LSHTM Research Online

A Combined Molecular, Cell and Structural Biology Approach Towards Characterising Malaria Alveolins

Michael Patrick Coghlan

Thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy

University of London

2020

Department of Infection Biology

Faculty of Infectious Tropical Diseases

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Dedication

For Emma, Forever Ago.

Acknowledgements

Without a doubt, my greatest thanks, deepest appreciation and boundless gratitude goes to my primary supervisor, boss, scientist-in-chief, Hans. Without your support and guidance, I genuinely would not have made it this far. A mentor, friend, colleague and supervisor all in one, you have kept me motivated during times when all seemed hopeless. Thank you for not giving up on me. If I become even half the competent scientist as you, I know I will have made it. It has been an honour and my privilege to have undertaken my doctoral training in the Dessens laboratory. It truly is a rarity to find such a solid lab group and supervisor to work with. It is thanks to you I will look back fondly on my time here, despite the terrible lab book keeping and number of errors I made along the way!

My secondary supervisor, Cara, has shown nothing but kindness and enthusiasm for my project, and welcomed me with open arms in to her laboratory and her group, training me in the world of structural biology and managing to make me feel like my fumbling around in the laboratory was never a hassle. Thank you for agreeing to take me under your wing. The Vaughan group and Rayne-Wolfson laboratory, have been imperative in ensuring that I did not lose my wit amongst the many preps and ÄKTA runs, so my thanks goes out to all members of the ISMB that have been a part of my experience. Special thanks goes to Dr Shomon Miah, a fellow LIDo student, whose friendship, guidance and expertise in structural biology is truly commendable.

Dr. Sadia Saeed, my greatest friend and biggest supporter throughout my entire time in the Dessens laboratory. Thank you for accompanying me along my journey and allowing me to do the same for you. Our friendship was key in making me feel settled in the laboratory and thank you for showing me that you can be a stellar scientist whilst life happens around it. Thank you for being an incredible mentor, for welcoming me in to your culture and enlightening me as to where my very distant cousins ended up in Kalash! Sarim and Ammar are incredibly lucky to have such a strong woman as their mother. I know I have found a friend for life in you.

Dr. Annie Tremp, an embodiment of what an excellent scientist is. Not only managing a successful career alongside a busy family life, you have shown endless compassion and a willingness to share your knowledge with me along my journey in completing my PhD, whilst offering constant reassurance when I felt things were going terribly wrong. It must not be easy coming back to work after maternity leave and settling back in; however, you did so seamlessly, and are a remarkable example of how strong a woman and mother you are to Lucas and Alex.

To Nadine Mogford, a friend and shoulder to cry on, you have been a rock throughout my entire journey, and the whole reason I was able to embark on it. Thank you for

Publications

Coghlan, M.P., Tremp, A.Z., Saeed, S., Vaughan, C.K., Dessens, J.T.
Distinct Functional Contributions by the Conserved Domains of the Malaria Parasite Alveolin IMC1h.
Front Cell Infect Microbiol. 9 (283).

Declaration

I have read and understood the LSHTM's definition of plagiarism and cheating. I declare that this thesis is my own work, and that I have acknowledged all results and quotations from the published or unpublished work of other people.

I have read and understood the LSHTM's definition and policy on the use of third parties (either paid or unpaid) who have contributed to the preparation of this thesis by providing copy editing and, or, proof reading services. I declare that no changes to the intellectual content or substance of this thesis were made as a result of this advice, and, that I have fully acknowledged all such contributions.

I have exercised reasonable care to ensure that the work is original and does not to the best of my knowledge break any UK law or infringe any third party's copyright or other intellectual property right.

Student signature

Date

Abstract

Intermediate filament (IF)-based cytoskeletal networks in metazoans have key roles in cell architecture and plasticity, and as mechanical stress absorbers. Much less is known about IFs in protozoans. Alveolins are a family of putative IF proteins found exclusively in apicomplexan parasites (causative agents of diseases such as malaria, toxoplasmosis, cryptosporidiosis), dinoflagellate algae and ciliates. All alveolins share functional domains that are characterized by possessing 12 amino-acid tandem repeats. These 'alveolin' modules resemble conserved domains found in other protozoan cytoskeletal proteins like articulins. The demonstrated essential nature of alveolins in malaria parasite development, their expression throughout the life cycle, and their absence in vertebrates makes them potentially attractive drug targets for malaria treatment, prophylaxis and transmission control. Moreover, such drugs could be active against a broad range of other apicomplexan parasites, as well as against related pathogenic protozoans. In this context, a better understanding of the core architecture of the *Plasmodium* alveolins and their assembly mechanisms is important.

This project set out to study the structural requirements of the alveolins and their conserved domains for assembly of the protein into the IF network, and for their functional contribution to cell shape, tensile strength and motility in live malaria parasites, using the *Plasmodium berghei* mouse malaria model. The results reveal, based on the ookinete and sporozoite-expressed alveolin IMC1h, that the 'alveolin' module is required for recruitment into the cortical cytoskeleton, consistent with the notion that it holds the properties for IF formation. In addition, the carboxy-terminal conserved domain of IMC1h, structurally unrelated to the 'alveolin' module, is implicated in facilitating parasite motility through direct or indirect interactions with the motility apparatus.

In addition, a structural biology approach was undertaken, aimed at determining the core atomic structure of the alveolins, with various techniques at hand to try and determine both tertiary and secondary structures formed by these proteins. Bioinformatic-based analyses indicated that the 'alveolin' module is structurally ordered, and adopts a predominantly β-strand architecture. High level expression, in soluble form, of various *P. berghei* alveolin domains in bacteria was achieved as amino-terminal fusions with the protein tag NusA. However, further purification of these recombinant alveolins was severely hampered by problems with solubility after cleavage of the NusA tag, or after concentration, resulting in protein precipitation. Whilst these problems have thus far precluded structural analyses by biophysical means, the observations could reflect actual physical properties of the alveolins and the way by which these molecules assemble in the cell into the insoluble IF network structure, possibly via the intermittent formation of shorter oligomers (protofilaments). Work is ongoing to optimise purification protocols.

Table of Contents

List of Figures

List of Tables

Abbreviations

µg	Microgram
µL	Microlitre
µm	Micrometer
µM	Micromolar
3D	Three dimensional
BLAST	Basic Local Align Search Tool
bp	Base pair
C-terminus/terminal (-C)	Carboxy-terminus/terminal
dH2O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylene diamine tetraacetic acid
EGFP	Enhanced green fluorescence protein
EM	Electron microscopy
FBS	Foetal bovine serum
g	Gram
g	G-force
gDNA	Genomic DNA
GFP	Green fluorescent protein
GM	Genetically modified
GST	Glutathione-S-transferase
h	Hours
HCl	Hydrochloric acid
hDHFR	Human dihydrofolate reductase
His	Histidine tag
HRP	Horseradish peroxidase
i.p.	Intraperitoneal
IF	Intermediate filament
IMC	Inner membrane complex
Kav	Proportion of pores available to the molecule
kb	kilobase
kbp	Kilobase pair
KCl	Potassium Chloride
kDa	Kilodalton
kg	Kilogram
KO	Knock-out
L	Litre
LB	Luria Broth
M	Molar
Mb	Megabase
Min	Minutes
mL	Millilitre
mm	Millimetre
mM	Millimolar

M_r	Relative molecular mass
mRNA	Messenger RNA
NaCl	Sodium chloride
NaOH	Sodium hydroxide
nm	Nanometer
N-terminus/terminal (-N)	Amino-terminus/terminal
NusA	N-utilisation substance protein A
OD_{600}	Optical density measured at a wavelength of 600nm
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PhIL	Photosensitised INA-labelled protein
psi	Pounds of force per square inch of area
PVDF	Polyvinylidene fluoride
RBC	Red blood cell
RNA	Ribonucleic acid
rpm	Revolutions per minute
RPMI	Roswell Park Memorial Institute
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
SOC	Super optimal broth with glucose
SPN	Subpellicular network
TE	Tris(hydroxymethyl)aminomethane ethylene diamine tetraacetic acid
TEV	Tobacco etch virus protease
*Tg*DHFR	*Toxoplasma gondii* dihydrofolate reductase
Tris	Tris(hydroxymethyl)aminomethane
u	Units
UTR	Untranslated region
UV	Ultraviolet
WT	Wildtype

Chapter 1

General Introduction

1.1 Introduction

1.1.1 Malaria

Malaria remains one of the most devastating parasitic diseases affecting the globe today. The estimated number of incidences of malaria in 2018 was 228 million, causing an estimated 405,000 deaths. Ten countries in sub-Saharan Africa and India together make up the large majority of malaria cases worldwide, while malaria cases are on the rise in the Americas, despite spending over US$ 2.7 billion globally on fighting the disease in 2018 (WHO Malaria Report, 2019).

The disease malaria is caused by the apicomplexan protozoan parasite *Plasmodium,* part of the phylum Apicomplexa*.* Apicomplexan organisms comprise of a number of significant parasites that burden both vertebrate and invertebrate species, including *Toxoplasma* and *Cryptosporidium*, causing toxoplasmosis and cryptosporidiosis respectively. Apicomplexans all have complex life cycles and present a wide variety of morphologically distinct shapes depending on genus and stage. Key morphological features of Apicomplexan parasites include the presence of an 'apical complex', containing polar rings and specialized secretory organelles such as rhoptries and micronemes, which are located at the anterior end of the infective life stages, as well as a unique cortical cytoskeletal structure known as the pellicle (Gould *et al*., 2008). The Apicomplexa are part of the superphylum Alveolata, which also includes ciliates and dinoflagellates. Alveolates are characterized by possessing a system of adjoining cytoskeletal components underneath the plasma membrane, known as alveoli, and the pellicle is the equivalent structure in *Plasmodium* (Gould *et al*., 2008).

Of the more than 100 identified species of malaria, five are known to infect humans. These are *Plasmodium (P.) falciparum, P. vivax, P. ovale* sp. (c*urtisi and wallikeri*)*, P.*

malariae and the newly emerging human pathogen of interest *P. knowlesi* (Kantele and Jokiranta, 2011, Calderaro *et al*., 2013, Fuehrer and Noedl, 2014)*.* Each species affects different areas of the globe and have different prevalence's in each area. *P. falciparum* is found predominantly in Africa and has the largest number of deaths associated with its infection. There are few antimalarial drugs available to treat the disease, and the increase in antimalarial drug resistance, including against the most widely used drug artemisinin, is cause for concern (Amato *et al*., 2018, Rosenthal, 2018, Uhlemann and Krishna, 2005). This emphasizes the need to increase our understanding of the parasite and therefore novel strategies for its control.

1.1.2 The *Plasmodium* life cycle

Plasmodium species have a complex life cycle involving different hosts, both the definitive host *Anopheles spp* and the intermediate vertebrate host (Aly *et al*., 2009). It involves different life stages with distinct morphological features, characteristic of apicomplexan parasites. A newly acquired infection will begin when a parasite-infected mosquito injects infective sporozoites during a mosquito bite (Fig. 1.1A). Sporozoites migrate to the liver through the blood, where they infect hepatocytes (Prudencio *et al*., 2006). Here, they develop in to a hepatic schizont, undergoing nuclear division known as schizogony (Cowman and Crabb, 2006). Each mature liver schizont can contain thousands of merozoites that are subsequently released in to the blood stream (Fig. 1.1B). These merozoites then invade red blood cells and enter the erythrocytic life stage (Fig. 1.1C). Once red blood cell (RBC) invasion has occurred, the merozoite develops in to a schizont which ruptures, releasing 8-32 merozoites (depending on species and quality of RBC), starting the erythrocytic cycle again upon merozoites

invasion. These asexual life cycle stages of the parasite are haploid (Cowman and Crabb, 2006).

In order for the parasite to complete its life cycle, during the asexual life stage, a small number of reinvading merozoites will develop in to sexual stage precursor cells named gametocytes in a process known as gametocytogenesis (Fig. 1.1D). At this point, an *Anopheles* mosquito is required to blood feed on the already infected intermediate host and ingest the female and male gametocytes. After entering the midgut of the mosquito, rapid division and egress of male gametocytes (microgametes) and escape of female gametocytes (macrogametes) from RBCs occurs due to environmental cues, in a process known as gametogenesis (Delves *et al.*, 2013). There are several contributing factors to gametogenesis, however major cues include a drop in temperature, a rise in pH, and xanthurenic acid, also known as gametocyte activating factor (Arai *et al.*, 2001). Each male gametocyte will produce eight microgametes, and a female gametocyte will produce a single macrogamete. The microgametes are motile and forcefully beat their flagella when leaving the RBC in a process known as exflagellation, facilitating their aim to penetrate a macrogamete (Fig. 1.1E). Once penetration and fertilisation occurs, the zygote develops in to a motile ookinete that is able to invade the midgut epithelium (Fig. 1.1F). On the basal lamina of the midgut wall, the ookinete transforms in to the oocyst. Over a period of 1-3 weeks, the oocyst grows in size with multiple rounds of endomitosis. A final cytokinesis step culminates in the formation of hundreds of sporozoites which are released from the oocyst, allowing them to invade the salivary glands (Fig. 1.1G). Here, they lie dormant waiting to be injected to a new host with the next blood feed (Aly *et al.*, 2009).
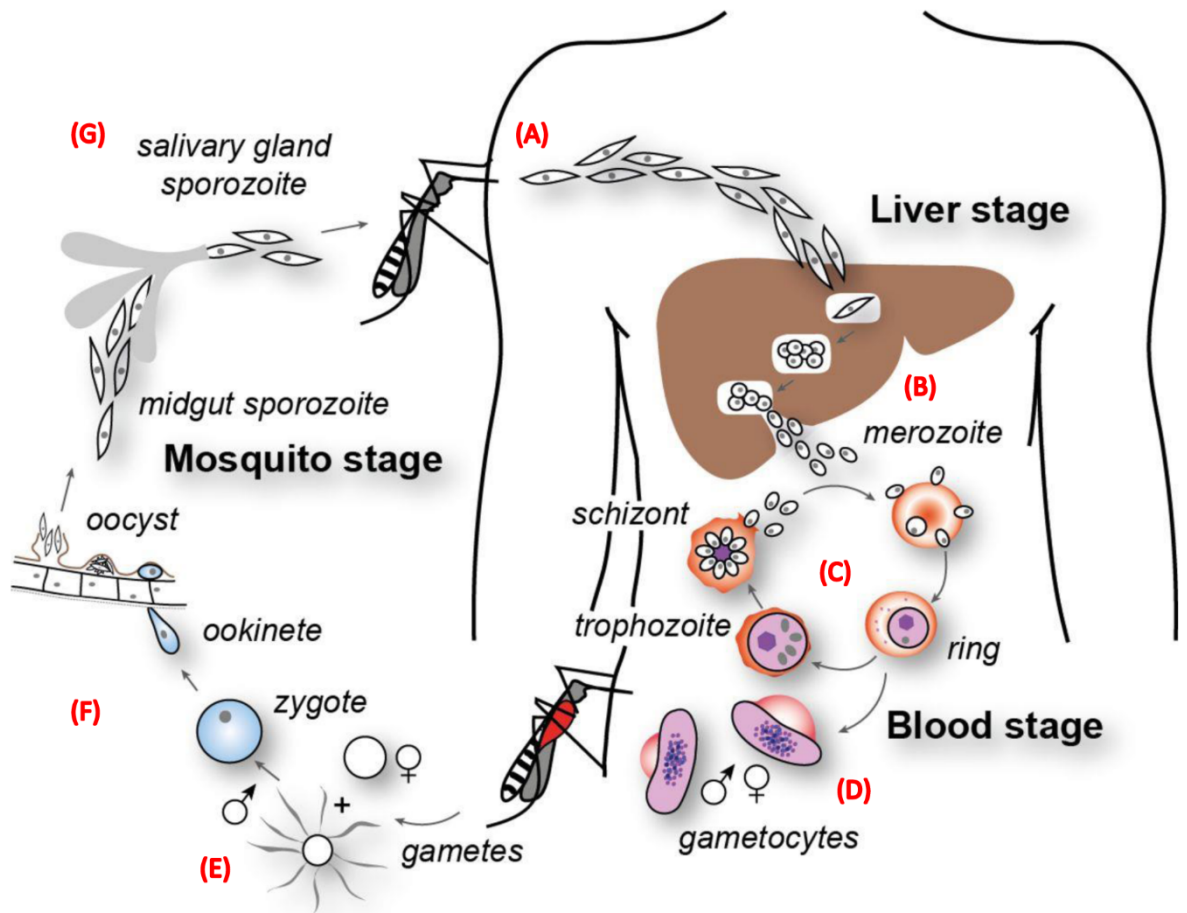
Fig 1.1 The malaria life cycle. The *Anopheles* mosquito bites and infects a human with sporozoites (A). Sporozoites then migrate to the liver where they pass through numerous hepatocytes before forming a liver schizont, releasing merozoites (B). Merozoites invade RBCs and grow in to trophozoites and schizonts, rupturing and repeating the cycle (C). Gametocytes are taken up by a feeding mosquito, causing gametogenesis (E). Upon fertilisation, the zygote grows in to an ookinete (F), and then an oocyst, releasing sporozoites that can invade the salivary gland (G). *Adapted from (Cowman et al., 2012).*

