INVESTIGATING A BIR-CONTAINING *PLASMODIUM FALCIPARUM* PROTEIN AND IDENTIFYING ITS BINDING PARTNERS

Dale Liebenberg

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 2018

DECLARATION

I, Dale Liebenberg, declare that this Thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted for any other degree or examination at any other University.

Dale Liebenberg

1st day of March 2018

In memory of my ever faithful shadow

Patches

2012 – 2017

RESEARCH OUTPUTS

Submitted Manuscript

Building a Protein Interaction Network for a *P. falciparum* BIR-Containing Protein using Phage Display Technology

Dale Liebenberg and Thérèsa L. Coetzer

- submitted to *Parasitology* (IF 2.713)

International Conferences

Identification of Binding Partners for a Putative *P. falciparum* Inhibitor of Apoptosis (IAP) Protein

Dale Liebenberg and Thérèsa L. Coetzer

Oral presentation at:

- The 6th Multilateral Initiative on Malaria (MIM) Pan-African Malaria Conference, Durban, KwaZulu-Natal, South Africa, October 2013

Study of a Putative *P. falciparum* Inhibitor of Apoptosis (IAP) Protein

Dale Liebenberg, Pierre M. Durand and Thérèsa L. Coetzer

Poster presentation at:

- The 23rd South African Society of Biochemistry and Molecular Biology (SASBMB) / Federation of African Societies of Biochemistry and Molecular Biology (FASBMB) Congress, Drakensburg, KwaZulu-Natal, South Africa, January 2012

Local Conferences

Using CRISPR-Cas9 to Manipulate the Genome of *P. falciparum*

Dale Liebenberg and Thérèsa L. Coetzer

Oral presentation at:

- The 25th South African Society of Biochemistry and Molecular Biology (SASBMB) Congress, East London, Eastern Cape, South Africa, July 2016

Poster presentations at:

- The Faculty of Health Sciences Research Day and Postgrad Expo, University of the Witwatersrand, Johannesburg, Gauteng, South Africa, September 2016
  - awarded Best Student Poster Presentation in the Molecular and Comparative Biosciences category
- The South African Medical Research Council's Office of Malaria Research (MOMR) 2016 Malaria Research Meeting, University of Pretoria, Pretoria, Gauteng, South Africa, July 2016

Can Protein-Protein Interactions Involved In Apoptosis Be Potential Drug Targets?

Dale Liebenberg and Thérèsa L. Coetzer

Oral presentation at:

- The South African Medical Research Council's Office of Malaria Research (MOMR) 2015 Malaria Research Meeting, Durban, KwaZulu-Natal, South Africa, August 2015

Novel Binding Partners for a *P. falciparum* Inhibitor of Apoptosis (IAP) Protein

Dale Liebenberg and Thérèsa L. Coetzer

Oral presentations at:

- The Faculty of Health Sciences Research Day and Postgrad Expo, University of the Witwatersrand, Johannesburg, Gauteng, South Africa, September 2014
- The Molecular Biosciences Research Thrust (MBRT) Research Day, University of the Witwatersrand, Johannesburg, Gauteng, South Africa, December 2014
  - awarded 3rd prize in the Oral Category

Poster presentation at:

- The 24th South African Society of Biochemistry and Molecular Biology (SASBMB) Congress, Worcester, Western Cape, South Africa, July 2014

Identification of Binding Partners for a Putative *P. falciparum* Inhibitor of Apoptosis (IAP) Protein

Dale Liebenberg and Thérèsa L. Coetzer

Poster presentation at:

- The Molecular Biosciences Research Thrust (MBRT) Research Day, University of the Witwatersrand, Johannesburg, Gauteng, South Africa, December 2013

Expression of a Putative *P. falciparum* Inhibitor of Apoptosis (IAP) Protein

Dale Liebenberg and Thérèsa L. Coetzer

Oral presentation at:

- The Faculty of Health Sciences Research Day and Postgrad Expo, University of the Witwatersrand, Johannesburg, Gauteng, South Africa, September 2012

Poster presentation at:

- The Molecular Biosciences Research Thrust (MBRT) Research Day, University of the Witwatersrand, Johannesburg, Gauteng, South Africa, December 2012

Study of a Putative *P. falciparum* Inhibitor of Apoptosis (IAP) Protein

Dale Liebenberg, Pierre M. Durand and Thérèsa L. Coetzer

Poster presentation at:

- The Molecular Biosciences Research Thrust (MBRT) Research Day, University of the Witwatersrand, Johannesburg, Gauteng, South Africa, December 2011

ABSTRACT

The majority of the worldwide malaria deaths are caused by the *Plasmodium falciparum* parasite and unfortunately parasite strains are emerging that are resistant to not only artemisinin, but also the partner drugs used in current antimalarial combination therapy. The intraerythrocytic *P. falciparum* life stage is characterised by exponential, asexual proliferation that could cause the premature death of the human host before the sexual gametocytes have had enough time to develop and be taken up by the mosquito vector to continue its lifecycle. It is hypothesised that *P. falciparum* maintains its population at a level low enough to allow for the transmission of these gametocytes by using a form of regulated cell death (RCD). The molecular members of this cell death pathway are currently unclear, but a putative *P. falciparum* inhibitor of apoptosis protein (PfIAP; PF3D7_0519600) has previously been identified. Metazoan IAP proteins play anti-apoptotic roles in cells by interacting and inhibiting pro-apoptotic caspases, but also perform other functions.

Analysis of the PfIAP protein using bioinformatic tools revealed that it contains one conserved baculoviral IAP repeat (BIR) domain and that this *P. falciparum* BIR domain is structurally similar to the BIR domains of various human IAP proteins. mRNA extracted from asexual *P. falciparum* parasites was used to construct a biotin-tagged phage display library, which was used in biopanning experiments with two regions of the PfIAP protein, expressed as recombinant GST-tagged proteins. Four binding partners were identified for the N-terminal BIR domain of the protein, while two proteins were identified as interacting partners for the C-terminal region of PfIAP. Of these, a double C2-like domain-containing protein (PfDOC2) and the high molecular weight rhoptry protein 3 (PfRhopH3) were expressed as recombinant His-tagged proteins and verified as PfIAP binding partners by *in vitro* binding assays. Transgenic *P. falciparum* parasites were generated expressing a GFP-tagged PfIAP BIR domain, which localised to the cytoplasm under both normal and high temperature conditions which mimic febrile malaria, a physiological trigger of RCD. Knockout experiments of the *pfiap* gene using the CRISPR-Cas9 genome editing tool suggested that this gene could be essential for the survival of asexual *P. falciparum* parasites. This study offers the first details of a putative *P. falciparum* inhibitor of apoptosis protein and suggests that it could have non-apoptotic roles in the parasite, given the diverse functions of the binding partners that comprise the PfIAP protein-protein network.

ACKNOWLEDGEMENTS

"The time will come when diligent research over long periods will bring to light things which now lie hidden. A single lifetime, even though entirely devoted to the sky, would not be enough for the investigation of so vast a subject... And so this knowledge will be unfolded only through long successive ages. There will come a time when our descendants will be amazed that we did not know things that are so plain to them... Many discoveries are reserved for ages still to come, when memory of us will have been effaced. Our universe is a sorry little affair unless it has in it something for every age to investigate... Nature does not reveal her mysteries once and for all."

*— **Seneca**, Natural Questions Book 7, c. first century*

ETHICS CLEARANCE

Ethics clearance was granted by the University of the Witwatersrand Human Research Ethics Committee (Medical). Clearance certificate number: M1703102.

TABLE OF CONTENTS

LIST OF FIGURES

List of Figures

List of Figures

LIST OF TABLES

NOMENCLATURE

A_{230}	Absorbance at 230 nm
A_{260}	Absorbance at 260 nm
A_{280}	Absorbance at 280 nm
aa	Amino Acid
ACD	Accidental Cell Death
ACT	Artemisinin-Based Combination Therapy
Ae. aegypti	*Aedes aegypti*
An. gambiae	*Anopheles gambiae*
ATG23	Autophagy-Related Protein 23
BAX	Bcl-2-Associated X Protein
Bcl-2	B Cell Lymphoma 2
Bcl-xL	B-Cell Lymphoma-Extra Large
BID	BH3 Interacting-Domain Death Agonist
BIR	Baculovirus IAP Repeat
BIRC	Baculoviral IAP Repeat Containing Protein
bp	Base Pairs
BSA	Bovine Serum Albumin
Cas	CRISPR-Associated Protein
CD	Cytosine Deaminase
cDNA	Complementary DNA
CHD1	Chromodomain-Helicase-DNA-Binding Protein 1
cIAP1/2	Cellular IAP1/2
CRISPR	Clustered Regulatory Interspaced Short Palindromic Repeat
crRNA	CRISPR RNA
dCas9	Dead Cas9
DD	Destabilizing Domain
DHODH	Dihydroorotate Dehydrogenase
DISC	Death-Inducing Signalling Complex
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DOC2	Double C2-Like Domain-Containing Protein
DSB	Double-Strand Break
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic Acid
FADD	Fas-Associated Protein With Death Domain
GFP	Green Fluorescent Protein
GOI	Gene of Interest
GPA/B	Glycophorin A/B
gRNA	Guide RNA
GST	Glutathione S-Transferase

HA	Haemagglutinin
HA_L/R	Homologous Arms_Left/Right
hDHFR	Human Dihydrofolate Reductase
HR	Homologous Recombination
HRP	Horseradish Peroxidase
HRP2	Histidine Rich Protein 2
HtrA2/Omi	High Temperature Requirement A/Omi Stress Regulated Endoprotease
IAP	Inhibitor of Apoptosis
IBM	IAP-Binding Motif
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
KAHRP	Knob-Associated, Histidine-Rich Protein
kDa	Kilodalton
LB	Luria Broth
LisH	Lissencephaly Type-1-Like Homology
MCA	Metacaspase
MDM2	Mouse Double Minute 2
MESA	Mature Parasite-Infected Erythrocyte Surface Antigen
MOI	Multiplicity of Infection
MOMP	Mitochondrial Outer Membrane Permeabilisation
mRNA	Messenger RNA
MSP1	Merozoite Surface Protein 1
NAIP	Neuronal Apoptosis-Inhibitory Protein
NHEJ	Non-Homologous End Joining
OD_{600}	Optical Density at 600 nm
P. falciparum	*Plasmodium falciparum*
PAM	Protospacer Adjacent Motif
PBS	Phosphate Buffered Saline
PBST	Tween20-PBS
PCD	Programmed Cell Death
PCR	Polymerase Chain Reaction
pfu	Plaque Forming Unit
pH	Potential of Hydrogen
Phyre2	Protein Homology/analogY Recognition Engine v2.0
pI	Isoelectric Point
PPI	Protein-Protein Interaction
pRBC	Parasitised Red Blood Cell
PS	Phosphatidylserine
PSAC	Plasmodial Surface Anion Channel
RBC	Red Blood Cell
RCD	Regulated Cell Death
RDT	Rapid Diagnostic Test
RhopH3	High Molecular Weight Rhoptry Protein 3

RING	Really Interesting New Gene
RNA	Ribonucleic Acid
rRNA	Ribosomal RNA
RT	Reverse Transcriptase
RVD	Repeat Variable Diresidue

S. cerevisiae	*Saccharomyces cerevisiae*
S. pyogenes	*Streptococcus pyogenes*
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
sgRNA	Single Guide RNA
SMAC/DIABLO	Second Mitochondria-Derived Activator of Caspases/Direct IAP Binding Protein with Low pI
SMART	Simple Modular Architecture Research Tool
snRNA	Small Nuclear RNA
SWI/SNF	Switch/Sucrose Non-Fermentable
SWIB	SWI/SNF Complex B

T. gondii	*Toxoplasma gondii*
T_a	Annealing Temperature
TAE	Tris, Acetic Acid and EDTA
TALE	Transcription Activator-Like Effector
TALEN	TALE Nuclease
TB	Terrific Broth
TBS	Tris Buffered Saline
TBST	Tween20-TBS
TE	Tris and EDTA
T_m	Melting Temperature
tracrRNA	Trans-Activating crRNA
tRNA	Transfer RNA

UBC	Ubiquitin Conjugating Enzyme
UTR	Untranslated Region

WHO	World Health Organisation

XIAP	X-Linked Inhibitor of Apoptosis Protein

YFCU	Yeast Cytosine Deaminase/Uracil Phosphoribosyl Transferase

ZF	Zinc Finger
ZFN	Zinc Finger Nuclease
Zn	Zinc

CHAPTER 1: Introduction

1.1 The Fight Against Malaria

There has been a steady decline in the number of malaria deaths over the past number of years (29% between 2010 and 2015). Africa recorded 90% of the estimated 212 million worldwide cases of malaria in 2015 and 92% of the estimated 429 000 deaths (World Health Organization 2016). Regrettably 70% of the global malaria deaths were of children younger than 5 years old (World Health Organization 2016). One of the goals of the World Health Organisation's Global Technical Strategy for Malaria 2016–2030 is to reduce the global incidence of and mortality from malaria by at least 90% compared to 2015 levels by 2030 (World Health Organization 2016). Funding is required to achieve this goal and in 2015 this amounted to an estimated US$ 2.9 billion, which unfortunately was less than the estimated required amount of US$ 6.4 billion (World Health Organization 2016). Achieving this Global Technical Strategy goal requires a multifaceted approach that confronts not only the *Plasmodium* parasite, but also the infected female *Anopheles* mosquito. Vector control methods primarily involve the distribution of insecticide-treated mosquito nets (ITNs) and protecting homes through indoor residual spraying (Benelli and Beier 2017). Due to the increase in insecticide resistance, particularly pyrethroids which is the only class used in ITNs, there is now a pressing need for new vector control tools (World Health Organization 2016; Benelli and Beier 2017).

Tackling the human malaria parasites: *Plasmodium falciparum*, *P. knowlesi*, *P. malariae*, *P. ovale* and *P. vivax* involves effective diagnostics and treatment of an infection, but ultimately immunisation against the parasite through the deployment of a vaccine. Rapid diagnostic tests (RDTs) allow for cheap and easy investigation of suspected malaria cases, but parasites in more than 10 countries with histidine rich protein 2 (HRP2) gene deletions can evade detection (World Health Organization 2016). Only one malaria vaccine has completed Phase 3 testing and despite showing only a 39% reduction in clinical incidents in young children, the WHO has recommended that the RTS,S/AS01 vaccine be deployed in a pilot project to gather more information on its feasibility, safety and impact on mortality (World Health Organization 2016).

The use of artemisinin-based combination therapy (ACT) is the primary method of treating someone infected with malaria, but resistance against artemisinin and its derivatives has emerged. The first artemisinin-resistant *P. falciparum* parasites were detected in Western Cambodia (Dondorp *et al.* 2009; Noedl *et al.* 2009). Since then there are areas of the Greater Mekong subregion (Cambodia, parts of China, Laos, Myanmar, Thailand, and Vietnam) where increasing artemisinin resistance is compounded by parasite resistance to the partner drugs, causing high failure rates of ACT (Dondorp *et al.* 2017; Woodrow and White 2017). The hallmark of artemisinin treatment is the fast clearance of ring stage parasites, so artemisinin resistance is best detected as the slow clearance of parasites from the peripheral blood (Figure 1.1.1). There is an urgent need to eliminate resistant *P. falciparum* from the Greater Mekong subregion, before the parasites become untreatable or spread to other parts of the world. To this end the deployment of new antimalarial compounds is required, as is the identification of novel parasite drug targets.
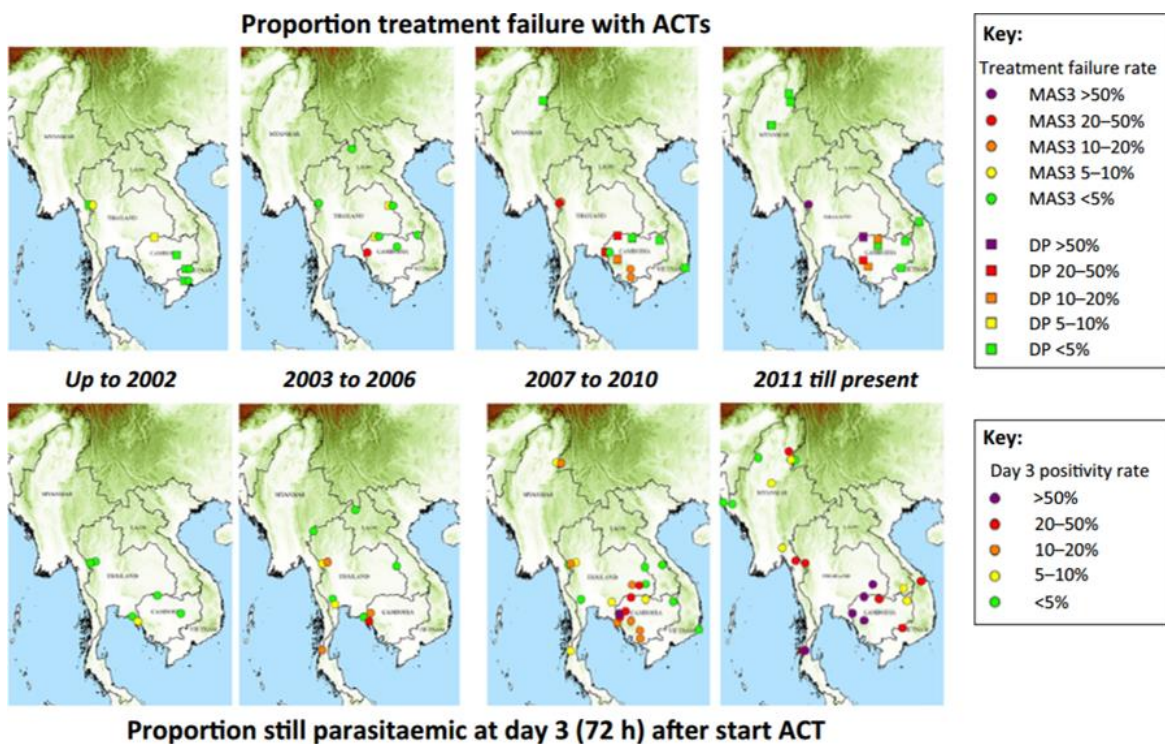
Figure 1.1.1: The Greater Mekong Subregion and Artemisinin Resistance
Adapted from Woodrow and White (2017).
Top) The percentage of treatment failure rates after treatment with artesunate–mefloquine (MAS3) or dihydroartemisinin–piperaquine (DP).
Bottom) The percentage of parasite positive patients 3 days after the initiation of artemisinin combination therapy (ACT), illustrating the current distribution of artemisinin resistance in *P. falciparum*.

1.2   The *Plasmodium* Life Cycle

All members of the genus *Plasmodium* are unicellular protists that cause malaria and have a complex life cycle involving a Dipteran insect vector and a vertebrate host (Figure 1.2.1). While asexual *Plasmodium* reproduction occurs in the human host, transmission between different humans and *Plasmodium* sexual reproduction can only occur in the *Anopheles* mosquito. When a female mosquito takes a blood meal from an infected person, both male and female gametocytes are ingested. In the mosquito midgut, the drop in temperature, exposure to xanthurenic acid and change in pH trigger the haploid gametocytes to mature into gametes (Sinden *et al.* 1998). Using their eight flagellae, the now motile male gametes fuse with female gametes to form diploid zygotes that develop into ookinetes. These motile ookinetes traverse the midgut epithelial cell wall and transform into oocysts. The ultimate rupture of the oocysts releases many sporozoites that actively migrate through the mosquito's haemocoel to the salivary glands. When the mosquito takes its next blood meal, these sporozoites are simultaneously injected into the new human host (Hafalla *et al.* 2011; Meibalan and Marti 2017).
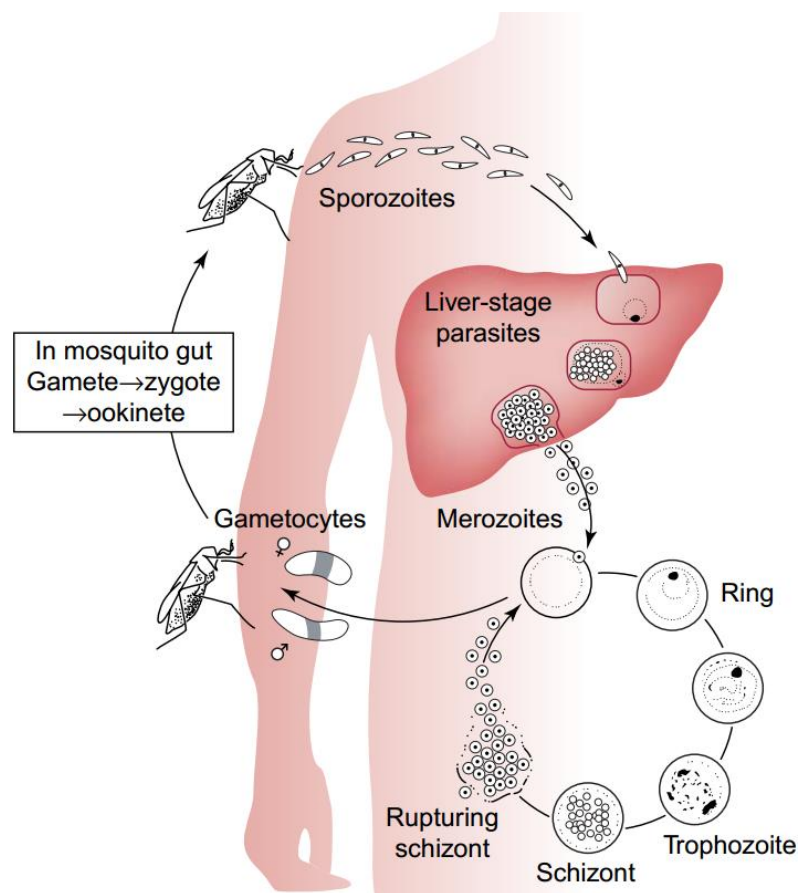
Figure 1.2.1: The Life Cycle of the *Plasmodium* Parasite (Good 2005)
Diagram showing the various stages of *Plasmodium* development in the mosquito vector and the human host.

Chapter 1: Introduction

The newly injected *Plasmodium* sporozoites travel to the human host's liver and invade hepatocytes. After replicating as hepatic schizonts, several thousand merozoites are then released into the vasculature to infect erythrocytes/red blood cells (RBCs). In *P. ovale* and *P. vivax* infections some hepatic parasites remain quiescent, only becoming infective at a later stage (White 2011). When inside the RBC, the merozoite becomes a ring stage parasite, develops into a trophozoite and then undergoes asexual replication forming a mature schizont (a process referred to as schizogony). This schizont ruptures the erythrocyte, releasing 16-32 merozoites that each invade a new RBC and begin the cycle again (Hafalla *et al.* 2011). For the *P. falciparum* parasite this ~48 hour intraerythrocytic process results in the clinical manifestations of the symptoms that are associated with the malaria disease (Miller *et al.* 2002).

For the malaria parasite to be able to undergo sexual reproduction in the mosquito vector, gametocytes are formed through the process of gametocytogenesis. *Plasmodium falciparum* gametocytogenesis can take approximately 10-12 days, which is considerably longer than other *Plasmodium* species. A *P. vivax* gametocyte requires 48 hours to develop, while the rodent malaria *P. berghei* requires little more than a day to produce a gametocyte (Josling and Llinás 2015; Meibalan and Marti 2017). For the majority of this time the developing *P. falciparum* gametocyte is sequestered in the bone morrow to avoid being removed by the spleen. While only a small proportion of asexual parasites develop into gametocytes, the commitment to sexual development occurs before schizogony as all the merozoites within a schizont either continue asexual development or differentiate into gametocytes (Bruce *et al.* 1990). The triggers involved in the commitment to gametocytogenesis are not defined, but environmental stresses like host immunity (Bousema *et al.* 2006) or high parasitaemia *in vitro* are associated with a higher rate of gametocytogenesis. This makes sense, since if the host is dying as a result of anaemia caused by a high parasitaemia, then it is in the best interest of the parasite to move to a new host (Josling and Llinás 2015).

The problem that arises for *P. falciparum* is that during the two weeks that are needed for gametocytogenesis, the exponential asexual reproduction could easily lead to the premature death of the host. Since the parasite cannot solely rely on the human immune system to remove infected RBCs at a high enough rate to counter the exponential asexual growth, it is hypothesised that some *P. falciparum* individuals undergo a form of cell suicide as a way of keeping the overall parasitaemia in the host at a manageable level (Deponte and Becker

2004). Cell death in multicellular organisms has been studied for decades and is the best available model for *P. falciparum* cell death.

1.3   The Death of Cells in Multicellular Organisms

1.3.1   Defining Regulated Cell Death

Over the course of time, from describing the loss of silk moth intersegmental muscles during development (Lockshin and Williams 1964) to the present, it has become clear that categorising the death of a cell as having occurred via apoptosis/programmed cell death (PCD) or through necrosis is an oversimplification. The Nomenclature Committee on Cell Death (NCCD) recommends that the death process be categorised as either accidental cell death (ACD) or regulated cell death (RCD; Galluzzi *et al.* 2015). ACD is the unpreventable, uncontrolled demise of a cell as a result of extreme physical, chemical or mechanical insults, while RCD is facilitated on a molecular level by genetic machinery and thus includes PCD. Such a classification is necessary as for example it has become evident that PCD does not always fail to induce an inflammatory response, an indicator typical of necrosis. PCD is proposed to now refer to RCD processes that occur as part of development or that preserve tissue homeostasis (Galluzzi *et al.* 2015).

RCD is a relatively new term and as yet to be fully adopted by all in the scientific world. Therefore this study will make use of the cell death terms (apoptosis, autophagy, necrosis etc.) used by the authors of published literature, but will use the term RCD when discussing the work described here.

1.3.2   The Apoptosis Pathway

Apoptosis is the best characterised of all the regulated cell death phenotypes and is instigated either through extracellular signals (such as environmental stress) or by internal cell signals (following infection with a pathogen for example), that both result in a proteolytic cascade involving caspases (Figure 1.3.1). This leads to the death of the cell through morphological changes that include DNA fragmentation, mitochondrial outer membrane permeabilisation (MOMP) and apoptotic body formation (Elmore 2007). Apoptosis is a complex process involving a range of different participants, one of these is the tumour suppressor protein p53. This protein is present in the cytosol but upon DNA damage translocates to the nucleus where it functions as a transcription factor and promotes the expression of a variety of pro-apoptotic proteins (Elmore 2007). A simplified description of the process of apoptosis and other RCD pathways follows.
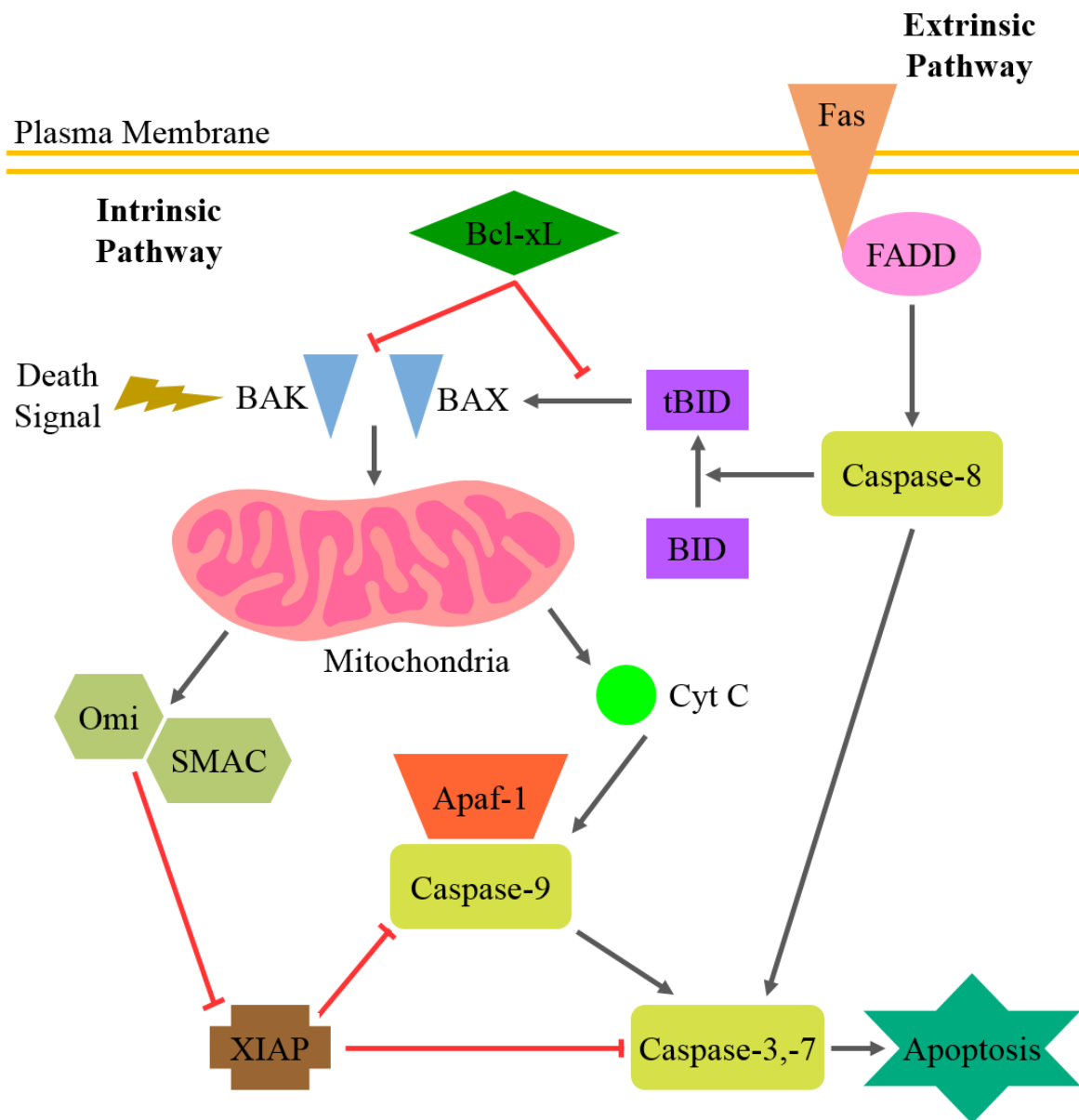
Figure 1.3.1: The Extrinsic and Intrinsic Pathways of Apoptosis
Diagram showing some of the major participants involved in metazoan apoptosis. Red line = protein inhibition. Apaf-1 = Apoptotic protease activating factor 1; BAK = Bcl-2 homologous antagonist/killer; BAX = Bcl-2-associated X protein; Bcl-xL = B-cell lymphoma-extra large; BID = BH3 interacting-domain death agonist; Cyt C = cytochrome c; FADD = Fas-associated protein with death domain; Omi = Omi stress regulated endoprotease; SMAC = second mitochondria-derived activator of caspases; XIAP = X-linked inhibitor of apoptosis protein. Based on Zhang *et al.* (2005); Ashkenazi (2008); Galluzzi *et al.* (2009); DeAlmagro and Vucic (2012); Czabotar *et al.* (2013); Ichim and Tait (2016).

Any of a multitude of cellular stressors, such as DNA damage, can initiate the intrinsic apoptosis pathway. This pathway is tightly regulated by the B cell lymphoma 2 (Bcl-2) protein family, which is comprised of both anti-apoptotic and pro-apoptotic members. Bcl-2-associated X protein (BAX) and Bcl-2 homologous antagonist/killer (BAK) are activated

in response to apoptotic stimuli and overcome the inhibitory effects of the anti-apoptotic Bcl-2 members (such as Bcl-xL). Translocation to the mitochondria by these activated pro-apoptotic proteins, leads to the loss of MOMP which results in the release of cytochrome c from the mitochondrial inter-membrane space into the cytoplasm. The proteins SMAC/DIABLO (second mitochondria-derived activator of caspases/direct IAP binding protein with low pI) and HtrA2/Omi (high temperature requirement A/Omi stress regulated endoprotease) are also released from the mitochondria at this stage. These two pro-apoptotic proteins sequester various inhibitor of apoptosis (IAP) proteins, allowing for the release of different caspases. The various caspase proteins are cysteinyl aspartate proteases that are present in the cytoplasm of the cell as inactive zymogens (Kumar 2007). Apoptotic protease activating factor 1 (Apaf-1) and the now free initiator pro-caspase-9 come together to form the apoptosome (along with cytochrome c) and activate effector caspases-3 and -7. This results in the death of the cell through various processes including the loss of cell skeleton integrity, chromatin condensation and fragmentation by nucleases and membrane blebbing. Active caspase-3, for example, cleaves the inhibitor of caspase-activated DNase (ICAD), which leads to the release of active CAD to mediate the fragmentation of DNA (Kumar 2007). The activation of caspases also brings about the accumulation of phosphatidylserine (PS) on the intact cell membrane outer leaflet, as opposed to normal conditions when it is retained exclusively on the inner leaflet of the membrane. This "eat me" signal is used to ensure the proper clearance of the apoptotic bodies by attracting the attention of phagocytes, thereby avoiding an inflammatory response (Elmore 2007; Ashkenazi and Salvesen 2014; Fuchs and Steller 2015).

The extrinsic death receptor pathway is initiated following the activation of transmembrane death receptors on the cell surface such as the Fas receptor. This leads to the recruitment of cytoplasmic proteins like FADD (Fas-associated protein with death domain), which form the death-inducing signalling complex (DISC). The initiator pro-caspase-8 interacts with DISC, is proteolytically cleaved and then directly activates caspase-3. This effector caspase is responsible for the death of the cell and the activation of caspase-7. The BH3 interacting-domain death agonist (BID) protein is also cleaved by activated caspase-8, after which the truncated BID protein triggers mitochondrial damage through the activation of the pro-apoptotic BAX (Elmore 2007; Ashkenazi and Salvesen 2014; Fuchs and Steller 2015).

Another phenotype of RCD is autophagy which is the mechanism by which superfluous or damaged proteins are removed. The targeted cytoplasmic components are engulfed by autophagosomes, which fuse with lysosomes to enable protein degradation and the resupply of the cell with building blocks. In a stress environment, autophagy has a pro-survival function by being able to delay apoptosis or suppress necrotic cell death (Mariño *et al.* 2014; Swart *et al.* 2016). Necrosis is the premature death of a cell that results in the release of cytokine interleukin-1 beta and activation of inflammatory responses. It is typically not associated with the activation of caspases or any of the morphological markers of apoptosis or autophagy (Elmore 2007; Tait *et al.* 2014). Regulated necrosis (necroptosis) on the other hand is a mode of cell death that involves death receptor signalling, the necrotic morphological features of death, but the inhibition of caspases (Vandenabeele *et al.* 2010; Ashkenazi and Salvesen 2014; Tait *et al.* 2014).

1.3.3   The Inhibitor of Apoptosis Family of Proteins

The inhibitor of apoptosis gene was first discovered in baculovirus infected *Spodoptera frugiperdato* insect cells to inhibit apoptosis (Crook *et al.* 1993) and homologues have since been identified in yeast, nematodes through to higher vertebrates. The inhibitor of apoptosis protein is characterised by the presence of one to three baculovirus IAP repeats (BIR) motifs that are found in the N-terminal region of the protein. Other functional regions that can be found in IAP proteins are the ~40 amino acid really interesting new gene (RING) domain or the caspase-associated recruitment domain (CARD).

There are eight IAP protein in humans with the smallest, survivin, only containing one BIR domain, while cellular IAP1 (cIAP1) and cIAP2 contain all of the mentioned domains (Figure 1.3.2). The neuronal apoptosis-inhibitory protein (NAIP) contains a unique nucleotide binding oligomerization domain (NOD) and a C-terminal leucine-rich repeat (LRR) domain, while the very large Apollon protein has one BIR domain and the C-terminal ubiquitin conjugating enzyme (UBC) domain. Since all of the proteins contain at least one BIR domain, they are also referred to as baculoviral IAP repeat containing proteins (BIRCs).
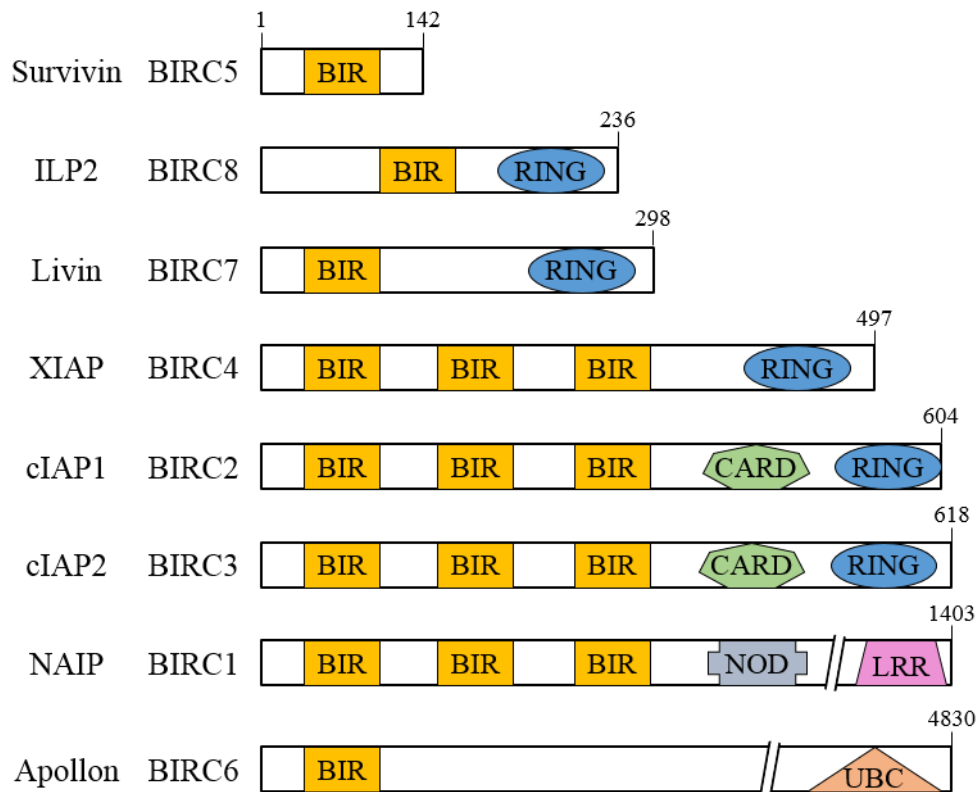
Figure 1.3.2: The Human Inhibitor of Apoptosis Proteins
Diagram showing the various domains of the eight human IAP proteins. BIR = baculovirus IAP repeat; CARD = caspase-associated recruitment domain; LRR = leucine-rich repeat; NOD = nucleotide binding oligomerization domain; RING = really interesting new gene; UBC = ubiquitin conjugating enzyme. Numbers = amino acid number. Adapted from Srinivasula and Ashwell (2008).

The BIR domain is a 70-80 amino acid region that contains the signature $CX_2CX_{16}HX_6C$ sequence, although the better defined $CX_2CX_6WX_3DX_5HX_6C$ pattern can often apply (C = cysteine; D = aspartic acid; H = histidine; W = tryptophan; X = any amino acid; Figure 1.3.3; Srinivasula and Ashwell 2008; Garg *et al.* 2016). The tertiary structure of this domain are a three-stranded β-sheet surrounded by a series of α-helices (Hinds *et al.* 1999). These structures form a hydrophobic core that allows for the coordination of a zinc ion by the three cysteines and one histidine (Figure 1.3.3 and Figure 1.3.4).
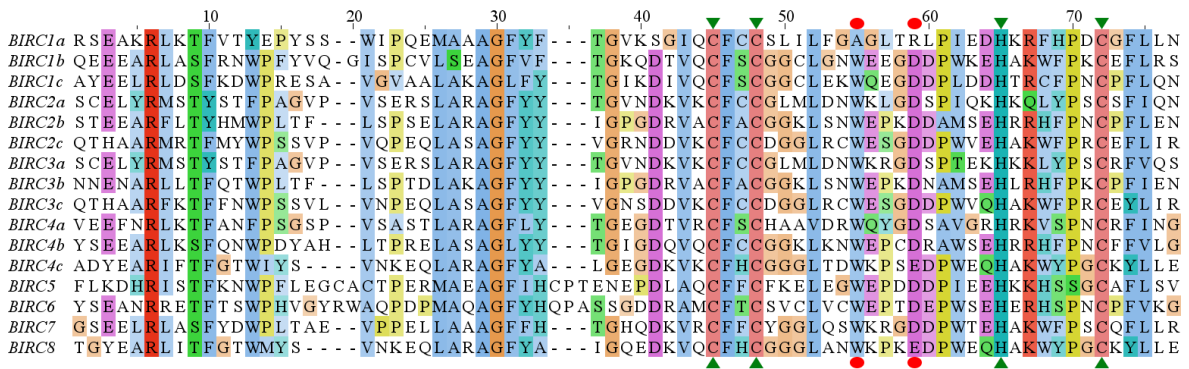
Figure 1.3.3: The Human BIR Domains

Amino acid alignment of the BIR domains of all of the human IAP proteins. Green triangles = zinc-chelating residues; red circles = additional zinc-binding residues. Amino acids are shaded using the Clustal X colour scheme.
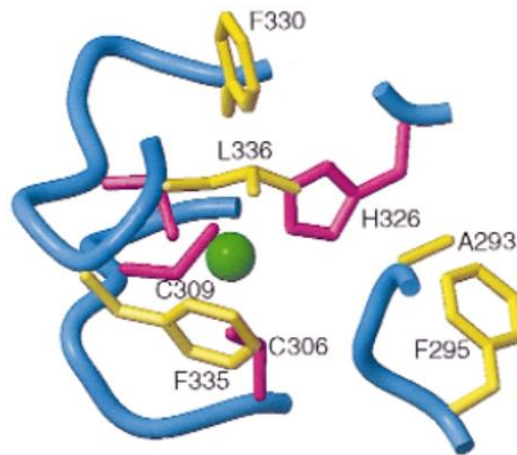
Figure 1.3.4: A Typical BIR Domain Zinc-Binding Fold (Hinds *et al.* 1999)

Visualisation of the cIAP1 BIR3 hydrophobic pocket used to chelate Zn^{2+}. Pink side chains = zinc-chelating residues (green triangles in Figure 1.3.3); yellow side chains = surrounding amino acids. Green ball = Zn^{2+} ion.

The binding of IAP proteins directly to caspases leads to either the inactivation of the caspase or caspase degradation and the presence of multiple BIR domains in each protein increases the range of partners that IAP can bind to (Srinivasula and Ashwell 2008). Human X-linked IAP (XIAP), along with cIAP1 and cIAP2, inhibit caspases-3, -7 and -9 but not caspase-8, via their BIR2 and BIR3 domains. Despite the significant amino acid sequence similarities between the BIR domains, there are structural differences between XIAP and cIAP1/2 as the latter show a weaker binding affinity to the caspases than the former (Deveraux *et al.* 1998). Subsequent work has shown that although the cIAPs contain the scaffold for caspase-binding, they either never gained or have since lost the specific residues required for caspase inhibition (Eckelman *et al.* 2006). XIAP is now regarded as the only mammalian IAP that directly inhibits caspase activity, while the mechanism by which the other IAP proteins

attenuate apoptosis could involve binding to pro-apoptosis proteins and targeting them for ubiquitin degradation (see below; Eckelman *et al.* 2006).

On the pro-apoptotic side are the mitochondrial proteins HtrA2/Omi and SMAC/DIABLO. These proteins are synthesised as precursors and trafficked to the mitochondria, where a tetrapeptide IAP-binding motif (IBM) is exposed. When apoptosis is triggered these pro-apoptotic proteins are released into the cytosol, where they use their IBM motifs to bind to IAP BIR2 and BIR3 domains, liberating the bound caspases (Kocab and Duckett 2016). The hydrophobic, BIR domain IBM-binding groove interacts with caspases or IBMs in a mutually exclusive manner – SMAC IBM peptides block the binding of caspase-9 to BIR3 (Srinivasula *et al.* 2001).

Another domain that most of the IAP proteins have is the RING domain and so IAP proteins are also classed as ubiquitin ligases. The 8 kDa ubiquitin is activated by the ubiquitin-activating enzyme (E1), transferred to the ubiquitin-conjugating enzyme (E2) and then promoted to move to lysine residues on the target protein by the ubiquitin-protein ligase (E3; Figure 1.3.5). This ubiquitination process results in a protein with a lysine-linked ubiquitin chain (the ubiquitin molecules bind to each other at K48), which is the signal for the protein to be degraded by a proteasome. Substrates of this E3 activity range from apoptotic and signalling molecules to IAP proteins themselves in an autoubiquitination fashion (Feltham *et al.* 2012).
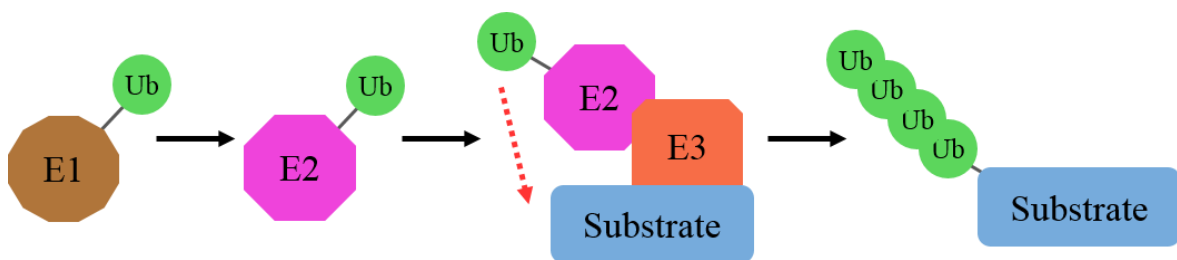
Figure 1.3.5: The Ubiquitin Pathway
Diagram showing the polyubiquitination of a protein substrate. E1 = ubiquitin-activating enzyme; E2 = ubiquitin-conjugating enzyme; E3 = ubiquitin-protein ligase; Ub = ubiquitin. Adapted from Feltham *et al.* (2012).

1.4 Regulated Cell Death in Unicellular Organisms

1.4.1 Unicellular RCD

For a multicellular organism to have a way of removing unwanted or damaged cells (via RCD mechanisms) is of clear evolutionary benefit. This same benefit is not as obvious in the context of a unicellular organism – using cellular machinery to kill the cell, results in the cell's death. This phenomenon does not agree with Darwinian ideas in which an organism will evolve ways to maximise its proliferation, not its death (Pollitt *et al.* 2010). Questions arise about not only how such processes would evolve and persist through generations, but why a unicellular organism would be willing to commit suicide? With respect to the initial part of the question, the apoptosis pathway is conserved in metazoans and many of the constituent proteins have been shown to also have non-death related functions. The hypothesis has been put forward that these cell death proteins evolved their function from proteins involved in normal cellular activities (Dick and Megeney 2013).

Multiple apoptotic markers have been seen in unicellular organisms such as *Saccharomyces cerevisiae*, many of whose apoptotic elements are similar to those in mammals (Deponte and Becker 2004; Kaczanowski *et al.* 2011). This list has grown to include parasitic protozoan organisms like *Leishmania* or *Trypanosoma* that show mammalian apoptosis characteristics (Kaczanowski *et al.* 2011; Proto *et al.* 2012).

The exponential *P. falciparum* asexual reproduction could result in premature host death during the two weeks required for gametocytogenesis to be completed. To counter this the parasite could limit its population in the host by reducing the number of merozoites produced each cycle, impairing the ability of merozoites to invade new RBCs, produce more gametocytes (decreasing the number of asexual participants) or actively reduce its population through cell death (Deponte and Becker 2004; Reece *et al.* 2009). There is evidence that *P. falciparum* exhibits density dependent processes for limiting uncontrolled growth of the parasites, akin to quorum sensing (Mutai and Waitumbi 2010). The use of cell suicide as a means of population control to ensure transmission to the next host could be altruistic, for the benefit of the colonial parasite line, perhaps undertaken by parasites with an inferior phenotype (Reece *et al.* 2011). The use of paracrine signals to activate cell death pathways in neighbours by commandeering proteins with important cellular roles is a means of population control employed by *Leishmania* and *Trypanosoma* (Dick and Megeney 2013).

1.4.2   The Phenotype of RCD in *P. falciparum*

Since Picot *et al.* (1997) linked the well-known "crisis form" morphological phenomenon of *P. falciparum* to apoptosis, various studies tried to definitively classify the cell death phenotype of *P. falciparum*, with mixed results. This work has been performed by looking at metazoan biochemical and morphological markers of cell death such as mitochondrial dysregulation, PS externalization on the surface of the host RBC, DNA laddering and cytoplasmic vacuolization. From these studies there is evidence that RCD does occur in *P. falciparum*, but the reported cell death phenotypes vary from apoptosis to necrosis or from apoptosis-like to autophagy-like. A cause of these discrepancies could be the technical diversity across these various studies (Engelbrecht *et al.* 2012). Another reason could be the fact that metazoan attributes are being enforced onto a unicellular organism that may go about RCD in a unique way. That said, until a definitive phenotype (unique or not) can be elucidated, comparing observations against multicellular RCD characteristics is a good starting point (Engelbrecht *et al.* 2012).

Depending on a range of different factors, various RCD phenotypes have been observed. Specific evidence for the occurrence of RCD in *P. falciparum*, with particular focus on the intraerythrocytic stages of the parasite's life cycle, was reviewed by Engelbrecht *et al.* (2012). Subsequent work includes the observation that an apoptosis-like cell death occurs in asexual *P. falciparum* cultures while under proteasome inhibition (Rathore *et al.* 2015). Ring stage *P. falciparum* parasites also exhibit an apoptosis-like cell death phenotype under heat stress, but late stage parasites exhibit an autophagy-like phenotype (Engelbrecht and Coetzer 2013). High parasitaemia in *P. falciparum* cultures resulted in a growth reduction, but the eventual recovery of the parasites. This decline in parasitaemia was brought about via an apoptosis-like phenotype of RCD with the collapse of the mitochondrial transmembrane potential and PS externalisation, followed by DNA fragmentation (Engelbrecht and Coetzer 2016). On the other hand exposure of *in vitro P. falciparum* cultures to natural sunlight resulted in the death of a portion of the parasite population (particularly the late stage parasites) and the induction of a unique form of RCD in which DNA fragmentation preceded mitochondrial hyperpolarization (Engelbrecht and Coetzer 2015). Haematological complications for the host during *P. falciparum* infection can be partially attributed to the fact that apoptosis can be induced in uninfected RBCs by plasma from *P. falciparum* (but not *P. vivax*) infected individuals (Totino *et al.* 2014; also noted by Engelbrecht and Coetzer 2016).

1.4.3 *P. falciparum* RCD Homologues

In spite of the observed RCD biochemical and morphological markers, to date no cell death protein network has been experimentally validated in *P. falciparum*. After looking for RCD genes within the *P. falciparum* genome using bioinformatics, such as a novel evolutionary rate-based alignment algorithm, some putative plasmodial homologues of metazoan proteins involved in RCD have been identified, including: ataxia-telangiectasia mutated, IAP, metacaspase (MCA1 - PF13_0289) and p53 (Wu *et al.* 2003; Nedelcu 2009; Coetzer *et al.* 2010; Figure 1.4.1). Laboratory experiments are now needed to investigate whether these proteins function like their metazoan counterparts, have unique *Plasmodium* roles related to RCD or are only involved in normal cellular functions.
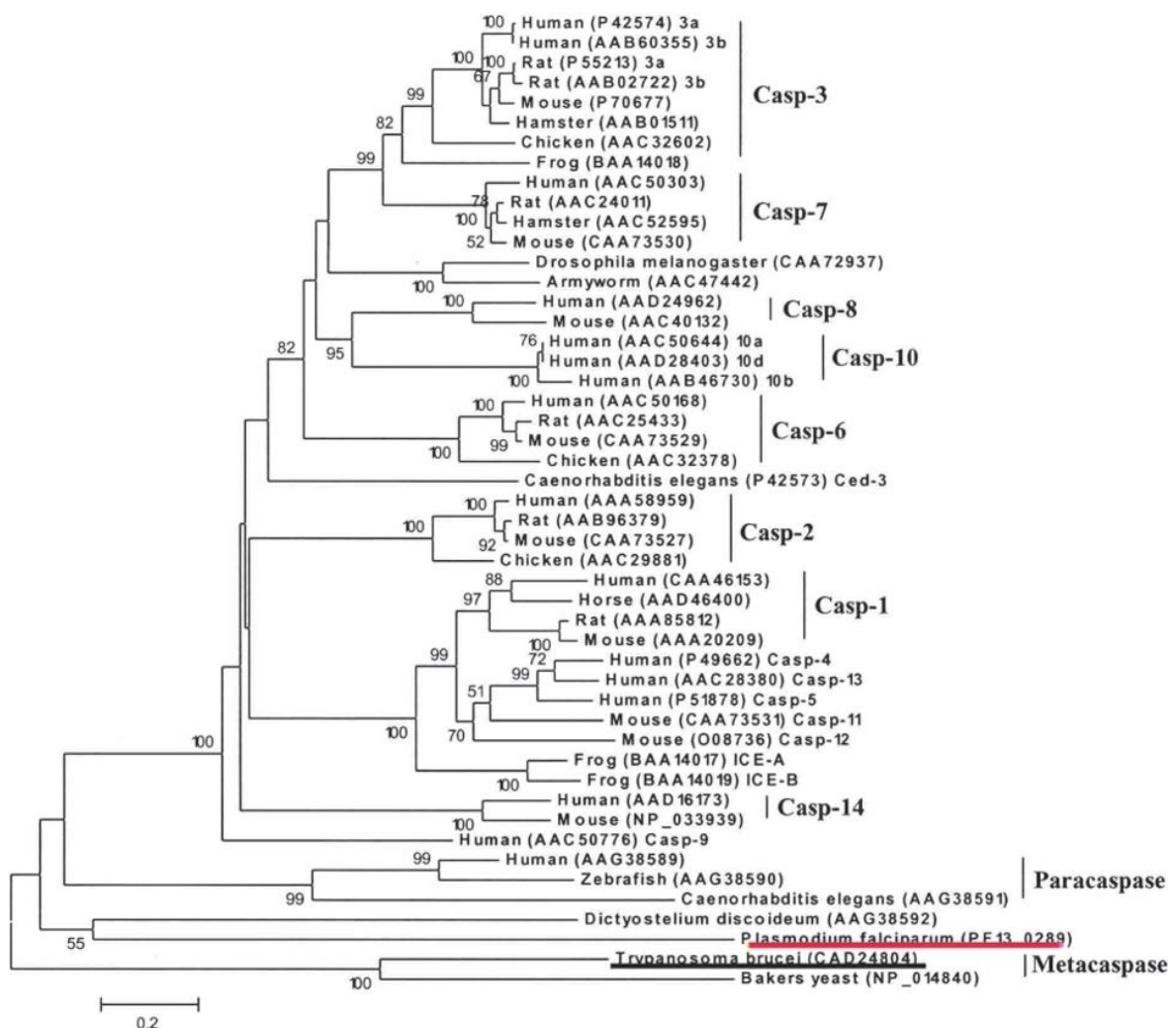
Figure 1.4.1: The Phylogenetic Tree of Various Caspase Proteins (Wu *et al.* 2003)
Phylogenetic tree of caspases, inferred by the neighbour-joining method based on the amino acid sequences with Poisson corrected distance. Bootstrap values of >50% are presented. Scale bar indicates the number of amino acid substitutions per site. Protozoan sequences are underlined.

While no caspases are present in *Plasmodium*, three related metacaspase proteins (MCA1-3) have been described in *P. falciparum*, *P. vivax* and *P. berghei*. MCA1 is the only protein that contains the required histidine and cysteine catalytic dyad (Meslin *et al.* 2011b). *Plasmodium falciparum* metacaspase expression was observed in all intraerythrocytic life stages, but its level was dependent on parasite density (Mutai and Waitumbi 2010). PfMCA1 has been shown to exhibit calcium-dependent, arginine-specific protease activity and induces yeast cell death while under oxidative stress, when expressed in *S. cerevisae* (Meslin *et al.* 2011a). The prodomain of PfMCA1 is removed through an autoprocessing mechanism when examined *in vitro* and this could be an important part of the protein's *in vivo* cell death functioning (Meslin *et al.* 2007). PbMCA1 (expressed in female gametocytes and the mosquito stages, but not in the asexual stages) was knocked out with no apparent loss-of-function phenotype, possibly because PbMCA2/3 offer some functional redundancy (LeChat *et al.* 2007).

The nuclear, anti-apoptotic mammalian mouse double minute 2 homolog (MDM2) contains, amongst others, a SWI/SNF complex B (SWIB)/MDM2 domain. This domain is used to bind to the N-terminal region of p53 (preventing p53 from binding to DNA), induce nuclear export and p53 degradation. The SWIB/MDM2 domain has also been identified elsewhere, for example in the SWItch/Sucrose Non-Fermentable (SWI/SNF) nucleosome remodelling complex. Two proteins containing SWIB/MDM2 domains have been identified in *P. falciparum* designated PfMDM2 and PfSWIB. PfMDM2 (PlasmoDB ID: PF3D7_0518200) localized to the mitochondria under normal and heat stress conditions and so is hypothesised to be involved in mitochondrial maintenance after not showing any RCD activity (Vieira and Coetzer 2016). PfSWIB (PF3D7_0611400) translocated from the cytoplasm to the nucleus in some late stage parasites in response to heat stress and so it is postulated that PfSWIB has a pro-survival function (Vieira and Coetzer 2016).

Analysis of a *P. falciparum* p53 homologue identified a putative DNA binding and a tetramerization domain within the protein. Tetramerization experiments using recombinant Pfp53 though were inconclusive and the protein failed to bind a canonical p53 DNA binding sequence (Vieira 2015).

In light of the recent nomenclature changes (Galluzzi *et al.* 2015), there is now indecision as to whether *Plasmodium* cell death is regulated or accidental (Sow *et al.* 2015). The academic need to classify cellular processes according to criteria, which can easy evolve with the increase of available knowledge, can be viewed as an unnecessary preoccupation in light of increasing *P. falciparum* artemisinin drug resistance (Sow *et al.* 2015). However, it could be well worth it to utilise the available information about *Plasmodium* RCD in the design of new drugs to kill the malaria parasite (Sow *et al.* 2015).

1.5 Study Objectives

With the ultimate goal of presenting a new cell death-related target to which an anti-malarial drug could be designed, this study aimed to characterise a putative *P. falciparum* inhibitor of apoptosis (PfIAP) protein (PF3D7_0519600) and investigate its role in RCD.

The objectives of this study were to:

- Analyse the PfIAP protein using available bioinformatic tools

- Express and purify various PfIAP domains as recombinant proteins

- Create a biotin-tagged *P. falciparum* cDNA phage display library

- Identify binding partners for the various PfIAP domains through biopanning experiments

- Verify these interactions by performing *in vitro* binding assays

- Determine the cellular localisation of the PfIAP BIR domain using transgenic parasites expressing a GFP tag

- Assess the effect that RCD stimulation has on the localisation of the GFP-tagged BIR domain

- Determine if the *pfiap* gene is essential for *P. falciparum* parasite survival through knockout experiments.

CHAPTER 2: The Characterisation of PfIAP

2.1 Introduction

2.1.1 Protein-Protein Interactions

The interactome of a cell represents all of the molecular interactions that occur in the cell. While this can refer to interactions between genes, it primarily refers to the physical interactions between molecules such as proteins. Characterising the protein-protein interactions (PPIs) of a cell is an important step in the understanding of a cell's biology. Although some of these proteins can operate in relative isolation, most proteins (over 80%) are predicted to work with others in complexes, with many PPIs forming part of the networks that are responsible for cellular structure or function (Berggård *et al.* 2007). These functions can range from cell-cycle control to signalling or transport. PPIs can be characterised in a number of ways by contrasting their structural and functional characteristics, for example as either being stable or transient. Stable interactions are typically associated with proteins that make up multi-subunit complexes, while transient interactions are used to control cellular processes in a temporary manner when certain conditions are met. The function of an unknown protein can be implied through PPI studies, where the unknown protein is associated with proteins with known functions.

2.1.1.1 Methods to Detect Protein-Protein Interactions

Methods to detect PPIs can be classified as either being *in vivo*, *in vitro* or *in silico*. *In vivo* techniques are performed using a whole living organism itself, while *in vitro* systems occur outside a living organism in a controlled environment. *In silico* predictions are made via computer simulations or through computer analysis that has been developed to support some of the experimental approaches (Rao *et al.* 2014).

The yeast two-hybrid method (Fields and Song 1989) is a classic *in vivo* example and allows for the direct recognition of an interaction between two proteins. A transcription factor is separated into two domains: a DNA-binding domain (BD) and an activation domain (AD). Both domains are required to be in close proximity to each other for the transcription factor to be able to activate a downstream reporter gene. The bait protein of interest is joined to the BD, while various prey proteins are fused to the AD. Only if the bait and prey proteins interact with each other will the transcription factor be able to activate the reporter gene. The yeast two-hybrid technique has been used to identify novel PPIs in the *P. falciparum* parasite (LaCount *et al.* 2005).

