Novel malaria parasite proteins involved in erythrocyte invasion

Claire Hollie Hastings

Division of Parasitology
MRC National Institute for Medical Research
The Ridgeway
Mill Hill, London
NW7 1AA

Department Infection and Immunity
University College London

This thesis is submitted in part fulfilment of the requiremnts of
University College London for the degree of Doctor of Philosophy

December 2012

I, Claire Hollie Hastings, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Claire H Hastings

December 2012

Abstract

Erythrocyte invasion is a key step in the *Plasmodium* life cycle. This process is tightly regulated, involving the sequential release of specialised apical secretory organelles – the micronemes, rhoptries and dense granules. These organelles contain proteins required for invasion and establishment of the parasitophorous vacuole, but most of the proteins remain uncharacterised.

The aim of this project was to uncover novel proteins with a role in invasion by the human malaria parasite *Plasmodium falciparum* merozoites. I identified proteins using the following selection criteria: a) expression in the schizont/merozoite form of the parasite; b) conservation across the genus; c) the presence of a signal peptide and d) one or more transmembrane (TM) domains. A list of 64 proteins was identified, and filtered further based on novelty, presence in the merozoite proteome, expression in other life cycle stages, and difficulty of study.

Five proteins were selected, and I produced recombinant protein and raised antibodies against three, which I used to identify the sub-cellular location of the protein within the parasite. The proteins appear to reside in either the rhoptries or the endoplasmic reticulum of the merozoite. Attempts were made to epitope-tag and delete all 3 genes, with a focus on one protein, the type IV Hsp40, PF11_0443. This protein contains two TM domains and is expressed during schizogony. By immunofluorescence it is present in the ER of early schizonts, before accumulating at the apex of merozoites in a rhoptry location. Immunoprecipitation experiments indicated that the protein binds known rhoptry proteins and other chaperones. The protein has been epitope-tagged but attempts to delete the gene by genetic recombination were unsuccessful. The gene is conserved in *Plasmodium* spp. and there are orthologues in higher eukaryotes, but it is absent from other Apicomplexa. Current studies are focused on the role of this protein in erythrocyte invasion.

Acknowledgements

Firstly I would like to thank my supervisor Dr. Tony Holder for accepting me into his lab and for continued guidance and support throughout my project. I would particularly like to express my sincere gratitude to Dr. Ellen Knüpfer for the constant guidance, lengthy discussions and frank advice; my time at the NIMR wouldn't have been the same without you. I would also like to thank Ellen for the use of her transfection constructs and antibodies; to thank Murina Grainger for the parasites she prepared; and to Dr. Steve Howell for his help and advice with mass spectrometry, all of which were crucial to the project's success. I would like to thank my thesis committee, Dr. Mike Blackman, Dr. Douglas Young and Dr. John Doorbar for their time and helpful advice and to acknowledge the Medical Research Council for funding this research.

I would like to thank Oniz Suleyman for her support in the early days of my project, for demonstrating new techniques and for all the fun times. To Dr. Rob Moon (aka Bobby, Grandad and Columbo), thanks for always being happy to listen, help out and give advice, even if it was sarcastic. To Andrea, thanks for your undying support, infectious enthusiasm and for always making time to listen. To the rest of the magnificent seven, Natalie, Mike and Matthew, thanks for the great discussions, great themed dinners and great times. I would like to thank Judith, Sola and Madhu for their pearls of wisdom and Robert Stallmach for his useful insights into my project. Thanks to Dave, Pete and Lasse from Molecular Structure for all your advice on protein biochemistry. To my fellow Ph.D students – Vicky, Lasse, Aylin and Joe – thanks for sticking around for the 4 years and being on hand for a trip to the bar when necessary. Azian and Ridzuan – the lab (and lunchtime!) wouldn't be the same without you and can't wait to visit you in Malaysia! Thanks to the NIMR volleyball and football teams, Fever and Sex Lethal, for giving me an excuse to take a break and enjoy the sunshine as well as keeping fit.

Thanks to my gran, dad, mother and brother for their encouragement and support in everything I've done. I would like to acknowledge that I have the best friends in the world, thank you for your constant support and friendship. To all the members, past and present, of the London Rockets and Sparks, thanks for the constant distraction and getting me through the low points in past four years, in particular J Loh, Jennie, Issy, Astrid, Tracy, Erin, Julia as well as Haslam, Hannah, Crockers, Anna, Steph, Spider and Nikki. If Carlsberg made friends...

To John, for your undying love, hugs, support and unchallenged ability to cheer me up, thank you doesn't seem enough.

Contents

List of figures

Index of Tables

Index of Appendices

List of Abbreviations

A, C, G, T	adenine, cytosine, guanine, thymine
ABRA	acid basic repeat antigen
AMA-1	apical membrane antigen-1
ATP	adenosine triphosphate
BC	before christ
BiP	immunoglobulin heavy chain binding protein
bp	base pairs
BSA	bovine serum albumin
BSD	blasticidin
CDPK	calcium dependent protein kinase
cm	centimetres
CSP	circumsporozoite protein
CYP	cyclophilin
DAPI	4',6-diamidino-2-phenylindole
DBL	duffy binding-like protein
dH$_2$0	distilled H$_2$0
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DDT	dichlorodiphenyltrichloroethane
dhfr	dihydrofolate reductase
DTT	dithiothreitol
EBA	erythrocyte binding antigen
EBL	erythrocyte binding ligand
EGF	epidermal growth factor
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol bis(2-aminoethyl ether)-*N*,*N*,*N*'*N*'-tetraacetic acid
ER	endoplasmic reticulum
g	grams
g	relative centrifugal force
GAP 40/45/50	gliding/glideosome associated protein 40/45/50
GFP	green fluorescene protein
GPI	glycosylphosphatidylinositol
h	hour
HCl	hydrochloric acid
HRP	horse radish peroxidise
Hsp	heat-shock protein
IFA	indirect immunofluorescence assay
IgG	immunoglobulin G
IMC	inner membrane complex
IRS	indoor residual spray
ITN	insecticide treated nets
IPTG	isopropyl-β-D-thiogalactopyranoside

kb	kilobase pairs
kDa	kilodaltons
LB	Luria-Bertani
LC-MS/MS	liquid chromatography-tandem mass spectrometry
M	molar
MCS	multiple cloning site
MES	2-(N-morpholino)ethanesulphonic acid
MESA	mature-parasite-infected erythrocyte surface antigen
mg	milligrams
µg	micrograms
ml	millilitres
µl	microlitres
µm	micrometres
mM	millimolar
µM	micromolar
MSP	merozoite surface protein
MTIP	myosin A tail interacting protein
MTRAP	merozoite thrombospondin-related anonymous protein
MyoA	myosin A
ng	nanograms
Ni-NTA	nickel-nitrilotriacetic acid
nM	nanomolar
OD	optical density
PBS	phosphate buffered saline
PBST	phosphate buffered saline with tween 20
PCR	polymerase chain reaction
PDI	protein disulphide isomerase
PEXEL	*Plasmodium* export element
PFA	paraformaldehyde
PfEMP1	*P. falciparum* erythrocyte membrane protein 1
PfRh	*P. falciparum* reticulocyte binding protein homologue
pi	post-invasion
PKA	protein kinase A
PKG	protein kinase G
PTRAMP	*Plasmodium* thrombospondin-related apical membrane protein
PV	parasitophorous vacuole
PVM	parasitophorous vacuole membrane
RAP	rhoptry-associated protein
RBC	red blood cell
RESA	ring expressed surface antigen
RhopH1/2/3	high molecular weight rhoptry protein 1/2/3
RNA	ribonucleic acid
ROM	rhomboid protease
RON	rhoptry neck protein
rpm	revolutions per minute

RPMI	Roswell Park Memorial Institute 1640
SDS	sodium dodecyle sulphate
SDS-PAGE	sodium dodecyle sulphate-polyacrylamide gel electrophoresis
SERA	serine repeat antigen
SSC	sodium chloridesodium citrate
SUB	subtilisin
TAE	Tris-acetate-EDTA
TK	thymidine kinase
TRAP	thrombospondin-related anonymous protein
TSR	thrombospondin repeat domain
UTR	untranslated region
UV	ultra violet
WT	wild type
WHO	World Health Organisation

1. Introduction

1.1 Malaria and global health

With an estimated 655,000 deaths, 216 million clinical episodes and over 3 billion people at risk in 2010 (WHO, 2011), malaria has a devastating impact on global health. Harrowing as these figures are, a recent study by Murray *et al.* (2012) suggests the actual number of annual deaths may be double that published in the WHO report, estimating the total number of global deaths due to malaria in 2010 at 1,238,000 (Murray *et al*., 2012). Malaria is primarily a disease of the tropics due to the habitual constraints of the definitive host and vector, the *Anopheles* mosquito, and its effects are more drastic in developing nations. An estimated 81% of clinical episodes and 91% of mortalities were in the African continent in 2010 however 99 countries worldwide present ongoing transmission (figure 1.1), including many regions across South America, India and South East Asia. Children under 5 years of age are most at risk of fatal outcome, accounting for 86% of deaths due to malaria (WHO, 2011). Other major risk groups include pregnant women and 'non-immune' adults or travellers. The disease varies in severity and can be categorised according to the symptoms the patient presents: uncomplicated malaria patients present fever and mild malaria anaemia; complicated and severe malaria symptoms additionally include severe anaemia, respiratory distress, metabolic acidosis, organ failure, mortality; and Cerebral Malaria, which is the most fatal presentation of the disease, and results in severe neurological symptoms, coma and death (WHO, 2000). As well as morbidity and mortality, there is a socio-economic cost associated with the disease. With the expense of treatment and the loss of a workforce through sick-leave, countries where malaria is endemic rarely attract foreign investment and face reduced economic growth as a result (Gallup and Sachs, 2001), which is precisely what these countries need to implement the necessary infrastructure which will ameliorate the burden of malaria.

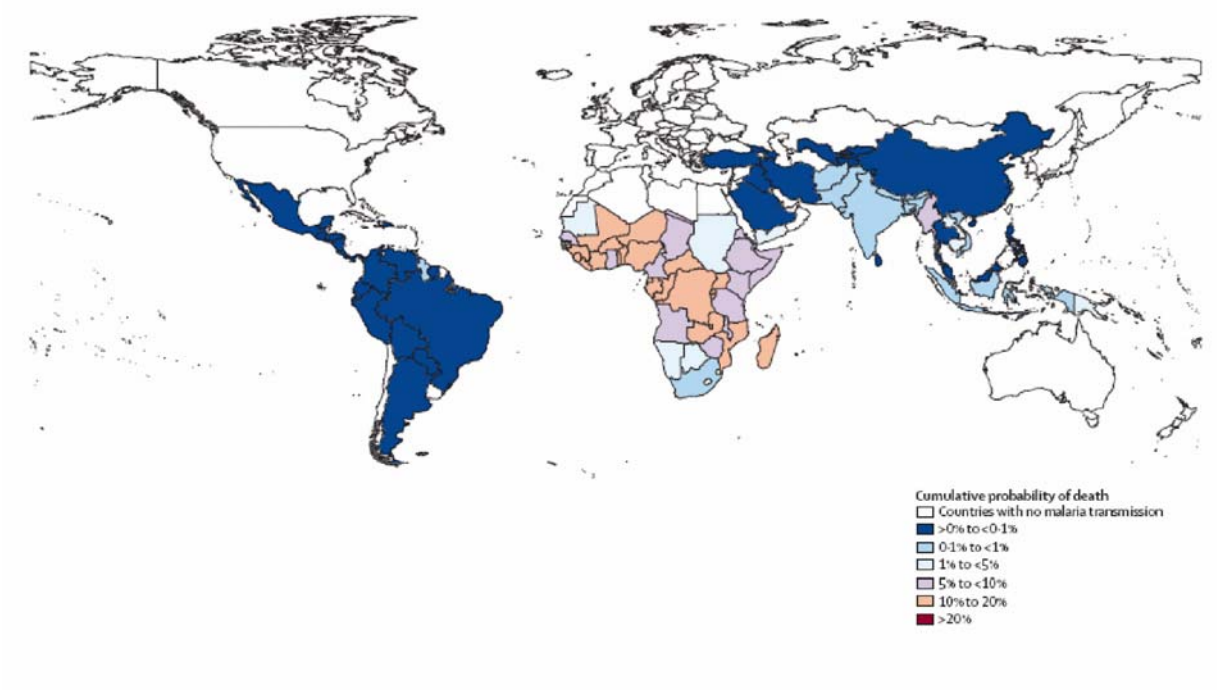


Figure 1.1: *Global likelihood of death due to malaria in 2010.* World map depicts the cumulative probability of death due to malaria in the absence of other cause from birth until aged 80 years in each country in 2010. Countries are colour coded based on percentage probability: 0-0.1% - deep blue; 0.1-1% - mid-blue; 1-5% - pale blue; 5-10% - pale purple; 10-20% - pale pink. No countries presented greater than 20% chance of death due to malaria. Countries with no transmission are pictured in white. While the majority of deaths occur in sub-Saharan Africa, large parts of South and South East Asia as well as Indonesia and south pacific islands face a substantial threat. Although South America experiences a large volume of cases of malaria, it is more usually caused by *P. vivax* which is rarely fatal. Adapted from (Murray *et al*., 2012).

1.2 Malaria: a history in brief

Malaria has been the scourge of the tropics for millennia and was most likely present in the earliest human populations (Brier, 2004). Paleopathological studies of mummified human remains from ancient Egypt identified several specimens exhibiting porotic hyperostosis (cranial thickening which is indicative of severe anaemia) and splenomegaly (enlarged spleen), common symptoms of malaria (Brier, 2004). As the field advanced, molecular biology techniques were able to validate the presence of malaria in ancient Egypt: a genetic study by Hawass *et al.* (2010) revealed King Tutankhamun (c. 1333-1324 BC) suffered from malaria although whether this caused his early death remains controversial (Hawass et al., 2010; Timmann and Meyer, 2010). Malaria is thought to have spread out of Africa along shipping routes to India, South East Asia, China and also parts of Europe. Hypocrates (460 – 370 BC) of ancient Greece and Celsus (25 BC – 54 AD) of ancient Rome both provided written accounts of the clinical presentations of what we now know as malaria. It was the Romans who named the disease '*mal aria*' meaning 'bad air', as they attributed the fevers to noxious vapours arising from the stagnant waters of the marshlands surrounding Rome. The draining of the marshes and consequent reduction of malaria episodes gave weight to their hypothesis but it would be nearly 2 millennia before the true cause of the disease was uncovered [reviewed in (Cunha and Cunha, 2008)].

Although several theories were put forward since the fall of the Roman Empire, few major scientific break-throughs were made regarding malaria until 1880, when a French clinician named Alfonse Laveran working in Algeria noted the presence of black pigment in the spleen and brain of his patients that died of malaria. He consequently studied the blood of malaria patients until finally witnessing the exflagellation of a gametocyte. This confirmed his suspicion that the disease was of parasitic origin, and not bacterial as had previously been suggested (Garnham, 1966). The observation of what was termed 'filariae of the blood' until later classified under the genus *Plasmodium*, as the causative agent of malaria, kick-started a new age of malaria research. The discovery of the parasite led to the question of transmission and research into the vector

responsible was top priority. Although suspected throughout history by doctors and philosophers from different nations, there was no evidence to link the occurrence of malaria with the mosquito. That was until Surgeon-Major Ronald Ross, aided in his observations by Dr. Patrick Manson, in 1897 identified the presence of the malarial parasite in the midgut of a particular and rare form of mosquito (later ascribed to the genus *Anopheles*) fed on the blood of malaria patients. The parasite was absent in the more common, grey coloured mosquito fed on malarial blood and also in *Anopheles* mosquitoes fed on healthy individuals. Ross also observed the presence of 'germinal rods' (sporozoites) in the salivary glands of mosquitoes and concluded the mosquito bite to be the method of host transfer of the parasite (Ross, 1897; Manson, 1898; Ross, 1898). Although Ross was able to demonstrate transmission from insect to host in birds with avian malaria, the first experimental evidence of mosquito-to-man transmission was provided in 1899 in Rome by Bignami and Grassi who were able to follow the parasite's lifecycle to completion and their experiments were repeated and confirmed by Dr. Patrick Manson in London a year later (Grassi and Noe, 1900; Manson, 1900). Since then, efforts in mosquito control including destruction of mosquito breeding grounds, distribution of bed nets and insecticide sprays, have resulted in a reduction of incidents of malaria in the developing world and eliminated malaria from much of the developed world where countries boast the infrastructure and capital to effectively control the spread of disease. Major break-throughs in scientific research such as the adaptation of *Plasmodium falciparum* to long term *in vitro* culture (Trager and Jensen, 1976), the sequencing of the *P. falciparum* genome (Gardner et al., 2002) and development of various animal models have aided drug and vaccine research immensely. Despite all this, malaria persists and devastates the tropics to this day.

1.3 How do you solve a problem like malaria?

Any strategy for control and eventual eradication of such a complex and multifaceted disease as malaria must encompass multiple and complementary methods, and should be approached with extreme caution. The WHO's Global Malaria Eradication Program,

adopted in 1955, saw the mass application of DDT to eliminate mosquitoes (WHO, 1956). The program succeeded in eliminating the disease from Europe, North America and Australia however, due to the publicly controversial and adverse environmental side effects as well as the eventual emergence of resistance to the insecticide, the programme was abolished in 1972. As a result of the failure to completely disrupt transmission in sub-Saharan Africa, malaria returned to equivalent or greater levels than before the control strategy was implemented.

With an injection of funding from the Bill and Melinda Gates Foundation, there has been a renewed drive for eradication of malaria in recent years from charities and other organisations, primarily focussing on controlling the mortality and morbidity of the disease. In 2011, after a 2 year consultation period, the newly established Malaria Eradication Research Agenda (malERA) set out to complement the ongoing work of charitable organisation and government initiatives but with the goal of reducing the reproductive rate of *Plasmodium* to less than 1 and ultimately removing it form the human population (Alonso et al., 2011) These programs exist to consolidate and strategise the best use of all the anti-malarial weapons in our armoury and to ensure new tools are being developed where needed. Current methods of malaria prevention and treatment include vaccines, anti-malarial drugs, vector control and improving infrastructure and monitoring.

1.3.1 Vaccines

Regardless of its aptitude for immune evasion, natural immunity to *Plasmodium* is acquired over time, increasing in efficiency with each exposure (Day and Marsh, 1991). This non-sterile immunity allowing a low level or parasitaemia to persist in the absence of symptoms is called premunition. Passive transfer of anti-*Plasmodium* antibodies from hyper-immune individuals ameliorates disease and confers protection to non-immune adults (Sabchareon et al., 1991). Taken together, this suggests a successful vaccine against the parasite is feasible and would represent a beneficial boost to our immune defence. Despite being commercially unattractive, a malaria vaccine could be one of the

most cost effective methods of eradicating malaria. Developing a successful vaccine is the major goal of several laboratories worldwide and has been for over 40 years; however this has proved to be more difficult than first anticipated. Vaccine candidates can be split into three groups: those aimed at blocking infection targeting the sporozoite stage parasites before they enter the host liver; disease blocking vaccines targeting blood stage parasites which are responsible for disease symptoms; and transmission-blocking vaccines that target the sexual stage of the parasite, preventing development in the mosquito. Transmission blocking vaccines are the least developed and ethically difficult to promote as the vaccine will not benefit the vaccinee in the short term as it will not protect from disease. It also requires high coverage of community participation to break the transmission cycle from the human to the mosquito. Disease blocking vaccines targeting the blood stage aim to prevent severe malaria and death without preventing the initial infection. A large number of potential antigens have been put forward as vaccine candidates for a disease blocking vaccine. Merozoite surface proteins, in particular merozoite surface protein (MSP)-1 were initially of interest due to their abundance. Other candidate antigens include MSP-2, -3, -4, -5, -8, -9, serine repeat antigens (SERAs), apical membrane antigen (AMA)-1, glutamate rich protein (GLURP), the erythrocyte binding antigens (EBAs) and the Rh family proteins. Although promising, due to polymorphisms, strain variability, multiple invasion pathways and the lack of correlation between animal models and protection in the field, few of these candidates have progressed in clinical trials. Infection blocking vaccines are by far the most successful to date. As early as 1967, Nussenzweig *et al*.demonstrated in mice that immunisation with irradiated sporozoite stage parasites conferred protection from subsequent infection (Nussenzweig *et al*., 1967). Today, the infection blocking vaccine RTS,S, produced by GlaxoSmithKline (GSK) and the Walter Reed Army Institute of Research, is the world's leading vaccine candidate. Comprised of the C-terminus of the circumsporozoite protein (CSP) fused to hepatitis B surface antigen in the form of virus like particles (VLPs), it is at present in phase III clinical trials however early results suggest it is delivering an efficacy of only around 50% in first time episodes (Agnandji *et al*., 2011).

1.3.2 Anti-malarial drugs

In a disease that has been around since antiquity, the discovery and development of effective remedies over time is to be expected. Quinine, extracted from the bark of the cinchona tree, was used by native South Americans to relieve fever and has been used in anti-malarial therapy since its introduction to Europe in the 17th century. Although effective, quinine produced undesirable side effects and after a long period of use, the parasite finally developed resistance in 1910 (Achan et al., 2011)). The drug chloroquine, a quinine derivative, was developed in the early 20th century and quickly replaced quinine as the drug-of-choice in the treatment of malaria. Not only was it an extremely effective anti-malarial, it was cheap to manufacture and therefore suitable for distribution amongst the developing world. Its success also led to its downfall, as heavy application eventually led to the emergence of resistance, first detected in *P. falciparum* around the Thailand-Cambodia border and in Colombia in the late1950s before spreading throughout the world over a 20 year period (Payne, 1987). Several drugs have been developed and distributed since then, however in time resistance has arisen to almost all of them. The current drug-of-choice in the treatment of malaria is artimisinin, derived from the white flowering plant *Artemisia annua*. It was first noted to display anti-malarial properties in 340 B.C. by the alchemist Ge Hong of the East Jin Dynasty. Over 2 millennia later, the active ingredient derived from this plant was identified and produced in tablet form. In order to delay the emergence of resistance to the drug, Professor Zhou Yiquing of the Chinese Academy of Military Medical Sciences developed the first artemisinin-based combination therapy (ACT), Coartem, which incorporates an artemisinin derivative, artemether in combination with lumefantrin (Weiyuan, 2009). Unfortunately the first case of artemisinin resistance was reported in Cambodia in 2009 (Dondorp *et al*., 2009). Although the WHO, local governments and charities endeavoured to prevent it, artemisinin resistance has reached Myanmar (Phyo *et al*., 2012), and authorities now fear its dispersal through to India and the rest of the world is inevitable.

1.3.3 Vector control

Vaccines boost natural immunity to a disease and drugs treat the disease which has already taken its toll on the body, but perhaps the most thorough method of preventing the problem of malaria is to remove the agent of its dispersal. A programme of indoor residual spraying (IRS) with pyrethroid insecticides in households combined with long lasting insecticide-impregnated nets (LLINs) has been employed in a number of malaria endemic countries (WHO, 2010). Although this may be effective in controlling malaria in areas of low-moderate transmission, areas of high transmission will require additional control mechanisms to prevent the disease (Control, 2011; malERA_Consultative_Group_on_Vector_Control, 2011). These control mechanisms are preventative only in areas where pyrethroid susceptible mosquitoes predominate, several species of *Anopheles* are now resistant to these insecticides (Ranson *et al*., 2009). Also, a number of Anopheline species feed and rest outdoors and are therefore not targeted by IRS and LLINs. Research and development of new insecticides as well as application strategies to include outdoor feeders should be a priority if malaria transmission is to be controlled in areas of high endemicity. In such areas, it would be prudent to employ additional, long term approaches to decrease or halt transmission. Genetic approaches are being developed to reduce the capacity of natural *Anopheles* populations to act as vector for the parasite. This includes vector replacement strategies using mosquitoes which are resistant to Plasmodium or those which are killed upon infection with *Plasmodium spp.* (Terenius *et al*., 2008). Other novel approaches to vector management target the mosquito environmentally (e.g. eliminating breeding grounds), biologically (e.g. niche competitors, insect pathogens, plasmodicidal symbiotic organisms) and chemically (e.g. new effective insect repellents,).

1.3.4 Infrastructure, Monitoring and Diagnostics

Frustrating though it may be to a scientist, the sad fact remains that the countries worst affected by malaria are those which lack the proper infrastructure to deal with it; a problem antagonised by tyrannic regimes, corruption and civil war. Something as simple as better housing design that limits mosquito entry would have a dramatic impact on transmission. Access to antimalarial drugs is another limiting factor. In rural areas of sub-

Saharan Africa patients are required to travel miles to reach a clinic and clinics struggle to keep up drug reserves due to problems with delivery. Patients require rapid diagnosis and application of treatment if they are to survive the infection; efficient health systems need to be in place across all malaria endemic countries in order to support a permanent block in malaria transmission. As the economies of the worst effected countries are unlikely to change while malaria is rife, charitable organisations and international aid must be relied upon to introduce the necessary infrastructure to combat malaria.

1.3.5 Combine and conquer?

Elimination is not achievable at the individual country level; if eradication was achieved in one country, any bordering countries with moderate malaria transmission would inevitably re-introduced malaria via migrants and travellers. Eradication must be a global effort, using strategic combinations of tools to permanently interrupt transmission of the parasite. Education and training regarding drug administration as well as schemes to ensure drug supplies remain stable will enable early diagnosis and treatment, limiting the transmission of sexual stage parasites to mosquitoes and therefore the spread of malaria. Attacking both indoor and outdoor feeding mosquitoes with insecticide will limit the transmission to humans however finding a long term solution to *Plasmodium* susceptible mosquitoes is important. In the event of insecticide or antimalarial drug resistance, a now non-immune population would be exposed to a pathogen they are no longer prepared to fight. Vaccines will be necessary here, as if malaria resurges in insecticide resistant mosquitoes or with the reintroduction of susceptible mosquitoes, vaccination will ensure a certain amount of immunity persists. Animal reservoirs are difficult to control and must also be eliminated if a complete block in transmission is to be achieved. Although western lowland gorillas in Africa are capable of transmitting *P. falciparum* evidence suggests this is not a regular occurrence. The simian malaria *P. knowlesi* is readily transmitted to humans in SE Asia from Macaques however there is limited evidence of human to human transmission. With thorough monitoring, it is hoped that these zoonoses will not present a significant threat to global malaria eradication.

Research into the parasite, vector and host biology is absolutely necessary to develop the highest quality tools to combat malaria. Thus far, the development of an effective vaccine remains elusive. With the continuing spread of drug and insecticide resistance, research into novel vaccine candidates is vital to the success of eradicating the disease. At present and until artemisinin resistance spreads, ACTs are the best method of treatment for those who have access or can afford it, however for the areas most in need cost is a major issue and inexpensive alternatives must be found. Current anti-malarial drugs are based a small number of chemical structures that act on a relatively few targets and further research should aim to expand this list. The ideal drug, as set out by the malERA initiative, should be: able to provide a single encounter radical cure and prophylaxis (SERCaP); effective against all human infecting species of *Plasmodium* including *P. vivax* long-lived hypnozoites; kill 100% of parasites in an infected person; suitable for mass administration (includes cost effectiveness and safety record); and able to be used as prophylaxis for a period that will see the effects outlive the mosquito stage of *Plasmodium* lifecycle (minimum of1 month post treatment) (malERA_Consultative_Group_on_Drugs, 2011). This mythical drug may seem like the stuff of fiction but with continuing research into the right target, who is to say that it will not be brought to life?

1.4 Introducing the protozoan problem parasite, Plasmodium

Malaria is caused by the protozoan micro-parasite *Plasmodium*, belonging to the phylum Apicomplexa, within the superphylum Alveolata (depicted in figure 1.2). This group of unicellular eukaryotes contains a number of perilous pathogens of medical and veterinary (and therefore economic) importance. One such Apicomplexan, *Toxoplasma gondii*, commonly infects humans following contagion from the faeces of feline house-pets or by consumption of infected livestock. Due to their penchant for under-cooked meats, up to 60% of the inhabitants of Southern Europe are positive for *T. gondii* (Pappas *et al.,* 2009). This infection has no obvious side effects unless contracted during pregnancy (Desmonts and Couvreur, 1974); or contracted by the immuno-compromised

and is therefore a major complication during the onset of acquired immunodeficiency syndrome (AIDS) (Luft *et al.*, 1993). *T. gondii* has been referred to as the 'model' Apicomplexan due to its superior genetic tractability compared with *Plasmodium* and is therefore easier to study (Kim and Weiss, 2004). Much of what we know of *Plasmodium* in terms of molecular biology was first discovered in *Toxoplasma*. Humans also play host to *Cryptosporidium spp.*, causing a severe gastroenteritis. *Neospora caninum* presents a very similar infection to *Toxoplasma* albeit with slightly different host preference – a canine instead of feline definitive host and where *Toxoplasma* is common is sheep, and *N. caninum* is prevalent in cattle and no infections have been recorded in humans unlike *Toxoplasma* (Dubey, 2003). Fellow coccidian *Eimeria* causes a gastro-intestinal infection in chickens (reviewed in (Allen and Fetterer, 2002). *Theileria* and *Babesia spp.* are a particular pest to agriculture, causing serious disease in cattle and resemble *Plasmodium* in as much as merozoite stage parasites infect erythrocytes, (Shaw, 2003; Yokoyama *et al*., 2006)

There are in existence approximately 450 species of *Plasmodium*, each infecting a diverse group of organisms (Perkins and Schall, 2002). Numerous genera of mammals, birds and reptiles are susceptible to particular species of *Plasmodium*. While avian malaria can be transmitted by several different genera of mosquitoes, *Anopheles* mosquito is solely responsible for transmission and spread of the parasite in mammals. In humans, *Anopheles gambiae, An. arabiensis, An. melas* and *An. merus* are the 4 main dominant vector species responsible for the majority of transmission in sub-Saharan Africa and consequently the most devastating species of mosquito in terms of human life (Sinka *et al.*, 2010). Four species of *Plasmodium* primarily infect humans; *P. vivax, P. ovale, P. malariae* and *P. falciparum*; the latter being responsible for the majority of fatalities due to malaria worldwide. The other species cause considerable morbidity however infections are rarely fatal although *P. vivax* malaria is particularly rife in South America and can cause recurrent bouts of disease from a single infection. A fifth human-infecting species, the simian malaria *P. knowlesi*, has been discovered in rural areas of Malaysia and other parts of SE Asia to infect humans and although only recently revealed to be a human pathogen, it has most likely been misdiagnosed in the past,

mistaken for *P. malariae* due to morphological similarities despite failing to test positive for *P. malariae* by PCR (Singh *et al*., 2004). Unlike the other *Plasmodium* spp.*, P. knowlesi* has a 24 hour lifecycle therefore parasite burden will increase more rapidly and febrile episodes associated with red blood cell rupture and parasite exit will become more frequent, resulting in a more aggressive infection. Although *P. knowlesi* research is in its infancy, early evidence suggests the mortality rate is comparable to that of *P. falciparum* of those positively diagnosed (Cox-Singh *et al*., 2008).

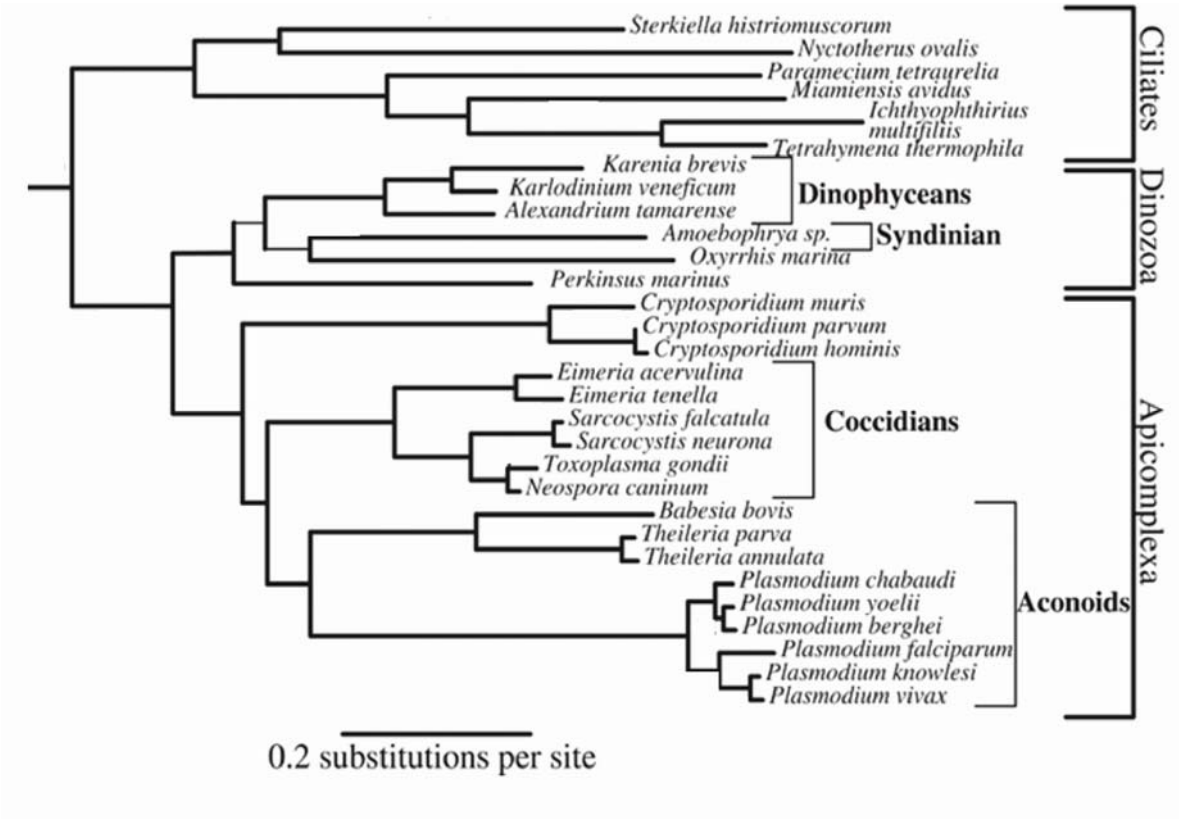


Figure 1.2: *Alveolate phylogenetic tree*. The Phylogenetic tree is based on amino acid sequence alignments from 17 ribosomal proteins. The 3 major alveolate lineages – Ciliates, Dinozoa and Apicomplexa are highlighted with brackets. Adapted from (Bachvaroff *et al*., 2011).

1.4.1 What's the problem with Plasmodium?

A parasite should not kill its host. It should keep its host alive for as long as possible in order to increase its transmission potential and spread, and therefore its success. So why is a disease that has been around since antiquity responsible for such a large number of deaths? *Homo sapiens* take at least 16 years to reach reproductive age therefore evolution on the side of the host is comparatively slow. Mutations in host haemoglobin are common in areas of high transmission however they, with the possible exception of the Haemoglobin E trait, confer a health disadvantage to the human host with varying severities of anaemia. Living with these haemoglobinopathies is almost as much of a burden as malaria itself, with 300,000 babies born with these genetic disorders every year (Weatherall *et al.*, 2010). Erythrocyte surface proteins are also subject to selection. Heterozygotes for band 3 mutation that causes ovalocytosis are protected from cerebral malaria however homozygotes are non-viable and result in miscarriage (Jarolim *et al*., 1991; Genton *et al*., 1995; Allen *et al*., 1999). Approximately 95% of the West African population are homozygous for a point mutation in the promoter of the Duffy antigen that results in a complete block of its expression. Given that the presence of this antigen is essential to permit *P. vivax* infection, West Africa is free from *Plasmodium vivax* malaria (Miller *et al*., 1976).

The laws of natural selection would suggest the parasite would evolve to become less virulent and although *P. vivax, P. ovale* and *P. malariae* cause significant morbidity they rarely result in death, in contrast to *P. falciparum* which is often fatal. *P. falciparum* must possess an additional virulence factor not present in the other species. One possible reason for this difference in disease severity is thought to be the phenomenon of cytoadherence of parasite-infected erythrocytes to the blood vessel endothelium. During the intra-erythrocytic stage of the *P. falciparum* lifecycle the parasite exports its own proteins onto the red blood cell surface and in the case of *P. falciparum* the major surface molecule is erythrocyte membrane protein (PfEMP)-1. The parasite uses this ligand to cytoadhere to the blood vessel endothelium, allowing it to avoid clearance by the spleen. In doing so *Plasmodium* prolongs its life but also clogs the microvasculature,

restricting blood flow to organs of the body. Parasites can cytoadhere to the brain endothelium which is believed to contribute to cerebral malaria. In the worst cases this results in coma and death. *P. falciparum* is capable of attachment to any of eleven endothelial receptors identified to date. This binding capacity varies between different strains and seems to be responsible for strain virulence - attachment to CD36 is common to all parasites and associated with uncomplicated malaria whereas binding of intercellular adhesion molecule (ICAM)-1 is elevated in severe and cerebral malaria (Newbold et al., 1997; Silamut et al., 1999; Ochola et al., 2011). PfEMP-1 also binds multiple uninfected red blood cells, cloaking itself from immune detection and shielding itself from immune clearance as well as keeping fresh rbc targets in close proximity. This phenomenon is known as rossetting and the ability to form rosettes is associated with severe disease (Rowe et al., 1995). *P. falciparum* contains around 60 copies of the *var* genes which encode PfEMP-1 and is able to switch which singular one is expressed on the erythrocyte surface, evading host immunity (Roberts *et al*., 1992; Smith *et al*., 1995). Antibodies to one PfEMP-1 molecule will not necessarily protect against another therefore natural acquired protective immunity will only build up after repeated exposure (Bull *et al*., 1998). Cytoadherence is clearly an important virulence determinant in *P. falciparum* but what about other human-infecting species? Although *P. vivax* appears to possess sub-telomeric multi-gene families (Carlton *et al*., 2008), there is limited evidence of cytoadherence in vivax malaria and the molecular basis for pathology is poorly understood (reviewed in (Costa et al., 2011). Due to lack of fatal cases, *in vitro* culture and incomplete genome sequence; there is a lack of evidence to prove or deny cytoadherence of *P. ovale* and *P. malariae* however cerebral symptoms have not been reported for these malarias. There is one variant multi-gene family in *P. knowlesi*: the schizont-infected cell agglutination variant antigen (SICAvar) genes (al-Khedery et al., 1999). *P. knowlesi* infected erythrocytes have been demonstrated to cytoadhere *ex vivo* and there has been one fatal case where parasites were found in the brain however no cerebral malaria symptoms were reported (Fatih et al., 2012). This, taken together with the fact *P. knowlesi* may result in a comparable number of fatalities as *P. falciparum*, implies there is a correlation between cytoadherence ability and parasite virulence.

1.4.2 *Plasmodium* Life Cycle

Plasmodium has a complex lifecycle involving the invertebrate definite host, the mosquito, and a vertebrate intermediate host which can range from reptiles and birds to mammals. Focusing on human infecting species of *Plasmodium*, the parasite is transferred to its host by the female *Anopheles* mosquito. The sporozoite form of the parasite, temporarily lodging in the mosquito salivary glands, is deposited onto the skin with the mosquito's saliva during a blood meal. The sporozoite uses gliding motility to traverse the various layers of the dermis and actively penetrates into the microvasculature (Amino *et al*., 2006). The parasite travels with the blood stream, stalling at the liver where it exits the blood vessel passing through sinusoidal cells of the liver capillaries, and glides through several layers of hepatocytes before selecting one to actively invade (Mota *et al*., 2002). Host cell invasion requires the secretion of proteins from apical secretory organelles - the micronemes and rhoptries – and proteins contained in these organelles are required for attachment to, active invasion of, and establishment within the hepatocyte, forming a parasitophorous vacuole (PV) upon entry, in which it resides throughout this stage of the lifecycle. The parasite undergoes several rounds of asexual replication in hepatocytes, producing a hepatic schizont where large numbers of merozoites are produced. Merozoite stage parasites are able to invade red blood cells but first must be released safely from the liver. Although it was initially thought that merozoites were release upon hepatic schizont rupture, evidence from the murine malaria model, *P. berghei* reveals the formation of merosomes - large merozoite-packed vesicles - which bud-off from the infected hepatocyte after the PV has broken down (Sturm et al., 2006). This has never been confirmed in any of the human malarias due to ethical and experimental difficulty. *P. vivax* and *P. ovale* produce a dormant stage hypnozoite during liver development that can result in relapse of the disease, however this is not seen in the other malarias (Krotoski et al., 1982; Cogswell et al., 1983).

Merozoites released from hepatocytes enter the blood stream where the parasite must rapidly invade a new host cell to escape immune detection (discussed in detail in section 1.5). The parasite selects which host cell to invade via interactions of merozoite surface

proteins (MSPs) with red blood cells (rbcs). Similar to invasion of hepatocytes by sporozoites, apical secretory organelles release proteins required for invasion onto the surface of the merozoite as well as into the rbc. A parasite molecular motor drives invasion, creating a PV upon entry through secretion of the contents of the rhoptries. *P. falciparum* develops within its vacuole over a period of 48 hours in visually discrete stages. The young trophozoite or 'ring' stage parasite, named after its appearance on a Giemsa-stained blood smear, describes the period of the intra-erythrocytic lifecycle immediately post-invasion (p.i.) and the following 24 hours where the parasite will begin feeding and establishing itself within the rbc, exporting a number of proteins and membranous structures into the rbc cytoplasm and onto the rbc surface. Once the parasite is established within its host, it undergoes a period of aggressive feeding and growth. This is typically around 24-32 hours post-invasion in *P. falciparum* and is known as the trophozoite stage. As increasing amounts of haemoglobin are digested, the haem by-products within the parasite's food vacuole form haemozoin crystals that darken the appearance of the infected cells. The parasite will also begin to replicate its DNA at this stage, ready for later cell division. Schizont stage parasites are initially observed at around 32-34 hours post-invasion when nuclear division takes place. Daughter cell segmentation occurs at 45-48 h, completing cell division and producing up to 32 (more typically 16) merozoites (Bannister and Mitchell, 2003). Completion of replication is followed by a process known as egress, in which both PV and rbc membranes must be broken down in order to release daughter merozoites into the blood stream. Prior to release, organelles referred to as the exonemes discharge the subtilisin-like serine protease (SUB)-1 into the parasitophorous vacuolar space (Yeoh et al., 2007) which is responsible for the necessary proteolytic processing of MSP-1, 6 and 7 precursor proteins (Koussis et al., 2009). This 'priming' of merozoite proteins is not sufficient for release; a cascade of kinases has proven to be essential in this process including calcium-dependant protein kinase 5 (CDPK5) potentially acting first (Dvorin et al., 2010), followed by cGMP-dependant protein kinase (PKG) (Taylor et al., 2009) however the initial source of increased calcium levels required to activate CDPK5 is undefined and the effector molecule responsible for membrane rupture remains elusive. Upon release into the blood stream, primed and ready merozoites quickly identify new target cells via there

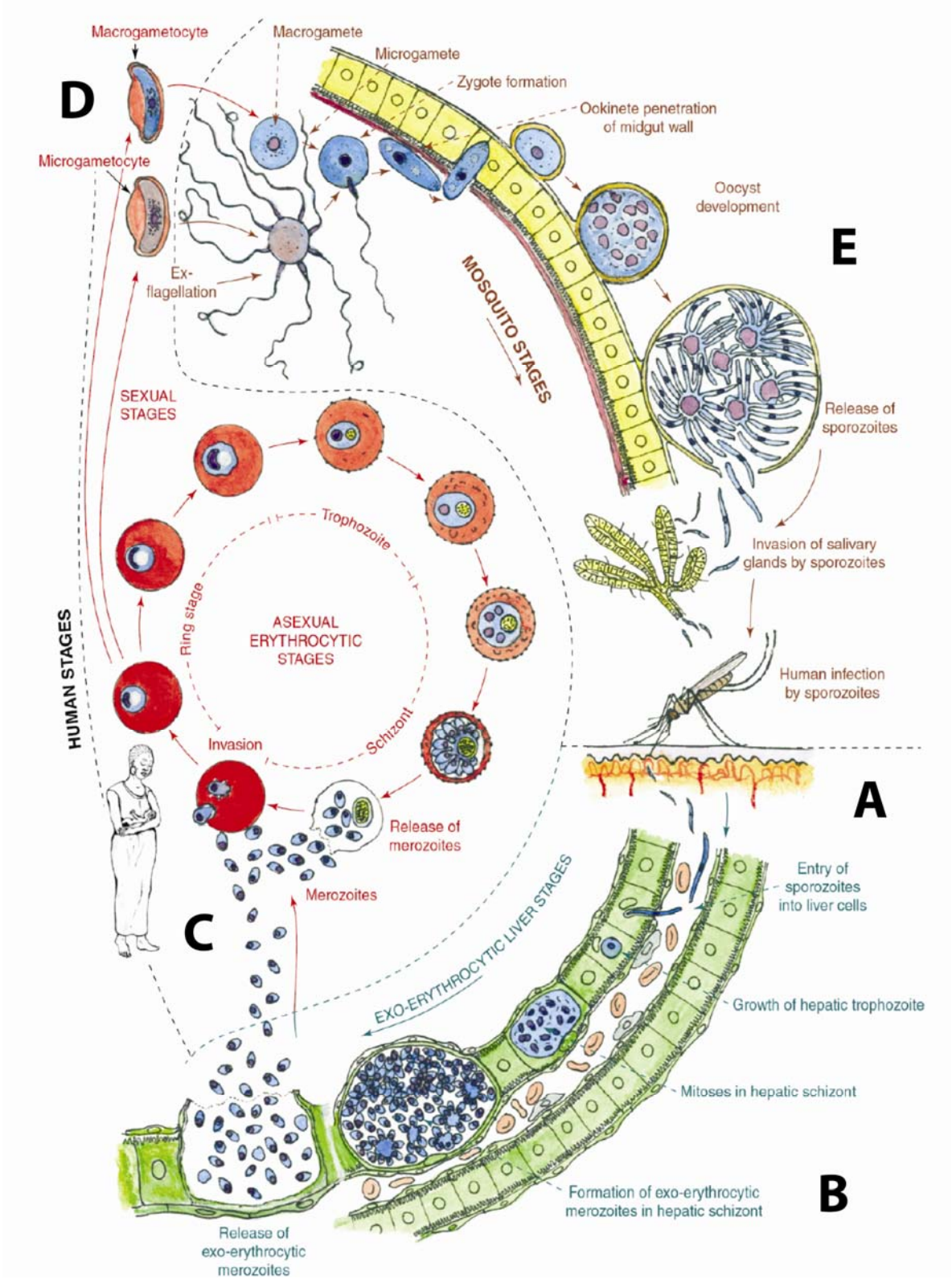
surface proteins and invade within 2 min of egress (Gilson and Crabb, 2009), completing one round of asexual development.

It is the blood stage of the *Plasmodium* lifecycle that is responsible for disease symptoms. The intra-erythrocytic cycle is synchronised to host circadian rhythm thus schizonts will rupture within hours of each other (Hawking et al., 1968). A bombardment of foreign material bursts into the blood stream resulting in an acute pro-inflammatory response and consequential fever. Fevers occur in multiples of 24h depending on the infecting species of *Plasmodium*. In more severe presentations of the disease, rbc rupture and diserythropoesis give rise to anaemia while cytoadherent infected rbcs obstruct bloods vessels, restricting blood flow to organs or parts of the brain (Marsh and Snow, 1997).

Continuation of the lifecycle following the trophozoite stage of intra-erythrocytic parasite development can take an alternative route to schizogony, producing female (macro) and male (micro) gametocytes. This commitment is prompted in the mother schizont, producing sexually committed merozoites which will invade the host erythrocyte as usual but intracellular development is now geared towards sexual stage production (Bruce et al., 1990; Silvestrini et al., 2000; Smith et al., 2000). *P. falciparum* gametocytes develop through stages I-V over a period of 8-12 days within the human host. Early gametocytes (stage I) are difficult to distinguish from the asexual trophozoite thus stage specific markers, such as the protein Pfs16 which is expressed 24 hours p.i. in sexually committed parasites, are extremely valuable (Bruce et al., 1994). Later stage gametocytes form the characteristic crescent shaped parasites, commonly associated with malaria in the late 19th century (Garnham, 1966). The 'trigger' for commitment to gametocytogenesis remains unclear although several stress stimuli from the surrounding environment such as host immune factors, haematological disruption, high parasite density, low nutrient levels and antimalarial drugs have all been implicated, reviewed in (Baker, 2010). Gametocytes are arrested at G0 until taken up by the *Anopheles* mosquitoes during blood meals where the change of environment (temperature, pH, salt concentration etc.) leads to rapid development in gametes.

Differentiated male and female gametes are released from the safety of the rbc, and are vulnerable to mosquito digestive enzymes and acids, as well as human immune factors injested with the blood meal. Gametes fuse to form a diploid zygote, which will develop into the invasive ookinete and traverse the midgut epithelium to escape the hostile environment it now finds itself in. This traversal is achieved by gliding motility and relies on microneme secretion for release of key adhesins essential for this process. The parasite traverses through various layers of the epithelium until reaching the basal lamina of the gut where it forms an oocyst (Vlachou et al., 2004). Sporogony occurs within the oocyst, developing and releasing thousands of sporozoites into the haemocoel (Canning and Sinden, 1973). Migration to, and active invasion of the salivary glands ensues, where sporozoites remain until deposition onto the skin of their human host, completing one round of the lifecycle. The lifecycle is summarised in figure 1.3 and each invasive stage is illustrated in figure 1.4.

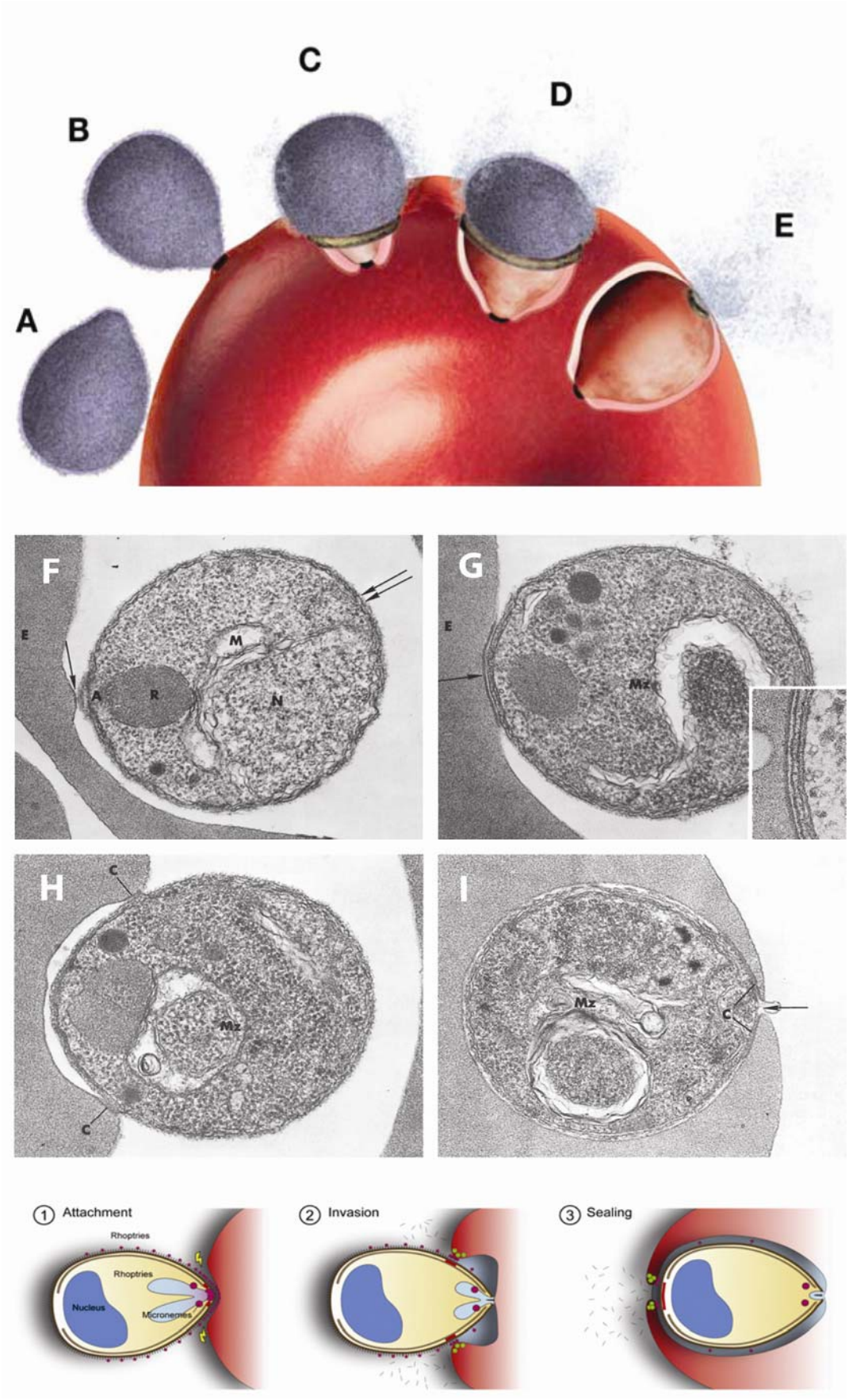
Figure 1.3 Plasmodium *lifecycle*. (A)Sporozoites are deposited in the skin during the mosquito's blood meal and from this site the parasite rapidly makes its way through the blood stream to the liver (B) where it invades and multiplies in hepatocytes, producing a large number of liver merozoites which eventually break free and re-enter the blood stream, where the parasite is now able to infect erythrocytes (C). The infecting merozoite develops through ring, trophozoite and schizont stages where it asexually produces up to 32 new daughter merozoites which are released back into the blood stream. Gametocytes are also produced in the blood stream and are taken up by *Anopheles* mosquito in a blood meal (D), where they are able to produce gametes that fuse and develop into ookinetes which traverse the midgut, forming an oocyst in the basal lamina. Thousands of sporozoites are produced which travel to and invade the salivary glands (E). These sporozoites will be deposited in the skin of the human source of the mosquito's next blood meal, continuing the cycle. Adapted from (Bannister and Mitchell, 2003).


Macrogametocyte
Macrogamete
Microgamete
Zygote formation
Ookinete penetration of midgut wall
D
Microgametocyte
Ex-flagellation
Oocyst development
E
MOSQUITO STAGES
SEXUAL STAGES
Release of sporozoites
Trophozoite
Invasion of salivary glands by sporozoites
Ring stage
ASEXUAL ERYTHROCYTIC STAGES
HUMAN STAGES
Human infection by sporozoites
Schizont
Invasion
A
Release of merozoites
Merozoites
C
Entry of sporozoites into liver cells
EXO-ERYTHROCYTIC LIVER STAGES
Growth of hepatic trophozoite
Mitoses in hepatic schizont
Formation of exo-erythrocytic merozoites in hepatic schizont
B
Release of exo-erythrocytic merozoites

1.5 Erythrocyte invasion

Plasmodium belongs to a group of Apicomplexa of the blood, order Haemosporidia. The 'choice' of the erythrocyte is a smart tactical move by the parasite. *Plasmodium* spends the majority of its stay in its mammalian host inside this immune-privileged cell and is therefore 'invisible' to a large faction of the immune system. Although a crucial step in propagation of the parasite, exiting the erythrocyte creates cell damage and immune 'danger' signals therefore it is critical that the parasite be removed from the treacherous extra-cellular environment of the blood stream and that re-invasion of a new erythrocyte occurs rapidly to limit the exposure of the parasite's surface proteins to the immune system, prolonging its life within the host. Control of erythrocyte invasion is tightly regulated and involves the sequential release of the apical organelles. Based on videomicroscopy of *P. knowlesi* (Dvorak et al., 1975) and several structural and electron microscopy studies (Ladda et al., 1969; Bannister et al., 1975; Aikawa et al., 1978), the invasion process is complete within minutes and can be broken down into distinct stages: initial attachment, apical reorientation, junction formation followed by rhoptry secretion and formation of the parasitophorous vacuole (PV) and finally active motor-driven entry into the host cell (illustrated in figure 1.5).

Figure 1.4: *Erythrocyte Invasion*. (A) The free merozoite initially forms a weak, reversible attachment with the erythrocyte surface *Erythrocyte Invasion* via merozoite surface proteins. (B, F) The merozoite reorientates so that the apical end of the parasite is in contact with the erythrocyte surface. Once the apical end is aligned perpendicular to the erythrocyte surface, a depression forms in the erythrocyte membrane (G) and a close irreversible attachment occurs, acting as a junction between parasite and host cells. (C, D, H) The junction then moves, by way of the actin/myosin motor, engulfing the parasite, shedding its surface proteins as the parasite enters the erythrocyte. During the junction's posterior progression, the merozoite's electron dense proteinacious coat is shed by proteolytic cleavage. (E+I) When the moving junction reaches the merozoite's posterior surface, the adhesion between the cells is removed and the parasite remains inside its own vacuole, separate from the erythrocyte cytoplasm. Transmission electron microscopy shows a thickening of the membrane around the formation of the junction. The junction remains intact until reaching the posterior end. Rhoptries are secreted into the rbc as the parasite enters, creating the parasitophorous vacuole in which the internalised parasite resides (adapted from (Aikawa et al., 1978; Cowman and Crabb, 2006; Riglar et al., 2011).


C
D
B
E
A
F
E
M
A
R
N
G
E
Mz
H
C
Mz
C
I
Mz
C
1 Attachment
Rhoptries
Rhoptries
Nucleus
Micronemes
2 Invasion
3 Sealing

1.5.1 Initial Attachment

Once released into the blood stream from their parent schizont, free merozoites must find a new erythrocyte host. The selection of host cell represents a specific molecular interaction as *P. knowlesi* merozoites attach to simian and human rbcs but not those of non-susceptible hosts such as chicken and guinea pig (Miller et al., 1977). The interaction is weak, reversible and can take place at any position of the merozoite surface (Dvorak et al., 1975; Bannister and Dluzewski, 1990). Merozoite surface proteins (MSPs) are undoubtedly involved in initial rbc binding although there is very limited evidence for it thus far. MSP-1, 2, 4, 5, 8 and 10 are attached to the parasite membrane by way of GPI-anchor. Interestingly MSP1, 4, 5, 8 and 10 all share one or more epidermal growth factor (EGF)-like domain at their C-terminus . EGF-domains are involved in protein-protein interactions in other biological system however this has yet to be demonstrated experimentally in *Plasmodium*. Merozoite surface protein (MSP)-1 is one of the most abundant protein on the surface of the merozoite. The MSP-1 protein precursor is c. 200 kDa, which is cleaved by the *P. falciparum* PfSUB-1 within the parasitophorous vacuole into 83, 30 and 38kDa fragments which are non-covalently linked to the C-terminal 42kDa fragment that remains attached to the membrane by its GPI anchor (Holder et al., 1987; McBride and Heidrich, 1987; Gerold et al., 1996). MSP-1 is found in a high molecular weight complex with soluble merozoite surface proteins MSP-6 and MSP-7 (Kauth et al., 2006). The soluble MSP-9 has been reported to associate with MSP-1(42), which in complex bind the rbc surface glycoprotein Band 3 (Goel et al., 2003; Li et al., 2004a). A study by Ranjan et al. (2010) suggests MSP-1 forms a complex with Rhoptry bulb proteins which have been described by some to have rbc binding properties (Ranjan et al., 2010). This is however unlikely due to temporal and spatial isolation within the merozoite. Although the MSPs are the most studied molecules of the merozoite, the role of these proteins and complexes has yet to be experimentally proven and the key event of host cell selection upon initial attachment remains poorly understood.

1.5.2 Apical reorientation

After attaching to the correct cell, the merozoite must re-orientate so that the apical end of the merozoite is in contact with the rbc surface. The molecular mechanisms behind this process are as yet unknown however Mitchell *et al* (2004) demonstrated that apical membrane antigen (AMA)-1 specific monoclonal antibodies are able to block re-orientation of *P.knowlesi* merozoites, suggesting an essential role for AMA-1 in this process (Mitchell et al., 2004). AMA-1 is conserved throughout Apicomplexa and was first described as part of the apical complex of blood stage merozoites (Peterson et al., 1989). Since then it has been shown to be essential for blood stage growth (Triglia et al., 2000). This protein is a type 1 membrane protein expressed in late schizogony, localising to a subset of the apical organelles called the micronemes (Healer et al., 2002; Bannister et al., 2003). These organelles secrete their contents from the apex prior to invasion and AMA-1 can be seen diffused out onto and diffused around the free merozoite cell surface by indirect immunofluorescence assay (IFA) (Healer et al., 2002). It was proposed by Mitchell *et al* (2004) that a gradient of AMA-1 existed that allowed the parasite to roll towards the apex where the highest concentration of AMA-1 is present although this has never been proven.

A role in apical re-orientation implies AMA-1 is capable of host cell binding and there is experimental evidence that demonstrates the erythrocyte adhesive properties of AMA1. The protein expressed in COS cells leads to rbc agglutination (Fraser et al., 2001) and AMA-1 derived peptides bind to human erythrocytes (Urquiza et al., 2000). Domain III of AMA-1 has been shown to bind to the rbc surface protein Kx (Kato et al., 2005), however AMAI derived from *Babesia divergens* bound to wild type and Kx null rbcs in equal measure (Montero et al., 2009). As AMA1 is expressed in zoites of all apicomplexa it would be surprising if the major role of AMA1 involved any kind of erythrocyte specific receptor. An alternative role for AMA1 in invasion is discussed in the following sections (1.5.4).

1.5.3 Interactions at the apex

With the apex in juxtaposition to the erythrocyte cell membrane, the parasite forms a closer interaction with its host. Upon reorientation, released micronemal proteins are able to interact with their host cell binding partners. The micronemal erythrocyte binding proteins of the Duffy-binding-like (DBL) superfamily are released onto the surface of the merozoite and appear to remain around the apical end of the parasite unlike their fellow microneme inhabitant, AMA-1, which is diffused around the surface (Singh *et al*. 2010). Proteins EBA-175, EBL-1 and EBA-140 (BAEBL) bind sialic acid residues of the rbc surface proteins glycophorin A (Sim et al., 1994), glycophorin B (Mayer et al., 2009) and glycophorin C (Maier et al., 2003), respectively. Other gene family members include the rbc binding protein EBA-181 (JESEBL), for which the receptor remains to be determined (Gilberger et al., 2003), and the transcribed pseudogene EBA-165 (PEBL) (Triglia *et al.,* 2001). A sixth member of this family, MAEBL, is expressed in merozoites and sporozoites however it possesses an atypical DBL-domain, more like the rbc binding region of AMA-1 (Kappe *et al*., 1998), and is essential for salivary gland, not erythrocyte invasion (Fu *et al*., 2005; Saenz *et al*., 2008).

A second family of invasion molecules are 5 members of the reticulocyte-binding-like (RBL) superfamily, the PfRh proteins. Unlike the EBA proteins the Rhs are located in the rhoptries/ rhoptry neck of the mature blood stages. PfRh1 has erythrocyte binding activity and localises to the apex of merozoites (Rayner *et al*., 2001), presumably to the rhoptry neck based on the presence of family members PfRh2a and PfRh2b which also exhibit erythrocyte binding capacity however the exact receptor has yet to be identified (Rayner *et al*., 2000; Taylor *et al*., 2002; Duraisingh *et al*., 2003). PfRh4 is also a rhoptry neck protein which binds to complement receptor 1 (CR-1) on the rbc surface (Tham *et al*., 2010). PfRh3 is not translated due to a frameshift mutation in the 5' end of the gene (Taylor *et al*., 2001). Interestingly, as also previously shown for the EBAs, it is possible to individually delete the Rh genes with the exception of PfRh5 (Duraisingh et al., 2003; Triglia et al., 2005; Baum et al., 2009a; Tham et al., 2010). The smallest of the Rh proteins lacking a transmembrane domain, PfRh5 has recently been shown to bind to the erythrocyte surface protein basigin (Crosnier *et al*., 2011).

The multiple host cell binding antigens imply that a certain amount of redundancy in the erythrocyte invasion process is required to safeguard successful invasion. If one invasion pathway is unavailable *P. falciparum* seems to be able to switch invasion pathways e.g. from sialic acid dependant to sialic acid independent (Dolan et al., 1990). In laboratory adapted strains of *P. falciparum*, genetic deletion of *eba-175* results in upregulation of *Rh4* (Stubbs et al., 2005). A large diversity of invasion phenotypes can be found in field isolates however the sialic acid dependant pathway appears to be the most prevalent invasion pathway, particularly in young children, based on dominant transcript expression of *eba175* and *eba140* (Gomez-Escobar et al., 2010). This redundancy in invasion recepters not only allows the parasite to invade a wide variety of human erythrocytes of all ages but facilitates immune evasion, since targeted deletion of *eba-175* and *Rh4* results in decreased susceptibility of parasites to growth inhibition from antibodies from semi-immune serum (Persson et al., 2008). It has been suggested that the binding of these antigens triggers commitment to invasion, in particular the binding of EBA-175 to glycophorin A or Rh4 to CR1 , after which rhoptry secretion will occur, even in the absence of a moving junction (Singh *et al.*, 2010; Riglar *et al*., 2011).

1.5.4 Junction formation

Protein-protein interactions here are not restricted to parasite-host, but in fact a parasite-parasite protein interaction may be the most important complex in the invasion process, forming an anchoring ring for a molecular motor, driving the internalisation of the merozoite. First discovered in *T. gondii* tachyzoites (Alexander *et al.*, 2005; Lebrun *et al*., 2005) and later shown in *P. falciparum* (Riglar *et al.*, 2011), on secretion from the rhoptry neck, proteins RON2, RON4 and RON5 are translocated across and inserted into the erythrocyte membrane where they follow the moving junction from apex to the posterior end of the invading zoite. The micronemal protein AMA-1 co-immunoprecipitates with RON2, RON4, RON5 and RON8 in tachyzoites (Alexander *et al*., 2006) - directly interacting with RON2 (Besteiro *et al*., 2009; Straub et al., 2009; Lamarque et al., 2011; Tyler and Boothroyd, 2011). A similar complex exists in *P. falciparum* without

the presence of RON8 which has no orthologue in *Plasmodium* (Cao et al., 2009; Richard et al., 2010). Antibodies and peptides that inhibit this complex formation also block invasion, highlighting its importance in host cell invasion (Collins et al., 2009; Richard et al., 2010; Lamarque et al., 2011). In light of this evidence the proposed model for moving junction interaction places RON4 on the cytoplasmic face of the erythrocyte membrane, interacting with transmembrane protein RON2, of which the C-terminus is to be found on the extracellular face of the erythrocyte membrane. This complex binds to AMA-1 which is anchored in the parasite membrane via a single transmembrane domain.

Although there are no established methods of genetic conditional knock-down in *P. falciparum*, AMA-1 cannot be deleted in blood stage parasites implying its function is essential (Triglia et al., 2000). Giovannini *et al.* (2011) show through conditional KO in *T. gondii* tachyzoites and *P. berghei* sporozoites that AMA-1 is not essential at these stages for invasion of host cells but invasion is impaired - tachyzoites invade at an angle and 'corkscrew' into the cell as opposed to their normal linear internalisation (Giovannini et al., 2011). RON4 is an essential component invasion of *T. gondii* tachyzoites and *P. berghei* sporozoites, and this dissociated phenotype implies separate roles in invasion. Importantly, AMA-1 was shown to be essential for *P. berghei* merozoite invasion of rbcs *in vivo* however, as the invasion process and machinery is shared among all apicomplexa zoites, a functionally diverse role is unlikely. AMA1, unlike RON4, seems to have its role on the parasite site of the invasion junction but might therefore be involved in stabilising correct orientation and distance between parasite and host cell rather than the junction itself. This stability may be more important in the fast-flowing, fluid environment of the blood stream.

1.5.5 The molecular motor – driving invasion

This junction moves with respect to the invading parasite, from the apex at which it is formed to the extreme posterior, engulfing the merozoite in rbc and parasite-derived membrane. This movement is powered by an actin-myosin motor which resides between the plasma membrane and an inner membrane complex (IMC), a series of

flattened membranous vesicles lining the interior of the cell membrane. The motor complex is highly conserved and is responsible for motility and invasion throughout the Apicomplexan family. The complex consists of membrane-anchored glideosome associated protein GAP50, GAP45, myosin A tail-interacting protein (MTIP), myosin A, F-actin and aldolase (Baum *et al*., 2006; Green *et al*., 2006b; Baum *et al*., 2008), illustrated in figure 1.5. An additional component of the motor complex, GAP40, has been discovered in *T. gondii* (Frenal et al., 2010), however its *Plasmodium* homologue remains undiscovered. Upon binding of the merozoite apex and rbc surface, an unknown trigger transmits the binding signal across the membrane and kick-starts the motor in which myosin A and F-actin interaction drive the parasite forward. In merozoites, the link between surface adhesion and activation of the motor remains to be defined. In sporozoites, thrombospondin related anonymous protein (TRAP) binds both hepatocyte membrane and aldolase (fructose-1,6-bisphosphate aldolase), an enzyme involved in glycolysis but that also is capable of binding filamentous F-actin and therefore linking host cell binding with the parasite's invasion machinery (Buscaglia et al., 2003). TRAP is a member of a family of thrombospondin repeat antigens which contain combinations of human thrombospondin type I-related (TSR) and von Willeband factor type A-related (vWA) adhesive domains. Family members include TRAP of *Babesia,* TRAP-C1 of *Cryptosporidium,* MIC2 of *Toxoplasma* and *Neospora,* and MIC1 and ETFP250 of *Eimeria* (Morahan et al., 2009). The *Plasmodium* merozoite equivalent of TRAP (MTRAP) is thought to fulfil this role and binds aldolase (Baum et al., 2006) however no red cell binding activity had ever been described until recently. A study by by Uchime *et al* (2012) showed for the first time *in vitro* that MTRAP but not fellow TSR domain-containing family member *Plasmodium* thrombospondin-related apical merozoite protein (PTRAMP), exhibited rbc binding activity, validating a potential role for this protein in linking rbc binding with activation of the motor (Uchime et al., 2012). Aldolase has been implicated in connecting the cytoplasmic tails of erythrocyte-binding surface molecules to actin, linking the surface interaction and triggering of the parasite molecular motor. It has been reported that aldolase is found on the cytoplasmic face of micronemes in sporozoites and *Toxoplasma* tachyzoites suggesting an interaction with MTRAP even before microneme release (Buscaglia et al., 2003; Jewett and Sibley, 2003).

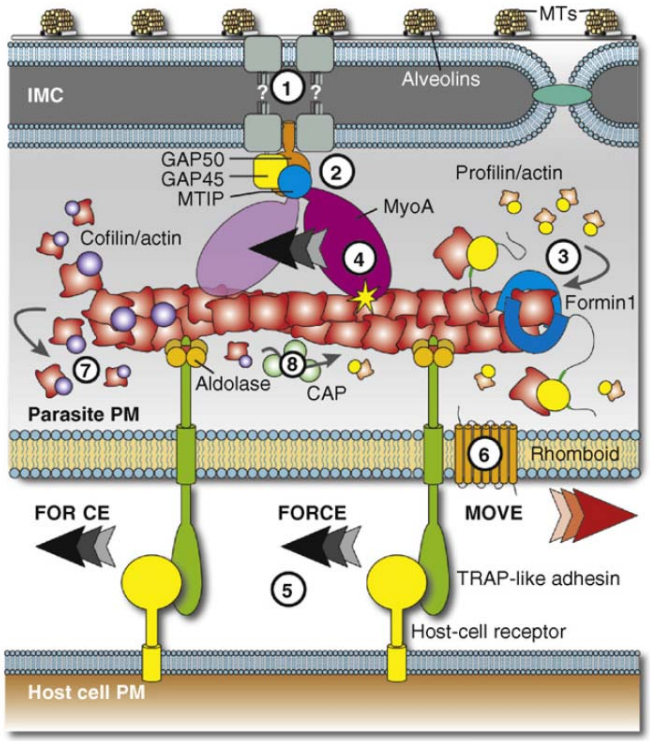


Figure 1.5: *The motor complex.* (1) MyoA anchored in detergent resistant membranes of the IMC and interacts with MTIP, GAP45 and GAP50 as well as IMC constituent alveolin or GAPMs (2). When signalled, formin and profilin polymerise actin (3) and MyoA drives actin (4) and TRAP (connected by aldolase) towards the merozoite posterior (5). Adhesions are shed to allow internalisation (6). Actin depolymerises and monomers are recycled (7,8) (adapted from (Baum et al., 2008))

1.5.6 A close shave

In order to facilitate smooth parasite entry into the rbc, its bulky surface proteins/protein complexes should be lost and interactions between parasite and host proteins must be uncoupled. This action can be seen by EM and is performed by two different families of proteases. The first is a member of the subtilisin-like serine protease family, PfSUB2, which is responsible for the juxta-membrane shedding of MSP-1, AMA-1 and PTRAMP (Harris et al., 2005; Green et al., 2006a). The 42 kDa fragment of MSP-1 is cleaved leaving a 19 kDa membrane-bound fragment to enter the parasitophorous vacuole (Blackman et al., 1990, 1991, 1996). A recent TEM study by Riglar *et al* (2011) demonstrated that cleavage by SUB2 and actual shedding may be temporally independent and occur at different sites. A second protease family have a role in merozoite membrane-protein shedding. Rhomboid proteases are involved in the disengagement of adhesive interactors between the rbc and the parasite. They are also serine proteases and cleave within the transmembrane domain of susceptible proteins (Freeman, 2008). ROM1 and

ROM4 cleave a number of adhesins involved in rbc attachment such as EBA-175, the cleavage of which by ROM4 is an essential process to the parasite (O'Donnell et al., 2006). Interestingly, although the majority of AMA-1 is shed by SUB2, a small amount of the protein is cleaved by a rhomboid protease (Howell et al., 2005), perhaps indicating a disparate role of AMA-1 concentrated at the junction and that observed around the circumference of the cell.

1.5.7 Parasitophorous vacuole formation and completion of invasion

As the merozoite enters the rbc, the electron-lucent, lipid-rich contents of the rhoptries are discharged into the erythrocyte, aiding the creation of a parasitophorous vacuole in which the merozoite is to reside throughout the intraerythrocytic stage of its lifecycle (Bannister et al., 1986). TEM shows the rhoptry associated protein (RAP)1moves to the apex mid-invasion, the position that rhoptry neck proteins previously occupied, before finally taking up residence in the PVM (Riglar et al., 2011). Rhoptry secretion is concomitant with merozoite entry into the erythrocyte through the junction. As the moving junction reaches the posterior of the merozoite, the PVM fuses so that the merozoite is contained within its own vacuole, out of the rbc cytosol, and the invasion process is complete (Lingelbach and Joiner, 1998). The dense granules remain intact throughout the invasion process but are secreted immediately following the completion of invasion so that the contents may fulfil their role in the PV, PVM or cytosol of the newly invaded rbc (Bannister et al., 1975; Torii et al., 1989; Culvenor et al., 1991; Riglar et al., 2011).

1.6 The Merozoite

The merozoite stage of the *Plasmodium* lifecycle is responsible for the obligatory process of erythrocyte invasion and entry into the lifecycle stage that gives rise to malaria disease symptoms. Understanding the molecular biology of this cell is therefore crucial in designing vaccines and new drug targets. *Plasmodium* is a eukaryotic pathogen and this is reflected in the merozoite cellular architecture (figure 1.6). The nucleus resides at

the posterior of the cell and contains the 14 chromosomes and 23Mb of DNA that make up the *Plasmodium* nuclear genome. The mitochondrial genome holds another 6kb while an Apicomplexan-specific organelle, the apicoplast, contains a 35kb circular genome (Gardner et al., 2002). All Apicomplexa, excepting *Cryptosporidium spp.*, possess this ancient relic of a red algal, secondary endosymbiont, the acquisition of which changed the course of their evolution. The resulting non-photosynthetic plastid organelle is reduced in size and activity, however it is an indispensible part of the zoite cellular architecture. It is involved in the biosynthesis of fatty acids, isoprenoids and the digestion of haem, providing the parasite with metabolites which are essential to its survival (reviewed by (Striepen, 2011)). Although in possession of its own genome, the majority of apicoplast proteins are encoded by the nuclear genome and must be trafficked through the classic eukaryotic secretory pathway. In order to reach the correct destination, proteins contain an N-terminal signal sequence for entry into the ER followed by a transit peptide which drives correct trafficking to the apicoplast via vesicular transport (Waller et al., 1998; Foth et al., 2003).

Surrounding the merozoite nucleus resides the membranous network of the endoplasmic reticulum (ER). Due to the cell's polarity this network can extend in the apical direction but is distinct from the apex. As with all eukaryotic cells, the merozoite ER plays host to newly synthesised transmembrane and soluble proteins destined for cellular compartments or secretion from the cell. Proteins require a classic signal sequence at the N-terminus in order to enter the ER and protein translocation machinery appears to be conserved, although this observation is reliant on sequence similarity and not experimental data (Tuteja, 2007). Proteins are retained in the ER by the conventional XDEL motif which retrieves escaped resident ER proteins from the Golgi (Elmendorf and Haldar, 1993; van Dooren et al., 2005). A Golgi apparatus of sorts exists and cis- and trans-Golgi compartments have been identified using the markers ERD2, Rab6 and GRASP respectively (Elmendorf and Haldar, 1993; de Castro et al., 1996; Van Wye et al., 1996; Noe et al., 2000). There is however limited evidence for the characteristic Golgi stacks seen in higher eukaryotes. The function of these stacks is to compartmentalise, facilitating processing and sorting of newly synthesised proteins and lipids. Since

Plasmodium has limited N-glycosylation and lacks an O-linked glycosylation pathway (Templeton et al., 2004) this level of compartmentalisation may not be necessary. One post-translational modification that does occur is the addition of glycosylphosphatidylinositol (GPI) to the cleaved C terminus of would-be surface proteins, to allow membrane insertion. The merozoite surface is coated in a thick layer of GPI-anchored proteins which include the abundant MSP-1 and are believed to facilitate attachment to red blood cells. Immediately underneath the merozoite membrane lies a series of flattened vesicles known as the inner membrane complex (IMC). This double membranous structure is common among many Chromalveloates including all Apicomplexa. On the cytosolic face of the IMC lies a network of cytoskeletal intermediate filament-like proteins and sub-pellicular microtubules, providing structural support to the cell. The sub-pellicular microtubules are also thought to be involved in the trafficking of secretory vesicles, particularly to the apical end of the cell (Bannister et al., 2003). One of the defining features of the phylum Apicomplexa is the presence of a unique set of organelles at the apex of the zoite form of the parasite. The protein contents of these apical organelles are arguably the most important proteins in the cell as their secretion is absolutely necessary to permit host cell invasion. These proteins, as well as surface proteins, are described in Table 1.1.

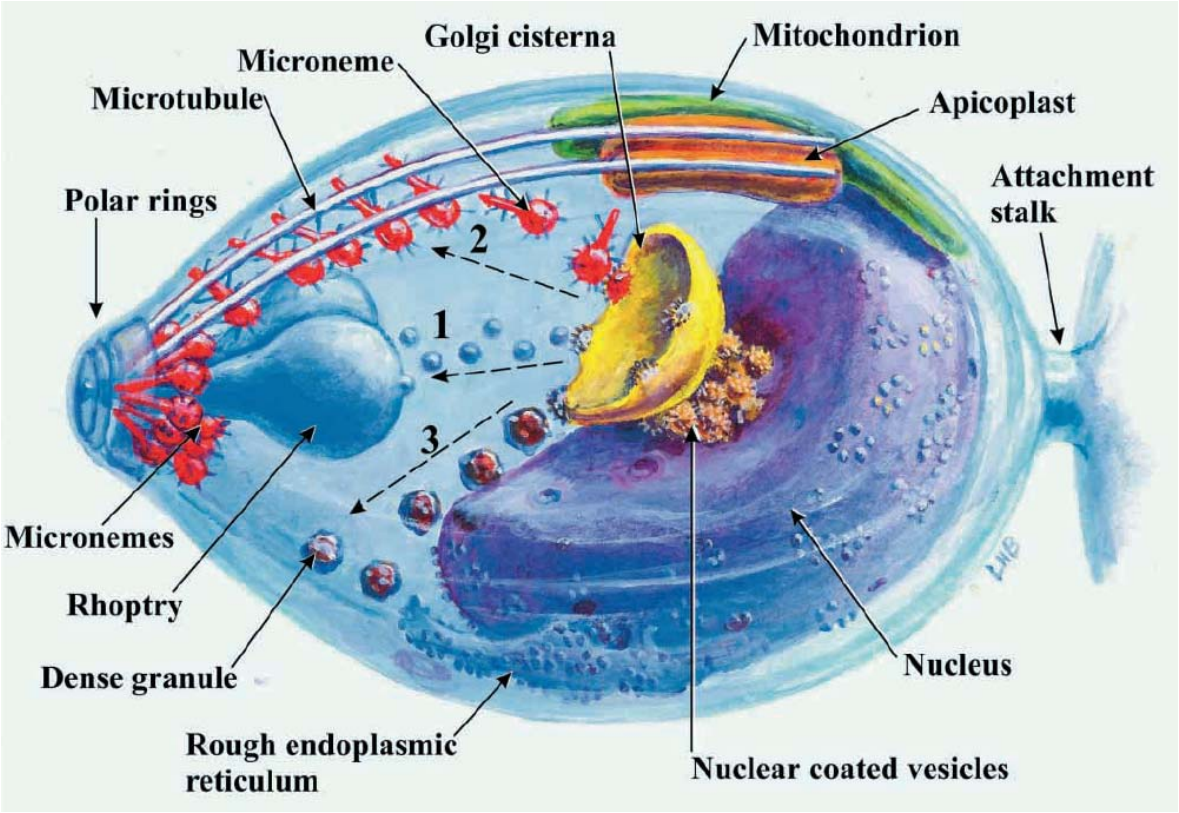


Figure 1.6: *The Merozoite*. All major organelles are highlighted. Note the polarity, with the secretory organelles at the apex and nucleus at the posterior. Trafficking routes from the ER and Golgi to the rhoptries, micronemes and dense granules are labelled 1, 2 and 3 respectively [adapted from (Bannister et al., 2003)].

