domain. The resulting crystals were shard-like and relatively larger than previously observed crystals (Figure 3.9). The crystals were mounted on a cryoloop and analysed at the Diamond Light Source.

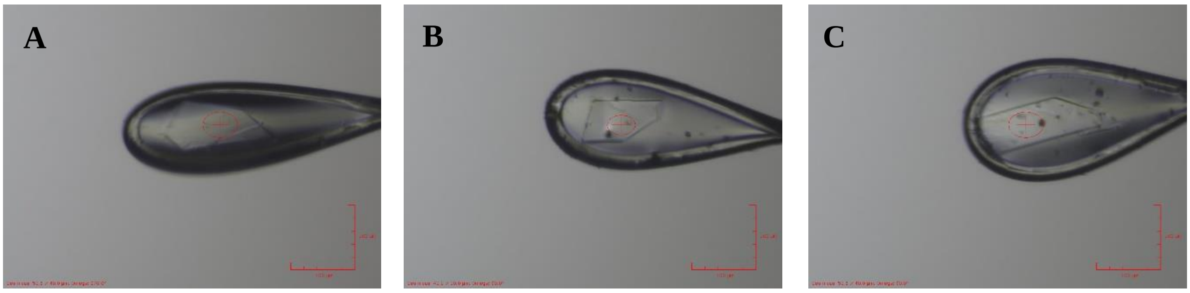


Figure 3.9: rHDP crystals mounted on a cryoloop under the Diamond Light Source Viewer.

A) and **B)** 0.06 M magnesium chloride hexahydrate, 0.06 M calcium chloride dihydrate, 0.1 M sodium HEPES, 0.1 M MOPS acid (pH 7.5), 37.5 % (v/v) precipitant mix 4 (25% v/v MPD, 25% PEG 1000, 25% w/v PEG 3350); **C)** 0.06 M magnesium chloride hexahydrate, 0.06 M calcium chloride dihydrate, 0.1 M imidazole, 0.1 M MES monohydrate acid (pH 6.5), 30 % (v/v) precipitant mix 2 (40 % v/v ethylene glycol, 20 % w/v PEG 8,000).

Analysis of these crystals showed poor diffraction data, which was not suitable for downstream processing, possibly due to disorder within the protein. This suggests further optimisation would be required to improve the quality of the crystals e.g. removal of the predicted N-terminal disordered region (Figure 3.5). Additional optimisation conditions could also include change in temperature, increased protein concentration as well as using an alternative crystallization method (Papageorgiou and Mattsson, 2014).

As the HDP crystals obtained did not exhibit adequate diffraction, *in silico* protein structure prediction of HDP was performed using 3 different protein structure prediction tools: RoseTTAFold, I-TASSER and the AlphaFold database. RoseTTAFold is a deep learning software tool that considers how amino acids interact with one another, patterns in protein sequences and the protein's possible 3D structure in order to build a model – known as a 'three-track' neural network (Baek *et al*., 2021). I-TASSER is a server that generates automated protein structure prediction by using structural templates from the PDB by a multiple threading approach and constructs models by iterative structure-based assembly simulations (Yang and Zhang, 2015). AlphaFold2 is a novel machine learning method that uses the physical and biological knowledge of proteins as well as multiple sequence alignments with homologues in its neural network to build highly accurate protein structure models (Jumper *et al*., 2021).

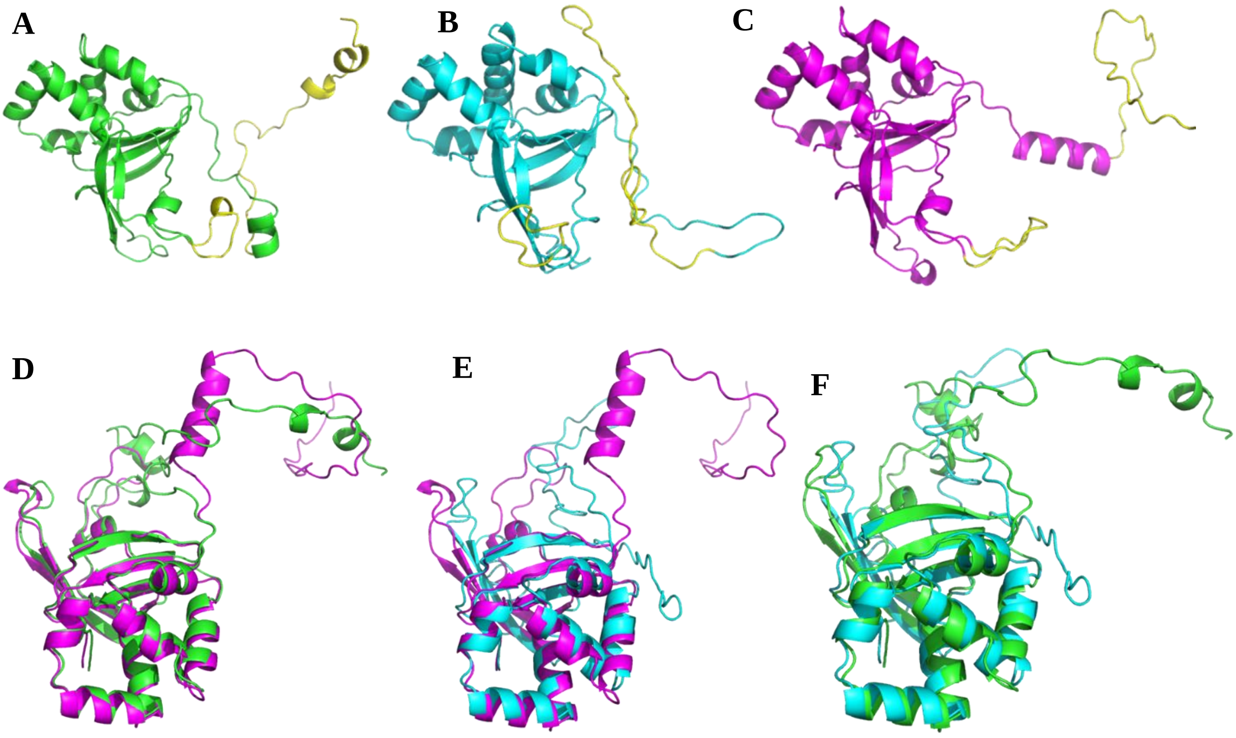


Figure 3.10: Predicted *Plasmodium falciparum* HDP tertiary structures.
Structures generated by **A)** RoseTTAFold, **B)** I-TASSER and **C)** AlphaFold Database. Yellow indicates predicted disordered regions within the protein. **D)** Aligned RoseTTAFold and AlphaFold structures, **E)** Aligned I-TASSER and AlphaFold structures, **F)** Aligned I-TASSER and RoseTTAFold structures.
*RoseTTAFold and I-TASSER models generated by Kyllen Dilsook

Comparison of the three generated models shows the models have a similar core structure and the highly disordered regions, amino acids 1-30 and 137-147, have no predicted features at all.

The model predicted by RoseTTAFold (Figure 3.10A) has a confidence score of 0.73 (where 0 is low confidence and 1.0 is very high confidence based on the local distance difference test [lDDT]). The lDDT score is a measure of the distance difference between the atoms in a model compared to its reference structure (Mariani *et al*., 2013). The relatively high confidence score observed suggests the model predicted is of moderate accuracy. The error per residue (Appendix Figure B1) shows high error in the expected regions amino acid residues 1- 45 and at approximately 137- 147 as well as at 165-175 and 203-205 – similar to the PSIPRED prediction (Figure 3.5C).

The model predicted by I-TASSER (Figure 3.10B) had a confidence score (C-score) of -1.55 where the score ranges between -5 (low confidence) and 2 (high confidence) based on the threading alignments and the convergence of the server's structural assembly refinement simulations (Roy *et al*., 2010). The server's comparison of the predicted structure with the PDB library showed PDB entry 5YJG – human periostin- had the closest structural similarity. Human periostin is a 90 kDa protein - containing four conserved fasciclin I domains - involved in remodelling the extracellular matrix by interacting with itself and other proteins (Liu *et al*., 2018). Ligand binding site prediction showed zinc as a possible ligand binding at amino acids 118, 122 and 202, however, the confidence score of this prediction was 0.10 over a 0-1 range where a higher score indicates a more reliable prediction thus the likelihood of zinc binding to HDP is low. Notably, despite the reported haem binding function of HDP, iron was not predicted to bind at any site on the structure.

The predicted AlphaFold model (Figure 3.10C) was evaluated by per-residue confidence scores between 0 and 100 where > 90 = very high confidence, 90-70 = confident, 70-50 = low confidence and <50 = very low confidence. The N-terminus of the protein (residues 1-26) and residues 137–147 were modelled with very low-low confidence as expected for disordered regions while the remaining structure was modelled with high to very high confidence.

The three HDP models were aligned using Maestro and the RMSD and alignment scores were determined. Alignment of the RoseTTAFold and AlphaFold structures had an RMSD score of 2.432° and an alignment score of 0.239, the I-TASSER and AlphaFold alignment had an RMSD score of 2.180° and an alignment score of 0.195 and the I-TASSER and RoseTTAFold structures had an RMSD score of 2.448° and an alignment score of 0.241. Overall, the AlphaFold and I-TASSER alignment showed these models have the highest similarity as they had the lowest RMSD and alignment score (lower is better).

Critical Assessment of Protein Structure Prediction (CASP) is considered the gold standard method for assessing the accuracy of structure prediction (Jumper *et al*., 2021). CASP14 (2020) evaluation of structure prediction methods showed AlphaFold structures were the most accurate, ranking 1st, I-TASSER was ranked 9th and RoseTTA was ranked 18th. Despite the different rankings, the generated models were highly similar and only the highly disordered protein regions were significantly different.

Previously Nakatani *et al*. (2014) reported a suggested mechanism for haemozoin formation by HDP (Figure 1.3) in which two haem molecules bind to His122 and His175 and His172 and His197 properly align the haem molecules for haemozoin growth to occur. The formation of β-hematin was described by Pagola *et al*. (2000). β-hematin dimer formation occurs through iron-oxygen bonds between two haem molecules in which the distance between the iron atom in one haem molecule and the oxygen atom in the second haem molecule is within 1.90 – 2.49 Å. Subsequently, β-hematin dimers are linked to each other by hydrogen bonds between oxygen atoms of adjacent dimers where the distance between the dimers is within 3.25 – 3.97 Å.

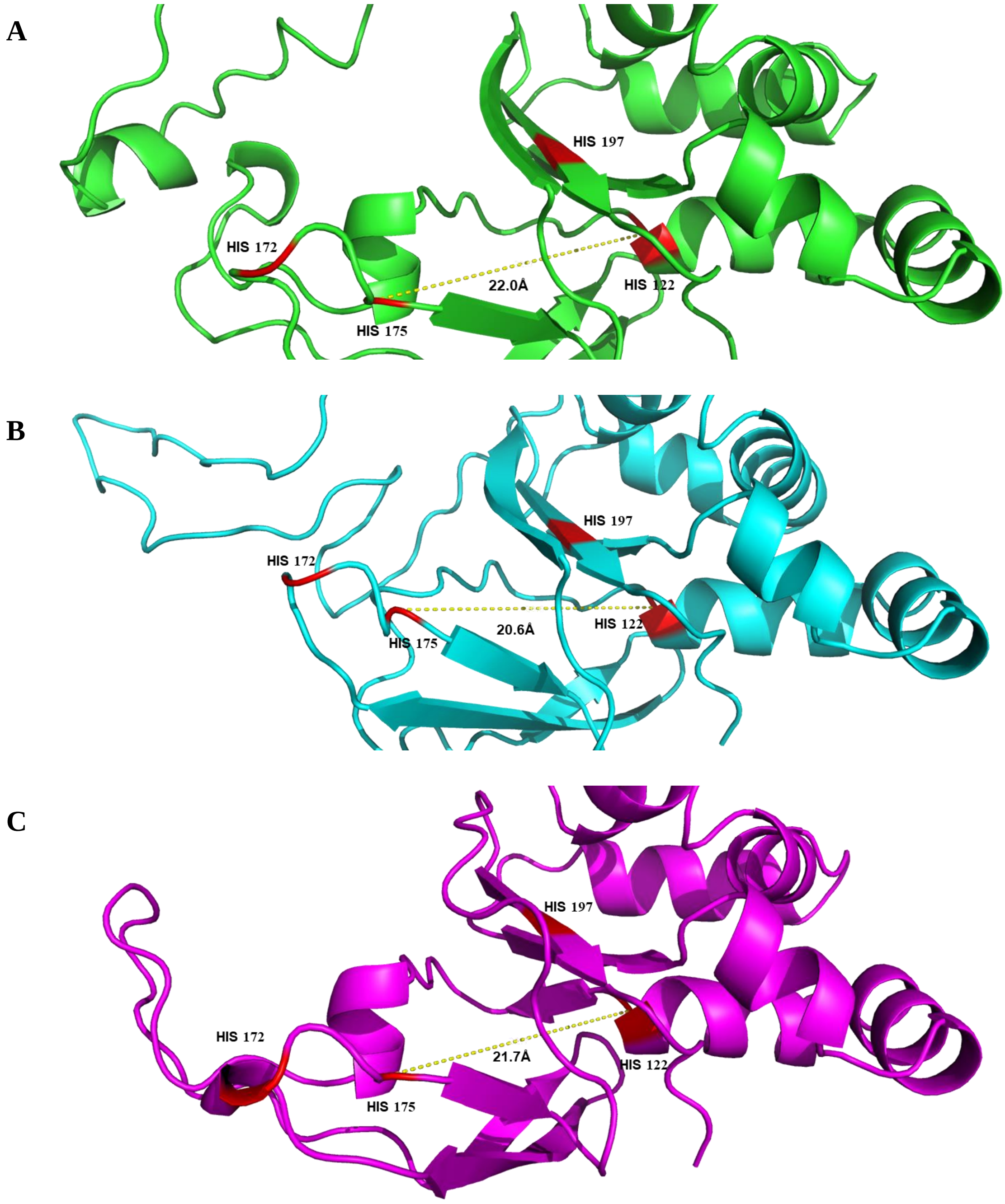


Figure 3.11 Distance between proposed haem binding residues (red) based on Nakatani *et al*. (2014) in HDP models.
A) RoseTTAFold model; **B)** I-TASSER model and **C)** AlphaFold model
*RoseTTAFold and I-TASSER models generated by Kyllen Dilsook

Comparing the HDP models to the Nakatani *et al*. (2014) proposed mechanism for haemozoin formation, the His122 and His175 positions are 22.0 Å, 20.6 Å and 21.7 Å apart in the RoseTTAFold, I-TASSER and AlphaFold models respectively (Figure 3.11) which indicates that haem molecules bound to these residues are not in a suitable position to facilitate haem dimer formation as proposed. However, it is important to consider the potential for conformational changes which may occur after initial haem binding that would allow for dimer formation to occur by folding the protein in a way that brings the bound haem molecules closer together.

Recently, a study by Matz *et al*. (2020) showed a lipocalin-like protein PV5 (PV-parasitophorous vacuole) controls haem crystallization and conditional knockdown of this protein in *P. falciparum* caused excessive multidirectional haemozoin crystal growth. However, PV5 appears to be non-essential as PV5-deficient parasites produce less haemozoin but still survive. The study suggests PV5 may bind haem or hematin dimers, thus reducing the level of haem supersaturation and moderating de novo haemozoin nucleation and promoting unidirectional crystal growth.

Knockdown studies of HDP indicate it is an essential protein for parasite survival during the erythrocytic stage of the parasite lifecycle (Jani *et al*., 2008). The core of the protein (residues 60-204) is a fasciclin I domain. Fasciclin I domains have been identified in numerous microorganisms, vertebrates and invertebrates (Moody and Williamson, 2013). Fasciclin domains are mostly found in proteins predicted to be in the extracellular environment and often play an important role in the function, structure and development of organisms (Liu *et al*., 2020). The presence of the fasciclin domain in HDP, prediction of its presence within the membrane as well as in the extracellular environment suggests the protein may play a role in cell adhesion and intracellular trafficking (as reported for mammalian fasciclin-containing proteins), cellular signalling (as hypothesised for plant fasciclin-containing proteins) or pathogenicity (as reported for fungi, archaea and eubacteria fasciclin-containing proteins) (Seifert, 2018). Considering this in the context of *Plasmodium falciparum*, it is possible *Pf*HDP plays a fasciclin-like role and facilitates invasion of the host cell and subsequent ingestion of haemoglobin in the parasite by interacting with other proteins in the parasite cytosol which drive these functions.

3.4 Conclusion

Purified haem detoxification protein was analysed for β-hematin formation activity, however, the results showed β-hematin formation was mediated by the SDS used in the assay rather than by HDP which supported findings by Roxanne Mohunlal (PhD Thesis, UCT 2019). It is unlikely elongation factor Tu 1 affected HDP activity in the assay as the observed amounts during purification were significantly less than the enriched HDP. In order to determine the possible function of HDP, an attempt was made to obtain single crystals so that x-ray diffraction could be used to determine the protein structure. Transmission electron microscopy of rHDP showed the protein does not form any tertiary complexes. HDP crystals obtained did not provide adequate diffraction data for structure determination. Despite the difficult and laborious nature of crystallography studies, it is possible that optimisation could improve the quality of crystals produced and thus improve the diffraction data. Optimisation could include removal of the disordered N-terminus region of HDP, using highly pure protein, using a different crystallisation method and possibly optimising around the initial crystal conditions. *In silico* methods were used to predict the possible structure of HDP using three different methods each of which predicted high disorder in the N-terminus of the protein. The structures predicted indicate it is unlikely the his-tag on the N-terminus is masked due to protein folding however it is possible the tag is masked by interactions with other proteins in the lysate. Evaluation of the models showed the four histidine residues predicted to facilitate haemozoin formation are not aligned in a way that would allow for haemozoin formation growth to occur. Further studies into the dynamics of the predicted structures could provide further insights into the feasibility of predicted/expected HDP characteristics such as the binding stoichiometry. Knockout studies previously reported indicate HDP is essential for parasite survival but the present data suggests this is unlikely due to haemozoin formation activity. Computational analysis of the HDP sequence shows it contains a fasciclin domain. As fasciclin-domain containing proteins often have similar functions, it is possible HDP is involved in functions similar to other fasciclin-domain containing proteins such as cell adhesion, pathogenicity or cellular signalling however, further biochemical studies are required.

Section B

Phosphatidylinositol 4-kinase

Chapter 4

Expression, Purification and Kinetic Characterisation of *Pv*PI4K-WT, *Pv*PI4K-F832A and *Pv*PI4K-C1327A

4.1 Introduction

Currently, the structure of *Plasmodium* PI4K remains unknown as the protein is difficult to express and crystallise. *Plasmodium* PI4K is significantly larger, contains extended loop regions within the kinase domain and has a larger regulatory domain compared to its closest homologue for which a structure is available - human PI4Kβ (PDB ID: 4D0L) (Burke *et al*., 2014; Fienberg *et al*., 2020). However, biochemical assays using recombinant protein have been reported and used to assess inhibitor potency. Full length *P. vivax* (*Pv*) PI4K has been successfully expressed in a baculovirus-insect cell expression system and used for *in vitro* inhibition studies (McNamara *et al*., 2013). More recently, a binding-based ELISA assay using streptavidin coated plates, biotinylated PI4P and a heavily truncated *Pf*PI4K construct expressed in yeast was reported (Sternberg and Roepe, 2020). *Plasmodium vivax* (*Pv*) PI4K and *Pf*PI4K share 97% sequence homology across the kinase domain, and their ATP binding sites are conserved thus the inhibitory potencies of ATP-competitive inhibitors are predicted to be similar (Fienberg *et al*., 2020).

Recently, a *Pf*PI4K homology model was built (Figure 4.1) using the human PI4Kβ crystal structure as a template (Fienberg *et al*., 2020). The homology model was validated using *Pf*PI4K inhibitors from the imidazopyridazine series (Cheuka *et al*., 2018), aminopyridine/pyrazine series (Younis *et al*., 2013) and the naphthyridine series (Kandepedu *et al*., 2018) and was subsequently used to rationalise inhibitor potency and selectivity for *Pf*PI4K over human homologues (Fienberg *et al*., 2020).

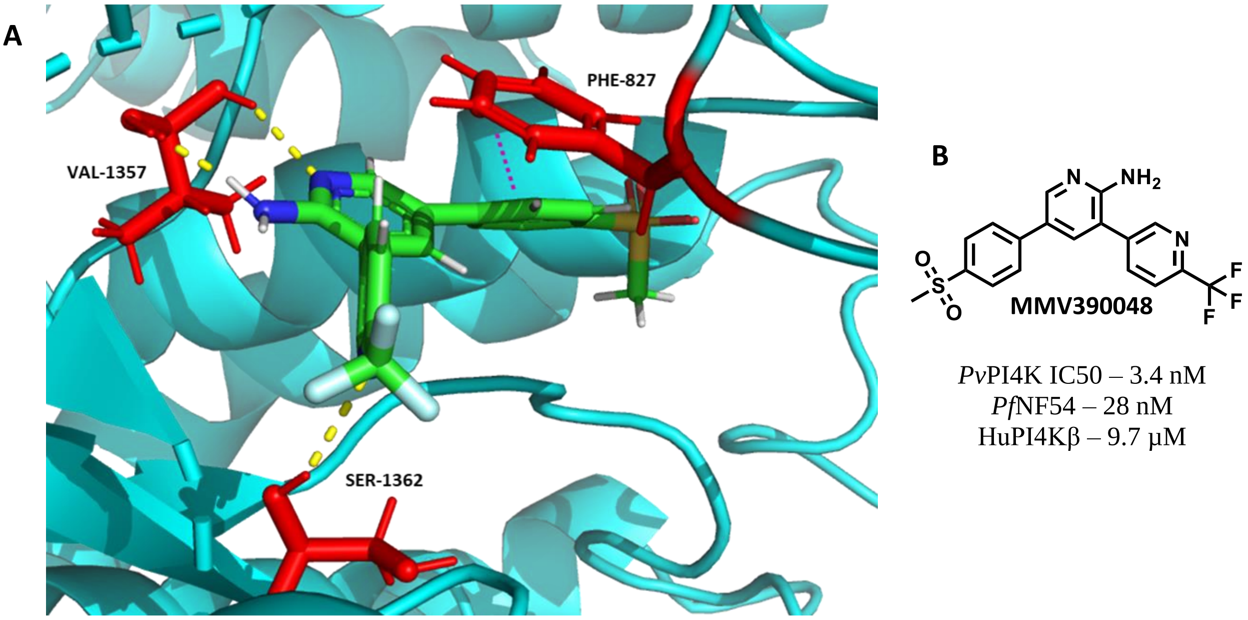


Figure 4.1: Known *Pf*PI4K inhibitor MMV390048 docked into *Pf*PI4K homology model. **(A)** *Pf*PI4K homology model built using human PI4Kβ (PDB ID: 4D0L) as a template. MMV390048 depicted in green. Red residues indicate amino acids which interact with the inhibitor. The yellow lines show hydrogen bonds, the pink line shows Pi-Pi stacking interaction. (B) 2D structure of MMV390048 and reported inhibition data (Paquet *et al*., 2017; Fienberg *et al*., 2020).

The ATP-binding site of the *Pf*PI4K homology model has relatively high similarity to the human PI4K orthologue - 48 % sequence identity between human PI4Kβ and *Pf*PI4K. Alignment of the *Pf*PI4K homology model with human PI4Kβ and comparison of MMV390048 docking into the ATP binding site (Figure 4.2) shows three conserved residues (K1308, V1357 and S1362) involved in inhibitor-enzyme interactions. The residues involved in inhibitor-enzyme interactions which are unique to *Plasmodium* are F827, Y1356 and S1365 (Fienberg *et al*., 2020). Fienberg *et al*. (2020) also found that these three residues are responsible for *Pf*PI4K selectivity over four human PI3K isoforms. F827 and Y1356 form Pi-Pi interactions with inhibitors thus multi-aromatic compounds have increased potency for *Pf*PI4K. The smaller serine group in *Pf*PI4K in comparison to glutamine in human PI4K – a larger, neutral, polar residue – allows for interactions between *Pf*PI4K and inhibitors with bulky groups e.g. the CF3 pyridyl of MMV390048 (Fienberg *et al*., 2020).

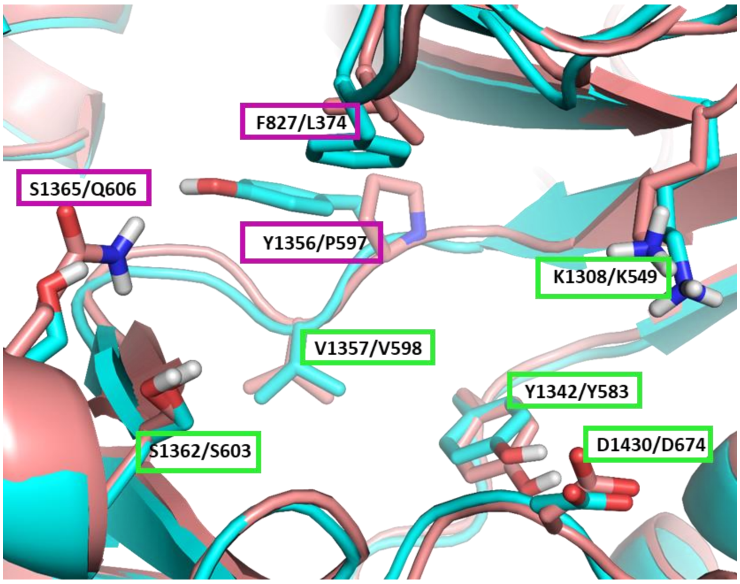


Figure 4.2: Alignment of MMV390048 docked into *Pf*PI4K homology model with human PI4Kβ crystal structure (Adapted from Fienberg *et al*., 2020).
The *Pf*PI4K homology model depicted in cyan, human PI4Kβ depicted in pink. Conserved residues highlighted in a green box; unique residues highlighted in a pink box. Residues in *Pf*PI4K and human PI4K shown respectively.

In this study, two residues unique to *Plasmodium*, F832 (L374 in humans) and C1327 (V602 in humans), were selected for mutation to evaluate key inhibitor-enzyme interactions in *Plasmodium* PI4K. F832 is thought to form key Pi-Pi interactions with inhibitors and C1327 is found on the periphery of the catalytic site and is a potential target for covalent inhibitors. The effect of these mutations on PI4K inhibition will be discussed in Chapter 5.

This chapter focuses on the expression and purification and kinetic characterisation of *Pv*PI4K-WT and two active site mutants *Pv*PI4K-F832A and *Pv*PI4K-C1327A.

4.2 Methods

*Pv*PI4K expression construct

Wild-type *Pv*PI4K (codon optimised for baculovirus expression) cloned into pFastBac-HTA, a donor vector for the Bac-to-Bac expression system for recombinant protein expression in insect cells (Figure 4.3), was gifted by John Burke (University of Victoria, Canada). The construct encodes full length *Pv*PI4K gene with an N-terminus 6X his-tag.

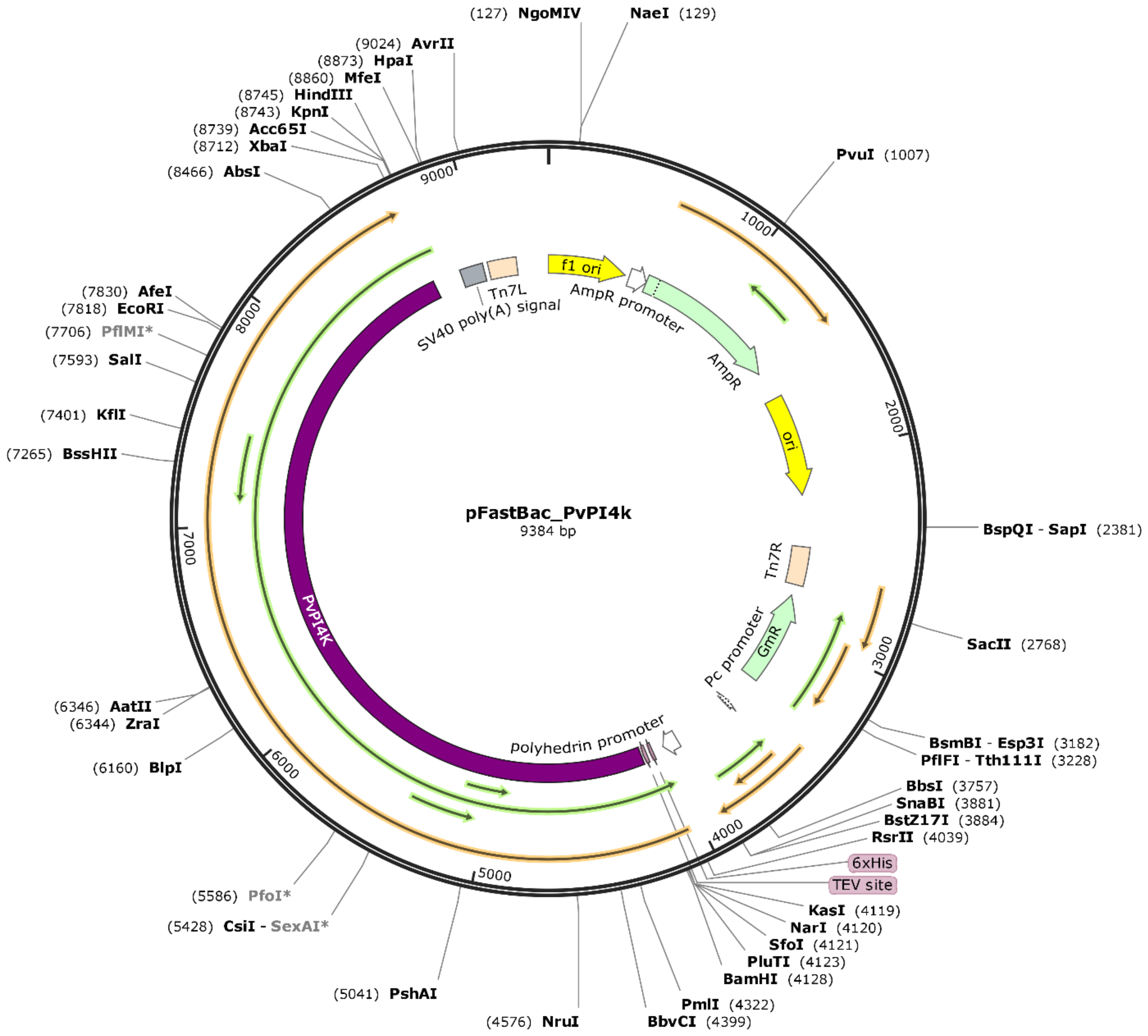


Figure 4.3 pFastBac-HTA vector containing *Pv*PI4K gene.

Site-directed mutagenesis

Two residues were selected for mutation based on the homology model of *Pf*PI4K (Fienberg *et al*., 2020): F827 and C1361. The corresponding residues in *Pv*PI4K are F832 and C1327.

4.2.1. Primer design

Primers were designed for inverse polymerase chain reaction (PCR) amplification of the *Pv*PI4K-pFastBac-HTA to introduce single-point mutations using SnapGene software. Primers were designed according to the following guidelines:

- ≤30 bp in length
- 40–60% GC content

- phosphorylated 5′ end

The primers were synthesised by Inqaba Biotechnical Industries, South Africa (Table 4.1). Lyophilised primers were reconstituted in nuclease-free water.

Table 4.1: Primers used for inverse PCR of *Pv*PI4K-pFastBac-HTA.

Name	Sequence (5′-3′)	Tm (℃)	GC content (%)
F832→A-F	GACGAGTGCAAGATCGCATTCAGCAAGAAG	64	50
F832→A-R	GTAGTTGAAGTGCAGGATCTGCAGGCCGGA	66	57
C1327→A-F	TACGTGAACGACACTGCATCCGTGGACTCC	68	57
C1327→A-R	CTCGATGATACCGGAGTTAGCACCGGTGAC	66	57

F, forward; R, reverse; Tm, primer melting temperature

4.2.2 Inverse PCR

PCR reactions were set up in a total reaction volume of 50 µL using Q5® Hot Start High-Fidelity DNA Polymerase (New England Biolabs) with 50 ng template DNA, 1× Q5® reaction buffer, 200 µM deoxyribonucleotide triphosphate (dNTP) mix, 0.5 µM of each primer, 0.02 units/µL Q5® Hot Start High-Fidelity DNA Polymerase, and nuclease-free water. A negative control without template DNA was also included. In a ProFlex PCR system (Life Technologies), PCR samples were heated at 98 ℃ for 30 s, followed by 35 cycles under the following conditions: 98 ℃ for 10 s, 66 ℃ for 30 s, 72 ℃ for 5 min, and a final step at 72 ℃ for 2 min. The PCR reactions were cooled at 4 ℃ and analysed via agarose gel electrophoresis (AGE).