There are a number of different approaches that can be taken when performing *in silico* PPI predictions. A structure-based method for example uses the structure (primary, secondary or tertiary) of the proteins involved to make predictions based on the known interactions of other proteins with similar structures. The gene expression approach on the other hand uses the idea that proteins, whose genes belong to the same expression cluster, are more likely to interact with each other than with proteins from another gene cluster (Rao *et al.* 2014). Other methods of PPI detection include: affinity chromatography, co-immunoprecipitation, protein microarrays and phage display. An attractive feature of using co-immunoprecipitation in live cells to identify a PPI is that the proteins are present in their native form and so should reflect natural interactions as the protein complex is captured from whole cell extract (Berggård *et al.* 2007). A protein array involves fixing a bait protein to a support surface, applying different prey proteins and determining which protein pair forms an interaction (Rao *et al.* 2014). Phage display was chosen for this study and is discussed in more detail below.

2.1.2   Phage Display

2.1.2.1   Phage Display Technology

From the first use of the filamentous bacteriophage f1 (Smith 1985), the PPI investigative technique of using phage display libraries has developed and diversified. Phage display technology involves incorporating foreign DNA into a gene encoding one of the coat proteins of the phage. Infection of a bacterial host with these modified phage results in these fusion peptides being expressed on the bacteriophage capsid surface. After exposing this library of peptides to a bait protein (a process termed biopanning), any resulting PPIs can be identified. This phenotype-genotype link is the power of phage display technology: the displayed peptide (phenotype) can be directly linked to the DNA sequence responsible for the peptide (genotype; Aghebati-Maleki *et al.* 2016).

Phage display systems can be classed as being either non-lytic or lytic. Vectors derived from filamentous phage (such as M13 or fd) are referred to as non-lytic. A frequently used technique is to present the peptide coupled to the minor coat protein pIII of M13 phage, although other capsid proteins (e.g. pVI and pVIII) are also used (Li and Caberoy 2010; Georgieva and Konthur 2011). Another approach involves the use of phagemids which are plasmid vectors that encode the fusion coat protein and phage packaging signal. A helper phage is needed to provide the wild type coat protein and the other genes required for the

assembly of the phage (Pande *et al.* 2010). While the M13 non-lytic phage display system is easy to manipulate there is a major drawback – the displayed peptide needs to be secreted through the *Escherichia coli* membrane which can result in bias towards peptides with a signal sequence, compromising the successful display of the desired proteins (Li 2012; Sundell and Ivarsson 2014). This is not a problem for lambda, T7 or other lytic phage as no protein secretion is required, instead the bacteriophage assemble within the cytoplasm of the bacteria and simply rupture the cell after replication (Danner and Belasco 2001), making these phage systems particularly suitable for the unbiased display of mammalian proteins (Li 2012). When comparing the diversity of 12-mer and disulphide-constrained 7-mer peptides displayed by both M13 and T7 phage display libraries, the T7 system outperformed the M13 system across numerous metrics (Krumpe *et al.* 2006).

T7 phage have numerous properties that make them attractive for use as a phage display vector including the fact that the bacteriophage are very easy to grow and that they replicate more rapidly than either filamentous or lambda bacteriophage (Rosenberg *et al.* 1996). The linear, double-stranded DNA T7 phage has an icosahedral capsid shell made up of 415 copies of the capsid protein (gene 10), attached to a conical tail (genes 11 and 12) and 6 tail fibres (gene 17) via the head-tail connector (gene 8) (Figure 2.1.1a; Rosenberg *et al.* 1996). The phage capsid can be composed entirely of protein 10A (344 amino acids; aa) or protein 10B (397 aa) or different ratios of these proteins, although capsid protein 10B (produced by a translational frameshift at amino acid 341) normally makes up only 10% of the capsid (Condron *et al.* 1991). The implication of this mosaic composition is that the T7 capsid is flexible and that external amino acids can be inserted into these capsid proteins. This fact has been commercialised by Novagen and released as the T7Select® phage display system. In the T7Select®10-3b vector the natural frameshift site has been removed and multiple cloning sites have been inserted, allowing for the cloning of the peptides to be displayed (Figure 2.1.1b). Only a single form of the capsid protein is now made, but because the size or sequence of the displayed peptide could negatively affect viral assembly, the bacteriophage are amplified in *E. coli* cells that contain a plasmid expressing multiple copies of the 10A protein (Rosenberg *et al.* 1996).
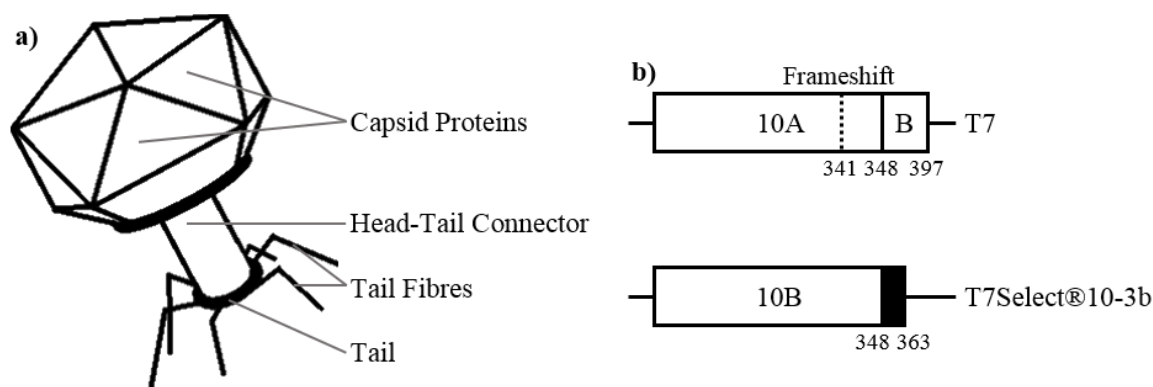
Figure 2.1.1: The T7 Bacteriophage used for Phage Display
a) Schematic representation of a T7 bacteriophage.
b) Diagram showing the T7Select®10-3b vector. The frameshift at amino acid 341, responsible for the 10B capsid protein, has been removed and a multiple cloning site has been included after amino acid 348 to allow for the insertion of foreign DNA. Numbers = amino acid number. Adapted from Rosenberg *et al.* (1996).

An unfortunate consequence of creating a phage display library from cDNA is that there is no guarantee that the sequence will be inserted into the phage vector in-frame. As an example, less than 10% of clones sequenced from a cDNA T7 phage display library were in-frame (Kalniņa *et al.* 2008). The effect of non-open reading frames is that they encode unnatural peptides which have no relevance to the real protein network being investigated (Li and Caberoy 2010). An advantage of using the T7Select® System to create the phage display library is that directional linkers are ligated to the cDNA prior to bacteriophage assembly. These linkers ensure that the inserts can only be introduced into the vector in one of three frames, as opposed to one of six frames. In an effort to further improve the likelihood that only in-frame peptides are expressed, various strategies have been proposed. In an approach that uses phagemids, the cDNA inserts are followed by a β-lactamase gene, flanked by two homologous *lox* sites. After ampicillin screening (only the in-frame inserts would result in antibiotic resistance), the β-lactamase gene was then removed using the Cre recombinase enzyme (Zacchi *et al.* 2003). A similar method involves the preselection of open reading frames using a C-terminal β-lactamase gene in a plasmid (Faix *et al.* 2004). It is necessary to remove the β-lactamase gene as it reduces the efficiency of foreign peptide display (Zacchi *et al.* 2003; Faix *et al.* 2004). A different approach involves tailing the cDNA insert with a biotinylation sequence which results in the expression of a C-terminal biotin tag if the insert is in-frame. Selection of these in-frame phage is then performed using immobilised streptavidin (Ansuini *et al.* 2002; Caberoy *et al.* 2009).

Phage display can be utilised for different purposes, using various sources of display peptides. In enzymology, phage libraries can determine substrate specificity or design modulators to the active site of an enzyme. Identifying the epitope of an antigen that binds an antibody of interest can be rapidly achieved using phage display. Random peptide libraries or peptides derived from cDNA can be used to identify novel binding partners of a protein of interest (Pande *et al.* 2010).

2.1.2.2 Phage Display in Malaria Research

Phage display is an effective tool that has been used in different ways to investigate the *Plasmodium* parasite (reviewed by Lanzillotti and Coetzer 2008). Investigations into host-parasite interactions using a *P. falciparum* cDNA phage display library were first performed in the Coetzer laboratory (Lauterbach *et al.* 2003) and identified thirteen *P. falciparum* proteins that bound to human erythrocyte membrane protein spectrin and protein 4.1 (Lauterbach *et al.* 2003; Lanzillotti and Coetzer 2004). Other *P. falciparum* cDNA phage display libraries have identified a novel interaction between the merozoite surface protein 1 (MSP1) and the erythrocyte glycophorin A (GPA) protein (Baldwin *et al.* 2015), a GPA-interacting peptide that partially inhibited merozoite invasion (Li *et al.* 2008) and the region within the erythrocyte-binding ligand protein that recognised glycophorin B (GPB) and partially inhibited merozoite invasion (Li *et al.* 2012). A *P. falciparum* cDNA phage display library has also been used in basic parasite biology studies to identify binding partners for both the *P. falciparum* MDM2 and SWIB proteins (Vieira and Coetzer 2016) and identify two novel apical membrane antigen-1 (AMA1) interacting proteins (Churchyard 2017).

Recent examples of other phage display work include the use of random peptide libraries to identify a 12-mer peptide that inhibited *P. berghei* sporozoite entry into Kupffer cell (Cha *et al.* 2015) and the female gamete peptide 1 which bound to the surface of female *P. berghei* gametes and interfered with oocyst formation (Vega-Rodriguez *et al.* 2015). Work on a *Plasmodium* vaccine has also utilised a peptide library to identify two *P. knowlesi* MSP1 epitopes that generated an immune response when used to immunise mice (Cheong *et al.* 2016), while an antibody phage display library was used to identify two unique human single chain antibody clones that bound recombinant *P. falciparum* HRP2 and could be used in the development of the next generation of RDTs (Leow *et al.* 2014).

Phage display libraries have proven to be a valuable tool in malaria research and so phage display technology was used in this study to identify novel binding partners of a putative PfIAP protein.

2.1.3   Objectives

The objectives of this chapter were to:

- Utilise bioinformatics tools to analyse a *P. falciparum* IAP protein and identify domains

- Express and purify selected PfIAP domains as recombinant proteins

- Create a biotin-tagged *P. falciparum* cDNA phage display library

- Identify PfIAP binding partners through biopanning experiments

- Express and purify these binding partners as recombinant proteins

- Verify these protein-protein interactions by employing *in vitro* binding assays

- Create a transgenic parasite line expressing a GFP-tag PfIAP BIR domain

- Determine the cellular localisation of PfIAPbir-GFP under normal and RCD simulating conditions using fluorescence microscopy

2.2   Materials and Methods

2.2.1   Bioinformatics

2.2.1.1   Gene Identification and Sequence Alignment

Using novel bioinformatics methods a putative *P. falciparum* inhibitor of apoptosis has been identified (Coetzer *et al.* 2010).   The *Plasmodium* database, PlasmoDB v34 (http://plasmodb.org/plasmo/; Aurrecoechea *et al.* 2009), lists this PfIAP protein as a putative zinc finger protein with the identifier PF3D7_0519600. Nucleotide and amino acid sequences of PfIAP and other *Plasmodium* proteins were obtained from PlasmoDB, while the UniProt protein database (http://www.uniprot.org/) was used to obtain all eukaryotic sequences.   Using the OrthoMCL database v5 (http://orthomcl.org/orthomcl/home.do; Li *et al.* 2003) *Plasmodium* orthologues within PlasmoDB were identified. The Simple Modular Architecture Research Tool (SMART; http://smart.embl.de/; Letunic *et al.* 2015) was used to identify the domains present in these proteins, while sequence alignments were performed using the European Molecular Biology Laboratory European Bioinformatics Institute MUltiple Sequence Comparison by Log-Expectation program (EMBL-EBI MUSCLE; http://www.ebi.ac.uk/Tools/msa/muscle/; Edgar 2004a; Edgar 2004b). Multiple sequence alignments were visually represented using Jalview v2.10 (Waterhouse *et al.*

2009).   Percentage identity and similarity was calculated using EMBOSS Needle (http://www.ebi.ac.uk/Tools/psa/emboss_needle/).

2.2.1.2   Analysis of Protein Structure

Secondary Structure. The secondary structure of the PfIAP protein was assessed using SWISS-MODEL (https://swissmodel.expasy.org/; Arnold *et al.* 2006; Bordoli *et al.* 2009; Biasini *et al.* 2014), the Protein Homology/analogY Recognition Engine v2.0 (Phyre2; http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index; Kelley *et al.* 2015), PSIPRED v3.3 (http://bioinf.cs.ucl.ac.uk/psipred/; Jones 1999; Buchan *et al.* 2013), JPred4 (http://www.compbio.dundee.ac.uk/jpred4; Drozdetskiy *et al.* 2015) and SSpro v5.2 (http://scratch.proteomics.ics.uci.edu/; Cheng *et al.* 2005). The probability of the presence of a particular secondary structure (α-helix or β-strand) within the PfIAP BIR domain was determined by comparing the various secondary structural predictions.

Tertiary Structure. Three-dimensional modelling of the PfIAP BIR domain was carried out using SWISS-MODEL (Arnold *et al.* 2006; Bordoli *et al.* 2009; Biasini *et al.* 2014) and Phyre2 (Kelley *et al.* 2015). Crystal structures of various human IAP domains were obtained from the Protein Data Bank (http://www.rcsb.org/). The quality of the generated 3D models was estimated using QMEAN (https://swissmodel.expasy.org/qmean/; Benkert *et al.* 2009; Benkert *et al.* 2011). Visualisation and colouring of the various 3D models was done using Jmol 14.8.0 (http://www.jmol.org/).

Protein Properties and Solubility Predictions. The isoelectric point (pI) and molecular weight of various *P. falciparum* proteins was calculated using the ExPASy compute pI/Mw tool (http://web.expasy.org/compute_pi/; Bjellqvist *et al.* 1993; Bjellqvist *et al.* 1994; Gasteiger *et al.* 2005), while Protein Calculator v3.4 (http://protcalc.sourceforge.net/) was used to determine the overall charge of proteins. The predicated solubility of the *P. falciparum* proteins that were expressed in as recombinant proteins was determined using PROSO   II   (http://mbiljj45.bio.med.uni-muenchen.de:8888/prosoII/prosoII.seam; Smialowski *et al.* 2012) and SOLpro (http://scratch.proteomics.ics.uci.edu/; Magnan *et al.* 2009).

2.2.2   *P. falciparum* Parasite Culture

All parasite culture work was carried out in a laminar flow hood following aseptic techniques to maintain sterility. All glassware was autoclaved, while culture medium and solutions were filter sterilised prior to use (0.8/0.2 μm Acrodisc® PF or Vacucap® 90 PF filter units). Chemical and equipment suppliers are available in APPENDIX E.

2.2.2.1   Preparation of Red Blood Cells

Fresh whole blood was collected from human volunteers in 6 mL acid citrate dextrose tubes. These tubes were centrifuged at 2 500 ×g for 10 minutes at 4°C (Eppendorf centrifuge 5702 R) before the plasma and buffy coat were aspirated. The packed RBCs were washed using phosphate buffered saline (PBS; 10 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl; pH 7.4). The RBC pellets were resuspended in a 1:1 ratio with incomplete RPMI-1640 medium (containing 25 mM HEPES, 0.05 mg/mL hypoxanthine and 0.02 mg/mL gentamycin) and stored at 4°C for up to two weeks.

2.2.2.2   Continuous Parasite Culture

Asexual *P. falciparum* 3D7 strain parasites were cultured *in vitro* using a modification of the method described by Trager and Jensen (1976), as serum was replaced with Albumax. Cultures were maintained at a 5% haematocrit in fresh RBCs and complete medium (incomplete medium with 0.5% (w/v) Albumax and 0.2% (w/v) $NaHCO_3$). Three different sized culture flasks (25 cm^3, 80 cm^3 or 175 cm^3) were used containing 5 mL, 20 mL and 35 mL of complete medium, respectively. Cultures were gassed with a mixture of 2% O_2, 5% CO_2 and 93% N_2 for 30 to 60 seconds, before being sealed and incubated at 37°C. Every 24 hours, spent medium was aspirated and replaced with fresh complete medium. The parasitaemia of the culture was assessed daily by a blood smear (stained using the RapiDiff Stain Set as per manufacturer's instructions) that was viewed using a Zeiss Axiostar microscope under oil immersion at 1 000X magnification. Parasitaemia was calculated as the percentage of infected RBCs out of the total number of erythrocytes, and parasite cultures were maintained at a parasitaemia of less than 8%.

2.2.2.3   Freezing of Parasite Cultures

Parasite cultures were prepared and stored in liquid nitrogen as described by Normark (2008), with slight modification of the components of the freezing solution. A medium (20 mL) parasite culture containing greater than 4% ring stage parasites was divided between

two 15 mL centrifuge tubes. After centrifuging at 1 400 ×g for 5 minutes at room temperature, the supernatant was aspirated and the RBC pellet was resuspended 1:1 with freezing solution (60% (v/v) glycerol in freezing PBS (83 mM Na_2HPO_4, 32 mM KH_2PO_4, 123 mM NaCl; pH 7.4)). The suspension was mixed, transferred (~1 mL) to a sterile 2 mL cryotube and allowed to incubate for 5 minutes at room temperature before being stored in liquid nitrogen.

2.2.2.4 Thawing of Frozen Parasite Cultures

Frozen parasite cultures were removed from liquid nitrogen storage and thawed using a modified form of the method described by Blomqvist (2008), with an extra NaCl and glucose wash step. A culture was warmed to 37°C, before one-tenth volume (~100 μL) of 12% (w/v) NaCl was added dropwise. The suspension was incubated for 5 minutes at room temperature, before being transferred to a 15 mL centrifuge tube containing 9 mL of 1.6% (w/v) NaCl. Following gentle mixing, the suspension was centrifuged at 400 ×g for 5 minutes at room temperature. The supernatant was removed and the RBC pellet was resuspended in nine volumes of 0.9% (w/v) NaCl and 0.2% (w/v) glucose. After centrifuging again and removing the supernatant, the volume of the pellet was made up to 750 μL with 1:1 diluted RBCs before being transferred to a small culture flask with 5 mL of complete medium (with 1% (w/v) Albumax for the next 7 days). The culture was then gassed and incubated at 37°C as described in 2.2.2.2.

2.2.2.5 Synchronising Parasite Cultures

Parasite cultures with predominantly ring stage parasites were synchronised by sorbitol treatment (Lambros and Vanderberg 1979). D-sorbitol lyses late stage infected RBCs due to their increased permeability when compared to early stage parasites.

A small (5 mL) parasite culture containing at least 3% ring stage parasites was transferred to a 15 mL centrifuge tube and centrifuged at 1 400 ×g for 5 minutes at room temperature. The supernatant was removed and ten volumes of 5% (w/v) D-sorbitol was added to the tube. After incubating for 5 minutes at 37°C, the tube was gently inverted before being centrifuged again. The supernatant was aspirated and the pellet was resuspended in complete medium with fresh 1:1 RBCs (5 mL culture with a haematocrit of 5%) before being transferred to a small culture flask. The culture was then gassed and incubated at 37°C as described in 2.2.2.2.

2.2.3 *P. falciparum* Genomic DNA Extraction

Genomic DNA was extracted from *P. falciparum* parasite cultures with a greater than 5% parasitaemia and a ratio of approximately 3:2 (trophozoite:rings), according to the hypotonic lysis method (Cowman *et al.* 2008).

A small (5 mL) parasite culture was transferred to a 15 mL centrifuge tube and centrifuged at 1 000 ×g for 5 minutes at 4°C. The supernatant was aspirated and the pellet was resuspended in 5 mL of PBS before being centrifuged again. After removing the supernatant, the pellet was disrupted by the addition of four volumes (~1 mL) of hypotonic solution (5 mM K_2HPO_4/KH_2PO_4; pH 7.4) which lysis the RBCs. One volume (~250 μL) of 18% (w/v) sodium dodecyl sulfate (SDS) was added and the solution was briefly vortexed before being incubated for 3 minutes at room temperature. Eight volumes of 1:1 phenol/chloroform were added to the tube, mixed briefly and centrifuged at 3 000 ×g for 10 minutes at 4°C. The top aqueous phase, containing the genomic DNA, was carefully collected and distributed between 1.5 mL Eppendorf tubes (~400 μL). To precipitate the DNA, one-tenth volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ice cold 100% ethanol were added to each aliquot. These tubes were mixed by inversion and incubated for 30 minutes or overnight at -70°C. The tubes were centrifuged at 16 000 ×g for 30 minutes at 4°C (Eppendorf centrifuge 5415 R) and the supernatant was decanted. One millilitre of cold 70% (v/v) ethanol was added and the tubes were centrifuged at 16 000 ×g for 5 minutes at 4°C. After removing the supernatant, the pellets were allowed to air-dry before they were dissolved and pooled into a final volume of 600 μL of Tris EDTA buffer (TE; 10 mM Tris-HCl, 1 mM EDTA; pH 8). RNase A (final concentration of 0.1 mg/mL) was added and allowed to incubate for 30 minutes at 37°C, before one volume of 1:1 phenol/chloroform was added. After mixing, the tube was centrifuged at 16 000 ×g for 5 minutes at 4°C and the aqueous phase was carefully transferred to a 1.5 mL Eppendorf tube. One volume of chloroform was added to the tube which was mixed and centrifuged again. The aqueous phase was transferred to a 2 mL Eppendorf tube in which the genomic DNA was again precipitated and washed. The air-dried pellet was resuspended in 25 μL of TE buffer and stored at -20°C.

Chapter 2: The Characterisation of PfIAP

2.2.3.1 DNA Analysis

Spectrophotometric Analysis. The concentration and purity of extracted DNA was determined using a NanoDrop® ND-1000 spectrophotometer (with integrated ND-1000 3.8.1 software) by measuring the light absorbed at 230 nm (A_{230}), 260 nm (A_{260}) and 280 nm (A_{280}).

Agarose Gel Electrophoresis. Extracted DNA samples were assessed by agarose gel electrophoresis. A 1% (w/v) agarose gel was prepared by dissolving agarose in 1x Tris acetate EDTA buffer (TAE; 40 mM Tris-acetate, 1 mM EDTA; pH 8) by boiling. After cooling, ethidium bromide (final concentration of 0.5 µg/mL) was added and the gel was cast. The samples, together with a reference MassRuler™ DNA Ladder Mix (80 - 10 000 bp), were loaded and the gel was resolved in 1x TAE buffer at 100V for 50 - 60 minutes. The gel was then visualised under ultraviolet (UV) light using the SynGene Gel Doc system and images were acquired with the GeneSnap 7.04 analysis software. Genomic DNA should resolve as a single band with a molecular weight greater than 10 kb, but will appear as a smear if the DNA is degraded.

2.2.4 Polymerase Chain Reaction (PCR)

2.2.4.1 PCR Primer Design

Plasmodium falciparum gene sequences were obtained from PlasmoDB (Aurrecoechea *et al.* 2009) and used to design PCR primers with the aid of Integrated DNA Technologies OligoAnalyzer 3.1 (Integrated DNA Technologies 2016). Endonuclease restriction sites were included at the 5' end of all primers to facilitate directional cloning into the vectors: pGEX-4T-2 (GE Healthcare, UK), pET-15b (Millipore, USA) and pARL2-GFP (donated by Dr. Jude Przyborski of the University of Marburg, Germany). pGEX-4T-2 and pET-15b reverse primers contained a stop codon prior to the restriction site and all oligonucleotides were synthesised by Inqaba Biotec™, South Africa. Primer sequences are available in APPENDIX A.

2.2.4.2 Polymerase Chain Reaction

To enhance the accuracy of all PCR amplification reactions prior to cloning, high fidelity DNA polymerases were used as directed by the manufacturer. Phusion Flash or High Fidelity PCR Mixes were used with 100 - 200 ng of *P. falciparum* genomic DNA and 0.2 µM of forward/reverse primers. All PCR reactions were carried out using an Eppendorf

Mastercycler Gradient Thermocycler and were divided into two sections: the first five cycles were performed at the lowest annealing temperature (T_a) of the *P. falciparum* specific regions of the primers, while the following 25 cycles were at the T_a of the full length primers. PCR cycling conditions are available in APPENDIX A.

Table 2.2.1: PCR Parameters for the Amplification of Various *P. falciparum* Genes

Domain Name	Vector	*P. falciparum* Specific T_a (°C)	Full Length Primer T_a (°C)	PCR Product (bp)
PfIAP1 (BIR Domain)	pGEX-4T-2	48.0	62.0	471
PfIAP2 (Middle Region)		48.0	62.0	561
PfIAP3 (Terminal Region)		48.0	62.0	984
PfDOC2 (Binding Region)	pET-15b	41.0	55.0	1 011
PfRhopH3 (Binding Region)		44.0	55.0	510
PfIAP1 (BIR Domain)	pARL2-GFP	48.0	55.0	468

PCR amplicons were then purified using either the QIAquick® PCR Purification or NucleoSpin® Gel and PCR Clean-Up Kits as per the manufacturers' instructions and eluted into 50 µL of nuclease-free water.

2.2.5 Isolation of Expression Plasmids

The pGEX-4T-2, pET-15b or pARL2-GFP vectors (available in APPENDIX B) were isolated either using an alkaline lysis method (Bimboim and Doly 1979) or by commercial kits. The GenElute™ Plasmid Miniprep or NucleoSpin® Plasmid Kits were used as per the manufacturers' instructions and plasmids were eluted into 50 µL of nuclease-free water.

A 5 mL Luria Broth (LB; 1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract, 10 mM Tris-HCl; pH 8.0) culture in a 50 mL Erlenmeyer flask, containing 100 µg/mL ampicillin was inoculated with a scraping of a glycerol stock of *E. coli* cells containing the desired plasmid and incubated at 250 rpm for 18 hours at 37°C (Labotec orbital shaker). The optical density (OD) of the bacterial culture was measured using a Thermo Biomate 5 Spectrophotometer at 600nm (OD_{600}) and if the OD was greater than 3.0, then 25 mL of LB medium was seeded with 500 µL of the initial culture (in a 100 mL Erlenmeyer flask, containing 100 µg/mL ampicillin). After incubating for 18 hours at 37°C, the culture was divided between two 15 mL centrifuge tube and centrifuged at 3 000 ×g for 10 minutes at 4°C.

The supernatant was aspirated and 500 µL of cold resuspension solution (50 mM glucose, 10 mM EDTA, 25 mM Tris; pH 8.0) was added to each tube. The pellets were disrupted by repeat pipetting and incubated for 5 minutes at room temperature. One millilitre of lysis solution (0.2 M NaOH, 1% (w/v) SDS) was added to each tube and the suspensions were

mixed by inversion, before being incubated for 5 minutes at room temperature. The SDS and NaOH lyse the bacterial cells and denature the high molecular weight chromosomal DNA (Bimboim and Doly 1979). Cold neutralisation solution (750 µL; 1.6 M potassium acetate and 3.2 M acetic acid; pH 4.8) was added to each tube, which was mixed by inversion and then incubated for 5 minutes on ice. During neutralisation the SDS-protein complexes and the high molecular weight RNA precipitate, while the chromosomal DNA renatures and forms an insoluble aggregate (Bimboim and Doly 1979). The samples were centrifuged at 3 000 ×g for 20 minutes at 4°C and the supernatants were divided between 2.0 mL Eppendorf tubes. One millilitre of 1:1 phenol/chloroform was added to each aliquot and the solutions were mixed by inversion, vortexed and then centrifuged at 16 000 ×g for 5 minutes at 4°C. The aqueous phases were transferred to new 2.0 mL Eppendorf tubes and the plasmid was precipitated as described in 2.2.3. The pellets were allowed to air-dry, before being resuspended in TE buffer and pooled to a final volume of 400 µL.

RNase A (final concentration of 0.1 mg/mL) was added to the sample, which was then incubated for 30 minutes at 37°C. One volume of 1:1 phenol/chloroform was added to the solution, which was mixed and centrifuged at 16 000 ×g for 5 minutes at 4°C. The aqueous phase was transferred to a new 1.5 mL Eppendorf tube and one volume of chloroform was added to the tube, which was mixed and centrifuged again. The aqueous phase was then transferred to a 2 mL Eppendorf tube and the vector was precipitated as described in 2.2.3. The pellet was allowed to air-dry and resuspended in 50 µL of TE buffer.

2.2.6   Cloning

2.2.6.1   Restriction Enzyme Digestion

Purified PCR amplicons and isolated expression vectors were digested for 30 minutes at 37°C using specific FastDigest® restriction endonucleases:

- pGEX-4T-2 vector and inserts - FastDigest® *Bam*HI and *Xho*I

- pET-15b vector and inserts   - FastDigest® *Bam*HI and *Nde*I

- pARL2-GFP vector and inserts - FastDigest® *Avr*II and *Xho*I

The plasmids were also simultaneously subjected to FastAP™ Thermosensitive Alkaline Phosphatase to be dephosphorylated. The treated PCR products and vectors were then purified using commercial kits (see 2.2.4.2) as per the manufacturers' instructions and eluted in 10 µL of nuclease-free water.

2.2.6.2 Ligation Reaction

The prepared PCR product and expression vector were quantified by agarose gel electrophoresis (2.2.3.1). The amount of insert and plasmid required for the ligation reaction was calculated using the In-Fusion® Molar Ratio Calculator (Clontech Laboratories 2016).

The ligation reaction was prepared using the Rapid DNA Ligation Kit as per the manufacturer's specifications, which included not exceeding 200 ng of total DNA per reaction and using a molar ratio of one vector unit to three insert units. The ligation was then carried out by incubating the reaction for 30 minutes at 16°C. To test the effectiveness of the vector dephosphorylation and restriction enzyme digestion, a control ligation reaction was also prepared that did not contain the PCR product.

2.2.6.3 Transformation of DH5α and XL10 *E. coli* Cells

In-house, chemically competent DH5α *E. coli* cells (provided by Dr Sonja Lauterbach) were transformed using half (10 μL) of the pGEX-4T-2 and pET-15b ligation reactions. A 100 μL *E. coli* aliquot was incubated with the ligation reaction for 30 minutes on ice, before being heat-shocked at 42°C for 90 seconds. The transformation mixture was then incubated for 5 minutes on ice, before being spread onto two 1.5% (w/v) agar plates (containing 100 μg/mL ampicillin) and incubated overnight at 37°C. The 'low' plate contained 20 μL of the transformation mixture, made to 80 μL with LB medium, while the remaining mixture was applied to the other 'high' plate.

Following ligation, the pARL2-GFP_PfIAPbir construct was used to transform commercial XL10 *E. coli* as per the manufacturer's instructions, with slight modification as NZY$^+$ broth was replaced with LB. Briefly, a 50 μL aliquot of XL10 cells was thawed on ice and then incubated with 2 μL of β-mercaptoethanol for 10 minutes. Ten microlitres of the ligation reaction was added and the cells were incubated for 30 minutes on ice. Heat-shock was performed at 42°C for 30 seconds, after which the transformation mixture was incubated for 2 minutes on ice. Four hundred and fifty microlitres of pre-warmed LB (42°C) was added and the cells were incubated for 1 hour at 37°C, before being plated on 1.5% (w/v) agar plates (containing 100 μg/mL ampicillin) and incubated overnight at 37°C.

Individual *E. coli* colonies, transformed with either the pGEX-4T-2, pET-15b or pARL2-GFP vector constructs, were picked from the transformation plates and suspended in 10 μL of nuclease-free water to prepare for an overnight LB grow and for colony PCR (see 2.2.6.4).

Five microlitres of the suspension was used to inoculate 2 mL of LB medium (in a 14 mL BD Falcon™ tube, containing 100 µg/mL ampicillin) and was incubated at 250 rpm for 18 hours at 37°C. Glycerol stocks of the individual colonies were created by adding 500 µL of each overnight grow to 500 µL of sterile 60% (v/v) glycerol and storing at -70°C. The remaining *E. coli* grow was used to perform a plasmid isolation using either the GenElute™ Plasmid Miniprep or NucleoSpin® Plasmid Kits as per the manufacturers' instructions, eluting into 50 µL of nuclease-free water.

2.2.6.4   Verification of Bacterial Transformation

Colony PCR. Five microlitres of *E. coli* suspension (2.2.6.3) was boiled at 94°C for 5 minutes to lyse the bacterial cells. DreamTaq Green PCR Master Mix was used as per the manufacturers' instruction, with the bacterial lysate providing the DNA template, to detect the presence of the inserts. Vector- or insert-specific primer details are available in APPENDIX A.

Restriction Enzyme Digestion. The previously isolated plasmid constructs were digested for 30 minutes at 37°C using restriction endonucleases:

- pGEX-4T-2 vector - FastDigest® *Bam*HI and *Xho*I
- pET-15b vector   - FastDigest® *Eco*RV
- pARL2-GFP vector - FastDigest® *Eco*RI

The resulting digestion patterns were analysed using agarose gel electrophoresis as described in 2.2.3.1. Vector constructs that demonstrated the expected banding pattern were sent to Inqaba Biotec™, South Africa for sequencing to verify the integrity of the desired insert.

2.2.7   Recombinant Protein Expression

2.2.7.1   Transformation of Rosetta™ 2 (DE3) *E. coli* Cells

Commercial Rosetta™ 2 (DE3) *E. coli* contain the chloramphenicol-resistant pRARE plasmid. This bacterial strain is used to enhance the expression of eukaryotic proteins by providing sevens transfer RNAs (tRNAs; AUA, AGG, AGA, CUA, CCC, GGA and CGG) that are rarely used in *E. coli* (Novagen 2004).

Rosetta™ 2 (DE3) cells were transformed with the pGEX-4T-2 and pET-15b vector constructs as per the manufacturer's instructions. After overnight growth on 1.5% (w/v) agar plates (containing 100 µg/mL ampicillin and 50 µg/mL chloramphenicol), single

colonies were picked and treated as described in 2.2.6.3 to create glycerol stocks. Colony PCR was utilised to verify the transformation of the *E. coli* as described in 2.2.6.4.

2.2.7.2   Recombinant Protein Expression and Purification

High-level protein expression was achieved using Overnight Express™ Instant TB Media, which has been designed for use with isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible expression plasmids. Components in the medium induce protein expression from a lac promoter and are metabolized differentially so as to promote increased cell mass and target protein yield, when compared with conventional IPTG induction protocols (Grabski *et al.* 2005).

A 5 mL LB culture (in a 50 mL Erlenmeyer flask, containing 100 μg/mL ampicillin and 50 μg/mL chloramphenicol) was inoculated with a glycerol scraping of Rosetta™ 2 (DE3) *E. coli* cells containing the desired plasmid construct and incubated at 250 rpm overnight at 37°C. Twenty millilitres of Overnight Express™ Instant TB Media was seeded with 500 μL of the initial culture (in 100 mL Erlenmeyer flasks) and incubated at 250 rpm for 22 hours at room temperature (Hoefer PR250 mini orbital shaker). The OD_{600} of each culture was measured and if greater than 3 the bacterial cells were centrifuged at 3 000 ×g for 10 minutes at 4°C, before being stored at -70°C. *E. coli* cells were also used to seed 20 mL of 2% (w/v) glucose in LB as an uninduced protein expression control. The presence of glucose means that the *lac* operon is not fully induced, preventing the expression of the recombinant proteins (Rosano and Ceccarelli 2014).

GST-Tagged Protein Purification. Glutathione S-transferase (GST) tagged proteins were purified using paramagnetic MagneGST™ Glutathione Particles by applying the following protocol (Promega 2014).

The bacterial pellet containing the pGEX-4T-2 plasmid was thawed for 5 minutes at 37°C before being resuspended in 1.5 mL of GST-Binding/Wash buffer (4.2 mM Na_2HPO_4, 2 mM KH_2PO_4, 140 mM NaCl, 10mM KCl; pH 7.2), containing 1.5μl of Protease Inhibitor Cocktail Set III and 1.5μl of DNase I. The bacterial suspension was frozen for 15 minutes at -70°C, thawed again and then transferred to a 15 mL centrifuge tube. Three cycles of sonication (30 seconds of 1.0 second pulses with 0.5 second rests; Bandelin Sonopuls HD 3100 with microtip MS 73) was used to lyse the bacterial cells, after which 75 μL was removed as the 'total' protein fraction and added to 75 μL of GST-Binding/Wash buffer.

The lysate was transferred to a 2.0 mL Eppendorf tube and centrifuged at 16 000 ×g for 20 minutes at 4°C. The supernatant was set aside (75 µL of which was removed as the 'soluble' protein fraction and added to 75 µL of GST-Binding/Wash buffer) and the pellet was resuspended using 1.5 mL of GST Binding/Wash buffer (75 µL was removed as the 'insoluble' protein fraction and added to 75 µL of GST-Binding/Wash buffer).

Ten microlitres of the MagneGST™ particle slurry was washed three times using 100 µL of the GST-Binding/Wash Buffer and a magnetic separator. The particles were resuspended in 100 µL of the GST-Binding/Wash buffer before being applied to the previously set aside soluble fraction. The mixture was incubated overnight at 4°C on an Intelli-mixer, after which the supernatant was removed (75 µL of which was removed as the 'unbound' protein fraction and added to 75 µL of GST-Binding/Wash buffer). The MagneGST™ particles were washed five times using 1.5 mL of the GST-Binding/Wash buffer (150 µL of each wash was removed as the 'wash' fraction), before the recombinant GST-tagged protein was eluted for 15 minutes at room temperature using 100 µL of the GST-Elution buffer (500 mM L-glutathione, 500 mM NaCl, 50 mM Tris-HCl; pH 8.1). The remaining MagneGST™ particles were resuspended in 100 µL of the GST-Binding/Wash buffer as the 'bead' fraction. All protein fractions were processed and visualised as described in 2.2.7.3 and 2.2.7.4.

His-Tagged Protein Purification. His-tagged proteins were purified using paramagnetic pre-charged MagneHis™ nickel particles (Promega 2009) according to the *GST-Tagged Protein Purification* protocol above with the following differences:

- His-Binding buffer (50 mM Na_2HPO_4/NaH_2PO_4, 150 mM NaCl; pH 8) containing 1.5µl of Protease Inhibitor Cocktail Set III and 1.5µl of DNase I was used to resuspend the thawed bacterial pellet.

- Twenty-five microlitres of MagneHis™ particle slurry was prepared using 100 µL of the His-Binding buffer, to which imidazole was added to a final concentration of 1 mM.

- Beads with bound proteins were washed using 1.5 mL of His-Wash buffer (His-Binding buffer containing: 35 mM imidazole for PfRhopH3 and 50 mM imidazole for PfDOC2) and eluted using 100 µL of His-Elution buffer (His-Binding buffer containing 0.5 M imidazole; pH 7.5).

2.2.7.3 SDS-PAGE Analysis

One volume of sample solution (40 µL of 5X suspension solution (50 mM Tris-HCl, 5 mM EDTA, 5% (w/v) SDS, 25% (w/v) sucrose; pH 8.0), 5 µL of dye mix (2.5% (w/v) sucrose, 0.5% (w/v) bromophenol blue) and 5 µL of β-mercaptoethanol) was added to three volumes of the protein fractions described in 2.2.7.2. These protein mixtures were vortexed, boiled for 5 minutes and then resolved using SDS-polyacrylamide gel electrophoresis (SDS-PAGE; Laemmli 1970).

A 6 cm 12% resolving gel (see Table 2.2.2) was cast using a Hoefer Mighty Small II SE 250 gel cassette and overlaid with a 2 cm 4% stacking gel. After sample loading, the gel was electrophoresed at a constant current of 20 mA (with a maximum voltage of 250 V) in Laemmli running buffer (25 mM Tris, 192 mM Glycine, 0.1% (w/v) SDS) and cooled to 4°C using a Labcon CPE 50 circulator. A human red blood cell membrane reference ladder, prepared in-house by Dr Kuben Naidoo, was also included.

Table 2.2.2: Laemmli SDS-PAGE Gel Set-Up

Component	12% Resolving Gel*	4% Stacking Gel*
30% (w/v) Acrylamide	4.0 mL (12%)	433 µL (4%)
1% (w/v) Bis-acylamide	1.1 mL (0.11%)	333 µL (0.1%)
4X Resolving Buffer (1.5 M Tris-HCl; pH 8.8)	2.5 mL (1X)	-
4X Stacking Buffer (0.5 M Tris-HCl; pH 6.8)	-	833 µL (1X)
10% (w/v) SDS	53.0 µL (0.05%)	6.7 µL (0.02%)
10% (w/v) Ammonium Persulfate	67.0 µL (0.07%)	67.0 µL (0.2%)
TEMED	5.0 µL (0.05%)	2.5 µL (0.08%)
Nuclease-Free Water	2.3 mL	1.63 mL
Total	**10.0 mL**	**3.3 mL**

*() = final concentration.

Following electrophoresis, the SDS-PAGE gel was stained using Coomassie Blue (0.05% (w/v) Coomassie Brilliant Blue R-250, 25% (v/v) isopropanol, 10% (v/v) acetic acid) overnight and then destained in 10% (v/v) acetic acid/10% (v/v) methanol, followed by 10% (v/v) acetic acid. The gel was photographed using the SynGene Gel Doc system. Protein quantification was performed through the use of a bovine serum albumin (BSA) standard. Known amounts of BSA (100 – 500 ng) were loaded onto the SDS-PAGE gel, together with the recombinant protein. The Coomassie Blue stained gel was analysed using densitometry and the SynGene Gel Doc system, with the GeneSnap 7.04 analysis software being used to quantify the known BSA standards. From these data, a linear BSA standard curve was constructed to determine the amount of expressed protein present on the SDS-PAGE gel.

2.2.7.4 Western Blot Analysis

The proteins of an unstained SDS-PAGE gel were transferred to Hybond-C Extra nitrocellulose membrane using a liquid transfer method at 35 V overnight at 4°C in transblot buffer (25 mM Tris, 192 mM Glycine, 0.1% (w/v) SDS, 20% (v/v) methanol; Towbin *et al.* 1979). Afterwards, the gel was stained using Coomassie Blue to detect if any proteins had failed to transfer to the membrane. The nitrocellulose membrane was washed in Tris buffered saline (TBS; 50 mM Tris-HCl, 0.9% (w/v) NaCl; pH 7.5) for five minutes and then stained with Ponceau S (1% (w/v) Ponceau S, 7% (v/v) glacial acetic acid) to verify that protein transfer had occurred. The non-permanent Ponceau S was washed away using water, before the membrane was rinsed in TBS for 10 minutes. The various bands of the Spectra™ Multicolor Broad Range Protein Ladder were then marked on the membrane.

The nitrocellulose membrane was blocked, to prevent non-specific antibody binding, by either incubation in a 3% (w/v) BSA Fraction V in TBS blocking solution (GST-tagged proteins) or a commercial anti-His horseradish peroxidase (HRP) conjugated blocking solution (His-tagged proteins) for one hour at room temperature on a shaking platform. A volume (25-50 mL) of either a 1:100 000 dilution of anti-GST HRP conjugated primary antibody in 1% (w/v) BSA in TBS or a 1:2 000 dilution of anti-His HRP conjugated primary antibody was used to cover the membrane and allowed to incubate for one hour at room temperature with gentle shaking. The membrane was then washed five times in 0.25% (v/v) Tween20-TBS (0.25% TBST) and once in TBS for five minutes each with vigorous shaking. A volume of 1.5 mL of SuperSignal™ West Pico Chemiluminescent Substrate was used to cover the membrane and allowed to incubate for 5 minutes at room temperature, before the chemiluminescent signal was detected using the SynGene Gel Doc system. After visualisation the membrane was stained using amido black (0.1% (w/v) amido black, 10% (v/v) acetic acid, 25% (v/v) isopropanol) for five minutes and then destained with 10% (v/v) acetic acid/10% (v/v) methanol, followed by 10% (v/v) acetic acid.

2.2.8 *P. falciparum* Phage Display Library Creation

A *P. falciparum* phage display library was created using the OrientExpress™ cDNA Synthesis and T7Select®10-3 Cloning Kits as per the manufacturer's instructions, although different suppliers were used during the reverse transcription and second-strand synthesis steps.

2.2.8.1   Total RNA Extraction

Total RNA was extracted from parasite cultures with a greater than 5% parasitaemia and a ratio of approximately 4:1 (trophozoite:rings). All instruments and surfaces were cleaned with RNaseZap™ prior to use to create an RNase-free environment. Parasites were isolated from RBCs by saponin lysis and RNA was extracted using TRI Reagent® as per the manufacturer's instruction. The use of TRI Reagent® separates the sample into three phases: a red organic phase, an interphase and a colourless upper aqueous phase (each containing proteins, DNA and RNA respectively).

Total RNA was extracted from a total of ten large (35 mL) parasite cultures, in batches of three/four flasks at a time. Each flask was divided between three 15 mL centrifuge tubes and centrifuged at 1 000 ×g for 10 minutes at 4°C. The supernatants were removed and each pellet (~500 µL) was resuspended in 1.5 mL of 0.05% (w/v) saponin in PBS, before the mixture was transferred to a 2 mL Eppendorf tube. After incubating for 10 minutes at room temperature, the tubes were centrifuged at 16 000 ×g for 5 minutes at 4°C. The supernatants were removed and the pellets were resuspended in 1.8 mL of 0.05% (w/v) saponin in PBS, before being incubated and centrifuged again. The resulting pellets were then washed with 1.8 mL of PBS. The parasite pellets were resuspended in 1.5 mL of TRI Reagent® and incubated for 5 minutes at room temperature, before 400 µL of chloroform was added. After mixing vigorously, the tubes were incubated for 15 minutes at room temperature and centrifuged at 12 000 ×g for 15 minutes at 4°C. The aqueous phase of each tube (~900 µL) was carefully transferred to a 2 mL Eppendorf tube and 900 µL of isopropanol was added, before the samples were mixed. After incubating for 10 minutes at room temperature, the tubes were centrifuged at 12 000 ×g for 10 minutes at 4°C. The RNA pellets were washed with 1.5 mL of 75% (v/v) ethanol and centrifuged at 12 000 ×g for 5 minutes at 4°C. After air-drying, the three pellets were dissolved using 25 µL of nuclease-free water by incubating each of pellet for   2 minutes at 60°C. The extracted, total RNA was then stored at -70°C.

RNA Analysis. The concentration and purity of each extracted RNA sample was determined by spectrophotometric analysis as described in 2.2.3.1, before the RNA samples were pooled together. An agarose gel (see 2.2.3.1) was used to visualise the three major bands of ribosomal RNA to assess if the sample was degraded.

Chapter 2: The Characterisation of PfIAP

2.2.8.2   mRNA Isolation

Messenger RNA (mRNA) was isolated from the extracted total *P. falciparum* RNA using the Dynabeads® mRNA Purification Kit. The kit relies on the binding of the poly(A) tail, found at the 3' end of mRNA strands, to the oligo $(dT)_{25}$ residues on the surface of the magnetic Dynabeads®.

Four hundred microlitres of 5 mg/mL Dynabeads® were washed using 200 µL of Binding Buffer (20 mM Tris-HCl, 1.0 M LiCl, 2 mM EDTA; pH 7.5) and then resuspended in 200 µL of Binding Buffer. The extracted total RNA was incubated for 2 minutes at 65°C to disrupt any secondary structures present, before being added to the Dynabeads® mixture. After incubating for 5 minutes at room temperature, the supernatant was set aside (as the 'unbound' fraction) and the beads were washed twice with 400 µL of Washing Buffer B (10 mM Tris-HCl, 150 mM LiCl, 1 mM EDTA; pH 7.5). Ten microlitres of 10 mM Tris-HCl (pH 7.5) was then applied to the beads, which were incubated for 2 minutes at 65°C to elute the mRNA. This elution step was repeated, before the 'unbound' fraction was reapplied to the beads and the process was repeated until eight elutions had been collected. The isolated mRNA was quantified as described in 2.2.3.1.

2.2.8.3   cDNA Synthesis

Isolated mRNA was reverse transcribed using anchored primers (5' $T_{10}VX$ 3'; V = A, G or C; X = A, T, G or C) in the presence of 5-methyl dCTP to protect any internal *Eco*RI or *Hind*III restriction sites from later restriction endonuclease digestion. First-strand cDNA synthesis was carried out using 4 µg of *P. falciparum* mRNA, as multiple 500 ng reactions.

Table 2.2.3: First-Strand Synthesis Reaction Set-Up

Component	Volume	Final Concentration	Manufacturer
Isolated mRNA	x µL	500 ng	-
12mer Anchored Primer Mix	0.5 µL	250 ng	-
10 mM Methylated dNTP Mix	1.0 µL	0.5 mM (in final reaction)	Fermentas
Nuclease-Free Water	y µL	-	-
Total	**14.0 µL**		

The first-strand mixture was incubated at 65°C for 5 minutes and then incubated on ice for 1 minute, before being used in the reverse transcription reaction.

Component	Volume	Final Concentration	Manufacturer
First-Strand Synthesis Mixture	14.0 µL	-	-
5X First-Strand Buffer	4.0 µL	1X	Invitrogen
100 mM DTT	1.0 µL	5 mM	Invitrogen
200 U/µL SuperScript™ III RT	1.0 µL	200 U	Invitrogen
Total	**20.0 µL**		

The reverse transcription reactions were carried out at 55°C for 1 hour, followed by incubation at 70°C for 15 minutes to inactivate the enzyme. Second-strand cDNA synthesis was then carried out for each of the 500 ng reverse transcription reactions.

Table 2.2.4: Second-Strand Synthesis Reaction Set-Up

Component	Volume	Final Concentration	Manufacturer
Reverse Transcription Mixture	20.0 μL	-	-
10X Reaction Buffer	8.0 μL	0.8X	Fermentas
10 mM Methylated dNTP Mix	1.0 μL	0.1 mM	Fermentas
5.0 U/μL RNase H	0.5 μL	2.5 U	Fermentas
10 U/μL DNA Polymerase I	2.5 μL	25 U	Fermentas
Nuclease-Free Water	68.0 μL	-	-
Total	**100.0 μL**		

The second-strand cDNA reaction mixtures were incubated at 15°C for 2 hours. The reaction mixtures were then combined, purified using the QIAquick® PCR Purification Kit as per the manufacturer's instructions and eluted into 21 μL of TE buffer.

2.2.8.4   cDNA End Modification

A single *Eco*RI/*Hind*III Directional Linker (GCTTGAATTCAAGC) was ligated to both sides of the double-stranded DNA (Figure 2.2.1). The Linker contains an internal *Eco*RI site, while a *Hind*III site is created at the 3' end of the *P. falciparum* insert by deriving two As from the cDNA. Digestion with *Eco*RI and *Hind*III thus allows for directional insertion of the *P. falciparum* dsDNA into the T7Select®10-3b Vector Arms. The T7Select®10-3b vector map is available in APPENDIX B.

```
      EcoRI                         HindIII
GCTTG AATTCAAGC  NNNNNNNNNAA A GCTTGAATTCAAGC
CGAAC TTAA GTTCG  NNNNNNNNNTTT CGA ACTTAAGTTCG
```

Figure 2.2.1: Ligation and Digestion of the *Eco*RI/*Hind*III Directional Linker
Diagram showing the ligation of the *Eco*RI/*Hind*III Directional Linker to the *P. falciparum* double-stranded cDNA insert and the subsequent digestion of the construct. Blue = *Eco*RI/*Hind*III Directional Linker; green = *P. falciparum* cDNA insert; red = *Eco*RI and *Hind*III restriction sites.

T4 DNA Polymerase was used to blunt the double-stranded cDNA ends.

Table 2.2.5: cDNA End Modification Set-Up

Component	Volume	Final Concentration	Manufacturer
Double-Stranded cDNA	20.0 μL	-	-
10X Flushing Buffer	3.0 μL	1X	Novagen
100 mM DTT	1.5 μL	5 mM	Novagen
1 mM dNTP Mixture	3.0 μL	0.1 mM	Novagen
2.5 U/μL T4 DNA Polymerase	0.6 μL	1.5 U	Novagen
Nuclease-Free Water	1.9 μL	-	-
Total	**30.0 μL**		

The blunting reaction was carried out at 11°C for 20 minutes, after which the reaction mixture was purified using the QIAquick® PCR Purification Kit as per the manufacturer's instructions and eluted into 10 µL of TE buffer. The *Eco*RI/*Hind*III Directional Linker was phosphorylated by a T4 polynucleotide kinase immediately before being ligated to the blunt-ended cDNA.

Table 2.2.6: *Eco*RI/*Hind*III Directional Linker Ligation Set-Up

Component	Volume	Final Concentration	Manufacturer
Blunt-Ended cDNA	10.0 µL	-	-
10X Ligation Buffer	2.0 µL	1X	Novagen
100 mM DTT	2.0 µL	10 mM	Novagen
1 mM ATP Mixture	2.0 µL	0.1 mM	Novagen
50 pmol/µL *Eco*RI/*Hind*III Directional Linker	2.0 µL	100 pmol	Novagen
10 U/µL T4 Polynucleotide Kinase	0.5 µL	5 U	Novagen

The phosphorylation reaction was incubated at 37°C for 5 minutes, cooled on ice for 1 minute before the T4 DNA ligase was added.

4 U/µL T4 DNA Ligase	2.0 µL	8 U	Novagen
Total	**20.5 µL**		

Ligation of the Directional Linkers occurred at 16°C for 20 hours, after which the ligase was inactivated by incubating at 70°C for 10 minutes. The reaction mixture was then allowed to cool to room temperature, before the construct was sequentially digested with *Hind*III and *Eco*RI restriction enzymes.

Table 2.2.7: Restriction Enzyme Reaction Set-Up

Component	Volume	Final Concentration	Manufacturer
DNA Insert	20.0 µL	-	-
10X *Hind*III Buffer	10.0 µL	1X	Novagen
20 U/µL *Hind*III	5.0 µL	100 U	Novagen
Nuclease-Free Water	65.0 µL	-	-

The initial *Hind*III digestion reaction was carried out at 37°C for 2 hours.

10X *Eco*RI Adjustment Buffer	10.0 µL	1X	Novagen
20 U/µL *Eco*RI	5.0 µL	100 U	Novagen
Total	**115.0 µL**		

The total digestion reaction was incubated at 37°C for 4 hours, before excess Directional Linker and small cDNA products were removed by size fractionation. Two millilitres of Gel Filtration Resin (Sepharose® 4B) was applied to an empty mini-column and allowed to pack under gravity flow. The mini-column was equilibrated four times using 1 mL of 1X Column Buffer (0.3 M sodium acetate), before the 115 µL of cDNA mixture was applied. After the DNA had settled into the resin, the mini-column was washed with 200 µL of 1X Column Buffer. Two hundred and fifty microlitres of 1X Column Buffer was applied to the mini-

column and the eluate was collected. This volume represents the void fraction of the mini-column and contains the largest cDNA fragments. These fragments were concentrated using the QIAquick® PCR Purification Kit as per the manufacturer's instructions and eluted into 10 µL of TE buffer. An aliquot of this DNA fragment solution was visualised using an agarose gel to determine the range of DNA sizes present.

2.2.8.5   Biotin Cassette Creation

To express biotinylated *P. falciparum* proteins as part of the T7 bacteriophage capsid, a biotinylation sequence (AviTag™) was inserted into the T7Select®10-3b vector (Figure 2.2.2a) based on the work done by Caberoy *et al.* 2009. When grown in *E. coli*, the lysine residue of the AviTag™ is biotinylated by the bacteria's natural BirA biotin-protein ligase, thus allowing for selection of bacteriophage via the unique biotin-streptavidin interaction. This should result in the selection of bacteriophage containing *P. falciparum* inserts that are in-frame.

Modification of the T7Select®10-3b Vector Arms to include a biotinylation sequence was generously carried out by Dr Sonja Lauterbach. The AviTag™ cassette was synthesised by GeneArt®, Invitrogen (Figure 2.2.2b). The stop codon between the *Eco*RI and *Hind*III restriction sites prevents the expression of the AviTag™ when no *P. falciparum* inserts are present in the T7 bacteriophage, while the restriction sites themselves facilitate the cloning of *P. falciparum* cDNA into the modified T7Select10-3b vector through the use of the Directional Linkers.

The AviTag™ cassette was digested with *Bam*HI and *Sal*I and inserted into the *Bam*HI and *Xho*I restriction sites of the T7Select®10-3b vector (Figure 2.2.2c; *Sal*I and *Xho*I result in the same sticky end pattern). The T7 bacteriophage were packaged and then propagated, following which the integrity of the AviTag™ cassette was verified by sequencing. The modified T7Select10-3b vector arms were then digested using *Eco*RI and *Hind*III to allow for the insertion of the prepared *P. falciparum* cDNA.

a)

T7SelectUP2

CGT**GCTAACTTCCAAGCGGACCA**GATTATCGCTAAGTACGCAATGGGCCACGGTGGTCTTCGCCCA
ArgAlaAsnPheGlnAlaAspGlnIleIleAlaLysTyrAlaMetGlyHisGlyGlyLeuArgPro

                                                        *Bam*HI *Eco*RI
GAAGCTGCAGGAGCTGTCGTATTCCAGTCAGGTGTGATGCTCGGGGATCCGAATTCTCCTGCAGGG
GluAlaAlaGlyAlaValValPheGlnSerGlyValMetLeuGlyAspProAsnSerProAlaGly

*Eco*RV   *Sal*I *Hind*III *Not*I   *Xho*I
ATATCCCGGGAGCTCGTCGACAAGCTTGCGGCCGCACTCGAG**TAA**CTAGTTAACCCCTTGGGGCCT
IleSerArgGluLeuValAspLysLeuAlaAlaAlaLeuGlu***......................

              T7SelectDown
C**TAAACGGGTCTTGAGGGGTT**TTT
....

b)

*Bam*HI *Eco*RI   *Hind*III   AviTag™
GGATCCGAATTCA**TAA**AAGCTTGGTAGTGGGAGCGGATTAAATGATATATTTGAGGCACAAAAAAT
  AviTag™   *Sal*I
TGAATGGCAT**TAA**GTCGAC

c)

T7SelectUP2

CGT**GCTAACTTCCAAGCGGACCA**GATTATCGCTAAGTACGCAATGGGCCACGGTGGTCTTCGCCCA
ArgAlaAsnPheGlnAlaAspGlnIleIleAlaLysTyrAlaMetGlyHisGlyGlyLeuArgPro

                                                *Bam*HI *Eco*RI   *Hind*III
GAAGCTGCAGGAGCTGTCGTATTCCAGTCAGGTGTGATGCTCGGGGATCCGAATTCA**TAA**AAGCTT
GluAlaAlaGlyAlaValValPheGlnSerGlyValMetLeuGlyAspProAsnSer***LysLeu

                            AviTag™
GGTAGTGGGAGCGGATTAAATGATATATTTGAGGCACAAAAAATTGAATGGCAT**TAA**GTCGAGTAA
GlySerGlySerGlyLeuAsnAspIlePheGluAlaGlnLysIleGluTrpHis***.........

                            T7SelectDown
CTAGTTAACCCCTTGGGGCCTC**TAAACGGGTCTTGAGGGGTT**TTT
....

Figure 2.2.2: The T7Select®10-3b Multiple Cloning Site and AviTag™ Cassette
a) Diagram showing the multiple cloning site of the T7Select®10-3b vector. Blue = forward and reverse primers used for sequencing reactions; red = various restriction sites; bold = STOP codon.
b) Diagram of the synthesised AviTag™ cassette. Green = AviTag™ sequence; red = various restriction sites; bold = STOP codons.
c) Diagram showing the multiple cloning site of the T7Select10-3b vector, now containing the AviTag™ biotinylation sequence. Blue = forward and reverse primers used for sequencing reactions; green = AviTag™ sequence; red = various restriction sites; bold = STOP codons.

2.2.8.6 Ligation and Viral Packaging

The optimal molar ratio for the ligation of the prepared cDNA inserts and the T7Select® Vector Arms is between 1:1 to 3:1 (insert:vector). The amount of vector arm used per reaction should be 0.02 pmol, while between 0.02-0.06 pmol of insert can be used (and is estimated based on the average length of the inserts after DNA synthesis).

Table 2.2.8: Insert and Vector Arm Ligation Set-Up

Component	Volume	Final Concentration	Manufacturer
cDNA Inserts	x μL	0.06 pmol	-
Digested Biotin-Tagged Vector Arms	y μL	0.02 pmol	-
10X Ligase Buffer	0.5 μL	1X	Novagen
10 mM ATP	0.5 μL	1 mM	Novagen
100 mM DTT	0.5 μL	10 mM	Novagen
2 U/μL T4 DNA Ligase	1.5 μL	0.6 U	Novagen
Nuclease-Free Water	z μL	-	-
Total	**5.0 μL**		

The ligation reaction was carried out at 16°C for 16 hours and then stored at 4°C. One 25 μL T7®Select Packaging Extract was thawed on ice, before the ligation reaction was added and the mixture was incubated at room temperature for 2 hours. The packaging reaction was terminated by the addition of 270 μL of LB medium. A plaque assay was performed to determine the number of recombinant T7 bacteriophage generated.