Table 1.1: *P.falciparum merozoite proteins important in erythrocyte invasion.*

Name	Features/Function
GPI-anchored Surface Proteins	
MSP-1	2 EGF domains, shedding upon invasion essential, cannot KO, putative Band 3 ligand
MSP-2	highly polymorphic, cannot KO
MSP-4	EGF domain, cannot KO
MSP-5	Non-essential paralogue of MSP4
MSP-10	cannot KO, surface AND apical staining, 2 EGF domains
Pf12	6-cys family member, similar to SAG of *T. gondii*
Pf38	6-cys family member, surface and apical location
Pf92	Cysteine-rich protein, cannot KO
Pf113	Putatively GPI anchored , putative surface protein, cannot KO
Peripheral Surface Proteins	
MSP-3	redundant member of MSP-3 family locus, coiled coil domain
MSP-6	MSP-3 family member, in complex with MSP-1+7 on the surface, not essential
MSP-7	non-essential in complex with MSP-1+6; deletion causes reduced invasion
MSP-9(ABRA)	essential, putative association with MSP-1 and potential Band 3 ligand
SERA5	serine repeat antigen, putative serine protease, high expression levels; essential
S-antigen	Cannot KO, function unknown
GLURP	glutamate-rich protein, targeting antibodies associated with protection from clinical symptoms in the field
H101	non-essential MSP3 family member
MSP11	non-essential MSP3 family member
Pf41	6-cys family member, apical surface protein
PfDBLMSP	redundant MSP-3 family member, DBL domain
Micronemal proteins	
AMA-1	essential merozoite invasion protein, antibodies inhibit apical reorientation/invasion, binds RON2 in complex with RON4 at moving junction
EBA-140	redundant erythrocyte binding protein, binds glycophorin C
EBA-175	redundant erythrocyte binding protein, binds glycophorin A
EBA-181	redundant erythrocyte binding protein, binds unknown rbc receptor
EBL-1	redundant erythrocyte binding protein, binds glycophorin B
MTRAP	essential rbc adhesin, also binds aldolase, putative motor function
SUB2	subtilisin-like protease or surface sheddase, cleaves MSP-1 and AMA-1 upon merozoite internalisation
Rhoptry Neck Proteins	
RH1	redundant erythrocyte binding antigen, receptor unknown
RH2a	redundant erythrocyte binding antigen, receptor unknown
RH2b	redundant erythrocyte binding antigen, receptor unknown
RH3	pseudogene
RH4	redundant erythrocyte binding antigen, binds to CR1 on rbc surface
RH5	essential invasion ligand, binds basigin on rbc surface

RON1	aka apical sushi protein (ASP), contains 'sushi' domain and GPI anchor
RON2	essential component of moving junction, target of invasion inhibitory antibodies, translocated to rbc surface, binds AMA-1
RON4	essential component of moving junction, target of invasion inhibitory antibodies, translocated to rbc cytoplasmic face
RON5	6 transmembrane protein, conserved in *Plasmodium spp,* homologue in *T. gondii*
RON6	cannot KO, soluble, cysteine rich domain, conserved in *T. gondii*
Rhoptry Bulb Proteins	
RAMA	Early expressed, GPI-anchored rhoptry protein, putative role in binding protein complexes
RAP1	non-essential member of low-molecular weight rhoptry protein complex
RAP2	non-essential member of low-molecular weight rhoptry protein complex, attaches to rbc surface upon aborted invasion
RAP3	non-essential member of low-molecular weight rhoptry protein complex
RhopH1(2)	high molecular weight rhoptry protein complex, redundant member, encoded by *clag* multigene family
RhopH1(3.1)	high molecular weight rhoptry protein complex, redundant member, encoded by *clag* multigene family
RhopH1(3.2)	high molecular weight rhoptry protein complex, redundant member, encoded by *clag* multigene family
RhopH1(9)	high molecular weight rhoptry protein complex, redundant member, encoded by *clag* multigene family
RhopH2	high molecular weight rhoptry protein complex, cannot KO
RhopH3	high molecular weight rhoptry protein complex, cannot KO
RON3	*P. falciparum* orthologue of *T. gondii* RON3 is located in the rhoptry bulb not neck, in complex with RON2+4 but not AMA-1

Table 1.1: *P.falciparum merozoite proteins important in erythrocyte invasion.* The table represents a summary of proteins of the merozoite surface and apex described in detail and referenced in the text or otherwise references can be found in (Cowman and Crabb, 2006) from which the table was adapted.

1.6.1 The apical organelles

The apical organelles are a defining feature of apicomplexa and their secretion facilitates motility and invasion of host cells. *Plasmodium* merozoite apical organelles are formed *de novo* during schizogony in the last 10-12 hours of the parasites intraerythrocytic development. In order to reach the apical organelles, proteins contain an N-terminal signal sequence in order to utilise the classical secretory pathway, trafficking through the ER and Golgi before reaching their apical destination.

1.6.1.1 Dense Granules

Dense granules form around the periphery of the Golgi but little is known of their trafficking other than timing of protein expression is key (Rug et al., 2004). Their name is derived from their appearance in electron microscopy – electron dense, granular vesicles which reside slightly behind the micronemes and rhoptries (Bannister et al., 1975). Less is known about the role of dense granules although their secretion appears to be coordinated and they diffuse laterally into the PV after invasion has taken place (Torii et al., 1989; Carruthers and Sibley, 1997; Riglar et al., 2011). As would be predicted from the timing of their secretion, proteins of the dense granules have a role in ring stage parasites, particularly in host cell remodelling. The best characterised dense granule protein is Ring-infected Erythrocyte Surface Antigen or RESA which is released from dense granules (Culvenor et al., 1991), exported into the rbc cytoplasm where it binds spectrin (Foley et al., 1991) and is involved in resistance to heat shock (Silva et al., 2005) and decreasing rbc deformability (Mills et al., 2007).

1.6.1.2 Rhoptries

The rhoptries are large, club or bulb shaped organelles which occur in pairs at the apex of the merozoite. Rhoptries are formed through the accumulation of vesicles trafficked from the Golgi (Jaikaria et al., 1993; Bannister et al., 2000). It seems the first 24 amino acids, including signal peptide, of RhopH2 is adequate to target to the rhoptries after transport through the secretory pathway (Ghoneim et al., 2007). A C-terminal tyrosine motif as well as a cytoplasmic dileucine motif has been described for the targeting of

rhoptry membrane proteins of *T. gondii* and although several *Plasmodium* rhoptry proteins also possess these motifs, a role in trafficking has not been elucidated (Hoppe et al., 2000; Ngo et al., 2003). Adaptor protein complex 1 (AP-1) and dynamin-related protein B (DrpB) have also been implicated in rhoptry targeting in *T. gondii* (Ngo et al., 2003; Breinich et al., 2009).There also appears to be a mechanism of secretory pathway-independent targeting to the rhoptries since the *P. falciparum* armadillo repeats-only (PfARO) protein can be targeted to rhoptry membranes by myristoylation and palmitoylation within the first 20 aa of the protein (Cabrera et al., 2012). Ghoneim *et al* (2007) suggest RhopH2 is post-translationally inserted into the ER, however the necessary machinery for this translocation has not been identified in *Plasmodium* (Tuteja, 2007). They also suggest a role for chaperones in identifying rhoptry targeting sequences however experimental evidence is lacking.

Each individual rhoptry is divided into 2 distinct compartments, the electron-dense rhoptry bulb and electron-lucent rhoptry neck. The thin neck regions of the rhoptries appear in close proximity at the extreme apical tip of this invasive cell (Bannister et al., 2000). The bulb compartment contains lipids which are responsible for the lucidity of the rhoptry bulb in TEM studies (Bannister et al., 1986). These lipids have a role in establishment of the PVM within the rbc.

Rhoptry exocytosis has been shown *in vitro* to occur following host cell interaction with EBA-175, EBA-140 or EBA-181 when previously elevated intracellular calcium levels are supposedly restored to basal levels (Singh et al., 2010), although the precise signalling mechanisms involved have yet to be demonstrated. The requirement for committed invasion as defined by binding of EBL- and also RBL- antigens was confirmed Riglar *et al* (2011), demonstrating by TEM, 3D SIM, IEM and IFA, rhoptry contents will be secreted irrespective of whether an intact junction forms at the apex, with an incomplete junction resulting in the secretion of rhoptry contents onto the surface and culture media but not into the host cell and without successful invasion. In parasites with intact junctions but chemically treated to block either the actin/myosin motor (cytochalasin D) or shedding (PMSF), rhoptry contents are secreted in 'whorls' into the rbc cytoplasm instead of

surrounding the invading merozoite (Miller et al., 1979; Bannister et al., 1986; Riglar et al., 2011).

Proteins contained in each compartment are separated based on temporal and functional differences since the rhoptry neck protein will be secreted first, followed by the rhoptry bulb contents (Bradley et al., 2005). Rhoptry neck proteins play an important part in erythrocyte attachment via the RH family (discussed in 1.5.3) and junction formation where RON2 and RON4 play essential roles (described in section 1.5.4). The specific role of other rhoptry neck proteins is less clear. The bulb is thought to contain protein constituents of the PV or PVM. Several rhoptry bulb proteins have been described to date, with perhaps the most studied being members of the high molecular weight rhoptry (or RhopH) protein complex. This complex consists of the proteins RhopH1, RhopH2 and RhopH3 and is formed in the schizont's secretory pathway, remains intact within the newly infected erythrocyte and is conserved in other *Plasmodium* species (Campbell et al., 1984; Holder and Freeman, 1984b; Cooper et al., 1988; Hienne et al., 1998). While RhopH2 and RhopH3 are encoded by single genes which are refractory to genetic deletion (Holder, unpublished data; Cowman et al., 2000), RhopH1 is encoded by at least 5 members of the *clag* (cytoadherence linked asexual gene) multi-gene family (Kaneko et al., 2001) and different clones express different *clag* genes - expression of *clag3.1* and *clag3.2* is mutually exclusive (Cortes et al., 2007). The first of these genes to be characterised, *clag9,* was implicated in cytoadherence (Trenholme et al., 2000), although the protein was demonstrated to localise and interact with the RhopH complex, which is present in the merozoite rhoptry bulb and translocated into the PVM upon invasion, where it remains throughout the ring stage (Hiller et al., 2003; Ling et al., 2003; Ling et al., 2004), calling this role into question. There has been much controversy regarding the role of rhoptry proteins since this complex also exhibits erythrocyte-binding properties and antibodies to the complex inhibit invasion (Sam-Yellowe and Perkins, 1991; Doury et al., 1994; Wang et al., 2006). Given that rhoptry bulb proteins are internal in the free merozoite and are passed through what is thought to be a molecular seal (the moving junction) during host cell invasion, it is not clear how the antibodies access the complex in order to hinder

merozoite entry. It is also unclear how the RhopH complex will come in contact with the rbc surface during 'successful' invasion, other than the portion of the rbc membrane that is internalised during the formation of the PVM where the complex resides in ring stage parasites (Ling et al., 2004). A similar situation is true of the low molecular weight rhoptry protein complex consisting of rhoprty associated proteins RAP1, RAP2 and RAP3. Although located in merozoite rhoptries and transferred into the PV during the invasion process, the complex can readily be observed on the surface of infected and non-infected erythrocytes (Sterkers et al., 2007; Awah et al., 2011). Destruction of uninfected red bloods cells is a major contributing factor to malaria-induced anaemia (Mendez et al., 2000). Erythrocytes coated with sticky rhoptry proteins are the target for immune-mediated clearance (Layez et al., 2005; Evans et al., 2006; Awah et al., 2009).
Since that we cannot assume that all invasion attempts are 'successful' and given that committed invasion is triggered by the binding of apical erythrocyte binding proteins resulting in the release of rhoptry contents irrespective of the presence of an intact moving junction (Singh et al., 2010; Riglar et al., 2011), it is probable that rhoptry proteins contact the rbc surface during aborted invasion. Deposition of these protein complexes onto erythrocytes that aren't invaded will tag them as 'foreign' nonetheless, rendering the rbc liable to immune destruction.

1.6.1.3 Micronemes

Micronemes are small, elongated vesicles with an electron dense membrane, found at the apex, in front of the rhoptry bulb as well as peripheral and slightly posterior to the rhoptry neck. The organelles bud off from the Golgi and are transported along sub-pellicular microtubules to the apical end of the merozoite (Bannister et al., 2003). Evidence from the ookinete microneme proteome suggests this trafficking may deliver motor components to the apex via attachment to micronemal proteins' cytoplasmic tails (Lal et al., 2009). This is consistent with reports of aldolase on the cytoplasmic face of micronemes in sporozoites and *Toxoplasma* tachyzoites (Buscaglia et al., 2003; Jewett and Sibley, 2003). Most micronemal proteins are single-pass transmembrane proteins with large extracellular domains and short cytoplasmic tails. In *T. gondii*, a tyrosine motif and patch of acidic residues in this C-terminal cytoplasmic tail targets transmembrane

proteins to the micronemes (Di Cristina et al., 2000), however this does not seem to be the case for the *P. falciparum* EBL family (Gilberger et al., 2003), which instead require a cysteine-rich region present in the ectodomain as well as correcting timing of expression (Treeck et al., 2006). Soluble microneme proteins have been described in ookinetes (Langer et al., 2000; Yuda et al., 2001) as well as coccidian apicomplexa (Eschenbacher et al., 1993; Fourmaux et al., 1996; Tomley et al., 1996). There is evidence from *Toxoplasma* that a family of micronemal EGF-domain containing proteins function as escorts, directing microneme soluble residents to the correct destination (Meissner et al., 2002), however no functional homologue has been discovered in *Plasmodium*.

Microneme exocytosis has been shown to occur upon egress, where the low potassium concentrations found in the blood stream trigger an intracellular rise in calcium, resulting in the micronemal protein AMA-1 being translocated to the surface (Singh et al., 2010). An intracellular calcium increase is also implicated in microneme release in *Toxoplasma*, with low intracellular potassium-induced release of calcium from intracellular stores resulting in the surfacing of invasion-essential micronemal protein MIC2 (Carruthers et al., 1999; Carruthers and Sibley, 1999; Lovett et al., 2002). Microneme membrane fusion that facilitates secretion is thought to occur at the rhoptry neck. Proteins contained in micronemes appear to have different specific roles however all micronemal proteins characterised to date contribute towards the same goal – the removal of the free merozoite from the inhospitable environment of the blood stream into the sanctuary of the erythrocyte. Despite being assigned a micronemal location, EBA-175 and AMA-1 do not co-localise. Upon release, AMA-1 is dispersed over the whole surface of the merozoite, colocalising with MSP-1, however EBA-175 is only ever found at the apical tip, suggesting some level of spatial compartmentalisation. This has been described in *Toxoplasma* tachyzoites with distinct proteins shown to localise only to the apex. Two distinct subsets of micronemes have been described in *P. falciparum*. One subset contains the protease SUB1 which is required for egress, hence these organelles have been termed the exonemes (Yeoh et al., 2007). Unlike classic micronemes they are believed to secrete their contents during schizogony by fusion with the plasmalemma and not using the apical organellar duct in a fashion reminiscent of dense granules. The

second subset, named mononemes, contain the rhomboid protease PfROM-1 and are distinct in location from other micronemal markers, as determined by fluorescent microscopy (Singh et al., 2007).

An ookinete microneme proteome was published by Lal *et al* (2009) however this has not been possible for other invasive stages due to difficulty in separating micronemes from rhoptries and dense granules, which are absent in ookinetes. In this study, a group of 24 proteins were classed as invasive or surface proteins. Other significant groups included chaperones and signalling molecules (and ribosomal contaminants), but by far the largest group of proteins produced by this study were unknown proteins, highlighting the importance of characterising new proteins in the search for new drug and vaccine candidates. Although this study involved ookinete micronemes, it is reasonable to presume the case is similar for merozoite micronemes and other organelles. Transcriptome data from *P. falciparum* merozoites shows a poignant induction of gene expression of known micronemal proteins late in schizont stage development. Of the 262 open reading frames (ORFs) that shared their expression pattern, 189 had no assigned function (Bozdech et al., 2003). Given that micronemal proteins are required for invasion and are the target of invasion-inhibitory antibodies, they represent important vaccine candidates. Characterisation of the unknown, late expressed proteins could uncover a whole host of new vaccine candidates or drug targets that are badly needed in the fight against malaria.

1.7 Project Aims

The aim of this project is to uncover novel proteins that play a part in the invasion of red blood cells by merozoites. Identifying components of this essential process will not only lead to a better biological understanding of the parasite, but may well identify potential drugs targets and perhaps even vaccine candidates.

Comparing the *P. falciparum* 3D7 genome with other sequenced genomes predicted the function of only c. 40% of the 5400 genes, leaving the remaining 60% with no identified orthologue in any other sequenced species and hence no assigned function (Gardner et al., 2002). In order to identify which of these proteins are important for invasion of rbcs, the list must be refined utilising the wealth of genomic, transcriptomic, proteomic and bioinformatic data available. I intend to characterise a small number of proteins of the malaria parasite *Plasmodium falciparum* using the following criteria for selection:

a) <u>The protein is expressed late during schizogony in the blood stage cycle</u>
From previous studies it appears that if we are to find proteins that are essential for invasion of erythrocytes, we need to look to apical organelles. AMA-1 is a micronemal protein essential for invasion. The selection process aims to identify proteins expressed at a similar time during the asexual cycle as AMA-1, in late schizogony just before merozoite release, in the hope they might be trafficked to the micronemes, rhoptries or the cell surface.

b) <u>The protein should have a signal peptide/anchor sequence</u>
Signal sequences direct emerging or newly synthesised proteins to the endoplasmic reticulum (ER), and into the eukaryotic secretory pathway. Proteins destined for the apical organelles utilise the secretory pathway in order to reach their destination. Therefore the proteins of interest to be studied must contain signal peptide/anchor sequences in order for them to be directed to the apical organelles.

c) The protein should have a transmembrane domain or a GPI anchor

Glycosylphosphatidylinositol (GPI) anchored proteins are abundant on the merozoite's surface and many are thought to have a role in host cell selection and attachment of the merozoite to the rbc. Recentlly MTRAP, a micronemal transmembrane protein has been identified as one of the ligands linking the adhesion receptor on the host cell surface with the parasite motor complex. Other as yet unidentified molecules however might be able to fulfil a similar function. Therefore we are looking into proteins containing transmembrane domain(s), in which the cytoplasmic tail could be involved in signalling pathways or directly in motor complex binding. Indeed many proteins known to be involved in invasion such as AMA-1 and the DBL-EBP and PfRH protein families, contain a single transmembrane domain.

d) The protein must be conserved across other *Plasmodium* species

The erythrocyte invasion process is common among all *Plasmodium* species, and indeed across all Apicomplexa, thus one would expect genes critical for invasion to be conserved throughout the genus.

Utilising a variety of experimental techniques, I aim to characterise the selected proteins, uncovering features that provide clues to their role within the parasite.

2. Materials and Methods

2.1 Molecular Biology Techniques

2.1.1 Protein Selection and Bioinformatic analysis

Proteins for study were selected by searching the *P. falciparum* genome using the database PlasmoDB (plasmodb.org) and the following parameters: maximal expression at 42 ± 2 hours post invasion, minimal expression 24 ± 8 hours, presence of a signal peptide, minimum of one transmembrane domain and presence in genomes of all other annotated *Plasmodium* spp. Signal peptide predictions were confirmed using the program SignalP (http://www.cbs.dtu.dk/services/SignalP/) and transmembrane regions verified using TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/). Protein sequence alignments were performed using ClustalW Multiple Alignment software.

2.1.2 Primers

Primers were designed based on the published genomic sequence of 3D7 available on PlasmoDB (plasmodb.org) and ordered from Sigma-Aldrich. Primer sequences are detailed in the table 1.

2.1.3 Polymerase Chain Reaction (PCR)

AmpliTaq DNA polymerase (Applied Biosystems) was used with 1 mM dNTPs, 2 mM $MgCl_2$ and 0.5 pM primer set to amplify gene fragments from genomic or cDNA for vector construction, PCR reactions were carried out in a Mastercycler Gradient PCR machine (Eppendorf). The amplified DNA was purified using QIAquick® PCR purification kit (Qiagen) according to manufacturer's instructions.

2.1.4 DNA-modifying enzymes

Restriction digests were carried out using endonucleases from New England Biolabs (NEB). Digested plasmid DNA was phosphatase treated using calf intestine alkaline phosphatase (Roche). Digested DNA fragments were purified using Qiaex® II gel extraction kit (150) or QIAquick® Gel Extraction kit (both from Qiagen) according to manufacturer's instructions. The resulting DNA fragments were ligated using T4 DNA

ligase (Roche). For ligation independent cloning into pET30- or pET32-Xa/LIC vectors, T4 DNA polymerase (Novagen) was used according to the manufacturers' instructions.

2.1.5 Purification of plasmid DNA

E.coli was grown in Luria-Bertani broth (LB) (Bertani, 1951). Plasmid DNA was purified using QIAprep® Spin Miniprep Kit or HiSpeed® Plasmid Maxiprep kit (both from Qiagen) according to the manufacturers' instructions. DNA yields were analysed using a Nanodrop spectrophotometer (Thermoscientific) or by running an aliquot on a 1% agarose (Roche) gel.

2.1.6 Nucleotide Sequencing

Plasmid DNA was sequenced by Geneservice Ltd, now renamed Source Bioscience, Cambridge.

2.1.7 Transfection vector construction

Parasite vectors for both single homologous genomic integration and episomal expression were based on the pHH3 transfection plasmid (Knuepfer and Holder, unpublished). These plasmids contain the Blasticidin resistance cassette. For C terminal epitope tags, regions of homology to the gene of interest were cloned into pHH3 using *EcoR*I and *Sac*II. FLAG and TY tags were introduced to the vector backbone on the primer and amplified with the region of homology. Codons were optimized for *P. falciparum* as based on previous research (Peixoto et al., 2003). Each tag was flanked with AvrII and SacII restriction sites. Truncation of functional gene product was attempted by single homologous recombination. Nucleotide regions of homology were amplified and cloned in between *Eco*RI and *Bam*HI sites of pHH3. This restriction enzyme digest removes the existing promoter and terminator region within pHH3 so that after homologous integration either no stable RNA would be produced or if expression of the gene occurred a truncated product would result in a functional knockout. Double homologous recombination vectors were based on the pHTK transfection vector (Duraisingh et al., 2002). This plasmid contains the human dihydrofolate reductase (*hdhfr*) selection cassette, conferring resistance to the antifolate drug WR99210 (Jacobus Pharmaceuticals). This vector was used to attempt to knock out PF11_0443 by double homologous recombination. A region form the 5'UTR and ORF were cloned into MCS 1

using *Sac*II and *Bgl*II and a region of 3'end of the ORF and 3'UTR were cloned into MCS 2 using *Eco*RI and *Avr*II. Vector backbones were generously provided by Ellen Knüpfer. Transfection vectors, cloning and transfection strategy are illustrated in figure 2.1. Primers used to amplify flanking regions are described in table 2.1.

2.1.8 Southern Blot

DNA preparation, separation and transfer

Genomic DNA was prepared from WT 3D7 and transgenic parasite lines and digested using suitable restriction enzymes. The resulting fragments were separated by agarose gel electrophoresis containing 0.5µg/ml ethidium bromide at 30V overnight. The gel was agitated in 0.25mM HCl for ten minutes, washed with H_20, then incubated in denaturing buffer (1.5M NaCl, 0.5M NaOH) for 1 hour, followed by 2x30 minute incubations with neutralising buffer (1.5M NaCl, 0.5M Tris, 1mM EDTA pH 8.0). The DNA was transferred onto a Gene Screen Plus membrane (Whatman) in 10x SSC (1xSSC: 150nM NaCl, 15mM Na Citrate pH7.0) by capillary transfer (Sambrook J., 1989). Following transfer the DNA on the membrane was denatured in 0.4M NaOH for 1 min before neutralisation, dried then UV crosslinked using a UV-Transilluminator (UVITEC Cambridge).

Hybridisation

The membrane was pre-hybridised in hybridisation buffer (6x SSC, 5x Denhardts [0.1% BSA, 0.1% Ficoll, 0.1% polyvinylpyrrolidone), 0.5% SDS, 0.2mg/ml salmon sperm DNA) rotating for 1 hour at 62°C. The hybridisation probe was labelled with α-[^{32}P] dATP (Perkin Elmer®) by random priming using the Decaprime II kit (Ambion) according to manufacturer's instructions. Unincorporated labelled nucleotides were removed using ProbeQuant™ G-50 Microcolumns (Amersham Biosciences). The membrane was incubated overnight at 62°C with the labelled probe in hybridisation buffer. The membrane was washed twice with 6x SSC , 0.1% SDS for 20 minutes at 62°C. The blot was visualised by autoradiography using Kodak BioMax™ MR film (Kodak®) following incubation at -80°C. If necessary, the probe was stripped off the membrane by boiling in 0.05x SSC, 0.01M EDTA, 0.1% SDS which allowed reprobing.

2.1.9 Site-Directed Mutagenesis (SDM)

SDM was used to introduce single amino acid changes into the transfection constructs with the intent to remove phosphorylation sites within the protein of interest. SDM was carried out using QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies) according to the manufacturer's instructions.

2.2 Recombinant proteins and Immunochemistry

2.2.1 Protein expression vectors

DNA fragments from genes of interest were amplified by PCR (section 2.1.3) and cloned into protein expression vectors pET30-Xa/LIC, pET32-Xa/LIC or pET46-Ek/LIC (Novagen) according to the manufacturers' instructions.

2.2.2 Competent cells and transformations

E. coli Giga single cells (Novagen) were used for propagation of plasmid DNA. *E. coli* BL21-*DE3* and BL21-*DE3 pLysS* (Stratagene) were used for protein expression. Transformations with plasmid DNA were carried out according to the manufacturers' instructions.

2.2.3 Expression of Recombinant Proteins

Protein expression vectors containing the correct sequence were transformed into either BL21 DE3 or BL21 (DE3) pLysS *E.coli* . Multiple colonies were picked and grown in 10ml of LB with antibiotic overnight at 37°C. To test expression, 100µl of overnight culture was added to 5 ml LB with antibiotic and grown to OD_{600} =0.6. Protein expression was induced using 1mM Isopropyl β-D-thiogalactopyranoside (IPTG) for 4hrs at 37°C or 22°C in an attempt to increase solubility. To rapidly screen for solubility, 2ml of culture was pelleted at 9600xg for 2 minutes. The pellet was lysed in 300 µl Bugbuster (Novagen) plus 1x complete, EDTA-free protease inhibitors (Roche) and 1µl Benzonase nuclease (Novagen) to reduce the viscosity and incubated for 20 minutes at room temperature with shaking. After centrifugation at 17000xg for 10 minutes at 4°C soluble (supernatant) and insoluble (pellet) proteins were collected and analysed by SDS-PAGE.

2.2.4Sodium Dodecyl Sulphate Polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed using pre-cast NUPAGE gels from Invitrogen in MES or MOPS running buffer (Invitrogen). Samples were run in either reducing or non-reducing two times Laemmli sample buffer (Sambrook J., 1989). Gels were stained with either coomassie blue (Sambrook J., 1989), the Mass Spectrometry-compatible Novex® Colloidal Blue Staining Kit (Invitrogen) or silver stain following a previously described method (Blum et al., 1987).

2.2.5 Protein purification

6xHis tagged proteins were purified on a Nickel resin. Bacterial lysate supernatant was incubated with Nickel-NTA beads (Qiagen) at 4°C for 1 hour. This was transferred into a disposable 10ml polypropylene column (Thermo-Scientific) and the resin washed in 50mM sodium dihydrogen orthophosphate, 300mM sodium chloride and 10mM imidazole, pH8.0. Protein was eluted by washing with the same buffer but containing 250mM imidazole. For insoluble protein, inclusion bodies were purified as described in the pET systems manual (Invitrogen), solubilsed in 8M urea and purified on a Nickel-NTA column by the same method but with the addition of 8M urea to the wash and elution buffers.

2.2.6 Antibody production

Antibodies to each recombinant protein were raised in either rabbits or rats by Harlan, UK, following standard protocol. The resulting antibodies and their working dilutions are listed in table 2.4.

2.2.7 Purification of antibodies by affinity chromatography

An affinity column was generated using CNBr activated sepharose 4B (Pharmacia). For each column, 0.5g of sepharose was reconstituted and washed in 1mM HCl, dried under vacuum and incubated with 20mg of purified recombinant protein in 0.1M sodium bicarbonate pH8.3, 0.5M sodium chloride, rotating overnight at 4°C. The resin was transferred into a column and the flow through tested by Bradford assay to ensure the protein had bound. Free CN-Br arms were blocked in 10 column volumes of 0.2M glycine pH8.0 then washed with 10 column volumes of 0.1M sodium bicarbonate pH8.3, 0.5M

NaCl followed by 0.1M sodium acetate pH4.0, 0.5M NaCl followed by 0.1M sodium bicarbonate pH8.3, 0.5M NaCl and finally in two column volumes of PBS. To bind antigen-specific antibodies, rabbit serum was diluted 1:1 in 10mM Tris pH7.5 and passed over the column 5 times. Then the column was washed with 20 column volumes of 10mM Tris pH7.5 and 10mM Tris pH7.5, 0.5mM NaCl. Antigen-specific antibodies were eluted with 15ml of 100mM glycine pH2.5 and collected in a tube containing 1 bed volume of 1M Tris pH8.0 to neutralise. The column was washed with 10mM Tris pH7.5 returning the column to pH7.5 and stored at 4°C in PBS containing sodium azide.

2.2.8 Antibody purification by immunoblot strip purification

Antigen solubilised in 8M urea was diluted 1:1 in reducing Laemmli sample buffer and resolved by SDS-PAGE in a 1 well NUPAGE 12% Bis-Tris polyacrylamide gel (Invitrogen).The protein was then transferred onto a Whatman Protran nitrocellulose transfer membrane in a western blotter at 13V overnight. The membrane was stained with Ponceau S and a strip containing the protein was cut out of the membrane, washed in PBS + 0.2% Tween 20 (PBST) and blocked in PBST containing 5% (w/v) non-fat milk powder for 1 hour. The membrane was then washed 3 times in PBST then incubated overnight in rabbit serum diluted 1:1 in PBS, rotating at 4°C. The membrane was again washed 3 times in PBST. To elute the bound antibodies, the membrane strip was placed into a 2 ml tube containing 100mM glycine pH2.5 for no more than 2 minutes after which the solution was neutralised by addition of 1/10 volume of 1M Tris pH8.0. The membrane was removed and washed 5 times in PBST. This was repeated several times. Purified antibodies were combined and concentrated using a Vivaspin 20 concentrator with 10kDa molecular weight cut off (Sartorius Biolab Products). The antibody concentration was estimated using the OD_{280} on a spectrophotometer.

2.2.9 IgG purification

To purify IgG, antibodies/serum were diluted in 10mM Tris pH7.5 and incubated with protein G sepharose 4B (Sigma-Aldrich) overnight, rotating at 4°C. The sepharose was transferred into a column and washed with 10 column volumes 100mM Tris pH8.0 then 10mM Tris pH8.0. IgG antibodies were eluted with 10 column volumes of 100mM glycine pH2.5 and collected in a tube containing 1 bed volume of 1M Tris pH8.0 to neutralise.

The column was washed with 10mM Tris pH7.5 until the column was pH7.5 and stored at 4°C in PBS containing sodium azide.

2.2.10 His-Antibody depletion

Similar to the affinity chromatography described above. CN-Br activated sepharose 4B (Pharmacia) was bound to the Txr and His-tag expressed as soluble protein from pET32-Xa/LIC (glycerol stock provided by Dr. Judith Green - NIMR) and antibodies were depleted of His-reacting antibodies by batch incubation with resign and collection of flowthrough.

2.2.11 Western Blot

Proteins from SDS-PAGE were transferred onto Whatman Protran nitrocellulose transfer membrane in a western blotter in transfer buffer (Invitrogen) for 3hrs at 35V or overnight at 13V. Blots were blocked in blocking buffer (5% milk powder in PBS, 0.2% Tween 20) for 1h or overnight. Primary antibodies were diluted in the blocking buffer and incubated at room temperature for 1 h. Blots were washed 3 times in PBST and incubated with HRP-linked secondary antibodies diluted in blocking buffer for 1h. The blots were then washed 3 times in PBST. Enhanced chemoluminescence (ECL) detection reagents (GE Healthcare) were used for visualisation on Biomax™ MR film (Kodak®).

2.2.12 Solubility Assay

Percoll-purified schizonts or early ring stage parasites were hypotonically lysed in 10 times the pellet volume of 10mM Tris pH8.0 plus 1x complete protease inhibitors (Roche) for 1 h at 4°C, with vortexing every 10 minutes. Parasite DNA was sheared using a 25G needle. Soluble and insoluble materials were separated by centrifugation at 100000xg for 30 minutes. The supernatant was collected and the pellet washed in the same volume of hypotonic lysis buffer to ensure all soluble material was removed. The pellet was resuspended in 0.1M sodium carbonate pH11 and incubated at 4°C for 1 h, with vortexing every 10 minutes. Again, to separate the soluble and insoluble membrane bound fractions, centrifugation at 100000xg for 30 minutes was performed. The pellet was washed in carbonate lysis buffer before solubilsation in Laemmli sample buffer and boiled for 5 min. The samples were then analysed by Western Blot.

2.2.13 Immunofluorescence Assay (IFA)

2.2.13.1 *Fixed IFA*

Thin films of *P. falciparum* stage specific parasite material were air dried then fixed in 4% paraformaldehyde for 30 minutes, permeabilised in 0.1%Triton X-100 in PBS for 10 minutes, blocked in 3% bovine serum albumin (BSA) in PBS and stored at 4°C. Primary antibodies, diluted in 3% BSA in PBS, were incubated for 1h at room temperature. The slide was washed 3 times in PBS before adding secondary antibodies (Invitrogen) in 3% BSA. For nuclear staining samples were stained with 5µg/µl of 4,6-diamidino-2-phenylindol (DAPI), then using Vectashield (Vector Labs) on a coverslip were mounted and the slide was stored at 4°C. Images were acquired using a Zeiss Axioplan 2 Imaging system (Carl Zeiss, Germany) and AxioVision 3.1 software.

2.2.13.2 *Unfixed or live IFA*

Highly synchronized, rupturing 3D7 schizonts were allowed to burst in RPMI 1640 medium, pelleted at 1500x*g* for 3 minutes at 4°C then resuspended in 3% BSA in PBS. Primary antibody was added at the correct dilution (see Table 2.1 for antibody dilutions) and incubated for 30 minutes at room temperature. The samples again were pelleted as before and resuspended in the correct dilution of secondary antibody in 3% BSA in PBS, incubating for 30 minutes after which samples were pelleted and resuspended in 50µl 3% BSA in PBS. 2µl of this suspension was transferred to glass slide and mounted with a coverslip. Images were acquired using a Zeiss Axioplan 2 Imaging system (Carl Zeiss, Germany) and AxioVision 3.1 software.

2.2.14 Immunoprecipitation of schizont stage parasites

Schizonts were lysed in 10 x pellet volume with lysis buffer (1% Triton X-100, 50mM Tris-HCl pH8.0, 5mM EDTA, 5mM EGTA, 150mM NaCl, 1x complete protease inhibitors [Roche]) and incubated on ice for 1hr, vortexing every 10 minutes. A 25G needle was used to shear the parasite DNA. The samples were centrifuged at 17000xg for 30mins at 4°C and the supernatant was transferred into a new tube. To reduce background and non-specific binding, the sample was pre-absorbed with 150µl of protein G sepharose (Sigma-Aldrich) and 20µl of pre-immune rabbit serum (Harlan) for 1 hour, rotating at 4°C. The samples were centrifuged at 17000xg at 4°C. The supernatant was collected and

split between 2 tubes; to the first 5µl of purified specific antibody was added and to the second 5µl of pre-immune serum. Incubation lasted overnight, rotating at 4°C. To pull-down the antibody/protein complexes, 50µl of protein G sepharose was added and rotated at 4°C for 1 hour. The samples were centrifuged at 17000x*g* for 2 minutes at 4°C. Supernatant was discarded and the pellet washed 6 times in lysis buffer with the addition of 300mM NaCl. The pellet was resuspended in 50µl reducing Laemmli sample buffer, boiled for 5 minutes and the supernatant analysed using SDS-PAGE.

2.2.15 Immunoprecipitation of ^{35}S radiolabelled schizonts

^{35}S labelled schizonts were prepared by M. Grainger. The immunoprecipitation protocol was identical to 2.2.14 above. After protein separation on SDS-PAGE the gel was fixed in 10% acetic acid and 25% isopropanol in water for 30 minutes, then the signal amplified using Amplify (GE Healthcare) for 30 minutes before drying the gel under vaccum at 70°C for 2 hours. To visualise, the dried gel was exposed to Biomax MR film (Kodak) for 1-21 days.

2.2.16 Large-scale Immunoprecipitation for Mass Spectrometry analysis

To identify binding partners 5ml of purified, pelleted schizont, (46 ± 2 hour p.i. prepared by M. Grainger) were lysed and prepared as in 2.2.15. The cleared lysate was firstly passed through a pre-immune antibody affinity column, after which the flowthrough was applied to a specific antibody affinity column. Both columns were washed with 20 column volumes of wash buffer 1 (lysis buffer:1% Triton X-100, 50mM Tris-HCl pH8.0, 5mM EDTA, 5mM EGTA, 150mM NaCl, 1x complete protease inhibitors [Roche]) followed by 20 column volumes of wash buffer 2 (lysis buffer minus NaCl: 1% Triton X-100, 50mM Tris-HCL pH8.0, 5mM EDTA, 5mM EGTA, 1x complete protease inhibitors [Roche]) before eluting protein from the column with 10 column volumes of 100mM glycine pH2.5, 0.01% Triton X-100. The eluate was neutralised by addition of 1/10 volume of 1M Tris-HCl pH 8.0. Samples were concentrated using Vivaspin 500 concentrator with 10kDa molecular weight cut off (Sartorius Biolab Products) and resolved by SDS-PAGE, staining with either Instant Blue™ or silver-nitrate (2.2.4). Visible bands were excised and analysed by Liquid Chromotography-tandem Mass Spectrometry (LC-MSMS) at PNAC, University of Cambridge.

2.2.17 Large-scale Immunoprecipitation with PF11_0443-FLAG schizonts

Approximately 3ml of purified (section 2.3.1), pelleted PF11_0443-schizonts, were lysed and prepared as in 2.2.15. The cleared lysate was bound in batch to anti-FLAG M2 affinity gel (Sigma Aldrich). The same volume of WT 3D7 schizonts were treated the same way and bound to FLAG affinity gel to control for non-specific binding. Both samples were applied to a spin column (Sartorius Biolab Products)and were washed with 20 column volumes of TBS (Tris Buffered Saline: 50 mM Tris, 150 mM NaCl) before eluting protein from the column with 5 column volumes of 150 nM 3xFLAG peptide in TBS. Samples were concentrated using Vivaspin 500 concentrator with 10kDa molecular weight cut off (Sartorius Biolab Products) and resolved by SDS-PAGE, staining with the fluorescent stain SYPRO® Ruby according to manufacturer's instructions. Visible bands were excised and analysed by Liquid Chromotography-tandem Mass Spectrometry (LC-MSMS) in house by Dr. Steve Howell, Division of Molecular Structure.

2.3 Parasite cell culture

2.3.1 Maintenance and synchronisation

P. falciparum 3D7 was cultured in RPMI-1640 with Albumax medium (GIBCO™), with the addition of 2 mM L-Glutamine, at 2-4% haematocrit and gassed with a mixture of 7% CO_2, 5% O_2, and 88% N_2(Trager and Jensen, 1976). Mature stage parasites were isolated on a 70% (v/v) Percoll (GE Healthcare) density gradient as described (Rivadeneira et al., 1983; Dluzewski et al., 1984). Further synchronisation using 5% D-sorbitol was performed as previously described (Trager and Jensen, 1976; Lambros and Vanderberg, 1979). Frozen stocks were prepared from 200-300µl of pelleted cultures (828xg [2100rpm, Heraeus Megafuge 16, Thermo Scientific], 4 minutes, room temperature) containing ~5% ring stage parasites by the drop-wise addition of 700µl Freezing solution (111mM NaCl, 166mM D-sorbitol, 28% glycerol). Samples were transferred to cryovials (Nunc™) and stored in liquid nitrogen. In order to revive frozen samples, stocks are thawed quickly in a 37°C water bath and transferred to a larger tube. One volume of thawing solution (3.5% NaCl) was added drop-wise to the stock, which was then centrifuged at 270xg (1200rpm in the same centrifuge as above) for 4 minutes at room

temperature and the supernatant discarded. The addition of one volume thawing solution and following centrifugation was repeated twice more, before adding warmed culture medium and transferring to a culture flask.

2.3.2 Preparation of parasite DNA

P.falciparum cell culture with parasitaemia > 5% of any stage was pelleted (828xg [2100rpm, Heraeus Megafuge 16, Thermo Scientific], 4 minutes, room temperature) and resuspended in 4 pellet volumes of room temperature Buffer A (50mM sodium acetate pH5.2, 100mM NaCl, 1mM EDTA). Cells were lysed by the addition of 1 pellet volume of 18% SDS, inverting the tube to mix, then incubated at room temperature for 2 minutes. An equal volume of phenol/chloroform (Sigma Aldrich) was added, mixed and the sample was centrifuged at 1028xg (2100rpm, Heraeus Megafuge 16, Thermo Scientific) for 14 minutes at room temperature. The parasite DNA was removed from the aqueous phase of the lysate to a 15ml Corex glass tube then precipitated with 2.5 pellet volumes of cold EtOH plus 1/10 volume of 3M sodium acetate, pH5.2, overnight at -20°C. The DNA was pelleted (12000xg, 15 minutes, 4°C) and the pellet air dried. The sample was then resuspended in 500ml TE (10mM Tris, 1mM EDTA, pH8.0) and the phenol/chloroform step repeated. The DNA was removed to a fresh Eppendorf tube and precipitated by addition of an equal volume of isopropanol plus 1/10 volume of 3M sodium acetate, pH5.2,. The DNA was pelleted (17000xg, 10 minutes, 4°C) then washed with 80% EtOH before air drying and resuspension in an appropriate volume of TE, depending on pellet size.

2.3.3 Preparation of parasite RNA and cDNA

P.falciparum 3D7 RNA was prepared as previously described (Kyes et al., 2000). In brief parasitized erythrocytes were pelleted (828xg [2100rpm, Heraeus Megafuge 16, Thermo Scientific], 4 minutes, room temperature) and resuspended in 10-20 volumes (10 volumes for early rings, 20 for late schizonts) of pre-warmed (to 37°C) TRIzol (Gibco Life Technologies) and agitated until clumps dissolved. Samples were stored at -80°C at this stage. To continue with RNA preparation, samples were thawed on ice and 2 pellet volumes of chloroform added. Samples are agitated vigorously for 15 seconds and incubated at room temperature for 3 minutes after which the sample was centrifuged at

4700xg (7000rpm, Heraeus Fresco 17, Thermo scientific) for 30 minutes at 4°C. The aqueous supernatant was removed to a fresh tube, avoiding the interface and possible DNA contamination. To this, 0.5 times the original TRIzol volume of ice cold isopropanol was added, mixed and left to incubated on ice for 2 hours before pelleting the RNA at 4°C, 12000xg for 30 minutes. The supernatant was removed and the pellet washed in 75% ethanol, pelleting as before. The RNA pellet was air dried before resuspending in a small volume of DEPC treated water. The concentration of the RNA was estimated by spectrophotometry. To remove salt and DNA contamination and concentrate the RNA, RNeasy® Minielute™ Cleanup kit (Qiagen) was used, following manufacturer's instructions. The Reverse Transcription System (Promega) was used to produce cDNA, according to the manufacturer's instructions.

2.3.4 Time-course Analysis

P. falciparum 3D7 blood stage parasites were tightly synchronised so that egress and re-invasion within the culture would complete in 1-2 hours. Upon re-invasion (time point zero) protein samples were taken and thin smears were made from the culture for Giemsa staining and use in indirect immunofluorescence assays. Further samples were taken every 4 hours until schizont egress/reinvasion which took place at approximately 46.5 h p.i. Each protein sample contained 50µl of pelleted cells and the number of cells per sample was counted using a haemocytometer. Parasitaemia was counted for each time point and averaged (as each time point came from the same starter culture) and this was used to calculate the number of parasites per sample. Due to the abundance of haemoglobin from the rbcs, samples were hypotonically lysed in 10mM Tris, pH8.0 and the hypotonic pellet was resuspended in 2x reducing Laemli sample buffer to a concentration of $2x10^5$ parasites/µl. Protein samples were separated by SDS-PAGE with $2x10^6$ parasites (10µl) per well. Hypotonically lysed rbcs were included in the western blot to control for parasite versus rbc specific band recognition. Protein expression was analysed by western blot using the polyclonal antibodies raised againt the recombinant proteins. A glycophorin A/B antibody was used as a loading control. Antibodies against the parasite ER resident protein BiP were used to confirm the presence of parasites in each lane and MSP-2 was used as a late stage marker. Images of Giemsa stained

parasites from each time point were taken to demonstrate parasite morphology at each stage.

2.3.5 Transfection

Ring-stage *P. falciparum* 3D7 parasites were transfected using previously described methods (Wu et al., 1996; Crabb et al., 1997). In brief, 100µg of transfection construct DNA was previously ethanol precipitated and dissolved in 15µl TE (10mM Tris, 1mM EDTA, pH8.0). On the day of transfection 385µl of sterile cytomix (120mM KCl, 0.15mM $CaCl_2$, 2mM EDTA, 5mM $MgCl_2$, 10mM K_2HPO_4/KH_2PO_4 pH7.6, 25mM Hepes pH 7.6) was added to the 100µg of DNA. 200 ul of ring stage parasitized RBCs (parasitaemia 8-10%) were collected by centrifugation at 800 xg and mixed with the dissolved DNA in cytomix before transferring to a Gene Pulser® Cuvette, 0.2cm (Bio-Rad) and electroporation at 0.31kV, 950µF, ∞ resistance. Transfected cells were immediately transferred to a dish with10ml RPMI-Albumax containing 3% rbcs. The appropriate drug for selection was added with the medium 24 hours post transfection and medium containing selection drug was replaced every day for 7 days, prior to replacing it every 48h (10nM WR99210; Jacobus Pharmaceuticals, Princeton, New Jersey, United States) for pHTK or pHH4 vectors, 2.5µg/ml Blasticidin-S (Calbiochem®) for the pHH3 vector – see transfection vector table, Table 2.2). Parasitaemia was checked every 2 days by thin blood smear stained with Giemsa and typically 2-4 weeks after transfection viable parasites appeared. Parasites were frozen as described in 2.3.1 and genomic DNA was prepared as in 2.3.2. For cultures containing episomal construct, drug pressure was continuous so the parasite maintains the plasmid DNA. For integration constructs parasites are cycled on-and-off drug. Drug selection pressure was removed for 3 weeks before reapplication, when frozen stocks and genomic samples were prepared. This was repeated for 3 to 5 cycles and integration status was analysed by PCR, Southern Blot or immunoblotting . For transfectants with double-crossover homologous recombination constructs (pHTK), after each drug cycle on WR99210 the culture was additionally put onto 10 µM gancicolvir (Sigma Aldrich). Any parasite containing the thymidine kinase (TK) cassette (parasites with episome or integration of a single flank) will be susceptible

to the drug. Integration was analysed by Southern blot. Any integrated lines were cloned by limiting dilution (Rosario, 1981; Kirkman et al., 1996).

Table 2.1: *List of primers*

	Primer	Purpose	Sequence
A	MAL13P1.94fwd	Protein expression	5′-GGTATTGAGGGTCGCCGCTTATGCTTGGAA GATTTAAGACATC-3′
B	MAL13P1.94rev	Protein expression	3′AGAGGAGAGTTAGAGCCTTATACTTTAATTAGATAT TCTTTTTCGTC-5′
C	PF11_0443fwd	Protein expression	5′- GGTATTGAGGGTCGCGAAGATTATGATGAA CATGATAAG-3′
D	PF11_0443rev	Protein expression	3′- AGAGGAGAGTTAGAGCCCTAATCGTTGTAA TTAAAACTTAAAC-5′
E	PF14_0045fwd	Protein expression	5′-GGTATTGAGGGTCGCAACAAAATATCGAGT GCTAGAGATGATATTC-3′
F	PF14_0045rev	Protein expression	3′-AGAGGAGAGTTAGAGCCTTAGGAACCTTAT TGTCTGTCAATACCTTC-5′
G	PF02_0040fwd (optimised gene)	Protein expression	5′-GGTATTGAGGGTCGCAATGAAATAAAAAAC GCAATGAGTTC-3′
H	PF02_0040rev (optimised gene)	Protein expression	3′- AGAGGAGAGTTAGAGCCTTATTTCAGTTTCA CCACATCTTC-5′
I	MAL13P1.94_KO flank1FWD	Transfection: Integration	5′-CATCATGAATTCGATTTATTATGAATATTTTA GAAAATGACATC-3′
J	MAL13P1.94_KO flank1REV	Transfection: Integration	3′-CATCATGGATCCAGCTAAATAATCCCCAATA GTCC-5′
K	MAL13P1.94_pH H3tagFWD	Transfection: Integration	5′-CATCATGAATTCTTATAGAATATAAACAGAA GAGGACATG-3′
L	MAL13P1.94_pH H3TYrev	Transfection: Integration	3′-CATCATCCGCGGATCTAATGGATCTTGGTT GGTGTGAACTTCCCTAGGTACTTTAATTAGATATTCTT TTTCGTCG-5′
M	MAL13P1.94_pH H3x3FLAGrev	Transfection: Integration	3′-CATCATCCGCGGTTTATCATCATCATCTTTG TAATCAATATCGTGATCTTTGTAATCACCATCGTGATC TTTGTAATCTACTTTAATTAGATATTCTTTTTCGTCG-5′
N	PF11_0443_KOfl ank1FWD	Transfection: Integration	5′-CATCATCCGCGGTTTATTATAATAAATGTATA TATTAGAAAAGG-3′
O	PF11_0443_KOfl ank1REV	Transfection: Integration	3′-CATCATAGATCTCTTGTACTTTTATTTACTAA TACGTTG-5′
P	PF11_0443_KOfl ank2FWD	Transfection: Integration	5′-CATCATGAATTCAATAGAAGAACAAATTGAA GAAGAGG-3′
Q	PF11_0443_KOfl ank2REV	Transfection: Integration	3′-CATCATCCTAGGTATTTACTCTTGTGTTGTC GAGC-5′
R	PF11_0443_pHH 3tagFWD	Transfection: Integration	5′-CATCATGAATTCATTAAAAAGGGTAATACAG AAATCATC-3′
S	PF11_0443_pHH 3tyREV	Transfection: Integration	3′- CATCATCCGCGGATCTAATGGATCTTGGTTGGTGT GAACTTCCCTAGGATCGTTGTAATTAAAACTTAAACC-5′
T	PF11_0443_pHH 3x3flagREV	Transfection: Integration	3′-CATCATCCGCGGTTTATCATCATCATCTTTGTAATC AATATCGTGATCTTTGTAATCACCATCGTGATCTTTGT AATCCCTAGGATCGTTGTAATTAAAACTTAAACC-5′

U	PF02_0040fwdKO	Transfection: Integration	5′-CATGAGGAATTCCCTATTCATTAGCTAG CCAAA-3′
V	PF02_0040revKO	Transfection: Integration	3′-CTCAGCGGATCCGTGATAATATTTGTAG GATGTATTTG-5′
W	PF02_0040fwdTAG	Transfection: Integration	5′-CATGAGGAATTCCCTATTCATTAGCTAG CCAAA-3′
X	PF02_0040revTYtag	Transfection: Integration	3′-CACTAGCCGCGGTTAATCTAATGGATCTTGGT TGGTGTGAACTTCCCTAGGTTTGCTTTTTTTCTTTT TCTTC-5′
Y	PF02_0040revFLAGtag	Transfection: Integration	3′-CACTAGCCGCGGTTATTTATCATCATCATCTTTGTA ATCAATATCGTGATCTTTGTAATCACCATCGTGATCTT TGTAATCCCTAGGTTTGCTTTTTTTCTTTTTCTTC-5′
Z	PF02_0040gfpFWD	Transfection: Episomal expression	5′-CATGGACCTAGGATGTTCATATTATACA TAATTTTTGTTGTTATTCC-3′
AA	PF02_0040gfpREV	Transfection: Episomal expression	3′-CATGGACCTAGGTTTGCTTTTTTTCTTTT TCTTCTTCGAATTC-5′
BB	MAL13P1.94_5_UTR	Diagnostic primer	5′-GTAGTTTTCTCAATAAATAGATATTATCA TGGC-3′
CC	MAL13P1.94_3_UTR	Diagnostic primer	3′-CTATAGAAAATATGCTCATTCCTGTCAT G-5′
DD	PF02_0040_5_UTR	Diagnostic primer	5′-AATTCCATAAAGCATTTCTCTTAAAATGT GTC-3′
EE	PF02_0040_3_UTR	Diagnostic primer	3′-TAGAAAATAAAAATAAGAAAATTAAATTA TGATTGAAGTGAC-5′
FF	PF11_0443diagnostic1FWD	Diagnostic primer	5′-ATCTTTTAATCCCTTATTAAATTGTTACAT ATC-3′
GG	PF11_0443diagnostic2FWD	Diagnostic primer	5′-TAGAAAAGGATTAAAAAAAGGGACATT G-3′
HH	PF11_0443FLAGsenseStoA	Site directed mutagenesis	5'-AAAACAAATAAGAATGAAGAAAAAAGGTTTA GCTTTTAATTACAACGATCCTAGGGATTACAAA-3'
II	PF11_0443FLAGantisenseStoA	Site directed mutagenesis	5'-TTTGTAATCCCTAGGATCGTTGTAATTAAAA GCTAAACCTTTTTTCTTCATTCTTATTTGTTTT-3'
JJ	PF02_0040signalanchorGFPfwd	Transfection: Episomal expression	5′-CTAGAATGTTCATATTATACATAATTTTT GTTGTTATTCCATTTTTAATATTTGCTTATATCATA TATAAAAATGTCC-3′

KK	PF02_0040signal anchorGFPrev	Transfection: Episomal expression	3'-CTAGGGACATTTTTATATATGATATAAGC AAATATTAAAAATGGAATAACAACAAAAATTATG TATAATATGAACATT-5'
LL	PF11_0443TM2f wd	Transfection: Episomal expression	5'-CATGGACCCGGGCAAGAAATTGTTATGATTA ATAATCAGAAGACTAAG-3'
MM	PF11_0443TM2f wd	Transfection: Episomal expression	5'-CATGGACCCGGGATCTTCATTTAATATATTA TATTTAATAACCCATTTTATATTCC-3'
NN	Hrp2 sequence 3 (E. Knüpfer)	Diagnosis	GATTCATTATTCTATATTTATAAGGAAGATTAC
OO	Hsp86 sequence 2 (E. Knüpfer)	Diagnosis	CTTAATTTCTCACGTTGTTAAAA

Legend: ____: Restriction enzyme site
Text: Mutagenesis site

Table 2.2: *Table of Transfection Constructs*

Construct name	Product	Purpose
pHH3MAL13P1.94ko	Truncated MAL13P1.94	Integration by single homologous recombination resulting in truncation of MAL13P1.94 and possible loss of function
pHH3MAL13P1.94ty	C terminally Ty1 tagged MAL13P1.94	Integration by single homologous recombination resulting in tagged MAL13P1.94 line
pHH3MAL13P1.94flag	C terminally 3x FLAG tagged MAL13P1.94	Integration by single homologous recombination resulting in tagged MAL13P1.94 line
pHH3MAL13P1.94_3′	WT 3D7 MAL13P1.94 gene	Integration by single homologous recombination demonstrating accessibility of genomic locus
pHTK_PF11_0443ko	Deleted PF11_0443	Integration of hDHFR cassette into the genomic locus by double homologous recombination resulting in the removal of PF11_0443
pHH3PF11_0443ty	C terminally Ty1 tagged PF11_0443	Integration by single homologous recombination resulting in PF11_0443-Ty1
pHH3PF11_0443flag	C terminally 3x FLAG tagged PF11_0443	Integration by single homologous recombination resulting in tagged PF11_0443 line
pHH3PF11_0443_3′	WT PF11_0443 gene	Integration by single homologous recombination demonstrating accessibility of genomic locus
pHH3PF02_0040ko	Truncated PF02_0040	Integration by single homologous recombination resulting in truncation of PF02_0040 protein and possible loss of function
pHH3PF02_0040ty	C terminally Ty1 tagged PF02_0040	Integration by single homologous recombination resulting in tagged PF02_0040
pHH3PF02_0040flag	C terminally 3x FLAG tagged PF02_0040	Integration by single homologous recombination resulting in tagged PF02_0040
pHH3PF02_0040_3′	WT PF02_0040 gene	Integration by single homologous recombination demonstrating accessibility of genomic locus
pHH3PF11_0443flagS316A	C terminally 3x FLAG tagged PF11_0443 S316A	Integration by single homologous recombination resulting in tagged PF11_0443 plus the exchange of serine 316 to alanine
pHH3PF11_0443gfp	PF11_0443-GFP	Integration by single homologous recombination resulting in tagged PF11_0443 line
pHH3MSP3prom_SP0040_gfp	GFP tagged PF02_0040 signal anchor	Episomal expression PF02_0040 signal anchor sequence fused to a GFP reporter

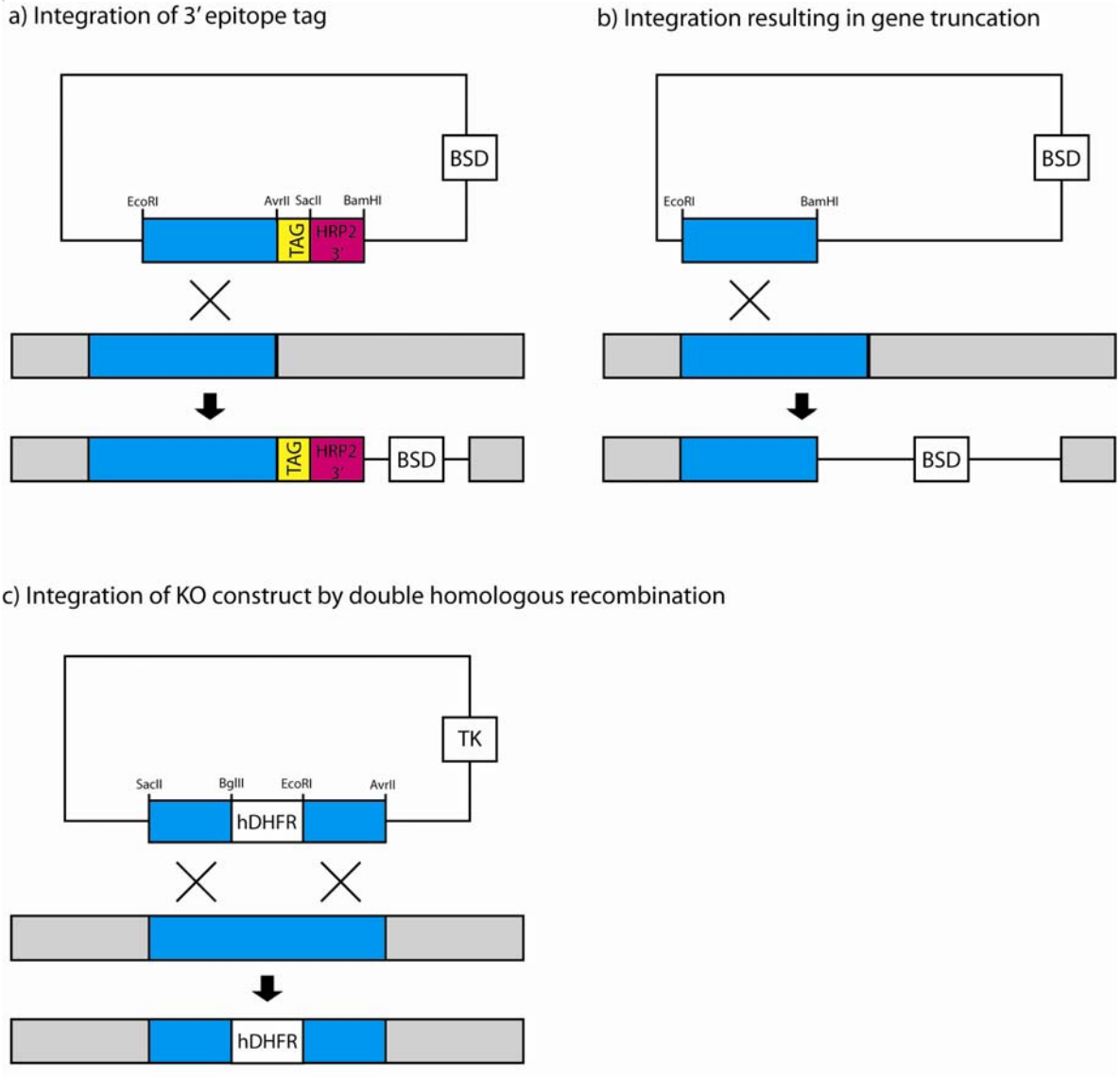


Figure 2.1: *Integration Strategy.* Parasite vectors for single homologous genomic integration (A+B) were based on the pHH3 transfection plasmid. A) Regions of homology to the gene of interest were cloned into pHH3 using *Eco*RI and *Sac*II. FLAG and TY tags were introduced to the vector backbone on the primer. Each tag was flanked with *Avr*II and *Sac*II restriction sites. B) Regions of homology were cloned using *Eco*RI and *Bam*HI. C) Double homologous recombination vectors were based on the pHTK transfection vector. Regions of homology from the N-terminus and 5' UTR region was inserted into multiple cloning site (MCS) 1 using *Sac*I and *Bgl*II and from the C-terminus and 3' UTR into MCS 2 using *Eco*RI and *Avr*II.

Table 2.3 Table of Protein Expression Constructs

Protein	Region (amino acids)	Vector	Purpose
MAL13P1.94	154 – 191	pET30-Xa/LIC	Raising antibodies
PF11_0443	237 – 321	pET30-Xa/LIC	Raising antibodies
PF14_0045	391 – 600	pET30-Xa/LIC	Raising antibodies
PF14_0045	391 – 600	pET32-Xa/LIC	Raising antibodies
PF02_0040	41 – 121	pET30-Xa/LIC	Raising antibodies
PF02_0040	28 – 276	pET30-Xa/LIC	Raising antibodies
PF02_0040	28 – 276	pET32-Xa/LIC	Raising antibodies
PF11_0443	34 – 98 (J domain)	pET30-Xa/LIC	Raising antibodies
PF11_0443	135 – 232	pET30-Xa/LIC	Raising antibodies

2.4 Table of Antibodies

Primary antibodies	Type	Animal	Target	Conc. IFA	Conc. Western	Supplier
αMAL13P1.94	Polyclonal	Rabbit	MAL13P1.94	1:200	1:200	This study
αPF11_0443	Polyclonal	Rabbit	PF11_0443	1:500	1:2000	This study
αPF14_0045	Polyclonal	Rabbit	PF14_0045	1:100	1:200	This study
α PF02_0040	Polyclonal	Rabbit	PF02_0040	1:1000	1:5000	This study
1E1	Monoclonal	Mouse	MSP1.19	1:3000	-	(Blackman et al., 1994)
αMSP1	Polyclonal	Rabbit	MSP1.83	1: 20000	-	(Arnot et al., 2008)
αRhopH2	Polyclonal	Rabbit	RhopH2	1:4000	-	(Ling et al., 2003)
61.3	Monoclonal	Mouse	RhopH2	1:200	-	(Holder and Freeman, 1984a)
24C6	Polyclonal	Mouse	RON4	1:1000	-	(Alexander et al., 2006)
αAMA-1	Polyclonal	Mouse	AMA-1	1:400	-	(Collins et al., 2007)
αAMA-1	Polyclonal	Rabbit	AMA-1	1:2000	-	(Arnot et al., 2008)
EBA-175VI	Monoclonal	Mouse	EBA-175	1:100	-	MR4
EBA-175VI	Polyclonal	Rabbit	EBA-175	-	1:500	MR4
MSP3	Polyclonal	Rabbit	MSP3	-	1:3000	(Mills et al., 2002)
113.1	Monoclonal	Mouse	MSP2	1:500	1:1000	(Gerold et al., 1996)
28.2	Monoclonal	Mouse	RESA	1:500		(Anders et al., 1986)
αBiP	Polyclonal	Rat	BiP	1:500		E. Knüpfer, unpublished
M2	Monoclonal	Mouse	FLAG tag	1:7500	1:20000	Sigma Aldrich
αTy1	Polyclonal	Rabbit	Ty1 tag	1:500	1:1000	antibodies online.com
Secondary antibodies						
αrabbit/mouse /ratAlexa Fluor 488	Polyclonal	Goat	Rabbit/mouse /rat Fc	1:5000	-	Invitrogen
αmouse/rabbit /rat Alexa Fluor 594	Polyclonal	Goat	Rabbit/mouse /rat Fc	1:5000	-	Invitrogen
αrabbit/mouse /rat HRP	Polyclonal	Goat	Rabbit/mouse /rat Fc	-	1:2500	Biorad

Chapter 3: The selection process and bioinformatic analysis of eligible proteins

3.1 Introduction

With the rapid spread of drug and insecticide resistance and the failure to discover an effective vaccine from the limited number of candidates, the urgent need for new antigens and drug targets is increasingly apparent. The completion of the *Plasmodium falciparum* 3D7 genome sequence gave the field of molecular parasitology the boost it needed to address these problems. Functional genomics has had a dramatic impact on the way in which stage-specifically expressed genes are identified and selected for further study. Genome comparisons revealed 60% of the *P. falciparum* genome was unique with little or no homology between *Plasmodium* genes and those from a wide range of previously sequenced organisms (Gardner et al., 2002). This great number of genes with unknown function could harbour the new *Plasmodium falciparum* drug or vaccine candidate. Utilising the wealth of genomic, transcriptomic, proteomic and bioinformatic data available we seek to identify novel *Plasmodium* antigens involved in invasion of host rbcs, a process which is thought to be an ideal target for intervention. Novel proteins involved in this process can be uncovered using features such as timing of expression from microarray or RNAseq based transcriptome analysis or presence in stage-specific proteomes.

3.1.1 P. falciparum transcriptome

Since the publication of the *P. falciparum* genome in 2002, a number of global transcriptome studies have been published. The first of these as published in 2003 and revealed the important relationship between expression timing and gene function. A DNA microarray study of the *P. falciparum* HB3 clone intra-erythrocytic stages (Bozdech et al., 2003) revealed that a large number of gene expression profiles exhibit periodicity and clear peaks and troughs of maximum and minimum expression within the 48 hour

cycle which corresponds to the point in the cycle the gene product is required. For example, the gene *ama1*, known to be involved in the erythrocyte invasion process, shows peak transcript expression 40 hours post invasion (p.i.). which coincides with the formation of the micronemes of which it is a part.

An alternative transcriptome by Le Roch *et al*. (2004) used high density oligonucleotide array analysis of RNA from 6 phases from the intra-erythrocytic cycle corresponding to the changes in distinct morphological stages of the intracellular parasite - early ring, late ring, early trophozoite, late trophozoite, early schizont and late schizont - as well as merozoites, sporozoites and gametocytes. Parasites used in the study were obtained using two alternative synchronization methods, sorbitol or temperature. Invasion related genes such as MSP-1, AMA-1, SUB2 and EBAs were highly expressed in the same cluster (cluster 15) suggesting genes expressed at a similar time may also function in invasion (Le Roch et al., 2004).

3.1.2 Proteome data

A number of high-throughput mass spectrometry based proteomic studies have been undertaken in the past decade. Florens *et al*. (2002) identified approximately 2400 proteins from sporozoites, trophozoites, merozoites and gametocytes using multi-dimensional protein identification technology (MudPIT)(Florens et al., 2002). The Leiden Malaria Group provided a mass spectrometry based merozoite proteome which is available on PlasmoDB (http://plasmodb.org). With advances in the field of mass spectrometry the data produced are more reliable and able to identify proteins of low abundance from every stage of the lifecycle including gametocytogenesis (Silvestrini et al., 2010); oocysts, oocyst-derived sporozoites, salivary gland sporozoites (Lasonder et al., 2008) and liver stages (Tarun et al., 2008), as well as late/rupturing schizonts (Bowyer et al., 2011) and proteomes from clinical samples (Acharya et al., 2009). Additional features such as the presence of phosphorylated serine and threonine residues can also be detected by mass spectrometry methods and the first phospho-proteome of blood stages has been described (Treeck et al., 2011; Lasonder et al., 2012) .

3.1.3 Comparative genomics

Although the sequencing of the *P. falciparum* 3D7 genome has provided valuable insights into the parasite's biology, comparison between the genomes of other *Plasmodium* spp. as well as other unicellular eukaryotic pathogens can supply important information regarding the evolution and physiology of the parasite. Additionally, it enhances our ability to assign function to a predicted protein through identification of orthologous proteins in other species. The genomes of several *Plasmodium* species have been published - *P. falciparum* IT (Parasite Genomics Group, 2011); *P. vivax* Sal-1 (Carlton et al., 2008); *P.knowlesi* strain H (Pain et al., 2008); *P. berghei* strain ANKA (Parasite Genomics Group, 2010b); *P.yoelii yoelii* str. 17XNL (Carlton et al., 2002); *P. chabaudi chabaudi* AS (Parasite Genomics Group, 2010a). Genomes of fellow apicomplexa - *T. gondii, Th. parva, Th. annulata, B. bovis, Cr. parvum, Cr. hominis, Neospora caninum* and *Eimeria tenella* are also available as well as a multitude of other organisms such as bacteria, yeast, fungi and higher eukaryotes such as plants and vertebrates. Comparison with *Toxoplasma* is particularly relevant as it is currently the best studied Apicomplexa due to its superior experimental capacity (Kim and Weiss, 2004).

3.1.4 Protein interaction studies

The yeast-2-hybrid system is a high-throughput method for identification of protein-protein interactions and was applied to *Plasmodium* by LaCount *et al.* (2005). The system applied to other organisms generated a large number of false positives, as was the case for *Drosophila* (Giot et al., 2003; Formstecher et al., 2005). Also, the use of this system in *Plasmodium* failed to identify a large number of well established interactions such as those between MSP-1, -6 and -7, thus the results from this work shall not be included in the bioinformatic analysis presented here (LaCount et al., 2005).

3.2 Results: the selection process and resulting proteins

The aim of this project was to characterise proteins of the malaria parasite *Plasmodium falciparum* involved in erythrocyte invasion. The published sequence, transcript and proteomic data for each gene in any of the sequenced species of the *Plasmodium* genome are housed in the *Plasmodium* database PlasmoDB (plasmodb.org). To identify proteins that fit the criteria (explained in section 1.7) I utilised the database's refined 'search' tool, taking advantage of published sequencing, transcript, proteomic and domain predictions.

Criterion a) Expressed during schizogony and in the merozoite

Transcription data from Bozdech *et al* (2003) estimated the peak transcript levels of AMA-1 at around 40-42 hours p.i. (data available on PlasmoDB). Therefore the first part of my search involved identifying proteins with a maximal expression of 42 ± 4 hours p.i. The second, following the same transcription profile, added the additional step of including minimum expression of 24 ± 8 hours. This resulted in a list of 566 genes to be processed with further stringency.

Criterion b) Presence of a signal peptide or anchor sequence

Signal sequences are required for entry into the secretory system, the first step for transport to apical organelles or the merozoite surface. Signal peptides are characterised by a stretch of hydrophobic residues followed by a 'stop transfer' motif and the presence of a signal peptidase cleavage site, whereas signal anchor sequences lack the cleavage site and remain in the ER membrane. The presence of a signal sequence can be predicted from the gene product's translated amino acid sequence using web-based algorithms such as SignalP (Nielsen et al., 1997), which was used in this case. After searching the 566 genes for those which translated into proteins containing signal sequences, 116 qualified for further scrutiny.

Criterion c) The proteins should possess one or more transmembrane domain(s)

From the predicted protein sequences of each gene, regions likely to span the cell membrane can be identified. This information is provided on the PlasmoDB database and using the TMHMM web-based algorithm (Sonnhammer et al., 1998; Krogh et al., 2001). Of the 166 genes with late stage transcription and a gene product with predicted secretory signal 64 also had one or more predicted transmembrane spanning regions.

Criterion d) Proteins should be conserved across *Plasmodium* spp.

Although the *P. falciparum* 3D7 was the first genome to be sequenced in 2002, several other species of *Plasmodium* have been sequenced since (*P. falciparum* IT; *P. vivax* Sal-1; *P.knowlesi* strain H; *P. berghei* strain ANKA; *P.yoelii yoelii* str. 17XNL; *P. chabaudi chabaudi*, referenced in 3.1.3 above) and these data were utilised to search for orthologous genes. Every one of the 64 predicted proteins possessed homologues in 2 or more of sequenced species and therefore failed to narrow the list any further.

This list of 64 genes and their putative proteins was filtered further based on available data such as transcript expression levels, expression in other life cycle stages, presence in schizont/merozoite proteomes, difficulty of study (low expression levels or a large number of transmembrane regions) and novelty. The full list of 64 proteins is detailed in Appendix A and includes reasons for consideration/dismissal. I selected 4 proteins initially for further study: MAL13P1.94, PF11_0443, PF14_0325, and PF02_0040, as well as an additional fifth protein of special interest to the lab, PF14_0045. This protein was not selected by the initial criteria as the peak transcript expression occurs later than 44 hours post invasion however from transcript and proteomic data it appears to be expressed in the merozoite.

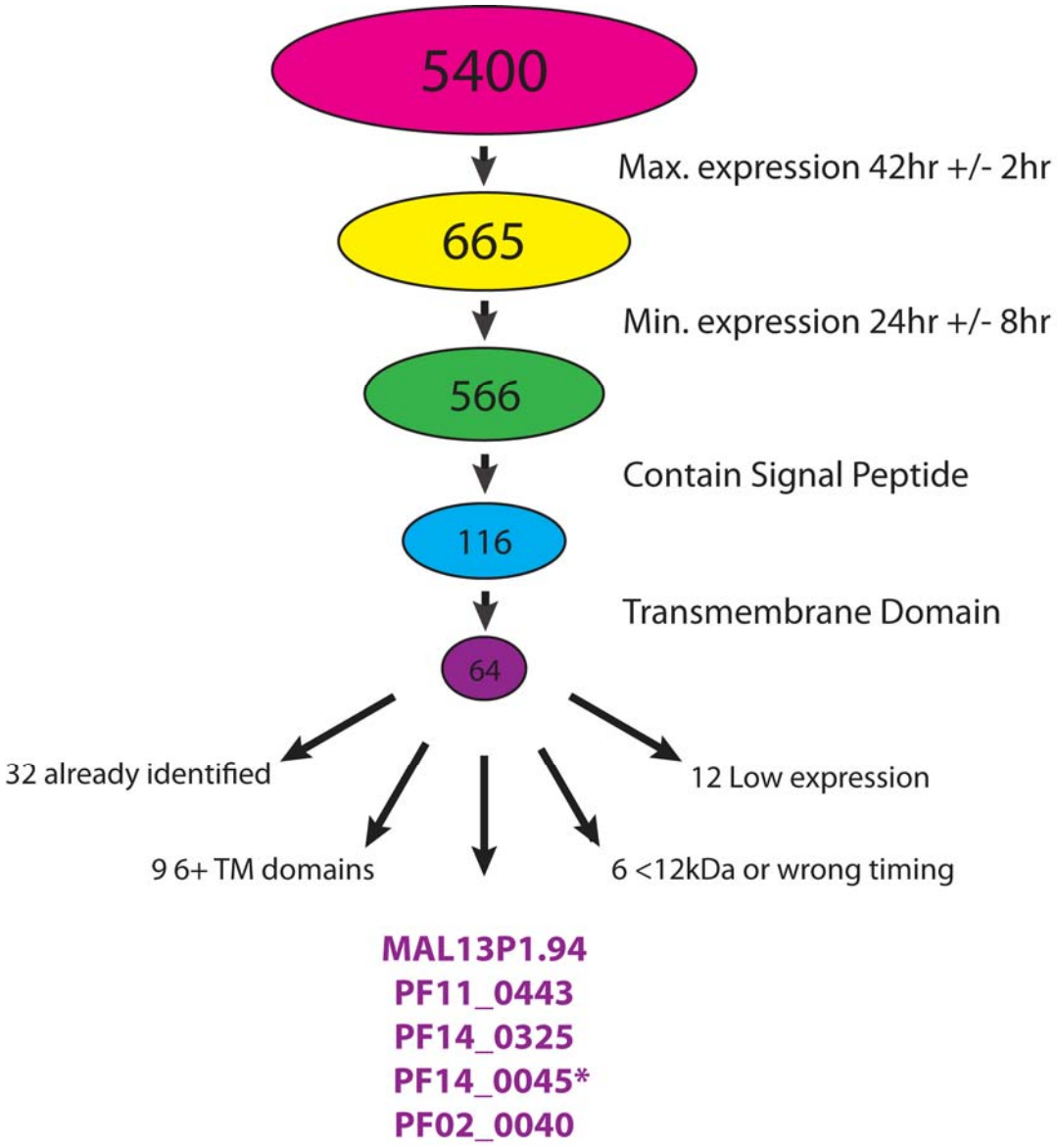


Figure 3.1. *The selection process.* Five proteins were selected using the following criteria to search the database PlasmoDB: a) the protein should be expressed during the blood stage – peak expression 42 hours post invasion (p.i.) ± 2 hours and then minimum expression 24 hours p.i. ± 8 hours; b) the translated protein should contain a signal peptide/anchor; c) one or more transmembrane domains and d) the protein should be conserved between *Plasmodium* species. The last step was unnecessary as all of the 64 proteins selected were conserved within the sequenced genomes of *Plasmodium* species. There were 32 selected proteins which were either already defined or under investigation. The list was then filtered further based on available data such as transcript expression levels, expression in other life cycle stages, presence in merozoite proteome and difficulty of study. * highlight gene chosen due to interest from the lab.

3.3 *In silico* analysis of individual proteins

After selection, the protein sequences were examined using numerous analysis programs and web-based algorithms, predicting the presence of some recognisable features within the selected proteins. Transmembrane region predictions were confirmed using TMHMM (Sonnhammer et al., 1998; Krogh et al., 2001), TMpred (K. Hofmann, 1993) and TopPred (von Heijne, 1992). Signal peptides were detected using the latest version of SignalP (SignalP4.0) (Petersen et al., 2011) which predicts signal peptidase cleavage sites only, whereas signal anchor regions were suggested using SignalP 2.0 which also predicts the presence of transmembrane regions within the signal sequence using the Hidden Markov Model (HMM) algorithm (Nielsen and Krogh, 1998). Any conserved domains were predicted from the protein sequence by Pfam (Finn et al., 2008), the conserved domain database (CDD) (Marchler-Bauer et al., 2011). SMART (simple modular architecture research tool) was used to confirm the presence of these domains as well as transmembrane regions and signal peptides (Schultz et al., 1998). Alignments were produced for each protein using the ClustalW algorithm, showing level of conservation between their other *Plasmodium* counterparts and orthologues in other species (Larkin et al., 2007). Orthologues were predicted and sequences were obtained from the OrthoMCL database (Li et al., 2003; Chen et al., 2006) and confirmed using the NCBI BLAST tool. For some proteins, this data mining produced interesting leads into the possible function of the protein. The selected proteins are introduced below and data recovered from *in silico* analysis discussed.

3.3.1 MAL13P1.94

Transcript data

The transcriptome data produced by Bozdech *et al*. (2003) shows peak expression of MAL13P1.94 mRNA in the HB3 clone at around 38 hours p.i. In accordance with this, the Le Roch study showed maximum expression of this protein in late schizonts in both sorbitol and temperature synchronised methods (Le Roch et al., 2004). Transcript

expression levels are comparable to those of AMA1 and therefore should be in reasonable abundance to visualise within the cell (illustrated in figure 3.1 A.).

Proteomic data

This protein was not present in any proteomic studies available at the time of selection of proteins for further study. However this was only based on two studies; Florens *et al.* (2002) and the Leiden Malaria Group merozoite proteome [(Khan, 2008), available on PlasmoDB]. Also the protein is relatively small in size and its firm attachment to the membrane via four transmembrane passes may have made it difficult to extract and solubilise. Since beginning this project 2 peptides from the extreme C-terminus were detected in a proteomic study by Treeck *et al* (2011) (figure 3.1 B.), confirming the protein presence in blood stages.

Bioinformatic predictions

At a predicted molecular weight of 22 kDa, MAL13P1.94 is a relatively small *Plasmodium* protein and yet is predicted to contain 4 transmembrane domains using TMHMM, TMpred and TopPred algorithms, their location is detailed in figure 3.1 C. The N-terminal transmembrane region overlaps with a signal sequence and is predicted to be a signal anchor region using SignalP 2.0. No conserved domains were predicted by either Pfam or CDD.

Orthologous proteins in other species

Orthologues of MAL13P1.94 in other species were identified from PlasmoDB and OrthoMCL-DB (Li et al., 2003), from the orthologue group OG5_145102 and confirmed using NCBI BLAST tool. The protein is highly conserved in *Plasmodium* spp. (*P. bergei*: PBANKA_141650; *P.*chabaudi: PCHAS_141830; *P.* knowlesi: PKH_141720; *P.* vivax: PVX_122720) with the notable absence of *P. yoelii,* although this is due to lack of proper sequence assembly rather than a notable omission from the genome. Using the tblastn algorithm on NCBI BLAST it appears the gene can be found split between contigs AABL01001120 and AABL01003528 (amino acids 1-112 and 141-192 respectively). MAL13P1.94 has orthologues in other genera, namely *Toxoplasma gondii*, *Neospora*

canium, Theileria annulata, Babesia bovis and several *Cryptosporidium* species but not outside the phylum of Apicomplexa. All of these proteins share similar domain architecture, with 4 transmembrane domains, the first of which is a signal anchor. The protein sequences were aligned using the ClustalW multiple alignment algorithm (Thompson et al., 1994) in order to visualise and calculate the sequence similarity between species (figure 3.2). *P. knowlesi* and *P. vivax* orthologues of MAL13P1.94 are highly conserved with identities of 73 and 72% respectively however outside the genus this is not the case, with sequence identities of 25% with *T. gondii* and only 15% for *B. bovis*. Despite this low identity these proteins share a similar predicted size and domain architecture, each containing 4 transmembrane regions.