4.2.3 Agarose gel electrophoresis (AGE)

PCR reactions were analysed via AGE on 0.8% (w/v) agarose gels, which were prepared by dissolving agarose (Sigma) in 1× TAE buffer (40 mM tris, 20 mM acetate, and 1 mM EDTA) and adding SYBR® Safe DNA gel stain (1× final concentration). Next, 6× gel loading dye was added to the DNA samples at a ratio of 1:5 loading dye:DNA. Electrophoresis was performed in 1× TAE buffer at 90 V for 1 h. A 1-kb DNA ladder (NEB) was used to estimate the size of the samples and the gels were visualised and photographed using a D-DiGit Gel Scanner (LI-COR).

4.2.4 PCR product purification and DpnI digestion

PCR products were purified using a Zymo DNA Clean and Concentrator -25 (Zymo Research) according to the manufacturer's instructions. The purified PCR products were treated with DpnI (NEB) according to the manufacturer's instructions. The digestion reaction was incubated at 37 °C for 30 min followed by 20 min at 80 °C.

4.2.5 DNA purification

Digested PCR products were analysed via AGE (section 4.2.3) and the DNA bands were excised and purified using a Monarch® DNA Gel Extraction Kit (NEB) according to the manufacturer's instructions. The concentration and purity of the purified PCR product were quantified using a Nanodrop 2000 Spectrophotometer (Thermo Fisher Scientific).

4.2.6 DNA ligation

Linearised PCR products were ligated using T4 DNA ligase (NEB) according to the manufacturer's instructions. Ligation reactions (20 µL) were set up using 33 ng linearised DNA and incubated at 16 °C overnight. Ligation reactions were stored at -20 °C.

4.2.7 Transformation in DH5α cells

Chemically competent DH5α cells were thawed on ice and 5 µL of the ligation reaction was added to 50 µL cells. A no-DNA negative control was also included. The cells were mixed briefly and incubated on ice for 25 min before heat-shocking at 42 °C for 2 min and incubating on ice for a further 2 min. Pre-heated LB broth without antibiotic was added to the cells and the culture was incubated at 37 °C with shaking at 100 rpm for 1 h. The transformation culture was plated onto LB-agar plates (12 g/L agar, 10 g/L tryptone, 10 g/L NaCl, and 5 g/L yeast extract) containing 100 µg/mL ampicillin. The plates were inverted and incubated overnight at 37 °C.

4.2.8 Plasmid isolation

Five colonies were selected for each mutant construct and inoculated into 5 mL LB broth containing 100 µg/mL ampicillin. The cultures were incubated overnight at 37 °C with shaking at 150–180 rpm. The cells were harvested via centrifugation at 6000 × *g* for 2 min. Plasmid DNA was isolated from the cells using a GeneJET Plasmid Miniprep kit (Thermo Fisher

Scientific) according to the manufacturer's instructions, quantified using a Nanodrop 2000 spectrophotometer, and submitted for sequencing.

4.2.9 Sequencing

Three samples of each mutant construct were submitted to Inqaba Biotechnical Industries (South Africa) for sequencing (Appendix Figure C1 and C2).

4.2.10 Long-term vector storage

Following verification via sequencing, the plasmids were transformed into DH5α cells as described in section 4.2.7 and a colony was grown in 5 mL LB broth containing 100 µg/mL ampicillin with shaking at 150–180 rpm. Glycerol was added to 500 µL of the overnight culture to a final concentration of 50% (v/v) and stored in sterile cryotubes at -80 ℃.

Recombinant protein expression in baculovirus-insect cell system

Expression of *Pv*PI4K-WT, *Pv*PI4K-F832A and *Pv*PI4K-C1327A was carried out using the Bac-to-Bac expression system based on previously described methods (Anderson *et al.*, 1995)

4.2.11 Transformation in DH10 cells

WT and mutant PI4K pFastBac-HTA expression constructs were isolated from DH5α cells as described in section 4.2.8 and transformed into chemically competent DH10 cells. Competent cells were thawed on ice and 1 µL plasmid DNA was added to 100 µL cells. The cells were mixed briefly and incubated on ice for 30 min before heat-shocking at 42 ℃ for 45 s and incubating on ice for a further 2 min. Room temperature LB without antibiotic was added to the cells and the cultures were incubated at 37 ℃ with shaking at 225 rpm for 4 h. The transformation cultures were plated onto LB-agar plates containing 10 µg/mL tetracycline, 50 µg/mL kanamycin, 7 µg/mL gentamicin, 1 mM IPTG, and 0.2 mg/mL X-Gal. The plates were inverted and incubated at 37 ℃ for 48 h. Two white colonies and one blue colony for each construct were selected and inoculated into 5 mL LB containing 10 µg/mL tetracycline, 50 µg/mL kanamycin, and 7 µg/mL gentamicin. The cultures were incubated overnight at 37 ℃ with shaking at 150–180 rpm. Glycerol was added to 500 µL of the overnight culture to a final concentration of 50% (v/v) and stored in sterile cryotubes at -80 ℃. The white colony stocks were streaked onto LB-agar plates containing 10 µg/mL tetracycline, 50 µg/mL kanamycin, 7 µg/mL gentamicin, 1 mM IPTG, and 0.2 mg/mL X-Gal for confirmation.

4.2.12 Bacmid isolation

A scraping of WT and mutant PI4K in the DH10 cells white colony glycerol stock was inoculated into 5 mL LB broth containing 10 µg/mL tetracycline, 50 µg/mL kanamycin, and 7 µg/mL gentamicin. The culture was incubated overnight at 37 °C with shaking at 150–180 rpm. Bacmid was isolated from the cells using a ZR Bac DNA Miniprep kit (Zymo Research) according to the manufacturer's instructions. Bacmid DNA was eluted in 40 µL nuclease-free water and quantified using a Nanodrop 2000 spectrophotometer.

Baculovirus expression system

4.2.13 Cell culture

Spodoptera frugiperda (Sf9) cells (Novagen, South Africa) were grown in suspension in Sf-900™ III serum-free complete medium (Gibco) at 28 °C in a shaking incubator at 400 rpm. Cell density and cell viability were determined using 0.4% Trypan blue (Gibco) and Bio-Rad cell counting slides (catalogue number: 1450011). The cells were maintained at a density of 1–5 × 10^6 cells/mL with viability ≥90%.

4.2.14 Transfection of bacmid DNA

Isolated bacmid DNA (section 4.2.13) was transfected into Sf9 insect cells using ExpiFectamine™ Sf Transfection Reagent (Gibco) according to the manufacturer's instructions and P0 virus was harvested 4 days post transfection.

4.2.15 Baculovirus amplification

Sf9 cells were sub-cultured to a density of 0.5 × 10^6 cells/mL (≥90% viability), placed in a T75 vented flask (10 mL culture per flask), and incubated at 27 °C for 30 min. Once the cells were approximately 50% confluent, 500 µL P0 virus (section 4.2.15) was added and the flask was incubated at 27 °C. Virus supernatant (P1) was harvested via centrifugation 6 days post infection. The culture medium was centrifuged at 1000 × *g* for 10 min in a Heraeus Megafuge 1.0R centrifuge and the supernatant was stored away from light at 4 °C. A 2-mL aliquot of P1 virus was stored at -80 °C. This was repeated for generation of P2 virus using 100 µL P1 virus for cell infection.

4.2.16 Small-scale expression test

In a 24-well deep-well plate, 3 mL Sf9 cells (1.0 × 10^6 cells/mL) was inoculated with 60 µL P2 virus (1:50). The cells were incubated at 28 °C with shaking at 400 rpm for 48 h before harvesting via centrifugation at 1000 × *g* for 10 min using a Heraeus Megafuge 1.0R centrifuge. The pellets were washed in 1× PBS at pH 7.4 (Sigma) and centrifuged at 1000 rpm for 10 min before storing at -80 °C. The pellets were resuspended in lysis buffer (25 mM HEPES at pH 7.5, 50 mM KCl, 20% (v/v) glycerol, 0.1% (v/v) Triton X-100, 1 mM 1,4-dithiothreitol (DTT), 150 mM NaCl, 6.25 µg/mL DNaseI, and 12.5 µg/mL RNase) and incubated with shaking at 80 rpm for 1 h on ice. The samples were sonicated on ice for 6 cycles (20 mA, 10 s on, 40 s off) and centrifuged at 13,000 rpm, 4 °C for 30 min. The soluble (supernatant) and insoluble (pellet) fractions were loaded on a 8% (w/v) SDS-PAGE gel and analysed via western blot using a monoclonal anti-polyhistidine antibody as described in Chapter 2, Section 2.2.13.

4.2.17 Large-scale protein expression

Large-scale expression cultures of *Pv*PI4K-WT, *Pv*PI4K-F832A and *Pv*PI4K-C1327A were produced for protein purification. Two 400-mL Sf9 cultures at a density of 1.0 × 10^6 cells/mL in 1-L vented flasks were inoculated with 8 mL P2 virus. The cells were incubated at 28 °C with shaking at 160 rpm for 48 h before harvesting via centrifugation in a Heraeus Megafuge 1.0R centrifuge at 1000 × *g* for 10 min. The cell pellets were washed in 1× PBS, centrifuged at 1000 × *g* for 10 min, and stored at -80 °C. The cell pellets were thawed on ice and resuspended in lysis buffer (20 mM Tris-HCl at pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 0.01% (v/v) Triton X-100, 10 mM 2-mercaptoethanol, and protease cocktail inhibitor (Roche catalogue number: 11836145001) and sonicated for 6 cycles with a microtip at power 5 (15 s on, 15 s off) using a Misonix Sonicator 3000. The lysates were centrifuged at 20,000 × *g*, 4 °C for 30 min and the soluble (supernatant) and insoluble (pellet) fractions were separated. The samples were then analysed via SDS-PAGE (section 2.2.3).

4.2.18 PI4K purification: His-tag affinity chromatography

WT and mutant PI4K was purified from the soluble fraction (filtered through a 0.45-µm filter) using a HisTrap™ HP 1-mL column (GE Healthcare) and an ÄKTA Explorer purification system (GE Healthcare) with fraction collection. The HiTrap™ HP column was equilibrated with five column volumes of equilibration buffer (20 mM Tris-HCl at pH 7.5, 500 mM NaCl, 5% glycerol, 20 mM imidazole, and 10 mM 2-mercaptoethanol) and the soluble fraction was loaded onto the column. The bound protein was eluted from the column using elution buffer (20 mM Tris-HCl at pH 7.5, 500 mM NaCl, 5% glycerol, 500 mM imidazole, and 10 mM 2-mercaptoethanol). Aliquots of the sample flow-through, column wash fractions, and elution fractions were collected and analysed via SDS-PAGE (section 2.2.3).

Purified PI4K was concentrated using a 100-kDa Amicon® Ultra centrifugal filter (Merck) at 4000 × *g* at 3-min intervals in an Allegra X-30R centrifuge until the volume was approximately 5 mL. Protein concentration was determined as described in section 2.2.7 using Ab 0.1% = 0.766 for PI4K - determined by Expasy ProtParam tool (Swiss Institute of Bioinformatics).

4.2.19 PI4K size-exclusion chromatography (SEC)

The second protein purification step was carried out on the same day using a HiLoad 16/600 Superdex 200 pg column attached to a Gilson chromatography system controlled by Uniprot software with 3 mL fraction collection. The column was pre-equilibrated with SEC buffer (20 mM HEPES at pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 10 mM 2-mercaptoethanol). The concentrated partially purified protein from section 4.2.18 was loaded onto the column and SEC fractions corresponding to the observed protein peaks were analysed by SDS-PAGE. Fractions corresponding to soluble *Pv*PI4K protein were pooled and concentrated to ≥ 1mg/mL using a 100-kDa Amicon® Ultra centrifugal filter (Merck). Single use aliquots were stored at -80 ˚C.

PI4K assay

4.2.20 Substrate preparation

L-α-phosphatidylinositol (PI; bovine liver, Sigma) was resuspended in 3% (v/v) octyl glucoside to a final concentration of 20 mg/mL and stored at -20 ℃.

4.2.21 Kinase assay

Purified WT and mutant PI4K enzyme activity was determined using an ADP-Glo™ kinase kit (Promega) according to the manufacturer’s instructions. In a 96-well plate, purified PI4K was diluted in assay buffer (25 mM HEPES at pH 7.5, 100 mM NaCl, 0.2% (v/v) Triton X-100, 1 mM DTT, 0.025 mg/mL BSA, and 6 mM $MgCl_2$). Substrate buffer (0.2 mg/mL PI and 20 µM ATP) was added to the enzyme at a 1:1 ratio and a no enzyme control reaction containing substrate buffer only was included. The reactions were incubated for 40 min at room temperature and aliquots were transferred into a 384-well plate. ADP-Glo™ reagent was added to the reaction at a 1:1 ratio and incubated at room temperature for 40 min. Kinase detection reagent was added to the reaction at a 1:2 ADP-Glo™ reagent:kinase detection reagent ratio. The reaction was incubated at room temperature for 40 min and luminescence was measured using an EnSpire multimode plate reader (PerkinElmer). Initial luminescence was corrected for background using the no enzyme control. ATP to ADP conversion was determined using a ADP:ATP standard curve prepared according to the manufacturer's instructions.

4.2.22 Determination of K_m

To determine the K_m of PI for *Pv*PI4K, the kinase assay was performed as described in section 4.2.21 in the presence of ATP at a final concentration of either 10 µM or 100 µM. The kinase reaction mix comprised 7.5 nM *Pv*PI4K final concentration and a 2-fold dilution of PI starting at 400 µM. Luminescent units were calculated using ATP-to-ADP conversion standard curves generated for the 10 and 100 ATP reactions respectively. ADP formed per minute was plotted against substrate concentration and K_m and V_max were calculated using the Michaelis-Menten equation on GraphPad prism.

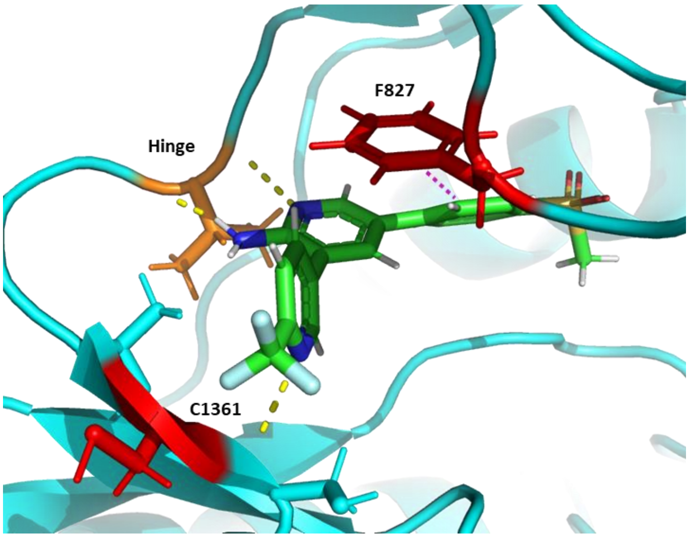
For K_m ATP, ATP-to-ADP conversion standard curves were generated at each ATP concentration used (two-fold dilution of ATP starting at 500 µM). The kinase assay was performed as described in section 4.2.21 and the reaction mix comprised 12.5 nM *Pv*PI4K-WT/*Pv*PI4K-C1327A final concentration (50 nM *Pv*PI4K-F832A final concentration), 100 µM PI, and a two-fold dilution of ATP starting at 500 µM. ATP formed (µM per minute) was plotted against ATP concentration and K_m and V_max were calculated using the Michaelis-Menten equation on GraphPad prism.

4.3 Results and Discussion

4.3.1 Mutation selection

As already indicated under Methods, two amino acids were selected for mutation based on docking studies performed using the *Pf*PI4K homology model (Stephen Fienberg, UCT) and site-directed mutagenesis was used to introduce point mutations at these residues in *Pv*PI4K..

A



B

***Pf*PI4K**	***Pv*PI4K (XP_001613270)**
F827→A	F832→A
C1361→A	C1327→A

Figure 4.4 Selected residues for mutation.
(**A**) Mutated *Pf*PI4K residues (red) in *Pf*PI4K homology model with docked MMV390048 (green). Pink line shows Pi-Pi stacking interaction, yellow lines show hydrogen bonds, orange residue shows hinge region. (**B**) Table showing corresponding *Pv*PI4K residues.

The *Pf*PI4K F827 residue is predicted to form Pi-Pi stacking interactions with aromatic rings of PI4K inhibitors (Figure 4.4). *Pf*PI4K C1361, situated on the periphery of the ribose binding site, is unique to *Plasmodium* PI4K, and is conserved across *Plasmodium* species (Kulkarni *et al*., 2020). This cysteine has been proposed as a target for covalent inhibitors (Kulkarni *et al*., 2020). The equivalent residues in *Pv*PI4K (F832 and C1327) were mutated to alanine in order to probe the functional role of the amino acid residues on PI4K-inhibitor interactions without altering the main amino acid chain conformation and without introducing major electrostatic or steric effects (Lefèvre *et al*., 1997). These mutant proteins were generated to provide insight into selectivity relative to related human lipid kinases as well as the action of targeted covalent inhibitors designed to target the cysteine in the parasite enzyme.

4.3.2 Recombinant protein expression

Recombinant *Pv*PI4K proteins were expressed in insect cells using the Bac-to-Bac® Baculovirus Expression System based on previously described methods (Anderson *et al*., 1995; McNamara *et* al., 2013). Following recombinant baculovirus generation and amplification to yield a P2 viral stock, small scale test expression was carried out a deep well plate to optimise expression conditions. Initially, four P2 virus to insect cell culture ratios (1:1000, 1:100, 1:30 and 1:10) and different expression times (48 h and 72 h) were evaluated for their effect on *Pv*PI4K-WT expression (Figure 4.5).

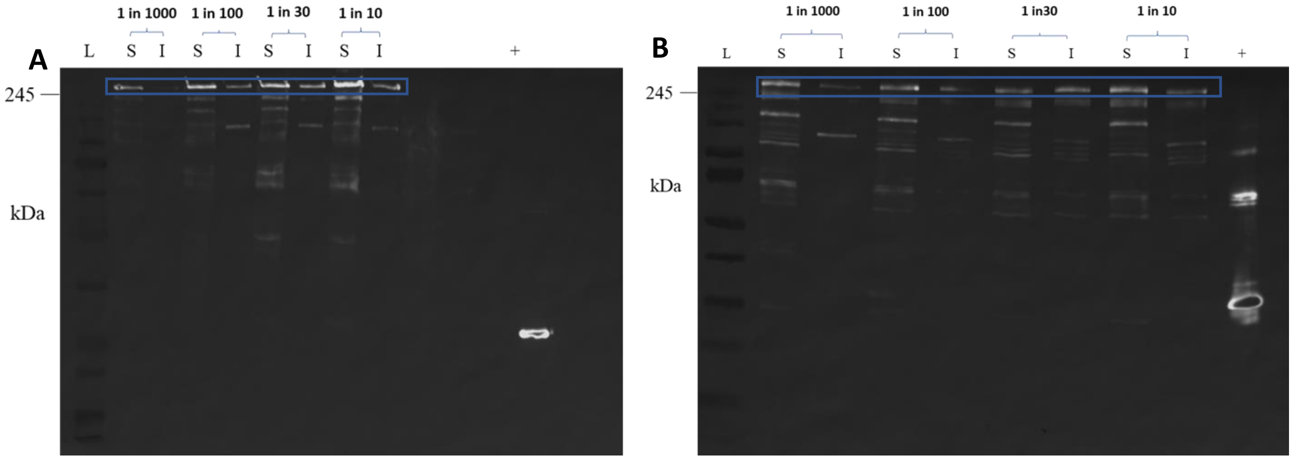


Figure 4.5: Optimisation of *Pv*PI4K-WT expression conditions.
Small-scale expression samples analysed via western blot using monoclonal anti-polyHistidine-peroxidase clone HIS-1antibody. M: Molecular mass marker; S: soluble fraction; I: insoluble fraction; +: Histidine tag positive control; Blue box shows expressed PI4K. (**A**) *Pv*PI4K-WT 48 h expression samples. (**B**) *Pv*PI4K-WT 72 h expression samples.

The expression conditions showed optimal 48 h expression time and good expression levels were observed at a virus to insect cell ratio at or above 1:100. The expression conditions used for subsequent *Pv*PI4K-WT and mutant expression were 48 h at a 1:50 virus to insect cell culture ratio.

Two batches of P2 viral stocks for *Pv*PI4K-WT, *Pv*PI4K-F832A and *Pv*PI4K-C1327A expression (Sample 1 and Sample 2) were evaluated in small scale test expression prior to bulk protein expression (Figure 4.6).

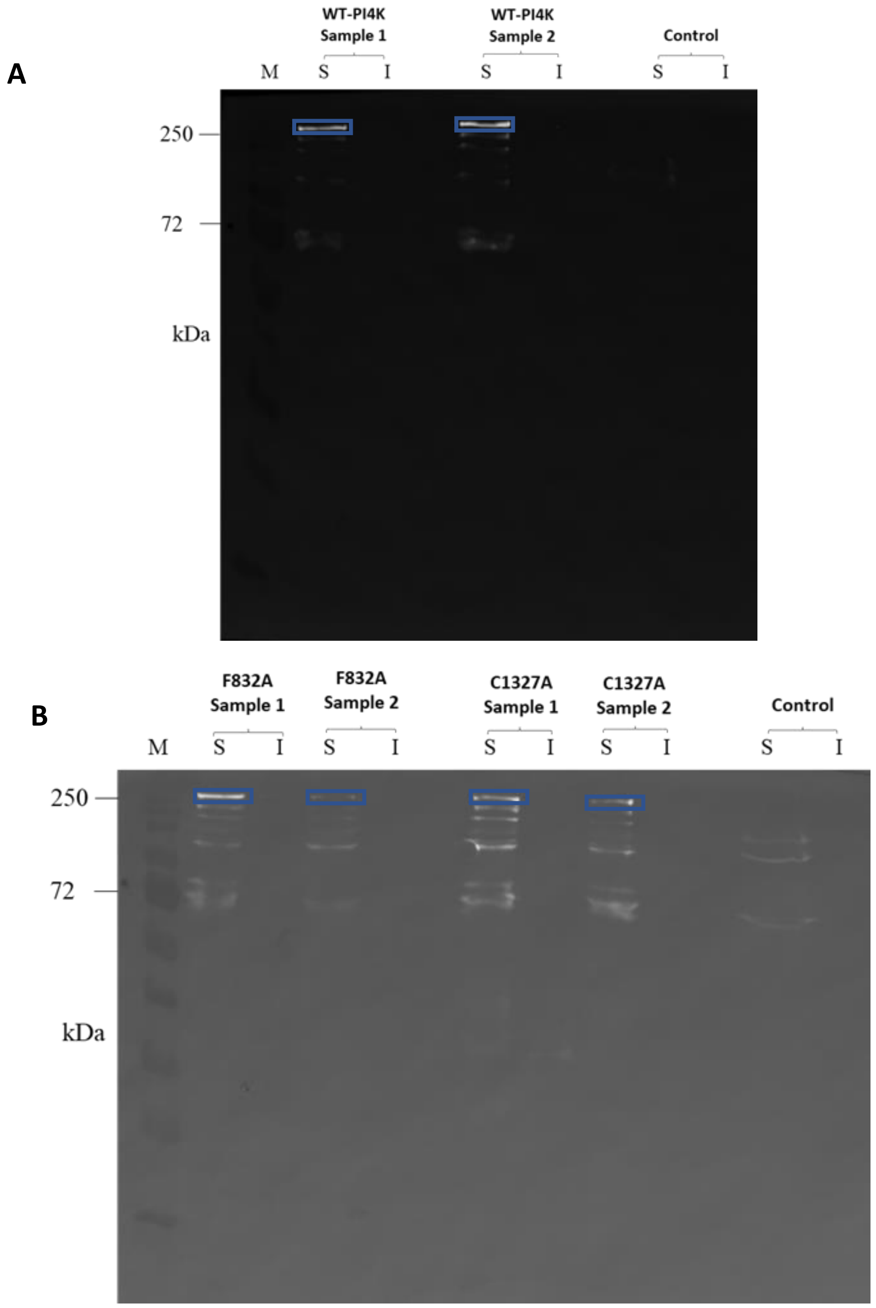


Figure 4.6: Western blot analysis of *Pv*PI4K-WT, *Pv*PI4K-F832A and *Pv*PI4K-C1327A small-scale expression.

Small-scale expression samples analysed by western blot using monoclonal anti-polyHistidine-peroxidase clone HIS-1antibody. M: Molecular mass marker; S: soluble fraction; I: insoluble fraction; Control: cells not infected with virus; blue box shows expressed PI4K. (**A**) *Pv*PI4K-WT small-scale expression samples. (**B**) *Pv*PI4K-F832A and *Pv*PI4K-C1327A small scale expression samples.

Figure 4.6 shows expression of the three enzymes. The bands corresponding to *Pv*PI4K proteins (169 kDa) were higher on the gel than expected (above 250 kDa marker band) based on the molecular weight of the protein, but this was consistent with what had previously been observed for *Pv*PI4K-WT (Figure 4.5). This may be due to the large size of the protein and the amino acid make up of the protein. A study by Shirai *et al*. (2008) found protein migration through an SDS-PAGE gel was affected by the hydrophobicity and isoelectric points of proteins- features determined by the amino acid content of proteins. Furthermore, post-translational modifications e.g. phosphorylation have been observed to result in slower migrating bands (Shirai *et al*., 2008).

Distinct bands were observed for the *Pv*PI4K-WT expression samples and *Pv*PI4K-F832A expression sample 1. Expression was also seen for *Pv*PI4K-F832A sample 2 and the *Pv*PI4K-C1327A samples, but the band intensity indicates lower expression levels for these samples. Sample 1 P2 virus for all three enzymes was used for large-scale expression of *Pv*PI4K-WT, *Pv*PI4K-F832A and *Pv*PI4K-C1327A.

4.3.3 Protein purification

A 400-mL insect cell culture inoculated with 8 mL *Pv*PI4K-WT P2 virus was incubated for 48 h and cells were collected. After cell lysis, the soluble fraction was filtered and loaded onto an equilibrated 1-mL HisTrap™ HP column. The eluted protein was concentrated using a 100-kDa Amicon® Ultra centrifugal filter prior to SEC using a HiLoad 16/600 Superdex 200 pg column. Samples collected during affinity chromatography were analysed via SDS-PAGE.

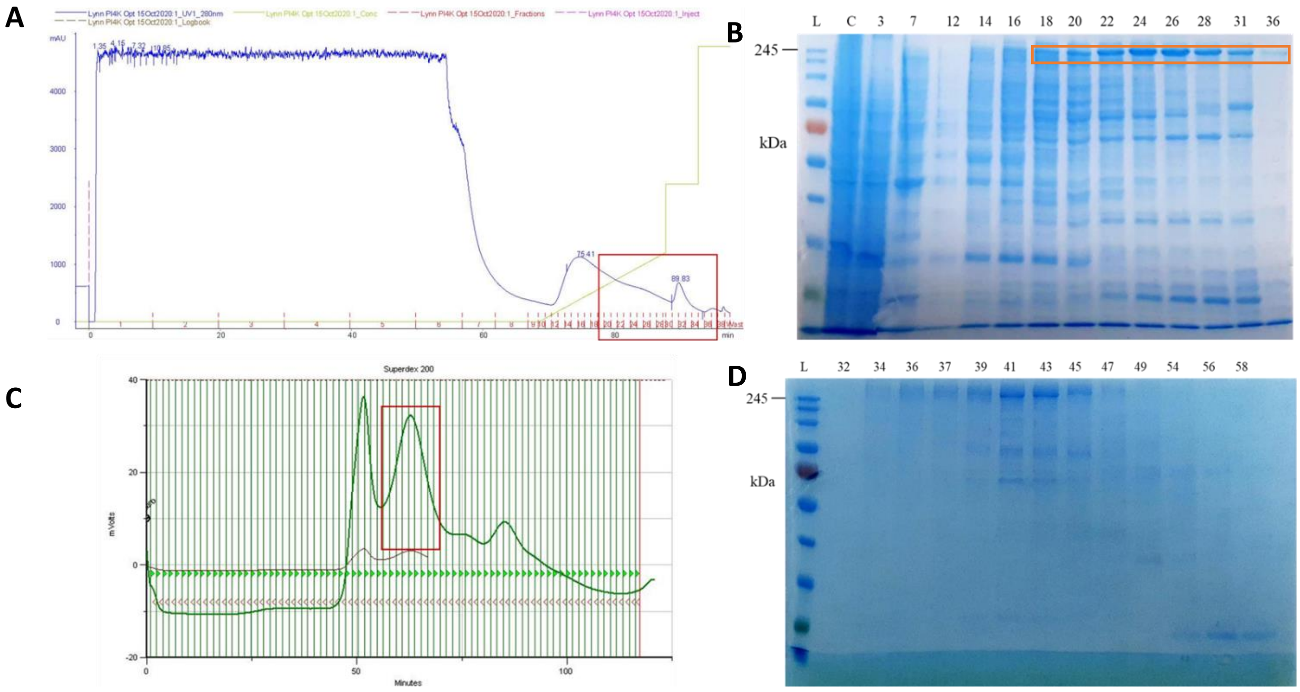


Figure 4.7: Two step purification of *Pv*PI4K-WT expressed using a baculovirus insect cell expression system.

(A)*Pv*PI4K-WT HisTrap™ HP purification profile, red box indicates the elution peak corresponding to *Pv*PI4K. **(B)** HisTrap™ HP purification fractions analysed by SDS-PAGE on an 8% gel. M: Molecular mass marker; S: soluble fraction; I: insoluble fraction; Fractions analysed: 3, 7, 12, 14, 16, 18, 20, 22, 24, 26, 28, 31, 36. Orange box indicates pooled elution fractions. (**C**) HiLoad 16/600 Superdex 200 pg size -chromatography profile; red box indicates the PI4K peak. (**D**) Size-exclusion chromatography fractions analysed via SDS-PAGE on an 8% gel. M: Molecular mass marker; Fractions analysed: 32, 34, 36, 37, 39, 41, 43, 45, 47, 49,54,56,58.

The gel (Figure 4.7B) showed distinct bands at the expected size based on the western blot and previous observations, indicating the enrichment of PI4K via affinity chromatography.

Size exclusion chromatography separates proteins based on size using porous beads, therefore the size of the pores determines which particles enter the beads and which particles do not (Giridhar *et al*., 2017). ‘Smaller’ molecules – within the pore size range – enter the pores and are retained thus elute later than ‘larger’ molecules. As protein conformation can also affect retention time e.g. the disordered regions and extended loop regions in *Pv*PI4K, separation using this technique only allows for an approximation the molecular weight based on elution time.

Two distinct peaks were observed in the SEC profile (Figure 4.7C), although the height of the peaks suggests low protein concentration. SDS-PAGE analysis of the size exclusion chromatography fractions showed distinct protein bands at the expected size for both peaks, indicating the protein may also aggregate and explaining its presence in the void volume peak (peak 1). Peak 2 was observed at 62.9 min elution time, which is approximately at 400 kDa – determined using HiLoad 16/600 Superdex 200 pg column calibration curve (Appendix Figure D1)- which suggests soluble *Pv*PI4K may elute as a dimer. Some contaminating proteins were observed in the peak 2 fractions (Figure 4.7D) albeit at very low concentration in comparison to the protein of interest (~ 80% purity). The peak 2 fractions were pooled and concentrated. Glycerol was added and aliquots at a concentration of 1.5 mg/mL were stored at -80 °C. The final protein yield was 0.54 mg from a 400 mL culture.

The affinity purification profile of *Pv*PI4K-F832A (Figure 4.8A) was similar to that of *Pv*PI4K-WT with slightly more contaminating proteins observed in the purification fractions (Figure 4.8B). Similar to the *Pv*PI4K-WT size exclusion profile, two peaks were observed for *Pv*PI4K-F832A (Figure 4.8C) but the aggregation peak (peak 1) was larger resulting in less separation between the peaks. This suggests that the mutant protein may be less stable than the wild-type protein. As expected, peak 2 had a similar elution time to *Pv*PI4K-WT (61.6 min). Analysis of the SEC fractions by SDS-PAGE (Figure 4.8D) shows distinct bands at the expected size with some contaminants as observed for *PvP*I4K-WT(~ 85% purity).. The final yield of *Pv*PI4K-F832A was comparable to *PvP*I4K-WT (0.57 mg from a 400 mL culture).