1.1.3 The 'zoite' stages

Plasmodium spp. have three motile and invasive life stages (also known as 'zoites'): the merozoite, the ookinete and the sporozoite (Fig 1.2). Despite their distinct sizes, shapes and host cell specificities, these zoites have common cellular architectures that facilitate motility, invasion and egress: the apical complex and the pellicle. The apical complex is responsible for aiding the zoite in interacting with the host cell and its invasion. The pellicle assists with the parasites gliding motility, a substrate-dependent movement that is essential for motion in and around the host environment. This substrate-dependent movement is supported by an actin-myosin motor, highly

9

conserved between apicomplexan parasites. This actomyosin network is located beneath the plasma membrane and is anchored into the inner membrane complex. Each invasive life stage of the parasite is specialised to invade different host cells (Sinnis and Coppi, 2007, Siden-Kiamos *et al.*, 2006).
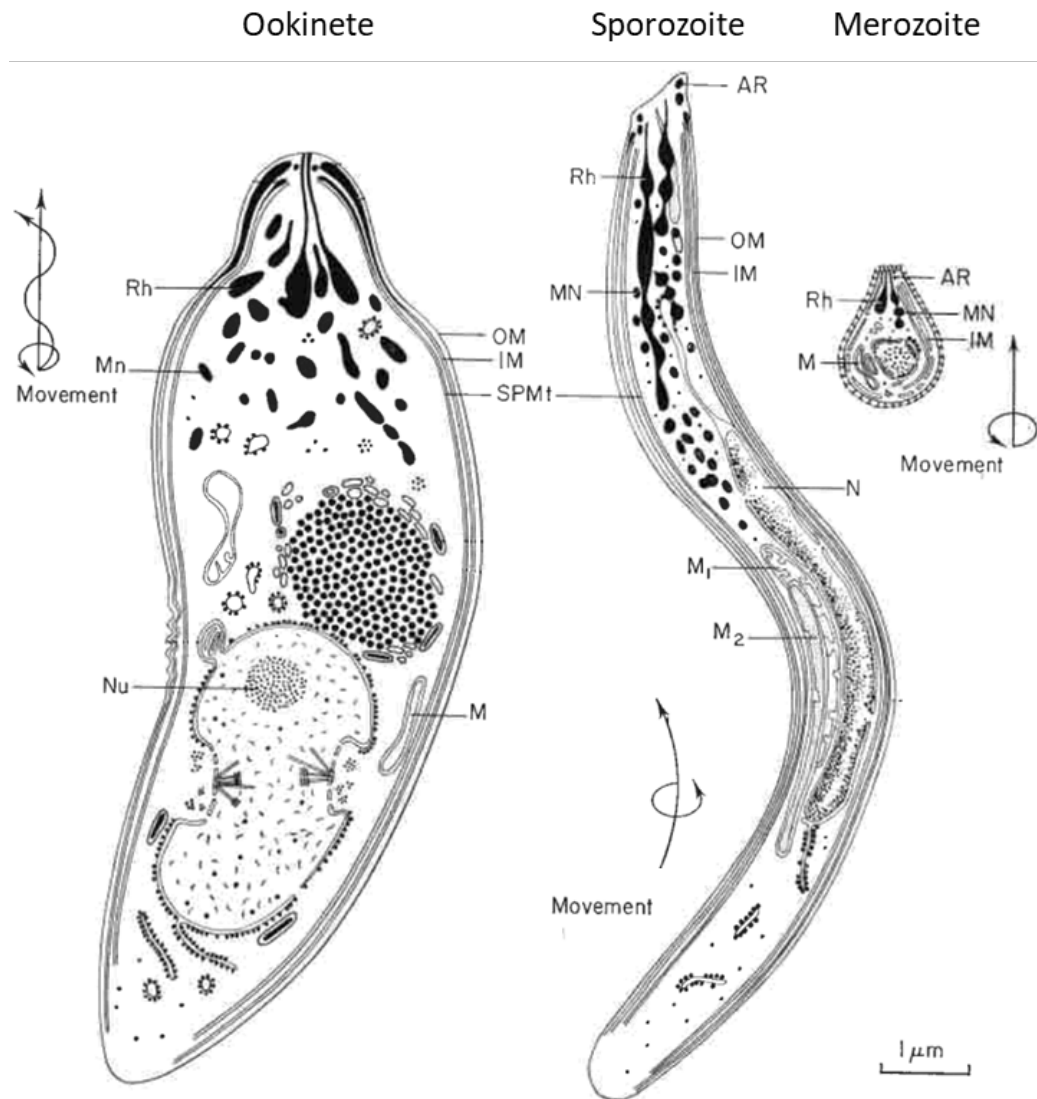
Fig 1.2 An ultrastructural summary of the invasive life stages of *Plasmodium* species. The ookinete, the merozoites and the sporozoite, with arrowed movements showing the direction and rotation of movement. Labels for the diagram include: Rh: rhoptries. Mn: micronemes. N: nucleus. Nu: nucleolus. M($_{1,2}$): mitochondrion. Mn: microneme. OM: Outer membrane. IM: Inner membrane. AR: apical ring. SPMt: Subpellicular microtubules. 1µm scale bar for size. Adapted from (Sinden, 1978).

1.1.4 The merozoite

The merozoite is the invasive life stage of the parasite that infects RBCs during asexual reproduction. Once an infection is established in hepatocytes and undergoes exoerythrocytic schizogony, merozoites are released in to the blood stream where they can invade and replicate within RBCs. During the brief extracellular movement of the merozoite before invading new RBCs, the merozoites are exposed to host innate and adaptive immune responses (Guevara Patino *et al*., 1997). The immune system may prevent a small number of merozoites from establishing an infection, however the high number of merozoites released from initial liver schizogony allows enough merozoites to reliably establish a host infection (Garnham, 1951, Belachew, 2018). Erythrocytic schizogony releases fewer merozoites than exoerythrocytic schizogony and shows no enlargement of RBCs upon infection, with up to 32 merozoites depending on the *Plasmodium* species (Antinori *et al*., 2012). The apical complex contributes to RBC invasion. After an interaction and attachment is established, the merozoite is able to reorientate itself so that the apical end of the parasite makes contact with the RBC. On invasion of the RBC, a parasitophorous vacuole is formed from which the parasite resides, growing and replicating to form a schizont, leading to release of more merozoites. On invasion, the merozoites thick coat, formed of thin filaments used for attachment, are cleaved from it (Bannister *et al*., 1986). There are a number of secretory organelles located in the apical complex that assist the entire process. These are: (1) micronemes, involved in gliding motility attachment and invasion; (2) rhoptries, involved in attachment, invasion, and the establishment of the parasitophorous vacuole; (3) dense granules for aiding in forming the parasitophorous vacuole, and (4) exonemes, involved in merozoite egress from an infected RBC (Cowman *et al*., 2012).

1.1.5 The ookinete

The ookinete is involved in host cell invasion and movement in the mosquito vector. *Plasmodium* infection of the vector constitutes a major population bottleneck for the parasite (Alavi *et al.*, 2003), and is therefore an attractive target for transmission prevention strategies. By preventing the development of the parasite in the *Anopheles* midgut, the vector will no longer be able to transmit the parasite to the human hosts. Upon fertilization, key changes take place during the differentiation of the spherical zygote into the elongated ookinete. Surface proteins on the parasite are changed to aid the ookinete in interacting with the mosquito environment, most notably P25 and P28 (Tomas *et al.*, 2001). The apical complex and pellicle form, and the zygote undergoes DNA replication, followed by meiotic division, resulting in a tetraploid nucleus with four haploid genomes (Sinden *et al.*, 1985). Once the mature ookinete is formed, it must escape the protease-rich midgut lumen before it is digested with the blood meal (Billingsley, 1990). The ookinete must penetrate the peritrophic matrix that envelops the blood meal, which is aided by the release of chitinases that break down this chitinous structure (Shahabuddin *et al.*, 1993). The ookinete must then invade and traverse cells in the midgut epithelium to reach the basal lamina, where transformation into the oocyst takes place.

1.1.6 The sporozoite

At the basal lamina, the oocyst grows and undergoes multiple steps of mitosis and replication of organelles within a short space of time to form up to thousands of sporozoites (Rosenberg and Rungsiwongse, 1991). Cytokinesis begins with the invagination of oocyst cytoplasm forming sporoblasts. These sporoblasts then form sporozoites using a budding process at the surface of the limiting membrane, a process

concomitant with pellicle formation (Matuschewski, 2006). As the budding sporozoite elongates, a haploid nucleus moves into it along with mitochondrion and apicoplast. Micronemes and rhoptries are formed at the apical region of the emerged sporozoite. Sporozoites become flexible and motile after their release from the sporoblast. On rupture of the oocyst, sporozoites are released into the haemolymph of the mosquito and migrate to the salivary glands of the insect (Smith and Jacobs-Lorena, 2010, Matuschewski, 2006). The sporozoites infect the glands via attachment and invasion of the epithelial cells using surface proteins. At this point, the sporozoites are ready to be injected along with the saliva into the vertebrate host when taking the next blood meal. Once in the vertebrate host, the sporozoites move into the blood stream via which they reach the liver. Upon reaching the liver sinusoids, they cross the endothelial layer to reach the hepatocytes. Surface proteins are again essential for this interaction, with circumsporozoite protein and thrombospondin-related adhesive protein providing interaction with glycoproteins on the surface of hepatocytes. Sporozoites will migrate through a number of hepatocytes before establishing an infection and forming a parasitophorous vacuole, starting exoerythrocytic schizogony (Mota *et al.*, 2001).

1.1.7 The pellicle

The pellicle is a trilaminar cortical structure that provides flexibility and strength for migration and invading the host cell. It contains three major components: the plasma membrane, the inner membrane complex (IMC), and the subpellicular network (SPN) consisting of intermediate filaments (Fig 1.3). A system of subpellicular microtubules run perpendicular beneath this subpellicular network, from the apical polar ring toward the posterior end of the parasite. The IMC lies directly underneath the plasma

membrane and consists of a system of flattened vacuoles derived from the endoplasmic reticulum. The IMC contains intramembranous particles that run alongside the subpellicular microtubules, suggesting that they could link these two structures. Tightly connected to the IMC on the cytoplasmic side is the SPN. The SPN was first characterised in the apicomplexan parasite *Toxoplasma gondii* (Mann and Beckers, 2001). The presence of the SPN was confirmed from detergent extraction of *T. gondii* tachyzoites, revealing a two dimensional network of interwoven intermediate filaments on the cytoplasmic surface of the IMC. The SPN maintains the shape of the cell and is resistant to detergent extraction, indicating that it is a membrane skeleton of the parasite, providing mechanical strength (Mann and Beckers, 2001). In *T. gondii*, variations in the stability of the subpellicular network were reported in mother and daughter cells (Harding and Meissner, 2014). Developing parasites had detergent-soluble SPN, whereas mature parasites had detergent-resistant SPN (Mann *et al*., 2002). Ookinetes have 40-60 microtubules, sporozoites 18-19, and merozoites as few as 2-3.
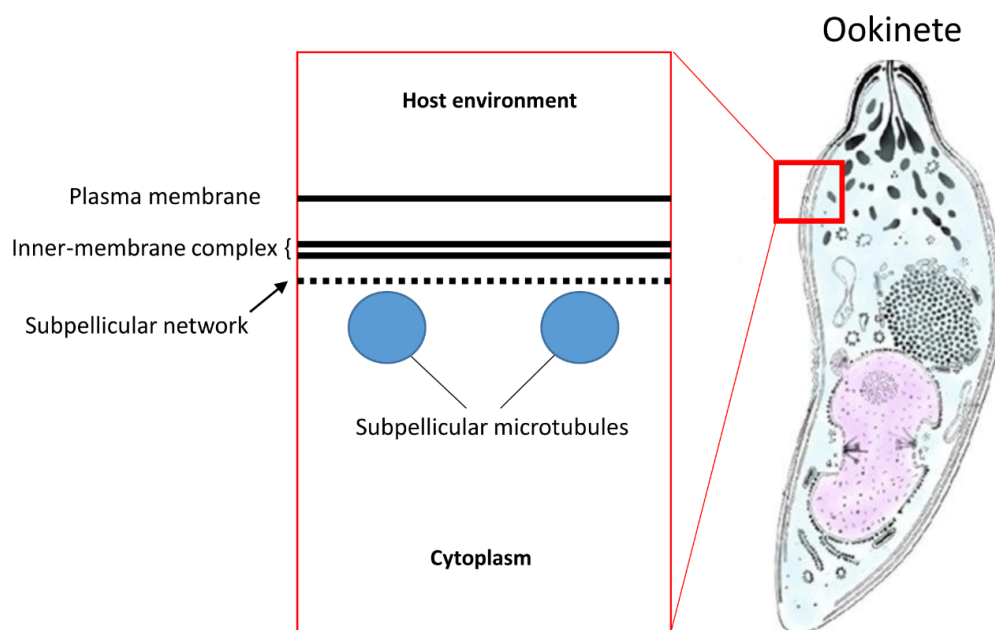
Fig 1.3 The pellicle trilaminar complex found in the motile zoite stages of malaria with the plasma membrane, inner-membrane complex and subpellicular network shown. Adapted from (Sinden, 1978).

14

1.1.8 The alveolins

Through investigating the SPN in *Toxoplasma gondii*, Mann and Beckers (2001) identified the protein TgIMC1 as a major constituent. This led to the identification of a family of structurally-related proteins in apicomplexan parasites, initially named IMC1 proteins (Khater *et al*., 2004). This protein family was later renamed 'alveolins' due to their restricted presence in alveolates (Gould *et al*., 2008). Thirteen alveolins have now been identified in *Plasmodium* (Table 1.1) (Al-Khattaf *et al*., 2015) and fourteen in *Toxoplasma* (Anderson-White *et al*., 2011). The structurally homologous domains of the alveolins are characterized by possessing tandem repeat sequences of twelve amino acids that are thought to constitute the filament-forming properties of these proteins similar to the coiled-coil domains of other metazoan intermediate filament proteins (Al-Khattaf *et al*., 2015), such as the keratins. Alveolins share sequence homology with articulins and plateins, cytoskeletal proteins of free-living protists that also possess 12 amino acid tandem sequences (Al-Khattaf *et al*., 2015). The average protein helix is 12 amino acids in length (Presta and Rose, 1988), further supporting the argument that these tandem repeats may form an α-helical coiled-coil secondary structure. The *Plasmodium* alveolins are differentially expressed in zoite stages throughout the life cycle (Table 1.1).

The alveolins have been characterised through a number of targeted gene disruptions in *P. berghei* (Khater *et al*., 2004, Tremp *et al*., 2014). Khater *et al* (2004) found that the structural orthologue to TgIMC1, PbIMC1a, is involved in the development of the sporozoite and essential for maintaining its shape, tensile strength and motility. When knocked out, it was found that the mutant sporozoites could not infect the salivary glands or be transmitted. This experiment highlighted the essential role the SPN and

alveolins have in development, motility and infectivity. Similar roles were later identified for the ookinete-expressed alveolin IMC1b (Tremp *et al*., 2008), and the ookinete and sporozoite-expressed IMC1h (Tremp and Dessens, 2011).