2.2.8.7 Plaque Assay

The total number of plaque forming units (pfu) was determined using a plaque assay. A 5 mL M9LB (22 mM KH_2PO_4, 22 mM Na_2HPO_4, 18.7 mM NH_4Cl, 1 mM $MgSO_4$, 0.4% (w/v) glucose in LB containing 50 μg/mL ampicillin) culture in a 50 mL Erlenmeyer flask was inoculated with a glycerol scraping of BLT5403 *E. coli* cells. The culture was incubated at 250 rpm at 37°C overnight before being used to seed a new flask with 5 mL of fresh M9LB medium. The BLT5403 *E. coli* were incubated at 37°C until an OD_{600} of 1.0 had been achieved. The newly packaged bacteriophage were serially diluted to a total volume of 1.0 mL using LB to a range of 10^{-3}-10^{-6}. One hundred microlitres of each phage dilution was added to 250 μL of the BLT5403 *E. coli* in M9LB. These phage/cell mixtures were each combined with 3 mL of melted top agarose (50°C; 1 % (w/v) tryptone, 0.5% (w/v) NaCl, 0.5% (w/v) yeast extract, 0.6% (w/v) agarose) and then applied to 1.5% (w/v) agar plates containing 50 μg/mL ampicillin. After allowing the top agarose to solidify, the plates were inverted and incubated for 3-4 hours at 37°C. The phage titre, library size and the efficiency of the packaging reaction were then determined by counting the resultant number of plaques:

$$phage\ titre\ \left(\frac{pfu}{mL}\right) = (number\ of\ plaques\ on\ plate) \times (dilution\ factor) \times 10$$

where 10 takes into account the 0.1 mL of the dilution plated

$$library\ size\ (pfu) = phage\ titre\ \left(\frac{pfu}{mL}\right) \times total\ sample\ volume\ (mL)$$

where the total sample volume is the 0.3 mL final packaging reaction volume

$$packaging\ efficiency\ \left(\frac{pfu}{\mu g}\right) = \frac{library\ size\ (pfu)}{amount\ of\ vector\ DNA\ used\ (ug)}$$

where 0.5 µg of vector DNA was used during the packaging reaction

2.2.8.8   Plate Lysate Amplification

Before the phage display library can be used in a biopanning experiment, the bacteriophage need to undergo a single round of amplification. Amplification is necessary as it allows the cloned sequences to be expressed and displayed on the surface of the bacteriophage particles. For a library consisting of less than 5×10^6 recombinants, the plate lysis method of amplification is preferred. The packaged bacteriophage (see 2.2.8.6) were added to 7 mL of fresh BLT5403 *E. coli* cells in LB (with an OD_{600} of 1.0). The phage/cell mixture was divided into 1 mL aliquots, to which 10 mL of melted top agarose was added. Each aliquot was poured onto a 150 mm 1.5% (w/v) agar plate containing 50 µg/mL ampicillin and spread evenly across the plate. After allowing the top agarose to solidify, the plates were inverted and incubated at room temperature overnight. Ten millilitres of phage extraction buffer (20 mM Tris-HCl, 100 mM NaCl, 6 mM $MgSO_4$; pH 8.0) was applied to each plate and incubated at 4°C for 3 hours. The amplified bacteriophage were harvested by collecting the phage in the extraction buffer and adding 500 µL of chloroform. The lysate was clarified by centrifugation at 3 000 ×g for 5 minutes at 4°C and the supernatant was collected and stored at 4°C. The titre of the amplified phage display library was determined by plaque assay using up to a 10^{-9} serial dilution (see 2.2.8.7), while glycerol stocks of the library were created by adding 0.1 volume of 80% (v/v) glycerol to phage aliquots and storing them at -70°C.

2.2.8.9   *P. falciparum* DNA Insert Analysis

Individual plaque plugs were selected from the plaque assay plates and placed in 50 µL of 10 mM EDTA; pH 8. The samples were vortexed briefly, incubated for 10 minutes at 65°C and then cooled to room temperature. PCR was then carried out using the DreamTaq Green PCR Master Mix as per the manufacturer's instructions, with 1 µL of the phage plug solution,

T7SelectUP2 and T7SelectDOWN primers (primer details are available in APPENDIX A), at an annealing temperature of 50°C. Bacteriophage with the biotin cassette, but not containing a *P. falciparum* insert are expected to produce a PCR product of 237 bp.

2.2.8.10   Streptavidin Selection

To isolate bacteriophage expressing a biotin tag, Streptavidin Magnetic Particles were used with modification. These particles allow for the isolation of a variety of biotin-labelled molecules, relying on the strong non-covalent interaction between streptavidin and biotin. To determine what volume of the bacteriophage library should be applied to the particles, the multiplicity of infection (MOI) calculation was applied:

$$volume\ required = \frac{library\ size\ (pfu) \times multiplicity}{amplified\ phage\ titre\ (\frac{pfu}{mL})}$$

A MOI of 1 means that each unique *P. falciparum* DNA insert is only represented by one bacteriophage in a given sample.

One hundred microlitres of streptavidin magnetic particles was washed three times with 500 μL of PBS. The calculated volume of bacteriophage at an MOI of 5 was added to 1.0 mL of 0.2% (v/v) Tween20-PBS (0.2% PBST), which was then applied to the streptavidin particles. The mixture was incubated for 1 hour at room temperature on an Intelli-mixer. The supernatant was removed and the particles were washed six times using 1.9 mL of 0.2% (v/v) PBST for 5 minutes at room temperature to remove any bacteriophage binding non-specifically. The biotin expressing bacteriophage were eluted off the streptavidin magnetic particles using 250 μL of 1% (w/v) SDS-PBS for 10 minutes at room temperature. The eluted bacteriophage were added to 50 mL of log phase BLT5403 *E. coli* in M9LB at an OD_{600} of 1.0. The cells were incubated at 250 rpm at 37°C until bacterial lysis was noted to allow for the amplification of the bacteriophage expressing biotin. Sodium chloride, at a final concentration of 0.5 M, was then added to the bacterial cells to aid in lysis before the culture was centrifuged at 10 000 ×g for 10 minutes at 4°C (Beckman Coulter Avanti® J-E centrifuge). The supernatant was collected and stored at 4°C before being used as the starting library in the biopanning experiments. The titre of the amplified biotin expressing phage display library was determined by plaque assay using up to a 10^{-8} serial dilution (see 2.2.8.7), while glycerol stocks of the library were created by adding 0.1 volume of 80% (v/v) glycerol to bacteriophage aliquots and storing them at -70°C.

2.2.8.11 Plaque Lift

Following Streptavidin Selection, bacteriophage expressing a biotin tag were visualised by means of a plaque lift. A plate from the Plaque Assay (see 2.2.8.7), with between 100 and 200 plaques, was chilled for 1 hour at 4°C to minimise the tendency of the top agarose to stick to nitrocellulose membrane. A membrane was gently applied to the plate for 1 minute at room temperature, during which time the edge of both the membrane and plate were marked. The membrane was then inverted and allowed to air-dry, before been blocked with 5% (w/v) BSA Fraction V in 0.1% (v/v) TBST for 30 minutes at room temperature. A solution of 1:100 000 biotin polyclonal antibody-HRP conjugate with 0.5% (w/v) BSA in 0.1% (v/v) TBST was then applied to the membrane for 30 minutes at room temperature. The membrane was washed three times using 0.1% (v/v) TBST for 5 minutes at room temperature. One and a half millilitres of SuperSignal® West Pico Chemiluminescent Substrate was applied to the membrane for 5 minutes at room temperature before the chemiluminescent signal was visualised using the SynGene Gel Doc system.

2.2.9 Biopanning with a Biotin-Tagged *P. falciparum* Phage Display Library

2.2.9.1 Biopanning

The streptavidin screened, biotin-tagged *P. falciparum* cDNA phage display library was used in biopanning experiments by probing against the recombinant GST-tagged PfIAP proteins (see 2.2.7.2). Using the titre of the amplified library, 1×10^7 pfu were prepared in a final volume of 500 μL of TBS. This was applied to ~8 μg of MagneGST™ bound, recombinant GST protein for 1 hour at room temperature on an Intelli-mixer. This background screening step serves to remove any of the bacteriophage that bind to either the GST protein or the MagneGST™ particles. The unbound bacteriophage were collected, mixed with ~8 μg of MagneGST™ bound PfIAP protein and incubated as described above. After removing the supernatant, the protein-bacteriophage complex was washed five times with 2 mL of 0.05% (v/v) TBST for 10 minutes at room temperature to remove any non-specific binders. The washed particles were added to 50 mL of log phase BLT5403 *E. coli* in M9LB and the bacteriophage were allowed to amplify overnight. The amplified phage were then collected as described in 2.2.8.10. A plaque assay (see 2.2.8.7) with serial dilutions up to 10^{-9} was carried out, before the bacteriophage were used for another round of biopanning (Figure 2.2.3). In total, four rounds of biopanning were performed for each of the GST-tagged PfIAP proteins. Each successive round of biopanning enriches for the bacteriophage that bind the strongest to the bait protein.
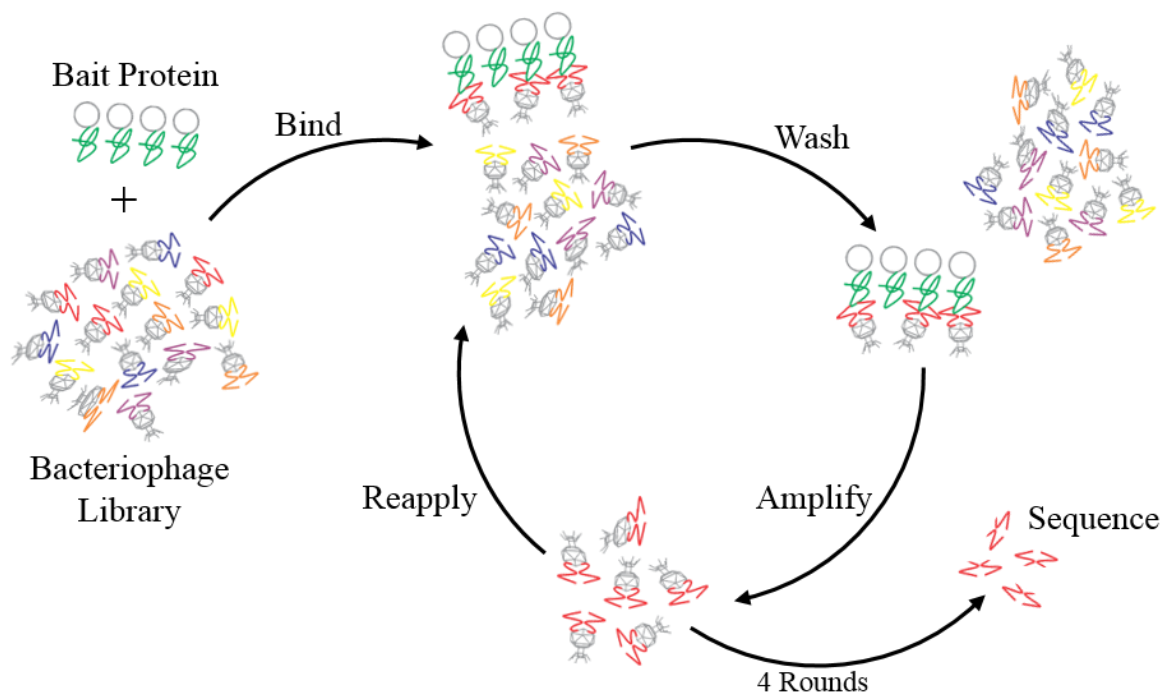
Figure 2.2.3: The Biopanning Process
Diagram showing how to perform biopanning with a phage display library. The bacteriophage library, expressing *P. falciparum* peptides, was applied to the immobilised GST-tagged 'bait' protein (green). Non-binding phage were washed away, while phage that bound because of protein-protein interactions (red) were amplified, before being reapplied to the 'bait'.

2.2.9.2 Identification of Interacting *P. falciparum* Peptides

Following biopanning, PCR was performed on individual bacteriophage as described in 2.2.8.9. PCR amplicons larger than 300 bp were sequenced by Inqaba Biotec™, South Africa. The identity of the *P. falciparum* binding peptides was determined by using the genomic data available in PlasmoDB (Aurrecoechea *et al.* 2009).

2.2.10 *In vitro* Binding Assay

To confirm the interaction of the identified binding partners with the GST-tagged PfIAP proteins, *in vitro* binding assays were performed with two of the interacting proteins. The recombinant His-tagged PfRhopH3 and PfDOC2 proteins, together with the GST-BIR protein, were prepared and eluted as described in 2.2.7.2. The GST-tagged protein was dialysed against three changes of TBS using a Slide-A-Lyzer® MINI Dialysis Unit (10 kDa molecular weight cut off) for 30 minutes at room temperature. The His-tagged proteins were likewise dialysed against the His-Binding buffer, before 0.5 µg aliquots of each protein were rebound to 5 µL of prepared MagneHis™ particles for 1 hour at 4°C. Increasing concentrations of GST-BIR were then exposed to the bound His proteins in a total volume

of 150 μL of TBS for 1 hour at 4°C on an Intelli-mixer. The unbound supernatant was removed, before the particles were washed three times with 1 mL of TBS for 5 minutes at room temperature. The bound protein complex was then assessed using SDS-PAGE as described in 2.2.7.3 and analysed using densitometry and the GeneSnap 7.04 and GraphPad Prism 6 software. A heat denatured (15 minutes at 70°C) GST-BIR control aliquot was included, as were various concentrations of recombinant GST protein, to account for any non-specific binding.

2.2.11 Creation of Transgenic PfIAPbir-GFP *P. falciparum*

2.2.11.1 Preparation of the PfIAPbir-GFP Construct

Plasmodium falciparum transfection requires more than 100 μg of plasmid construct, which was prepared as described below. A 5 mL LB culture (containing 100 μg/mL ampicillin, in a 50 mL Erlenmeyer flask) was inoculated with a glycerol scraping of XL10 *E. coli* cells containing the pARL2-GFP_PfIAPbir plasmid construct (see 2.2.6) and incubated at 250 rpm for 8 hours at 37°C. Five hundred millilitres of LB medium (containing 100 μg/mL ampicillin) was seeded with 1 mL of the initial culture, split into two 250 mL aliquots (in 1 L Erlenmeyer flasks) and incubated for 16 hours as described above (Thermo Scientific Forma incubated/refrigerated floor orbital shaker). The OD_{600} of each aliquot was measured and when greater than 2 the bacterial cells were centrifuged at 6 000 ×g for 15 minutes at 4°C (Beckman Coulter Avanti® J-E centrifuge).

Large-scale plasmid purification was performed using the NucleoBond® Xtra Maxi Plus Plasmid Kit according to the manufacturer's specifications and eluted into 1 mL of the provided Tris Buffer (5 mM Tris-HCl; pH 8.5). A 1:10 dilution of the eluted plasmid construct was linearized with FastDigest® *Bam*HI and then quantified by agarose gel electrophoresis as described in 2.2.3.1. An elution volume equivalent to ≥150 μg of plasmid construct was purified by adding an equal volume of 1:1 phenol/chloroform. After mixing, the tube was centrifuged at 16 000 ×g for 5 minutes at 4°C and the aqueous phase was transferred to a 1.5 mL Eppendorf tube. One volume of chloroform was added to the tube which was mixed and centrifuged again. The aqueous phase was transferred to a 2 mL Eppendorf tube and the plasmid construct was precipitated as described in 2.2.3. The pellet was allowed to air-dry before being resuspended in 30 μL of TE buffer.

2.2.11.2 Transfection of Parasite Cultures

The ~150 µg of purified pARL2-GFP_PfIAPbir construct in 30 µL of TE buffer was added to 370 µL of pre-warmed cytomix (37°C; 120 mM KCl, 0.15 mM CaCl$_2$, 2 mM EGTA, 5 mM MgCl$_2$, 10 mM K$_2$HPO$_4$/KH$_2$PO$_4$, 25 mM HEPES; pH 7.6; Wu *et al.* 1995). Two hundred microlitres of fresh, unwashed, packed RBCs were added to the plasmid construct cytomix solution, which was transferred to a 2 mm Bio-Rad Gene Pulser® electroporation cuvette. The mixture was electroporated using a Bio-Rad Gene Pulser Xcell™ electroporation system configured to 310 V and 950 µF, with an expected time range for the electroporation of   12 to 14 msec. The contents of the cuvette were then transferred to a small culture flask using 5 mL of complete medium, before 50 µL of ~5% ring stage parasitised RBCs (pRBCs) were added (final parasitaemia ~1%). The culture was then gassed and incubated at 37°C as described in 2.2.2.2. The following day the culture medium was changed as described in 2.2.2.2, but from 48 hours post-transfection the positive selection WR99210 compound was added to the culture medium (10 µM stock in DMSO; final concentration of 2 nM). The pARL2-GFP plasmid contains the human dihydrofolate reductase (*hdhfr*) cassette which confers resistance to the WR99210 compound. The transfected parasite culture was monitored daily and maintained under WR99210 pressure for one week. Subsequently, the culture medium (containing the WR99210 compound) was changed three times a week, with the addition of 100 µL of fresh, washed RBCs suspended in incomplete medium in a 1:1 ratio once a week. Daily culturing, with the continued pressure of the WR99210 compound, was resumed once transgenic parasites were microscopically detectable. Glycerol stocks of the transfected parasites were created from a ~4% ring stage parasite culture, as described in 2.2.2.3.

To verify transfection, parasite DNA was extracted from a culture with a parasitaemia greater than 4% mixed stage parasites (Tirasophon *et al.* 1991; Vu *et al.* 1995). Three millilitres of culture was transferred to a 15 mL centrifuge tube and 1.5 mL of lysis solution (34 mM NaCl, 1% (v/v) Triton X-100, 1.2 mM EDTA) added. The tube was briefly vortexed before being centrifuged at 3 000 ×g for 10 minutes at 4°C. The supernatant was aspirated and the parasite pellet was resuspended in 100 µL of Tris buffer (10 mM Tris-HCl, 50 mM KCl; pH 8.3) before being transferred to a 1.5 mL Eppendorf tube. After a brief vortex, the tube was centrifuged at 16 000 ×g for 5 minutes at 4°C and the supernatant was aspirated. The pellet was again resuspended in 100 µL of Tris buffer, vortexed and centrifuged as above. After the supernatant was aspirated, the pellet was resuspended in 100 µL of boiling

buffer (10 mM Tris-HCl, 50 mM KCl, 3 mM MgCl$_2$; pH 8.8) and boiled for 10 minutes. Two hundred microlitres of 1:1 phenol/chloroform were added to the tube, mixed briefly and centrifuged at 16 000 ×g for 5 minutes at 4°C. The top aqueous phase was transferred to a 1.5 mL Eppendorf tube and one volume of chloroform was added to the tube which was briefly vortexed and centrifuged again. The aqueous phase (~80 µL) was transferred to a new 1.5 mL Eppendorf tube and stored at -20°C. A PCR reaction, using pARL2-GFP specific primers (pARL2 Fwd/Rev; available in APPENDIX A), at a T_a of 63°C was carried out and the PCR amplicons were visualised by agarose gel electrophoresis.

2.2.11.3   Visualisation of PfIAPbir-GFP in Live Transgenic Parasites

Three hundred microlitres of synchronised transgenic parasite culture (~5% parasitaemia) was removed from its culture flask and added to 700 µL of incomplete medium. To visualise the nucleus either DAPI (5 mg/mL stock in water; final concentration of 0.2 µg/mL) or Hoechst 33258 pentahydrate (1 mg/mL stock in water; final concentration of 6 µg/mL) was added to the parasite mixture which was incubated for 5 minutes at room temperature or 2 hours at 37°C, respectively. The parasite suspension was centrifuged at 600 ×g for 3 minutes at 4°C. The supernatant was aspirated, the pRBC pellet was resuspended in incomplete medium and centrifuged at 600 ×g for 3 minutes at 4°C. The pRBC pellet was washed again, before being resuspended in 200 µL of incomplete medium.

Live parasites were visualised by placing 5 µL of the stained parasite suspension on a microscope slide with a cover slip and viewing at 1 000X magnification using an Olympus BX41 microscope. Filters U-MWB2 (emission above 510 nm) and U-MWU2 (emission above 410 nm) were used to visualise GFP and stained nucleic material, respectively, while images were captured using an Olympus DP72 camera with CellSense Dimensions 1.7 software.

Induction of RCD. The effect of fever in malaria patients was mimicked by exposing synchronised transgenic parasite cultures (~5% parasitaemia) of both ring and trophozoite stages to 41°C for 2 hours, after which time the cultures were returned to 37°C. Images of live, stressed parasite cultures were captured 30 minutes, 2 and 24 hours post heat-shock.

2.3   Results

2.3.1   Characteristics of PfIAP and its Orthologues

Scrutiny of the PF3D7_0519600 entry in PlasmoDB v9.1 showed that *pfiap* is a single exon gene (Figure 2.3.1) with two identified domains: a C3HC zinc finger-like domain (aa: 50-168) and a β-sandwich domain (aa: 378-726; although this domain has since been refined to amino acid positions: 378-437 – PlasmoDB v34). SMART was used to identify the location of the BIR domain in the PfIAP protein (aa: 53-128), which corresponds to what is now identified as an IAP repeat (aa: 52-127) in PlasmoDB v34. The PF3D7_0519600 protein lacks the other domains (e.g. RING domain) found in some members of the BIRC family of proteins.
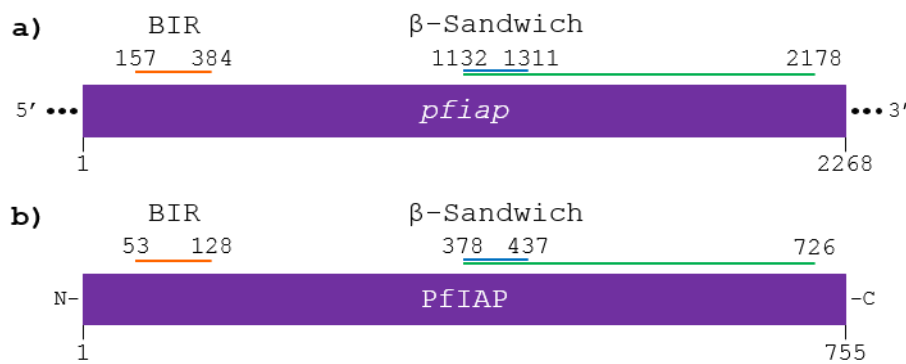
Figure 2.3.1: Domain Structure of PF3D7_0519600
a) Diagram of the *pfiap* gene. Blue = location of the β-sandwich domain (PlasmoDB v34); green = location of the β-sandwich domain (PlasmoDB v9.1); orange = location of the BIR domain; purple = *pfiap* gene. Numbers = nucleotide number.
b) Diagram of the PfIAP protein. Blue = location of the β-sandwich domain (PlasmoDB v34); green = location of the β-sandwich domain (PlasmoDB v9.1); orange = location of the BIR domain; purple = PfIAP protein. Numbers = amino acid number.

One orthologue of the *pfiap* gene was found in other *Plasmodium* species including: *P. berghei* (PBANKA_123440), *P. chabaudi* (PCHAS_123500), *P. knowlesi* (PKNH_1013500), *P. malariae* (PmUG01_10025000), *P. ovale* (PocGH01_00238900), *P. reichenowi* (PRCDC_0518700), *P. vivax* (PVX_080265) and *P. yoelii* (PY17X_1237800). The *Plasmodium* orthologues all contained one conserved BIR domain which, along with the *iap* gene, is not present in other Protozoa (*Toxoplasma gondii*, *Trichomonas vaginalis*, *Leishmania*, *Entamoeba* etc.). Analysis of the amino acid alignment of these eight *Plasmodium* proteins with PfIAP (Figure 2.3.2) showed that these *Plasmodium* orthologues were 34.8% identical (31.2%-38.1%) and 52.4% similar (50.7%-54.4%) to PfIAP. *Plasmodium reichenowi* was an outlier with 86.5% identity (and 89.4% similarity)

to PfIAP, but this can be explained by that fact these two species are closely related to each other. A high degree of similarity was also observed at the beginning and at the very C-terminal end of the proteins. The green and blue boxes (Figure 2.3.2) show the previous (PlasmoDB v9.1) and new (PlasmoDB v34) locations of the predicted β-sandwich domain of the *P. falciparum* protein, which is not predicted for any of the other orthologues. β-sandwich domains are composed of two opposing antiparallel β-sheets (Kister *et al.* 2006). Another unique feature of the *P. falciparum* orthologue is that it is the only protein that has a 62 amino acid low complexity region, made up of repeating sequences of the asparagine, glutamine and histidine amino acids.

```
P.falciparum MSESSALSTSVKDTNKCDSSNCDKEELVDINKSFINIE-EENTKDITNYLKDKISFMKRIETFKAWVRKPAHLSSLILARNGYVCQDEFVIRCEICNCKYIYE
P.reichenowi MSESSALSTSVKDTYKCDSSNCDKEELIDINKSFINIE-DENTKDITNYLKDKISFMKRIETFKAWVRKPAHLSSLILARNGYVCQDEFVIRCEICNCKYIYE
P.knowlesi   MNDPLP-----SEATGVDRLEPQNEELINVRKTFLGEEEDEYTKSSRQYLRDQENYVKRIKTYNVWVKKPAHLSSLLLARSGFICENETTIRCEMCKCKFVYE
P.vivax      MNGPCP-----GEAAGVDPREAPNEELINVRKTFLEGEEEDEYTKSSRRYLQDENYFNRIKTYTLWVKKPHLCSLLLARNGFICENETTVRCELCRCKFVYQ
P.ovale      ---------MEKQPEKSYAEEKNPGELINVGKSFLNEEPNDFAKYVKKYLYDEVSYANRVNTFSVWVNKPPHLCSLILARNGFVCKNENAIKCEMCGCKYVYK
P.malariae   MSNPVEENESVGRDIATSEVKLVTEELIDVNKKFVNEEEEEYIKDSKRYLYDEVSYINRINTFKIWVRKPTHLCSLILARNGFICESECVIKCEMCKCKYVYE
P.chabaudi   MSSSGDQSENSGEKIKLIKEENENEELISVENRFLNDEKDGHIKDVREYINNKSYINRLNTYNKWISKPVHLCSLILARNGFVCKNESTIECEMCQCKYIYE
P.berghei    MNSSGDQIENADKKIKLIKGENEKEELISVENRFLNDEKDDHIKDVKEYINNKSYINRLNTYNKWINKPVHLCSLILARNGFVCKNESTIECEICQCKYIYE
P.yoelii     MNSSGDKIESDGKKIKLIKDENEKEELISVENRFLNDEKDNIKDVKEYINNKSYINRLNTYNKWINKPVHLCSLILARNGFVCKNESTIECENCQCKYIYE
                  :    **.:: : *: * : *   *: ::.:.:*::*:. *: ** **.**.***.*::*:.* .:: ** * **::*:

P.falciparum KDTYSMYTRINDLCLLHLHNCPWRNNLMDISIFDLDKKSLTKEELLKEYEDNIVSLKNNLLFVPLINIKKTLHDLIFILKKFLENNMNKKRFSIFDHYVKIFK
P.reichenowi KDTYSMYTRINDLCLLHLDNCPWRNNLMDISIFHLDKKSLTKEELLREYEDNIVSLKNNLLFVPLINIKKTLHDLIFILKKFLENNMNKKRFTIFDHYVKIFK
P.knowlesi   KDTYSMYTRINDLCLLHLKNCPWGKLMDLQILHLDGESLKRENLLREYQNVTTLLQNGTHDVPYIDIKKTISDFLLIVKKYLTNEDSKKKFLLFNSYVELFK
P.vivax      KDTYSMYTKINDLCLLHLKNCPWGKLMDLQILRLDGDSLQRDNLLREYQNVSTFLQSGTNDVPYVDIKKTIGDLLLIVKKHLANEDSKKKFLLFSSYVELFK
P.ovale      IGTYSIYTKINNLSLLHKDNCPWKKHIMHLSFFKLDNISLTRSEILKEYEKNTTILKNVLNKIPLINIKGIIHDLILIIKKHLRNDKSKENFLIFNSQVQFFK
P.malariae   KNTYSMYTRINDLCLLHLDNCPWKNKIMDLSVFKLDDKSLSRQKILEEYENNTKLIKSSLVKIPLINIKKTINDLILIVKKHLDDDISKKKFLIFNSHVEFFK
P.chabaudi   KGAYSMYTRINDLCLLHTNSCPWKKLIDLSFFRLDNNSLERIELLREYKHNCEILRRSLLHIPMINIKKTISDMILIIKKHVHNDKTKKKFTIFNSYLDFFK
P.berghei    KGTYSMYTRINDLCLLHTDNCLWKNKLIDLSFFKLDSNSLDRVELLSEYKNNCEILRSLLYIPMINIKKTINDIILVIKKHVQNDKTKKKFTIFNSYLDFFK
P.yoelii     KGTYSMYTRINDLCLLHSDNCPWKNKLIDLSFFKLDNNSLEGVKLLREYKNNCEILKRSLLYIPMINIKKTINDIILVIKKHVQNDKTKKKFTIFNSYLEFFK
             .:**.**.**:*.***  .*.** :: :..: ** **  ::* **:    :.    :* ::**  : *:::::**.: :: .*:.* :*.  ..:**

P.falciparum EEYIHIFLNNKNIIDKGIEIIKNKYPLYEKYIVDINYLNNLNFNMNNVTSFINILSDIC-LDDSFCDE-QDDINMYFKIVQKLKNYQYENVIFKLLALFGWS
P.reichenowi EEYIHIFLNNKNIIDKGIEIIKNKYPLYEKYIVDINYLNNLNFNMNNVTSFINILSDIC-LDDSFCDE-QDDINRYFKIVQKLKNYQYENINIFKLLALFGWS
P.knowlesi   SHYVEIFQSHKNLVDRGIQLIGKKCKHLKNHLVDEAFLKSLTYDPLDVHTYINIVGDANDRTQKLSHP-NEDLNMYFKIVQKLKNCEYEDINIYKLIALLGWT
P.vivax      THYVEIFKNHKKLVDRGMHLIQSNYNHLKTPLVDEAFLNSLTYDPLDVHTYINIVADENRKAERLTHP-DADLHMYFKIVNQMKDFQYEDVNIYKLIALLGWT
P.ovale      KHFVEIFLYNDDVIQDGINFVKNFKELQNCIVDLKYLENLNFDIFDIKNYINILSGVN-SSELLSNDLREDINMYLKIVCKMRDYDYFNVCIFKAISLFGWT
P.malariae   KHFIQIFLNNKSVIDEGIEKIKKNYKHLEKNIIDMNYLNNLHFDLFDVTNYINILSDDEYSTSLLIDE-KEDINMYYIVNKMKNYNYEHVNIFKIIALLGWT
P.chabaudi   NNFIPIFLKNENIIQKSIDYIKQNKDMEKYIIDADYIKNLNFLFDINNFINILSDTS-EPPQLTHE-KEDINAYIKIMHYLRTYEYEDINLFKIISLLGWT
P.berghei    NNFIPIFLKNENIIQKSIDYIKKNYKDMEKYIVDSDYIKNLNFLFNVNNFINILSDTS-AAPELENE-KEDINAYIKIMHYLRKYEYDDINLFKVISLLGWT
P.yoelii     NNFIPIFLKNENIIQKSMDYIKQNYKDMEKYIIDADYIKNLKFNLFNINNFINILSDIE-TLAQLPHE-KEDINAYVKIIHHLKKYEYEDINIFKIISLLGWT
             :: ** :..:::.:: : .: ::*  :::.* ::  :.:***:.. : *:: * **:  :.  :* :  ::* ::*:**:

P.falciparum YKSPQDDEH----TQILQCKYCFREINILNYSHFSFKQNKLSLHKDIFASHQKDVLQYILKNINILLEQRKKENQNQNQHQNQHQNQHQNQHQNQHQNQHQNQ
P.reichenowi YKSQQEEHT-----QILQCKYCFREINILNYSYFSFKENKLSLHKDIFSSHQKDVLQYILKKINILLEQRKKEN------------------------------
P.knowlesi   YKTSPEGEDPDVSSHILYCRYCFREVNLADYSTLKIMHDLDLFKTIDIEPSGEAPKQSLGGARDALERLTGLADGWIETDEEEESQEG--------------
P.vivax      YRDSPQGEDPNGDSHILYCRYCFREVNLGDYSTLSRRHNRMDLFKAIDVEPPGEVANESVDDARHALERLTGLAHAWVEVEQKEANEKETAEM----------
P.ovale      CKENSTDLD--SSRYIIVCKYCCREVNICHYSHFDRNTKEMVLFKDVNPVHEKEVRGYLTNSDRDR--SIGGDDSKGVDAVQRLIDMVIGQSNKAEKGESKKG
P.malariae   CKDNPYHSK-KYSSHILLCRYCREVDISNYSFLSSTDNELHLFKDYDADKAKDIMEDLENKKRHRKERRENANAHGMNVMQRIFNLISGYVDKG--------
P.chabaudi   YKNETQNNS---KEHIISCKYCYREVNISDYSYFSTNELLIEYKHNCLLIEVNSNDMDYFLSGSSKNKSANQEKDEKTV---DPILNQILTLVNGHFD----------
P.berghei    CKDEIKNNS---KEQIISCKYCYREVNISDYSYFSLNDKKYNLFIEVDSNDMDYFLSDPSNNKSMNQDKNEKAT---NIILDKILTIVNGHFD----------
P.yoelii     YKDEIQNNS---KEQIISCKYCYREVNISDYSYFSPDNKYNLFIQVDSNDMDYFLSETSKNKSINQEQNEKTI---NLILDEILTLVNGHFD----------
                    *: *.** **::: **  :      : *

P.falciparum HQNQHQNQHQNQHQNQHQNQHQNQHQNQHQNQKDTLEYIPEHTILSKMSDRLRETYDK-YYIDIARNNKNKEQKQNEEKENNEVNKEVYIERNNNKEETSEKE
P.reichenowi --------------------------------EKGTLEYIPEHTILSKMCDRLRDTYDK-YYIDIARNNKKKEQKKNEEKENNQVNKEAYIERNNNKEETNEEE
P.knowlesi   --------------------------------EEGEEKVDEERDGEVRMSAGEEARGDEAEVEETEKGVVEKEQEE--------------IKGVEEKGGEPLGE
P.vivax      --------------------------------EAAEMEAAEKEAPEKEVEEKEAVEKEEAEMEAAEKETAEKEQEENPVEEKNGGEKERKEPSGEEKNDGEQEG
P.ovale      ESKKGEVEK---------------------DEVKDEVEKDEVKDEMEKDEVEKDEVKDEMEKDEVEKDEVKDEMEKDEVEKDEMEKDEVEKDEMEKGE
P.malariae   --------------------------------GESAIEDGKEEDGKEEDGKEEDGKEEDGKEEDGKEEDGKEEDGKEEDGKEKDGKEKDGQEEG
P.chabaudi   --------------------------------EAGEIIY---EEVEDQLSKNRHHNKDEDKSMDDTEK----------------------IEPNDEKENTSDEK
P.berghei    --------------------------------EEGEIIY---DEIEKQLYKNVNKEDKD-KSIDISK-NENKIVVDAEK-----------IEPINEKENVSEEK
P.yoelii     --------------------------------QAGEIIY---EEMENQLGKNVNNKDDD-KSIDV-RGNEKETVVDAEK-----------IEPINEKEASEKEA
                                                  ..                                                            :*

P.falciparum ENSKNEKFNRSNKASQSDSYEENAEYSQS---------------------------------D----GITNTESNFENSQKEREGFSFKWKFKNLFSSQSIH
P.reichenowi ENSKNEKFNRSDKGSQSDSYEENAEYSQS---------------------------------D----GITNTESNLESSQKEREGFSFKWKFKNLFSSQRVH
P.knowlesi   GNKSNEKEGNGNVSEDANIAPNEEGGTKD---------------------------------LPDEGKSEDQENAEEQEGFSFKWSIKQFFLPKKST
P.vivax      EEPPVEEENDGKKEEEEPPREENNGGEKEVENPVDETNGGEKEGEPPVEEINGGEKKGNENASDDAKGDPNEGGGTKDDPAEEQEGFSFKWSIKQFFLPKKST
P.ovale      MEKGEAEKGEAEKGEVEKDEVEKDEVEKD---------------------------------EMEKGEAEKGEVKYDQASGNKEGLAFRWKLKDLFLSKKNY
P.malariae   GNKAGAVKDDDDRVDNGKEAAPRD----------------------------------------------DSTISEMENQRLTKDTES-SLKWTLKHFFLPKKRS
P.chabaudi   EKS-------------------------------------------------------------------NESTTKQRESFSFKSTLKGIFLSKKSS
P.berghei    ENK-------------------------------------------------------------------NENEIKHKESFSFKSTLKGIFLSKKTN
P.yoelii     SEKEVSEKEVSEKE-------------------------------------------------------ASDEKEVSGENEIKQKESFSFKSTLKGIFLSKKTN
                 :                            :                                       *. ::. ::*   :* .:

P.falciparum QDVNDKHNDKLSCEEDREHILEDMKEIIFLDKMENNKADNENNNNNNNNNNNTYSNSNDSNCNDSNCNDSNSNNDDNNVENHLKNYLKKCIGEMCLNICPYLEN
P.reichenowi QDVNDKHNDKLSCEEDREHILEDMKEIIFLDKME----NNKADSENNNNNNNTYSNSNSKDSNDSNSNDDNNVENHLKNYLKKCIGEMCLNICPYLEN
P.knowlesi   ENEDKNENP---------------------------DDNSPSECSPTSVNQGGLAQNKEGNAPCARKR---------IKKKVQLENILSRVFLEISRYTED
P.vivax      DKGDKNQSA---------------------------DDASPRECPPTSVNQGELTPNRKRNSPCARKR---------RKKKVQLKDILSRVFLEVSSYTEN
P.ovale      DNIKDGKDE----------------------------NSTPRENSLDESPVSKKETKLIEKKL-------YMNKTFEENSLSKTFLDICTHVEN
P.malariae   KNLKENEED----------------------------SNTSLSKDMVANNKIEANCSEYKL-------YKVKLSLKDSMSKLFLEICSYMDD
P.chabaudi   DNLKEDETD----------------------------DSELEDGKEKNKDNEK-------------NENKIKQASSLNEAMLELCKYTDD
P.berghei    DNLKEYKSN----------------------------DNGLEGENGKNNIEIKDNDDEK-------NENKIEQASSLNEAMLKLCKYTDD
P.yoelii     DNLKEDGKK----------------------------NNNNDNNDNNDDNNDNNDNNNDNNNDNNND-----NNNDNKQASSLNEAMLKVCKYIDD
                         .    . ..    .  :   *..       ::   *::.   :

P.falciparum KNYFIKSLYDYYVLL-TSEMSTTHINENA--------------------IKNSFIIRPFNVVENHRIYCPYITEDLYLFSKITKLIFELLILEFERKYVFK
P.reichenowi QNYFIKSLYDYYVLL-TSEMSTTNINENA--------------------IKNAFIIRPFNVVENHRIYCPYITEDLYLFSKITKLIFELLILEFERKYVFK
P.knowlesi   ENLFSKAEHDCYVLRGGPHLAHLDKERTTADIGTDGGNYTDNTVGSDETAGTTRKRTIRAFNLIENHRAFCPYMTEDLYSFSKITKLLFELVMSEFQRRYVMR
P.vivax      EILFSKAANDCYVLRGGPHLAYLDKEPTPAEDGTDGANCTGNTAECAETAGNTQKRTIRAFNLIENHRAFCPYMTEDLYSFSKITRLLFELVTSEFQRRYVMR
P.ovale      EPYFLKTVYDYYVLE-DSHMTHSSEKDTNEEK------------------DKKLYAIRPFNTVENHRLYCPYMMEDLYSFSKITKLFFELLMSDFQRKYMFK
P.malariae   DKFFIKTLYDYYTLT-TPPMEYINDEGNTNK------------------EKKCYYIRPFNAIENHRLYCPYITEDLYAFSKITKLLFELLMSEFERKFEFK
P.chabaudi   DKYFLKGKSNYYILK-ESIQENLNKKI-----------------------SNKIYAIRLFNLIENHRVYCPYIMEDLYGFSKITKLFFELLMSEFQRMYMFK
P.berghei    DKYFLKGKSNYYILK-EFIQENLDDNQIKKQKITKEQEKREKI--------TNKIYAIRLFNLIENHRVYCPYVMEDLYGFSKITKLFFELLISEFQRMYMFK
P.yoelii     DKYFFKGKSNYYILK-ESIQENLDKNQIKKQKI-----------------TNKIYAIRLFNLIENHRVYCPYIMEDLYGFSKITKLLFELLMAEFQRMYMFK
             . ** :** :   .. ** ** :*::.:***.:***.**** *****.*:***: :*:* .::
```

Figure 2.3.2: *Plasmodium* IAP Alignment
Amino acid alignment of the nine BIR-containing *Plasmodium* proteins. Blue box = β-sandwich domain (PlasmoDB v34); green box = β-sandwich domain (PlasmoDB v9.1); orange box = BIR domain. * = identical; : = conserved; . = semi-conserved.

Each of the nine *Plasmodium* proteins only has one BIR domain (orange box in Figure 2.3.3), that shows 64.7% identity (57.9%-72.4%) and 81.4% similarity (77.6%-88.2%) to PfIAP (except for *P. reichenowi* which shows 98.7% identity). The four important zinc ion chelating amino acid residues are conserved between the nine studied *Plasmodium* IAP proteins and all follow the same sequence layout (green triangles in Figure 2.3.3). This $CX_2CX_{23}HX_3C$ (C = cysteine; H = histidine; X = any amino acid) sequence of the *Plasmodium* BIR domains is different to the general $CX_2CX_{16}HX_6C$ sequence of human and other species' BIR domains (green triangles in Figure 2.3.4; Srinivasula and Ashwell 2008). Some amino acids are identical between the BIR domain of *P. falciparum* PF3D7_0519600 and the BIR domains of members of the human IAP family of proteins (Figure 2.3.4). The PfIAP BIR domain appears to have the highest degree of identity and similarity with the third BIR domain of the human NAIP protein (BIRC1c; Table 2.3.1), although the XIAP BIR domains are the preferred templates for the PfIAP BIR 3D models (see below).
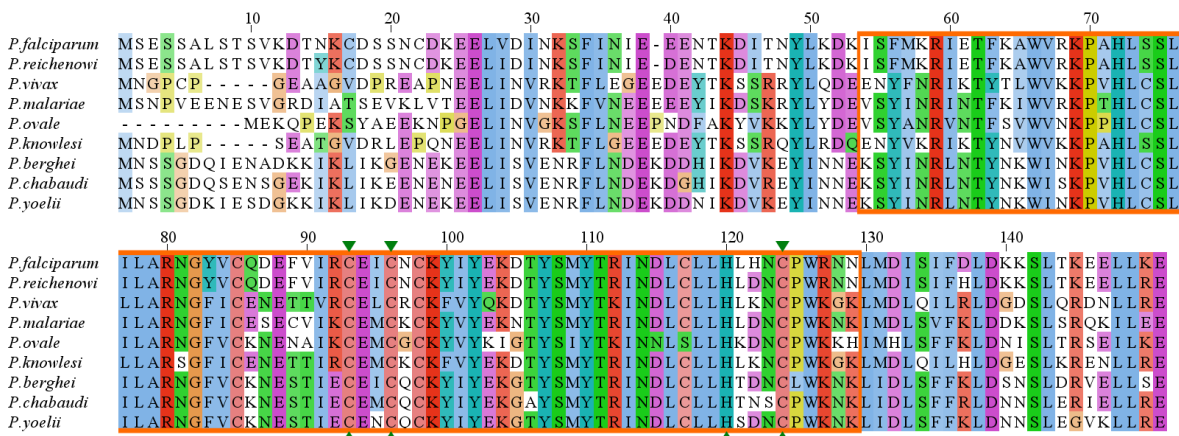
Figure 2.3.3: The BIR Domains of the *Plasmodium* IAP Orthologues
Amino acid alignment of the N-terminal end of the nine *Plasmodium* orthologues. Green arrows = inferred zinc-chelating residues; orange box = BIR domain.
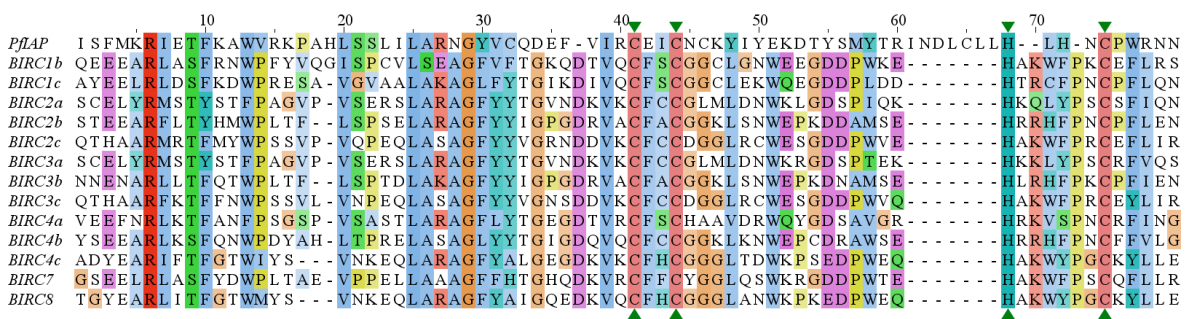
Figure 2.3.4: Human IAP Proteins and PfIAP
Amino acid alignment of the BIR domains of some of the human IAP proteins and the *P. falciparum* IAP protein. Green triangles = Zn^{2+}-chelating residues. a,b,c = the first, second and third BIR domains of the given human BIRC protein respectively.

Table 2.3.1: Amino Acid Comparison Between PfIAP BIR and Some Human BIR Domains

a,b,c = the first, second and third BIR domains of the given human BIRC protein respectively.

IAP Protein and BIR Domain	Identity (%)	Similarity (%)	IAP Protein and BIR Domain	Identity (%)	Similarity (%)
BIRC1b	16.3	31.5	BIRC3c	14.7	23.5
BIRC1c	25.9	46.9	BIRC4a	21.2	42.5
BIRC2a	11.9	16.1	BIRC4b	15.8	24.8
BIRC2b	24.7	36.5	BIRC4c	16.2	24.2
BIRC2c	14.7	20.2	BIRC7	14.6	25.2
BIRC3a	11.9	16.1	BIRC8	16.2	24.2
BIRC3b	24.1	37.3			

The consensus of various secondary structure algorithms predicts that the BIR domain of the PfIAP protein consists of three α-helices and three β-strands (Figure 2.3.5). This prediction is similar to various predictions for the three BIR domains of the human XIAP protein - three α-helices and two/three β-strands. These structures are present in the predicted 3D structure of the PfIAP BIR domain (Figure 2.3.6a) that was modelled on the human XIAP BIR1 domain (Figure 2.3.6b) using SWISS-MODEL. These two models look very similar: both have two α-helices, an antiparallel β-sheet made up of three β-strands, followed by a third α-helix. A zinc ion (grey dot) is shown to be chelated by four amino acids (blue residues) towards the C-terminal end of the XIAP BIR1 domain (Figure 2.3.6b). The corresponding amino acids (green triangles in Figure 2.3.4) in the PfIAP model (blue residues in Figure 2.3.6a) are in a similar spatial arrangement, but do not form a pocket of the exact same shape. This discrepancy can be explained by the inaccuracy of this PfIAP 3D model, with a QMEAN score less than -3 (the closer the QMEAN is to 1, the better the prediction). This discrepancy is also evident in the Phyre2 derived model (Figure 2.3.6c) in which only 62% of residues are modelled at >90% accuracy, primarily using the BIR3 domain of XIAP (PDB: 2VSL chain A), with a QMEAN score smaller than -4. The amino acid residues that were not accurately modelled occur after the β-sheet, which explains why the third α-helix of the Phyre2 model is longer and more rigid than the helices of the other model and the XIAP crystal structure, and has no zinc ion pocket (blue residues in Figure 2.3.6c). Notwithstanding the uncertainty associated with these two PfIAP models, they do conform to the general 3D structure that is expected in a BIR domain. The hydrophobic, IBM-binding groove of the human BIR domain is created by the conformational arrangement of the last strand of the β-sheet and the following α-helix (Herman *et al.* 2009). This same location in the PfIAP BIR domain contains a number of nonpolar amino acid residues that are similar in number and position to those found in the human BIR domain, suggesting that these domains may share the same hydrophobic characteristics.

a)

```
        53                                                                    128
        |                                                                      |
PfIAP   ISFMKRIETFKAWVRKPAHLSSLILARNGYVCQDEFVIRCEICNCKYIYEKDTYSMYTRINDLCLLHLHNCPWRNN
JPred4  --HHHHHHHH----------HHHHH---EE----EEEEEE--EEEEE----HHHHHHHHHHHHH---------
Phyre2  --HHHHHHHHHH--------HHHHHHH-------HHEEEE----EEEEE---HHHHHHHHHHHHHHHHH-----
PSIPRED -HHHHHHHHHHHH------HHHHHH--------EEEE----EEEEE----HHHHHHHHHH---------
SSpro   --HHHHHHHHHHH---HHHHHHHHH---EE---EEEEEEE---EEEEE----HHHHHHHHHHH---------
SWISS   --HHHHHHH-----------HHHHHH--EEE-----EEE----EE-----------HHHHHHHH---HHH--
        --HHHHHHHHHHH-------HHHHHH---EE----EEEEE----EEEEE----HHHHHHHHHHHH---------
```

b)
```
ISFMKRIETFKAWVRKPAHLSSLILARNGYVCQDEFVIRCEICNCKYIYEKDTYSMYTRINDLCLLHLHNCPWRNN
```

Figure 2.3.5: PfIAP BIR Domain Secondary Structure Prediction
a) Various secondary structure predictions of the PfIAP BIR domain using online tools. E = β-strand; H = α-helix. Numbers = amino acid number
b) Diagram showing the average secondary structure prediction of the BIR domain of the PfIAP protein. Gold arrow = β-strand; pink block = α-helix.

Figure 2.3.6: 3D Models of the PfIAP BIR Domain
a) Two views of the predicted 3D structure of the BIR domain of the PfIAP protein by SWISS-MODEL. Blue = inferred zinc-chelating residues; gold = β-strand; pink = α-helix. QMEAN = -3.49.
b) Visualisation of the crystal structure of the XIAP BIR1 domain (PDB: 4OXC chain A). Blue = zinc-chelating residues; gold = β-strand; grey dot = Zn^{2+} ion; pink = α-helix.
c) Two views of the predicted 3D structure of the PfIAP BIR domain by Phyre2. Blue = inferred zinc-chelating residues; gold = β-strand; pink = α-helix. QMEAN = -4.46.

2.3.2 Cloning and Expression of Recombinant PfIAP Proteins

2.3.2.1 *P. falciparum* Genomic DNA Extraction

Genomic DNA was extracted from *P. falciparum* 3D7 parasites with a concentration of 586 ng/μL and subsequently used for PCR. Spectrophotometric analysis revealed that the DNA had an A_{260}:A_{280} ratio of 1.89, which is indicative of DNA free from RNA or any contaminating proteins that tend to lower the absorbance value by absorbing 280 nm light. The A_{260}:A_{230} ratio of 1.94 was greater than the A_{260}:A_{280} ratio as expected and so is another

measure of the DNA purity (NanoDrop 2008). Agarose gel electrophoresis resolved this DNA sample as a single, high molecular weight band with no contaminating bands.

2.3.2.2 Creation of Various PfIAP Constructs

Three regions of the *pfiap* gene (BIR, Middle and Terminal; Figure 2.3.7a) were amplified by PCR as single bands of the expected size and then inserted into the pGEX-4T-2 vector (Figure 2.3.7b). The predicted β-sandwich domain (PlasmoDB v34) consists of a series of amino acids repeats and so this unstructured low complexity region (Aravind *et al.* 2003) was not selected for protein expression. The BIR region encompasses the area around the identified BIR domain, the Middle region covers the part of the gene between the BIR domain and the predicted β-sandwich domain, while the Terminal region represents the 3' end of the gene that follows the predicted β-sandwich domain. The linearized pGEX-4T-2 plasmid appeared slightly larger than expected, but sequencing did confirm that the GST coding sequence of the plasmid was as expected.
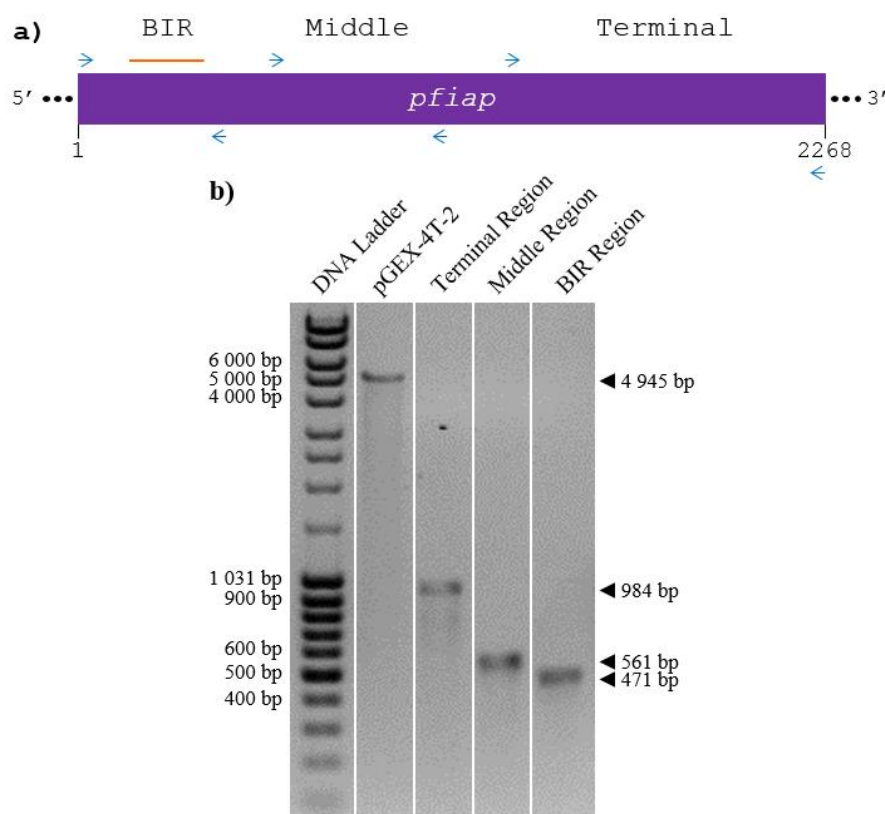

Figure 2.3.7: PCR Amplification of Three PfIAP Regions and the pGEX-4T-2 Plasmid
a) Diagram showing the locations of the three regions of the *pfiap* gene (BIR, Middle and Terminal) amplified by PCR for cloning into the pGEX-4T-2 vector. Blue = the forward and reverse PCR primers; orange = BIR domain; purple = *pfiap* gene. Numbers = nucleotide number. Amplified regions: BIR = 1 - 450 bp, Middle = 580 - 1 119 bp and Terminal = 1 297 - 2 259 bp.
b) Agarose gel (1%) showing three *Bam*HI/*Xho*I digested PfIAP PCR amplicons and the linearized pGEX-4T-2 vector prior to ligation. ◄ = expected band size.

After transformation, bacterial colonies were assessed for the presence of the three PfIAP constructs using colony PCR and restriction enzyme digestion of the extracted plasmids (Figure 2.3.8). The expected PCR amplicon sizes and digestion patterns were obtained for all three of the PfIAP clones, one sample of each was sent for sequencing. The sequencing results for the Middle and Terminal constructs were as expected, but an error was detected in the BIR region sequencing result. The expected codon sequence TTA was instead CTA (Figure 2.3.9), but as this mutation is silent and still allows for the translation of a leucine amino acid, the plasmid construct was retained.
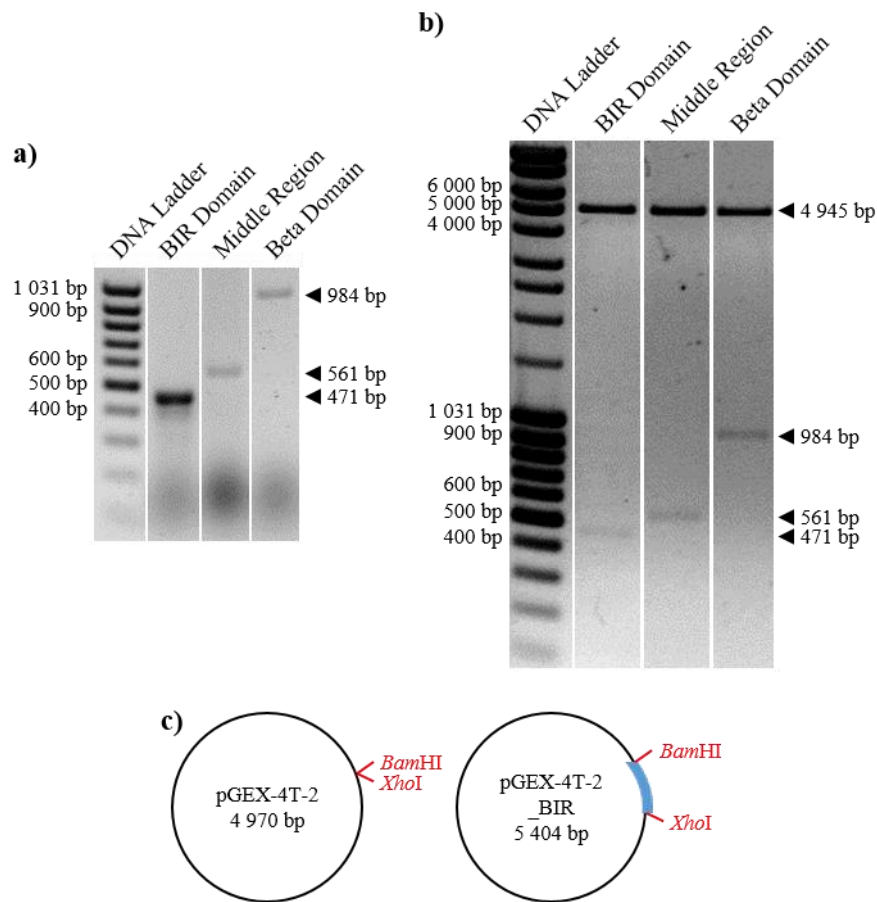

Figure 2.3.8: Verification of the Insertion of PfIAP Amplicons into the pGEX-4T-2 Vector.
a) Agarose gel (1%) showing colony PCR of the three PfIAP regions after cloning. ◄ = expected band size.
b) Agarose gel (1%) of the three pGEX-4T-2 plasmids after digestion with *Bam*HI/*Xho*I, resulting in two bands each. ◄ = expected band size.
c) Vector diagrams of the pGEX-4T-2 plasmid and the plasmid containing the PfIAP BIR region with *Bam*HI/*Xho*I cloning restriction sites. Blue = inserted BIR domain.

```
Wild Type:
        TyrThrArgIleAsnAspLeuCysLeuLeuHisLeu
    5'...TATACACGCATAAATGATTTATGCTTATTACATTTA...3'

BIR Actual:
    5'...TATACACGCATAAATGATCTATGCTTATTACATTTA...3'
        TyrThrArgIleAsnAspLeuCysLeuLeuHisLeu
```

T A T A C A C G C A T A A A T G A T C T A T G C T T A T T A C A T T T A

Figure 2.3.9: Sequencing Chromatogram of the pGEX-4T-2_BIR Plasmid
a) Extract of the sequencing chromatogram of the PfIAP BIR construct present in the pGEX-4T-2_BIR vector. Purple = the identified T→C substitution.

2.3.2.3   Expression and Purification of Recombinant GST-Tagged PfIAP

Rosetta™ 2 (DE3) *E. coli*, transformed with the different pGEX-4T-2 constructs, were grown in Overnight Express™ Instant TB Media to induce the expression of GST-tagged PfIAP proteins. It was not possible to purify the PfIAP Middle region (referred to as GST-Middle) as, in six different experiments, it only expressed as an insoluble protein when compared to uninduced *E. coli* (Figure 2.3.10), as predicted using the PROSO II and SOLpro online tools (solubility probabilities of 35% and 24% respectively). The protein appeared slightly smaller than expected, but this cannot be accurately judged as the induced fractions have not travelled the same distance through the polyacrylamide gel as the ladder (the dye front is not at the same level across the gel). Virtually no purified, GST-Middle protein was detected by Western blot (data not shown).
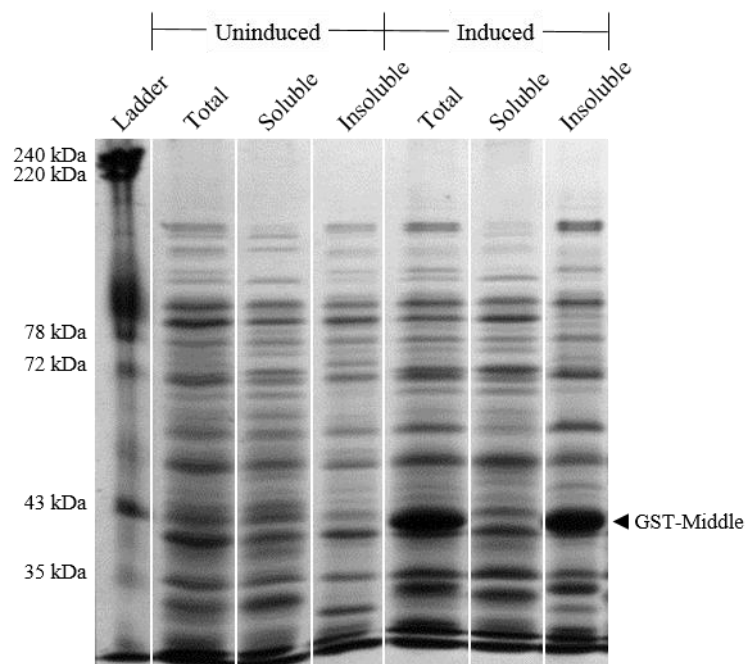
Figure 2.3.10: Expression of the Insoluble GST-Tagged PfIAP Middle Region
Coomassie-stained SDS-polyacrylamide gel (12%) of induced and uninduced GST-Middle aliquots
(2.5 µL of 1.5 mL). Ladder = human red cell membrane ladder. ◄ GST-Middle: predicted size =
48 kDa; observed size = ~42 kDa.