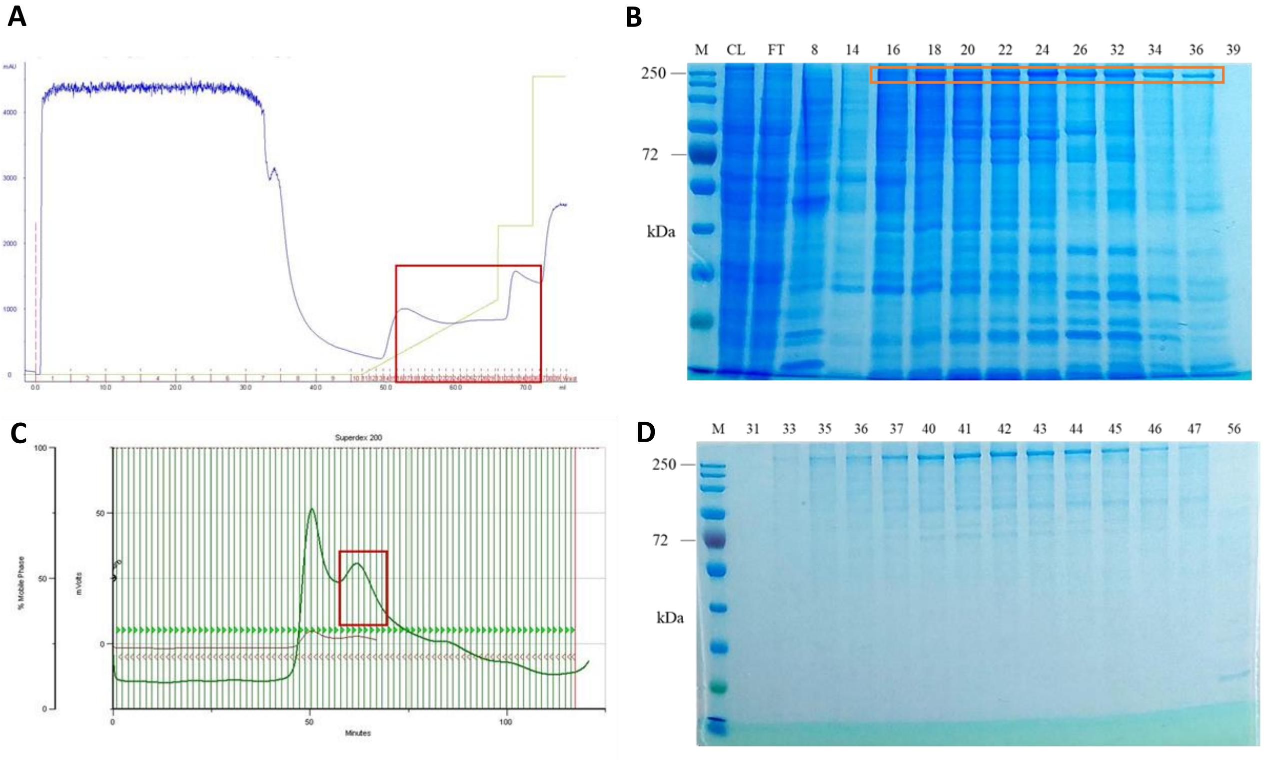


Figure 4.8: Two step purification of *Pv*PI4K-F832A expressed using a baculovirus insect cell expression system.

(A)*Pv*PI4K-F832A HisTrap™ HP purification profile, red box indicates the elution fractions. **(B)** HisTrap™ HP purification fractions analysed by SDS-PAGE on an 8% gel. M: Molecular mass marker; CL: cleared lysate; FT: flow-through fraction; Fractions analysed: 8, 14, 16, 18, 20, 22, 24, 26, 32, 34, 36, 39. Orange box indicates pooled elution fractions. (**C**) HiLoad 16/600 Superdex 200 pg size -chromatography profile; red box indicates the PI4K peak. (**D**) Size-exclusion chromatography fractions analysed via SDS-PAGE on an 8% gel. M: Molecular mass marker; Fractions analysed: 31, 33, 35, 36, 37, 40, 41, 42, 43, 44, 45, 46, 47, 56.

The *Pv*PI4K-C1327A HisTrap™ and SEC purification profiles (Figure 4.9A and Figure 4.9C respectively) were very similar to *Pv*PI4K-F832A. Elution of *Pv*PI4K-C1327A from the HisTrap™ column occurred slightly later than observed for *Pv*PI4K-F832A. As expected, peak 2 had a similar elution time to *Pv*PI4K-F832A (61.7 min) Analysis of the SEC fractions by SDS-PAGE (Figure 4.9D) shows distinct bands at the expected size (~ 90% purity) but the size and intensity of the bands suggests less protein is present. This was confirmed as the final protein yield for *Pv*PI4K-C1327A was 0.2 mg from a 400 mL culture.

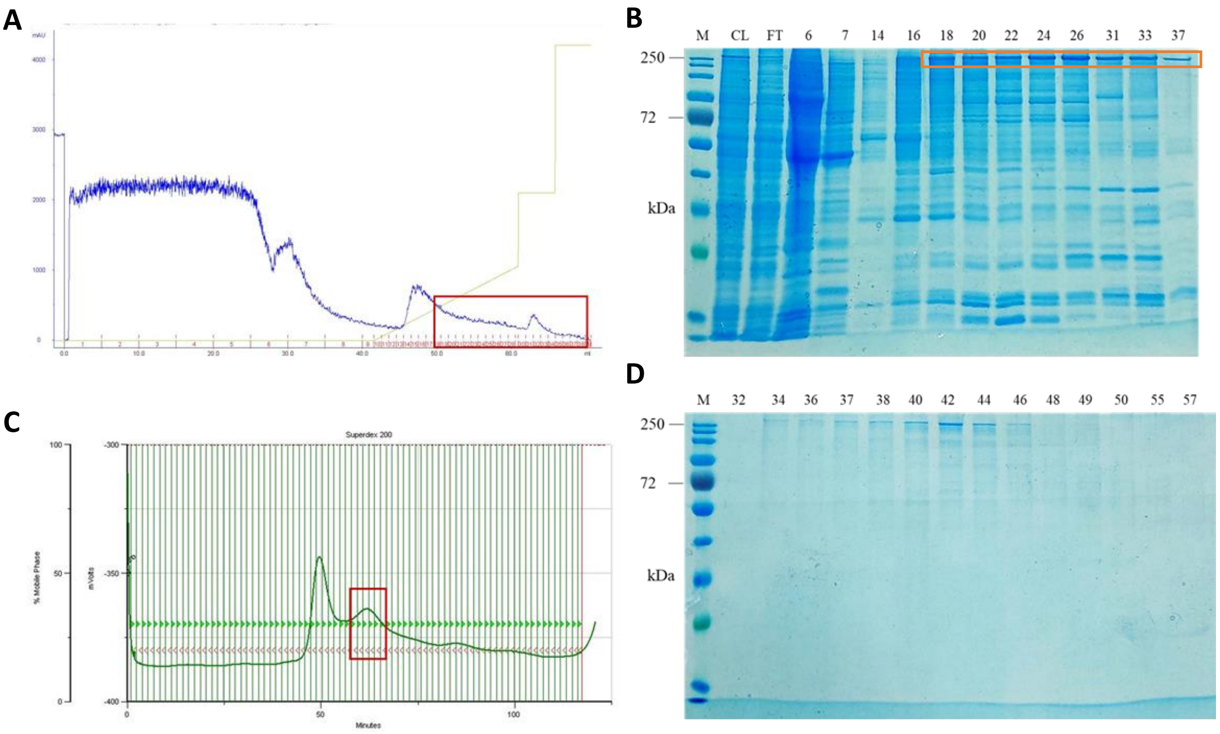


Figure 4.9: Two step purification of *Pv*PI4K-C1327A expressed using a baculovirus insect cell expression system.

(A) *Pv*PI4K-C1327A HisTrap™ HP purification profile, red box indicates the elution fractions. **(B)** HisTrap™ HP purification fractions analysed by SDS-PAGE on an 8% gel. M: Molecular mass marker; CL: cleared lysate; FT: flow-through fraction; Fractions analysed: 6, 7, 14, 16, 18, 20, 22, 24, 26, 31, 33, 37. Orange box indicates pooled elution fractions. (**C**) HiLoad 16/600 Superdex 200 pg size -chromatography profile; red box indicates the PI4K peak. (**D**) Size-exclusion chromatography fractions analysed via SDS-PAGE on an 8% gel. M: Molecular mass marker; Fractions analysed: 32, 34, 36, 37, 38, 40, 42, 44, 46, 48, 49, 50, 55, 57.

4.3.4 Kinase assays

The purified enzymes were evaluated for activity based on previously described methods (McNamara *et al*., 2013) in an endpoint assay using the ADP-Glo™ kinase assay kit to measure ADP formation.

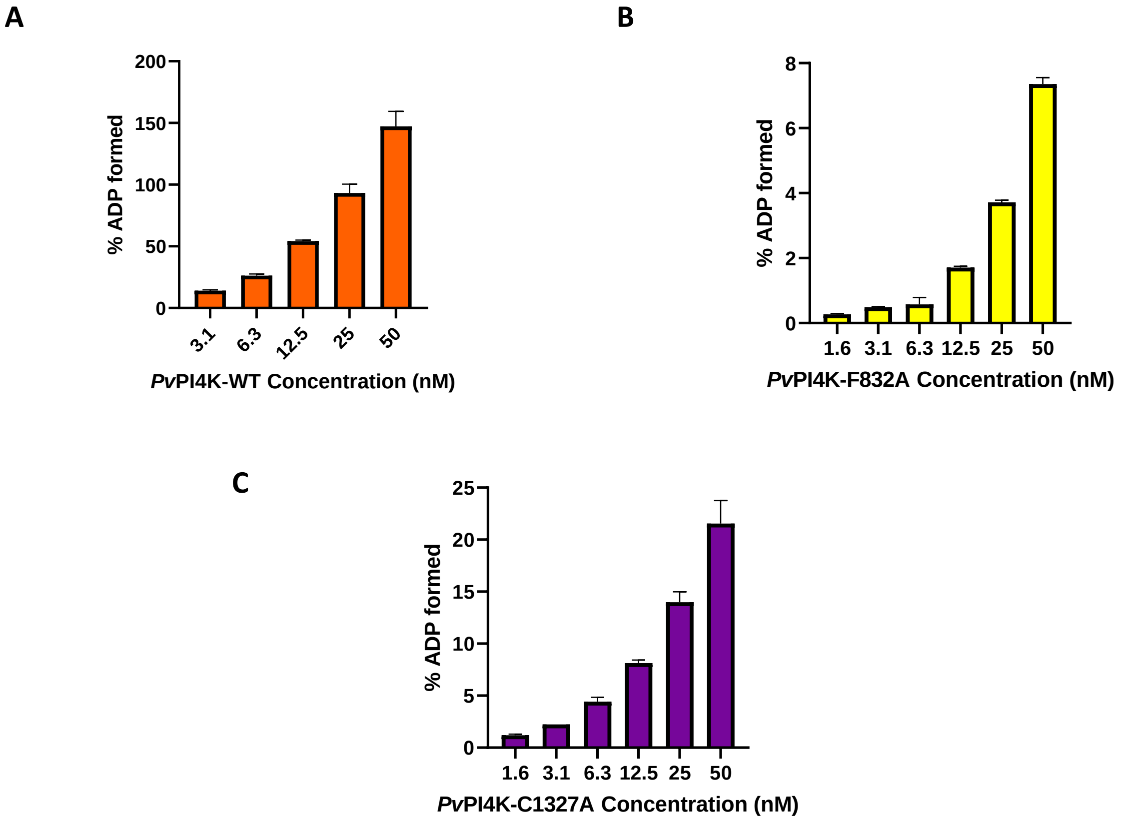


Figure 4.10: Evaluation of *Pv*PI4K enzyme activity.
(**A**) *Pv*PI4K-WT, (**B**) *Pv*PI4K-F832A and (**C**) *Pv*PI4K-C1327A activity at concentrations of ranging from 1.5 - 50 nM enzyme, 0.1 mg/mL PI substrate, and 10 µM ATP. The % ADP formed was corrected for background signal using the no enzyme control reaction.

The data in Figure 4.10 indicated ADP formation was dependent on the enzyme concentration in the assay. Figure 4.10A shows purified *Pv*PI4K-WT displayed higher activity than *Pv*PI4K-F832A and *Pv*PI4K-C1327A mutants (Figure 4.10B and Figure 4.10C respectively). It is likely the low enzyme activity of *Pv*PI4K-F832A and *Pv*PI4K-C1327A compared to *Pv*PI4K-WT is due to the introduced mutations which may affect protein stability and optimal protein folding.

Protein concentration for subsequent assays was selected based on these results to yield a good signal-to-background ratio. When working with 10 µM ATP in the reaction, an ATP-to-ADP conversion in the 5% - 20% range displays good quality assay data (Z' factor 0.82 – 0.90) (Promega, catalogue number V9102), while minimising effects of substrate depletion. Downstream assays were performed at enzyme concentrations yielding ATP-to-ADP conversion within this range over the 40-minute reaction time i.e. 7nM for *Pv*PI4K-WT, 50 nM for *Pv*PI4K-F832A and 12.5 nM *Pv*PI4K-C1327A.

The effect of different PI concentrations on *Pv*PI4K-WT activity is shown in Figure 4.11. Increased enzyme activity was observed with increased substrate concentration indicating enzyme activity is dependent on the substrate concentration.

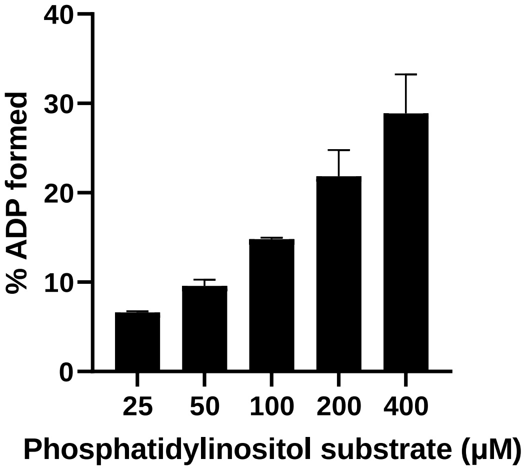


Figure 4.11: *Plasmodium vivax* phosphatidylinositol 4-kinase (*Pv*PI4K) enzyme assay substrate concentration effect on kinase activity.
The effect of PI concentration on the kinase assay was determined by performing the assay at different substrate concentrations. A two-fold serial dilution (400 µM starting concentration) was incubated with 7 nM *Pv*PI4K-WT and 10 µM ATP at room temperature for 40 min. The kinase assay was performed as previously described.

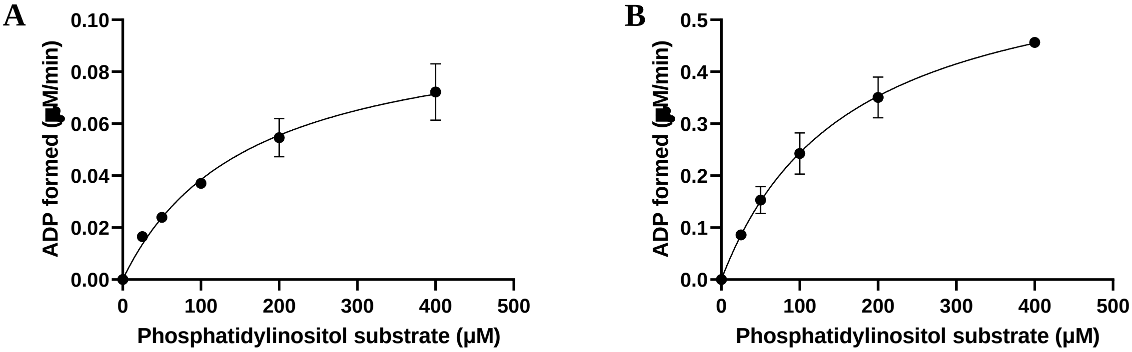
4.3.5 *Pv*PI4K enzyme kinetics

The rate at which most enzymes convert substrate to product varies hyperbolically with substrate concentration at a fixed enzyme concentration (Wilson, 2010). This relationship can be shown in a mathematical equation known as the Michaelis-Menten equation shown below where v_0 is initial rate, V_{max} is the maximal velocity of the enzyme, [S] is the substrate concentration and K_m is the Michaelis constant – a measure of the affinity of the enzyme for the substrate (Robinson, 2015).

$$v_0 = \frac{Vmax\,[S]}{Km+[S]}$$

The K_m of the lipid substrate PI (K_m^{PI}) was determined by performing the assay at 10 µM and 100 µM ATP using a two-fold serial dilution of the substrate prepared in assay buffer (0.4 mg/mL starting concentration). The amount of ADP produced in the assay correlates with phosphorylation of PI substrate by PI4K. The amount of ADP produced in the kinase reaction was estimated using ATP-to-ADP conversion standard curves (Appendix Figure E1 and Figure E2**)**. The Michaelis-Menten model was fitted to the data using non-linear regression and K_{m} and V_{max} values were determined (Figure 4.12).

The $K_{\mathrm{m}}^{\mathrm{PI}}$ at 10 µM ATP and 100 µM ATP was relatively constant, which was also expected as K_{m} of the PI substrate should be independent of ATP concentration. The data show that $K_{\mathrm{m}}^{\mathrm{PI}}$ is approximately 160 µM. As expected, V_{max} increased with an increase in ATP concentration (Figure 4.12).



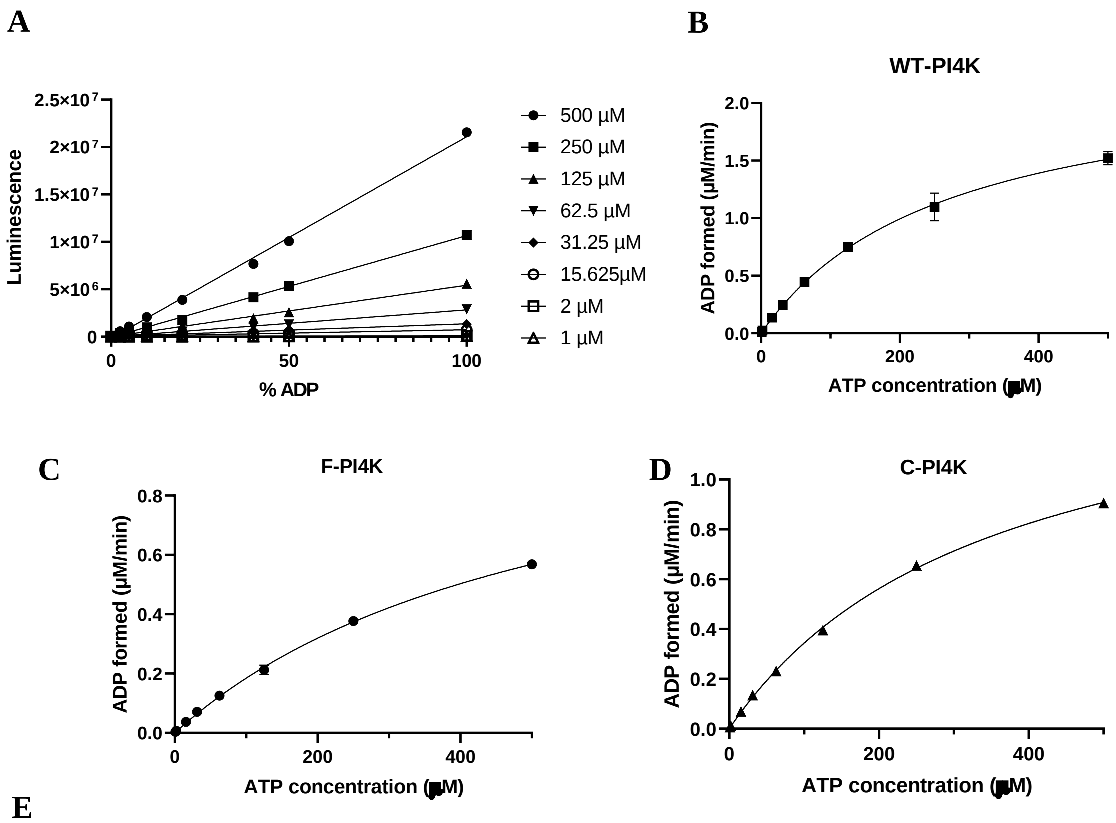
V_{max}	**0.09 ±0.02 µM/min**
K_{m}	**160 ± 33 µM**

V_{max}	**0.64 ± 0.04 µM/min**
K_{m}	**162 ±27 µM**

Figure 4.12: *Plasmodium vivax* phosphatidylinositol 4-kinase (*Pv*PI4K) L-α-phosphatidylinositol (PI) K_{m} determination.

$K_{\mathrm{m}}^{\mathrm{PI}}$ was determined by performing the assay using a two-fold serial dilution of PI and 7 nM *Pv*PI4K-WT at **(A)** 10 µM ATP and **(B)** 100 µM ATP. The kinase assay was performed as previously described. Mean values and standard deviations are based on n=2 independent experiments.

Determining the $K_{\mathrm{m}}^{\mathrm{ATP}}$ is important as many kinase inhibitors are ATP-competitive, which means that their IC_{50} values will be dependent on the ATP concentration and the affinity of ATP for the active site of the kinase in question. The *Pv*PI4K $K_{\mathrm{m}}^{\mathrm{ATP}}$ was determined by performing the assay at 100 µM PI, 7 nM *Pv*PI4K, and a two-fold serial dilution of ATP (500 µM starting concentration). The amount of ADP produced in the kinase reaction was estimated using ATP-to-ADP conversion standard curves (Figure 4.13A, Appendix Table E1). The Michaelis-Menten model was fitted to the data using non-linear regression and $K_{\mathrm{m}}^{\mathrm{ATP}}$ and K_{cat} values were determined.



E

Enzyme	**K_m^{ATP} (µM)**	**K_{cat} (min^{-1})**	**(K_{cat}/K_m^{ATP})**
*Pv*PI4K-WT	301 ± 28	125 ± 54	0.4
*Pv*PI4K-F832A	484 ± 42	38 ± 11	0.07
*Pv*PI4K-C1327A	405 ± 53	141 ± 58	0.3

Figure 4.13: *Plasmodium vivax* phosphatidylinositol 4-kinase (*Pv*PI4K) K_m^{ATP} determination.

K_m^{ATP} was determined by performing the assay using a two-fold serial dilution of ATP, 100 µM PI, and *Pv*PI4K. The kinase assay was performed as previously described. **(A)** ATP-to-ADP conversion standard curves were generated at each ATP concentration used. **(B)** *Pv*PI4K-WT ATP Michaelis-Menten plot. **(C)** *Pv*PI4K-F832A ATP Michaelis-Menten plot. **(D)** *Pv*PI4K-C1327A ATP Michaelis-Menten plot. **(E)** Determined K_m^{ATP}, *Kcat* and K_{cat}/K_m^{ATP} values for *Pv*PI4K-WT, *Pv*PI4K-F832A and *Pv*PI4K-C1327A. Mean values and standard deviations are based on n=2 independent experiments.

The data shows that K_m^{ATP} for *Pv*PI4K-WT is 301 µM, which is higher than reported K_m^{ATP} values for *Pf*PI4KIIIβ N-CAT and Human PI4KIIIβ (79 and 90 µM, respectively; Sternberg and Roepe, 2020). The difference in K_m^{ATP} between *Pv*PI4K-WT and *Pf*PI4KIIIβ N-CAT may be due to *Pf*PI4KIIIβ N-CAT being a heavily truncated construct (*Pv*PI4K-WT in this study is full length) and the use of a different assay technique to determine the K_m^{ATP}. The K_m^{ATP} values for *Pv*PI4K-F832A and *Pv*PI4K-C1327A were slightly higher but comparable to the wild-type enzyme.

The data in Figure 4.13E shows *Pv*PI4K-WT and *Pv*PI4K-C1327A have a similar K_{cat} and a similar catalytic efficiency (K_{cat}/ K_m^{ATP}). *Pv*PI4K-F832A displayed a significantly lower K_{cat}

and consequently catalytic efficiency, demonstrating that the F832A mutation has a significant effect on either the enzyme's catalytic activity or stability (leading to a lower proportion of active protein in the preparation).

As most kinase inhibitors are ATP-competitive, it is important to note the kinase K_m^{ATP}, particularly when studying inhibitor potency as the IC_{50} (the concentration at which kinase activity is 50% inhibited) is dependent of the level of ATP competition under assay conditions (Knight and Shokat, 2005). The K_m^{ATP}, IC_{50} and K_i (the intrinsic affinity of the inhibitor) are related to each other by the Cheng-Prusoff equation:

$$IC_{50} = K_i\ (1+ [ATP] / K_m^{ATP})$$

The above equation shows that at low ATP concentration i.e. lower than K_m^{ATP}, IC_{50} is approximately equal to K_i, when the ATP concentration is higher than K_m^{ATP}, the IC_{50} will increase with increase in ATP.

4.4 Conclusion

Site-directed mutagenesis was used to introduce mutations at F832 and C1327 in *Pv*PI4K-WT. The three enzymes *Pv*PI4K-WT, *Pv*PI4K-F832A and *Pv*PI4K-C1327A were successfully expressed and purified. Kinetic characterisation of *Pv*PI4K-WT revealed that K_m^{PI} was 160 µM and K_m^{ATP} was 301 µM. *Pv*PI4K-F832A and *Pv*PI4K-C1327A had similar K_m^{ATP} values, 484 µM and 405 µM respectively, indicating the mutations did not have a major effect on the enzyme's ATP binding affinity. This information was used to determine optimal PI4K enzyme assay parameters for future inhibitor studies. The catalytic efficiency of the three enzymes was calculated revealing *Pv*PI4K-WT and *Pv*PI4K-C1327A had relatively similar catalytic efficiency, while *Pv*PI4K-F832A had a significantly lower catalytic efficiency indicating the mutation has a significant effect on enzyme activity and/or stability.

Chapter 5

Evaluation of *Plasmodium* phosphatidylinositol 4-kinase ATP-competitive and covalent inhibitors

5.1 Introduction

In this chapter, *in vitro* inhibition assays utilising purified recombinant *Pv*PI4K wild-type and mutant proteins were established to study ATP-competitive and targeted covalent inhibitors. Following miniaturisation and validation of the *Pv*PI4K kinase inhibition assay, a range of inhibitors were tested to gain insight into their mechanism of inhibition and key protein-inhibitor interactions.

5.1.1 ATP-competitive inhibitors

Most kinase inhibitors approved for clinical use are Type I inhibitors (Zhao and Bourne, 2020). Type I inhibitors – also known as ATP-competitive inhibitors – function by binding to the ATP-binding pocket of the kinase while the enzyme is in its ‘active’ conformation (Bhullar *et al*., 2018). ATP-competitive inhibitors typically acquire their selectivity by extending into other regions proximal to the ATP adenine binding site including the DFG motif, the hydrophobic back pocket, the P-loop and the front pocket regions of the kinase (Zhao and Bourne, 2020). ATP-competitive inhibitors tested in this study included i) representative *Plasmodium* PI4K compounds from the aminopyridine/pyrazine (SFK40) chemical series and ii) *Plasmodium* PI4K hits identified from the Medicines for Malaria Venture (MMV) Pathogen Box.

Plasmodium PI4K inhibitors from the aminopyridine/pyrazine (SFK40) series

Three known *Plasmodium* phosphatidylinositol 4-kinase **(**PI4K) ATP-competitive inhibitors from the 3,5-diaryl-2-aminopyridine/pyrazine (SFK40) series (Figure 5.1) were evaluated to elucidate the importance of the Pi-Pi stacking interactions with phenylalanine 827 and how disruption of these interactions may affect inhibitor potency.

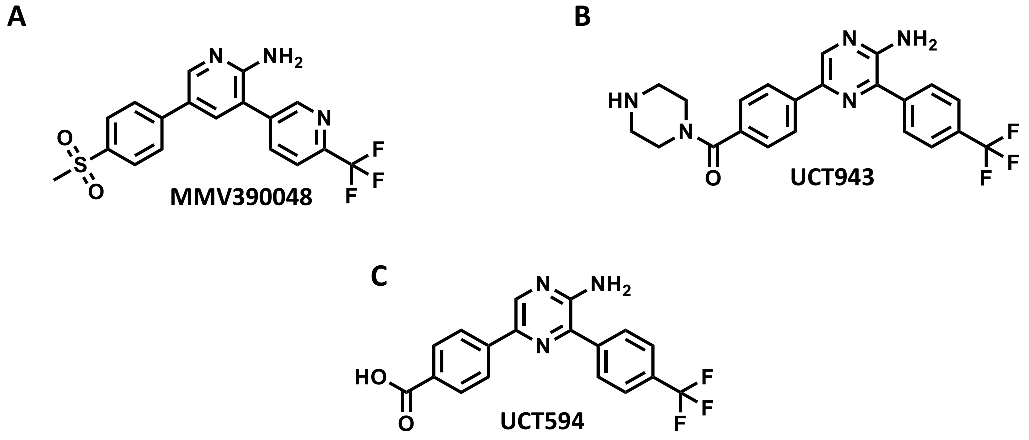
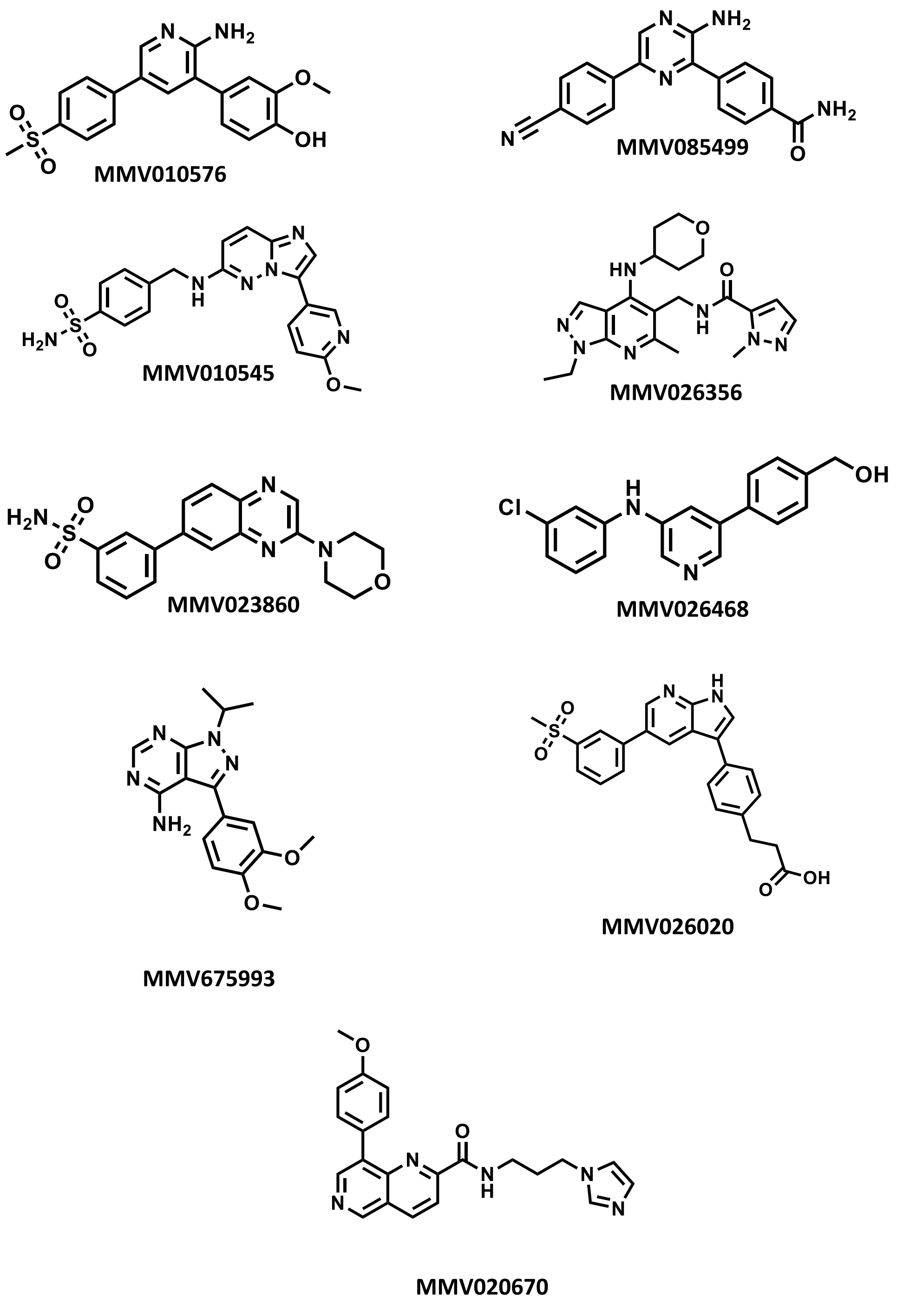


Figure 5.1: Known *Plasmodium* phosphatidylinositol 4-kinase ATP-competitive inhibitors.
A) MMV390048. **B**) UCT943. **C**) UCT594

Plasmodium PI4K hits from the Medicines for Malaria Venture (MMV) Pathogen Box

The Medicines for Malaria Venture (MMV) Pathogen Box – a collection of 400 drug-like compounds active against the causative agents of various neglected tropical diseases – was screened by the Jacquin Niles' group at Massachusetts Institute of Technology (MIT) to identify potential *Plasmodium* PI4K inhibitors. Screening of the MMV pathogen box using a cell-based *Plasmodium* PI4K conditional knockdown (cKD) assay, which utilizes the *Pf*DOZI-TetR system and is based on an increase in parasite sensitivity to inhibitors upon knockdown of PI4K, resulted in 9 hits (Figure 5.2). These hits, which included inhibitors (MMV010576 and MMV085499) from the aminopyridine/pyrazine (SFK40) series were assessed for *in vitro Pf*PI4K inhibition to validate direct interaction with the target and docking studies were used to predict the binding poses for these compounds.