Table 1.1 *A list of all 13 identified alveolin proteins and the currently identified zoite life stages where they are expressed. Adapted from (Al-Khattaf et al., 2015).*

	LIFE STAGE EXPRESSED		
GENE	Merozoite	Ookinete	Sporozoite
pbimc1a			x
pbimc1b		x	
pbimc1c	x	x	x
pbimc1d		x	
pbimc1e	x	x	x
pbimc1f			
pbimc1g	x	x	x
pbimc1h		x	x
pbimc1i			
pbimc1j			
pbimc1k			
pbimc1l			
pbimc1m			

A phylogenetic analysis of the alveolins detail the evolutionary relationship to one another. Results from Al-Khattaf *et al* (2015) suggest that, throughout the 13 alveolin domains, there is a significant difference in the sequences between two separate alveolin domains, termed alveolin domain 1 (red) and alveolin domain 2 (green) (Fig 1.5 A and B). Both domains are defined by the presence of a tandem repeat periodicity, however there are primary structure differences that split them in to two clades. Whilst the presence of alveolin domain 1 is present in all proteins (except IMC1d, proven to be functionally redundant (Al-Khattaf *et al.,* 2015)), the alveolin domain 2 is present in 4 out of the 13 alveolins. The differential expression of these alveolins, as detailed in table 1.1, provide an interesting narrative alongside the contributions the individual alveolin domains have to the overall function of the

proteins. For example, IMC1a is a sporozoite-specific alveolin, only expressed during this life stage. Conversely, IMC1b is an ookinete-specific alveolin. IMC1c, IMC1e and IMC1g, however, are expressed in all three zoite stages of the malaria lifecycle.
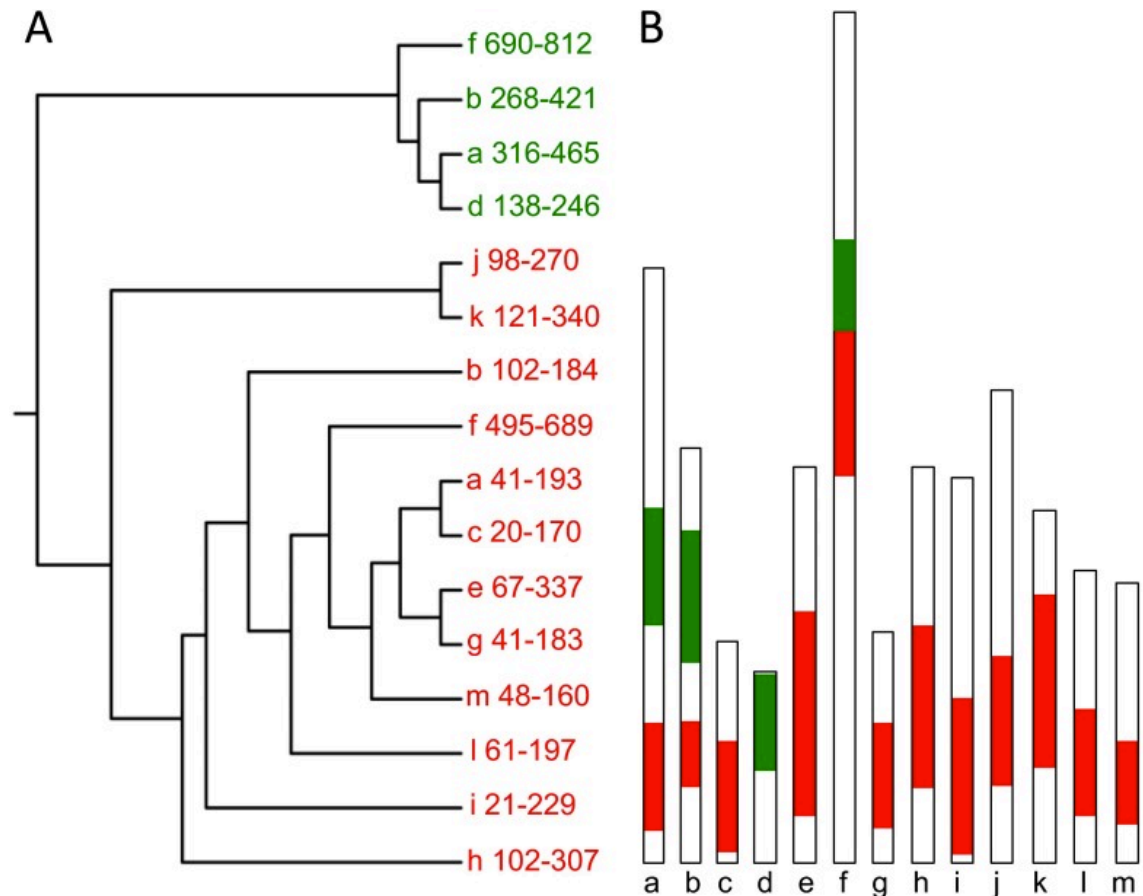
Fig 1.4 A) Phylogenetic analysis of the alveolins *Pb*IMC1a – h, with numbers corresponding to amino acid coordinates of the conserved domains within the protein. The analysis shows two distinct clades of domains, type 1 (red) and type 2 (green). B) a schematic diagram of the 13 alveolins showing the relative positions of either/both the type 1 and type 2 domains. Adapted from Al-Khattaf *et al,*. 2015.

Sequence analysis of the 13 alveolins has shown that IMC1e is the most recent common ancestor of the alveolins. Using a similarity matrix from BLAST scores, IMC1e was ranked as being the most representative of the alveolins as a whole, detecting the most number of alveolins in BLAST homology searches. This is in comparison with IMC1d, the most divergent alveolin, only detecting four alveolins in BLAST searches. This was also reflected in other genera across the Apicomplexa phylum, for example the *Plasmodium* IMC1e sequence was able to detect 13 out of 14 *Toxoplasma* IMC

17

proteins (Anderson-White *et al,.* 2011), whilst IMC1d detected only five (Al-Khattaf *et al.,* 2015).

The phylogenetic clade split between the alveolin domain 1 and alveolin domain 2 is further highlighted through a primary sequence analysis of the domains in IMC1b showing periodicity within the primary sequence (Fig 1.5).
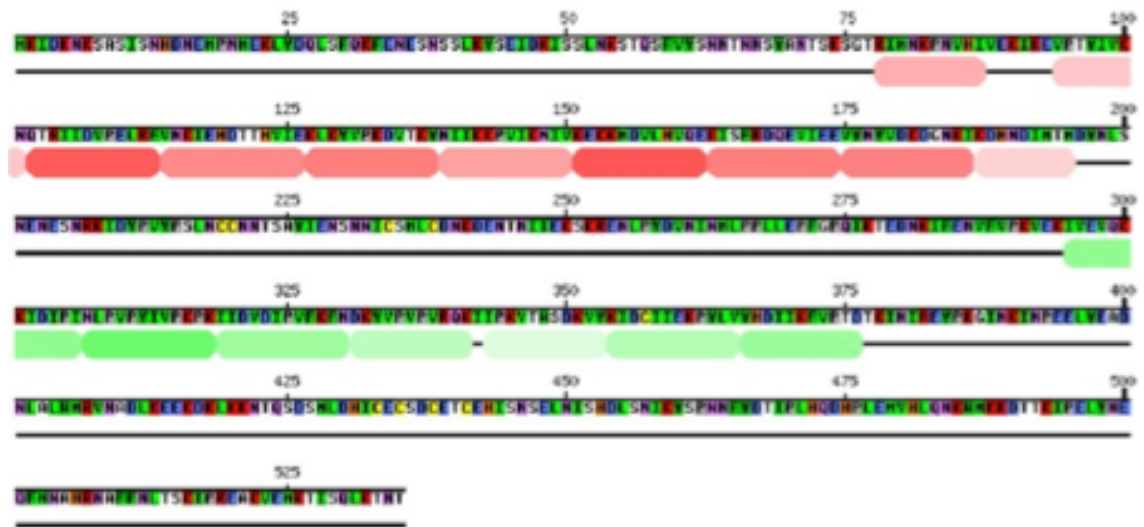
Fig 1.5 The primary sequence of IMC1b showing the 12 amino acid tandem repeat periodicity identification present in alveolins by the programme HHrepID. Alveolin domain 1 is in red, alveolin domain 2 is in green. The red and green sections coincide with the highly conserved domains that define an alveolin domain. Similar results are also seen in closely related articulins and plateins. Adapted from (Al-Khattaf *et al,.* 2015).

The areas shown with red (domain 1) and green (domain 2) highlight the 12 amino acid periodicity, where there is a repetition of primary sequence pattern across both domains. This further adds to the argument that the secondary structure of these proteins allow for intramolecular interactions between alveolins, as is seen with intermediate filaments.

The alveolin domains are highly conserved throughout *Plasmodium* species, with defined borders surrounding the areas that are shown to have tandem repeat periodicity (Fig 1.6).

```
Pb.ANKA     MKIDKNKSASISNHDNEMPNMEKLYDQLSFQKFENESNSSLKYSE--IDKI--SSLNKST
Pchas       MKTDKNKSTNISKYDNEMPNMEKLYDQLSFQKFENESSSSLKYSD--IEKM--SGLNKSI
Pyoelii     MKIDKNKSASISNHDNEMPNMEKLYDQLSFPKFEDESNSSLKYSE--IDKI--SSLNKTI
Pkn.H       MQTTKSKSISALSSESGTTNFSK-------HREMERPSAAKNGA--DEEQFALSLKRSI
Pviv        MDTTKSKSSSALSSESGTANLPK-------HREMERTAATKNGA--DGEQSALSLNRSI
Preic.CDC   ----------MSHYNPEQINFVN------------NHRNNVKEGYFYKNNLLKEHINKM-
Pf.3D7      ----------MSHYNPEQLNFVN------------NHRNNVKEGYFYKNNLLKENNNKM-
Povale      MEATKSKCTFISNNKEENAKFRKNYEKLIFQQLETETNTPISYSN--HGERLPASFTKSV
Pmalariae   -----------------------MKFQNFNIEPRTPLKYGG--KEGRSLSSLRGSI
                                                     :         .  .

Pb.ANKA     QSFVYSNNTNNSYANTSKSGTRIMNKPNVHIVEKIKEVPTYIVKNQT RIIDVPELRFVNK
Pchas       QSSIYSNNTSNSYANTPKSGTRIMNKSNVQIVEKIKEVPTYIVKNQT RIINVPEVRFVNK
Pyoelii     QSSIYSNNTNNSYANTPKSGTRIMNKPNVHIVEKIKEVPTYIVKNQT RIIDVPEVRFVNK
Pkn.H       SPFTTD----RSYANGQTKDTKLIDSANFDFAGKAVDTKTYITQNKT KVIEVPELRFIDK
Pviv        PPFSRD----RTYANGQTKETKLIDSANHDFAGKAIDTQTYIPHNKT KVIEVPELRFIDK
Preic.CDC   --ESSKNCIMVPLINKPSHSENNFSKNNID-ENINEEKSKYQFEYQN KIIQVPELKYVDK
Pf.3D7      --ESSKNCIMVPLINNPSHSENNFSKNNRD-ENINEEKSKYQFEYQN KIIQVPELKYVDK
Povale      PPP---SCHSS-IVYTSIKETQIRNKPNIEFVEKIKEVPSYVYQSKN RVIEIPEVTFVDK
Pmalariae   PISTSSSCSSTCAPNTKEKETKIIQKQNAELVEKIREVPNYIFHNKT KVIEVPELRFIDK
                      .  .. * .        :    .*  . .::*::**: :::*

Pb.ANKA     IEHDTTHVIEKLKYVPKDVTKYNIIKKPVIKNIVKEKKMDVLHVQEKISFRDQEVIEEVY
Pchas       VEHDTVEVIEKLKYVPKEVTKYNIIKKPVIKNIVKEKKMDVLHVQEKISFRDQEVVEEVY
Pyoelii     IEHDTVNVIEKLKYVPKDVTKYNIIKKPVIKNIVKEKKMDVLHIQEKISFRDQEVVEEVY
Pkn.H       IEYDP-FVIEKLRYVPKQVTKYNIIEKPVIKNIVSEKKVDVLYVQEKISFKDQEIVEEVY
Pviv        IEYDP-FVIEKLRYVPKQVTKYNIIEKPIIKNIVTDRKVDVLYVQEKISFKDQEIVEEVY  ALV1
Preic.CDC   IVYDP-VIIEKVKYVPKDVIKYNIIKKPVIKNIITEKKVDILQVKEKISFKEEEIVEDVY
Pf.3D7      MVYDP-VIIEKVKYVPKEVIKYNIIEKPVIKNIITEKKVDVLQVKEKISFKEEEIVEDVY
Povale      IQYDP-VVVEKLKYVPKEVTKYNIIEKPVIKNIITEKKVDVLHVQEKISFKDQEIEEDVY
Pmalariae   IEYDP-QVIEKIKYVPKEVTKYNIIEKKVINNIVTEKKVDVLQVQEKIRFKDQEIVEEVY
            : :*    ::**::****:* *****:* :*:**.::*:*:* ::*** *:::*::*:**

Pb.ANKA     NYVDKD GNKIKDHNDIMTMDYNLSNENESNRK-IDYPVYPSLNC----------------
Pchas       NYVDED GNQIKDSQDIMTMDYILSNENESNKK-RECPVYPSLNC----------------
Pyoelii     NYVDKD GNQIKDHNDIMTMDYIQSNENESNRK-TDCPVYSSLNC----------------
Pkn.H       NYVDKD NNRILADQGSPSLDRSALPITEMTQGEWDSSALSPHVNSDIR--------RNSI
Pviv        NYVDKD NQRILADQGGPSPSRAVLPLGGTTRGECDSSASCPYINGDII--------RNSM
Preic.CDC   NYVDKD LNTKWNESQYDNEMYKDLTKKKNYIG-----------QNHLL----P-------
Pf.3D7      NYVDKD LNTKWNESQYDNEMYKDLTKKKNYIG-----------QNHLL----P-------
Povale      HYVDKD NNTYVKDHGSSTKIKRSLVEGDMHRK-FDYPAYQPARSSNTNGANVPNVGNYGS
Pmalariae   HYVDKD NNRFINKQGISTENYTSSKDKGVNEK-FDQLTYQYINN----------------
            :***:* :                        .

Pb.ANKA     --CNNT--SAYIENSNNICSMLCDNKDENTNIIEKSK---------RENLPYDVNINMLP
Pchas       --CNNT--SAYVENSNNICSMAYNNKIDNSNIIGISK---------GETLPYDVDLNVLP
Pyoelii     --CNNT--SAYIENSNNICSMVCNNKDENTNIIEKSK---------RENLPYDVDLNMLP
Pkn.H       HNSNNHN-----------------------------------DNRNSMGGASDGVLP
Pviv        HHCNNHN-----------------------------------DSRSSMGGAPDGVLP
Preic.CDC   --------------NNINKMG---------------------HINDCTYKHINNNIT
Pf.3D7      --------------NNINKMG---------------------HINDRTYKHINNNIT
Povale      NICNNHDNAMYEMHANDVCEMHDSDVYDMQDNAVYDRHDNAVYNMYDNDVFDNQCSRALP
Pmalariae   -------------------------------------NINTSNDIDYTKSCGSQLP
                                                                     :

Pb.ANKA     PLLEPFGPQIKTEDNKIFENVFVPKVEKIVEVQKKIDIPINLPVPYIVPKPKIIDVDIPV
Pchas       PLLEPFGPQIKTEDNKIFENVFVPKVEKVVEVQKKIDIPINLPVPYIVPKPKIIDVDIPV
Pyoelii     PLLEPFGPQIKTEENKIFENVFVPKVEKVVEVQKKIDIPINLPVPYIVPKPKIIDVDIPV
Pkn.H       SLLEPFGPQVNITENKIFENVFIPKVEKVVKNKIDIPINLPVPYIVPKPKIVDVDIPV
Pviv        PLLEPFGPQVNVTENKIFENVFVPKVEKVVEVKKKIDIPINLPIPYIVPKPKIVDVDIPV
Preic.CDC   NLLEPFGPQIDVEENKIIENVFVPNVEKVIEVNKKIDIPINLPVPYIVPKPKIIDVDVPV
Pf.3D7      NLLEPFGPQIDVEENKIIENVFVPNVEKVIEVNKKIDIPINLPVPYIVPKPKIIDVDVPV
Povale      PLLEPFGPQVNEDNKTFENVFVPKVEKVVEVKKKIDIPINLPVPYIVPKPKIIDVDVPV
Pmalariae   PLMEPFGPQVYLGENKTFENIFIPKVEKVVEVKKKIDIPINLPIPYIVPKPKVIDVDVPV
             *:*******:   :** :**:*:*:***::**::*********:********::***:**
```