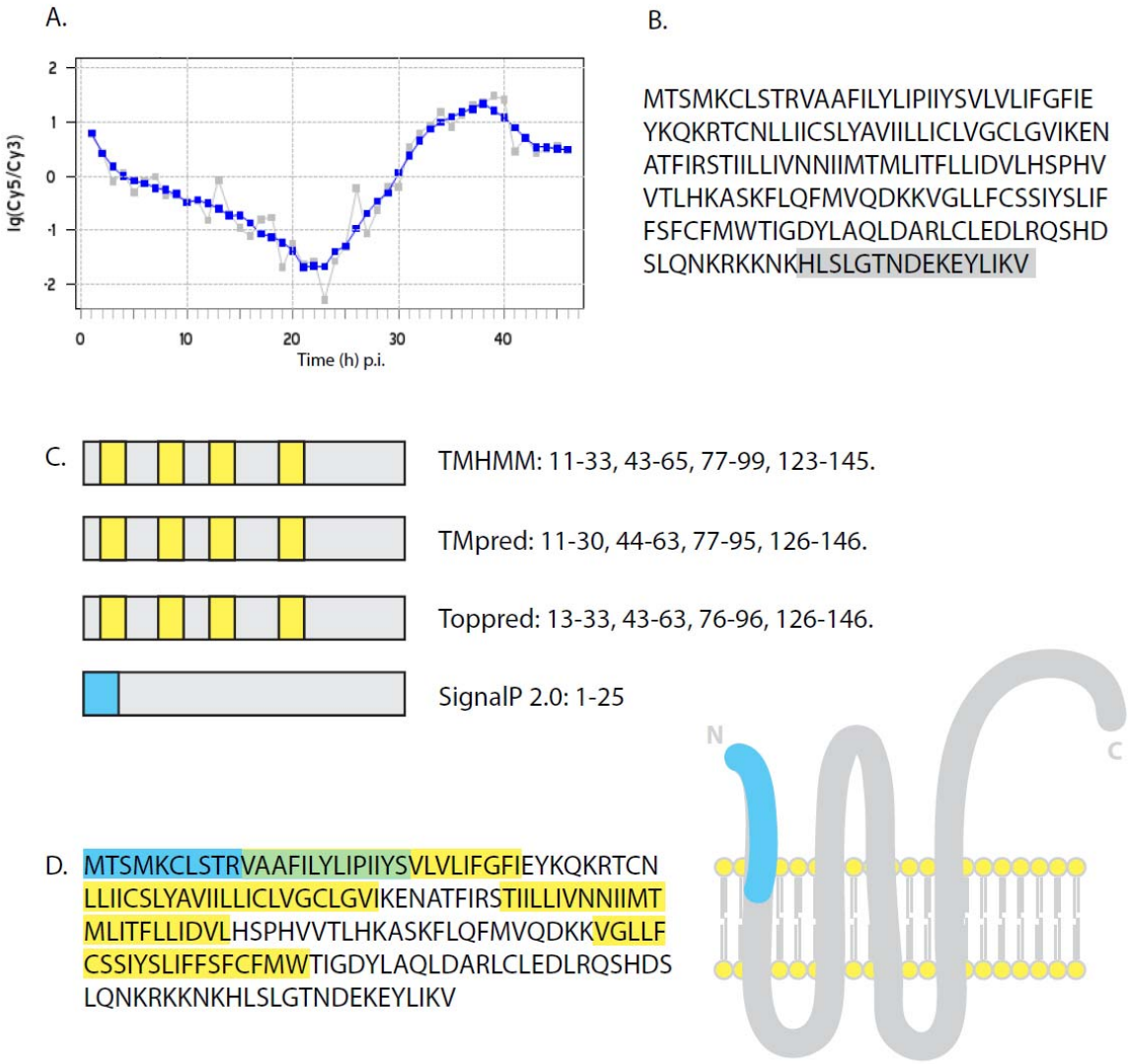


Figure 3.2: *MAL13P1.94 transcription profiles, proteomic data and predicted domain architecture.* (A) The transcriptome from Bozdech *et al* (2003) shows a maximum transcript level of *mal13p1.94* at 38 h p.i. Two peptides from the C-terminus of the protein were recovered from schizont stage parasites in the proteomic study by Treeck *et al* (2011). These peptides have been mapped to the predicted protein sequence and highlighted in grey (B). This 22kDA protein is predicted to contain 4 transmembrane domains, the first of which a signal anchor region (C). The domains were mapped onto the predicted protein sequence and illustrated in (D).

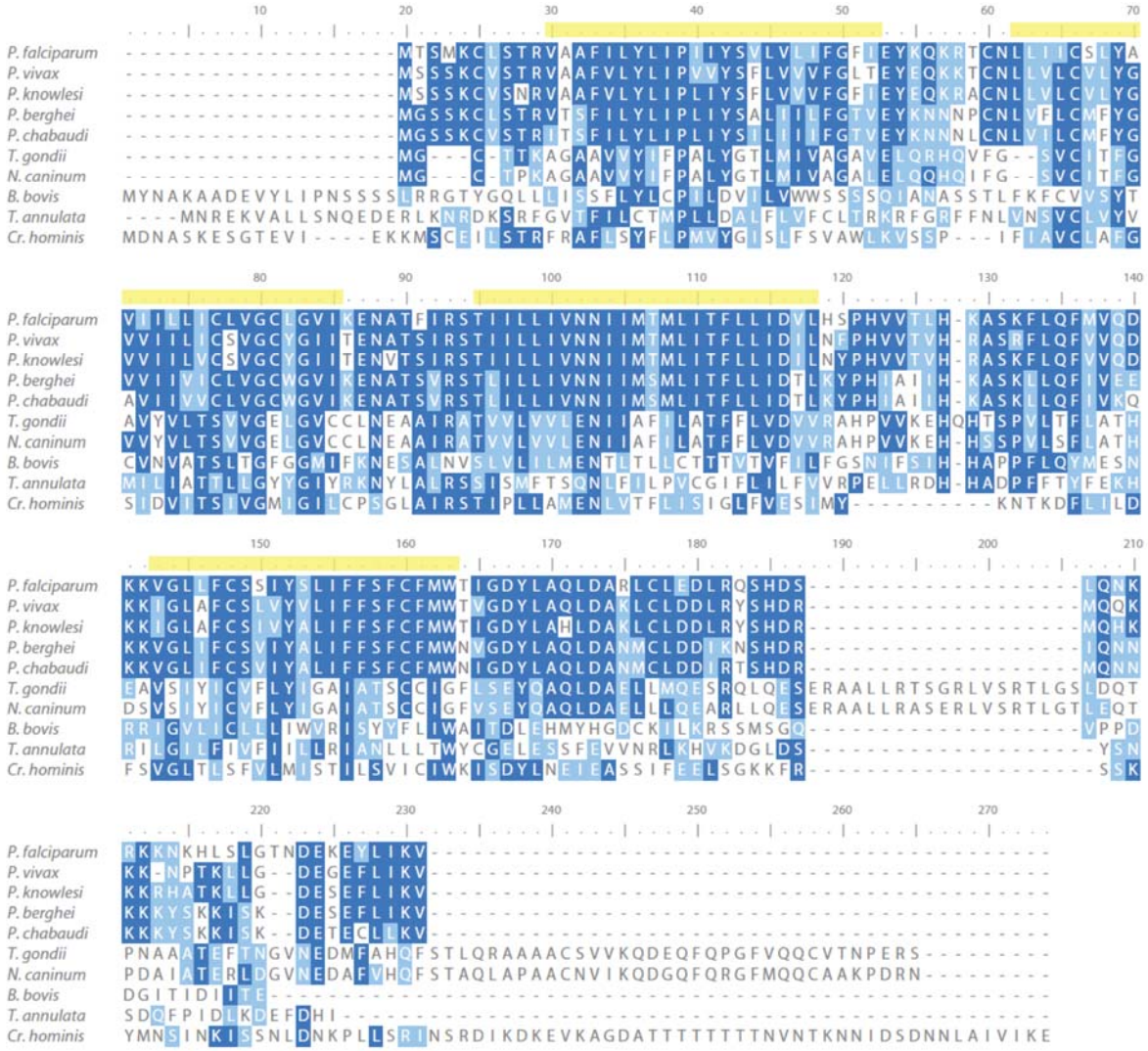


Figure 3.3: *Multiple sequence alignment of MAL13P1.94 orthologues.* Sequences from *Plasmodium* spp. and orthologues from Apicomplexa of MAL13P1.94 were aligned using the ClustalW multiple alignment algorithm. Residues identical in over 50% of the chosen sequences are shaded in dark blue and those which are functionally similar are highlighted in light blue. Transmembrane domains are shown in yellow based on the *Plasmodium* sequences, regions may be different in other Apicomplexa.

3.3.2 PF11_0443

3.3.2.1 Transcript data

The transcriptome data produced by Bozdech *et al*. (2003) shows peak expression of PF11_0443 mRNA in the HB3 clone at around 33 hours p.i. In accordance with this, the Le Roch study showed maximum expression of this protein in early schizonts in temperature synchronised cultures however expression is maintained at a high level in late schizonts. Using the sorbitol method the transcript levels peak in late schizonts; this highlights a discrepancy between the two synchronisation methods and between transcriptomes. Expression levels are comparable that of to AMA-1and therefore should be in reasonable abundance to ease further experimental analysis (Le Roch et al., 2003). These data are depicted in figure 3.3 A.

3.3.2.2 Proteomic data

This protein was not detected in the proteomic study by Florens *et al.* (2003) however the Leiden Malaria Group merozoite proteome (available on PlasmoDB) identified 3 peptide matches, confirming the gene *PF11_0443* encodes a protein which is expressed in the erythrocyte invasive form of the parasite. Since then peptides have been detected in several proteomic studies from schizont and merozoite stages, covering the full length of the protein with the exception on the extreme N terminus, illustrated in figure 3.3 B (Bowyer et al., 2011; Siverstrini et al., 2010; Treeck and Saunders et al., 2011).

3.3.2.3 Bioinformatic predictions

PF11_0443 has a predicted molecular weight of 39kDa with 2 or 3 predicted membrane spanning regions depending on the algorithm (illustrated in figure 3.3 C). Although 2 of the 3 algorithms used assigned a transmembrane domain to the N-terminus of the protein, SignalP 4.0 identified a signal peptide cleavage site after the cysteine at position 22 with a probability of 0.967 and therefore it is likely this region is not a signal anchor region but a classical hydrophobic signal peptide. The fact the N-terminus is absent from the previously described proteomes gives weight to this likelihood. The second predicted transmembrane was predicted by all 3 algorithms and the third by 2 out of 3,

with the third just below the threshold score necessary for transmembrane prediction. The reason for this discrepancy may be a proline at position 217 which is traditionally a helix-breaker and may disrupt the helical transmembrane domain. Although it is likely this is a transmembrane domain, experimental confirmation is required. The most distinctive feature of this protein is the presence of a DnaJ domain at the N terminus of the protein between residues 36 and 92, making this protein a member of the Hsp40 or J protein family. A study by Botha *et al* (2007) identified 43 Hsp40s present in *P. falciparum* 3D7 and assigned this protein to a group of type IV J proteins. This means it lacks any of the previously recognised features which would allow it to achieve the primary function of an Hsp40 which is to bind to Hsp70. The role of Hsp40s in *Plasmodium* is discussed in chapter 5.1. Botha *et al* (2007) suggested that PF11_0443 is a resident of the apicoplast based on the protein's charged N-terminal sequence which is similar to an apicoplast transit peptide however analysis using the PlasmoAP algorithm (Foth et al., 2003) could not confirm this.

3.3.2.4 Orthologous proteins in other species

Orthologues of PF11_0443 in other species were identified from PlasmoDB and OrthoMCL-DB, from the orthologue group OG5_131379 (Li et al., 2003) and confirmed using NCBI BLAST tool. The protein is highly conserved in *Plasmodium* spp. (*P. berghei*: PBANKA_090580; *P.*chabaudi: PCHAS_070700; *P.* knowlesi: PKH_094070; *P. vivax*: PVX_092765; *P. yoelii*: PY01970). Significantly, PF11_0443 has no obvious orthologues in other apicomlexa and in fact almost all of the non-*Plasmodium* orthologous proteins appear to be in multi-cellular eukaryotes (listed in table 3.1). The existence of orthologous proteins in non-Apicomplexa and non-plant species provides further evidence against a role for this protein in the apicoplast. The protein sequences were aligned using the ClustalW multiple alignment algorithm (Thompson et al., 1994) in order to visualise and calculate the sequence similarity between species (figure 3.4). *P. falciparum* is 59% identical to *P. vivax* or *P. knowlesi* and slightly less at 50% and 52% for the rodent malarias, *P. berghei* and *P. chabaudi* respectively. This identity drops by up to 30% when comparing *P. falciparum* with metazoa for example PF11_0443 shares 20% identity with *Homo sapiens*. The orthologues all share similar a transmembrane domain

architecture as well as an N-terminal DnaJ domain. Importantly the non-*Plasmodium* orthologues all share an HPD motif within their J domain, the significance of this will be discussed in chapter 5. They also share a higher degree of similarity at the C-terminus with 4 conserved tryptophan residues. Tryptophan is a relatively rare amino acid to *P. falciparum* due to its single guanine rich codon suggesting these residues are necessary for the protein's function. These features are illustrated on the PF11_0443 orthologue alignment in figure 3.5.

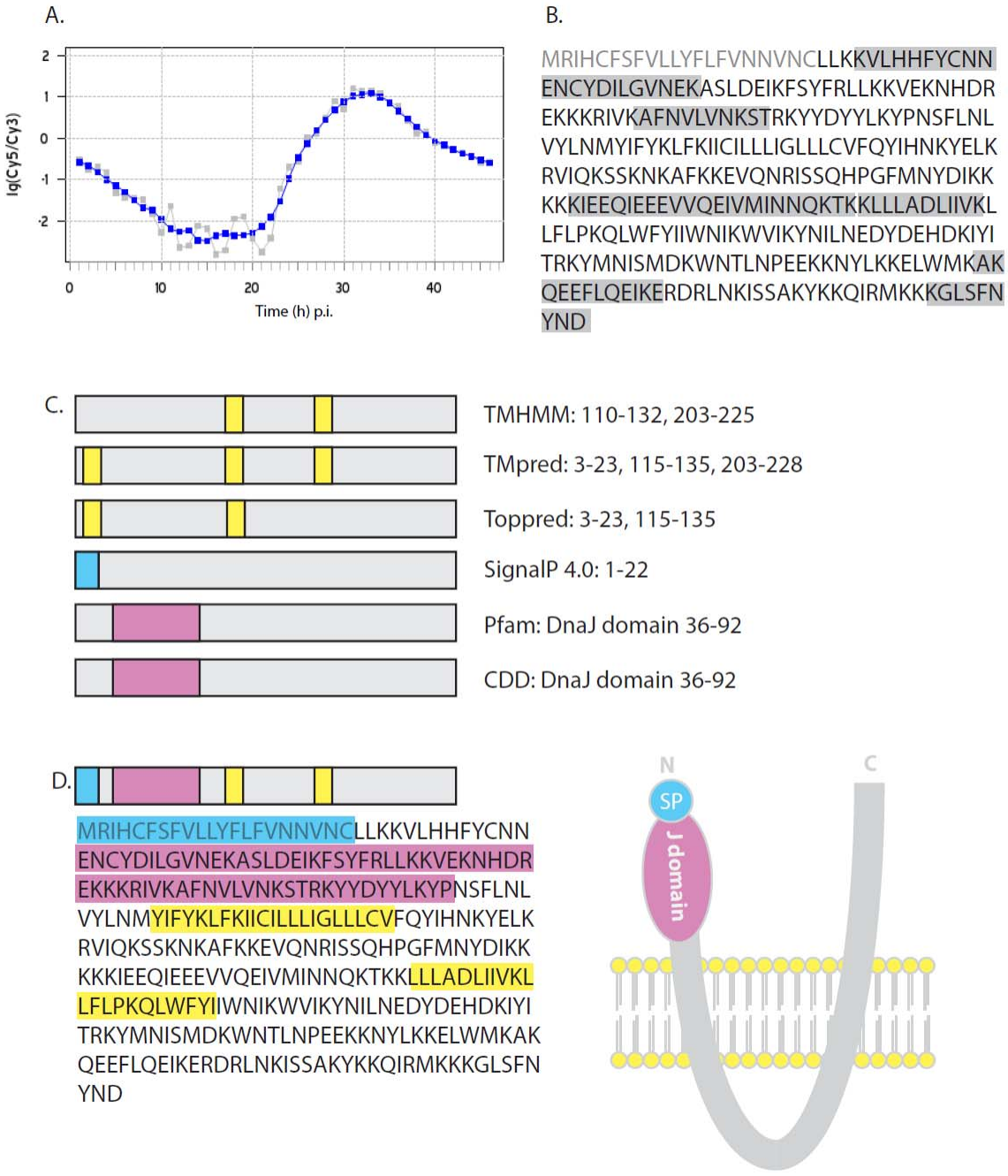
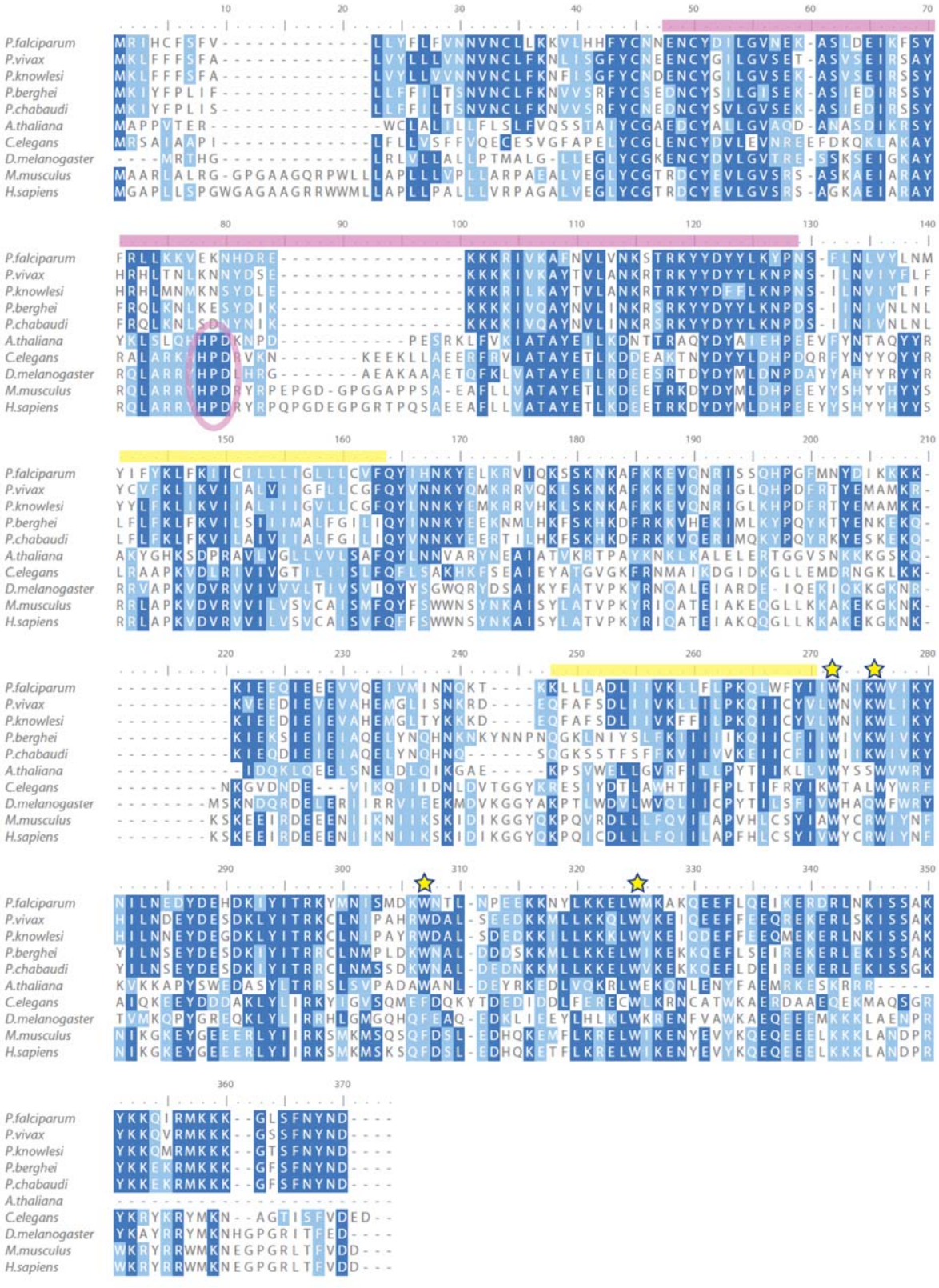


Figure 3.4: *PF11_0443 transcription profile, proteomic and bioinformatic data summary.* (A) The transcriptome from Bozdech *et al.* (2003) shows a maximum transcript level of *pf11_0443* at 33 h p.i. Several peptides have been recovered from schizont and merozoite stage parasites in a number of proteomic studies. These peptides have been mapped to the predicted protein sequence highlighted above in (B). This 39kDa protein is predicted to contain either 2 or 3 transmembrane domains and a signal peptide (C). Domains were mapped onto the predicted protein sequence and illustrated in (D).

Table 3.1: *Species which possess orthologues of PF11_0443.*

Aedes aegypti	*Dictyostelium discoideum AX4*	*Ricinus communis*
Anopheles gambiae str. PEST	*Drosophila melanogaster*	*Rattus norvegicus*
Arabidopsis thaliana	*Danio rerio*	*Schistosoma mansoni*
Brugia malayi	*Gallus gallus*	*Schistosoma mansoni*
Bombyx mori	*Homo sapiens*	*Trichoplax adhaerens*
Caenorhabditis briggsae AF16	*Monosiga brevicollis MX1*	*Tetraodon nigroviridis*
Caenorhabditis elegans	*Monodelphis domestica*	*Tetraodon nigroviridis*
Caenorhabditis elegans	*Mus musculus*	*Thalassiosira pseudonana CCMP1335*
Ciona intestinalis	*Nematostella vectensis*	*Takifugu rubripes*
Canis lupus familiaris	*Nematostella vectensis*	*Tetrahymena thermophila SB210*
Culex pipiens	*Ornithorhynchus anatinus*	*Encephalitozoon cuniculi GB-M1*
Culex pipiens	*Oryza sativa Japonica Group*	*Ricinus communis*

Figure 3.5 *Multiple sequence alignment of PF11_0443 orthologues*. Sequences from *Plasmodium* spp. as well as a varied selection of higher eukaryotic sequences of PF11_0443 orthologues were aligned using the ClustalW algorithm. Residues identical in over 50% of the chosen sequences are shaded in dark blue and those which are functionally similar are highlighted in light blue. The region predicted to contain the DnaJ domain is indicated by a purple line, with the conserved HPD motif circled in the same colour. Conserved tryptophans are indicated with a yellow star and transmembrane domains are shown in yellow.


P.falciparum
P.vivax
P.knowlesi
P.berghei
P.chabaudi
A.thaliana
C.elegans
D.melanogaster
M.musculus
H.sapiens

3.3.3 PF14_0325

Transcript data

The transcriptome data produced by Bozdech *et al*. (2003) shows peak expression of PF14_0325 mRNA in the HB3 clone at around 38 hours p.i. (illustrated in figure 3.6 A). In accordance with this, the Le Roch study showed maximum expression of this gene in late schizonts in both sorbitol and temperature synchronised methods (Le Roch et al., 2003). Absolute expression levels are comparable to that of AMA-1. Interestingly PF14_0325 expression was detected (at equivalent levels to the peak expression in late schizonts) in sporozoites, the form of the parasite responsible for invading hepatocyes.

Proteomic data

This protein has not been detected in any proteomic studies available either at the time of selection or present day. The cause for this could be the protein's relatively small size or the protein's abundance or even that it is not translated at all.

Bioinformatic predictions

At the time of selection, PF14_0325 possessed 2 different annotations, each of which contains a different 5' end exon assembly. I was able to identify by PCR from cDNA which of the two models was correct (Appendix B) and subsequently the annotation on PlasmoDB was also corrected. The gene encodes a protein of approximately 33kDa with a signal peptide at the N-terminus, cleaved after residue 22. This signal peptide was not predicted using either SignalP2.0 or 4.0 but via SMART, which utilised an undisclosed version of SignalP. The *P. vivax* homologue of this protein possesses a highly predicted signal peptide and after sequence comparison it looks as though the negative charge of the aspartic acid at position 15 is the reason for the discrepancy. Position 15 is a glycine residue in *P. vivax* and *P. knowlesi*. It is unlikely this is due to a sequencing error as the other sequenced *P. falciparum* genomes also contain an aspartic acid instead of a glycine. This particular residue's codon is split by an intron and it is possible that the location of the intron is incorrect. There is a predicted transmembrane region at the extreme C terminus of the protein however there are no positively charged residues

after the membrane region (the last residue is an isoleucine) therefore the protein could not be anchored in the membrane without the addition of glycosylphosphatidylinositol (GPI) membrane-anchor. GPI-anchored proteins are common in all eukaryota including *Plasmodium,* with proteins such as MSP-1 and -2 (Gerold et al., 1996) and CSP (Dame et al., 1984) among their number. These proteins require an N-terminal signal peptide for entry into the ER lumen and a GPI-anchor signal sequence at the C-terminus which is recognised by the transamidase complex (Dailey & Bulleid, 2003). This complex catalyses the proteolytic removal of the hydrophobic C-terminal region (often predicted as a transmembrane region) between residue ω and ω+1, before the addition of a membrane-bound GPI (Eisenhaber et al., 2001). GPI prediction algorithms GPIpred, GPI-SOM and Big-π (based on protozoa) all failed to predict PF14_0325 to contain the suitable signal required for the addition of a GPI-anchor. Fortunately and somewhat bizarrely, the Big-π algorithm based on metazoa did predict GPI-anchorage, with cleavage predicted to take place after the asparagine residue 265 (ω) with a p-value of 0.00081. Generally prediction programmes have been designed using metazoa sequences and have not been tailored towards protozoa such as *P. falciparum*. Gilson *et al.* (2006) developed an alternative algorithm, GPI-HMM, based on the hidden Markov model and using known GPI-anchored *P. falciparum* proteins to identify as yet unidentified *P. falciparum* GPI-anchored proteins. Although their algorithm is not released for general use, their published list of 30 proteins most likely to be GPI-anchored featured PF14_0325 at position 29 with a p-value of 0.015. All 16 of the characterised proteins on this list contain GPI-anchored and none of the uncharacterised proteins have been shown to be without one, validating the model, however biochemical classification of the protein is necessary to provide a definitive answer.
The coiled-coil detection algorithm 'Coils' predicts a coiled-coil domain in the centre region of the protein (illustrated in figure 3.6). These domains are often involved in protein-protein interaction.

Orthologous proteins in other species

Orthologues of PF14_0325 in other species were identified from PlasmoDB and OrthoMCL-DB, from the orthologue group OG5_150579 (Li et al., 2003) and confirmed

using NCBI BLAST tool. The protein is highly conserved in *Plasmodium* spp. (*P. bergei*: PBANKA_101040; *P.chabaudi*: PCHAS_101130; *P. knowlesi*: PKH_131490; *P.* vivax: PVX_084815; *P. yoelii*: PY03139). PF14_0325 has orthologues in other genera, namely *Neospora canium, Theileria annulata* and *Babesia bovis* but not outside the phylum of Apicomplexa. The so called orthologue of *N. caninum* appears to be a dramatically different protein both in size and in amino acid content however there is one small region near the C-terminus which seems to be conserved. Other than this divergent orthologous protein from *N. caninum,* the protein appears only in Apicomplexa of the blood, which may give a clue to its function. Protein sequences were aligned using the ClustalW multiple alignment algorithm (Thompson et al., 1994) in order to visualise and calculate the sequence similarity between species (figure 3.6). The first 119 amino acids were removed from the *N. caninum* sequence as it was not aligned with any other protein sequences and this protein is over one hundred amino acids larger than the other proteins. All of the sequences contain signal peptides but none are predicted to have GPI anchors. Due to the unreliable nature of GPI anchor predictions this does not necessarily prove their absence. Although the C-termini of the non-*Plasmodium* spp. do not appear sufficiently hydrophobic, asparagine 265 which is the ω site on *P. falciparum* is completely conserved throughout the orthologous species thus the *in silico* evidence is inconclusive. Sequence identity between the group members is not all that high even between Plasmodia since PF14_0325 shares 50% identity with *P. vivax* and *P. knowlesi*, 28% with *Th. annulata* and 17% with *B. bovis*. A region of 50-60 residues nearing the C-terminus is much more conserved with the presence of a number of conserved motifs which are perhaps necessary for a function. The asparagine-rich coiled coil region in the centre of PF14_0325 is not conserved in the other proteins.

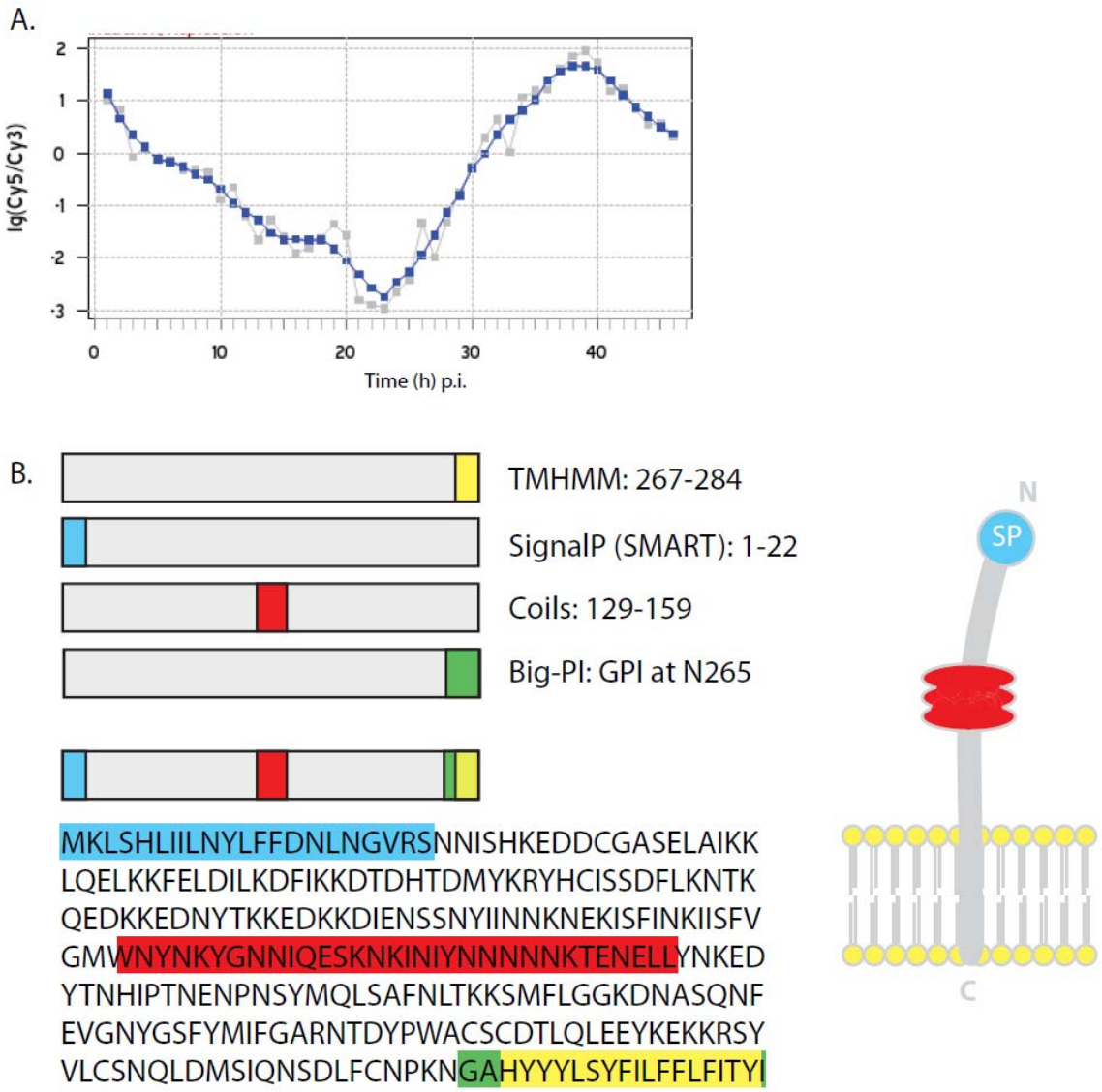


Figure 3.6: *PF14_0325 transcription profile and domain architecture summary*. (A) The transcriptome from Bozdech *et al* (2003) shows a maximum transcript level of *pf14_0325* at 38 h p.i. This 33kDa protein is predicted to contain a signal peptide and a coiled coil domain between residues 129 and 159. A transmembrane domain is predicted between residues 267 and 284 however this hydrophobic stretch is actually part of the GPI-achor signal, as predicted by Big- π. Domains were mapped onto the predicted protein sequence and illustrated in (B).

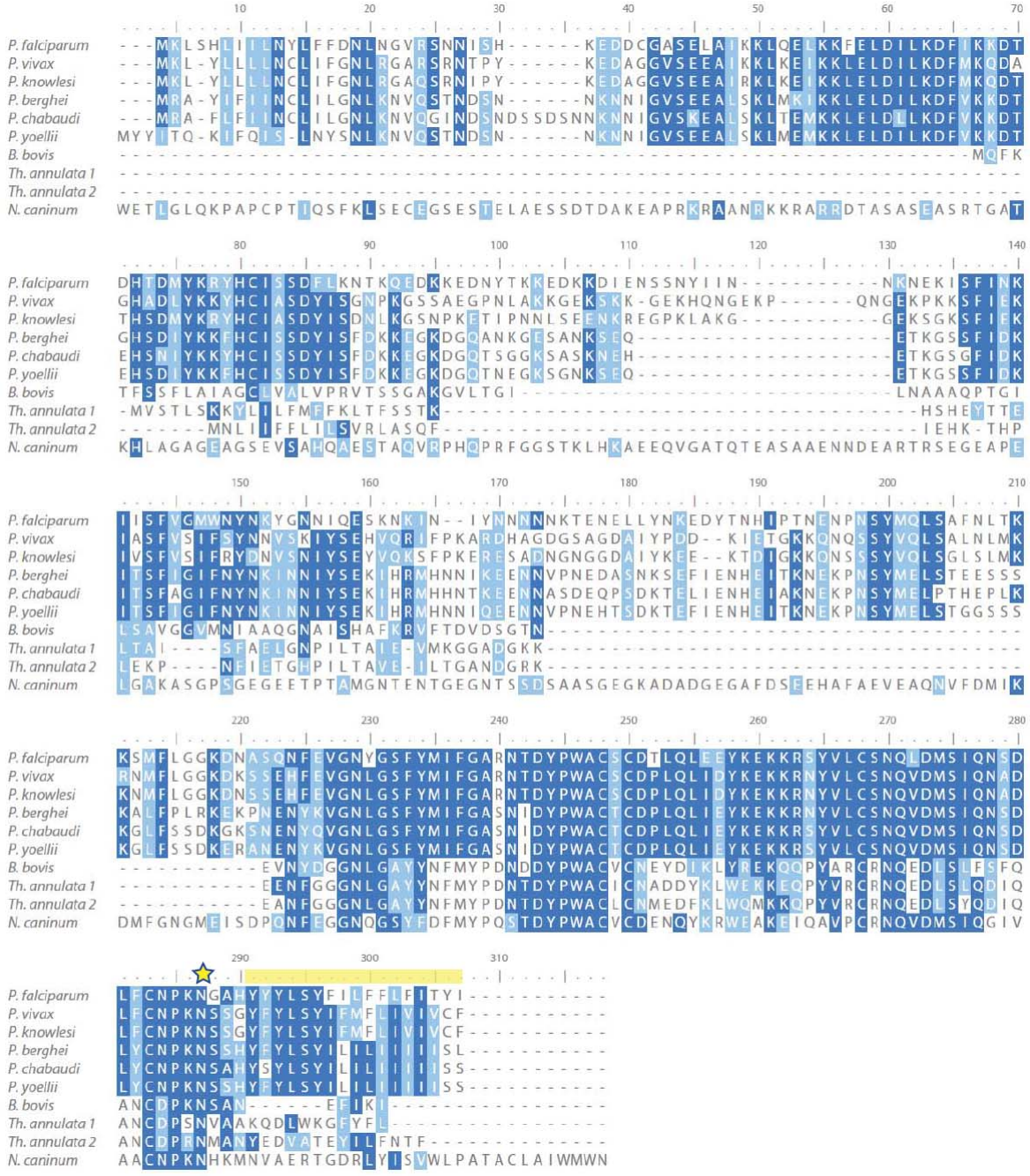


Figure 3.7 *Multiple sequence alignment of PF14_0325 orthologues.* Sequences from *Plasmodium* spp. and orthologues from Apicomplexa of PF14_0325 were aligned using the ClustalW multiple alignment algorithm. Residues identical in over 50% of the chosen sequences are shaded in dark blue and those which are functionally similar are highlighted in light blue. Transmembrane domains are shown in yellow based on the *Plasmodium* sequences, regions may be different in other Apicomplexa. The ω is highlighted with a yellow star. The first 119 amino acids were removed from the *N. caninum* sequence as it was not aligned with any other protein sequences. *Th. annulata* contains 2 orthologous proteins, labelled here as 1 and 2.

3.3.4 PF14_0045

Transcript data

With peak expression at 45 hours p.i. (Bozdech *et al*., 2003) this protein is expressed rather later than the previous 3 proteins (figure 3.7 A). In accordance with this, the Le Roch study showed maximum expression of this protein in merozoites (Le Roch et al., 2003). As with the other proteins, absolute expression levels are comparable to that of AMA-1 (illustrated in figure 10). This protein is also expressed in sporozoites, suggesting a conserved invasion related role. However, transcript levels remain relatively high and persistent during early to middle of the ring stage, suggesting a possible function early on after invasion of a new host cell.

Proteomic data

Florens *et al.* (2003) found one peptide hit for this protein but only in the sporozoite form of the parasite. Fortunately the Leiden Malaria Group merozoite proteome (available on PlasmoDB) identified 4 peptide matches, confirming the gene *PF14_0045* encodes a protein which is expressed in both human invasive forms of the parasite. No other hits were detected in subsequent proteomes.

Bioinformatic predictions

This 114 kDa protein enters the secretory pathway by way of signal sequence at the N-terminus from residue 1 to 31, predicted by SignalP 2.0. This is a signal anchor region, with a transmembrane region predicted between amino acids 13-34 by TMHMM, TMpred and TopPred. The protein contains cysteine-rich patches suggesting these regions are particularly folded and also a coiled-coil region, predicted between 391 and 427 by the Coils algorithm (illustrated in figure 3.7 B).

Orthologous proteins in other species

Orthologues of PF14_0045 in other species were identified from PlasmoDB and OrthoMCL-DB, from the orthologue group OG5_205512 (Li et al., 2003) and confirmed using NCBI BLAST tool. This protein appears to be unique to *Plasmodium* with no orthologues present in other sequenced species. Moreover, orthologues are only present in the human and simian malarias *P. knowlesi* (PKH_134210) and *P. vivax* (PVX_086195) suggesting a role specific to primate malaria. The protein sequences were aligned using the ClustalW multiple alignment algorithm (Thompson et al., 1994) in order to visualise and calculate the sequence similarity between species (figure 3.9). Sequence identity between PF14_0045 and its *P.vivax* and *P. knowlesi* orthologues is 33 and 34% respectively. Both orthologous proteins possess the coiled coil region as well as an additional coiled domain at the C-terminus. Although they do contain a signal sequence, it is predicted to be a signal peptide, not anchor, with cleavage predicted to take place after residue 26. No other transmembrane regions exist implying the protein is soluble in these species. The asparagine rich regions of PF14_0045, which are absent from *P. knowlesi* and *P.vivax,* are partly responsible for lack of sequence identity. These regions are thought to be common in *P. falciparum* due to the AT bias of the genome (codons AAT or AAC). Lysine, coded by AAA or AAG, is also a common amino acid for this reason.

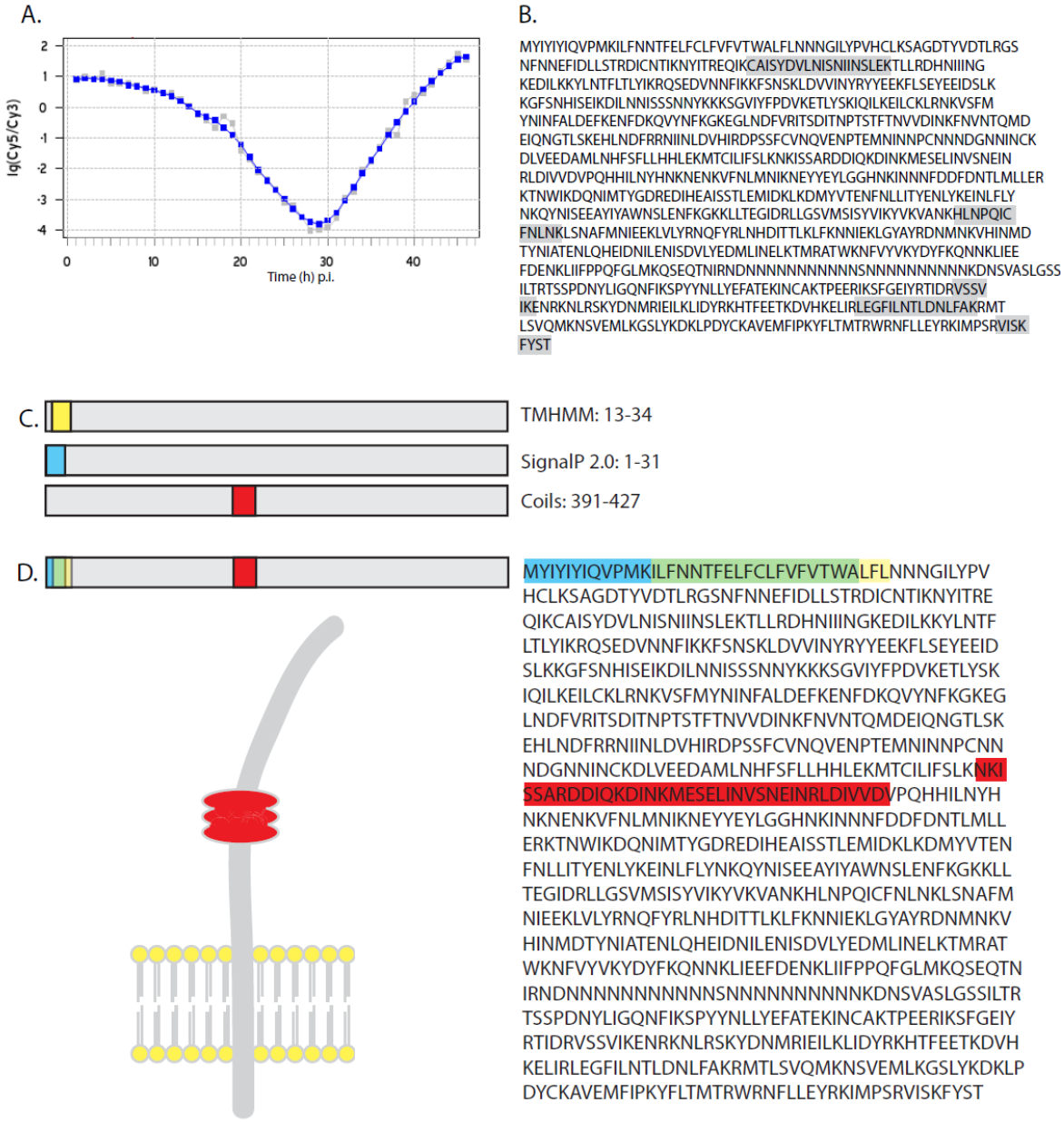
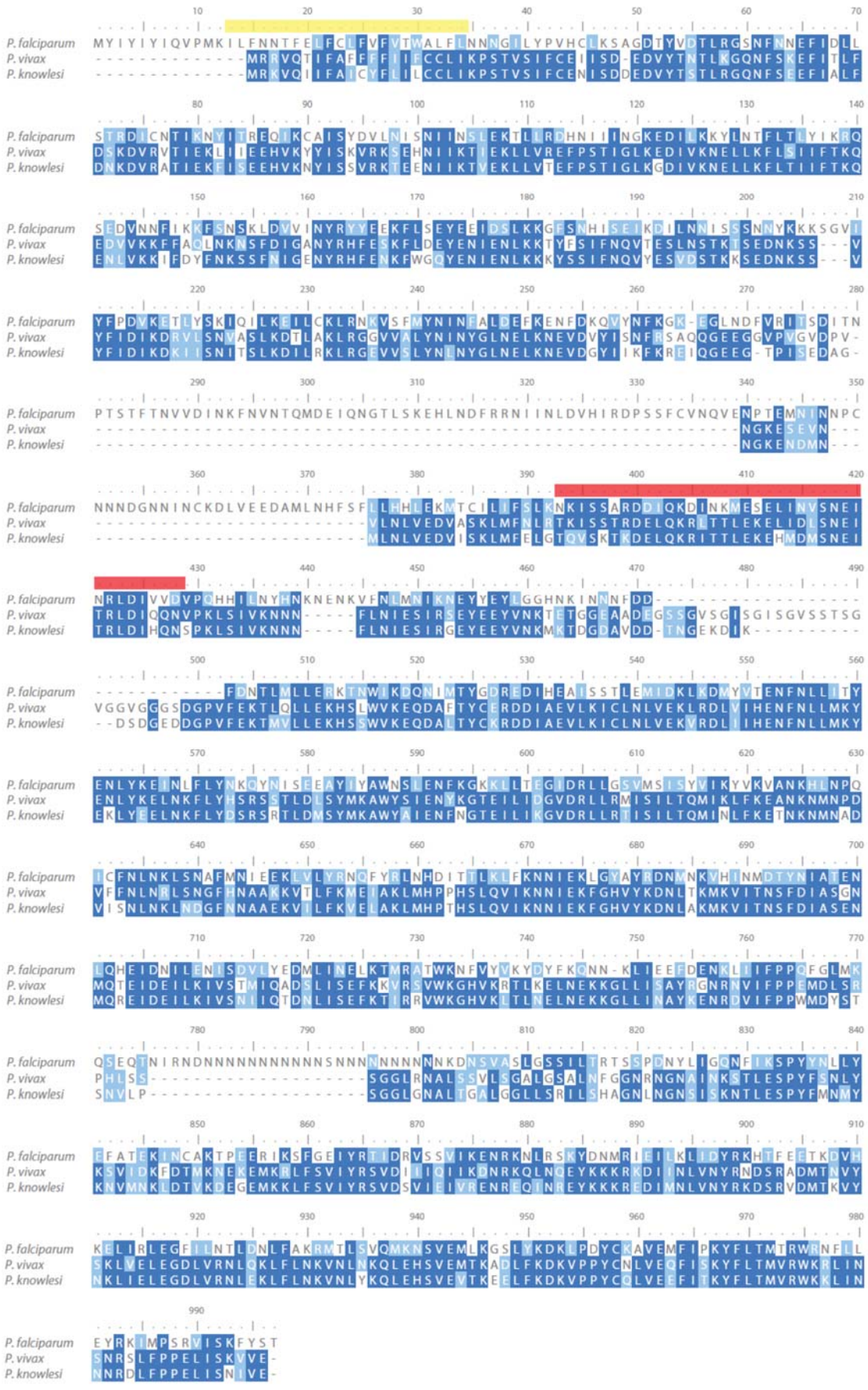


Figure 3.8: *PF14_0045 transcription profile, proteomic data and predicted domain architecture*. (A) The transcriptome from Bozdech *et al* (2003) shows a maximum transcript level of *pf14_0045* at 45 h p.i. Peptides from various regions of the protein were recovered from schizont and merozoite stage parasites in several proteomic studies. These peptides have been mapped to the predicted protein sequence above and highlighted in (B). This 114kDA protein is predicted to contain 1 transmembrane domain, a signal anchor and a coiled coil region (C). The domains were mapped onto the predicted protein sequence and illustrated in (D).

Figure 3.9: *Multiple sequence alignment of PF14_0045 orthologues.* Orthologue sequences from *P. falciparum, P. vivax* and *P. knowlesi* were aligned using the ClustalW multiple alignment algorithm. Residues identical in over 50% of the chosen sequences are shaded in dark blue and those which are functionally similar are highlighted in light blue. Transmembrane domains are shown in yellow based on the *P. falciparum* sequence, regions may be different in other species. The coiled coil region is highlighted in red.


P. falciparum
P. vivax
P. knowlesi

3.3.5 PF02_0040

Transcript data

The transcriptome data produced by Bozdech *et al*. (2003) shows peaks expression of PF02_0040 mRNA in the HB3 clone at around 40 hours p.i. In accordance with this, the Le Roch study showed maximum expression of this protein in late schizonts in both sorbitol and temperature synchronised methods. (Le Roch et al., 2003). As with the other proteins selected, absolute expression levels are comparable with AMA-1 (illustrated in figure 3.10). It is also expressed in sporozoites and gametocyte stages of the intraerythrocytic cycle.

Proteomic data

This protein was detected in merozoites in the proteome by Florens *et al.* (2003) and the Leiden Malaria Group merozoite proteome (available on PlasmoDB) with substantial peptide coverage, confirming the gene *PF02_0040* encodes a protein which is expressed in the erythrocyte invasive form of the parasite. Since then, peptides have been detected in every proteomic study listed on PlasmoDB, covering the full length of the protein with the exception on the extreme N terminus. Although the majority of peptides were predicted from schizont stages (Gerold et al., 1996; Bowyer et al., 2011; Treeck et al., 2011), the protein was also detected in trophozoites and gametocytes (Silvestrini et al., 2010), illustrated in figure 3.9 B.

Bioinformatic predictions

PF02_0040 encodes a protein with a molecular weight of 33 kDa. Like PF14_0045, the protein is predicted to contain 1 signal anchor transmembrane region, as predicted by SignalP 2.0 from amino acid 1-18 and TMHMM, TMpred and TopPred between residues 2 and 21. Classically, transmembrane regions are hydrophobic stretches of around 20 amino acids and either side of the membrane there are charged, hydrophilic residues which strengthen the alpha helical positioning of the domain in the membrane bi-layer. There are lysine residues immediately following the membrane spanning region at the C-terminal end of the signal anchor however the extreme N-terminus lacks any charge.

The positive charge on one side interacting with the negatively charged phospholipid head may be sufficient to hold the protein in place or it is also possible that an intra-membrane interaction anchors the protein. Biochemical validation is required before any conclusions can be drawn.

The Coils algorithm predicts a coiled coil region between amino acids 212-274 (depicted in figure 3.10 C+D). Pfam predicts the presence of a domain named after the yeast secretory pathway mutant *SEC66* (aka *SEC71*) which is part of the pre-protein post-translational translocase which inserts signal sequence-containing proteins into the ER. This domain was predicted from amino-acids 89 to 255 with an e-value of 0.00011 which is significant. A search of non-redundant protein sequences using the blastp algorithm with the PF02_0040 protein sequence did not identify the yeast protein. However using the amino acid sequence of the *T. gondii* orthologue of PF02_0040 (see below) Sec66 from *Saccharomyces* is pulled up in a blastp search. Due to the AT rich nature of the *P. falciparum* genome, amino acids which are coded for by adenosine- or thymidine-rich codons will be preferentially utilised. This has a knock-on effect on protein sequence, and this could be responsible for the lack of detectable sequence identity between *P. falciparum* and yeast proteins. Further experimental evidence is required to confirm PF02_0040 is a functional orthologue of the prototype Sec66 protein. This is discussed further in Chapter 6.

Orthologous proteins in other species

Orthologues of PF02_0040 in other species were identified from PlasmoDB and OrthoMCL-DB, from the orthologue group OG5_142839 (Li et al., 2003) and confirmed using NCBI BLAST tool. The protein is highly conserved in *Plamsodium* spp. (*P. bergei*: PBANKA_030200; *P.chabaudi:* PCHAS_030420; *P. knowlesi*: PKH_041580; *P.* vivax: PVX_003640; *P. yoelii*: PY00894). PF02_0040 has orthologues in other genera, namely *Toxoplasma gondii*, *Neospora canium, Theileria annulata* and several *Cryptosporidium* species from the phylum of Apicomplexa, as well as the eukaryotic marine phytoplankton *Thalassiosira pseudonana* and *Phytophora ramorum,* the oomycete plant pathogen responsible for sudden oak death. The protein sequences were aligned using

the ClustalW multiple alignment algorithm (Thompson et al., 1994) in order to visualise and calculate the sequence similarity between species (figure 3.10). PF02_0040 is 70% identical to its *P. vivax* and *P. knowlesi* orthologues however this level of identity is not shared outside the genus. Identity is 22, 21 and 20% with *T. gondii, N. caninum* and *Cr. hominus* respectively. Despite this diversity, all proteins in the orthologue group are predicted to contain a Sec66 domain. All contain predicted signal sequences except *Tha. pseudomonas* however only *Plasmodium* and *Cryptosporidium* contain predicted transmembrane regions. This is surprising considering the prototype Sec66 of *Saccharomyces* is a transmembrane protein. Although PF02_0040 is predicted to contain a transmembrane region, due to a lack of any charged amino acid at the N-terminus of the protein it is unclear whether the protein will be able to remain in the membrane whilst anchored only at one face, presumably through interaction of lysine at position 22 with the negatively charged phospholipid head. No orthologues outside *Plasmodium* spp. contain a C terminal coiled coil domain.

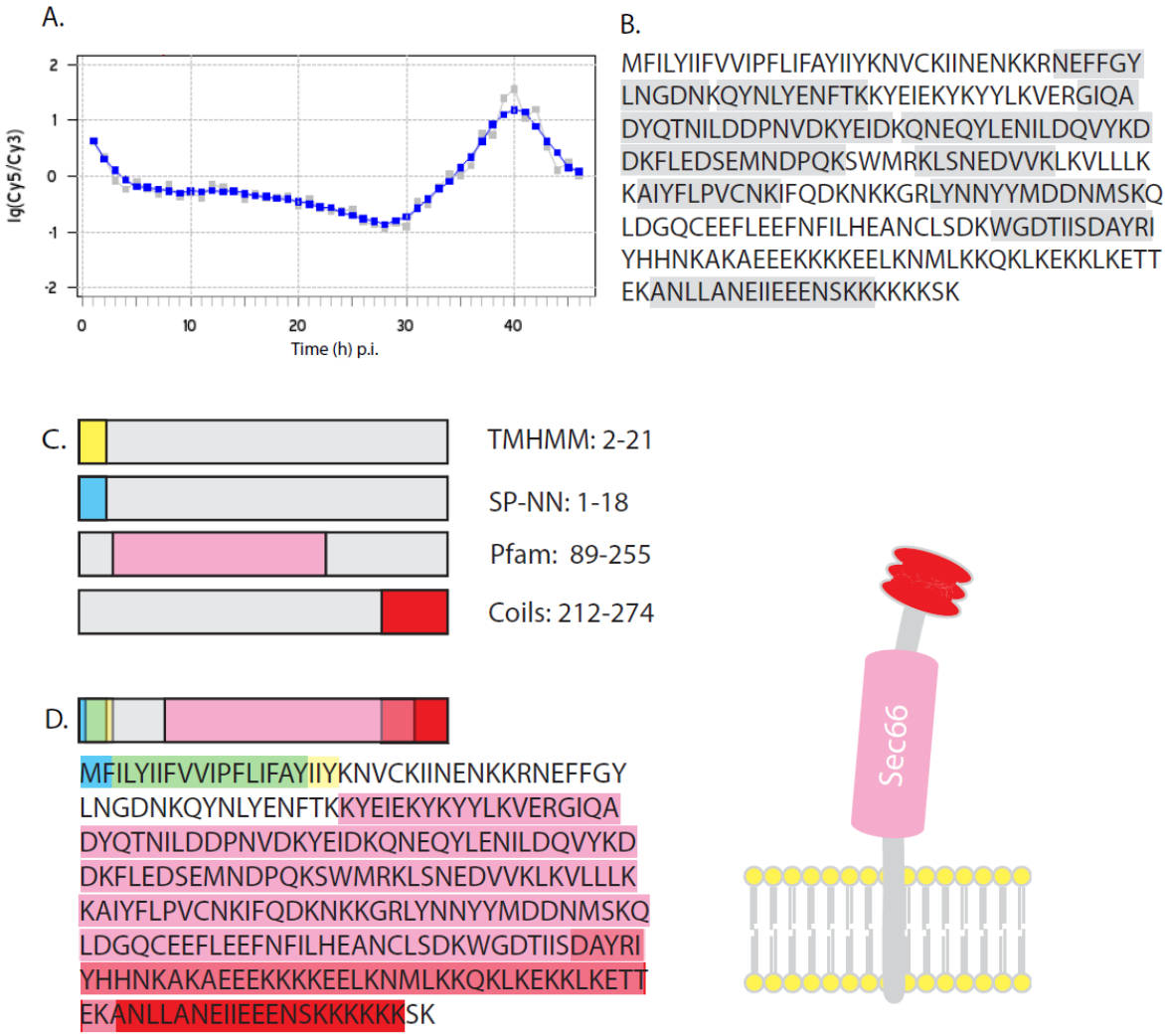


Figure 3.10: *PF02_0040 transcription profile, proteomic and bioinformatic data summary.* (A) The transcriptome from Bozdech *et al* (2003) shows a maximum transcript level of *pf02_0040* at 40 h p.i. Several peptides have been recovered from multiple stages of parasites in a number of proteomic studies. These peptides have been mapped to the predicted protein sequence and are highlighted above in (B). This 33 kDa protein is predicted to contain one transmembrane domain, a signal anchor, as well as a coiled coil region and a domain that bears homology to Sec66 (C). Domains were mapped onto the predicted protein sequence and illustrated in (D).

Figure 3.11 *Multiple sequence alignment of PF02_0040 orthologues.* Sequences from *Plasmodium* spp. and orthologues from Apicomplexa of PF02_0040 were aligned using the ClustalW multiple alignment algorithm. Residues identical in over 50% of the chosen sequences are shaded in dark blue and those which are functionally similar are highlighted in light blue. Transmembrane domains are shown in yellow, coiled coil domain in red and Sec66 domain in pink. Domains are highlighted based on the *Plasmodium* sequences, regions may be different in other Apicomplexa.

Chapter 4: Selected proteins, their expression in bacteria and production of antibodies

4.1 Introduction

Specific antibodies are an essential tool for many of the experiments involved in protein characterisation such as western blot, immunofluorescence and immunoprecipitation. It was therefore necessary to express fragments of the selected proteins for polyclonal antibody production.

Regions within the proteins used for recombinant expression were selected based on hydrophilicity predicted from a Kyte-Doolittle hydropathy plot (Kyte and Doolittle, 1982), hydrophobic patches were avoided. Regions rich in cysteine residues were avoided where possible since correct protein folding may require correct disulfide bond formation not possible in *E.coli*. Initially chosen regions were used in a blastp search against the whole *P. falciparum* 3D7 genome to avoid possible cross-reactivity of the antibodies raised to other *Plasmodium* proteins.. Recombinant protein molecular weight, isoelectric point (pI) and extinction coefficients were calculated by ProtParam (Gasteiger E., 2005). The size and location of chosen regions with respect to the domain architecture of each protein are highlighted in figure 4.1 A.

4.2 Strategy

4.2.1 Ligation-independent cloning

In order to express the regions of each protein depicted in figure 4.1 A, a ligation-independent cloning (LIC) method was adopted, using the Novagen pET vector system. Primers designed to amplify the chosen coding regions contained 5′ or 3′ overhangs which after incubation of the PCR product with T4 DNA polymerase, the 3′-5′ exonuclease activity allowed annealing to occur within the multiple cloning site of the pET vector. In cases where the region for recombinant expression extended between

multiple exons, cDNA was used for amplification. Sizes of expected PCR products are indicated in Table 4.1.

All LIC inserts created were initially annealed into the pET30-Xa/LIC vector which encodes an N-terminal His- and S-tag, allowing identification and affinity purification of recombinant protein from *E. coli* proteins. It also encodes a factor Xa protease cleavage site for specific removal of the N-terminal tags from the recombinant protein. In cases where recombinant protein was not initially soluble, pET32-Xa/LIC was used as an alternative. This vector contains an additional thioredoxin (Txr) tag which aids solubility. These vectors were used to chemically transform *E. coli* Giga Singles competent cells and vector DNA was prepared from multiple clones. To screen for inserts of the correct size, plasmid DNA was digested using *Kpn*I and *Xho*I restriction enzymes, which cleave either side of the multiple cloning site but do not cleave the amplified gene fragment, followed by gel electrophoresis. The DNA sequences of inserts of positive clones were confirmed by commercial DNA sequencing before transformation into expression competent *E. coli*.

4.2.2 Protein expression and solubility

The preparations of cloned plasmid DNA were used to transform *E. coli* BL21 (DE3) pLysS protein expression cells. If solubility was low, a number of different cell types were tested. For largely insoluble proteins, BL21 (DE3) Tf2 cells containing an additional plasmid which encodes the chaperones GroEL/GroES and trigger factor, were used in order to promote protein folding and hopefully solubility. A small volume of culture was used for induction of protein expression in the first instance to test the expression levels and solubility of the protein. Typically, protein expression was induced using 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at an optical density reading at 600nm of 0.6 for 4 hours at 22°C however for less soluble proteins this was amended to 18°C overnight. Soluble and insoluble protein fractions were prepared and resolved under reducing conditions by SDS-PAGE and visualised by Coomassie blue stain. A large number of *E.coli* proteins are visible by Coomassie so to confirm the suspected protein is

in fact the His-tagged protein of interest a western blot was performed on an identical SDS-PAGE gel, with an anti-His antibody. Protein expression conditions for each protein are summarised in Table 4.1.

4.2.3 Protein purification and quantification

Protein expression was increased to 100 ml culture in order to test purification methods. Taking advantage of the hexa-His tag's affinity for nickel, the tagged recombinant protein was bound to nickel-agarose (Ni-NTA), washed and competitively eluted with increasing concentrations of imidazole so that an optimum wash and elution concentration could be determined. Once optimised, protein production was increased to 2 litres of culture or enough to produce a sufficient quantity for antibody production. Purified protein was buffer-exchanged into PBS and concentration was estimated by comparison with BSA standards on SDS-PAGE and calculated using the extinction coefficient and the absorbance at 280 nm.

4.2.4 Rabbit polyclonal antibody production and affinity purification

The concentration of the pure recombinant protein in PBS was adjusted to 0.4 mg/ml or 200µg in 500µl per immunisation, to be diluted 1:1 with adjuvant. Rabbits were immunised at Harlan Biosciences following standard protocol. Six immunisations were required per protein for one rabbit therefore a minimum of 1.2 mg of protein is required for antibody production. Animals to be used were pre-screened prior to inoculation to avoid previously existing cross reacting antibodies. Pre-screening was carried out by western blot against whole rbc, schizont and merozoite lysates resolved by SDS-PAGE. Selected rabbits were screened after the third and fourth inoculations of recombinant protein to ensure the antibody response was sufficient to recognise a protein of the predicted size. Terminal bleeds were affinity column purified with recombinant protein coupled to CNBr (for methods see section 2.2.7). As a number of future experiments require recognition by secondary antibody, the IgG fraction of the affinity purified protein was enriched using protein G sepharose (2.2.9). Antibody was dialysed into PBS and concentrated to 1mg/ml and stored in 50µl aliquots at 4°C or -20°C until needed.

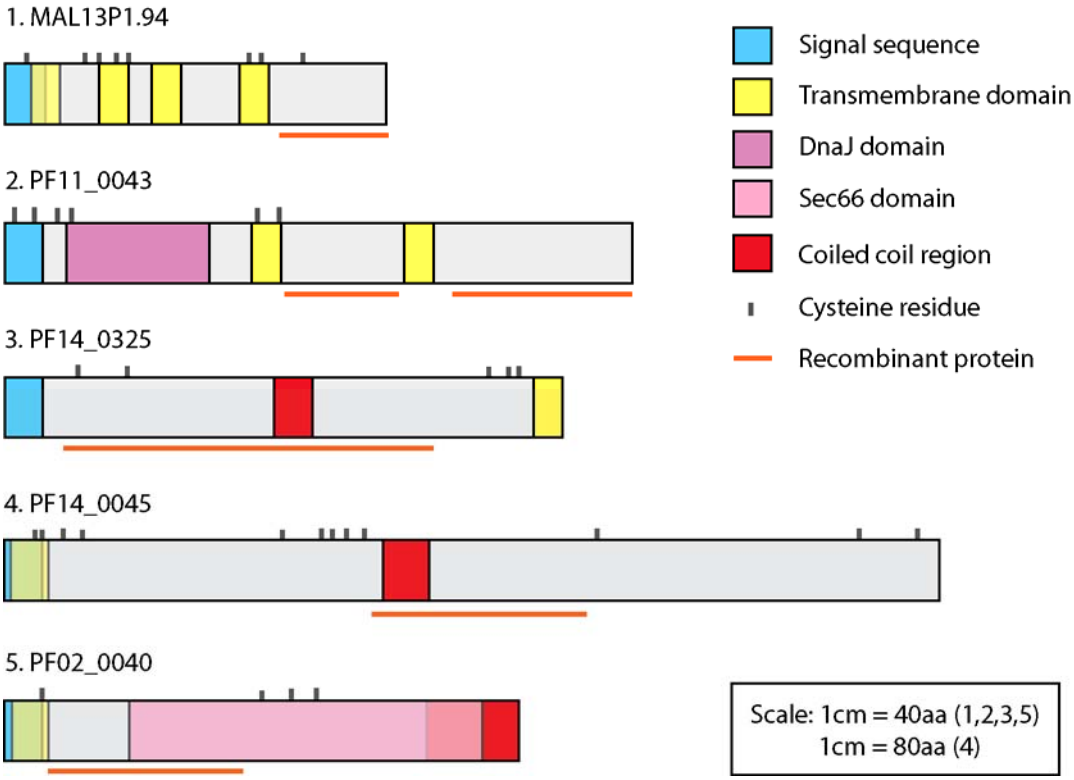


Figure 4.1: *Illustration of protein fragments selected for recombinant expression.* Schematic representation of each selected protein. Signal peptide, transmembrane domains and other domains of interest are shown colour-coded above (key in figure). Regions selected for recombinant protein expression and antibody production are underlined in orange.

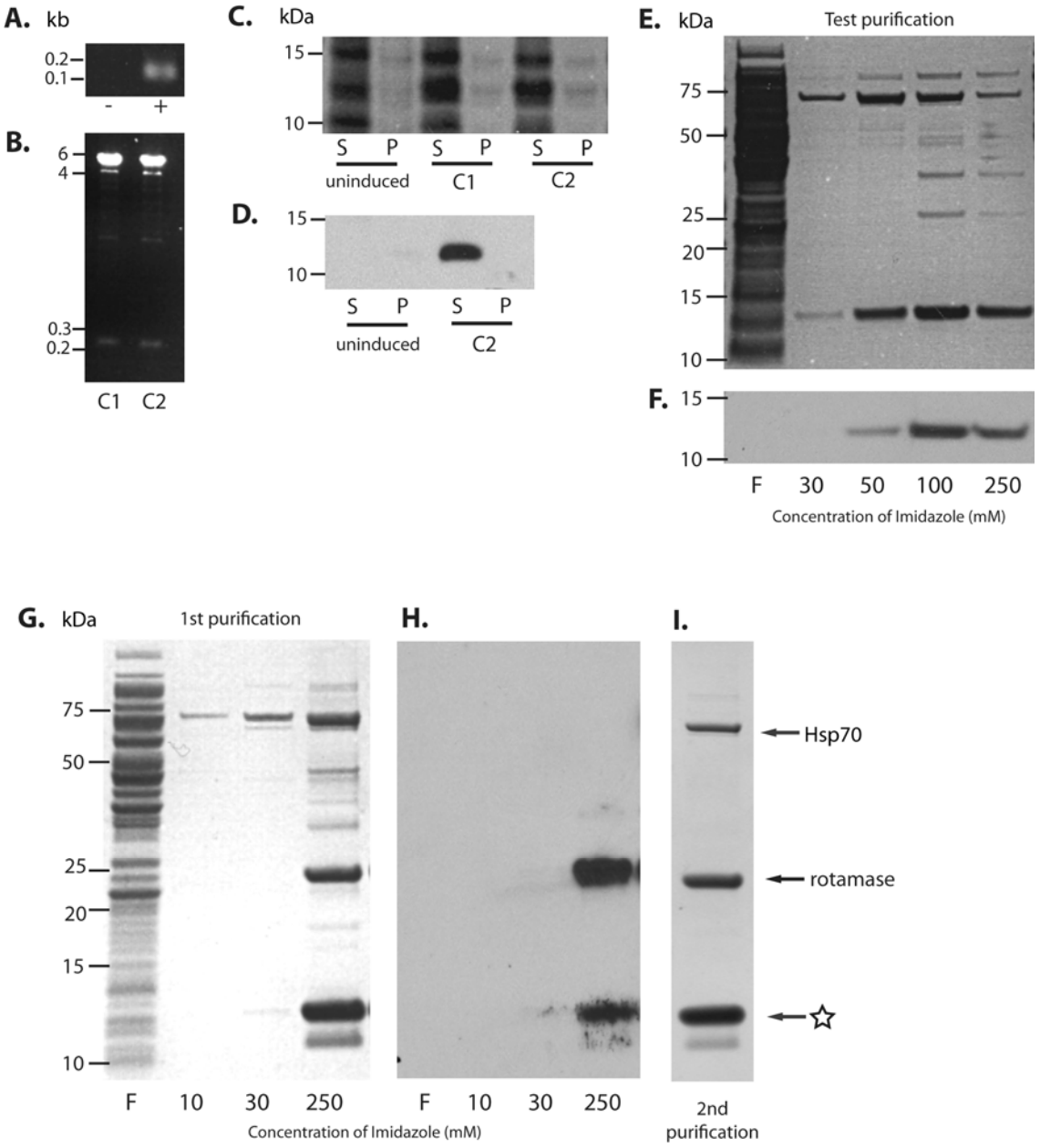
4.3 Recombinant expression results

4.3.1 MAL13P1.94

A 5kDa fragment from the C-terminus of MAL13P1.94 was selected for expression (illustrated in figure 4.1). Ideally the fragment would be longer however this was the longest stretch of hydrophilic amino acids not predicted to be associated with membrane in this 22 kDa 4 TM protein. The region was amplified from *P. falciparum* 3D7 genomic DNA, incorporating pET30-Xa/LIC overhangs required for LIC (figure 4.2 A). After annealing to the expression vector and transformation into *E. coli* BL21 (DE3) pLysS expression cells, a 10kDa soluble protein was expressed (including the ~5kDa tag) from a 1ml test culture at 22°C for 4 hours (figure 4.2 C+D). Test purification using increasing amounts of imidazole as a Histidine competitor for Ni-NTA revealed a small amount of the protein began to elute from the column at 30mM imidazole but more so at 50 and 100mM (E+F). Since the protein was bound to the nickel column in 10mM imidazole, washing the column with any less than 30mM imidazole would fail to remove non-specific binding proteins from the column therefore 30mM imidazole was used in the column wash buffer and 250mM imidazole in the elution buffer for future purifications. Expression culture was scaled-up to 2 litres and after 2 rounds of Ni-NTA purification, produced ample protein for antibody production (figure 4.2 G+I). Unfortunately, a relatively large amount of an additional 25 kDa protein was also expressed and purified alongside the recombinant protein of interest. Curiously this protein also seems to be His-tagged as it is recognised by rabbit anti-His antibody (figure 4.2 H). This protein was not present in the smaller test cultures. In order to establish the 25 kDa protein's identity, the band was excised from the Coomassie-stained SDS-PAGE gel and prepared for MALDI-TOF mass spectrometry analysis. This revealed the protein to be an FKBP-type peptidyl-prolyl cis-trans isomerase (rotamase) of bacterial origin that possesses a pseudo-His tag within its amino acid sequence. The binding of *E. coli* proteins is a common occurrence with immobilised metal affinity chromatography (IMAC). This particular contaminant, rotamase or SlyD, has been found not to bind to TALON™ which is a cobalt resin and can be used to purify His-tagged proteins (Bolanos-Garcia and Davies, 2006). Unfortunately MAL13P1.94 also does not bind to cobalt so this cannot be used as a method of purification. When preparing the expression culture, *E.coli* was

grown to an O.D. of 0.6 at 600nm at 37°C before induction and incubation at 22°C for expression. To investigate whether rotamase was up-regulated due to heat shock in large scale cultures, the culture was cooled down slowly, first to 32°C, then 28°C before reaching 22°C. This did reduce expression levels of the contaminant however there was still a considerable amount of the contaminating protein present in the eluant. Since the rotamase contamination only occurred in large scale (4x500ml) culture, I used less resin in the hope that the recombinant protein would preferentially bind to the nickel resin above rotamase however this was not successful. In the end, the protein cocktail containing recombinant MAL13P1.94, rotamase and a small amount of Hsp70 was used for antibody production (figure 4.2 I). There are predicted FKBP-type *cis-trans* isomerases in *P. falciparum* however the protein sequences are divergent and antibodies would not be expected to cross-react. Western blot analysis of test bleeds showed rabbit sera was reactive with two bands in schizonts, one at ~50kDa and one at ~25kDa, the latter of which is likely to be MAL13P1.94. Two bands were again present in merozoites however this time a 15kDa band appeared along with the 25kDa band (presumed to be MAL13P1.94) but not the 50kDa protein. Terminal bleed serum was affinity and IgG purified however the banding pattern persisted (figure 4.7 A). Further investigation into this protein is detailed in chapter 7.

Figure 4.2. *MAL14P1.94 recombinant protein expression.* The DNA sequence for the chosen C-terminal region of MAL13P1.94 was amplified using specific primers from *P. falciparum* 3D7 genomic DNA, incorporating pET30-Xa/LIC overhangs. The PCR product (149bp) was analysed by gel electrophoresis where a band between 0.1 and 0.2 kb was visible (A; -, + without and with template DNA). The gene fragment was annealed into the pET30-Xa/LIC expression vector and transformed into *E. coli Gigasingles*. Plasmid DNA was prepared from resulting clones and screened for inserts by digestion with *Kpn*I and *Xho*I before resolving digested fragments by gel electrophoresis (B). DNA size is shown in kb. Clones containing inserts with correct sequences were tested for soluble recombinant protein expression. A band of around 12 kDa is visibly increased in the soluble fractions (S) compared to the un-induced control when resolved by SDS-PAGE, molecular weight is shown in kDa (C). No additional bands were visible in the pellet fractions (P). Solubility was confirmed by western blot using anti-His antibody (D). Protein purification conditions were optimised using a 100 ml culture induced with 1mM IPTG for 4h at 22°C and elution of recombinant protein from a nickel-agarose column was tested using 30, 50, 100 and 250 mM imidazole. Eluants were resolved by SDS-PAGE and stained with Coomassie blue (E) or probed with an anti-His antibody (F.). Column flow-through (F) is shown in the left-most lane. A small amount of protein was eluted from the column with 30mM imidazole however the majority was eluted at a concentration ≥50mM. Protein production was increased, expressing from 2 litres of transformed *E.coli*. The soluble protein fraction was bound to nickel-agarose in 10 mM imidazole and washed with 10 and 30 mM imidazole buffer before eluting in 10 ml of 250 mM imidazole. Samples were taken and resolved by SDS-PAGE (G). The 25 kDa band appeared to be His-tagged, shown here on a western blot (H). Purification was repeated to increase the purity of the preparation although 3 additional contaminant proteins could not be removed (I). MALDI-TOF mass spectrometry analysis revealed the 25kDa protein to be rotamase and the identity of the ~70kDa protein as Hsp70. The band just below MAL13P1.94 (indicated with a star) is degraded recombinant protein.


A.
kb
0.2
0.1
-
+
B.
6
4
0.3
0.2
C1
C2
C.
kDa
15
10
S
P
S
P
S
P
uninduced
C1
C2
D.
15
10
S
P
S
P
uninduced
C2
E.
kDa
Test purification
75
50
25
20
15
10
F.
15
10
F
30
50
100
250
Concentration of Imidazole (mM)
G.
kDa
1st purification
75
50
25
20
15
10
F
10
30
250
Concentration of Imidazole (mM)
H.
F
10
30
250
I.
Hsp70
rotamase
2nd
purification

4.3.2 PF11_0443

C-terminal protein fragment (PF11_0443CT)

Initially, a hydrophillc 11kDa fragment from the C-terminus of PF11_0443 was selected for expression (figure 4.1). The region was amplified from *P. falciparum* 3D7 genomic DNA, incorporating pET30-Xa/LIC overhangs required for LIC (figure 4.3 A). After annealing to the expression vector and transformation into *E. coli* BL21 (DE3) pLysS expression cells, a 17kDa soluble protein was expressed (including the ~5kDa tag) from a 1ml test culture at 22°C for 4 hours (figure 4.3 C+D). Test purification revealed the protein began to elute from the column at 50mM imidazole, thus a lower concentration of 30mM imidazole was used to wash the nickel column and 250mM for elution (figure 4.3 E+F). Expression culture was scaled-up to 2 litres which, after 2 purification steps, produced ample protein for antibody production (figure 4.3 G). As with MAL13P1.94, rotamase was co-purified with recombinant PF11_0443 using nickel and similar attempts to reduce the contaminant were unsuccessful. Western blot analysis of affinity and IgG purified antibodies showed rabbit sera was reactive with a single ~37kDa band in both schizonts and merozoites but not rbcs (figure 4.7 B). These antibodies were used in further experiments, detailed in chapter 5.

"Middle" region protein fragment (PF11_0443M)

As the project progressed it became desirable to produce an antibody of a different species for immunoprecipitation studies and also from the "other side" of the membrane for topology identification. For this purpose the "middle" section of the protein was selected, between the 2 membrane-spanning regions from residues 136-202 (figure 4.1). Although initially insoluble when expressed from pET30-Xa/LIC in *E.coli* BL21 (DE3) pLysS, when re-transformed into BL21 (DE3) Tf2 cells co-expressing chaperones GroEL/ES and trigger factor at 18°C O/N the protein was around 20% soluble and 2 litres of culture produced sufficient protein for rat antibody production (figure 4.4). As with the C-terminal fragment, antibodies raised recognised a single 37 kDa band in western blots of both schizonts and merozoites but failed to recognise any rbc proteins (figure 4.7 C).

Figure 4.3. *PF11_0443 C-terminal recombinant protein expression.* A DNA fragment encoding the chosen C-terminal region was amplified using specific primers from *P. falciparum* 3D7 genomic DNA, incorporating pET30-Xa/LIC overhangs required for LIC. The PCR product (284 bp) was analysed by gel electrophoresis where a band between 0.2 and 0.3 kb was visible (A). The gene fragment was annealed into the pET30-Xa/LIC expression vector and transformed into *E. coli Gigasingles*. Plasmid DNA was prepared from resulting clones and screened for inserts by digestion with *Kpn*I and *Xho*I before resolving digested fragments by gel electrophoresis (B). DNA size is shown in kb. Clones containing correct sequence inserts were tested for soluble recombinant protein expression. A band of around 17k Da is visibly increased in the soluble fractions (S) compared to the un-induced control when resolved by SDS-PAGE, molecular weight is shown in kDa (C) No additional bands were visible in the pellet fractions (P). Solubility was confirmed by western blot using anti-His antibody (D). Protein purification conditions were optimised using a 100 ml expression culture by binding to a nickel-agarose column and using 30, 50, 100 and 250 mM concentrations of imidazole for washing/elution. Eluants were resolved by SDS-PAGE and stained with Coomassie blue (E) or probed with an anti-His antibody (F.). Column flow-through (F) is shown in the left-most lane. Recombinant PF11_0443CT was eluted at a concentration ≥50mM. Protein production was increased, expressing from 2 litres of transformed *E.coli*. The soluble protein fraction was bound to nickel-agarose in 10mM imidazole and washed with 10 and 30mM imidazole buffer before eluting in 10ml of 250mM imidazole. Samples were taken and resolved by SDS-PAGE (G). Purification was repeated to increase the purity of the preparation although 3 additional contaminant proteins could not be removed. MALDI-TOF mass spectrometry analysis revealed the 25kDa protein to be rotamase and the identity of the ~70kDa protein as Hsp70. The band just below PF11_0443 (indicated with a star) was deemed to be degraded recombinant protein.