Although some of the PfIAP BIR region (GST-BIR) expressed as insoluble protein
(insoluble fraction; Figure 2.3.11a), sufficient amounts of soluble GST-tagged protein was
purified using MagneGST™ particles. At least 8 µg of eluted protein was obtained per
20 mL *E. coli* culture. As seen in the SDS-polyacrylamide gel and the anti-GST Western
Blot (Figure 2.3.11) the purified protein was of the expected size of 44 kDa, although other
smaller bands were detected. One of these bands is the same size as the GST positive control
protein (Figure 2.3.11b) and may be endogenous *E. coli* GST. It is assumed that the
<44 kDa, GST-positive bands are recombinant proteins that suffered premature mRNA
termination during translation. Two elutions were performed using 500 mM glutathione and
500 mM NaCl, but this was not sufficient to remove all of the GST-tagged protein from the
MagneGST™ particles (beads fraction in Figure 2.3.11; elution 2 not shown).
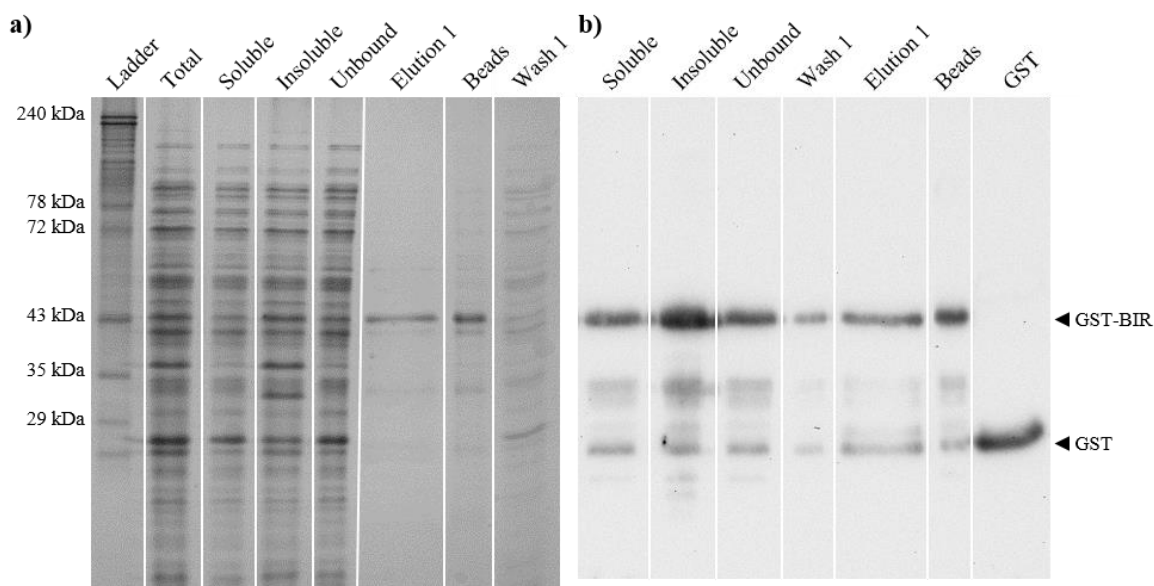
Figure 2.3.11: Expression and Purification of GST-Tagged PfIAP BIR Domain
a) Coomassie-stained SDS-polyacrylamide gel (12%) of the purification of GST-BIR. Total, soluble and insoluble *E. coli* lysates, unbound = 1.0 μL of 1.5 mL; elution 1, protein remaining on beads = 10.0 μL of 50 μL; wash 1 = 15.0 μL of 1.5 mL. Ladder = human red cell membrane ladder. ◄ GST-BIR: predicted/observed size = 44 kDa.
b) Anti-GST western blot of the purification of GST-BIR. Soluble and insoluble *E. coli* lysates, unbound = 2.5 μL of 1.5 mL; wash 1 = 5.0 μL of 1.5 mL; elution 1 = 10.0 μL of 100 μL; protein remaining on beads = 5.0 μL of 100 μL. ◄ GST-BIR: predicted/observed size = 44 kDa; GST: known size = 26 kDa.

The GST-Terminal protein (PfIAP Terminal region), which was predicted to have a 50% to 80% chance of being insoluble by PROSO II and SOLpro, expressed as both soluble and insoluble proteins and was successfully purified. When analysed on a SDS-polyacrylamide gel the GST-tagged protein appeared to have a larger molecular mass than expected (~74 kDa vs 65 kDa; Figure 2.3.12a). This discrepancy in the expected migration pattern has previously been documented for proteins characterised by regions of low complexity, a number of which are present in the expressed region (Figure 2.3.13; Tompa 2003). Other bands are present in the eluates but are not visible on the Western Blot (Figure 2.3.12b), implying that these bands are not truncated forms of the protein but rather other *E. coli* proteins. For each 20 ml *E. coli* grow ~10 μg of soluble GST-Terminal protein was obtained, but more than this remained bound to the magnetic beads despite using a high salt concentration elution buffer (500 mM glutathione + 500 mM NaCl; beads fraction; Figure 2.3.12a).
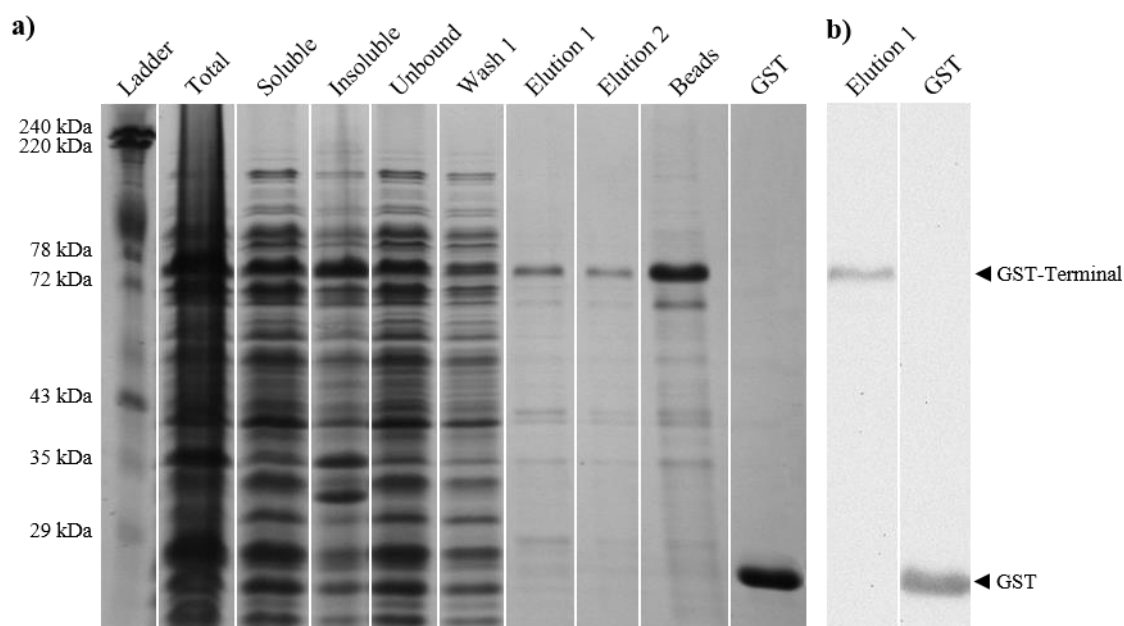
Figure 2.3.12: Expression and Purification of GST-Tagged PfIAP Terminal Region
a) Coomassie-stained SDS-polyacrylamide gel (12%) of the purification of GST-Terminal. Total, soluble and insoluble *E. coli* lysates, unbound = 2.5 μL of 1.5 mL; wash 1 = 5.0 μL of 1.5 mL; elution 1, elution 2, protein remaining on beads = 10.0 μL of 50 μL. Ladder = human red cell membrane ladder. ◄ GST-Terminal: predicted size = 65 kDa; observed size = ~74 kDa; GST = 26 kDa.
b) Anti-GST western blot of the purified GST-Terminal. Elution 1 = 10.0 μL of 100 μL. ◄ GST-Terminal: predicted size = 65 kDa; observed size = ~74 kDa; GST: known size = 26 kDa.
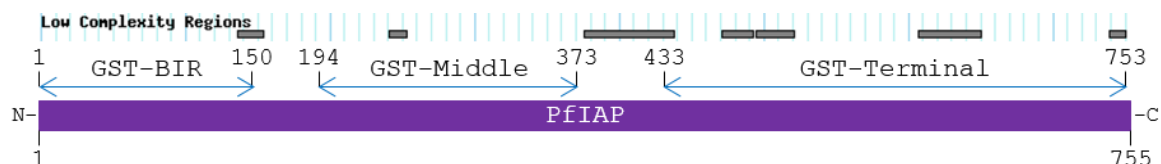

Figure 2.3.13: Low Complexity Regions of PfIAP
Diagram showing regions of low complexity in the PfIAP protein. Blue = the regions expressed as GST-tagged proteins; purple = PfIAP protein. Numbers = amino acid number. Adapted from PlasmoDB (Aurrecoechea *et al.* 2009).

2.3.3   *P. falciparum* Phage Display Library Creation

2.3.3.1   *P. falciparum* mRNA Isolation

The combined amount of total RNA extracted from ten 35 mL, mixed stage *P. falciparum* culture flasks was 923 μg. The A_{260}:A_{280} ratio of the various RNA extractions ranged from 1.45 to 2.13. Since the expected value for pure RNA is >2 (NanoDrop 2008) this indicates that some of the total RNA samples were contaminated, most likely with the protein containing phase of the TRI Reagent®. Agarose gel electrophoresis showed the predominant 28S and 18S and the small 5S rRNA subunits as expected (Figure 2.3.14).

Dynabeads® were used to isolate 9 µg of mRNA or 1% of the total available RNA, which is acceptable as a typical eukaryotic cell has between 1-5% mRNA (Lodish *et al.* 2000).
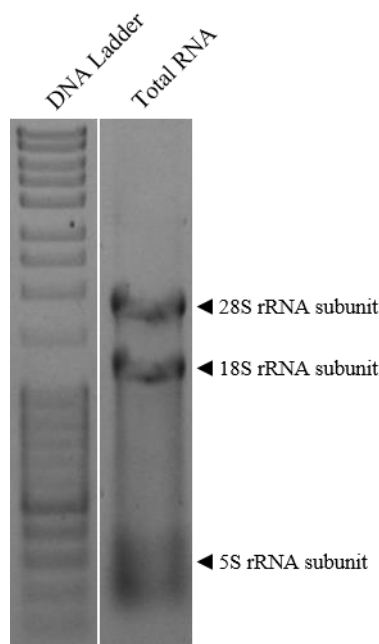

Figure 2.3.14: Extracted *P. falciparum* RNA
Agarose gel (1%) showing a sample of total RNA extracted from *P. falciparum* parasites.

2.3.3.2   *P. falciparum* Phage Display Library Construction

The isolated mRNA was reverse transcribed, before *Eco*RI/*Hind*III Directional Linkers were ligated to the *P. falciparum* cDNA. The average size of the DNA fragments present was determined by visualising an aliquot of the sample by agarose gel electrophoresis (~300 bp – ~800 bp; data not shown). These fragments were then inserted into the T7 vector arms containing the biotin cassette and packaged. A plaque assay was performed to determine the number of recombinant T7 bacteriophage generated by the packaging reaction (Table 2.3.2). Using 0.5 µg of vector DNA for the 0.3 mL packaging reaction resulted in the creation of 7.32×10^5 bacteriophage (or plaque forming units).

Table 2.3.2: Biotin-Tagged *P. falciparum* Phage Display Library Numbers

Phage Titre (pfu/ml)	Library Size (pfu)	Packaging Efficiency (pfu/µg)
2.44×10^6	7.32×10^5	1.46×10^6

A plate amplification step was used to amplify the number of bacteriophage present in the library as less than 5×10^6 phage were present in the library. The titre of the phage display library increased to 6.49×10^{11} pfu/mL after plate amplification. PCR amplification of randomly selected plaques from the library showed a range of sizes from 237 bp (an empty cassette) to 550 bp (Figure 2.3.15).
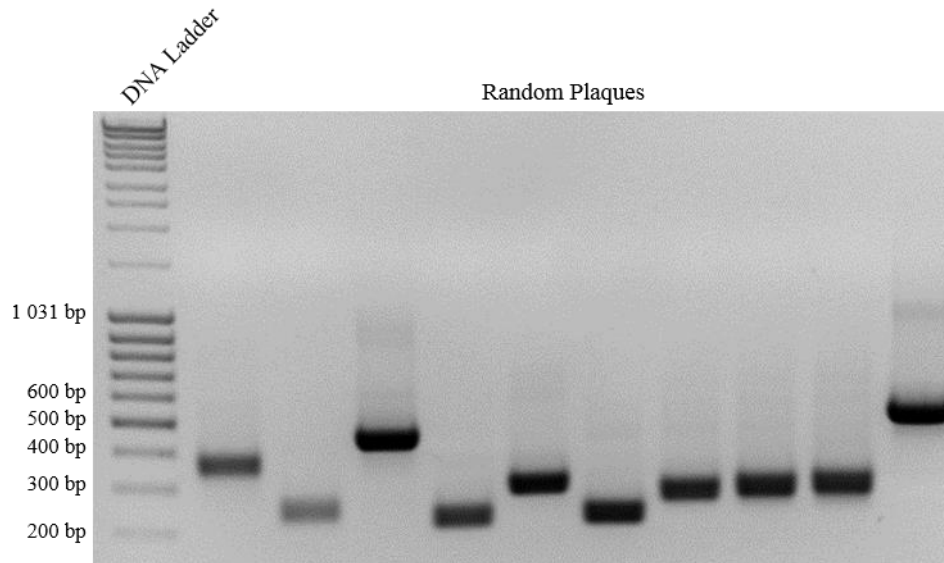
Figure 2.3.15: PCR Amplification of Random Plaques from the Biotin-Tagged *P. falciparum* Phage Display Library

Agarose gel (1%) showing the PCR products of a random selection of bacteriophage taken after plate amplification of the initial library.

Streptavidin magnetic particles were used to select bacteriophage from the newly created phage display library expressing the biotinylated protein and therefore in-frame *P. falciparum* peptides. Phage were applied to these particles so that a multiplicity of infection of 5 was achieved (5 bacteriophage representing each unique *P. falciparum* insert). Following streptavidin isolation, the biotin phage were amplified to a titre of 1.86×10^{10} pfu/mL.

2.3.4 Identification of Novel PfIAP Binding Partners

The *P. falciparum* phage display library was biopanned against both the GST-BIR and GST-Terminal recombinant proteins. After four rounds of biopanning, a plaque lift using an anti-biotin antibody was performed (Figure 2.3.16). Biotin-positive plaques were then selected for PCR using primers specific to the T7 phage arms (Figure 2.3.17). A range of differently sized PCR amplicons over 300 bp in size were obtained. The presence of multiple, similarly sized bands (not shown) suggests enrichment and amplification of the same specific phage which is what is expected after four rounds of biopanning. A number of the larger phage inserts were sequenced (see APPENDIX C for an example) and the *P. falciparum* inserts were then translated to their amino acid sequences. The identity of these perspective binding partners was determined by using search and alignment facilities of PlasmoDB (Table 2.3.3).

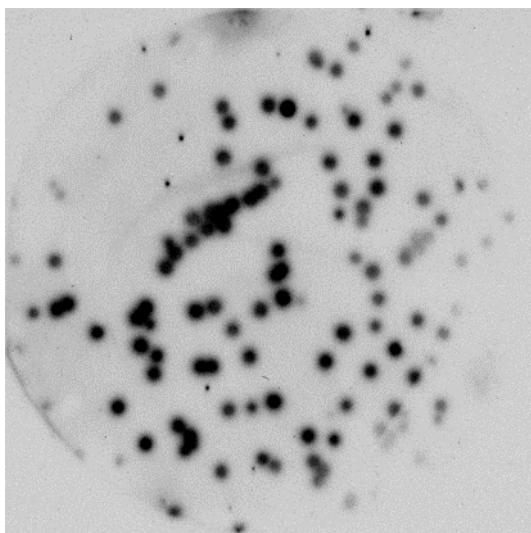**Figure 2.3.16: Plaque Lift Detecting Bacteriophage Expressing Biotin**
A plaque lift using an anti-biotin antibody on a 10^{-6} dilution plaque assay plate of the phage display library after streptavidin screening.
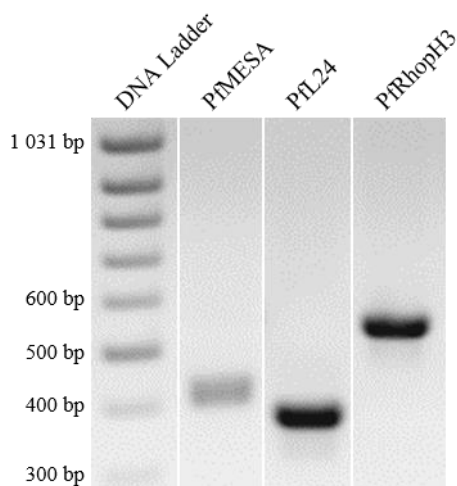

Figure 2.3.17: PCR of Three Selected Plaques after Four Rounds of Biopanning
Agarose gel (1%) showing the PCR products of some of the bacteriophage plaques after biopanning that were selected for sequencing.

Table 2.3.3: GST-Tagged PfIAP Binding Partners*

PlasmoDB Name	Gene I.D.	Peptide Length (aa)
GST-BIR		
High molecular weight rhoptry protein 3 (RhopH3)	PF3D7_0905400	101
60S ribosomal protein L24, putative	PF3D7_1309100	50
Double C2-like domain-containing protein (DOC2)	PF3D7_1243900	49
LisH domain-containing protein, putative (LisH)	PF3D7_1303400	34
GST-Terminal		
Chromodomain-helicase-DNA-binding protein 1 homolog, putative (CHD1)	PF3D7_1023900	65
Mature parasite-infected erythrocyte surface antigen (MESA)	PF3D7_0500800	63

*Additional details are available in APPENDIX D.

Four unique binding partners were identified for the GST-BIR domain, while GST-Terminal had two unique binding partners (Figure 2.3.18). PfRhopH3 is in the family of proteins that includes PfRhopH1/Clag and PfRhopH2, while PF3D7_1309100 is the L24 component of the 60S ribosomal subunit. PfDOC2 contains two C2 domains that act as Ca^{2+} binding motifs and one Lissencephaly type-1-like homology (LisH) motif is present at the N-terminal end of PF3D7_1303400. PfCHD1 has a helicase domain and motifs that are involved in chromatin organization, while the C-terminal end of PfMESA contains a DnaJ molecular chaperone homology domain.
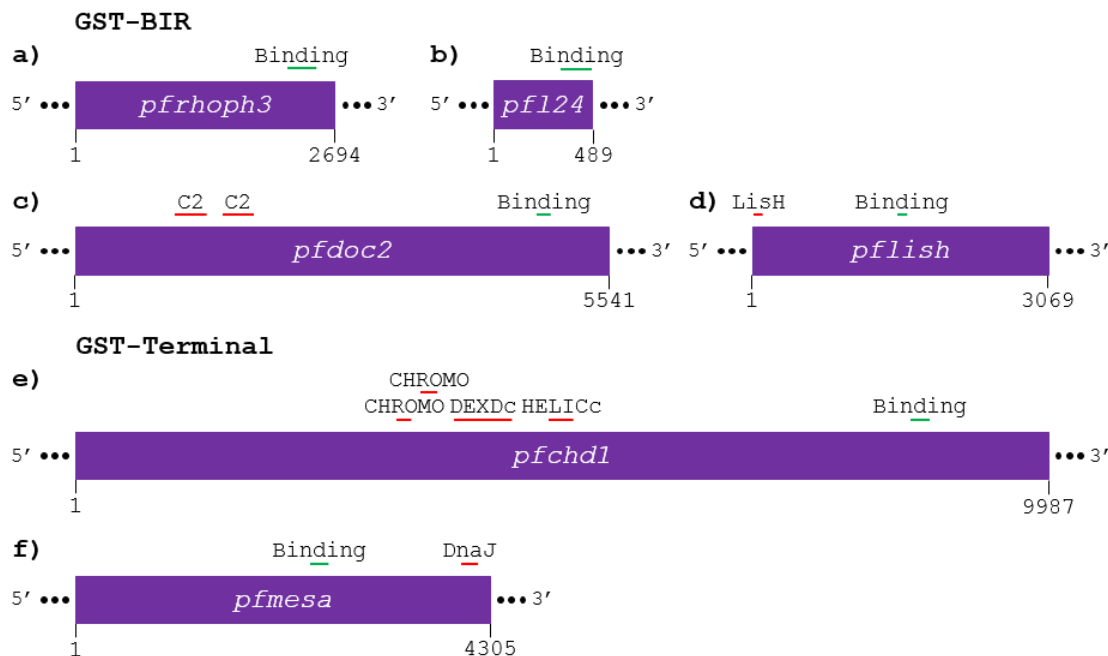
Figure 2.3.18: The PfIAP Binding Partners Identified Using a Phage Display Library
Diagram showing the gene coding sequence (purple) and the region that bound to the GST-PfIAP proteins (green). Red = known gene domains. Numbers = nucleotide number.

a) The *pfrhoph3* (PF3D7_0905400) gene.   **d)** The *pflish* (PF3D7_1303400) gene.
b) The *pfl24* (PF3D7_1309100) gene.   **e)** The *pfchd1* (PF3D7_1023900) gene.
c) The *pfdoc2* (PF3D7_1243900) gene.   **f)** The *pfmesa* (PF3D7_0500800) gene.

No common binding motif was present across the binding regions of the four GST-BIR partners, nor across the two GST-Terminal binding partners (Figure 2.3.19a). Analysis of the amino acid composition of these binding regions revealed that the five charged amino acids (D, E, H, K and R) made up 52 – 63% of the total amino acids present in these peptides (except for PfRhopH3 with 35%; Figure 2.3.19b and Table 2.3.4). The GST-BIR and GST-Terminal proteins only contain 33 – 34% charged amino acids, with a fairly equal distribution of positive and negative residues.
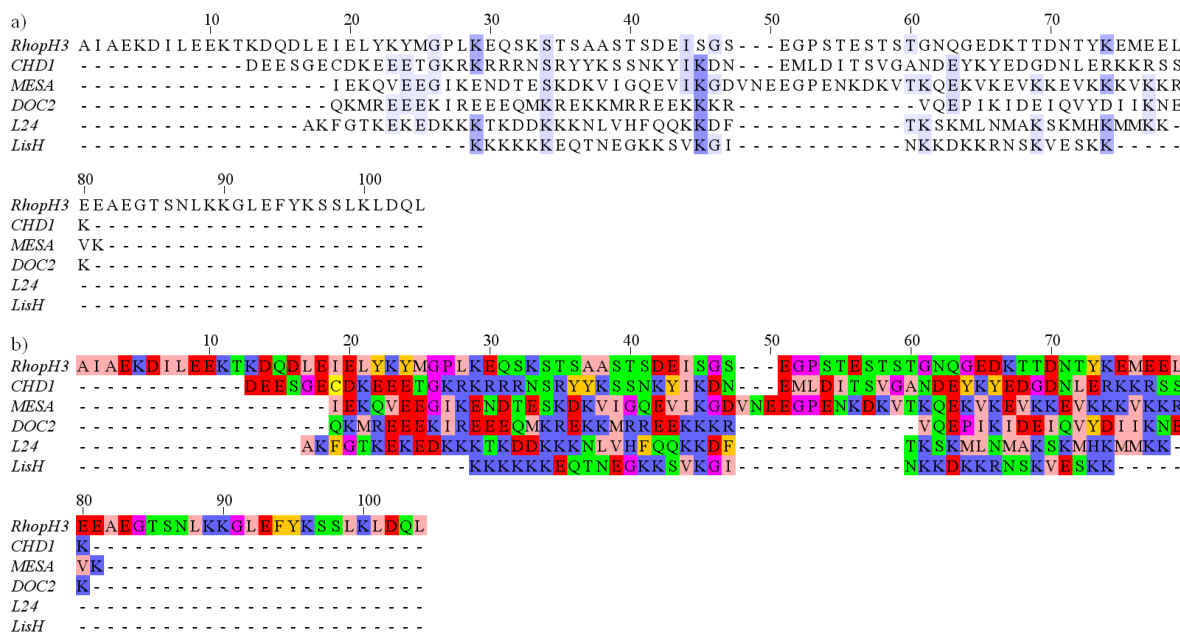
Figure 2.3.19: Amino Acid Alignment of the Binding Regions of the PfIAP Partner Proteins
Amino acid alignment of the six PfIAP binding regions.
a) BLOSUM62 colour scheme. Percentage identity across all the sequences determines the shade of blue.
b) Zappo colour scheme. Blue = positive; green = hydrophilic; magenta = conformationally special; pink = aliphatic/hydrophobic; red = negative; yellow = aromatic.

Table 2.3.4: Percentage of Charged Amino Acids Present in Each Binding Peptide

Binding Partner	Basic Residues (%)			Acidic Residues (%)		Total Charged Residues (%)
	Arg	His	Lys	Asp	Glu	
GST-BIR						
PfDOC2	12.2	0.0	22.5	4.1	24.5	63.3
PfLisH	2.9	0.0	47.1	2.9	8.8	61.7
PfL24	0.0	4.0	38.0	8.0	4.0	54.0
PfRhopH3	0.0	0.0	11.9	6.9	16.8	35.6
GST-Terminal						
PfMESA	1.6	0.0	28.6	6.3	19.0	55.5
PfCHD1	10.8	0.0	15.4	10.8	15.4	52.4

The secondary structures of the recombinant PfIAP proteins revealed that α-helices are the predominantly predicted structure with only some β-strands predicted to be present (Figure 2.3.20 and Figure 2.3.21). The same observation is true when looking at the secondary structures of the six binding regions (Figure 2.3.20 and Figure 2.3.21). Across both groups of proteins the charged amino acids are evenly distributed, with most of the α-helices containing both positive and negative residues (Figure 2.3.20 and Figure 2.3.21). Biopanning occurred at a pH of 7.5, which means that both GST-tagged PfIAP proteins were predicted to be negatively charged, while the binding peptides are positively charged at this pH (except for PfCHD1 which is neutral and PfRhopH3 which is negative). However, these predicted charges do not take into account the 3D conformation of the peptides, which could alter the overall charge. The lack of binding motif and the predominant presence of charged basic and acidic residues imply that these interactions occurred via ionic bonds between the α-helices of the respective proteins.

Figure 2.3.20: Secondary Structures of GST-BIR and the Binding Regions of its Partners
Diagrams showing the predicted secondary structure of GST-BIR and its binding partners. Gold arrow = β-strand; pink block = α-helix; gold/pink shapes = both α-helix and β-strand predictions. Blue = positive amino acids; red = negative amino acids.

GST-Terminal
```
QHQNQKDTLEYIPEHTILSKMSDRLRETYDKYYIDIARNNKNKEQKQNEEKENNEVNKEVYIERNNNKEETSEKE

ENSKNEKFNRSNKASQSDSYEENAEYSQSDGITNTESNFENSQKEREGFSFKWKFKNLFSSQSIHQDVNDKHNDK

LSCEEDREHILEDMKEIIFLDKMENNKADNENNNNNNNNNNNTYSNSNDSNCNDSNCNDSNSNNDDNNVENHLKNY

LKKCIGEMCLNICPYLENKNYFIKSLYDYYVLLTSEMSTTHINENAIKNSFIIRPFNVVENHRIYCPYITEDLYL

FSKITKLIFELLILEFERKYV
```

PfCHD1
```
DEESGECDKEEETGKRKRRRNSRYYKSSNKYIKDNEMLDITSVGANDEYKYEDGDNLERKKRSSK
```

PfMESA
```
IEKQVEEGIKENDTESKDKVIGQEVIKGDVNEEGPENKDKVTKQEKVKEVKKEVKKKVKKRVK
```

Figure 2.3.21: Secondary Structures of GST-Terminal and the Binding Regions of its Partners
Diagrams showing the predicted secondary structure of GST-Terminal and its binding partners. Gold arrow = β-strand; pink block = α-helix; gold/pink shapes = both α-helix and β-strand predictions. Blue = positive amino acids; red = negative amino acids.

2.3.5   Cloning and Expression of Recombinant Binding Partners

2.3.5.1   Cloning of His-Tagged Binding Partners

Since the BIR domain is the principal PfIAP region, it was important to verify the novel protein interactions with this domain. Three of the interesting GST-BIR binding partners (PfDOC2, PfLisH and PfRhopH3) were selected and cloned into the pET-15b plasmid. They were expressed as N-terminal His-tagged recombinant proteins, but sufficient amounts of purified PfLisH could not be obtained to perform a binding assay, so only the other two binding partners selected for further investigation.

pfdoc2 consists of six exons and the sequence that was selected for PCR amplification is towards the 3' end of the gene and includes the identified binding region. Almost all of the last exon (out of seven) of the *pfrhoph3* gene, which contains the binding region, was selected for PCR amplification (Figure 2.3.22).
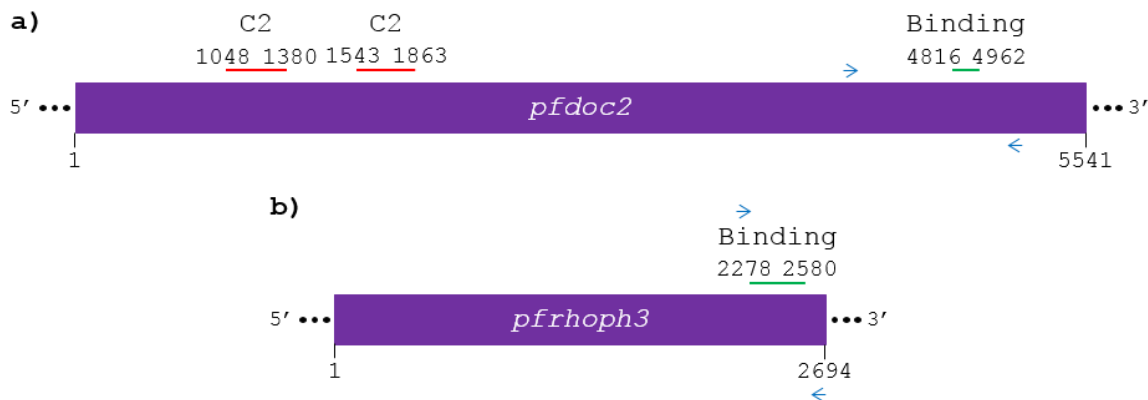
Figure 2.3.22: Binding Partners of PfIAP BIR
Diagram showing the location of the PCR primers on either side of the binding region identified from the phage display library. Blue = the forward and reverse PCR primers; green = the binding region; purple = gene coding sequence. Numbers = nucleotide number.
a) The *pfdoc2* (PF3D7_1243900) gene. Red = the location of the C2 domains. Amplified region = 4 210 - 5 193 bp.
b) The *pfrhoph3* (PF3D7_0905400) gene. Amplified region = 2 209 - 2 694 bp.

The two binding regions were PCR-amplified as single bands of the expected size (Figure 2.3.23), cloned into the pET-15b vector and used to transform DH5α *E. coli* cells.
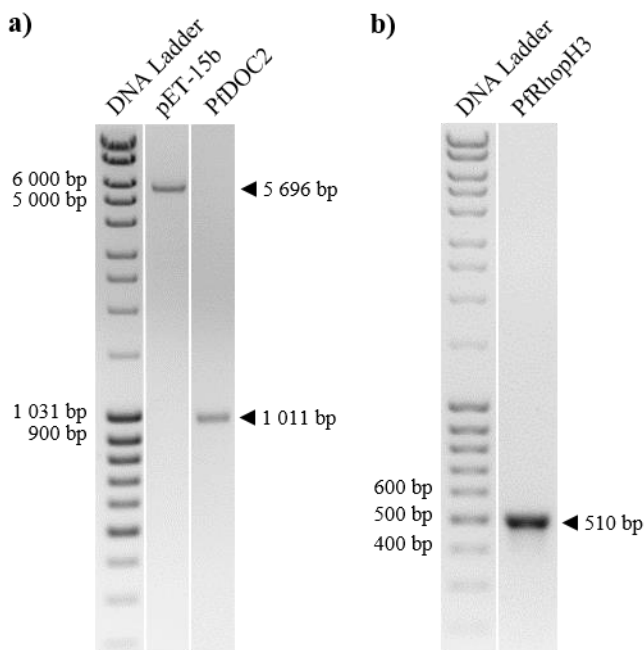
Figure 2.3.23: PCR Amplicons of the Binding Partners and pET-15b Plasmid prior to Ligation
Agarose gel (1%) showing PCR amplicons after digestion with *Bam*HI/*Nde*I and the linearized pET-15b vector. ◄ = expected band size.

Colony PCR, using pET-15b specific primers which result in amplicons larger than the initial cloned region, was used to confirm the presence of the inserts after transformation (Figure 2.3.24a), as was digestion of the new pET-15b constructs with *Eco*RV (Figure 2.3.24b) that

resulted in the expected digestion patterns. Sequencing of one sample of each insert verified that the cloning was error-free. These verified plasmids were used to transform Rosetta™ 2 (DE3) *E. coli*, which allowed for expression of the recombinant binding proteins.
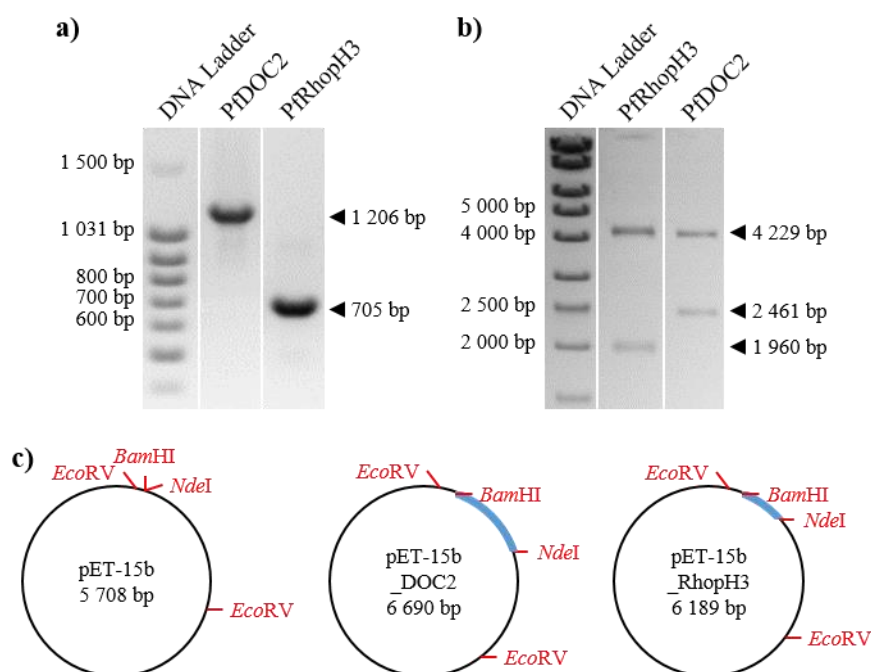

Figure 2.3.24: Verification of the Insertion of Binding Amplicons into the pET-15b Vector.
a) Agarose gel (1%) showing colony PCR of the two binding partners after cloning. ◄ = expected band size.
b) Agarose gel (1%) of the two pET-15b plasmids after digestion with *Eco*RV. ◄ = expected band size.
c) Vector diagrams of the pET-15b plasmid and the two plasmids containing the binding partners with *Bam*HI/*Nde*I and *Eco*RV restriction sites. Blue = inserted binding regions.

2.3.5.2 Expression of Recombinant His-Tagged Binding Partners

Overnight Express™ Instant TB Media was used to induce the expression of the His-tagged binding proteins. His-DOC expressed easily as a soluble protein, as predicted by the online tool SOLpro (85%), while PROSO II predicted that the protein would only have a 48% chance of being soluble. The protein was larger than the expected molecular size of 44 kDa (Figure 2.3.25a), which could be explained by the presence of low complexity regions in the expressed region (Figure 2.3.26). Slightly truncated His-DOC was also purified, as verified by the anti-His western blot (Figure 2.3.25b). Approximately 25 µg of the His-tagged protein was obtained per 20 mL *E. coli* culture.
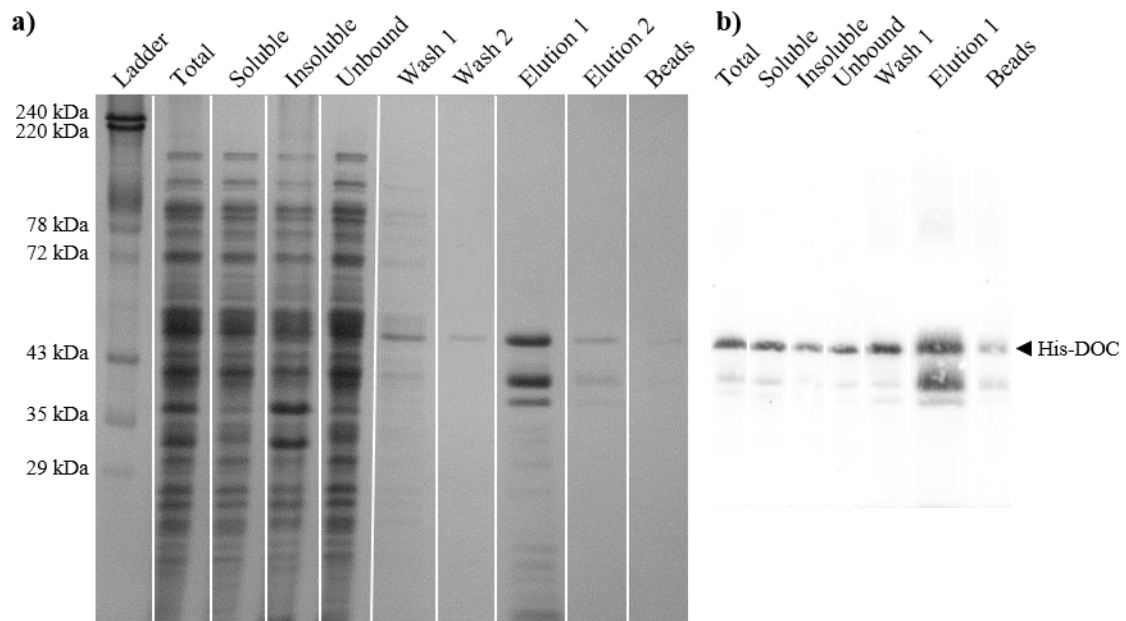
Figure 2.3.25: Expression and Purification of His-Tagged PfDOC2 Binding Region
a) Coomassie-stained SDS-polyacrylamide gel (12%) of the purification of His-DOC. Total, soluble and insoluble *E. coli* lysates, unbound = 1.5 µL of 1.5 mL; wash 1, wash 2 = 15.0 µL of 1.5mL; elution 1, elution 2, protein remaining on beads = 10.0 µL of 100 µL. Ladder = human red cell membrane ladder. ◄ His-DOC: predicted size = 44 kDa; observed size = ~50 kDa.
b) Anti-His western blot of the purification of His-DOC. Total, soluble and insoluble *E. coli* lysates, unbound = 1.0 µL of 1.5 mL; wash 1 = 15.0 µL of 1.5 mL; elution 1, protein remaining on beads = 10.0 µL of 100 µL. ◄ His-DOC: predicted size = 44 kDa; observed size = ~50 kDa.

Figure 2.3.26: Low Complexity Regions of PfDOC2
Diagram showing regions of low complexity in the PfDOC2 protein. Blue = the region expressed as a His-tagged protein; purple = PfDOC2 protein. Numbers = amino acid number. Adapted from PlasmoDB (Aurrecoechea *et al.* 2009).

As seen when comparing the soluble and insoluble fractions of His-Rhop (Figure 2.3.27a), the protein is soluble (as predicted by both online tools) but resolved at a higher molecular weight on the SDS-polyacrylamide gel than was expected likely due to the presence of two regions of low complexity (Figure 2.3.28). At least 15 µg of eluted protein was obtained per 20 mL *E. coli* grow. An anti-His western blot confirmed that this protein, and the slightly smaller truncated proteins, were His-tagged proteins (Figure 2.3.27b).
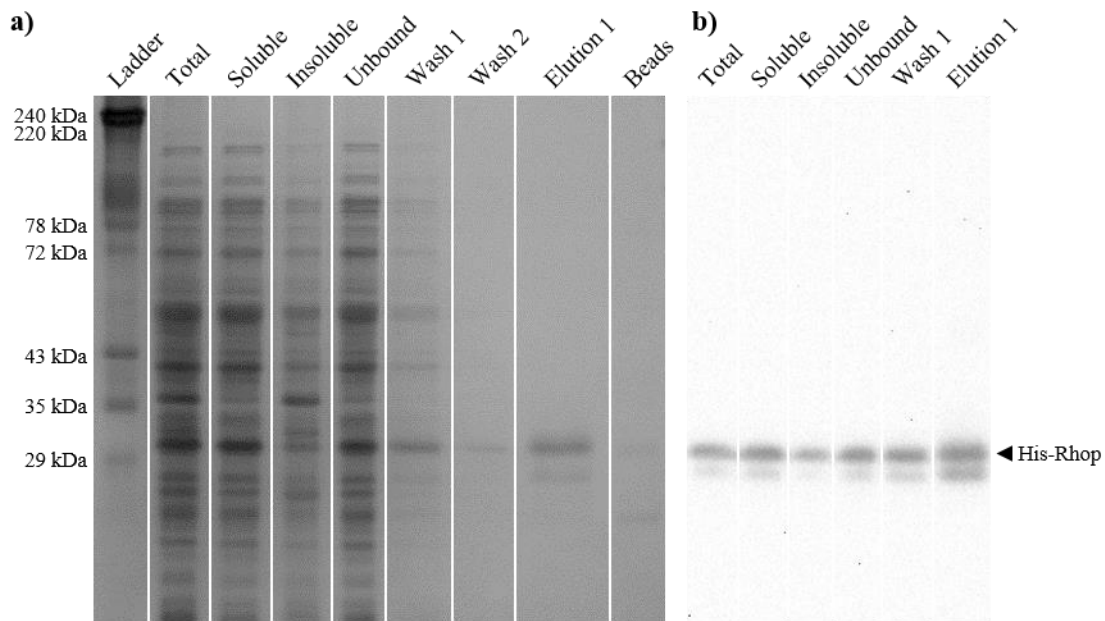
Figure 2.3.27: Expression and Purification of His-Tagged PfRhopH3 Binding Region
a) Coomassie-stained SDS-polyacrylamide gel (12%) of the purification of His-Rhop. Total, soluble and insoluble *E. coli* lysates, unbound = 1.0 μL of 1.5 mL; wash1, wash 2 = 15.0 μL of 1.5mL; elution 1, protein remaining on beads = 10.0 μL of 100 μL. Ladder = human red cell membrane ladder. ◀ His-Rhop: predicted size = 21 kDa; observed size = ~30 kDa.
b) Anti-His western blot of the purification of His-Rhop. Total, soluble and insoluble *E. coli* lysates, unbound = 1.0 μL of 1.5 mL; wash 1 = 15.0 μL of 1.5 mL; elution 1 = 10.0 μL of 100 μL. ◀ His-Rhop: predicted size = 21 kDa; observed size = ~30 kDa.

Figure 2.3.28: Low Complexity Regions of PfRhopH3
Diagram showing regions of low complexity in the PfRhopH3 protein. Blue = the region expressed as a His-tagged protein; purple = PfRhopH3 protein. Numbers = amino acid number. Adapted from PlasmoDB (Aurrecoechea *et al.* 2009).

2.3.6 *In vitro* Binding Assays between GST-BIR and its Binding Partners

As a means of verifying the biopanning results and to eliminate the possibility of having identified a false binding partner, two and three *in vitro* binding assays were performed using the His-Rhop and His-DOC proteins respectively. A constant amount of both of the His-tagged binding partners was immobilised on MagneHis™ particles, before being subjected to increasing amounts of the GST-BIR protein. As seen in the representative SDS-polyacrylamide gels and verified graphically (Figure 2.3.29 and Figure 2.3.30), there was a

dose-dependent association between the proteins that reached saturation, with PfIAP binding PfDOC2 and PfRhopH3 with a stoichiometry of approximately 1:3 and 1:5 respectively. Controls ensured the specificity of these binding results. Heat denatured GST-BIR showed substantially reduced binding to both the immobilized binding partners (the last lane in both Figure 2.3.29a and Figure 2.3.30a), while free GST showed minimal interaction with the His-tagged proteins (not shown).

Figure 2.3.29: Binding Assay between GST-BIR and His-DOC Binding Partner
a) Coomassie-stained SDS-polyacrylamide gel (12%) of the *in vitro* binding assay between GST-BIR and His-DOC. Lane 2 = ~0.5 μg free GST-BIR; lane 3 = ~0.5 μg free His-DOC; lanes 4-8 = 0.5 μg immobilised His-DOC/0.2 – 1.0 μg GST-BIR; lane 9 = 0.5 μg immobilised His-DOC/0.5 μg denatured GST-BIR. Ladder = human red cell membrane ladder. ◄ GST-BIR: observed size = ~44 kDa; His-DOC: observed size = ~50 kDa.
b) Graphs showing the dose-dependent binding of GST-BIR to immobilised His-DOC, following densitometry of the SDS-polyacrylamide gel.
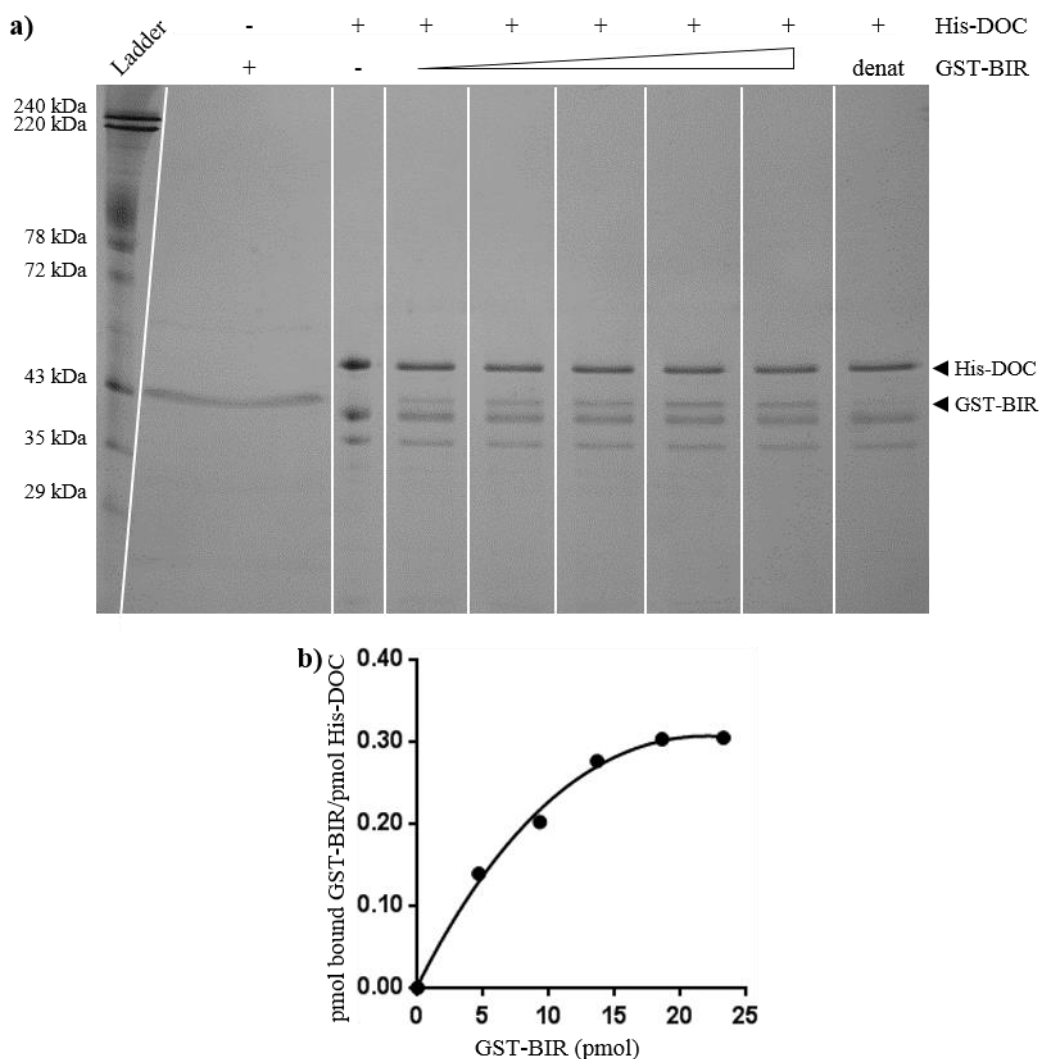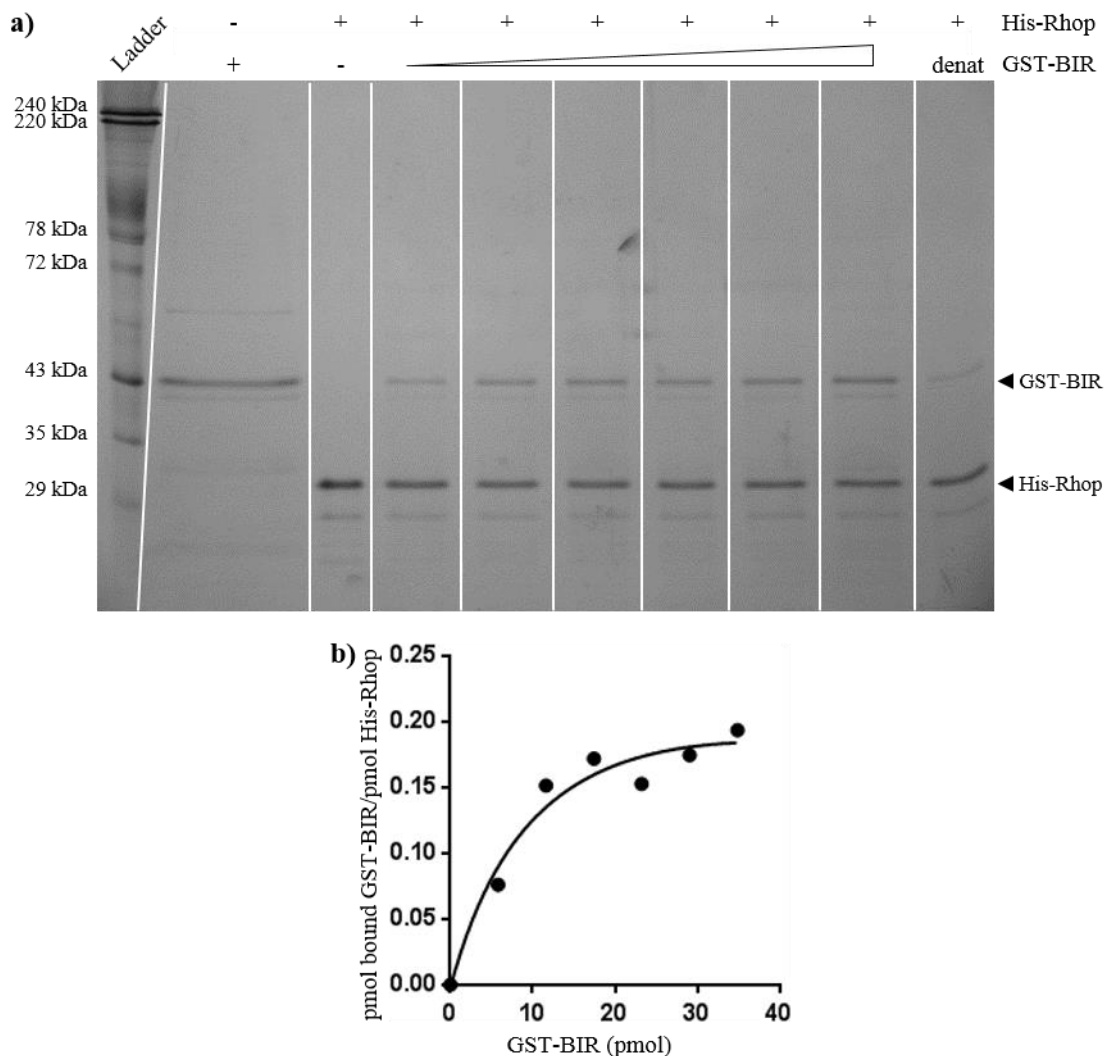
Figure 2.3.30: Binding Assay between GST-BIR and His-Rhop Binding Partner
a) Coomassie-stained SDS-polyacrylamide gel (12%) of the *in vitro* binding assay between GST-BIR and His-Rhop. Lane 2 = ~0.5 µg free GST-BIR; lane 3 = ~0.5 µg free His-Rhop; lanes 4-8 = 0.5 µg immobilised His-Rhop/0.25 – 1.5 µg GST-BIR; lane 9 = 0.5 µg immobilised His-Rhop/0.5 µg denatured GST-BIR. Ladder = human red cell membrane ladder. ◀ GST-BIR: observed size = ~44 kDa; His-Rhop: observed size = ~30 kDa.
b) Graphs showing the dose-dependent binding of GST-BIR to immobilised His-Rhop, following densitometry of the SDS-polyacrylamide gel.

2.3.7 Episomal Expression of PfIAPbir-GFP

2.3.7.1 Creation of the pARL2-GFP_PfIAPbir Vector

Determining the localisation of the PfIAP protein in the *P. falciparum* parasite is an important step in the characterisation of the protein. The 5' region of the *pfiap* gene containing the BIR domain (Figure 2.3.31a) was amplified by PCR, without a STOP codon, as a single band of the expected size and prepared for ligation into the pARL2-GFP vector (Figure 2.3.31b).
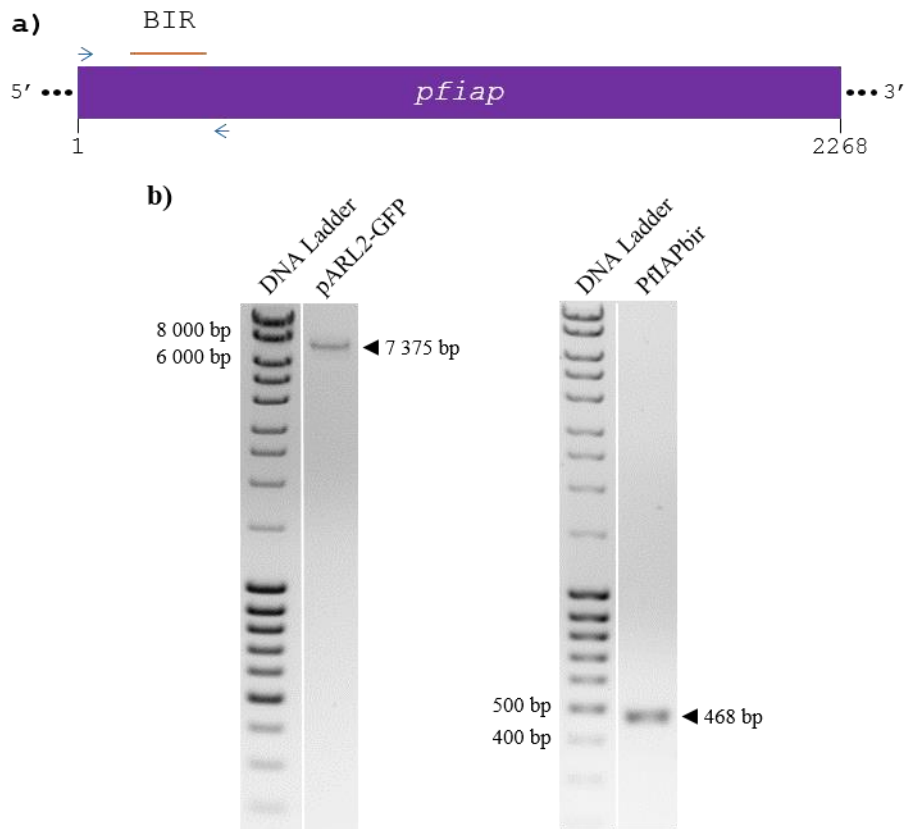
Figure 2.3.31: The pARL2-GFP Plasmid and PfIAPbir PCR Amplicon
a) Diagram showing the location of the region of the *pfiap* gene amplified by PCR for cloning into the pARL2-GFP vector. Blue = the forward and reverse PCR primers; orange = BIR domain; purple = *pfiap* gene. Numbers = nucleotide number.
b) Agarose gel (1%) showing the *Avr*II/*Xho*I digested pARL2-GFP vector and PfIAPbir PCR amplicon prior to ligation. ◄ = expected band size.

After *E. coli* transformation, the pARL2-GFP_PfIAPbir construct was verified by colony PCR and restriction enzyme digestion. PCR was performed using pARL2-GFP specific primers, which generate an amplicon larger than the initial cloned region, and resulted in a strong band around 600 bp (Figure 2.3.32a). Digestion analysis of the new pARL2-GFP_PfIAPbir vector, using *Eco*RI, showed two bands at the expected sizes (Figure 2.3.32b). DNA sequencing of the insert indicated that the PfIAPbir sequence was error-free and suitable for *P. falciparum* transfection.
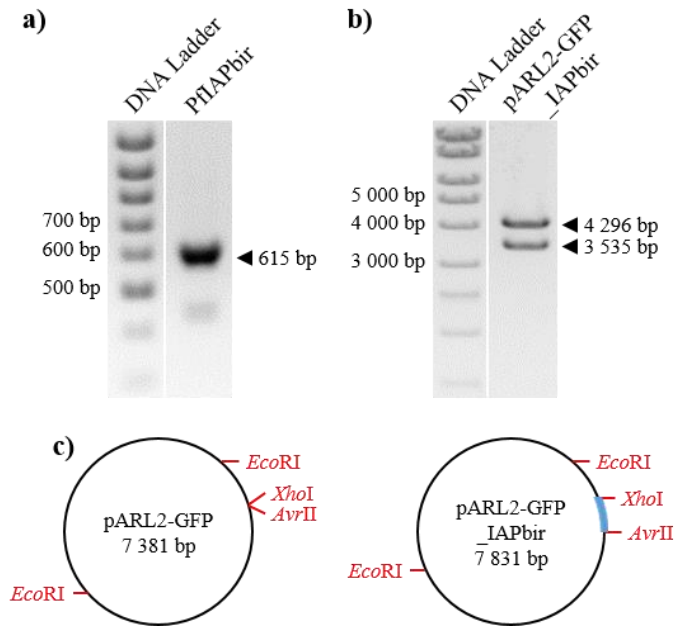
Figure 2.3.32: Verification of the Insertion of PfIAPbir into the pARL2-GFP Vector.
a) Agarose gel (1%) showing colony PCR of PfIAPbir after cloning into the pARL2-GFP vector.
◄ = expected band size.
b) Agarose gel (1%) of the pARL2-GFP_IAPbir plasmid following cloning, after digestion with
*Eco*RI. ◄ = expected band size.
c) Vector diagrams of the pARL2-GFP plasmid and the plasmid containing the PfIAPbir insert with
*Avr*II/*Xho*I and *Eco*RI restriction sites. Blue = inserted BIR domain.

2.3.7.2   Generation of PfIAPbir-GFP *P. falciparum* Transgenic Parasites

The pARL2-GFP_PfIAPbir construct (Figure 2.3.33) was used to transfect ring stage
P. falciparum 3D7 parasites and transgenic parasites were detected 32 days after
electroporation. The GFP-tagged protein expressed in parasites containing the pARL2-
GFP_PfIAPbir plasmid will be referred to as PfIAPbir-GFP. A control *P. falciparum*
pARL2-GFP transgenic line was donated by Dr Belinda Bezuidenhout and only expressed
the GFP protein.
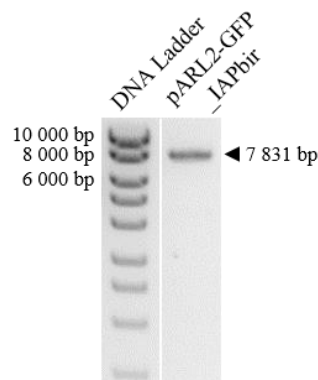
Figure 2.3.33: The pARL2-GFP_IAPbir Vector prior to Parasite Transfection
Agarose gel (1%) showing the *Bam*HI-linearized pARL2-GFP_PfIAPbir plasmid before parasite
transfection. ◄ = expected band size.