19

```
Pb.ANKA     FKFNDKYVPVPVRQKIIPKVTWSDKVYKIDCIIEKPYLVYHDIIKFVPTDTKINIREYPK
Pchas       FKFNDKYVPVPVRQKIIPKITWSDKVYKVDCVIEKPYLVYHDIIKFVPTDTKISVREYPK
Pyoelii     FKFNDKYVPVPVRQKIIPKVTWSDKVYKVDCIIEKPYLVYHDIIKFVPTDTKINVREYPK
Pkn.H       FKFNDKYVPVPIRQRIIPKVTWTDKVYKVDCVVEKPYLVYHDIVKIVPTDTKISVREYPK
Pviv        FKFNDKYVPVPVRQRIIPKVTWTDRVYKVDCVVEKPYLVYHDIVKIVPTDTQISVREYPK     ALV2
Preic.CDC   FKFNDKYVPVPVSKKIIPKITWTDKIYQVDCLIEKPYLVYHNIIKMVPTDSKITVREYPK
Pf.3D7      FKFNDKYVPVPVSKKIIPKITWTDKIYQVDCLIEKPYLVYHNIIKMVPTDSKITVREYPK
Povale      FKFNDKYIPVPVRQKIIPKVTWSDKVYQVDCVIEKPYLVYHDIIKFVPCDSKIMVREYPR
Pmalariae   FKFNDTYVPVPVRQRIIPKVTWSDKVYKVDCLVEKPYLVYHNIIKVVPTDSKITIREYPK
            *****.*:***: ::*****:**:*::*::**:*********:*:*.** *::* :****:

Pb.ANKA     GINKINPEELYEADNLALWMRVNADLKEEKDKLKKNT--------------QSDSMLDH
Pchas       GINKINPEELYEADNLALWMRVNADLKEEKDNLKKNT--------------QTESTIDH
Pyoelii     GINKINPEELYEADNLALWMRVNADLKEEKDNLKKSA--------------QPDIILDH
Pkn.H       GITKINPEELYEVDNLALWMRVNADLKEEKDALKKTQ---------------DNFSDH
Pviv        GITKINPEELYEVDNLALWMRVNADLKEEKDALKKTQ---------------DNCSDR
Preic.CDC   GIKKINPEELYEVDNLALWMRVNADLKQEHDQMKNEKYETNKKKGKGETEELDDNILSDH
Pf.3D7      GIKKINPEELYEVDNLALWMRVNADLKQEHDQMKNEKYETNKKKGKGETEQLDDNISSDH
Povale      GITKINQDELYEVDNLALWMRVNADLKEEKDKL--------------------NLSDH
Pmalariae   GINKINTEELYEADNLALWMRVNADLKHEKDKLKKTQ---------------DSLSDH
            **.*** :****.**************.*:* :                       *:

Pb.ANKA     ICECSDCE-TCEHISN-SELNISHD-LSNIKYSPNNFYDTIPLHQDHPLEMVHLQNKWMK
Pchas       ICECSDCE-TCEHISN-SELNISHD-LFNTKSSPNNFSDTIPLHQDHPLEMVHLQNKWMK
Pyoelii     ICECSDCE-TCEHISN-SELNISHD-LSNIKSSPNNFSDTIPLHQDHPLEMVHLQNKWMK
Pkn.H       TCECSNSD-SFEECSSTTELNSYQDGATTIKSSNENVLDTLPIHPGHPLEIVHLQNKWIK
Pviv        TCECSIAD-SFEECSSTSELNSSQDGATTIKSSNENVLDTLPIHPGHPLEMVHLQNKWMQ    ALV3
Preic.CDC   TCECSESYETYGKLSN-GEFNSSNE-ETTIKSSNENILDTLPLHPGHPLEFIHLQNKWIS
Pf.3D7      TCECSESYETYEKLSN-EEFNSSNE-ETTIKSSNENILDTLPLHPGHPLEFIHLQNKWIN
Povale      TCECSDCD-SYEEVSN-TELNSSQD-VSTIKSSNENLSNTMPIHPGHPLEIIHLQNKWIK
Pmalariae   TCECSDCD-TYEQNSN-SELNSSQD-ATTIKSSNENLLSTLPLHPGHPLEITHLQNKWIK
             ****  .   :   .  *.   *:*  ::    .  * *  :*.  .*:*:*  .****: ******:.

Pb.ANKA     KDTTKIPELYNEQFMNAHRNAFFNLTSKIPREAKVEMKTISQLKTNT-------------
Pchas       QDATKIPELYNEQFMNAHRNAFFNLTSKIPREAKVEMKTIAQLKTNT-------------
Pyoelii     QDTTKMPELYNEQFMNAHRNAFFNLTSKIPREAKVEMKTISQLKTNT-------------
Pkn.H       QDTTRMQELYQDGFFDAHRNAMFNLGTRIPREAEIEVRQIAQLQRFGEKREYDN------
Pviv        QDTTRMMELYQDGFFDAHRNAVFNLATRIPREAEIEVRQISQLQRRMEKRECGTD-----
Preic.CDC   QDTTNIPDMYDQKYLDAHRNAVFNLTTQMPREAEVEAKQLLYIQKKLQQEETL-------
Pf.3D7      QDTTNIPDMYDQKYLDAHRNAVFNLTTQMPREAEVEAKQLLYIQKKLQQEETL-------
Povale      QDTTKTSELYSEQFLEAHKNAIFNLSTRLPREAEIDIKQIFAKPKEDR------------
Pmalariae   QETTRIPELYEEKFLDANRNAVFNLSTQIPREAEIEAKQMFQMQKQMDKRELNTISGKLF
            :::*.  ::*.: :::*::**.*** :::****::: : :
```

Fig 1.6 A Clustal Omega alignment of various *Plasmodium* species IMC1b primary sequences, and the regions of high conservation between the alveolin domains, denoted with ALV 1, for domain type 1, ALV 2 for domain type 2, and ALV 3 for a third conserved domain. Produced using software from (Madeira *et al*., 2019)

A Clustal Omega (Madeira *et al*., 2019) multiple amino acid alignment further highlights the structural similarities between the alveolins across *Plasmodium* species. The presence of a third highly conserved domain, domain 3, shows no tandem repeat periodicity as seen in domain types 1 and 2, and is therefore not a true alveolin. This third domain is found in both IMC1b and IMC1h. Its contribution to the function of the protein is not known.

1.1.9 *Plasmodium* motility

Plasmodium utilise gliding motility to migrate and egress through various tissues in the zoite stages of its life cycle, in both invertebrate and vertebrate hosts. It is a substrate dependent form of locomotion, requiring a substrate from which the parasite is able to inversely propel itself against the direction of force (Heintzelman, 2006, Baum *et al*., 2008). Throughout these motile life stages, the glideosome proteins are conserved, highlighting the importance of these structures for the motility of the parasite. These include proteins such as the myosin motor, myosin light chains, glideosome associated proteins (GAP40, GAP45, GAP50, GAPM) and parasite adhesins (for example TRAP, thrombospondin-related anonymous protein in sporozoites) (Fig 1.7). The fluid-mosaic model of lipid bilayers allows for explanation of movement of the parasite. This allows for the physical interaction of the parasite adhesin and the host substrate to be maintained whilst the parasite glides. Due to the amount of force required to move the parasite, it has been suggested that the alveolins, through the SPN, provide a solid structure for the complex to allow movement, much like the pivot point of an oar whilst rowing (Baum *et al*., 2008). This hypothesis is consistent with the demonstrated reduction in zoite motility of alveolin knockout parasites (Khater *et al*., 2004, Tremp *et al*., 2008, Tremp and Dessens, 2011).
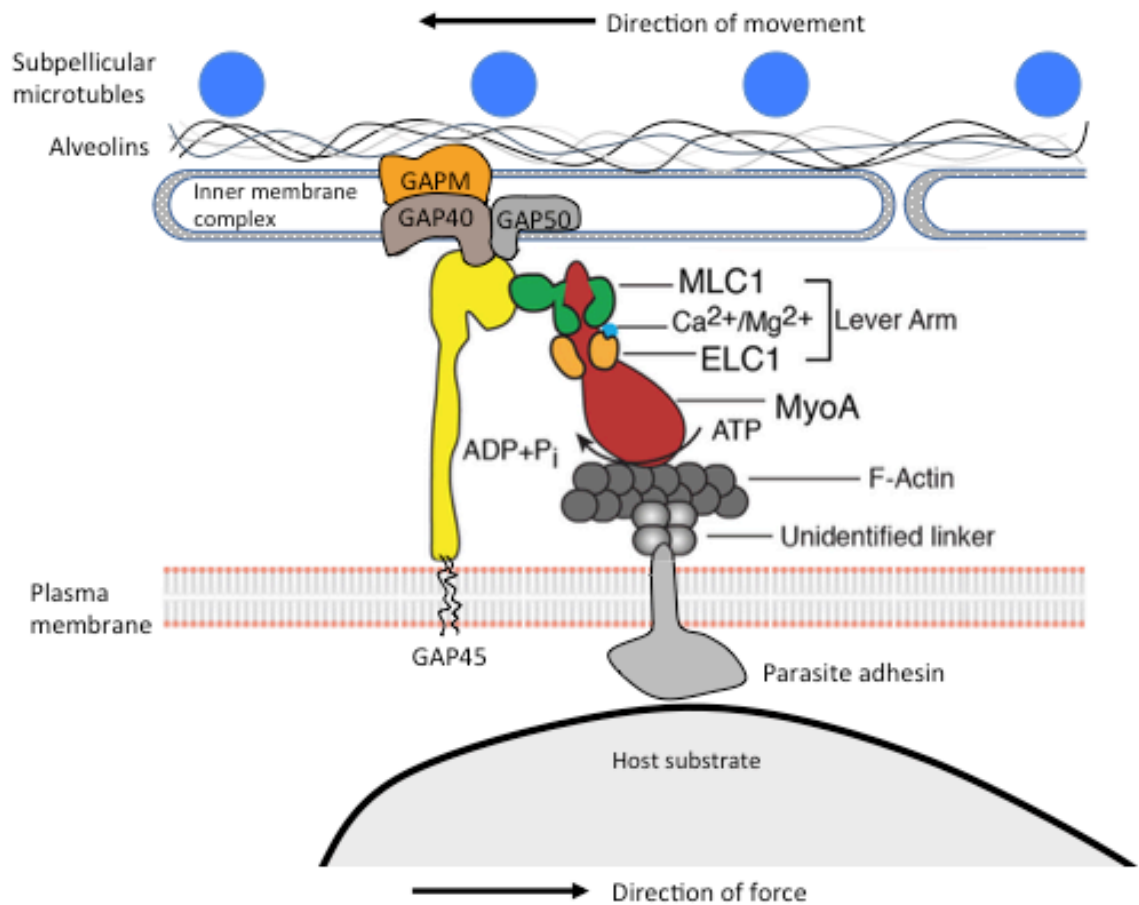
Fig. 1.7 The substrate dependent gliding motility model for Apicomplexan parasite movement. The parasite adhesin is a transmembrane protein that interacts with host substrate, attached through its C-terminal domain to the actomyosin motor domain. The myosin motor domain (MyoA), essential light chain (ELC1) and myosin light chain (MLC) are stabilised by the presence of GAP45, and anchored in place by GAP40 and GAP40. The complex interacts with the rigid alveolin-rich subpellicular network through the GAPM protein. The MyoA protein is secured with ELC1 and MLC1 in the lever-arm domain. MyoA drives ATP-dependent, barbed end directed force, which, through F-actin, propels the secreted parasite adhesin inversely to the direction of movement through the fluidity of the outer plasma membrane. Adapted from (Williams *et al.*, 2015, Baum *et al.*, 2006).

1.1.10 Intermediate filaments

Intermediate filaments are proteins that provide a structural role in the cell, largely giving mechanical strength to cells and tissues. Intermediate filaments are classed as such as they tend to have a diameter of ~10nm, which is intermediate between the diameters of the two other main components of the cytoskeleton, actin (around 7nm) and microtubules (around 25nm) (Herrmann *et al.*, 2009). Another feature of intermediate filaments is that they are amorphous, with no polarity. Intermediate filaments are less dynamic than microtubules or actin filaments. There are six groups

of intermediate filaments currently classified in eukaryotes. Type I and II are keratins, type III include vimentin and desmin. Type IV include neurofilament proteins, type V include nuclear lamins, and type VI includes nestin (Hyder *et al.*, 2011, Herrmann *et al.*, 2009).

Despite there being little homology regarding size and amino acid sequence, both between and across these groups of intermediate filaments, they all share a common secondary structure organisation, most notably a central helical rod domain that can form coiled-coils. Archetypal coiled-coil domains are made up of a heptad-repeat structure that contain an *a-b-c-d-e-f-g* conformation (Fig 1.8). Residues at positions *a* and *d* that are sequentially hydrophobic, allowing for the orientation of the residues to be buried within the structure. Residues at positions *g* and *e* are charged positively and negatively respectively, interacting with each other and forming non-covalent bonds (Hicks *et al.*, 1997). The coiled-coil domains are thought to be important for the filament-forming properties of the intermediate filaments. This offers an interesting parallel with the alveolins, which also possess tandem repeat structures, albeit of a different periodicity.
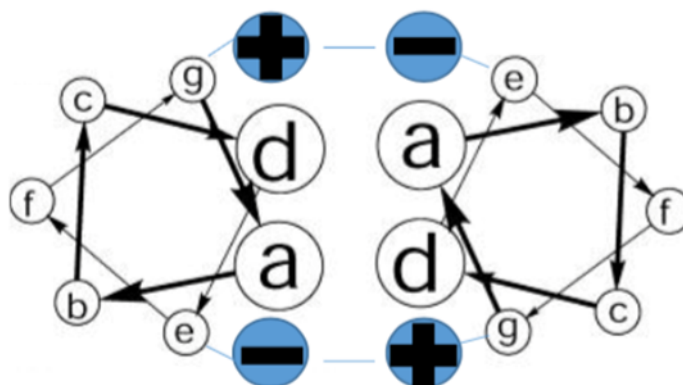
Fig 1.8 The heptad repeat structure. Residues *a* − *g* are indicated with arrows indicating the primary sequence, forming a coiled-coil with two parallel helices. Residue *g* is positively charged, whereas residue e is negatively charged (blue circles), forming an ionic bond (blue middle line).

Filament forming β-sheets are seen within β-keratins of avian feathers, however the structure and sequence of these filaments do not correspond with alveolin primary sequences or the 12 amino acid tandem periodicity (Fraser and Parry, 2008), with further evidence suggesting periodicity between two types of keratin, scale and feather keratins, have structural similarity (Gregg *et al.*, 1984). With the filaments readily produced from natural feather growth, X-ray diffraction patterns were readily produced directly from the feather rachis (Fraser *et al.*, 1971), limiting the methodological approach for characterising these proteins *in vivo.* However, intermediate filaments ability to form simultaneously rigid and flexible filaments remains remarkable. Their ability to form dimers by associating laterally and in an antiparallel fashion allow them to form these stable structures (Robert *et al.*, 2016).

Insoluble by their filament-forming nature, expression and purification of intermediate filament proteins provide a set of challenges that differ to soluble proteins. These insoluble proteins have had their atomic structures resolved through harsh denaturing conditions. They are able to be solubilised by cells through various kinases that reduce them down to their characteristic coiled-coil complexes (Herrmann and Aebi, 2016). Vimentin has been characterised through extensive efforts over a number of years (Nicolet *et al.*, 2010). Though challenging, results were obtained through using multiple fragments of vimentin, between 60-100 residues each, and were expressed in *Esherichia coli* using an N-terminal 6×His tag and a TEV protease cleavage site (Nicolet *et al.*, 2010). Proteins were expressed and induced using common techniques and purified using ion-exchange chromatography and size exclusion chromatography in highly denaturing buffers (Nicolet *et al.*, 2010). These methods provide a basis for structural studies of the alveolins.