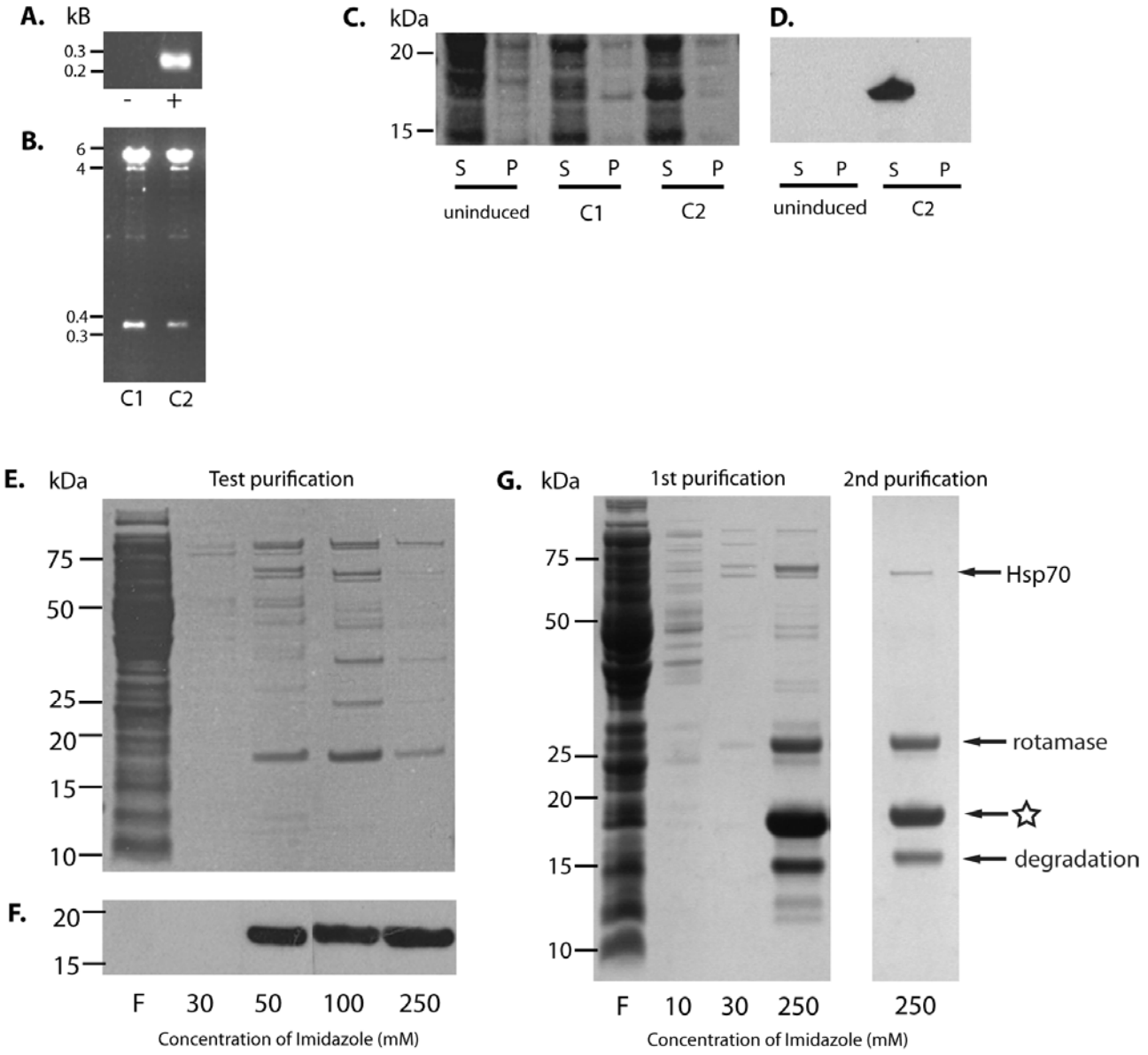
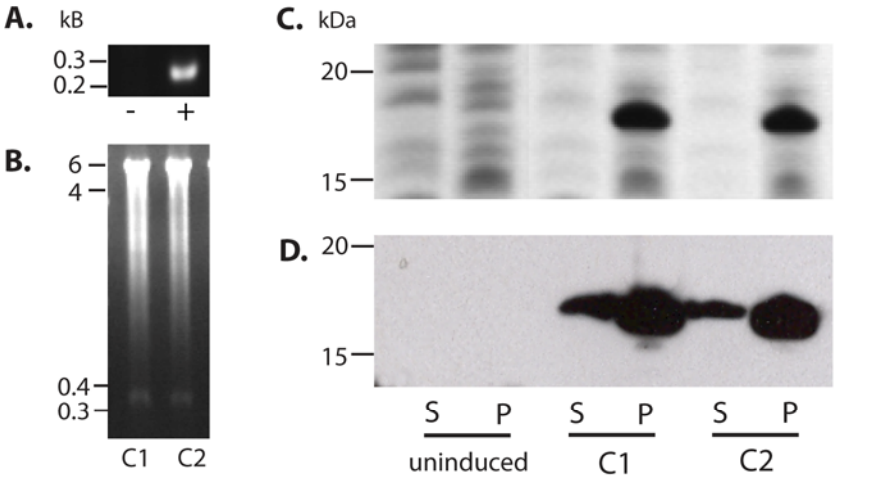
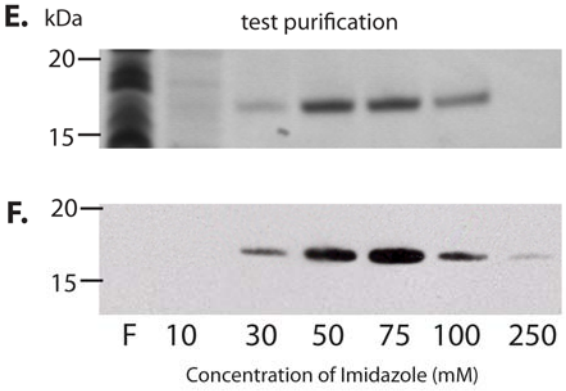
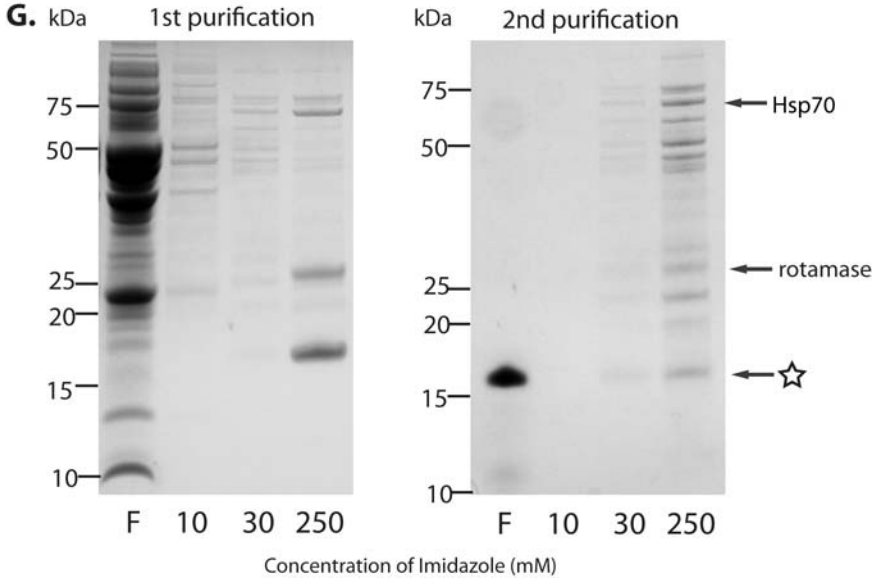

A.
kB
0.3
0.2
- +
B.
6
4
0.4
0.3
C1 C2
C.
kDa
20
15
S P S P S P
uninduced C1 C2
D.
S P S P
uninduced C2
E.
kDa
Test purification
75
50
25
20
15
10
F.
20
15
F 30 50 100 250
Concentration of Imidazole (mM)
G.
kDa
1st purification
2nd purification
75
50
25
20
15
10
Hsp70
rotamase
degradation
F 10 30 250 250
Concentration of Imidazole (mM)

Figure 4.4: *PF11_0443 "middle" section recombinant protein expression.* The chosen region of PF11_0443 was amplified using specific primers from *P. falciparum* 3D7 genomic DNA, incorporating pET30-Xa/LIC overhangs required for LIC. The PCR product (233 bp) was analysed by gel electrophoresis where a band between 0.2 and 0.3 kb was visible (A.). The gene fragment was annealed into the pET30-Xa/LIC expression vector and transformed into *E. coli Gigasingles*. Plasmid DNA was prepared from resulting clones and screened for inserts by digestion with *Kpn*I and *Xho*I before resolving digested fragments by gel electrophoresis (B.). DNA size is shown in kb. Clones containing correctly sequenced inserts were tested for soluble recombinant protein expression after transformation into E. coli BL21 DE3 Tf2. A band of around 17 kDa is visible by Coomassie blue stain (C.) in the pellet (P) fraction only, however an identical gel is probed with anti-His antibody (D.), it is evident there is a proportion of protein present in the soluble fraction (S). Molecular weight is shown in kDa. Soluble protein purification conditions were optimised from a 100 ml culture testing elution from a nickel-agarose column using 30, 50, 100 and 250 mM concentrations of imidazole. Eluants were resolved by SDS-PAGE and stained with Coomassie blue (E) or probed with an anti-His antibody (F.). Column flow-through (F) is shown in the left-most lane. A small amount of protein was eluted from the column with 30mM imidazole however the majority was eluted at concentration ≥50mM. Protein production was increased, expressing from 2 litres of transformed *E.coli*. The soluble protein fraction was bound to nickel-agarose in a 10mM imidazole solution and washed with 10 and 30mM imidazole buffer before eluting in 10ml of 250mM imidazole. Samples were taken and resolved by SDS-PAGE (G). Purification was repeated to increase the purity of the preparation. On repetition, the recombinant protein of interest did not re-bind the column and is present in the flow-through however the contaminant proteins did bind to the column, thus negatively purifying the protein preparation.


A.
kB
0.3
0.2
- +
B.
6
4
0.4
0.3
C1 C2
C.
kDa
20
15
D.
20
15
S P S P S P
uninduced C1 C2


E.
kDa
test purification
20
15
F.
20
15
F 10 30 50 75 100 250
Concentration of Imidazole (mM)

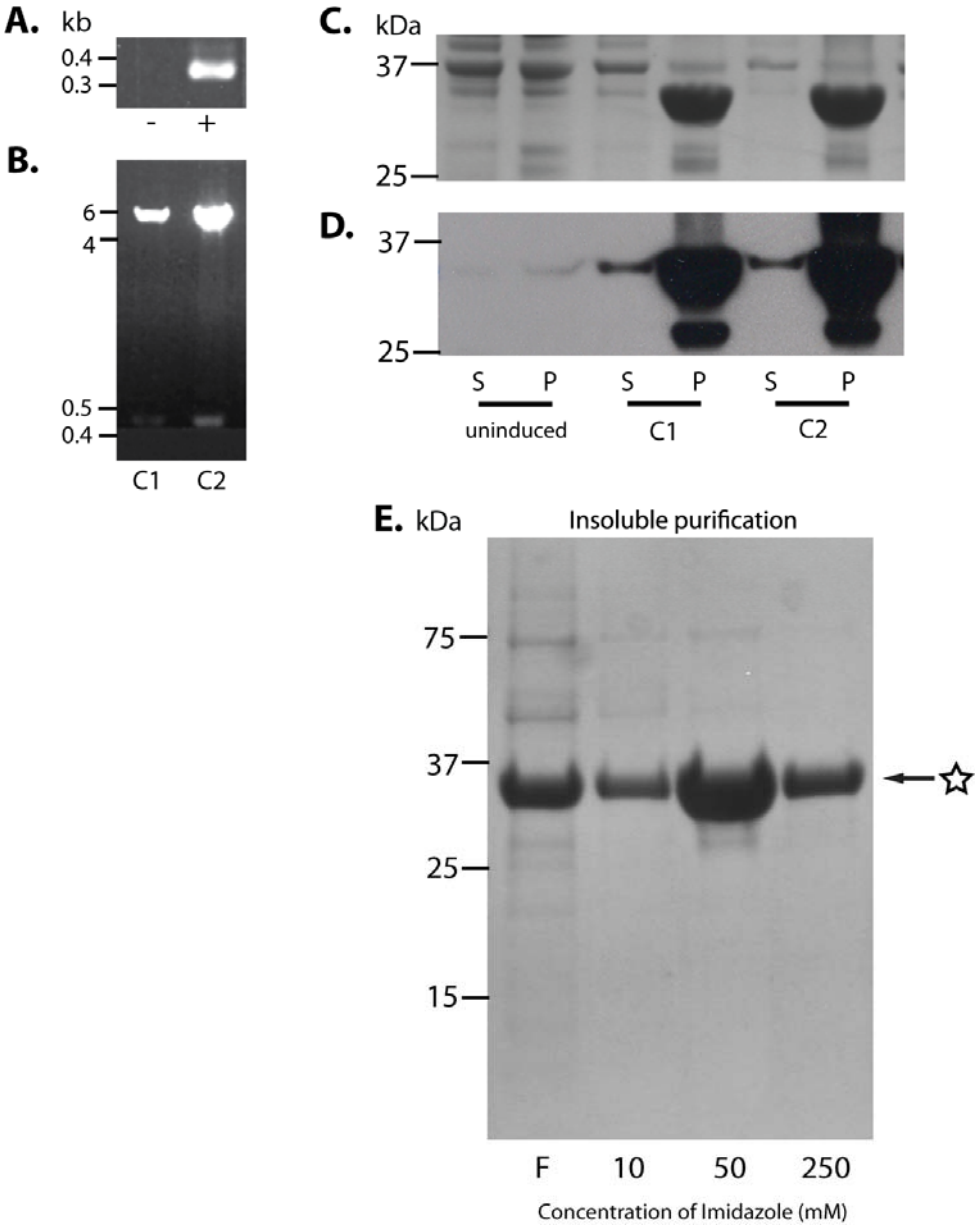

G.
kDa
1st purification
75
50
25
20
15
10
F 10 30 250
kDa
2nd purification
75
50
25
20
15
10
Hsp70
rotamase
F 10 30 250
Concentration of Imidazole (mM)

4.3.3 PF02_0040

Initially a 14kDa fragment from PF02_0040 was selected for expression from the region immediately following the signal anchor from residues 29-139 (figure 4.1). The region was amplified from *P. falciparum* 3D7 cDNA from schizont stage parasites 39 h p.i. (figure 4.5 A). The PCR fragment was annealed into pET30-Xa/LIC after treatment with T4 DNA polymerase to generate single strand complementary overhangs. However the resulting protein was almost completely insoluble and showed low expression levels. Protein expressed using the pET32-Xa/LIC vector was also largely insoluble irrespective of which expression cells and temperature of expression was tested. Attempts to increase protein solubility by incorporating an N-terminal glutathione-S-transferase (GST)-tag were also unsuccessful, as was changing the region expressed to include the whole protein minus the signal anchor (residues 42-276). The low levels of expression may have been due to the AT-rich sequence of this protein. To overcome this, an *E.coli* codon-optimised gene was constructed by GeneArt® and new primers designed based on the new sequence to incorporate residues 29-139 into pET32-Xa/LIC. Although this dramatically increased expression levels, the protein remained insoluble (figure 4.5 C+D). Due to the nature of antibody-antigen recognition, it is desirable to raise antibodies against a soluble, and therefore presumed correctly folded, protein however it is possible to raise antibodies against insoluble aggregates and often these are more immunogenic. Bearing this in mind and owing to the relatively high levels of expression, the recombinant protein fragment was expressed from pET32-Xa/LIC at 37°C for 4 hours. Protein was extracted from inclusion bodies from 2 litres of cell culture, solubilised in 8M urea and purified over a nickel column (figure 4.5 E). The eluted protein was dialysed into PBS where, as expected, it precipitated. The concentration of the aggregate was estimated from the band size produced by SDS-PAGE compared to a know standard. Western blot analysis of test bleeds revealed that the rabbit sera was reactive with a band just under the 37 kDa marker in both schizonts and merozoites but not rbcs. Estimated molecular weight for Pf02_0040 is 33 kDa,. Unfortunately, this was not the only strong band on the blot. The serum also recognised a doublet running close to the 20 kDa marker. Terminal bleed serum was affinity purified from protein blotted onto nitrocellulose membrane, a

procedure which helped to purify the antibody but did not remove the 20 kDa doublet even when used at a dilution of 1:10000 on western blot (figure 4.7). In a further attempt to obtain soluble recombinant protein to use in affinity purification of these antibodies, I buffer-exchanged of PF02_0040 recombinant protein into room temperature PBS using a a PD-10 de-salting column. This kept the protein in solution. The recombinant protein again precipitated when transferred to 4°C thus binding of the protein to CNBr sepharose in order to form an affinity column was done at room temperature for 2 hours. Affinity purified antibodies were further IgG purified to be used in future experiments however the 20 kDa bands did persist. PF02_0040 is discussed in chapter 7.

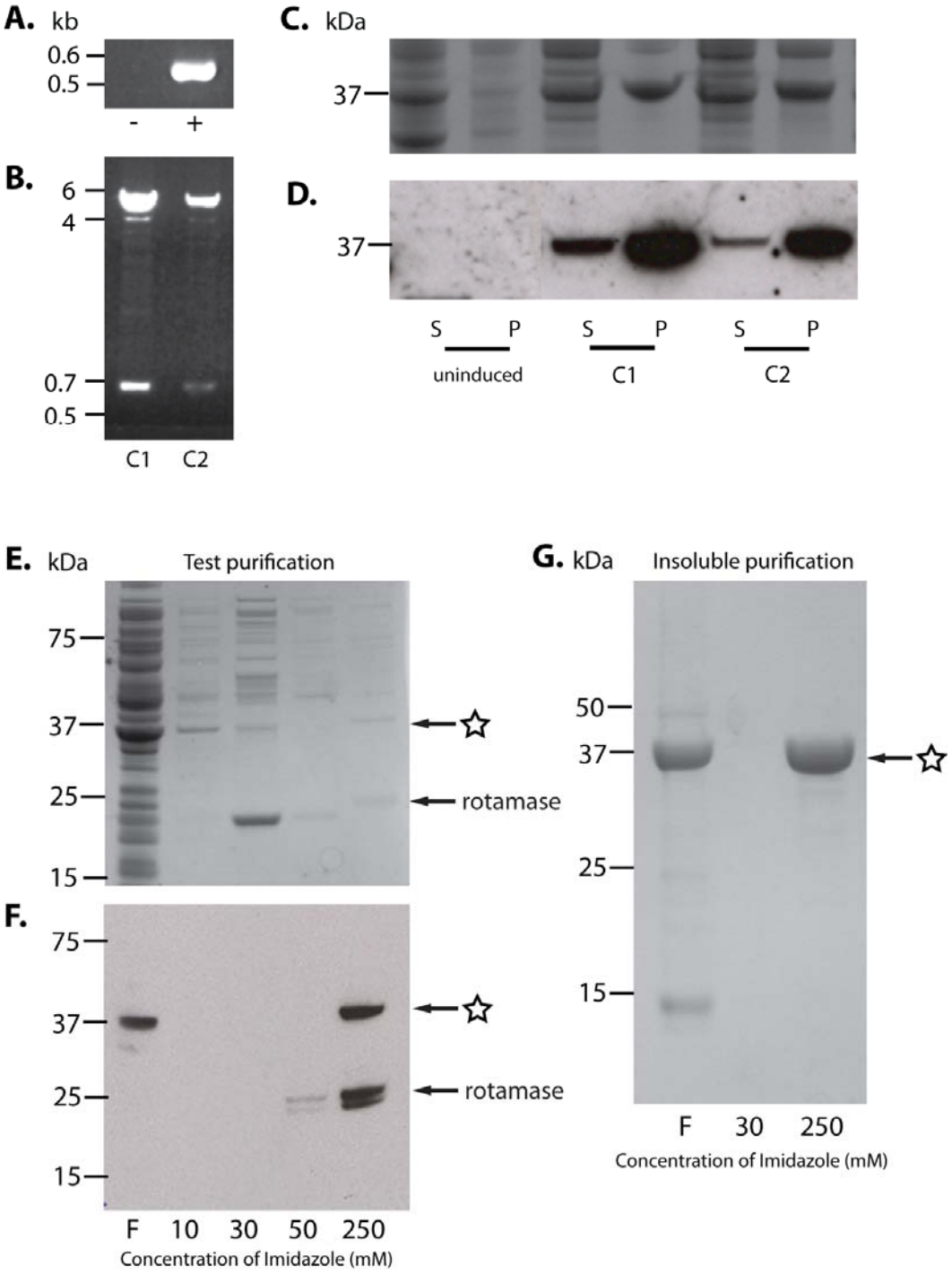
Figure 4.5. *PF02_0040 recombinant protein expression.* The DNA fragment encoding the chosen region for protein expression was amplified using primers G and H (Chapter 2, Table 2.1) from a plasmid containing a PF02_0040 codom optimised for *E. coli* expression, incorporating pET32-Xa/LIC overhangs required for LIC. The PCR product (366 bp) was analysed by gel electrophoresis where a band between 0.3 and 0.4 kb was visible (A). The gene fragment was annealed into the pET32-Xa/LIC expression vector and transformed into *E. coli Gigasingles*. Plasmid DNA was prepared from resulting clones and screened for inserts by digestion with *Kpn*I and *Xho*I before resolving digested fragments by gel electrophoresis (B). DNA size is shown in kb. Clones containing inserts with correct DNA sequence were tested for protein expression after transformation into BL21 DE3 pLysS. A band of around 37 kDa is visible by Coomassie blue stain (C) in the pellet (P) fraction only. Although western blot of an equivalent gel probed with anti-His antibody (D) shows a thin band in the soluble fractions (S), it is most likely the result of over-spill from the abundant pellet fraction. Molecular weight is shown in kDa. Due to its abundance, the insoluble protein fraction was prepared from 1l of culture and bound to a nickel-agarose column. The column was washed with 10 mM and 50 mM imidazole buffer before eluting with 10 ml of 250 mM imidazole containing buffer. Samples were taken and resolved by SDS-PAGE (E). Most of the protein eluted from the column with 50 mM imidazole wash buffer and as this eluant contained a large concentration of protein and was free of contaminating proteins, this fraction was used after dialysis for antibody production.


A.
kb
0.4
0.3
- +
B.
6
4
0.5
0.4
C1 C2
C.
kDa
37
25
D.
37
25
S P S P S P
uninduced C1 C2
E.
kDa
Insoluble purification
75
37
25
15
F 10 50 250
Concentration of Imidazole (mM)

4.3.4 PF14_0045

A 22 kDa fragment from the centre of PF14_0045 was selected containing the predicted coiled coil domain from residues 391-577, avoiding the cysteine-rich areas (figure 4.1). A DNA fragment encoding this region was amplified from *P. falciparum* 3D7 genomic DNA (figure 4.6 A). Initially this fragment was annealed into pET30-Xa/LIC and standard expression conditions tested. Only approximately 10% of the protein was soluble. Changing expression vector to pET32-Xa/LIC also resulted in a largely insoluble protein irrespective of which expression cells and temperature of expression was used (figure 4.5 C+D), although expression levels were higher than when expressed from pET30-Xa/LIC. Attempted test purification of soluble material from a 100 ml culture resulted in very low amounts of purified recombinant protein which eluted with a number of contaminant proteins including rotamase (figure 4.5 E+F). This lack of abundance of soluble protein made it impractical to produce sufficient quantities for antibody production. Attempts to increase protein solubility by incorporating a glutathione-S-transferase (GST)-tag were also unsuccessful. The recombinant protein was therefore purified from inclusion bodies of a 2 liter expression culture of BL21DE3pLysS transformed with pET32-Xa-PF14_0045 expression construct induced at 37°C for 4 hours with 1mM IPTG. The protein extract was solubilised in 8M urea and purified over a Ni-NTA column (figure 4.5 G). The eluted protein was dialysed into PBS where, as expected, it precipitated. Concentration was estimated comparing intensity of Coomassie stained bands separated by SDS-PAGE with that of a marker of known concentration. Western blot analysis of test bleeds revealed rabbit sera was reactive with the same 20 kDa doublet observed with PF02_0040 antibodies. Only a very weak band was observed at approximately 100 kDa, close to the predicted size of PF14_0045 at 110kDa.Terminal bleed antibodies were affinity purified by binding to recombinant protein blotted onto nitrocellulose membrane. This failed to remove antibody specificities to the cross reacting bands (figure 4.7). PF14_0045 recombinant protein was not soluble after buffer-exchange using a PD-10 desalting column and therefore no further affinity purification was attempted.

Figure 4.6. *PF14_0045 recombinant protein expression.* The chosen DNA region was amplified using specific primers from *P. falciparum* 3D7 genomic DNA, incorporating pET32-Xa/LIC overhangs required for LIC. The PCR product (593 bp) was analysed by gel electrophoresis where a band between 0.5 and 0.6 kb was visible (A). The gene fragment was annealed into the pET32-Xa/LIC expression vector and transformed into *E. coli Gigasingles*. Plasmid DNA was prepared from resulting clones and screened for inserts by digestion with *Kpn*I and *Xho*I before resolving digested fragments by gel electrophoresis (B). DNA size is shown in kb. Clones containing inserts with correct DNA sequence were tested for soluble recombinant protein expression. A band of the correct size (38kDa) is visible by Coomassie blue stain (C.) in the pellet (P) fraction only, however from identical gel probed with anti-His antibody (D.), it is evident there is a proportion of protein present in the soluble fraction (S). Molecular weight is shown in kDa. Soluble protein purification conditions were optimised from a 100 ml culture testing elution from a nickel-agarose column using 30, 50, 100 and 250 mM concentrations of imidazole. Eluants were resolved by SDS-PAGE and stained with Coomassie blue (E) or probed with an anti-His antibody (F.). Column flow-through (F) is shown in the left-most lane. PF14_0045 did not wash off the column until eluted with 250mM imidazole and not all of the protein was eluted. Protein expression is relatively low compared to the other recombinant proteins attempted in this thesis. However when production was increased to 2 l, no soluble protein was detectable by Coomassie stain. Instead, the insoluble protein fraction from a 2 l culture was bound to nickel-agarose and washed with 30 mM imidazole buffer before eluting in 10 ml of 250 mM imidazole containing buffer. Samples were taken and resolved by SDS-PAGE (G). The eluted protein sample was free of contaminating proteins and of sufficient quantity for antibody production.


A.
kb
0.6
0.5
-
+
B.
6
4
0.7
0.5
C1
C2
C.
kDa
37
D.
37
S
P
S
P
S
P
uninduced
C1
C2
E.
kDa
Test purification
75
37
25
15
rotamase
F.
75
37
25
15
rotamase
F
10
30
50
250
Concentration of Imidazole (mM)
G.
kDa
Insoluble purification
50
37
25
15
F
30
250
Concentration of Imidazole (mM)

4.3.5 PF14_0325

A 22kDa fragment from residue 21-210, immediately following the signal peptide of PF14_0325 was selected for expression (figure 4.1). The region was amplified from *P. falciparum* 3D7 cDNA from schizont stage parasites 39 h p.i. Despite annealing to the expression vectors pET30- and pET32-Xa/LIC vectors and expression at 37°C, 22°C, 18°C for 2, 4, 8 hours and O/N in various *E.coli* expression cell types, there was very little expression only detectable in the pellet fraction by western blot after long (30 minute) exposure. Altering the region expressed to a 27 kDa fragment between amino acids 44-277 had no effect on the expression levels. Due to this lack of success, the characterisation of PF14_0325 was not taken any further.

Table 4.1 *Summary of recombinant protein expression*.

	MAL13P1.94	PF11_0443 CT	PF14_0443 M	PF14_0045	PF02_0040
PCR insert size (bp)	149	284	233	593	366
Expression vector	pET30	pET30	pET30	pET32	pET32
MW of recomb. protein (kDa)	5	11	8	22	14
MW plus tag (kDa)	10	17	13	38	30
Expression cells	BL21(DE3) pLysS	BL21(DE3) pLysS	Tf2	BL21(DE3)	BL21(DE3)
Expression conditions	22°C, 4 h	22°C, 4 h	18°C, O/N	18°C, O/N	18°C, O/N
Soluble?	Yes	Yes	Yes	No	No
Purification conditions (Imidazole conc.)	Wash 50mM elute 250mM	Wash 50mM elute 250mM	Wash 50mM elute 250mM	Wash 50mM elute 250mM	Wash 50mM elute 250mM
Final conc. from culture (mg/l)	5	10	1	N/A	N/A

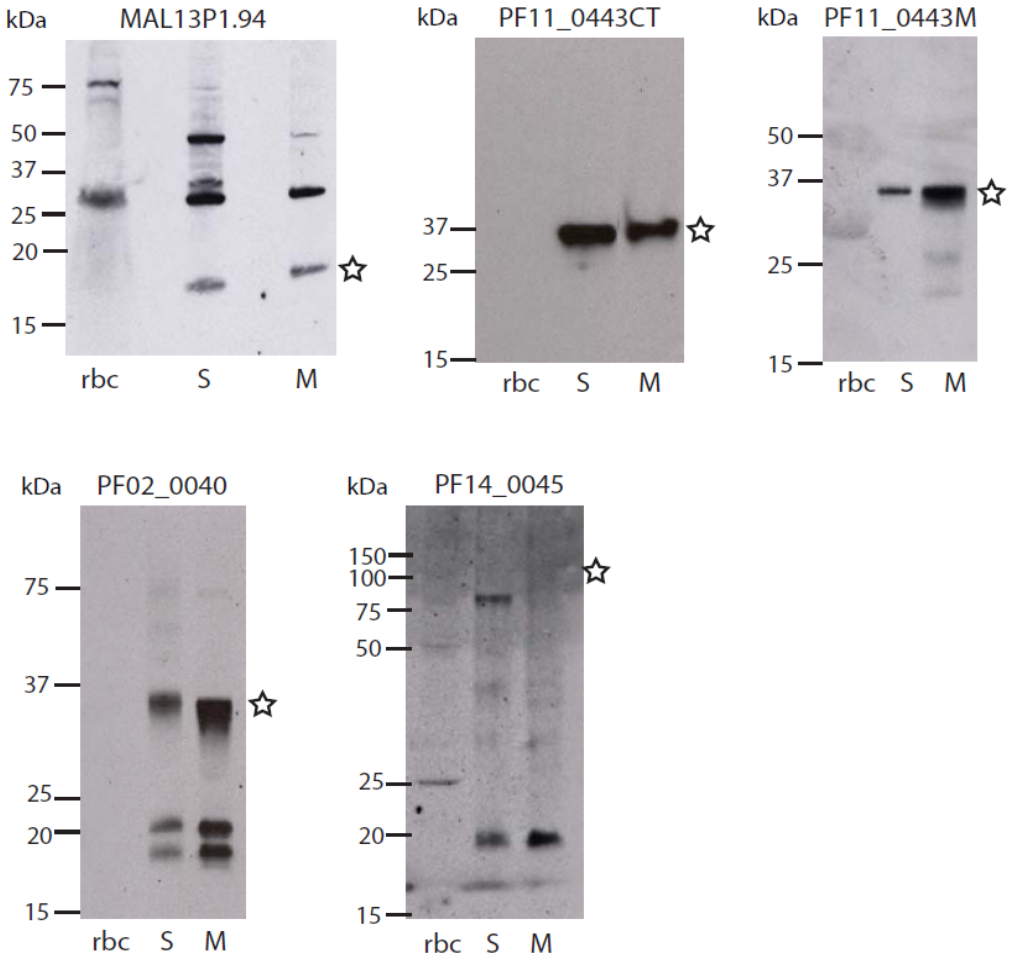


Figure 4.7. *Western blots on parasite material using purified antibodies against all 4 proteins of interest.* Antibodies raised against the aforementioned recombinant proteins were affinity purified and tested for reactivity against rbc, schizont and merozoite protein preparations resolved by SDS-PAGE under reducing conditions by Western blot. The approximate expected molecular weights are highlighted with a star. MAL13P1.94 antibodies recognise a ~27 kDa band in both schizonts and merozoites with additional 50 kDa band in schizonts and 15kDa band in merozoites. Antibodies raised against the C terminus of PF11_0443 identify a single ~37 kDa band in both schizont and merozoite preparations, as do PF11_0443M antibodies. PF02_0040 antibodies recognise a band of expected size, ~33kDa, in schizonts and merozoites, however 2 additional bands at ca. 20 kDa are also apparent. PF14_0045 antibodies recognised a faint band just above the 75 kDa marker, not the predicted 114kDa, but the more predominant band was of around 20 kDa. N.B. Although loading of similar amounts of proteins was attempted the protein preparations were not of equal concentrations, hence different intensities between schizont and merozoite bands can be detected; the purpose of these westerns was merely to test antibody reactivity/specificity not quantification. Detailed, controlled westerns are presented in separate protein chapters 5, 6 and 7.

4.4 Discussion

Protein expression was attempted for all of the 5 selected proteins with varying levels of success. Although soluble protein fragments are desirable, in cases where it was not possible to produce sufficient quantities of soluble protein, the decision was taken to use insoluble protein for antibody production instead. Although not folded, protein aggregates can also be immunogenic and immunogenicity cannot be guaranteed, irrespective of immunogen solubility.

Antibodies to MAL13P1.94 detected bands of 50, 25 and 17 kDa bands in schizonts and merozoite lysates. The 25 kDa band appears to be rbc-specific whereas the 50 and 17 kDa proteins are parasite-specific. Further work is required to address their identity. Antibody cross-reactivity was always a potential problem and this was limited at the sequence level utilising blastp to ensure the sequence was not similar to any other *Plasmodium* or human proteins. Polyclonal antibodies however have the potential to recognise both linear but also conformational epitopes and therefore cross-reactivity cannot be completely prevented by design.

Recombinant protein used to immunise rabbits contained the selected region from the protein of interested as well as the tag used in purification. Protease cleavage using factor Xa to separate target protein from the affinity-tag was not attempted as His tags are not particularly immunogenic and previous antibodies raised against fusion proteins using the same vectors resulted in specific antibodies without complications. The fact that PF11_0443 antibodies recognise a single band of approximately 37 kDa which is close to its estimated molecular weight of 39 kDa and the absence of other cross-contaminating bands suggest that the affinity tags do not generate cross reacting antibodies by themselves. Hence it is unlikely that additional bands seen by western blot probed with MAL13P1.94 antibodies are caused by histidine reactive antibodies. This same argument is valid for the contaminating *E. coli* protein rotamase which was present in both Pf11_0443 and MAL13P1.94 recombinant protein purifications; rotamase does not appear to produce antibodies reactive with any *P. falciparum* protein.

Because antibodies raised against both PF02_0040 and PF14_0045 both exhibited reactivity to an unspecific 20 kDa, it is likely that this reactivity is caused by antibodies generated against the affinity tag. Unlike the previous fusion proteins both of these proteins were expressed using the pET32-Xa/LIC vector which contains not only a His-tag but an additional thioredoxin tag to enhance solubility of the fusion protein. Thioredoxins are disulphide oxidoreductases and are present in all kingdoms of life including mammals and apicomplexa so it is feasible that this is what the rabbit sera is recognise by western blot (Holmgren, 1985). In an attempt to deplete these thioredoxin-specific antibodies, a pET32-XaLIC vector was obtained in which a stop codon was introduced by site directed mutagenesis immediately following the thioredoxin tag [kind gift of Dr Judith Green]. This protein was highly soluble and 10 mg were used on an affinity column to negatively deplete the PF02_0040 antisera of thioredoxin-specific antibodies. Column depletion was repeated 3 times over fresh columns however this was insufficient and the antibodies remained. Despite the presence of the ~20kDa bands, PF02_0040 antibodies identify a specific band at 33 kDa and are effective at a titre of 1:20000 on Western blot.

After expressing recombinant proteins and generating antisera against 4 of the 5 selected proteins a decision was made to concentrate on the most promising candidates based on resources generated. Therefore the PF14_0325 and PF14_0045 projects were terminated. No recombinant protein based on PF14_0325 amino acid sequence could be expressed in E. coli and antibodies generated against recombinant PF14_0045 failed to recognise a band of the expected size.

On the other hand, good reagents to three of the proteins have been successfully generated and therefore we are focussing on a detailed characterisation of MAL13P1.94, PF11_0443 and PF02_0040 and their role in the blood stage life cycle of *P. falciparum*. Each shall be described in detail and results discussed in chapters 5, 6 and 7 respectively.

5. Localisation and functional characterisation of the *Plasmodium* type IV Hsp40, PF11_0443

5.1 Introduction

5.1.1 The J protein/Hsp40 family

J proteins or Hsp40s are a family of co-chaperones which act in conjunction with Hsp70s. Multiple J proteins can function with a single Hsp70, either by binding and targeting protein substrates to Hsp70 or targeting Hsp70 activity directly to protein substrates within specific cellular locations (Liberek et al., 1991; Laufen et al., 1999). The bacterial chaperones DnaJ/ DnaK are the prototypical Hsp40/Hsp70 partnership and members of these families exist in every form of life. The pair classically function in binding short, unfolded hydrophobic sequences and promote correct protein folding during heat shock but family members are also involved in protein complex assembly, disassembly or protein translocation across membranes (Walsh et al., 2004). Binding of an Hsp40+substrate complex stabilises Hsp70 in a substrate-bound state via interaction of the Hsp40's J domain with the ATPase domain of the Hsp70, thereby stimulating ATP hydrolysis (Greene et al., 1998; Laufen et al., 1999). This highly conserved N-terminal J domain contains 4 helices: helix I and IV are shorter and are thought to stabilise the close interaction between anti-parallel helices II and III which form a tightly packed 'finger' that interacts with its Hsp70 binding groove (Pellecchia et al., 1996). The region between these two helices contains a protruding tripeptide HPD motif which is essential for Hsp70 binding and activation of ATP hydrolysis of Hsp70 (Tsai and Douglas, 1996), and mutation of these residues leads to a loss of function in the yeast J protein Sec63 (Feldheim et al., 1992). The J domain is one of four domains that determine the structure and function of the Hsp40. The second is a Glycine/Phenylalanine (GF) rich region which regulates substrate binding to Hsp70 (Pellecchia et al., 1996; Mayer et al., 1999). A third characteristic domain is a cysteine rich zinc finger domain which folds upon binding of 2 zinc ions and is thought to stabilise the structure of Hsp40 or the interaction with

substrate (Banecki et al., 1996; Szabo et al., 1996). The C-terminal domain is less conserved and is generally involved in substrate specificity (Banecki et al., 1996). Hsp40s are classified into 3 subgroups in most organisms depending on the presence/absence of certain features (figure 5.1). Type I Hsp40s possess all the features described above; type II Hsp40s lack the zinc finger domain and type III proteins lack both zinc finger and GF region, containing only the J domain and the divergent C-terminal domain (Cheetham and Caplan, 1998). The ability of type III Hsp40s to bind to unfolded protein has never been shown and these proteins are thought to posses more diverse roles such as recruitment of an Hsp70 to a particular site. This is exemplified by the ER J protein Sec63 which recruits BiP to the ER protein translocon (Corsi and Schekman, 1997). A review of Hsp40s in *P. falciparum* by Botha *et al* (2007), has described a fourth class of J protein. These type IV Hsp40s contain an incomplete J domain, with 4 predicted helices but lack the HPD motif suggesting a more diverse role [reviewed in (Botha et al., 2007)].

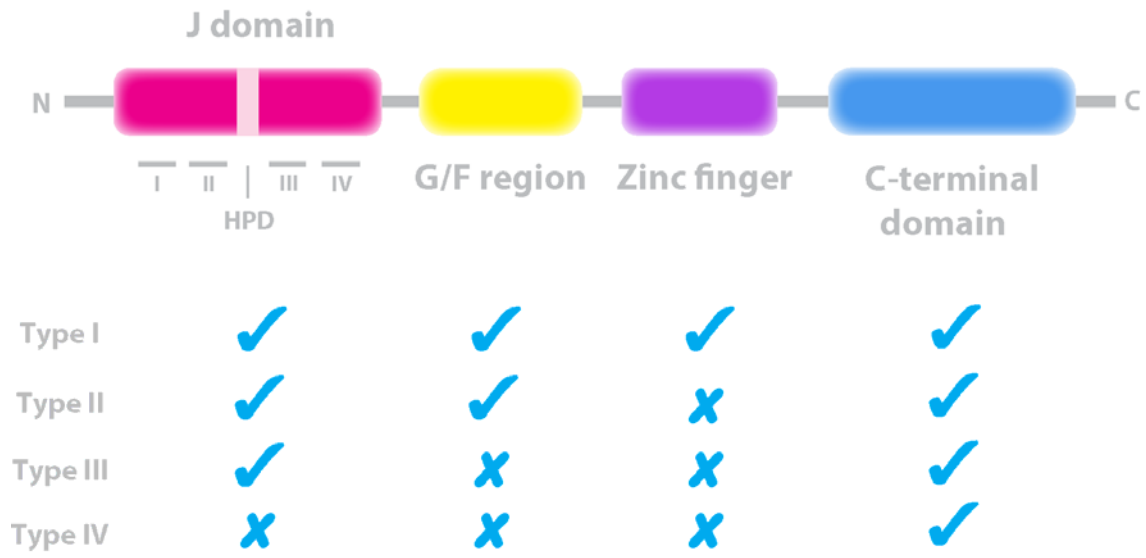


Figure 5.1: *Prototypical J protein structure.* J proteins are comprised of 4 domains. The conserved N-terminal J domain contains 4 helices with a crucial HPD motif between turns II and III. The Glycine/Phenylalanine (GF) rich region regulates substrate binding to Hsp70 and the cysteine rich zinc finger is capable of binding 2 zinc ions and is thought to stabilise the structure of Hsp40 or the interaction with substrate. The less conserved C-terminal domain is involved in substrate specificity. Hsp40s can be classified into 4 types all of which contain the less conserved C-terminal domain: type I possess all the features described above; type II lack the zinc finger domain; type III lacks both zinc finger and GF region, containing only the J domain; and type IV Hsp40s contain the J domain but lack the HPD motif suggesting a more diverse role.

5.1.2 J proteins in P. falciparum

Compared to the number of Hsp70s (of which there are 6) in *P. falciparum,* the family of Hsp40s is over-represented. According to Botha et al (2007), there are 43 predicted within the *P. falciparum* 3D7 genome, with only 2 classical type I, 9 type II which are thought to be capable of stimulating ATP hydrolysis by Hsp70, 20 type III and 12 type IV putative J proteins (Botha et al., 2007). Nineteen of these Hsp40s, from types II, III and IV, are predicted to contain *Plasmodium* export element (PEXEL) motifs which enable protein export into the erythrocyte cytosol (Marti et al., 2004). Of the 19 PEXEL-

containing Hsp40s, eleven are classified as type IV leaving only PF11_0443 as the only type IV J protein not to contain a PEXEL motif. None of the Hsp70s of *P. falciparum* contain PEXEL motifs therefore the PEXEL-containing Hsp40s are likely to have a different role (unless they are able to interact with host derived Hsp70s). Ring-infected surface antigen (RESA) is one of the 11 exported type IV J proteins. RESA is found in the dense granules and exported into the rbc cytoplasm where it has been shown to bind cytoskeletal protein spectrin (Foley et al., 1991), and is implicated in protection against thermally induced unfolding of this protein (Da Silva et al., 1994). These data are supported by knock-out studies where deletion of RESA resulted in susceptibility to heat shock (Silva et al., 2005), thus presenting clear evidence that this divergent class of type IV Hsp40s may still function in response to heat shock, despite a lack of exported Hsp70 co-chaperones.

5.1.3 PF11_0443 orthologues

PF11_0443 has orthologues in 36 different sequenced species all from higher eukaryotes (listed in Chapter 3, Table 3.1), with two exceptions – *Dictyostelium discoideum* and *Tetrahymena thermophila*. These were identified from PlasmoDB and OrthoMCL-DB, from the orthologue group OG5_131379 (Li et al., 2003) and confirmed using NCBI BLAST tool – a select alignment is present in Chapter 3, figure 3.5. Curiously, *Plasmodium* is the only Apicomplexan to contain a member of this orthologue group and it is only expressed in blood stages. This may imply a role for PF11_0443 in interaction between *Plasmodium* and it's specific host cell. Fellow Haemosporidia *Thieleria* and *Babesia* do not appear to possess an orthologue however one major difference between *Plasmodium* and other Haemosporidia is that *Plasmodium* requires an intact PV to be present throughout the lifecycle, *Thieleria* and *Babesia* do not. This would suggest a role for the protein in PV formation/stabilisation however this speculation requires experimental validation.

5.1.3.1 Clues from other species

As mention above, PF11_0443 has many orthologues in other species (Table 3.1, figure 3.5). One of these is Dmel\CG7872 from the fruit fly *Drosophila melanogaster* – one the most extensively studied multicellular organisms. The online database Flybase (flybase.org) was used to obtain information regarding the *Drosophila* orthologue of PF11_0443, Dmel\CG7872. Transcription data reveal that this protein is strongly expressed in the larval stage salivary glands at around ten-fold compared to other organs (Chintapalli et al., 2007). Salivary glands in the larva stage of fly development are responsible for the large scale secretion of a glue-like substance that allows the larva to adhere to a single position while developing into an adult fly. This suggests a role for this protein in the secretory pathway. Although the salivary glands as a whole are not essential in the laboratory environment, it is reasonable to assume that larva lacking these organs would be at a selective disadvantage in the field.

Caenorhabditis elegans is another widely studied organism due to its genetic amenability. PF11_0443 has two orthologues in this species one encoding a ca.40 kDa protein DNJ-2 and the other encodes a truncated form which will not be discussed here. Information regarding *dnj-2* (WBGene00001020) was retrieved from wormbase.org. *C. elegans* is used as a model to study the human neuro-degenerative disorder, Parkinson's Disease (PD). In one such study, the targeted knockdown of *dnj-2* resulted in mislocalisation and aggregation of the transgene alpha-synuclein suggesting DNJ-2 has a role in protein targeting and inhibiting protein aggregation (Hamamichi et al., 2008).

The human orthologue of PF11_0443 is classified as DNAJC25, a type lll J protein. Computational analysis identified high expression levels (determined by relative number of expressed sequence tags or ESTs per tissue) of DNAJ25 in the bladder, pancreas and salivary glands (Hageman and Kampinga, 2009) all of which are organs specialise in secretion.

Data from other species offer a potential insight into the function of PF11_0443, however since there are few orthologues of this protein in more closely related

organisms, it is likely that *Plasmodium* has retained or acquired PF11_0443 for a specific function which may differ to the function of the protein in metazoa. Thorough experimental analysis is required in order to characterise and uncover the function of PF11_0443 within *Plasmodium*.

5.2 PF11_0443 is an integral membrane protein in *P. falciparum* schizonts

PF11_0443 contains multiple predicted transmembrane domains. It is therefore highly likely that PF11_0443 is an integral protein, nevertheless experimental confirmation is required. *P. falciparum* 3D7 schizonts were first lysed in high pH 0.1M sodium carbonate buffer in order to separate soluble and any peripheral membrane proteins from those with membrane anchors. Western blot analysis of the protein extracts revealed that, as predicted, PF11_0443 resides within the carbonate pellet fraction, confirming that it is indeed an integral membrane protein. SERA-5 is a soluble protein within the PV of schizonts and peripherally associated with the cell surface in merozoites. Antibodies to SERA-5 and membrane anchored MSP-2 were used to control for loading carbonate supernatant and the pellet fractions, respectively.

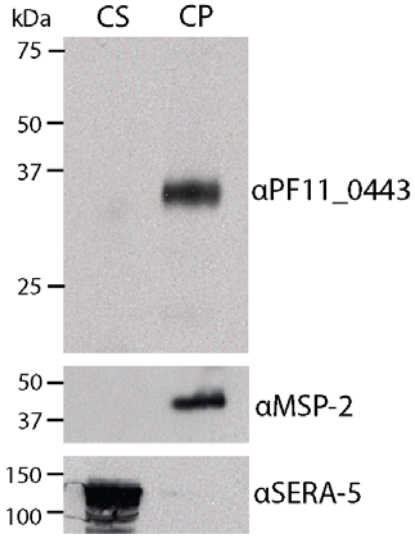


Figure 5.2: *Western blot analysis of carbonate lysed schizonts. P. falciparum* 3D7 schizonts were treated with a high pH carbonate buffer to separate soluble and peripheral membrane proteins from the integral membrane protein fraction. Samples from carbonate supernatant and carbonate pellet were separate by SDS-PAGE and subject to western blot. SERA-5 (peripheral membrane) and MSP-2 (membrane-anchored) antibodies were used to detect these proteins as loading controls. The blot was probed with PF11_0443CT antibodies to reveal this protein as integral to the membrane.

5.3 PF11_0443 is expressed from 36 h p.i. in schizonts and merozoites but is transferred into ring stages.

Bioinformatic analysis revealed PF11_0443 peak transcript levels occur at around 33h p.i. but this is not necessarily indicative of the commencement of protein expression. Proteomic data place PF11_0443 in schizonts and merozoites (as discussed in chapter 3.2.3) however there are no proteomic data available that provide evidence for this protein in ring stage parasites. In order to uncover the onset and length of PF11_0443 protein expression in the *P. falciparum* blood stage cycle it was necessary to complete a time-course experiment, sampling one complete cycle at 4 h intervals, beginning with newly invaded rings (time point zero) through to late/rupturing schizonts. Samples were subject to western blot analysis, along with a rbc control, probing with specific PF11_0443CT rabbit polyclonal antibodies for protein detection (figure 5.2 A) (for method please refer to section 2.3.4). Equal loading of rbc lysates in each well was visualised by probing with an anti-glycophorin A/B antibody. The blot was also probed with an antibody against the ER constituent parasite protein BiP to confirm the presence of parasite material in each of the lanes from the time-course (Kumar et al., 1988; Kumar et al., 1991). MSP-2, which is expressed in schizonts (Wickham et al., 2003), was used as a late stage control. Probing time course samples with PF11_0443CT antibody by western blot revealed that protein expression began at 36 h p.i. in early schizogony and increased during the following 4 hours. Protein levels were then maintained until egress however PF11_0443 was not visible in ring stage parasites by western blot suggesting either the C-terminus was lost perhaps by shedding upon entry into the host cell or rapidly degraded within the parasite (figure 5.2 A).

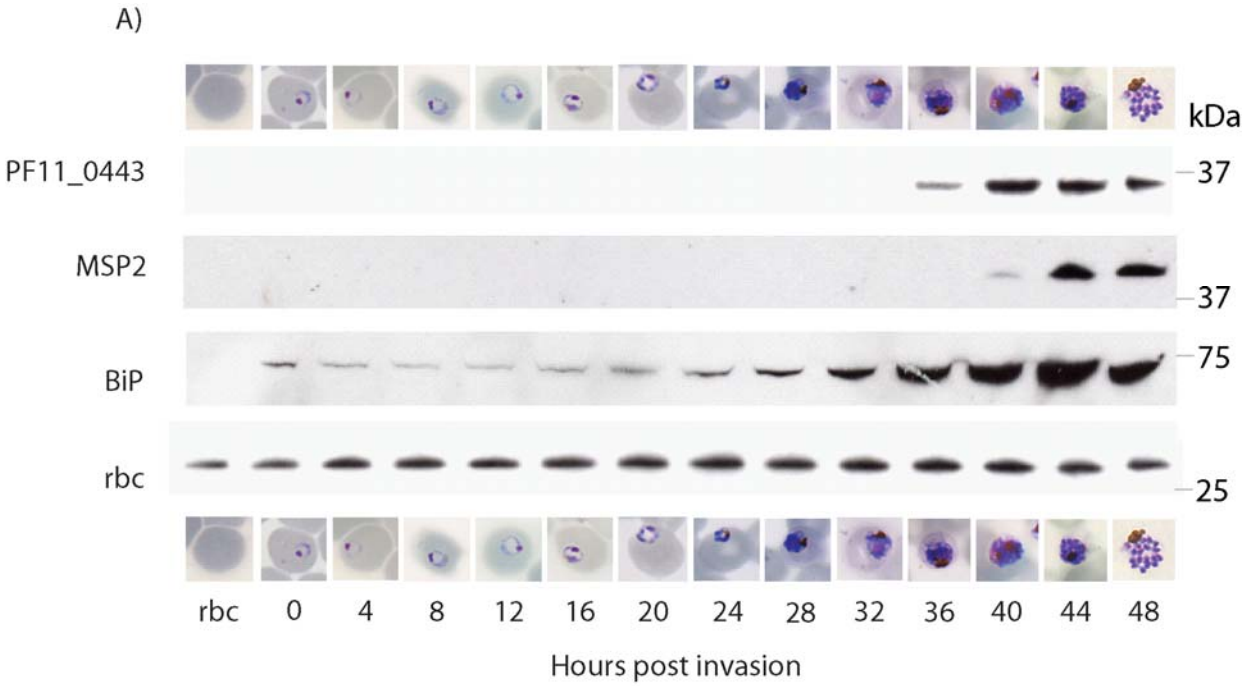
Based on the western blot results, PF11_0443CT antibodies were used in indirect immunofluorescence (IFA) experiments using parasite slides from each time point beginning at 36h p.i. (figure 5.2 B). At the beginning of protein expression, which from the western blot is around 36h and around the start of schizogony, PF11_0443 appears peripheral to the nucleus which would imply that perhaps it is to be found within the ER however distinct foci are clearly visible. As this protein contains predicted

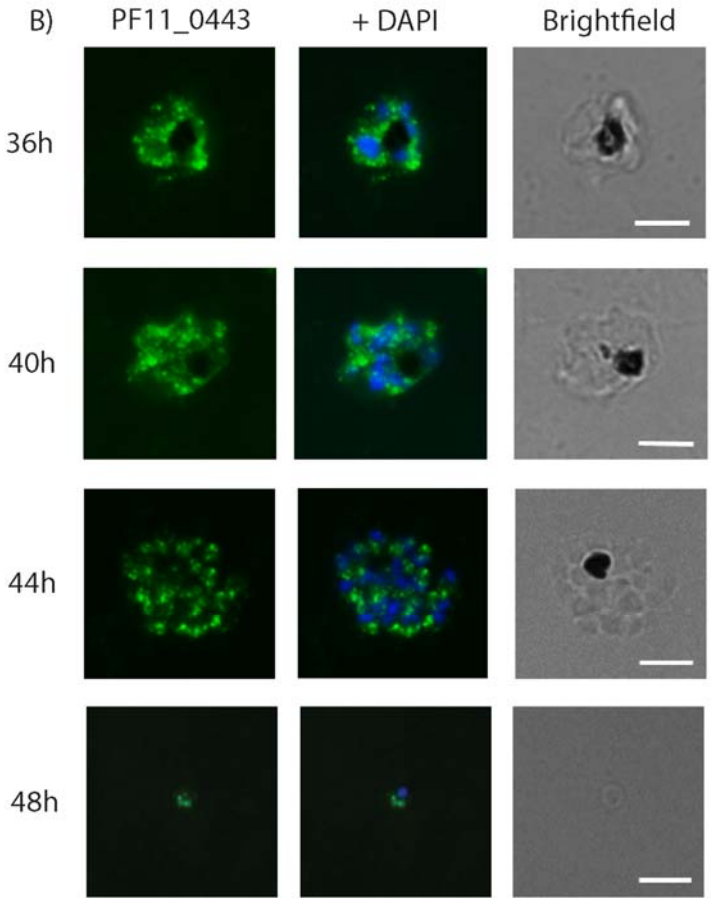
transmembrane domains these foci might be ER-derived transport vesicles or ER subdomains. This punctate staining was also detected at 40h p.i. schizonts although the staining moves from peri-nuclear to what appears to be foci throughout the schizont cytoplasm. Interestingly, as the schizont matures and undergoes cell division (44h p.i.) , PF11_0443 appears to associate with the individual merozoites located anterior to the nucleus. This pattern is similar to that exhibited by residents of the apical organelles. Co-localisation IFA studies with known organelle markers was attempted to pinpoint the intracellular location of PF11_0443.

Figure 5.3: *Expression of PF11_0443 throughout the lifecycle.*

A) Protein samples of synchronous parasite culture taken every 4 hours for one complete synchronised blood-stage cycle were hypotonically lysed and the pellet fraction was solubilised in Laemmli SDS reducing sample buffer and separated by SDS-PAGE. $2x10^6$ rbcs were loaded per well. As a control an equal number of uninfected rbcs prepared identically were loaded per well.. Expression of PF11_0443 was analysed by western blot using the previously generated PF11_0443CT rabbit polyclonal antibody. A glycophorin A/B antibody was used as a loading control (rbc). Antibodies against the parasite ER resident protein BiP were used to confirm the presence of parasite material in each lane, and against MSP-2 were used as a schizont marker. Giemsa stained parasites from each time point are presented as an example to demonstrate parasite morphology for each stage.

B) Thin smears of parasite culture from each time point were fixed in paraformaldehyde and permeabilised with Triton X-100, and used in IFAs to detect the presence of PF11_0443 in the parasite using specific antibodies. PF11_0443 is detected by a fluorescent secondary antibody (shown in green) and the nucleus is stained with DAPI (blue).A single merozoite is shown for the 48-h time point. Scale bar on the brightfield image represents 5µm.


A)
kDa
PF11_0443
37
MSP2
37
BiP
75
rbc
25
rbc 0 4 8 12 16 20 24 28 32 36 40 44 48
Hours post invasion


B)
PF11_0443
+ DAPI
Brightfield
36h
40h
44h
48h

5.4 PF11_0443 co-localises with apical organelles in late schizonts and merozoites.

A clear change was observed in the localisation of PF11_0443 from numerous small foci close to the nucleus in early schizonts to larger foci located apical in late schizonts. To determine which apical organelle PF11_0443 was located in, it was necessary to co-localise with antibodies to protein markers of known sub-cellular location. In order to do this, thin smears of *P. falciparum* 3D7 were fixed and permeabilised and subjected to indirect immunofluorescence using PF11_0443CT polyclonal rabbit antibodies. Although PF11_0443middle antibodies strongly recognised a band of the correct size in western blot based assays, they produced a weak fluorescence pattern in IFAs and were not used to identify protein location.

The fluorescence pattern was compared to that produced by antibodies against proteins of already determined location, encompassing a number of locations within the cell. PF11_0443 is present in early schizogony where it appears to reside within the ER as shown by colocalisation with BiP (Kumar et al., 1991) (figure 5.4). As schizogony progresses and daughter cells begin to segregate, the location of PF11_0443 appears to change from the periphery of the nucleus to a more punctuate pattern common among proteins of the apical organelles. Focussing on late schizonts, the protein appears to co-localise best with AMA-1 (Healer et al., 2002; Bannister et al., 2003) and EBA-175 (Sim et al., 1992), both micronemal proteins and not MSP-1 (Holder and Freeman, 1984a) (figure 5.4). There is some overlap of fluorescence between PF11_0443 and RhopH2 of the rhoptry bulb (Holder et al., 1985) in schizonts and with the rhoptry neck protein RON-4 (Alexander et al., 2006) in merozoites. Due to their close proximity within the merozoite and the fact that the diameter of fluorescence is around 200-300nm, it is often difficult to distinguish between micronemal and rhoptry locations. There also appears to be a small amount of residual PF11_0443 associated with the periphery of the nucleus, presumably the ER. In free merozoites there is at least partial colocalisation with both AMA-1 and EBA-175 (figure 6.3). Interestingly AMA-1 and EBA-175 do not co-localise with each other and are therefore thought to reside in different subsets of micronemes –

AMA-1 is released onto and diffuses across the plasma membrane upon release whereas EBA-175 only ever appears at the apex of merozoites. EBA-175 is generally seen as a single point of fluorescence in merozoites however PF11_0443 is clearly visible in multiple points of fluorescence at the apex of the cell (in comparison with the nucleus at the posterior). As there are multiple micronemes within the cell, this pattern could indicate PF11_0443 is present in each of them.

Micronemes and rhoptries are not the only organelles present at the apex of merozoites. The protein SUB-1 is present in a subset of the micronemes – the exonemes – which are released into the PV of schizont stage parasites before merozoite release. PF11_0443 does not co-localise with this subset (data not shown), and is still present in released merozoites. The dense granules are apical organelles of spherical nature which are 140 – 120 nm and therefore smaller than the two rhoptries but larger than micronemes. PF11_0443 does not co-localise with the dense granule constituent protein RESA (figure 5.5). Ruling out colocalisation with markers of these other organelles, the micronemes or rhoptries remain the most likely destination for PF11_0443.

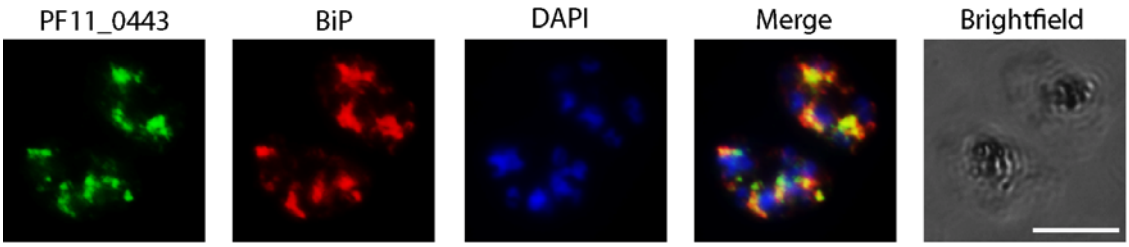
Micronemal proteins such as AMA-1 and EBA-175 are released onto the merozoite surface and shed upon invasion (Harris et al., 2005; O'Donnell et al., 2006) and consequently these proteins can be detected in culture supernatants by western blot. PF11_0443 was not detected in culture supernatants by western blot (data not shown). To address where PF11_0443 is after egress and reinvasion we analysed ring stage parasites by western blot and IFA. Although PF11_0443 was not detected in rings by western blot (figure 5.3), use of the PF11_0443CT antibody in an IFA revealed this protein can be detected in early rings by this method (figure 5.5). It is important to note that antibodies used here were used at a more concentrated level in ring stage IFAs than in schizont or merozoite IFAs and the camera exposure times were doubled when capturing the images implying PF11_0443 is present at a lower level. This apparent reduction could explain why this protein was not visible by western blot using ring stage lysates. Presumably the protein is degraded within the newly invaded rbc when no longer needed.

Indirect immunofluorescence has identified PF11_0443 as a possible micronemal protein however carry-over of a small amount of this protein into early rings calls this localisation into question as other micronemal proteins are shed during invasion. The expression from 36 h onwards when apical organelles are formed and the presence within the ER in early schizonts together with the putative divergent Hsp40 function suggests a possible role for this protein in protein sorting or trafficking. Further investigation is required to determine the protein's function within the *P. falciparum* intra-erythrocytic cycle.

Figure 5.4: *Indirect immunofluorescence assay of early and late schizonts.*

Thin smears of early and late schizonts were fixed and permeabilised as described in 2.2.13.1 and used in IFAs to detect the presence of PF11_0443 in the parasite using C-terminal specific rabbit polyclonal antibodies. The primary anti-PF11_0443CT antibody was detected by a fluorescent secondary antibody (AlexaFluor 488 – green) and the nucleus is stained with DAPI (blue). Compartments were identified using previously characterised antibodies to proteins of known location followed by secondary antibodies labelled with AlexaFluor 594 (red) and images were merged to determine comparative location (BiP: ER; AMA-1 and EBA-175: microneme; RhopH2: Rhoptry bulb; MSP-1: parasite surface). Antibodies and dilutions are described in chapter 2, table 2.4. Areas of co-localisation appear yellow. Scale bar on the bright-field image represents 5µm.

Early Schizonts



Late Schizonts

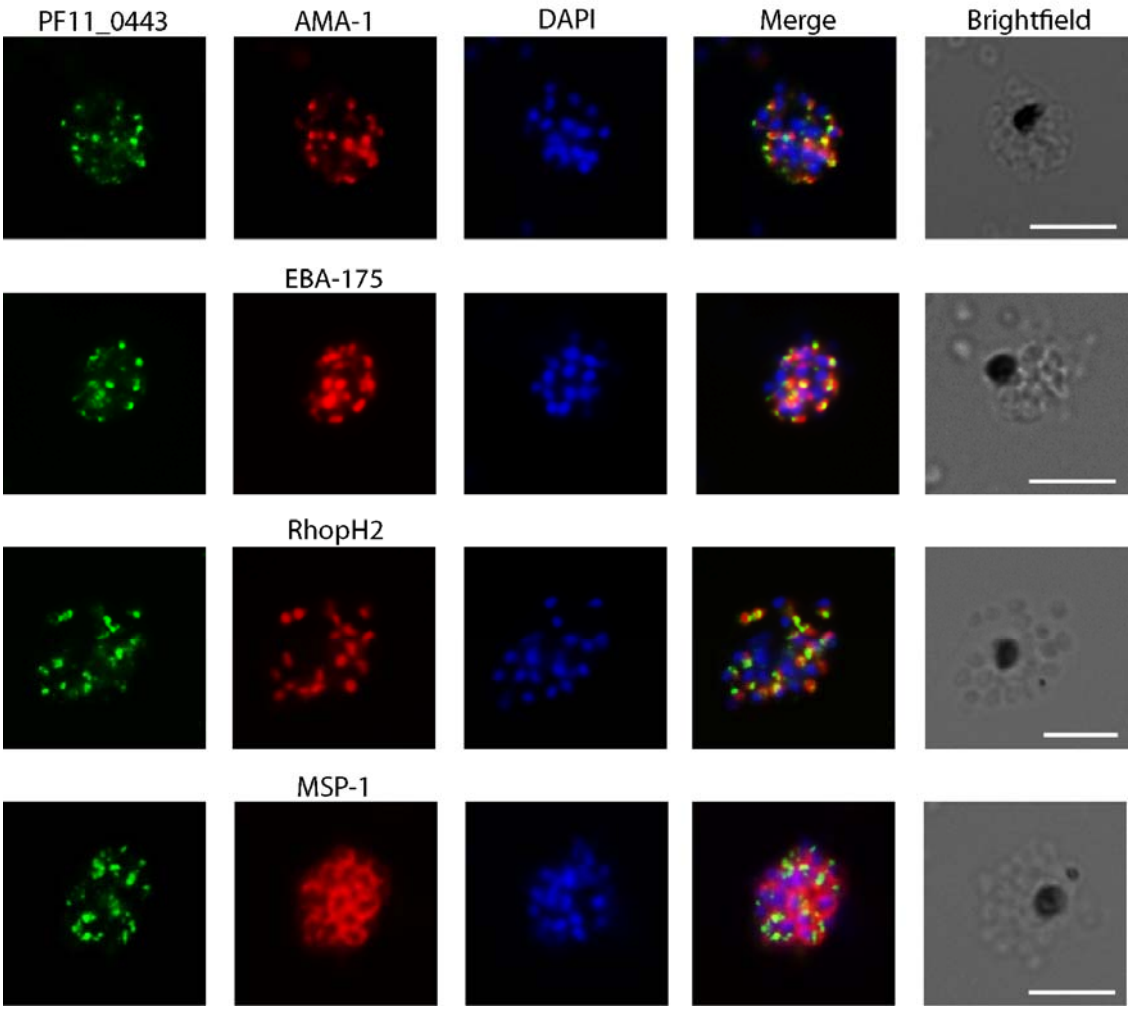
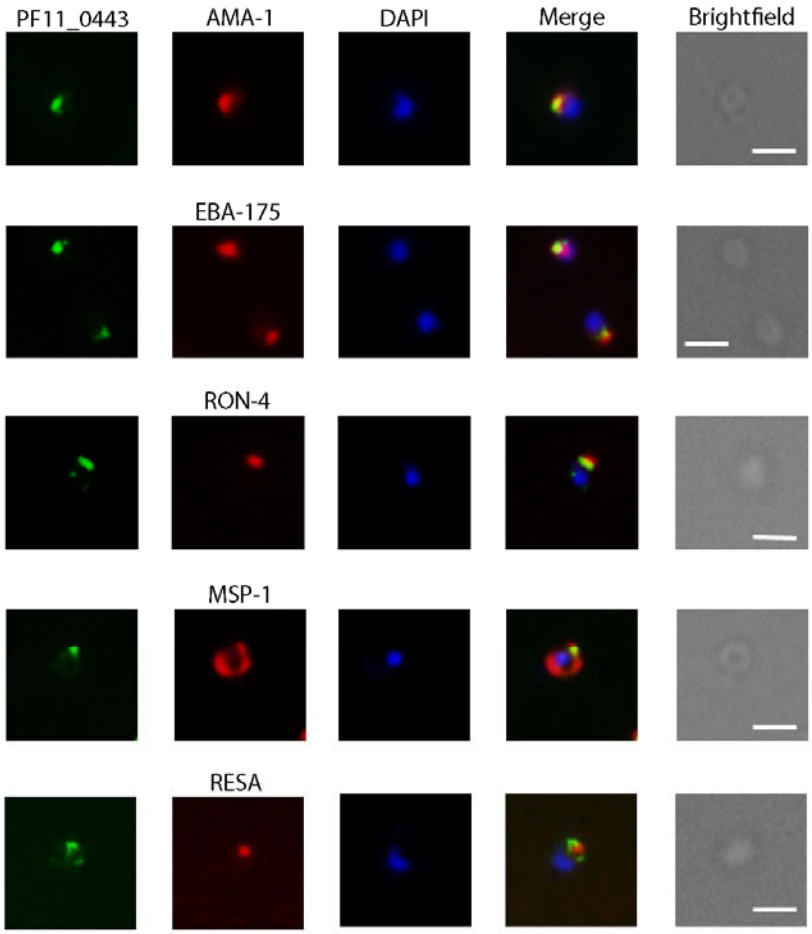


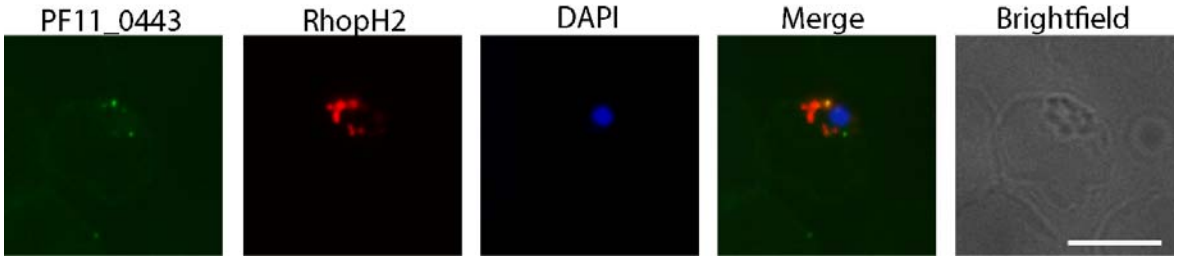
Figure 5.5: *Indirect immunofluorescence assay of merozoites and ring stages.*

Thin smears of late schizonts and merozoites as well as early rings were fixed and permeabilised as described in 2.2.13.1 and used in IFAs to detect the presence of PF11_0443 in the parasite using C-terminal specific rabbit polyclonal antibodies. Primary anti PF11_0344CT antibody was detected by a fluorescent secondary antibody (AlexaFluor 488 – green) and the nucleus is stained with DAPI (blue). Compartments were identified using previously characterised antibodies to proteins of known location followed by secondary antibodies labelled with AlexaFluor 594 (red) and images were merged to determine comparative location (AMA-1 and EBA-175: microneme; RON4: rhoptry neck; RhopH2: Rhoptry bulb; MSP-1: surface; RESA: dense granules). Antibodies and dilutions are described in table 2.4. Areas of co-localisation appear yellow. Scale bar on the bright-field image represents 5µm in ring-infected rbc and 1µm in merozoite images.

Merozoites



Rings



5.5 PF11_0443 is not released onto the surface of the merozoite prior to invasion.

Contents of the micronemes contain proteins such as AMA1 and EBA175 which are necessary for merozoite entry into rbcs. It is therefore necessary that the contents are released onto the surface of the merozoite prior to invasion to allow for contact with host cell proteins. If PF11_0443 resides in the micronemes, a part of the protein must end up on the surface of the merozoite. The protein sequence of PF11_0443 was used in several topology algorithms in an attempt to predict which sections of the protein would end up on the surface however, of the 4 algorithms, 2 predicted the C-terminus and J-domain as external and 2 predicted these domains to be cytoplasmic with the middle section of the protein on the outside of the cell. An experimental approach was therefore required.

Antibodies against both the C-terminal and the middle region of PF11_0443 were used in a live or un-fixed immunofluorescence assay of newly ruptured schizonts in the absence of fresh rbcs in order to visualise free merozoites (methods described in 2.2.13.2). Cells were incubated with primary antibodies to PF11_0443CT, PF11_0443middle, MSP-2 which served as a positive control for surface staining and MTIP – a marker of the IMC (Baum et al., 2006; Green et al., 2006b), to ensure the cells were not permeable to antibody. As visible from figure 5.6, although positive control MSP-2 showed clear surface fluorescence in the absence of any staining from MTIP labelled parasites, antibodies to both the middle and C-terminal sections of PF11_0443 failed to be detected on the parasite surface.

There are several possible explanations of this negative result. The first is that PF11_0443 is not a micronemal protein and therefore would not be expected to be released onto the surface prior to invasion, despite a convincing co-localisation by IFA. However other micronemal proteins such as EBA 175 are not observed on the surface of free merozoites in the absence of rbcs whereas AMA 1 is readily observed on the surface of merozoites (E. Knuepfer, unpublished data). This provides further evidence for a sub-division of the

micronemes both in content and in discharge conditions. Also, as mentioned in the previous section, PF11_0443middle antibody has lower affinity than the C-terminal-specific rabbit polyclonal antibody. To compensate it was used at a higher concentration but perhaps this weaker affinity was responsible for the lack of detection on the merozoite surface. A final possibility is that both the C-terminus and the middle section are on the same side of the membrane and that the second transmembrane domain of slightly weaker prediction is actually not a transmembrane spanning region after all. Looking at the sequence of this region (Chapter 3, figures 3.4 and 3.5), there are a couple of hydrophilic and charged lysine residues as well as the helix-breaker proline in the middle of the region which would interrupt a transmembrane helix. Further investigation into the validity of the TM prediction is required to rule this out as the reason for the negative result.

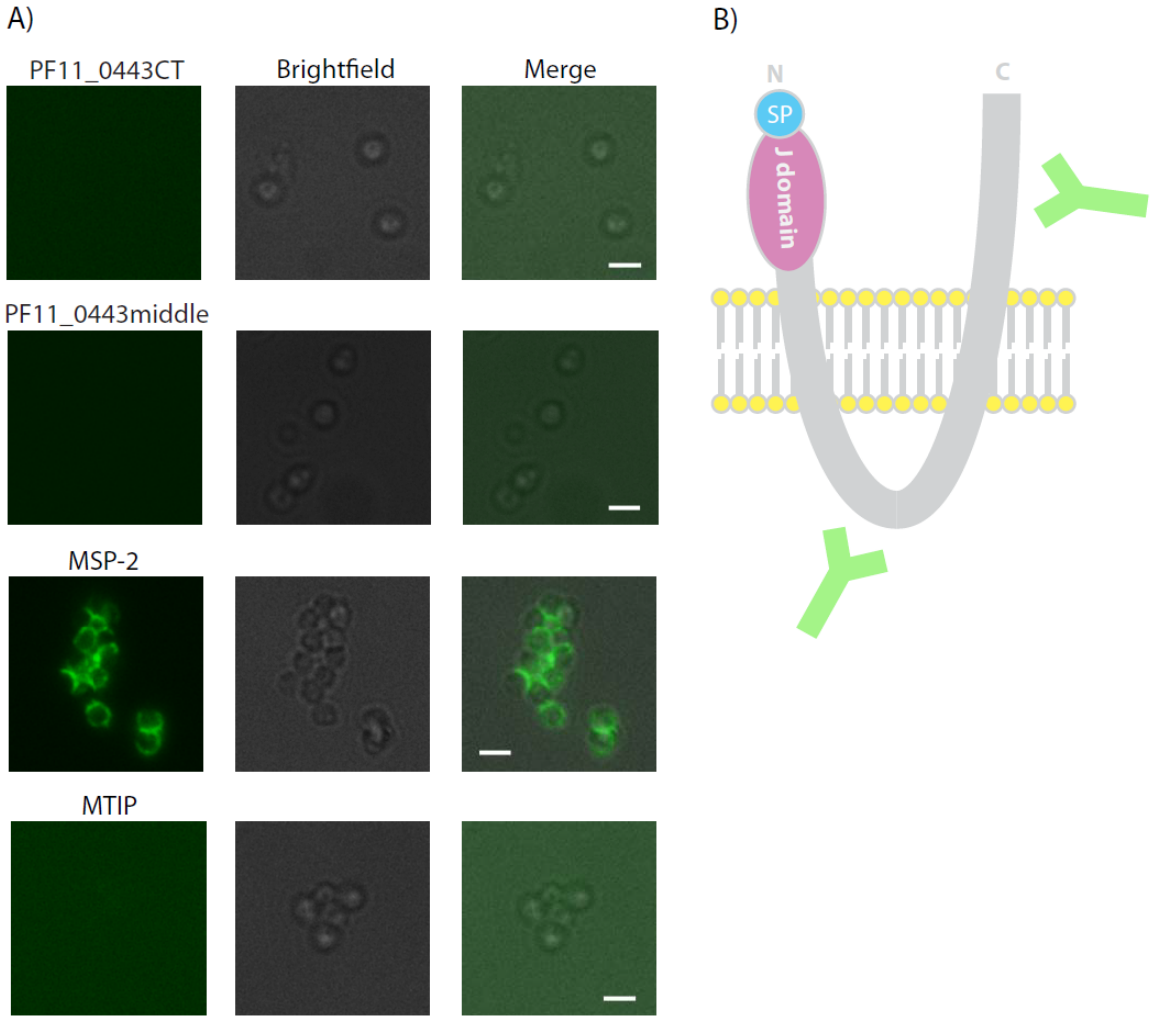


Figure 5.6: *Unfixed IFA with PF11_0443 middle and C-terminal antibodies.* A) Synchronised *P. falciparum* 3D7 schizonts were allowed to release free merozoites in the absence of rbcs and incubated with PF11_0443CT, PF11_0443middle, MSP-2 or MTIP antibodies followed by the fluorescent secondary antibody (AlexaFluor 488 – green) (methods described in 2.2.13.2). B) Schematic illustrates the predicated domain architecture for PF11_0443 and highlights regions to which antibodies were raised. This negative IFA result calls into question the validity of the second transmembrane prediction.

5.6 Immunoprecipitation experiments using protein-specific polyclonal rabbit antibody pulls down PF11_0443 as well as potential interaction partners

In the quest to determine the function of PF11_0443, the identification of interaction partners was attempted by immunoprecipitation followed by LcMS/MS analysis. Initially, PF11_0443-specific antibody was used to immuno-precipitate protein from ^{35}S-labelled schizonts (methods described in chapter 2.2.15). Resolution of this radio-labelled protein sample by SDS-PAGE and subsequent exposure to film revealed a band of ca 37 kDa which is consistent with the predicted and observed size of PF11_0443 by western blot. Other precipitated bands in the antibody IP and absent from the pre-immune IgG control IP were visible at ca 39kDa band and at ca 250kDa (figure 5.7 A). In order to separate and visualise larger bands, labelled protein samples were resolved on a lower percentage, 4-12% Tris-acetate gel (figure 5.7 B). A band of 37kDa, presumably representing PF11_0443, was visible, as well as 2 unique bands greater than 250kDa and one at approximately 130kDa. Parasites were labelled at approximately 42h p.i. for 2 h therefore only proteins being expressed at this time will incorporate ^{35}S, anything expressed later or that has completed expression for this cycle would not have been labelled therefore it is possible that additional binding partners exist. Proteins require a number of cysteine or methionine residues in order to incorporate detectable ^{35}S label, any proteins lacking in these amino acids would also not be detected here.

In order to identify possible binding partners, a protein preparation from 3ml pelleted schizonts was passed over a column of PF11_0443 specific IgG bound to activated CNBr (method described in 2.2.26). Unbound proteins were washed off with 300mM NaCl before eluting bound proteins in low pH buffer containing 100mM glycine. The elution was concentrated to 50 µl, of which 30µl was loaded onto a gel and eluted proteins were separated by SDS-PAGE. Initially the gel was stained with Instant Blue® (figure 5.7 C) however only faint bands were visible so an addition gel was prepared using the remaining 20µl of sample and silver-stained using a mass spectrometry-compatible

method (Blum, 1987) (figure 5.7 D). Visible bands were excised and analysed by Lc MS/MS analysis (PNAC, University of Cambridge). Results from mass spectrometry analysis are summarised in appendix A. Only proteins regarded as significant had a MOWSE (Molecular Weight SEarch) score of around 200 and over. The MOWSE score is based on the probability that the observed match is a random event: the lower the probability that the match is due to chance, the higher the score.

The experiment successfully precipitated PF11_0443 which, in-keeping with its apparent molecular weight observed in previous western blots, was primarily found in band 1 (figure 5.7 C), although peptides were recovered from other bands. PF11_0443 was not recovered from rings stages and here the only significant hit was the ER Hsp70 BiP (band 6, figure 5.7 D). Unfortunately, due to the role of Hsp70s in binding unfolded proteins, they are common contaminants in immunoprecipitation experiments. This is particularly problematic considering the function of type IV Hsp40 may well involve binding this family of proteins. Abundant proteins are also common contaminants and it is likely this is the reason for the presence of MSP-1 and SERA-5. Rhoptry proteins have also been found to non-specifically bind to the matrix however four rhoptry proteins have been identified here – RhopH2, RhopH3, RAP1 and RAP3 and this could be significant. In *T. gondii,* dynamin-related protein B (DrpB) has been implicated in rhoptry targeting (Breinich et al., 2009) and this experiment co-precipitated a dynamin-like protein which could perhaps function in a similar way although this has never been verified in *Plasmodium.*

Interestingly micronemal proteins AMA-1 and EBA-181 also were co-precipitated. These proteins are not common non-specific binders and could be the result of specific interaction. Other proteins of interest include Hsp60, Hsp101, the DnaJ containing proteins, IMC resident GAP50, protein phosphatases and conserved proteins of unknown function. The presence of a FKBP-type isomerase in the IP eluant is likely to be due to the presence of antibodies on the column that recognise the bacterial FKBP-type isomerase rotamase, which was present in the recombinant protein fraction used to immunise the rabbit that produced the PF11_0443CT antibodies. Interestingly two

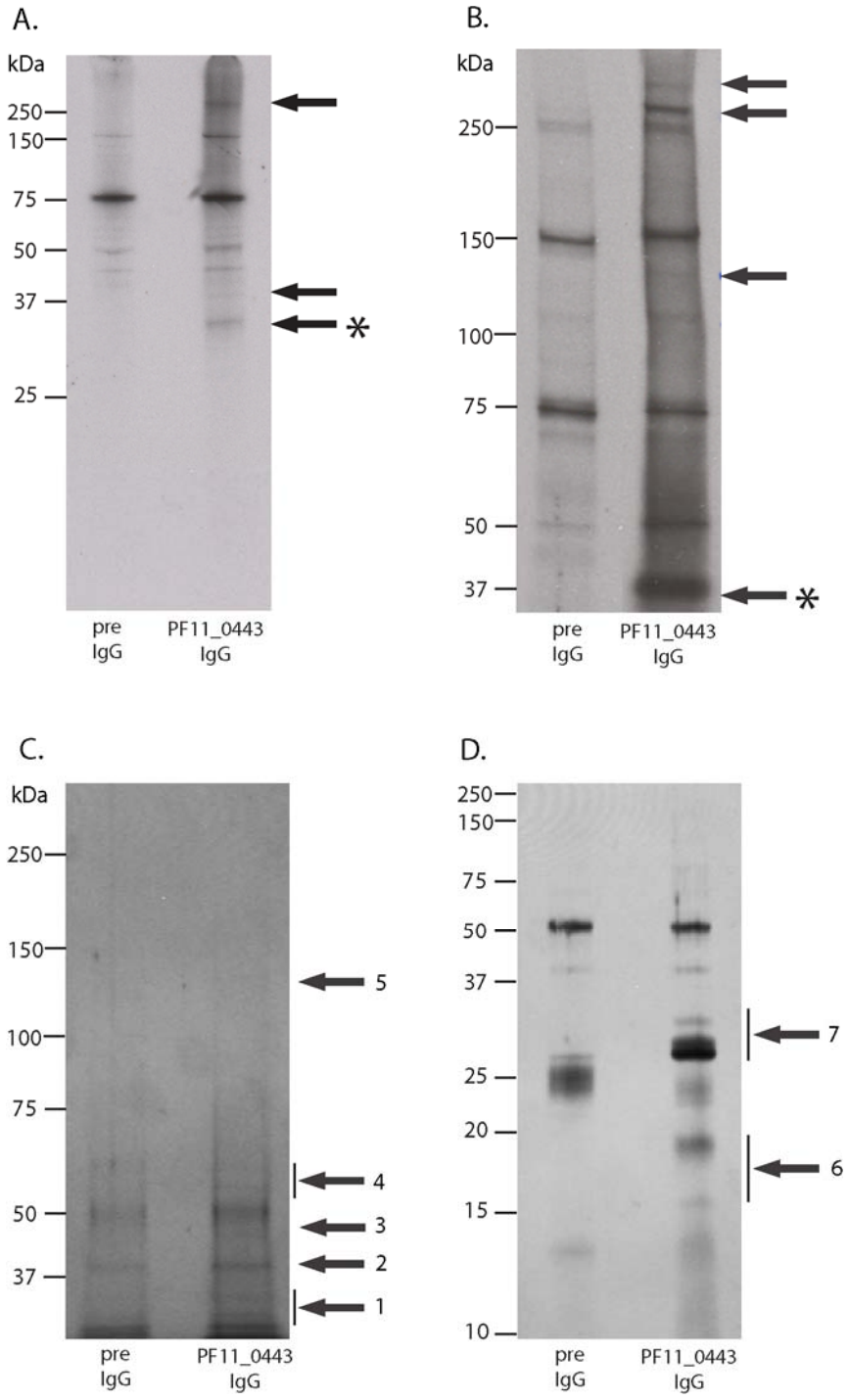
protein disulphide isomerases (PDIs) were co-precipitated. These proteins have long been established to form and preserve disulphide bridges in the ER but they are also present in the *P. berghei* ookinete microneme proteome (Lal et al., 2009), where the authors suggest it may be required for the conservation of protein structure. In *P. falciparum,* PDI exhibits punctate staining in schizonts staining which could correspond to micronemes (Mouray et al., 2007) and PDIs are also present in micronemes of *N. caninum* (Naguleswaran et al., 2005) and on the surface of *T. gondii* (Meek et al., 2002), however it has not been determined whether it is secreted through the micronemes in this particular organism. As ookinetes do not possess rhoptries, it is not clear whether this protein is present in this organelle also. The ookinete microneme proteome also uncovered a large number of chaperones, eight of which were DnaJ domain-containing proteins. Although PF11_0443 is not among them as it is not expressed at this stage and it is not clear how many of these are contaminants; this microneme proteome suggests potential role for the J protein family in these organelles.

Equivalent bands eluted from the pre-immune column were not initially analysed as it was thought necessary to first determine whether or not the experiment was successful. Although PF11_0443 was clearly identified, a vast array of other proteins appeared to co-precipitate, calling in to question the quality/strength of this method for identifying binding partners. Immunoprecipitation of an epitope tagged protein may lead to more specific results therefore the relevance of pre-immune analysis was postponed until other avenues of binding partner identification were explored.