MMV010576

MMV085499

MMV010545

MMV026356

MMV023860

MMV026468

MMV675993

MMV026020

MMV020670

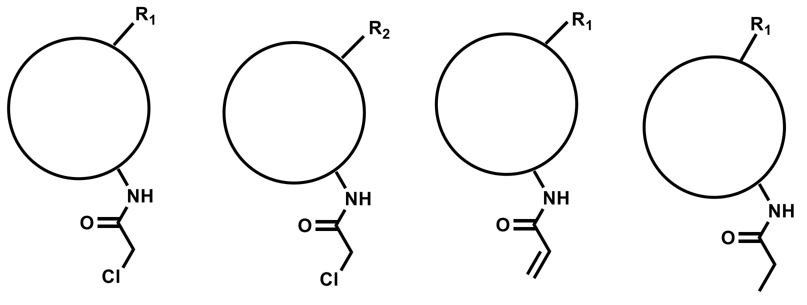
Figure 5.2: Medicines for Malaria Venture Pathogen Box hit compounds.

5.1.2 *Plasmodium* PI4K targeted covalent inhibitors

Recently, interest in developing Type VI covalent inhibitors has increased. This is in an attempt to address drug resistance and improve the duration of action, selectivity and efficacy (Sanderson, 2013) of kinase inhibitors. Covalent kinase inhibitors typically consist of an ATP-competitive inhibitor scaffold – which interacts with the ATP-binding region of the kinase –

attached to a reactive electrophilic group – known as a warhead – which reacts with a cysteine in the kinase active site, creating an irreversible enzyme-inhibitor complex (Martinez III *et al*., 2020). An important aspect of developing a covalent kinase inhibitor is achieving the correct balance between efficacy, selectivity and reactivity (Ghosh *et al*., 2019). This involves considering the binding potency of the reversible ligand and the reactivity of the attached warhead (Gehringer and Laufer, 2018). The reactivity of the warhead should be high enough to form a covalent bond but not too high in order to minimise the risk of off-target interactions (Gehringer and Laufer, 2018). As previously mentioned, Kulkarni *et al*. (2020) used bioinformatic analysis to identify and propose *Pf*PI4K C1361, situated on the periphery of the ribose binding site, as a target for covalent inhibitors as it is unique to *Plasmodium* PI4K, conserved across *Plasmodium* species and is absent in the human ortholog. Covalent inhibitors bind to the target and form an inhibitor-target complex and any unbound drug clears rapidly from the body (Kulkarni *et al*., 2020). The bound drug maintains the pharmacological effect while free drug is cleared, potentially reducing off-targets effects related to reversible binding thus reducing the risk of resistance arising due to continuous drug treatment – an issue commonly faced in malaria. However, it is important to consider toxicity related to potential covalent modification of off-targets- challenges often faced when developing covalent inhibitors (Aljoundi *et al*., 2020).

*Pv*PI4K inhibitors synthesised as part of a MMV/Merck/H3D drug discovery project (Figure 5.3) were used to establish an evaluation strategy for covalent inhibitors. These included putative targeted covalent inhibitors (TCIs) with a chloroacetamide (MMV1848400 and MMV190872) and acrylamide (MMV1792459) warheads and a matched pair lacking a warhead serving as a control (MMV1793498).



MMV1848400 MMV1902872 MMV1792459 MMV1793498

Figure 5.3: *Plasmodium* phosphatidylinositol 4-kinase targeting covalent inhibitors.

Once proof of principle was established, additional TCIs based on a different chemotype (Prevo *et al*., 2012) were designed, synthesized and assessed. GlaxoSmithKline Kinobead data – in which promiscuous kinase inhibitors immobilised on Sepharose beads are used to find potential kinase targets from a cell lysate –screening human kinase inhibitors for whole cell antiplasmodium activity revealed that human serine/threonine protein kinase ART (ataxia telangiectasia and Rad3-related protein) inhibitor VE-821 inhibits *Plasmodium* PI4K. Given the structural similarity VE-821 shares with *Plasmodium* PI4K inhibitors from the aminopyridine/pyrazine (SFK40) series, analogues were designed (Figure 5.4) in a repositioning approach by Ferdinand Ndubi (UCT). This was done towards improving *Plasmodium* PI4K selectivity and potency by improving compound binding within the ATP-binding site and adding a warhead to facilitate covalent bond formation.

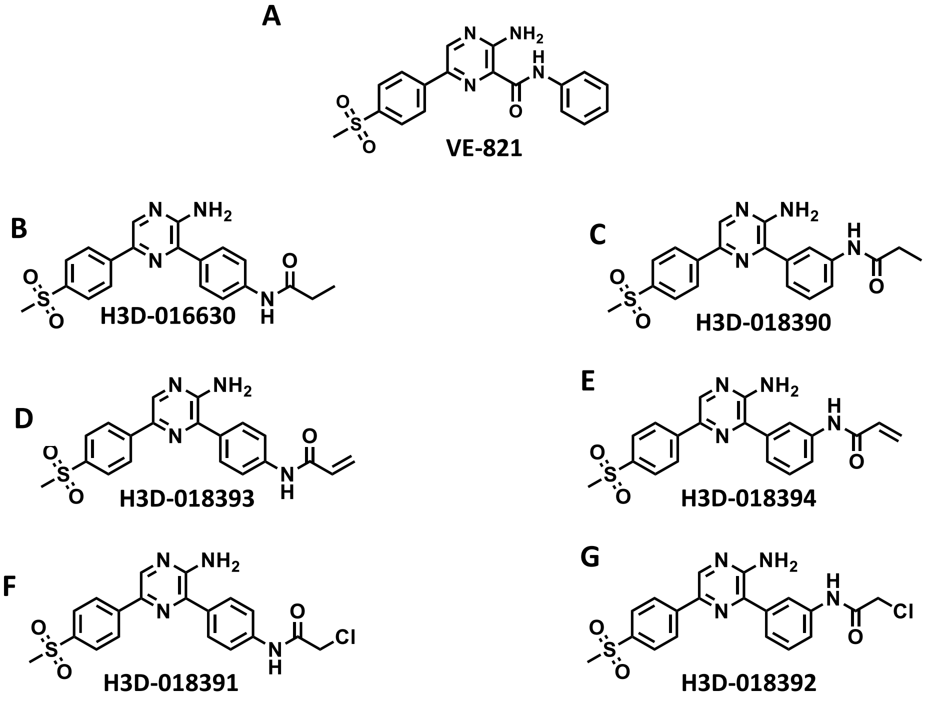


Figure 5.4: VE-821 and designed *Plasmodium* phosphatidylinositol 4-kinase targeting covalent inhibitor analogues.
A) VE-821. **B** and **C**) VE-821 competitive (control) analogue inhibitors. **D** and **E**) VE-821 acrylamide warhead analogues. **F** and **G**) VE-821 chloroacetamide warhead analogues.

This chapter focuses on evaluation of competitive *Plasmodium* PI4K inhibitors using inhibition assays and *in silico* docking and establishing a *Plasmodium* PI4K covalent inhibitor evaluation strategy using the ADP-Glo™ kinase assay and mass spectrometry.

5.2 Methods

5.2.1 Inhibition assay

The kinase assay described in section 4.2.23 was adapted to a miniaturised format. The kinase reaction mix contained 10 nM *Pv*PI4K-WT or 25 nM *Pv*PI4K-F832A or 7nM *Pv*PI4K-C1327A (final concentration), 100 µM PI, 10 µM ATP, inhibitor (three-fold serial dilution) and 1% DMSO. The three-fold inhibitor dilution was prepared in DMSO and diluted in assay buffer (25 mM HEPES at pH 7.5, 100 mM NaCl, 0.2% (v/v) Triton X-100, 1 mM DTT, 0.025 mg/mL BSA, and 6 mM $MgCl_2$) to 1.5X final concentration in a 96-well plate (15 µM highest concentration). DMSO only was used in the negative control reaction (100% activity) and 20 µM MMV390048 was used in the positive control reaction (100% inhibition). 2 µL inhibitor was transferred to a 384-shallow well plate and 0.5 µL enzyme (at 6X final concentration) and incubated for approximately 5 min (enzyme-inhibitor pre-incubation). 0.5 µL substrate buffer (at 6X final concentration containing 600 µM PI and 60 µM ATP) was added using a Mantis® Liquid Handler (Formulatrix®). The kinase reaction was incubated for 40 min at 22 °C. 2 µL ADP-Glo™ reagent was added to the reaction and incubated at 22 °C for 40 min. 2 µL kinase detection reagent was added to the reaction. The reaction was incubated at 22 °C for 40 min and luminescence was measured using an EnSpire multimode plate reader (PerkinElmer). ATP to ADP conversion was estimated using a standard curve prepared according to the manufacturer's instructions and the data was normalised according to negative (no inhibitor) and positive (100% inhibition) controls. IC_{50} values were calculated using GraphPad Prism (nonlinear regression). (H3D numbers and their corresponding MMV numbers for all inhibitors evaluated can be found in Appendix F).

5.2.2 *In silico* docking

Docking of compounds into the *Plasmodium falciparum* phosphatidylinositol 4-kinase homology model was done by Stephen Fienberg (Drug Discovery and Development Centre) using the model generated in Fienberg *et al*. (2020).

5.2.3 Covalent inhibitor time-course assay

Kinase assays were performed as previously described in section 4.2.23 at 7 enzyme-inhibitor pre-incubation time points – 5 min, 30 min, 2 hrs, 6 hrs, 12 hrs, 18 hrs and 24 hrs. A three-fold dilution of the inhibitor was prepared in DMSO and diluted in assay buffer (25 mM HEPES at pH 7.5, 100 mM NaCl, 0.2% (v/v) Triton X-100, 1 mM DTT, 0.025 mg/mL BSA, and 6 mM $MgCl_2$) to 1.5X final concentration (15 µM highest concentration). DMSO was used in the negative control reaction (100 % activity) in lieu of inhibitor and 2 µM MMV390048 was used in the positive control reaction (100% inhibition). A 'master mix' of containing 48 µL 1.5X inhibitor and 12µL 60 nM *Pv*PI4K-WT starting concentration or 42 nM *Pv*PI4K-C1327A starting concentration was prepared in a 96-well plate, sealed and incubated at 22 °C. At each time point, 2.5 µL of the master mix was transferred to a white 384-shallow well plate and the assay was performed as described in section. 5.2.1.

5.2.4 Enzyme-inhibitor reaction for mass spectrometry

*Pv*PI4K-WT was diluted to 0.75 mg/mL in assay buffer (25 mM HEPES at pH 7.5, 100 mM NaCl, 0.2% (v/v) Triton X-100, 1 mM DTT, 0.025 mg/mL BSA, and 6 mM $MgCl_2$). A master mix containing 32 µL assay buffer and 20 µL 100 µM inhibitor was prepared. Enzyme-inhibitor reactions were set up containing 2.9 µM *Pv*PI4K-WT final concentration and 13 µM inhibitor final concentration. The reactions were incubated on ice at 5 time points – 5 min, 10 min, 15min 30 min and 2hr. The reactions were quenched by the addition of iodoacetamide (IAA) to a final concentration of 5mM, incubation in the dark at RT for 30 min and subsequent addition of 0.5% formic acid. The samples were stored at -80 °C.

5.2.5 Mass Spectrometry sample preparation

To each reaction, 40 µl 50 mM Tris-1mM DTT (pH 8.5) followed by 5.11 µl 1M Tris (pH 8.5) was added. Trypsin and chymotrypsin were added to a final concentration of 100 ng each. The samples were incubated at 37 °C for 16 hrs. The samples were cleaned up using Evotips® (Evosep Biosystems) according to the manufactures protocol.

5.2.6 Mass Spectrometry analysis

Approximately 200 ng digested *Pv*PI4K was loaded to a dionex RS300 LC using a 20 cm, 75 µm ID column packed inhouse with 1.9 µm Reprosil-Pur C18 beads (Dr. Maisch, Ammerbuch, Germany) coupled to a Q-Exactive mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). Peptides were separated using a 10 minute linear gradient from 5% solvent B (0.1% FA, ACN) to 30% Solvent A (2% ACN, 0.1% FA) at 300 nL/min followed by a 5 minute washout at 80 % B while being maintained at 40^o C.

Mass spectra were collected on a Q Exactive mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) operating in positive mode, with data-dependent acquisition and a top-10 method. Intensity threshold for MS^2 ion selection was 1.3e4 with charge exclusion of $z = 1$ and $z > 5$. Peptides were ionised by electrospray ionisation, and MS spectra were acquired at a resolution of 70,000 for MS1 and 17,500 for MS/MS. Automated gain control (AGC) target was set to 1e6 with a maximum integration time of 30 ms (MS1) and 1e5 with a maximum IT of 80 ms (MS/MS). MS1 scan range was 300–1,750 Da, and peptide fragmentation was performed using higher-energy collision dissociation (HCD) and setting the energy to 28 Normalised Collision Energy (NCE).

The raw data files were processed with the MaxQuant software (version 1.6.14) using the Andromeda search engine to search MS/MS data against the *Pv*PI4K sequence and available Spodoptera frugiperda sequences retrieved from UniProt with common contaminants. Carbamidomethylation of cysteine and oxidation of methionine were set as variable modifications. Additionally, the expected mass shifts induced by cysteine covalent modification by the respective covalent inhibitors were calculated and entered into MaxQuant and set as variable cysteinc modifications. Trypsin/P and Chymotrypsin were selected as the proteases with 5 missed cleavages allowed. Precursor mass tolerance was 4.5 ppm and MS/MS mass tolerance was 20 ppm for HCD fragmentation data analysis. The false discovery rate (FDR) for peptides and proteins was set to 1%. For all other parameters, the default settings were used.

5.3 Results and Discussion

5.3.1 Inhibition Assay Validation

The kinase assay was adapted to a miniaturised format for inhibition assays. The robustness of the assay was validated by performing the assay using only negative and positive controls (DMSO only and 2 µM MMV390048 respectively) in the plate format shown in Figure 5.5A. The kinase assay was performed as described in section 5.2.1 and the average luminescence signals and error was determined Figure 5.5B.

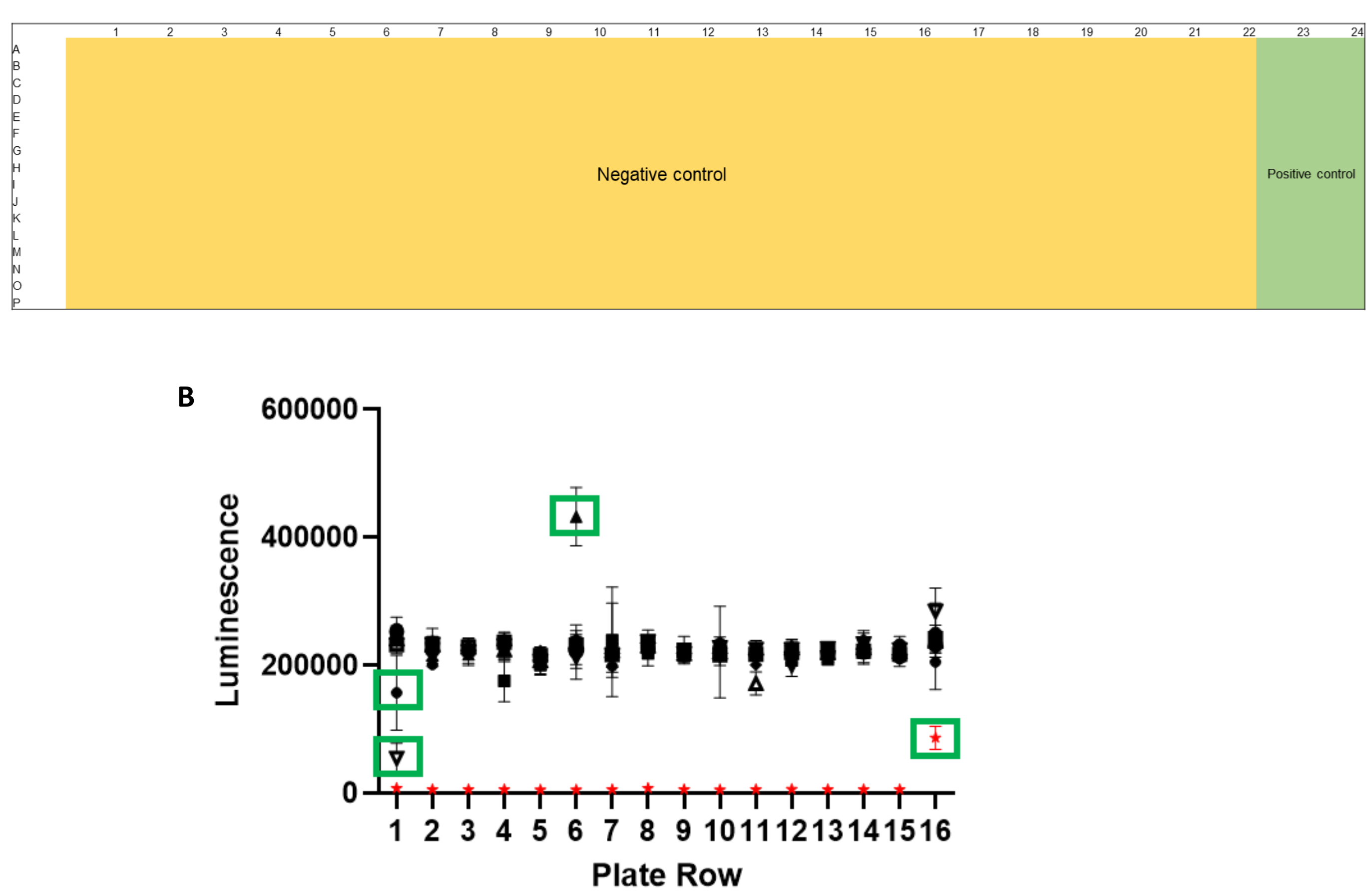


Figure 5.5: Validation of miniaturised ADP-Glo™ kinase assay.
A) Inhibition assay plate format. **B**) Luminescence for positive and negative control reactions. Negative control (black points) - 0% *Pv*PI4K-WT inhibition; Positive control (red points) – 100% *Pv*PI4K-WT inhibition. Green boxes highlight outliers. Representative plot of 2 separate experiments shown above with technical duplicates.

Two separate validation assays were performed and the data was comparable. Figure 5.5B shows good separation between the no inhibition (negative control) and 100 % inhibition (positive control). The signal-to-noise ratio – which compares the assay signal to background noise – was calculated (Table 5.1) and showed there was an 11-fold difference between the assay signal and background noise indicating high confidence in the data obtained using this assay format. The coefficient of variation (CV) is a measure of variability defined by the standard deviation of a set of measurements divided by the mean, often presented as a percentage (Canchola *et al*., 2017). According to Schultheiss and Stanton (2009), an intra-assay CV of less than 10 % indicates good reliability. In the validation assay, the CV for the negative control was 4.7 % and the CV for the positive control was 9.4 % indicating the data produced in this assay is reliable. Further evaluation of the assay quality was determined by calculating the Z-factor. The Z-factor is a measure of the degree of separation between the positive control and the negative control (Zhang *et al*., 1999). According to Zhang *et al*. (1999), the Z-factor value reflects the quality of the assay. A Z-factor of 1 – the maximum value – indicates an 'ideal' assay, a Z-factor between 0.5 and 1 indicates an 'excellent assay', a Z-factor between 0 and 0.5 is 'marginal' and 0 is not a good assay for screening (Zhang *et al*., 1999). The miniaturised kinase assay Z-factor was determined to be 0.7, thus the assay can be classified as an 'excellent' assay with high reproducibility and reliability.

Table 5.1 Summary of miniaturised kinase assay validation values.

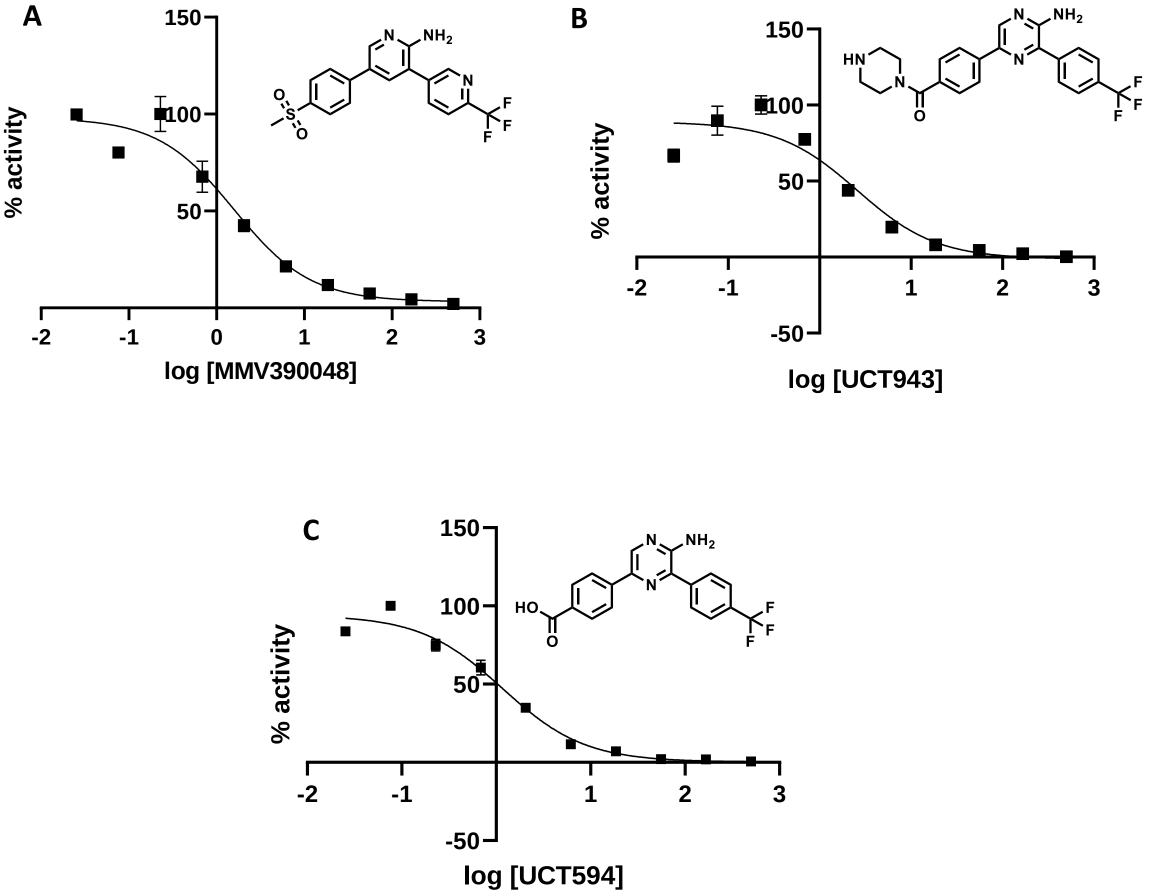
	Luminescence	
	Average	**Standard deviation**
Positive Control	5956.55	967.8
Negative control	222947.6	19514.6
Positive control CV (%)	9.4	
Negative control CV (%)	4.7	
Signal-to-Noise ratio	11	
Z-factor	0.7	

CV-coefficient of variation

5.3.2 Evaluation of ATP-competitive *Plasmodium* phosphatidylinositol 4-kinase (PI4K) inhibitors

Aminopyridine/pyrazine (SFK40) series

Inhibition of *Pv*PI4K-WT was studied using three known PI4K inhibitors from the SFK40 series of compounds: MMV390048, UCT943 and UCT594. Inhibition assays were performed using 7 nM *Pv*PI4K-WT, 100 µM PI, 10 µM ATP, and a three-fold dilution of inhibitor. An ATP concentration << K_m (301 µM) was used so that IC_{50} approximates K_i.



	MMV390048	UCT943	UCT594
***Pv*PI4K-WT IC_{50} (nM)**	2.2 ± 0.6	2.1 ± 0.5	1.3 ± 0.1
***Pf*NF54 IC_{50} (nM)[*]**	28	5	21

Figure 5.6 Inhibition data for known PI4K inhibitor compounds against *Pv*PI4K-WT. Inhibition assays were performed using 7 nM *Pv*PI4K-WT, 100 µM PI, 10 µM ATP. Representative IC_{50} curves are shown for **(A)** MMV390048 **(B)** UCT943 and **(C)** UCT594. Mean values and standard deviations are based on n=2 independent experiments with technical duplicates. [*]Reported values.

The data in Figure 5.6 shows MMV390048, UCT943 and UCT594 are potent *Pv*PI4K inhibitors with IC_{50} values of 2.2 nM, 2.1 nM, and 1.3 nM, respectively. MMV390048 inhibition was similar to previously reported values for *Pv*PI4K (3.4 nM; Paquet *et al.*, 2017). MMV390048 was also very potent against an engineered *Pf*PI4K construct (*Pf*PI4KIIIβ N-CAT) - IC_{50} = 1.32 nM (Sternberg and Roepe, 2020). However, *Pv*PI4K IC_{50} values for

UCT943 and UCT594 were ~10-fold lower than previously reported (23 nM and 24 nM, respectively (Brunschwig *et al*., 2018; Fienberg *et al*., 2020). Previously reported assays for UCT943 and UCT594 were carried out using an ATP concentration of 100 µM and the Transcreener® ADP2 FI Assay for ADP detection. Based on the Cheng-Prusoff equation, the difference in ATP concentration used in the assay cannot account for the ~10-fold difference in IC_{50} values. Since *Pv*PI4K has been identified as the primary target for all three compounds, the *Pv*PI4K IC_{50} data here correlates better with the potent asexual blood stage antiplasmodium activity reported for these compounds (*Pf*NF54 IC_{50} for MMV390048, UCT943 and UCT594 of 28 nM, 5 nM and 21 nM respectively (Paquet *et al*., 2017; Brunschwig *et al*., 2018). The difference in IC_{50} values in the biochemical assay and the whole cell assay may also be due to slow inhibitor uptake into the parasite, possible accumulation of the inhibitor and differences in ATP concentrations in the whole cell assay compared to the concentration used in the biochemical assay, all of which could affect inhibitor IC_{50}.

The low *Pv*PI4K IC_{50} values observed for these inhibitors (i.e. lower than the theoretical minimum IC_{50} of 3.5 nM based on half the enzyme concentration) indicates that although 7 nM of the purified *Pv*PI4K protein was used in the assay, the concentration of active *Pv*PI4K was lower. Due to the potency of the inhibitors, it is likely that the 'assay wall' effect is occurring in which the IC_{50} observed is limited by the kinase concentration used in the assay thus the 'true' IC_{50} (and corresponding Ki) may be lower than what was observed for these inhibitors (Hafenbradl *et al*., 2011). The 'assay wall' is a limiting factor when evaluating highly potent inhibitors as assays reach a point in optimisation where the enzyme concentration cannot feasibly be reduced any further.

As previously mentioned, the relationship between an ATP-competitive inhibitor and its ligand can be defined by the Cheng-Prusoff equation under classical conditions:

$$IC_{50} = K_i\ (1+ [ATP] / K_m^{ATP}$$

Inhibitors can be confirmed as ATP-competitive inhibitors using the Cheng-Prusoff equation as an increase in ATP concentration results in an increase in $IC_{50.}$ Using the equation and the determined *Pv*PI4K-WT K_m^{ATP} (301 µM), the expected shift for *Pv*PI4K-WT was determined to be 2.66 when the ATP concentration in the assay was increased from 10 µM to 500 µM, provided the assay wall is not reached (i.e assay is performed under classical conditions). *Pv*PI4K IC_{50} assays for MMV390048, UCT943 and UCT594, which will henceforth be referred to as ‘potent SFK40 *Plasmodium* PI4K inhibitor compounds’, were performed at 10 µM and 500 µM ATP in order to determine the IC_{50} shift (Figure 5.7 and Table 5.2).

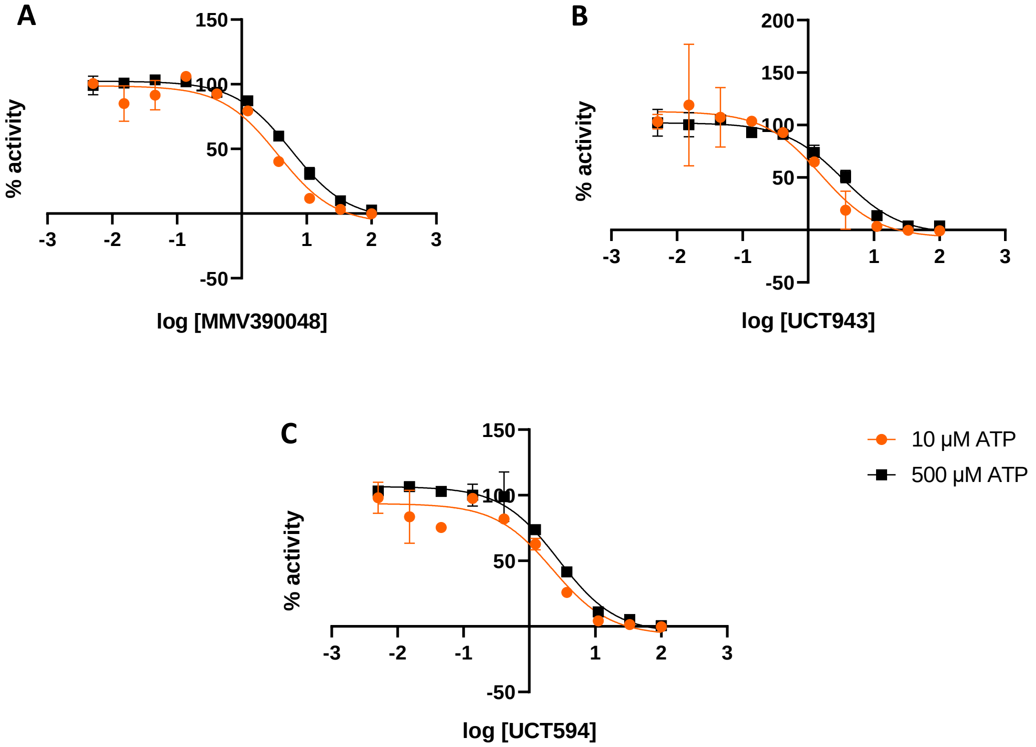


Figure 5.7: ATP shift evaluation of potent SFK40 *Plasmodium* PI4K inhibitor compounds against *Pv*PI4K-WT.
Inhibition assays were performed using 7 nM *Pv*PI4K-WT, 100 µM PI, 10 µM ATP and 500 µM ATP. The assays were performed in parallel using the same dilution series. **(A)** MMV390048 **(B)** UCT943 and **(C)** UCT594. Mean values and standard deviations are based on n=2 independent experiments with technical duplicates.