DNA was extracted from both the pARL2-GFP and pARL2-GFP_PfIAPbir transgenic parasites. PCR analysis, using the pARL2-GFP specific primers, was used to verify the presence of the episomal vectors in each of the transgenic lines (Figure 2.3.34). DNA was also extracted from wild type *P. falciparum* 3D7 parasites as a control, but as expected this did not produce a PCR amplicon (data not shown).
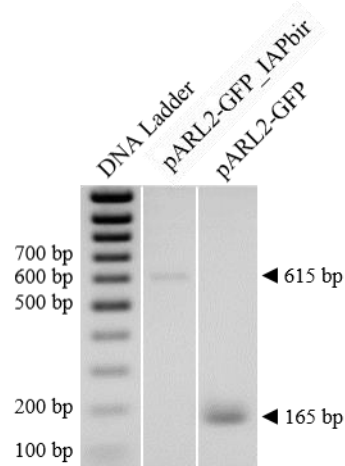
Figure 2.3.34: PCR Verification of the PfIAPbir-GFP Parasite Line
Agarose gel (1%) showing the presence of the pARL2-GFP_PfIAPbir plasmid by PCR using pARL2-GFP specific primers after parasite transfection. ◄ = expected band size.

2.3.7.3   Visualisation of Episomally Expressed PfIAPbir-GFP

Fluorescent images of live *P. falciparum* 3D7 pARL2-GFP transgenic parasites only expressing the GFP protein were captured. Across all three stages of the asexual life cycle the GFP protein was visible, as its expression was being driven by the *crt* promoter. As expected, the GFP protein displayed a cytoplasmic localisation (Figure 2.3.35; VanWye and Haldar 1997). Similarly, transgenic *P. falciparum* pARL2-PfIAPbir-GFP parasites expressed PfIAPbir-GFP in the cytoplasm of ring, trophozoite and schizont parasites, although the appearance of the signal was more punctate (Figure 2.3.36).

77
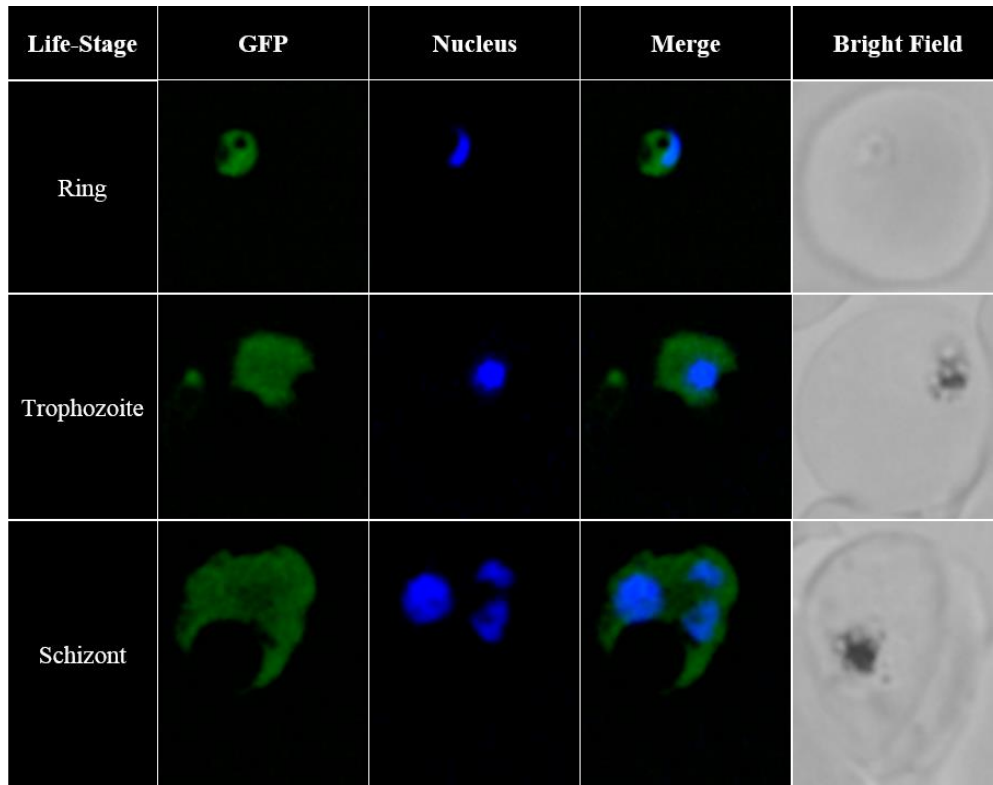
Figure 2.3.35: Fluorescence Microscopy Images of Live *P. falciparum* Transgenic Parasites Expressing the pARL2-GFP Protein

Localisation of the GFP protein in asexual *P. falciparum* transgenic parasites. Hoechst dye was used to stain ring stage nucleic acid, while DAPI was used to stain trophozoite and schizont nuclei.
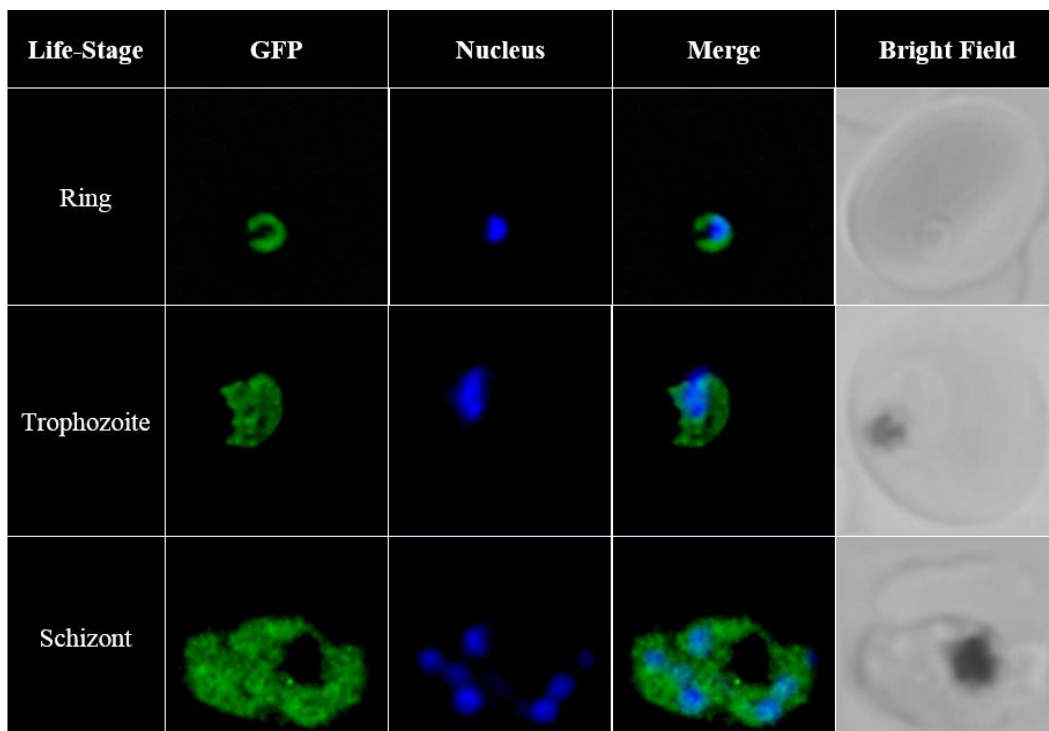
Figure 2.3.36: Fluorescence Microscopy Images of Live PfIAPbir-GFP *P. falciparum* Transgenic Parasites

Localisation of the fluorescent PfIAPbir-GFP protein in asexual *P. falciparum* transgenic parasites. Hoechst dye and DAPI were used to stain ring stage and trophozoite/schizont nuclei respectively.

Chapter 2: The Characterisation of PfIAP

2.3.7.4   The Effect of Heat-Shock on the Localisation of PfIAPbir-GFP

Febrile illness is the most recognisable symptom of malaria infection and has documented consequences for the malaria parasite, including differential gene expression and the induction of various RCD-like phenotypes (Oakley *et al.* 2007; Engelbrecht and Coetzer 2013). Synchronised ring and trophozoite transgenic parasite cultures were subjected to 41°C for 2 hours in order to simulate the effects of RCD on the parasites (Engelbrecht and Coetzer 2013). The localisation of both the GFP and PfIAPbir-GFP fluorescent proteins remained unchanged at various time points after heat-shock (Figure 2.3.37, Figure 2.3.38 and Figure 2.3.39). A large percentage of dead parasites (those lacking any intracellular movement) was observed 24 hours after heat-shock and those surviving parasites displayed retarded development as they remained in their initial life stage for at least 24 hours before progressing (as previously noted by Engelbrecht and Coetzer 2013).

Hours After Heat-Shock	GFP	Nucleus	Merge	Bright Field
0,5				
2,0				
24				
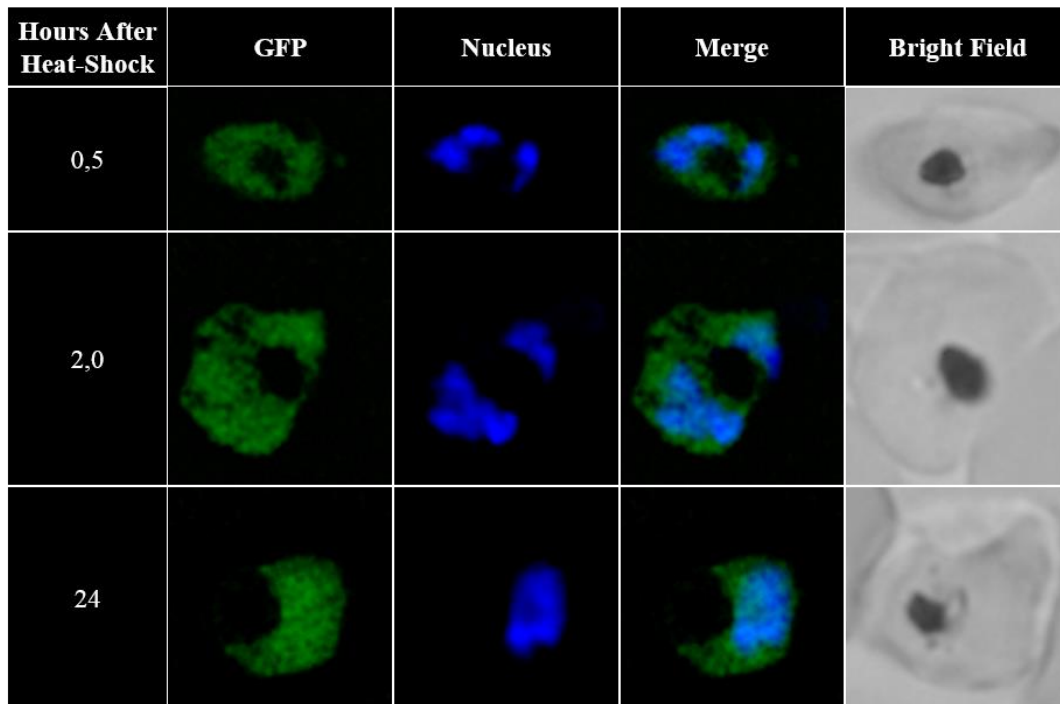
Figure 2.3.37: Fluorescence Microscopy Images of Heat-Shocked *P. falciparum* Transgenic Schizont Parasites Expressing the pARL2-GFP Protein
Heat-shock (41°C, 2 hours) had no effect on the localisation of the GFP protein in live, asexual *P. falciparum* transgenic parasites. DAPI was used to stain schizont nuclei.

Hours After Heat-Shock	GFP	Nucleus	Merge	Bright Field
2,0				
24				
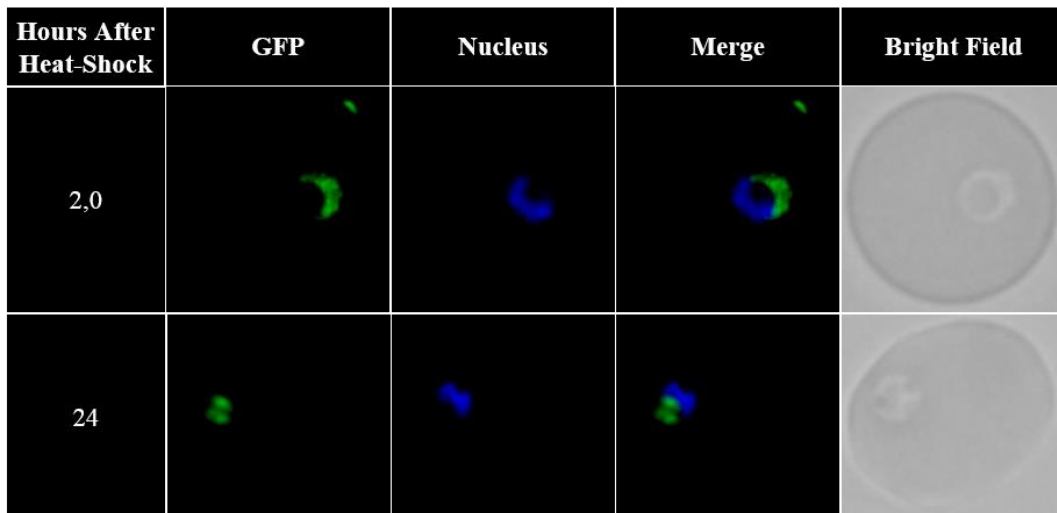
Figure 2.3.38: Fluorescence Microscopy Images of Heat-Shocked PfIAPbir-GFP *P. falciparum* Transgenic Ring Stage Parasites

The localisation of PfIAPbir-GFP in live, ring stage *P. falciparum* transgenic parasites after heat-shock (41°C, 2 hours) was unchanged. Hoechst dye was used to stain ring stage nucleic acid.

Hours After Heat-Shock	GFP	Nucleus	Merge	Bright Field
0,5				
24				
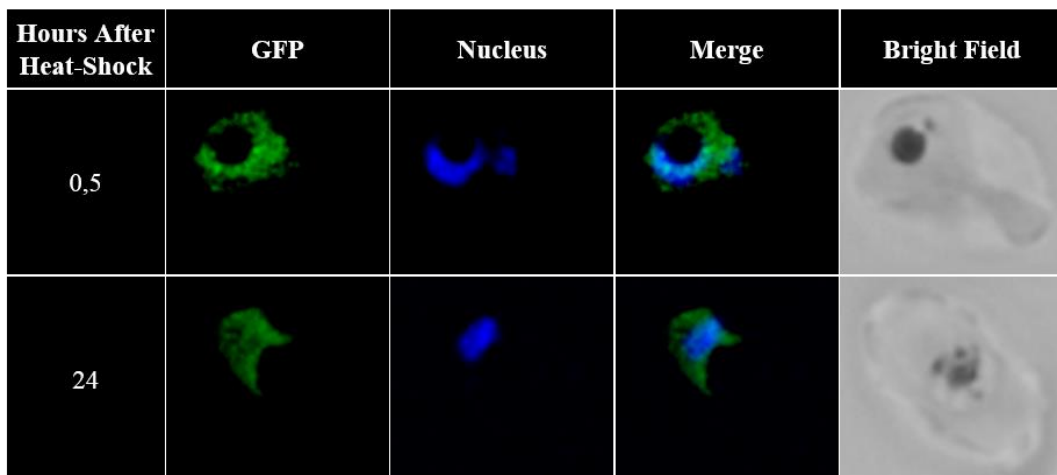
Figure 2.3.39: Fluorescence Microscopy Images of Heat-Shocked PfIAPbir-GFP *P. falciparum* Transgenic Trophozoite Parasites

Heat-shock (41°C, 2 hours) had no effect on the localisation of PfIAPbir-GFP in live, trophozoite stage *P. falciparum* transgenic parasites. DAPI was used to stain late stage nuclei.

2.4 Discussion

In light of the growing *P. falciparum* drug resistance (Dondorp *et al.* 2017), there is a need to identify novel drug targets. Characterising the ~50% hypothetical *P. falciparum* genes (Webster and McFadden 2014) is an important step in this process. PF3D7_0519600 is currently annotated in PlasmoDB as a putative zinc finger protein, but in this study is referred to as a *P. falciparum* IAP protein, following the identification of an N-terminal BIR domain. Biopanning experiments with recombinant PfIAP regions and a newly created biotin-tagged, *P. falciparum* cDNA phage display library have resulted in the identification of six novel PfIAP binding partners.

2.4.1   Recombinant Expression of PfIAP and Other *P. falciparum* Proteins

Five protein regions were expressed during the course of this study as recombinant proteins, using *E. coli* as the expression host. A bacterial protein expression system is a common and convenient tool but is not without challenges, especially when trying to express *P. falciparum* proteins. This is primarily a consequence of the fact that *E. coli* is a prokaryotic organism, while *P. falciparum* is eukaryotic. Differences between these two cell types include how proteins are synthesised: eukaryotic protein synthesis occurs in the cytoplasm, while protein synthesis in prokaryotes is initiated even before mRNA transcription is complete and is referred to as transcription–translation coupling (Gowrishankar and Harinarayanan 2004).

Another hazard is codon bias that occurs when the frequency of specific codons in the foreign DNA is different to the codons normally used by the host organism. The range of tRNAs available in a cell closely reflects the codon bias of its mRNA (Sahdev *et al.* 2007; Rosano and Ceccarelli 2014). A consequence of this is that during synthesis of a recombinant protein these low-abundant tRNA become depleted, resulting in either the misincorporation of an amino acid or the truncation of the peptide (Rosano and Ceccarelli 2014). Three of the proteins expressed in this study (GST-BIR, His-DOC and His-Rhop) were purified as both whole and truncated proteins, something that is not uncommon when expressing *P. falciparum* proteins using an *E. coli* system (Flick *et al.* 2004; Mehlin *et al.* 2006). Some of the identified rare codons (Sahdev *et al.* 2007), particularly those coding for arginine and isoleucine, are present in these proteins. As a way of limiting the effect of codon bias, Rosetta™ 2 (DE3) *E. coli* cells containing the pRARE plasmid were used for recombinant protein expression. This plasmid provides tRNAs for the amino acids alanine (CUA), arginine (AGA, AGG and CGG), cysteine (CCC), glycine (GGA) and isoleucine (AUA). Apart from codon bias, another contributing factor to protein truncation could be the fact that these *P. falciparum* proteins contain short stretches of the same amino acid (particularly lysine). This repetition could cause the bacterial translation machinery to stall for a moment, resulting in the production of a truncated recombinant protein. The predicted β-sandwich domain (PlasmoDB v34) consists of asparagine, glutamine and histidine amino acid repeats and was not expressed as a recombinant protein. This series of amino acids repeats could create a situation where the supply of the required tRNAs is simply exhausted in the bacterial host, leading to the incomplete translation of the recombinant domain.

Plasmodium falciparum proteins often aggregate and form insoluble inclusion bodies (Flick *et al.* 2004). All of the proteins expressed in the study had some level of protein detectable in the insoluble fraction, although GST-Middle was wholly insoluble. The factors affecting the solubility of expressed proteins include the size of the protein and its pI, although these factors cannot be compiled into a rule that can be universally applied to all proteins (Mehlin *et al.* 2006). For example, although proteins with a pI between 6 and 8 are favoured for protein expression (Mehlin *et al.* 2006), GST-Middle has a pI of 6.33 and was the only insoluble protein in this study, while the other proteins all had pI values outside of the given range. The predicted 48 kDa size of GST-Middle was between that of the other two soluble GST-tagged proteins, inferring that perhaps there is something else about this protein that resulted in it expressing as an insoluble protein. An easy to employ technique that often improves the solubility of expressed proteins is to grow the bacterial cells at a lower temperature. The hydrophobic interactions involved in inclusion body formation are temperature dependent and so a low temperature helps with the correct folding and increased stability of the protein (Sahdev *et al.* 2007). In this study, all protein expression cultures were incubated at room temperature, but it did not aid in the expression of soluble GST-Middle protein. Other approaches that can result in improved protein solubility include the use of different affinity tags (haemagglutinin (HA), maltose binding protein, thioredoxin etc.) or a different expression system, such as the yeast *S. cerevisiae* or insect cells (Birkholtz *et al.* 2008). Since the Middle region does not contain a recognisable domain, no further experiments were conducted with this protein.

For three of the four purified proteins in this study (GST-Terminal, His-DOC and His-Rhop), the expected molecular weights of the proteins were different to the molecular weights calculated from the SDS-polyacrylamide gels. This can be attributed to the regions of low complexity present in these recombinant proteins. Low complexity regions are characterised by one or a few amino acid repeat sequences and are present in more than 90% of *P. falciparum* proteins, including highly conserved housekeeping genes (Pizzi and Frontali 2001; Aravind *et al.* 2003). As a consequence of their unusual amino acid composition, intrinsically unstructured proteins bind less SDS than would be expected and so they migrate slower through the SDS-polyacrylamide gel, resulting in apparent molecular weights up to 1.8 times higher than calculated (Tompa 2003). The observed molecular weights of the three affected recombinant proteins in this study were 1.1-1.4 times higher than their calculated molecular weights. These unstructured soluble low complexity regions tend to be located

on the external surface of the folded protein and do not appear to interfere with the functioning of the rest of the protein (Aravind *et al.* 2003). Although no function has been assigned to low complexity regions, it is thought that they may be involved in antigen diversification (Ferreira *et al.* 2003; Cortes *et al.* 2005) or used as a means of eliciting an ineffective immune response (Anders 1986).

2.4.2   The Pros and Cons of Using a Phage Display Library

Most cellular proteins interact with other proteins, forming networks that facilitate the functioning of a cell (Berggård *et al.* 2007). From amongst the protein techniques available to detect these PPIs, this study utilised a newly constructed phage display library to identify novel binding partners for the BIR-containing PfIAP. Phage display libraries are easy to use, allow for the simultaneous expression of many different peptides and can be applied to a diverse range of targets (Lanzillotti and Coetzer 2008). *Plasmodium falciparum* phage display libraries were first successfully created in the Coetzer laboratory (Lauterbach *et al.* 2003) and used to identify *P. falciparum* interactions with the erythrocyte proteins spectrin and protein 4.1 (Lauterbach *et al.* 2003; Lanzillotti and Coetzer 2004).

The binding partners that are identified do not necessarily represent all of the bait protein's PPIs. Although the cDNA library created in this study was sourced from mixed stage, asexual *P. falciparum* cultures, the mRNA that was extracted may not represent all of the proteins that are expressed by the parasite during this life stage. During early and late trophozoite development the parasite synthesises many new proteins, leading to a drastic increase in the amount of mRNA present. The parasite cultures used in this study were biased towards late stage parasites as this helped facilitate the isolation of a sufficient amount of mRNA. Taken together, these factors mean that ring stage specific mRNA is likely to be present at a lower concentration then trophozoite mRNA and therefore could be under represented. Even at each discreet time point of the *P. falciparum* life cycle, some mRNAs are only transcribed at a low copy number and so they could easily be lost amongst the more abundant transcripts. *Plasmodium* parasites undergo many phenotypic changes during their life cycle, each of which is brought about by a different transcriptome. A phage display library constructed from only asexual *P. falciparum* parasites will inherently not be able to facilitate the identification of PPIs from other points of the parasite's development, such as the sexual gametocytes and mosquito stages. Therefore it is important to remember that the protein network constructed as part of this study is not the complete PfIAP interactome.

The oligo (dT)$_{25}$ beads that were used to isolate the parasite mRNA, rely on the 3' poly(A) tail of the mRNA. The complication that arises is that the *P. falciparum* genome is very AT-rich (~80%; Gardner *et al.* 2002) so any noncoding or ribosomal sequences that contain a series of adenosine repeats would also be isolated. The consequence of this *Plasmodium* quirk is the unfortunate presentation of short peptides that would not normally be present *in vivo* but that could still bind to the bait PfIAP protein (Lanzillotti and Coetzer 2008). In this study, bacteriophage expressing short sequences of unnatural lysine amino acids (codons AAA and AAG) were frequently detected.

A different consideration when using phage display libraries is the bait which is used during the biopanning experiments. A potential consequence of using a recombinant protein, expressed in bacterial cells, is that the protein may not have folded correctly (Sahdev *et al.* 2007), which could hamper the protein-binding ability of the bait. Since the natural conformation of the PfIAP protein is unknown, it was not possible to verify that the recombinant domains were correctly folded. A related issue is that peptides that are expressed on the surface of the bacteriophage may not be presented in their native conformation. This study investigated two regions of PfIAP, which were expressed as two separate recombinant proteins. While there is no reason that these individual regions would not interact with other proteins, particularly the BIR domain, it is also feasible that the whole protein may be required for certain protein interactions. Unfortunately it was not practical to express the whole PfIAP protein in this study, nor was it possible to use natively expressed PfIAP protein, isolated from *P. falciparum* parasites, due to the lack of PfIAP-specific antibodies required for protein isolation.

The actual biopanning procedure itself also requires some thought. The immobilisation of the recombinant PfIAP regions on magnetic beads, or the actual GST tag itself, could physically hinder binding of the phage peptides to the bait protein. The bacteriophage could also bind the GST tag, although this concern is diminished by the fact that the library underwent initial pre-screening using only the GST protein. The stringency of the wash steps following bait/phage library binding could inadvertently dislodge binding partners that form weak or transient connections. Between each successive round of biopanning, the bound bacteriophage are amplified. This provides opportunity for some phage (practically those with smaller inserts) to replicate at a faster rate than others, and so be in a position to outcompete them during the next round of biopanning (Ansuini *et al.* 2002). This event was

observed during this study. Of the bacteriophage plaques that were amplified by PCR, the number of amplicons that were >400 bp in size were vastly outnumbered by the amplicons of ~300 bp (each of which codes for a peptide about 20 amino acids in length). Since biopanning experiments are performed *in vitro*, peptides that would not normally bind to the bait protein *in vivo* now have the opportunity to interact, which can lead to false positives.

The validity of any newly identified binding partner needs to be confirmed using another biological technique. This study made use of an *in vitro* binding assay as its means of binding partner validation, although other techniques such as surface plasmon resonance may be used (Berggård *et al.* 2007). Heat denatured GST-BIR and free GST protein were employed as controls during the *in vitro* binding assays. The dose-dependent interactions that reached saturation between GST-BIR and the immobilised binding partner proteins were truly as a result of the increasing amounts of GST-BIR that were applied, since the recombinant GST and the heat-denatured BIR proteins showed only minimal and substantially reduced binding respectively.

2.4.3   The PfIAP BIR Domain Localises to the Cytoplasm

Expressing a protein of interest with a fluorescent tag from an episomal plasmid construct is a useful method for determining the subcellular localisation of the protein within a cell. In this study the N-terminal region of the PfIAP protein, containing the BIR domain, was expressed with a GFP tag in live, transgenic *P. falciparum* parasites. PfIAPbir-GFP was observed to have a cytoplasmic pattern that was dispersed throughout the parasite during all three stages of asexual development. The pARL2-GFP vector uses the *crt* promoter to drive the expression of the GFP-tagged protein, therefore the persistent expression of PfIAPbir-GFP is not unexpected. Given that *pfiap* mRNA is detected during many stages of the parasite's life cycle, prominently during the ring and trophozoite stages (Figure 2.4.1a), and that the PfIAP protein is detected throughout the asexual life cycle (Figure 2.4.1b), the use of an episomal plasmid with the *crt* promotor that allows for ubiquitous protein expression was appropriate.
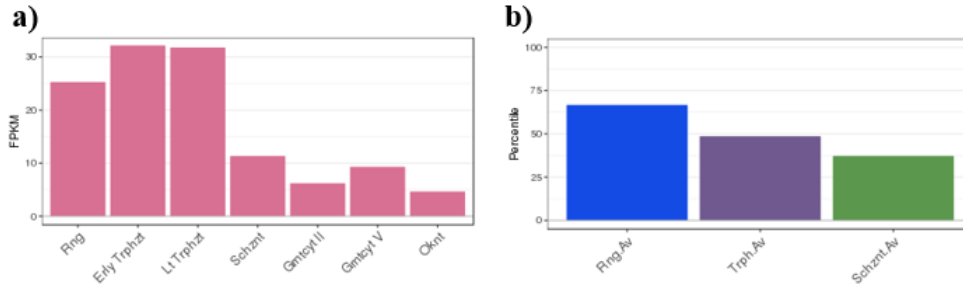
Figure 2.4.1: PfIAP mRNA and Protein Expression Levels (PlasmoDB)
Adapted from PlasmoDB (Aurrecoechea *et al.* 2009).
a) RNA was collected from synchronised *P. falciparum* parasites at seven different life stages (x-axis) and sequenced (López-Barragán *et al.* 2011). Transcript levels are indicated as fragments per kilo-base of exon model per million mapped reads (FPKM; y-axis).
b) Protein was harvested from synchronised *P. falciparum* parasites at three asexual life stages (x-axis) and quantified (Pease *et al.* 2013). Protein abundance percentile values (y-axis).

The various human anti-apoptotic, BIR-containing proteins have been shown to localise to a range of cellular components. cIAP1 (BIRC2) is a nuclear protein, while cIAP2 (BIRC3) is both cytoplasmic and nuclear (Samuel *et al.* 2005). The mitochondrial and cytoplasmic fractions of survivin (BIRC5) are thought to function as an apoptotic suppressor, while the nuclear pool of the protein is believed to regulate cell division (Garg *et al.* 2016). Full length livin (BIRC7) is exclusively found in the cytoplasm (Nachmias *et al.* 2007), as is the prominent anti-apoptotic XIAP (BIRC4; Liston *et al.* 2001). PlasmoDB (Aurrecoechea *et al.* 2009) indicates that the cellular compartment of PfIAP is the nucleus, therefore the observed cytoplasmic pattern of PfIAPbir-GFP may not be representative of the true subcellular localisation of the PfIAP protein. Numerous attempts were made to clone the whole *pfiap* gene into the pARL2-GFP vector, but none of the expected banding patterns were observed after restriction enzyme digestion.

It should be possible through single-crossover homologous recombination to insert a GFP coding sequence at the 3' end of the *iap* gene, allowing for the endogenous expression of GFP-tagged, full-length PfIAP protein. This will allow for the monitoring of the spatiotemporal expression of the whole PfIAP protein, either by direct fluorescence or through immunofluorescence assays (using an anti-GFP antibody). A GFP-tagged PfIAP would also allow for co-immunoprecipitation to be performed in live transgenic parasite cultures, followed by mass spectrometry, which would provide another method of identifying PfIAP binding partners. Co-localisation and co-immunoprecipitation studies

could also be performed on wild type parasites using anti-PfIAP antibodies, generated in mice or rabbits by immunisation with the recombinant PfIAP proteins.

PfIAPbir is a truncated form of the whole protein and could therefore be missing a trafficking motif that is necessary for the correct localisation of the protein. An example of this is provided by the cytoplasmic XIAP, which mainly localised in the nucleus when the C-terminal RING domain was removed (Cao *et al.* 2014). The nuclear localisation of cIAP1 on the other hand is dependent on its BIR domains (Samuel *et al.* 2005), while caspase-mediated cleavage of the cytoplasmic livin results in translocation of the truncated protein to the Golgi apparatus (Nachmias *et al.* 2007).

2.4.4 Heat Shock Does Not Affect the Localisation of PfIAPbir

The activation of RCD is a stressful event and results in biochemical changes within a cell. In human cells, the use of cytotoxic drugs to induce apoptosis results in the redistribution of XIAP from the cytoplasm into nuclear inclusions (Nowak *et al.* 2004). Also the nuclear XIAP-associated factor 1 binds to and inhibits XIAP by sequestering it in the nucleus, allowing apoptosis to proceed (Liston *et al.* 2001). RCD in *P. falciparum* caused the temporary translocation of PfSWIB from the cytoplasm to the nucleus, before the protein returned to the cytoplasm (Vieira and Coetzer 2016). In the current study RCD was stimulated by incubating parasites at 41°C for 2 hours, mimicking the high temperatures associated with febrile malaria illness (Engelbrecht and Coetzer 2013). The cytoplasmic localisation of PfIAPbir-GFP remained unchanged at various time points after heat-shock, irrespective of whether ring or trophozoite stage parasites had been exposed to heat-shock. This could be because the truncated protein lacks the signal sequence necessary for it to respond to heat stress, or it could be that the cytoplasmic localisation of the PfIAP BIR domain is already where the domain is required to be upon RCD activation.

2.4.5 *Plasmodium* IAP Proteins are Similar but PfIAP has Distinctive Structural Features

Each of the *Plasmodium* IAP proteins contain one BIR domain, which all follow the same general amino acid sequence, and a C-terminal end that is conserved across the *Plasmodium* genus. When compared to the human BIRC family of proteins, possessing one BIR domain is not uncommon, although some of the human IAP proteins contain three domains. None of the *Plasmodium* IAP proteins contain any of the recognisable C-terminal domains (RING

or UBC domains for example) that are associated with metazoan IAP proteins. This means that the *Plasmodium* IAPs cannot be classed as ubiquitin ligases, unlike some of the human IAP proteins (Feltham *et al.* 2012) and so makes it unlikely that PfIAP would be able to ubiquitinate other proteins without the help of an interacting partner.

PfIAP is the only *Plasmodium* species that contains what has been identified as a β-sandwich domain (Aurrecoechea *et al.* 2009), which is not present in any of the human BIRC proteins. These domains are common in proteins and have the characteristic structure of two opposing antiparallel β-sheets (Kister *et al.* 2006). The refined location of this β-sandwich domain (PlasmoDB v34) corresponds to a 62 residue, low complexity region of repeating amino acids: $NQNQ(NQHQ)_{56}NQ$ (H = histidine; N = asparagine; Q = glutamine). Structural predictions of this β-sandwich region failed to produce any meaningful results (data not shown), which is not unexpected as low complexity regions are regarded as unstructured (Aravind *et al.* 2003). This implies that the identification of this region by PlasmoDB as a β-sandwich domain using the Hidden Markov Model SUPERFAMILY library (Gough *et al.* 2001) may not be correct. This low complexity region is unique to *P. falciparum* and is not even present in the primate *P. reichenowi* IAP orthologue, which is by far the most similar (89.4%) of all the *Plasmodium* PfIAP orthologues. Why such a stretch of amino acids would be present in the *P. falciparum* protein is uncertain given the conserved nature of the rest of this gene throughout the *Plasmodium* genus. It could be required as a binding motif or structural component of the PfIAP protein for some function that none of the other *Plasmodium* IAP proteins participate in, perhaps something related to *P. falciparum* RCD. Or this stretch of amino acids could act as a decoy against the host's immune system (Anders 1986), reflecting the importance of the PfIAP protein when compared to the other *Plasmodium* species.

2.4.6   The *Plasmodium iap* Gene is Unique amongst Apicomplexa, Limiting what can be Suggested about the Protein

The *iap* gene was not identified in other members of the Apicomplexa phylum or the wider protozoan group of unicellular eukaryotic organisms. Protozoa like *Leishmania* or *Trypanosoma* experience cell death (Proto *et al.* 2012), but a BLAST search was unable to identify any BIR-containing proteins within these genomes. This implies that either these organisms undergo RCD in a way that does not need the participation of an IAP protein or the *Plasmodium* BIRC proteins have been adapted to the parasite's life cycle in its vertebrate

and non-vertebrate hosts. Alternatively the *Plasmodium* IAP proteins may have a role in the parasite that is not related to RCD (whether solely or in addition to its RCD function) and so is needed by all *Plasmodium* species during the course of normal cellular functioning. Such an uncharacteristic role is not unprecedented. There are numerous examples of human proteins (including IAP proteins) that are involved in regulating RCD, but also have non-apoptotic functions (Galluzzi *et al.* 2012). This introduces the question about the evolutionary origin of the BIR domain as it is present in metazoans where RCD is a hallmark of multi-cellularity, but not in the unicellular organisms that are more closely related to *Plasmodium*. Did *Plasmodium* acquire the *iap* gene from its human host or was the gene present in the last common ancestor of the Apicomplexa and the other protozoa simply lost the gene as there was no selective pressure to keep a protein that was not being used?

2.4.7   Why is *iap* Present in Numerous *Plasmodium* Species?

It is currently unclear why all five of the human *Plasmodium* species, one species infecting primates and three species infecting rodents possess the proposed anti-apoptotic IAP protein given the intraerythrocytic differences between the species, which means that each species experiences a different population density pressure and need for RCD. *Plasmodium falciparum* sexual gametocytes mature over the course of 10-12 days (Meibalan and Marti 2017), while asexual reproduction involves a merozoite invading an erythrocyte and developing for 48 hours, before releasing up to 32 merozoites to repeat the cycle (Josling and Llinás 2015). A consequence of this exponential asexual reproduction is likely the premature death of the human host before the sexual gametocytes have had enough time to develop and so the use of RCD by *P. falciparum* to control its population number makes sense.

For the other *Plasmodium* species infecting humans the population pressures during the intraerythrocytic stages are not as severe and sexual development is much quicker than for *P. falciparum*, so the need for RCD is less clear. *Plasmodium malariae* parasites produce fewer merozoites per cycle and require 72 hours for asexual development so the parasitaemia experienced by patients is usually lower than that of *P. falciparum* infections (Collins and Jeffery 2007; Kerlin and Gatton 2013). *Plasmodium vivax* and *P. ovale* both only invade immature red blood cells, reticulocytes, which is the likely reason why these infections are characterised by low parasitaemia and low pathogenicity, as compared to *P. falciparum*. The asexual cycles of *P. vivax* and *P. ovale* both take about two days to complete with fewer than

20 merozoites being released per schizont and only 48-49 hours are required for gametocyte development (McKenzie *et al.* 2002; Collins and Jeffery 2005). Although the red blood cell preference of *P. knowlesi* is currently unknown (Kerlin and Gatton 2013), the release of up to 16 merozoites every 24 hours (the shortest asexual cycle of any human and non-human primate species) means that there is the potential for a very rapid increase in parasitaemia to the detriment of the infected human, but this is offset by the short time period of 48-49 hours needed for gametocytes to develop (Collins 2012). The rodent *P. berghei* and *P. yoelii* parasites, with asexual cycles of less than one day, produce gametocytes in the very short time of 24-27 hours (Meibalan and Marti 2017).

Available transcriptomic data (Figure 2.4.2) show that the *P. berghei* and *P. vivax* IAP orthologues have very similar expression profiles to PfIAP (with peak expression during the trophozoite stage). Despite this, it is evident that the differences in the intraerythrocytic parasite stages between the various *Plasmodium* species means that each *Plasmodium* species experiences a different population pressure and has different RCD requirements, yet all of the *Plasmodium* species possess the *iap* gene. The invertebrate *Plasmodium* life stage involves the development of ookinetes and their transformation into oocysts in the mosquito vector. It is during this ookinete stage that apoptosis-like cell death has been detected in both *P. berghei* and *P. falciparum* (Al-Olayan *et al.* 2002; Arambage *et al.* 2009; Ali *et al.* 2010). Available data show that both *P. berghei* and *P. falciparum iap* mRNA is detected during this time. The existence of the conserved IAP protein across the *Plasmodium* genus could be because, irrespective of their individual needs for intraerythrocytic RCD, IAP is required by all the *Plasmodium* species for RCD during the mosquito stage, or the protein might play another unrelated role in the parasite.
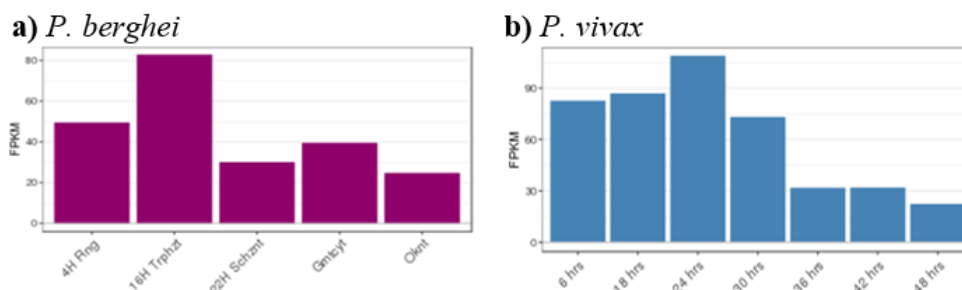
Figure 2.4.2: mRNA Expression Levels of *Plasmodium* IAP Orthologues (PlasmoDB)
Adapted from PlasmoDB (Aurrecoechea *et al.* 2009). RNA was collected from *Plasmodium* parasites at different time points (x-axis) and sequenced. Transcript levels are indicated as fragments per kilo-base of exon model per million mapped reads (FPKM; y-axis).
a) *P. berghei* IAP orthologue (PBANKA_1234400; Otto *et al.* 2014).
b) *P. vivax* IAP orthologue (PVX_080265; Zhu *et al.* 2016)**.**

Chapter 2: The Characterisation of PfIAP

2.4.8 Could PfIAP Perform the Same Functions as the Human IAP Proteins?

Each of the *Plasmodium* IAP proteins is currently annotated in PlasmoDB as putative zinc finger (ZF) proteins. In eukaryotic genomes, ZFs are amongst the largest transcription factor families but in addition to their DNA binding roles, zinc fingers have been shown to participate in protein assembly, lipid binding, RNA packaging and the regulation of apoptosis (Laity *et al.* 2001). Human BIR motifs use the CCHC (C = cysteine; H = histidine) motif to coordinate zinc and structurally resemble the classical C_2H_2 zinc finger (Laity *et al.* 2001). Although the conserved $CX_2CX_{23}HX_3C$ (X = any amino acid) pattern of the *Plasmodium* BIR domains is slightly longer than the sequence attributed to human BIR domains ($CX_2CX_{16}HX_6C$), the predicted *Plasmodium* BIR domain structure is in keeping with the known structures of the various human BIR domains. Visualising the conserved zinc-chelating residues in the PfIAP 3D models, makes it possible to locate the important hydrophobic pocket in this *Plasmodium* domain (see Figure 2.3.6).

Given the structural similarities between the human and *P. falciparum* BIR domains, it is possible that this domain is involved in biological functions analogous to those performed by the human BIR domains – perhaps inhibiting the caspase-like metacaspase proteins present in *P. falciparum* (Meslin *et al.* 2007; no caspases are present in *Plasmodium*). However, the following examples show that the various human IAP proteins also have non-apoptotic roles and are involved in other cellular processes. The BIR3 domain of XIAP can promote cellular copper retention by binding COMMD1 (copper metabolism gene MURR1 domain containing 1), a protein involved in copper efflux to ensure proper homeostasis, triggering its ubiquitination and degradation. XIAP is also able to bind free copper itself, resulting in a conformational change in the protein that alters its ability to successfully inhibit caspase-3, an effector caspase (Burstein *et al.* 2004; Mufti *et al.* 2007). In addition to anti-apoptotic functions, survivin is known to play an essential role in cell division by forming a chromosomal passenger complex with other proteins, mediated through its BIR domain (Kelly *et al.* 2010; Wang *et al.* 2010; Yamagishi *et al.* 2010). Although NAIP has been shown to attenuate apoptosis, its main role is emerging to be in the regulation of innate immune signalling (Finlay *et al.* 2017). Likewise, the cIAP1 and cIAP2 proteins are important constituents of the tumor necrosis factor receptor pathway and play a role in NF-κB signalling (Kocab and Duckett 2016).

Given the various functions each of the human IAP homologues can perform, and the fact that some of these roles are not directly transferrable from the human system to that of the *Plasmodium* parasite, the function of the PfIAP protein cannot be inferred from its structure alone.

2.4.9 Building a PfIAP Protein-Protein Network to Potentially Ascertain its Function

Using the results of the biopanning experiments in this study it is possible to create a protein-protein network centred on PfIAP. This is likely to only be a partial protein network, since the phage display library used is not representative of the full *P. falciparum* transcriptome (see 2.4.2). Previous high throughput yeast two-hybrid experiments failed to identify any binding partners for PfIAP (LaCount *et al.* 2005).

2.4.9.1 Binding Partners of the PfIAP BIR Domain

In the current study the PfIAP BIR region bound four *P. falciparum* proteins (Figure 2.4.3).
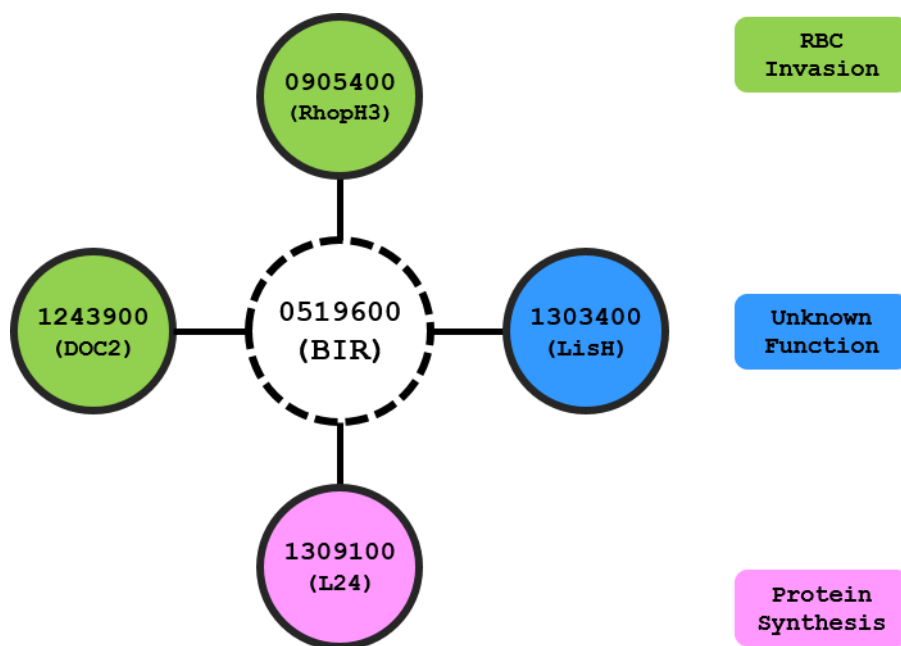
Figure 2.4.3: Primary PfIAP BIR Domain Protein Network
Diagram showing the PfIAP BIR domain interactions as determined by biopanning in this study. Numbers = PlasmoDB gene identifiers. Colours = Gene Ontology biological process available on PlasmoDB (Ashburner *et al.* 2000; Aurrecoechea *et al.* 2009). Detailed protein information can be found in APPENDIX D.

The LisH Protein. The phage peptide that bound GST-BIR is located in the middle of PF3D7_1303400 which has one N-terminal LisH motif and no other identified domains. The eukaryotic LisH motif plays a role in regulating transcription (Wei *et al.* 2003).

Bioinformatic analysis has identified 114 intracellular eukaryotic proteins that contain this domain, with many of these proteins also possessing other, unrelated C-terminal domains (Emes 2001). This group of proteins has many functions, including being involved in protein-protein interactions (Gerlitz *et al.* 2005; Mikolajka *et al.* 2006). Bioinformatics strongly predicts that PfLisH localises to the *P. falciparum* nucleus (Vieira and Coetzer 2016), which correlates with a role in transcription. Through phage display and binding studies PfLisH has been verified as a binding partner of PfMDM2, although the biological function of this interaction is uncertain as PfMDM2 localises to the *P. falciparum* mitochondria (Vieira and Coetzer 2016). The mRNA and protein expression profiles of PfIAP and PfLisH are similar (Figure 2.4.1 and Figure 2.4.4), with mRNA being detected for both proteins during asexual, gametocyte and ookinete life stages. Both PfIAP and PfLisH are more prevalent during the ring and trophozoite stages, but are still present during the schizont life stage. If PfLisH has a similar regulatory role as it eukaryotic counterparts, its interaction with PfIAP is feasible, given that both proteins are predicted to localise to the nucleus. This could provide a mechanism by which PfIAP can influence the expression of other proteins, either by up- or down-regulating proteins involved in RCD or perhaps proteins involved in other cellular processes that are unknown at this time.

Figure 2.4.4: PfLisH mRNA and Protein Expression Levels (PlasmoDB)
Adapted from PlasmoDB (Aurrecoechea *et al.* 2009).
a) RNA was collected from synchronised *P. falciparum* parasites at seven different life stages (x-axis) and sequenced (López-Barragán *et al.* 2011). Transcript levels are indicated as fragments per kilo-base of exon model per million mapped reads (FPKM; y-axis).
b) Protein was harvested from synchronised *P. falciparum* parasites at three asexual life stages (x-axis) and quantified (Pease *et al.* 2013). Protein abundance percentile values (y-axis).

The Protein Synthesis Component. PfL24 (PF3D7_1309100) is only mentioned in the literature with respect to the *P. falciparum* 80S ribosome, where it is a constituent of the Pf60S subunit (Wong *et al.* 2014). This protein is located in the cytoplasm and is present throughout the *P. falciparum* asexual life cycle (Figure 2.4.5), particularly during the ring and trophozoite stages. This makes sense as it is during these life stages that the parasite is

growing rapidly and requires the synthesis of many proteins. While there is currently no indication that PfIAP is directly involved in protein synthesis, PfIAP could be present with PfL24 as part of a larger protein complex, one that perhaps controls which/when proteins are synthesised. The interaction of the cytoplasmic PfL24 with the nuclear PfIAP might be inconsistent, but the PlasmoDB prediction is based on the fact that this protein is currently listed as a zinc finger protein. While these types of proteins, which includes the IAP family, can be nuclear proteins, many of them do have additional roles outside of the nucleus (Laity *et al.* 2001). As noted previously (see 2.4.3), the human BIRC proteins are distributed to various locations throughout the cell and as such PfIAP may localise to more than one area of the parasite or may translocate to the cytoplasm under stress conditions.
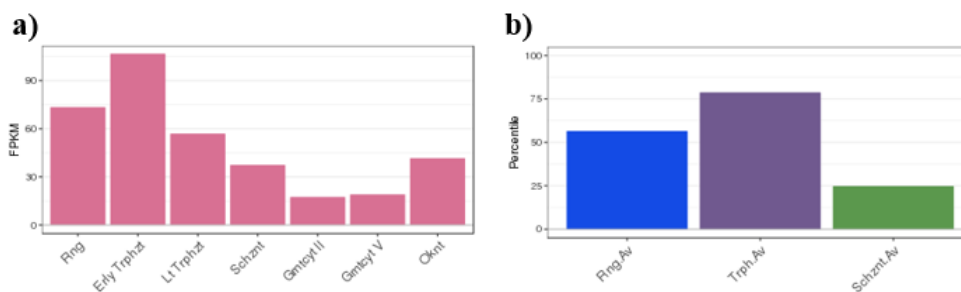

Figure 2.4.5: PfL24 mRNA and Protein Expression Levels (PlasmoDB)
Adapted from PlasmoDB (Aurrecoechea *et al.* 2009).
a) RNA was collected from synchronised *P. falciparum* parasites at seven different life stages (x-axis) and sequenced (López-Barragán *et al.* 2011). Transcript levels are indicated as fragments per kilo-base of exon model per million mapped reads (FPKM; y-axis).
b) Protein was harvested from synchronised *P. falciparum* parasites at three asexual life stages (x-axis) and quantified (Pease *et al.* 2013). Protein abundance percentile values (y-axis).

The Invasion Conspirators - PfDOC2 and PfRhopH3. The Ca^{2+}-triggered release of cellular components (such as neurotransmitters) is regulated by a range of different proteins, including those with double calcium lipid-binding domains, DOC2 proteins (Friedrich *et al.* 2010). Egress of merozoites from erythrocytes and the subsequent release of invasion proteins are important *P. falciparum* exocytotic events. PfDOC2 (PF3D7_1243900) has two calcium binding sites and is conserved across the Apicomplexa phylum (Farrell *et al.* 2012). Only when this protein is associated with membranous structures will calcium bind to it, which is consistent with other double C2 domains (Jean *et al.* 2014). PfDOC2 displayed a punctate staining pattern throughout the cytoplasm in schizont stage parasites (Jean *et al.* 2014). Conditional knockdown of the PfDOC2 protein had no effect on the asexual development or egress of the parasite, but did negatively affect RBC invasion as a consequence of incomplete microneme discharge (Farrell *et al.* 2012). This adds to a body

of evidence showing that, although both merozoite egress and the release of invasion proteins from the micronemes are calcium-dependent processes, they utilise different mechanisms (Soni *et al.* 2017). The PfDOC2 protein has been detected at the same level in all three asexual life stages (Jean *et al.* 2014). While PfDOC2 is involved in red blood cell invasion, it could have other cellular functions and as such could have the opportunity to interact with PfIAP.

PfRhopH3 (PF3D7_0905400) plays a role in both erythrocyte invasion and parasite growth and is part of the high molecular weight, multi-protein RhopH complex (Cooper *et al.* 1988). It is released from the rhoptry during erythrocyte invasion and contributes to merozoite invasion either through interaction with red blood cell receptors (such as the anion exchanger band 3; Baldwin *et al.* 2014) and/or by helping with parasite internalization (Ito *et al.* 2017). Following invasion, the multi-protein PfRhopH complex is introduced into the parasitophorous vacuole and then possibly uses the *Plasmodium* translocon of exported proteins (PTEX) complex to move to the erythrocyte cytosol. From here the complex is inserted into the RBC membrane where it is involved in the formation of the plasmodial surface anion channel (PSAC; Ito *et al.* 2017), which is an ion and nutrient channel in the host cell membrane and is part of the new permeability pathways that are created following merozoite invasion. The single copy *pfrhoph3* gene is essential for *P. falciparum* survival (Ito *et al.* 2017) and conditional truncation of PfRhopH3 results in parasites with a reduced invasive capacity and altered growth as a consequence of defective nutrient import (Sherling *et al.* 2017). *pfrhoph3* mRNA is prominent during the late trophozoite and schizont life stages (Figure 2.4.6a) which correlates with its role in invasion, but the PfRhopH3 protein is present throughout the asexual stages (Figure 2.4.6b), in line with the protein being present in the PSAC. It is unlikely that PfIAP would interact with PfRhopH3 while the protein is being stored in the rhoptry, but PfIAP could interact with PfRhopH3 after invasion, before it is exported out of the parasite to the host cell membrane. The other members of the RhopH complex are the multi-gene family PfRhopH1/Clag (PfCLAG2, 3.1, 3.2, 8 and 9) and the single gene *pfrhoph2*. When PfCLAG3 is knocked out, growth defects consistent with poor nutrient uptake are observed (Comeaux *et al.* 2011), which is expected since PfCLAG3 has been implicated with the PSAC (Nguitragool *et al.* 2011). PfRhopH2 is also involved in PSAC formation, is essential for *P. falciparum* survival (Ito *et al.* 2017) and alters parasite growth if its expression is knocked down using a *glmS* riboswitch (Counihan *et al.* 2017).
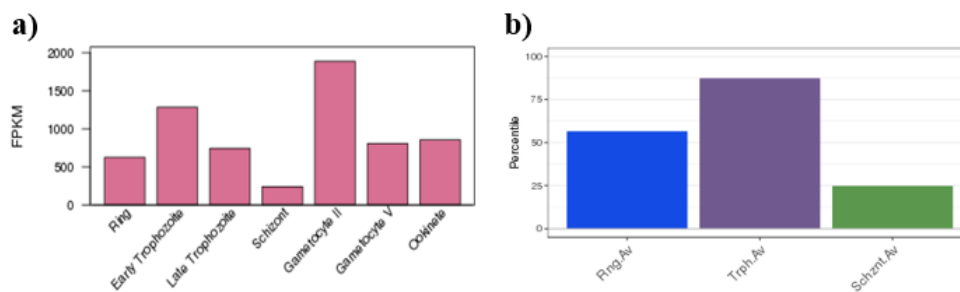
Figure 2.4.6: PfRhopH3 mRNA and Protein Expression Levels (PlasmoDB)

Adapted from PlasmoDB (Aurrecoechea *et al.* 2009).

a) RNA was collected from synchronised *P. falciparum* parasites at seven different life stages (x-axis) and sequenced (López-Barragán *et al.* 2011). Transcript levels are indicated as fragments per kilo-base of exon model per million mapped reads (FPKM; y-axis).

b) Protein was harvested from synchronised *P. falciparum* parasites at three asexual life stages (x-axis) and quantified (Pease *et al.* 2013). Protein abundance percentile values (y-axis).

2.4.9.2   The C-Terminal Binding Partners of PfIAP

The recombinant Terminal region of PfIAP bound two *P. falciparum* proteins (Figure 2.4.7).
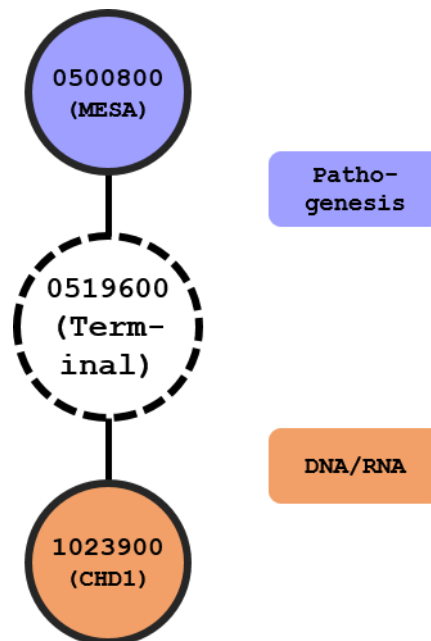
Figure 2.4.7: Primary PfIAP C-Terminal Protein Network

Diagram showing the C-terminal PfIAP interactions as determined by biopanning in this study. Numbers = PlasmoDB gene identifiers. Colours = Gene Ontology biological process available on PlasmoDB (Ashburner *et al.* 2000; Aurrecoechea *et al.* 2009). Detailed protein information can be found in APPENDIX D.

The Transcription Participant. One of the proteins that bound to the Terminal region of PfIAP is the *P. falciparum* chromodomain-helicase-DNA-binding protein 1 (PF3D7_1023900). A common role of the human CHD proteins is in ATP-dependent

chromatin remodelling (Hall and Georgel 2007). Human CHD1 in particular is involved in transcriptional repression and RNA splicing (Tai *et al.* 2003), keeping the double-stranded DNA helix open and transcriptionally active, and is critical to the pluripotency of embryonic stem cells (Gaspar-Maia *et al.* 2009). Experimental evidence showed that *P. falciparum* CHD1 is found in the parasite nucleus (Volz *et al.* 2010) and its expression profiles show that the most protein is present during the trophozoite life stage (Figure 2.4.8), which is understandable as this is when the *P. falciparum* parasite is undergoing rapid growth that requires transcription and synthesis of new proteins. As with the PfLisH protein, the interaction of PfIAP with a protein that is likely to be involved in transcription means that PfIAP may be able to use this interaction to influence the expression of other proteins, possibly ones involved in parasite RCD.
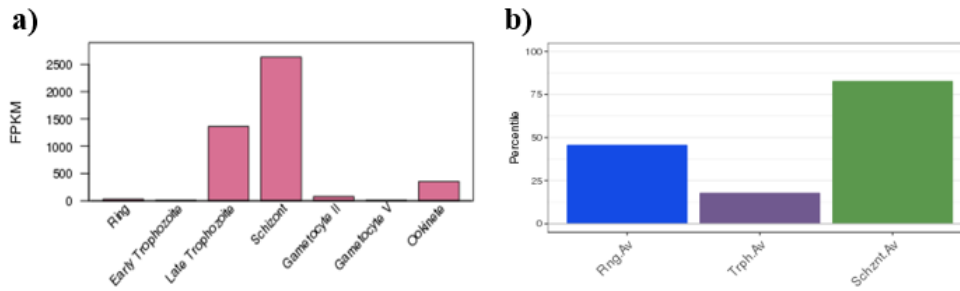

Figure 2.4.8: PfCHD1 mRNA and Protein Expression Levels (PlasmoDB)
Adapted from PlasmoDB (Aurrecoechea *et al.* 2009).
a) RNA was collected from synchronised *P. falciparum* parasites at seven different life stages (x-axis) and sequenced (López-Barragán *et al.* 2011). Transcript levels are indicated as fragments per kilo-base of exon model per million mapped reads (FPKM; y-axis).
b) Protein was harvested from synchronised *P. falciparum* parasites at three asexual life stages (x-axis) and quantified (Pease *et al.* 2013). Protein abundance percentile values (y-axis).