Figure 5.7: *Immunoprecipitation studies using PF11_0443 antibodies.*

(A+B) PF02_0040 antibodies were used to immuno-precipitate interacting proteins from 50µl of pelleted ^{35}S labelled schizonts (methods described in chapter 2.2.15). Co-precipitating proteins were separated by SDS-PAGE under reducing conditions with 12% Bis-Tris (A) and a 4-8% Tris-Acetate (B) gels and exposed to film. A band of the predicted size of PF11_0443, approximately 37kDa, was observed and highlighted with * and unique precipitated bands were indicated with an arrow.

(C+D) Purified schizonts were lysed and passed over a column of 10mg of pre-immune rabbit IgG followed by that containing affinity purified anti-PF11_0443 rabbit polyclonal antibodies conjugated to activated CNBr (methods described in chapter 2.2.16). Co-precipitating proteins were resolved under reducing conditions by SDS-PAGE on 4-8% Tris Acetate (C) and 12% Bis-Tris (D) gels and visualised by Instant Blue (C) and silver stain (D) respectively. Unique bands compared to the control were excised (labelled 1-7) and the protein complement of each was identified by LC-MS/MS.


A.
kDa
250
150
75
50
37
25
*
pre IgG
PF11_0443 IgG
B.
kDa
250
150
100
75
50
37
*
pre IgG
PF11_0443 IgG
C.
kDa
250
150
100
75
50
37
5
4
3
2
1
pre IgG
PF11_0443 IgG
D.
250
150
75
50
37
25
20
15
10
7
6
pre IgG
PF11_0443 IgG

5.8 Epitope tagging of the *PF11_0443* locus by targeted homologous recombination

Incorporation of a sequence coding for an epitope tag into the genomic locus of a gene is a useful tool for characterisation of the resulting protein. Often there are commercial monoclonal and polyclonal antibodies available which have been tried and tested for reactivity and specificity for the tag antigen. Due to the problems with immunoprecipitation seen using the polyclonal antibodies generated here, incorporating an epitope tag into the *PF11_0443* locus would provide a useful tool for further characterisation of the protein.

One of the most common epitope tags is green fluorescent protein (GFP). Fluorescent tags such as this allow direct visualisation of the endogenous protein using a fluorescent microscope, without the need for antibodies. This can be used in video-microscopy to track the life history of the protein, from translation to degradation or shedding, with trafficking to cellular location in between.

One potential problem with the fluorescent proteins is that they are relatively large compared to other tags. GFP is 26kDa and may not be suitable to add-on to the C-terminus as the protein is only predicted to be 39 kDa and this would effectively double its molecular weight. For this reason, smaller tags were initially selected to incorporate into the C-terminus of PF11_0443. Since the N-terminus contains a signal peptide which is cleaved upon translocation into the ER, tagging the N-terminus is not possible. One of the chosen tags is the 3xFLAG tag. FLAG™ (Sigma-Aldrich) is an 8 aa, aspartic acid-rich tag with an enterokinase cleavage site incorporated between the C-terminus of the endogenous protein and the tag sequence. Sigma-Aldrich produces a multitude of products useful in protein characterisation such as mono- and polyclonal antibodies; antibodies directly conjugated to HRP for western blots or to fluorophores for direct immunofluorescence; and immunoprecipitation kits which include monoclonal antibody M2 conjugated to an agarose resin and 3xFLAG peptide for specific elution. The 3xFLAG tag is 21aa long and was chosen over a single copy of FLAG due to the

increased sensitivity it gives in terms of recognition by the αFLAG monoclonal M2 antibody which will detect the presence of 3xFLAG peptide at femtomolar concentrations. The second eptiope tag chosen was the TY-1 tag, a 10aa tag first described in *Trypanosomas brucei* (Bastin et al., 1996), used commonly in *Toxoplasma gondii* and also used previously in *P.falciparum* (Treeck et al., 2009).

In order to introduce these tags onto the C-terminus of the protein, a region of homology of 540 bp from the 3' end of the *PF11_0443* locus (minus the stop-codon) was amplified by PCR using primer combinations R/S and R/T (see chapter 2, table 2.1). Ty1 and FLAG epitope tag sequences (including a stop codon at the 3' end) were included in reverse primers S and T respectively and used to amplify the region of homology, thereby incorporating the tag sequences. This region of homology plus tag sequence was cloned into the multiple cloning site of the pHH3 vector via the *Eco*RI/*Sac*II restriction sites generating pHH3_PF11_0443_Ty1 and pHH3_PF11_0443_3xFLAG. An *Avr*II site was included between the gene and tag sequences to allow the use of these vectors to epitope tag other genes. This strategy is summarised in figure 5.8 A. The pHH3 transfection vector which includes the Blasticidin resistance cassette as a selectable marker is depicted in appendix D. The same region of homology of PF11_0443 was in addition cloned into the pHH3 vector via *Eco*RI/*Sac*II without the presence of an epitope tag to serve as a 3' replacement vector.

Inserts were initially sub-cloned into the pGEM T-easy vector and transformed into *E. coli* TOP10 cells. Colonies were selected and sequenced before excising the correct inserts using *Eco*RI and *Sac*II. Inserts were gel-purified and ligated into the pHH3 vector and transformed into *E. coli* TOP10. Colonies were re-sequenced and one construct with correct DNA sequence was propagated and plasmid DNA purified for transfection. For analysis purposes the construct was subject to digestion with *Eco*RI and *Sac*II to confirm the presence of the correct sized (540 bp) insert prior to transfection (figure 5.8 B)

Complete and purified plasmids pHH3_PF11_0443_3xFLAG, pHH3_PF11_0443_Ty1 and pHH3_PF11_0443_3' were transfected into ring-stage parasites and selected for

presence of episome by 2.5µg/ml of Blasticidin applied to cultures . Stock samples of these cultures were cryopreserved and genomic DNA was prepared. Transfection cultures were taken off drug for 3 weeks before re-application of drug which is required for enrichment of the parasite population that contains integrated versus episomal copies of the transfection vector. This was repeated for up to 3 cycles.

A diagnostic PCR on genomic DNA prepared from drug cycle 2 was designed to test for integration events using a forward primer (a) from the *PF11_0443* 5'UTR region (primer N, table 2.1) and reverse primers from either (b) the *PF11_0443* 3'UTR region (primer Q, table 2.1) to identify the wild type locus of 1293 bp or (c) from the *HRP2* 3'UTR contained within pHH3_PF11_0443_Ty1 construct (primer NN, table 2.1) to detect integrated plasmid of 1345 bp(figure 5.8 A). The upper bands on the diagnostic PCR gel represent correct integration of the constructs pHH3_PF11_0443_3xFLAG, pHH3_PF11_0443_Ty1 and pHH3_PF11_0443_3' into the *PF11_0443* locus by drug cycle 2, the lower band is most likely a result of mispriming (figure 5.8 C).

To test whether these integrated tags after drug cycle 2 are expressed in parasites, protein samples were prepared from 3D7 WT and tagged lines and subject to western blot analysis with either tag-specific or PF11_0443M antibodies (figure 5.8 D+E). PF11_0443M recognised a single band in WT 3D7 parasites however a clear doublet was visible in both of the tagged lines illustrating 2 distinct parasites populations with the lower band representative of WT parasites and the higher band representing epitope tagged parasites, with the tag responsible for the small increase in molecular weight (3xFLAG – 39 kDa; Ty1 - 41 kDa). When both WT 3D7 and tagged proteins were probed with Ty-1 or FLAG antibodies no band was visible in the WT line and only a single band appeared in the tagged lines. Parasites containing a Ty-1 or FLAG-tagged copy of PF11_0443 were used in IFA (figure 5.9). Antibodies to either tag revealed the epitope-tagged PF11_0443 exhibited a similar cellular location to the wild type protein, with what appears to be ER staining pattern in early schizonts and a more punctuate and apical location in late, segmented schizonts and merozoites, however colocalisation with known markers is required to confirm this.

These initial tests of FLAG and TY-1 specific antibodies by western blot and IFA revealed the αFLAG M2 monoclonal antibody to be the better reagent in terms of sensitivity and commercial availability, therefore all further experiments were performed with the 3xFLAG-tagged line. The mixed population culture was cloned by limiting dilution and clones screened by PCR using primer combination R/S and R/T (mentioned above) to ensure there was no wild type *PF11_0443* present. Although WT and integrated DNA was present in the uncloned drug cycle 3 parasite line, clones a and c contained only integrated DNA (figure 5.10 A). A Southern blot was designed to confirm integration into the correct locus (figure 5.10 B). Genomic DNA taken each drug cycle was digested by *Pac*I and *Bgl*II before separating resulting fragments by gel electrophoresis and blotting onto nitrocellulose membrane. The membrane was hybridised with a radiolabelled DNA probe encompassing the region of homology used in the construct design. The resulting banding pattern from exposure to film revealed the presence of wild type DNA with a band of 2560 bp (WT) in cycle 0 as well as the presence of both integrated (2789 bp) and episomal construct DNA (E) (5985 bp) in cycle 3. In the 3 cloned lines tested (clones a, c and e), this episomal band is lost and only the integration band at 2789 remains. There is a fourth band present which does not correspond to WT, episome or integration into the *PF11_0443*locus in both C3 and clone E which may correspond to integration of the construct into another region of the genome. Random integration of the blasticidin cassette has been observed by other lab members using this construct to epitope tag other genes.

Colocalisation of PF11_0443-FLAG with known markers of sub-cellular compartments confirmed that this protein is present in early schizonts where it co-localises with the rhoptry bulb marker RhopH2. The pattern is indicative of presence in the ER, however an ER marker antibody of the correct species was not available. In late schizonts, FLAG antibodies co-localise with RhopH2 suggesting PF11_0443-FLAG is present in rhoptries. Whereas previously in section 5.4 it was shown that WT PF11_0443 colocalised best with micronemal markers, only partial colocalisation was observed with AMA1 in merozoites. Again using the FLAG-tagged PF11_0443 no colocalisation was observed with MSP1 but

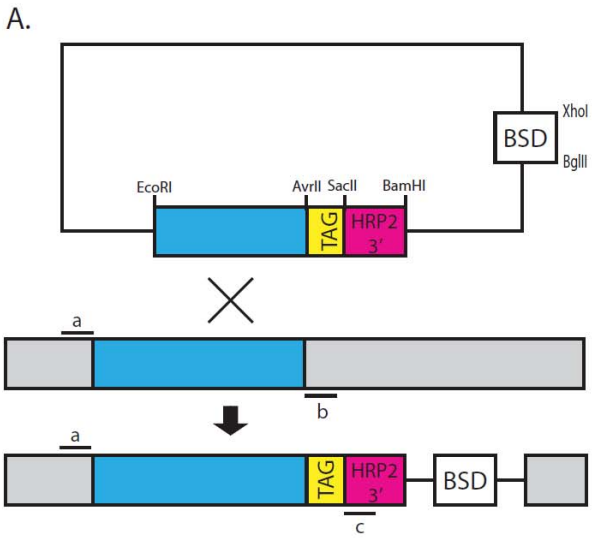
colocalisation with the rhoptry neck protein RON4 could not be tested as both antibodies are mouse derived. Colocalising PF11_0443 CT antibodies with anti FLAG antibodies shows only modest overlap.

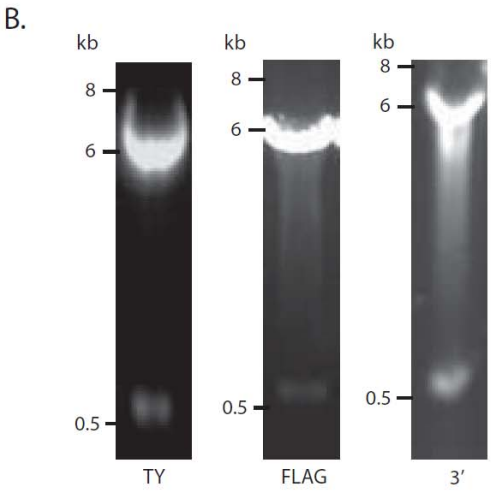
The difference in fluorescence pattern could be due to experimental difficulties, since the resolution of light microscopy may not be sensitive enough to distinguish between the different locations. It is currently unclear whether PF11_0443 in 3D7 parasites resides within rhoptries or micronemes since IFAs using WT parasites did not uncover a definitive location for this protein. It is therefore unknown whether the localisation of the FLAG-tagged PF11_0443 in rhoptries is depicting the same localisation as the wild type protein or whether C-terminal tagging resulted in the mislocalisation of this protein. Introduction of a tag to the C-terminus of PF11_0443 may have caused a change in the proteins sub-cellular location, perhaps by interfering with trafficking signals. If the latter proves to be true, perhaps the protein is not essential, as a change in location suggests the protein's function in its original location is unnecessary. It is also possible that the increased sensitivity of the anti-FLAG antibodies has led to the discovery of PF11_0443's location within the cell.

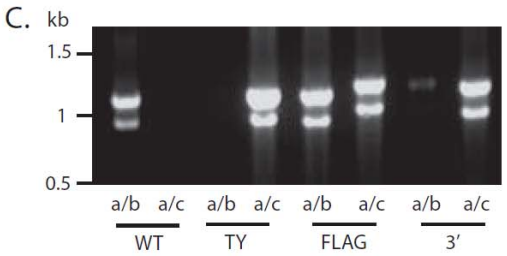
Definitive localisation studies using epifluorescense is difficult when trying to detect proteins located in closely assembled organelles. Only electronmicroscopy studies with newly produced polyclonal antibodies will clarify the discrepancies definitively.

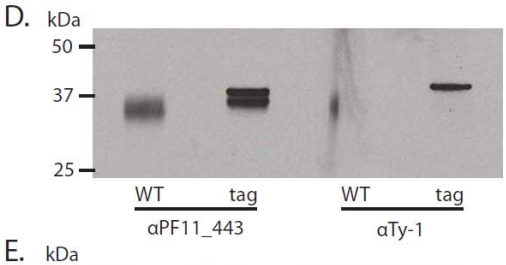
The *PF11_0443* locus has proven to be an accessible locus, with all 3 vectors integrating into the locus by drug cycle 2. These tags could be detected using Ty-1 and FLAG specific antibodies by western blot and IFA of tagged parasite lines. Due to the superior sensitivity of the FLAG antibody, this line was cloned and clone C was chosen to be used as a tool for further characterisation of PF11_0443. Incorporation of FLAG onto the C-terminus of this protein did not result in any observable invasion phenotype, with invasion rates comparable to the WT 3D7 line (data not shown).

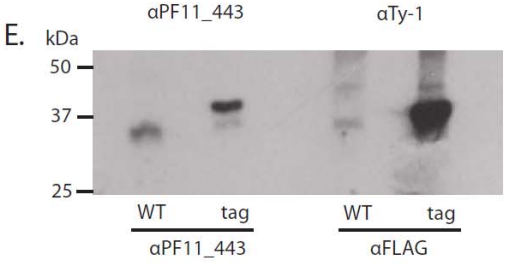
Figure 5.8: *Introduction of 3' epitope tags to the gene* PF11_0443. A 540 bp region of homology of *PF11_0443* was amplified using primers R/S and R/T (chapter 2, table 2.1) and introduced into the pHH3 transfection vector along with TY-1 and 3xFLAG epitope tag via *Eco*RI and *Sac*II restriction sites (A)The pHH3 vector contains the Blasticidin resistance cassette and a copy of the *P. falciparum hrp2* gene 3' untranslated region (UTR). Restriction digest of the complete construct (5985bp) using *Eco*RI and *Sac*II shows the 540bp insert and 5445bp vector backbone (B). The plasmid was transfected into *P. falciparum* 3D7 and drug cycled to promote loss of episomal copies of the construct. A diagnostic PCR was designed to test for integration events using a forward primer from the *PF11_0443* 5'UTR region (a: primer N, table 2.1) and reverse primers from either the *PF11_0443* 3'UTR region (b: primer Q, table 2.1) to identify the wild type (WT) locus or the *HRP2* 3'UTR region from the pHH3_PF11_0443_Ty1 or 3xFLAG constructs (c: primer NN, table 2.1) to detect integrated plasmid. PCR revealed the presence of integrated Ty-1, 3xFLAG and 3' replacement constructs into the *PF11_0443* locus (C). Protein samples from tagged lines as well as WT parasites were resolved by SDS-PAGE and subject to western blot analysis. Blots were probed with PF11_0443M rat polyclonal antibodies and either Ty1 (D) or FLAG antibodies (E). Both Ty-1 and FLAG antibodies recognise a band of the predicted 39 and 41 kDa respectively in the tagged line but not WT line.


A.
BSD
XhoI
BglII
EcoRI
AvrII
SacII
BamHI
TAG
HRP2
3′
a
b
a
TAG
HRP2
3′
BSD
c


B.
kb
8
6
0.5
TY
kb
8
6
0.5
FLAG
kb
8
6
0.5
3′


C.
kb
1.5
1
0.5
a/b a/c a/b a/c a/b a/c a/b a/c
WT
TY
FLAG
3′


D.
kDa
50
37
25
WT tag
αPF11_443
WT tag
αTy-1


E.
kDa
50
37
25
WT tag
αPF11_443
WT tag
αFLAG

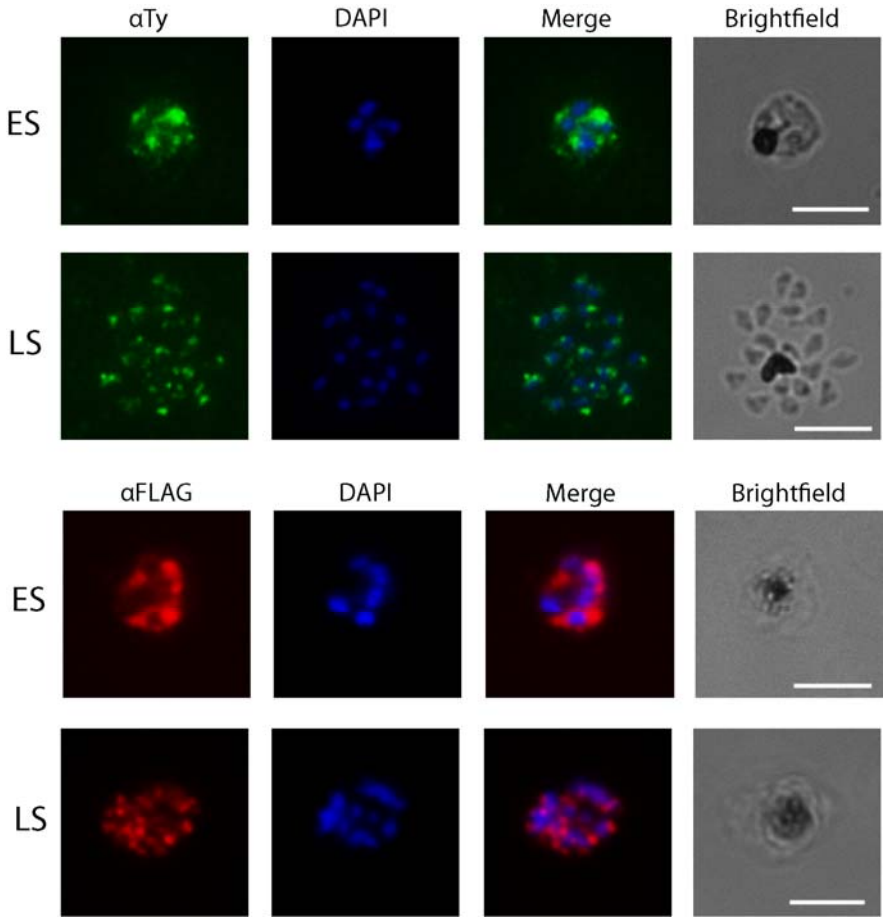


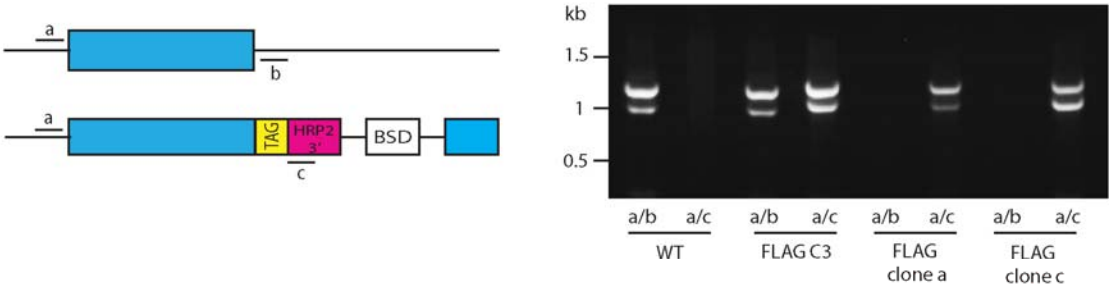
Figure 5.9: *IFAs using anti-Ty1 and FLAG antibodies.* Thin smears of synchronised parasite culture containing either early or late schizonts were fixed in paraformaldehyde and permeabilised with Triton X-100 (as described in 2. 2.13.) and used in IFAs to detect the presence of Ty-1 or FLAG tagged PF11_0443 in the parasite using specific antibodies. Tags were detected by a fluorescent secondary antibody (Ty-1: AlexaFluor 488 shown in green; FLAG: AlexaFluor 594 shown in red) and the nucleus is stained with DAPI (blue). Scale bar on the brightfield image represents 5μm.

Figure 5.10: *Cloning and Southern blot of 3xFLAG-tagged* PF11_0443.

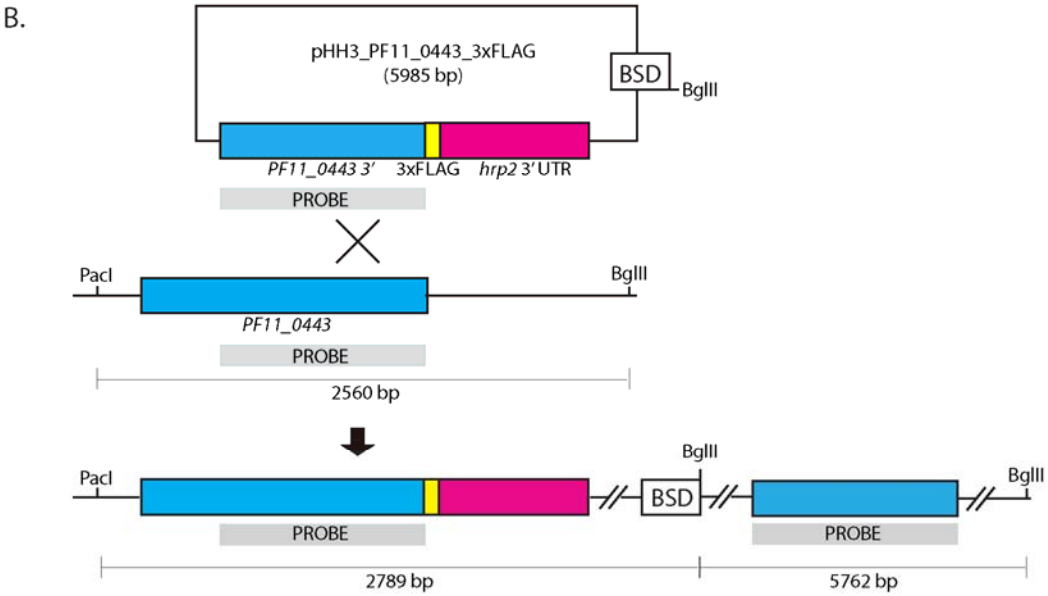
A) The integrated PF11_0443-3xFLAG line was cloned by limiting dilution (for methods see section 2.3.5.6). DNA was prepared from two of the clones and screened by PCR for presence of integrated 3xFLAG (using primers N [a]+NN [c], table 2.1) and for the absence of WT *PF11_0443* (using primers N [a]+Q [b]).

B) +C) A Southern blot was designed to test for integration of the pHH3_PF11_0443_3xFLAG construct in to the *PF11_0443* locus (B). Genomic DNA taken each drug cycle was digested by *Pac*I and *Bgl*II before separating resulting fragments by gel electrophoresis and blotting onto nitrocellulose membrane. The membrane was hybridised with a radiolabelled DNA probe identical to the region of homology used in the construct. The resulting banding pattern from exposure to film revealed the presence of wild type DNA with a band of 2560 bp (WT) in cycle 0 as well as the presence of both integrated (2789 bp) and episomal construct (E) (5985 bp) in cycle 3. In the 3 cloned lines tested (clones a, c and e), this episomal band is lost and only the integration band at 2789 remains. There is a fourth band (X) present which does not correspond to WT, episome or integration into the *PF11_0443*locus in both C3 and clone E which may correspond to integration of the construct into another region of the genome.

A.


a
b
a
TAG
HRP2 3'
BSD
c
kb
1.5
1
0.5
a/b a/c
WT
a/b a/c
FLAG C3
a/b a/c
FLAG clone a
a/b a/c
FLAG clone c

B.


pHH3_PF11_0443_3xFLAG
(5985 bp)
BSD
BglII
PF11_0443 3'
3xFLAG
hrp2 3' UTR
PROBE
PacI
BglII
PF11_0443
PROBE
2560 bp
PacI
BglII
BSD
BglII
PROBE
PROBE
2789 bp
5762 bp

C.

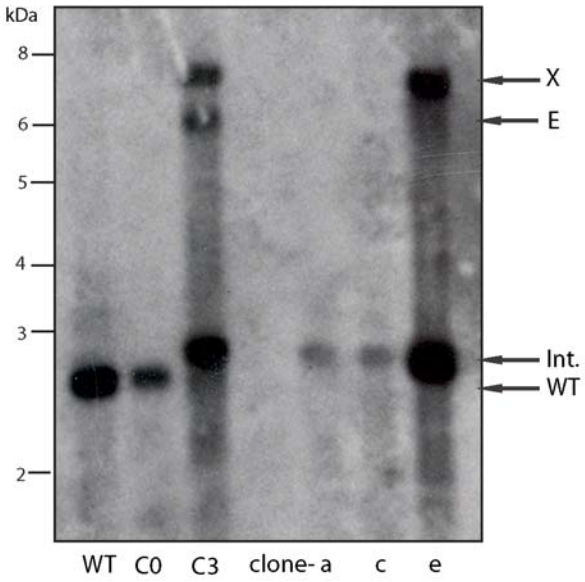
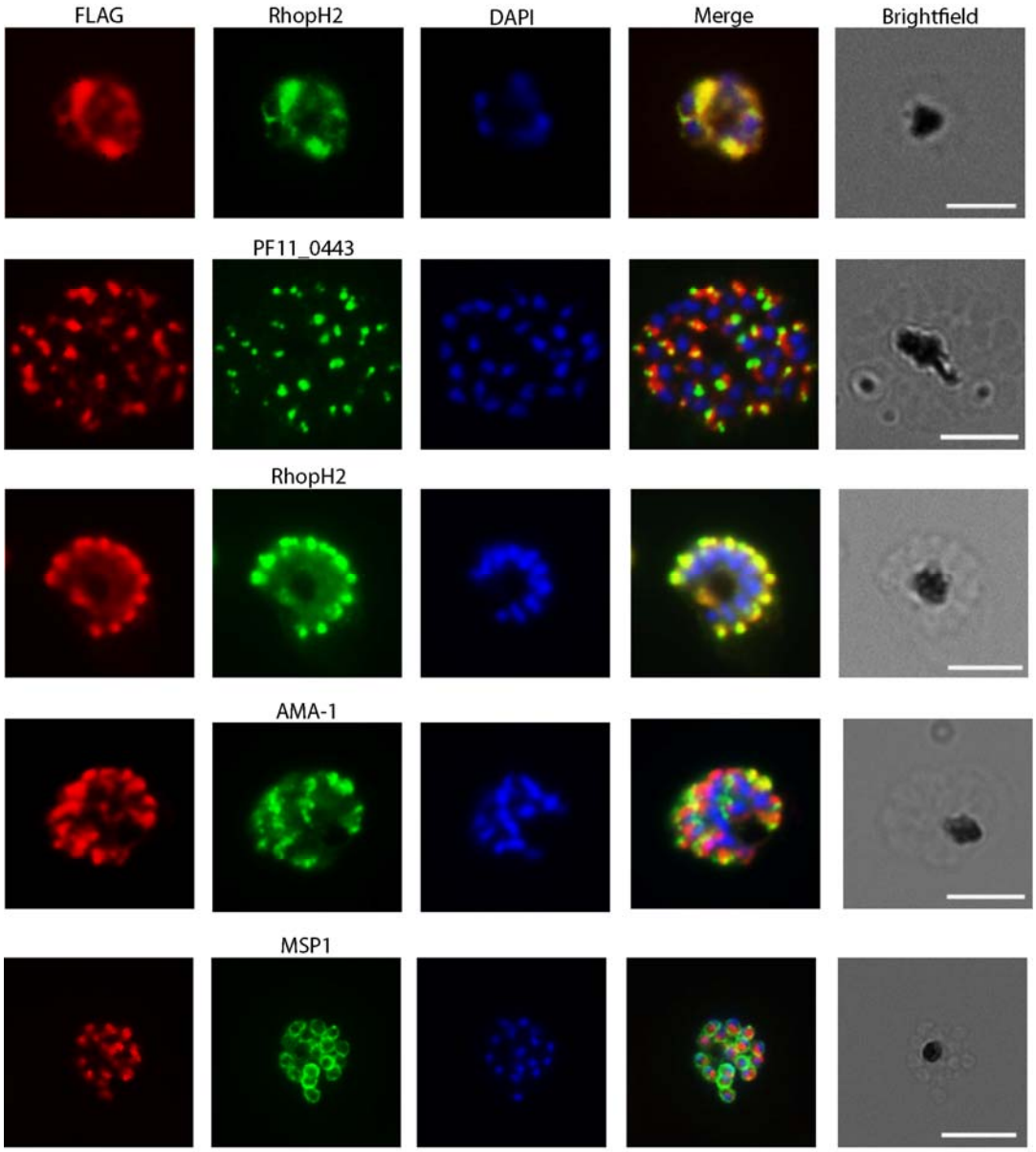
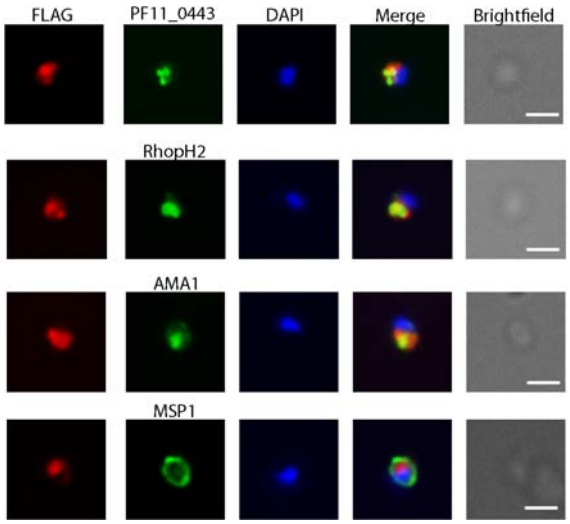

kDa
8
6
5
4
3
2
X
E
Int.
WT
WT C0 C3 clone- a c e

Figure 5.11: *FLAG tag PF11_0443 colocalises with RhopH2 in early, late schizonts and merozoites.* Thin smears of synchronised parasite culture containing schizonts and merozoites were fixed in paraformaldehyde and permeabilised with Triton X-100 (methods described in section 2. 2.13.) and used in IFAs to detect PF11_0443-FLAG in the parasite using the anti-FLAG M2 monoclonal antibody, labelled by a fluorescent secondary antibody AlexaFluor 594 (shown in red). Compartments were identified using previously characterised antibodies to proteins of known location followed by secondary antibodies labelled with AlexaFluor 488 (green) and images were merged to determine comparative location (AMA-1: microneme; RhopH2: Rhoptry bulb; MSP-1: surface). The PF11_0443CT antibody was used in combination with anti FLAG antibodies to show colocalisation. Antibodies and dilutions are described in chapter 2, table 2.4. The nucleus is stained with DAPI (blue). Scale bar on the brightfield image represents 5µm.


FLAG
RhopH2
DAPI
Merge
Brightfield
PF11_0443
RhopH2
AMA-1
MSP1


FLAG
PF11_0443
DAPI
Merge
Brightfield
RhopH2
AMA1
MSP1

5.9 Immunoprecipitation with anti-FLAG antibodies uncovers potential binding partners of PF11_0443.

Immunoprecipitation experiments using rabbit polyclonal antibodies identified a number of potential binding partners of PF11_0443. With an integrated 3xFLAG tag line now established, FLAG antibodies could be used to pull down PF11_0443-FLAG and any binding partners and therefore confirm the previously found partners. Protein samples were prepared from 3 ml of pelleted schizonts and passed over a column of FLAG M2 monoclonal antibody coupled to agarose beads (Sigma-Aldrich) (method described in 2.2.26). Unbound proteins were washed off with TBS before eluting bound proteins in a solution of 3xFLAG peptide at a final concentration of 150 nM in TBS (TBS or PBS?). The same experiment was performed using 3ml of WT 3D7 purified schizont pellet to control for unspecific binding to and elution from the FLAG affinity column. The elutions were concentrated to 30 µl and loaded onto a gel, separating eluted proteins by SDS-PAGE. The gel was stained with the fluorescent protein stain, Sypro Ruby (Invitrogen) (figure 5.12 A). This fluorescent stain was used in place of silver stain and is at least as sensitive but is also easier to remove from the protein during band preparation for mass spectrometry. Due to its increased sensitivity a large number of proteins were visible, however unique bands were identified and excised along with their equivalent control IP spot and analysed by Lc MS/MS in house by Dr. Steve Howell (Division of Molecular Structure). Results from mass spectrometry analysis are detailed in Appendix B. The table lists detected proteins' accession number, gene description, number of peptide matches, percentage peptide coverage and the band in which they were discovered. The table is ordered by descending MOWSE score. Proteins with lower scores but discovered in the PF11_0443_FLAG IP and absent from the control were also included as the lower score may be swayed by abundance, with proteins that are expressed at relatively low levels resulting in lower score and less percentage coverage. PF11_0443 is highlighted in blue, the next highest scoring protein in pink, other DnaJ domain-containing proteins are shown in purple and rhoptry proteins in yellow.

Looking at the results table, the strongest band with a high score of 2116.20 and coverage of 44.24%, labelled as number 9 (figure 5.12 A), was the protein of interest – PF11_0443-FLAG, confirming successful expression of the C-terminal epitope tag. The next highest scoring band comes from Hsp60 (PF10_0153) present in band 5. With a score of 1188.61 and coverage of 48.97%, this protein is far higher scoring than any of the other eluted proteins and no peptides were detected from it in the control lane. The same protein was present in the previous IP using antigen specific polyclonal antibodies, albeit at a lower level and with a score of 167 and 16% peptide coverage. Interestingly this protein was shown by IEM to reside within the mitochondrion in *Plasmodium* blood stages (Das et al., 1997), and is located exclusively within this organelle (Sato et al., 2003). The possibility of PF11_0443 residing within the mitochondria has not been investigated by IFA however the mitochondrion does not colocalise with apical organelles and PF11_0443 lacks any mitochondrial targeting sequence, as predicted by the PlasMit algorithm (Bender, 2002). For these reasons it is unlikely that the Hsp60 PF10_0153 is a binding partner of PF11_0443. Equally unlikely to bind PF11_0443 is the third highest scoring protein on the list, regulator of chromosome condensation MAL7P1.38. This protein is found in the nucleus [reviewed in (Frankel and Knoll, 2009)] and from IFAs using both specific polyclonal and FLAG monoclonal antibodies it is clear that PF11_0443 is not. Three conserved proteins of unknown function were identified as potential binding partners however since their function is unknown it does not aid the discovery of function of PF11_0443. Instead they should be kept in mind for future analysis. A *Plasmodium* –specific guanine nucleotide exchange factor (GEF) was co-precipitated and although these protein display a range of functions, including involvement in vesicular transport, the specific role of this protein is not known. Erythrocyte membrane protein (PfEMP)-3 is exported out into rbc cytsol in ring and trophozoite stages where it resides in Maurer's clefts (Pasloske et al., 1993; Knuepfer et al., 2005). This expression does not coincide with PF11_0443 and therefore these proteins are unlikely to function together. Rhoptry neck protein 4 (RON4) was also co-precipitated with PF11_0443-FLAG. Unlike the other co-precipitants, this protein is expressed in schizonts and merozoites and transferred into rings at the tight junction. It's also present at the apex, and although the best observed colocalisation of

PF11_0443 was with micronemal protein, the rhoptry neck is in close enough proximity for overlap to be observed with this sub-cellular compartment also. Another rhoptry protein observed is RhopH3, a protein of the rhoptry bulb. This protein was also co-precipitated with PF11_0443CT antibodies. There was a notable absence of any micronemal proteins, such as those observed in the previous pull down, AMA-1 and EBA-181. If PF11_0443 was to reside in the micronemes one might expect it to bind proteins present in these organelles.

In order to improve upon this IP it is important to reduce the background non-specific binding to the FLAG resin. In the previous experiment 500µl of FLAG-agarose was used in each precipitation. A western blot of the parasite lysates used before and after binding to the FLAG affinity column showed the complete removal of PF11_0443-FLAG from the column flow through (data not shown) so in an attempt to reduce background binding, less FLAG agarose could be used. This experiment was then repeated but using 100µl of FLAG agarose and washing with 300mM NaCl in place of TBS in the hope that this will decrease background binding. In the previous IP, the column was washed with TBS which contains a physiological concentration of NaCl since it was presumed that FLAG antibody binding and eluting with 3xFLAG peptide would only select specific interactors. Also washing with increased salt can disrupt protein complexes meaning results only represent salt-resistant complexes and not necessarily those that exist within the parasite. Due to the reduced volume of beads, binding and washing was performed in batch. The eluted protein was resolved by SDS-PAGE under reducing conditions however the protein gel was only allowed to run for 1.5 cm before staining with Instant Blue® and each lane was dissected into 6 equal bands 2.5 x 2.5 mm in preparation for analysis by LcMS/MS (as before, performed in house by Dr. Steve Howell). This difference in band preparation allows the analysis of the full protein complement instead of the subjective process of selecting and excising particular bands which can result in over-looked binding partners. Results of the mass spectrometry analysis are summarised in table 5.3, below and in detail in Appendix C. As before, the table lists detected proteins' accession number, gene description, number of peptide matches, percentage peptide coverage and the band in which they were discovered.

The table is ordered by descending MOWSE score. Proteins with lower scores but discovered in the PF11_0443_FLAG IP and absent from the control were also included as were proteins with high scores in the FLAG specific pull down but also present in the WT control IP, albeit at a reduced level. PF11_0443 is highlighted in blue, the next highest scoring protein in pink, other DnaJ domain-containing proteins are shown in purple and rhoptry proteins are in yellow. PF11_0443 was detected with a MOWSE score of 440.92, 5 fold less than with the previous IP. Only 2 other proteins obtained a higher score, however these were none specific binders, determined by presence in the WT control IP at an equivalent level to the FLAG IP. PF11_0443 co-precipitated a peptidyl-propyl *cis-trans* isomerase (PPIase) or cycophilin named CYP24 or CYP25 based on its predicted molecular weight of 24.9 kDa. This protein was present in band 9 (figure 5.12 B), and with a score of 421.56 and coverage of 33%, it is a strong binding partner candidate. These proteins display PPIase activity that is crucial in folding of proteins that require switching of peptidyl-propyl bonds from a *trans* to *cis* conformation during assembly and transport (Fischer and Aumuller, 2003). Cyclophilins are also known as immunophilins as they are targets of the immunosuppressive drug, cyclosporin A (CsA), the binding of which leads to the inhibition of protein phosphatase calcineurin in T-lymphocytes, a component of the calcium signalling pathway (Matsuda and Koyasu, 2000). CsA is also a potent anti-malarial (Bell et al., 1994). Reddy (1995) showed CYP25 is capable of PPIase activity and is sensitive to CsA (Reddy, 1995) however this finding could not be replicated in a more recent study where CYP25 was instead shown to exhibit chaperone activity, in particular the prevention of protein aggregation (Marin-Menendez et al., 2012). CYP25 is expressed throughout the lifecycle but it is possible that it could be recruited by PF11_0443 in late stages for a specific chaperoning role. Interestingly the initial IP using PF11_0443CT polyclonal antibodies pulled down the FKBP PPIase which is also an immunophilin, so what was originally attributed to the potential presence of anti-rotamase antibodies due to an unclean protein preparation may be a specific interaction with PF11_0443. Another cycophilin, CYP19A, is present in the ookinete microneme protein (Lal et al., 2009), providing further evidence for a role for these proteins in the apical organelles.

Both 40S and 60S ribosomal subunits featured heavily in this experiment. The nucleoproteins are common contaminants of pull down experiments however from experience with other immunoprecipitation experiments, including the two PF11_0443 related IPs preceeding this one, these proteins are always also seen in the control. Although PF11_0443 is predominantly observed in the ER in early schizonts, it is certainly apical in late schizonts – ribosomes are not. It therefore seems unlikely that this interaction is specific, particularly since the majority of ribosomal proteins were also observed in the control IP with WT parasites.

One of the highest scoring co-precipitating proteins in this experiment was actin depolymerising factor (ADF)-1. This protein has been shown to exclusively bind monomeric actin and its function is essential to *Plasmodium* blood stages (Schuler et al., 2005). ADF1 is cytosolic and recruited to the plasma membrane via its capacity to bind phosphatidyl inositol derivatives so it is hard to see what purpose the binding of PF11_0443 would serve.

In this second attempt, the rhoptry neck protein RON2 was co-precipitated but not its binding partner RON4, the reverse of what was observed in FLAG IP attempt 1. Moreover, rhoptry neck proteins, RON3 and RON5 were also co-precipitated. Although RON3 was present in the wild type control, the peptide coverage was less than half that observed in the FLAG specific IP and the score for per band was considerably less in the control IP. RON5 was only present in the FLAG specific IP however the score per band peaked at 79. This would normally be considered low however, since the score for the PF11_0443 were also lower, at 440.92, therefore it would be expected that all interactors will achieve even lower scores in comparison. Rhoptry bulb protein CLAG9 was also co-precipitated at similar levels to RON5.

PF11_0443-FLAG co-precipitated 2 other DnaJ domain containing proteins in this experiment. There is a precedent for J protein interactions in yeast where the mitochondrial pre-sequence-associated motor complex component J protein PAM18 has been shown to bind to J-like proteins (that lack an HPD) via their respective J

domains. PAM16 functions here in tethering PAM18 to the translocon and regulates the ability of PAM18 to stimulate ATP hydrolysis of the Hsp70 Ssc1(Frazier et al., 2004; Li et al., 2004b; D'Silva et al., 2005). The first of the co-precipitated J proteins is an uncharacterised type III J protein, PF08_0032, from which no functional clues can be derived. Like PF11_0443, it is also membrane bound however it appears to be expressed at a different stage of the intraerythrocytic cycle, with peak expression occurring in late rings and trophzoites and rapidly declining in schizonts (Bozdech et al., 2003), therefore it is improbable that they interact. The second is a soluble type II J protein, Pfj4, which has been shown to reside within the nucleus and cytoplasm and binds PfHsp70-1 (Pesce et al., 2008), and therefore unlikely to bind PF11_0443 due to spatial incompatibility.

Importantly this second attempt at the FLAG IP showed the presence of Hsp60 PF10_0153 in both wild type and FLAG specific pull downs, suggesting that it might be unspecific interactions that retained the protein on the matrix. This experiment also revealed that aldolase interacts nonspecifically with the column as it is present in both wild type and PF11_0443-FLAG IPs despite being present at high levels in the IP using PF11_0443CT polyclonal antibody.

Despite providing a more specific method of immunoprecipitation than use of polyclonal antibodies coupled to CNBr, there are many problems with the system. Surprisingly there was a large amount of unspecific binding to the FLAG agarose. Many eluted proteins were present in both WT 3D7 and PF11_0443-FLAG line extracts despite eluting with 3xFLAG peptide at physiological pH. Also, on the second FLAG IP attempt, a large proportion of PF11_0443-FLAG remained on the column after elution with 3xFLAG, as determined by boiling the FLAG-agarose beads in reducing sample buffer after 3XFLAG peptide elution and resolving the proteins by SDS-PAGE. This is perhaps due to a strong affinity of the FLAG M2 antibody to 3xFLAG and the 150nM concentration of 3xFLAG peptide solution used for elution was not high enough.

Immunoprecipitation studies using both anti PF11_0443 CT and the anti-FLAG M2 monoclonal antibody affinity gel has revealed a multitude of potential binding partners

some of which were also found eluted from the control columns. CYP25 is one of these potential interaction partners of PF11_0443. The presence of rhoptry proteins in the PF11_0443-FLAG elute suggests this protein may reside within this compartment. However unless PF11_0443 can be detected in a reverse IP no definitive conclusions can be drawn. If PF11_0443 is working in a chaperone secretory role one would expect it to interact with a substantial number of proteins. To try and attempt clarification of PF11_0443's function, analysis of phenotype from transgenic knockout or truncation lines is required.

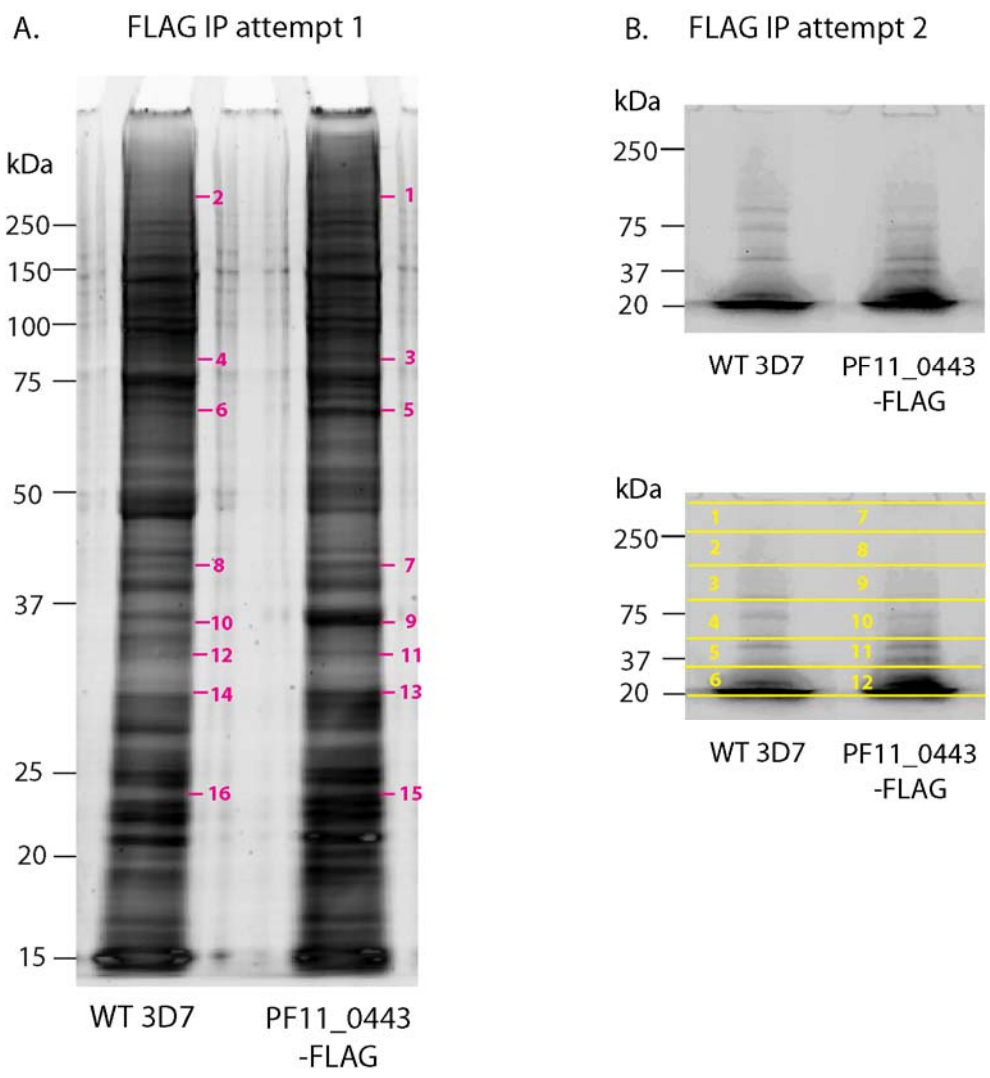


Figure 5.12: *Immunoprecipitation with anti-FLAG antibodies to uncover potential binding partners of PF11_0443*. Purified late schizonts from the PF11_0443-FLAG line were lysed in buffer containing 1% Triton X-100. The cleared lysate was bound in batch to anti-FLAG M2 affinity gel (methods described in chapter 2.2.26). Co-precipitating proteins were eluted using 150 nM 3xFLAG peptide in TBS and resolved under reducing conditions by SDS-PAGE. During the first attempt (A), proteins were allowed to run the full length of the gel and were visualised by Instant Blue. Unique bands compared to the control were excised (labelled 1-16) and the protein complement of each was identified by LC-MSMS. This experiment was repeated (B), however the gel was run 1.5 cm and both the control and PF11_0443-FLAG lanes were divided in to 6 approximately equal sized bands (WT control: 1-6; PF11_0443-FLAG: 7-12)before analysis by LcMS/MS.

5.10 Unsuccessful attempts to knock out *PF11_0443* suggest an essential role for this gene within blood stages.

Although there is a lack of functional data regarding the many othologues of PF11_0443, there is evidence from *Drosophila melanogaster* (flybase.org) and *C. elegans* (wormbase.org/species/c_elegans) that the corresponding genes in these species can produce viable knock-outs in a laboratory environment. To ascertain whether this is also the case for the *P. falciparum* version of the gene, the production of a *PF11_0443* knock-out line was attempted. A double homologous recombination strategy was employed utilising the pHTK vector (figure 5.13 A), which contains 2 multiple cloning sites that flank the human dihydrofolate reductase cassette, conferring resistance to WR99210 upon integration. The construct also contains the thymidine kinase cassette which confers susceptibility to ganciclovir. A region of 480 bp from the 5′ UTR and ORF of *PF11_0443* was amplified by PCR using primers N and O (sequences listed in chapter 2, table 2.1) and introduced to the construct via the first flank multiple cloning site (F1) using the restriction sites of *Sac*II and *Bgl*II. A second region of homology from the 3′ end and UTR of *PF11_0443* was amplified by PCR using primers P and Q (chapter 2, table 2.1) and introduced into the second multiple cloning site (F2) utilising *Eco*RI and *Avr*II sites to produce the integration transfection vector pHTK_PF11_0443_KO.

Inserts were initially sub-cloned into the pGEM T-easy vector and transformed into *E. coli* TOP10 cells. Colonies were selected and sequenced before excising the correct sequence inserts using *Sac*II and *Bgl*II (F1) or *Eco*RI and *Avr*II (F2). Inserts were gel-purified and ligated into the pHH3 vector and transformed into *E. coli* TOP10. Colony plasmid DNA was screened by restriction digest using *Sac*II and *Bgl*II (F1) or *Eco*RI and *Avr*II (F2) (figure 5.13 B) and re-sequenced before selecting one correctly sequenced vector to propagate and purify for transfection.

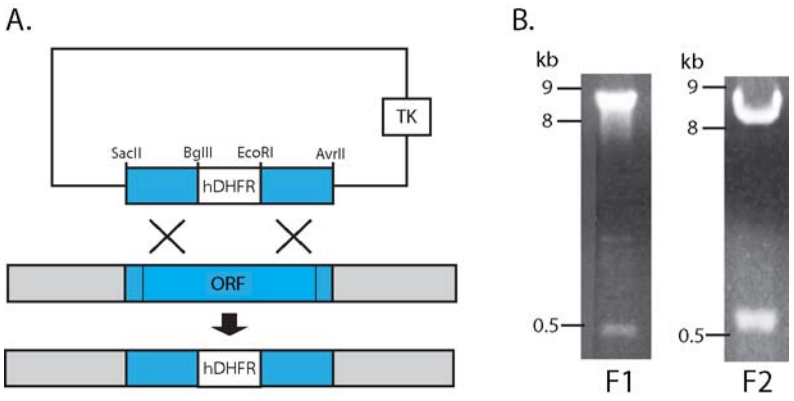


Figure 5.13: PF11_0443 *KO strategy and vector construction.* A double homologous recombination strategy was employed for the KO of *PF11_0443* which would result in the complete removal of the gene. The transfection vector pHTK was manipulated to include 480 bp from the 5' UTR and ORF of *PF11_0443*, amplified by PCR using primers N and O (sequence listed in chapter 2, table 2.1) and introduced to the construct's flank 1 (F1) multiple cloning site using the enzymes *Sac*II and *Bgl*II. A 540 bp fragment was amplified using primers P and Q (table 2.1) and introduced into flank 2 using *Eco*RI and *Avr*II. Digestion with these enzymes from the completed plasmid excised resulted in the correct size bands from flanks 1 and 2 (B).

As with the epitope tag constructs, the complete and purified pHTK_PF11_0443_KO was transfected into ring-stage parasites. Selection for presence of episome was achieved by the addition of 5 nM of WR to transfection cultures until parasitaemia recovered to around 10%. Stock samples of these cultures were taken and stored in liquid nitrogen and genomic DNA was prepared in order to test for presence of integrants. Transfection cultures were taken off drug for 3 weeks before its re-application, which was repeated initially 3 cycles or in the absence of integration, 5 cycles. A Southern blot was designed to analyse potential integration events (figure 5.14 A). Genomic DNA taken each drug cycle was digested by *Pac*II and *Bgl*II before separating resulting fragments by gel electrophoresis and blotting onto nitrocellulose membrane. The membrane was probed with a radiolabelled oligomer of the region of homology. The first knock-out attempt resulted in no integration throughout 5 cycles; this was repeated with an additional 2 transfections, both of which resulted in 3' integration. On these occasions (one of which

is presented in figure 5.14 B below), the resulting banding pattern of Southern blot on exposure to film revealed the presence of wild type DNA with a band of 2560 bp (figure 5.14) in the control lane containing 3D7 DNA, as well as at drug cycles 1 and 2 before disappearing completely. As the WT band disappears, a band of between 6 and 8 kb appears (labelled 'int'), corresponding to the 7218 bp fragment produced upon digest in the presence of integration of flank 2 into the 3' end of the locus. A band corresponding to the episome (labelled 'E') is present throughout the drug cycling until cycle 5 on WR99210 and ganciclovir when only the band corresponding to 3' integration is observed.

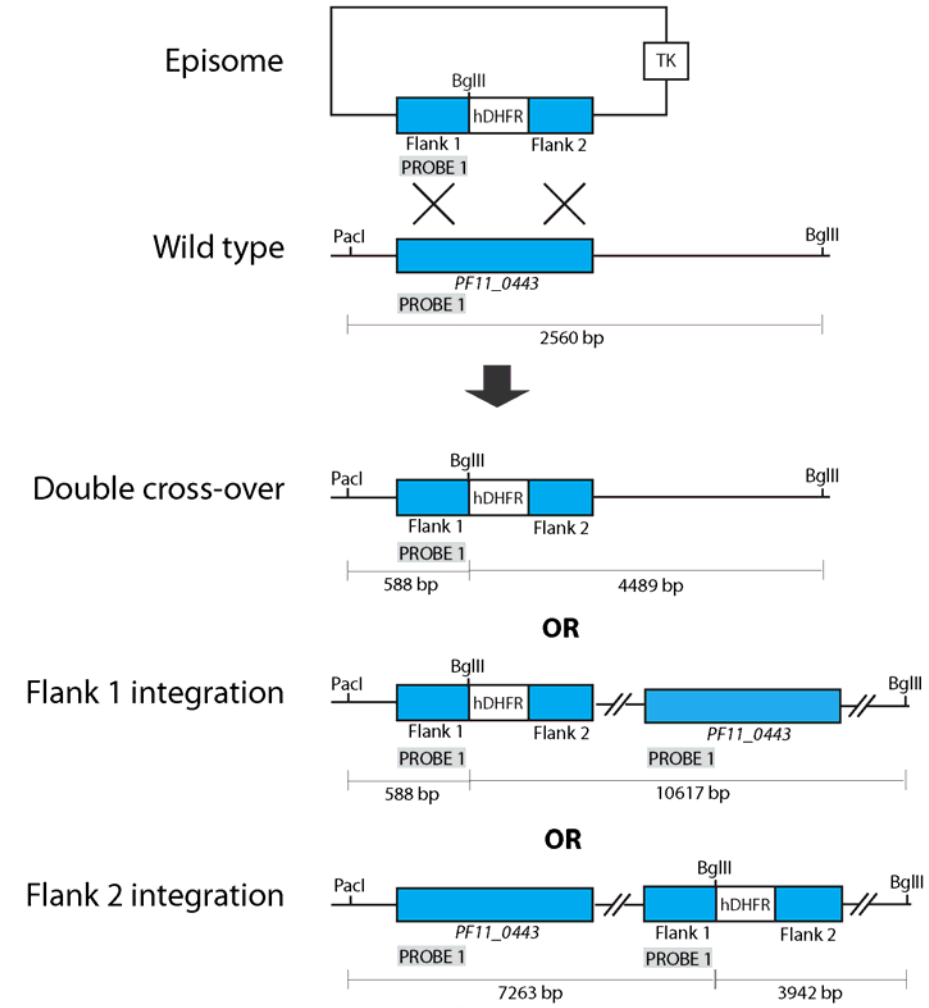
Integration of only the second flank may not result in a knock out since a full copy of the gene will be present only minus the 3' UTR. The absence of 3' UTR could mean termination of translation is suspended or RNA stability affected due to lack of poly A tail. In order to test whether PF11_0443 is still expressed in the flank 2 integrated line and at the same level as the wild type protein, both lines were synchronised to approximately 1.5 h and protein samples were prepared from late schizonts. Samples were separated by SDS-PAGE under reducing conditions and subject to western blot analysis. Bands produced by exposure to film from wild type and flank 2 integrated lines were compared using ImageJ software. Blots were also probed with a BiP antibody to ensure equal loading and with an MSP2 antibody to ensure the protein samples were of late schizonts. From figure 5.15 A it appears that PF11_0443 is expressed in the flank 2 integrated line and to a similar level as the wild type. Using ImageJ software (Collins, 2007) to quantify band intensity, it was clear that a greater number of KO parasites were loaded compared to wild type at a ratio of 1.1 as calculated by comparison with BiP expression in each line. MSP2 levels were used as a stage-specific control and it appears parasites loaded in the KO lane were more mature than the wild type parasites at a ratio of 1.3. There was 1.2 times more PF11_0443 in the KO lane compared to the wildtype. Therefore the expression of PF11_0443 in the KO construct integrated line was 91.5% that of the wildtype. An IFA of flank 2 integrated parasites using the specific rabbit polyclonal PF11_0443CT antibody revealed that the protein is expressed in every

parasite in that particular field and this was representative of the entire slide (figure 5.15 B).

Three unsuccessful knock-out attempts of *PF11_0443*, despite a clearly accessible locus, suggest an essential role for this gene within the parasite. Although flank 2 was seen to integrate in 2 of these attempts, only a small reduction in *PF11_0443* expression ocurred, which is well within the error of the quantification method used and not considered as a genuine phenotype. However the region of homology of flank 1 of 480 bp is relatively short and contains a component of AT rich 5'UTR sequence therefore the possibility that the lack of integration is due to this technical hurdle in homologous recombination cannot be ruled out.

Figure 5.14: *Attempted KO of PF11_0443 by double homologous recombination results in integration of flank 2 only*. A Southern blot was designed to confirm this result (A). Genomic DNA taken each drug cycle was digested by *Pac*I and *Bgl*II before separating resulting fragments by gel electrophoresis and blotting onto nitrocellulose membrane. The membrane was probed with a radiolabelled oligomer of the region of homology (B). The wild type DNA band of 2560 bp is present in the control lane containing 3D7 DNA, as well as in drug cycles 1 and 2 before disappearing completely. As the WT band disappears, a band of between 6 and 8 kb appears (labelled 'int'), corresponding to the 7218 bp fragment produced upon digest in the presence of integration of flank 2 into the 3' end of the locus. A band corresponding to the episome (labelled 'E') is present through out the drug cycles until cycle 5 on WR99210 and ganciclovir when only the band corresponding to 3' integration is present.

A.


Episome
TK
BglII
hDHFR
Flank 1
Flank 2
PROBE 1
Wild type
PacI
BglII
PF11_0443
PROBE 1
2560 bp
Double cross-over
PacI
BglII
hDHFR
Flank 1
Flank 2
PROBE 1
588 bp
4489 bp
OR
Flank 1 integration
PacI
BglII
hDHFR
Flank 1
Flank 2
PF11_0443
PROBE 1
PROBE 1
588 bp
10617 bp
OR
Flank 2 integration
PacI
BglII
PF11_0443
hDHFR
Flank 1
Flank 2
PROBE 1
PROBE 1
7263 bp
3942 bp

B.

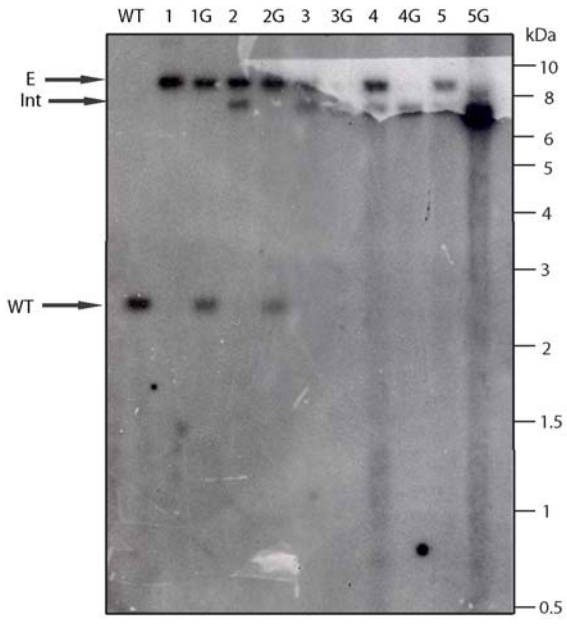
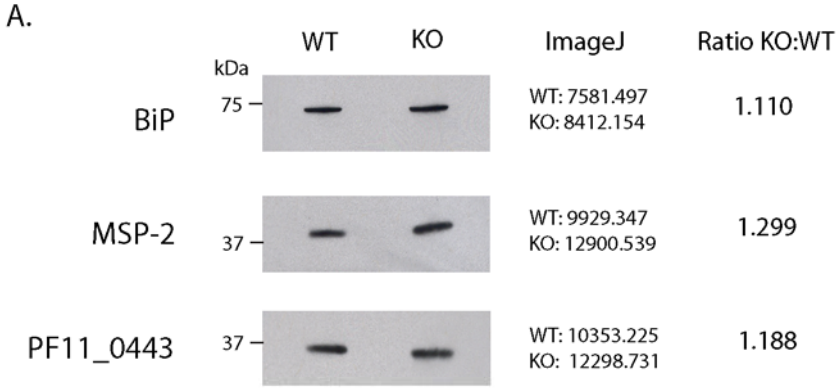

WT 1 1G 2 2G 3 3G 4 4G 5 5G
kDa
10
8
6
5
4
3
2
1.5
1
0.5
E
Int
WT

Figure 5.15: *PF11_0443 KO construct flank 2 integration had a small effect on PF11_0443 expression.*

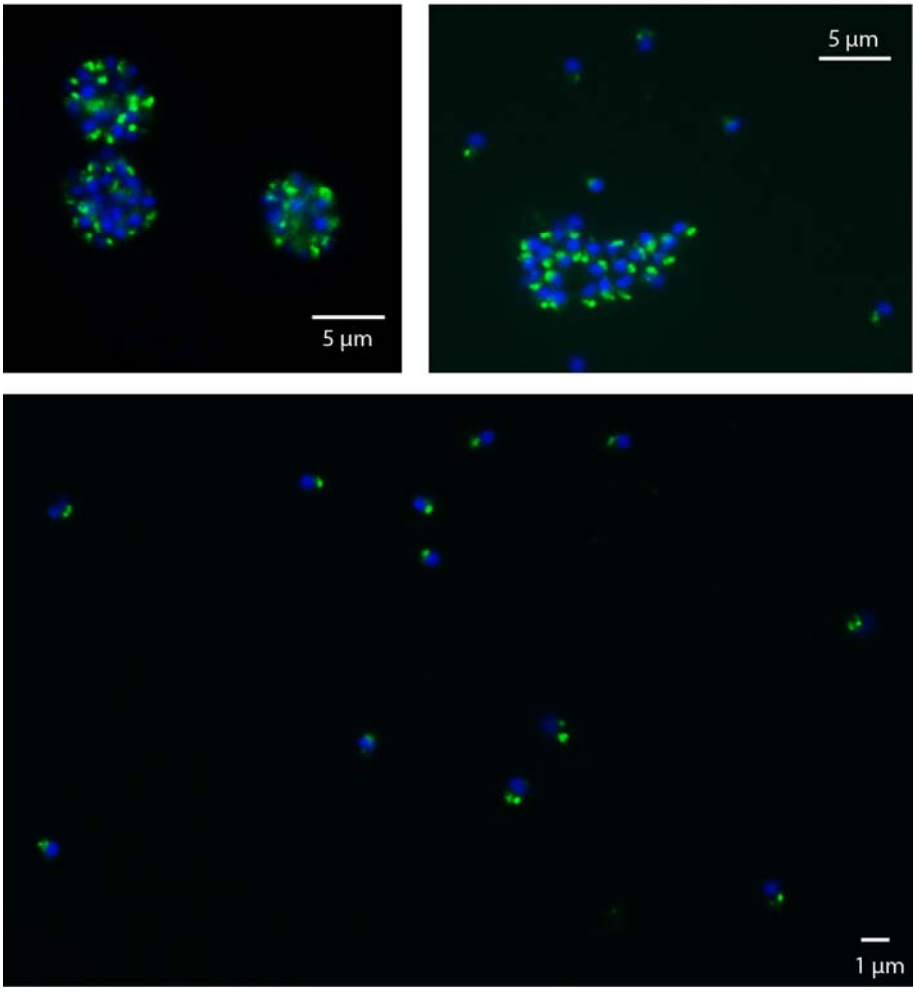
A) Protein samples of purified late schizonts from wild type and flank 2 integrant (knock-out attempt 2, drug cycle 5 on WR99210 and ganciclovir) parasite lines, synchronised to approximately 1.5 hrs, were resuspended in reducing SDS sample buffer and separated by SDS-PAGE with approximately $2x10^6$ parasites loaded per well. Expression of PF11_0443 was analysed by western blot using the specific rabbit polyclonal antibody PF11_0443CT. Antibodies recognising the parasite ER resident protein BiP were used to analyse equality of loading between lanes and a monoclonal antibody against MSP-2 was used as stage-specific marker, ensuring the schizonts loaded into control and integrant lanes were the same level of maturity. Molecular weight markers for each western blot are shown in kDa on the left hand side of the figure. The intensity of each band was analysed by ImageJ software (Collins, 2007) and the ratio of KO to wild type was calculated. This revealed that expression in the flank 2 integrated line was 91.45% that of the WT.

B) Thin smears of flank 2 integrated parasite culture from synchronised late schizonts and merozoites were fixed and permeabilised and used in IFAs to detect the presence of PF11_0443 in the parasite using specific rabbit polyclonal antibodies PF11_0443CT. PF11_0443 localisation is depicted in green (achieved by secondary antibody AlexaFluor488) and the nucleus is stained with DAPI, shown in blue. The length of representative scale bars are stated on individual images. PF11_0443 can be detected in every parasite and the images taken were representative of all observed fields.

A.


WT
KO
ImageJ
Ratio KO:WT
kDa
BiP
75
WT: 7581.497
KO: 8412.154
1.110
MSP-2
37
WT: 9929.347
KO: 12900.539
1.299
PF11_0443
37
WT: 10353.225
KO: 12298.731
1.188

B.


5 µm
5 µm
1 µm

5.11 Discussion

The data presented here show that PF11_0443 is an integral membrane protein expressed from early schizogony from 36 h post invasion and maintained until early ring stages. In early schizonts this protein appears within the ER, as shown by IFA where it co-localises with ER resident Hsp70 BiP. The early expression and presence within the ER in schizonts suggests a role for PF11_0443 in protein sorting and trafficking, particularly given the change in location upon daughter cell segmentation. In these late schizonts PF11_0443 moves to the apical end of the merozoite where it appears to co-localise with micronemal proteins AMA-1 and EBA-175. As there are multiple micronemes within the merozoite and AMA-1 and EBA-175 do not co-localise with each other this pattern could indicate PF11_0443 is present in each of them. This is consistent with the fact there are multiple foci of fluorescence observed by IFA with PF11_0443CT antibodies.

One potential problem with this finding is that PF11_0443 was not detectable on the surface of the free merozoite where micronemal proteins are secreted immediately prior to invasion. Also, micronemal protein such as AMA-1 and EBA-175 are shed upon invasion (Harris et al., 2005; O'Donnell et al., 2006), however this protein was not detectable in culture supernatants by western blot. Although PF11_0443 was not detected in rings by western blot, the protein was detected by IFA at this stage immediately following erythrocyte invasion. Antibodies were used here in more concentrated form for ring stage IFAs than the schizont or merozoite IFAs and camera exposure times were doubled when taking the images, implying PF11_0443 is present at a lower level. This apparent reduction could explain why this protein was not visible by western blot using ring stage lysates (figure 5.3).

The fact that PF11_0443 is carried through into rings suggests PF11_0443 resides within a different apical organelle such as the dense granules or rhoptry neck/bulb. Due to their close proximity within the merozoite and the fact the limit of the resolution of light microscopy, it is often difficult to distinguish between the rhoptry neck and micronemes. Although an IFA using a RESA antibody did not show colocalisation with

PF11_0443, the antibody was not sensitive and gave rise to an unexpected pattern – only one point of fluorescence, despite its presence in multiple dense granules, as seen by immuno-electron microscopy (IEM) (Culvenor et al., 1991). IEM would have provided stronger evidence to determine the sub-cellular location of PF11_0443 however attempts at this method using PF11_0443CT rabbit polyclonal antibodies were unsuccessful.

Without IEM, the ambiguous apical location cannot be pin-pointed to a particular organelle and the sub-cellular location of PF11_0443 must be verified by another means. For this reason the introduction of a C-terminal epitope tag into the genomic locus of the protein was undertaken, so that the antibodies against the tag could be used to determine protein location. The aspartic acid rich FLAG epitope tag was chosen due to its small size and commercially available tools such as monoclonal antibodies and affinity resin. Antibodies to this tag have also been used in IEM in the past (Wada and Kanwar, 1998). Both 3xFLAG and Ty1 epitope tags were successfully integrated into the *PF11_0443* locus by cycle 2, highlighting its accessibility to genetic manipulation. Both tags were visible by western blot using tag specific antibodies however due to the superior sensitivity of the FLAG antibody, this tagged line was chosen for use in future experiments and was subsequently cloned. PF11_0443-FLAG confirmed what was already seen by IFA using PF11_0443CT antibodies – that this protein is present in the ER in early schizonts but this location changes upon daughter cell segregation, as in late schizonts PF11_0443-FLAG is seen in an apical location. Colocalisation with organelle markers show PF11_0443 co-localises with RhopH2 in schizonts and merozoites, but not with micronemal marker AMA1, as had previously been seen by IFA using wild type parasites. Since the IFAs using wild type parasites were inconclusive due to colocalisation with markers of multiple organelles, it is possible that the superior sensitivity of the anti-FLAG M2 monoclonal antibody has produced the actual localisation of both tagged and wild type PF11_443. Although the colocalisation of WT PF11_0443 with micronemal proteins was strong, PF11_0443 was not released onto the surface or shed into culture supernatants, as seen with other characterised micronemal proteins; instead it is transferred into the newly invaded rbc, a fact which supports a

rhoptry location. To address the question of subcellular location IEM should be undertaken in the future using the anti-FLAG M2 monoclonal antibody on the PF11_0443-FLAG parasite.

The multi-foci, punctate staining observed by immunofluorescence using anti-PF11_0443 or FLAG antibodies is reminiscent of the small vesicle-like structures or 'J-dots' which are exported into the rbc cytoplasm and contain the type II J proteins PFA0660w and PFE0055c. These highly motile vesicles are thought to be involved in the transport of parasite proteins through the rbc cytoplasm. Similar structures are observed using antibodies to the human orthologue of PF11_0443, DNAJC25. Appendix F shows a fluorescent image of DNAJC25 from the skin derived cell line A-431 from antibodies produced by the Human Protein Atlas [www.proteinatlas.org (Uhlen et al., 2010)]. These vesicles appear in both the nucleus (excluding nucleolus) and the cytoplasm. Although the function of both human and parasite proteins are unknown, the presence of J proteins in small vesicular structures appears to be common throughout eukaryotes and supports a common role in vesicular transport.

Although clearly membrane bound, it is not clear which orientation PF11_0443 exists. Chaperone activity may be required inside apical vesicles in order to prevent protein aggregation or may facilitate formation of protein complexes therefore one might expect the J domain to be present inside these vesicles. Equally PF11_0443 may interact with molecules involved in vesicular transport and therefore the J domain and C-terminus would be present on the cytoplasmic side of the membrane. The topology of PF11_0443 will be addressed in the future by proteinase K digestion, as previously described (Feldheim et al., 1992; Cabrera et al., 2012).

If PF11_0443 is to reside within the rhoptries then a rhoptry targeting motif should be present. Although dileucine and tyrosine based motifs have been described in *Toxoplasma* for the rhoptry-targeting of membrane proteins (Hoppe et al., 2000; Ngo et al., 2003), evidence for these particular motifs is lacking in *Plasmodium*. Looking at the PF11_0443 protein sequence (Chapter 3, figure 3.4), there is a potential tyrosine based

motif at the C-terminus. Additionally, a dileucine motif exists immediately following the predicted signal peptididase cleavage site. This dileucine motif has to be cytoplasmic in order to facilitate targeting to the rhoptries (Hoppe et al., 2000; Ngo et al., 2003), therefore if this is a targeting motif then the N-terminus and J domain would be present of the cytoplasmic face of the ER and rhoptries. Experimental manipulation of these potential targeting motifs would establish whether they are required for targeting PF11_0443 to the rhoptries.

Ghoneim *et al* (2007) identified the first 24 aa of PfRhopH2 as sufficient to target this protein to the rhoptries however the first 20 of these amino acids are predicted to contain a signal peptide and will be cleaved upon translocation into the ER. The authors suggest a role for chaperones in identifying rhoptry targeting sequences however experimental evidence is lacking. Since signal peptides are hydrophobic and chaperones such as Hsp40 and Hsp70 recognise unfolded hydrophobic stretches, perhaps PF11_0443 has a role here.

Although identifying the protein's location is important, determining the function of PF11_0443, in the absence of any data from orthologous proteins this is difficult. Uncovering binding partners in order to establish a function is one approach and initial immunoprecipitation experiments using PF11_0443CT rabbit polyclonal antibodies coupled to CNBr did reveal a number of candidates. However verification was lacking. Immunoprecipitation studies using the anti-FLAG M2 monoclonal antibody affinity gel has revealed, amonst others, the cycophilin CYP25 as a potential interaction partner. This protein has been shown to possess chaperone activity in the form of prevention of protein aggregation *in vitro* (Marin-Menendez et al., 2012), which is consistent with data from the *C. elegans* orthologue, dnj-2; the RNAi knockdown of this protein results in a protein aggregation mutant (Hamamichi et al., 2008). CYP25 is also present in *C. elegans,* and all other species where PF11_043 orthologues were detected as well as other Apicomplexa in which PF11_0443 was not. CYP25 does not contain a signal peptide which implies it is cytosolic and therefore could not function to prevent protein aggregation inside any of the apical organelles. Since PF11_0443 is membrane bound, it

may function to recruit CYP25 to a specific membrane where required. A member of the cyclophilin family, CYP19A, is present in the microneme proteome, indicating a role for these chaperones in the apical organelles (Lal et al., 2009). The result presented here needs to be replicated and confirmed by IP followed by western blot and using anti CYP25 as well as anti-FLAG antibodies before coming to definite conclusions.

The *P. berghei* microneme proteome identified 23 members of the vesicular trafficking network and an additional 23 chaperones, one, as discussed above was CYP19A, and a further 8 belonged to the J protein family (Lal et al., 2009). Although there is no available rhoptry proteome (ookinetes don't have rhoptries), this data suggests there is a role for chaperones in the apical organelles.

PF11_0443-FLAG also co-precipitated a number of rhoptry proteins, in particular rhoptry neck proteins; providing further evidence that PF11_0443 may be a resident of the rhoptries. To confirm this finding, immunoprecipitation experiments should be performed using RON2 or RON4 antibodies to ascertain whether the interaction is reciprocal. The FLAG IP did not pull down protein disulphide isomerase (PDI) which is a putative ookinete micronemal protein (Lal et al., 2009), and was co-precipitated in the IP using PF11_0443CT antibodies. Unfortunately this protein is a common contaminant in IP experiments in the lab and therefore reciprocal IP experiments are required to confirm this protein as a binding partner for PF11_0443.

The introduction of this C-terminal tag did not appear to confer a phenotype on the parasite since there was no difference in parasite invasion rates (data not shown). It is therefore presumed that the tag does not prevent protein-protein interactions, validating the use of the tagged line in IP experiments. However perhaps the protein is not essential and the phenotype was too subtle to be observed. Evidence from PF11_0443 orthologues suggest a non-essential role since the corresponding genes of *Drosophila melanogaster* (flybase.org) and *C. elegans* (wormbase.org/species/c_elegans) produce viable knock-outs in a laboratory environment. To ascertain whether this is true for PF11_0443, production of a knock-out line was attempted by double homologous

recombination. The first KO attempt resulted in no integration of either flank after 5 drug cycles. The transfection was repeated in triplicate and attempts 2 and 3 resulted in integration of flank 2 at the 3' end of the gene. Because integration of only this flank would result in the presence of a full copy of the gene but without the full 3'UTR, PF11_0443 could still be expressed and this was in fact the case, with only a small reduction in expression level compared to wild type, lying within the margin of error of the experiment. Although the locus is clearly accessible, there was never any specific control for integration of N-terminal region of *PF11_0443*. Also the region of homology of flank 1 is shorter than that of flank 2 (480 bp) and contains a component of AT rich 5'UTR sequence therefore the possibility that the lack of integration is due to this technical difficulty in homologous recombination cannot be ruled out.

A conditional KO would be beneficial here to establish a possible phenotype derived from the removal of the protein, however the capacity to generate a conditional KO is not fully established in *P. falciparum* and RNAi knock down, so commonly used in other organisms, is not available for the study of *Plasmodium* as it lacks the protein machinery necessary for RNAi based gene regulation (Baum et al., 2009b).

The fact that *Plasmodium* has retained this gene at all suggests it has a function important to the parasite, but it is not necessary in other Apicomplexa. So what could be the function? Host cell choice seems to be the most obvious answer. The only reason a parasite as well adapted to its environment as *Plasmodium* would hold on to this protein is if it had a specific need for it that it couldn't get from its host. Although this protein is not likely to be expressed in rbcs it is also unlikely that the other apicomplexa recruit and internalise a specific host chaperone. Also, there are two other erythrocyte-dwelling apicomplexa – *Theileria* and *Babesia* and they also do not appear to harbour a copy of this gene. There are in fact several major differences between *Plasmodium* and its fellow Haemosporidia, in both the mode of invasion and nature of habitation within erythrocytes. *Theileria* merozoites are not polarised and lack micronemes; preferring instead to invade at any orientation following irreversible attachment of their surface proteins (Shaw, 2003). *Babesia,* possess merozoites that are more similar to that of

Plasmodium and invasion mechanisms appear conserved due to the presence of rhoptries and micronemes (Potgieter and Els, 1977). Interestingly, both *Babesia* and *Theileria* form a PV that rapidly disintegrates upon the completion of rbc invasion, unlike the PV of *Toxoplasma* and *Plasmodium* which require an intact PV throughout the asexual cycle (Sam-Yellowe, 1996). If one speculates that PF11_0443 were to have a unique and necessary function, it could be linked in some way to the parasitophorous vacuole. Since PF11_0443 levels diminish in ring stages, perhaps their role is linked to the rhoptries since proteins from both the bulb and the neck are to be found there post invasion. Although theoretical, this hypothesis, along with the experimental evidence presented here, suggests that PF11_0443 is to be found in the rhoptries of late schizonts and merozoites.

This study has identified a schizont-stage with possible chaperone function, present in the ER in early schizonts until segregation of daughter merozoites brings about a change in localisation resulting in translocation of the protein to the apex, where it appears to be present in the rhoptries. The signal dictating this change in location remains to be determined. A recent phosphoproteomic study detected the phosphorylation of serine 316, 5 residues from the C-terminus of PF11_0443 (Treeck et al., 2011) and it is possible that this signal is responsible for the change in location. Initial work has begun to substitute this serine with alanine and this integration has been successful, implying phosphorylation of this residue is not essential for parasite growth, but further study of this phosphorylation is required.

The initial aim of this study was to identify and characterise a protein involved in the invasion of erythrocytes by *Plasmodium falciparum* merozoites. PF11_0443 is a (putatively) essential protein present at the apex of merozoites and although the precise location has not been elucidated, the protein has met the project criteria. Further work is required to identify definitive binding partners and uncover the function of this protein within the parasite.

Chapter 6. Localisation and functional characterisation of PF02_0040

6.1 Introduction

6.1.1 ER targeting and translocation

Proteins destined for an extracellular existence or residence in a membrane-enclosed compartment must enter the ordered labyrinth of the endoplasmic reticulum. In order to enter the ER, newly synthesised proteins require a signal sequence at the N-terminus. This sequence is recognised by components of the protein translocation machinery which grant the signal containing protein access to the ER lumen (if soluble) or permit insertion into the ER membrane. There are two mechanisms of entry into the ER: co-and post-translation translocation. Each translocation pathway is illustrated in figure 6.1. The co-translational translocation of proteins into the ER (figure 6.1 a) involves recognition and interaction of the signal sequence and signal recognition particle (SRP) at the exit site of a ribosome during translation, causing a temporary arrest in elongation (Meyer and Dobberstein, 1980a, b; Walter and Blobel, 1981b). The complex of SRP, polypeptide and ribosome docks with the membrane associated SRP-receptor (SRPR) in the ER membrane (Walter and Blobel, 1981a), and can now be translocated into the ER lumen/membrane through the Sec61 complex comprised of α, β and γ subunits that form a pore (Stirling et al., 1992; Wirth et al., 2003). Completion of translocation involves the recruitment of the ER luminal Hsp70 immunoglobulin heavy-chain-binding protein (commonly known as BiP) by its Hsp40 partner, membrane bound Sec63 (Feldheim et al., 1992; Corsi and Schekman, 1997). BiP gates the Sec61 channel (Alder et al., 2005) and serves as a molecular ratchet (Tyedmers et al., 2003). The signal peptide is removed by signal peptidase unless the peptide is a transmembrane region (signal anchor), in which case no cleavage will take place. Translocation is also possible without the SRP and SRPR. The post-translational translocation pathway (figure 6.1 b) is SRP-independent and instead recognition of signal sequence is thought to be performed by a heterotrimeric complex of membrane bound Sec62, Sec63 and Sec71, and soluble Sec72 before

translocation through the Sec61 complex (Deshaies et al., 1991; Feldheim et al., 1993; Fang and Green, 1994; Wittke et al., 2000; Harada et al., 2011). Signal sequences recognised by the Sec71/Sec72 complex of the post-translational translocon are often less hydrophobic than those bound by SRP (Ng et al., 1996).