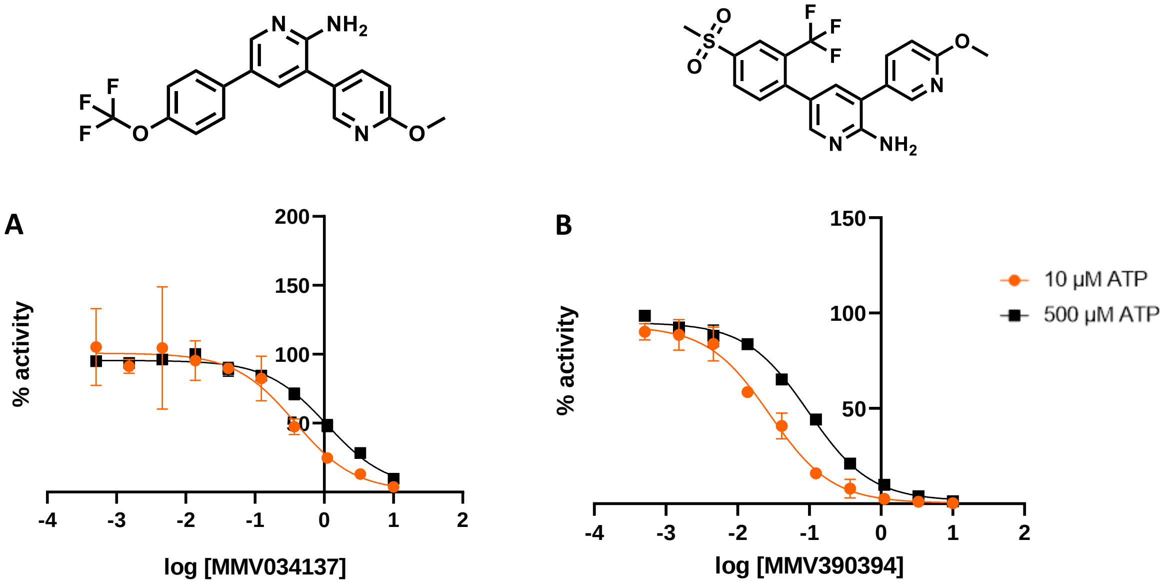
Table 5.2: IC_{50} shift determination for known PI4K inhibitor compounds against *Pv*PI4K-WT.

	IC_{50} nM (10 µM ATP)	**IC_{50} nM (500 µM ATP)**	**IC_{50} fold change**
MMV390048	2.4 ± 0.7	4.5 ± 1.3	1.9
UCT943	1.49 ± 0.05	2.8 ± 0.5	1.9
UCT594	1.7 ± 0.5	2.2 ± 0.5	1.3

Mean values and standard deviations are based on n=2 independent experiments.

The data in Figure 5.7 and Table 5.2 shows a small but consistent shift in the IC$_{50}$ values of MMV390048, UCT943 and UCT594 at 10 µM ATP and 500 µM ATP – likely due to the high potency of these inhibitors (assay wall).

The calculated expected ATP fold change for ATP-competitive inhibitors targeting *Pv*PI4K-WT is very low (2.66) due to the enzyme's high K_m^{ATP} (301 µM) making it challenging to confirm this experimentally. The ATP shift assays were repeated using less potent PI4K inhibitors from the same series, MMV034137 (*Pf*NF54 IC$_{50}$ 2.7 µM) and MMV390394 (*Pf*NF54 IC$_{50}$ 0.95 µM) (Paquet *et al*., 2017), to determine whether a shift would be observed.

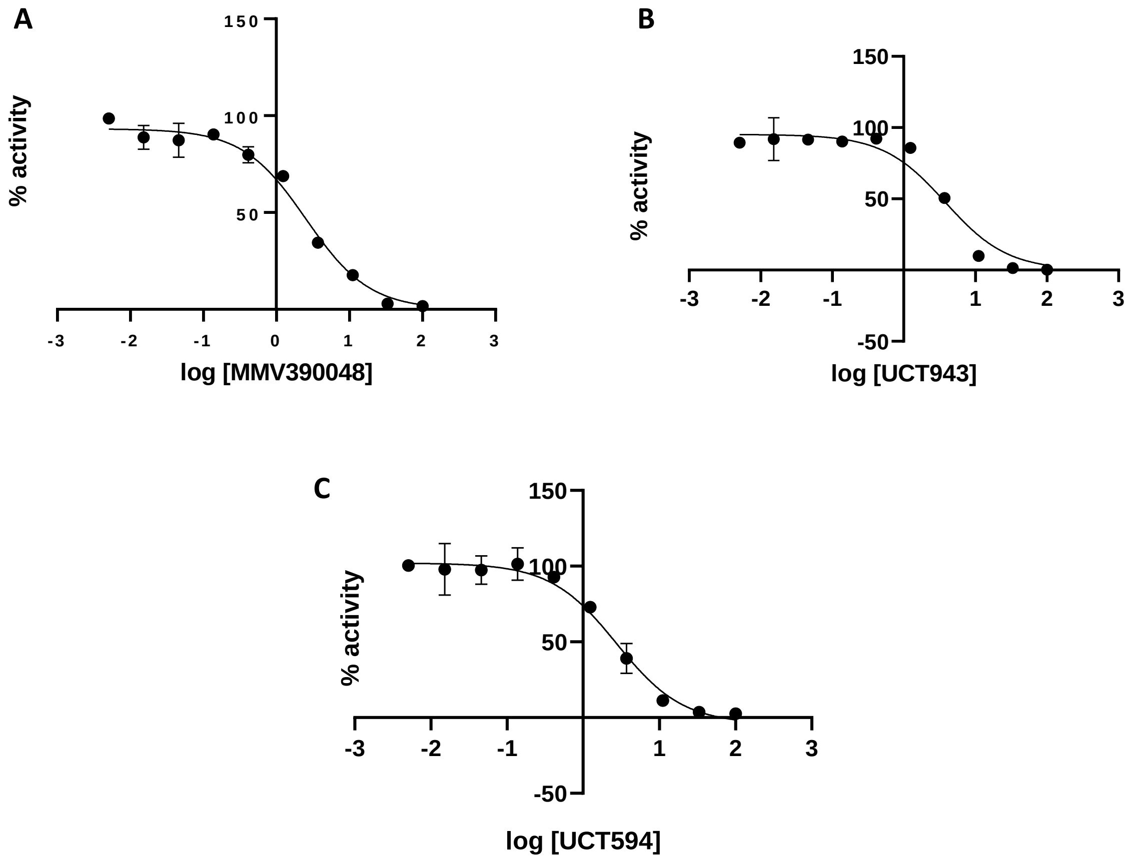


	IC$_{50}$ (10 µM ATP) µM	IC$_{50}$ (500 µM ATP) µM	IC$_{50}$ shift
MMV034137	0.36 ± 0.02	1.01 ± 0.08	2.8
MMV390394	0.0286 ± 0.0006	0.104 ± 0.008	3.4

Figure 5.8: ATP shift evaluation of MMV034137 and MMV390394 *Plasmodium* PI4K inhibitor compounds.
Inhibition assays were performed using 7 nM *Pv*PI4K-WT, 100 µM PI in the presence of10 µM ATP and 500 µM in parallel. **(A)** MMV390394 and **(B)** MMV034137. Mean values and standard deviations are based on n=2 independent experiments, with technical duplicates.

The data in Figure 5.8 shows a significant shift – closer to the predicted shift calculated using the Cheng-Prusoff equation. Overall, although the ATP shift assay can be used to determine whether an inhibitor is ATP-competitive or non-competitive, if the Km^{ATP} of the enzyme is high, the expected shift will be very small within the ATP concentration range compatible with the assay (up to 500 µM) and thus it can be difficult to observe a clear shift. Additionally, clear shifts are difficult to observe when evaluating potent inhibitors due to the tight binding assay conditions resulting in an assay wall (where $K_i^{app} \leq$ enzyme concentration).

Inhibition of *Pv*PI4K-F832A was evaluated using three known *Plasmodium* PI4K inhibitors: MMV390048, UCT943 and UCT594 in order to determine the extent to which disruption of the Pi-Pi stacking interaction predicted to occur between the phenylalanine 832 residue and aromatic rings found in these compounds affects inhibitor potency. The phenylalanine 832 residue is unique to *Plasmodium* thus assessing the importance of the Pi-Pi interaction between PI4K and inhibitors could provide useful information for guided inhibitor design that exploits this interaction. Inhibition assays were performed using 50 nM *Pv*PI4K-F832A, 100 µM PI, 10 µM ATP, and a three-fold dilution of inhibitor.



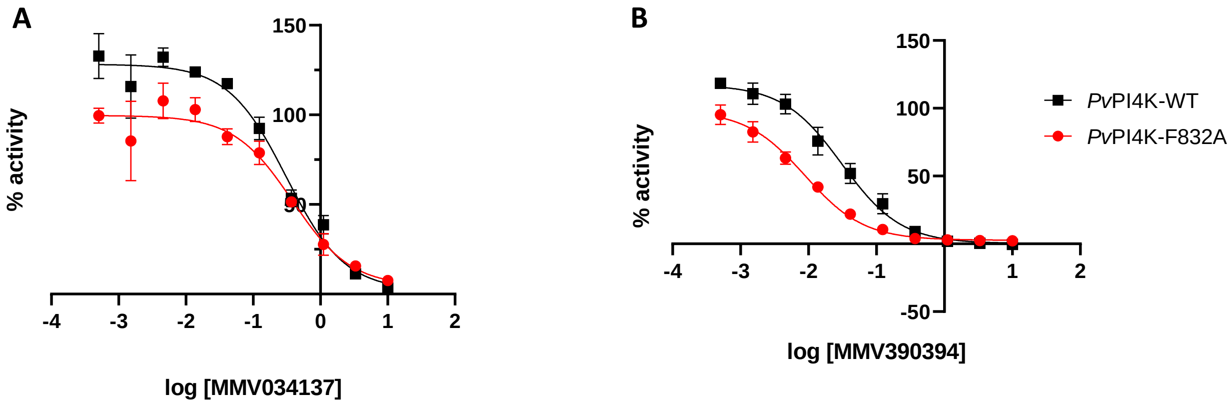
	MMV390048	UCT943	UCT594
***Pv*PI4K-F832A IC$_{50}$ (nM)**	3.0 ± 0.5	4.1 ± 0.3	3.1 ± 0.4

Figure 5.9: Inhibition data for potent SFK40 *Plasmodium* PI4K inhibitor compounds against *Pv*PI4K-WT and *Pv*PI4K-F832A.

Inhibition assays were performed using 50 nM *Pv*PI4K-F832A, 100 µM PI, 10 µM ATP, and **(A)** MMV390048 **(B)** UCT943 and **(C)** UCT594. Mean values and standard deviations are based on n=2 independent experiments with technical duplicates.

The data in Figure 5.9 shows MMV390038, UCT943 and UCT594 are potent inhibitors of *Pv*PI4K -F832A with IC$_{50}$ values of 3.0 nM, 4.1 nM and 3.1 nM respectively. The IC$_{50}$ values observed indicate only a fraction of the protein used in the assay is active which suggests the reduced activity observed for the mutant is likely also due to decreased levels of active protein as well as reduced catalytic activity. The IC$_{50}$ values for *Pv*PI4K -F832A inhibition (Figure 5.9) were similar to those observed for *Pv*PI4K -WT i.e. within 3-fold (Figure 5.6) although the assay wall makes it difficult to assess the effect of the mutation on these potent inhibitors. The specific activity of *Pv*PI4K-WT is 0.1 µmols/min/mg compared to that of *Pv*PI4K -F832A which is 0.02 µmols/min/mg whereas the *Pv*PI4K-C1327A specific activity was comparable to *Pv*PI4K-WT (0.08 µmols/min/mg). The low IC_{50} values (~3-4 nM) observed for the PvPI4K -F832A, comparable to the wild-type enzyme for potent inhibitors, despite the high enzyme concentration required for the assay (~50 nM), suggests that the lower specific activity is a result of a large proportion of inactive protein in the sample. As the K_m^{ATP} values for *Pv*PI4K-WT and *Pv*PI4K -F832A are similar, the IC_{50} data for these enzymes is comparable despite the differences in specific activity and low IC_{50}'s can be determined for both *Pv*PI4K -WT and *Pv*PI4K -F832A.

Less potent *Plasmodium* PI4K inhibitors from the same series (MMV034137 and MMV390394) were evaluated for activity against *Pv*PI4K -F832A.



	*Pv*PI4K-WT IC_{50} (µM)	*Pv*PI4K-F832A IC_{50} (µM)	IC_{50} shift
MMV034137	0.29 ± 0.02	0.44 ± 0.06	1.5
MMV390394	0.029 ± 0.003	0.007 ± 0.001	4.1

Figure 5.10: Inhibition data for MMV034137 and MMV390394 against *Pv*PI4K-WT and *Pv*PI4K-F832A.

Inhibition assays were performed using 7nM *Pv*PI4K-WT/50 nM *Pv*PI4K-F832A, 100 µM PI, 10 µM ATP, and **(A)** MMV034137 and **(B)** MMV390394 in parallel. Representative curves shown. Mean values and standard deviations are based on n=2 independent experiments with technical duplicates.

The data in Figure 5.10 shows a slight shift for MMV034137 (1.5-fold difference) and a significant shift for MMV390394 (4.1-fold difference). MMV034137 data suggests disrupting the Pi-Pi interaction between *Pv*PI4K and *Plasmodium* PI4K inhibitors containing aromatic rings does not significantly affect inhibitor potency as predicted. Unexpectedly, MMV390394 was more potent on *Pv*PI4K-F832A than on the wildtype. Docking of MMV390394 into the *Pf*PI4K homology model -done by Stephen Fienberg, UCT- (Figure 5.11) showed the trifluoromethyl (CF_3) group on the inhibitor clashes with the phenylalanine residue in the wildtype, accounting for the increase in potency resulting from the phenylalanine to alanine mutation.

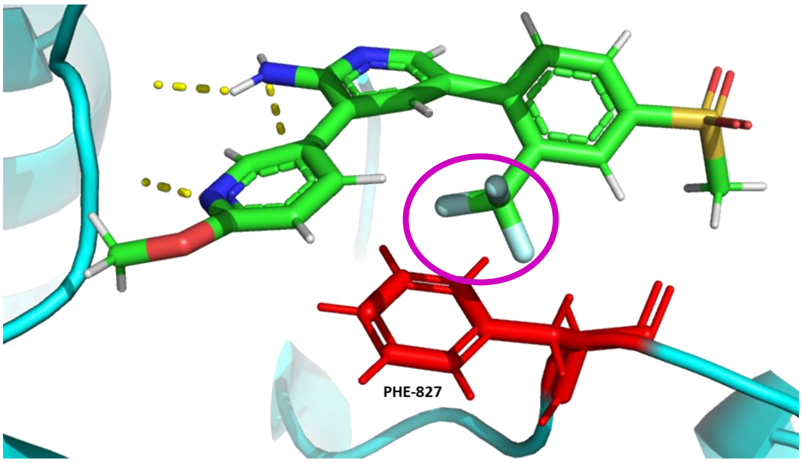


Figure 5.11: ATP-competitive inhibitor MMV390394 docked into *Plasmodium falciparum* phosphatidylinositol 4-kinase homology model.
ATP-competitive inhibitor MMV390394 (green) docked into the *Pf*PI4K homology model. F827 clashing CF_3 group highlighted in pink circle.

Plasmodium PI4K hits from the MMV Pathogen Box

As already mentioned, nine MMV Pathogen Box hits were identified by collaborators at MIT as putative PI4K inhibitors with distinct chemotypes and were assessed for *in vitro* inhibitory activity against *Pv*PI4K. These hits were identified using a cell-based cKD screen in which *Plasmodium* PI4K expression in parasites was controlled by varying the anhydrotetracycline (aTc) concentration in the culture conditions. High aTc conditions resulted in PI4K expression levels comparable with wild-type parasites, low aTc conditions resulted in low PI4K expression i.e. conditional knockdown of PI4K. Putative PI4K inhibitors were identified based on the increased sensitivity of parasites to compounds (decrease in IC_{50}) upon PI4K knockdown.

Of the 9 pathogen box hits evaluated (Table 5.3), most were potent to moderate *Pv*PI4K inhibitors with enzyme IC_{50} values less than 25 nM – except MMV026468 which had an IC_{50} value of 461 nM. The data observed validated the cKD hits as PI4K inhibitors indicating this cKD screening approach is a valuable phenotypic screening technique for identifying potential PI4K inhibitors.

Table 5.3: ***In vitro Plasmodium vivax*** **phosphatidylinositol 4-kinase inhibition activity of the MMV Pathogen Box hits.**

Compound	**High aTc/Low aTC (EC_{50} shift)****	***Pv*****PI4K-WT IC_{50} (nM)**
MMV010545	14.4	6 ± 1
MMV026468	8.8	461 ± 198
MMV026020	2.3	7 ± 3
MMW026356	17.0	5.2 ± 0.2
MMV020670	7.1	24.1 ± 0.4
MMV023860	n.d*	16 ± 6
MMV675993	n.d*	3.9 ± 0.2
MMV010576	n.d*	4 ± 2
MMV085499	16.0	1.7 ± 0.6

Mean values and standard deviations are based on n=2 independent experiments. EC_{50} data generated by Jacquin Niles lab (MIT) *n.d shift clearly observed but value could not be determined. **High aTc-regular PI4K expression cell line EC_{50}. Low aTC-low PI4K expression cell line EC_{50}.

Docking of these inhibitors (Table 5.3) into the *Pf*PI4K homology model (Stephen Fienberg, UCT) provided further insights. As expected and/or unsurprisingly MMV010576 and MMV085499 were found to have a similar binding mode to potent PI4K inhibitors from the SFK40 series e.g. MMV390048 while MMV010545 was found to have a similar binding mode to KDU691 (McNamara *et al*., 2013) another potent PI4K inhibitor with the option to optimise for improved binding. MMV023860 had a binding mode inverse to that observed for BQR695 (McNamara *et al*., 2013) also with the option of improved binding with further optimisation. Docking of MMV026356 showed no binding site pose in the ATP binding site of the *Pf*PI4K homology model suggesting the compound may have a different binding mode. MMV020670 had a poor binding pose in *Pf*PI4K which may explain its moderate potency while weak inhibitor MMV026468 had some similar binding motifs to those observed in other PI4K inhibitors but no unique interactions. Potent inhibitor MMV675993 had a binding mode similar to MLN0128 (a potent mammalian target of rapamycin (mTOR) inhibitor) while MMV026020 was also potent and displayed strong binding in the hinge region, front pocket region and back pocket region of the kinase. The relatively high potency and unique scaffolds observed for some of these inhibitors provides a starting point for the development of new PI4K inhibitors.

Another residue that could be mutated to probe for compound selectivity and potency is *Pf*PI4K Y1356. The tyrosine residue adds aromaticity to the hinge region which increases the potency of compounds that contain multiple aromatic groups as the Y1356 residue can form Pi-Pi stacking interactions with the hinge binding aromatic group (Fienberg *et al.*, 2020). Additionally, Y1356 could be used to probe for selectivity as the Y1356 position in HuPI3Ks contains nonaromatic lipophilic amino acids thus excluding Pi-Pi stacking interactions with human PI3K off-targets (Fienberg *et al.*, 2020). The ribose pocket S1365 residue has been shown to be important for compound selectivity and potency thus mutation of this residue to alanine or its corresponding amino acid in potential human PI3K off-targets could provide valuable insights into the potency and more importantly, selectivity, of MMV390048 and other SFK40 analogues.

Overall, the results show mutation of F832 to alanine had a minimal effect on inhibition of *Pv*PI4K suggesting Pi-Pi interactions with this residue are not key for potent inhibition for this series of compounds.

5.3.3 Establishing a *Plasmodium* PI4K targeted covalent inhibitor evaluation strategy

Plasmodium PI4K targeted covalent inhibitors (TCIs) designed to specifically target *Pv*PI4K-C1327 were evaluated using time-course inhibition assays to monitor the effect of enzyme-inhibitor incubation time on inhibitor potency.

Four tool inhibitor compounds, were used to establish the time-course assay for evaluating TCIs. The tool compounds used were MMV1793498 – a competitive inhibitor with no warhead, MMV1792459 – a TCI derivative of MMV1793498 containing an acrylamide warhead, MMV1848400 – a TCI derivative of MMV1793498 containing a chloroacetamide warhead and MMV1902872 – a TCI containing a chloroacetamide warhead but composed of a different ATP-competitive inhibitor scaffold Figure 5.3). Standard *Pv*PI4K inhibition assays (5 min enzyme-inhibitor preincubation time) were performed using 7 nM *Pv*PI4K-WT and 10 nM *Pv*PI4K-C1327A (Table 5.4).

Table 5.4: Targeted covalent inhibitor IC_{50} values from standard inhibition assay.

Compound	***Pv*PI4K-WT IC_{50}**	***Pv*PI4K-C1327A IC_{50}**
MMV1793498	79 ± 35	24 ± 11
MMV1792459	102 ± 33	23 ± 3
MMV1848400	30 ± 11	40 ± 17
MMV1902872	30 ± 20	20 ± 14.5

Mean IC50 ± standard deviation from n=2.

Data from the standard inhibition assay showed MMV1793498 and MMV1792459 were slightly more potent for *Pv*PI4K-C1327A compared to *Pv*PI4K-WT (3.2 fold and 4.4 fold respectively) while MMV1848400 and MMV1902872 IC_{50} values for *Pv*PI4K-WT and *Pv*PI4K-C1327A were comparable, indicating these inhibitors have a similar reversible binding potency at 5 min pre-incubation for both *Pv*PI4K-WT and *Pv*PI4K-C1327A.

Time-course inhibition assays were performed as described in section 5.2.3 over 24 hrs using *Pv*PI4K-WT and *Pv*PI4K-C1327A. The stability of the enzymes was determined over the time course of the assay in the absence of inhibitor (DMSO control) (Figure 5.12). The enzymes remained sufficiently active to ensure sufficient signal to background readings over the 24 hr time course but a notable decrease in activity was observed after 6 hrs.

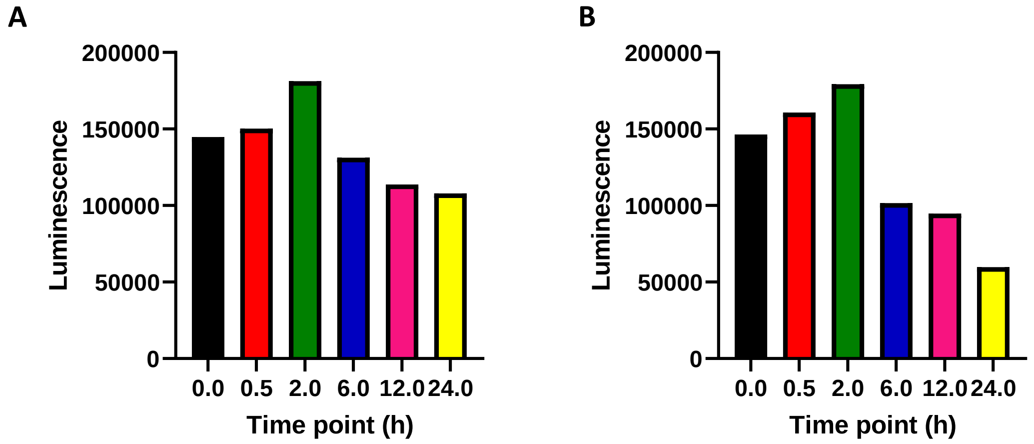


Figure 5.12: Stability of A) *Pv*PI4K-WT and B) *Pv*PI4K-C1327A over 24 hour incubation period.

Luminescence units represent mean enzyme activity for the no inhibitor (DMSO only) controls which show 100% enzyme activity.

The TCI time-course inhibition data (Figure 5.13 and Figure 5.14) showed no time-dependent shift in IC_{50} for MMV1793498 (derivative lacking a warhead) in the wildtype enzyme or the C1327A mutant assay as expected. A clear decrease in IC_{50} was observed for MMV1792459, MMV1848400 and MMV1902872 with increase in pre-incubation time for the wildtype enzyme but no decrease was observed for the C1327A mutant. These observations support C1327-dependent covalent inhibition. MMV1848400 showed atypical IC_{50} curves (Figure 5.13C) with shallow Hill slopes (0.4-0.7) while the other TCI compounds had Hill slopes ranging from 0.8 to 1.4. The reason for this discrepancy is unclear.

The rate of covalent bond formation – determined by the time point where a significant decrease in IC_{50} was observed over the 24 hr pre-incubation period (Figure 5.14) – was comparable for the three TCI tool compounds. Comparing the matched pairs MMV1792459 and MMV1848400, the data suggests the chloroacetamide and acrylamide warheads have similar reactivity. Additionally, changing the ATP-competitive region of the TCI (MMV1902872) did not significantly affect potency when compared to its chloroacetamide warhead counterpart MMV1848400.

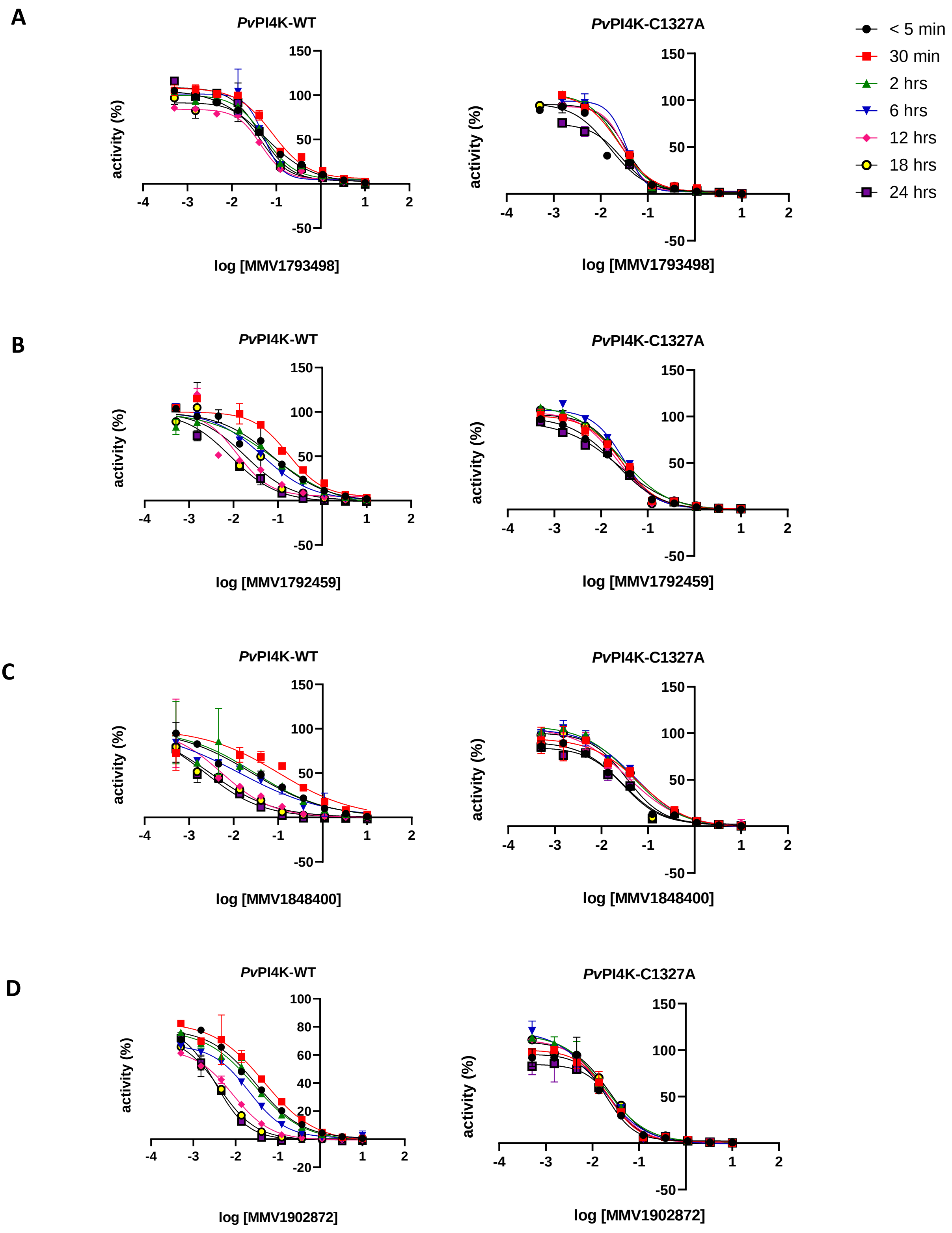


Figure 5.13: Time-dependent IC$_{50}$ plots for *Pv*PI4K-WT and *Pv*PI4K-C1327 for the TCI tool compounds.

A) MMV1793498, **B**) MMV1792459, **C**) MMV1848400 and **D**) MMV1902872. Time points represent the enzyme-inhibitor pre-incubation period. Error bars represent standard deviation for the range of data. Top and bottom of the curves constrained to zero and 100 % where necessary.

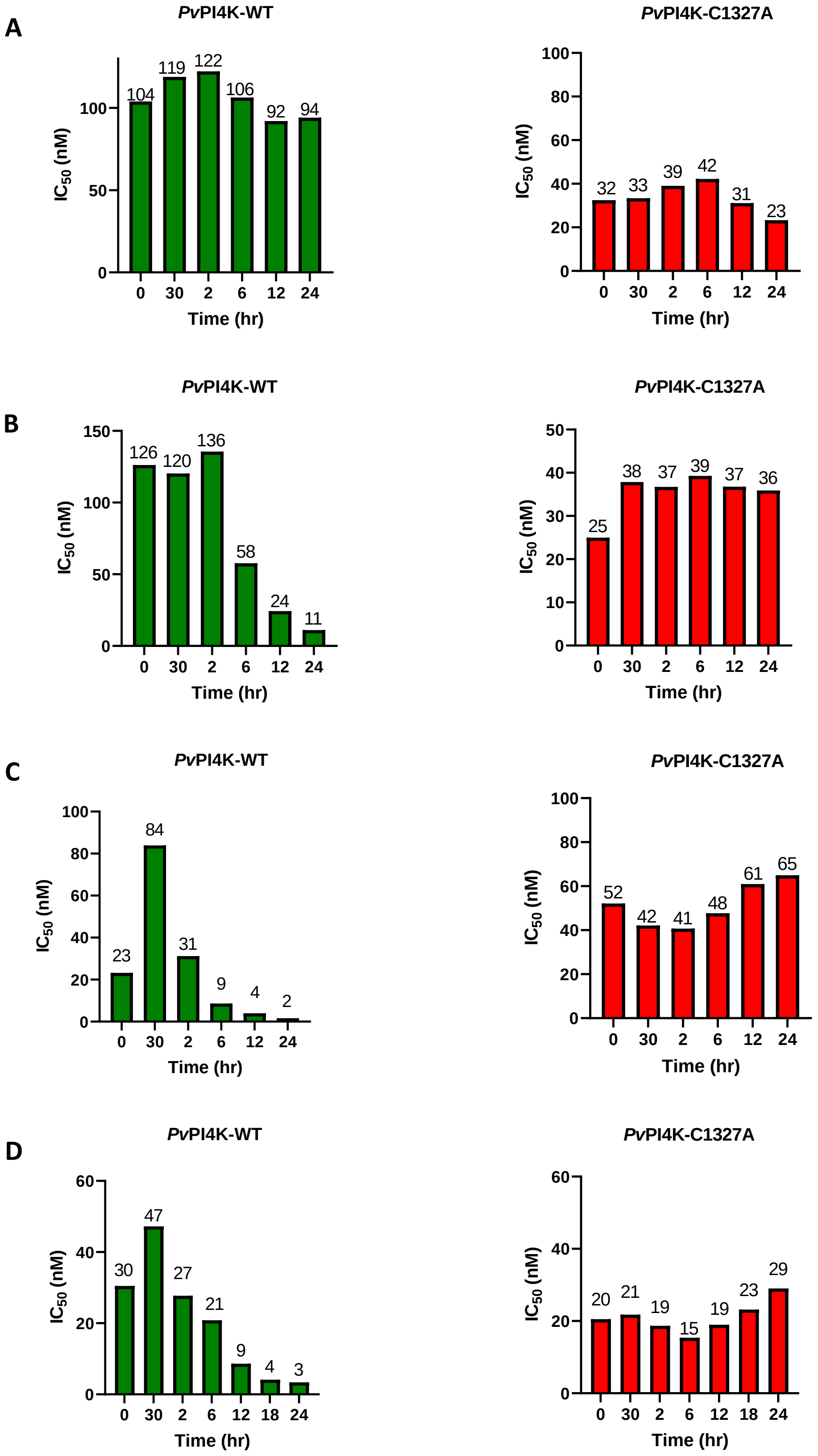


Figure 5.14: Bar graphs showing IC_{50} change over 24 hours for *Pv*PI4K-WT (green) and *Pv*PI4K-C1327 (red) using the TCI tool compounds.