1.1.11 Recombinant protein expression systems

Structural biology has provided a significant advancement in understanding the biomolecular structure of various proteins essential for cellular processes. This discipline of biology involves studying anything from the structure and function of the protein, to its localisation or protein interactions. Using cellular protein expression systems, large amounts of a protein of interest can potentially be obtained. Both eukaryotic and prokaryotic protein expression systems are widely used in research today, with each system having various strengths and weaknesses. The selection of each system depends on the type of protein, the requirements of post-translational modifications and other modifications for functional activity, and the desired yield. Bacterial expression systems (*E. coli*) are relatively reliable in producing proteins for structural analysis, and are scalable for producing large yields quickly. It has a low cost and simple culture conditions, allowing for optimal expression. Disadvantages of this system include a lack of post-translational modifications of the protein that would be seen in eukaryotic systems, and there may be a need for protein-specific optimisation steps. There are a number of hurdles to tackle when using bacteria to overexpress a desired protein. As energy and cellular resources are being spent to overexpress a protein the bacteria may not deem beneficial, a number of proteases may be deployed to recycle spent protein resources (Paraskevopoulou and Falcone, 2018). The addition of protein tags such as glutathione-S-transferase (GST) partly protect against such proteolytic degradation. Limiting the amount of proteases innately expressed in bacteria can counteract this. Overexpression can also lead to high protein insolubility, with aggregated protein being produced which is useless for structural studies. The N-utilisation substance protein A (NusA) tag confers stability and increases solubility of its bound protein (De Marco *et al*., 2004, Dummler *et al*., 2005, Turner *et al*., 2005).

Furthermore, protein tags can be used for the purpose of affinity-based protein purification, the best example perhaps being the polyhistidine (His) tag, which is small and unlikely to affect protein folding, and can be used for ion-exchange chromatography on Nickel or Cobalt cation columns. Larger tags such as GST or NusA have to be cleaved off the fusion protein before embarking on structural analysis. For this purpose, specific protease cleavage sites are introduced at the protein-tag junction. The tobacco etch virus (TEV) protease cleavage site (`ENLYFQS`) is commonly used, as it is small and can be efficiently cleaved by TEV protease *in vitro* (cleaving between the glutamine and serine residues).

1.1.12 *Plasmodium berghei* as a research model

The ethics and feasibility of investigating the whole life cycle of human *Plasmodium* species makes it difficult to study life stages that cannot be cultured in the laboratory. *P. berghei* is a closely related species that infects rodents, and has been widely validated as a model for human malaria, allowing for investigation in to all aspects of the life cycle. As they are the same genus, they share many of the same structures and cellular biology. Further similarities between *P. berghei* and *P. falciparum* include their genome organisation, size and synteny. The *P. berghei* genome size 18.5Mb as opposed to *P. falciparum's* 23.3Mb, and both have 14 chromosomes (Otto *et al*., 2014).

It is possible to safely, simply and efficiently manipulate the entire *P. berghei* life cycle with both definitive and intermediate hosts. It is also possible to scale up and study specific life cycle stages due to reliable *in vitro* protocols, particularly for the ookinete. Indeed, *P. berghei* has been widely used as a model for malaria due to the ability to reliably and efficiently cultivate its ookinetes *in vitro,* and historically has allowed for

detailed studies on gametogenesis, fertilisation and zygote development (Rosales-Ronquillo and Silverman, 1974, Weiss and Vanderberg, 1977). As a high number of the *Plasmodium* alveolin repertoire are expressed in the ookinete, this is a key life stage for many investigations throughout this thesis, highlighting the suitability of *P. berghei* as the chosen malaria model. *P. berghei* is easily maintained with one of the human malaria vectors *Anopheles stephensi*, and allows for efficient and stable gene targeting techniques using selectable markers. The easily manipulated genetics of the parasite make it an attractive experimental organism. The ability to transfect *P. berghei* merozoites is advantageous, as transfecting life cycle stages that are still within red blood cells decreases the efficiency of transfecting DNA across multiple membranes to reach the parasite nucleus.

Whilst reminiscent of human malarial infection, it is worth considering that the complexity and spectrum of disease seen in humans is not replicated within mouse models of malaria, due to the number of strategies the parasite has developed to evade host immune response (Craig *et al.*, 2012).

1.2 Aims and objectives

This project aims to determine the core architecture and atomic structure of the alveolins, as well as the underlying mechanisms for their assembly into the cortical cytoskeleton of malaria parasites. The essential nature of the alveolins in malaria parasite development, their expression throughout the life cycle, and their absence in vertebrates makes them potentially attractive drug targets for malaria treatment, prophylaxis and transmission control, for example by inhibiting filament formation. Moreover, such drugs may be active against a broad range of other apicomplexan parasites, as well as against related pathogenic alveolates. In this context, a better understanding of the core architecture of the alveolins, their assembly mechanisms, and domain-specific contributions to alveolin function is vital. The following hypothesis-driven questions will be addressed:

1: Do the conserved alveolin domains have a crucial role within the proteins innate functions in motility, tensile strength, and recruitment to the SPN? This will be addressed through structure-function studies using transgenic parasite lines expressing modified forms of the alveolins IMC1b and IMC1h.

2: Do alveolins form α-helical coiled-coil domains similar to the coiled-coil regions of metazoan intermediate filament proteins that are required for IF formation? This will be addressed using recombinant expression in *E. coli*, followed by purification and downstream structural analyses, of conserved domains from IMC1e, IMC1g and IMC1c.

Chapter 2

Materials and Methods

This chapter describes the general methods used throughout these results chapters. Specific protocols related to particular chapters are described in each one separately. Chemicals used were purchased from Sigma Aldrich unless otherwise specified.

2.1. In-Fusion cloning

Rather than using traditional restriction enzyme digestion combined with T4 DNA ligase to introduce PCR fragments into plasmids, In-Fusion cloning was used (In-Fusion Kit, Clontech/Takara Bio). To achieve In-Fusion, forward and reverse PCR primers were designed to contain at least 15 bases of homology with the sequences flanking the targeted insertion site in the cloning plasmid. Using these primers, amplification of the DNA was performed, and purified PCR product was introduced into a linearized cloning vector in the In-Fusion cloning reaction. The In-Fusion enzyme generates single-stranded regions of homology and then fuses the DNA insert to the vector. The added advantage of In-Fusion is that cloning is directional.

Plasmid vectors were linearized by utilising specific restriction endonuclease enzymes. The QIAquick gel extraction kit (Qiagen) was used to clean up amplified PCR product and linearized plasmids. To prepare the In-Fusion cloning reaction, the linearized plasmid and insert were typically mixed together at an approximately 1:1 molar ratio in a 10µL reaction that contains 1µL In-Fusion enzyme and its appropriate buffer. Reactions were carried out at 37°C for 15min, followed by 15min at 50°C.

2.2. Site-directed mutagenesis

In order to create mutations or deletions at specific sites in the plasmids, a PCR-based site-directed mutagenesis method was used. The number of PCR cycles was kept low (typically 5) to minimize the undesired expansion of potential errors generated by the

PCR. To compensate, a higher starting amount of template plasmid DNA was used (typically 1µg). The sense and antisense primers were designed to anneal to complementary strands of the same circular template plasmid DNA, to enable the targeted modification. Additionally, the first 16 nucleotides of each primer were each other's reverse complement, to permit the re-circularization of the linear PCR product by In-Fusion. The PCR reaction was performed using Advantage HD polymerase mix (Clontech/Takara Bio) with ~1µg of DNA template, 10µM of each primer and 2.5mM of each dNTP. Following PCR, the reaction was subjected to *Dpn*I digestion overnight at 37°C to digest the methylated template DNA. The linearized PCR product was then purified using the QIAquick gel extraction kit (Qiagen) and circularized by In-Fusion.

2.3. Bacterial transformation and selection

Bacterial transformation was carried out with NEB 5-alpha competent cells (New England Biolabs). Briefly, frozen competent cells were thawed on ice followed by incubation with the DNA on ice for 30min. Cells were heat-shocked for 45s at 42°C and placed on ice for 1min. To allow expression of the antibiotic resistance, cells were resuspended in 450µL super optimal broth with glucose (SOC) medium and incubated at 37°C for 1h with shaking. The cell culture was then plated onto Luria Broth (LB) agar plates supplemented with 100µg/mL ampicillin. Colonies were picked and grown in 5mL LB culture containing 100µg/mL ampicillin and incubated at 37°C overnight. Plasmid DNA was extracted and purified from the overnight culture using the Wizard Plus SV Minipreps DNA purification system (Promega) and analysed by diagnostic restriction enzyme digestion.

2.4. Parasite maintenance

P. berghei ANKA clone 234 parasites were maintained as cryopreserved stabilates or by mechanical blood passage and mosquito transmission. For cryopreservation parasitized blood was mixed with an equal volume of cryopreservation solution (20% dimethyl sulfoxide (DMSO), 10% fetal bovine serum (FBS), 70% RPMI 1640) and slowly frozen to -80°C before transfer to liquid nitrogen for long-term storage. As rodent hosts, CD1 mice (typically 20-25g, female) were used. Animals were infected by intraperitoneal (i.p.) injection of fresh or cryopreserved parasitized blood, or by sporozoite-infected mosquito bites. The levels of infection were monitored and observed by light microscopic examination of Giemsa-stained thin blood films.

2.5. Mosquito maintenance and infection

A. stephensi mosquitoes were maintained in cages at ~27°C, at high relative humidity, under a 12h light/dark cycle and fed on 10% glucose solution. Adult insects were fed on human blood in membrane feeders (kept at 37°C) twice a week to induce egg laying. Naïve mosquitoes were infected with *P. berghei* by feeding on anaesthetized, gametocytaemic mice. For optimal development of the parasite, experimental mosquitoes were kept at 19-21°C at high relative humidity and with 10% glucose for feeding. Transmission of sporozoites (biteback) was carried out with sporozoite-infected mosquitoes (about three weeks post-infection) and naïve mice.

2.6. Ookinete culture and purification

To induce reticulocytosis, mice were treated with phenylhydrazine (~10µL/g body weight of a 6mg/mL solution in PBS) followed by injection of 107–108 parasites three days later. Three days following infection, parasitized blood was harvested and mixed with 10 volumes of ookinete medium (RPMI 1640 containing 25mM Hepes and L-

glutamine, 2g/L sodium bicarbonate, 1/100 volume of 10,000u/mL penicillin/10mg/mL streptomycin solution (Gibco), 50mg/L hypoxanthine, 100mM xanthurenic acid and 20% FBS, pH7.4). Cultures were incubated at 19-21°C for 18-24h to allow ookinete development.

For ookinete purification, cells were collected by centrifugation at 800xg for 5min at 4°C, the pellet resuspended in ice-cold 0.17M ammonium chloride and incubated on ice for 30min to lyse the unparasitized RBCs. Parasites were collected by centrifugation at 800xg for 10min at 4°c with low deceleration and washed twice in PBS (500xg centrifugation). The number of ookinetes was estimated by haemocytometer count.

2.7. Generation of genetically modified parasite lines

For the generation of transgenic *P. berghei* parasite lines, purified schizonts were transfected followed by drug selection and limiting dilution cloning, according to described methods (Janse *et al.*, 2006). To achieve double homologous crossover recombination, targeting vectors were digested with specific enzymes to remove the backbone of the vector and column purified.

Blood was harvested from *P. berghei*-infected mice with parasitaemia between 1% and 4%, and added to 100mL filter-sterilized schizont culture medium (RPMI 1640 containing 20mM Hepes and L-glutamine (Sigma R73880), 0.8g/L sodium bicarbonate, 1/200 volume of neomycin solution (Sigma N1142) and 20% FBS, pH7.2). The culture was equilibrated with a 5% CO_2, 10% O_2, 85% N_2 gas mixture (BOC) and the flask sealed airtight, followed by incubation at 36°C with gentle shaking (60–80 revolutions per minute) for 20h. Schizonts were purified by centrifugation for 30min at 500xg at room temperature (no brake) on a 50% Histodenz (in PBS) cushion (100% Histodenz contains 27.6% [w/v] Histodenz powder in 5mM Tris-HCl pH 7.2, 3mM KCl, 0.3mM

EDTA). The interface layer containing the schizonts was harvested and centrifuged for 8min at 450x*g*. The schizont pellet was gently resuspended in complete nucleofector solution (Basic Parasite Nucleofector kit, Amaxa), and for each transfection 100μL was mixed with DNA (1-5μg in 10-12μL dH2O or TE buffer). Electroporation was carried out with programme U-033 of the NucleofectorTM II Device (Amaxa). Directly post electroporation, approximately 250μL of naïve blood (previously incubated at 37°C) was added to the parasites, and the mixture was kept at 37°C for 30min to permit merozoite invasion of naive RBCs. Finally, the parasite/RBC mixture was injected i.p. into naïve mice.

One day after transfection, pyrimethamine selection was started (supplied in drinking water at a dose equal to 10mg/kg/day) until patent parasitaemia was observed (usually 7-10 days post-inoculation). When parasitaemia reached around 1%, limiting dilution cloning was carried out. Parasites were diluted in RPMI to an estimated 0.3 parasites/inoculum, and typically 10 naive animals were inoculated by i.p. injection. Infected mice were identified one week later by Giemsa staining of thin blood films. Clonal/isogenic parasite populations were checked by diagnostic PCR for correct integration of the transgene in the target locus, and for the absence of the unmodified target allele.

2.8. Western blot analysis

Protein samples were run through pre-cast NuPAGE Bis-Tris SDS gels (Invitrogen) using an XCell SureLock Mini-Cell system (Invitrogen). Before loading, samples were heated at 70°C for 10min in the presence or absence of 1% reducing agent (Invitrogen). The fractionated protein was then transferred onto a polyvinylidene fluoride (PVDF) membrane by electroblotting in an XCell II Blot Module (Invitrogen). Membranes were

blocked for 1h at room temperature in PBS supplemented with 0.1% Tween 20 and 5% skimmed milk. Membranes were incubated either overnight at 4°C or for 1h at room temperature with specific antibody at the appropriate dilution (outlined in table 2.1). Typically horse radish peroxidase (HRP) conjugated antibodies were used to visualise the signal by chemiluminescence, using ECL Western Blotting Substrate (Pierce) and X-ray film (CL-XPosure™ Film, Pierce).

Table 2.1. *Antibodies and concentrations used for western blot analysis.*

Primary goat polyclonal antibody to GFP conjugated to horseradish peroxidase (Abcam ab6663)	1:5000
Anti-Pbs21 monoclonal antibody (13.1, Tirawanchai and Sinden, 1990)	1:5000
Horse radish peroxidase conjugated goat anti-mouse IgG (Sigma A4461)	1:5000

2.9. Light microscopy

Standard light microscopic analysis of parasite and mosquito samples, Giemsa-stained blood films and haemocytometer counts were undertaken using an Olympus CX41 microscope. An Olympus SZ microscope was used for mosquito dissections. For assessment of fluorescence, live or fixed parasite samples were assessed, and images captured, on a Zeiss LSM510 inverted confocal microscope, or on a Zeiss Axioplan-2 fluorescent microscope with Retiga 2000R CCD camera system and Volocity software.

2.10 Sporozoite size measurements

Images of individual midgut sporozoites were captured by microscopy on Zeiss LSM510 inverted laser scanning confocal microscope. Using the Zeiss LSM image browser software, the circumference was measured of the sporozoite was measured. This gave a value equivalent to the occupied surface area, and was calculated as a measure of cell size. Statistical analysis was carried out using Student's t-test.

2.11 Osmotic shock and viability assay

Unpurified ookinetes were subjected to a hypo-osmotic shock of 0.5 × normal osmotic strength through adding an equal volume of water. Sporozoites were released from oocyst-infected midguts at 15 days post-infection in a Dounce homogeniser, and were subjected to 0.33 × normal osmotic strength by adding two equal volumes of water. After 5 minutes, normal osmotic conditions were restored by adding an appropriate amount of 10 × PBS.

Cell viability was recorded using fluorescence microscopy in the presence of 0.5% propidium iodide and 1% Hoechst 33258. Ookinetes whose nuclei stained positive for both propidium iodide and Hoechst were scored as non-viable, whereas ookinetes whose nuclei stained positive only for Hoechst were scored as viable. Values were normalized to 100% viability in untreated cells.

2.12 Motility assay

Aliquots of unpurified ookinete cultures were added to equal volumes of Matrigel (BD Biosciences) on ice, mixed thoroughly, placed onto a microscope slide, and covered with a cover slip. After sealing with nail varnish, the Matrigel was allowed to set at room temperature for 30 min before analysis. Sporozoites were gently released from around 20 salivary glands in 200 μL RPMI medium in a Dounce homogenizer on ice. Following addition of an equal volume of RPMI supplemented with 20% foetal bovine serum, the sporozoites were transferred to an Eppendorf tube and collected by centrifugation in a swing-out rotor for 10 min at 1,000 × g at 4°C, followed by removal of excess supernatant.

Aliquots of resuspended sporozoites were mixed with an equal volume of Matrigel on ice, before transfer to microscope slides. After sealing the cover slip with nail varnish,

the Matrigel was allowed to set at room temperature for 30 min before analysis. Cells were examined and time-lapse images taken on a Zeiss Axioplan II microscope. Movies were analysed with ImageJ using the Manual Tracking plugin. Statistical analysis was carried out using ANOVA and Tukey's multiple comparison.