Note: Sec71 was characterised independently by two different laboratories investigating two different mutations in *Saccharomyces cerevisiae* – *SEC66* (Feldheim et al., 1993) and *SEC71* (Fang and Green, 1994). Although Sec66 was published first, the authors suggest the protein should be termed Sec71 in the event that the genes were shown to be identical which was achieved by Fang & Green (1994). For this reason the protein shall hence-forth be termed Sec71. The protein domain name according to the Pfam database (pfam.sanger.ac.uk) remains Sec66.

6.1.2 Translocation across the kingdoms

Much of our knowledge of eukaryotic cell biology comes from work performed in yeast, and the secretory pathway is no different. However with the increasing prevalence of whole genome sequencing, more information is becoming available from other organisms. *Saccharomyces* possess the machinery for and utilises both the SRP-dependant and independent pathways at an apparent equal frequency, but this is not the situation in other organisms. *Homo sapiens* appear not to utilise the SRP-independent pathway as they do not possess Sec71 or Sec72 orthologues (Zimmermann and Blatch, 2009). Interestingly, much of our recent knowledge of the ER protein translocon comes from work done by Goldschmidt *et al* (2008) from their work on the protozoan parasite *Trypanosoma brucei*, the cause of African sleeping sickness. Although this organism does not possess Sec72, it does contain an orthologue of Sec71 and this was shown to be essential by RNA interference. Furthermore, *SEC71* silenced parasites showed a significant reduction of GPI-anchored proteins that was not observed in *SRP68* silenced cells. Like *Plasmodium*, the *T. brucei* cell surface is predominated by a GPI-anchored protein – variant surface glycoprotein (VSG) – and the parasite could not tolerate a reduction in its production (Goldshmidt et al., 2008).

Very little is known regarding the secretory pathway of *P. falciparum*. The only member of the translocation machinery characterised is PfSec61 (Couffin et al., 1998). A study by Tuteja in 2007 bioinformatically identified the components of the translocon complex in *P. falciparum* and although they uncovered orthologues of the co-translational translocation pathway, putative Sec71 or Sec72 proteins were absent suggesting *Plasmodium* does not utilise the post-translational translocation pathway (Tuteja, 2007). Due to the A/T rich nature of the *P. falciparum* genome, there is often little sequence identity between proteins, despite sharing a homologous function.

The protein selection criteria in this study uncovered a potential orthologue of Sec71 in *P. falciparum*.

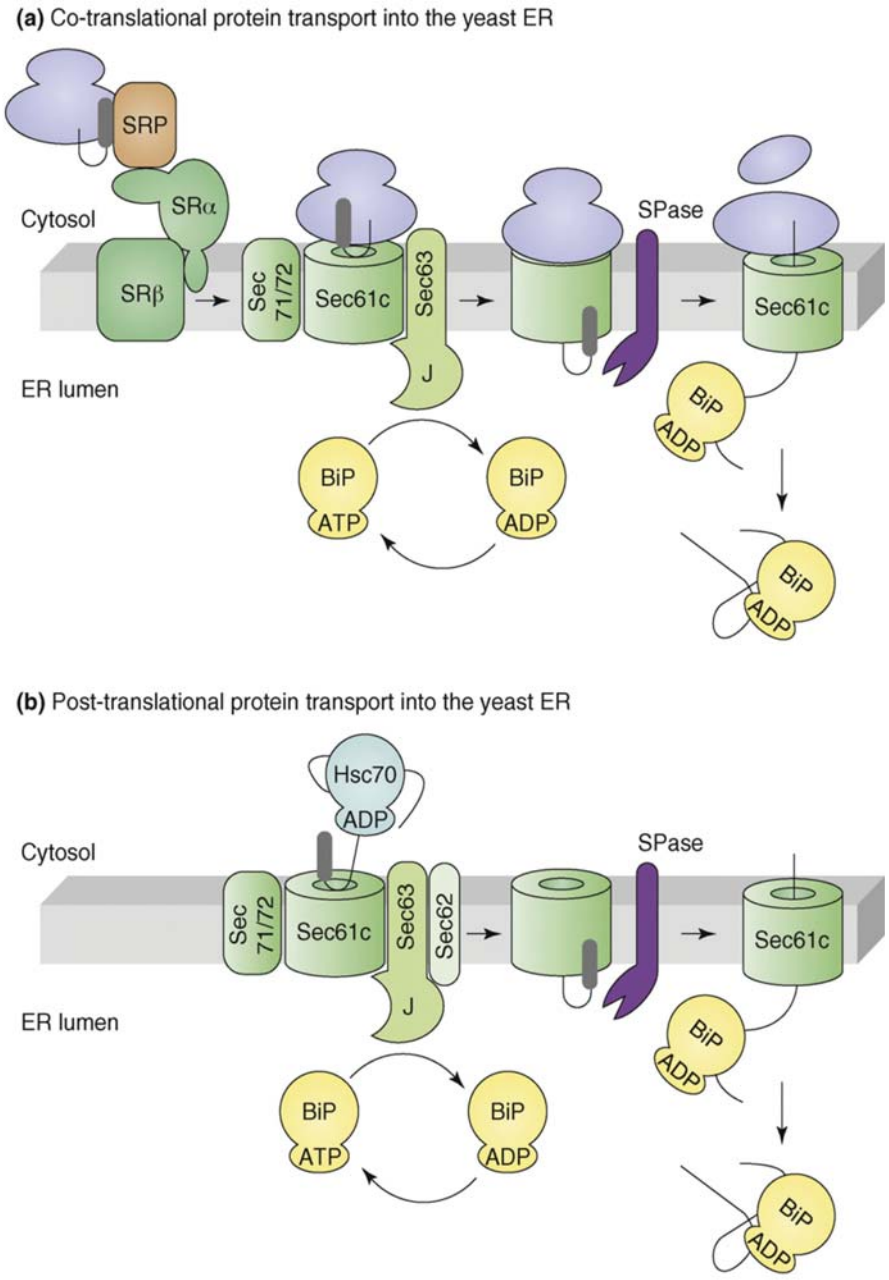


Figure 6.1: *Co- and post-translational translocation into the endoplasmic reticulum.* a) The signal sequence of a protein is recognised by the signal recognition particle (SRP) during translation at the exit site of a ribosome. The complex docks with SRP-receptor (SRPR) and is translocated into the ER lumen/membrane through the Sec61 complex. BiP is recruited to the complex by Sec63 and gates the Sec61 channel (Alder et al, 2005) and serves as a molecular ratchet (Tyedmers, et al, 2003). The signal peptide is removed by signal peptidase unless the peptide is a transmembrane region (signal anchor), in which case no cleavage will take place. b) Recognition of signal sequence in the post-translational translocation pathway is performed by a complex of Sec71 and Sec72 [adapted from (Zimmermann and Blatch, 2009)].

6.2 Cellular fractionation reveals PF02_0040 is an integral membrane protein

PF02_0040 contains a predicted signal anchor region at the N-terminus however the protein lacks any charge in front of the hydrophobic membrane-spanning region therefore it is not obvious how this protein is retained within the membrane. To test whether PF02_0040 is a soluble protein or whether it is a peripheral or integral membrane protein, *P. falciparum* 3D7 schizonts were treated with a high pH carbonate buffer to separate any soluble and peripheral membrane proteins from those with membrane anchors. Western blot analysis of the protein extracts revealed PF02_0040 resides within the carbonate pellet fraction and therefore must be an integral membrane protein, despite the lack of charge preceding the signal anchor region at the N-terminus. SERA-5 is a soluble protein, peripherally associated with the merozoite cell surface and antibodies to this protein were used to control for loading of carbonate supernatant samples. Antibodies recognising membrane anchored MSP-2 control for the carbonate pellet.

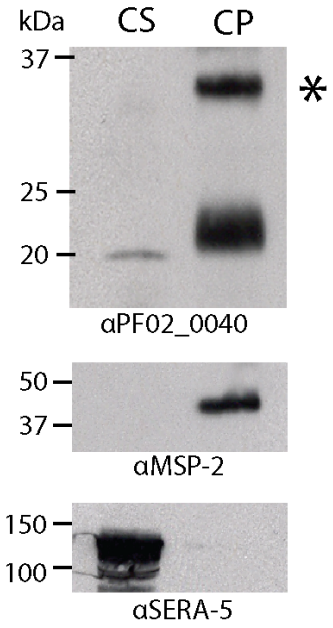


Figure 6.2: *Western blot analysis of carbonate lysed schizonts. P. falciparum* 3D7 schizonts were treated with a high pH carbonate buffer to separate soluble and peripheral membrane proteins from the integral membrane protein fraction. Samples from carbonate supernatant and carbonate pellet were separate by SDS-PAGE and subject to western blot. SERA-5 (peripheral membrane) and MSP-2 (membrane-anchored) antibodies were used as loading controls. The blot was probed with PF02_0040 antibodies to reveal this protein as an integral to the membrane (highlighted with *). The antibodies cross-react with a protein between 20 and 25 kDa.

6.3 PF02_0040 is present throughout the P. falciparum erythrocytic cycle

In the previous chapter it was shown that peak transcript levels of *PF11_0443* occurred around the same time that protein expression was initiated however this may not be the case for *PF02_0040*. Peak transcript levels for this gene are detected at 42h p.i. however the protein has been detected in trophozoites stages (Silvestrini et al., 2010), which are typically observed between 24 and 32h p.i. In order to better understand the protein expression profile of PF02_0040 in the *P. falciparum* blood stage cycle, PF02_0040 polyclonal antibodies were used to probe a western blot of parasite material separated by SDS-PAGE from a highly synchronous culture, sampled every 4 hour from invasion to egress (figure 6.3) (for method please refer to section 2.3.4). Rbcs were also included in the western blot to ensure antibodies were parasite specific and an anti-glycophorin A/B antibody was used to ensure equal loading of cells in each well. The blot was also probed with an antibody against the ER constituent parasite protein BiP to confirm the presence of parasite material in each of the lanes containing time-course samples (Kumar et al., 1988; Kumar et al., 1991). MSP-2, which is expressed in schizonts (Wickham et al., 2003), was used as late stage control. Probing the time-course with the PF02_0040 antibody revealed the protein to be present throughout the lifecycle in a similar pattern to that seen with the ER protein BiP and not just in late stages, as the publish transcript data would suggest, but throughout the erythrocytic cycle. This is in agreement with the prediction of the presence of a Sec66 domain within the protein and the possibility that PF02_0040 is a functional homologue of the yeast prototype from this family.

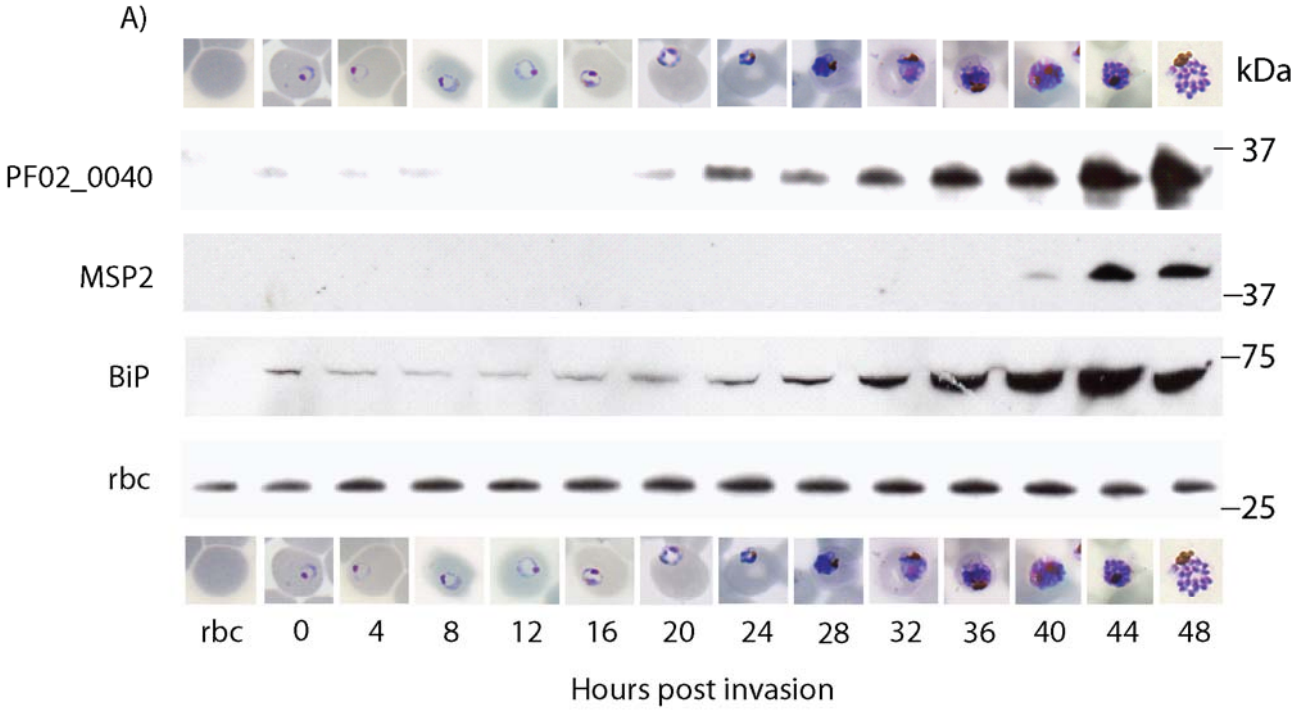
PF02_0040 antibodies were also used in indirect immunofluorescence assays of each time point and a selection of images covering the complete intra-erythrocytic cycle is displayed in figure 6.3 B. Indirect immunofluorescence reveals a peripheral location in early ring stage parasites, 4h pi which looks to be close to the parasite surface. A similar pattern is observed in trophozoites (28h pi) although the ring-like staining is more diffuse. In late schizonts, staining appears to be more tightly associated with the nucleus, a pattern which is often associated with the ER. Co-localisation studies are

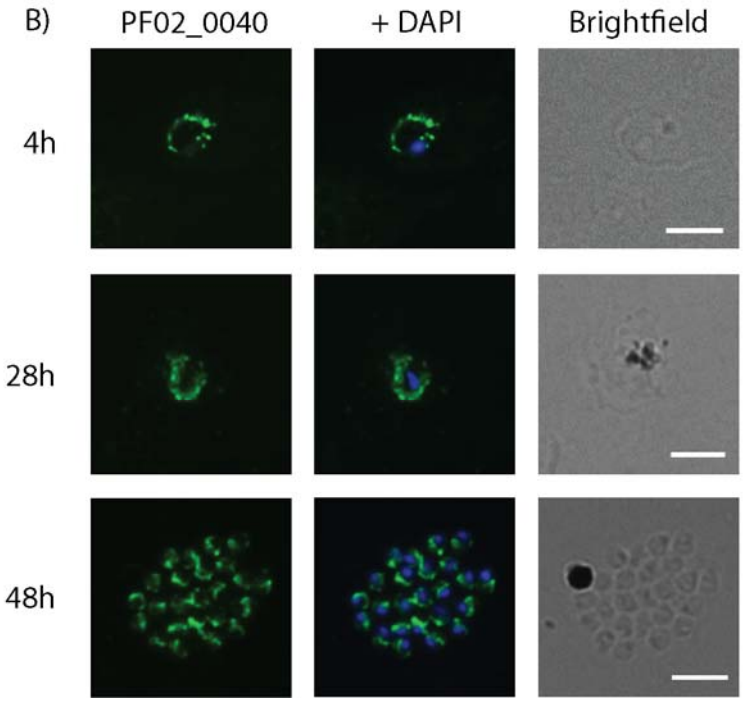
required to rule out other possible localisations surrounding the nucleus such as the IMC or the parasite cell surface.

Figure 6.3: *Expression of PF02_0040 throughout the lifecycle.*

A) Protein samples taken every 4 hours for one complete synchronised blood-stage cycle were hypotonically lysed and separated by SDS-PAGE with $2x10^6$ parasites loaded per well. Hypotonically lysed rbcs were included in the western blot to control for parasite versus rbc specific band recognition. Expression of PF02_0040 was analysed by western blot using the specific polyclonal antibody. A glycophorin A/B antibody was used as a loading control. Antibodies against the parasite ER resident protein BiP were used to confirm the presence of parasites in each lane and a monoclonal antibody against MSP-2 was used as schizont marker. Images of Giemsa-stained parasites from each time point are shown to demonstrate parasite morphology at each stage. Molecular weight markers for each western blot are shown in kDa on the right hand side of the figure.

B) Thin smears of parasite culture from each time point were fixed and permeabilised and used in IFAs to detect the presence of PF02_0040 in the parasite using specific rabbit polyclonal antibodies. PF02_0040 localisation is depicted in green and the nucleus is stained with DAPI and shown in blue. Scale bar on the bright-field image represents 5µm.


A)
kDa
PF02_0040
37
MSP2
37
BiP
75
rbc
25
rbc 0 4 8 12 16 20 24 28 32 36 40 44 48
Hours post invasion


B)
PF02_0040
+ DAPI
Brightfield
4h
28h
48h

6.4 PF02_0040 co-localises with ER resident protein BiP

PF02_0040 contains a predicted Sec66 domain which would indicate that the protein is present within the ER membrane however experimental conformation is required. Localisation was determined by indirect immunofluorescence using polyclonal PF02_0040 rabbit antibodies in comparison to proteins of known organellar location. In ring stage parasites the protein is present around the periphery and is located between the nucleus and the PVM marker RhopH2 (Hiller et al., 2003) (figure 6.4). PF02_0040 partially co-localises with both the 19kDa fragment of parasite surface protein MSP-1 (Holder and Freeman, 1984a) and the ER marker BiP in these newly invaded parasites (Kumar et al., 1988; Kumar et al., 1991) (figure 6.4). It is surprising that the ER should be found around the periphery of the cell; this is perhaps due to the spatial limitation imposed by the concave morphology of the rbc. As the parasite develops within its host MSP-1.19 is translocated from the surface of the plasmalemma to the food vacuole (Dluzewski et al., 2008) whereas the ER appears to remain associated with the nucleus. From figure 6.4, MSP-1.19 is clearly visible in the food vacuole (around the haemozoin) in the earliest stages of schizogony whereas PF02_0040 is associated with the newly divided nuclei. The IMC, visualised using the resident protein GAP45, forms around the periphery of early schizonts (Yeoman et al., 2011; Ridzuan et al., 2012) however PF02_0040 is already associated with individual nuclei and appears to co-localise well with BiP suggesting this protein is to be found within the ER. In late schizonts where cytokinesis occurs and individual merozoite are approaching full maturity, PF02_0040 is seen to reside inside of the plasmalemma, depicted by merozoite surface protein MSP-1 and also inside of the IMC protein GAP45. It does not appear to co-localise with cytoplasmic protein calmodulin but at least partially colocalises with BiP (figure 6.5). This pattern can also be observed in individual merozoites.

These colocalisation experiments reveal a single location for PF11_0443 however the same antibodies used here recognised additional bands between 20 and 25 kDa by western blot, as seen in figure 6.2. It is important that the identity of these bands are

uncovered to ensure the sub-cellular location observed by IFA is that of PF02_0040 and not the other protein/proteins recognised by the polyclonal antibodies.

Indirect immunofluorescence has identified PF02_0040 as an ER resident protein. The nature of this ER retention is not clear since there is no obvious ER retention sequence present at the C-terminus. ER association is in-keeping with the possibility that the protein is an orthologue of yeast Sec71 however further investigation is required to determine the protein's function within the *P. falciparum* intra-erythrocytic cycle.

Figure 6.4: *Indirect immunofluorescence assay of ring and early schizont stage P. falciparum parasites.*

Thin smears of parasite culture 4h post invasion (ring stage) and 32h p.i. (early schizont stage parasites) were fixed and permeabilised as described in 2.2.13.1 and used in IFAs to detect the presence of PF02_0040 in the parasite using specific rabbit polyclonal antibodies. Primary antibodies were detected by a fluorescently labelled secondary antibody (AlexaFluor 488 – green) and the nucleus is stained with DAPI (in blue). Intracellular compartments were identified using previously characterised antibodies to proteins of known location followed by secondary antibodies labelled with AlexaFluor 594 (red) and images were merged to determine comparative locations (MSP-1: cell surface; RhopH2: PV; Bip: ER; GAP45: IMC). Specific antibody targets and dilutions used are listed in table 2.4. Areas of co-localisation appear yellow. Antibodies and dilutions are listed in chapter 2, table 2.4. Scale bar on the bright-field image represents 5µm.

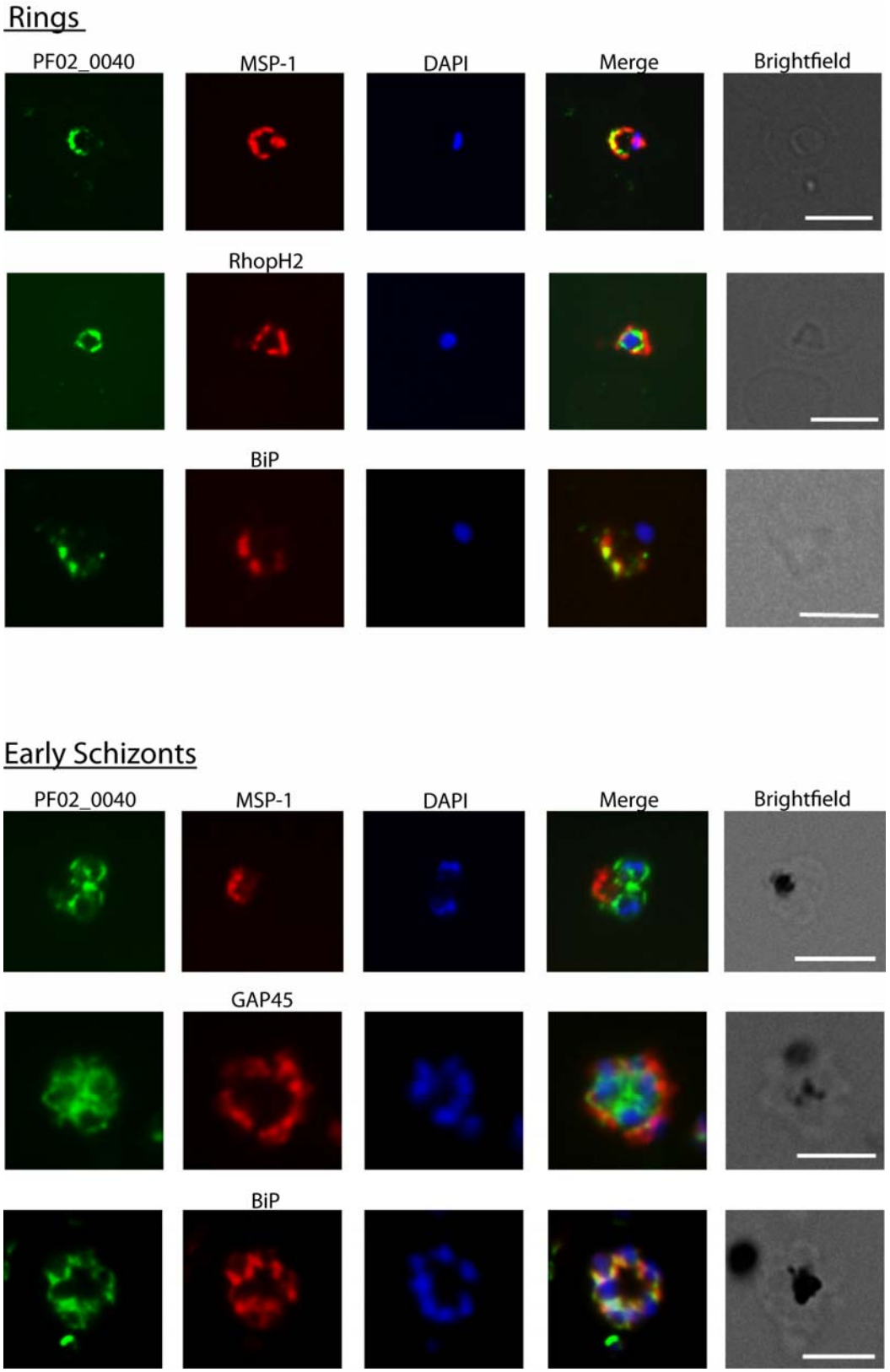
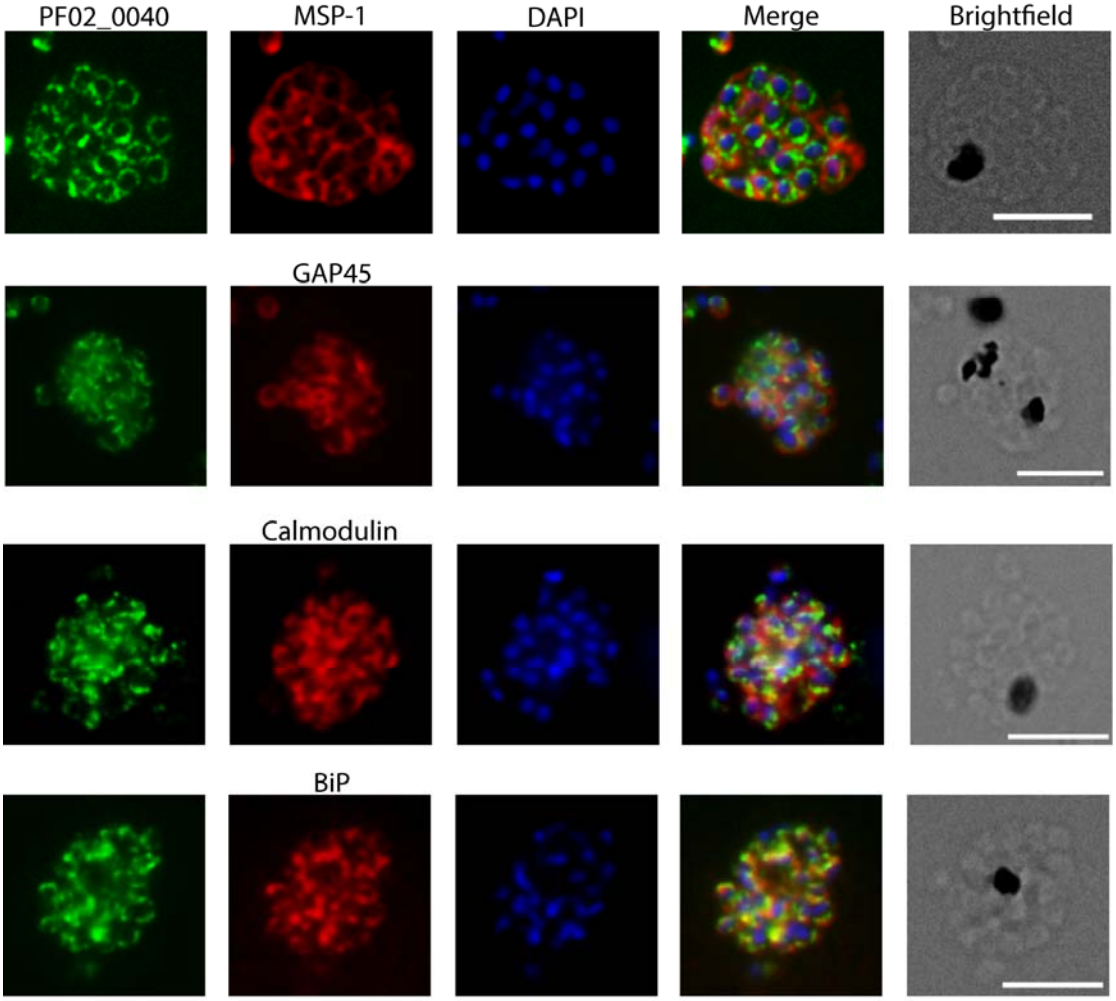

Rings
PF02_0040
MSP-1
DAPI
Merge
Brightfield
RhopH2
BiP
Early Schizonts
PF02_0040
MSP-1
DAPI
Merge
Brightfield
GAP45
BiP

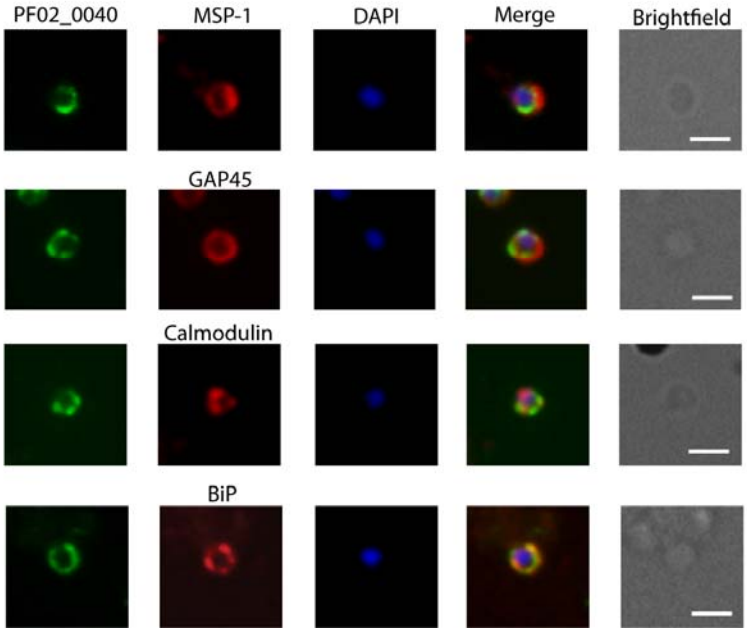
Figure 6.5: *Indirect immunofluorescence assay of late schizonts and merozoites.*

Thin smears of late schizonts and merozoites were fixed and permeabilised as described in 2. 2.13.1 and used in IFAs to detect the presence of PF02_0040 in the parasite using specific rabbit polyclonal antibodies. Primary antibodies were detected by a fluorescent secondary antibody (AlexaFluor 488 – green) and the nucleus is stained with DAPI (in blue). Compartments were identified using previously characterised antibodies to proteins of known location followed by secondary antibodies labelled with AlexaFluor 594 (red) and images were merged to determine comparative locations (MSP-1: surface; GAP45: IMC; Calmodulin: cytoplasm; BiP: ER). Areas of co-localisation appear yellow. Antibodies and dilutions are described in chapter 2, table 2.4. Scale bar on the bright-field image represents 5µm in schizont or 1µm in merozoite images.

Late Schizonts



Merozoites



6.5 Immunoprecipitation experiments using protein-specific polyclonal rabbit antibody pull down PF02_0040 and thioredoxin but no obvious interaction partners

In the quest to determine the function of PF02_0040, it is necessary to find out what other molecules the protein interacts with. If this protein is a true orthologue of Sec71 then it might be expected to bind components of the ER protein translocon such as Sec61/62/63. Initially, PF02_0040-specific antibody was used to immunoprecipitate protein from ^{35}S-labelled schizonts (methods described in chapter 2.2.15). Resolution of this radio-labelled protein sample by SDS-PAGE and subsequent exposure to film revealed these antibodies pulled down a band of around 33kDa which corresponds to the predicted size of PF02_0040 however no other bands were obvious when compared to the pre-immune IgG control IP. The ~20kDa bands observed by western were not visible but this may be due to the time point at which the parasites were labelled - if no new protein was being expressed at around 42h p.i. then the proteins would not be labelled. Also, the proteins would require a number of cysteine or methionine residues in order to incorporate the ^{35}S labelled amino acids.

In order to confirm the identity of the 33kDa band, the ~20kDa bands and possible binding partners, a protein preparation from 3ml pelleted schizonts was passed over a column of PF02_0040 specific IgG bound to activated CNBr sepharose (methods described in section 2.2.16). Unbound proteins were washed off using a 10 mM Tris pH8.0 containing 300mM NaCl before eluting bound proteins in low pH buffer containing 100mM glycine. The elution was concentrated to 50 µl using protein spin concentrators with 10 kDa MW cutoff , of which 30µl was loaded onto a Bis-Tris gel and eluted proteins were separated by SDS-PAGE. Initially the gel was stained with Instant Blue™ however this failed to show up any bands so the gel was stained with silver nitrate using a mass spectrometry-compatible method (Blum, 1987) (figure 6.6 B). The same experiment was performed with early ring stage parasites ~4 h p.i. (figure 6.6 C) Band of approximately 1x4mm were excised and analysed by Lc MS/MS analysis (PNAC,

University of Cambridge). Results from mass spectrometry analysis are summarised in table 6.1, below.

Bands D from the schizont IP and H from ring IP were predominated by the protein of interest, PF02_0040, confirming the specificity of the antibodies. As suspected, thioredoxin (bands a and f) and thioredoxin-related protein (bands b, c, d, e) featured heavily in the eluate, verifying that the ~20kDa contaminating bands are a consequence of antibodies raised against the pET32-Xa/LIC recombinant protein tag. Bands b and d also contained elongation factor 1 alpha which are involved in conveying tRNAs to ribosomes however this protein is a common contaminant in IP experiments in the lab. The ribosome 60S subunit was also precipitated but this is also a common contaminant in IP experiments with *P. falciparum*. Although it may make sense for components of the co-translational translocation pathway to bind translation machinery, it doesn't seem logical that the post-translational pathway components would do the same since their function takes place after the translation has been terminated. Merozoite surface proteins MSP-1 and MSP-9 were also co-precipitated however these proteins are abundant and therefore also abundantly found in pull down experiments involving *P. falciparum* blood stages.

Equivalent bands from the pre-immune column were not initially analysed as it was thought necessary to first determine whether or not the experiment was successful. Although PF02_0040 was detected and contaminant bands seen by western were identified, no other relevant proteins, in particular components of the ER protein translocon, were identified therefore the relevance of preimmune analysis was not justifiable. Unfortunately this experiment failed to identify any binding partners of PF02_0040.

Band	Protein identified from significant peptide hits
a	Thioredoxin 1 (356)
b	EF-1-α (364), thioredoxin related protein (TxrRP) (310), PF02_0040 (278), SERA 5 (221)
c	TxrRP (302), MSP-9 (220)
d	PF02_0040 (876), MSP-1 (548), EF-1-α (471), TxrRP (376), enolase (350), aspartic proteinase (247), PF14_0567 (227), SERA5 (211)
e	TxrRP (346), aldolase (227)
f	Thioredoxin 1 (202)
g	60S ribosomal protein L12 (329)
h	PF02_0040 (383)
i	No significant hits

Table 6.1: *Summary of schizont and ring IP results.* Bands related to those highlighted on the silver-stained gel picture in figure 6.6 B and C. Significant peptide hits are defined as those with a MOWSE score over 200 and with 2 or more peptide hits per protein per band. The individual scores are shown in brackets in blue of each protein.

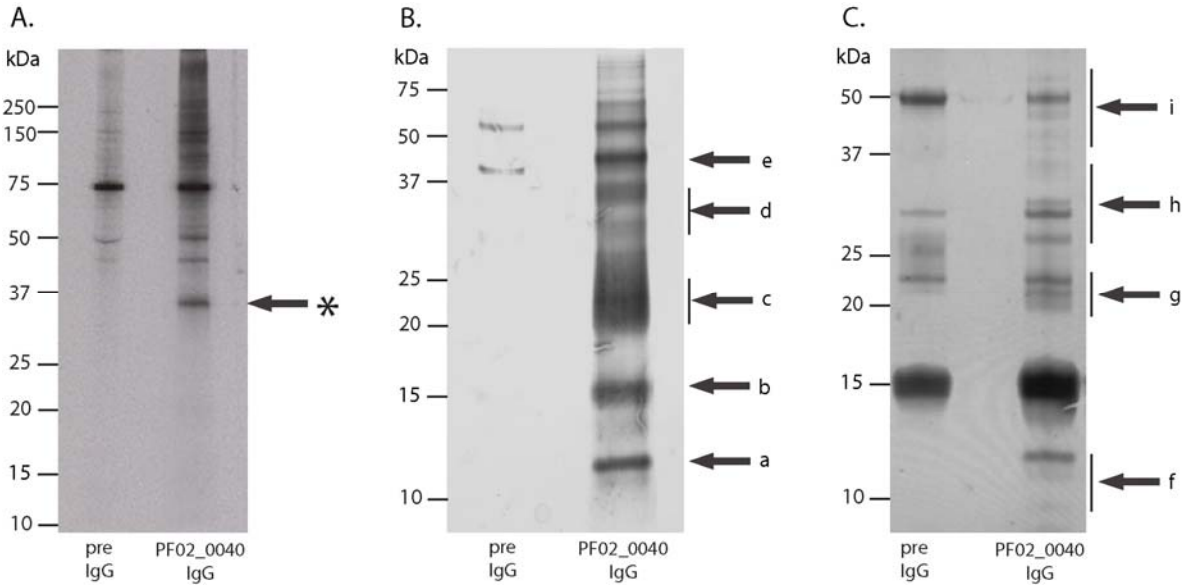


Figure 6.6: *Immunoprecipitation studies using PF02_0040 antibodies.* A) PF02_0040 antibodies were used to immuno-precipitate interacting proteins from 50 µl of pelleted ^{35}S labelled schizonts (methods described in chapter 2.2.15). Co-precipitating proteins were separated by SDS-PAGE and exposed to film. A band of the predicted size of PF02_0040, approximately 33kDa, was observed and highlighted with *. Purified schizonts (B) and synchronised early rings of ~10% parasitaemia (C) were lysed and passed over a column of 10mg of pre-immune rabbit IgG follwed by that containing affinity purified anti-PF02_0040 rabbit polyclonal antibodies conjugated to activated CNBr sepharose (methods described in chapter 2.2.16). Co-precipitating proteins were resolved by SDS-PAGE and visualised by silver stain. Unique bands compared to the control were excised (labelled a-f) and the protein complement of each was identified by Lc-MS/MS.

6.6 Epitope tagging of the PF11_0443 locus by targeted homologous recombination

As discussed in chapter 5.7, incorporating an epitope tag into the genomic locus of a gene is a useful tool for characterisation of the resulting protein. Due to the limitations of immunoprecipitation seen with polyclonal antibodies, incorporating an epitope tag into the *PF02_0040* locus may increase the experimental capacity of the protein. As with PF11_0443, the smaller epitope tags 3xFLAG and TY-1 tags were selected to incorporate into the *PF02_0040* locus. Again, these tags are to be introduced into the C-terminus since the protein contains a signal sequence.

PF02_0040 is comprised of 5 exons and 4 introns concentrated around the 3' end of the gene. Since these introns contain A/T rich, repetitive sequences, including them in the region of homology could allow integration into any number of A/T rich, repetitive sequences in the genome. For this reason, the decision was taken to omit these sequences from the region of homology and amplify a region of 881 bp from cDNA which included the entire gene, minus the stop codon and with the addition of a region of the 5' UTR so that the largest consecutive region of homology was 436 bp. The sequence was amplified by PCR using primer combinations W/X and W/Y (sequence detailed in chapter 2, table 2.1). Ty1 and FLAG epitope tag sequences (including a stop codon at the 3' end) were included in reverse primers X and Y respectively, thereby incorporating the tags at the 3' end of the *PF02_0040* exon sequence. This region of homology plus tag sequence was cloned into the multiple cloning site of the pHH3 vector via the *Eco*RI/*Sac*II restriction sites generating pHH3_PF02_0040_Ty1 and pHH3_PF02_0040_3xFLAG. An *Avr*II site was included between the gene and tag sequences to allow the use of these vectors to epitope tag other genes. The pHH3 transfection vector which includes the Blasticidin resistance cassette as a selectable marker is depicted in appendix D. The same region of homology of PF02_0040 was in addition cloned into the pHH3 vector via EcoRI/SacII without the presence of an epitope

tag to serve as a 3' replacement vector which upon successful integration would swap the endogenous 3'UTR with hrp2 3'UTR.

Inserts were initially sub-cloned into the pGEM T-easy vector and transformed into *E. coli* TOP10 cells. Colonies were selected and sequenced before excising the inserts of correct sequence using *Eco*RI and *Sac*II. Inserts were gel-purified and ligated into the pHH3 vector and transformed into *E. coli* TOP10. Colonies were re-sequenced and one correct DNA sequence was propagated and plasmid DNA purified for transfection. For analysis purposes the construct was subject to digestion with *Eco*RI and *Sac*II to confirm the presence of the correct sized (881 bp) insert prior to transfection (figure 7.6 B)

Unfortunately due to complications in preparation and time constraints, the pHH3_PF02_0040_3' construct were never completed or transfected. Complete and purified plasmids pHH3_PF02_0040_3xFLAG, and pHH3_PF02_0040TY were transfected into ring-stage parasites and selected for presence of episome by 2.5µg/ml of Blasticidin applied to cultures. Stock sample of these cultures were cryo-preserved and genomic DNA was prepared. Transfection cultures were taken off drug for 3 weeks before re-application of drug which is required for enrichment of the parasite population that contains integrated versus episomal copies of the transfection vector. This was repeated for 5 cycles. A diagnostic PCR was designed to test for integration events using a forward primer from the *PF02_0040* 5'UTR region (primer DD, table 2.1) and reverse primers from either the *PF02_0040* 3'UTR region (primer EE, table 2.1) to identify the wild type locus or from the *HRP2* 3'UTR contained within the pHH3 constructs (primer NN, table 2.1) to detect integrated plasmid (figure 6.7 A). As shown in Figure 6.7 C, this diagnostic PCR revealed no integration of Ty1 or 3xFLAG epitope tag after 5 cycles however western blot of Ty1-tagged versus WT PF02_0040 containing parasites revealed Ty1 antibodies detected a band of the correct size of 37kDa in Ty1-tagged but not the WT 3D7 line (figure 6.7 D). This discrepancy is perhaps due to an incompatibility of the locus and episome derived primers DD and NN (table 2.1). Further analysis by Southern blot analysis showed no integration into the *PF02_0040* locus (figure 6.8). This would imply that integration did not occur and PF02_0040 should not contain a Ty1 tag at its C-

terminus. If this is true, the specific 37kDa band observed by western blot would have been brought about by integration into another locus, which expresses a protein of similar molecular weight. The chance of this happening seems very small indeed. Alternatively, there is a small possibility the protein plus tag is being expressed from the episome, however since there is no 3' UTR contained within the construct this is unlikely, but not impossible. It would be impossible if there was no promoter region however in order to increase the number of base pairs in the region of homology, 56 bp of the promoter region immediately upstream of the ATG start codon was included in the region of homology. It is therefore possible (although unlikely) that expression is episomal. The fact that PF02_0040_3xFLAG shows no detectable expression by western blot (figure 6.7 E) despite detection with a very sensitive antibody (Chapter 5) makes the argument of episomal expression even more unlikely.

The FLAG M2 monoclonal antibody did not detect 3xFLAG protein by western blot (figure 6.7 E) which is in agreement with the diagnostic PCR (figure 6.7 C) and Southern blot (data not shown).

Integration of Ty1 into *PF02_0040* would demonstrate that the locus is amenable to genetic manipulation however conflicting evidence from DNA and protein analysis leaves the question of ability to integrate unanswered.

Figure 6.7: *Epitope tagging of the* PF02_0040 *locus.* A 881 bp region of homology of *PF02_0040* was amplified using primers W/X and W/Y respectively (chapter 2, table 2.1) introduced into the pHH3 transfection vector along with TY-1 and 3xFLAG epitope tag via *Eco*RI and *Sac*II restriction sites (A). The pHH3 vector contains the Blasticidin resistance cassette and a copy of the *P. falciparum hrp2* gene 3' untranslated region (UTR). Restriction digest of the complete construct (6287bp) using *Eco*RI and *Sac*II shows the 881bp insert and 5406bp vector backbone (B). The plasmid was transfected into *P. falciparum* 3D7 and drug cycled to promote loss of episomal copies of the construct. A diagnostic PCR was designed to test for integration events using a forward primer from the *PF02_0040* 5'UTR region (a: primer DD, table 2.1) and reverse primers from either the *PF02_0040* 3'UTR region (b: primer EE, table 2.1) to identify the wild type (WT) locus or from the *HRP2* 3'UTR from the pHH3_PF02_0040_Ty1 or 3xFLAG constructs (c: primer NN, table 2.1) to detect integrated plasmid. PCR revealed the presence of WT DNA only and a faint band of correct size (bp) for Ty1 crossover into the PF02_0040 locus (C). Protein samples from tagged lines as well as WT parasites were resolved by SDS-PAGE and subject to western blot analysis. Blots were probed with PF02_0040 rabbit polyclonal antibodies and either Ty1 (D) or FLAG antibodies (E). Ty1 antibodies recognise a band of 33 kDa in the tagged line but not WT. FLAG antibodies failed to detect the expression of 3xFLAG in the line containing pHH3_PF02_0040_3xFLAG.

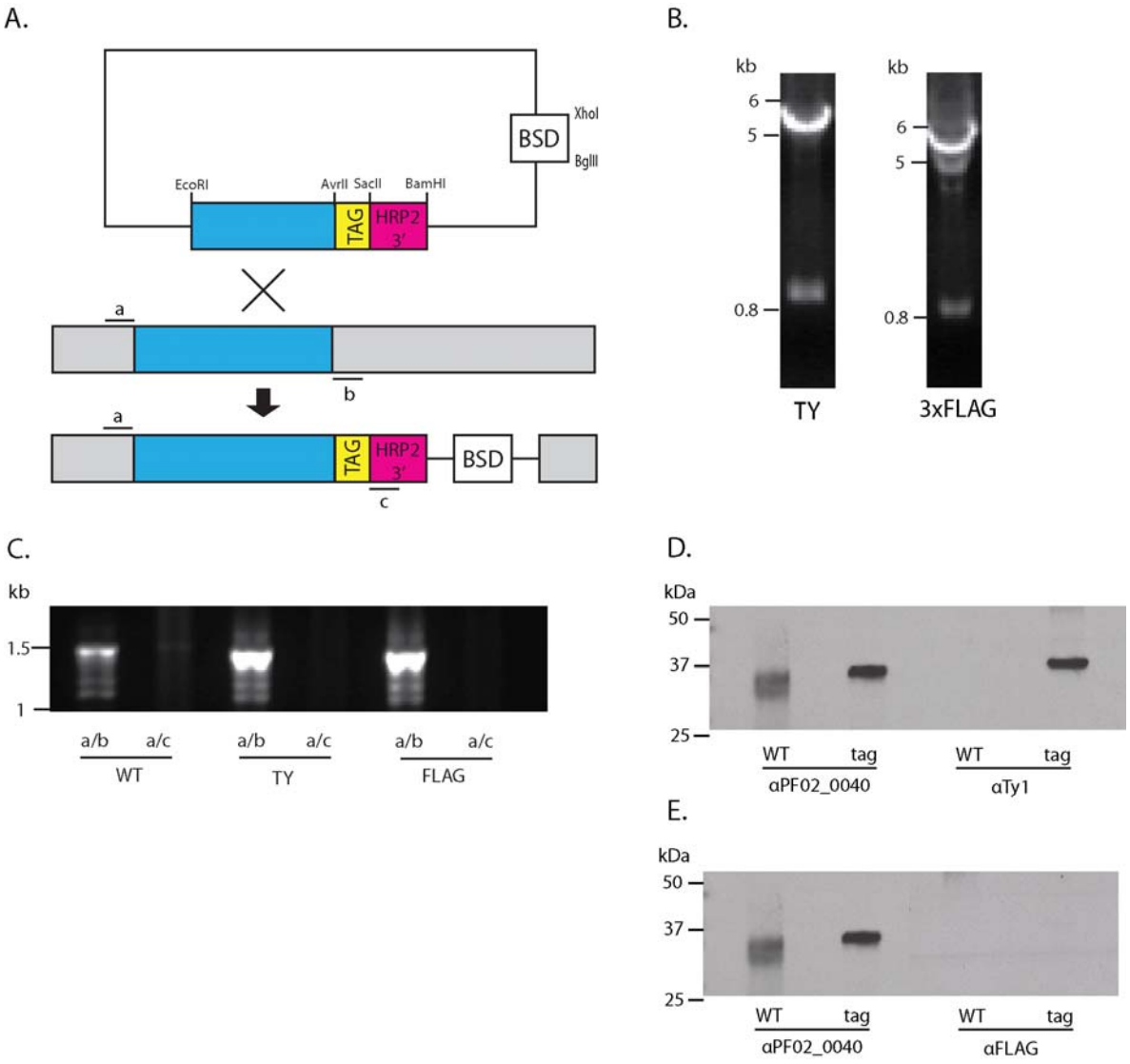
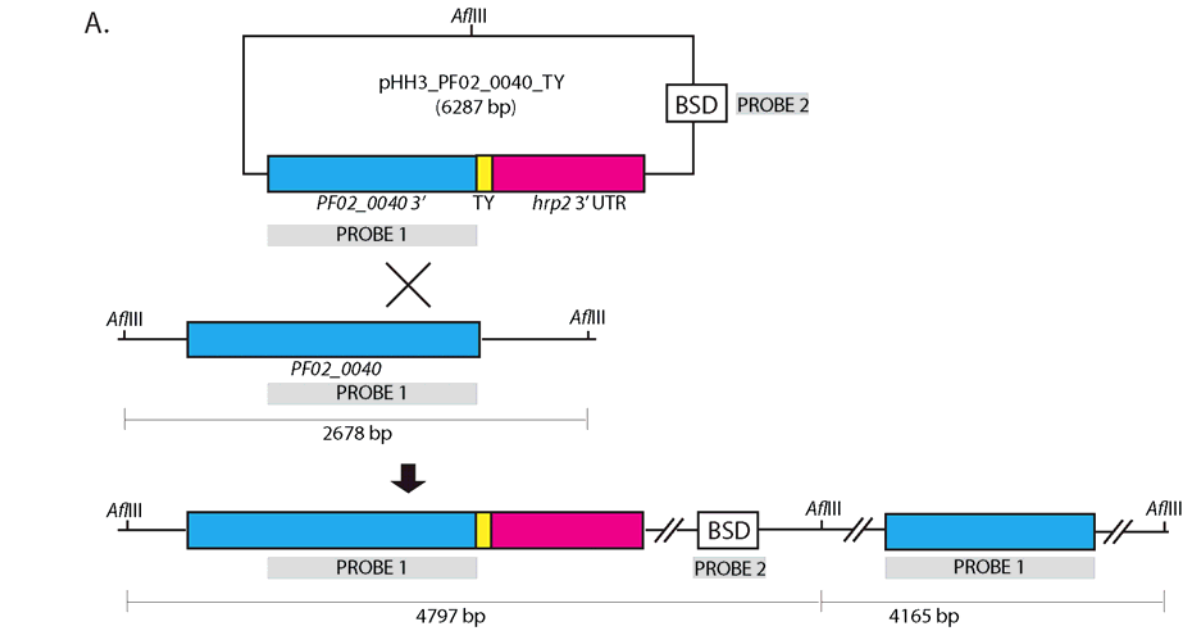
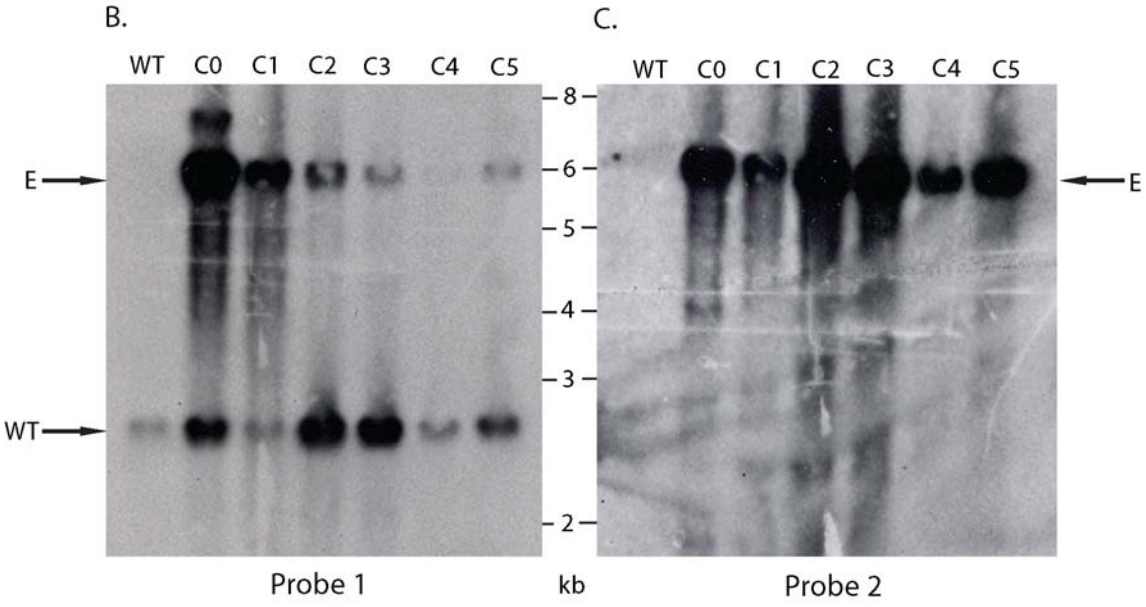

A.
BSD
XhoI
BglII
EcoRI
AvrII SacII
BamHI
TAG
HRP2
3′
a
b
c
B.
kb
6
5
0.8
TY
3xFLAG
C.
kb
1.5
1
a/b
a/c
WT
TY
FLAG
D.
kDa
50
37
25
WT
tag
αPF02_0040
αTy1
E.
αFLAG

Figure 6.8: *Southern blot of attempted 3′ epitope tagging of* PF02_0040 *with Ty1*.
A Southern blot was designed to test for integration of the pHH3_PF02_0040_Ty1 construct in to the *PF02_0040* locus (A). Genomic DNA taken each drug cycle was digested by *Afl*III before separating resulting fragments by gel electrophoresis and blotting onto nitrocellulose membrane. The membrane was probed with a radiolabelled oligomer of either (B) the region of homology (probe 1) or (C) the blasticidin ORF of the pHH3_PF02_0040_Ty1 construct (probe 2). The resulting banding pattern from exposure to film revealed the presence of wild type with a band of 2678 bp (WT) and a band just above the 6 kb markers corresponding to the 6287 bp episome (E). These bands persist throughout the 5 drug cycles and suggest no integration into the *PF02_0040* has taken place.


A.
AflIII
pHH3_PF02_0040_TY
(6287 bp)
BSD
PROBE 2
PF02_0040 3′
TY
hrp2 3′ UTR
PROBE 1
AflIII
AflIII
PF02_0040
PROBE 1
2678 bp
AflIII
AflIII
AflIII
BSD
PROBE 1
PROBE 2
PROBE 1
4797 bp
4165 bp


B.
WT C0 C1 C2 C3 C4 C5
E
WT
8
6
5
4
3
2
kb
Probe 1
C.
WT C0 C1 C2 C3 C4 C5
E
Probe 2

6.7 Unsuccessful attempts to knock out PF02_0040 by truncation suggest an essential role for this gene within blood stages

Sec66 is not essential in yeast growing at 22°C but is when growth temperature is increased to 37°C (Feldheim et al., 1993; Fang and Green, 1994). Based on this one might expect to be able to knock-out *PF02_0040*. However the transfer from insect to mammalian host already imposes such a temperature change. Indeed the protein has not been detected in the mosquito stages of the *Plasmodium* lifecycle (Lasonder et al., 2008). Alternatively a knock out may be viable ay 37°C but not when growth temperature is increased as would be experienced during febrile episodes in malaria patients.

To test these hypotheses, a construct was designed to integrate into the 5′ end of the *PF02_0040* which, if successful, would express a truncated version of the protein with 149 of the 276 amino acids removed from the C-terminus. A complete or partial reduction of expression can be expected due to the lack of a 3′ UTR and terminator sequence that would be imposed upon integration of the plasmid (see figure 6.9 for KO strategy). Unfortunately, due to the AT rich nature of the 5′ and 3′ UTRs and the presence of 4 introns in the 3′ end of the gene, a full knock out of *PF02_0040* by double homologous recombination was not possible. It was hoped however that successful integration would disrupt function and either exhibit a phenotype, or in the case that *PF02_0040* is an essential gene, result in non-viable parasites and therefore integration would not be possible.

The integration transfection vector pHH3_PF02_0040_KO was manipulated to include 486 bp from the 5′ UTR and ORF of *PF02_0040*, amplified by PCR using primers U and V (sequences listed in chapter 2, table 2.1) and introduced to the construct's multiple cloning site using the enzymes *Eco*RI and *Bam*HI. The construct contains the blasticidin resistance cassette to be used as a selectable marker, conferring blasticidin resistance to successfully transfected parasites. Inserts were initially sub-cloned into the pGEM T-easy

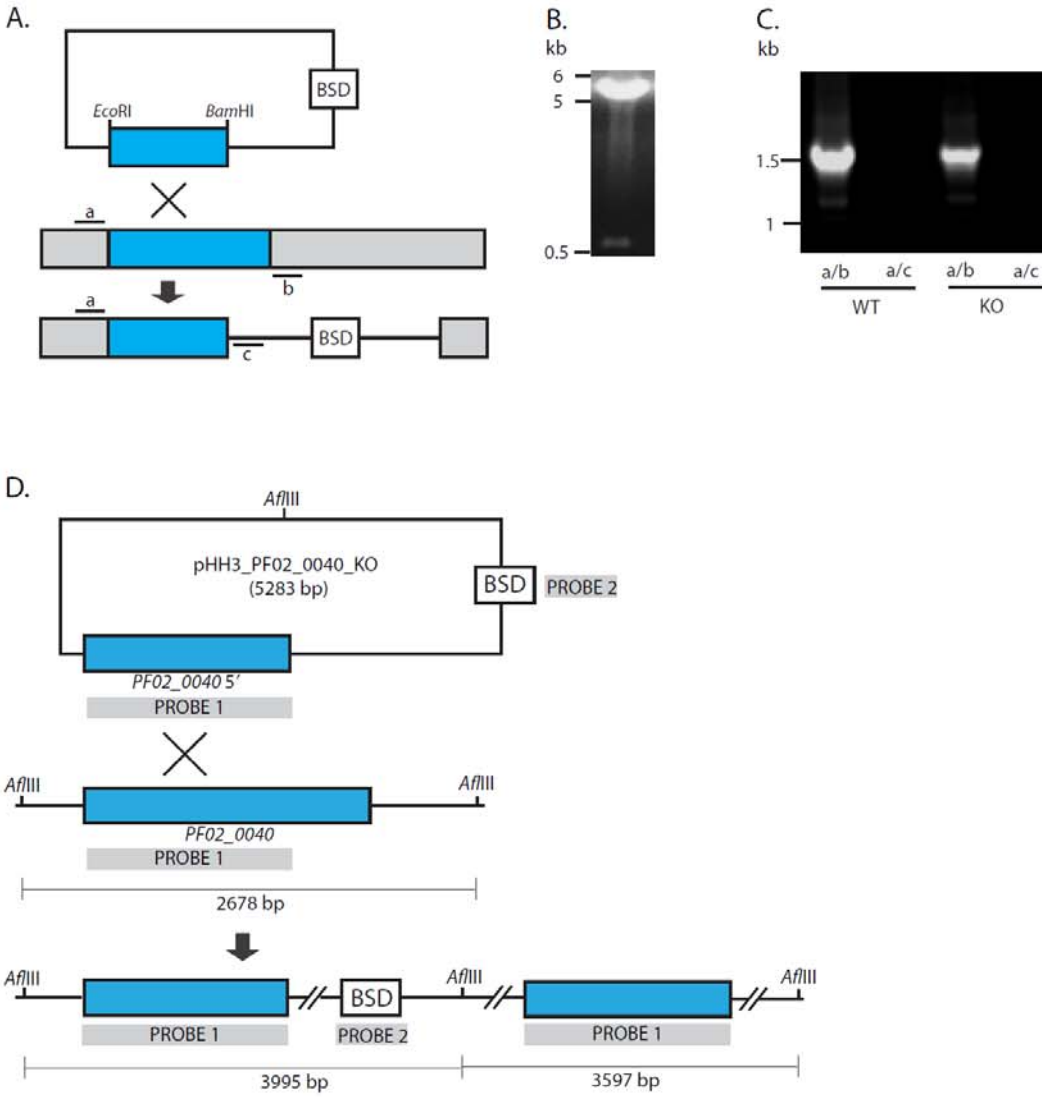
vector and transformed into *E. coli* TOP10 cells. Colonies were selected and sequenced before excising the correct sequence inserts using *Eco*RI and *Bam*HI. Inserts were gel-purified and ligated into the pHH3 vector and transformed into *E. coli* TOP10. Colony plasmid DNA was screened by restriction digest using *Eco*RI and *Bam*HI (figure 6.9 B) and re-sequenced before selecting one correctly sequenced vector to propagate and purify for transfection.

As with the epitope tag constructs, complete and purified pHH3_Pf02_0040_KO was transfected into ring-stage parasites. Selection for presence of episome was achieved by the addition of 2.5µg/ml of Blasticidin to transfection cultures until parasitaemia was recovered to around 10%. Stock sample of these cultures were taken and stored in liquid nitrogen and genomic DNA was prepared in order to test for presence of integrants. Transfection cultures were taken off drug for 3 weeks before its re-application which was repeated until integration occurred or for 5 cycles in the absence of any detectable integration. A diagnostic PCR was designed to test for integration events using a forward primer from the *PF02_0040* 5'UTR region (a: primer DD, table 2.1) and reverse primers from either the *PF02_0040* 3'UTR region (b: primer EE, table 2.1) to identify the wild type locus or from the *hsp86* promoter sequence of the pHH3_PF02_0040_KO vector back-bone (c: primer OO) to detect integrated plasmid (figure 6.9 A). No integration was detected by this method within 5 drug cycles (figure 6.9 C). A Southern blot was designed to confirm this negative result. The genomic DNA prepared after each cycle was subject to digest by restriction enzyme *Afl*III before separating the resulting fragments by gel electrophoresis and blotting onto a nitrocellulose membrane. In order to ascertain whether integration had occurred, the membrane was probed with a radiolabelled oligomer of either the region of homology (probe 1) or the blasticidin ORF of the pHH3_PF02_0040_KO construct (probe 2). The resulting banding pattern from exposure to film revealed the presence of wild type with a band of 2678 bp (figure 6.9 E) and a band between the 5 and 6 kb markers corresponding to the 5283 bp episome (figure 6.9 E+F) throughout the 5 cycles. In accordance with the diagnostic PCR (figure 6.9 C), no bands of 3995 or 3597 bp corresponding to an integration event were detected. Instead, the 5.2kDa episome band disappears in cycles 4 and 5 when probed

with probe 1 and is replaced by 2 bands of approximately 2.9 and 6.5 kDa, seen using probe 2 suggesting the episome has integrated elsewhere in the genome. This attempted KO by truncation was repeated in triplicate however none of the attempts resulted in integration into the *PF02_0040* genomic locus (data not shown).

The data presented here suggests that *PF02_0040* cannot be knocked-out and by implication, is an essential gene. Caution must be taken in drawing this conclusion. The region of homology of 486 bp is relatively short and contains a component of AT rich 5'UTR sequence therefore the possibility that the lack of integration is due to this technical difficulty in homologous recombination cannot be ruled out.

Figure 6.9: *Attempted KO of* PF02_0040 *by truncation.* A) A single cross-over strategy was employed for the KO of *PF02_0040* which would result in a truncated version of the gene with no 3'UTR. Primers a, b and c (listed in table 2.1 as primers DD, EE and OO) were designed to detect the integrated and wild type (WT) locus. The transfection vector pHH3 was manipulated to include 486 bp from the 5' UTR and ORF of *PF02_0040*, amplified by PCR using primers U and V (sequence listed in chapter 2, table 2.1) and introduced to the construct's multiple cloning site using the enzymes *Eco*RI and *Bam*HI. Digesion with these enzymes from the completed plasmid excised a band of 881 bp (B). After transfection of *P. falciparum* 3D7, the positive parasites were cycled on/off blasticidin for 5 cycles, and the diagnostic PCR was used to test for integration events using primer combinations a and b (detection of WT) or a and c (to detect integration of plasmid). These PCRs from WT and transfected 3D7 gDNA revealed the presence of the WT locus but not of any integration events (C). A Southern blot was designed to confirm this result (D). Genomic DNA taken each drug cycle was digested by *Afl*III before separating resulting fragments by gel electrophoresis and blotting onto nitrocellulose membrane. The membrane was probed with a radiolabelled oligomer of either (E) the region of homology (probe 1) or (F) the blasticidin ORF of the pHH3_PF02_0040_KO construct (probe 2). The resulting banding pattern from exposure to film revealed the presence of wild type with a band of 2678 bp (WT) throughout the 5 cycles and a band between 5 and 6 kb markers corresponding to the 5283 bp episome (E) in the first three drug cycles. The two bands labelled X seen when probing the digested DNA with the blasticidin probe illustrate an integration event but do not correspond to the expected size of integration into the *PF02_0040* locus.


A.
EcoRI
BamHI
BSD
a
b
c
B.
kb
6
5
0.5
C.
kb
1.5
1
a/b
a/c
a/b
a/c
WT
KO
D.
AflIII
pHH3_PF02_0040_KO
(5283 bp)
BSD
PROBE 2
PF02_0040 5'
PROBE 1
PF02_0040
2678 bp
3995 bp
3597 bp
E.
Probe 1.
kb
8
6
5
4
3
2
E
WT
WT
C1
C2
C3
C4
C5
F.
Probe 2.
X
E
X

6.8 Discussion

PF02_0040 is a putative orthologue of the ER protein translocon component Sec71, an integral membrane protein of the ER in yeast (Feldheim et al., 1993; Fang and Green, 1994), and it appears that the *P. falciparum* protein is no different owing to its presence in the pellet of parasites lysed in a high pH carbonate buffer (figure 6.2). PF02_0040 contains a predicted signal anchor region at the N-terminus however the protein lacks any charge in front of the hydrophobic membrane-spanning region therefore it was not clear if this region alone is sufficient to sustain membrane integration. Since treatment with a high pH carbonate buffer should remove any peripheral membrane proteins, this result confirms that PF02_0040 is an integral membrane protein. TheSec71 proteins of *Saccharomyces* and *Trypanosoma* also contain a signal anchor domain but this is approximately 30 aa downstream of the N-terminus with clear charged residues either side of the membrane spanning region (Goldshmidt et al., 2008). Although one might assume that to fulfil the same role as the prototype Sec71, the *Plasmodium* orthologue should also reside within the membrane however the putative *Toxoplasma* orthologue of Sec66 does not contain a predicted membrane spanning region which could indicate a diverse role of the protein between genera. Not all organisms appear to require Sec71 to function and indeed *Homo sapiens* do not possess the *SEC71* gene.

The ER is a fundamental part of the eukaryotic cell architecture and is certainly present throughout the blood stage of the lifecycle of *Plasmodium*. However from earlier transcription data, PF02_0040 appeared to be upregulated in late blood stages (Bozdech et al., 2003). If this protein is an ER resident then one might expect its presence throughout the lifecycle. Western blot analysis of parasite protein samples taken at regular intervals throughout the lifecycle revealed that PF02_0040 is in fact present throughout the blood stages, with a similar pattern of expression to the ER protein BiP (figure 6.3 A). This finding is now backed up by recent proteomic data, with peptide hits from ring (Oehring & Woodcroft et al., unpublished – available on plasmodb.org), trophozoite and even stage V gametocytes (Silvestrini et al., 2010), as well as schizonts and merozoites (Florens et al., 2002; Silvestrini et al., 2010; Bowyer et al., 2011; Treeck et

al., 2011). The protein could also be detected by IFA at ring, trophozoite and schizont stages as shown in this thesis (figure 6.3 B). PF02_0040 must be expressed de novo in late stages since metabolic labelling in schizonts at 42 h p.i. resulted in a band of the protein's size (33 kDa). This may simply be due to the need for ER expansion ready for segregation of developing daughter cells but also may be due to the requirement for specific recognition of certain proteins to facilitate their export via different route.

To identify the protein's subcellular location, rabbit polyclonal antibodies were used to colocalise PF02_0040 with known protein markers of parasite organelles by IFA. Like Sec71, PF02_0040 was to be found in the ER throughout the lifecycle, as determined by colocalisation with ER resident Hsp70, BiP. The partial colocalisation between GAP45 and PF02_0040 in merozoites is likely to be a product of the close proximity of all compartments in a cell that is approx. 1µm in diameter at its longest point (from anterior to posterior). PF02_0040 also co-localises with MSP-1.19 in rings however it is known that MSP-1 is internalised at some point during ring development (Dluzewski et al., 2008), so this may represent the beginning of its transit. Also, the cellular architecture of ring stage parasites is a little unclear. One model ("the Big Gulp"), based on EM data, predicts that the parasite takes on a cup formation upon entry into the erythrocyte before sealing at the open end to enclose a vacuole that will be later filled with the bi-product of haeme degradation (Elliott et al., 2008). This explains why a 'ring' is observed in early trophozoites before intensive feeding and DNA replication brings about the morphological change indicative of the late trophozoite, however why the ER then moves from a peri-nuclear location to one that extends to the ring periphery is unclear.

Comparison to BiP by IFA did suggest that PF02_0040 is an ER resident however the antibodies used to come to this conclusion clearly recognise 2 bands between 20 and 25 kDa, one of which is membrane associated, as shown in figure 6.2. From LcMS/MS analysis these bands were identified as thioredoxin 1, a soluble protein found in the cytoplasm (Kanzok et al., 2000) and thioredoxin related protein (PF13_0272), a putative 2 transmembrane containing protein of unknown sub-cellular location (plasmodb.org). No colocalisation was seen with cytoplasmic protein calmodulin ruling out any cross-

reactivity with thioredoxin 1 however thioredoxin related protein is predicted to reside within a membrane bound compartment which may be the ER. Even if this is the case, the antibodies used in IFA appear to only label the ER, indicating both proteins are present in this organelle. It is also possible that these proteins are only recognised under reducing conditions when linear epitopes are exposed, in which case what we see by IFA is the location of PF02_0040 only.

PF02_0040 appears to reside within the ER however, upon analysis of the primary amino acid sequence the nature of this ER retention is unclear. The secretory pathway is continuous between the ER, Golgi and secretory vesicles, transporting cargo to organelles and the cell's exterior; therefore proteins of the ER membrane will be lost in transit if not for a signal that directs these proteins back to the ER from the Golgi complex. Soluble ER proteins are brought back to the ER via recognition of a C-terminal KDEL motif in mammals (Munro and Pelham, 1987) or HDEL motif in yeast (Pelham et al., 1988) by the protein ERD2 (Lewis et al., 1990). *Plasmodium* contains an orthologue of this retrieval protein (Elmendorf and Haldar, 1993), although it appears only the 'DEL' is important for recognition (Kumar et al., 1991; Kumar and Zheng, 1992) . A C-terminal di-lysine motif, KKXX and KXKXX, has been shown to be involved in ER retrieval of type-I integral membrane proteins (Nilsson et al., 1989; Jackson et al., 1993; Gaynor et al., 1994) and that this retrieval is mediated by the coatomer complex (Cosson and Letourneur, 1994; Letourneur et al., 1994). PF02_0040 contains a version of this retrieval sequence (KKSK) however the *P. vivax* and *P. knowlesi* homologues have no such motif. Surprisingly, neither does the yeast or trypanosome Sec71 despite evidence suggesting Sec71 to continuously cycled between the ER and Golgi and that this process is dependant on the coatamer complex (Sato et al., 1997). The *Saccharomyces* protein instead contains a δ-COP1-binding retrieval sequence (δL) from residues 147-166, 43 amino acids from the C-terminus for the protein which is characterised by the presence of acidic residues followed by an essential aromatic residue (Cosson et al., 1998). This region appears relatively conserved in orthologues in *T. brucei* and *N. crassa* (Goldshmidt et al., 2008). PF02_0040 contains a potential sequence of this nature (EEFLEEFNFILHEANCLSDKW) however experimental evidence is lacking and out-with the

scope of this study. The transmembrane domain of ER proteins has also been implicated in retrieval from the Golgi, as is the case for ER resident Sec21 (Sato et al., 1996) and cytochrome b5, which requires a transmembrane domain of less than 22 aa for retrieval from the Golgi as well an as yet undefined luminal motif for retention within the ER (Honsho et al., 1998).

So it seems there are a number of ER retrieval/retention sequences and to uncover which is the functional one in the case of PF02_0040 is no trivial undertaking. In order to investigate whether the signal anchor region of PF02_0040 is responsible for ER retention, this region was amplified and incorporated into the pHH3_MSP3prom_GFPstop episomal expression construct which would allow the expression of the PF02_0040 signal anchor fused to soluble GFP. This would serve a dual purpose in establishing if the region is sufficient for membrane anchorage and in determining if this region sufficient for ER retention or if and additional/alternative retention sequence is required. Although this vector was produced and transfected, data is currently unavailable.

If the Ty1 tag had successfully integrated and was expressed then it is possible to rule out a retrieval signal which requires localisation directly at the C-terminus of the protein. Unfortunately, genetic evidence suggests integration did not take place therefore one might suggest that the Ty-1 C-terminal tag was not a stable parasites line, although further investigation is required.

Although location was identified with reasonable certainty by IFA, PF02_0040's binding partners were not successfully identified using specific rabbit polyclonal antibodies. A lack of sensitivity of the antibodies combined with a potentially weak interaction with other members of the translocon components and a weak and temporal interaction with proteins to be imported post-translationally resulted in only abundant proteins such as MSPs and SERA5, or 'sticky' proteins (common contaminants) such as EF-1-α and ribosomal subunits to be identified by affinity chromatography followed by LcMS/MS analysis. Apart from one peptide and an insignificant hit from Sec61 beta subunit (which

Sec71 has not been documented to bind), none of the translocon components were co-precipitated with PF02_0040. There is a possibility that Triton X-100 is not a strong enough detergent to remove this complex from its tight membrane association therefore the rest of the complex (other than PF02_0040) is not being solubilised. Moreover, if the detergent is strong enough to remove the translocon complex from the membrane, it may also be strong enough to disrupt the interaction between individual translocon components. Unfortunately, all yeast IP experiments involving the translocon complex use this detergent in their lysis buffer. The translocon complex in yeast was precipitated from purified ER microsomes, a difficult technique that has not been achieved in *Plasmodium*. In an additional attempt to precipitate the complex from *P. falciparum*, purified schizonts were treated with the cross-linker Dithiobis[succinimidyl propionate] (DSP) before lysis and subject to immunoprecipitation (methods described in chapter 2.2.17). This IP was unsuccessful and no protein bands were observed after sample separation by SDS-PAGE and subsequent staining with Coomassie or silver-nitrate (data not shown). It is unfortunate that the 3xFLAG tag was not introduced into the locus, since from the previous chapter it appears to be a promising tool for the purpose of immunoprecipitation in particular. Ty1 antibodies are not as sensitive as the anti-FLAG M2 monoclonal antibody (Sigma-Aldrich) and perhaps not suitable for immunoprecipitation.

The data presented in section 6.7 suggests that *PF02_0040* cannot be knocked-out and by implication, is an essential gene. Caution should however be used here as it is possible that other factors influenced the ability of the construct to integrate. One such factor is accessibility of the locus to the construct – the gene may be subject to modification that obstructs integration. To rule this out the 3' replacement control would be required to demonstrate locus accessibility and should be attempted in the future.

In *T. brucei,* Goldschmidt *et al*., (2008) used RNAi to knock-down *SEC71* and show a significant reduction of GPI-anchored proteins present in these parasites. Like *Plasmodium*, GPI-anchored proteins cover the parasite cell surface and one in particular

– the variant surface glycoprotein VSG. This protein is essential for parasite survival and the parasites could not tolerate a reduction in its presence on the surface and it is this that is thought to cause the lethal consequence of silencing Sec71. Consequently, Sec71 of *T. brucei* is now a major potential drug target and compounds are being developed to inhibit its function. The fact that it has no human homologue makes it particularly attractive. If PfSec71 has a similar function then it too could serve as a target for novel anti-malarial drugs.

Alternatively, PF02_0040 and the post-translational secretory pathway may function in an Apicomplexan or *Plasmodium* specific manner. As described in the previous chapter, the targeting of RhopH2 appears to involve specific recognition of its signal peptide for correct targeting to the rhoptries (Ghoneim et al., 2007), and given the role of Sec71 in other species, a role in recognition of this alternative and specific set of signal peptides is possible but requires experimental validation. Since rhoptry proteins are important invasion organelles, a role in rhoptry targeting would also present PF02_0040 as an attractive drug target.

This study has shown that PF02_0040 is a membrane protein of the ER. Deciphering whether it is a functional homologue of Sec71 requires further characterisation, and co-precipitating other components of the translocon complex would increase the validity of this hypothesis, as would the study of a conditional knock-out. Unfortunately this is not achievable at present but hopefully in the future the technology will allow this analysis and the characterisation of what could be an important and much-needed new drug target.

7. MAL13P1.94 expression timing, localisation and functional characterisation

7.1 Introduction

MAL13P1.94 is predicted to encode a 22 kDa protein with 4 transmembrane domains which, unlike the other two genes under investigation in this thesis (PF11_0443 and PF02_0443) appears to be exclusive to apicomplexa. Considering that the invasion process is relatively conserved throughout apicomplexa and apical organelles are pivotal in this process MAL13P1.94 is an attractive protein to study.
Although much of what we know of *Plasmodium* host cell invasion was derived from findings in *T. gondii,* there is little information regarding this protein in the literature or on the *Toxoplasma* database ToxoDB (toxodb.org). Interestingly and in accordance with MAL13P1.94, the only proteomic data confirming the presence of this protein in *Toxoplasma* was from the proteomic study of Treeck *et al.* (2011), finding one peptide from the C-terminus. Due to this lack of information from orthologous proteins in other species, characterisation of MAL13P1.94 will hopefully provide further insights into *Plasmodium* host cell invasion and further our knowledge of this process in other apicomplexa.

7.2 MAL13P1.94 is an integral membrane protein in P. falciparum schizonts

MAL13P1.94 contains 4 predicted transmembrane domains therefore it is highly likely this protein is an integral membrane protein, nevertheless experimental confirmation is required. *P. falciparum* 3D7 schizonts were first hypotonically lysed in a high pH 0.1M sodium carbonate buffer to separate soluble and peripheral membrane proteins from those with membrane anchors. Western blot analysis of the protein extracts revealed MAL13P1.94, with an apparent molecular weight of 17 kDa, resides within the carbonate pellet fraction, confirming that it is indeed an integral membrane protein. Other bands seen by immunoblot of approximately 30 and 50 kDa are equivalent to rbc proteins and

crossreact with the antibodies produced against MAL13P1.94 (see chapter 4, figure 4.7 A). SERA-5 is a soluble protein, peripherally associated with the merozoite cell surface and antibodies to this protein were used to control for correct solubilisation of peripheral membrane proteins by carbonate buffer extraction, while anti-MSP2 antibodies were used to detect the GPI-anchored merozoite surface protein in the carbonate pellet fraction.

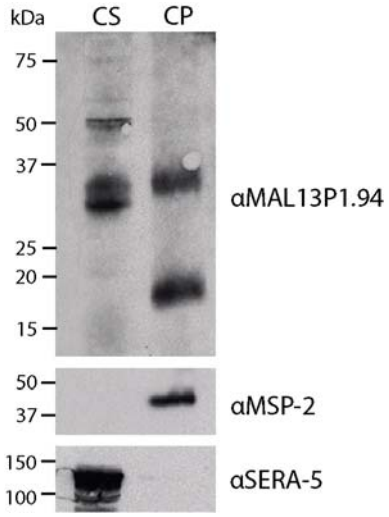


Figure 7.1: *Western blot analysis of carbonate lysed schizonts. P. falciparum* 3D7 schizonts were lysed in 0.1 M sodium carbonate buffer, pH11.0 to separate soluble and peripheral membrane proteins from the integral ones. Samples from carbonate supernatant and carbonate pellet were separated by SDS-PAGE and subject to western blot. SERA-5 (peripheral membrane) and MSP-2 (membrane-anchored) antibodies served as controls. The blot was probed with MAL13P1.94 antibodies and revealed this protein to be integral to the membrane.

7.3 MAL13P1.94 is expressed from 40h p.i. in schizonts and merozoites.

Bioinformatic analysis revealed MAL13P1.94 peak transcript levels occur around 38h p.i. but this is not necessarily indicative of the commencement of protein expression. Proteomic data places MAL13P1.94 in schizonts however, due to the lack of a ring stage proteome, it is unclear if it is carried through into ring stages. In order to discover at what point and for how long the protein MAL13P1.94 is present in the *P. falciparum* blood stage cycle it was necessary to complete a time-course experiment, sampling one complete cycle at 4 h intervals for protein detection (for method please refer to chapter 2.3.4). Samples were taken every 4 h starting with newly invaded rings (time point zero) through to late/rupturing schizonts and subjected to western blot analysis, including a rbc control (figure 7.3 A). Equal loading of rbcs in each well was confirmed by probing with an anti-glycophorin A/B antibody. The blot was also probed with an antibody against the parasite ER constituent protein BiP to confirm the presence of parasite material in each of the lanes containing material from the time-course (Kumar et al., 1988; Kumar et al., 1991). MSP-2, which is expressed in schizonts (Wickham et al., 2003), was used as a late stage control. Probing the time course with MAL13P1.94 antibody revealed protein expression began approximately 40 h p.i. during schizogony and increased during the following 4 hours. Protein levels were maintained until egress however the protein is not visible in ring stage parasites suggesting either the C-terminus was lost perhaps by shedding upon entry into the host cell or the protein is rapidly degraded within the parasite (figure 7.2 A).