A) MMV1793498, **B**) MMV1792459, **C**) MMV1848400 and **D**) MMV1902872.

To confirm covalent bond formation at C1327, 5 µg recombinant *Pv*PI4K was incubated with excess MMV1792459, MMV1848400 and MMV1902872 for 5 min, 10 min,15 min 30 min and 2 hr. The reactions were quenched by the addition of IAA to block any unmodified cysteine residues by forming S-carboxyamidomethyl-cysteine and subsequent addition of 0.5 % formic acid. The pH in the samples was increased followed by digestion with trypsin and chymotrypsin prior to mass spectrometry.

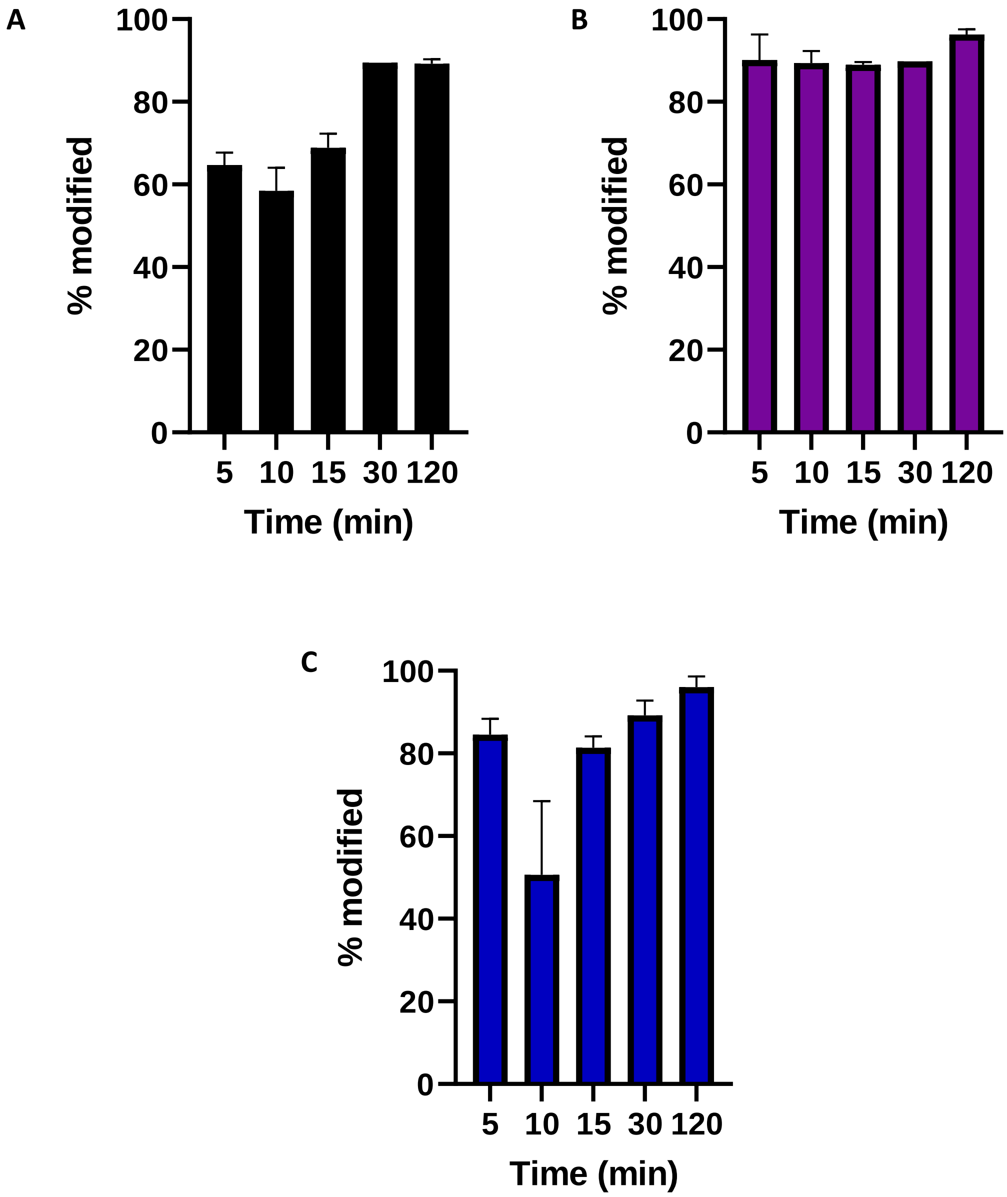


Figure 5.15: Modification of recombinant *Pv*PI4K C1327 by tool TCI compounds.
A) MMV1792459, **B**) MMV1848400 and **C**) MMV1902872. Time points represent the enzyme-inhibitor pre-incubation period. Error bars represent standard deviation for the range of data. % modification indicates the proportion of C1327 present in the sample with covalently bound inhibitor compared to 'control' carbamidimethylated peptide.

Mass spectrometry showed that 90.3 % of the total signal acquired corresponded to *Pv*PI4K suggesting 90% purity of the recombinant *Pv*PI4K. The combination of trypsin and chymotrypsin digestion resulted in 234 peptides which equates to 74.3 % unique *Pv*PI4K sequence coverage. MMV1792459 resulted in 64 - 69% C1327 modification at the earlier time points (5-15 mins) reaching a maximum of ~90 % by 30 minutes (Figure 5.15A). In contrast, MMV1902872 (Figure 5.15C) resulted in almost complete modification (~90 %) at the earliest time point. This data suggests the chloroacetamide warhead forms a covalent bond with the target residue in *Pv*PI4K faster than the acrylamide warhead. A dip in modification and large variation between technical replicates was observed for the 10 min time point in the MMV1902872 which is likely due to experimental error in the sample preparation stage.

Modification of other cysteine residues by inhibitors was also assessed at each time point. For MMV1792459 and MMV1902872, only one cysteine – C1327- was modified in *Pv*PI4K, indicating selectivity of these compounds for the target cysteine although 5 and 3 modifications in MMV1902872 and MMV1792459 respectively were also detected in insect cell contaminant proteins and BSA (Appendix G), but with very low confidence (PEP scores in the 0.018 - 0.031 range in comparison to the high confidence score of the target cysteine 7.76×10^{-32}), suggesting that these are likely false positives although this needs to be confirmed. Six cysteine residues in *Pv*PI4K (C1327 and 5 off-target residues) were modified by MMV1848400 suggesting this compound is less selective (more reactive) although the off-target modifications observed occurred at a significantly lower intensity compared to the C1327 modification (Table 5.5). Additionally, similar to MMV1902872, 5 modifications were also detected in insect cell contaminant proteins and BSA with low confidence (Appendix G).

Table 5.5: Table showing modified cysteine residues in recombinant *Pv*PI4K by MMV1848400.

Cysteine position in *Pv*PI4K	**Sequence**	**Intensity (%)**	**Conserved in *Pf*PI4K**
694	FLPEFNCTF	3.9	No
781	ECLSPPQSGQR	1.2	No
1327	VNDTCSVDSLK	89.8	Yes
1469	IPCFANGTQF	2.8	No
1477	CIDSLKER	0.11	Yes
1494	AVDVCIQR	0.70	Yes

Intensity % -ratio of modified peptide signal intensity vs unmodified peptide signal intensity. Intensity values are summary of all time points.

Comparison of the modified sites in *Pv*PI4K to the *Pf*PI4K homology model showed that, of the 5 cysteine residues observed with non-specific modifications in *Pv*PI4K, 3 residues were conserved in *Pf*PI4K (C1469 C1477 and C1494) while the remaining 2 were not conserved and C694 and C781 are not in the kinase domain. Notably, the 3 conserved residues were not near the ATP binding site of the kinase domain (Figure 5.16).

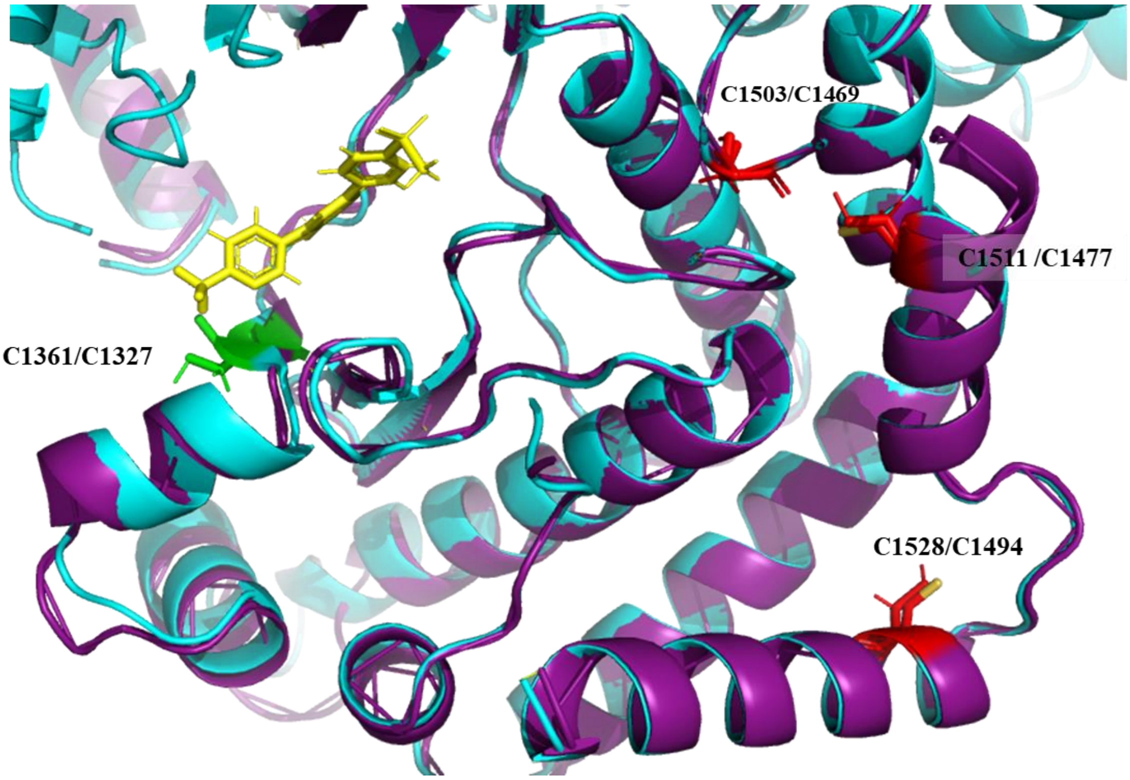


Figure 5.16: Modified conserved cysteine residues in *Pf*PI4K and *Pv*PI4K.
Aligned *Pf*PI4K(cyan) and *Pv*PI4K(purple) homology models. Green residues highlight targeted cysteine. Red residues highlight cysteine residues with non-specific modifications. Labelling-*Pf*PI4K position/*Pv*PI4K position. *Pv*PI4K model generated by SWISS-MODEL. Docked compound (yellow): MMV390048.

The data shows successful modification of the target cysteine with the tool TCIs confirming covalent bond formation between the compounds and C1327. Furthermore, the data suggests MMV1792459 and MMV1902872 are site specific while MMV1848400 is not. The results also indicate that the chloroacetamide warhead is more reactive than the acrylamide warhead as shown when comparing matched pairs MMV1848400 and MMV1792459. This result is not surprising and was expected due to the more electrophilic nature of the methylene carbon attached to the chloro group relative to the terminal olefin carbon. It is possible the high reactivity of the chloroacetamide warhead is the reason for the increased non-specificity of the compound. The increased specificity observed for MMV1902872 suggests that altering the ATP-competitive binding region of the inhibitor decreased the number of off-target cysteine residues modified. While the mass spectrometry data confirms covalent modification of the target cysteine, it is important to consider off-target modifications and their potential adverse effects. In addition, given that the targeted cysteine residue is not essential for PI4Kβ function based on activity data using the mutated enzyme there is the possibility that the parasite could mutate this residue to evade covalent PI4Kβ modification by the drug. The likelihood of this will be assessed in future in vitro resistance studies.

Evaluating aminopyridine/pyrazine (SFK40) series TCI analogues

TCI compounds designed were designed (Ferdinand Ndubi, UCT) based on the addition of a warhead to *aminopyridine/pyrazine (SFK40) series* analogues were evaluated using the TCI time-course inhibition assay. An initial 3-point time-course assay was performed on the analogues with the warhead attached in the para position: H3D-016630 (competitive control), H3D-018393 (acrylamide warhead) and H3D-018391 (chloroacetamide warhead). Time-course inhibition assays were performed as described in section 5.2.3 at 5 min, 30 min and 2 hrs using *Pv*PI4K-WT and *Pv*PI4K-C1327A (Figure 5.17 and Table 5.6).

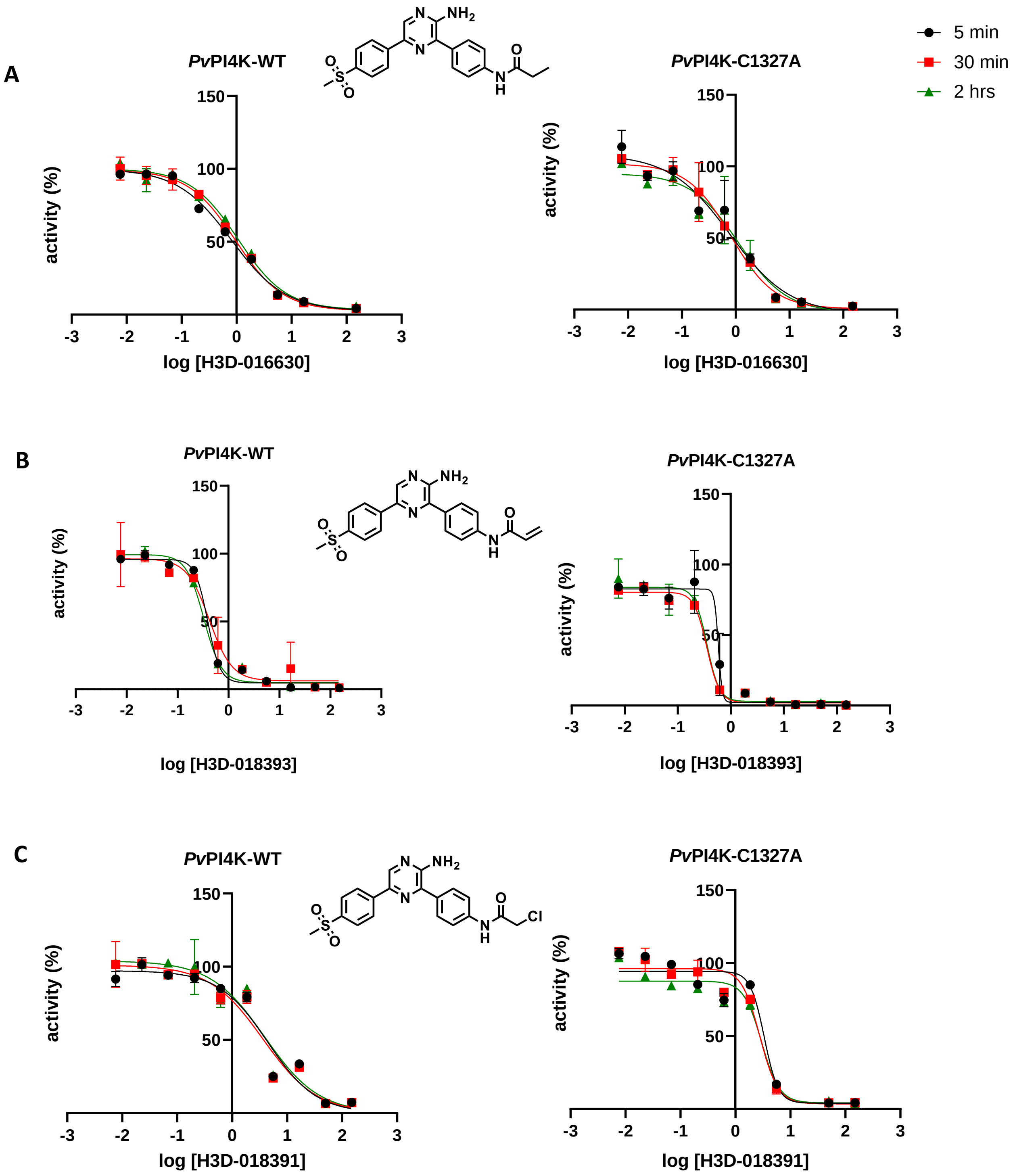


Figure 5.17: Time-dependent IC$_{50}$ for *Pv*PI4K-WT and *Pv*PI4K-C1327A for VE-821 para position reverse amide analogues.

Time points represent *Pv*PI4K-WT/*Pv*PI4K-C1327A enzyme-inhibitor pre-incubation prior to the addition of ATP and phosphatidylinositol substrate. **A**) H3D-016630, **B**) H3D-018393 and **C**) H3D-018391. Error bars represent standard deviation for the range of data.

Table 5.6: Table showing time-dependent IC_{50} for *Pv*PI4K-WT and *Pv*PI4K-C1327A for VE-821 para position reverse amide analogues.

Compound		**IC_{50} at 5 min**	**IC_{50} at 30 min**	**IC_{50} at 2h**
H3D-016630	*Pv*PI4K-WT	~ 1 nM		
	*Pv*PI4K-C1327A	~ 1 nM		
H3D-018393	*Pv*PI4K-WT	~ 1 nM		
	*Pv*PI4K-C1327A	~ 1 nM		
H3D-018391	*Pv*PI4K-WT	4.4 nM	3.6 nM	3.9 nM
	*Pv*PI4K-C1327A	3.3 nM	2.8 nM	2.9 nM

The time-course inhibition data shows potent inhibition of *Pv*PI4K by the TCI analogues with both the competitive and acrylamide warhead containing inhibitors displaying IC_{50} values of 1nM. There was no observed difference in the IC_{50} values with increased enzyme-inhibitor pre-incubation time for the *Pv*PI4K-WT or *Pv*PI4K-C1327A. IC_{50} values below 2 nM cannot be accurately measured due to the assay wall, thus covalent bond formation between these compounds and *Pv*PI4K-WT cannot be shown for very potent inhibitors using this method.

The TCI analogues in which the warhead is in the meta position, H3D-018390 (competitive control), H3D-018394 (acrylamide warhead) and H3D-018392 (chloroacetamide warhead), were also evaluated using the TCI time-course inhibition assay (Figure 5.18 and Figure 5.19).

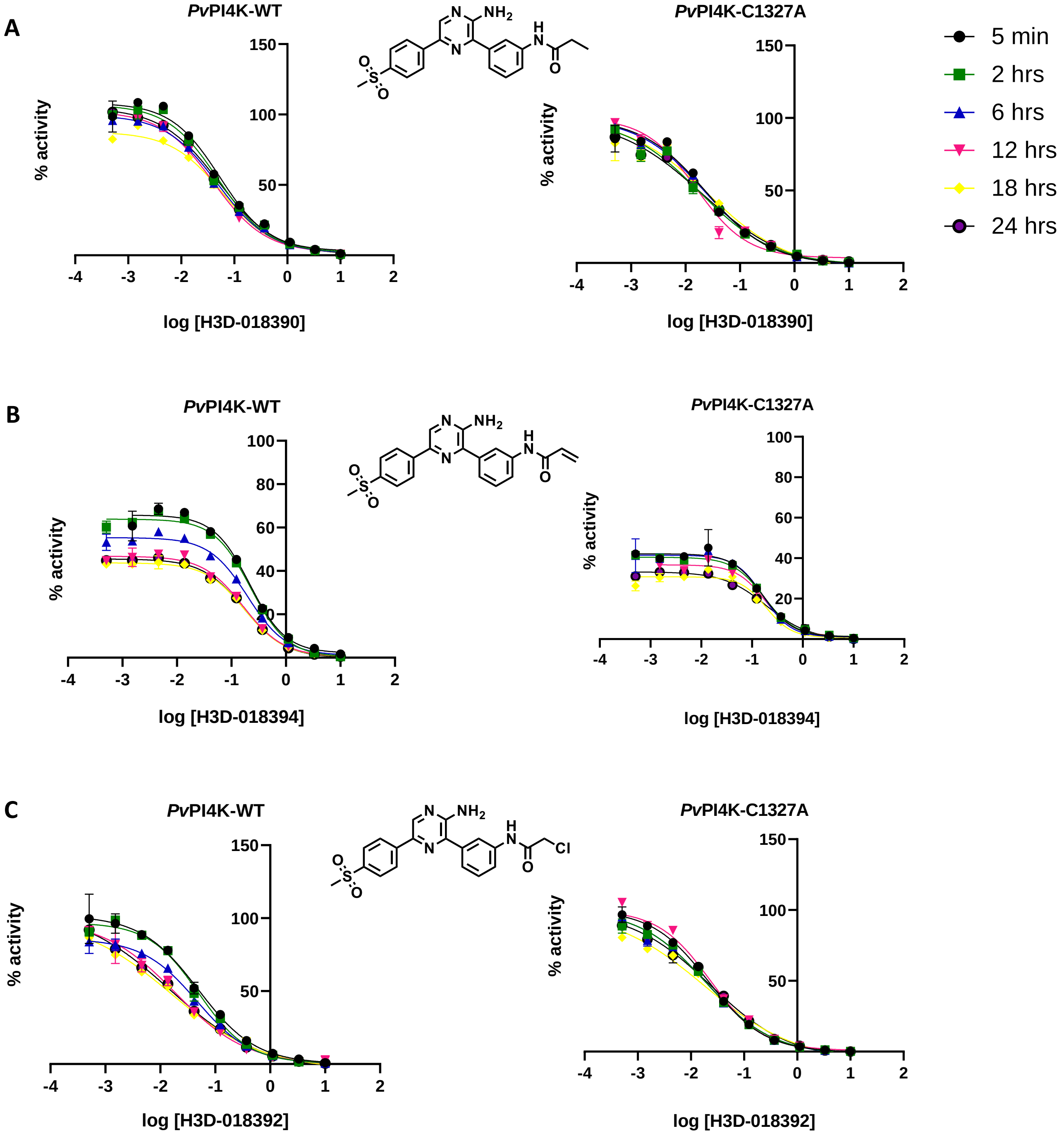


Figure 5.18: Time-dependent IC50 for *Pv*PI4K-WT and *Pv*PI4K-C1327A for VE-821 meta position reverse amide analogues.

Time points represent *Pv*PI4K-WT/*Pv*PI4K-C1327A enzyme-inhibitor pre-incubation prior to the addition of ATP and phosphatidylinositol substrate. **A**) H3D-018390, **B**) H3D-018394 and **C**) H3D-018392. Error bars represent standard deviation for the range of data.

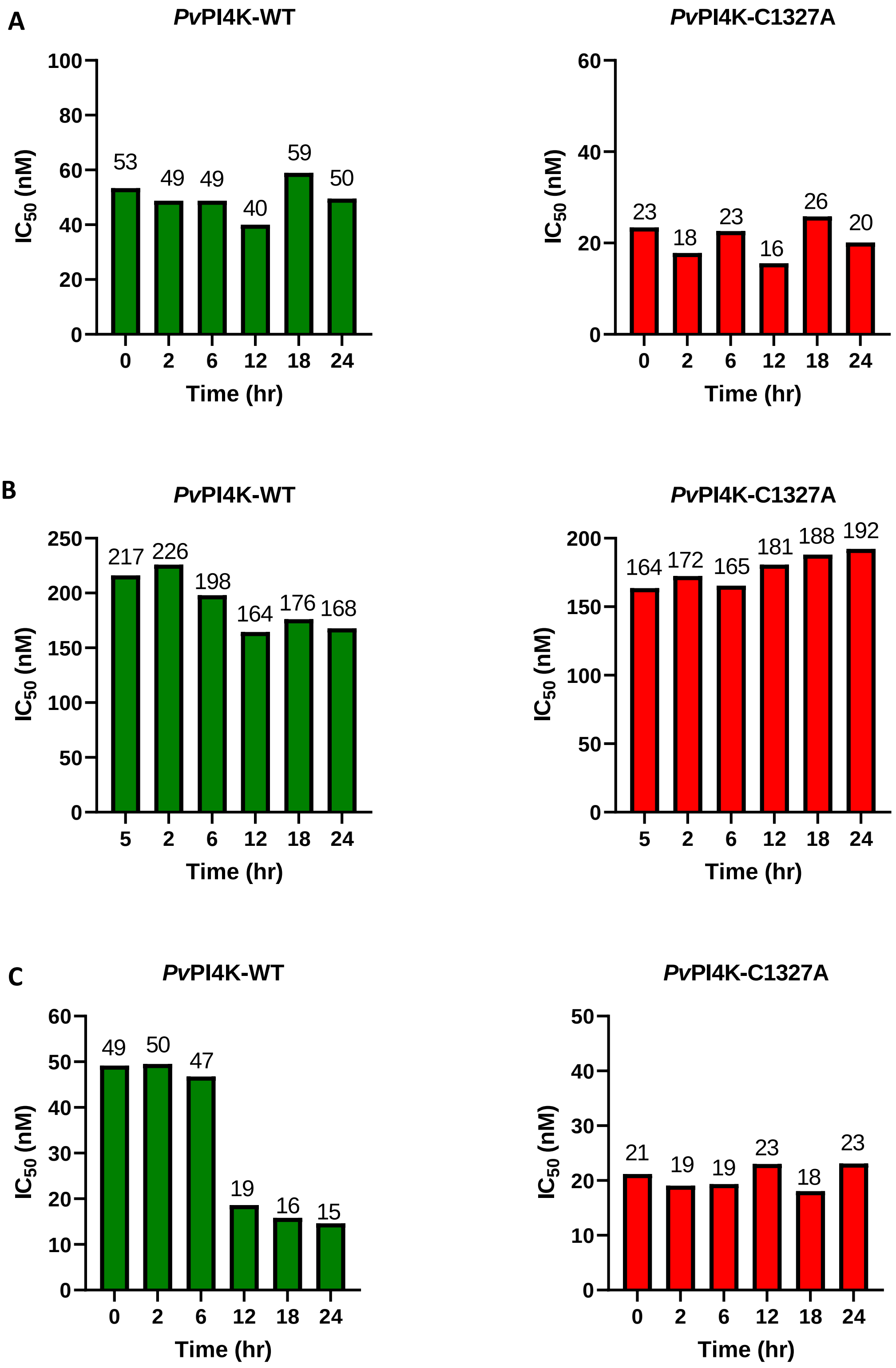


Figure 5.19: Bar graphs showing IC_{50} change over 24 hours for *Pv*PI4K-WT (green) and *Pv*PI4K-C1327 (red) using the VE-821 meta position reverse amide analogues.
Time points represent *Pv*PI4K-WT/*Pv*PI4K-C1327A enzyme-inhibitor pre-incubation prior to the addition of ATP and phosphatidylinositol substrate. **A**) H3D-018390, **B**) H3D-018394 and **C**) H3D-018392.

No change in the IC_{50} was observed over the 24 hr pre-incubation period for the competitive control (H3D-18390) as expected while a decrease in IC_{50} was observed for the chloroacetamide warhead analogue (H3D-18392) starting at 12 hrs – slightly later than was observed for the chloroacetamide tool compound (MMV1848400) but a decrease at the same time point as was seen for the second chloroacetamide tool compound (MMV1902872). The data for the acrylamide warhead analogue (H3D-18394) suggests covalent bond formation between the compound and *Pv*PI4K-WT did not occur as the IC_{50} values remained relatively similar over the 24 hr pre-incubation period. Additionally, the acrylamide warhead-containing analogue was observed to be approximately 3 times less potent than the competitive control, suggesting that the presence of the acrylamide group in this position may affect the binding of the compound in the enzyme active site resulting in decreased inhibition. Mass spectrometry experiments for the analogues, H3D-018391, H3D-018392, H3D-018393 and H3D-018394 are ongoing.

Overall, successful inhibition of *Pv*PI4K-WT by C1327-targeting TCIs was observed. The data suggests that covalent bond formation is more rapid for the chloroacetamide analogues than the acrylamide warhead analogues and furthermore, the binding of the ATP-competitive region of the inhibitor in the enzyme active site was an important factor in the TCI inhibition activity observed and possibly affected TCI specificity.

5.4 Conclusion

In summary, miniaturisation of the ADP-Glo™ kinase assay was successful and evaluation of known potent aminopyridine/pyrazine (SFK40) series compounds showed the ‘assay wall’ is a limiting factor for potent PI4K inhibitors. The high K_m^{ATP} of PI4K results in a small shift which can be difficult to observe in the ATP shift assay when evaluating potent inhibitors.

A *Pv*PI4K-F832A assay was set up to complement the *Pv*PI4K-WT assay in evaluating competitive inhibition, particularly in compounds that form Pi-Pi stacking and hydrogen bonds with F832 in the enzyme active site. However, the assay is limited particularly when evaluating very potent compounds. Additionally, mutation of F832 to alanine had a minimal effect on competitive inhibition of *Pv*PI4K suggesting other factors must be responsible for the selectivity of MMV390048 and other SFK40 analogues. Site directed mutagenesis of other residues such as Y1356 and S1365 could be used to confirm this.

A TCI time-course assay was successfully established that can be used to monitor C1327-dependent covalent inhibition with time. However, it is difficult to assess potent TCIs using this strategy. The mass spectrometry experiments not only allow for confirmation of covalent bond formation regardless of compound potency and binding speed, but also provide further insights such as possible off-targets within the enzyme or other proteins in the reaction as well as the levels of modified enzyme compared to unmodified enzyme. The TCI time-course assay and mass spectrometry data in this study showed the acrylamide and chloroacetamide warheads successfully bound to C1327 and are good candidates for use in designing C1327 targeting TCIs. The chloroacetamide warhead was shown to be highly reactive compared to the acrylamide warhead resulting in increased non-specificity. However, changing the ATP-competitive binding region of the inhibitor was shown to reduce this non-specificity. The TCI-time course assay and mass spectrometry can be used in combination to evaluate TCIs to confirm covalent bond formation, binding speed as well as compound specificity for the target residue hence the data obtained can be used inform further SAR studies.

Chapter 6

Conclusions and Future work

Section A - *Plasmodium* Haem Detoxification Protein

The aim of this study was to optimise expression and purification of soluble *Pf*HDP using an *E.coli* expression system and characterise the protein function and structure. Recombinant expression of *Pf*HDP was evaluated in ArcticExpress cells, C41 cells and BL21 cells. Expression of soluble *Pf*HDP was successfully optimised in the ArcticExpress system using TB media supplemented with 1 % glucose and 5-ALA however, separation of *Pf*HDP from chaperone proteins co-expressed to help protein folding at low temperatures – Cpn60 and Cpn10 – was challenging and purification was unsuccessful. Soluble *Pf*HDP expressed in the C41 and ArcticExpress cells did not bind to the Ni^{2+}-NTA chromatography column while *Pf*HDP purified under denaturing conditions did bind to the Ni^{2+}-NTA chromatography column. This suggests that the N-terminal His-tag was masked by a binding partner or buried in the native protein preparation. Another potential explanation for the lack of binding to the column is that the his-tag has been cleaved. However, since the his-tag is observed in the denatured protein (by purification using a Ni-NTA column) and detection of *Pf*HDP by western blot, this appears to be the less likely explanation unless cleavage happens during the purification process.

Refolded HDP purified under denaturing conditions was evaluated using the β-hematin formation assay. The assay results showed β-hematin formation was mediated by SDS used in the assay not HDP, supporting extensive studies by Roxanne Mohunlal (PhD Thesis, UCT 2019). In order to determine what the role of HDP may be, structure determination studies were performed. Based on PSIPRED analysis of the *Pf*HDP sequence, the N-terminus residues 1-35 were predicted to be highly disordered. HDP crystals obtained by crystallography were small and did not provide adequate diffraction data for structure determination thus *in silico* modelling was used to predict the *Pf*HDP structure.

Three *in silico* structure prediction methods were used to generate *Pf*HDP structures: RoseTTAFold, I-TASSER and AlphaFold. The structures predicted were similar (RMSD of backbone atoms = 2.18-2.45) with the largest difference between the structures observed at the highly disordered N-terminus region. Further evaluation of the models showed the four histidine residues in HDP predicted to facilitate haemozoin formation in the parasite vacuole

(His122, His175, His172 and His197) are not positioned in a way that would allow for haemozoin formation to occur by the previously proposed mechanism.