2.13 Bioinformatics tools

The program PrDOS (Ishida and Kinoshita, 2007) was used to look at the disorder probability of proteins. The program Quick2D (MPI Bioinformatics Toolkit, https://toolkit.tuebingen.mpg.de) was used to make secondary structure predictions.

2.14 Protein expression

E. coli strain BL21*(DE3) was used for protein expression. This strain is suitable for protein expression as it is deficient in Lon and OmpT proteases. BL21*(DE3) cells were co-transformed with pRARE$_2$ plasmid that provides rare amino acids to enhance protein synthesis, which was selected for with 100μg/mL chloramphenicol. When transformed with a plasmid containing a T7 inducible promotor region, protein expression can be induced through the addition of IPTG.

BL21 cells were co-transformed with the pETM6T1 and pRARE$_2$ constructs of choice and selected with 100μL/mL kanamycin and chloramphenicol. Typically, transformed bacteria were grown in a starter culture with 200mL sterile LB broth with 100μL/mL kanamycin and chloramphenicol, and 10mL of the starter culture was added to separately prepared sterile 1L LB flasks. Flasks were shaken at 200rpm at 37°C for ~4 hours, until an OD$_{600}$ of 0.6 - 0.8 was reached. This OD$_{600}$ indicates that the bacteria are entering log phase of growth, which has been shown to be the optimal time to induce for expression with 100μM IPTG. Cells were pelleted at 4,000rpm for 1 hour. Pellets

were resuspended in 10mL of cold lysis buffer (150mM NaCl, 50mM Tris-HCl, pH8) and stored at -80°C.

2.15 Protein purification

To release protein from bacterial cells, pellets were defrosted in a 20% ethanol-ice bath. Once defrosted, the cells were sonicated at 100% amplitude in 2 second pulses for 2 minutes on ice to ensure heat dissipation. This was repeated 3 times before cells were further lysed using an Emulsiflex C3 with a pressure beating around 10,000 psi. Both processes ensured that bacterial cells were sufficiently lysed to release expressed protein. The lysate was ultracentrifuged for 1 hour at 20,000rpm to separate the soluble and insoluble fractions. The insoluble fraction was discarded and the soluble fraction was added to an equilibrated batch column containing 2mL of TALON metal affinity resin, and was incubated at 4°C rolling for at least 1 hour. This immobilizes the His-tagged NusA protein. The column was rinsed 5 times with wash buffer A (150mM NaCl, 50mM Tris-HCl, pH8) to remove unbound proteins. The batch-column was then rinsed with wash buffer B (150mM NaCl, 50mM Tris-HCl pH8, 20mM imidazole) to elute any non-specific protein that may have bound due to a negative charge. Protein was eluted off the column with 10mL elution buffer (150mM NaCl, 50mM Tris-HCl pH8, 500mM imidazole). The eluate was incubated overnight at 4°C with TEV protease at a final concentration of 0.1mg/mL, whilst being dialysed against wash buffer A to remove imidazole. The protein-TEV solution was then added to a HiTrap Q Fast Flow anion exchange column (GE Healthcare), equilibrated with 5 column volumes of buffer A (50mM Tris-HCl pH8), 5 column volumes of charge buffer (1M NaCl, 50mM Tris-HCl pH8) and then 10 column volumes of buffer A. The pH of the buffer was set at 8, because the isoelectric point (pI) for the alveolins is around pH7, whilst that of NusA is

around pH5. Accordingly, at pH8 the NusA and alveolin portions of the fusion protein were expected to behave very differently on an ion exchange column facilitating their separation. On occasion ion exchange was used in conjunction with a HisTrap Fast Flow column (GE Healthcare)

If desired, proteins were purified and concentrated by size exclusion chromatography, using a Superdex 75 (s75) column (GE Healthcare). Protein was loaded on to the column using an ÄKTA device. The s75 column is efficient in separating proteins with molecular weights between M_r 3,000 and 70,000.

Chapter 3

Structure-function analysis of the alveolin IMC1b

3.1 Introduction

Alveolins are grouped by virtue of possessing conserved domains made up of tandem repeat sequences of a 12 amino acid periodicity (Al-Khattaf *et al.*, 2015). The presence of such modules in all alveolin proteins suggests that they carry the potential filament-forming properties of these molecules. Besides these shared modules, a subset of alveolins with demonstrated roles in motility (i.e. IMC1a, IMC1b and IMC1h) possess additional conserved domains that are structurally unrelated to the 'alveolin' module or any other known protein domains. The roles of the latter in protein function remain poorly understood. To address this question, and to gain a better understanding of the contribution of different conserved domains in alveolin function more generally, a series of structure-function analyses of the ookinete-expressed alveolin IMC1b were carried out.

IMC1b contains a type 1 alveolin domain (ALV1), a type 2 alveolin domain (ALV2) and a third, structurally unrelated conserved domain (ALV3) (Fig. 3.2A). IMC1b has a close structural paralogue named IMC1a, which is expressed in sporozoites where it carries out functionally equivalent roles (Khater *et al.*, 2004). Because IMC1b is dispensable for blood stage parasite development, its coding sequence can be successfully disrupted (Tremp *et al.*, 2008). This revealed that IMC1b null mutant (IMC1b-KO) parasites exhibit clear and measurable phenotypes with regards to ookinete morphology, tensile strength and motility (Tremp *et al.*, 2008). For these reasons, IMC1b constituted a suitable alveolin for assessing the roles of different alveolin modules by structure-function analysis in live parasites using allelic replacement via homologous recombination.

3.2 Materials & Methods

3.2.1 Plasmid constructs

Plasmid pLP-IMC1b/EGFP was previously constructed for the allelic replacement of *imc1b* with a full-length GFP-tagged version via double crossover homologous recombination (Tremp *et al.*, 2008) (Fig 3.1). This plasmid contains one *Hind*III site near the start of the coding sequence (at amino acid position 30) and a second *Hind*III site at the IMC1b/GFP junction. Recombinant alveolin sequences were PCR amplified from *P. berghei* genomic DNA with specific primers (Table 3.1). PCR conditions were optimised to reflect the primer sequences, with an annealing temperature of 55°C and an extension time of 2 minutes (1 min per kb) for 25 cycles. PCR products encoding the various IMC1b- and IMC1a-specific parts were introduced by In-Fusion into *Hind*III-digested pLP-IMC1b/EGFP (Fig. 3.2A) giving rise to plasmids pLP-IMC1b/ALV1, pLP-IMC1b/ALV2, pLP-IMC1b/ALV3, pLP-IMC1b/ALV1+2, pLP-IMC1b/ALV1-3 and pLP-IMC1b/IMC1a.
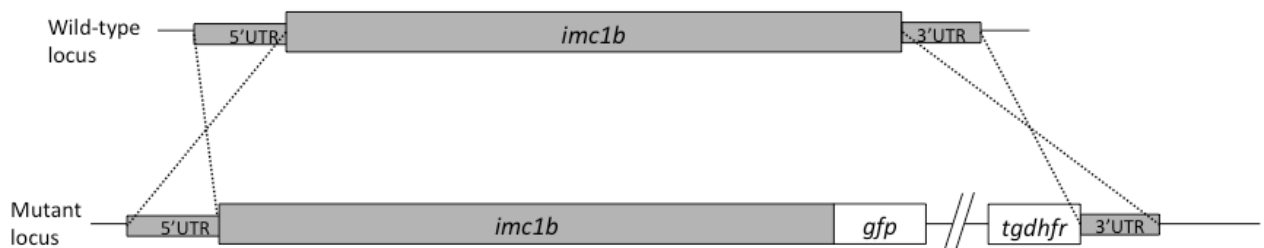
Fig 3.1 The double homologous recombination strategy. The wildtype locus containing the innate 5'UTR, *imc1b* and 3'UTR has a homologous sequence between the introduced mutant locus, that has attached to it a *gfp* reporter and *tgdhfr*, a drug selectable marker. Upon DNA replication, a double strand break is introduced that allows for the homologous flanking regions to doubly recombine, integrating the mutant locus in to the parasites genome.

Plasmid pLP-IMC1b/EGFP was used as the DNA template for a PCR-based site directed mutagenesis to remove sequences encoding specific domains of IMC1b, using specific primer pairs (Table 3.2). Following PCR, template plasmid DNA was digested with *Dpn*I

and the PCR fragment circularized by In-Fusion, to give plasmids pLP-IMC1bΔ1, pLP-IMC1bΔ2, pLP-IMC1bΔ3 and pLP-IMC1bΔ1+2.

Table 3.1. *PCR primer sequences used to amplify IMC1b- and IMC1a-specific sequences*

Region Amplified	Forward Primer	Reverse Primer
IMC1b ALV1	ATGATCAATTAAGCTCAACAAGAATAATTG ACGTTCCAG	GAGGGCCCCTAAGCTATTTCCATC TTTATCAACATAATTATAAAC
IMC1b ALV2	ATGATCAATTAAGCTCACTTTTAGAGCCTT TTGGCCC	GAGGGCCCCTAAGCTCTTTTTTAA TTTATCTTTTTCTTCTTTTAAA
IMC1b ALV3	ATGATCAATTAAGCTCAAAATATAGTCCTA ACAATTTTTA TGATACTATACC	GAGGGCCCCTAAGCTCTCTACTTT GGCTTCACGTGG
IMC1b ALV1-ALV2	ATGATCAATTAAGCTCAACAAGAATAATTG ACGTTCCAG	GAGGGCCCCTAAGCTCTTTTTTAA TTTATCTTTTTCTTCTTTTAAA
IMC1b ALV1- ALV3	ATGATCAATTAAGCTCAACAAGAATAATTG ACGTTCCAG	GAGGGCCCCTAAGCTCTCTACTTT GGCTTCACGTGG
IMC1a	ATGATCAATTAAGCTCAATGTTTGATCCAT GTAAAATAAATAG	GAGGGCCCCTAAGCTCTTATCTTG ATTACAAAAATAATTACAACA TTTG

Table 3.2. *PCR primer sequences used in site-directed mutagenesis to delete IMC1b-specific sequences.*

Region Deleted	Forward Primer	Reverse Primer
IMC1b ALV1	TTAAAAATGGAAATAAAATTAAAGACCA CAACG	TTATTTCCATTTTTAACAATATAAGT TGGTACTTCCTTTATT
IMC1b ALV2	TACCTCCAAACACACAAAGTGATAGTAT GCTAGATCA	TGTGTGTTTGGAGGTAGCATATTAAT ATTTACATCATATG
IMC1b ALV3	TATCTAATATGAAAACCATTTCTCAATT GAAA	GTTTTCATATTAGATAAATCATGTGA AATATTTAGTTCGG
IMC1b ALV1+ALV2	TTAAAAATAACACACAAAGTGATAGTAT GCTAGATCA	TGTGTGTTATTTTTAACAATATAAGT TGGTACTTCCTTTATT
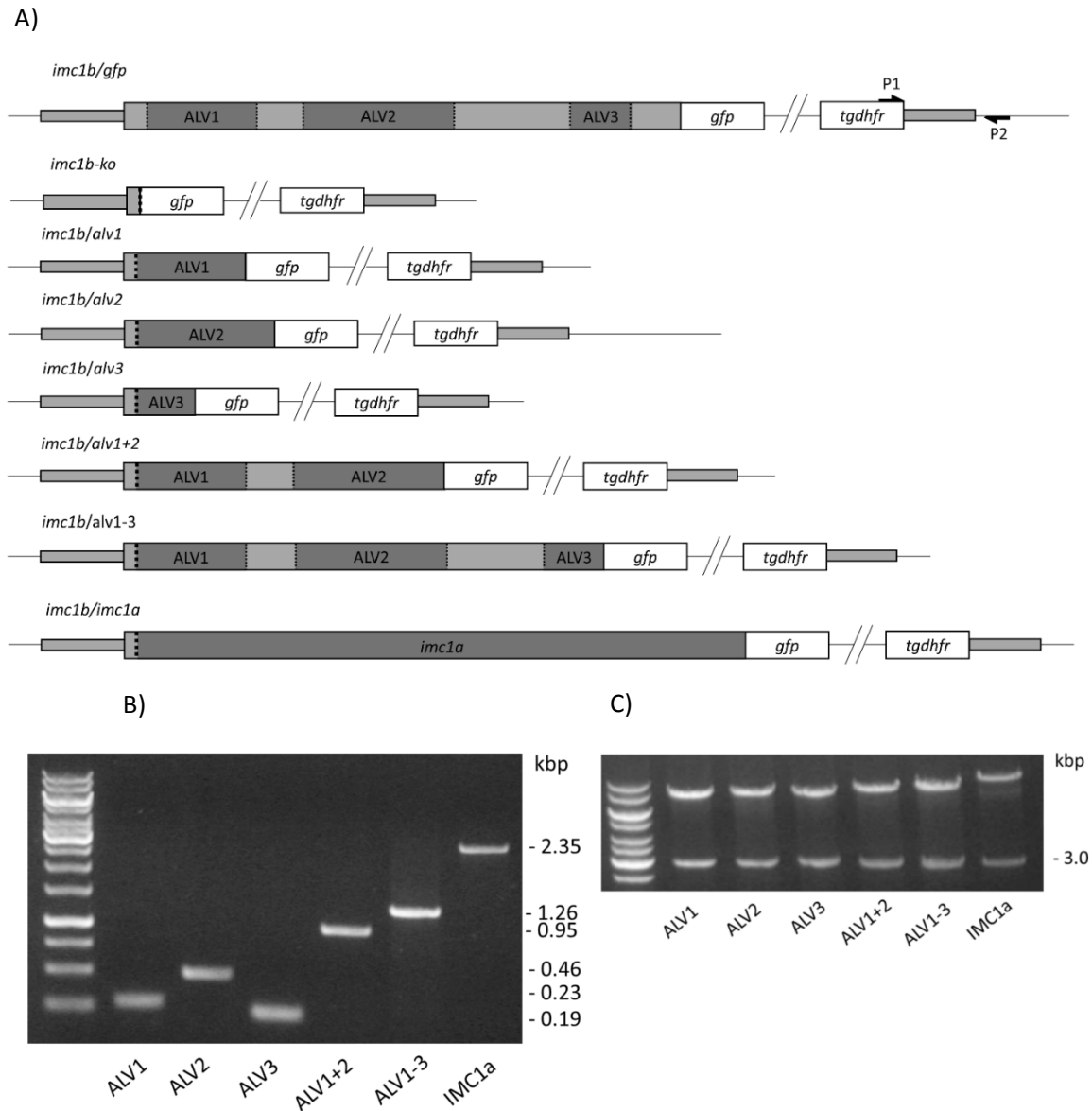
Fig. 3.2. DNA constructs for generating IMC1b/IMC1a knock-in parasite lines. A) Predicted structure of the modified imc1b alleles in the genetically modified parasite lines after transfection with various DNA constructs following 'knock-in' strategy 1. The two alleles at the top depict those in the previously generated parasite lines IMC1b/EGFP and IMC1b-KO (Tremp *et al.*, 2008). The relative positions of primers P1 and P2 are shown for amplification of a 1.2kb DNA fragment across the 3'-integration site. Sizes of domains are not to scale. B) PCR fragments corresponding to the inserts. Sizes are shown on the right hand side. C) Double digestion of plasmids with *Kpn*I and *Sac*II drops out an approximately 3kb fragment corresponding to the plasmid backbone.

3.2.2 Transfection and genotyping

Prior to transfection, plasmids were doubly digested with *Kpn*I and *Sac*II to remove the

plasmid backbone (Fig 3.2C), allowing targeting into the native *imc1b* locus by double

crossover homologous recombination. Correct integration of the recombinant alleles

44

and selectable marker genes into the *imc1b* locus was carried out by diagnostic PCR of genomic DNA, with primer pair P1 (TgDHFRCAS3') (TCGTGGGCTACGTCCCGCAC) and P2 (IMC1b3'diag) (TGGTTATATTTTCATTTTGAATTAATAATATATG), or P3 (TgDHFR-F) (AGAGGGGCATCGGCATCAAC) and P2, designed to PCR amplify across the 3'-integration site. Diagnostic PCR across the 5'-integration site was achieved by diagnostic PCR with primer pair P4 (IMC1b5'diag) (CGCCGTTCTAAATCAACTGA) and P5 (EGFP-R) (GTGCCCATTAACCATCACC), or P4 and P6 (IMC1b-delta1-R) (CGTTGTGGTCTTTAATTTTATTTCC).