The Red Cell Membrane Connection. The *P. falciparum* mature parasite-infected erythrocyte surface antigen (PfMESA; PF3D7_0500800) is the first *P. falciparum* protein to have its cytoskeleton-binding partner identified (Lustigman *et al.* 1990). MESA interacts with the N-terminal 30 kDa domain of the phosphorylated form of protein 4.1, competing with the cellular protein p55 (Waller *et al.* 2003). This interaction occurs adjacent to the PEXEL/HT motif, which is how PfMESA is exported to the RBC cytosol (Black *et al.* 2008). This protein is expressed during the intraerythrocytic trophozoite stage (Figure 2.4.2) and colocalises with the knobs that develop on the erythrocyte surface, but is not required for the formation of these knobs nor is it involved in their associated cytoadherence (Petersen *et al.* 1989). There is no indication that PfIAP travels to the host cell membrane, therefore PfIAP

would have to interact with PfMESA prior to its export from the parasite, although it is currently unclear what function PfIAP would perform binding to PfMESA.
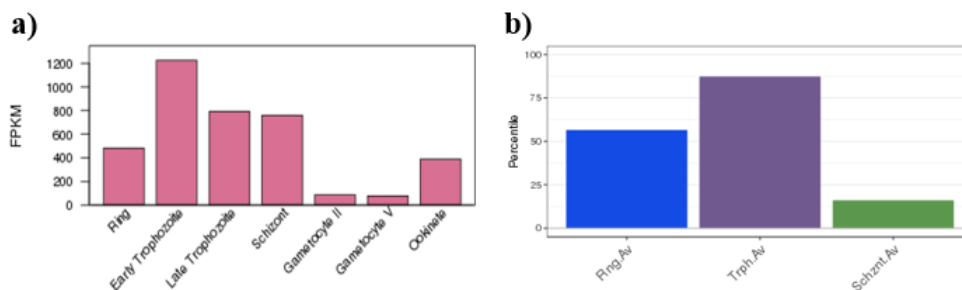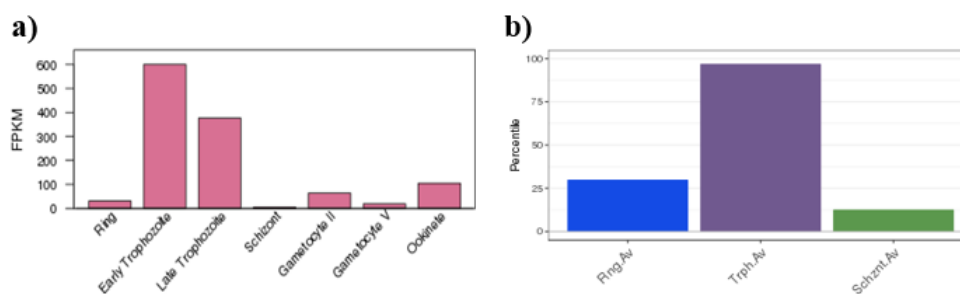

Figure 2.4.9: PfMESA mRNA and Protein Expression Levels (PlasmoDB)
Adapted from PlasmoDB (Aurrecoechea *et al.* 2009).
a) RNA was collected from synchronised *P. falciparum* parasites at seven different life stages (x-axis) and sequenced (López-Barragán *et al.* 2011). Transcript levels are indicated as fragments per kilo-base of exon model per million mapped reads (FPKM; y-axis).
b) Protein was harvested from synchronised *P. falciparum* parasites at three asexual life stages (x-axis) and quantified (Pease *et al.* 2013). Protein abundance percentile values (y-axis).

2.4.9.3   Expanding the PfIAP Protein Network

The PfCHD1, PfL24, PfMESA and PfRhopH3 binding partners all interact with other proteins as identified by yeast two-hybrid studies (LaCount *et al.* 2005). By incorporating these new proteins it is possible to expand the PfIAP protein-protein network (Figure 2.4.10) to include proteins involved in various biological processes, ranging from metabolism to pathogenesis. A number of these proteins are interacting partners of more than one of the PfIAP binding proteins, pointing to PfIAP being involved in a network of highly interconnected *P. falciparum* proteins. The wide range of Gene Ontology annotations (Ashburner *et al.* 2000) that can be applied to proteins identified in both biopanning and yeast two-hybrid experiments is an indication of alternative functions of both PfIAP and its binding partners (LaCount *et al.* 2005).

The interaction of PfIAP with any of its primary partners could indicate that this binding has a biological function. Alternatively these binding partners could form complexes with PfIAP and their own interacting proteins, providing PfIAP with the opportunity to interact with other proteins. For example, one of the PfL24 binding partners is the putative autophagy-related protein 23 (PfATG23) which is required for selective *P. falciparum* autophagy (Navale *et al.* 2014) and is an RCD protein that PfIAP would conceivably interact with. The cell death participants in *P. falciparum* are currently unknown, therefore any of these secondary proteins with undefined functions could have RCD-related roles and be of significance to the functioning of PfIAP. At this stage it is not possible to say whether these

primary and secondary interactions are stable or transient protein-protein connections. It is also not clear whether these interactions represent the action of PfIAP on these binding partners or if it is these associated proteins that influencing PfIAP. Further biochemical and molecular biology experiments involving PfIAP and its binding partners are required before these interactions can be truly elucidated.

Figure 2.4.10: The Expanded PfIAP Protein Network

Diagram showing the PfIAP protein (dashed outline), its partner proteins (dark outline) and their associated binding partners as determined by yeast two-hybrid experiments (LaCount *et al.* 2005). * = PfMDM2 binds PfLisH (Vieira and Coetzer 2016). Numbers = PlasmoDB gene identifiers. Dashed lines = Proteins that appear multiple times. Colours = Gene Ontology biological process available on PlasmoDB (Ashburner *et al.* 2000; Aurrecoechea *et al.* 2009). Detailed protein information can be found in APPENDIX D.

Given the previously described need for RCD in *P. falciparum*, the BIR-containing PfIAP protein may be expected to have an anti-apoptotic function in the parasite – that is, function in such a way as to prevent an excessive number of *P. falciparum* parasites from dying after the activation of RCD in response to population pressure during gametocyte development. This could involve binding to and inhibiting the actions of the available metacaspases or other cell death proteins. Metacaspase proteins would thus have been expected amongst the identified binding partners, although technical reasons for this lack of association might apply (see 2.4.2). Other reasons could be that these two protein families do not interact and that PfIAP is involved in the regulation of *P. falciparum* RCD via a different pathway/proteins. In addition to this role, or perhaps as its sole function, the PfIAP protein could be involved with everyday processes in the parasite, since its binding partners are involved in DNA transcription, RBC invasion and other functions.

Identifying the binding partners of a protein of interest and elucidating its protein-protein network is only a piece of the PfIAP puzzle. Another question of primary concern when characterising an unknown protein is whether it is important in the cell and if the gene is essential for survival of the parasite.

CHAPTER 3: Manipulation of the *P. falciparum* Genome

3.1   Introduction

The whole *P. falciparum* genome is distributed between 14 nuclear chromosomes, a 35 kb apicoplast genome and a 6 kb mitochondrial genome (Gardner *et al.* 2002). At the time of publication of the *P. falciparum* genome ~60% of predicted genes were hypothetical, more than a decade later that number has only decreased to ~50% (Webster and McFadden 2014). When investigating the function of an unknown gene/protein many different approaches can be employed. Conditional disruption of a gene locus can for example be achieved using either the Cre-*loxP* or FLP–*frt* systems that utilise a recombinase enzyme to trigger a recombination event between two sequence specific regions (DeKoning-Ward *et al.* 2015). If the protein has many different roles in the cell though, then simply removing the gene will mean that only the earliest non-redundant downstream function of the protein will be observed (Webster and McFadden 2014). As a way to circumvent this it is possible to affect the transcriptional activation of a gene, interfere with the produced mRNA or influence the protein itself. RNA interference (RNAi) is a popular method of gene silencing, but unlike in mammalian systems, *Plasmodium* species seem to lack the necessary machinery for RNAi to be an effective means of regulating mRNA levels (Baum *et al.* 2009). A technique that has been shown to work in *P. falciparum* involves inserting the *glmS* ribozyme sequence at the 3' of the gene of interest (GOI). The addition of exogenous glucosamine triggers ribozyme self-cleavage and the conditional knockdown of the associated mRNA. The stability of the synthesised protein can also be disrupted using tools such as the destabilizing domain (DD) system. The FK506-binding protein is fused to the C-terminal end of the protein of interest and will cause this protein to be degraded if the synthetic Shield-1 ligand is not present to stabilise the DD (Webster and McFadden 2014; DeKoning-Ward *et al.* 2015).

Deleting a GOI and observing the phenotypic effects is an important step in characterising an unknown gene. Performing such a genetic manipulation can be achieved quicker and more precisely using one of the gene editing tools that have been developed recently.

3.1.1 Engineered Genome Editing Techniques

Genome editing methods modify GOIs by using site-specific DNA nucleases to cleave the genome, triggering the activation of cellular repair machinery which can then be directed to 'fix' the genome in a number of different ways, such as introducing a mutation into the sequence or deleting a whole gene (Lee and Fidock 2014).

A potentially hazardous event for any cell is the occurrence of a double-strand break (DSB) in a chromosome. This can be a consequence of normal cellular processes, such as the action of reactive oxygen species for example, or it can be experimentally induced using ionizing radiation, chemical mutagens and DNA nucleases (Lee *et al.* 2014). A single DSB can lead to genetic instability, for example chromosome translocation or even cell death, and as such organisms have developed DNA repair machinery. DSBs though can also have benefits, like immune response diversification, and thus the fidelity of DSB repair need not be perfect (Lee *et al.* 2014). One of these mechanisms is non-homologous end joining (NHEJ) and is the potentially "error-prone" pathway. It simply ligates the broken DNA ends together, without the use of a homologous template, which means that the process can introduce indels (insertions or deletions) into the DNA sequence. An alternative is the "error-free" homologous recombination (HR) pathway which utilizes a homologous template, whether present on the homologous chromosome or a donor plasmid, to repair the broken DNA strand (Lee and Fidock 2014; Lee *et al.* 2014). No *Plasmodium* canonical NHEJ homologues have been identified (Straimer *et al.* 2012), although *T. gondii* does make use of a functional NHEJ pathway (Fox *et al.* 2009). Given the close evolutionary relationship between the two Apicomplexa this suggests that *Plasmodium* has lost its NHEJ machinery (Lee *et al.* 2014), similar to other eukaryotic pathogens such as *Encephalitozoon cuniculi* (Gill and Fast 2007) and *Giardia lamblia* (Manning *et al.* 2011).

3.1.1.1 Zinc Finger Nucleases and TALENs

A zinc finger nuclease (ZFN) is a metalloenzyme that is able to bind to and cleave DNA (Figure 3.1.1a). Each ZF is about 30 amino acids in size and generally binds three base pairs. Therefore 3-6 ZF units are generally combined together to produce a DNA-binding array with 9-18 base pairs of specificity. This is joined to a DNA-cleaving module in the form of the restriction endonuclease *Fok*I. Dimerization of two *Fok*I domains is required for cleavage of double-stranded DNA and so two ZFNs must be simultaneously employed, one for each DNA strand. This increases the specificity of the ZFN binding as not only are

two DNA-binding events required, but the ZFNs also need to be in the correct orientation (Carlson *et al.* 2012; Gaj *et al.* 2013). ZFNs were the first customised nucleases to be widely employed for genome editing in various organisms, including *P. falciparum* (Straimer *et al.* 2012; McNamara *et al.* 2013), *P. vivax* (Moraes Barros *et al.* 2015) and *P. berghei* (Singer *et al.* 2015). They have also been used in the mosquito *Aedes aegypti* to disrupt pathways and behaviours as part of dengue control measures (Reegan *et al.* 2016).
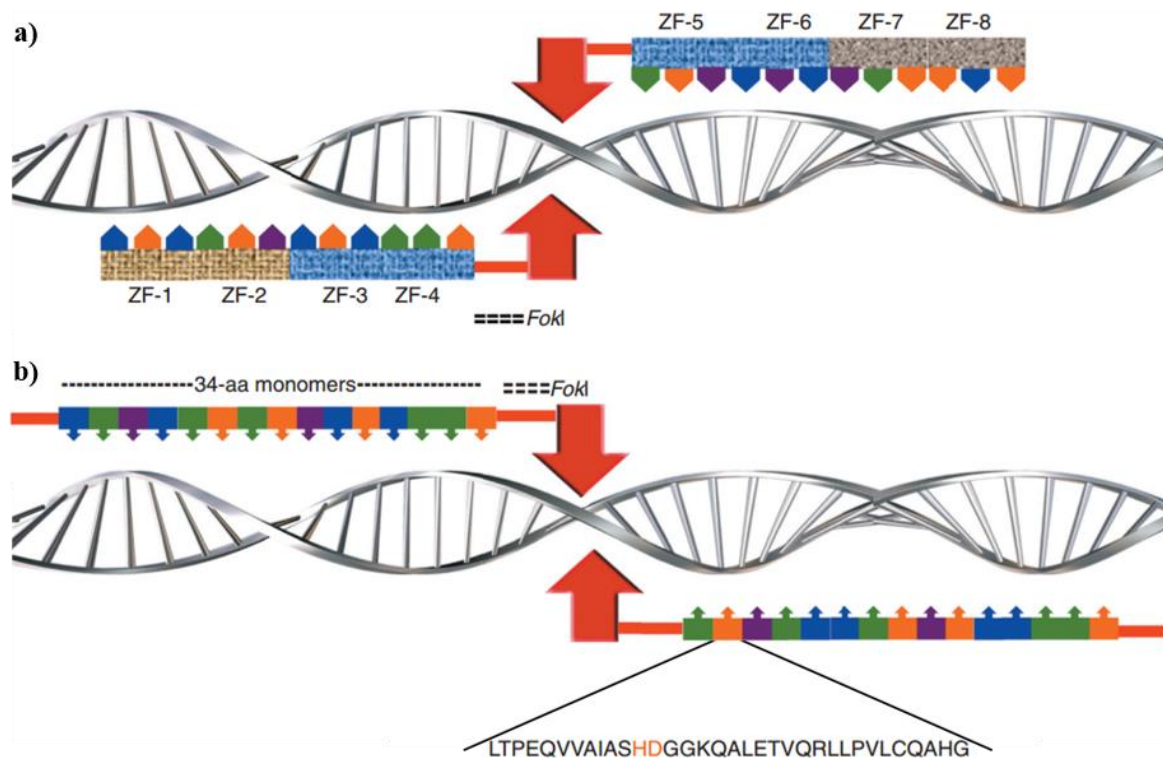
Figure 3.1.1: The Architecture of ZFNs and TALENs
Adapted from Carlson *et al.* (2012).
a) Schematic representation of a zinc-finger nuclease (ZFN). Each zinc finger (ZF) is able to bind three base pairs (coloured triangles). These units are combined to form a DNA-binding array, which is attached to the *Fok*I restriction endonuclease (red arrow).
b) Schematic representation of a transcription activator-like effector nuclease (TALEN). The TALE polypeptide consists of DNA-binding monomers (coloured triangles), which are composed of a conserved amino acid sequence. The 12th and 13th repeat variable diresidue (RVD; orange) dictates which DNA base the monomer will bind to. A TALEN is formed through the fusing of the TALE with a *Fok*I restriction endonuclease (red arrow).

Transcription activator-like effectors (TALEs) are DNA-binding domains, derived from the plant pathogen *Xanthomonas*, which are composed of a series of conserved 33-35 amino acid repeats (Figure 3.1.1b). These monomers only differ at the 12th and 13th positions, referred to as the repeat variable diresidue (RVD). There is a simple relationship between each RVD combination and its target DNA base, and this is how the TALE is able to identify a specific DNA sequence. TALE nucleases (TALENs) have been engineered by fusing the

*Fok*I restriction endonuclease to the DNA-binding domain and like ZFNs, the binding of two TALEN units is required for a double-stranded break to occur (Gaj *et al.* 2013; Guha *et al.* 2017). No experimental data has been published involving *Plasmodium* and TALENs, but *in silico* work has resulted in the design of TALENs that would target the *Plasmepsin V* gene of *P. falciparum* (Cermak *et al.* 2011). Work in dengue and malaria control on *Ae. aegypti* and *Anopheles gambiae* mosquitoes respectively has been performed using TALENs (Reegan *et al.* 2016).

3.1.2   CRISPR-Cas9

Clustered regulatory interspaced short palindromic repeats (CRISPRs) were first discovered in *E. coli* (Ishino *et al.* 1987) and together with the CRISPR-associated protein (Cas) form the only known prokaryotic adaptive immune system (Hille and Charpentier 2016). This RNA-based, defensive mechanism involves a group of Cas genes, that encode the Cas proteins, which are located near the CRISPR locus. This locus acts as the immunological memory and contains short repetitive sequences (repeats) that are separated by hypervariable sequences (protospacers), which have been obtained from previously invading pathogens (viruses and plasmids; Amitai and Sorek 2016). The molecular defence mechanism of the CRISPR-Cas system can be divided into three stages: adaptation, expression and maturation, and interference (Figure 3.1.2). During the adaptation stage, foreign DNA is recognised and Cas proteins acquire a new protospacer sequence, that is inserted into the CRISPR locus. Expression and maturation refer to the transcription of the CRISPR array into noncoding, precursor RNAs which are then processed into smaller CRISPR RNAs (crRNAs). Each crRNA, containing a single protospacer and part of the repeat sequence, then forms a complex with a Cas protein. The interference phase involves the Cas-crRNA recognising specific invading DNA, due to complementary base-pairing between the crRNA and the target, and cleaving it to produce a double-stranded break (Amitai and Sorek 2016; Cui and Yu 2016; Hille and Charpentier 2016).
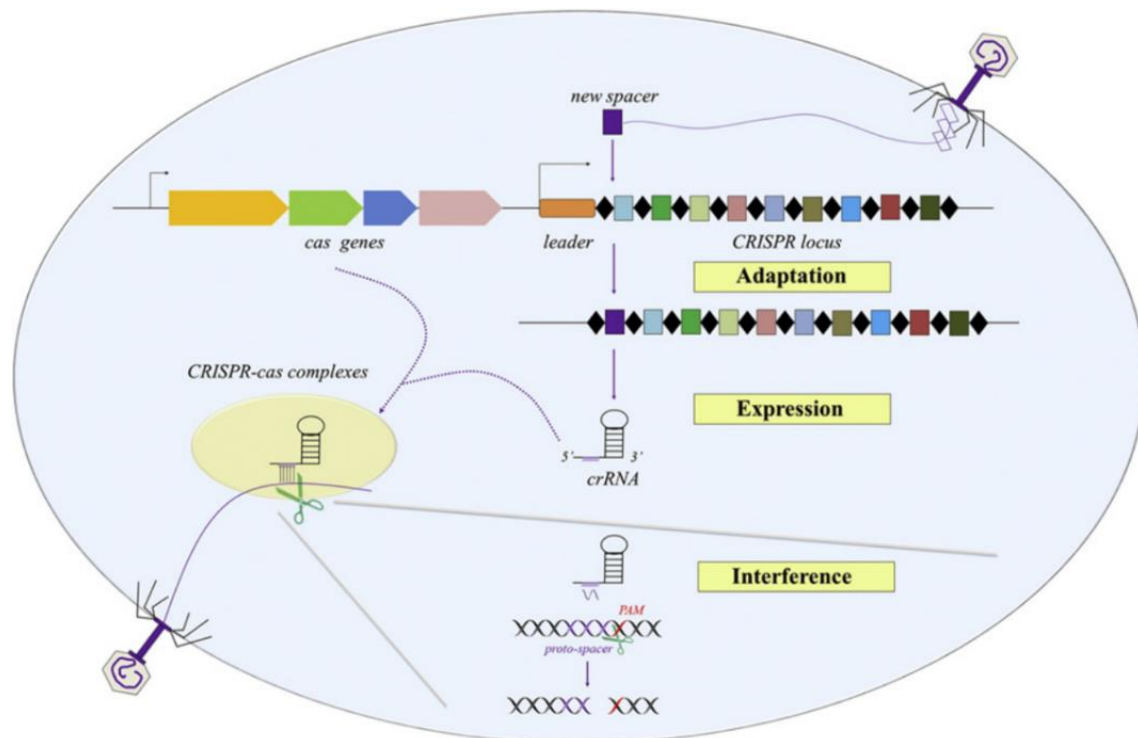
Figure 3.1.2: The Process of Type II CRISPR-Derived Immunity in Bacteria (Cui and Yu 2016) The CRISPR locus consists of protospacers (coloured blocks) separated by repeat sequences (black diamond), with the Cas genes nearby. During the adaptation phase foreign DNA is fragmented and a new protospacer (purple square) is inserted into the CRISPR locus. The expression stage involves the transcription and processing of protospacers and repeats into mature CRISPR RNAs (crRNAs). During the interference phase the Cas-crRNA complex targets and cleaves invading DNA by recognising the nucleotides complementary to the protospacer sequence.

In recent years CRISPR has been adapted for use as a gene editing tool in a variety of different organisms including bacteria, *Drosophila* and human cells. CRISPR-Cas type II has become the prominent system that is employed and makes use of the well characterised Cas9 endonuclease from *Streptococcus pyogenes*. This protein contains two active domains, the HNH and RuvC domains, which cleave the target and non-target DNA strands respectively. The Cas9 protein forms a complex with two RNA sequences: the trans-activating crRNA (tracrRNA) and the crRNA (Figure 3.1.3). The tracrRNA is a scaffold sequence that binds the Cas9 protein, while the crRNA contains the ~20 nucleotide "user-designed" guide RNA (gRNA) sequence which targets the complex to a specific gene locus. In engineered systems, these two RNA sequences are expressed as a single guide RNA (sgRNA) construct. The specificity of this system is due to the fact that the gRNA needs to be located upstream of a protospacer adjacent motif (PAM) sequence that is specific for each of the bacterial Cas9 proteins, in the case of *S. pyogenes* Cas9 it is NGG. Upon base-pairing between the gRNA and the target, the Cas9 nuclease creates a double-stranded break in the

DNA 3 bp upstream of the PAM sequence. The domains of the Cas9 protein can be mutated to produce "dead" Cas9 (dCas9), which is nuclease-deficient but still retains its DNA-binding characteristics. Fusing effector domains, such as the negative transcriptional KRAB (Krüppel-associated box) domain, with the dCas9 protein can allow for silencing or activation of the targeted gene, as opposed to cleavage (Doetschman and Georgieva 2017; Guha *et al.* 2017; Singh *et al.* 2017).
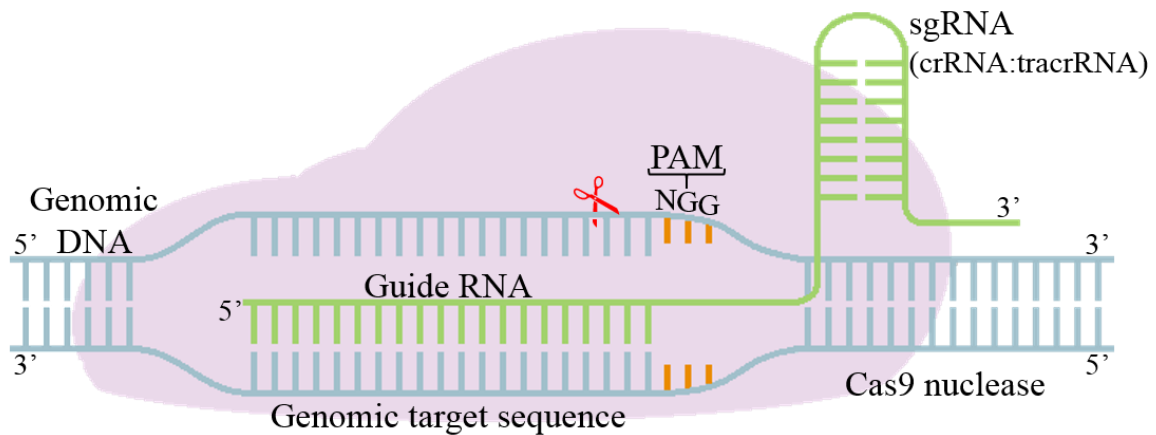
Figure 3.1.3: The CRISPR-Cas9 Gene Targeting System
Diagram showing the binding of the CRISPR-Cas9 complex to a target gene locus. The Cas9 protein (purple) forms a complex with the engineered single guide RNA (sgRNA; green), which is composed of the CRISPR RNA (crRNA) and the trans-activating crRNA (tracrRNA). The guide RNA (gRNA) targets the CRISP-Cas9 complex by binding to the genomic target sequence (blue), immediately upstream of the NGG protospacer adjacent motif (PAM) sequence. The Cas9 nuclease then cleaves the DNA (red scissors), creating a double-stranded break upstream of the PAM sequence. Adapted from Ceasar *et al.* (2016).

3.1.2.1   CRISPR-Cas9 and Malaria

CRISPR-Cas9 in P. falciparum. The CRISPR-Cas9 technology has recently been adapted for use in *Plasmodium*. In the first description of CRISPR-Cas9 in *P. falciparum*, Ghorbal *et al.* (2014) expressed the Cas9 nuclease and the sgRNA off separate plasmids using the U6 small nuclear RNA (snRNA) regulatory elements to recruit RNA polymerase III. These plasmids contained positive and negative drug selection cassettes, while the sgRNA plasmid also had the DNA templates for HR. Two non-essential genes, an integrated GFP gene and the knob-associated, histidine-rich protein (KAHRP) gene, were successfully disrupted in three weeks by the insertion of a selectable marker within their sequences (Ghorbal *et al.* 2014). Modifications to the coding sequences of the putative gene silencing *orc1* and artemisinin resistance kelch propeller domain were also obtained. Whole-genome

sequencing of parental and mutant lines did not reveal any off-target activity. *In silico* analysis of the *P. falciparum* genome predicted that there is a CRISPR-Cas9 target site on average every 35 base pairs, meaning virtually any gene locus could be targeted by CRISPR-Cas9 (Ghorbal *et al.* 2014). Two plasmids with drug selection cassettes, one containing the Cas9 nuclease and the sgRNA (each controlled using separate T7 promotors) and the other with the donor template and the T7 RNA polymerase, were used to disrupt the *P. falciparum* KAHRP and erythrocyte binding antigen 175 genes within ~4-6 weeks (Wagner *et al.* 2014).

CRISPR-Cas9 has been used in subsequent studies investigating *P. falciparum* drug resistance. It was shown that mutations in the cyclic amine resistance locus are sufficient to cause resistance to the imidazolopiperazine class of antimalarial compounds (LaMonte *et al.* 2016). The CRISPR-Cas9 system has confirmed that point mutations in the multidrug resistance-1 gene mediate resistance to the piperazine-containing compound ACT-451840 (Ng *et al.* 2016) and verified the mutations in the cleavage and polyadenylation specificity factor subunit 3 gene that are responsible for resistance to the benzoxaborole compound AN3661 (Sonoiki *et al.* 2017). This system has been used to introduce a bulky gatekeeper residue into the *P. falciparum* calcium-dependent protein kinase 1 in order to assess the activity of the protein (Bansal *et al.* 2016). CRISPR-Cas9 methods involving deletion of the *P. falciparum var2csa* intron have shown that, while the *var* intron may play a role in in the regulation and switching of the members of the *var* gene family, it is not required for silencing or activation of the *var2csa* gene (Bryant *et al.* 2017).

Modifications have been made to the *P. falciparum* CRISPR-Cas9 system allowing for "marker-free" gene editing (Lu *et al.* 2016; Mogollon *et al.* 2016). This technology is also being incorporated with other gene editing tools like the Cre-*loxP* recombination system (Knuepfer *et al.* 2017). A plasmid-free method, which involves transfecting *P. falciparum* parasites with the Cas9 protein, the guide RNA and the repair template, has also been developed (Crawford *et al.* 2017).

CRISPR-Cas9 in P. yoelii. Gene deletion, reporter knock-in and nucleotide replacement was performed in *P. yoelii* using a single plasmid. It contained the HR templates, the sgRNA and expressed off the same promotor the positive drug selection cassette and Cas9 protein, separated by the viral "ribosome skip" 2A peptide (Zhang *et al.* 2014). Genes important for

P. yoelii ookinete motility have been sequentially disrupted using a CRISPR-Cas9 mediated system (Zhang *et al.* 2017).

CRISPR-Cas9 in Anopheles. CRISPR-Cas9 has been used to manipulate *An. gambiae* and *An. stephensi* mosquitoes with the goal, for example, of controlling their population number or generating mosquitoes that are not able to transmit the malaria parasite (McLean and Jacobs-Lorena 2016; Reegan *et al.* 2016).

The studies described above indicate that gene editing is a useful technique to manipulate the *P. falciparum* genome and so CRISPR-Cas9 was used in this study to delete the *pfiap* gene.

3.1.3 Objectives

The objective of this chapter was to:

- Use CRISPR-Cas9 to facilitate knockout experiments to determine if the *pfiap* gene is essential for *P. falciparum* parasite survival

3.2 Materials and Methods

The preparation of the vector constructs described in this section follows the general protocol described in 2.2.4 - 2.2.6 with respect to the PCR amplification of gene regions, insertion of these amplicons into plasmids, bacterial transformation and the treatment of the resultant *E. coli* colonies.

3.2.1 CRISPR-Cas9 Plasmid Preparation

Two CRISPR-Cas9 vectors were adapted for use in *P. falciparum* by Ghorbal *et al.* (2014). The pUF1-Cas9 plasmid (Figure 3.2.1a) contains an engineered *S. pyogenes* Cas9 endonuclease (with a nuclear localisation signal) under the control of plasmodial regulatory heat shock protein elements. Resistance to DSM1, a *P. falciparum* dihydroorotate dehydrogenase (PfDHODH) inhibitor, is achieved through the presence of the drug-selectable marker *ydhodh*. The pL6-GOI plasmid (Figure 3.2.1b) makes use of *P. falciparum* U6 snRNA regulatory elements for the transcription of the sgRNA, for which there is a cloning site. Multiple cloning sites are present on either side of the positive selection *hdhfr* cassette to facilitate the insertion of the donor DNA templates. The vector also contains the suicide-gene yeast cytosine deaminase/uracil phosphoribosyl transferase (*yfcu*) which allows for the removal of parasites that contain the pL6-GOI plasmid. Cytosine deaminase (CD)

converts a harmless 5-fluorocytosine metabolite into 5-fluorouracil, which inhibits RNA synthesis and the enzyme thymidylate synthase (TS). Uracil phosphoribosyl transferase efficiently converts 5-fluorouracil into further metabolites that also inhibit RNA synthesis and TS. This suicide gene codes for a bifunctional chimeric protein that is 1 000-fold more potent than CD itself (Maier *et al.* 2006).
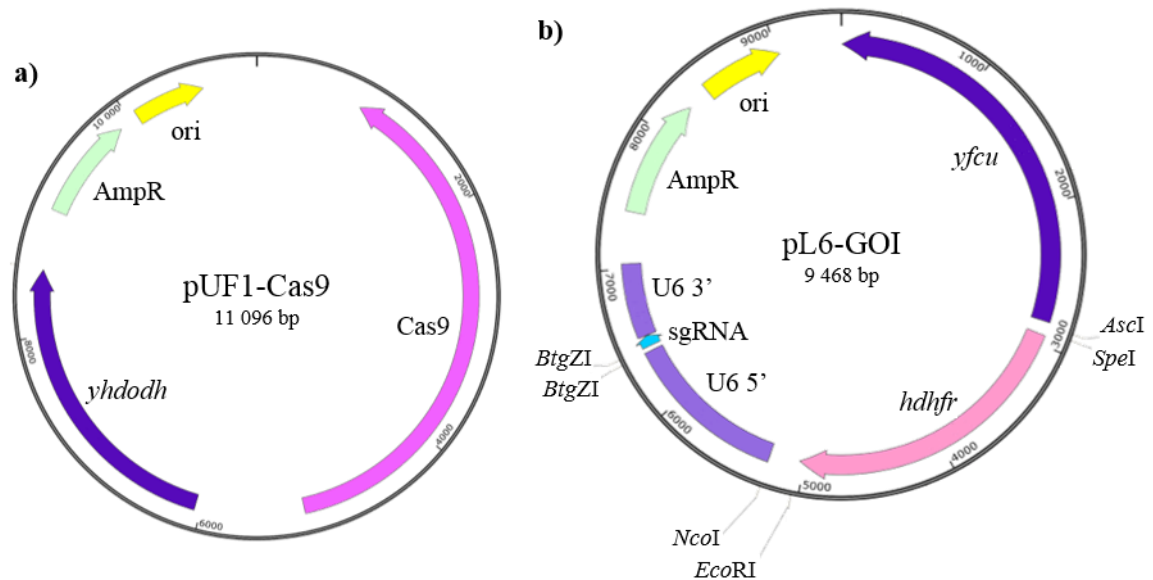
Figure 3.2.1: Plasmid Maps of the pUF1-Cas9 and pL6-GOI Vectors
a) A basic vector map of the pUF1-Cas9 plasmid, details are available in APPENDIX B.
b) A basic vector map of the pL6-GOI plasmid, details are available in APPENDIX B.

The pUF1-Cas9 and pL6-GOI plasmids were kindly donated by Dr Ghorbal and used to transform in-house, chemically competent XL10 *E. coli* cells (provided by Dr Sonja Lauterbach). A 50 µL aliquot of *E. coli* cells was incubated with either 25 ng of the pUF1-Cas9 vector or 25 ng of the pL6-GOI vector for 30 minutes on ice, before being heat-shocked at 42°C for 90 seconds. After transforming *E. coli* with the two samples, glycerol stocks of the plasmids were created from the resulting colonies.

3.2.2   Deletion of the *pfiap* Gene

3.2.2.1   Construction of the pL6-DEL Vector

Knockout of the *pfiap* gene required the creation of the intermediate pL6-DEL vector which involved the insertion of *pfiap* homologous arms into the pL6-GOI plasmid (Figure 3.2.2).
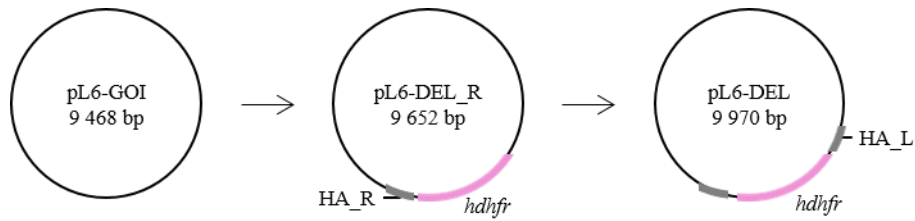
Figure 3.2.2: Diagram of the Steps in the Creation of the pL6-DEL Plasmid
Diagram showing the process followed to generate the pL6-DEL vector. Grey = left/right *pfiap* homologous arms (HA_L/R); pink = human dihydrofolate reductase cassette (*hdhfr*).

The 5' and 3' untranslated regions (UTRs) of the *pfiap* gene were amplified by PCR using DreamTaq Green PCR Master Mix (Figure 3.2.3) to create the left and right homologous arms (HA_L/R) respectively. The PCR amplicons were created with the appropriate restriction sites (DEL_L Fwd – *Asc*I, DEL_L Rev – *Spe*I, DEL_R Fwd – *Eco*RI, DEL_R Rev – *Nco*I) to allow for the sequential insertion of the homologous arms, on either side of the *hdhfr* cassette. Detailed primer sequences are available in APPENDIX A.

Figure 3.2.3: Location of the Knockout Homologous Arms of the *pfiap* Gene
Diagram showing the two *pfiap* UTRs amplified by PCR as homologous arms. Blue = the forward and reverse PCR primers; grey = *pfiap* UTRs; purple = *pfiap* gene. Numbers = nucleotide number.

Table 3.2.1: PCR Parameters for the Amplification of *pfiap* Homologous Arms

Region Name	Vector	*P. falciparum* Specific T_a (°C)	Full Length Primer T_a (°C)	PCR Product (bp)
DEL_R	pL6-GOI	45.0	56.0	396
DEL_L	pL6-DEL_R	47.0	59.0	376

The pL6-GOI vector was extracted from XL10 *E. coli* and together with the DEL_R amplicon, was prepared for cloning by digestion with FastDigest® *Eco*RI and *Nco*I restriction enzymes, before being ligated together. A 100 µL aliquot of in-house competent XL10 *E. coli* cells were transformed, after which successful transformation was verified using the FastDigest® *Xba*I restriction enzyme and sequencing of the new pL6-DEL_R plasmid construct. The DEL_L PCR amplicon was then cloned into the pL6-DEL_R vector using FastDigest® *Asc*I and *Spe*I restriction enzymes, followed by restriction enzyme digestion (FastDigest® *Bam*HI) and sequencing to verify the new pL6-DEL plasmid construct.

3.2.2.2 Creation of the pL7-DEL Vector

The second step in creating the knockout vector involved inserting the selected gRNA into the pL6-DEL plasmid (Figure 3.2.4).
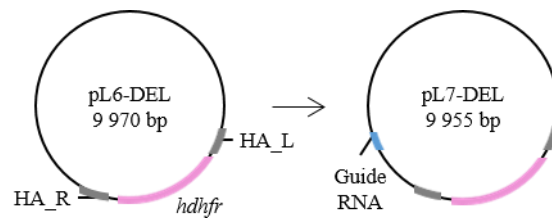
Figure 3.2.4: Diagram of the Steps in the Creation of the pL7-DEL Plasmid
Diagram showing the generation the pL7-DEL vector. Blue = guide RNA; grey = left/right *pfiap* homologous arms (HA_L/R); pink = human dihydrofolate reductase cassette (*hdhfr*).

The Cas9 endonuclease will only bind to the target DNA if the unique ~20 bp guide sequence is immediately upstream of a three nucleotide PAM site. The PAM sequence (GGG) used for the knockout of the *pfiap* gene is on the antisense strand towards the end of the BIR domain (Figure 3.2.5a). The 20 bp preceding the PAM site were selected (Figure 3.2.5b) and synthesised by Integrated DNA Technologies, USA as complementary 5'→3' and 3'→5' oligonucleotides (DEL (+)/DEL (-); available in APPENDIX A). A G nucleotide was included at the 5' end of the selected sequence (this is the preferred initiation nucleic acid and so increases the transcription efficiency of the U6 promoter; Ma *et al.* 2014), while 15 bp homologous to the pL6-DEL vector were included on both the 5' and 3' ends of the sequence (Figure 3.2.5c and d). Even though these oligonucleotides are DNA they will be referred to as the guide RNA, as the sequence will ultimately be transcribed as part of the sgRNA.

Figure 3.2.5: The *pfiap* Knockout Guide RNA

a) Diagram showing the *pfiap* knockout PAM site. Grey = *pfiap* UTRs; purple = *pfiap* gene; red = the location of the PAM site. Numbers = number of nucleotides.

b) Diagram showing the PAM site and the selected guide RNA used for *pfiap* knockout. Green = the selected guide sequence; red = the location of the PAM site. Numbers = nucleotide number.

c) Diagram showing the complementary guide RNA sequences annealed together. Green = the selected guide sequence; grey = 15 bp pL6-GOI homologous regions.

d) Diagram showing the pL7-DEL plasmid with the inserted guide RNA sequence and the 15 bp pL6-DEL homologous regions synthesised as part of the (+)/(-) oligonucleotides.

These guide oligonucleotides were resuspended to a final concentration of 100 μM using nuclease-free water, before they were annealed by means of a step-down PCR program (Ghorbal *et al.* 2014).

Table 3.2.2: Guide RNA Reaction Set-Up

Component	Volume	Final Concentration	Manufacturer
10X NEBuffer 2	2.22 μL	1X	New England Biolabs
DEL (+)	10.00 μL	45 μM	-
DEL (-)	10.00 μL	45 μM	-
Total	**22.22 μL**		

95,0 95,0
10:00 0:01
- 0,6 °C 85,0 85,0
0,6 °C/s 1:00 0:01
- 0,6 °C 75,0 75,0
0,6 °C/s 1:00 0:01
- 0,6 °C 65,0 65,0
0,6 °C/s 1:00 0:01
- 0,6 °C 55,0 55,0
0,6 °C/s 1:00 0:01
- 0,6 °C 45,0 45,0
0,6 °C/s 1:00 0:01
- 0,6 °C 35,0 35,0
0,6 °C/s 1:00 0:01
- 0,6 °C 25,0
0,6 °C/s 1:00

17 cycles   17 cycles   17 cycles   17 cycles   17 cycles   17 cycles   17 cycles
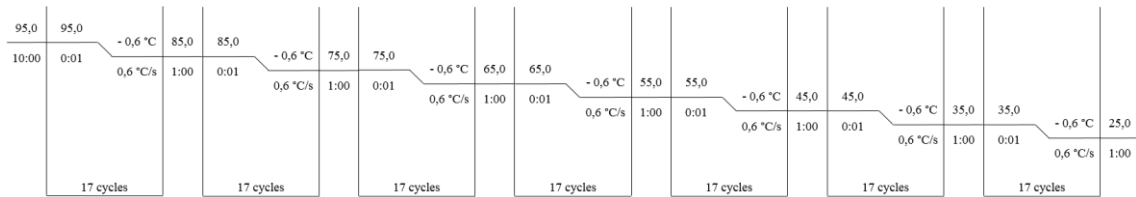
Figure 3.2.6: Guide RNA PCR Annealing Program
Diagram showing PCR reaction conditions used during the guide RNA annealing reaction.

The annealed guide RNA was stored on ice and then diluted to 0.5 µM with ice-cold 5 mM Tris-HCl; pH 8. The pL6-DEL vector was extracted from XL10 *E. coli* and digested using *Btg*ZI for 2 hours at 60°C before being purified using the NucleoSpin® Gel and PCR Clean-Up Kit as per the manufacturer's instructions and eluted in 10 µL of nuclease-free water. The guide RNA was cloned into the digested pL6-DEL vector (which had been quantified by agarose gel electrophoresis) using Gibson Assembly as per the manufacturer's instructions (Gibson *et al.* 2009). The Gibson Assembly Master Mix contains a 5' exonuclease which removes some of the exposed 5' nucleotides from all of the fragments. The homologous regions of the fragments anneal and a DNA polymerase extends the 3' ends, before a DNA ligase seals the exposed nicks (Figure 3.2.7). Cloning using Gibson Assembly has the advantage that only the vector, but not the insert, needs to be digested prior to cloning and can be used to assemble multiple fragments in one reaction.

a)

<pre>
                                          BtgZI   BtgZI
5'...TAAGTATATAA|TATT|CCTAGGAACTCATCGCTCGCGATGCTGCCCGACA|GTTTTAGAGCTAGAA..3'
3'...ATTCATATATTATAA|GGATCCTTGAGTAGCGAGCGCTACGACGGGCTGT|CAAA|ATCTCGATCTT..5'
</pre>

<pre>
                                   │ BtgZI
                                   ↓
5'...TAAGTATATAA   GTTTTAGAGCTAGAA..3'
3'...ATTCATATATTATAA   ATCTCGATCTT..5'
</pre>

b)

pL6-DEL
<pre>
5'...TCATATTAAGTATATAA   GTTTTAGAGCTAGAAATAGCA...3'
3'...AGTATAATTCATATATTATAA   ATCTCGATCTTTATCGT...5'
</pre>

                                   +

Guide RNA
<pre>
      5' TAAGTATATAATATTGCCATAAGATTATTTCTCCATGTTTTAGAGCTAGAA 3'
      3' ATTCATATATTATAACGGTATTCTAATAAAGAGGTACAAAATCTCGATCTT 5'
</pre>

                                   │ Exonuclease removes 5' ends
                                   ↓

<pre>
5'...TCATATTAAGTATATAA   CA...3'
3'...AG   ATCTCGATCTTTATCGT...5'
</pre>

<pre>
      5'   ATAAGATTATTTCTCCATGTTTTAGAGCTAGAA 3'
      3' ATTCATATATTATAACGGTATTCTAATAAAGAGG   5'
</pre>

                                   │ DNA fragments anneal
                                   ↓

<pre>
5'...TCATATTAAGTATATAA   ATAAGATTATTTCTCCATGTTTTAGAGCTAGAA   CA...3'
3'...AG   ATTCATATATTATAACGGTATTCTAATAAAGAGG   ATCTCGATCTTTATCGT...5'
</pre>

                                   │ DNA polymerase & ligase
                                   ↓

pL7-DEL
<pre>
5'...TCATATTAAGTATATAATATTGCCATAAGATTATTTCTCCATGTTTTAGAGCTAGAAATAGCA...3'
3'...AGTATAATTCATATATTATAACGGTATTCTAATAAAGAGGTACAAAATCTCGATCTTTATCGT...5'
</pre>

Figure 3.2.7: Insertion of the Guide RNA into the pL6-DEL Vector
a) Diagram showing the digestion of the pL6-DEL vector with the *Btg*ZI restriction enzyme. Red underlined = the *Btg*ZI recognition sites; red cut = the *Btg*ZI cleavage sites
b) Diagram showing the creation the pL7-DEL plasmid using Gibson Assembly. A 5' exonuclease removes base pairs from the 5' ends, which allows the homologous regions to anneal before the missing base pairs are replaced using a DNA polymerase to extend the 3' ends and a DNA ligase to seal the nicks (Gibson *et al.* 2009). Green = the selected guide sequence; orange = pL6-DEL plasmid; grey = 15 bp pL6-DEL homologous regions.

Table 3.2.3: Gibson Assembly Reaction Set-Up

Component	Volume	Final Concentration	Manufacturer
2X Gibson Assembly Master Mix	5.0 μL	1X	New England Biolabs
Annealed Guide RNA	1.5 μL	75 nM	-
Digested pL6-DEL Vector	x μL	110 ng	-
Nuclease-Free Water	y μL	-	-
Total	**10.0 μL**		

The Gibson Assembly reaction was incubated for 30 minutes at 50°C, after which 5 μL of the reaction mixture was used to transform a 100 μL aliquot of in-house XL10 *E. coli* cells (see 3.2.2.1). Successful transformation and creation of the pL7-DEL vector construct was

verified using the FastDigest® *Xba*I restriction enzyme, followed by plasmid sequencing to ensure the construct was error-free.

3.2.2.3   Transfection of *P. falciparum* Parasite Cultures

The pL7-DEL and pUF1-Cas9 plasmid constructs were prepared as described in 2.2.11.1, except that the constructs were prepared from one litre LB cultures (four 250 mL aliquots) and quantified after linearization with FastDigest® *Eco*RI. Elution volumes equivalent to 150 µg of pL7-DEL and pUF1-Cas9 plasmid were purified and the plasmids were each resuspended in 15 µL of TE buffer. The two plasmid solutions were combined and added to 370 µL of pre-warmed cytomix, before being used to transfect a *P. falciparum* culture (see 2.2.11.2). This transfection experiment was performed three times, each time the transfected culture was maintained under 2 nM WR99210 selection pressure for more than two months.

3.3   Results

3.3.1   Knocking Out the *pfiap* Gene

3.3.1.1   Creation of the pL7-DEL Plasmid Construct

The 396 bp 3' UTR of the *pfiap* gene was cloned into the *Eco*RI/*Nco*I site of the pL6-GOI vector (Figure 3.3.1).
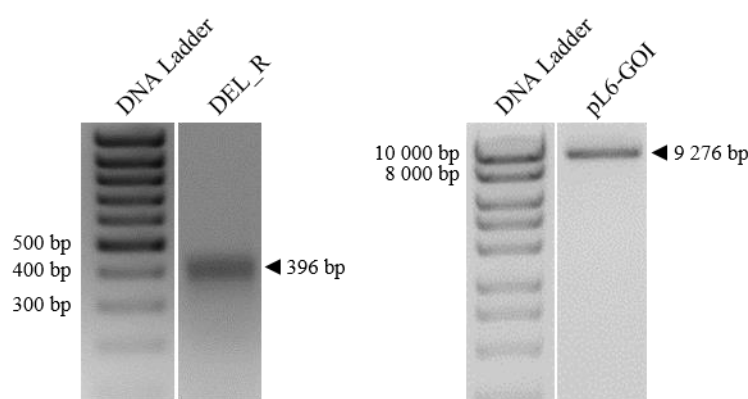
Figure 3.3.1: DEL_R PCR Amplicon and pL6-GOI Plasmid prior to Ligation
Agarose gel (1%) showing the *Eco*RI/*Nco*I-digested DEL_R PCR amplicon and pL6-GOI vector.
◄ = expected band size.

Digestion of the newly created pL6-DEL_R plasmid using *Xba*I resulted in the expected banding pattern on the agarose gel (Figure 3.3.2a). Sequence analysis of one sample confirmed that the expected DEL_R arm was present in the pL6-DEL_R vector, but did reveal that two nucleotides (AT) were absent from the middle of the insert (Figure 3.3.3). The DEL_R fragment was generated by PCR using the DreamTaq DNA polymerase which

is a low fidelity enzyme that does not have any proofreading ability (3'→5' exonuclease activity). It is therefore understandable why there are errors in this fragment, particularly in a stretch of DNA that contains repeating AT nucleotides. The DreamTaq enzyme was used because PCR of the DEL _R arm using various high fidelity polymerase enzymes was unsuccessful. Since this deletion occurred in the *pfiap* 3' UTR homologous arm and represented less than 0.5% of the nucleotides in the UTR, it was decided that the effect of these mutations would be negligible and that the recombination event would still be able to occur.

Figure 3.3.2: Verification of the Insertion of DEL_R into the pL6-GOI Vector
a) Agarose gel (1%) of the pL6-DEL_R plasmid following cloning, after digestion with *Xba*I.
◄ = expected band size.
b) Vector diagrams of the pL6-GOI plasmid and the plasmid containing the 3'UTR DEL_R insert with *Eco*RI/*Nco*I and *Xba*I restriction sites. Grey = the 3' URT homologous insert.

Figure 3.3.3: Sequencing Chromatogram of the pL6-DEL_R Plasmid
Extract of the sequencing chromatogram of the DEL_R insert present in the pL6-DEL_R vector.
Purple = the location of the missing AT nucleotides.

The 5' UTR of the *pfiap* gene (DEL_L) was PCR-amplified as a single amplicon, digested with *Asc*I and *Spe*I and inserted into the prepared pL6-DEL_R plasmid to create the pL6-DEL construct (Figure 3.3.4).
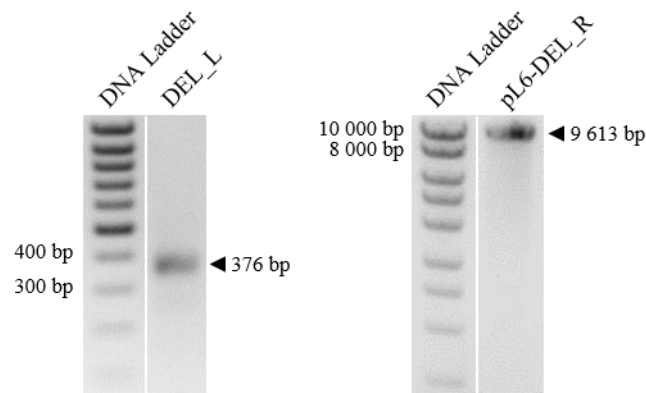
Figure 3.3.4: Digested DEL_L PCR Amplicon and pL6-DEL_R Plasmid
Agarose gel (1%) showing the DEL_L PCR amplicon and pL6-DEL_R vector, digested with *Asc*I/*Spe*I, prior to ligation. ◀ = expected band size.

Following *E. coli* transformation, digestion of the pL6-DEL plasmid with *Bam*HI and *Eco*RI showed the expected banding pattern (Figure 3.3.5a) Digestion of the initial pL6-DEL_R plasmid would have resulted in a band of expected size 1 247 bp as opposed to the observed band of size 1 563 bp. Two mutations were identified after sequencing one sample insert: an AT nucleotide pair were missing, and an A had been replaced by a G (Figure 3.3.6). The use of the DreamTaq enzyme could explain the discrepancies between the expected and achieved sequences, but as these minor mutations should not interfere with the homologous recombination event the pL6-DEL plasmid was taken forward.
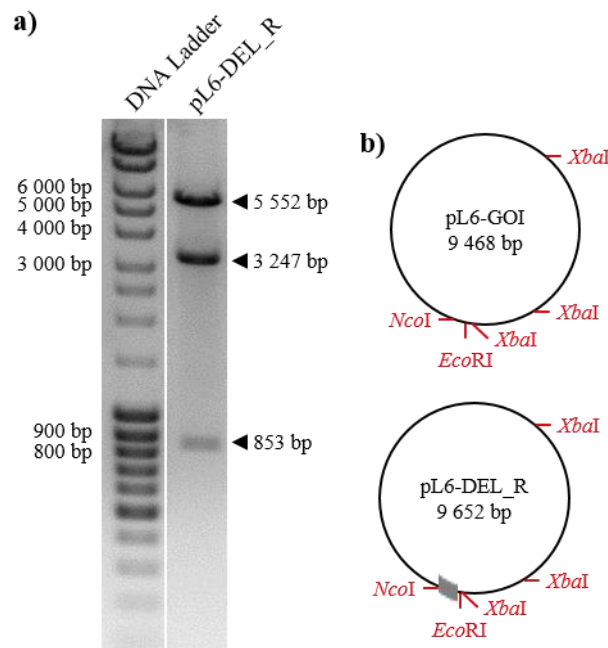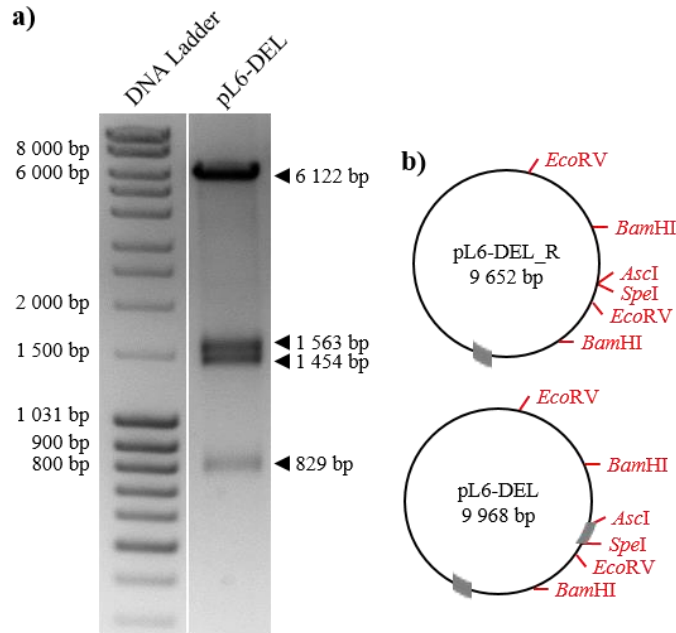
Figure 3.3.5: Verification of the Insertion of DEL_L into the pL6-DEL_R Vector

a) Agarose gel (1%) of the pL6-DEL plasmid after digestion with *Bam*HI/*Eco*RV. ◄ = expected band size.

b) Vector diagrams of the pL6-DEL_R plasmid and the pL6-DEL plasmid containing both DEL_R & DEL_L inserts with *Asc*I/*Spe*I and *Bam*HI/*Eco*RV restriction sites. Grey = the 5' & 3' URT homologous inserts.
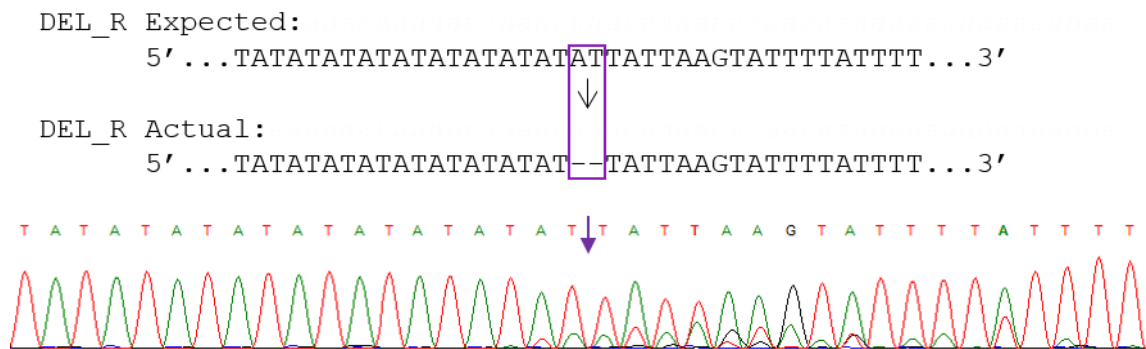

Figure 3.3.6: Sequencing Chromatogram of the pL6-DEL Plasmid

a) Extract of the sequencing chromatogram of the DEL_L construct present in the pL6-DEL vector. Purple = the identified A→G substitution.

b) Extract of the sequencing chromatogram of the DEL_L insert present in the pL6-DEL vector. Purple = the location of the missing AT nucleotides.

The pL6-DEL vector was digested with *Btg*ZI and placed in a Gibson Assembly reaction with the annealed DEL guide RNA oligonucleotides to create the pL7-DEL vector. Following *E. coli* transformation, digestion of the newly created plasmid with *Xba*I resulted in what was assumed to be the expected banding pattern (Figure 3.3.7a). It was not possible to detect a 20 bp insert in the 5 539 bp largest band, but sequencing of one sample of the pL7-DEL vector verified that the guide RNA was present and error-free.
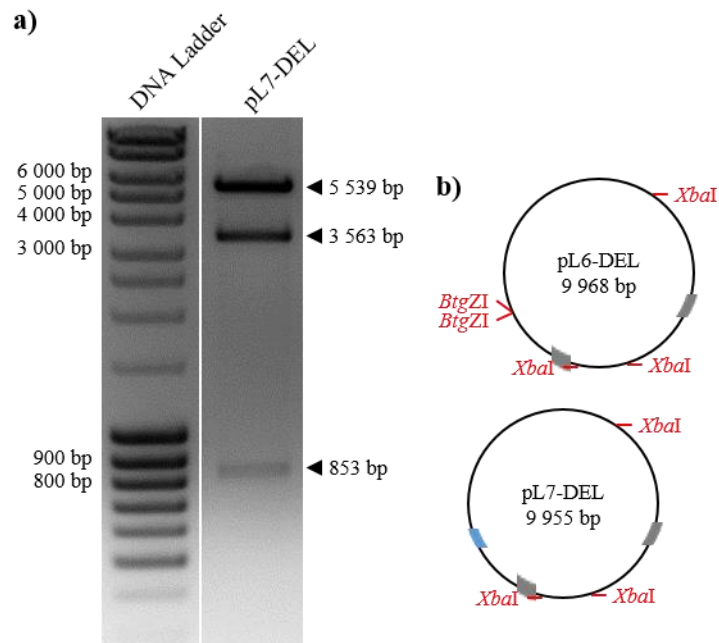
Figure 3.3.7: Verification of the Insertion of the DEL gRNA into the pL6-DEL Vector
a) Agarose gel (1%) of the pL7-DEL plasmid with *Xba*I. ◄ = expected band size.
b) Vector diagrams of the pL6-DEL plasmid and the plasmid containing the DEL gRNA insert with *Btg*ZI and *Xba*I restriction sites. Blue = the guide RNA insert; grey = the URT homologous inserts.

3.3.1.2   Knocking Out *pfiap* by Transfecting *P. falciparum* with pL7-DEL

The circular pL7-Del and pUF1-Cas9 plasmids (quantified in Figure 3.3.8) were used to transfect ring stage *P. falciparum* 3D7 parasites, with constant 2 nM WR99210 pressure after 48 hours, on three different occasions. After more than 60 days of monitoring, no parasites were detected in any of the transfected culture flasks.
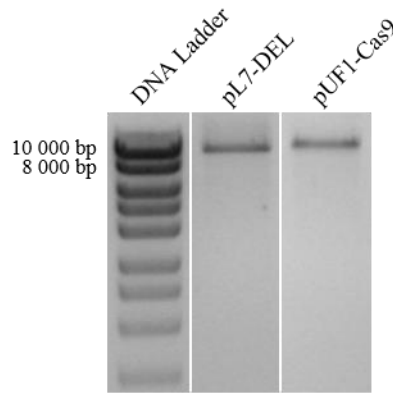
Figure 3.3.8: Linearized pL7-DEL and pUF1-Cas9 prior to Parasite Transfection
Agarose gel (1%) showing the *Eco*RI-linearized pL7-DEL and pUF1-Cas9 plasmids before parasite transfection. Expected sizes: pL7-DEL = 9 955 bp; pUF1-Cas9 = 11 096 bp.

3.4 Discussion

A key part of characterising an unknown protein is determining if the protein plays a critical role in the cell, during the life stage under investigation. To this end, the CRISPR-Cas9 genome editing tool was utilised in this study in attempts to disrupt the *pfiap* gene, as a means of gathering more information about the PfIAP protein.

3.4.1 Genome Editing in *Plasmodium* and the Advantage of CRISPR-Cas9

Following the introduction of transfection techniques in the mid-1990s and their implementation in the analysis of the *P. falciparum* genome, electroporation has emerged as a tool that provides consistent transfection reliability. The trade-off is that transfection experiments require large amounts of the plasmid to be transfected and the timeframe for recovery is variable (up to three months in some cases), with the need for constant monitoring for the re-emergence of drug-resistant parasites (Lee *et al.* 2014; Webster and McFadden 2014). Although there has been movement in the development of genetic tools for the investigation of *Plasmodium* genes, difficulties exist and these hinder progress. One problem is the study of genes that are critical to the parasite's survival, as the disturbance of such genes is lethal. Another is the low transfection efficiency that is experienced and how this limits the high-throughput screening of multiple genes. A third is the scarceness of positive selectable markers, which restricts the number of consecutive genetic operations that can be performed on the same parasites (Webster and McFadden 2014; DeKoning-Ward *et al.* 2015).

The primary reason for the long delay between transfection of a *P. falciparum* culture and the appearance of genetically modified parasites is the fact that traditional gene manipulation techniques rely on randomly occurring, stochastic HR events to introduce the modification into the GOI. The benefit of employing an engineered gene editing tool is that HR is no longer a rare event as the selected nuclease actively creates a DSB at the target location. This forces the cell to repair the damaged DNA, thereby facilitating the introduction of the desired change into the genome at an accelerated rate (Lee and Fidock 2014; DeKoning-Ward *et al.* 2015).

Several tools have been developed to improve the efficiency of genomic manipulation experiments, including ZFNs and TALENs. Although the specificity associated with ZFN binding is quite high, a hindrance to its use is the arduous process of engineering the user-defined ZF arrays which are context-dependent and this impacts the reproducibility and scale to which they can be used (Sander and Joung 2014; Eid and Mahfouz 2016). The simple, modular design is a benefit to using the TALEN system, but the highly repetitive composition of the multiple domain sequences and the size of the nuclease can ultimately lead to difficulties (Sander and Joung 2014; Ceasar *et al.* 2016). The CRISPR-Cas9 system is comparatively simpler to use, not requiring a new protein to be engineered for each GOI. The targeting gRNA constructs are easy to make, the system is cost-effective and can be used to edit multiple genes simultaneously (DeKoning-Ward *et al.* 2015; Ceasar *et al.* 2016; Eid and Mahfouz 2016). Caution though must be employed when using the CRISPR-Cas9 system as off-target effects have been noted. It is likely that this unintended Cas9 binding to DNA can be mediated by the perfect complementation of only the ten guide nucleotides proximal to the PAM site, out of the twenty total gRNA bases (Kuscu *et al.* 2014). In *P. falciparum* CRISPR-Cas9 has been used to realise genome modifications within three to six weeks (Ghorbal *et al.* 2014; Wagner *et al.* 2014).