It was initially hypothesised that protein folding at the disordered N-terminus region of the protein was masking the his-tag in soluble *Pf*HDP preventing its binding to Ni^{2+}-NTA chromatography columns. Alternatively, the loose disordered region may be binding to small proteins in the lysate thus masking the his-tag. To prevent this from occurring, alternative purification tags could be used such as GST (Glutathione S-transferase), which is larger than the his-tag, or a Strep-II tag which is a small tag (8 amino acids) but does not affect protein folding or expression. Thus, these alternative tags are good options for studying functional proteins (Kimple, Brill and Pasker, 2013).

As the aim of this study was to determine the function of *Pf*HDP, full length *Pf*HDP was used despite numerous challenges encountered. Future work could include the truncation of HDP by removing residues 1-35, which may result in a more stable protein, and improved purification and protein crystallisation. Further optimisation of the crystallography experiments could also include using a different crystallisation method and further optimisation of crystallisation conditions i.e. pH, temperature and a wider range of protein:precipitant ratios.

Published knockout studies have shown that HDP is essential for parasite survival. The HDP sequence contains a fasciclin domain. Thus, it is possible HDP may have a similar function as other fasciclin-domain containing proteins e.g. pathogenicity and cell adhesion. However, further biochemical studies in this area are required. Studies on the function of fasciclin domain containing proteins in various species i.e. human/mammalian, bacterial, plant and fungal have predominantly been *in vivo*, often involving fluorescence (for localisation), cell adhesion assays and knockdown studies to evaluate the effect on the predicted function e.g. cell adhesion, cell invasion, plant stem development and root development (Seifert, 2018).

Section B - *Plasmodium* phosphatidylinositol 4-kinase

The aim of this study was to use *Plasmodium* PI4K inhibition assays and mass spectrometry to evaluate ATP-competitive inhibitors and C1327 targeting covalent inhibitors to support PI4K target-based drug discovery research.

Site-directed mutagenesis was used to introduce mutations F832A and C1327A into *Pv*PI4K-WT to probe key protein inhibitor interactions within the ATP-binding site. The specific activity of *Pv*PI4K-WT was significantly higher than that of *Pv*PI4K-F832A whereas the specific activity of *Pv*PI4K-C1327 was comparable to that of *Pv*PI4K-WT. Kinetic characterisation of the enzymes showed that the mutations did not have a significant effect on the K_m^{ATP}. *Pv*PI4K-WT and *Pv*PI4K-C1327 displayed similar catalytic efficiency while *Pv*PI4K-F832A had a significantly lower catalytic efficiency indicating that the mutation had a detrimental effect on the enzyme stability and/or catalytic activity. Despite the difference in the specific activity of the enzymes, the similar K_m^{ATP} for *Pv*PI4K-WT, *Pv*PI4K-F832A and *Pv*PI4K-C1327 indicated inhibition data for these enzymes was comparable.

Miniaturised *in vitro* kinase inhibition assays using *Pv*PI4K-WT and *Pv*PI4K-F832A enzymes were used to evaluate ATP-competitive inhibitors predicted to form Pi-Pi stacking interactions and hydrogen bonds with the F832 residue in the enzyme active site. The data showed mutation of the F832 residue had minimal effect on *Pv*PI4K inhibition suggesting that other unique residues (or a combination of unique residues) within the ATP-binding site are responsible for the potency and selectivity of MMV390048 and other related SFK40 series compounds, relative to the human orthologue. However, the assays cannot predict IC_{50} values lower than 1 nM due to the assay wall effect, making accurate IC_{50} fold changes difficult to determine for very potent compounds.

Future site-directed mutagenesis studies to probe the basis for the potency and selectivity of *Plasmodium* PI4K could include mutating other ATP-site residues unique to *Plasmodium* PI4K, e.g. Y1356 (Y1342 in *Pv*PI4K) and S1365 (S1331 in *Pv*PI4K). Y1356 is another residue predicted to affect PI4K inhibitor potency and selectivity by forming Pi-Pi stacking with aromatic compounds and S1365 is predicted as the probable origin of selectivity over human phosphoinositide kinases (Figure 4.2). *Plasmodium falciparum* PI4K Y1356 adds aromaticity to the hinge region and increases the potency of multiaromatic compounds due to Pi-Pi stacking interactions with the inhibitor compared to lipophilic interactions with P597 in HuPI4K (Fienberg *et al*., 2020). The analogous position for S1365 is Q606 in HuPI4K – a large neutral polar residue which tends to clash with appropriately positioned large hydrophobic substituents in *Pf*PI4K inhibitors exemplified by the CF3-pyridyl in MMV390048 while S1365 tends to form hydrogen bonds hence the selectivity over human PI4K (Fienberg *et al*., 2020). Additionally, the S1365 analogous position in HuPI3K varies considerably with glutamine in HuPI3Kα (Q859), aspartic acid in HuPI3Kβ (D856), lysine in HuPI3Kγ (K890) and asparagine in HuPI3Kδ (N839) (Fienberg *et al*., 2020). As many of these amino acid residues are large, compared to serine, the likelihood of inhibitor off-target effects is decreased due to potential steric clashes between the inhibitors and the amino acids. It is likely that interactions between inhibitors and more than one unique residue are responsible for selectivity and potency e.g. Pi-Pi stacking interactions between an inhibitor and both F827 and Y1356. Thus, mutating the unique residues in combination would be beneficial. The unique residues i.e. F827, Y1356 and S1365 could also be mutated to the residues found in human PI4K i.e. Y1356 to proline, S1365 to glutamine, and F827 to leucine as this could provide useful data for compound selectivity over human PI4K to guide SAR. Additionally, all the unique residues could be simultaneously mutated to the human counterparts, creating a chimeric protein which could provide valuable insights into inhibitor selectivity over human PI4K.

The TCI enzyme assay was used to monitor C1327-dependent covalent bond formation. Mass spectrometry was incorporated to confirm covalent bond formation between the TCI and the target cysteine residue. Furthermore, mass spectrometry was able to detect modification of any other cysteine residues in *Pv*PI4K i.e. nonspecific modifications. Taken together, the TCI enzyme assay and mass spectrometry confirmed covalent binding of chloroacetamide and acrylamide warheads to *Pv*PI4K C1327, demonstrating proof-of-principle and confirming the modelling predictions. As expected, the chloroacetamide warhead was found to be more reactive compared to the acrylamide warhead and as a result, more off-target modifications were observed. However, this non-specificity could be mediated by changing the ATP-competitive binding region of the compound. Similar to the *Pv*PI4K-F832A assay, the TCI enzyme assay was limited when evaluating potent TCIs as the assay wall effect is observed at the earliest time point, making it unfortunately impossible to observe covalent inhibition over time. Fortunately, covalent bond formation for potent TCIs can be evaluated using the mass spectrometry experiments.

The next steps will be to carry out mass spectrometry experiments using lower inhibitor concentrations (at 1 × IC_{50}) and enzyme concentrations to distinguish more clearly between the rate of covalent bond formation for potent inhibitors. Additionally, the mass spectrometry experiments can be further adapted for high-throughput analysis using 96-well plates and liquid dispensers for faster and increased data output.

Lastly, another potential area of interest is TCIs targeting an alternative *Plasmodium* PI4K residue – the catalytic lysine K1308 (K1274 in *Pv*PI4K) – a strategy which has been employed for targeting human kinases such as HuPI3Kδ (Dalton *et al*., 2018). This residue is highly conserved and is essential for kinase activity thus there is a reduced risk of resistance mutations arising at this position. However, the residue is also conserved across the entire kinase superfamily thus it would be important to have highly selective inhibitors as starting points to avoid off-target effects.

References

Aldred, E., 2009. *Pharmacology*. Edinburgh: Churchill Livingstone/Elsevier.

Aljoundi, A., Bjij, I., El Rashedy, A. and Soliman, M., 2020. Covalent Versus Non-covalent Enzyme Inhibition: Which Route Should We Take? A Justification of the Good and Bad from Molecular Modelling Perspective. *Protein J.*, 39(2), pp.97-105.

Anderson, D., Harris, R., Polayes, D., Ciccarone, V., Donahue, R., Gerard, G. and Jessee, J., 1995. Rapid generation of recombinant baculovirus and expression of foreign genes using the BAC-to-BAC Baculovirus expression system. *Focus*, 17(2), pp.53-58.

Arendse, L., Wyllie, S., Chibale, K. and Gilbert, I., 2021. *Plasmodium* Kinases as Potential Drug Targets for Malaria: Challenges and Opportunities. *ACS Infect. Dis.*, 7(3), pp.518-534.

Ariey, F., Witkowski, B., Amaratunga, C., Beghain, J., Langlois, A., Khim, N., Kim, S., Duru, V., Bouchier, C., Ma, L., Lim, P., Leang, R., Duong, S., Sreng, S., Suon, S., Chuor, C., Bout, D., Ménard, S., Rogers, W., Genton, B., Fandeur, T., Miotto, O., Ringwald, P., Le Bras, J., Berry, A., Barale, J., Fairhurst, R., Benoit-Vical, F., Mercereau-Puijalon, O. and Ménard, D., 2013. A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature*, 505(7481), pp.50-55.

Ashley, E. and Phyo, A., 2018. Drugs in Development for Malaria. *Drugs*, 78(9), pp.861-879.

Ashley, E., Pyae Phyo, A. and Woodrow, C., 2018. Malaria. *Lancet*, 391(10130), pp.1608-1621.

Avery, V., Bashyam, S., Burrows, J., Duffy, S., Papadatos, G., Puthukkuti, S., Sambandan, Y., Singh, S., Spangenberg, T., Waterson, D. and Willis, P., 2014. Screening and hit evaluation of a chemical library against blood-stage *Plasmodium falciparum*. *Malar. J.*, 13(190).

Ayala-Aguilera, C., Valero, T., Lorente-Macías, Á., Baillache, D., Croke, S. and Unciti-Broceta, A., 2021. Small Molecule Kinase Inhibitor Drugs (1995–2021): Medical Indication, Pharmacology, and Synthesis. *J. Med. Chem.*, 65(2), pp.1047-1131.

Baca, A. and Hol, W., 2000. Overcoming codon bias: A method for high-level overexpression of *Plasmodium* and other AT-rich parasite genes in Escherichia coli. *Int. J. Parasitol.*, 30(2), pp.113-118.

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G., Wang, J., Cong, Q., Kinch, L., Schaeffer, R., Millán, C., Park, H., Adams, C., Glassman, C., DeGiovanni, A., Pereira, J., Rodrigues, A., van Dijk, A., Ebrecht, A., Opperman, D., Sagmeister, T., Buhlheller, C., Pavkov-Keller, T., Rathinaswamy, M., Dalwadi, U., Yip, C., Burke, J., Garcia, K., Grishin, N., Adams, P., Read, R. and Baker, D., 2021. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557), pp.871-876.

Baillie, T., 2016. Targeted Covalent Inhibitors for Drug Design. *Angew. Chem. Int. Ed.*, 55(43), pp.13408-13421.

Baker, D. and Sali, A., 2001. Protein Structure Prediction and Structural Genomics. *Science*, 294(5540), pp.93-96.

Balikagala, B., Fukuda, N., Ikeda, M., Katuro, O., Tachibana, S., Yamauchi, M., Opio, W., Emoto, S., Anywar, D., Kimura, E., Palacpac, N., Odongo-Aginya, E., Ogwang, M., Horii, T. and Mita, T., 2021. Evidence of Artemisinin-Resistant Malaria in Africa. *N. Engl. J. Med.*, 385(13), pp.1163-1171.

Balla, T., 2013. Phosphoinositides: Tiny Lipids With Giant Impact on Cell Regulation. *Physiol. Rev.*, 93(3), pp.1019-1137.

Baskin, J., Wu, X., Christiano, R., Oh, M., Schauder, C., Gazzerro, E., Messa, M., Baldassari, S., Assereto, S., Biancheri, R., Zara, F., Minetti, C., Raimondi, A., Simons, M., Walther, T., Reinisch, K. and De Camilli, P., 2015. The leukodystrophy protein FAM126A (hyccin) regulates PtdIns(4)P synthesis at the plasma membrane. *Nat. Cell Bio.*, 18(1), pp.132-138.

Belete, T., 2020. Recent Progress in the Development of New Antimalarial Drugs with Novel Targets. *Drug Des. Devel. Ther.*, Volume 14, pp.3875-3889.

Bernhardt, T., Roof, W. and Young, R., 2002. The Escherichia coli FKBP-type PPIase SlyD is required for the stabilization of the E lysis protein of bacteriophage phiX174. *Mol. Microbiol.*, 45(1), pp.99-108.

Bhatwa, A., Wang, W., Hassan, Y., Abraham, N., Li, X. and Zhou, T., 2021. Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in Escherichia coli and Strategies to Address Them for Industrial Applications. *Front. Bioeng. and Biotechnol.*, 9.

Bhullar, K., Lagarón, N., McGowan, E., Parmar, I., Jha, A., Hubbard, B. and Rupasinghe, H., 2018. Kinase-targeted cancer therapies: progress, challenges and future directions. *Mol. Cancer*, 17(1).

Biamonte, M., Wanner, J. and Le Roch, K., 2013. Recent Advances in Malaria Drug Discovery. *Bioorg. Med. Chem. Lett.*, 23(10), pp.2849-2843.

Bornhorst, B. and Falke, J., 2000. Purification of Proteins Using Polyhistidine Affinity Tags. *Methods Enzymol.*, (326), pp.245-254.

Boura, E. and Nencka, R., 2015. Phosphatidylinositol4-kinases: Function, structure, and inhibition. *Exp. Cell. Res.*, 337, pp.136-145.

Braun, P. and LaBaer, J., 2003. High throughput protein production for functional proteomics. *Trends Biotechnol.*, 21(9), pp.383-388.

Bray, P., Janneh, O., Raynes, K., Mungthin, M., Ginsburg, H. and Ward, S., 1999. Cellular Uptake of Chloroquine Is Dependent on Binding to Ferriprotoporphyrin IX and Is Independent of NHE Activity in Plasmodium falciparum. *J. Cell. Biol.*, 145(2), pp.363-376.

Brunschwig, C., Lawrence, N., Taylor, D., Abay, E., Njoroge, M., Basarab, G., Le Manach, C., Paquet, T., Cabrera, D., Nchinda, A., de Kock, C., Wiesner, L., Denti, P., Waterson, D., Blasco, B., Leroy, D., Witty, M., Donini, C., Duffy, J., Wittlin, S., White, K., Charman, S., Jiménez-Díaz, M., Angulo-Barturen, I., Herreros, E., Gamo, F., Rochford, R., Mancama, D., Coetzer, T., van der Watt, M., Reader, J., Birkholtz, L., Marsh, K., Solapure, S., Burke, J., McPhail, J., Vanaerschot, M., Fidock, D., Fish, P., Siegl, P., Smith, D., Wirjanata, G., Noviyanti, R., Price, R., Marfurt, J., Silue, K., Street, L. and Chibale, K., 2018. UCT943, a Next-Generation Plasmodium falciparum PI4K Inhibitor Preclinical Candidate for the Treatment of Malaria. *Antimicrob. Agents Chemother.*, 62(9).

Burke, J., 2018. Structural Basis for Regulation of Phosphoinositide Kinases and Their Involvement in Human Disease. *Mol. Cell*, 71(5), pp.653-673.

Burke, J., Inglis, A., Perisic, O., Masson, G., McLaughlin, S., Rutaganira, F., Shokat, K. and Williams, R., 2014. Structures of PI4KIII complexes show simultaneous recruitment of Rab11 and its effectors. *Science*, 344(6187), pp.1035-1038.

Burrows, J., Duparc, S., Gutteridge, W., Hooft van Huijsduijnen, R., Kaszubska, W., Macintyre, F., Mazzuri, S., Möhrle, J. and Wells, T., 2017. New developments in anti-malarial target candidate and product profiles. *Malar. J.*, 16(1).

Burrows, J., Hooft van Huijsduijnen, R., Möhrle, J., Oeuvray, C. and Wells, T., 2013. Designing the next generation of medicines for malaria control and eradication. *Malar. J.*, 12(1).

Canchola, J., Tang, S., Hemyari, P., Paxinos, E. and Marins, E., 2017. Correct Use of Percent Coefficient of Variation (%CV) Formula for Log-Transformed Data. *MOJ Proteom. Bioinform.*, 6(3), pp.316-317.

Cheuka, P., Lawrence, N., Taylor, D., Wittlin, S. and Chibale, K., 2018. Antiplasmodial imidazopyridazines: structure–activity relationship studies lead to the identification of analogues with improved solubility and hERG profiles. *MedChemComm*, 9(10), pp.1733-1745.

Chugh, M., Sundararaman, V., Kumar, S., Reddy, V., Siddiqui, W., Stuart, K. and Malhotra, P., 2013. Protein complex directs hemoglobin-to-hemozoin formation in Plasmodium falciparum. *PNAS*, 110(14), pp.5392-5397.

Conrad, M. and Rosenthal, P., 2019. Antimalarial drug resistance in Africa: the calm before the storm? *Lancet Infect. Dis.*, 19(10), pp.e338-e351.

Cowell, A. and Winzeler, E., 2019. Advances in omics-based methods to identify novel targets for malaria and other parasitic protozoan infections. *Genome Med.*, 11(63).

Cowell, A. and Winzeler, E., 2019. The genomic architecture of antimalarial drug resistance. *Brief. Funct. Genomics*, 18(5), pp.314-328.

Cowman, A., Healer, J., Marapana, D. and Marsh, K., 2016. Malaria: Biology and Disease. *Cell*, 167(3), pp.610-624.

Cui, L., Mharakurwa, S., Ndiaye, D., Rathod, P. and Rosenthal, P., 2015. Antimalarial Drug Resistance: Literature Review and Activities and Findings of the ICEMR Network. *Am. J. Trop. Med. Hyg.*, 93(3), pp.57-68.

Dalton, S., Dittus, L., Thomas, D., Convery, M., Nunes, J., Bush, J., Evans, J., Werner, T., Bantscheff, M., Murphy, J. and Campos, S., 2018. Selectively Targeting the Kinome-

Conserved Lysine of PI3Kδ as a General Approach to Covalent Kinase Inhibition. *J. Am. Chem. Soc.*, 140(3), pp.932-939.

Daskum, A., Chessed, G., A. Qadeer, M. and Mustapha, T., 2020. Antimalarial Chemotherapy, Mechanism of Action and Resistance to major Antimalarial Drugs in Clinical Use: A Review. *Microbes Infect. Dis.*, 2(1), pp.130-142.

de Villiers, K. and Egan, T., 2021. Heme Detoxification in the Malaria Parasite: A Target for Antimalarial Drug Development. *Acc. Chem. Res.*, 54(11), pp.2649-2659.

Deng, H., Jia, Y. and Zhang, Y., 2018. Protein structure prediction. *Int. J. Mod. Phys. B*, 32(18).

Duffy, P. and Gorres, J., 2020. Malaria vaccines since 2000: progress, priorities, products. *npj Vaccines*, 5(1).

Dumon-Seignovert, L., Cariot, G. and Vuillard, L., 2004. The toxicity of recombinant proteins in *Escherichia coli:* a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression Purif.*, 37(1), pp.203-206.

Egan, T., 2008. Recent advances in understanding the mechanism of hemozoin (malaria pigment) formation. *J. Inorg. Biochem.*, 102(5-6), pp.1288-1299.

Egan, T., Mavuso, W. and Ncokazi, K., 2000. The Mechanism of β-Hematin Formation in Acetate Solution. Parallels between Hemozoin Formation and Biomineralization Processes. *Biochemistry*, 40(1), pp.204-213.

Elliott, D., McIntosh, M., Hosgood, H., Chen, S., Zhang, G., Baevova, P. and Joiner, K., 2008. Four distinct pathways of hemoglobin uptake in the malaria parasite *Plasmodium falciparum*. *PNAS*, 105(7), pp.2463-2468.

Esposito, D. and Chatterjee, D., 2006. Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol.*, 17(4), pp.353-358.

Fabbro, D., 2014. 25 Years of Small Molecular Weight Kinase Inhibitors: Potentials and Limitations. *Mol. Pharmacol.*, 87(5), pp.766-775.

Fidock, D., Nomura, T., Talley, A., Cooper, R., Dzekunov, S., Ferdig, M., Ursos, L., bir Singh Sidhu, A., Naudé, B., Deitsch, K., Su, X., Wootton, J., Roepe, P. and Wellems, T., 2000. Mutations in the P. falciparum Digestive Vacuole Transmembrane Protein PfCRT and Evidence for Their Role in Chloroquine Resistance. *Mol. Cell*, 6(4), pp.861-871.

Fienberg, S., Eyermann, C., Arendse, L., Basarab, G., McPhail, J., Burke, J. and Chibale, K., 2020. Structural Basis for Inhibitor Potency and Selectivity of *Plasmodium falciparum* Phosphatidylinositol 4-Kinase Inhibitors. *ACS Infect. Dis.*, 6(11), pp.3048-3063.

Flannery, E., Chatterjee, A. and Winzeler, E., 2013. Antimalarial drug discovery — approaches and progress towards new medicines. *Nat. Rev. Microbiol.*, 12(11), pp.849-862.

Forte, B., Ottilie, S., Plater, A., Campo, B., Dechering, K., Gamo, F., Goldberg, D., Istvan, E., Lee, M., Lukens, A., McNamara, C., Niles, J., Okombo, J., Pasaje, C., Siegel, M., Wirth, D., Wyllie, S., Fidock, D., Baragaña, B., Winzeler, E. and Gilbert, I., 2021. Prioritization of Molecular Targets for Antimalarial Drug Discovery. *ACS Infect. Dis.*, 7(10), pp.2764-2776.

Furano, A., 1975. Content of elongation factor Tu in Escherichia coli. *PNAS.*, 72(12), pp.4780-4784.

Gehringer, M. and Laufer, S., 2018. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.*, 62(12), pp.5673-5724.

Ghosh, A., Samanta, I., Mondal, A. and Liu, W., 2019. Covalent Inhibition in Drug Discovery. *ChemMedChem*, 14(9), pp.889-906.

Giridhar, G., Manepalli, R. and Apparao, G., 2017. Size-Exclusion Chromatography. *Thermal and Rheological Measurement Techniques for Nanomaterials Characterization*, pp.51-65.

Goldberg, D., 2013. Complex nature of malaria parasite hemoglobin degradation. *PNAS*, 110(14), pp.5283-5284.

Gower, C., Chang, M. and Maly, D., 2014. Bivalent inhibitors of protein kinases. *Crit. Rev. Biochem. Mol. Biol.*, 49(2), pp.102-115.

Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, M.B., Bray, J., Gileadi, O., … Gunsalus, K.C., 2008. Protein production and purification. *Nat Methods*, 5**,** pp.135–146.

Gupta, P., Mehrotra, S., Sharma, A., Chugh, M., Pandey, R., Kaushik, A., Khurana, S., Srivastava, N., Srivastava, T., Deshmukh, A., Panda, A., Aggarwal, P., Bhavesh, N., Bhatnagar, R., Mohmmed, A., Gupta, D. and Malhotra, P., 2017. Exploring Heme and Hemoglobin Binding Regions of *Plasmodium* Heme Detoxification Protein for New Antimalarial Discovery. *J. Med. Chem.*, 60(20), pp.8298-8308.

Hain, A. and Bosch, J., 2013. Autophagy in *Plasmodium*, A multifuctional pathway? *Comput. Struct. Biotechnol. J.*, 8(11), p.e201308002.

Hartinger, D., Heinl, S., Schwartz, H., Grabherr, R., Schatzmayr, G., Haltrich, D. and Moll, W., 2010. Enhancement of solubility in Escherichia coli and purification of an aminotransferase from Sphingopyxis sp. MTA144 for deamination of hydrolyzed fumonisin B1. *Microb. Cell Fact.*, 9(1), pp.62.

Harvey, K., Jarocki, V., Charles, I. and Djordjevic, S., 2019. The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. *Front. Microbiol.*, 10.

Hassett, M. and Roepe, P., 2018. PIK-ing New Malaria Chemotherapy. *Trends Parasitol.*, 34(11), pp.925-927.

Hering, J., Missel, J., Zhang, L., Gunnarsson, A., Castaldo, M., Pedersen, P., Ek, M., Gourdon, P. and Snijder, H., 2020. The rapid “teabag” method for high-end purification of membrane proteins. *Sci. Rep.*, 10(1).

Hovlid, M. and Winzeler, E., 2016. Phenotypic Screens in Antimalarial Drug Discovery. *Trends Parasitol.*, 32(9), pp.697-707.

Hsieh, P. and Vaisvila, R., 2013. Protein Engineering: Single or Multiple Site-Directed Mutagenesis. *Methods Mol. Biol.*, 978, pp.173-186.

Jani, D., Nagarkatti, R., Beatty, W., Angel, R., Slebodnick, C., Andersen, J., Kumar, S. and Rathore, D., 2008. HDP—A Novel Heme Detoxification Protein from the Malaria Parasite. *PLoS Pathog.*, 4(4).

Jones, D. and Cozzetto, D., 2014. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*, 31(6), pp.857-863.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S., Ballard, A., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A., Kavukcuoglu, K., Kohli, P. and Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), pp.583-589.

Kandepedu, N., Gonzàlez Cabrera, D., Eedubilli, S., Taylor, D., Brunschwig, C., Gibhard, L., Njoroge, M., Lawrence, N., Paquet, T., Eyermann, C., Spangenberg, T., Basarab, G., Street, L. and Chibale, K., 2018. Identification, Characterization, and Optimization of 2,8-Disubstituted-1,5-naphthyridines as Novel *Plasmodium falciparum*Phosphatidylinositol-4-kinase Inhibitors with in Vivo Efficacy in a Humanized Mouse Model of Malaria. *J. Med. Chem.*, 61(13), pp.5692-5703.

Kanev, G., de Graaf, C., de Esch, I., Leurs, R., Würdinger, T., Westerman, B. and Kooistra, A., 2019. The Landscape of Atypical and Eukaryotic Protein Kinases. *Trends Pharmacol. Sci.*, 40(11), pp.818-832.

Kimple, M., Brill, A. and Pasker, R., 2013. Overview of Affinity Tags for Protein Purification. *Curr. Protoc. Protein Sci.*, 73(1).

Knapp, S., 2018. New opportunities for kinase drug repurposing and target discovery. *Br. J. Cancer,* 118(7), pp.936-937.

Knight, Z. and Shokat, K., 2005. Features of selective kinase inhibitors. *Chemistry & Biology*, 12(6), pp.621-637.

Krainer, F., Capone, S., Jäger, M., Vogl, T., Gerstmann, M., Glieder, A., Herwig, C. and Spadiut, O., 2015. Optimizing cofactor availability for the production of recombinant heme peroxidase in Pichia pastoris. *Microb. Cell Fact.*, 14(4) (2015), doi:10.1186/s12934-014-0187-z.

Kram, K. and Finkel, S., 2015. Rich Medium Composition Affects Escherichia coli Survival, Glycation, and Mutation Frequency during Long-Term Batch Culture. *Appl. Environ. Microbiol.*, 81(13), pp.4442-4450.

Kulkarni, S., Urbahns, K. and Spangenberg, T., 2020. Targeted Covalent Inhibitors for the Treatment of Malaria? *ACS Infect. Dis.*, 6(11), pp.2815-2817.

Kumar, S., Guha, M., Choubey, V., Maity, P. and Bandyopadhyay, U., 2007. Antimalarial drugs inhibiting hemozoin (β-hematin) formation: A mechanistic update. *Life Sci.*, 80(9), pp.813-828.

Laemmli, U., 1970. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, 227(5259), pp.680-685.

Lee, P., Yeoh, Y. and Low, T., 2022. A recent update on small-molecule kinase inhibitors for targeted cancer therapy and their therapeutic insights from mass spectrometry-based proteomic analysis. *FEBS J.* https://doi.org/10.1111/febs.16442.

Lefevre, F., Rémy, M. and Masson, J., 1997. Alanine-stretch scanning mutagenesis: a simple and efficient method to probe protein structure and function. *Nucleic Acids Res.*, 25(2), pp.447-448.

Liu, E., MacMillan, C., Shafee, T., Ma, Y., Ratcliffe, J., van de Meene, A., Bacic, A., Humphries, J. and Johnson, K., 2020. Fasciclin-Like Arabinogalactan-Protein 16 (FLA16) Is Required for Stem Development in Arabidopsis. *Front. Plant Sci.*, 11.

Liu, J., Zhang, J., Xu, F., Lin, Z., Li, Z. and Liu, H., 2018. Structural characterizations of human periostin dimerization and cysteinylation. *FEBS Lett.*, 592(11), pp.1789-1803.

Mariani, V., Biasini, M., Barbato, A. and Schwede, T., 2013. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*, 29(21), pp.2722-2728.

Martinez, R., Defnet, A. and Shapiro, P., 2020. Avoiding or Co-Opting ATP Inhibition: Overview of Type III, IV, V, and VI Kinase Inhibitors. *Next Generation Kinase Inhibitors*, pp.29-59.

Martinez, R., Defnet, A. and Shapiro, P., 2020. *Next Generation Kinase Inhibitors*. Cham: Springer International Publishing, pp.29-59.

Matz, J., Drepper, B., Blum, T., van Genderen, E., Burrell, A., Martin, P., Stach, T., Collinson, L., Abrahams, J., Matuschewski, K. and Blackman, M., 2020. A lipocalin mediates unidirectional heme biomineralization in malaria parasites. *PNAS*, 117(28), pp.16546-16556.

McNamara, C., Lee, M., Lim, C., Lim, S., Roland, J., Nagle, A. and Winzeler, E. (2013). Targeting Plasmodium PI(4)K to eliminate malaria. *Nature*., 504, pp. 248-253.

McPhail, J. and Burke, J., 2020. *Druggable Lipid Signaling Pathways*. Cham: Springer International Publishing, pp.203-222.

McPherson, A. and Gavira, J., 2013. Introduction to protein crystallization. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, 70, pp.2-20.

Menard, D. and Dondorp, A., 2017. Antimalarial Drug Resistance: A Threat to Malaria Elimination. *CHS Perspect. Med.*, 7.

Milani, K., Schneider, T. and Taraschi, T., 2015. Defining the Morphology and Mechanism of the Hemoglobin Transport Pathway in *Plasmodium falciparum* -Infected Erythrocytes. *Eukaryot. Cell*, 14(4), pp.415-426.

Miroux, B. and Walker, J., 1996. Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *J. Mol. Biol.*, 260(3), pp.289-298.

Miyake, R., Kawamoto, J., Wei, Y., Kitagawa, M., Kato, I., Kurihara, T. and Esaki, N., 2007. Construction of a Low-Temperature Protein Expression System Using a Cold-Adapted Bacterium, Shewanella sp. Strain Ac10, as the Host. *Appl. Environ. Microbiol.*, 73(15), pp.4849-4856.

Modi, V. and Dunbrack, R., 2019. Defining a new nomenclature for the structures of active and inactive kinases. *PNAS*, 116(14), pp.6818-6827.

Mondal, J., Tiwary, P. and Berne, B., 2016. How a Kinase Inhibitor Withstands Gatekeeper Residue Mutations. *J. Am. Chem. Soc.*, 138(13), pp.4608-4615.

Moody, R. and Williamson, M., 2013. Structure and function of a bacterial Fasciclin I Domain Protein elucidates function of related cell adhesion proteins such as TGFBIp and periostin. *FEBS Open Bio*, 3(1), pp.71-77.

Müller, G., Klebl, B. and Hamacher, M., 2011. *Protein kinases as drug targets*. Weinheim, Germany: Wiley-VCH.

Muralidharan, V. and Goldberg, D., 2013. Asparagine Repeats in *Plasmodium falciparum* Proteins: Good for Nothing? *PLoS Pathog.*, 9(8), p.e1003488.

Mustière, R., Vanelle, P. and Primas, N., 2020. Plasmodial Kinase Inhibitors Targeting Malaria: Recent Developments. *Molecules*, 25(24), p.5949.

Naing, C., Whittaker, M., Htet, N., Aye, S. and Mak, J., 2019. Efficacy of antimalarial drugs for treatment of uncomplicated falciparum malaria in Asian region: A network meta-analysis. *PLOS ONE*, 14(12), p.e0225882.

Nakagawa, T., Goto, K. and Kondo, H., 1996. Cloning, Expression, and Localization of 230-kDa Phosphatidylinositol 4-Kinase. *J. Biol. Chem.*, 271(20), pp.12088-12094.

Nakatani, K., Ishikawa, H., Aono, S. and Mizutani, Y., 2013. Heme-binding properties of heme detoxification protein from *Plasmodium falciparum*. *Biochem. Biophys. Res. Commun.*, 439(4), pp.477-480.