3.3 Results

The first strategy for the structure-function analysis of IMC1b was to generate a series of genetically modified *P. berghei* parasite lines in which specific IMC1b domains C-terminally tagged with GFP, as well as the entire alveolin IMC1a tagged with GFP, were introduced into the native *imc1b* locus by allelic replacement. This 'knock-in' strategy effectively replaces the native IMC1b protein with the recombinant GFP-tagged IMC1b domain (or the alveolin IMC1a), allowing assessment of their ability to assemble into the ookinete's pellicle (looking at cortical GFP fluorescence), as well as their ability to rescue the IMC1b knockout phenotypes (looking at cell shape, motility, tensile strength and infectivity) in live parasites.

PCR amplification of the various parts of *imc1b* and *imc1a* with the specific primer pairs (Table 3.1) gave rise to DNA fragments of the expected sizes (Fig. 3.2B). Successful In-Fusion cloning of these fragments into *Hind*III-digested pLP-IMC1b/EGFP was verified by double digestion with *Kpn*I and *Sac*II, removing the 3kb plasmid backbone (Fig. 3.2C).

Pyrimethamine resistant parasite populations were recovered 7-10 days post-transfection, indicating that successful integration had occurred. This would result in a series of parasite lines expressing recombinant *imc1b* loci as depicted in Fig. 3.2A. Successful integration was corroborated by diagnostic PCR across the 3' integration site with primers P1 and P2 (Fig. 3.2A), which amplified a 1.2kb fragment of the expected size in all transfections (Fig. 3.3A). Following this, ookinete cultures were set up from the uncloned parasite populations to determine if GFP-positive ookinetes were present. Green fluorescence was observed in five out of the six transfections (Fig. 3.3B). Because the transfection knocking in the *imc1a* sequence showed no green

46

fluorescence, it was decided to sequence the corresponding DNA construct. This revealed the presence of a frame shift in a stretch of adenine repeats within the open reading frame, explaining why no GFP fluorescence was observed.

Given the positive assays for integration in the transfections, parasites were subjected to a second round of drug selection to further enrich for parasites carrying the integrated drug resistance gene, followed by limiting dilution cloning to isolate isogenic populations of the transgenic parasite lines. Clonal parasite populations were then harvested and tested again for integration using the same diagnostic PCR. Surprisingly, clones of all the transfections tested negative for integration, suggesting that they constituted parental parasites. Subsequent diagnostic PCR of the second round, uncloned parasite populations indicated integration of the selectable marker had diminished or was lost entirely (Fig. 3.3C), pointing to a loss of parasites with integrated resistance and explaining why dilution cloning only yielded parental parasites. This trend continued after a third round of selection (data not shown). Similar outcomes were obtained after a second and third set of independent transfections followed by drug selection and dilution cloning (data not shown). This was entirely unexpected as sequential rounds of drug selection should enrich the parasite population for the desired, stably integrated transgenes and not the wildtype population.
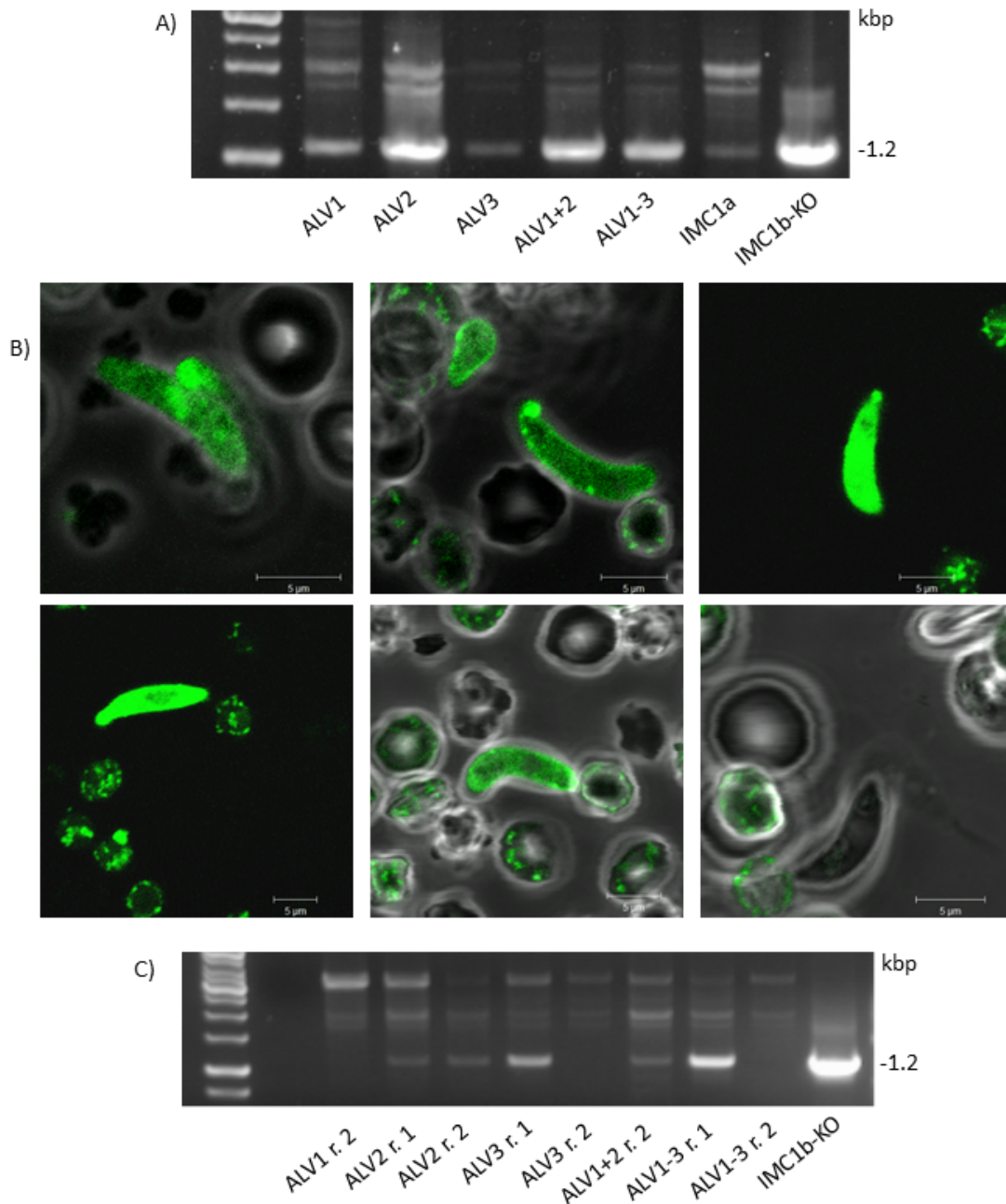
Fig. 3.3. Genotyping of strategy 1 transfections. A) Diagnostic PCR across the 3'-integration site from parasite genomic DNA after first round of drug selection. A fragment of 1.2kb diagnostic for integration is amplified in all transfections. Genomic DNA from clonal parasite line IMC1b-KO (Tremp *et al*. 2008) was used as a positive control. B) Confocal microscopy of ookinetes cultured from uncloned parasite populations after 1 round of drug selection. Green fluorescent ookinetes are present in all but one transfection. C) Diagnostic PCR across the 3' integration site from parasite genomic DNA after second round of drug selection. The 1.2kb fragment diagnostic for integration is lost or reduced.

The problems with the first approach led to a different strategy, in which the function of IMC1b lacking specific domains was assessed, rather than assessing the function of specific IMC1b domains as in approach 1. Using this 'domain knock-out' approach, the contribution of individual modules can be assessed by looking at what the alveolin

48

protein can no longer do when these modules are lacking. This strategy had been successfully used in the Dessens laboratory to study the contribution of certain conserved domains of the ookinete-expressed crystalloid proteins LAP1 and LAP3 (Carter *et al*., 2008, Tremp *et al*., 2017, Saeed *et al*., 2015). Whilst this method used essentially the same allelic replacement strategy as the knock-in strategy, the sequences available for crossover recombination were longer, increasing the chances of achieving stable integration.

Successful site-directed mutagenesis of the parental plasmid to generate the four 'domain knockout' constructs was verified by diagnostic *Hind*III digest (Fig. 3.4A). *Hind*III drops out the entire *imc1b* coding sequence (bar the first 30 amino acids) and this fragment is reduced in size by the corresponding domain deletions in the four different constructs (Fig. 3.4A). Selected plasmids were successfully doubly-digested with *Kpn*I/*Sac*II to remove the plasmid backbone, followed by transfection into purified schizonts. This would result in a series of parasite lines expressing recombinant *imc1b* loci as depicted in Fig. 3.4B.
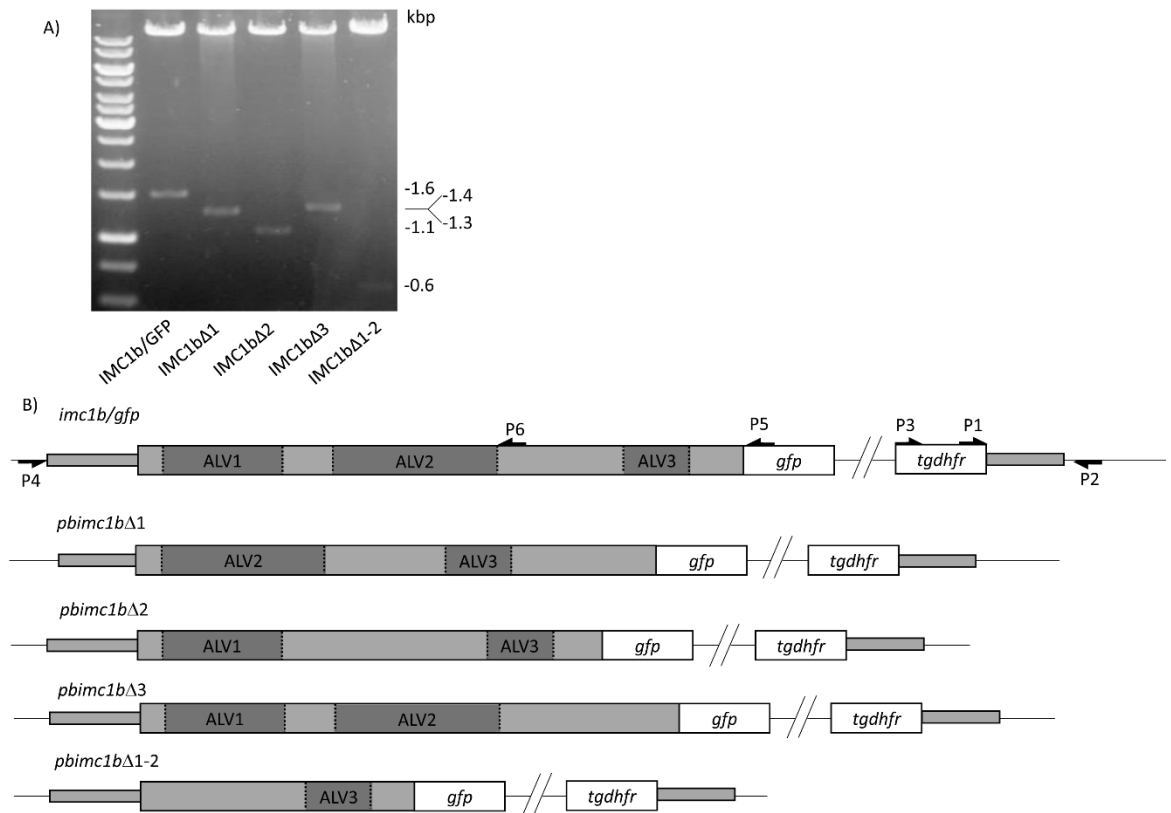
Fig. 3.4. Strategy 2 predicted modified alleles. A) Diagnostic *Hind*III digest of plasmids pLP-IMC1b/EGFP, pLP-IMC1bΔ1, pLP-IMC1bΔ2, pLP-IMC1bΔ3 and pLP-IMC1bΔ1+2. Note the sizes of the *Hind*III fragment corresponding to the IMC1b coding sequence differ with respect to the introduced domain deletions. DNA fragment sizes are indicated on the right hand side. B) Predicted structure of the modified *imc1b* alleles in the genetically modified parasite lines after transfection with various DNA constructs following 'domain deletion' strategy 2. The allele at the top depicts that in the previously generated parasite line IMC1b/EGFP (Tremp *et al*., 2008). The relative positions of primers P1-P6 are indicated. Sizes of domains are not to scale.

A first series of transfections resulted in drug-resistant parasite populations for IMC1bΔ1 and IMC1bΔ3 only (2 out of 4). Diagnostic PCR across the 3'-integration site with primers P3 and P2 (Fig. 3.4B) gave rise to a 2.9kb fragment in the IMC1bΔ1 transfection (Fig. 3.5A), indicating that integration had occurred. This was supported by diagnostic PCR across the 5'-integration site with primers P4 and P5 (Fig. 3.4B) resulting in a product of 2.2kb in the control parasite (IMC1b/GFP), and a somewhat smaller product in the IMC1bΔ1 transfection(Fig. 3.5A), indicating that the sequence encoding domain 1 was indeed lacking. However, following dilution cloning, all the clones obtained (7 in 7) possessed a full-length *imc1b::gfp* allele, as is present in parasite line IMC1b/GFP (Fig. 3.5B). Further scrutiny of the IMC1bΔ1 transfection with

a reverse primer closer to domain 1 (primer P6, Fig. 3.4B) indicated that a mixture of parasites were in fact present, the majority of which possessed the full-length *imc1b* coding sequence (Fig. 3.5C).
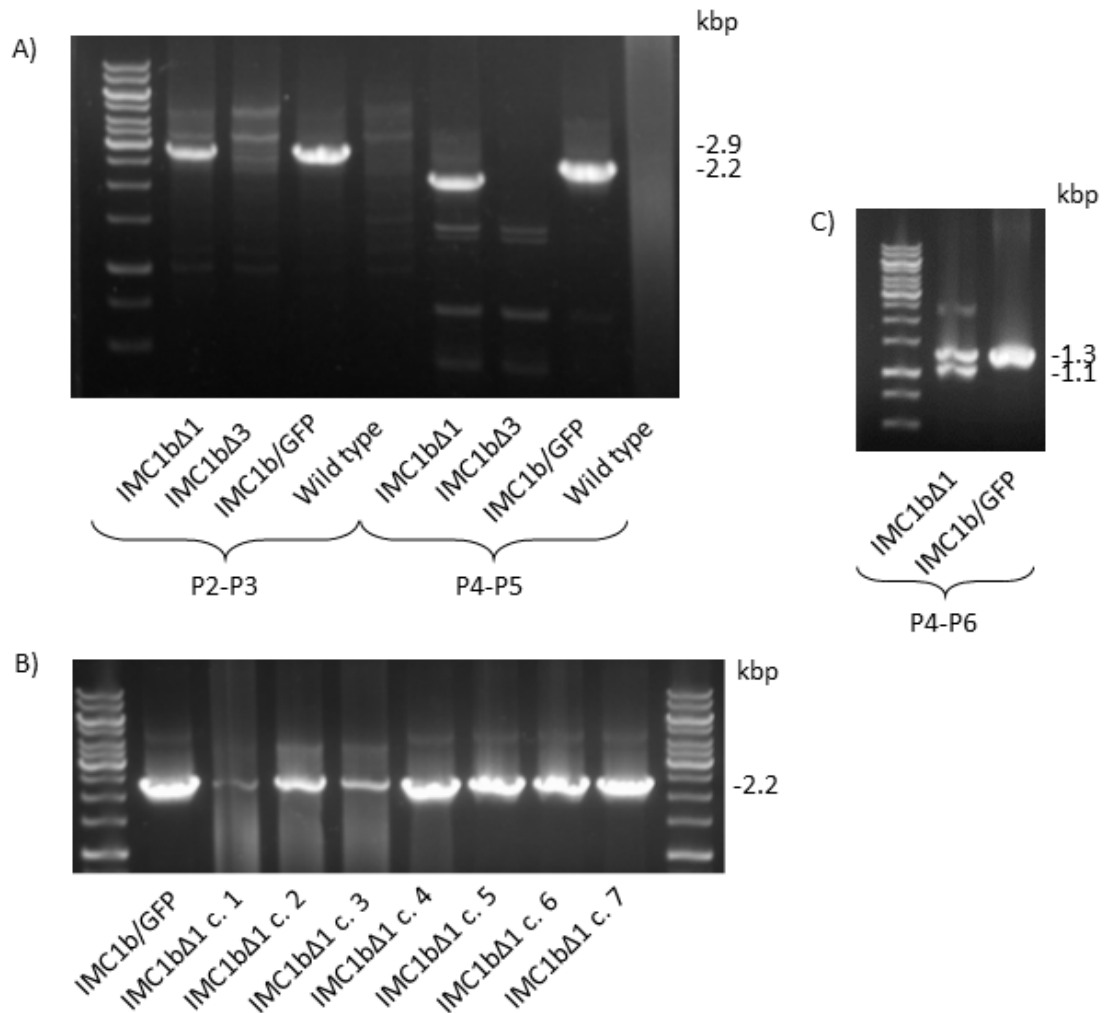
Fig. 3.5. Genotyping of strategy 2 transfections series 1. A) Diagnostic PCR with primer pairs P2/P3 and P4/P5 for integration across the 3'- and 5'-integration sites, respectively, of gDNA from IMC1bΔ1 and IMC1bΔ3 transfections. IMC1b/GFP is used as a positive control, while parental parasites are used as negative control. B) Clonal parasite populations after dilution cloning of IMC1bΔ1 possess full-length imc1b::gfp alleles. IMC1b/GFP is used as a positive control. C) Diagnostic PCR of transfection IMC1bΔ1 with primer pair P4/P6, showing a doublet amplified from full-length imc1b::gfp and imc1bΔ1::gfp. IMC1b/GFP is used as a full-length control.