3.4.2 The *pfiap* Gene may be Essential

Three separate *P. falciparum* cultures were each co-transfected with the pL7-DEL and pUF1-Cas9 plasmid constructs and maintained under positive selection pressure for more than two months. No parasites were detected in any of these transfected cultures which would mean that a homologous recombination event had occurred, replacing the *pfiap* gene with the *hdhfr* cassette that provides resistance to WR99210, but that the lack of the *pfiap* gene and the subsequent absence of the PfIAP protein was detrimental to the survival of the

P. falciparum parasite. This implies that the *pfiap* gene is essential in *P. falciparum* asexual parasites. In mice models *iap* deletion has resulted in various phenotypes including being embryonically lethal when survivin or Apollon (BIRC6), for example, were knocked out (Silke and Vaux 2015), so it is not unexpected that the deletion of *pfiap* has a detrimental effect on the parasite.

By itself, the absence of evidence is not sufficient to prove a negative result, so alternative options to account for the lack of viable *P. falciparum* parasites may also be considered. Technical reasons could include that electroporation of the RBCs with the plasmid containing the *hdhfr* cassette was not successful, and that the subsequent application of the WR99210 compound simply killed all of the parasites, or that any viable parasites died as a result of suboptimal culturing. These explanations are unlikely as *P. falciparum* parasites have been successfully transfected with a pARL2-GFP vector (see 2.3.7) and it is not likely that a parasite culturing error would have occurred on three separate occasions. Performing a transfection using a CRISPR-Cas9 vector designed to target an established, non-essential gene (such as PfEBA175) could act as a culturing and transfection technique control.

Transfected parasites were maintained in culture for more than 60 days, therefore viable *pfiap* knockout parasites should have been detected since the CRISPR-Cas9 system should reduce the time for transfected parasites to re-emerge when compared to conventional methods (Ghorbal *et al.* 2014). During this time there was only selection pressure for the pL7-DEL plasmid, not for the pUF1-Cas9 vector in the form of the PfDHODH inhibitor DSM1, which means that the transfected parasites would eventually lose the pUF1-Cas9 plasmid. This transient expression of the Cas9 protein though should still be sufficient to induce a DSB as cleavage and repair of a target sequence has been shown to occur within the first few growth cycles following transfection (Ghorbal *et al.* 2014).

Not all PAM sites and guide RNAs function with the same level of efficiency (Doench *et al.* 2014) and so it is possible that the guide RNA used during these experiments was not optimal. *pfiap* deletion experiments can be attempted after identifying other PAM sites within the BIR domain and cloning their associated gRNA into the pL6-DEL vector. These new targeting sequences might be more effective than the gRNA used in this study, but while a suboptimal gRNA might have decreased the chances that the Cas9 nuclease would bind the target site and create a DSB, it does not mean that the recombinant knockout event could

not occur. The only difference between the pL6-GOI vector used in this study and the pCC-1 vector, which is the vector that has traditionally been used in knockout experiments and relies solely on spontaneous recombination events, is that the pL6-GOI plasmid contains the sgRNA sequence and its associated promotor. This means that the pL7-DEL vector with the *pfiap* homologous arms, which is always present due to the constant positive drug pressure, would still be able to participate in a spontaneously occurring recombination event and replace the *pfiap* gene with the *hdhfr* cassette. This point also covers any scenario in which some of the *P. falciparum* parasites might only have been transfected with the pL7-DEL plasmid and not successfully co-transfected with both pL7-DEL and PUF1-Cas9 vectors.

3.4.3 Alternative Approaches for Manipulating *pfiap*

The absence of *P. falciparum* parasites following *pfiap* knockout transfection does not conclusively prove that *pfiap* is essential for *P. falciparum* intraerythrocytic development. Unfortunately there is no ideal strategy that can be followed to investigate and interfere with an essential asexual gene and there is not a globally accepted number of attempts that should be undertaken before it is concluded that the GOI is indispensable (Webster and McFadden 2014).

As part of building a case that a GOI is essential, experiments can be done to show that the gene is targetable. This can involve single-crossover homologous recombination to, for example, insert a HA coding sequence at the 5' start of the GOI or a GFP sequence at the 3' end of the gene to produce HA- and GFP-tagged proteins respectively, that can be used in subsequent molecular experiments. Once *pfiap* has been shown to be targetable with no loss-of-function, then a single-crossover modification of the coding region could be performed to show that it is the disruption of the gene itself that results in no viable parasites. An additional step may involve the complementation system (Sultan *et al.* 2001). It should be possible to knock out the genomic *pfiap* gene in an established transgenic *P. falciparum* line with an episomal plasmid that expresses the PfIAP protein. Once the gene deletion has been confirmed in viable parasites, then negative selection can be used to remove the episomal plasmid. If any parasites survive this step, then it means that *pfiap* is not an essential gene as no copies of the gene would be present in the *P. falciparum* culture.

Approaches not focusing on the removal of *pfiap* can also be employed in the quest to obtain information about the function of PfIAP. Examples of these techniques include conditional

inactivation of *pfiap* transcription, mRNA knockdown or the fusion of the Shield-controlling destabilizing domain to the PfIAP protein (Webster and McFadden 2014). Future *in vivo* work employing any number of molecular techniques should be able to provide more information about the function of this putative *P. falciparum* inhibitor of apoptosis protein.

CHAPTER 4: Concluding Remarks

While the willingness of a unicellular organism to commit suicide is hard to comprehend, the need for a *P. falciparum* population to be able to control its number by sacrificing some individual parasites is easier to understand. The exponential, asexual growth of the parasite during its 48 hour intraerythrocytic life stage is capable of killing the human host before the 12 days required for sexual gametocyte development have passed. Various studies have shown that some form of RCD occurs in *P. falciparum*, but there is a lack of consensus about the phenotype of this cell death. No RCD protein network has been elucidated, but a few possible participants, such as the metacaspase family of proteins, have been investigated. This study adds to the pool of available *Plasmodium* cell death knowledge by investigating a putative *P. falciparum* IAP protein.

Metazoan IAP proteins are primarily known as anti-apoptotic proteins as they are able to interact with and inhibit various pro-apoptotic caspases, although they also perform other cellular roles. The PfIAP protein contains one N-terminal BIR domain, the characteristic domain of its metazoan IAP homologues. This *P. falciparum* domain contains the conserved amino acids required for zinc ion binding and is structurally similar to human BIR domains, containing the expected α-helices and β-sheet conformations. This suggests that PfIAP could have the same function as the human IAP proteins, but given the wide range of their roles in the cell nothing definite can be extrapolated to the function of PfIAP. One orthologue of PfIAP is present in various other *Plasmodium* species, including all the human parasites. Given the intraerythrocytic differences and the reduced population pressure in these other species, it is unclear why they would need this proposed anti-apoptotic IAP protein to be expressed during the RBC life stage, unless IAP plays other roles in the parasite that pertain to normal cellular functioning.

These non-apoptotic roles of PfIAP, suggested by both the human proteins and the *Plasmodium* orthologues, are supported by the wide spectrum of PfIAP binding partners that were identified during the biopanning experiments. The functions of these proteins range from invasion to transcription, nutrient intake to protein synthesis and they localise to various sites in the parasite, some are even exported to the host cell membrane. The *in vivo* relevance of these interactions still needs to be confirmed experimentally, as does the localisation of the full length PfIAP protein. Of the PfIAP binding partners and their

associated interacting partners, PfATG23 is the only protein that has an obvious connection to RCD, although there are many unknown proteins in this protein-protein network.

The inference of the CRISPR-Cas9 knockout experiments is that the *pfiap* gene could be essential for *P. falciparum* survival during the intraerythrocytic life stage. This effect is conceivable given the expansive PfIAP centred protein-protein network and the fact that it covers numerous biological functions. To build a more complete picture of the role of the *pfiap* gene additional genomic manipulation, such as tagging the gene, is required. At the same time further information could be obtained about the PfIAP protein by generating, for example, a parasite line in which the PfIAP protein can be conditionally knocked down.

The findings of this study strongly suggest that PfIAP has a non-apoptotic role in *P. falciparum*, as well as a potential RCD function that is currently undefined. Future work investigating PfIAP and the various components of its protein-protein network is needed before a true role can be assigned to PfIAP.

APPENDIX A: Primers

Table A 1: PCR Cycling Protocols

Step	DreamTaq PCR Master Mix / High Fidelity PCR Enzyme Mix		Phusion Flash PCR Master Mix		Number of Cycles
	Temperature (°C)	Time	Temperature (°C)	Time	
Initial Denaturation	94/95	3 min	98	10 sec	1
Denaturation	94/95	30 sec	98	1 sec	5
Annealing	*P. falciparum* Specific T_a	30 sec	*P. falciparum* Specific T_a	5 sec	
Extension	68	1 min/kb	72	15 sec/kb	
Denaturation	94/95	30 sec	98	1 sec	25
Annealing	Full Length Primer T_a	30 sec	Full Length Primer T_a	5 sec	
Extension	68	1 min/kb	72	15 sec/kb	
Final Extension	68	5-10 min	72	1 min	1

Table A 2: PCR Primers for *P. falciparum* Specific Region Amplification

Protein	Region	Vector	Primer Name	Primer Sequence (5' - 3')*	PCR Product (bp)
PfIAP	BIR Domain	pGEX-4T-2	IAP1 Fwd	*TTT*GGATCCATGAGCGAATCAAGTGCCT	471
			IAP1 Rev	*TTT*CTCGAG**TTA**TTCTTTAAGTAATTCTTCTTTTGT	
		pARL2-GFP	BIR Fwd	*TTT*CTCGAGATGAGCGAATCAAGTGCCTTGAGT	468
			BIR Rev	*TTT*CCTAGGTTCTTTAAGTAATTCTTCTTTTGTT	
	Middle Region	pGEX-4T-2	IAP2 Fwd	*TTT*GGATCCTTTAGCATTTTTGATCATTATGT	561
			IAP2 Rev	*TTT*CTCGAG**TTA**TTTCCTTTGTTCTAATAAGATGT	
	Terminal Region	pGEX-4T-2	IAP3 Fwd	*AAA*GGATCCCAACATCAAAATCAAAAGGATAC	984
			IAP3 Rev	*TTT*CTCGAG**TCA**TACATATTTTCTTTCAAATTCCA	
	5' UTR	pL7-DEL	DEL_L Fwd	*GTCTAA*GGCGCGCCATATAATTCAACATATCTGTC	376
			DEL_L Rev	*GTCGAC*ACTAGTCCTTCATTTGAATATTAAAAAAGG	
	3' UTR	pL7-DEL	DEL_R Fwd	*GACGTC*GAATTCAATAAAAAAAATAAAAGCAATCA	396
			DEL_R Fwd	*GACATA*CCATGGGAAAATTTGCTACATATTTTCT	
PfDOC2	Binding Region	pET-15b	DOC2 Fwd	*TCTGTTGCG*CATATGCAATTTCAAAAGGAAAAAATTCGT	1 011
			DOC2 Rev	*TTT*GGATCC**TTA**CTTTGAATAAATTTGATTCAAATCT	
PfRhopH3	Binding Region	pET-15b	RhopH3 Fwd	*TCTGTTGCG*CATATGTATGAAGAAACTGTAGATGAAGA	510
			RhopH3 Rev	*TTT*GGATCC**TTA**CAATTCATTTTCAGAAGTAAAGGT	

* Bold = STOP codon; bold & underlined = intentional mutation to PCR product; italics = overlapping region; italics and underlined = extra nucleotides for restriction enzyme efficiency; underlined = restriction enzyme site

Table A 3: PCR Primers for Vector Specific Region Amplification

Vector	Primer Name	Primer Sequence (5' - 3')	PCR Product (bp)
pARL2-GFP	pARL2 Fwd	CCGTTAATAATAAATACACGCAG	165
	pARL2 Rev	CCATCTAATTCAACAAGAATTGGGACAAC	
pET-15b	T7 Promoter Fwd	TAATACGACTCACTATAGGG	224
	T7 Terminator Rev	GCTAGTTATTGCTCAGCGGT	
T7 Phage Arms	T7SelectUP2 Fwd	GCTAACTTCCAAGCGGACCA	237
	T7SelectDOWN Rev	GCTAGTTATTGCTCAGCGGT	

Table A 4: CRISPR-Cas9 Targeting Sequences

Vector	Sequence Name	Targeting Sequence (5' - 3')*
pL7-DEL	DEL (+)	*TAAGTATATAATATTG*CCATAAGATT ATTTCTCCAT*GTTTTAGAGCTAGAA*
	DEL (-)	*TTCTAGCTCTAAAAC*ATGGAGAAATA ATCTTATGG*CAATATTATATACTTA*

* Italics = overlapping region; italics & underlined = addition 'G' base

APPENDIX B: Vector Maps
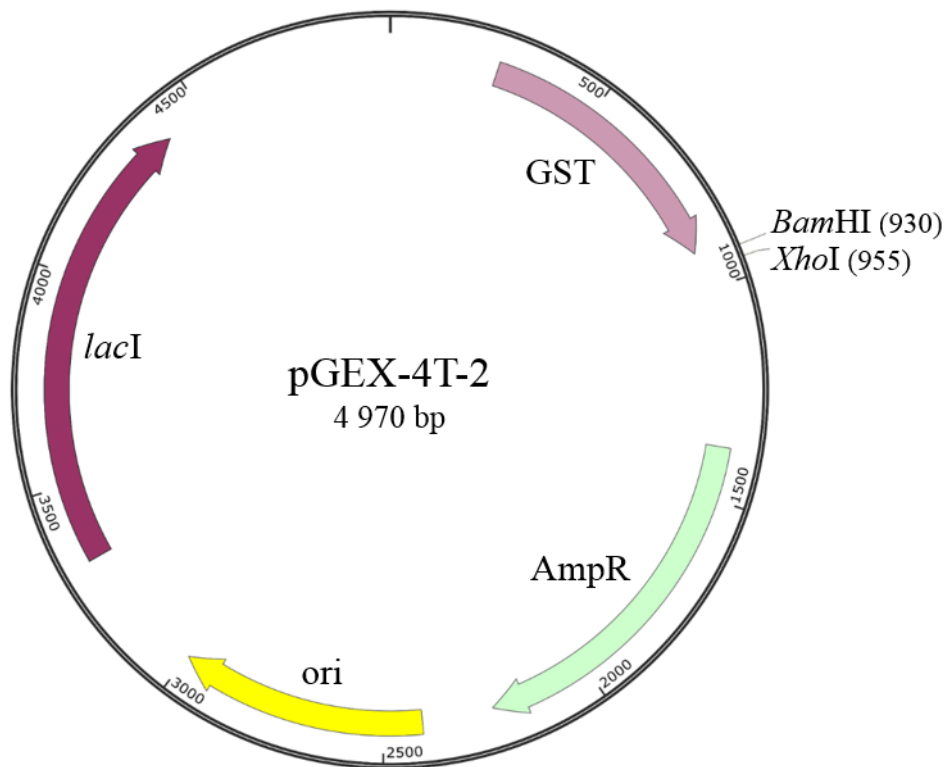
Figure B 1: pGEX-4T-2 Vector Map

A vector map of the pGEX-4T-2 plasmid with various restriction enzymes digestion sites marked. AmpR = confers resistance to ampicillin and related antibiotics; GST = glutathione S-transferase N-terminal tag; *lac*I = *lac* repressor that binds to the *lac* operator to inhibit transcription, but can be relieved; ori = origin of replication.
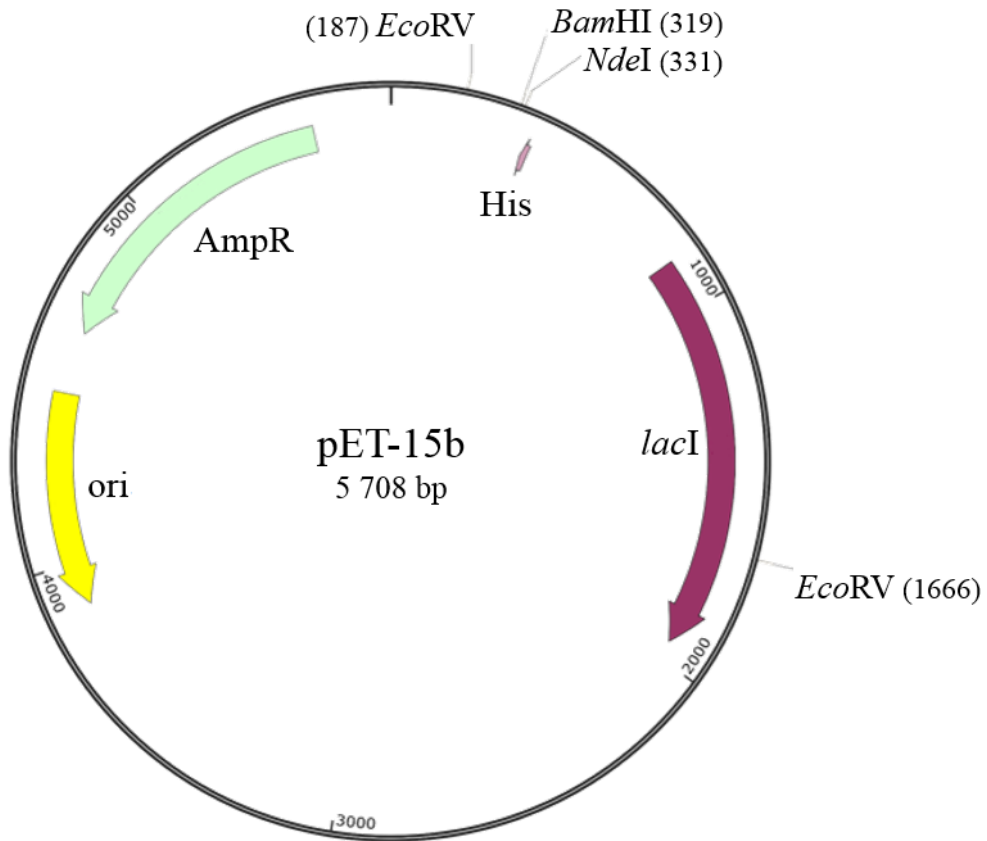
Figure B 2: pET-15b Vector Map

A vector map of the pET-15b plasmid with various restriction enzymes digestion sites marked. AmpR = confers resistance to ampicillin and related antibiotics; His = 6× His amino acid N-terminal tag; *lac*I = *lac* repressor that binds to the *lac* operator to inhibit transcription, but can be relieved; ori = origin of replication.
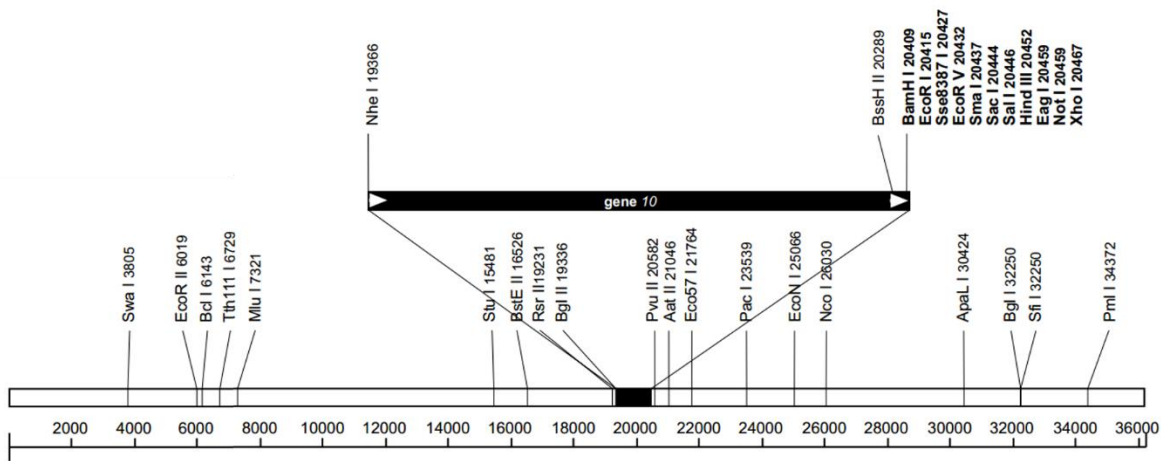
Figure B 3: T7Select®10-3b (Novagen) Bacteriophage Vector Map
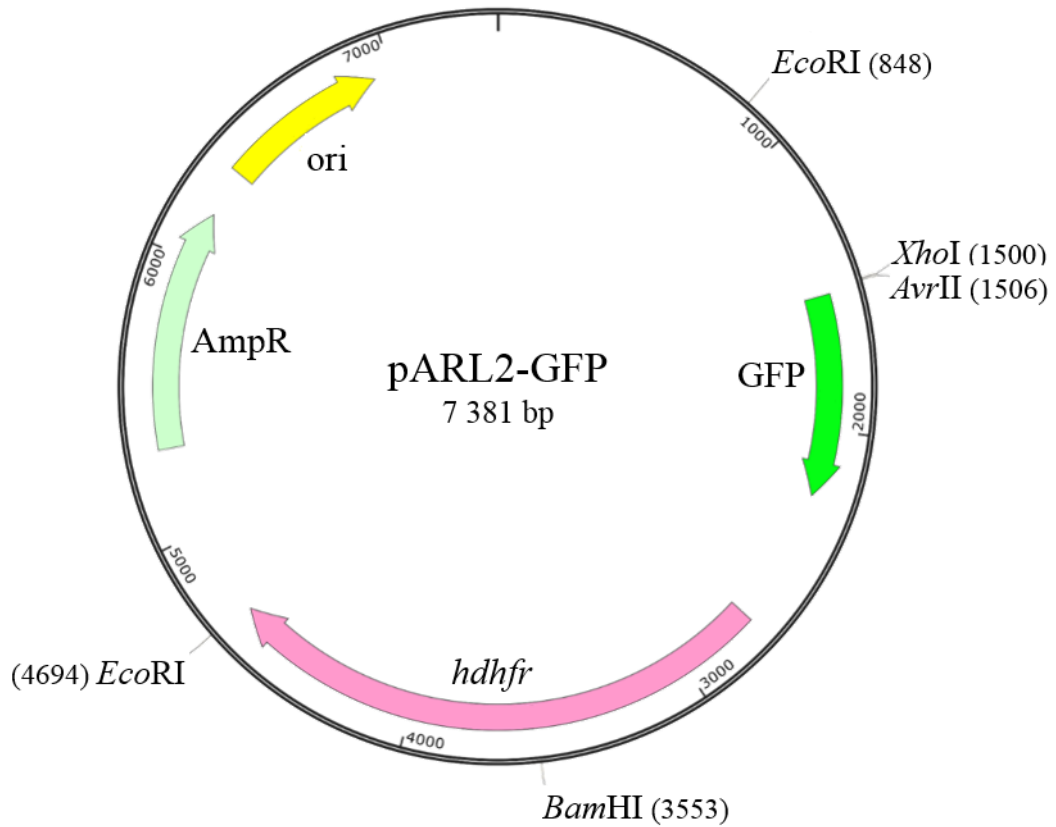
Figure B 4: pARL2-GFP Vector Map

A vector map of the pARL2-GFP plasmid with various restriction enzymes digestion sites marked. AmpR = confers resistance to ampicillin and related antibiotics; GFP = green fluorescent N-terminal tag; *hdhfr* = human dihydrofolate reductase cassette which confers resistance to the WR99210 compound; ori = origin of replication.
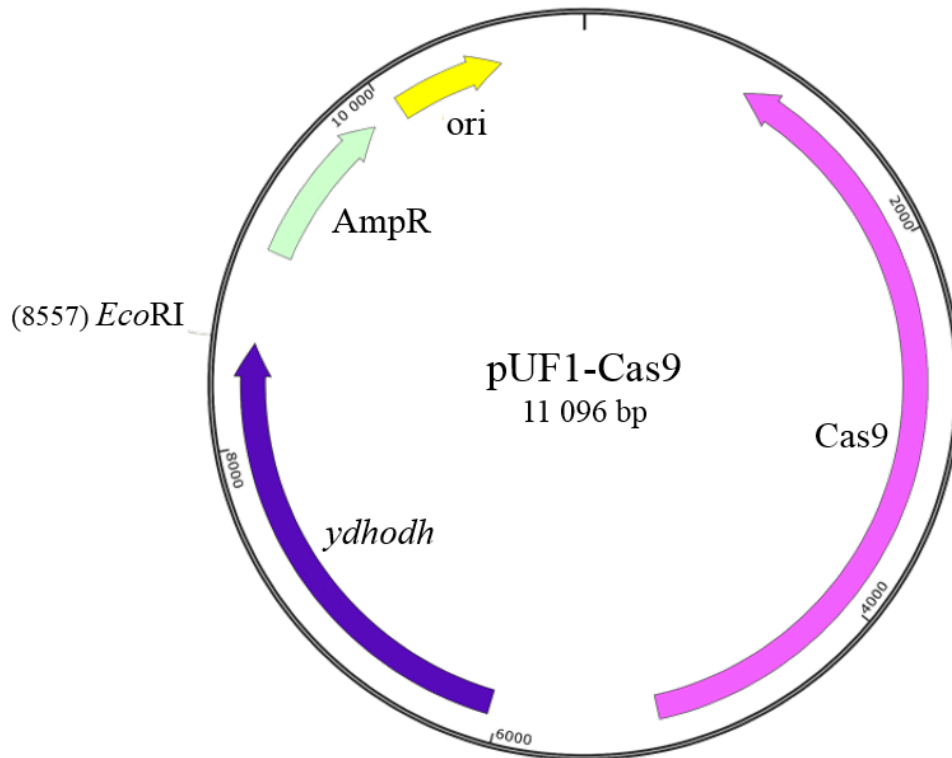
Figure B 5: pUF1-Cas9 Vector Map

A vector map of the pUF1-Cas9 plasmid with the *Eco*RI restriction enzyme digestion site marked. AmpR = confers resistance to ampicillin and related antibiotics; Cas9 = engineered *S. pyogenes* Cas9 endonuclease under the control of plasmodial regulatory elements; ori = origin of replication; *ydhodh* = drug-selectable marker which gives resistance to a *P. falciparum* dihydroorotate dehydrogenase (PfDHODH) inhibitor, DSM.
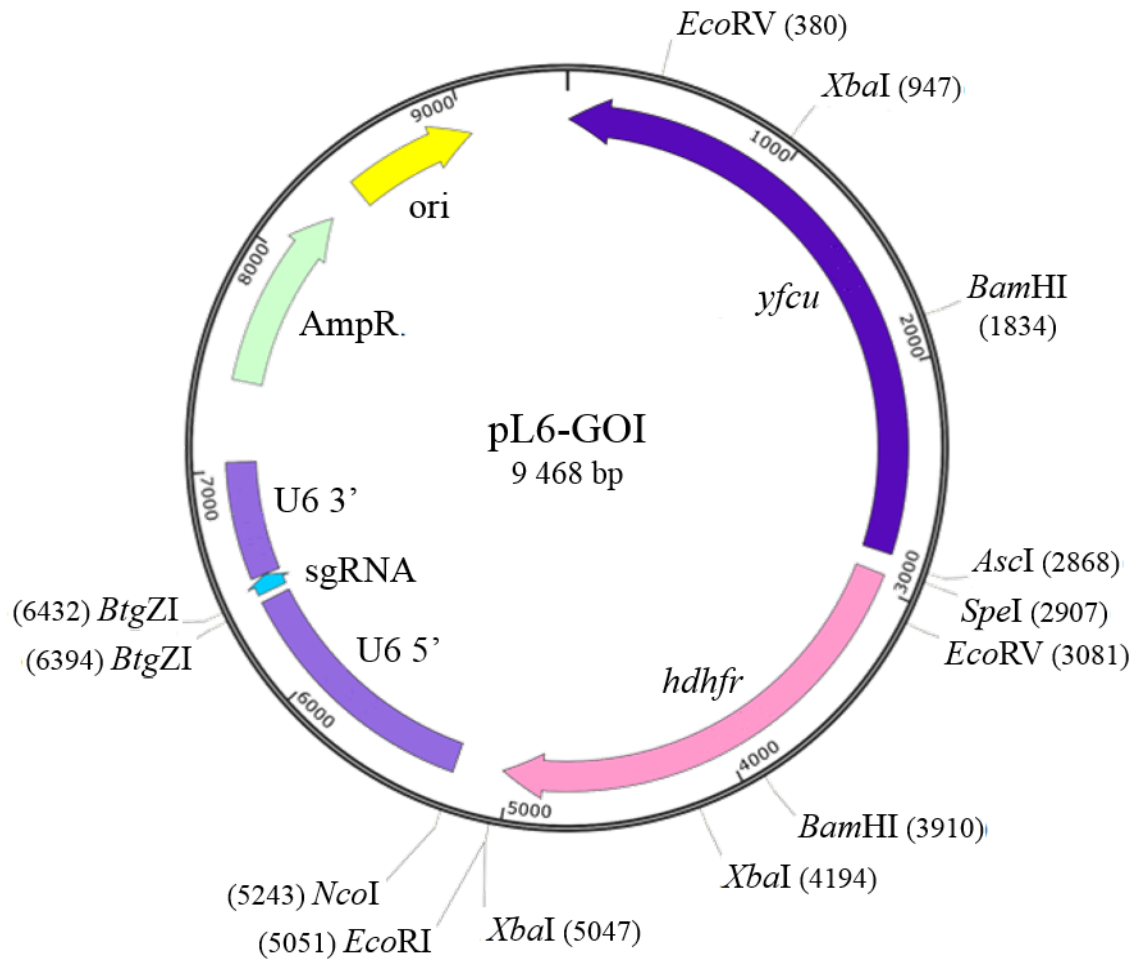
Figure B 6: pL6-GOI Vector Map

A vector map of the pL6-GOI plasmid with various restriction enzymes digestion sites marked. AmpR = confers resistance to ampicillin and related antibiotics; *hdhfr* = human dihydrofolate reductase cassette which confers resistance to the WR99210 compound; ori = origin of replication; U6 = *P. falciparum* small nuclear RNA regulatory elements that allow for the transcription of the sgRNA; sgRNA = single guide RNA that is made of the trans-activating crRNA (tracrRNA) that binds to the Cas9 endonuclease and the CRISPR RNA (crRNA) which contains the 20 bp guide RNA; *yfcu* = the suicide-gene cassette, yeast cytosine deaminase/uracil phosphoribosyl transferase, results in the conversion of harmless 5-fluorocytosine into 5-fluorouracil which kills the parasite.

APPENDIX C: Sequencing Results
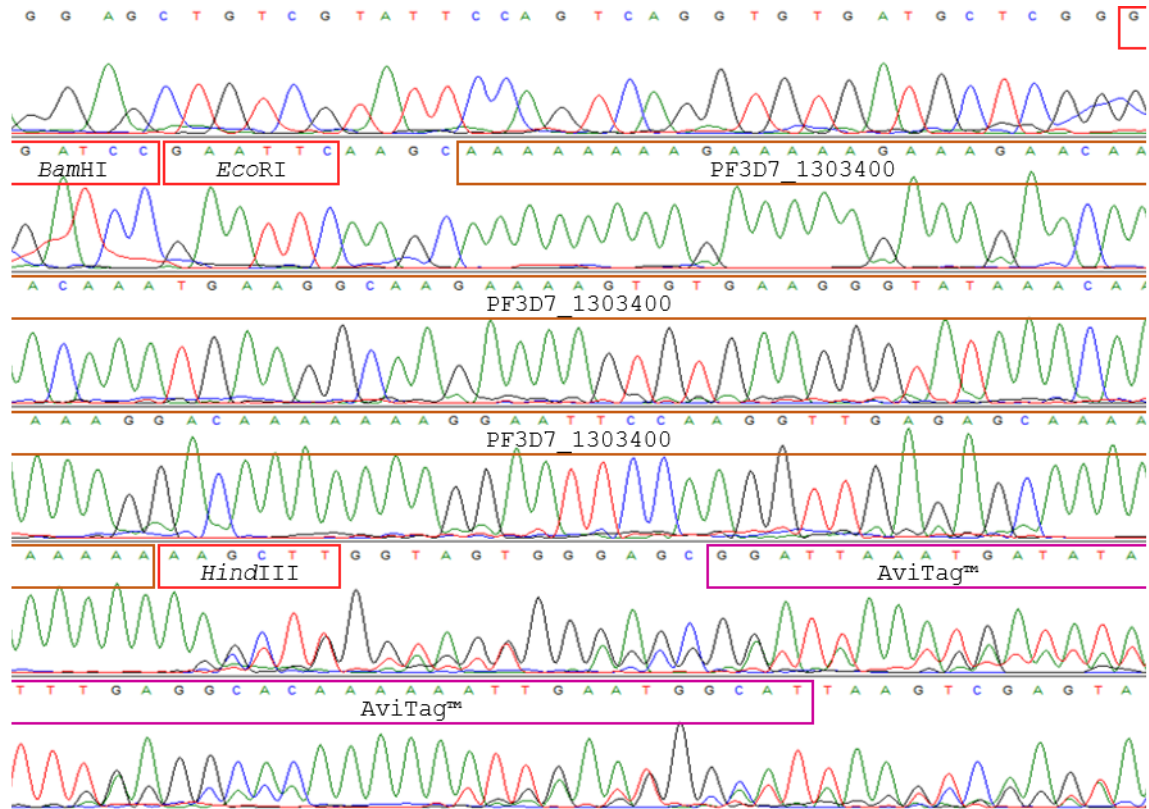
Figure C 1: Identification of a PfIAP BIR Domain Binding Partner
The sequencing chromatogram of the identified PF3D7_1303400 binding partner. Orange = the *P. falciparum* peptide insert; purple = the AviTag™ biotin cassette; red = restriction enzyme sites used during the creation of the phage display library.

APPENDIX D: PfIAP Binding Partners

Table D 1: PfIAP Binding Partner Details

	60S ribosomal protein L24, putative	Chromodomain-helicase-DNA-binding protein 1 homolog, putative (CHD1)	Double C2-like domain-containing protein (DOC2)
PlasmoDB Name	60S ribosomal protein L24, putative	Chromodomain-helicase-DNA-binding protein 1 homolog, putative (CHD1)	Double C2-like domain-containing protein (DOC2)
Gene I.D.	PF3D7_1309100	PF3D7_1023900	PF3D7_1243900
Intron/Exon	2 introns/3 exons	0 introns/1 exon	5 introns/6 exons
Gene Size	489 bp	9 987 bp	5 541 bp
Protein Size	162 aa	3 328 aa	1 846 aa
Binding Region (aa positions)	113 – 162	2 891 – 2 955	1 606 – 1 654
Binding Sequence	AKFGTKEKEDKKKTK DDKKKNLVHFQQKKD FTKSKMLNMAKSKMH KMMKK	DEESGECDKEEETGKR KRRRNSRYYKSSNKYI KDNEMLDITSVGAND EYKYEDGDNLERKKR SSK	QKMREEEKIREEEQM KREKKMRREEKKKRV QEPIKIDEIQVYDIIKNE K

	High molecular weight rhoptry protein 3 (RhopH3)	LisH domain-containing protein, putative (LisH)	Mature parasite-infected erythrocyte surface antigen (MESA)
PlasmoDB Name	High molecular weight rhoptry protein 3 (RhopH3)	LisH domain-containing protein, putative (LisH)	Mature parasite-infected erythrocyte surface antigen (MESA)
Gene I.D.	PF3D7_0905400	PF3D7_1303400	PF3D7_0500800
Intron/Exon	6 introns/7 exons	0 introns/1 exon	1 intron/2 exons
Gene Size	2 694 bp	3 069 bp	4 305 bp
Protein Size	897 aa	1 022 aa	1 434 aa
Binding Region (aa positions)	760 – 860	505 – 538	814 – 876
Binding Sequence	AIAEKDILEEKTKDQD LEIELYKYMGPLKEQS KSTSAASTSDEISGSEG PSTESTSTGNQGEDKT TDNTYKEMEELEEAE GTSNLKKGLEFYKSSL KLDQL	KKKKKKEQTNEGKKS VKGINKKDKKRNSKV ESKK	IEKQVEEGIKENDTES KDKVIGQEVIKGDVNE EGPENKDKVTKQEKV KEVKKEVKKKVKKRV K

Table D 2: The expanded PfIAP protein network combining biopanning data from this study and available yeast two-hybrid data (LaCount *et al.* 2005).

Numbers = PlasmoDB gene identifiers. Classification = Consensus of the Gene Ontology biological processes available on PlasmoDB (Ashburner *et al.* 2000; Aurrecoechea *et al.* 2009).

| PlasmoDB I.D.	PlasmoDB Name	Classification
GST-BIR		
PF3D7_0905400	**High Molecular Weight Rhoptry Protein 3 (RhopH3)**	**Erythrocyte Invasion**
PF3D7_0930300	Merozoite Surface Protein 1 (MSP1)	Erythrocyte Invasion
PF3D7_0108000	Proteasome Subunit Beta Type-3, Putative	Protein Metabolism
PF3D7_0625200	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function
PF3D7_0813400	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1227600	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1243900	**Double C2-Like Domain-Containing Protein (DOC2)**	**Erythrocyte Invasion**
PF3D7_1303400	**LisH Domain-Containing Protein, Putative (LisH)**	**Unknown Function**
PF3D7_1309100	**60S Ribosomal Protein L24, Putative**	**Protein Synthesis**
PF3D7_1203700	Nucleosome Assembly Protein (NAPL)	DNA/RNA
PF3D7_0936800	*Plasmodium* Exported Protein (PHISTc), Unknown Function	Exported Protein
PF3D7_1126700	Autophagy-Related Protein 23, Putative (ATG23)	RCD
PF3D7_1236100	Clustered-Asparagine-Rich Protein (CARP)	Unknown Function

PlasmoDB I.D.	PlasmoDB Name	Classification
GST-Terminal		
PF3D7_0500800	**Mature Parasite-Infected Erythrocyte Surface Antigen (MESA)**	**Pathogenesis**
PF3D7_0303300	DNA-Directed RNA Polymerases I, II, And III Subunit RPABC2, Putative (RPB6)	DNA/RNA
PF3D7_0803100	U3 Small Nucleolar RNA-Associated Protein 14, Putative (UTP14)	
PF3D7_1014600	Transcriptional Coactivator ADA2 (ADA2)	
PF3D7_1216900	DNA-Binding Chaperone, Putative	
PF3D7_1321700	Splicing Factor 1 (SF1)	
PF3D7_1428100	WW Domain-Binding Protein 11, Putative	
PF3D7_0620400	Merozoite Surface Protein 10 (MSP10)	Erythrocyte Invasion
PF3D7_0929400	High Molecular Weight Rhoptry Protein 2 (RhopH2)	
PF3D7_0930300	Merozoite Surface Protein 1 (MSP1)	
PF3D7_1228600	Merozoite Surface Protein 9 (MSP9)	
PF3D7_1401200	*Plasmodium* Exported Protein, Unknown Function	Exported Protein
PF3D7_0935600	Gametocytogenesis-Implicated Protein (GIG)	Gametocytogenesis
PF3D7_0702900	Centrin, Putative	Metabolism
PF3D7_1434500	Dynein-Related AAA-Type ATPase, Putative	
PF3D7_1462800	Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH)	
PF3D7_0207600	Serine Repeat Antigen 5 (SERA5)	Pathogenesis
PF3D7_0302500	Cytoadherence Linked Asexual Protein 3.1 (CLAG3.1)	
PF3D7_1221000	Histone-Lysine N-Methyltransferase, H3 Lysine-4 Specific (SET10)	
PF3D7_0302100	Serine/Threonine Protein Kinase (SRPK1)	Protein Metabolism
PF3D7_0501300	Skeleton-Binding Protein 1 (SBP1)	
PF3D7_1129200	26S Proteasome Regulatory Subunit RPN7, Putative (RPN7)	
PF3D7_1145400	Dynamin-Like Protein (DYN1)	
PF3D7_1355500	Serine/Threonine Protein Phosphatase 5 (PP5)	
PF3D7_0710600	60S Ribosomal Protein L34 (RPL34)	Protein Synthesis
PF3D7_0828500	Translation Initiation Factor eIF-2B Subunit Alpha, Putative	
PF3D7_0730900	EMP1-Trafficking Protein (PTP4)	Unknown Function
PF3D7_0813400	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_0825500	Protein KRI1, Putative (KRI1)	
PF3D7_0903600	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1325800	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1433400	Conserved *Plasmodium* Membrane Protein, Unknown Function	
PF3D7_1023900	**Chromodomain-Helicase-DNA-Binding Protein 1 Homolog, Putative (CHD1)**	**DNA/RNA**
PF3D7_0603800	Centrosomal Protein CEP76, Putative (CEP76)	DNA/RNA
PF3D7_0623100	Nuclear Polyadenylated RNA-Binding Protein NAB2, Putative (NAB2)	
PF3D7_0802100	Transcription Factor With AP2 Domain(s) (ApiAP2)	
PF3D7_1207100	Small Subunit rRNA Processing Factor, Putative	
PF3D7_1216900	DNA-Binding Chaperone, Putative	
PF3D7_1235300	CCR4-NOT Transcription Complex Subunit 4, Putative (NOT4)	
PF3D7_0930300	Merozoite Surface Protein 1 (MSP1)	Erythrocyte Invasion
PF3D7_1228600	Merozoite Surface Protein 9 (MSP9)	
PF3D7_1335100	Merozoite Surface Protein 7 (MSP7)	
PF3D7_0702100	*Plasmodium* Exported Protein (PHISTb), Unknown Function, Pseudogene	Exported Protein
PF3D7_0935500	*Plasmodium* Exported Protein, Unknown Function (GEXP22)	
PF3D7_1148800	*Plasmodium* Exported Protein (hyp11), Unknown Function	
PF3D7_0419900	Phosphatidylinositol 4-Kinase, Putative	Metabolism
PF3D7_1017400	Phosphomannomutase, Putative (PMM)	
PF3D7_1449400	Crossover Junction Endonuclease MUS81, Putative (MUS81)	
PF3D7_0207600	Serine Repeat Antigen 5 (SERA5)	Pathogenesis
PF3D7_1149000	Antigen 332, DBL-Like Protein (Pf332)	

Appendix D

PlasmoDB I.D.	PlasmoDB Name	Classification
PF3D7_0608500	Proteasome Subunit Alpha Type-2, Putative	
PF3D7_0728600	RING Zinc Finger Protein, Putative	
PF3D7_0803700	Tubulin Gamma Chain (g-tub)	Protein Metabolism
PF3D7_0810500	Protein Phosphatase PPM7, Putative (PPM7)	
PF3D7_0826100	E3 Ubiquitin-Protein Ligase, Putative	
PF3D7_1118500	Nucleolar Protein 56, Putative (NOP56)	
PF3D7_1302700	ATP-Dependent RNA Helicase DHR1, Putative	Protein Synthesis
PF3D7_1342000	40S Ribosomal Protein S6	
PF3D7_0202500	Early Transcribed Membrane Protein 2 (ETRAMP2)	
PF3D7_0317300	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_0407700	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_0410000	Erythrocyte Vesicle Protein 1 (EVP1)	
PF3D7_0418300	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_0510100	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_0532100	Early Transcribed Membrane Protein 5 (ETRAMP5)	
PF3D7_0617200	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_0619000	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_0721000	Conserved *Plasmodium* Membrane Protein, Unknown Function	
PF3D7_0725300	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_0910200	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1013400	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1014900	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function
PF3D7_1021700	Conserved *Plasmodium* Membrane Protein, Unknown Function	
PF3D7_1029900	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1112100	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1138700	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1140200	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1238500	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1338600	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1359600	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1406200	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1446500	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1451200	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1452400	Conserved *Plasmodium* Protein, Unknown Function	
PF3D7_1467600	Conserved *Plasmodium* Protein, Unknown Function	

Table D 3: The expanded PfIAP protein network with Gene Ontology classifications.
Numbers = PlasmoDB gene identifiers. Biological Process/Molecular Function/Location = Gene Ontology labels available on PlasmoDB (Ashburner *et al.* 2000; Aurrecoechea *et al.* 2009).

PlasmoDB I.D.	PlasmoDB Name	Biological Process	Molecular Function	Location
GST-BIR				
PF3D7_0905400	**High Molecular Weight Rhoptry Protein 3 (RhopH3)**	**Erythrocyte Invasion**	**Protein Binding**	**Organelle**
PF3D7_0108000	Proteasome Subunit Beta Type-3, Putative	Protein Degradation	Protein Activity	Cytoplasm Nucleus
PF3D7_0625200	Conserved *Plasmodium* Protein, Unknown Function	DNA/RNA	Nucleotide Binding	Organelle
PF3D7_0813400	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Nucleotide Binding	Nucleus
PF3D7_0930300	Merozoite Surface Protein 1 (MSP1)	Erythrocyte Invasion	Protein Binding	Membrane
PF3D7_1227600	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Unknown
PF3D7_1243900	**Double C2-Like Domain-Containing Protein (DOC2)**	**Erythrocyte Invasion**	**Protein Binding**	**Host Membrane**
PF3D7_1303400	**LisH Domain-Containing Protein, Putative (LisH)**	**Unknown Function**	**Protein Binding**	**Unknown**
PF3D7_1309100	**60S Ribosomal Protein L24, Putative**	**Protein Synthesis**	**Structure**	**Cytoplasm**
PF3D7_0936800	*Plasmodium* Exported Protein (PHISTc), Unknown Function	Unknown Function	Unknown Function	Exported Host Cytoplasm Host Membrane

PlasmoDB I.D.	PlasmoDB Name	Biological Process	Molecular Function	Location
PF3D7_1126700	Autophagy-Related Protein 23, Putative (ATG23)	RCD	Unknown Function	Unknown
PF3D7_1203700	Nucleosome Assembly Protein (NAPL)	DNA/RNA	Nucleotide Binding Protein Binding	Cytoplasm Nucleus
PF3D7_1236100	Clustered-Asparagine-Rich Protein (CARP)	Unknown Function	Nucleotide Binding Protein Binding	Unknown
GST-Terminal				
PF3D7_0500800	**Mature Parasite-Infected Erythrocyte Surface Antigen (MESA)**	**Pathogenesis**	**Protein Binding**	**Host Membrane**
PF3D7_0207600	Serine Repeat Antigen 5 (SERA5)	Pathogenesis	Protein Activity Protein Binding	Host Cytoplasm
PF3D7_0302100	Serine/Threonine Protein Kinase (SRPK1)	Protein Modification	Protein Activity	Cytoplasm Nucleus
PF3D7_0302500	Cytoadherence Linked Asexual Protein 3.1 (CLAG3.1)	Pathogenesis	Protein Binding	Host Membrane Organelle
PF3D7_0303300	DNA-Directed RNA Polymerases I, II, And III Subunit RPABC2, Putative (RPB6)	DNA/RNA	Nucleotide Binding	Nucleus
PF3D7_0501300	Skeleton-Binding Protein 1 (SBP1)	Protein Transport	Protein Binding	Membrane
PF3D7_0620400	Merozoite Surface Protein 10 (MSP10)	Erythrocyte Invasion	Protein Binding	Host Membrane
PF3D7_0702900	Centrin, Putative	Metabolism	Substrate Binding	Cytoplasm
PF3D7_0710600	60S Ribosomal Protein L34 (RPL34)	Protein Synthesis	Structure	Cytoplasm
PF3D7_0730900	EMP1-Trafficking Protein (PTP4)	Unknown Function	Unknown Function	Unknown
PF3D7_0803100	U3 Small Nucleolar RNA-Associated Protein 14, Putative (UTP14)	DNA/RNA	Unknown Function	Unknown
PF3D7_0813400	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Nucleotide Binding	Nucleus
PF3D7_0825500	Protein KRI1, Putative (KRI1)	Unknown Function	Substrate Binding	Unknown
PF3D7_0828500	Translation Initiation Factor eIF-2B Subunit Alpha, Putative	Protein Synthesis	Nucleotide Binding	Nucleus
PF3D7_0903600	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Protein Activity Protein Binding	Cytoplasm
PF3D7_0929400	High Molecular Weight Rhoptry Protein 2 (RhopH2)	Erythrocyte Invasion	Protein Binding	Host Membrane Organelle
PF3D7_0930300	Merozoite Surface Protein 1 (MSP1)	Erythrocyte Invasion	Protein Binding	Host Membrane
PF3D7_0935600	Gametocytogenesis-Implicated Protein (GIG)	Gametocytogenesis	Unknown Function	Unknown
PF3D7_1014600	Transcriptional Coactivator ADA2 (ADA2)	DNA/RNA	Nucleotide Binding	Membrane
PF3D7_1129200	26S Proteasome Regulatory Subunit RPN7, Putative (RPN7)	Protein Degradation	Protein Activity Protein Binding	Cytoplasm Nucleus
PF3D7_1145400	Dynamin-Like Protein (DYN1)	Protein Transport	Protein Activity	Cytoplasm Host Cytoplasm Host Membrane Membrane Organelle
PF3D7_1216900	DNA-Binding Chaperone, Putative	DNA/RNA	Nucleotide Binding	Unknown
PF3D7_1221000	Histone-Lysine N-Methyltransferase, H3 Lysine-4 Specific (SET10)	Pathogenesis	Protein Activity Protein Binding	Nucleus
PF3D7_1228600	Merozoite Surface Protein 9 (MSP9)	Erythrocyte Invasion	Unknown Function	Host Cytoplasm Membrane
PF3D7_1321700	Splicing Factor 1 (SF1)	DNA/RNA	Nucleotide Binding	Nucleus
PF3D7_1325800	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Unknown
PF3D7_1355500	Serine/Threonine Protein Phosphatase 5 (PP5)	Protein Modification	Protein Activity	Cytoplasm Nucleus
PF3D7_1401200	*Plasmodium* Exported Protein, Unknown Function	Unknown Function	Unknown Function	Exported
PF3D7_1428100	WW Domain-Binding Protein 11, Putative	DNA/RNA	Unknown Function	Cytoplasm Membrane
PF3D7_1433400	Conserved *Plasmodium* Membrane Protein, Unknown Function	Unknown Function	Protein Binding	Cytoplasm Membrane

PlasmoDB I.D.	PlasmoDB Name	Biological Process	Molecular Function	Location
PF3D7_1434500	Dynein-Related AAA-Type ATPase, Putative	Metabolism	Protein Activity	Membrane
PF3D7_1462800	Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH)	Metabolism	Protein Binding	Cytoplasm Membrane
PF3D7_1023900	**Chromodomain-Helicase-DNA-Binding Protein 1 Homolog, Putative (CHD1)**	**DNA/RNA**	**Nucleotide Binding**	**Nucleus**
PF3D7_0202500	Early Transcribed Membrane Protein 2 (ETRAMP2)	Unknown Function	Unknown Function	Host Membrane
PF3D7_0207600	Serine Repeat Antigen 5 (SERA5)	Pathogenesis	Protein Activity Protein Binding	Host Cytoplasm
PF3D7_0317300	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Protein Activity Protein Binding	Cytoplasm
PF3D7_0407700	Conserved *Plasmodium* Protein, Unknown Function	Transcription	Unknown Function	Nucleus
PF3D7_0410000	Erythrocyte Vesicle Protein 1 (EVP1)	Unknown Function	Unknown Function	Host Cytoplasm
PF3D7_0418300	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Unknown
PF3D7_0419900	Phosphatidylinositol 4-Kinase, Putative	Metabolism	Protein Activity	Unknown
PF3D7_0510100	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	RNA Binding	Unknown
PF3D7_0532100	Early Transcribed Membrane Protein 5 (ETRAMP5)	Unknown Function	Protein Binding	Host Membrane
PF3D7_0603800	Centrosomal Protein CEP76, Putative (CEP76)	DNA/RNA	Protein Binding	Organelle
PF3D7_0608500	Proteasome Subunit Alpha Type-2, Putative	Protein Degradation	Endopeptidase Activity	Cytoplasm Nucleus
PF3D7_0617200	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Protein Activity Protein Binding	Cytoplasm
PF3D7_0619000	Conserved *Plasmodium* Protein, Unknown Function	Metabolism	Protein Binding	Cytoplasm
PF3D7_0623100	Nuclear Polyadenylated RNA-Binding Protein NAB2, Putative (NAB2)	DNA/RNA	Nucleotide Binding	Unknown
PF3D7_0702100	*Plasmodium* Exported Protein (PHISTb), Unknown Function, Pseudogene	Unknown Function	Metal Ion Binding	Exported
PF3D7_0721000	Conserved *Plasmodium* Membrane Protein, Unknown Function	Unknown Function	Unknown Function	Unknown
PF3D7_0725300	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Cytoplasm
PF3D7_0728600	RING Zinc Finger Protein, Putative	Protein Degradation	Protein Binding	Unknown
PF3D7_0802100	Transcription Factor With AP2 Domain(s) (ApiAP2)	DNA/RNA	Nucleotide Binding	Unknown
PF3D7_0803700	Tubulin Gamma Chain (g-tub)	Protein Transport	Protein Binding	Organelle
PF3D7_0810500	Protein Phosphatase PPM7, Putative (PPM7)	Protein Modification	Protein Activity	Cytoplasm Nucleus
PF3D7_0826100	E3 Ubiquitin-Protein Ligase, Putative	Protein Degradation	Protein Activity	Unknown
PF3D7_0910200	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Nucleus
PF3D7_0930300	Merozoite Surface Protein 1 (MSP1)	Erythrocyte Invasion	Protein Binding	Host Membrane
PF3D7_0935500	*Plasmodium* Exported Protein, Unknown Function (GEXP22)	Unknown Function	Substrate Binding	Unknown
PF3D7_1013400	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Substrate Binding	Cytoplasm
PF3D7_1014900	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Protein Activity Protein Binding	Cytoplasm
PF3D7_1017400	Phosphomannomutase, Putative (PMM)	Metabolism	Protein Activity	Cytoplasm
PF3D7_1021700	Conserved *Plasmodium* Membrane Protein, Unknown Function	Metabolism	Protein Activity	Organelle
PF3D7_1029900	Conserved *Plasmodium* Protein, Unknown Function	DNA/RNA	Nucleotide Binding	Nucleus
PF3D7_1112100	Conserved *Plasmodium* Protein, Unknown Function	Protein Modification	Protein Activity	Cytoplasm Membrane Nucleus
PF3D7_1118500	Nucleolar Protein 56, Putative (NOP56)	Protein Synthesis	Unknown Function	Nucleus
PF3D7_1138700	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Protein Activity Protein Binding	Cytoplasm

PlasmoDB I.D.	PlasmoDB Name	Biological Process	Molecular Function	Location
PF3D7_1140200	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Unknown
PF3D7_1148800	*Plasmodium* Exported Protein (hyp11), Unknown Function	Unknown Function	Unknown Function	Unknown
PF3D7_1149000	Antigen 332, DBL-Like Protein (Pf332)	Pathogenesis	Protein Binding	Exported Host Membrane
PF3D7_1207100	Small Subunit rRNA Processing Factor, Putative	DNA/RNA	Nucleotide Binding	Cytoplasm Nucleus
PF3D7_1216900	DNA-Binding Chaperone, Putative	DNA/RNA	Nucleotide Binding	Unknown
PF3D7_1228600	Merozoite Surface Protein 9 (MSP9)	Erythrocyte Invasion	Unknown Function	Host Cytoplasm Membrane
PF3D7_1235300	CCR4-NOT Transcription Complex Subunit 4, Putative (NOT4)	DNA/RNA	Nucleotide Binding Protein Binding	Unknown
PF3D7_1238500	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Protein Activity Protein Binding	Cytoplasm
PF3D7_1302700	ATP-Dependent RNA Helicase DHR1, Putative	Protein Synthesis	Nucleotide Binding	Nucleus
PF3D7_1335100	Merozoite Surface Protein 7 (MSP7)	Erythrocyte Invasion	Protein Binding	Host Membrane
PF3D7_1338600	Conserved *Plasmodium* Protein, Unknown Function	Protein Synthesis	Unknown Function	Membrane
PF3D7_1342000	40S Ribosomal Protein S6	Protein Synthesis	Structure	Cytoplasm
PF3D7_1359600	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Cytoplasm
PF3D7_1406200	Conserved *Plasmodium* Protein, Unknown Function	DNA/RNA	Protein Activity	Nucleus
PF3D7_1446500	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Cytoplasm
PF3D7_1449400	Crossover Junction Endonuclease MUS81, Putative (MUS81)	Metabolism	Nucleotide Binding	Organelle
PF3D7_1451200	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Unknown
PF3D7_1452400	Conserved *Plasmodium* Protein, Unknown Function	Unknown Function	Unknown Function	Cytoplasm Nucleus
PF3D7_1467600	Conserved *Plasmodium* Protein, Unknown Function	Metabolism	Unknown Function	Cytoplasm

Appendix D

APPENDIX E: Chemical and Equipment Suppliers

Chemical, Equipment or Kit	Manufacture or Supplier
Acid Citrate Dextrose Tubes	BD Vacutainer, UK
Acrodisc® PF 0.8/0.2 μm Filter Unit	Pall Life Sciences, USA
Acrylamide	Promega, USA
Agar	Merck, Germany
Agarose	Sigma-Aldrich, USA
Albumax II	Life Technologies, USA
Ammonium Chloride	Saarchem, RSA
Ammonium Persulfate	Promega, USA
Ampicillin	Roche, Germany
Anti-GST HRP Conjugated Primary Antibody	Amersham Biosciences, UK
Anti-His HRP Conjugated Blocking Solution	Qiagen, Germany
Anti-His HRP Conjugated Primary Antibody	Qiagen, Germany
Avanti® J-E Centrifuge	Beckman Coulter, USA
Axiostar Plus Transmitted Light Microscope	Zeiss, Germany
Biotin Polyclonal Antibody-HRP Conjugate	Thermo Fisher Scientific, USA
Bis-Acrylamide	Sigma-Aldrich, USA
Bovine Serum Albumen (BSA)	Pierce, USA
Bromophenol Blue	Merck, Germany
BtgZI	New England Biolabs, USA
BX41 Microscope	Olympus, Japan
Calcium Chloride	Merck, Germany
Centrifuge 5415 R	Eppendorf, Germany
Centrifuge 5702 R	Eppendorf, Germany
Chloramphenicol	Roche, Germany
Chloroform	Merck, Germany
Coomassie Brilliant Blue R-250	BDH, UK
CPE 50 Circulator	Labcon, RSA
Cryotubes	Nunc, Denmark
Culture Flasks	Thermo Fisher Scientific, USA
DAPI	Sigma-Aldrich, USA
Dipotassium Phosphate	Merck, Germany
Disodium Phosphate	Saarchem, RSA
Dithiothreitol (DTT)	Invitrogen, USA
DMSO	BDH, UK
DNA Polymerase I	Fermentas, USA
DNaseI	Fermentas, USA
DP72 Camera	Olympus, Japan
DreamTaq Green PCR Master Mix	Thermo Fisher Scientific, USA
Dynabeads® mRNA Purification Kit	Invitrogen, USA
EDTA	Merck, Germany
EGTA	BDH, UK
Eppendorf Tubes	Eppendorf, Germany
Ethanol	Merck, Germany
Ethidium Bromide	Sigma-Aldrich, USA
Falcon™ Round Bottom Tubes	Becton Dickinson, USA
FastAP™ Thermosensitive Alkaline Phosphatase	Thermo Fisher Scientific, USA
FastDigest® Restriction Endonuclease	Thermo Fisher Scientific, USA
Filter Tips	QSP, USA

Chemical, Equipment or Kit	Manufacture or Supplier
Forma Incubated/Refrigerated Floor Orbital Shaker	Thermo Fisher Scientific, USA
Gas Mixture	Afrox, RSA
Gel Doc System	Syngene, UK
Gene Pulser Xcell™ Electroporation System	Bio-Rad Laboratories, USA
Gene Pulser® Electroporation Cuvette	Bio-Rad Laboratories, USA
GenElute™ Plasmid Miniprep Kit	Sigma-Aldrich, USA
Gentamycin	Sigma-Aldrich, USA
Gibson Assembly® Master Mix	New England Biolabs, USA
Glacial Acetic Acid	Merck, Germany
Glucose	Merck, Germany
Glutathione, Reduced	Sigma-Aldrich, USA
Glycerol	Merck, Germany
Glycine	Sigma-Aldrich, USA
HEPES	Merck, Germany
High Fidelity PCR Enzyme Mix	Thermo Fisher Scientific, USA
Hoechst 33258 Pentahydrate	Invitrogen, USA
Hybond-C Extra Nitrocellulose Membrane	Amersham Biosciences, UK
Hypoxanthine	Sigma-Aldrich, USA
Imidazole	Sigma-Aldrich, USA
Intelli-Mixer	ELMI, Latvia
Isopropanol	Merck, Germany
Laminar Flow Hood	Labotec, RSA
Lithium Chloride	Sigma-Aldrich, USA
MagneGST™ Protein Purification System	Promega, USA
MagneHis™ Protein Purification System	Promega, USA
Magnesium Chloride	Merck, Germany
Magnesium Sulphate	Merck, Germany
MassRuler™ DNA Ladder Mix	Fermentas, USA
Mastercycler Gradient Thermocycler	Eppendorf, Germany
Methanol	Merck, Germany
Methylated dNTP Mix	Fermentas, USA
Microscope Slides	Thermo Fisher Scientific, USA
Mighty Small II SE250 Gel Cassette	Hoefer Scientific Instruments, USA
Monopotassium Phosphate	Merck, Germany
Monosodium Phosphate	Merck, Germany
NanoDrop® 1000 Spectrophotometer	Thermo Fisher Scientific, USA
Nuclease Free Water	Fermentas, USA
Nucleobond® Xtra Maxi Plus Plasmid Kit	Macherey-Nagel, Germany
Nucleospin® Gel and PCR Clean-Up Kit	Macherey-Nagel, Germany
Nucleospin® Plasmid Kit	Macherey-Nagel, Germany
Orbital Shaker	Labotec, RSA
OrientExpress™ cDNA Synthesis Kit	Novagen, USA
Overnight Express™ Instant TB Medium	Novagen, USA
Petri Dishes	Costar, USA
pH Meter	Beckman Coulter, USA
Phenol	Merck, Germany
Phusion Flash PCR Master Mix	Thermo Fisher Scientific, USA
Ponceau S	Sigma-Aldrich, USA
Potassium Acetate	Merck, Germany
Potassium Chloride	Merck, Germany
PR250 Mini Orbital Shaker	Hoefer Scientific Instruments, USA

Chemical, Equipment or Kit	Manufacture or Supplier
Protease Inhibitor Cocktail Set III	Calbiochem, USA
QIAquick® PCR Purification Kit	Qiagen, Germany
Rapid DNA Ligation Kit	Roche, Germany
Rapi-Diff Stain Kit	Diagnostic Media Products, RSA
RNase A	Thermo Fisher Scientific, USA
RNase H	Fermentas, USA
RNaseZap™	Ambion, USA
Rosetta™ 2 (DE3) Competent Cells	Novagen, USA
RPMI Culture Medium	GibcoBRL, USA
Saponin	USB, USA
Slide-A-Lyzer® Mini Dialysis Unit	Pierce, USA
Sodium Acetate	Merck, Germany
Sodium Bicarbonate	PAL Chemicals, UK
Sodium Chloride	Merck, Germany
Sodium Dodecyl Sulfate (SDS)	Merck, Germany
Sodium Hydroxide	Merck, Germany
Sonopuls HD 3100 Ultrasonic Homogenizer	Bandelin Electronic, Germany
Sorbitol	Sigma-Aldrich, USA
Spectra™ Multicolor Broad Range Protein Ladder	Fermentas, USA
Streptavidin Magnetic Particles	Roche, Germany
Sucrose	Merck, Germany
SuperScript™ III Reverse Transcriptase	Invitrogen, USA
SuperSignal® West Pico Chemiluminescent Substrate	Thermo Fisher Scientific, USA
T7Select®10-3 Cloning Kit	Novagen, USA
TEMED	Promega, USA
Thermo Biomate 5 Spectrophotometer	Thermo Fisher Scientific, USA
TRI Reagent®	Sigma-Aldrich, USA
Tris	Sigma-Aldrich, USA
Triton-X	BDH, UK
Tryptone [Pancreatic Digest of Casein]	Merck, Germany
Tween20	Calbiochem, USA
U-25ND25 Neutral Density Filter	Olympus, Japan
U-MWB2 Filter	Olympus, Japan
U-MWU2 Filter	Olympus, Japan
Vacucap® 90 PF 0.8/0.2 µm Filter Unit	Pall Life Sciences, USA
Water Bath	Lauda, Germany
XL10-Gold Ultracompetent Cells	Stratagene, USA
Yeast Extract	Oxoid, UK
β-Mercaptoethanol	Merck, Germany

APPENDIX F: Ethics Clearance Certificate

R14/49 Mr Dale Liebenberg

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M1703102

NAME:
(Principal Investigator)   Mr Dale Liebenberg

DEPARTMENT: Molecular Medicine and Haematology
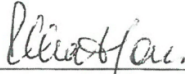University of the Witwatersrand

PROJECT TITLE: The Role of Baculovirus Inhabitor of Apopotosis
Protein Repeat-Containing Plasmodium Falciparum
Protein and its Binding Partners in Programmed
Cell Death

DATE CONSIDERED: Adhoc

DECISION: Approved unconditionally

CONDITIONS: Sub-Study (M130569)

SUPERVISOR: Prof Theresa Coetzer

APPROVED BY: _____
Professor P. Cleaton-Jones, Chairperson, HREC (Medical)

DATE OF APPROVAL: 12/04/2017

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary in Room 10004,10th floor, Senate House/2nd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand. I/We fully understand the the conditions under which I am/we are authorised to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. in this case, the study was initially review in March and will therefore be due in the month of March each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

_____
Principal Investigator Signature   Date   13/04/2017

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

APPENDIX G: Plagiarism Report

Supervisor: Prof TL Coetzer   Signature

Candidate: Dale Liebenberg   Signature

DLiebenbergPhDThesis.pdf

ORIGINALITY REPORT

15%	**12**%	**4**%	**9**%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	**Submitted to University of Witwatersrand** Student Paper	**5**%
2	**phenoplasm.org** Internet Source	**1**%
3	**Submitted to University of the Witwatersrand** Student Paper	**1**%
4	Alam, Mahmood M., Lev Solyakov, Andrew R. Bottrill, Christian Flueck, Faiza A. Siddiqui, Shailja Singh, Sharad Mistry, Maria Viskaduraki, Kate Lee, Christine S. Hopp, Chetan E. Chitnis, Christian Doerig, Robert W. Moon, Judith L. Green, Anthony A. Holder, David A. Baker, and Andrew B. Tobin. "Phosphoproteomics reveals malaria parasite Protein Kinase G as a signalling hub regulating egress and invasion", Nature Communications, 2015. Publication	**<1**%
5	**oro.open.ac.uk** Internet Source	**<1**%

REFERENCES

Aghebati-Maleki, L., Bakhshinejad, B., Baradaran, B., Motallebnezhad, M., Aghebati-Maleki, A., Nickho, H., Yousefi, M. and Majidi, J. (2016) Phage display as a promising approach for vaccine development. *Journal of Biomedical Science* **23**(1) 66.