In addition to western blot analysis, MAL13P1.94 antibodies were used in indirect immunofluorescence experiments carried out on each time point at which the protein was seen to be expressed by western blot (figure 7.2 B). Probing with a fluorescent secondary antibody revealed the presence of distinct foci associated with individual merozoites which reside in an anterior position in comparison to the nucleus. As MAL13P1.94 is an integral membrane protein, it must be associated with the membrane of an organelle or vesicle and this pattern is similar to that exhibited by residents of the

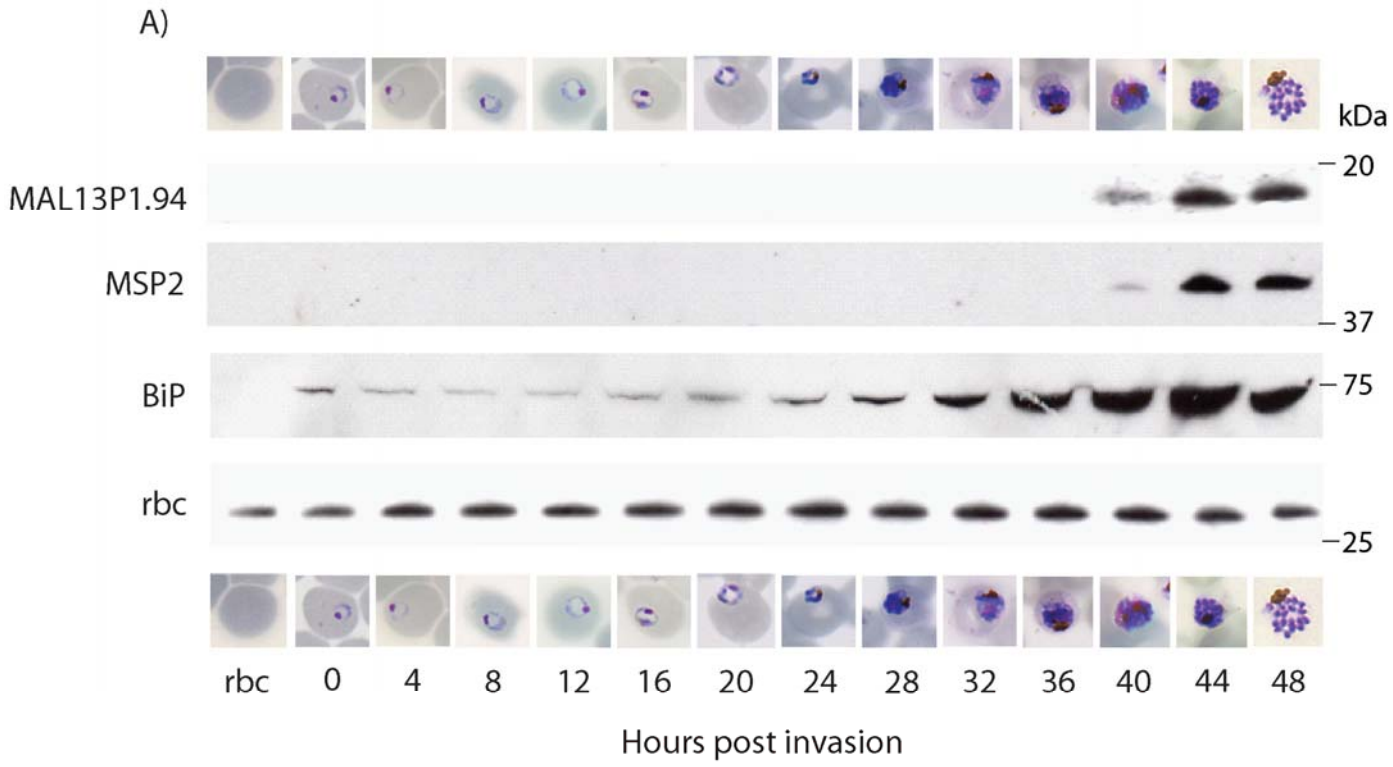
apical organelles. Although the protein is detectable from 40h p.i. by western blot, sufficient levels of fluorescence were not detectable by immunofluorescence until late in schizogony (figure 7.2 B). Co-localisation IFA studies with known organelle markers would give a clearer indication of the intracellular location MAL13P1.94 and this is addressed in the next section.

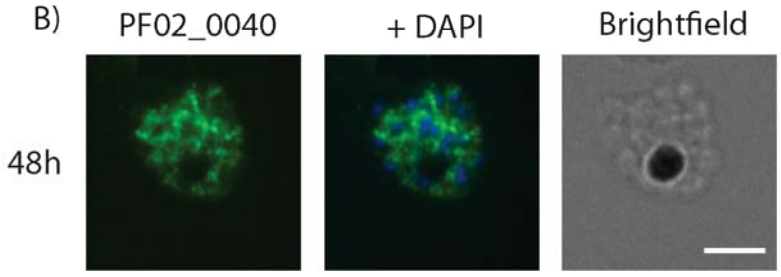
Figure 7.2: *Expression of MAL13P1.94 throughout the lifecycle.*

A) Samples of parasitized and uninfected rbcs were taken every 4 hours for one complete synchronised blood-stage cycle and hypotonically lysed , the pellet fraction solubilised in Laemmli SDS sample buffer and separated by SDS-PAGE ($2x10^6$ parasites loaded per well). Pellets of hypotonically lysed rbcs were included in the western blot to control for parasite versus rbc band recognition. Expression of MAL13P1.94 was analysed by western blot using the previously generated rabbit polyclonal anti MAL13P1.94. A specific antibody was used to detect glycophorin A/B antibody as a loading control. Antibodies against the parasite ER resident protein BiP were used to confirm the presence of parasites in each lane and anti MSP-2 was used as a late stage marker. Giemsa stained parasites from each time point are shown to demonstrate parasite morphology at each stage.

B) Thin smears of parasite culture from each time point were fixed in paraformaldehyde, permeabilised with Triton X-100 and used in IFAs to detect the presence of MAL13P1.94 in the parasite using specific antibodies. MAL13P1.94 is detected by a fluorescent secondary antibody (shown in green) and the nucleus is stained with DAPI (blue). Scale bar on the bright field image represents 5µm.

.


A)
kDa
MAL13P1.94
MSP2
BiP
rbc
20
37
75
25
rbc 0 4 8 12 16 20 24 28 32 36 40 44 48
Hours post invasion


B)
PF02_0040
+ DAPI
Brightfield
48h

7.4 MAL13P1.94 is to be found in the rhoptry neck of late schizonts and merozoites from where it is transferred to the PVM in early rings.

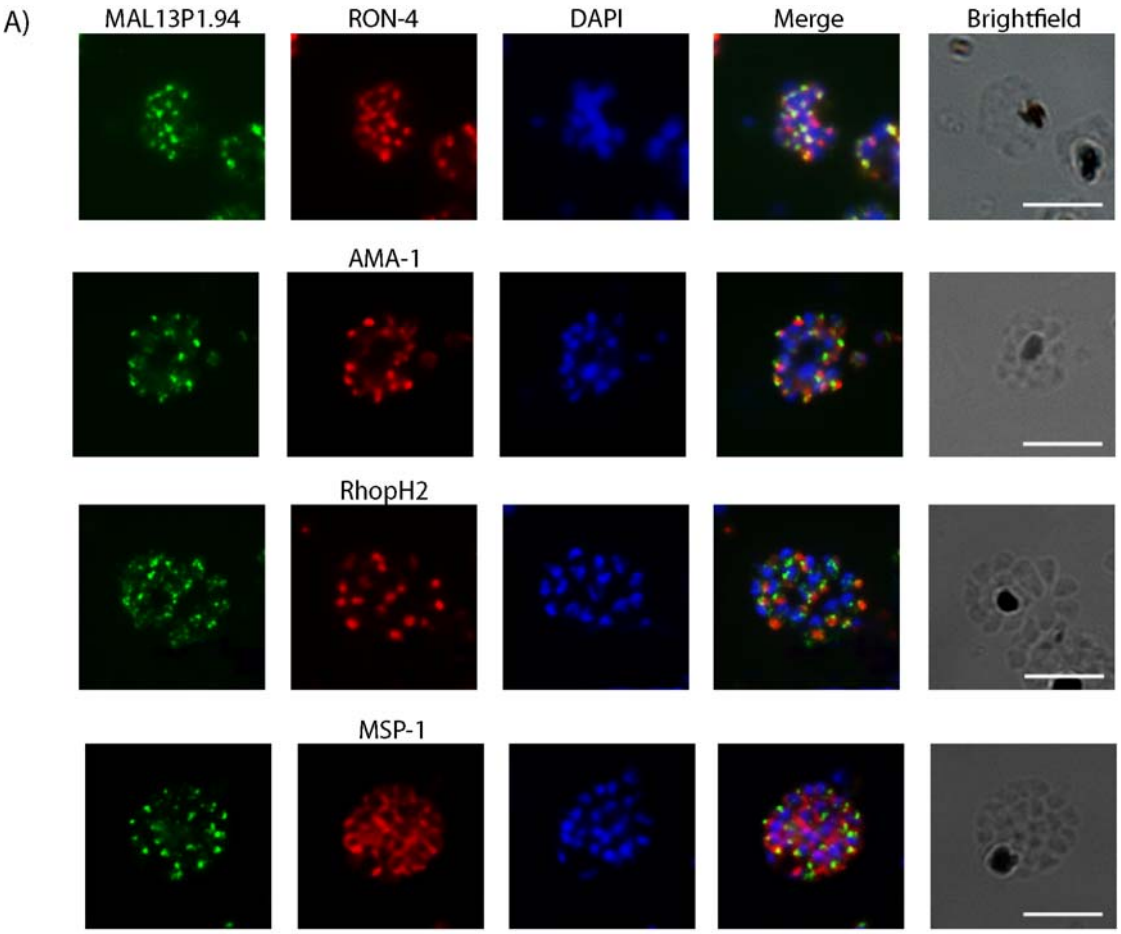
From the western blot analysis of the time-course protein samples it appears that MAL13P1.94 is expressed late in schizogony but it is not clear where it resides within the cell. In order to uncover the cellular location of this protein, thin smears of *P. falciparum* 3D7 were fixed and permeabilised and subject to indirect immunofluorescence assay using specific rabbit polyclonal antibodies. Fluorescence was compared to that produced by antibodies against proteins of already determined location, encompassing a number of locations within the cell. Focussing on later stages, MAL13P1.94 is clearly visible in multiple points of fluorescence at the apex of the cell (in comparison with the nucleus at the posterior) and appears to co-localise with rhoptry neck protein RON-4 (Alexander et al., 2006) in both schizonts (figure 7.4 A) and merozoites (figure 7.4 B), suggesting a rhoptry neck localisation. There is a similarity between the pattern of this protein and the micronemal protein AMA-1 (Healer et al., 2002; Bannister et al., 2003) in schizonts however the resulting fluorescence pictures from the assay does not overlap. Looking at micronemes in merozoites, there is some overlap between MAL13P1.94 and microneme resident protein EBA-175 (Sim et al., 1992). Due to their close proximity, size and the resolution limit of light microscopy, it is often difficult to distinguish between the rhoptry neck and micronemes. Although MAL13P1.94 does not co-localise with rhoptry bulb protein RhopH2 (Holder et al., 1985) in schizonts and in merozoites this protein appears at a single point at the apical tip, overlapping to some extent with RhopH2. However the rhoptry bulb protein covers a larger area and the parasite appears to be orientated at an angle, therefore MAL13P1.94 could well be closer to the anterior. Some overlapping fluorescence is seen between MAL13P1.94 and surface protein MSP-1 (Holder and Freeman, 1984a), however only in a small section which is more an indication of the organelles proximity to the cell membrane than evidence of the protein's presence on the surface.

From the time-course data presented in figure 7.3, MAL13P1.94 was not detected in ring stage parasites by western blot however rhoptry neck and bulb proteins such as RON4 and RhopH2 are transferred into rings upon invasion. To test whether this is true of MAL13P1.94, thin smears of early ring stage parasites were probed with specific antibodies and compared with RhopH2 which is transferred to the outer face of the PVM upon invasion (Hiller et al., 2003). Figure 7.4 C demonstrates that MAL13P1.94 is indeed present in rings and although the pattern of fluorescence differs slightly from RhopH2, they appear to be on the same plane suggesting MAL13P1.94 is transferred to the PVM (Ling et al., 2003) or the plasma membrane of the ring stage parasite. It is important to note that antibodies were used at a higher concentration in ring stage IFAs than in the schizont or merozoite IFAs and the camera exposure time was doubled when taking the images implying MAL13P1.94 is present at a lower level or more diffuse. This apparent reduction of MAL13P1.94 in ring stages could explain why this protein was not visible by western blot in the lysates ofring stage parasites.

In summary, indirect immunofluorescence has identified MAL13P1.94 as a possible rhoptry neck protein. Further investigation is required to determine the protein's function within the *P. falciparum* intra-erythrocytic cycle.

Figure 7.3: *Indirect immunofluorescence of late schizonts and merozoites.*

Thin smears of late schizonts (A), merozoites (B) and ring stage parasites (C) were fixed and permeabilised as described in section 2. 2.13.1 and used in IFAs to detect the presence of MAL13P1.94 in the parasite using specific rabbit polyclonal antibodies. Primary antibodies were detected by a fluorescently labelled secondary antibody (AlexaFluor 488 – green) and the nucleus is stained with DAPI (in blue). Compartments were identified using previously characterised antibodies to proteins of known location followed by secondary antibodies labelled with AlexaFluor 594 (red) and images were merged to determine comparative location (AMA-1 and EBA-175: micronemes; RON4: rhoptry neck; RhopH2: Rhoptry bulb [schizonts and merozoites] or PVM [ring]; MSP-1: surface). Antibodies and dilutions are listed in Table 2.4. Areas of co-localisation appear yellow. Scale bar on the bright-field image represents 5µm in schizont or 1µm in merozoite images.


A)
MAL13P1.94
RON-4
DAPI
Merge
Brightfield
AMA-1
RhopH2
MSP-1
B)
MAL13P1.94
RON-4
DAPI
Merge
Brightfield
EBA-175
RhopH2
MSP-1
C)
MAL13P1.94
RhopH2
DAPI
Merge
Brightfield

7.5 Protein-specific polyclonal rabbit antibody precipitates a band of the correct size for MAL13P1.94 and co-precipitates other proteins of unknown identity.

In order to determine the function of MAL13P1.94, it is desirable to establish which proteins it interacts with. Although conserved between apicomplexa, including *T. gondii*, there is no information regarding the function of MAL13P1.94's orthologues in the literature. With the aim of identifying potential binding partners, MAL13P1.94-specific antibodies were used to immuno-precipitate protein from ^{35}S-labelled schizonts (methods described in chapter 2.2.15). Resolution of this radio-labelled protein sample by SDS-PAGE and subsequent exposure to film revealed bands of approximately 18, 22, 28, 35, 100, 150 and 200 kDa. The 18 or 22 kDa bands most likely correspond to MAL13P1.94 since its predicted size is 22 kDa and its apparent molecular weight determined in previous western blots is 17 kDa (figure 4.7 A and figure 7.3 A).

In order to identify possible binding partners, a protein preparation from 3 ml of purified schizonts was passed over a column of MAL13P1.94 specific IgG bound to activated CNBr (methods described in section 2.2.16). Unbound proteins were washed off using a Tris buffer containing 300mM NaCl before eluting bound proteins in low pH buffer containing 100mM glycine. The elution was concentrated to 50 µl, of which 30µl was loaded onto a gel and eluted proteins were separated by SDS-PAGE. Staining the SDS-PAGE gel either with Instant Blue™ or silverstain (Blum et al., 1996) resulted in no clear bands. This could be in part due to the low affinity of the antibodies used or due to low protein levels of MAP13P1.94 in parasites. Due to time constraints, no further immunoprecipitation experiments were attempted with MAL13P1.94 polyclonal antibodies.

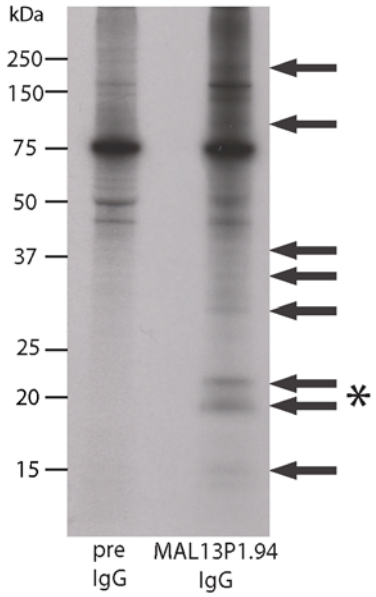


Figure 7.4: *Immunoprecipitation of radio-labelled schizonts with MAL13P1.94 polyclonal antibodies*. MAL13P1.94 rabbit polyclonal antibodies were used to immuno-precipitate interacting proteins from 50 µl of pelleted ^{35}S-labelled schizonts. Co-precipitating proteins were separated by SDS-PAGE and exposed to film. Two bands of the predicted size of MAL13P1.94, approximately 22kDa, were observed and highlighted with * and unique additional precipitated bands are indicated with arrows.

7.6 Epitope tagging of the MAL13P1.94 locus by targeted homologous recombination

As discussed in chapter 5.7, incorporating an epitope tag into the genomic locus of a gene is a useful tool for characterisation of the resulting protein. Due to the low affinity of the polyclonal antibodies raised as well as the problems of cross-reactivity of this antiserum with parasite and host proteins, incorporating an epitope tag into the *MAL13P1.94* locus is required as a tool for characterisation. The small epitope tags 3xFLAG and Ty1 were chosen and I attempted to introduce them onto the C-terminus of the protein since the protein contains a signal sequence that should not be interrupted at the N-terminus.

In order to introduce these tags onto the C-terminus of the protein, a region of homology of 480 bp from the 3′ end of the *MAL13P1.94* locus (minus the stop-codon) was amplified by PCR using primer combinations K/L and K/M (see Chapter 2, Table 2.1). Ty1 and FLAG epitope tag sequences (including a stop codon at the 3′ end) were included in reverse primers L and M respectively and used to amplify the region of homology, thereby incorporating the tag sequences. This region of homology plus tag sequence was cloned into the multiple cloning site of the pHH3 vector via the *Eco*RI/*Sac*II restriction sites generating pHH3_MAL13P1.94_Ty1 and pHH3_MAL13P1.94_3xFLAG. An *Avr*II site was included between the gene and tag sequences to allow the use of these vectors to epitope tag other genes. The pHH3 transfection vector which includes the Blasticidin resistance cassette as a selectable marker is depicted in Appendix D. The same region of homology of MAL13P1.94 was in addition cloned into the pHH3 vector via *Eco*RI/*Sac*II without the presence of an epitope tag to serve as a 3′ replacement vector.

Inserts were initially sub-cloned into the pGEM T-easy vector and transformed into *E. coli* TOP10 cells. Colonies were selected and sequenced before excising the correct inserts using *Eco*RI and *Sac*II. Inserts were gel-purified and ligated into the pHH3 vector and

transformed into *E. coli* TOP10. Colonies were re-sequenced and one construct with correct DNA sequence was propagated and plasmid DNA purified for transfection. For analysis purposes the construct was subject to digestion with *Eco*RI and *Sac*II to confirm the presence of the correct sized (480 bp) insert prior to transfection (figure 7.6 B)

Unfortunately due to complications in preparation and time constraints, pHH3_MAL13P1.94_3xFLAG and pHH3_MAL13P1.94_3' constructs were never completed or transfected. Complete and purified pHH3_MAL13P1.94TY was transfected into ring-stage parasites and selected for presence of episome by 2.5µg/ml of Blasticidin applied to cultures. Stock samples of these cultures were cryopreserved and genomic DNA was prepared. Transfected cultures were taken off drug for 3 weeks before re-application of drug, which is required for enrichment of the parasite population that contains integrated versus episomal copies of the transfection vector. This was repeated for up to 4 cycles. A diagnostic PCR was designed to test for integration using a forward primer from the *MAL13P1.94* 5'UTR region (primer BB, table 2.1) and reverse primers from either the *MAL13P1.94* 3'UTR region (primer CC, table 2.1) to identify the wild type locus or from the *HRP2* 3'UTR contained within pHH3_MAL13P1.94_Ty1 construct (primer NN, table 2.1) to detect integrated plasmid (figure 7.6 A). As shown in Figure 7.6C integration occurred during drug cycle 2 however MAL13P1.94_Ty1 parasites were not detected by western blot, despite previous success using the αTy1 antibodies in western blot analysis of other Ty1-tagged lines. This discrepancy is perhaps due to only a small number of parasites in the population containing the epitope tag but might also be a problem with the batch of Ty1 antibody used. The western blot will be repeated with an alternative batch of Ty1 antibody, however this is out-with the time restrictions of the project. Additionally, a Southern blot of genomic DNA taken from each cycle would provide definitive evidence to support or deny the presence of integrated construct. If successful, attempts shall be made in introduce FLAG- or GFP- tag the protein for use in immunofluorescence assays and immunoprecipitation experiments.

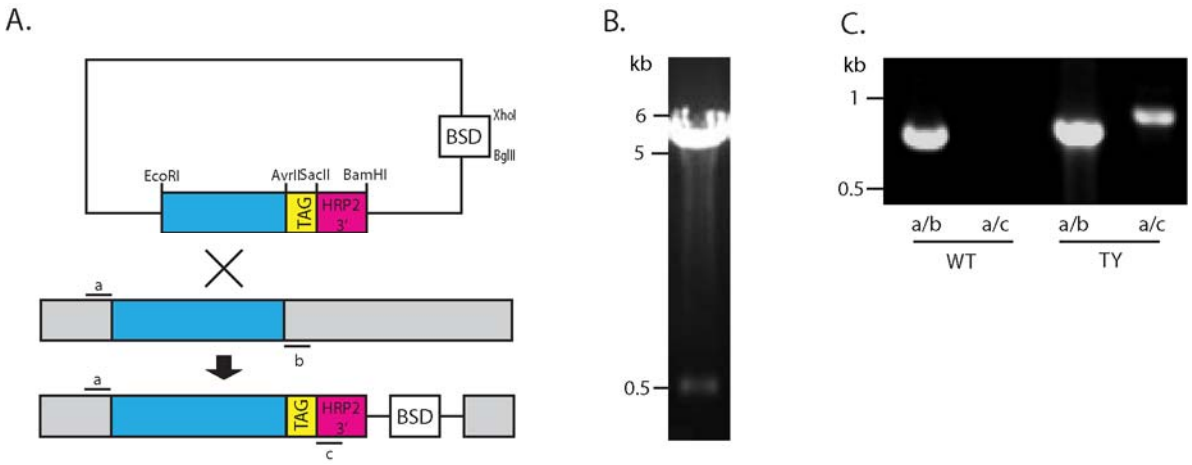


Figure 7.5: *Epitope tagging of the* MAL13P1.94 *locus with Ty1.* A 480bp region of homology from the 3′ end of the *MAL13P1.94* gene was amplified using primers K/L (sequence listed in chapter 2, table 2.1) and introduced into the pHH3 transfection vector along with a TY-1 epitope tag by way of *Eco*RI and *Sac*II restriction sites. An *Avr*II site was included between the region of homology and the Ty1 tag sequence to enable the use of this vector in epitope-tagging other genes (A). pHH3 contains the Blasticidin resistance cassette and a copy of the *P. falciparum hrp2* gene 3′ untranslated region (UTR). Restriction digest of the complete construct (5891bp) using *Eco*RI and *Sac*II shows the 480bp insert and 5411bp vector backbone (B). The plasmid was transfected into *P. falciparum* 3D7 and drug cycled to promote loss of episomal copies of the construct. A diagnostic PCR was designed to test for integration events using a forward primer from the *MAL13P1.94* 5′UTR region (a: primer BB, table 2.1) and reverse primers from either the *MAL13P1.94* 3′UTR region (b: primer CC, table 2.1) to identify the wild type (WT) locus or from the *HRP2* 3′UTR from the pHH3_MAL13P1.94_TY construct (c: primer NN, table 2.1) to detect integrated plasmid. PCR revealed the presence of both WT and integrated DNA (C).

7.7 Unsuccessful attempts to knock out MAL13P1.94 by truncation suggests an essential role for this gene within blood stages.

One way to uncover a protein's function is to study the effect of the loss of protein by observing the phenotype of a knockout parasite line compared to the wildtype. As discussed previously, conditional knock-downs are not well established for *P. falciparum* therefore attempting a full knock out is the best alternative method available to analyse the effects of a protein's absence or to establish whether or not it is essential. If the protein is essential during asexual blood stages then a knock out will not be possible and no integration will take place; if not essential then there is potential for identifying the gene's function through phenotypic analysis of the genetically modified parasites.

Unfortunately, due to the small size of the gene and the AT rich nature of the 5' and 3' UTRs, a full knock out of *MAL13P1.94* by double homologous recombination was not possible. Instead gene truncation was attempted using a construct designed to integrate into the 5' end of the gene. If successful, the expressed protein would lack the last 43 amino acids from the C-terminus of MAL13P1.94 and a complete or partial reduction of expression can be expected due to the lack of a 3' UTR and terminator sequence (see figure 7.7 for KO strategy). These 43 residues from the C-terminal end of the parasite represent the most significant portion of the protein not predicted to be incorporated into a membrane and it was therefore hoped that successful integration would disrupt function and exhibit a phenotype, or in the case that MAL13P1.94 is an essential gene, would result in non-viable parasites and therefore integration would not be possible.

The integration transfection vector pHH3_MAL13P1.94_KO was manipulated to include 474bp from the 5' UTR and ORF of *MAL13P1.94*, amplified by PCR using primers I/J (sequences listed in table 2.1) and introduced into the construct's multiple cloning site using the enzymes *Eco*RI and *Bam*HI. The construct contains the blasticidin resistance cassette to be used as a selectable marker, conferring blasticidin resistance to

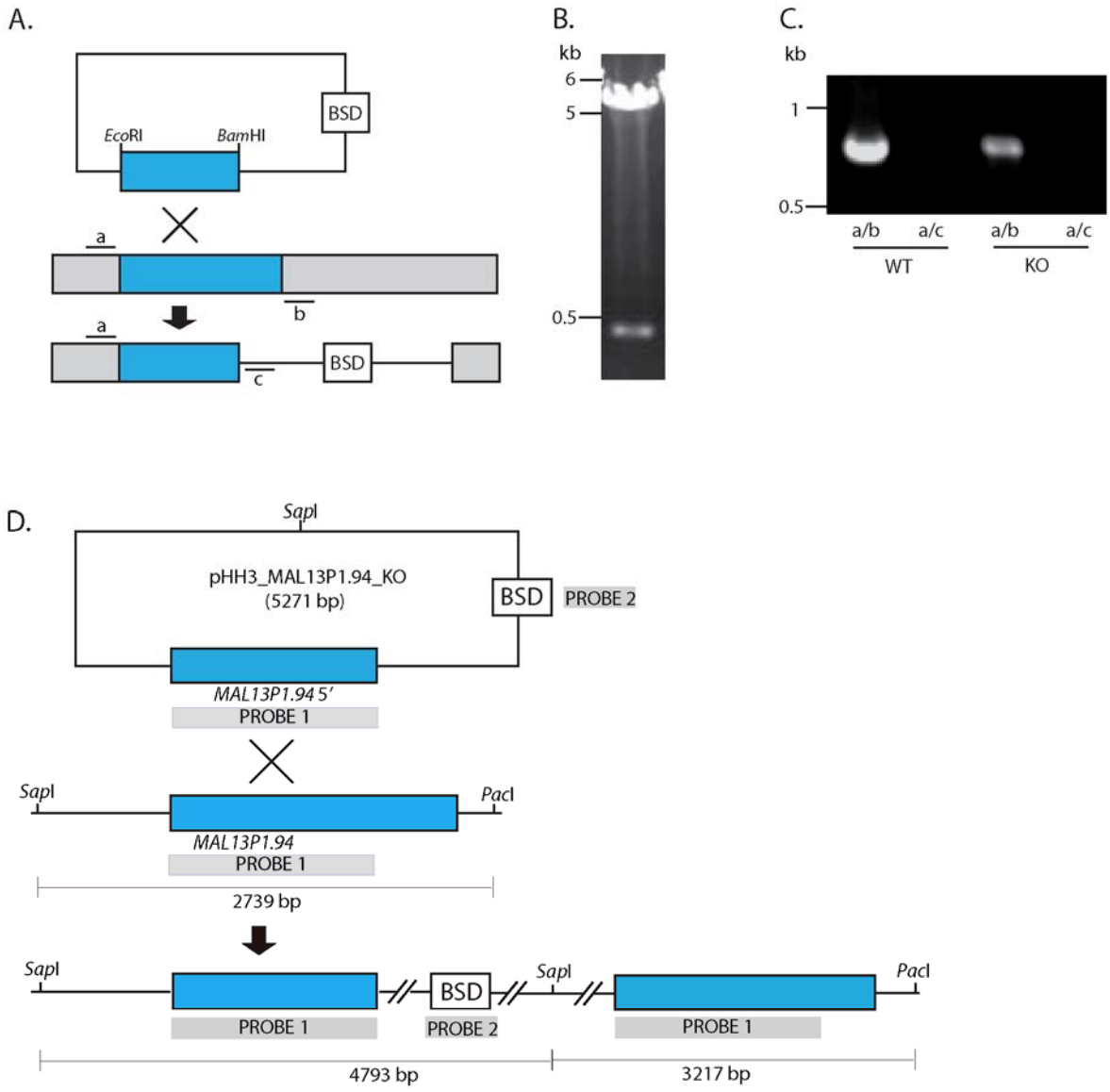
successfully transfected parasites. Inserts were initially sub-cloned into the pGEM T-easy vector and transformed into *E. coli* TOP10 cells. Colonies were selected and sequenced before excising the correct sequence inserts using *Eco*RI and *Bam*II. Inserts were gel-purified and ligated into the pHH3 vector and transformed into *E. coli* TOP10. Colony plasmid DNA was screened by restriction digest using *Eco*RI and *Bam*II (figure 7.7 B) and re-sequenced before selecting one correct sequence vector to propagate and purify for transfection.

As with the epitope TY-tag construct, complete and purified pHH3_MAL13P1.94_KO was transfected into ring-stage parasites. Selection for presence of episome was achieved by the addition of 2.5µg/ml of Blasticidin to transfection cultures. Stock samples of these cultures were cryopreserved and genomic DNA was prepared. Transfected cultures were taken off drug for 3 weeks before its re-application, which will enrich the proportion of the parasite population that contains integrated versus episomal copies of the transfection vector. This was repeated for 5 cycles. A diagnostic PCR was designed to test for integration events using a forward primer from the *MAL13P1.94* 5'UTR region (a: primer BB, table 2.1) and reverse primers from either the *MAL13P1.94* 3'UTR region (b: primer CC, table 2.1) to identify the wild type locus or from the *hsp86* sequence of the pHH3_MAL13P1.94_KO vector back-bone (c: primer OO) to detect integrated plasmid (figure 7.7 A). No integration was detected by this method within 5 drug cycles (figure 7.7 C). A Southern blot was designed which confirmed this negative result. For this procedure, gDNA was subject to digestion by restriction enzymes *Sap*I and *Pac*I before separating the resulting fragments by gel electrophoresis and blotting onto a nitrocellulose membrane. In order to ascertain whether integration occurred, the membrane was probed with a radiolabelled DNA probe homologous to either blasticidin ORF or the 474 bp fragment of MAL13P1.94 cloned into the pHH3_MAL13P1.94_KO construct. The resulting banding pattern from exposure to film revealed the presence of wild type DNA with a band of 2739 bp (figure 7.7 E) and a band between 5 and 6 kb markers corresponding to the 5271 bp episome (figure 7.7 E+F) throughout the 5 cycles. In accordance with the diagnostic PCR (figure 7.7C), no bands of 4793 or 3217 bp, corresponding to an integration event, were detected. Instead, the

5.2kDa episome band disappears in cycle 5 when probed with the region of homology and a band of 5.5 kDa is detected by the blasticidin probe. Although similar to the episome band, there is a clear difference between the bands, suggesting the episome has integrated elsewhere in the genome. This attempted KO by truncation was repeated in duplicate however neither attempt was successful (data not shown).

The data presented here suggests that *MAL13P1.94* cannot be knocked-out and by implication, is an essential gene.

Figure 7.6: *Attempted KO of* MAL13P1.94 *by truncation.* A) A single cross-over strategy was employed for the KO of *MAL13P1.94* which would result in a truncated version of the gene with no 3'UTR. Primers a (*MAL13P1.94* 5'UTR region), b (*MAL13P1.94* 3'UTR region), c (*hsp86* sequence of pHH3_MAL13P1.94_KO back-bone) were designed to use for diagnosis of intergration versus the presence of wild type sequence. The transfection vector pHH3_MAL13P1.94_KO was manipulated to include 474bp from the 5' UTR and ORF of *MAL13P1.94*, amplified by PCR and introduced into the construct's multiple cloning site using the enzymes *Eco*RI and *Bam*HI. Digestion with these enzymes from the completed plasmid excised a band of the correct size (B). After transfection, the positive parasites were drug cycled for 5 cycles, and a diagnostic PCR was designed to test for integration events using primer combinations a and b (detection of wild type) or a and c (to detect integration of plasmid). The PCR revealed presence of the wild type locus but not of any integration events (C). A Southern blot was designed to confirm this result (D). Genomic DNA taken each drug cycle was digested by *Sap*I and *Pac*I before separating resulting fragments by gel electrophoresis and blotting onto nitrocellulose membrane. The membrane was probed with either (E) a radiolabelled DNA probes corresponding to the 474 nucleotide fragment of the MAL13P1.94 region of homology (probe 1) or (F) the blasticidin ORF (probe 2). The resulting banding pattern from exposure to film revealed the presence of wild type DNA with a band of 2739 bp (WT) and a band between the 5 and 6 kb markers corresponding to the 5271 bp episome (E) throughout the 5 cycles. Band X does not correspond to the size expected from intergration into the *MAL13P1.94* locus.


A.
BSD
EcoRI
BamHI
a
b
a
c
BSD
B.
kb
6
5
0.5
C.
kb
1
0.5
a/b
a/c
a/b
a/c
WT
KO
D.
SapI
pHH3_MAL13P1.94_KO
(5271 bp)
BSD
PROBE 2
MAL13P1.94 5'
PROBE 1
SapI
PacI
MAL13P1.94
PROBE 1
2739 bp
SapI
BSD
SapI
PacI
PROBE 1
PROBE 2
PROBE 1
4793 bp
3217 bp
E.
Probe 1.
kDa
6
5
4
3
2
E
WT
WT
C1
C2
C3
C4
C5
F.
Probe 2.
kDa
6
5
4
3
2
X
E
WT
C1
C2
C3
C4
C5

7.8 Discussion

MAL13P1.94 is 22 kDa integral membrane protein expressed from 40 h p.i. in blood stage *P. falciparum* 3D7. This protein appears apical by immunofluorescence and co-localises with known rhoptry neck protein RON4. Overlapping fluorescence is also seen with micronemal proteins however the pattern of fluorescence is different in schizont stages. Due to their close proximity within the merozoite and the fact that the limit of the resolution of light microscopy (diameter of light omitted by the fluorophore) is around 200-300nm, it is often difficult to distinguish between the rhoptry neck and micronemes. MAL13P1.94 does not co-localise with rhoptry bulb marker RhopH2 in schizonts, and is anterior to RhopH2 in merozoites. Rhoptry neck and bulb proteins are transferred into rings upon invasion, however by western blot MAL13P1.94 was not detected in ring stages. Indirect immunofluorescence revealed that MAL13P1.94 is in fact present in rings (figure 7.4 C) although it was necessary to use a higher concentration of antibody to stain ring stage parasites compared to late schizont and merozoite stages. This suggests that there is less protein present and perhaps this is why no band was not detected by western blot on ring stage lysates. Micronemal proteins are normally shed upon invasion and therefore antibodies to the extracellular portion of the proteins do not detect the newly invaded rings in IFAs. Given that MAL13P1.94 can be detected in ring stages gives weight to the hypothesis that this protein is present in the rhoptry neck, however the possibility that antibodies were raised against an intracellular region which would not be shed upon invasion cannot be ruled out since the topology of the protein has not been addressed in this study. RhopH2 was used as a marker for the PVM in ring stages and although MAL13P1.94 and RhopH2 did not completely co-localise, these proteins seem to be in close apposition. It is possible that MAL13P1.94 is maintained in the plasmalemma of the ring parasite or that its distribution along the PVM is not homogenous.

Since beginning this study, a global transcriptome study analysing the effects of growth perturbations by a set of growth inhibiting compounds at schizonts stages assigned MAL13P1.94 as a protein with a putative role in invasion and episomal expression of a

GFP-tagged MAL13P1.94 also places this protein at the apex of merozoites by IFA although no colocalisation attempts were published (Hu et al., 2010).

From the western blot of schizonts lysed in carbonate buffer, we saw that MAL13P1.94 behaves like an integral membrane protein. This result was as expected due to the presence of 4 putative transmembrane domains within the protein. The antibodies raised against MAL13P1.94 cross-reacted with 2 additional proteins: an approximately 30kDa protein that is also detected in erythrocytes (chapter 4, figure 4.7) and a parasite specific 50 kDa protein. The same antibodies were used in the IFAs and therefore caution should be taken when analysisng the results. Despite these cross-reacting bands, the pattern observed by IFA is consistent with a single, apical location. No rbc cross-reactivity was observed by IFA and in fact no reactivity at all was observed in lifecycle stages other than segmented schizonts, merozoites and early rings. The 50 kDa protein is present in the carbonate supernatant and was absent from the hypotonically lysed schizonts from the time-course westerns. Since completely soluble, it is possible that the 50 kDa protein is lost during paraformaldehyde fixation, as is the case for haemoglobin. Antibodies recognise globular as well as linear epitopes and when proteins are separated by SDS-PAGE under reducing conditions, linear epitopes are exposed that are not necessarily available to the antibody when the protein is in its native state. This is perhaps responsible for the discrepancy between western blot and immunofluorescence assays.

To definitively assign a sub-cellular location to MAL13P1.94 we tried to localise the protein using immuno-electron microscopy (IEM) on late schizont stages. Unfortunately, attempts to confirm location by IEM using MAL13P1.94 rabbit polyclonal antibodies were unsuccessful.

In order to confirm the results obtained with the specific polyclonal antibody raised in this study we attempted to introduce an epitope tag at the C-terminus of MAL13P1.94. Epitope tags are designed to be immunogenic and commercial antibodies are available. GFP is a particularly attractive epitope tag, however thought unsuitable for MAL13P1.94

since its size of 27 kDa would more than double the size of the wild type protein and is therefore more likely to affect the protein's function. Small epitope tags such as Ty1 and FLAG have previously been used in Plasmodium with some success (Treeck et al., 2009; Wider et al., 2009). At first glance, attempts to tag MAL13P1.94 with the Ty1 epitope tag seemed successful since the diagnostic PCR revealed the presence of integrated episome. Unfortunately, the anti-Ty1 antibodies failed to detect the presence of the tagged protein in purified late schizonts by western blot or IFA. As discussed earlier this might be due to only a small subpopulation in the culture being actually tagged. However time constraints have not allowed further attempts to enrich for the integrated population and cloning of the tagged line. As seen with PF11_0443, epitope tags can also be useful tools in immunoprecipitation studies; unfortunately this was not possible here.

The lack of identification of any binding partners makes it impossible to predict the function of this protein. A successful knock-out would have allowed the study of any phenotype and perhaps deduce the protein's function. The data presented in section 7.7 suggest that *MAL13P1.94* cannot be knocked-out and by implication, is an essential gene. Caution should however be used here as it is possible that other factors influenced the ability of the construct to integrate. One such factor is accessibility of the locus to the construct – the gene may be silenced or subject to modification that obstructs cross-over events. The fact that the pHH3_MAL13P1.94_Ty1 construct was able to integrate into the MAL13P1.94 genomic locus demonstrates however that the locus is accessible. The regions of homology in the case of the epitope tag and the truncation construct differ which may mean that the genomic region used in the KO strategy is incompatible with homologous recombination. An alternative control could include the precise region of homology used, followed directly by a re-codonised copy of the remaining sequence of the gene, and a 3'UTR/terminator sequence to test if the region is compatible with recombination without altering the *MAL13P1.94* gene product. Even with this approach there is still the somehow remote possibility that the re-codonised sequence is not compatible with *P. falciparum* and expression is consequently reduced or lost so no integration of the control would not necessarily equate to inaccessibility of

the region of homology; however this would be something to attempt in the future. Another limiting factor for this experiment was the size of the available region of homology. From work in the Division it has generally been noted that integration will not occur with regions of homology less than 400 bp and that although 500+ bp is what is generally used, longer regions integrate in an earlier drug cycle and have even been shown to integrate where shorter gene fragments have failed. Although 474bp is adequate, perhaps if a longer gene fragment was accessible then a different result would be had. Unfortunately this was not possible due to the 30 base stretch of poly-T and high AT percentage in the flanking UTRs which would prove difficult to maintain in bacterial culture and would increase the chance of random integration into one of the many other regions in the genome with over 90% AT content.

In order to take work on this protein forward, it is essential that an epitope tag is integrated into the genomic locus. FLAG or GFP would be particularly beneficial due to the abundance of commercially available tools such as monoclonal antibodies and affinity columns which increase the experimental capacity of the protein and would aid functional studies.

Proteomic data from Treeck & Sanders *et al* (2011) identified phosphorylated residues at the C-terminus of the protein, namely S179 and T182 (S*LGT* where *=phosphate addition). Interestingly phosphorylated residues are also present in the *Toxoplasma* homologue (TGME49_054070) in a similar motif at residues T178 and S181 (T*LGS*) with an additional phosphorylated threonine at position 185. Uncovering the purpose of this phosphorylation could reveal the function of the protein within both organisms. The fact that the C-terminus is phosphorylated suggests that it is on the cytoplasmic side of the membrane, therefore this protein has no real extracellular domain of sufficient stature characteristic of possible vaccine targets. Nevertheless the role of this protein may be mechanistic with phosphorylation/dephosphorylation acting in a regulatory capacity and therefore MAL13P1.94 may yet be interesting to study with regards to a role in the invasion process.

Although we did not characterize MAL13P1.94 fully this study has been able to locate this protein to the apical end of the parasite and therefore confirm the validity of the selection criteria laid out in section 1.8. Due to the protein's size and the presence of 4 membrane spanning regions, the only hydrophilic exposed stretches of amino acids available to raise antibodies against are very small. For this reason generating an epitope tagged transgenic parasite line is highly desirable. We have attempted with some success to generate a Ty-1 tagged line suggesting that such a line might be viable; laying the groundwork for future attempts to tag and characterize this protein.

8. Project summary

The aim of this project was to uncover novel proteins that play a part in the invasion of red blood cells by merozoites. In order to identify such proteins, the *Plasmodium falciparum* genome was explored, using bioinformatic data that suggested expression in the late blood stages. Five proteins were initially selected for study based on conservation in the *Plasmodium* genus, the presence of a signal sequence and of one or more transmembrane domains and transcription data confirming expression of candidate genes in schizogony. The set of criteria used to search for these proteins was deemed successful due to the presence of previously identified invasion related proteins including AMA-1, MTRAP and multiple members of the EBA, Rh and MSP families.

Although initially five proteins were selected for further study, characterisation of the full set was never intended. Starting with this number allowed room for experimental difficulty so that if one or two proteins were not compatible with being taken forward in some way, the other proteins would ensure the project's success. The initial hurdle to overcome was protein solubility and it was at this stage that work on PF14_0325 was terminated due to difficulty in recombinant protein expression in *E. coli*. This protein was attractive due to the presence of transcripts in both schizonts and merozoites, implying the protein might have a conserved role in invasion. However PF14_0325 has a putative GPI-anchor and it would therefore not have been possible to integrate a tag onto the C- or N-terminus. Without protein- or epitope tag-antibodies, characterisation of this protein would be difficult. Perhaps this will be achieved in the future, particularly if a usable conditional knock-out method can be established for *P. falciparum.*

The protein PF14_0045 was also selected for further study and although recombinant protein was expressed and antibodies were raised, the antibodies did not recognise a band of the correct size in merozoites. Consequently the protein remains uncharacterised. Perhaps if the initial strategy used had focused on genetic manipulation to tag the protein instead of production of specific polyclonal antibodies

as tools for characterisation then this protein may have featured more heavily in this thesis. However this could be something to focus on in the future.

Having discounted two selected proteins, three remained to be characterised. MAL13P1.94 was found to be an integral membrane protein present at the apex of merozoites and found in early rings. Colocalisation by IFA strongly suggested this protein is present in the rhoptry neck, however confirmation by IEM was not successful. Antibodies to a protein are an extremely important tool in characterisation studies but unfortunately the antibodies raised against MAL13P1.94 recognised additional, apparently non-specific bands. Although an alternative set of antibodies was raised against a slightly different protein sequence, the same contamination problem persisted. Due to the presence of three transmembrane domains and a signal anchor in this protein, only a small region at the C-terminus was suitable for protein expression in terms of size and hydrophilicity, therefore antibody options were limited. Integrating an epitope tag into the locus of MAL13P1.94 would go some way to solving this problem. Attempts to tag MAL13P1.94 with the Ty1 epitope tag seemed successful since the diagnostic PCR revealed the presence of integrated DNA. Unfortunately, the anti-Ty1 antibodies failed to detect the tagged protein in purified late schizonts by western blot or IFA. This might be due to only a small subpopulation in the culture expressing tagged protein or low levels of tagged protein in each cell, however time constraints have not allowed further attempts to enrich for the integrated population. Cloning of the tagged line is something that will be attempted in the future. Integration into the locus proves it is accessible and therefore the fact that a construct designed to knock-out MAL13P1.94 failed to integrate over 5 drug cycles in 3 attempts suggests the protein is essential. Although this provides little insight into function, as *Plasmodium* genetics advances a conditional knock-out method may become available that will allow the role of MAL13P1.94 to be uncovered. A study by Hu *et al* (2010) that was published during the course of this project analysed the effects of growth perturbations by a set of growth inhibiting compounds on schizonts stage transcript levels and assigned MAL13P1.94 as a protein with a putative role in invasion. Also, episomal expression of a GFP-tagged MAL13P1.94 places this protein at the apex of merozoites by IFA although no

colocalisation attempts were published (Hu et al., 2010). These data confirm what was already observed by IFA using anti-MAL13P1.94 polyclonal antibodies, and further validate the selection criteria since MAL13P1.94 appears to be a novel protein with a role in invasion, even if the precise role remains elusive.

Uncovering the function of a protein without conserved domains is an arduous undertaking. *Plasmodium* is relatively well studied in comparison to other pathogens due to the funding it attracts from its medical relevance [*Toxoplasma* is also relatively well studied however more so in cellular processes important for virulence and host cell invasion than more general molecular cell biology]; therefore little can be learned from bioinformatics unless the protein has orthologues of known function. This was certainly the case for MAL13P1.94, but in the beginning PF02_0040 appeared to have no known orthologues other than in eukaryotic pathogens, particularly if the data from OrthoMCL (available on plasmodb.org). PF02_0040 was discovered to contain a Sec66 domain 2 years into the project by BLAST search using the protein sequence of the *T. gondii* orthologue. Due to the increased power of protein domain prediction algorithms, this domain is now predicted by Pfam with an e-value of 0.00011 which is significant but lower than normally observed between protein family members . For example the J domain of PF11_0443 has an e-value of $1.5x10^{-8}$ (Finn et al., 2008) and this information has been added to PlasmoDB following its latest update from May of this year. Due to the AT rich nature of the *P. falciparum* genome, amino acids which are coded for by A- or T-rich codons are preferentially utilised and this could contribute to the lack of detectable sequence identity between *P. falciparum* and, for example, yeast proteins.

The identification of the Sec66 domain implies PF02_0040 is a homologue of ER resident protein Sec66 (aka Sec71) of *Saccharomyces*. If this protein feature had been detectable at the beginning of the project then perhaps PF02_0040 would not have made the shortlist of protein for further study, because any protein selected by the criteria containing an obvious ER retention signal, such as an XDEL motif, was discounted. At the start of the project it was assumed that only receptor/ligand interaction at the surface are important for invasion, a fact which is now clear not to be the case.

Specific antibodies raised against PF02_0040 revealed this protein is an integral membrane protein present throughout the intra-erythrocytic lifecycle of *P. falciparum,* where is resides within the ER, as shown by colocalisation with parasite ER resident BiP. Unfortunately these antibodies were unable to immuno-precipitate other members of the ER translocase therefore an equivalent role for this protein between *Plasmodium* and yeast was not verified. An IP using antibodies to an epitope tag may have improved upon this result however this was not possible as the 3xFLAG integration construct for this protein did not integrate. Targeted genetic deletion attempts of *PF02_0040* were unsuccessful over 5 drug cycles, which would suggest this protein is essential. However the possibility that failure to integrate an epitope tag into the locus may represent a technical difficulty in recombination and gene accessibility cannot be ruled out. This issue of accessibility will be addressed in the near future by attempted integration of a 3′ replacement construct containing genomic sequence from *PF02_0040* without the presence of a C-terminal tag. The issue of ER retention will also be addressed since PF02_0040 contains no obvious retention sequence such an XDEL motif which is present in other ER resident proteins.

Despite PF02_0040's presence in the ER, the protein may still have an impact on invasion. Targeting of RhopH2 appears to involve specific recognition of its signal peptide for correct targeting to the rhoptries (Ghoneim et al., 2007), and given the role of Sec71 in the post-translational translocation of proteins into the ER in other species, a role in recognition of a specific set of signal peptides is possible. Since rhoptry proteins are important in these invasion organelles, a role here would also present PF02_0040 as an attractive drug target. Sec71 of *T. brucei* is already being developed as a drug target due to its indispensible role in translocating GPI-anchored proteins into the secretory pathway (Ghoneim et al., 2007). Since Sec71 is absent in humans, PF02_0040 would also be an attractive drug target if the proteins share similar role. However in the absence of any conditional knock-down, this is impossible to predict. Nonetheless there is definite potential for this protein to impact upon erythrocyte invasion by merozoites, despite residence in the ER and homology to conserved eukaryotic secretory components.

The final protein selected is also conserved in other eukaryotes. PF11_0443 is a putative type IV J domain-containing protein or Hsp40. Although a role in invasion for this protein would not necessarily be predicted, transcript data revealed this protein is only expressed in blood stage schizonts (Bozdech et al., 2003) and there are other J proteins present in dense granules [such as ring-expressed surface antigen (RESA) (Culvenor et al., 1991)] and *P. berghei* ookinete micronemes (Lal et al., 2009). PF11_0443 was found to be an integral membrane protein expressed from 36 h pi. The protein is present in the ER of early schizonts where it remains until segregation of daughter merozoites in what is termed late schizonts or segmenters. At this stage it appears to translocate to the apex, but it is not clear whether it is present in rhoptries or micronemes. A rhoptry location is more likely as it is transferred into the early ring stage, albeit at a lower level than seen in schizonts and merozoites and it appears to be degraded at this stage as it is not visible until the commencement of the next round of schizogony. IFAs using 3xFLAG epitope-tagged PF11_0443 also supported a rhoptry location although only IEM would provide a definitive answer to the question of subcellular location. Unfortunately this approach was unsuccessful and so far attempting to repeat it using the 3xFLAG tagged line and anti-FLAG antibodies has not been possible. Immunoprecipitation studies pulled-down a number of potential binding partners including a number of chaperones, however reciprocal immunoprecipitation is required to confirm these interactions as real. As with the other two proteins, targeted genetic deletion of PF11_0443 was not possible over 5 drug cycles in 3 attempts. On this occasion a double homologous recombination strategy was employed and the 3' region of homology (flank 2) did integrate into the *PF11_0443* locus, but this expression had little effect since the protein was still expressed at 90% of WT levels. The PF11_0443 locus is clearly accessible, however without a control for integration into the exact region used in flank 1, the possibility that the lack of integration is due to a technical difficulty in homologous recombination cannot be ruled out.

The apical organelles have specifically evolved to facilitate invasion and the presence of PF11_0443 in these organelles implies this protein has a role in erythrocyte invasion.

Since there is a clear distinction between the subcellular location of PF11_0443 in early schizonts and PF11_0443 in late schizonts, perhaps the protein has a role in the trafficking of late-expressed blood stage proteins of the secretory pathway out to the apical organelles. From IFA and IP data, the organelles in question are most likely the rhoptries. Experimental validation of this putative function is required and will be undertaken in the future. If PF11_0443 has an essential role in this process then it could also represent an attractive anti-malarial drug target.

8.1 Concluding remarks

Invasion of erythrocytes involves a multitude of molecular processes: from signal peptide recognition which appears important in rhoptry protein targeting (Ghoneim et al., 2007); to signalling pathways involving kinases and phosphatases; extensive proteolytic processing of protein precursors (Blackman, 2004); and cell motility by way of an actin myosin motor (Baum et al., 2006). Receptor/ligand interactions at the host/parasite interface represent only a small subset of these processes, yet are the most studied due to their vaccine antigen potential. Blocking the processes involved in erythrocyte invasion will result in a block in the parasite's lifecycle and therefore they are also attractive drug targets. Although none of the proteins detailed in this study appears to function via direct interaction with the host, they have shown potential involvement in other processes important for invasion and could represent desperately needed new drug targets.

9. Bibliography

Achan, J., Talisuna, A.O., Erhart, A., Yeka, A., Tibenderana, J.K., Baliraine, F.N., Rosenthal, P.J., D'Alessandro, U., 2011. Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. Malar J 10, 144.

Acharya, P., Pallavi, R., Chandran, S., Chakravarti, H., Middha, S., Acharya, J., Kochar, S., Kochar, D., Subudhi, A., Boopathi, A.P., Garg, S., Das, A., Tatu, U., 2009. A glimpse into the clinical proteome of human malaria parasites *Plasmodium falciparum* and *Plasmodium vivax*. Proteomics Clin Appl 3, 1314-1325.

Agnandji, S.T., Lell, B., Soulanoudjingar, S.S., Fernandes, J.F., Abossolo, B.P., Conzelmann, C., Methogo, B.G., Doucka, Y., Flamen, A., Mordmuller, B., Issifou, S., Kremsner, P.G., Sacarlal, J., Aide, P., Lanaspa, M., Aponte, J.J., Nhamuave, A., Quelhas, D., Bassat, Q., Mandjate, S., Macete, E., Alonso, P., Abdulla, S., Salim, N., Juma, O., Shomari, M., Shubis, K., Machera, F., Hamad, A.S., Minja, R., Mtoro, A., Sykes, A., Ahmed, S., Urassa, A.M., Ali, A.M., Mwangoka, G., Tanner, M., Tinto, H., D'Alessandro, U., Sorgho, H., Valea, I., Tahita, M.C., Kabore, W., Ouedraogo, S., Sandrine, Y., Guiguemde, R.T., Ouedraogo, J.B., Hamel, M.J., Kariuki, S., Odero, C., Oneko, M., Otieno, K., Awino, N., Omoto, J., Williamson, J., Muturi-Kioi, V., Laserson, K.F., Slutsker, L., Otieno, W., Otieno, L., Nekoye, O., Gondi, S., Otieno, A., Ogutu, B., Wasuna, R., Owira, V., Jones, D., Onyango, A.A., Njuguna, P., Chilengi, R., Akoo, P., Kerubo, C., Gitaka, J., Maingi, C., Lang, T., Olotu, A., Tsofa, B., Bejon, P., Peshu, N., Marsh, K., Owusu-Agyei, S., Asante, K.P., Osei-Kwakye, K., Boahen, O., Ayamba, S., Kayan, K., Owusu-Ofori, R., Dosoo, D., Asante, I., Adjei, G., Adjei, G., Chandramohan, D., Greenwood, B., Lusingu, J., Gesase, S., Malabeja, A., Abdul, O., Kilavo, H., Mahende, C., Liheluka, E., Lemnge, M., Theander, T., Drakeley, C., Ansong, D., Agbenyega, T., Adjei, S., Boateng, H.O., Rettig, T., Bawa, J., Sylverken, J., Sambian, D., Agyekum, A., Owusu, L., Martinson, F., Hoffman, I., Mvalo, T., Kamthunzi, P., Nkomo, R., Msika, A., Jumbe, A., Chome, N., Nyakuipa, D., Chintedza, J., Ballou, W.R., Bruls, M., Cohen, J., Guerra, Y., Jongert, E., Lapierre, D., Leach, A., Lievens, M., Ofori-Anyinam, O., Vekemans, J., Carter, T., Leboulleux, D., Loucq, C., Radford, A., Savarese, B., Schellenberg, D., Sillman, M., Vansadia, P., 2011. First results of phase 3 trial of RTS,S/AS01 malaria vaccine in African children. N Engl J Med 365, 1863-1875.

Aikawa, M., Miller, L.H., Johnson, J., Rabbege, J., 1978. Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. J Cell Biol 77, 72-82.

al-Khedery, B., Barnwell, J.W., Galinski, M.R., 1999. Antigenic variation in malaria: a 3' genomic alteration associated with the expression of a *P. knowlesi* variant antigen. Mol Cell 3, 131-141.

Alder, N.N., Shen, Y., Brodsky, J.L., Hendershot, L.M., Johnson, A.E., 2005. The molecular mechanisms underlying BiP-mediated gating of the Sec61 translocon of the endoplasmic reticulum. J Cell Biol 168, 389-399.

Alexander, D.L., Mital, J., Ward, G.E., Bradley, P., Boothroyd, J.C., 2005. Identification of the moving junction complex of *Toxoplasma gondii*: a collaboration between distinct secretory organelles. PLoS Pathog 1, e17.

Alexander, D.L., Arastu-Kapur, S., Dubremetz, J.F., Boothroyd, J.C., 2006. *Plasmodium falciparum* AMA1 binds a rhoptry neck protein homologous to TgRON4, a component of the moving junction in *Toxoplasma gondii*. Eukaryot Cell 5, 1169-1173.

Allen, P.C., Fetterer, R.H., 2002. Recent advances in biology and immunobiology of Eimeria species and in diagnosis and control of infection with these coccidian parasites of poultry. Clin Microbiol Rev 15, 58-65.

Allen, S.J., O'Donnell, A., Alexander, N.D., Mgone, C.S., Peto, T.E., Clegg, J.B., Alpers, M.P., Weatherall, D.J., 1999. Prevention of cerebral malaria in children in Papua New Guinea by southeast Asian ovalocytosis band 3. Am J Trop Med Hyg 60, 1056-1060.

Alonso, P.L., Brown, G., Arevalo-Herrera, M., Binka, F., Chitnis, C., Collins, F., Doumbo, O.K., Greenwood, B., Hall, B.F., Levine, M.M., Mendis, K., Newman, R.D., Plowe, C.V., Rodriguez, M.H., Sinden, R., Slutsker, L., Tanner, M., 2011. A research agenda to underpin malaria eradication. PLoS Med 8, e1000406.

Amino, R., Thiberge, S., Martin, B., Celli, S., Shorte, S., Frischknecht, F., Menard, R., 2006. Quantitative imaging of Plasmodium transmission from mosquito to mammal. Nat Med 12, 220-224.

Anders, R.F., Shi, P.T., Scanlon, D.B., Leach, S.J., Coppel, R.L., Brown, G.V., Stahl, H.D., Kemp, D.J., 1986. Antigenic repeat structures in proteins of *Plasmodium falciparum*. Ciba Found Symp 119, 164-183.

Arnot, D.E., Cavanagh, D.R., Remarque, E.J., Creasey, A.M., Sowa, M.P., Morgan, W.D., Holder, A.A., Longacre, S., Thomas, A.W., 2008. Comparative testing of six antigen-based malaria vaccine candidates directed toward merozoite-stage *Plasmodium falciparum*. Clin Vaccine Immunol 15, 1345-1355.

Awah, N., Balogun, H., Achidi, E., Mariuba, L.A., Nogueira, P.A., Orlandi, P., Troye-Blomberg, M., Gysin, J., Berzins, K., 2011. Antibodies to the *Plasmodium falciparum* rhoptry protein RAP-2/RSP-2 in relation to anaemia in Cameroonian children. Parasite Immunol 33, 104-115.

Awah, N.W., Troye-Blomberg, M., Berzins, K., Gysin, J., 2009. Mechanisms of malarial anaemia: potential involvement of the *Plasmodium falciparum* low molecular weight rhoptry-associated proteins. Acta Trop 112, 295-302.

Bachvaroff, T.R., Handy, S.M., Place, A.R., Delwiche, C.F., 2011. Alveolate phylogeny inferred using concatenated ribosomal proteins. J Eukaryot Microbiol 58, 223-233.

Baker, D.A., 2010. Malaria gametocytogenesis. Mol Biochem Parasitol 172, 57-65.

Banecki, B., Liberek, K., Wall, D., Wawrzynow, A., Georgopoulos, C., Bertoli, E., Tanfani, F., Zylicz, M., 1996. Structure-function analysis of the zinc finger region of the DnaJ molecular chaperone. J Biol Chem 271, 14840-14848.

Bannister, L., Mitchell, G., 2003. The ins, outs and roundabouts of malaria. Trends Parasitol 19, 209-213.

Bannister, L.H., Butcher, G.A., Dennis, E.D., Mitchell, G.H., 1975. Studies on the structure and invasive behaviour of merozoites of *Plasmodium knowlesi*. Trans R Soc Trop Med Hyg 69, 5.

Bannister, L.H., Mitchell, G.H., Butcher, G.A., Dennis, E.D., 1986. Lamellar membranes associated with rhoptries in erythrocytic merozoites of *Plasmodium knowlesi*: a clue to the mechanism of invasion. Parasitology 92 (Pt 2), 291-303.

Bannister, L.H., Dluzewski, A.R., 1990. The ultrastructure of red cell invasion in malaria infections: a review. Blood Cells 16, 257-292; discussion 293-257.

Bannister, L.H., Hopkins, J.M., Fowler, R.E., Krishna, S., Mitchell, G.H., 2000. A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages. Parasitol Today 16, 427-433.

Bannister, L.H., Hopkins, J.M., Dluzewski, A.R., Margos, G., Williams, I.T., Blackman, M.J., Kocken, C.H., Thomas, A.W., Mitchell, G.H., 2003. *Plasmodium falciparum* apical membrane antigen 1 (PfAMA-1) is translocated within micronemes along subpellicular microtubules during merozoite development. J Cell Sci 116, 3825-3834.

Bastin, P., Bagherzadeh, Z., Matthews, K.R., Gull, K., 1996. A novel epitope tag system to study protein targeting and organelle biogenesis in *Trypanosoma brucei*. Mol Biochem Parasitol 77, 235-239.

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T.W., Green, J.L., Holder, A.A., Cowman, A.F., 2006. A conserved molecular motor drives cell invasion and gliding motility across malaria life cycle stages and other apicomplexan parasites. J Biol Chem 281, 5197-5208.

Baum, J., Gilberger, T.W., Frischknecht, F., Meissner, M., 2008. Host-cell invasion by malaria parasites: insights from *Plasmodium* and *Toxoplasma*. Trends Parasitol 24, 557-563.

Baum, J., Chen, L., Healer, J., Lopaticki, S., Boyle, M., Triglia, T., Ehlgen, F., Ralph, S.A., Beeson, J.G., Cowman, A.F., 2009a. Reticulocyte-binding protein homologue 5 - an essential adhesin involved in invasion of human erythrocytes by *Plasmodium falciparum*. Int J Parasitol 39, 371-380.

Baum, J., Papenfuss, A.T., Mair, G.R., Janse, C.J., Vlachou, D., Waters, A.P., Cowman, A.F., Crabb, B.S., de Koning-Ward, T.F., 2009b. Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. Nucleic acids research 37, 3788-3798.

Bell, A., Wernli, B., Franklin, R.M., 1994. Roles of peptidyl-prolyl cis-trans isomerase and calcineurin in the mechanisms of antimalarial action of cyclosporin A, FK506, and rapamycin. Biochem Pharmacol 48, 495-503.

Bender, A., 2002. PlasMit. In, Prediction of mitochondrial transit peptides from *Plasmodium falciparum*.

Besteiro, S., Michelin, A., Poncet, J., Dubremetz, J.F., Lebrun, M., 2009. Export of a *Toxoplasma gondii* rhoptry neck protein complex at the host cell membrane to form the moving junction during invasion. PLoS Pathog 5, e1000309.

Blackman, M.J., Scott-Finnigan, T.J., Shai, S., Holder, A.A., 1994. Antibodies inhibit the protease-mediated processing of a malaria merozoite surface protein. J Exp Med 180, 389-393.

Blackman, M.J., 2004. Proteases in host cell invasion by the malaria parasite. Cell Microbiol 6, 893-903.

Blum, H., Beier, H., Gross, H.J., 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis 8, 93-99.

Bolanos-Garcia, V.M., Davies, O.R., 2006. Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. Biochim Biophys Acta 1760, 1304-1313.

Botha, M., Pesce, E.R., Blatch, G.L., 2007. The Hsp40 proteins of *Plasmodium falciparum* and other apicomplexa: regulating chaperone power in the parasite and the host. The international journal of biochemistry & cell biology 39, 1781-1803.

Bowyer, P.W., Simon, G.M., Cravatt, B.F., Bogyo, M., 2011. Global profiling of proteolysis during rupture of *P. falciparum* from the host erythrocyte. Mol Cell Proteomics.

Bozdech, Z., Llinas, M., Pulliam, B.L., Wong, E.D., Zhu, J., DeRisi, J.L., 2003. The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS Biol 1, E5.

Bradley, P.J., Ward, C., Cheng, S.J., Alexander, D.L., Coller, S., Coombs, G.H., Dunn, J.D., Ferguson, D.J., Sanderson, S.J., Wastling, J.M., Boothroyd, J.C., 2005. Proteomic analysis of rhoptry organelles reveals many novel constituents for host-parasite interactions in *Toxoplasma gondii*. J Biol Chem 280, 34245-34258.

Breinich, M.S., Ferguson, D.J., Foth, B.J., van Dooren, G.G., Lebrun, M., Quon, D.V., Striepen, B., Bradley, P.J., Frischknecht, F., Carruthers, V.B., Meissner, M., 2009. A dynamin is required for the biogenesis of secretory organelles in *Toxoplasma gondii*. Curr Biol 19, 277-286.

Brier, B., 2004. Infectious diseases in ancient Egypt. Infect Dis Clin North Am 18, 17-27.

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

Bruce, M.C., Carter, R.N., Nakamura, K., Aikawa, M., Carter, R., 1994. Cellular location and temporal expression of the *Plasmodium falciparum* sexual stage antigen Pfs16. Mol Biochem Parasitol 65, 11-22.

Bull, P.C., Lowe, B.S., Kortok, M., Molyneux, C.S., Newbold, C.I., Marsh, K., 1998. Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. Nat Med 4, 358-360.

Buscaglia, C.A., Coppens, I., Hol, W.G., Nussenzweig, V., 2003. Sites of interaction between aldolase and thrombospondin-related anonymous protein in plasmodium. Mol Biol Cell 14, 4947-4957.

Cabrera, A., Herrmann, S., Warszta, D., Santos, J.M., John Peter, A.T., Kono, M., Debrouver, S., Jacobs, T., Spielmann, T., Ungermann, C., Soldati-Favre, D., Gilberger, T.W., 2012. Dissection of Minimal Sequence Requirements for Rhoptry Membrane Targeting in the Malaria Parasite. Traffic.

Campbell, G.H., Miller, L.H., Hudson, D., Franco, E.L., Andrysiak, P.M., 1984. Monoclonal antibody characterization of *Plasmodium falciparum* antigens. Am J Trop Med Hyg 33, 1051-1054.

Canning, E.U., Sinden, R.E., 1973. The organization of the ookinete and observations on nuclear division in oocysts of *Plasmodium berghei.* Parasitology 67, 29-40.

Cao, J., Kaneko, O., Thongkukiatkul, A., Tachibana, M., Otsuki, H., Gao, Q., Tsuboi, T., Torii, M., 2009. Rhoptry neck protein RON2 forms a complex with microneme protein AMA1 in *Plasmodium falciparum* merozoites. Parasitol Int 58, 29-35.

Carlton, J.M., Angiuoli, S.V., Suh, B.B., Kooij, T.W., Pertea, M., Silva, J.C., Ermolaeva, M.D., Allen, J.E., Selengut, J.D., Koo, H.L., Peterson, J.D., Pop, M., Kosack, D.S., Shumway, M.F., Bidwell, S.L., Shallom, S.J., van Aken, S.E., Riedmuller, S.B., Feldblyum, T.V., Cho, J.K., Quackenbush, J., Sedegah, M., Shoaibi, A., Cummings, L.M., Florens, L., Yates, J.R., Raine, J.D., Sinden, R.E., Harris, M.A., Cunningham, D.A., Preiser, P.R., Bergman, L.W., Vaidya, A.B., van Lin, L.H., Janse, C.J., Waters, A.P., Smith, H.O., White, O.R., Salzberg, S.L., Venter, J.C., Fraser, C.M., Hoffman, S.L., Gardner, M.J., Carucci, D.J., 2002. Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. Nature 419, 512-519.

Carlton, J.M., Adams, J.H., Silva, J.C., Bidwell, S.L., Lorenzi, H., Caler, E., Crabtree, J., Angiuoli, S.V., Merino, E.F., Amedeo, P., Cheng, Q., Coulson, R.M., Crabb, B.S., Del Portillo, H.A., Essien, K., Feldblyum, T.V., Fernandez-Becerra, C., Gilson, P.R., Gueye, A.H., Guo, X., Kang'a, S., Kooij, T.W., Korsinczky, M., Meyer, E.V., Nene, V., Paulsen, I., White, O., Ralph, S.A., Ren, Q., Sargeant, T.J., Salzberg, S.L., Stoeckert, C.J., Sullivan, S.A., Yamamoto, M.M., Hoffman, S.L., Wortman, J.R., Gardner, M.J., Galinski, M.R., Barnwell, J.W., Fraser-Liggett, C.M., 2008. Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. Nature 455, 757-763.

Carruthers, V.B., Sibley, L.D., 1997. Sequential protein secretion from three distinct organelles of *Toxoplasma gondii* accompanies invasion of human fibroblasts. Eur J Cell Biol 73, 114-123.

Carruthers, V.B., Giddings, O.K., Sibley, L.D., 1999. Secretion of micronemal proteins is associated with *Toxoplasma* invasion of host cells. Cell Microbiol 1, 225-235.

Carruthers, V.B., Sibley, L.D., 1999. Mobilization of intracellular calcium stimulates microneme discharge in *Toxoplasma gondii*. Mol Microbiol 31, 421-428.

Cheetham, M.E., Caplan, A.J., 1998. Structure, function and evolution of DnaJ: conservation and adaptation of chaperone function. Cell stress & chaperones 3, 28-36.

Chen, F., Mackey, A.J., Stoeckert, C.J., Jr., Roos, D.S., 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Res 34, D363-368.

Chintapalli, V.R., Wang, J., Dow, J.A., 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. Nat Genet 39, 715-720.
Cogswell, F.B., Krotoski, W.A., Hollingdale, M.R., Gwadz, R.W., 1983. Identification of hypnozoites and tissue schizonts of *Plasmodium vivax* and *P. cynomolgi* by the immunoperoxidase method. Am J Trop Med Hyg 32, 1454-1455.
Collins, C.R., Withers-Martinez, C., Bentley, G.A., Batchelor, A.H., Thomas, A.W., Blackman, M.J., 2007. Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1. J Biol Chem 282, 7431-7441.
Collins, C.R., Withers-Martinez, C., Hackett, F., Blackman, M.J., 2009. An inhibitory antibody blocks interactions between components of the malarial invasion machinery. PLoS Pathog 5, e1000273.
Collins, T.J., 2007. ImageJ for microscopy. Biotechniques 43, 25-30.
Control, m.C.G.o.V., 2011. A research agenda for malaria eradication: vector control. PLoS medicine 8, e1000401.
Cooper, J.A., Ingram, L.T., Bushell, G.R., Fardoulys, C.A., Stenzel, D., Schofield, L., Saul, A.J., 1988. The 140/130/105 kilodalton protein complex in the rhoptries of *Plasmodium falciparum* consists of discrete polypeptides. Mol Biochem Parasitol 29, 251-260.
Corsi, A.K., Schekman, R., 1997. The lumenal domain of Sec63p stimulates the ATPase activity of BiP and mediates BiP recruitment to the translocon in *Saccharomyces cerevisiae*. J Cell Biol 137, 1483-1493.
Cortes, A., Carret, C., Kaneko, O., Yim Lim, B.Y., Ivens, A., Holder, A.A., 2007. Epigenetic silencing of *Plasmodium falciparum* genes linked to erythrocyte invasion. PLoS Pathog 3, e107.
Cosson, P., Letourneur, F., 1994. Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. Science 263, 1629-1631.
Cosson, P., Lefkir, Y., Demolliere, C., Letourneur, F., 1998. New COP1-binding motifs involved in ER retrieval. Embo J 17, 6863-6870.
Costa, F.T., Lopes, S.C., Ferrer, M., Leite, J.A., Martin-Jaular, L., Bernabeu, M., Nogueira, P.A., Mourao, M.P., Fernandez-Becerra, C., Lacerda, M.V., del Portillo, H., 2011. On cytoadhesion of *Plasmodium vivax*: raison d'etre? Mem Inst Oswaldo Cruz 106 Suppl 1, 79-84.
Couffin, S., Hernandez-Rivas, R., Blisnick, T., Mattei, D., 1998. Characterisation of PfSec61, a *Plasmodium falciparum* homologue of a component of the translocation machinery at the endoplasmic reticulum membrane of eukaryotic cells. Mol Biochem Parasitol 92, 89-98.
Cowman, A.F., Crabb, B.S., 2006. Invasion of red blood cells by malaria parasites. Cell 124, 755-766.
Cox-Singh, J., Davis, T.M., Lee, K.S., Shamsul, S.S., Matusop, A., Ratnam, S., Rahman, H.A., Conway, D.J., Singh, B., 2008. *Plasmodium knowlesi* malaria in humans is widely distributed and potentially life threatening. Clin Infect Dis 46, 165-171.
Crabb, B.S., Triglia, T., Waterkeyn, J.G., Cowman, A.F., 1997. Stable transgene expression in Plasmodium falciparum. Mol Biochem Parasitol 90, 131-144.
Culvenor, J.G., Day, K.P., Anders, R.F., 1991. *Plasmodium falciparum* ring-infected erythrocyte surface antigen is released from merozoite dense granules after erythrocyte invasion. Infect Immun 59, 1183-1187.
Cunha, C.B., Cunha, B.A., 2008. Brief history of the clinical diagnosis of malaria: from Hippocrates to Osler. J Vector Borne Dis 45, 194-199.
D'Silva, P.R., Schilke, B., Walter, W., Craig, E.A., 2005. Role of Pam16's degenerate J domain in protein import across the mitochondrial inner membrane. Proceedings of the National Academy of Sciences of the United States of America 102, 12419-12424.

Da Silva, E., Foley, M., Dluzewski, A.R., Murray, L.J., Anders, R.F., Tilley, L., 1994. The *Plasmodium falciparum* protein RESA interacts with the erythrocyte cytoskeleton and modifies erythrocyte thermal stability. Molecular and biochemical parasitology 66, 59-69.

Dame, J.B., Williams, J.L., McCutchan, T.F., Weber, J.L., Wirtz, R.A., Hockmeyer, W.T., Maloy, W.L., Haynes, J.D., Schneider, I., Roberts, D., et al., 1984. Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. Science 225, 593-599.

Das, A., Syin, C., Fujioka, H., Zheng, H., Goldman, N., Aikawa, M., Kumar, N., 1997. Molecular characterization and ultrastructural localization of *Plasmodium falciparum* Hsp 60. Mol Biochem Parasitol 88, 95-104.

Day, K.P., Marsh, K., 1991. Naturally acquired immunity to *Plasmodium falciparum*. Immunol Today 12, A68-71.

de Castro, F.A., Ward, G.E., Jambou, R., Attal, G., Mayau, V., Jaureguiberry, G., Braun-Breton, C., Chakrabarti, D., Langsley, G., 1996. Identification of a family of Rab G-proteins in *Plasmodium falciparum* and a detailed characterisation of pfrab6. Mol Biochem Parasitol 80, 77-88.

Deshaies, R.J., Sanders, S.L., Feldheim, D.A., Schekman, R., 1991. Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. Nature 349, 806-808.

Desmonts, G., Couvreur, J., 1974. Toxoplasmosis in pregnancy and its transmission to the fetus. Bull N Y Acad Med 50, 146-159.

Di Cristina, M., Spaccapelo, R., Soldati, D., Bistoni, F., Crisanti, A., 2000. Two conserved amino acid motifs mediate protein targeting to the micronemes of the apicomplexan parasite *Toxoplasma gondii*. Mol Cell Biol 20, 7332-7341.

Dluzewski, A.R., Ling, I.T., Rangachari, K., Bates, P.A., Wilson, R.J., 1984. A simple method for isolating viable mature parasites of *Plasmodium falciparum* from cultures. Trans R Soc Trop Med Hyg 78, 622-624.

Dluzewski, A.R., Ling, I.T., Hopkins, J.M., Grainger, M., Margos, G., Mitchell, G.H., Holder, A.A., Bannister, L.H., 2008. Formation of the food vacuole in *Plasmodium falciparum*: a potential role for the 19 kDa fragment of merozoite surface protein 1 (MSP1(19)). PLoS One 3, e3085.

Dolan, S.A., Miller, L.H., Wellems, T.E., 1990. Evidence for a switching mechanism in the invasion of erythrocytes by Plasmodium falciparum. J Clin Invest 86, 618-624.

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., Lindegardh, N., Socheat, D., White, N.J., 2009. Artemisinin resistance in *Plasmodium falciparum* malaria. N Engl J Med 361, 455-467.

Doury, J.C., Bonnefoy, S., Roger, N., Dubremetz, J.F., Mercereau-Puijalon, O., 1994. Analysis of the high molecular weight rhoptry complex of *Plasmodium falciparum* using monoclonal antibodies. Parasitology 108 (Pt 3), 269-280.

Dubey, J.P., 2003. Review of *Neospora caninum* and neosporosis in animals. Korean J Parasitol 41, 1-16.

Duraisingh, M.T., Triglia, T., Cowman, A.F., 2002. Negative selection of *Plasmodium falciparum* reveals targeted gene deletion by double crossover recombination. Int J Parasitol 32, 81-89.

Duraisingh, M.T., Triglia, T., Ralph, S.A., Rayner, J.C., Barnwell, J.W., McFadden, G.I., Cowman, A.F., 2003. Phenotypic variation of *Plasmodium falciparum* merozoite proteins directs receptor targeting for invasion of human erythrocytes. EMBO J 22, 1047-1057.

Dvorak, J.A., Miller, L.H., Whitehouse, W.C., Shiroishi, T., 1975. Invasion of erythrocytes by malaria merozoites. Science 187, 748-750.

Dvorin, J.D., Martyn, D.C., Patel, S.D., Grimley, J.S., Collins, C.R., Hopp, C.S., Bright, A.T., Westenberger, S., Winzeler, E., Blackman, M.J., Baker, D.A., Wandless, T.J., Duraisingh, M.T., 2010. A plant-like kinase in *Plasmodium falciparum* regulates parasite egress from erythrocytes. Science 328, 910-912.

Elliott, D.A., McIntosh, M.T., Hosgood, H.D., 3rd, Chen, S., Zhang, G., Baevova, P., Joiner, K.A., 2008. Four distinct pathways of hemoglobin uptake in the malaria parasite Plasmodium falciparum. Proc Natl Acad Sci U S A 105, 2463-2468.

Elmendorf, H.G., Haldar, K., 1993. Identification and localization of ERD2 in the malaria parasite *Plasmodium falciparum*: separation from sites of sphingomyelin synthesis and implications for organization of the Golgi. Embo J 12, 4763-4773.

Eschenbacher, K.H., Klein, H., Sommer, I., Meyer, H.E., Entzeroth, R., Mehlhorn, H., Ruger, W., 1993. Characterization of cDNA clones encoding a major microneme antigen of *Sarcocystis muris* (Apicomplexa) cyst merozoites. Mol Biochem Parasitol 62, 27-36.

Evans, K.J., Hansen, D.S., van Rooijen, N., Buckingham, L.A., Schofield, L., 2006. Severe malarial anemia of low parasite burden in rodent models results from accelerated clearance of uninfected erythrocytes. Blood 107, 1192-1199.

Fang, H., Green, N., 1994. Nonlethal sec71-1 and sec72-1 mutations eliminate proteins associated with the Sec63p-BiP complex from S. cerevisiae. Mol Biol Cell 5, 933-942.

Fatih, F.A., Siner, A., Ahmed, A., Woon, L.C., Craig, A.G., Singh, B., Krishna, S., Cox-Singh, J., 2012. Cytoadherence and virulence - the case of Plasmodium knowlesi malaria. Malar J 11, 33.

Feldheim, D., Rothblatt, J., Schekman, R., 1992. Topology and functional domains of Sec63p, an endoplasmic reticulum membrane protein required for secretory protein translocation. Mol Cell Biol 12, 3288-3296.

Feldheim, D., Yoshimura, K., Admon, A., Schekman, R., 1993. Structural and functional characterization of Sec66p, a new subunit of the polypeptide translocation apparatus in the yeast endoplasmic reticulum. Molecular biology of the cell 4, 931-939.

Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L., Bateman, A., 2008. The Pfam protein families database. Nucleic Acids Res 36, D281-288.

Fischer, G., Aumuller, T., 2003. Regulation of peptide bond cis/trans isomerization by enzyme catalysis and its implication in physiological processes. Reviews of physiology, biochemistry and pharmacology 148, 105-150.

Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., Grainger, M., Haynes, J.D., Moch, J.K., Muster, N., Sacci, J.B., Tabb, D.L., Witney, A.A., Wolters, D., Wu, Y., Gardner, M.J., Holder, A.A., Sinden, R.E., Yates, J.R., Carucci, D.J., 2002. A proteomic view of the *Plasmodium falciparum* life cycle. Nature 419, 520-526.

Foley, M., Tilley, L., Sawyer, W.H., Anders, R.F., 1991. The ring-infected erythrocyte surface antigen of *Plasmodium falciparum* associates with spectrin in the erythrocyte membrane. Mol Biochem Parasitol 46, 137-147.

Formstecher, E., Aresta, S., Collura, V., Hamburger, A., Meil, A., Trehin, A., Reverdy, C., Betin, V., Maire, S., Brun, C., Jacq, B., Arpin, M., Bellaiche, Y., Bellusci, S., Benaroch, P., Bornens, M., Chanet, R., Chavrier, P., Delattre, O., Doye, V., Fehon, R., Faye, G., Galli, T., Girault, J.A., Goud, B., de Gunzburg, J., Johannes, L., Junier, M.P., Mirouse, V., Mukherjee, A., Papadopoulo, D., Perez, F., Plessis, A., Rosse, C., Saule, S., Stoppa-Lyonnet, D., Vincent, A., White, M., Legrain, P., Wojcik, J., Camonis, J., Daviet, L., 2005. Protein interaction mapping: a *Drosophila* case study. Genome Res 15, 376-384.

Foth, B.J., Ralph, S.A., Tonkin, C.J., Struck, N.S., Fraunholz, M., Roos, D.S., Cowman, A.F., McFadden, G.I., 2003. Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. Science 299, 705-708.

Fourmaux, M.N., Achbarou, A., Mercereau-Puijalon, O., Biderre, C., Briche, I., Loyens, A., Odberg-Ferragut, C., Camus, D., Dubremetz, J.F., 1996. The MIC1 microneme protein of *Toxoplasma gondii* contains a duplicated receptor-like domain and binds to host cell surface. Mol Biochem Parasitol 83, 201-210.

Frankel, M.B., Knoll, L.J., 2009. The ins and outs of nuclear trafficking: unusual aspects in apicomplexan parasites. DNA Cell Biol 28, 277-284.

Fraser, T.S., Kappe, S.H., Narum, D.L., VanBuskirk, K.M., Adams, J.H., 2001. Erythrocyte-binding activity of *Plasmodium yoelii* apical membrane antigen-1 expressed on the surface of transfected COS-7 cells. Mol Biochem Parasitol 117, 49-59.

Frazier, A.E., Dudek, J., Guiard, B., Voos, W., Li, Y., Lind, M., Meisinger, C., Geissler, A., Sickmann, A., Meyer, H.E., Bilanchone, V., Cumsky, M.G., Truscott, K.N., Pfanner, N., Rehling, P., 2004. Pam16 has an essential role in the mitochondrial protein import motor. Nat Struct Mol Biol 11, 226-233.

Freeman, M., 2008. Rhomboid proteases and their biological functions. Annu Rev Genet 42, 191-210.

Frenal, K., Polonais, V., Marq, J.B., Stratmann, R., Limenitakis, J., Soldati-Favre, D., 2010. Functional dissection of the apicomplexan glideosome molecular architecture. Cell Host Microbe 8, 343-357.