Nakatani, K., Ishikawa, H., Aono, S. and Mizutani, Y., 2014. Identification of Essential Histidine Residues Involved in Heme Binding and Hemozoin Formation in Heme Detoxification Protein from *Plasmodium falciparum*. *Sci. Rep.*, 4, 6137.

Noedl, H., Se, Y., Schaecher, K., Smith, B., Socheat, D. and Fukuda, M., 2008. Evidence of Artemisinin-Resistant Malaria in Western Cambodia. *N. Engl. J. Med.*, 359(24), pp.2619-2620.

Nugent, T. and Jones, D., 2009. Transmembrane protein topology prediction using support vector machines. *BMC Bioinf.*, 10(1).

Pagola, S., Stephens, P., Bohle, D., Kosar, A. and Madsen, S., 2000. The structure of malaria pigment β-haematin. *Nature*, 404(6775), pp.307-310.

Pagola, S., Stephens, P., Bohle, D., Kosar, A. and Madsen, S., 2000. The structure of malaria pigment β-haematin. *Nature*, 404(6775), pp.307-310.

Papageorgiou, A. and Mattsson, J., 2014. Protein Structure Validation and Analysis with X-Ray Crystallography. In: N. Labrou, ed., *Protein Downstream Processing: Design, Development and Application of High and Low-Resolution Methods*, 1st ed. New Jersey: Humana Press, pp.397-403.

Paquet, T., Le Manach, C., Cabrera, D., Younis, Y., Henrich, P., Abraham, T., … Chibale, K., 2017. Antimalarial efficacy of MMV390048, an inhibitor of *Plasmodium* phosphatidylinositol 4-kinase. *Sci. Transl. Med.*, 9, eaad9735.

Penzo, M., de las Heras-Dueña, L., Mata-Cantero, L., Diaz-Hernandez, B., Vazquez-Muñiz, M., Ghidelli-Disse, S., Drewes, G., Fernandez-Alvaro, E. and Baker, D., 2019. High-throughput screening of the *Plasmodium falciparum* cGMP-dependent protein kinase identified a thiazole scaffold which kills erythrocytic and sexual stage parasites. *Sci. Rep.*, 9(1).

Phillips, M., Burrows, J., Manyando, C., Hooft van Huijsduijnen, R., Van Voorhis, W. and Wells, T., 2017. Malaria. *Nat. Rev. Dis. Primers.*, 3(17050).

Prasad, S., Khadatare, P. and Roy, I., 2011. Effect of Chemical Chaperones in Improving the Solubility of Recombinant Proteins in Escherichia coli. *Appl. Environ. Microbiol.*, 77(13), pp.4603-4609.

Prevo, R., Fokas, E., Reaper, P., Charlton, P., Pollard, J., McKenna, W., Muschel, R. and Brunner, T., 2012. The novel ATR inhibitor VE-821 increases sensitivity of pancreatic cancer cells to radiation and chemotherapy. *Cancer Biol. Ther.*, 13(11), pp.1072-1081.

Robinson, P., 2015. Enzymes: principles and biotechnological applications. *Essays in Biochemistry*, 59, pp.1-41.

Rohrbach, P., Sanchez, C., Hayton, K., Friedrich, O., Patel, J., Sidhu, A., Ferdig, M., Fidock, D. and Lanzer, M., 2006. Genetic linkage of pfmdr1 with food vacuolar solute import in *Plasmodium falciparum*. *EMBO J.*, 25(13), pp.3000-3011.

Rondeau, J. and Schreuder, H., 2015. Protein Crystallography and Drug Discovery. In: C. Wermuth, D. Aldous, P. Raboisson and D. Rognan, ed., *The Practice of Medicinal Chemistry*, 4th ed. Academic Press, pp.511-516.

Rosano, G. and Ceccarelli, E., 2014. Recombinant protein expression in Escherichia coli: advances and challenges. *Front. Microbiol.*, 5, 172.

Rosenthal, P., 2021. Has artemisinin resistance emerged in Africa? *Lancet Infect. Dis.*, 21(8), pp.1056-1057.

Roskoski Jr., R., 2022. Properties of FDA-approved small molecule protein kinase inhibitors: A 2022 update. *Pharmacol. Res.*, 175, p.106037.

Roskoski, R., 2015. A historical overview of protein kinases and their targeted small molecule inhibitors. *Pharmacol. Res.*, 100, pp.1-23.

Roskoski, R., 2016. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacol. Res.*, 103, pp.26-48.

Roskoski, R., 2020. Properties of FDA-approved small molecule protein kinase inhibitors: A 2020 update. *Pharmacol. Res.*, 152, p.104609.

Ross, L. and Fidock, D., 2019. Elucidating Mechanisms of Drug-Resistant *Plasmodium falciparum*. *Cell Host Microbe*, 26(1), pp.35-47.

Roy, A., Kucukural, A. and Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.*, 5(4), pp.725-738.

Russo Krauss, I., Merlino, A., Vergara, A. and Sica, F., 2013. An Overview of Biological Macromolecule Crystallization. *Int. J. Mol. Sci*, 14(6), pp.11643-11691.

Sanderson, K., 2013. Irreversible kinase inhibitors gain traction. *Nat. Rev. Drug Discovery*, 12(9), pp.649-651.

Santos, G. and Torres, N., 2013. New Targets for Drug Discovery against Malaria. *PLoS ONE*, 8(3), p.e59968.

Schultheiss, O. and Stanton, S., 2009. *Methods in social neuroscience*. New York: Guilford Press, pp.28-29.

Seifert, G., 2018. Fascinating Fasciclins: A Surprisingly Widespread Family of Proteins that Mediate Interactions between the Cell Exterior and the Cell Surface. *Int. J. Mol. Sci.*, 19(6), p.1628.

Shahinas, D., Folefoc, A. and Pillai, D., 2013. Targeting *Plasmodium falciparum* Hsp90: Towards Reversing Antimalarial Resistance. *Pathogens*, 2(1), pp.33-54.

Shchemelinin, I., Sefc, L. and Necas, E. (2006). Protein Kinases, Their Function and Implication in Cancer and Other Diseases. *Folia Biol.*, 52, pp.81-101.

Shears, S., 2009. Diphosphoinositol Polyphosphates: Metabolic Messengers? *Mol. Pharmacol.*, 76(2), pp.236-252.

Shen, C., 2019. Gene Expression: Translation of the Genetic Code. *Diagnostic Molecular Biology*, pp.87-116.

Shirai, A., Matsuyama, A., Yashiroda, Y., Hashimoto, A., Kawamura, Y., Arai, R., Komatsu, Y., Horinouchi, S. and Yoshida, M., 2008. Global Analysis of Gel Mobility of Proteins and Its Use in Target Identification. *J. Biol. Chem.,* 283(16), pp.10745-10752.

Sinha, S., Sarma, P., Sehgal, R. and Medhi, B., 2017. Development in Assay Methods for in Vitro Antimalarial Drug Efficacy Testing: A Systematic Review. *Front. Pharmacol.*, 8.

Sleator, R. and Walsh, P., 2010. An overview of *in silico* protein function prediction. *Arch. Microbiol.*, 192(3), pp.151-155.

Sorensen, H. and Mortensen, K., 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microb. Cell Fact.*, 4(1).

Spriestersbach, A., Kubicek, J., Schäfer, F., Block, H., & Maertens, B., 2015. Purification of His-Tagged Proteins. *Methods Enzymol.*, *559*, 1–15.

Srinivasan, A. and Krupa, A., 2004. A genomic perspective of protein kinases in *Plasmodium falciparum*. *Proteins Struct. Funct. Bioinf.*, 58(1), pp.180-189.

Srivastava, A., Nagai, T., Srivastava, A., Miyashita, O. and Tama, F., 2018. Role of Computational Methods in Going beyond X-ray Crystallography to Explore Protein Structure and Dynamics. *Int. J. Mol. Sci.*, 19(11), p.3401.

Sternberg, A. and Roepe, P., 2020. Heterologous Expression, Purification, and Functional Analysis of the *Plasmodium falciparum* Phosphatidylinositol 4-Kinase IIIβ. *Biochem.*, 59(27), pp.2494-2506.

Stevens, R., 2000. Design of high-throughput methods of protein production for structural biology. *Structure*, 8(9), pp.R177-R185.

Strelow, J., 2017. A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS Discovery*, 22(1), pp.3-20.

Sullivan, D., Gluzman, I. and Goldberg, D., 1996. *Plasmodium* Hemozoin Formation Mediated by Histidine-Rich Proteins. *Science*, 271(5246), pp.219-222.

Tachioka, M., Sugimoto, N., Nakamura, A., Sunagawa, N., Ishida, T., Uchiyama, T., Igarashi, K. and Samejima, M., 2016. Development of simple random mutagenesis protocol for the protein expression system in Pichia pastoris. *Biotechnol. Biofuels*, 9(1).

Talapko, J., Škrlec, I., Alebić, T., Jukić, M. and Včev, A., 2019. Malaria: The Past and the Present. *Microorganisms*, 7(6), pp.179-196.

Talevich, E., Tobin, A., Kannan, N. and Doerig, C., 2012. An evolutionary perspective on the kinome of malaria parasites. *Philos. Trans. R. Soc. B: Biol. Sci.*, 367(1602), pp.2607-2618.

Tawk, L., Chicanne, G., Dubremetz, J., Richard, V., Payrastre, B., Vial, H., Roy, C. and Wengelnik, K., 2010. Phosphatidylinositol 3-Phosphate, an Essential Lipid in Plasmodium, Localizes to the Food Vacuole Membrane and the Apicoplast. *Eukaryot. Cell*, 9(10), pp.1519-1530.

Terpe, K., 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.*, 72(2), pp.211-222.

The Lancet, 2021. Malaria vaccine approval: a step change for global health. *Lancet*, 398(10309), p.1381.

Treiber, D. and Shah, N., 2013. Ins and Outs of Kinase DFG Motifs. *Chem. Biol.*, 20(6), pp.745-746.

Troxell, B. and Hassan, H., 2013. Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria. *Front. Cell. Infect. Microbiol.*, 3.

Tse, E., Korsik, M. and Todd, M., 2019. The past, present and future of anti-malarial medicines. *Malar. J.*, 18(93).

Uwimana, A., Umulisa, N., Venkatesan, M., Svigel, S., Zhou, Z., Munyaneza, T., Habimana, R., Rucogoza, A., Moriarty, L., Sandford, R., Piercefield, E., Goldman, I., Ezema, B., Talundzic, E., Pacheco, M., Escalante, A., Ngamije, D., Mangala, J., Kabera, M., Munguti, K., Murindahabi, M., Brieger, W., Musanabaganwa, C., Mutesa, L., Udhayakumar, V., Mbituyumuremyi, A., Halsey, E. and Lucchi, N., 2021. Association of Plasmodium falciparum kelch13 R561H genotypes with delayed parasite clearance in Rwanda: an open-label, single-arm, multicentre, therapeutic efficacy study. *Lancet Infect. Dis.*, 21(8), pp.1120-1128.

Walker, K., 2016. Site-Directed Mutagenesis. *Encyclopedia of Cell Biology*, 1, pp.122-127.

Ward, P., Equinet, L., Packer, J. and Doerig, C., 2004. Protein kinases of the human malaria parasite *Plasmodium falciparum*: the kinome of a divergent eukaryote. *BMC Genomics*, 5(1).

Watsierah, C., Jura, W., Oyugi, H., Abong'o, B. and Ouma, C., 2010. Factors determining anti-malarial drug use in a peri-urban population from malaria holoendemic region of western kenya. *Malar. J.*, 9(1).

Weiss, G., Watanabe, C., Zhong, A., Goddard, A. and Sidhu, S., 2000. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *PNAS*, 97(16), pp.8950-8954.

Wicht, K., Mok, S. and Fidock, D., 2020. Molecular Mechanisms of Drug Resistance in *Plasmodium falciparum* Malaria. *Annu. Rev. Microbiol.*, 74(1), pp.431-454.

Wilson, K. and Walker, J., 2010. *Principles and Techniques of Biochemistry and Molecular Biology*. 7th ed. Cambridge: Cambridge Univ. Press, pp.459-461; pp.584-585.

Wingfield, P., 2001. Protein Precipitation Using Ammonium Sulfate. *Curr Protoc Protein Sci.*, Appendix 3(Appendix 3F).

Wong, K., Meyers, R. and Cantley, L., 1997. Subcellular Locations of Phosphatidylinositol 4-Kinase Isoforms. *J. Biol. Chem.*, 272(20), pp.13236-13241.

World Health Organization. 2022. *World Malaria Report 2021*. World Health Organization.

Wu, P., Nielsen, T. and Clausen, M., 2016. Small-molecule kinase inhibitors: an analysis of FDA-approved drugs. *Drug Discov. Today*, 21(1), pp.5-10.

Yang, J. and Zhang, Y., 2015. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.*, 43(W1), pp.W174-W181.

Younis, Y., Douelle, F., González Cabrera, D., Le Manach, C., Nchinda, A., Paquet, T., Street, L., White, K., Zabiulla, K., Joseph, J., Bashyam, S., Waterson, D., Witty, M., Wittlin, S., Charman, S. and Chibale, K., 2013. Structure–Activity-Relationship Studies around the 2-Amino Group and Pyridine Core of Antimalarial 3,5-Diarylaminopyridines Lead to a Novel Series of Pyrazine Analogues with Oral in Vivo Activity. *J. Med. Chem.*, 56(21), pp.8860-8871.

Zhang, H., Cao, X., Tang, M., Zhong, G., Si, Y., Li, H., Zhu, F., Liao, Q., Li, L., Zhao, J., Feng, J., Li, S., Wang, C., Kaulich, M., Wang, F., Chen, L., Li, L., Xia, Z., Liang, T., Lu, H., Feng, X. and Zhao, B., 2021. A subcellular map of the human kinome. *ELife*, 10.

Zhang, J., Chung, T. and Oldenburg, K., 1999. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screening*, 4(2), pp.67-73.

Zhang, Y., 2009. Protein structure prediction: when is it useful? *Curr. Opin. Struct. Biol.*, 19(2), pp.145-155.

Zhao, Z. and Bourne, P., 2020. *Next Generation Kinase Inhibitors*. Cham: Springer International Publishing, pp.13-28.

Appendices

Appendix A: Recombinant *Plasmodium falciparum* Haem Detoxification Protein expression vector and construct

MKNRFYYNLIIKRLYTRSGGLRKPQKVTNDPESINRKVYWCFEHKPVKRTIINLIYSH
NELKIFSNLLNHPTVGSSLIHELSLDGPYTAFFPSNEAMQLINIESFNKLYNDENKLSEF
VLNHVTKEYWLYRDLYGSSYQPWLMYNEKREAPEKLRNLLNNDLIVKIEGEFKHCN
HSIYLNGSKIIRPNMKCHNGVVHIVDKPIIF

Number of amino acids: 205

Protein size: 25 kDa

Figure A1: *Plasmodium falciparum* HDP protein sequence

Figure A2: pColdI vector map

Appendix B: RoseTTAFold analysis of *Plasmodium falciparum* Haem Detoxification Protein predicted model

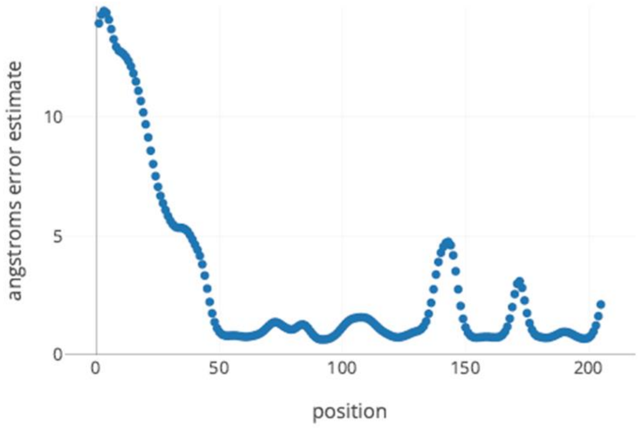


Figure B1: RoseTTAFold model error per residue

Error prediction for each residue position on the Haem Detoxification Protein structure predicted by RoseTTAFold.

Appendix C: Cloning and Sequencing of *Plasmodium vivax* phosphatidylinositol 4-kinase (*Pv*PI4K) F832A and C1327A mutants

Alignment statistics for match #1

Score	Expect	Identities	Gaps	Strand
1875 bits (1015)	0.0	1092/1126(97%)	23/1126(2%)	Plus/Plus

```
Query  46   CATCGTGGCTGCTTGCACCAACGTGTTCTCCATGACCGGCCTGTGCATCCCCCTCGAGTG  105
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  10   CATCGTGGCTGCTTGCACCAACGTGTTCTCCATGACCGGCCTGTGCATCCCCCTCGAGTG  69

Query  106  CCTGTCCCCTCCACAGTCCGGACAGCGTACCGAGCCCACCCCTTCCTCCAACGGCTCCAA  165
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  70   CCTGTCCCCTCCACAGTCCGGACAGCGTACCGAGCCCACCCCTTCCTCCAACGGCTCCAA  129

Query  166  CTCCAACGGTAGCTCCAACATCTCCGGTTCCTCTGGTTCCTCCGGCTCCGGCCTGCAGAT  225
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  130  CTCCAACGGTAGCTCCAACATCTCCGGTTCCTCTGGTTCCTCCGGCTCCGGCCTGCAGAT  189

Query  226  CCTGCACTTCAACTACGACGAGTGCAAGATCTTCTTCAGCAAGAAGCGTGCTCCCTACCT  285
            |||||||||||   |||||||||||||||||   ||||||||||||||||||||||||||
Sbjct  190  CCTGCACTTCA---ACGACGAGTGCAAGATCGCATTCAGCAAGAAGCGTGCTCCCTACCT  246

Query  286  GCTCATGTTCGAGGTGGCCGACCTGGACGAGGACATCTCCCACATCCCCGACGGCCTCTT  345
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  247  GCTCATGTTCGAGGTGGCCGACCTGGACGAGGACATCTCCCACATCCCCGACGGCCTCTT  306

Query  346  CTACCCCTCCTCCGTGGGTGGCGACGAGGCTGCTGGCGAGGAAGAGGAAGAGGCTGGACA  405
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  307  CTACCCCTCCTCCGTGGGTGGCGACGAGGCTGCTGGCGAGGAAGAGGAAGAGGCTGGACA  366

Query  406  GACCGCTTCCGAAACCGCTGGCGAGGCTGCCGACGGTGCTACCCGTGGTGTCACCGACGA  465
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  367   GACCGCTTCCGAAACCGCTGGCGAGGCTGCCGACGGTGCTACCCGTGGTGTCACCGACGA  426

Query  466   GCCCACCTCCCAGCCTACCCCTCCACCTCGTCGCAAGGGTCACCGTATCGAGCGCAACTT  525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  427   GCCCACCTCCCAGCCTACCCCTCCACCTCGTCGCAAGGGTCACCGTATCGAGCGCAACTT  486

Query  526   CTGCTCCTCCAAGAACTACGGCCTGTTCAACGAGAACTCCTCCAGCTACGACTACGAGCC  585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  487   CTGCTCCTCCAAGAACTACGGCCTGTTCAACGAGAACTCCTCCAGCTACGACTACGAGCC  546

Query  586   CCGTGGTGGTGTCCACGACGGCAAGCGCaaaaaaaagaaaaagaagaagaagCGCCGTTC  645
             |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Sbjct  547   CCGTGGTGGTGTCCACGACGGC-AGCGCAAAAAAAAGAAAAAGAAGAAGAAGCGCCGTTC  605

Query  646   CGAGTCCGGCGAGGACGACGAAGGTGGCGCTCACACCGGACAGTCCGGTTCCGAGGTGGA  705
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  606   CGAGTCCGGCGAGGACGACGAAGGTGGCGCTCACACCGGACAGTCCGGTTCCGAGGTGGA  665

Query  706   CCGTGGCAAGGAAGAGAGGGGCAAGGGCGCTCAGGGCAAGAAAACCACCGGCGCTGACCT  765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  666   CCGTGGCAAGGAAGAGAGGGGCAAGGGCGCTCAGGGCAAGAAAACCACCGGCGCTGACCT  725

Query  766   GGACGCCCTGAAGATGGAACACATGTACGTGTACAACTGCATCGTGAACGACCTGCGCAA  825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  726   GGACGCCCTGAAGATGGAACACATGTACGTGTACAACTGCATCGTGAACGACCTGCGCAA  785

Query  826   GGAAAACCTGATCAGCTTCCCATCCGGCGAAGAGGACTCCCTGTCCATCATCCGCAAGTG  885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  786   GGAAAACCTGATCAGCTTCCCATCCGGCGAAGAGGACTCCCTGTCCATCATCCGCAAGTG  845

Query  886   CATCGGCATGCGCGCTGACGAAGAGGACGAGG-AAGGCGGAGACACCGAGGGCGGCGACA  944
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Sbjct  846   CATCGGCATGCGCGCTGACGAAGAGGACGAGGAAAGGCGGAGACACCGAGGGCGGCGACA  905

Query  945   CTAACGGTGGCGACCCAGGCGACCCTTCCGAGCCTTCCAACCTGGGACAGGGTGGTCCTG  1004
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Sbjct  906   CTAACGGTGGCGACCCAGGCGACCCTTCCGAGCCTTCCAACCTGGGACAGGGT-GTCCTG  964

Query  1005  GCTGGGGTGCTCCCCCTGCTGAAGAGGGTCCCCGTGAAGCTACCAAGTCCTTCCTGGTCA  1064
             ||||||||||| ||||||||||||||||||||||||| |||| ||||||||||||| |||
Sbjct  965   GCTGGGGTGCT-CCCCTGCTGAAGAGGGTCCCCGTGA-GCTA-CAAGTCCTTCCTGATCA  1021

Query  1065  ACGCTCGTTCCGCTTCCATGCCCAACTACCTGTCCTCTTCCCTGGAACGTTGCAACATCT  1124
             ||||||| || ||| | ||| || |||||||||||||| || || | ||| ||| |||||
Sbjct  1022  ACGCTCGATC-GCT-C-ATGACC-ACTACCTGTCCTCT-CC-TG-A-CGT-GCA-CATCT  1071

Query  1125  CCTCCAACGACAACGACTCCAACGAAGTCAAGACCTCCGTGGAATC  1170
             | |||| |||| |||||||| |||| ||||  |||||||  |||||
Sbjct  1072  C-TCCA-CGAC-ACGACTCCGACGA-GTCARAACCTCCGKKGAATC  1113
```

Figure C1: Pairwise alignment of the *Pv*PI4K-F832 mutant construct against wild-type *Pv*PI4K using BLAST

The *Pv*PI4K-F832A construct was sequenced and compared to the wild-type *Pv*PI4K sequence using BLAST. Query-*Pv*PI4K-WT; Subject-*Pv*PI4K-F832A. Mutated residues highlighted in yellow.

Alignment statistics for match #1

Score	Expect	Identities	Gaps	Strand
1317 bits (713)	0.0	725/732(99%)	3/732(0%)	Plus/Plus

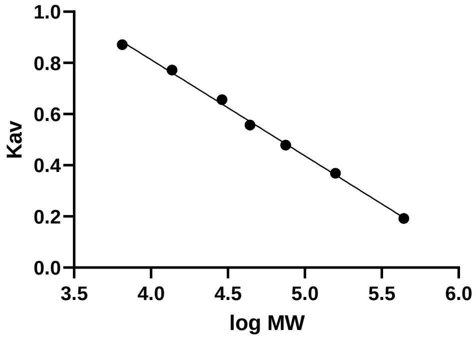
```
Query  51   GCT-GGCTTCCCAGCTGATCCGTCAGTTCAAGATCATCTTCGAAAACGCTGGCCTGCCCC  109
            ||| || |||||||||||||||||||||||||||||||||||||||||||| | ||||||
Sbjct  13   GCTGGGYTTCCCAGCTGATCCGTCAGTTCAAGATCATCTTCGAAAACGCTGSCYTGCCCC  72

Query  110  TGTGGCTGCGCCCCTACGAGATCCTGGTCACCGGTGCTAACTCCGGTATCATCGAGTACG  169
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  73   TGTGGCTGCGCCCCTACGAGATCCTGGTCACCGGTGCTAACTCCGGTATCATCGAGTACG  132

Query  170  TGAACGACACTTGC-TCCGTGGACTCCCTGAAGCGCAAGTTCGGTGCTGACTCCATCTCC  228
            ||||||||||| || |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  133  TGAACGACACT-GCATCCGTGGACTCCCTGAAGCGCAAGTTCGGTGCTGACTCCATCTCC  191

Query  229  ACCATCTTCAACGTGGTGTTCGCTGACTACATCTTCGAGGCCAAGAAGAACTTCATCGAG  288
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  192  ACCATCTTCAACGTGGTGTTCGCTGACTACATCTTCGAGGCCAAGAAGAACTTCATCGAG  251

Query  289  TCCCACGCTGCTTACTCCCTCGTGTCTTACCTGCTGCAAGTGAAGGACCGTCACAACGGC  348
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  252  TCCCACGCTGCTTACTCCCTCGTGTCTTACCTGCTGCAAGTGAAGGACCGTCACAACGGC  311

Query  349  AACCTGCTGCTGGACTCCGACGGTCACCTGATCCACATCGACTACGGTTTCATGCTGACC  408
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  312  AACCTGCTGCTGGACTCCGACGGTCACCTGATCCACATCGACTACGGTTTCATGCTGACC  371

Query  409  AACTCCCCCGGAAACGTGAACTTCGAGACTTCCCCATTCAAGCTCACCCAAGAGTACCTG  468
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  372  AACTCCCCCGGAAACGTGAACTTCGAGACTTCCCCATTCAAGCTCACCCAAGAGTACCTG  431

Query  469  GACATCATGGACGGCGAGAAGTCCGAGAACTACGAATACTTCCGTCGTCTGATCGTGTCC  528
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  432  GACATCATGGACGGCGAGAAGTCCGAGAACTACGAATACTTCCGTCGTCTGATCGTGTCC  491

Query  529  GGTTTCCTCGAGGCTCGCAAGCACTCCGAAGAGATCATCCTGTTCGTCGAGCTGATGATG  588
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  492  GGTTTCCTCGAGGCTCGCAAGCACTCCGAAGAGATCATCCTGTTCGTCGAGCTGATGATG  551

Query  589  CCCGCTCTGAAGATCCCCTGCTTCGCTAACGGCACCCAGTTCTGCATCGACTCCCTCAAG  648
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  552  CCCGCTCTGAAGATCCCCTGCTTCGCTAACGGCACCCAGTTCTGCATCGACTCCCTCAAG  611

Query  649  GAACGTTTCATGACCAACCTGGCTGTGGACGTGTGCATCCAGCGTATCAACGCTCTGATC  708
            |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  612  GAACGTTTCATGACMAACCTGGCTGTGGACGTGTGCATCCAGCGTATCAACGCTCTGATC  671

Query  709  GAGTCCTCCATCAACAACTTCCGTTCCGTGCAGTACGACTACTTCCAGCGCATCACCAAC  768
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  672  GAGTCCTCCATCAACAACTTCCGTTCCGTGCAGTACGACTACTTCCAGCGCATCACCAAC  731

Query  769  GGCATCATGTAA  780
            ||||||||||||
Sbjct  732  GGCATCATGTAA  743
```

Figure C2: Pairwise alignment of the *Pv*PI4K-C1327A mutant construct against wild-type *Pv*PI4K using BLAST

The *Pv*PI4K-C1327A construct was sequenced and compared to the wild-type *Pv*PI4K sequence using BLAST. Query-*Pv*PI4K-WT; Subject-*Pv*PI4K-C1327A. Mutated residues highlighted in yellow.

Appendix D: HiLoad 16/600 Superdex 200 pg column calibration curve

Calibration of the size exclusion column was performed by Kaylene Baron (UCT). A mixed sample containing elements of known molecular weight was analysed on the HiLoad 16/600 Superdex 200 pg column. The elution time and molecular weight of the samples was used to generate a calibration curve.



Equation	Y = -0.3758*X + 2.316

Figure D1: Calibration curve for the Superdex 200 pg size exclusion column

Calibration curve generated when analysing the elution time of control samples with known molecular weight.

Appendix E: ADP-Glo™ kinase assay ATP-to-ADP conversion standard curves and equations

A standard curve was generated corresponding the luminescence signal observed to the conversion of ATP to ADP. The standard curve prepared according to the manufacturer's instructions. Standard solutions of 100 µM ATP and ADP were prepared and combined in different ratios to simulate the different concentrations of ATP and ADP at various percent conversion conditions. This was repeated using 10 µM ATP and ADP.

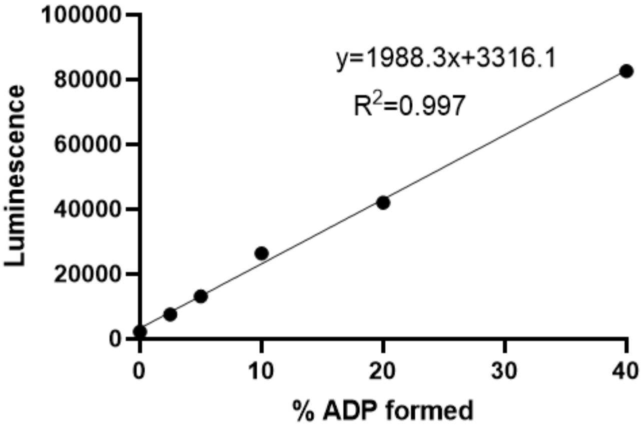


Figure E1: ATP to ADP conversion standard curve at 10 µM ATP

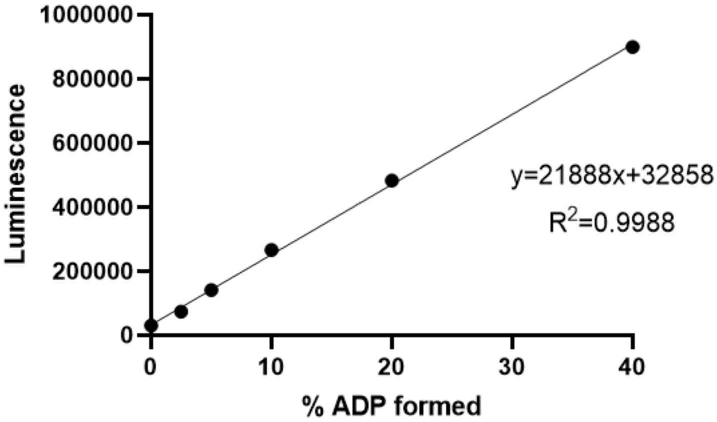


Figure E2: ATP to ADP conversion standard curve at 100 µM ATP

Table E1: ATP to ADP conversion standard curve equations

ATP and ADP concentration (µM)	Standard curve equation
500	Y = 349115X + 1676551
250	Y = 199519X + 669237
125	Y = 107761X + 322974
100	Y = 21888X + 32852
62.5	Y = 58166X + 137392
31.25	Y = 28957X + 72106
15.625	Y = 17516X + 824
10	Y= 1988,3X + 3316,1
2	Y = 1322X + 977,4
1	Y = 624,4X + 1295

Appendix F: *Plasmodium* PI4K inhibitors evaluated in this project

H3D number	**MMV number**
H3D-002054	MMV390048
H3D-009499	MMV010545
H3D-009614	MMV026468
H3D-009794	MMV026020
H3D-009605-02-01	MMW026356
H3D-009613	MMV020670
H3D-009579	MMV023860
H3D-009632	MMV675993
H3D-001911	MMV010576
H3D-002027	MMV085499
H3D-002002	MMV034137
H3D-002090	MMV390394
H3D-006984	MMV1793498
H3D-006666	MMV1792459
H3D-012859	MMV1848400
H3D-017017	MMV1902872
H3D-016630	-
H3D-018390	-
H3D-018391	-
H3D-018392	-
H3D-018393	-
H3D-018394	-
-	🟩
H3D-002182	UCT943
H3D-002610	UCT594

Appendix G: Table showing off-target modified proteins by tool TCI compounds

Compound	Protein ID	PEP Score
MMV1792459	Q91F75	0.031
	Q91G65	0.026
MMV1848400	Q0E589	0.032
	Q91FU0	0.037
	Q91FW7	0.018
	Q91G65	0.026
	CON__ENSEMBL:ENSBTAP00000018574	0.026
MMV1902872	O55741	0.031
	Q91F86	0.019
	Q91FR7	0.032
	Q9QSK1	0.029
	CON__Q29443/ CON__Q0IIK2	0.024

Appendix H: TMHMM analysis of the *Pf*HDP sequence