Al-Olayan, E.M., Williams, G.T. and Hurd, H. (2002) Apoptosis in the malaria protozoan, *Plasmodium berghei*: A possible mechanism for limiting intensity of infection in the mosquito. *International Journal for Parasitology* **32**(9) 1133–1143.

Ali, M., Al-Olayan, E.M., Lewis, S., Matthews, H. and Hurd, H. (2010) Naturally occurring triggers that induce apoptosis-like programmed cell death in *Plasmodium berghei* ookinetes. *PLoS ONE* **5**(9) e12634.

Amitai, G. and Sorek, R. (2016) CRISPR–Cas adaptation: insights into the mechanism of action. *Nature Reviews Microbiology* **14**(2) 67–76.

Anders, R.F. (1986) Multiple cross-reactivities amongst antigens of *Plasmodium falciparum* impair the development of protective immunity against malaria. *Parasite Immunology* **8**(6) 529–539.

Ansuini, H., Cicchini, C., Nicosia, A., Tripodi, M., Cortese, R. and Luzzago, A. (2002) Biotin-tagged cDNA expression libraries displayed on lambda phage: a new tool for the selection of natural protein ligands. *Nucleic Acids Research* **30**(15) e78.

Arambage, S.C., Grant, K.M., Pardo, I., Ranford-Cartwright, L. and Hurd, H. (2009) Malaria ookinetes exhibit multiple markers for apoptosis-like programmed cell death *in vitro*. *Parasites & Vectors* **2**(1) 32.

Aravind, L., Iyer, L.M., Wellems, T.E. and Miller, L.H. (2003) *Plasmodium* biology. *Cell* **115**(7) 771–785.

Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**(2) 195–201.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M. and Sherlock, G. (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**(1) 25–29.

Ashkenazi, A. (2008) Directing cancer cells to self-destruct with pro-apoptotic receptor agonists. *Nature Reviews Drug Discovery* **7**(12) 1001–1012.

Ashkenazi, A. and Salvesen, G. (2014) Regulated cell death: signaling and mechanisms. *Annual Review of Cell and Developmental Biology* **30**(1) 337–356.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Miller, J.A., Nayak, V., Pennington, C., Pinney, D.F., Roos, D.S., Ross, C., Stoeckert, C.J., Treatman, C. and Wang, H. (2009) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research* **37**(Database) D539–D543.

Baldwin, M., Yamodo, I., Ranjan, R., Li, X., Mines, G., Marinkovic, M., Hanada, T., Oh, S.S. and Chishti, A.H. (2014) Human erythrocyte band 3 functions as a receptor for the sialic acid-independent invasion of *Plasmodium falciparum*. Role of the RhopH3–MSP1 complex. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1843**(12) 2855–2870.

Baldwin, M.R., Li, X., Hanada, T., Liu, S.-C. and Chishti, A.H. (2015) Merozoite surface protein 1 recognition of host glycophorin A mediates malaria parasite invasion of red blood cells. *Blood* **125**(17) 2704–2711.

Bansal, A., Ojo, K.K., Mu, J., Maly, D.J., Van Voorhis, W.C. and Miller, L.H. (2016) Reduced activity of mutant calcium-dependent protein kinase 1 is compensated in *Plasmodium falciparum* through the action of protein kinase G. *mBio* **7**(6) e02011-16.

Baum, J., Papenfuss, A.T., Mair, G.R., Janse, C.J., Vlachou, D., Waters, A.P., Cowman, A.F., Crabb, B.S. and DeKoning-Ward, T.F. (2009) Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. *Nucleic Acids Research* **37**(11) 3788–3798.

Benelli, G. and Beier, J.C. (2017) Current vector control challenges in the fight against malaria. *Acta Tropica* **174** 91–96.

Benkert, P., Biasini, M. and Schwede, T. (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**(3) 343–350.

Benkert, P., Künzli, M. and Schwede, T. (2009) QMEAN server for protein model quality estimation. *Nucleic Acids Research* **37**(suppl_2) W510–W514.

Berggård, T., Linse, S. and James, P. (2007) Methods for the detection and analysis of protein–protein interactions. *PROTEOMICS* **7**(16) 2833–2842.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L. and Schwede, T. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* **42**(W1) W252–W258.

Bimboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**(6) 1513–1523.

Birkholtz, L.M., Blatch, G., Coetzer, T.L., Hoppe, H.C., Human, E., Morris, E.J., Ngcete, Z., Oldfield, L., Roth, R., Shonhai, A., Stephens, L. and Louw, A.I. (2008) Heterologous expression of plasmodial proteins for structural studies and functional annotation. *Malaria Journal* **7**(1) 197.

Bjellqvist, B., Basse, B., Olsen, E. and Celis, J.E. (1994) Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis* **15**(3–4) 529–539.

Bjellqvist, B., Hughes, G.J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J.C., Frutiger, S. and Hochstrasser, D. (1993) The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* **14**(10) 1023–31.

Black, C.G., Proellocks, N.I., Kats, L.M., Cooke, B.M., Mohandas, N. and Coppel, R.L. (2008) *In vivo* studies support the role of trafficking and cytoskeletal-binding motifs in the interaction of MESA with the membrane skeleton of *Plasmodium falciparum*-infected red blood cells. *Molecular and Biochemical Parasitology* **160**(2) 143–147.

Blomqvist, K. (2008) Thawing of glycerolyte-frozen parasites with NaCl. In K. Moll *et al.*, eds. *Methods in Malaria Research*. Paris: Malaria Research and Reference Reagent Resource Center.

Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J. and Schwede, T. (2009) Protein structure homology modeling using SWISS-MODEL workspace. *Nature Protocols* **4**(1) 1–13.

Bousema, J., Drakeley, C. and Sauerwein, R. (2006) Sexual-stage antibody responses to *P. falciparum* in endemic populations. *Current Molecular Medicine* **6**(2) 223–229.

Bruce, M.C., Alano, P., Duthie, S. and Carter, R. (1990) Commitment of the malaria parasite *Plasmodium falciparum* to sexual and asexual development. *Parasitology* **100 Pt 2** 191–200.

Bryant, J.M., Regnault, C., Scheidig-Benatar, C., Baumgarten, S., Guizetti, J. and Scherf, A. (2017) CRISPR/Cas9 genome editing reveals that the intron is not essential for var2csa gene activation or silencing in *Plasmodium falciparum*. *mBio* **8**(4) e00729-17.

Buchan, D.W.A., Minneci, F., Nugent, T.C.O., Bryson, K. and Jones, D.T. (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research* **41**(W1) W349–W357.

Burstein, E., Ganesh, L., Dick, R.D., van De Sluis, B., Wilkinson, J.C., Klomp, L.W.J., Wijmenga, C., Brewer, G.J., Nabel, G.J. and Duckett, C.S. (2004) A novel role for XIAP in copper homeostasis through regulation of MURR1. *The EMBO Journal* **23**(1) 244–254.

Caberoy, N.B., Zhou, Y. and Li, W. (2009) Can phage display be used as a tool to functionally identify endogenous eat-me signals in phagocytosis? *Journal of Biomolecular Screening* **14**(6) 653–661.

Cao, Z., Li, X., Li, J., Luo, W. and Huang, C. (2014) X-linked inhibitor of apoptosis protein (XIAP) lacking RING domain localizes to the nuclear and promotes cancer cell anchorage-independent growth by targeting the E2F1 / Cyclin E axis. *Oncotarget* **5**(16) 7126–7137.

Carlson, D.F., Fahrenkrug, S.C. and Hackett, P.B. (2012) Targeting DNA with fingers and TALENs. *Molecular Therapy - Nucleic Acids* **1**(1) e3.

Ceasar, S.A., Rajan, V., Prykhozhij, S. V., Berman, J.N. and Ignacimuthu, S. (2016) Insert, remove or replace: A highly advanced genome editing system using CRISPR/Cas9. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1863**(9) 2333–2344.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N. V, Bogdanove, A.J. and Voytas, D.F. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* **39**(12) e82–e82.

Cha, S.J., Park, K., Srinivasan, P., Schindler, C.W., van Rooijen, N., Stins, M. and Jacobs-Lorena, M. (2015) CD68 acts as a major gateway for malaria sporozoite liver infection. *The Journal of Experimental Medicine* **212**(9) 1391–403.

Cheng, J., Randall, A.Z., Sweredoski, M.J. and Baldi, P. (2005) SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Research* **33**(Web Server) W72–W76.

Cheong, F.W., Fong, M.Y. and Lau, Y.L. (2016) Identification and characterization of epitopes on *Plasmodium knowlesi* merozoite surface protein-142 (MSP-142) using synthetic peptide library and phage display library. *Acta Tropica* **154** 89–94.

Churchyard, A. (2017) *Trafficking of Plasmodium falciparum invasion proteins to the parasite micronemes*. PhD Thesis - University of the Witwatersrand.

Clontech Laboratories (2016) In-Fusion® Molar Ratio Calculator. Available at: http://bioinfo.clontech.com/infusion/molarRatio.do [Accessed February 2016].

Coetzer, T.L., Durand, P.M. and Nedelcu, A.M. (2010) Genomic evidence for elements of a programmed cell death pathway in *Plasmodium*: Exploiting programmed parasite death for malaria control? *Blood* **116** 4226.

Collins, W.E. (2012) *Plasmodium knowlesi*: A malaria parasite of monkeys and humans. *Annual Review of Entomology* **57**(1) 107–121.

Collins, W.E. and Jeffery, G.M. (2007) *Plasmodium malariae*: Parasite and disease. *Clinical Microbiology Reviews* **20**(4) 579–592.

Collins, W.E. and Jeffery, G.M. (2005) *Plasmodium ovale*: Parasite and disease. *Clinical Microbiology Reviews* **18**(3) 570–581.

Comeaux, C.A., Coleman, B.I., Bei, A.K., Whitehurst, N. and Duraisingh, M.T. (2011) Functional analysis of epigenetic regulation of tandem RhopH1/clag genes reveals a role in *Plasmodium falciparum* growth. *Molecular Microbiology* **80**(2) 378–390.

Condron, B.G., Atkins, J.F. and Gesteland, R.F. (1991) Frameshifting in gene 10 of bacteriophage T7. *Journal of Bacteriology* **173**(21) 6998–7003.

Cooper, J.A., Ingram, L.T., Bushell, G.R., Fardoulys, C.A., Stenzel, D., Schofield, L. and Saul, A.J. (1988) The 140/130/105 kilodalton protein complex in the rhoptries of *Plasmodium falciparum* consists of discrete polypeptides. *Molecular and Biochemical Parasitology* **29**(2–3) 251–60.

Cortes, A., Mellombo, M., Masciantonio, R., Murphy, V.J., Reeder, J.C. and Anders, R.F. (2005) Allele specificity of naturally acquired antibody responses against *Plasmodium falciparum* apical membrane antigen 1. *Infection and Immunity* **73**(1) 422–430.

Counihan, N.A., Chisholm, S.A., Bullen, H.E., Srivastava, A., Sanders, P.R., Jonsdottir, T.K., Weiss, G.E., Ghosh, S., Crabb, B.S., Creek, D.J., Gilson, P.R. and DeKoning-Ward, T.F. (2017) *Plasmodium falciparum* parasites deploy RhopH2 into the host erythrocyte to obtain nutrients, grow and replicate. *eLife* **6**.

Cowman, A., Crabb, B., Maier, A., Tonkin, C., Healer, J., Gibson, P. and DeKoning-Ward, T.F. (2008) Preparation of *P. falciparum* genomic DNA. In K. Moll *et al.*, eds. *Methods in Malaria Research*. Paris: Malaria Research and Reference Reagent Resource Center, pp. 287–288.

Crawford, E.D., Quan, J., Horst, J.A., Ebert, D., Wu, W. and DeRisi, J.L. (2017) Plasmid-free CRISPR/Cas9 genome editing in *Plasmodium falciparum* confirms mutations conferring resistance to the dihydroisoquinolone clinical candidate SJ733. *PLoS ONE* **12**(5) e0178163.

Crook, N.E., Clem, R.J. and Miller, L.K. (1993) An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *Journal of Virology* **67**(4) 2168–2174.

Cui, Y. and Yu, L. (2016) Application of the CRISPR/Cas9 gene editing technique to research on functional genomes of parasites. *Parasitology International* **65**(6) 641–644.

Czabotar, P.E., Lessene, G., Strasser, A. and Adams, J.M. (2013) Control of apoptosis by the BCL-2 protein family: implications for physiology and therapy. *Nature Reviews Molecular Cell Biology* **15**(1) 49–63.

Danner, S. and Belasco, J.G. (2001) T7 phage display: A novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. *Proceedings of the National Academy of Sciences* **98**(23) 12954–12959.

DeAlmagro, M.C. and Vucic, D. (2012) The inhibitor of apoptosis (IAP) proteins are critical regulators of signaling pathways and targets for anti-cancer therapy. *Experimental Oncology* **34**(3) 200–211.

DeKoning-Ward, T.F., Gilson, P.R. and Crabb, B.S. (2015) Advances in molecular genetic systems in malaria. *Nature Reviews Microbiology* **13**(6) 373–387.

Deponte, M. and Becker, K. (2004) *Plasmodium falciparum* - do killers commit suicide? *Trends in Parasitology* **20**(4) 165–169.

Deveraux, Q.L., Roy, N., Stennicke, H.R., Van Arsdale, T., Zhou, Q., Srinivasula, S.M., Alnemri, E.S., Salvesen, G.S. and Reed, J.C. (1998) IAPs block apoptotic events induced by caspase-8 and cytochrome c by direct inhibition of distinct caspases. *The EMBO Journal* **17**(8) 2215–2223.

Dick, S.A. and Megeney, L.A. (2013) Cell death proteins: An evolutionary role in cellular adaptation before the advent of apoptosis. *BioEssays* **35**(11) 974–983.

Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J. and Root, D.E. (2014) Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nature Biotechnology* **32**(12) 1262–1267.

Doetschman, T. and Georgieva, T. (2017) Gene editing with CRISPR/Cas9 RNA-directed nuclease. *Circulation Research* **120**(5) 876–894.

Dondorp, A.M., Nosten, F., Yi, P., Das, D., Phyo, A.P., Tarning, J., Lwin, K.M., Ariey, F., Hanpithakpong, W., Lee, S.J., Ringwald, P., Silamut, K., Imwong, M., Chotivanich, K., Lim, P., Herdman, T., An, S.S., Yeung, S., Singhasivanon, P., Day, N.P.J., Lindegardh, N., Socheat, D. and White, N.J. (2009) Artemisinin resistance in *Plasmodium falciparum* malaria. *New England Journal of Medicine* **361**(5) 455–467.

Dondorp, A.M., Smithuis, F.M., Woodrow, C. and von Seidlein, L. (2017) How to contain artemisinin- and multidrug-resistant falciparum malaria. *Trends in Parasitology* **33**(5) 353–363.

Drozdetskiy, A., Cole, C., Procter, J. and Barton, G.J. (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Research* **43**(W1) W389–W394.

Eckelman, B.P., Salvesen, G.S. and Scott, F.L. (2006) Human inhibitor of apoptosis proteins: why XIAP is the black sheep of the family. *EMBO Reports* **7**(10) 988–994.

Edgar, R.C. (2004a) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**(1) 113.

Edgar, R.C. (2004b) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**(5) 1792–1797.

Eid, A. and Mahfouz, M.M. (2016) Genome editing: the road of CRISPR/Cas9 from bench to clinic. *Experimental & Molecular Medicine* **48**(10) e265.

Elmore, S. (2007) Apoptosis: A review of programmed cell death. *Toxicologic Pathology* **35**(4) 495–516.

Emes, R.D. (2001) A new sequence motif linking lissencephaly, Treacher Collins and oral-facial-digital type 1 syndromes, microtubule dynamics and cell migration. *Human Molecular Genetics* **10**(24) 2813–2820.

Engelbrecht, D. and Coetzer, T.L. (2016) *Plasmodium falciparum* exhibits markers of regulated cell death at high population density *in vitro*. *Parasitology International* **65**(6) 715–727.

Engelbrecht, D. and Coetzer, T.L. (2015) Sunlight inhibits growth and induces markers of programmed cell death in *Plasmodium falciparum in vitro*. *Malaria Journal* **14**(1) 378.

Engelbrecht, D. and Coetzer, T.L. (2013) Turning up the heat: heat stress induces markers of programmed cell death in *Plasmodium falciparum in vitro*. *Cell Death and Disease* **4**(12) e971.

Engelbrecht, D., Durand, P.M. and Coetzer, T.L. (2012) On programmed cell death in *Plasmodium falciparum*: Status quo. *Journal of Tropical Medicine* **2012** 646534.

Faix, P.H., Burg, M.A., Gonzales, M., Ravey, E.P., Baird, A. and Larocca, D. (2004) Phage display of cDNA libraries: enrichment of cDNA expression using open reading frame selection. *BioTechniques* **36**(6) 1018–1029.

Farrell, A., Thirugnanam, S., Lorestani, A., Dvorin, J.D., Eidell, K.P., Ferguson, D.J.P., Anderson-White, B.R., Duraisingh, M.T., Marth, G.T. and Gubbels, M.-J. (2012) A DOC2 protein identified by mutational profiling is essential for Apicomplexan parasite exocytosis. *Science* **335**(6065) 218–221.

Feltham, R., Khan, N. and Silke, J. (2012) IAPS and ubiquitylation. *IUBMB Life* **64**(5) 411–418.

Ferreira, M.U., Ribeiro, W.L., Tonon, A.P., Kawamoto, F. and Rich, S.M. (2003) Sequence diversity and evolution of the malaria vaccine candidate merozoite surface protein-1 (MSP-1) of *Plasmodium falciparum*. *Gene* **304** 65–75.

Fields, S. and Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340**(6230) 245–246.

Finlay, D., Teriete, P., Vamos, M., Cosford, N.D.P. and Vuori, K. (2017) Inducing death in tumor cells: roles of the inhibitor of apoptosis proteins. *F1000Research* **6** 587.

Flick, K., Ahuja, S., Chene, A., Bejarano, M.T. and Chen, Q. (2004) Optimized expression of *Plasmodium falciparum* erythrocyte membrane protein 1 domains in *Escherichia coli*. *Malaria Journal* **3**(1) 50.

Fox, B.A., Ristuccia, J.G., Gigley, J.P. and Bzik, D.J. (2009) Efficient gene replacements in *Toxoplasma gondii* strains deficient for nonhomologous end joining. *Eukaryotic Cell* **8**(4) 520–529.

Friedrich, R., Yeheskel, A. and Ashery, U. (2010) DOC2B, C2 domains, and calcium: A tale of intricate interactions. *Molecular Neurobiology* **41**(1) 42–51.

Fuchs, Y. and Steller, H. (2015) Live to die another way: modes of programmed cell death and the signals emanating from dying cells. *Nature Reviews Molecular Cell Biology* **16**(6) 329–344.

Gaj, T., Gersbach, C.A. and Barbas, C.F. (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in biotechnology* **31**(7) 397–405.

Galluzzi, L., Blomgren, K. and Kroemer, G. (2009) Mitochondrial membrane permeabilization in neuronal injury. *Nature Reviews Neuroscience* **10**(7) 481–494.

Galluzzi, L., Bravo-San Pedro, J.M., Vitale, I., Aaronson, S.A., Abrams, J.M., Kroemer, G., *et al.* (2015) Essential versus accessory aspects of cell death: recommendations of the NCCD 2015. *Cell Death and Differentiation* **22**(1) 58–73.

Galluzzi, L., Kepp, O., Trojel-Hansen, C. and Kroemer, G. (2012) Non-apoptotic functions of apoptosis-regulatory proteins. *EMBO Reports* **13**(4) 322–330.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.-S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M.A., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and Barrell, B. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**(6906) 498–511.

Garg, H., Suri, P., Gupta, J.C., Talwar, G.P. and Dubey, S. (2016) Survivin: a unique target for tumor therapy. *Cancer Cell International* **16**(1) 49.

Gaspar-Maia, A., Alajem, A., Polesso, F., Sridharan, R., Mason, M.J., Heidersbach, A., Ramalho-Santos, J., McManus, M.T., Plath, K., Meshorer, E. and Ramalho-Santos, M. (2009) Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* **460**(7257) 863–868.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. In J. M. Walker, ed. *The Proteomics Protocols Handbook*. Paris: Humana Press.

Georgieva, Y. and Konthur, Z. (2011) Design and screening of M13 phage display cDNA libraries. *Molecules* **16**(12) 1667–1681.

Gerlitz, G., Darhin, E., Giorgio, G., Franco, B. and Reiner, O. (2005) Novel functional features of the LIS-H domain: Role in protein dimerization, half-life and cellular localization. *Cell Cycle* **4**(11) 1632–1640.

Ghorbal, M., Gorman, M., Macpherson, C.R., Martins, R.M., Scherf, A. and Lopez-Rubio, J.-J. (2014) Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nature Biotechnology* **32**(8) 819–821.

Gibson, D.G., Young, L., Chuang, R., Venter, J.C., Hutchison, C. a and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**(5) 343–345.

Gill, E.E. and Fast, N.M. (2007) Stripped-down DNA repair in a highly reduced parasite. *BMC Molecular Biology* **8**(1) 24.

Good, M. (2005) Vaccine-induced immunity to malaria parasites and the need for novel strategies. *Trends in Parasitology* **21**(1) 29–34.

Gough, J., Karplus, K., Hughey, R. and Chothia, C. (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *Journal of Molecular Biology* **313**(4) 903–919.

Gowrishankar, J. and Harinarayanan, R. (2004) Why is transcription coupled to translation in bacteria? *Molecular Microbiology* **54**(3) 598–603.

Grabski, A., Mehler, M. and Drott, D. (2005) The Overnight Express Autoinduction System: High-density cell growth and protein expression while you sleep. *Nature Methods* **2**(3) 233–235.

Guha, T.K., Wai, A. and Hausner, G. (2017) Programmable genome editing tools and their regulation for efficient genome engineering. *Computational and Structural Biotechnology Journal* **15** 146–160.

Hafalla, J.C., Silvie, O. and Matuschewski, K. (2011) Cell biology and immunology of malaria. *Immunological Reviews* **240**(1) 297–316.

Hall, J.A. and Georgel, P.T. (2007) CHD proteins: a diverse family with strong ties. *Biochemistry and Cell Biology* **85**(4) 463–476.

Herman, M.D., Moche, M., Flodin, S., Welin, M., Trésaugues, L., Johansson, I., Nilsson, M., Nordlund, P. and Nyman, T. (2009) Structures of BIR domains from human NAIP and cIAP2. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **65**(11) 1091–1096.

Hille, F. and Charpentier, E. (2016) CRISPR-Cas: biology, mechanisms and relevance. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**(1707) 20150496.

Hinds, M.G., Norton, R.S., Vaux, D.L. and Day, C.L. (1999) Solution structure of a baculoviral inhibitor of apoptosis (IAP) repeat. *Nature Structural Biology* **6**(7) 648–651.

Ichim, G. and Tait, S.W.G. (2016) A fate worse than death: apoptosis as an oncogenic process. *Nature Reviews Cancer* **16**(8) 539–548.

Integrated DNA Technologies (2016) OligoAnalyzer 3.1. Available at: https://eu.idtdna.com/calc/analyzer [Accessed February 2016].

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. and Nakata, A. (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology* **169**(12) 5429–5433.

Ito, D., Schureck, M.A. and Desai, S.A. (2017) An essential dual-function complex mediates erythrocyte invasion and channel-mediated nutrient uptake in malaria parasites. *eLife* **6**.

Jean, S., Zapata-Jenks, M.A., Farley, J.M., Tracy, E. and Mayer, D.C.G. (2014) *Plasmodium falciparum* double C2 domain protein, PfDOC2, binds to calcium when associated with membranes. *Experimental Parasitology* **144**(1) 91–95.

Jones, D.T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology* **292**(2) 195–202.

Josling, G.A. and Llinás, M. (2015) Sexual development in *Plasmodium* parasites: knowing when it's time to commit. *Nature Reviews Microbiology* **13**(9) 573–587.

Kaczanowski, S., Sajid, M. and Reece, S.E. (2011) Evolution of apoptosis-like programmed cell death in unicellular protozoan parasites. *Parasites & Vectors* **4**(1) 44.

Kalniņa, Z., Siliņa, K., Meistere, I., Zayakin, P., Rivosh, A., Abols, A., Leja, M., Minenkova, O., Schadendorf, D. and Linē, A. (2008) Evaluation of T7 and lambda phage display systems for survey of autoantibody profiles in cancer patients. *Journal of Immunological Methods* **334**(1–2) 37–50.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**(6) 845–858.

Kelly, A.E., Ghenoiu, C., Xue, J.Z., Zierhut, C., Kimura, H. and Funabiki, H. (2010) Survivin reads phosphorylated histone H3 threonine 3 to activate the mitotic kinase Aurora B. *Science* **330**(6001) 235–239.

Kerlin, D.H. and Gatton, M.L. (2013) Preferential invasion by *Plasmodium* merozoites and the self-regulation of parasite burden. *PLoS ONE* **8**(2) e57434.

Kister, A.E., Fokas, A.S., Papatheodorou, T.S. and Gelfand, I.M. (2006) Strict rules determine arrangements of strands in sandwich proteins. *Proceedings of the National Academy of Sciences* **103**(11) 4107–4110.

Knuepfer, E., Napiorkowska, M., van Ooij, C. and Holder, A.A. (2017) Generating conditional gene knockouts in *Plasmodium* − a toolkit to produce stable DiCre recombinase-expressing parasite lines using CRISPR/Cas9. *Scientific Reports* **7**(1) 3881.

Kocab, A.J. and Duckett, C.S. (2016) Inhibitor of apoptosis proteins as intracellular signaling intermediates. *The FEBS Journal* **283**(2) 221–231.

Krumpe, L.R.H., Atkinson, A.J., Smythers, G.W., Kandel, A., Schumacher, K.M., McMahon, J.B., Makowski, L. and Mori, T. (2006) T7 lytic phage-displayed peptide libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide libraries. *Proteomics* **6**(15) 4210–4222.

Kumar, S. (2007) Caspase function in programmed cell death. *Cell Death and Differentiation* **14**(1) 32–43.

Kuscu, C., Arslan, S., Singh, R., Thorpe, J. and Adli, M. (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nature Biotechnology* **32**(7) 677–683.

LaCount, D.J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J.R., Schoenfeld, L.W., Ota, I., Sahasrabudhe, S., Kurschner, C., Fields, S. and Hughes, R.E. (2005) A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* **438**(7064) 103–107.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**(5259) 680–685.

Laity, J.H., Lee, B.M. and Wright, P.E. (2001) Zinc finger proteins: new insights into structural and functional diversity. *Current Opinion in Structural Biology* **11**(1) 39–46.

Lambros, C. and Vanderberg, J.P. (1979) Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *The Journal of Parasitology* **65**(3) 418.

LaMonte, G., Lim, M.Y.-X., Wree, M., Reimer, C., Nachon, M., Corey, V., Gedeck, P., Plouffe, D., Du, A., Figueroa, N., Yeung, B., Bifani, P. and Winzeler, E.A. (2016) Mutations in the *Plasmodium falciparum* cyclic amine resistance locus (PfCARL) confer multidrug resistance. *mBio* **7**(4) e00696-16.

Lanzillotti, R. and Coetzer, T.L. (2004) Myosin-like sequences in the malaria parasite *Plasmodium falciparum* bind human erythrocyte membrane protein 4.1. *Haematologica* **89**(10) 1168–71.

Lanzillotti, R. and Coetzer, T.L. (2008) Phage display: a useful tool for malaria research? *Trends in Parasitology* **24**(1) 18–23.

Lauterbach, S.B., Lanzillotti, R. and Coetzer, T.L. (2003) Construction and use of *Plasmodium falciparum* phage display libraries to identify host parasite interactions. *Malaria Journal* **2**(1) 47.

LeChat, L., Sinden, R.E. and Dessens, J.T. (2007) The role of metacaspase 1 in *Plasmodium berghei* development and apoptosis. *Molecular and Biochemical Parasitology* **153**(1) 41–47.

Lee, A.H., Symington, L.S. and Fidock, D.A. (2014) DNA repair mechanisms and their biological roles in the malaria parasite *Plasmodium falciparum*. *Microbiology and Molecular Biology Reviews* **78**(3) 469–486.

Lee, M.C. and Fidock, D.A. (2014) CRISPR-mediated genome editing of *Plasmodium falciparum* malaria parasites. *Genome Medicine* **6**(8) 63.

Leow, C., Jones, M., Cheng, Q., Mahler, S. and McCarthy, J. (2014) Production and characterization of specific monoclonal antibodies binding the *Plasmodium falciparum* diagnostic biomarker, histidine-rich protein 2. *Malaria Journal* **13**(1) 277.

Letunic, I., Doerks, T. and Bork, P. (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Research* **43**(D1) D257–D260.

Li, L., Stoeckert, C.J. and Roos, D.S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research* **13**(9) 2178–2189.

Li, W. (2012) ORF phage display to identify cellular proteins with different functions. *Methods* **58**(1) 2–9.

Li, W. and Caberoy, N.B. (2010) New perspective for phage display as an efficient and versatile technology of functional proteomics. *Applied Microbiology and Biotechnology* **85**(4) 909–919.

Li, X., Chen, H., Khan, A.A., Lauterbach, S.B., Lanzillotti, R., Rai, P.R., Kane, R.S., Coetzer, T.L. and Chishti, A.H. (2008) Receptor-based identification of an inhibitory peptide against blood stage malaria. *Biochemical and Biophysical Research Communications* **376**(3) 489–493.

Li, X., Marinkovic, M., Russo, C., McKnight, C.J., Coetzer, T.L. and Chishti, A.H. (2012) Identification of a specific region of *Plasmodium falciparum* EBL-1 that binds to host receptor glycophorin B and inhibits merozoite invasion in human red blood cells. *Molecular and Biochemical Parasitology* **183**(1) 23–31.

Liston, P., Fong, W.G., Kelly, N.L., Toji, S., Miyazaki, T., Conte, D., Tamai, K., Craig, C.G., McBurney, M.W. and Korneluk, R.G. (2001) Identification of XAF1 as an antagonist of XIAP anti-caspase activity. *Nature cell biology* **3**(2) 128–33.

Lockshin, R.A. and Williams, C.M. (1964) Programmed cell death. Endocrine potentiation of the breakdown of the intersegmental muscles of silkmoths. *Journal of Insect Physiology* **10**(4) 643–649.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D. and Darnell, J. (2000) Section 11.6: Processing of rRNA and tRNA. In *Molecular Cell Biology*. New York: W. H. Freeman.

López-Barragán, M.J., Lemieux, J., Quiñones, M., Williamson, K.C., Molina-Cruz, A., Cui, K., Barillas-Mury, C., Zhao, K. and Su, X. (2011) Directional gene expression and antisense transcripts in sexual and asexual stages of *Plasmodium falciparum*. *BMC Genomics* **12**(1) 587.

Lu, J., Tong, Y., Pan, J., Yang, Y., Liu, Q., Tan, X., Zhao, S., Qin, L. and Chen, X. (2016) A redesigned CRISPR/Cas9 system for marker-free genome editing in *Plasmodium falciparum*. *Parasites & Vectors* **9**(1) 198.

Lustigman, S., Anders, R.F., Brown, G. V. and Coppel, R.L. (1990) The mature-parasite-infected erythrocyte surface antigen (MESA) of *Plasmodium falciparum* associates with the erythrocyte membrane skeletal protein, band 4.1. *Molecular and Biochemical Parasitology* **38**(2) 261–270.

Ma, H., Wu, Y., Dang, Y., Choi, J.-G., Zhang, J. and Wu, H. (2014) Pol III promoters to express small RNAs: Delineation of transcription initiation. *Molecular Therapy - Nucleic Acids* **3** e161.

Magnan, C.N., Randall, A. and Baldi, P. (2009) SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics* **25**(17) 2200–2207.

Maier, A.G., Braks, J.A.M., Waters, A.P. and Cowman, A.F. (2006) Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in *Plasmodium falciparum* for targeted gene deletion by double crossover recombination. *Molecular and Biochemical Parasitology* **150**(1) 118–121.

Manning, G., Reiner, D.S., Lauwaet, T., Dacre, M., Smith, A., Zhai, Y., Svard, S. and Gillin, F.D. (2011) The minimal kinome of *Giardia lamblia* illuminates early kinase evolution and unique parasite biology. *Genome Biology* **12**(7) R66.

Mariño, G., Niso-Santano, M., Baehrecke, E.H. and Kroemer, G. (2014) Self-consumption: the interplay of autophagy and apoptosis. *Nature Reviews Molecular Cell Biology* **15**(2) 81–94.

McKenzie, F.E., Jeffery, G.M. and Collins, W.E. (2002) *Plasmodium vivax* blood-stage dynamics. *Journal of Parasitology* **88**(3) 521–535.

McLean, K.J. and Jacobs-Lorena, M. (2016) Genetic control of malaria mosquitoes. *Trends in Parasitology* **32**(3) 174–176.

McNamara, C.W., Lee, M.C.S., Lim, C.S., Lim, S.H., Roland, J., Nagle, A., Simon, O., Yeung, B.K.S., Chatterjee, A.K., McCormack, S.L., Manary, M.J., Zeeman, A.-M., Dechering, K.J., Kumar, T.R.S., Henrich, P.P., Gagaring, K., Ibanez, M., Kato, N., Kuhen, K.L., Fischli, C., Rottmann, M., Plouffe, D.M., Bursulaya, B., Meister, S., Rameh, L., Trappe, J., Haasen, D., Timmerman, M., Sauerwein, R.W., Suwanarusk, R., Russell, B., Renia, L., Nosten, F., Tully, D.C., Kocken, C.H.M., Glynne, R.J., Bodenreider, C., Fidock, D.A., Diagana, T.T. and Winzeler, E.A. (2013) Targeting *Plasmodium* PI(4)K to eliminate malaria. *Nature* **504**(7479) 248–253.

Mehlin, C., Boni, E., Buckner, F.S., Engel, L., Feist, T., Gelb, M.H., Haji, L., Kim, D., Liu, C. and Mueller, N. (2006) Heterologous expression of proteins from *Plasmodium falciparum*: Results from 1000 genes. *Molecular and Biochemical Parasitology* **148**(2) 144–160.

Meibalan, E. and Marti, M. (2017) Biology of malaria transmission. *Cold Spring Harbor Perspectives in Medicine* **7**(3) a025452.

Meslin, B., Barnadas, C., Boni, V., Latour, C., Monbrison, F. De, Kaiser, K. and Picot, S. (2007) Features of apoptosis in *Plasmodium falciparum* erythrocytic stage through a putative role of PfMCA1 metacaspase-like protein. *The Journal of Infectious Diseases* **195**(12) 1852–1859.

Meslin, B., Beavogui, A.H., Fasel, N. and Picot, S. (2011a) *Plasmodium falciparum* metacaspase PfMCA-1 triggers a z-VAD-fmk inhibitable protease to promote cell death. *PLoS ONE* **6**(8) e23867.

Meslin, B., Zalila, H., Fasel, N., Picot, S. and Bienvenu, A.-L. (2011b) Are protozoan metacaspases potential parasite killers? *Parasites & Vectors* **4**(1) 26.

Mikolajka, A., Yan, X., Popowicz, G.M., Smialowski, P., Nigg, E.A. and Holak, T.A. (2006) Structure of the N-terminal domain of the FOP (FGFR1OP) protein and implications for its dimerization and centrosomal localization. *Journal of Molecular Biology* **359**(4) 863–875.

Miller, L.H., Baruch, D.I., Marsh, K. and Doumbo, O.K. (2002) The pathogenic basis of malaria. *Nature* **415**(6872) 673–679.

Mogollon, C.M., van Pul, F.J.A., Imai, T., Ramesar, J., Chevalley-Maurel, S., de Roo, G.M., Veld, S.A.J., Kroeze, H., Franke-Fayard, B.M.D., Janse, C.J. and Khan, S.M. (2016) Rapid generation of marker-free *P. falciparum* fluorescent reporter lines using modified CRISPR/Cas9 constructs and selection protocol. *PLoS ONE* **11**(12) e0168362.

Moraes Barros, R.R., Straimer, J., Sa, J.M., Salzman, R.E., Melendez-Muniz, V.A., Mu, J., Fidock, D.A. and Wellems, T.E. (2015) Editing the *Plasmodium vivax* genome, using zinc-finger nucleases. *Journal of Infectious Diseases* **211**(1) 125–129.

Mufti, A.R., Burstein, E. and Duckett, C.S. (2007) XIAP: Cell death regulation meets copper homeostasis. *Archives of Biochemistry and Biophysics* **463**(2) 168–174.

Mutai, B.K. and Waitumbi, J.N. (2010) Apoptosis stalks *Plasmodium falciparum* maintained in continuous culture condition. *Malaria Journal* **9**(Suppl 3) S6.

Nachmias, B., Lazar, I., Elmalech, M., Abed-El-Rahaman, I., Asshab, Y., Mandelboim, O., Perlman, R. and Ben-Yehuda, D. (2007) Subcellular localization determines the delicate balance between the anti- and pro-apoptotic activity of Livin. *Apoptosis* **12**(7) 1129–1142.

NanoDrop (2008) *NanoDrop 1000 Spectrophotometer V3.7 User's Manual*, Wilmington: Thermo Fisher Scientific.

Navale, R., Atul, Allanki, A.D. and Sijwali, P.S. (2014) Characterization of the autophagy marker protein Atg8 reveals atypical features of autophagy in *Plasmodium falciparum*. *PLoS ONE* **9**(11) e113220.

Nedelcu, A.M. (2009) Comparative genomics of phylogenetically diverse unicellular eukaryotes provide new insights into the genetic basis for the evolution of the programmed cell death machinery. *Journal of Molecular Evolution* **68**(3) 256–268.

Ng, C.L., Siciliano, G., Lee, M.C.S., de Almeida, M.J., Corey, V.C., Bopp, S.E., Bertuccini, L., Wittlin, S., Kasdin, R.G., Le Bihan, A., Clozel, M., Winzeler, E.A., Alano, P. and Fidock, D.A. (2016) CRISPR-Cas9-modified pfmdr1 protects *Plasmodium falciparum* asexual blood stages and gametocytes against a class of piperazine-containing compounds but potentiates artemisinin-based combination therapy partner drugs. *Molecular Microbiology* **101**(3) 381–393.

Nguitragool, W., Bokhari, A.B., Pillai, A., Rayavara, K., Sharma, P., Turpin, B., Aravind, L. and Desai, S.A. (2011) Malaria parasite clag3 genes determine channel-mediated nutrient uptake by infected red blood cells. *Cell* **145**(5) 665–677.

Noedl, H., Socheat, D. and Satimai, W. (2009) Artemisinin-resistant malaria in Asia. *New England Journal of Medicine* **361**(5) 540–541.

Normark, J. (2008) Freezing of patient isolates and strains with glycerolyte. In K. Moll *et al.*, eds. *Methods in Malaria Research*. Paris: Malaria Research and Reference Reagent Resource Center.

Novagen (2004) *Competent Cells*, Germany: EMD Biosciences.

Nowak, D., Boehrer, S., Brieger, A., Kim, S.-Z., Schaaf, S., Hoelzer, D., Mitrou, P.S., Weidmann, E. and Chow, K.U. (2004) Upon drug-induced apoptosis in lymphoma cells X-linked inhibitor of apoptosis (XIAP) translocates from the cytosol to the nucleus. *Leukemia & Lymphoma* **45**(7) 1429–1436.

Oakley, M.S.M., Kumar, S., Anantharaman, V., Zheng, H., Mahajan, B., Haynes, J.D., Moch, J.K., Fairhurst, R., McCutchan, T.F. and Aravind, L. (2007) Molecular factors and biochemical pathways induced by febrile temperature in intraerythrocytic *Plasmodium falciparum* parasites. *Infection and Immunity* **75**(4) 2012–2025.

Otto, T.D., Böhme, U., Jackson, A.P., Hunt, M., Franke-Fayard, B., Hoeijmakers, W.A.M., Religa, A.A., Robertson, L., Sanders, M., Ogun, S.A., Cunningham, D., Erhart, A., Billker, O., Khan, S.M., Stunnenberg, H.G., Langhorne, J., Holder, A.A., Waters, A.P., Newbold, C.I., Pain, A., Berriman, M. and Janse, C.J. (2014) A comprehensive evaluation of rodent malaria parasite genomes and gene expression. *BMC Biology* **12**(1) 86.

Pande, J., Szewczyk, M.M. and Grover, A.K. (2010) Phage display: Concept, innovations, applications and future. *Biotechnology Advances* **28**(6) 849–858.

Pease, B.N., Huttlin, E.L., Jedrychowski, M.P., Talevich, E., Harmon, J., Dillman, T., Kannan, N., Doerig, C., Chakrabarti, R., Gygi, S.P. and Chakrabarti, D. (2013) Global analysis of protein expression and phosphorylation of three stages of *Plasmodium falciparum* intraerythrocytic development. *Journal of Proteome Research* **12**(9) 4028–4045.

Petersen, C., Nelson, R., Magowan, C., Wollish, W., Jensen, J. and Leech, J. (1989) The mature erythrocyte surface antigen of *Plasmodium falciparum* is not required for knobs or cytoadherence. *Molecular and Biochemical Parasitology* **36**(1) 61–65.

Picot, S., Burnod, J., Bracchi, V., Chumpitazi, B.F.F. and Ambroise-Thomas, P. (1997) Apoptosis related to chloroquine sensitivity of the human malaria parasite *Plasmodium falciparum*. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **91**(5) 590–591.

Pizzi, E. and Frontali, C. (2001) Low-complexity regions in *Plasmodium falciparum* proteins. *Genome Research* **11**(2) 218–229.

Pollitt, L.C., Colegrave, N., Khan, S.M., Sajid, M. and Reece, S.E. (2010) Investigating the evolution of apoptosis in malaria parasites: the importance of ecology. *Parasites & Vectors* **3**(1) 105.

Promega (2014) *MagneGST^{TM} Protein Purification System*, Madison: Promega Corporation.

Promega (2009) *MagneHis^{TM} Protein Purification System*, Madison: Promega Corporation.

Proto, W.R., Coombs, G.H. and Mottram, J.C. (2012) Cell death in parasitic protozoa: regulated or incidental? *Nature Reviews Microbiology* **11**(1) 58–66.

Rao, V.S., Srinivas, K., Sujini, G.N. and Kumar, G.N.S. (2014) Protein-protein interaction detection: Methods and analysis. *International Journal of Proteomics* **2014**(ii) 1–12.

Rathore, S., Datta, G., Kaur, I., Malhotra, P. and Mohmmed, A. (2015) Disruption of cellular homeostasis induces organelle stress and triggers apoptosis like cell-death pathways in malaria parasite. *Cell Death and Disease* **6**(7) e1803.

Reece, S.E., Pollitt, L.C., Colegrave, N. and Gardner, A. (2011) The meaning of death: Evolution and ecology of apoptosis in protozoan parasites. *PLoS Pathogens* **7**(12) e1002320.

Reece, S.E., Ramiro, R.S. and Nussey, D.H. (2009) Plastic parasites: sophisticated strategies for survival and reproduction? *Evolutionary Applications* **2**(1) 11–23.

Reegan, A.D., Ceasar, S.A., Paulraj, M.G., Ignacimuthu, S. and Al-Dhabi, N.A. (2016) Current status of genome editing in vector mosquitoes: A review. *BioScience Trends* **10**(6) 424–432.

Rosano, G.L. and Ceccarelli, E.A. (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology* **5**(APR) 172.

Rosenberg, A., Griffin, G., Studier, F.W., McCormick, M., Berg, J. and Mierendorf, R. (1996) T7 Select phage display system: a powerful new protein display system based on bacteriophage T7. *inNovations* **6**(6) 1–6.

Sahdev, S., Khattar, S.K. and Saini, K.S. (2007) Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry* **307**(1–2) 249–264.

Samuel, T., Okada, K., Hyer, M., Welsh, K., Zapata, J.M. and Reed, J.C. (2005) cIAP1 localizes to the nuclear compartment and modulates the cell cycle. *Cancer Research* **65**(1) 210–218.

Sander, J.D. and Joung, J.K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology* **32**(4) 347–355.

Sherling, E.S., Knuepfer, E., Brzostowski, J.A., Miller, L.H., Blackman, M.J. and Ooij, C. van (2017) The *Plasmodium falciparum* rhoptry protein RhopH3 plays essential roles in host cell invasion and nutrient uptake. *eLife* **6**.

Silke, J. and Vaux, D.L. (2015) IAP gene deletion and conditional knockout models. *Seminars in Cell & Developmental Biology* **39** 97–105.

Sinden, R.E., Morris, H.R., Billker, O., Lindo, V., Panico, M., Etienne, A.E., Paxton, T., Dell, A. and Rogers, M. (1998) Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito. *Nature* **392**(6673) 289–292.

Singer, M., Marshall, J., Heiss, K., Mair, G.R., Grimm, D., Mueller, A.-K. and Frischknecht, F. (2015) Zinc finger nuclease-based double-strand breaks attenuate malaria parasites and reveal rare microhomology-mediated end joining. *Genome Biology* **16**(1) 249.

Singh, V., Braddick, D. and Dhar, P.K. (2017) Exploring the potential of genome editing CRISPR-Cas9 technology. *Gene* **599** 1–18.

Smialowski, P., Doose, G., Torkler, P., Kaufmann, S. and Frishman, D. (2012) PROSO II - a new method for protein solubility prediction. *The FEBS Journal* **279**(12) 2192–2200.

Smith, G. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**(4705) 1315–1317.

Soni, R., Sharma, D., Rai, P., Sharma, B. and Bhatt, T.K. (2017) Signaling strategies of malaria parasite for its survival, proliferation, and infection during erythrocytic stage. *Frontiers in Immunology* **8** 349.

Sonoiki, E., Ng, C.L., Lee, M.C.S., Guo, D., Zhang, Y.-K., Zhou, Y., Alley, M.R.K., Ahyong, V., Sanz, L.M., Lafuente-Monasterio, M.J., Dong, C., Schupp, P.G., Gut, J., Legac, J., Cooper, R.A., Gamo, F.-J., DeRisi, J., Freund, Y.R., Fidock, D.A. and Rosenthal, P.J. (2017) A potent antimalarial benzoxaborole targets a *Plasmodium falciparum* cleavage and polyadenylation specificity factor homologue. *Nature Communications* **8** 14574.

Sow, F., Nyonda, M., Bienvenu, A.-L. and Picot, S. (2015) Wanted *Plasmodium falciparum*, dead or alive. *Microbial Cell* **2**(7) 219–224.

Srinivasula, S.M. and Ashwell, J.D. (2008) IAPs: What's in a name? *Molecular Cell* **30**(2) 123–135.

Srinivasula, S.M., Hegde, R., Saleh, A., Datta, P., Shiozaki, E., Chai, J., Lee, R.-A., Robbins, P.D., Fernandes-Alnemri, T., Shi, Y. and Alnemri, E.S. (2001) A conserved XIAP-interaction motif in caspase-9 and Smac/DIABLO regulates caspase activity and apoptosis. *Nature* **410**(6824) 112–116.

Straimer, J., Lee, M.C.S., Lee, A.H., Zeitler, B., Williams, A.E., Pearl, J.R., Zhang, L., Rebar, E.J., Gregory, P.D., Llinás, M., Urnov, F.D. and Fidock, D.A. (2012) Site-specific genome editing in *Plasmodium falciparum* using engineered zinc-finger nucleases. *Nature Methods* **9**(10) 993–998.

Sultan, A.A., Thathy, V., DeKoning-Ward, T.F. and Nussenzweig, V. (2001) Complementation of *Plasmodium berghei* TRAP knockout parasites using human dihydrofolate reductase gene as a selectable marker. *Molecular and Biochemical Parasitology* **113**(1) 151–156.

Sundell, G.N. and Ivarsson, Y. (2014) Interaction analysis through proteomic phage display. *BioMed Research International* **2014** 1–9.

Swart, C., Du Toit, A. and Loos, B. (2016) Autophagy and the invisible line between life and death. *European Journal of Cell Biology* **95**(12) 598–610.

Tai, H.H., Geisterfer, M., Bell, J.C., Moniwa, M., Davie, J.R., Boucher, L. and McBurney, M.W. (2003) CHD1 associates with NCoR and histone deacetylase as well as with RNA splicing proteins. *Biochemical and Biophysical Research Communications* **308**(1) 170–176.

Tait, S.W.G., Ichim, G. and Green, D.R. (2014) Die another way - non-apoptotic mechanisms of cell death. *Journal of Cell Science* **127**(10) 2135–2144.

Tirasophon, W., Ponglikitmongkol, M., Wilairat, P., Boonsaeng, V. and Panyim, S. (1991) A novel detection of a single *Plasmodium falciparum* in infected blood. *Biochemical and Biophysical Research Communications* **175**(1) 179–184.

Tompa, P. (2003) Intrinsically unstructured proteins evolve by repeat expansion. *BioEssays* **25**(9) 847–855.

Totino, P.R.R., Magalhães, A. das D., Alves, E.B., Costa, M.R.F., de Lacerda, M.V.G., Daniel-Ribeiro, C.T. and Ferreira-da-Cruz, M. de F. (2014) *Plasmodium falciparum*, but not *P. vivax*, can induce erythrocytic apoptosis. *Parasites & Vectors* **7**(1) 484.

Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences* **76**(9) 4350–4354.

Trager, W. and Jensen, J. (1976) Human malaria parasites in continuous culture. *Science* **193**(4254) 673–675.

Vandenabeele, P., Galluzzi, L., Vanden Berghe, T. and Kroemer, G. (2010) Molecular mechanisms of necroptosis: an ordered cellular explosion. *Nature Reviews Molecular Cell Biology* **11**(10) 700–714.

VanWye, J.D. and Haldar, K. (1997) Expression of green fluorescent protein in *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* **87**(2) 225–229.

Vega-Rodriguez, J., Perez-Barreto, D., Ruiz-Reyes, A. and Jacobs-Lorena, M. (2015) Targeting molecular interactions essential for *Plasmodium* sexual reproduction. *Cellular Microbiology* **17**(11) 1594–1604.

Vieira, W.A. (2015) *Investigating the Molecular Participants of Programmed Cell Death in Plamodium falciparum*. PhD Thesis - University of the Witwatersrand.

Vieira, W.A. and Coetzer, T.L. (2016) Localization and interactions of *Plasmodium falciparum* SWIB/MDM2 homologues. *Malaria Journal* **15**(1) 32.

Volz, J., Carvalho, T.G., Ralph, S.A., Gilson, P., Thompson, J., Tonkin, C.J., Langer, C., Crabb, B.S. and Cowman, A.F. (2010) Potential epigenetic regulatory proteins localise to distinct nuclear sub-compartments in *Plasmodium falciparum*. *International Journal for Parasitology* **40**(1) 109–121.

Vu, T.T., Tran, V.B., Phan, N.T., Le, T.T., Luong, V.H., O'Brien, E. and Morris, G.E. (1995) Screening donor blood for malaria by polymerase chain reaction. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **89**(1) 44–47.

Wagner, J.C., Platt, R.J., Goldfless, S.J., Zhang, F. and Niles, J.C. (2014) Efficient CRISPR-Cas9–mediated genome editing in *Plasmodium falciparum*. *Nature Methods* **11**(9) 915–918.

Waller, K.L., Nunomura, W., An, X., Cooke, B.M., Mohandas, N. and Coppel, R.L. (2003) Mature parasite-infected erythrocyte surface antigen (MESA) of *Plasmodium falciparum* binds to the 30-kDa domain of protein 4.1 in malaria-infected red blood cells. *Blood* **102**(5) 1911–1914.

Wang, F., Dai, J., Daum, J.R., Niedzialkowska, E., Banerjee, B., Stukenberg, P.T., Gorbsky, G.J. and Higgins, J.M.G. (2010) Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. *Science* **330**(6001) 231–235.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G.J. (2009) Jalview version 2: A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**(9) 1189–1191.

Webster, W.A.J. and McFadden, G.I. (2014) From the genome to the phenome: Tools to understand the basic biology of *Plasmodium falciparum*. *Journal of Eukaryotic Microbiology* **61**(6) 655–671.

Wei, Y., Jin, J. and Harper, J.W. (2003) The cyclin E/Cdk2 substrate and cajal body component p220NPAT activates histone transcription through a novel LisH-like domain. *Molecular and Cellular Biology* **23**(10) 3669–3680.

White, N.J. (2011) Determinants of relapse periodicity in *Plasmodium vivax* malaria. *Malaria Journal* **10**(1) 297.

Wong, W., Bai, X., Brown, A., Fernandez, I.S., Hanssen, E., Condron, M., Tan, Y.H., Baum, J. and Scheres, S.H. (2014) Cryo-EM structure of the *Plasmodium falciparum* 80S ribosome bound to the anti-protozoan drug emetine. *eLife* **3**.

Woodrow, C.J. and White, N.J. (2017) The clinical impact of artemisinin resistance in Southeast Asia and the potential for future spread. *FEMS Microbiology Reviews* **41**(1) 34–48.

World Health Organization (2016) *World Malaria Report 2016*, Geneva, Switzerland: World Health Organization.

Wu, Y., Sifri, C.D., Lei, H.H., Su, X.Z. and Wellems, T.E. (1995) Transfection of *Plasmodium falciparum* within human red blood cells. *Proceedings of the National Academy of Sciences* **92**(4) 973–977.

Wu, Y., Wang, X., Liu, X. and Wang, Y. (2003) Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite. *Genome Research* **13**(4) 601–616.

Yamagishi, Y., Honda, T., Tanno, Y. and Watanabe, Y. (2010) Two histone marks establish the inner centromere and chromosome bi-orientation. *Science* **330**(6001) 239–43.

Zacchi, P., Sblattero, D., Florian, F., Marzari, R. and Bradbury, A.R.M. (2003) Selecting open reading frames from DNA. *Genome Research* **13**(5) 980–990.

Zhang, C., Gao, H., Yang, Z., Jiang, Y., Li, Z., Wang, X., Xiao, B., Su, X., Cui, H. and Yuan, J. (2017) CRISPR/Cas9 mediated sequential editing of genes critical for ookinete motility in *Plasmodium* yoelii. *Molecular and Biochemical Parasitology* **212** 1–8.

Zhang, C., Xiao, B., Jiang, Y., Zhao, Y., Li, Z., Gao, H., Ling, Y., Wei, J., Li, S., Lu, M., Su, X. -z., Cui, H. and Yuan, J. (2014) Efficient editing of malaria parasite genome using the CRISPR/Cas9 system. *mBio* **5**(4) e01414-14.

Zhang, N., Hartig, H., Dzhagalov, I., Draper, D. and He, Y.W. (2005) The role of apoptosis in the development and function of T lymphocytes. *Cell Research* **15**(10) 749–769.

Zhu, L., Mok, S., Imwong, M., Jaidee, A., Russell, B., Nosten, F., Day, N.P., White, N.J., Preiser, P.R. and Bozdech, Z. (2016) New insights into the *Plasmodium vivax* transcriptome using RNA-Seq. *Scientific Reports* **6**(1) 20498.

References