Fu, J., Saenz, F.E., Reed, M.B., Balu, B., Singh, N., Blair, P.L., Cowman, A.F., Adams, J.H., 2005. Targeted disruption of maebl in *Plasmodium falciparum*. Mol Biochem Parasitol 141, 113-117.

Gallup, J.L., Sachs, J.D., 2001. The economic burden of malaria. The American journal of tropical medicine and hygiene 64, 85-96.

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., 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., Barrell, B., 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419, 498-511.

Garnham, P.C.C., 1966. Malaria parasites and other haemosporidia. Blackwell Scientific, Oxford,.

Gasteiger E., H.C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A, 2005. Protein Identification and Analysis Tools on the ExPASy Server. Humana Press. .

Gaynor, E.C., te Heesen, S., Graham, T.R., Aebi, M., Emr, S.D., 1994. Signal-mediated retrieval of a membrane protein from the Golgi to the ER in yeast. J Cell Biol 127, 653-665.

Genton, B., al-Yaman, F., Mgone, C.S., Alexander, N., Paniu, M.M., Alpers, M.P., Mokela, D., 1995. Ovalocytosis and cerebral malaria. Nature 378, 564-565.

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

Ghoneim, A., Kaneko, O., Tsuboi, T., Torii, M., 2007. The *Plasmodium falciparum* RhopH2 promoter and first 24 amino acids are sufficient to target proteins to the rhoptries. Parasitol Int 56, 31-43.

Gilberger, T.W., Thompson, J.K., Triglia, T., Good, R.T., Duraisingh, M.T., Cowman, A.F., 2003. A novel erythrocyte binding antigen-175 paralogue from *Plasmodium falciparum* defines a new trypsin-resistant receptor on human erythrocytes. J Biol Chem 278, 14480-14486.

Gilson, P.R., Crabb, B.S., 2009. Morphology and kinetics of the three distinct phases of red blood cell invasion by *Plasmodium falciparum* merozoites. Int J Parasitol 39, 91-96.
Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C.A., Finley, R.L., Jr., White, K.P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R.A., McKenna, M.P., Chant, J., Rothberg, J.M., 2003. A protein interaction map of *Drosophila melanogaster*. Science 302, 1727-1736.
Giovannini, D., Spath, S., Lacroix, C., Perazzi, A., Bargieri, D., Lagal, V., Lebugle, C., Combe, A., Thiberge, S., Baldacci, P., Tardieux, I., Menard, R., 2011. Independent roles of apical membrane antigen 1 and rhoptry neck proteins during host cell invasion by apicomplexa. Cell Host Microbe 10, 591-602.
Goel, V.K., Li, X., Chen, H., Liu, S.C., Chishti, A.H., Oh, S.S., 2003. Band 3 is a host receptor binding merozoite surface protein 1 during the Plasmodium falciparum invasion of erythrocytes. Proc Natl Acad Sci U S A 100, 5164-5169.
Goldshmidt, H., Sheiner, L., Butikofer, P., Roditi, I., Uliel, S., Gunzel, M., Engstler, M., Michaeli, S., 2008. Role of protein translocation pathways across the endoplasmic reticulum in Trypanosoma brucei. J Biol Chem 283, 32085-32098.
Gomez-Escobar, N., Amambua-Ngwa, A., Walther, M., Okebe, J., Ebonyi, A., Conway, D.J., 2010. Erythrocyte invasion and merozoite ligand gene expression in severe and mild *Plasmodium falciparum* malaria. J Infect Dis 201, 444-452.
Grassi, B., Noe, G., 1900. The Propagation of the Filariae of the Blood Exclusively by means of the Puncture of Peculiar Mosquitos. Br Med J 2, 1306-1307.
Green, J.L., Hinds, L., Grainger, M., Knuepfer, E., Holder, A.A., 2006a. *Plasmodium* thrombospondin related apical merozoite protein (PTRAMP) is shed from the surface of merozoites by PfSUB2 upon invasion of erythrocytes. Mol Biochem Parasitol 150, 114-117.
Green, J.L., Martin, S.R., Fielden, J., Ksagoni, A., Grainger, M., Yim Lim, B.Y., Molloy, J.E., Holder, A.A., 2006b. The MTIP-myosin A complex in blood stage malaria parasites. J Mol Biol 355, 933-941.
Greene, M.K., Maskos, K., Landry, S.J., 1998. Role of the J-domain in the cooperation of Hsp40 with Hsp70. Proc Natl Acad Sci U S A 95, 6108-6113.
Hageman, J., Kampinga, H.H., 2009. Computational analysis of the human HSPH/HSPA/DNAJ family and cloning of a human HSPH/HSPA/DNAJ expression library. Cell stress & chaperones 14, 1-21.
Hamamichi, S., Rivas, R.N., Knight, A.L., Cao, S., Caldwell, K.A., Caldwell, G.A., 2008. Hypothesis-based RNAi screening identifies neuroprotective genes in a Parkinson's disease model. Proc Natl Acad Sci U S A 105, 728-733.
Harada, Y., Li, H., Wall, J.S., Li, H., Lennarz, W.J., 2011. Structural studies and the assembly of the heptameric post-translational translocon complex. J Biol Chem 286, 2956-2965.
Harris, P.K., Yeoh, S., Dluzewski, A.R., O'Donnell, R.A., Withers-Martinez, C., Hackett, F., Bannister, L.H., Mitchell, G.H., Blackman, M.J., 2005. Molecular identification of a malaria merozoite surface sheddase. PLoS Pathog 1, 241-251.
Hawass, Z., Gad, Y.Z., Ismail, S., Khairat, R., Fathalla, D., Hasan, N., Ahmed, A., Elleithy, H., Ball, M., Gaballah, F., Wasef, S., Fateen, M., Amer, H., Gostner, P., Selim, A., Zink, A., Pusch, C.M., 2010. Ancestry and pathology in King Tutankhamun's family. JAMA 303, 638-647.
Hawking, F., Worms, M.J., Gammage, K., 1968. Host temperature and control of 24-hour and 48-hour cycles in malaria parasites. Lancet 1, 506-509.

Healer, J., Crawford, S., Ralph, S., McFadden, G., Cowman, A.F., 2002. Independent translocation of two micronemal proteins in developing *Plasmodium falciparum* merozoites. Infect Immun 70, 5751-5758.

Hienne, R., Ricard, G., Fusai, T., Fujioka, H., Pradines, B., Aikawa, M., Doury, J.C., 1998. *Plasmodium yoelii*: identification of rhoptry proteins using monoclonal antibodies. Exp Parasitol 90, 230-235.

Hiller, N.L., Akompong, T., Morrow, J.S., Holder, A.A., Haldar, K., 2003. Identification of a stomatin orthologue in vacuoles induced in human erythrocytes by malaria parasites. A role for microbial raft proteins in apicomplexan vacuole biogenesis. J Biol Chem 278, 48413-48421.

Holder, A.A., Freeman, R.R., 1984a. The three major antigens on the surface of *Plasmodium falciparum* merozoites are derived from a single high molecular weight precursor. J Exp Med 160, 624-629.

Holder, A.A., Freeman, R.R., 1984b. Characterization of a high molecular weight protective antigen of *Plasmodium yoelii*. Parasitology 88 (Pt 2), 211-219.

Holder, A.A., Freeman, R.R., Uni, S., Aikawa, M., 1985. Isolation of a *Plasmodium falciparum* rhoptry protein. Mol Biochem Parasitol 14, 293-303.

Holder, A.A., Sandhu, J.S., Hillman, Y., Davey, L.S., Nicholls, S.C., Cooper, H., Lockyer, M.J., 1987. Processing of the precursor to the major merozoite surface antigens of *Plasmodium falciparum*. Parasitology 94 (Pt 2), 199-208.

Holmgren, A., 1985. Thioredoxin. Annu Rev Biochem 54, 237-271.

Honsho, M., Mitoma, J.Y., Ito, A., 1998. Retention of cytochrome b5 in the endoplasmic reticulum is transmembrane and luminal domain-dependent. J Biol Chem 273, 20860-20866.

Hoppe, H.C., Ngo, H.M., Yang, M., Joiner, K.A., 2000. Targeting to rhoptry organelles of *Toxoplasma gondii* involves evolutionarily conserved mechanisms. Nature cell biology 2, 449-456.

Howell, S.A., Hackett, F., Jongco, A.M., Withers-Martinez, C., Kim, K., Carruthers, V.B., Blackman, M.J., 2005. Distinct mechanisms govern proteolytic shedding of a key invasion protein in apicomplexan pathogens. Mol Microbiol 57, 1342-1356.

Hu, G., Cabrera, A., Kono, M., Mok, S., Chaal, B.K., Haase, S., Engelberg, K., Cheemadan, S., Spielmann, T., Preiser, P.R., Gilberger, T.W., Bozdech, Z., 2010. Transcriptional profiling of growth perturbations of the human malaria parasite *Plasmodium falciparum*. Nat Biotechnol 28, 91-98.

Jackson, M.R., Nilsson, T., Peterson, P.A., 1993. Retrieval of transmembrane proteins to the endoplasmic reticulum. J Cell Biol 121, 317-333.

Jaikaria, N.S., Rozario, C., Ridley, R.G., Perkins, M.E., 1993. Biogenesis of rhoptry organelles in *Plasmodium falciparum*. Mol Biochem Parasitol 57, 269-279.

Jarolim, P., Palek, J., Amato, D., Hassan, K., Sapak, P., Nurse, G.T., Rubin, H.L., Zhai, S., Sahr, K.E., Liu, S.C., 1991. Deletion in erythrocyte band 3 gene in malaria-resistant Southeast Asian ovalocytosis. Proc Natl Acad Sci U S A 88, 11022-11026.

Jewett, T.J., Sibley, L.D., 2003. Aldolase forms a bridge between cell surface adhesins and the actin cytoskeleton in apicomplexan parasites. Mol Cell 11, 885-894.

K. Hofmann, W.S., 1993. TMBASE - A database of membrane spanning protein segments. In, Biol. Chem. , Hoppe-Seyler, pp.

374,166

Kaneko, O., Tsuboi, T., Ling, I.T., Howell, S., Shirano, M., Tachibana, M., Cao, Y.M., Holder, A.A., Torii, M., 2001. The high molecular mass rhoptry protein, RhopH1, is encoded by members of the clag multigene family in *Plasmodium falciparum* and *Plasmodium yoelii*. Mol Biochem Parasitol 118, 223-231.

Kanzok, S.M., Schirmer, R.H., Turbachova, I., Iozef, R., Becker, K., 2000. The thioredoxin system of the malaria parasite *Plasmodium falciparum*. Glutathione reduction revisited. J Biol Chem 275, 40180-40186.

Kappe, S.H., Noe, A.R., Fraser, T.S., Blair, P.L., Adams, J.H., 1998. A family of chimeric erythrocyte binding proteins of malaria parasites. Proc Natl Acad Sci U S A 95, 1230-1235.

Kato, K., Mayer, D.C., Singh, S., Reid, M., Miller, L.H., 2005. Domain III of *Plasmodium falciparum* apical membrane antigen 1 binds to the erythrocyte membrane protein Kx. Proc Natl Acad Sci U S A 102, 5552-5557.

Kauth, C.W., Woehlbier, U., Kern, M., Mekonnen, Z., Lutz, R., Mucke, N., Langowski, J., Bujard, H., 2006. Interactions between merozoite surface proteins 1, 6, and 7 of the malaria parasite *Plasmodium falciparum*. J Biol Chem 281, 31517-31527.

Khan, D.S., 2008. Leiden Malaria Group merozoite proteome. In, Leids Universitair Medisch Centrum, Leiden.

Kim, K., Weiss, L.M., 2004. *Toxoplasma gondii*: the model apicomplexan. Int J Parasitol 34, 423-432.

Kirkman, L.A., Su, X.Z., Wellems, T.E., 1996. *Plasmodium falciparum:* isolation of large numbers of parasite clones from infected blood samples. Exp Parasitol 83, 147-149.

Knuepfer, E., Rug, M., Klonis, N., Tilley, L., Cowman, A.F., 2005. Trafficking determinants for PfEMP3 export and assembly under the *Plasmodium falciparum*-infected red blood cell membrane. Mol Microbiol 58, 1039-1053.

Koussis, K., Withers-Martinez, C., Yeoh, S., Child, M., Hackett, F., Knuepfer, E., Juliano, L., Woehlbier, U., Bujard, H., Blackman, M.J., 2009. A multifunctional serine protease primes the malaria parasite for red blood cell invasion. EMBO J.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305, 567-580.

Krotoski, W.A., Collins, W.E., Bray, R.S., Garnham, P.C., Cogswell, F.B., Gwadz, R.W., Killick-Kendrick, R., Wolf, R., Sinden, R., Koontz, L.C., Stanfill, P.S., 1982. Demonstration of hypnozoites in sporozoite-transmitted *Plasmodium vivax* infection. Am J Trop Med Hyg 31, 1291-1293.

Kumar, N., Syin, C.A., Carter, R., Quakyi, I., Miller, L.H., 1988. *Plasmodium falciparum* gene encoding a protein similar to the 78-kDa rat glucose-regulated stress protein. Proc Natl Acad Sci U S A 85, 6277-6281.

Kumar, N., Koski, G., Harada, M., Aikawa, M., Zheng, H., 1991. Induction and localization of *Plasmodium falciparum* stress proteins related to the heat shock protein 70 family. Mol Biochem Parasitol 48, 47-58.

Kumar, N., Zheng, H., 1992. Nucleotide sequence of a *Plasmodium falciparum* stress protein with similarity to mammalian 78-kDa glucose-regulated protein. Mol Biochem Parasitol 56, 353-356.

Kyes, S., Pinches, R., Newbold, C., 2000. A simple RNA analysis method shows var and rif multigene family expression patterns in *Plasmodium falciparum*. Mol Biochem Parasitol 105, 311-315.

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J Mol Biol 157, 105-132.

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., Hughes, R.E., 2005. A protein interaction network of the malaria parasite *Plasmodium falciparum*. Nature 438, 103-107.

Ladda, R., Aikawa, M., Sprinz, H., 1969. Penetration of erythrocytes by merozoites of mammalian and avian malarial parasites. J Parasitol 55, 633-644.

Lal, K., Prieto, J.H., Bromley, E., Sanderson, S.J., Yates, J.R., 3rd, Wastling, J.M., Tomley, F.M., Sinden, R.E., 2009. Characterisation of *Plasmodium* invasive organelles; an ookinete microneme proteome. Proteomics 9, 1142-1151.

Lamarque, M., Besteiro, S., Papoin, J., Roques, M., Vulliez-Le Normand, B., Morlon-Guyot, J., Dubremetz, J.F., Fauquenoy, S., Tomavo, S., Faber, B.W., Kocken, C.H., Thomas, A.W., Boulanger, M.J., Bentley, G.A., Lebrun, M., 2011. The RON2-AMA1 Interaction is a Critical Step in Moving Junction-Dependent Invasion by Apicomplexan Parasites. PLoS Pathog 7, e1001276.

Lambros, C., Vanderberg, J.P., 1979. Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. J Parasitol 65, 418-420.

Langer, R.C., Hayward, R.E., Tsuboi, T., Tachibana, M., Torii, M., Vinetz, J.M., 2000. Micronemal transport of *Plasmodium* ookinete chitinases to the electron-dense area of the apical complex for extracellular secretion. Infect Immun 68, 6461-6465.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947-2948.

Lasonder, E., Janse, C.J., van Gemert, G.J., Mair, G.R., Vermunt, A.M., Douradinha, B.G., van Noort, V., Huynen, M.A., Luty, A.J., Kroeze, H., Khan, S.M., Sauerwein, R.W., Waters, A.P., Mann, M., Stunnenberg, H.G., 2008. Proteomic profiling of *Plasmodium* sporozoite maturation identifies new proteins essential for parasite development and infectivity. PLoS Pathog 4, e1000195.

Lasonder, E., Treeck, M., Alam, M., Tobin, A.B., 2012. Insights into the Plasmodium falciparum schizont phospho-proteome. Microbes Infect 14, 811-819.

Laufen, T., Mayer, M.P., Beisel, C., Klostermeier, D., Mogk, A., Reinstein, J., Bukau, B., 1999. Mechanism of regulation of hsp70 chaperones by DnaJ cochaperones. Proc Natl Acad Sci U S A 96, 5452-5457.

Layez, C., Nogueira, P., Combes, V., Costa, F.T., Juhan-Vague, I., da Silva, L.H., Gysin, J., 2005. *Plasmodium falciparum* rhoptry protein RSP2 triggers destruction of the erythroid lineage. Blood 106, 3632-3638.

Le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S.F., Williamson, K.C., Holder, A.A., Carucci, D.J., Yates, J.R., 3rd, Winzeler, E.A., 2004. Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. Genome Res 14, 2308-2318.

Lebrun, M., Michelin, A., El Hajj, H., Poncet, J., Bradley, P.J., Vial, H., Dubremetz, J.F., 2005. The rhoptry neck protein RON4 re-localizes at the moving junction during *Toxoplasma gondii* invasion. Cell Microbiol 7, 1823-1833.

Letourneur, F., Gaynor, E.C., Hennecke, S., Demolliere, C., Duden, R., Emr, S.D., Riezman, H., Cosson, P., 1994. Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. Cell 79, 1199-1207.

Lewis, M.J., Sweet, D.J., Pelham, H.R., 1990. The ERD2 gene determines the specificity of the luminal ER protein retention system. Cell 61, 1359-1363.

Li, L., Stoeckert, C.J., Jr., Roos, D.S., 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13, 2178-2189.

Li, X., Chen, H., Oo, T.H., Daly, T.M., Bergman, L.W., Liu, S.C., Chishti, A.H., Oh, S.S., 2004a. A co-ligand complex anchors Plasmodium falciparum merozoites to the erythrocyte invasion receptor band 3. J Biol Chem 279, 5765-5771.

Li, Y., Dudek, J., Guiard, B., Pfanner, N., Rehling, P., Voos, W., 2004b. The presequence translocase-associated protein import motor of mitochondria. Pam16 functions in an antagonistic manner to Pam18. The Journal of biological chemistry 279, 38047-38054.

Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C., Zylicz, M., 1991. *Escherichia coli* DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. Proc Natl Acad Sci U S A 88, 2874-2878.

Ling, I.T., Kaneko, O., Narum, D.L., Tsuboi, T., Howell, S., Taylor, H.M., Scott-Finnigan, T.J., Torii, M., Holder, A.A., 2003. Characterisation of the rhoph2 gene of *Plasmodium falciparum* and Plasmodium yoelii. Mol Biochem Parasitol 127, 47-57.

Ling, I.T., Florens, L., Dluzewski, A.R., Kaneko, O., Grainger, M., Yim Lim, B.Y., Tsuboi, T., Hopkins, J.M., Johnson, J.R., Torii, M., Bannister, L.H., Yates, J.R., 3rd, Holder, A.A., Mattei, D., 2004. The *Plasmodium falciparum* clag9 gene encodes a rhoptry protein that is transferred to the host erythrocyte upon invasion. Mol Microbiol 52, 107-118.

Lingelbach, K., Joiner, K.A., 1998. The parasitophorous vacuole membrane surrounding *Plasmodium* and *Toxoplasma*: an unusual compartment in infected cells. J Cell Sci 111 (Pt 11), 1467-1475.

Lovett, J.L., Marchesini, N., Moreno, S.N., Sibley, L.D., 2002. *Toxoplasma gondii* microneme secretion involves intracellular Ca(2+) release from inositol 1,4,5-triphosphate (IP(3))/ryanodine-sensitive stores. J Biol Chem 277, 25870-25876.

Luft, B.J., Hafner, R., Korzun, A.H., Leport, C., Antoniskis, D., Bosler, E.M., Bourland, D.D., 3rd, Uttamchandani, R., Fuhrer, J., Jacobson, J., et al., 1993. Toxoplasmic encephalitis in patients with the acquired immunodeficiency syndrome. Members of the ACTG 077p/ANRS 009 Study Team. N Engl J Med 329, 995-1000.

Maier, A.G., Duraisingh, M.T., Reeder, J.C., Patel, S.S., Kazura, J.W., Zimmerman, P.A., Cowman, A.F., 2003. *Plasmodium falciparum* erythrocyte invasion through glycophorin C and selection for Gerbich negativity in human populations. Nat Med 9, 87-92.

malERA_Consultative_Group_on_Drugs, 2011. A research agenda for malaria eradication: drugs. PLoS Med 8, e1000402.

malERA_Consultative_Group_on_Vector_Control, 2011. A research agenda for malaria eradication: vector control. PLoS medicine 8, e1000401.

Manson, P., 1898. SURGEON-MAJOR RONALD ROSS'S RECENT INVESTIGATIONS on the MOSQUITO-MALARIA THEORY. Br Med J 1, 1575-1577.

Manson, P., 1900. Experimental Proof of the Mosquitomalaria Theory. Br Med J 2, 949-951.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., Bryant, S.H., 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res 39, D225-229.

Marin-Menendez, A., Monaghan, P., Bell, A., 2012. A family of cyclophilin-like molecular chaperones in *Plasmodium falciparum*. Mol Biochem Parasitol 184, 44-47.

Marsh, K., Snow, R.W., 1997. Host-parasite interaction and morbidity in malaria endemic areas. Philos Trans R Soc Lond B Biol Sci 352, 1385-1394.

Marti, M., Good, R.T., Rug, M., Knuepfer, E., Cowman, A.F., 2004. Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science 306, 1930-1933.

Matsuda, S., Koyasu, S., 2000. Mechanisms of action of cyclosporine. Immunopharmacology 47, 119-125.

Mayer, D.C., Cofie, J., Jiang, L., Hartl, D.L., Tracy, E., Kabat, J., Mendoza, L.H., Miller, L.H., 2009. Glycophorin B is the erythrocyte receptor of *Plasmodium falciparum* erythrocyte-binding ligand, EBL-1. Proc Natl Acad Sci U S A 106, 5348-5352.

Mayer, M.P., Laufen, T., Paal, K., McCarty, J.S., Bukau, B., 1999. Investigation of the interaction between DnaK and DnaJ by surface plasmon resonance spectroscopy. J Mol Biol 289, 1131-1144.

McBride, J.S., Heidrich, H.G., 1987. Fragments of the polymorphic Mr 185,000 glycoprotein from the surface of isolated *Plasmodium falciparum* merozoites form an antigenic complex. Mol Biochem Parasitol 23, 71-84.

Meek, B., Back, J.W., Klaren, V.N., Speijer, D., Peek, R., 2002. Protein disulfide isomerase of Toxoplasma gondii is targeted by mucosal IgA antibodies in humans. FEBS letters 522, 104-108.

Meissner, M., Reiss, M., Viebig, N., Carruthers, V.B., Toursel, C., Tomavo, S., Ajioka, J.W., Soldati, D., 2002. A family of transmembrane microneme proteins of *Toxoplasma gondii* contain EGF-like domains and function as escorters. J Cell Sci 115, 563-574.

Mendez, F., Carrasquilla, G., Munoz, A., 2000. Risk factors associated with malaria infection in an urban setting. Trans R Soc Trop Med Hyg 94, 367-371.

Meyer, D.I., Dobberstein, B., 1980a. Identification and characterization of a membrane component essential for the translocation of nascent proteins across the membrane of the endoplasmic reticulum. J Cell Biol 87, 503-508.

Meyer, D.I., Dobberstein, B., 1980b. A membrane component essential for vectorial translocation of nascent proteins across the endoplasmic reticulum: requirements for its extraction and reassociation with the membrane. J Cell Biol 87, 498-502.

Miller, L.H., Mason, S.J., Clyde, D.F., McGinniss, M.H., 1976. The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy. N Engl J Med 295, 302-304.

Miller, L.H., Haynes, J.D., McAuliffe, F.M., Shiroishi, T., Durocher, J.R., McGinniss, M.H., 1977. Evidence for differences in erythrocyte surface receptors for the malarial parasites, *Plasmodium falciparum* and *Plasmodium knowlesi.* J Exp Med 146, 277-281.

Miller, L.H., Aikawa, M., Johnson, J.G., Shiroishi, T., 1979. Interaction between cytochalasin B-treated malarial parasites and erythrocytes. Attachment and junction formation. J Exp Med 149, 172-184.

Mills, J.P., Diez-Silva, M., Quinn, D.J., Dao, M., Lang, M.J., Tan, K.S., Lim, C.T., Milon, G., David, P.H., Mercereau-Puijalon, O., Bonnefoy, S., Suresh, S., 2007. Effect of plasmodial RESA protein on deformability of human red blood cells harboring *Plasmodium falciparum.* Proc Natl Acad Sci U S A 104, 9213-9217.

Mills, K.E., Pearce, J.A., Crabb, B.S., Cowman, A.F., 2002. Truncation of merozoite surface protein 3 disrupts its trafficking and that of acidic-basic repeat protein to the surface of *Plasmodium falciparum* merozoites. Mol Microbiol 43, 1401-1411.

Mitchell, G.H., Thomas, A.W., Margos, G., Dluzewski, A.R., Bannister, L.H., 2004. Apical membrane antigen 1, a major malaria vaccine candidate, mediates the close attachment of invasive merozoites to host red blood cells. Infect Immun 72, 154-158.

Montero, E., Rodriguez, M., Oksov, Y., Lobo, C.A., 2009. *Babesia divergens* apical membrane antigen 1 and its interaction with the human red blood cell. Infect Immun 77, 4783-4793.

Morahan, B.J., Wang, L., Coppel, R.L., 2009. No TRAP, no invasion. Trends Parasitol 25, 77-84.

Mota, M.M., Hafalla, J.C., Rodriguez, A., 2002. Migration through host cells activates *Plasmodium* sporozoites for infection. Nat Med 8, 1318-1322.

Mouray, E., Moutiez, M., Girault, S., Sergheraert, C., Florent, I., Grellier, P., 2007. Biochemical properties and cellular localization of *Plasmodium falciparum* protein disulfide isomerase. Biochimie 89, 337-346.

Munro, S., Pelham, H.R., 1987. A C-terminal signal prevents secretion of luminal ER proteins. Cell 48, 899-907.

Murray, C.J., Rosenfeld, L.C., Lim, S.S., Andrews, K.G., Foreman, K.J., Haring, D., Fullman, N., Naghavi, M., Lozano, R., Lopez, A.D., 2012. Global malaria mortality between 1980 and 2010: a systematic analysis. Lancet 379, 413-431.

Naguleswaran, A., Alaeddine, F., Guionaud, C., Vonlaufen, N., Sonda, S., Jenoe, P., Mevissen, M., Hemphill, A., 2005. *Neospora caninum* protein disulfide isomerase is involved in tachyzoite-host cell interaction. Int J Parasitol 35, 1459-1472.
Newbold, C., Warn, P., Black, G., Berendt, A., Craig, A., Snow, B., Msobo, M., Peshu, N., Marsh, K., 1997. Receptor-specific adhesion and clinical disease in *Plasmodium falciparum*. Am J Trop Med Hyg 57, 389-398.
Ng, D.T., Brown, J.D., Walter, P., 1996. Signal sequences specify the targeting route to the endoplasmic reticulum membrane. J Cell Biol 134, 269-278.
Ngo, H.M., Yang, M., Paprotka, K., Pypaert, M., Hoppe, H., Joiner, K.A., 2003. AP-1 in *Toxoplasma gondii* mediates biogenesis of the rhoptry secretory organelle from a post-Golgi compartment. J Biol Chem 278, 5343-5352.
Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10, 1-6.
Nielsen, H., Krogh, A., 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB 6, 122-130.
Nilsson, T., Jackson, M., Peterson, P.A., 1989. Short cytoplasmic sequences serve as retention signals for transmembrane proteins in the endoplasmic reticulum. Cell 58, 707-718.
Noe, A.R., Fishkind, D.J., Adams, J.H., 2000. Spatial and temporal dynamics of the secretory pathway during differentiation of the *Plasmodium yoelii* schizont. Mol Biochem Parasitol 108, 169-185.
Nussenzweig, R.S., Vanderberg, J., Most, H., Orton, C., 1967. Protective immunity produced by the injection of x-irradiated sporozoites of plasmodium berghei. Nature 216, 160-162.
O'Donnell, R.A., Hackett, F., Howell, S.A., Treeck, M., Struck, N., Krnajski, Z., Withers-Martinez, C., Gilberger, T.W., Blackman, M.J., 2006. Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite. J Cell Biol 174, 1023-1033.
Ochola, L.B., Siddondo, B.R., Ocholla, H., Nkya, S., Kimani, E.N., Williams, T.N., Makale, J.O., Liljander, A., Urban, B.C., Bull, P.C., Szestak, T., Marsh, K., Craig, A.G., 2011. Specific receptor usage in *Plasmodium falciparum* cytoadherence is associated with disease outcome. PLoS One 6, e14741.
Pain, A., Bohme, U., Berry, A.E., Mungall, K., Finn, R.D., Jackson, A.P., Mourier, T., Mistry, J., Pasini, E.M., Aslett, M.A., Balasubrammaniam, S., Borgwardt, K., Brooks, K., Carret, C., Carver, T.J., Cherevach, I., Chillingworth, T., Clark, T.G., Galinski, M.R., Hall, N., Harper, D., Harris, D., Hauser, H., Ivens, A., Janssen, C.S., Keane, T., Larke, N., Lapp, S., Marti, M., Moule, S., Meyer, I.M., Ormond, D., Peters, N., Sanders, M., Sanders, S., Sargeant, T.J., Simmonds, M., Smith, F., Squares, R., Thurston, S., Tivey, A.R., Walker, D., White, B., Zuiderwijk, E., Churcher, C., Quail, M.A., Cowman, A.F., Turner, C.M., Rajandream, M.A., Kocken, C.H., Thomas, A.W., Newbold, C.I., Barrell, B.G., Berriman, M., 2008. The genome of the simian and human malaria parasite *Plasmodium knowlesi*. Nature 455, 799-803.
Pappas, G., Roussos, N., Falagas, M.E., 2009. Toxoplasmosis snapshots: global status of *Toxoplasma gondii* seroprevalence and implications for pregnancy and congenital toxoplasmosis. Int J Parasitol 39, 1385-1394.
Parasite Genomics Group, W.T.S.I. 2010a, posting date. *Plasmodium chabaudi* AS strain whole genome sequencing project. [Online.]
Parasite Genomics Group, W.T.S.I. 2010b, posting date. The genome of *Plasmodium berghei* ANKA [Online.]
Parasite Genomics Group, W.T.S.I. 2011, posting date. *P. falciparum* IT strain whole genome sequencing project. [Online.]

Pasloske, B.L., Baruch, D.I., van Schravendijk, M.R., Handunnetti, S.M., Aikawa, M., Fujioka, H., Taraschi, T.F., Gormley, J.A., Howard, R.J., 1993. Cloning and characterization of a *Plasmodium falciparum* gene encoding a novel high-molecular weight host membrane-associated protein, PfEMP3. Mol Biochem Parasitol 59, 59-72.

Payne, D., 1987. Spread of chloroquine resistance in *Plasmodium falciparum*. Parasitol Today 3, 241-246.

Pelham, H.R., Hardwick, K.G., Lewis, M.J., 1988. Sorting of soluble ER proteins in yeast. Embo J 7, 1757-1762.

Pellecchia, M., Szyperski, T., Wall, D., Georgopoulos, C., Wuthrich, K., 1996. NMR structure of the J-domain and the Gly/Phe-rich region of the *Escherichia coli* DnaJ chaperone. J Mol Biol 260, 236-250.

Perkins, S.L., Schall, J.J., 2002. A molecular phylogeny of malarial parasites recovered from cytochrome b gene sequences. J Parasitol 88, 972-978.

Persson, K.E., McCallum, F.J., Reiling, L., Lister, N.A., Stubbs, J., Cowman, A.F., Marsh, K., Beeson, J.G., 2008. Variation in use of erythrocyte invasion pathways by *Plasmodium falciparum* mediates evasion of human inhibitory antibodies. J Clin Invest 118, 342-351.

Pesce, E.R., Acharya, P., Tatu, U., Nicoll, W.S., Shonhai, A., Hoppe, H.C., Blatch, G.L., 2008. The *Plasmodium falciparum* heat shock protein 40, Pfj4, associates with heat shock protein 70 and shows similar heat induction and localisation patterns. The international journal of biochemistry & cell biology 40, 2914-2926.

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

Peterson, M.G., Marshall, V.M., Smythe, J.A., Crewther, P.E., Lew, A., Silva, A., Anders, R.F., Kemp, D.J., 1989. Integral membrane protein located in the apical complex of *Plasmodium falciparum*. Mol Cell Biol 9, 3151-3154.

Potgieter, F.T., Els, H.J., 1977. The fine structure of intra-erythrocytic stages of *Babesia bigemina*. Onderstepoort J Vet Res 44, 157-168.

Ranjan, R., Chugh, M., Kumar, S., Singh, S., Kanodia, S., Hossain, M.J., Korde, R., Grover, A., Dhawan, S., Chauhan, V.S., Reddy, V.S., Mohmmed, A., Malhotra, P., 2010. Proteome analysis reveals a large merozoite surface protein-1 associated complex on the *Plasmodium falciparum* merozoite surface. J Proteome Res 10, 680-691.

Ranson, H., Abdallah, H., Badolo, A., Guelbeogo, W.M., Kerah-Hinzoumbe, C., Yangalbe-Kalnone, E., Sagnon, N., Simard, F., Coetzee, M., 2009. Insecticide resistance in *Anopheles gambiae*: data from the first year of a multi-country study highlight the extent of the problem. Malar J 8, 299.

Rayner, J.C., Galinski, M.R., Ingravallo, P., Barnwell, J.W., 2000. Two *Plasmodium falciparum* genes express merozoite proteins that are related to *Plasmodium vivax* and *Plasmodium yoelii* adhesive proteins involved in host cell selection and invasion. Proc Natl Acad Sci U S A 97, 9648-9653.

Rayner, J.C., Vargas-Serrato, E., Huber, C.S., Galinski, M.R., Barnwell, J.W., 2001. A *Plasmodium falciparum* homologue of *Plasmodium vivax* reticulocyte binding protein (PvRBP1) defines a trypsin-resistant erythrocyte invasion pathway. J Exp Med 194, 1571-1581.

Reddy, G.R., 1995. Cloning and characterization of a *Plasmodium falciparum* cyclophilin gene that is stage-specifically expressed. Mol Biochem Parasitol 73, 111-121.

Richard, D., MacRaild, C.A., Riglar, D.T., Chan, J.A., Foley, M., Baum, J., Ralph, S.A., Norton, R.S., Cowman, A.F., 2010. Interaction between *Plasmodium falciparum* apical membrane antigen 1 and the rhoptry neck protein complex defines a key step in the erythrocyte invasion process of malaria parasites. J Biol Chem 285, 14815-14822.

Ridzuan, M.A., Moon, R.W., Knuepfer, E., Black, S., Holder, A.A., Green, J.L., 2012. Subcellular location, phosphorylation and assembly into the motor complex of GAP45 during *Plasmodium falciparum* schizont development. PLoS One 7, e33845.

Riglar, D.T., Richard, D., Wilson, D.W., Boyle, M.J., Dekiwadia, C., Turnbull, L., Angrisano, F., Marapana, D.S., Rogers, K.L., Whitchurch, C.B., Beeson, J.G., Cowman, A.F., Ralph, S.A., Baum, J., 2011. Super-resolution dissection of coordinated events during malaria parasite invasion of the human erythrocyte. Cell Host Microbe 9, 9-20.

Rivadeneira, E.M., Wasserman, M., Espinal, C.T., 1983. Separation and concentration of schizonts of *Plasmodium falciparum* by Percoll gradients. J Protozool 30, 367-370.

Roberts, D.J., Craig, A.G., Berendt, A.R., Pinches, R., Nash, G., Marsh, K., Newbold, C.I., 1992. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. Nature 357, 689-692.

Rosario, V., 1981. Cloning of naturally occurring mixed infections of malaria parasites. Science 212, 1037-1038.

Ross, R., 1897. On some Peculiar Pigmented Cells Found in Two Mosquitos Fed on Malarial Blood. Br Med J 2, 1786-1788.

Ross, R., 1898. PIGMENTED CELLS in MOSQUITOS. Br Med J 1, 550-551.

Rowe, A., Obeiro, J., Newbold, C.I., Marsh, K., 1995. *Plasmodium falciparum* rosetting is associated with malaria severity in Kenya. Infect Immun 63, 2323-2326.

Rug, M., Wickham, M.E., Foley, M., Cowman, A.F., Tilley, L., 2004. Correct promoter control is needed for trafficking of the ring-infected erythrocyte surface antigen to the host cytosol in transfected malaria parasites. Infect Immun 72, 6095-6105.

Sabchareon, A., Burnouf, T., Ouattara, D., Attanath, P., Bouharoun-Tayoun, H., Chantavanich, P., Foucault, C., Chongsuphajaisiddhi, T., Druilhe, P., 1991. Parasitologic and clinical human response to immunoglobulin administration in falciparum malaria. Am J Trop Med Hyg 45, 297-308.

Saenz, F.E., Balu, B., Smith, J., Mendonca, S.R., Adams, J.H., 2008. The transmembrane isoform of *Plasmodium falciparum* MAEBL is essential for the invasion of Anopheles salivary glands. PLoS One 3, e2287.

Sam-Yellowe, T.Y., Perkins, M.E., 1991. Interaction of the 140/130/110 kDa rhoptry protein complex of *Plasmodium falciparum* with the erythrocyte membrane and liposomes. Exp Parasitol 73, 161-171.

Sam-Yellowe, T.Y., 1996. Rhoptry organelles of the apicomplexa: Their role in host cell invasion and intracellular survival. Parasitol Today 12, 308-316.

Sambrook J., G., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Sato, K., Sato, M., Nakano, A., 1997. Rer1p as common machinery for the endoplasmic reticulum localization of membrane proteins. Proc Natl Acad Sci U S A 94, 9693-9698.

Sato, M., Sato, K., Nakano, A., 1996. Endoplasmic reticulum localization of Sec12p is achieved by two mechanisms: Rer1p-dependent retrieval that requires the transmembrane domain and Rer1p-independent retention that involves the cytoplasmic domain. J Cell Biol 134, 279-293.

Sato, S., Rangachari, K., Wilson, R.J., 2003. Targeting GFP to the malarial mitochondrion. Mol Biochem Parasitol 130, 155-158.

Schuler, H., Mueller, A.K., Matuschewski, K., 2005. A *Plasmodium* actin-depolymerizing factor that binds exclusively to actin monomers. Mol Biol Cell 16, 4013-4023.

Schultz, J., Milpetz, F., Bork, P., Ponting, C.P., 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proceedings of the National Academy of Sciences of the United States of America 95, 5857-5864.

Shaw, M.K., 2003. Cell invasion by Theileria sporozoites. Trends Parasitol 19, 2-6.

Silamut, K., Phu, N.H., Whitty, C., Turner, G.D., Louwrier, K., Mai, N.T., Simpson, J.A., Hien, T.T., White, N.J., 1999. A quantitative analysis of the microvascular sequestration of malaria parasites in the human brain. Am J Pathol 155, 395-410.

Silva, M.D., Cooke, B.M., Guillotte, M., Buckingham, D.W., Sauzet, J.P., Le Scanf, C., Contamin, H., David, P., Mercereau-Puijalon, O., Bonnefoy, S., 2005. A role for the *Plasmodium falciparum* RESA protein in resistance against heat shock demonstrated using gene disruption. Mol Microbiol 56, 990-1003.

Silvestrini, F., Alano, P., Williams, J.L., 2000. Commitment to the production of male and female gametocytes in the human malaria parasite *Plasmodium falciparum*. Parasitology 121 Pt 5, 465-471.

Silvestrini, F., Lasonder, E., Olivieri, A., Camarda, G., van Schaijk, B., Sanchez, M., Younis Younis, S., Sauerwein, R., Alano, P., 2010. Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. Mol Cell Proteomics 9, 1437-1448.

Sim, B.K., Toyoshima, T., Haynes, J.D., Aikawa, M., 1992. Localization of the 175-kilodalton erythrocyte binding antigen in micronemes of *Plasmodium falciparum* merozoites. Mol Biochem Parasitol 51, 157-159.

Sim, B.K., Chitnis, C.E., Wasniowska, K., Hadley, T.J., Miller, L.H., 1994. Receptor and ligand domains for invasion of erythrocytes by Plasmodium falciparum. Science 264, 1941-1944.

Singh, B., Kim Sung, L., Matusop, A., Radhakrishnan, A., Shamsul, S.S., Cox-Singh, J., Thomas, A., Conway, D.J., 2004. A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. Lancet 363, 1017-1024.

Singh, S., Plassmeyer, M., Gaur, D., Miller, L.H., 2007. Mononeme: a new secretory organelle in *Plasmodium falciparum* merozoites identified by localization of rhomboid-1 protease. Proc Natl Acad Sci U S A 104, 20043-20048.

Singh, S., Alam, M.M., Pal-Bhowmick, I., Brzostowski, J.A., Chitnis, C.E., 2010. Distinct External Signals Trigger Sequential Release of Apical Organelles during Erythrocyte Invasion by Malaria Parasites. PLoS Pathog 6, e1000746.

Sinka, M.E., Rubio-Palis, Y., Manguin, S., Patil, A.P., Temperley, W.H., Gething, P.W., Van Boeckel, T., Kabaria, C.W., Harbach, R.E., Hay, S.I., 2010. The dominant *Anopheles* vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic precis. Parasit Vectors 3, 72.

Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S., Pinches, R., Newbold, C.I., Miller, L.H., 1995. Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell 82, 101-110.

Smith, T.G., Lourenco, P., Carter, R., Walliker, D., Ranford-Cartwright, L.C., 2000. Commitment to sexual differentiation in the human malaria parasite, *Plasmodium falciparum*. Parasitology 121 (Pt 2), 127-133.

Sonnhammer, E.L., von Heijne, G., Krogh, A., 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB 6, 175-182.

Sterkers, Y., Scheidig, C., da Rocha, M., Lepolard, C., Gysin, J., Scherf, A., 2007. Members of the low-molecular-mass rhoptry protein complex of *Plasmodium falciparum* bind to the surface of normal erythrocytes. J Infect Dis 196, 617-621.

Stirling, C.J., Rothblatt, J., Hosobuchi, M., Deshaies, R., Schekman, R., 1992. Protein translocation mutants defective in the insertion of integral membrane proteins into the endoplasmic reticulum. Mol Biol Cell 3, 129-142.

Straub, K.W., Cheng, S.J., Sohn, C.S., Bradley, P.J., 2009. Novel components of the Apicomplexan moving junction reveal conserved and coccidia-restricted elements. Cell Microbiol 11, 590-603.
Striepen, B., 2011. The apicoplast: a red alga in human parasites. Essays Biochem 51, 111-125.
Stubbs, J., Simpson, K.M., Triglia, T., Plouffe, D., Tonkin, C.J., Duraisingh, M.T., Maier, A.G., Winzeler, E.A., Cowman, A.F., 2005. Molecular mechanism for switching of *P. falciparum* invasion pathways into human erythrocytes. Science 309, 1384-1387.
Sturm, A., Amino, R., van de Sand, C., Regen, T., Retzlaff, S., Rennenberg, A., Krueger, A., Pollok, J.M., Menard, R., Heussler, V.T., 2006. Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. Science 313, 1287-1290.
Szabo, A., Korszun, R., Hartl, F.U., Flanagan, J., 1996. A zinc finger-like domain of the molecular chaperone DnaJ is involved in binding to denatured protein substrates. Embo J 15, 408-417.
Tarun, A.S., Peng, X., Dumpit, R.F., Ogata, Y., Silva-Rivera, H., Camargo, N., Daly, T.M., Bergman, L.W., Kappe, S.H., 2008. A combined transcriptome and proteome survey of malaria parasite liver stages. Proc Natl Acad Sci U S A 105, 305-310.
Taylor, H.M., Triglia, T., Thompson, J., Sajid, M., Fowler, R., Wickham, M.E., Cowman, A.F., Holder, A.A., 2001. *Plasmodium falciparum* homologue of the genes for Plasmodium vivax and Plasmodium yoelii adhesive proteins, which is transcribed but not translated. Infect Immun 69, 3635-3645.
Taylor, H.M., Grainger, M., Holder, A.A., 2002. Variation in the expression of a *Plasmodium falciparum* protein family implicated in erythrocyte invasion. Infect Immun 70, 5779-5789.
Taylor, H.M., McRobert, L., Grainger, M., Sicard, A., Dluzewski, A.R., Hopp, C.S., Holder, A.A., Baker, D.A., 2009. The malaria parasite cGMP-dependent protein kinase plays a central role in blood stage schizogony. Eukaryot Cell.
Templeton, T.J., Iyer, L.M., Anantharaman, V., Enomoto, S., Abrahante, J.E., Subramanian, G.M., Hoffman, S.L., Abrahamsen, M.S., Aravind, L., 2004. Comparative analysis of apicomplexa and genomic diversity in eukaryotes. Genome Res 14, 1686-1695.
Terenius, O., de Oliveira, C.D., Pinheiro, W.D., Tadei, W.P., James, A.A., Marinotti, O., 2008. 16S rRNA gene sequences from bacteria associated with adult *Anopheles darlingi* (Diptera: Culicidae) mosquitoes. J Med Entomol 45, 172-175.
Tham, W.H., Wilson, D.W., Lopaticki, S., Schmidt, C.Q., Tetteh-Quarcoo, P.B., Barlow, P.N., Richard, D., Corbin, J.E., Beeson, J.G., Cowman, A.F., 2010. Complement receptor 1 is the host erythrocyte receptor for *Plasmodium falciparum* PfRh4 invasion ligand. Proc Natl Acad Sci U S A 107, 17327-17332.
Timmann, C., Meyer, C.G., 2010. Malaria, mummies, mutations: Tutankhamun's archaeological autopsy. Trop Med Int Health 15, 1278-1280.
Tomley, F.M., Bumstead, J.M., Billington, K.J., Dunn, P.P., 1996. Molecular cloning and characterization of a novel acidic microneme protein (Etmic-2) from the apicomplexan protozoan parasite, *Eimeria tenella*. Mol Biochem Parasitol 79, 195-206.
Torii, M., Adams, J.H., Miller, L.H., Aikawa, M., 1989. Release of merozoite dense granules during erythrocyte invasion by *Plasmodium knowlesi*. Infect Immun 57, 3230-3233.
Trager, W., Jensen, J.B., 1976. Human malaria parasites in continuous culture. Science 193, 673-675.
Treeck, M., Struck, N.S., Haase, S., Langer, C., Herrmann, S., Healer, J., Cowman, A.F., Gilberger, T.W., 2006. A conserved region in the EBL proteins is implicated in microneme targeting of the malaria parasite *Plasmodium falciparum*. J Biol Chem 281, 31995-32003.
Treeck, M., Zacherl, S., Herrmann, S., Cabrera, A., Kono, M., Struck, N.S., Engelberg, K., Haase, S., Frischknecht, F., Miura, K., Spielmann, T., Gilberger, T.W., 2009. Functional analysis of the

leading malaria vaccine candidate AMA-1 reveals an essential role for the cytoplasmic domain in the invasion process. PLoS Pathog 5, e1000322.
Treeck, M., Sanders, J.L., Elias, J.E., Boothroyd, J.C., 2011. The phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites' boundaries. Cell Host Microbe 10, 410-419.
Trenholme, K.R., Gardiner, D.L., Holt, D.C., Thomas, E.A., Cowman, A.F., Kemp, D.J., 2000. clag9: A cytoadherence gene in *Plasmodium falciparum* essential for binding of parasitized erythrocytes to CD36. Proc Natl Acad Sci U S A 97, 4029-4033.
Triglia, T., Healer, J., Caruana, S.R., Hodder, A.N., Anders, R.F., Crabb, B.S., Cowman, A.F., 2000. Apical membrane antigen 1 plays a central role in erythrocyte invasion by *Plasmodium* species. Mol Microbiol 38, 706-718.
Triglia, T., Thompson, J.K., Cowman, A.F., 2001. An EBA175 homologue which is transcribed but not translated in erythrocytic stages of *Plasmodium falciparum*. Mol Biochem Parasitol 116, 55-63.
Triglia, T., Duraisingh, M.T., Good, R.T., Cowman, A.F., 2005. Reticulocyte-binding protein homologue 1 is required for sialic acid-dependent invasion into human erythrocytes by *Plasmodium falciparum*. Mol Microbiol 55, 162-174.
Tsai, J., Douglas, M.G., 1996. A conserved HPD sequence of the J-domain is necessary for YDJ1 stimulation of Hsp70 ATPase activity at a site distinct from substrate binding. J Biol Chem 271, 9347-9354.
Tuteja, R., 2007. Unraveling the components of protein translocation pathway in human malaria parasite *Plasmodium falciparum*. Arch Biochem Biophys 467, 249-260.
Tyedmers, J., Lerner, M., Wiedmann, M., Volkmer, J., Zimmermann, R., 2003. Polypeptide-binding proteins mediate completion of co-translational protein translocation into the mammalian endoplasmic reticulum. EMBO reports 4, 505-510.
Tyler, J.S., Boothroyd, J.C., 2011. The C-terminus of *Toxoplasma* RON2 provides the crucial link between AMA1 and the host-associated invasion complex. PLoS Pathog 7, e1001282.
Uchime, O., Herrera, R., Reiter, K., Kotova, S., Shimp, R.L., Jr., Miura, K., Jones, D., Lebowitz, J., Ambroggio, X., Hurt, D.E., Jin, A.J., Long, C., Miller, L.H., Narum, D.L., 2012. Analysis of the conformation and function of the *Plasmodium falciparum* merozoite proteins MTRAP and PTRAMP. Eukaryot Cell.
Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., Wernerus, H., Bjorling, L., Ponten, F., 2010. Towards a knowledge-based Human Protein Atlas. Nat Biotechnol 28, 1248-1250.
Urquiza, M., Suarez, J.E., Cardenas, C., Lopez, R., Puentes, A., Chavez, F., Calvo, J.C., Patarroyo, M.E., 2000. *Plasmodium falciparum* AMA-1 erythrocyte binding peptides implicate AMA-1 as erythrocyte binding protein. Vaccine 19, 508-513.
van Dooren, G.G., Marti, M., Tonkin, C.J., Stimmler, L.M., Cowman, A.F., McFadden, G.I., 2005. Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of *Plasmodium falciparum*. Mol Microbiol 57, 405-419.
Van Wye, J., Ghori, N., Webster, P., Mitschler, R.R., Elmendorf, H.G., Haldar, K., 1996. Identification and localization of rab6, separation of rab6 from ERD2 and implications for an 'unstacked' Golgi, in *Plasmodium falciparum*. Mol Biochem Parasitol 83, 107-120.
Vlachou, D., Zimmermann, T., Cantera, R., Janse, C.J., Waters, A.P., Kafatos, F.C., 2004. Real-time, in vivo analysis of malaria ookinete locomotion and mosquito midgut invasion. Cell Microbiol 6, 671-685.
von Heijne, G., 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J Mol Biol 225, 487-494.

Wada, J., Kanwar, Y.S., 1998. Characterization of mammalian translocase of inner mitochondrial membrane (Tim44) isolated from diabetic newborn mouse kidney. Proc Natl Acad Sci U S A 95, 144-149.
Waller, R.F., Keeling, P.J., Donald, R.G., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A.F., Besra, G.S., Roos, D.S., McFadden, G.I., 1998. Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. Proc Natl Acad Sci U S A 95, 12352-12357.
Walsh, P., Bursac, D., Law, Y.C., Cyr, D., Lithgow, T., 2004. The J-protein family: modulating protein assembly, disassembly and translocation. EMBO Rep 5, 567-571.
Walter, P., Blobel, G., 1981a. Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein. J Cell Biol 91, 551-556.
Walter, P., Blobel, G., 1981b. Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. J Cell Biol 91, 557-561.
Wang, T., Fujioka, H., Drazba, J.A., Sam-Yellowe, T.Y., 2006. Rhop-3 protein conservation among Plasmodium species and induced protection against lethal P. yoelii and P. berghei challenge. Parasitol Res 99, 238-252.
Weatherall, D.J., Williams, T.N., Allen, S.J., O'Donnell, A., 2010. The population genetics and dynamics of the thalassemias. Hematology/oncology clinics of North America 24, 1021-1031.
Weiyuan, C., 2009. Ancient Chinese anti-fever cure becomes panacea for malaria. Bull World Health Organ 87, 743-744.
WHO, 1956. WHO Expert Committee on Malaria, sixth Report. In.
WHO, 2000. Severe falciparum malaria. Transactions of the Royal Society of Tropical Medicine and Hygiene 94.
WHO, 2010. World Malaria Report 2010. In, World Health Organization.
WHO, 2011. World Malaria Report 2011. In: WHO (Ed).
Wickham, M.E., Thompson, J.K., Cowman, A.F., 2003. Characterisation of the merozoite surface protein-2 promoter using stable and transient transfection in *Plasmodium falciparum*. Mol Biochem Parasitol 129, 147-156.
Wider, D., Peli-Gulli, M.P., Briand, P.A., Tatu, U., Picard, D., 2009. The complementation of yeast with human or *Plasmodium falciparum* Hsp90 confers differential inhibitor sensitivities. Mol Biochem Parasitol 164, 147-152.
Wirth, A., Jung, M., Bies, C., Frien, M., Tyedmers, J., Zimmermann, R., Wagner, R., 2003. The Sec61p complex is a dynamic precursor activated channel. Molecular cell 12, 261-268.
Wittke, S., Dunnwald, M., Johnsson, N., 2000. Sec62p, a component of the endoplasmic reticulum protein translocation machinery, contains multiple binding sites for the Sec-complex. Mol Biol Cell 11, 3859-3871.
Wu, Y., Kirkman, L.A., Wellems, T.E., 1996. Transformation of *Plasmodium falciparum* malaria parasites by homologous integration of plasmids that confer resistance to pyrimethamine. Proc Natl Acad Sci U S A 93, 1130-1134.
Yeoh, S., O'Donnell, R.A., Koussis, K., Dluzewski, A.R., Ansell, K.H., Osborne, S.A., Hackett, F., Withers-Martinez, C., Mitchell, G.H., Bannister, L.H., Bryans, J.S., Kettleborough, C.A., Blackman, M.J., 2007. Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes. Cell 131, 1072-1083.
Yeoman, J.A., Hanssen, E., Maier, A.G., Klonis, N., Maco, B., Baum, J., Turnbull, L., Whitchurch, C.B., Dixon, M.W., Tilley, L., 2011. Tracking Glideosome-associated protein 50 reveals the development and organization of the inner membrane complex of *Plasmodium falciparum*. Eukaryot Cell 10, 556-564.

Yokoyama, N., Okamura, M., Igarashi, I., 2006. Erythrocyte invasion by *Babesia* parasites: current advances in the elucidation of the molecular interactions between the protozoan ligands and host receptors in the invasion stage. Vet Parasitol 138, 22-32.
Yuda, M., Yano, K., Tsuboi, T., Torii, M., Chinzei, Y., 2001. von Willebrand Factor A domain-related protein, a novel microneme protein of the malaria ookinete highly conserved throughout *Plasmodium* parasites. Mol Biochem Parasitol 116, 65-72.
Zimmermann, R., Blatch, G.L., 2009. A novel twist to protein secretion in eukaryotes. Trends Parasitol 25, 147-150.

Appendix A: *PF11_0443rabbit polyclonal antibody immunoprecipitation LcMS/MS results*

Band	Gene ID	Protein	kDa (gel)	kDa (MS)	Score	% coverage	Comments
1	PFI1475w	MSP-1	25-37	197217	1049	16.27	N and C terminus
	PFB0340c	SERA 5		113294	786	26.18	
	PF11_0443			39678	711	48.6	
	PF14_0102	RAP1		90467	588	35.68	
	PF08_0054	hsp70		74754	546	15.81	C terminus only
	PFF0435w	ornithine aminotransferase		46938	519	26.81	
	PF11_0351	hsp70		71945	515	16.44	C terminus only
	PF11_0099	Pfj2		62747	499	24.26	C terminal half
	PF11_0344	AMA-1		61606	478	21.22	N terminal half
	PF10_0348	MSPDBL1		80997	426	15.93	
	PFI0265c	RhopH3		105587	422	14.83	
	PFD0240c	6-cysteine protein		43802	399	30.42	
		hyp17		144218	344	11.88	
	PF14_0077	plasmepsin II		51847	332	24.94	
	PF10_0115	QF122 antigen		132002	330	10.1	
	PFD1130w	conserved protein		43432	306	26.24	
	MAL13P1.336	conserved protein		78582	268	11.09	
	PFL2275c	FKBP-type isomerase		35147	258	25.33	
	PFF1025c	SNO glutamine amidotransferase		33391	238	23.26	
	PF14_0567	conserved protein		40556	226	21.18	
	PFL0300c	protein phosphatase		36120	217	15.46	
	PF14_0660	protein phosphatase		42697	217	23.46	
2	PF14_0425	fructose-1,6-bisphosphate aldolase	37kDa	40348	1123	62.33	
	PFI1475w	MSP-1		197217	715	13.72	N and C terminus only
	PF11_0099	Pfj2		62747	714	31.48	C terminal half
	PFD0240c	6-cysteine protein		43802	597	42.33	
	PFF0435w	ornithine aminotransferase		46938	455	25.12	
	PF10_0348	MSPDBL1		80997	432	19.51	
	PF11_0443			39678	423	42.37	
	PFE0080c	RhopH2		47017	348	22.11	
	PFI0265c	RhopH3		105587	342	13.15	N terminus only
	PF11_0175	hsp101		103038	327	12.47	N terminus only
	PF11_0086	MIF4G domain containing protein		383293	320		
	PFE0075c	RAP3		47232	318	16.25	
	PFF1415c	DNAJ domain protein		45040	287	25.53	
	PF07_0129	acyl-coA synthase		93624	283	9.49	
	PF14_0281	Plasmepsin IX		74820	259	15.15	
	PF14_0075	Plasmepsin IV		37159	253	20.71	
	PF13_0272	Thioredoxin-related protein		24314	193	31.25	
	PFI0880c	GAP50		44804	179	15.15	
3	PFI1475w	MSP-1	50kDa	196746	660	9.71	N- terminus only
		hyp17		144218	507	10.37	
	PF11_0344	AMA-1		69741	504	22.51	
	MAL8P1.17	disulphide isomerase		55808	379	26.71	
	PF14_0694	disulphide isomerase		66989	194	10.67	
	PF11_0443			39678	168	15.26	
4	PFI1475w	MSP-1	60kDa	197217	545	10.06	N-terminus only
	PF14_0102	RAP-1		90467	536	22.51	
	PFA0125c	EBA-181		183259	361	7.34	aa's 180-720
	PF11_0443			39678	271	21.5	
	PF11_0465	dynamin-like protein		96712	192	8	
	PF10_0153	hsp-60		62911	167	16.03	
5	PFI1445w	RhopH2	150kDa	163587	1573	27	
6	PFI0875w	BiP	15-20kDa	30696	227	9.2	C terminus only
7		No significant hits					

Appendix B: *PF11_0443-FLAG immunoprecipitation attempt 1 LcMS/MS results.* Proteins are listed according to peptide matches in descending MOWSE score. PF11_0443 is highlighted in blus, the next hightest scoring hit in pink, rhoptry proteins in yellow and DnaJ domain containing proteins in purple.

Results of FLAG IP attempt 1

Accession	Description	MW (kDa)	Coverage (%)	# Peptides	Band	Score
PF11_0443	DnaJ protein, putative	39.3	44.24	27	9	2,116.20
PF10_0153	heat shock protein 60	62.5	48.97	29	5	1,188.61
MAL7P1.38	regulator of chromosome condensation	79.0	22.13	12	3	321.60
PFL1875w	conserved Plasmodium protein	61.3	15.09	6	5	307.46
PF14_0407	guanine nucleotide exchange factor	404.3	2.98	8	1	279.49
PF13_0357	conserved Plasmodium protein	100.0	10.00	7	3	272.61
PFB0095c	erythrocyte membrane protein 3	273.5	39.57	10	1	262.21
MAL13P1.352	conserved Plasmodium protein	131.3	5.15	5	3	251.05
PF11_0168	RON4	135.5	7.33	8	5	206.15

Proteins with lower scores absent fromWT control IP

Accession	Description	MW (kDa)	Coverage (%)	# Peptides	Band	Score
PF11_0084	conserved Plasmodium protein	84.2	7.57	4	3	184.93
MAL7P1.108	phosphoinositide-binding protein	258.3	3.55	7	1	182.36
MAL7P1.204	conserved Plasmodium protein	196.4	3.17	5	3	172.82
PFF1030w	mRNA binding Pumilio-homology domain protein	94.2	6.05	4	3	166.45
PF13_0061	ATP synthase gamma chain, mitochondrial precursor	35.7	14.15	3	11	159.41
PFI0265c	RhopH3	104.8	6.02	5	3	150.41
PFL1385c	MSP-9	86.6	4.85	3	3	149.80
PF14_0386	adaptor complexes medium subunit family	87.1	7.09	4	3	126.68
PF08_0067	ubiquitin	44.9	13.40	4	11	126.44
PFL1545c	chaperonin, cpn60	81.4	6.96	4	3	113.21
PFL1070c	endoplasmin homolog precursor (Hsp90)	95.0	4.14	3	3	111.43
PF14_0350	N-acetyltransferase, putative	64.1	3.86	2	5	107.91
PF11_0380	DnaJ protein, putative	44.5	6.50	2	11	95.59

Appendix C: *PF11_0443-FLAG immunoprecipitation attempt 2 LcMS/MS results.* Proteins are listed according to peptide matches in descending MOWSE score. PF11_0443 is highlighted in blue, the next hightest scoring hit in pink, rhoptry proteins in yellow and DnaJ domain containing proteins in purple.

Results of FLAG IP attempt 2

Accession	Description	MW (kDa)	Coverage (%)	# Peptides	Band	Score
PF3D7_1143200	DnaJ protein, putative	39.3	30.84	18	11	440.92
PF3D7_0804800	peptidyl-prolyl cis-trans isomerase (CYP24)	24.9	33.64	9	12	421.56
PF3D7_1341200	60S ribosomal protein L18, putative	21.7	40.76	9	12	262.2
PF3D7_0503400	actin-depolymerizing factor 1 (ADF1)	13.7	22.95	2	12	246.54
PF3D7_1004000	60S ribosomal protein L13, putative	23.7	39.60	10	12	220.66
PF3D7_0322900	40S ribosomal protein S3A, putative	30.0	39.31	10	12	215.41
PF3D7_1452000	rhoptry neck protein 2 (RON2)	249.3	4.61	10	9	210.84
PF3D7_0823800	DnaJ protein, putative	76.6	13.13	9	11	202.49
PF3D7_1361900	proliferating cell nuclear antigen (PCNA)	30.6	12.04	3	12	177.29
PF3D7_0706000	conserved Plasmodium protein	145.4	4.31	6	9	176.65
PF3D7_1331800	60S ribosomal protein L23, putative	15.0	11.51	2	12	164.28
PF3D7_1455200	conserved Plasmodium protein	24.0	29.76	5	12	164.2
PF3D7_0513300	purine nucleoside phosphorylase (PNP)	26.8	16.73	3	12	137.53
PF3D7_0711400	sin3 associated polypeptide p18-like protein	88.4	8.10	6	12	135.19
PF3D7_1438900	thioredoxin peroxidase 1 (Trx-Px1)	21.8	40.00	6	12	132.69
PF3D7_0415900	60S ribosomal protein L15, putative	24.1	27.32	5	12	131.91
PF3D7_0627700	transportin	132.8	4.80	5	9	119.28
PF3D7_1120100	phosphoglycerate mutase, putative (PGM1)	28.8	8.00	2	12	117.53
PF3D7_1238100	calcyclin binding protein, putative	26.6	18.86	4	12	113.16
PF3D7_1211400	heat shock protein DNAJ homologue Pfj4	28.0	16.80	5	12	112.56
PF3D7_1451800	sortilin, putative	102.2	3.35	3	10	107.53

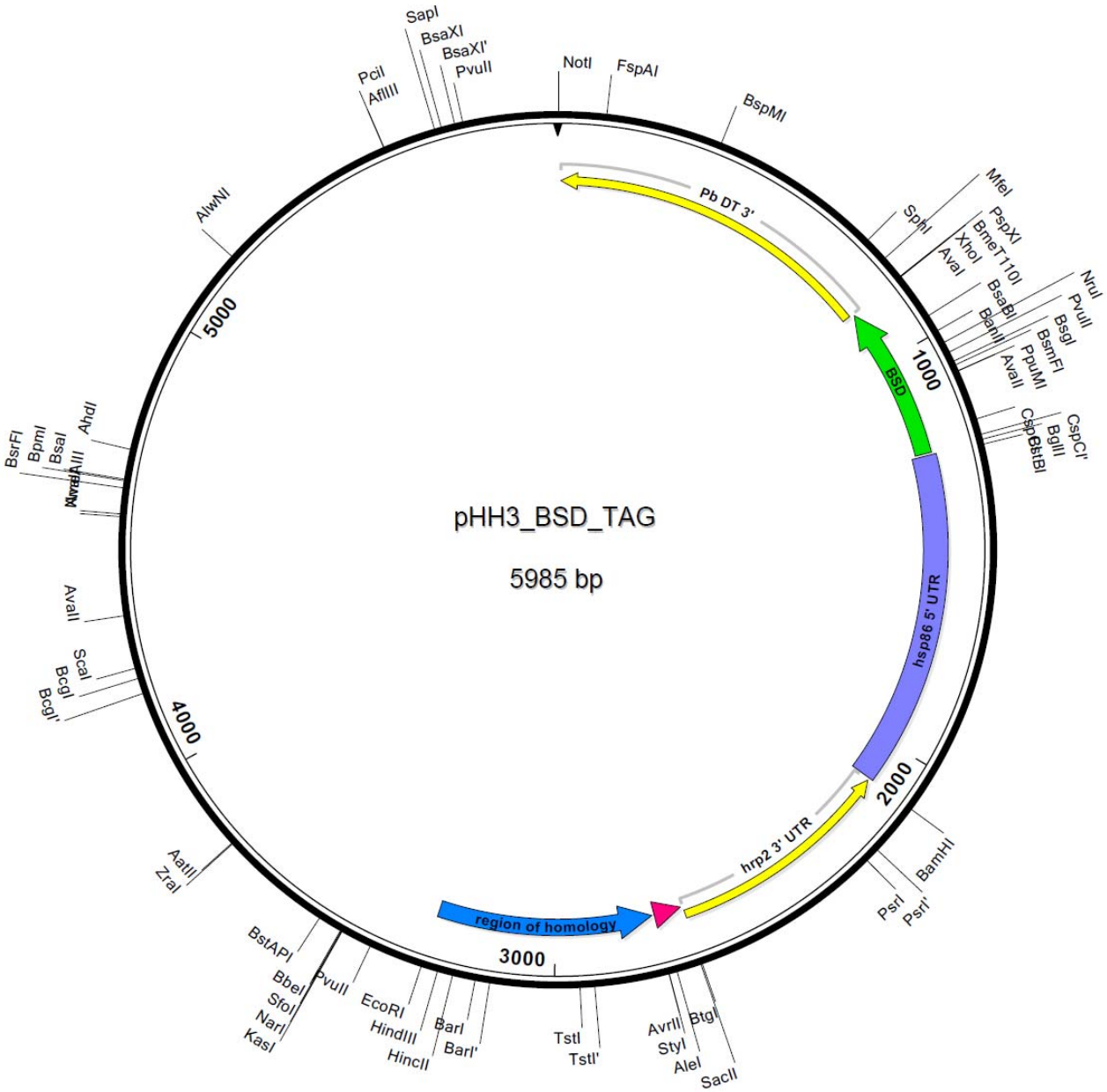
Proteins of interest with low scores

Accession	Description	MW (kDa)	Coverage (%)	# Peptides	Band	Score
PF3D7_0817700	rhoptry neck protein 5 (RON5)	133.6	5.71	6	9,10,11,12	73,79,60,57

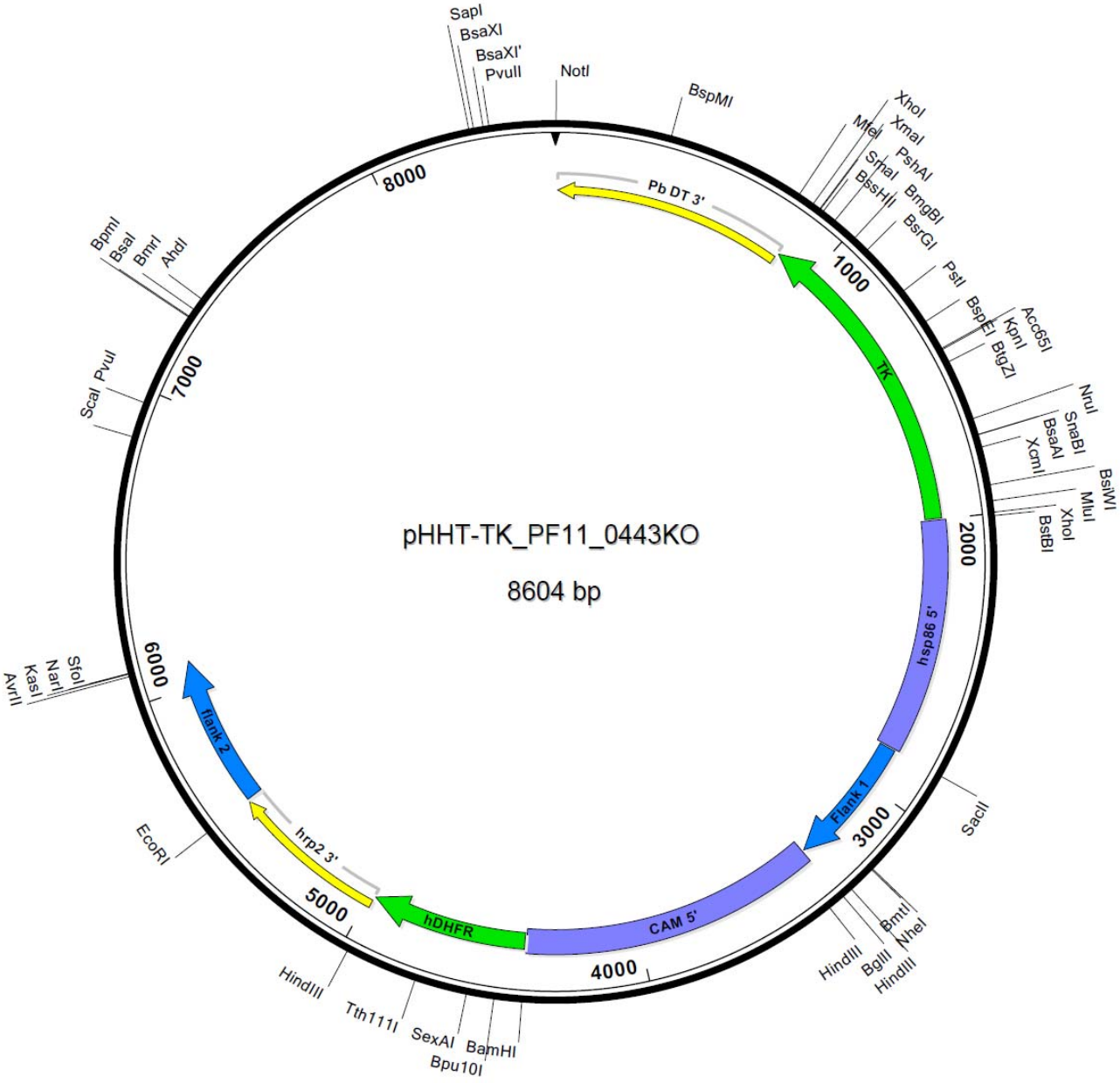
Protein with high score but WT presence

Accession	Description	MW (kDa)	Coverage (%)	# Peptides	Bands	Score	WT Band	WT Score
PF3D7_0720900	conserved Plasmodium protein	31.1	30.94	12	11,12	411,245	5	270.9
PF3D7_1252100	rhoptry neck protein 3 (RON3)	263.0	11.11	27	9,10,11	449,68,88	3	381.44
PF3D7_0935800	CLAG9	160.3	7.69	11	9,10,11	157,140,85	3	124

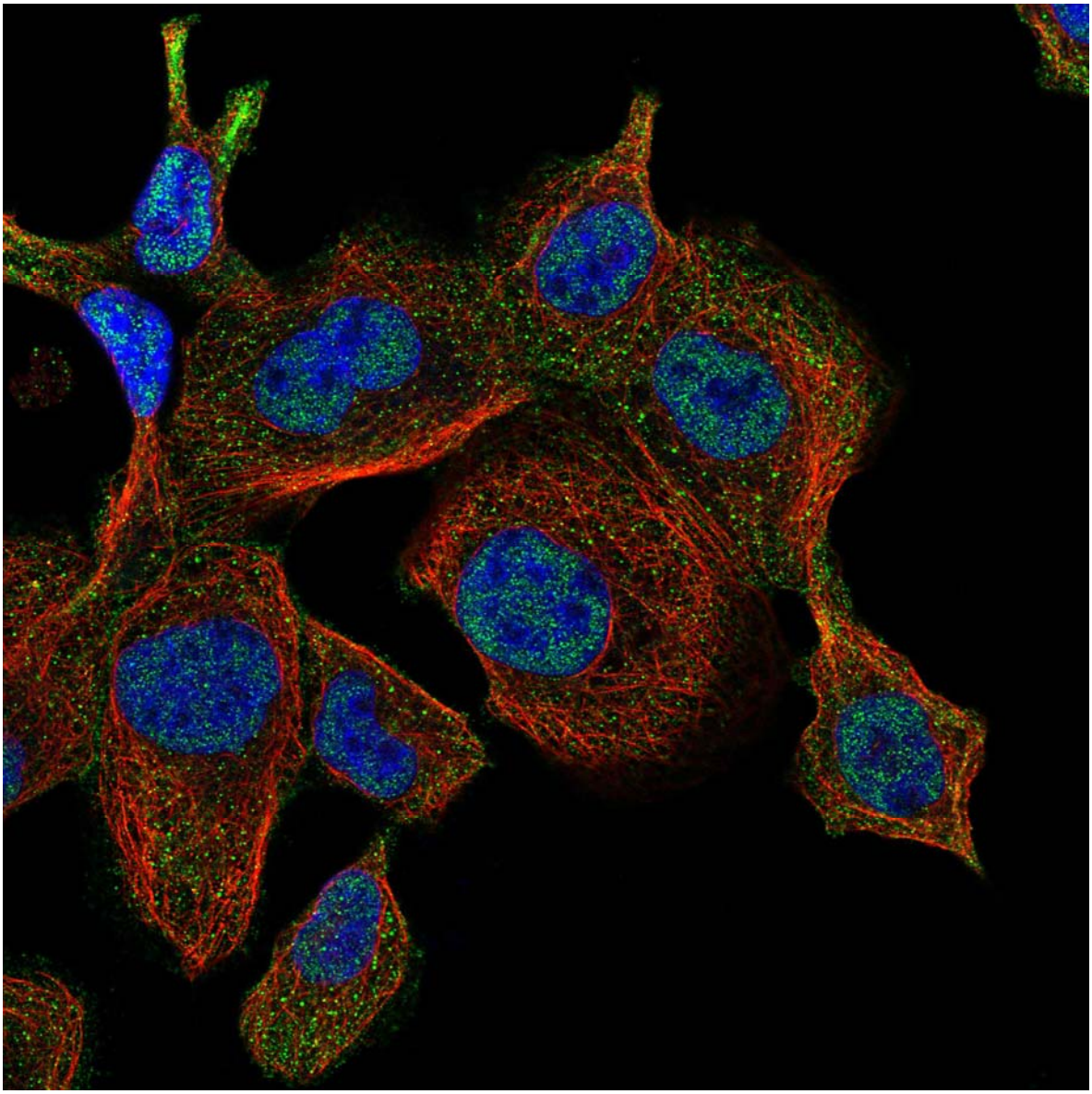
Appendix D: *Vector map for the epitope-tagging construct pHH3.* The diagram shows a vector map of an example pHH3 construct used to epitope tag genes presented in this thesis. All unique restriction sites and important features are indicated, the pink arrow hightlights the region containing either the Ty1 or 3xFLAG epitope tag. For KO by truncation, the tag and 3′ UTR were removed and the region of homology for MAL13P1.94 and PF02_0040 was cloned in between *Eco*RI and *Bam*HI restriction sites.



Appendix E: *Vector map of pHTK_PF11_0443_KO*. The diagram shows the vector map of the construct used in an attempt to KO PF11_0443 by double homologous recombination. All unique restriction sites and important features are indicated.



Appendix F: *Localisation of DNAJC25 in human skin cell line A-431*. Image taken from The Human Protein Atlas website: www.proteinatlas.org (Uhlen et al., 2010). DNAJC25 from the skin cell line A-431 is stained with antibody HPA019122 (green), microtubules are probed with anti-tubulin in red and the nucleus is stained with DAPI and shown in blue. The confocal image presented represents a single optical section of the cells.



Appendix G: *Protien selected by the project criteria before choosing five for further study.*

Gene	Product Description	TM	Molecular Weight	Orthologues	Contender?	Comments
PFA0125c	erythrocyte binding antigen-181	2	181151	6	No	Already characterised
PFA0265c	conserved Plasmodium protein, unknown function	2	16856	4	No	No vivax, small
PFB0060w	rifin	1	39920	0	No	Already characterised
PFB0194w	conserved Plasmodium protein, unknown function	1	33415	4	Yes	CANDIDATE
PFB0305c-a	merozoite surface protein 5	2	30974	2	No	Already characterised
PFC0581w	co-chaperone p23, putative	1	32611	5	No	wrong function
PFD0110w	reticulocyte-binding protein homologue 1	1	357784	0	No	Already characterised
PFD0295c	apical sushi protein, ASP	1	85461	5	No	Already characterised
PFD0930w	CGI-141 protein homolog, putative	3	15891	5	No	too many TMs for a small protein
PFD1100c	conserved Plasmodium protein, unknown function	1	58351	5	No	In gamaetocytes
PFD1150c	reticulocyte binding protein homolog 4, Rh4	2	205850	0	No	Already characterised
PFD1155w	erythrocyte binding antigen-165	1	160266	6	No	Already characterised
PFE1130w	conserved protein, unknown function	8	55565	5	No	too many TMs
PFE1445c	conserved Plasmodium protein, unknown function	2	82906	5	No	Poor expression
PFE1515w	conserved Plasmodium membrane protein, unknown function	11	173268	5	No	Large with too many TMs (hard to work with)
PFF0615c	6-cysteine protein, putative	1	39435	5	?	
PFF0995c	Merozoite surface protein 10, MSP10	2	61381	6	No	Already characterised
MAL7P1.119	conserved Plasmodium protein, unknown function	1	87880	6	No	RALP1 (Gilberger lab), low expression
MAL7P1.176	erythrocyte binding antigen 175	1	174589	6	No	Already characterised
MAL8P1.135	conserved Plasmodium membrane protein, unknown function	9	122107	5	No	too many TMs for a small protein
PF08_0104	rifin	2	37838	0	No	Already characterised
PF08_0057	conserved Plasmodium protein, unknown function	1	6159	5	No	too small
PF08_0008	conserved Plasmodium protein, unknown function	2	85251	5	No	GAMA (Holder lab)
PFI1475w	merozoite surface protein 1 precursor	1	195728	6	No	Already characterised

Appendix G continued

Gene	Product Description	TM	Molecular Weight	Orthologues	Contender?	Comments
PFI1560c	conserved Plasmodium membrane protein, unknown function	4	147394	6	No	No longer appears in search results
PF10_0082	conserved Plasmodium membrane protein, unknown function	10	154751	5	No	too many TMs
PF10_0166	conserved Plasmodium protein, unknown function	1	36405	5	Yes	low expression
PF10_0168-a	golgi re-assembly stacking protein 2	1	68411	5	No	Already characterised
PF10_0177b	conserved Plasmodium protein, unknown function	2	118420	5	Yes	
PF10_0281	merozoite TRAP-like protein, MTRAP	1	58085	4	No	Already characterised
PF10_0295	conserved Plasmodium protein, unknown function	4	51393	5	Yes	
PF10_0399	rifin	1	36463	0	No	Already characterised
PF11_0035	Plasmodium exported protein, unknown function	1	53488	0	No	Its exported into rbc therefore not a role in invasion
PF11_0040	early transcribed membrane protein 11.2, etramp11.2	2	10231	3	No	Already characterised
PF11_0261	conserved Plasmodium protein, unknown function	1	34708	5	Yes	low expression
PF11_0286	conserved Plasmodium protein, unknown function	1	51186	6	Yes	low expression
PF11_0344	apical membrane antigen 1, AMA1	1	72042	5	No	Already characterised
PF11_0381	subtilisin-like protease 2	1	154801	6	No	Already characterised
PF11_0443	DNAJ protein, putative	2	39362	6	Yes	CANDIDATE
PF11_0466	transporter, putative	6	99350	5	No	
PFL0065w	conserved Plasmodium protein, unknown function	2	12268	2	No	too small
PFL0870w	Thrombospondin-related apical membrane protein	1	41261	5	No	Already characterised
PFL1945c	early transcribed membrane protein 12, ETRAMP12	2	11729	0	No	Already characterised
PFL2205w	conserved Plasmodium protein, unknown function	2	57944	5	Yes	very low expression
PFL2410w	conserved Plasmodium membrane protein, unknown function	9	121332	5	No	too many TMs
PFL2510w	chitinase	1	42787	6	No	Already characterised

Appendix G continued

Gene	Product Description	TM	Molecular Weight	Orthologues	Contender?	Comments
MAL13P1.60	erythrocyte binding antigen-140	1	140597	6	No	Already characterised
MAL13P1.49	conserved Plasmodium protein, unknown function	2	16746	5	Yes	Dubious expression data
MAL13P1.94	conserved Plasmodium membrane protein, unknown function	4	22051	5	Yes	CANDIDATE
PF13_0223	conserved protein, unknown function	1	12259	5	No	too small
MAL13P1.206	Na+ -dependent Pi transporter, sodium-dependent phosphate transporter	10	75296	7	No	Already characterised
PF13_0247	transmission blocking target antigen precursor	1	51487	5	No	Already characterised
PF13_0265	conserved Plasmodium protein, unknown function	1	133582	5	No	Higher expression in rings than schizonts
MAL13P1.342	conserved Plasmodium protein, unknown function	1	77875	5	Yes	low expression
PF14_0016	early transcribed membrane protein 14.1, etramp14.1	2	11427	0	No	Already characterised
PF14_0211	Ctr copper transporter domain containing protein, putative	4	18752	5	No	wrong function
PF14_0369	copper transporter putative	2	27151	5	No	wrong function
PF14_0372	conserved Plasmodium protein, unknown function	3	220749	8	Yes	low expression
PF14_0440	conserved Plasmodium membrane protein, unknown function	8	143163	7	No	too many TMs
PF14_0495	rhoptry neck protein 2	1	249490	5	No	Already characterised
PF14_0572	conserved Plasmodium membrane protein, unknown function	4	23055	5	Yes	
PF14_0607	conserved Plasmodium membrane protein, unknown function	7	127138	6	No	too many TMs
PF14_0691	conserved Plasmodium membrane protein, unknown function	7	64421	6	No	too many TMs
PF14_0735	probable protein, unknown function	1	38449	0	Yes	low expression