In a second series of transfections, drug resistant parasites were obtained for all four transfections, but only IMC1bΔ1 and IMC1bΔ3 populations appeared to have integration according to diagnostic PCR (Fig. 3.6A). In this case, however, diagnostic PCR across the 5'-integration site failed to support integration (Fig. 3.6B) and, following dilution cloning, only clones with an unmodified *imc1b* allele were obtained (Fig. 3.6C). A third series of transfections resulted in drug-resistant parasites in three transfections, but no integration was observed in any (Fig. 3.6D).
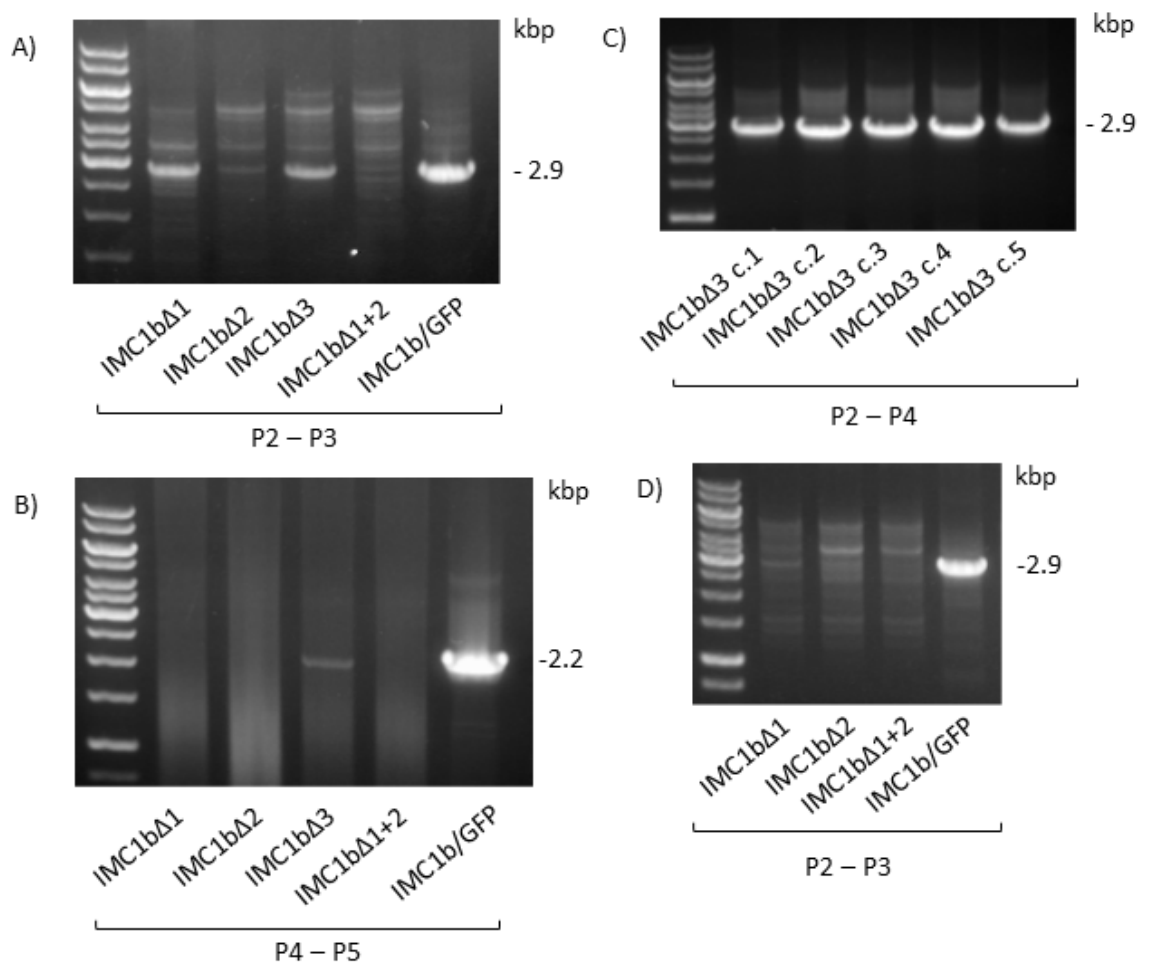
Fig. 3.6. Genotyping of strategy 2 transfections series 2 and 3. A) Diagnostic PCR with primer pairs P2/P3 for integration across the 3'-integration sites of gDNA from IMC1bΔ1, IMC1bΔ2, IMC1bΔ3 and IMC1bΔ1+2 transfections series 2. IMC1b/GFP is used as a positive control. B) Diagnostic PCR with primer pair P4/P5 for integration across 5'-integration sites of gDNA from IMC1bΔ1 and IMC1bΔ3 transfections. IMC1b/GFP is used as a positive control. C) Diagnostic PCR with primer pair P2/P4 of clonal parasite populations after dilution cloning of IMC1bΔ3 and IMC1bΔ3 amplifies an approximately 3.0kb fragment indicative of the unmodified *imc1b* allele. D) Diagnostic PCR with primer pairs P2/P3 for integration across the 3'-integration sites of gDNA from IMC1bΔ1, IMC1bΔ2, and IMC1bΔ1+2 transfections series 3. IMC1b/GFP is used as a positive control.

3.4 Discussion

IMC1b is expressed exclusively in the ookinete life cycle stage of the parasite. Through tagging and knockout of IMC1b, previous studies have shown that the full length protein localises to the pellicle of the parasite, consistent with its predicted cytoskeletal role as an intermediate filament protein (Tremp *et al*., 2008). Furthermore, knockout of IMC1b gives rise to misshapen ookinetes with reduced tensile strength, motility and infectivity that can be quantified through a number of established assays (Tremp *et al*., 2008). IMC1b and its structural paralogue IMC1a possess a conserved type 1 (ALV1) and type 2 (ALV2) domain, as well as an unrelated conserved domain near the C-terminus (ALV3). Although the first two domains do not share much homology at the primary sequence level, they are structurally similar in terms of possessing tandem repeats of a 12 amino acid periodicity, and in terms of their amino acid composition (64% and 66%, respectively, composed of the amino acids proline, isoleucine, valine, aspartate, glutamate and lysine) (Al-Khattaf *et al*., 2015). While this suggests that these two domains have functionally equivalent roles, this has not yet been experimentally shown. Moreover, the role of the third conserved (non-alveolin) ALV3 domain remains elusive. This study set out to shed light on these issues through a structure-function analysis of IMC1b.

The first experimental strategy used in this study was to replace the native *imc1b* allele with different parts of the gene, tagged with GFP, in a 'knock-in' approach. This would allow both the subcellular localisation and functional contribution of the alveolin domains to SPN targeting, shape, tensile strength and motility to be determined in live parasites. The different variations of transgene domains were chosen to assess the effects of different IMC1b modules, independently or combined, on the protein's

function. In particular this would compare the contributions of the type 1 (ALV1) vs. type 2 (ALV2) alveolin domains, as well as shed light on the role of the non-alveolin (ALV3) domain. In addition, IMC1a was chosen to be expressed under the IMC1b promoter in order to test the hypothesis that it can complement the function of IMC1b in the ookinete. In terms of experimental design, it is possible that the 30 N-terminal amino acids of IMC1b upstream of the IMC1b domains could interfere with correct expression. However, this is unlikely as these amino acids are naturally present in full-length IMC1b. It is also possible that the C-terminal GFP module could interfere with the function of the upstream domains. This is again unlikely, because full-length IMC1b can be GFP-tagged without apparent adverse effect on its function (Tremp *et al*., 2008). Nonetheless, due to the shorter sequences of the domains being expressed and the different proximity of the N-terminal 30 amino acids and C-terminal GFP module to the expressed domains, the folding of the fusion proteins could potentially be affected differently compared to full-length IMC1b::GFP.

Because *imc1b* had previously been successfully targeted for tagging and knockout, no problems were anticipated with our allelic replacement strategy. Initial observations of the uncloned parasite populations using confocal microscopy showed positive, green fluorescent protein being expressed. This indicated that the DNA constructs were accurate with no frame shifts, with exception of the IMC1a construct as was later shown. However, the observations were unexpected in that almost every green fluorescent ookinete appeared to be morphologically normal (Fig. 3.3B); it was expected that at least some of the allelic replacements would have resulted in an abnormal shape phenotype.

When unable to clone the desired transgenic parasites from second and third round drug selected populations, obtaining only wildtype parasites, further questions were raised about the success of the transfections. Indeed, genomic DNA analysis of the parasites from second and third rounds of drug selection showed that the population of parasites containing an integrated selectable marker gene, as was seen after the first round of drug selection, decreased with time. This indicated that the transgenes were in fact not stably integrated, and instead were being expressed episomally. Parasites with episomal resistance are normally outgrown by those with integrated resistance genes, because only part of the progeny inherit the episomes. This also explained why only clonal populations with unmodified *imc1b* alleles were obtained by dilution cloning, as well as the 'normal' shape of the GFP-positive ookinetes.

Besides the three times the transfections were repeated in order to maximise the chances of pulling out a successfully integrated transgenic parasite, an effort was also made to try and clone out parasite populations directly from the first round of drug selected parasites, containing the apparent highest proportion of successfully integrated DNA from the correct size fragment being amplified. This was also not fruitful, probably because the proportion of parasites with episomal resistance to those with integrated resistance was too high. It was therefore not justifiable to further repeat dilution cloning experiments as it was wasting animal units needlessly.

Double homologous crossover recombination relies on two separate pieces of DNA having large stretches of homologous DNA sequence. During replication, double-stranded breaks are formed in the gDNA of the parasite, where host cell DNA repair mechanisms are recruited to repair the break, and through regions of high sequence homology, the DNA repair mechanisms utilise the transfected DNA fragment for

integration. This process must happen at both the 5' and 3' end of the DNA sequence introduced in order for successful, stable integration to occur. On balance, our collective data indicate that no stable integration was obtained. This raises two questions. First, why was it so difficult to achieve stable integration of the recombinant alleles into the *imc1b* locus, given that this had worked successfully in the past using the same DNA construct as the vehicle for homologous recombination? Perhaps targeting *imc1b* is more difficult than was originally thought. There could also be small differences in the experimental setups that may account for less efficient transfection on this occasion. The amounts of residual circular plasmid in the doubly digested DNA samples used for transfection may also differ (circular DNA is the main source of episomally maintained resistance), although extra care was taken to digest the DNA to completion even including a third enzyme (*Not*I) to minimise any circular DNA remaining. It is also possible that the successful crossover events into the *imc1b* locus were accompanied by unforeseen recombination events in a different part of the genome that caused a fitness loss causing these parasites to be outcompeted by those carrying episomes.

The second question raised is what caused the (false) positive diagnostic PCR for integration? The same diagnostic PCR was used reliably in other experiments (Tremp and Dessens, 2011, Tremp *et al*., 2008), suggesting it is a good diagnostic indicator of integration. It is possible that the forward primer may have annealed to the plasmid-resident *tgdhfr* sequence and synthesised a sense strand, whilst the reverse primer could have annealed to the genomic DNA region downstream of the integration site, and produced an antisense strand that would have significant overlap with the amplified sense strand, producing template DNA from which oligonucleotide primers

could further amplify the diagnostic 1.2kb DNA fragment. Whilst we cannot rule out this possibility, it is worth noting that there had not previously been any issues with 'false positive' diagnostic PCR in the Dessens laboratory.

Given the persistence of the problems with the 'knock in' strategy used here, it was decided to change to the second approach, whereby the native *imc1b* allele was replaced with an *imc1b* allele lacking specific modules. Using this strategy, only 1 in 12 transfections led to verifiable stable integration, indicating that the *imc1b* locus is actually quite difficult to target. In the one experiment where we obtained stable integration, we ended up with a population that contained the desired IMC1bΔ1 parasites, but which were mixed with parasites encoding the full-length IMC1b::GFP. This happens when crossover at the 5'-end occurs downstream of the mutated region. We were therefore unable to clone out the desired parasite line.

Chapter 4

Distinct Functional Contributions by the Conserved Domains of the Malaria Parasite Alveolin IMC1h

RESEARCH PAPER COVER SHEET

SECTION A – Student Details

Student ID Number	1604906	Title	Mr
First Name(s)	Michael Patrick		
Surname/Family Name	Coghlan		
Thesis Title	A Combined Molecular, Cell and Structural Biology Approach Towards Characterising Malaria Alveolins		
Primary Supervisor	Dr. Johannes Dessens		

SECTION B – Paper Already Published

Where was the work published?	Frontiers in Cellular and Infection Microbiology
When was the work published?	07/2019
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	

Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or author) to include this work.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach further sheet if necessary)	Although this is a co-authored research paper, the work included was primarily my own. Annie Tremp and Sadia Saeed assisted with animal work and mosquito dissections. Johannes Dessens helped design the experiments and provided guidance and support with experimental work and preparation of the manuscript. All co-authors contributed to manuscript revision.

SECTION E

Student Signature	
Date	

Supervisor Signature	
Date	

Distinct Functional Contributions by the Conserved Domains of the Malaria Parasite Alveolin IMC1h

Michael P. Coghlan[1,2], Annie Z. Tremp[1], Sadia Saeed[1], Cara K. Vaughan[2] and Johannes T. Dessens[1]*

[1] Department of Infection Biology, London School of Hygiene and Tropical Medicine, London, United Kingdom, [2] Institute of Structural and Molecular Biology, School of Biological Sciences, Birkbeck, London, United Kingdom

Invasive, motile life cycle stages (zoites) of apicomplexan parasites possess a cortical membrane skeleton composed of intermediate filaments with roles in zoite morphogenesis, tensile strength and motility. Its building blocks include a family of proteins called alveolins that are characterized by conserved "alveolin" domains composed of tandem repeat sequences. A subset of alveolins possess additional conserved domains that are structurally unrelated and the roles of which remain unclear. In this structure-function analysis we investigated the functional contributions of the "alveolin" vs. "non-alveolin" domains of IMC1h, a protein expressed in the ookinete and sporozoite life cycle stages of malaria parasites and essential for parasite transmission. Using allelic replacement in *Plasmodium berghei,* we show that the alveolin domain is responsible for targeting IMC1h to the membrane skeleton and, consequently, its deletion from the protein results in loss of function manifested by abnormally-shaped ookinetes and sporozoites with reduced tensile strength, motility and infectivity. Conversely, IMC1h lacking its non-alveolin conserved domain is correctly targeted and can facilitate tensile strength but not motility. Our findings support the concept that the alveolin module contains the properties for filament formation, and show for the first time that tensile strength makes an important contribution to zoite infectivity. The data furthermore provide new insight into the underlying molecular mechanisms of motility, indicating that tensile strength is mechanistically uncoupled from locomotion, and pointing to a role of the non-alveolin domain in the motility-enhancing properties of IMC1h possibly by engaging with the locomotion apparatus.

Keywords: *Plasmodium*, cytoskeleton, ookinete, sporozoite, motility, transmission

INTRODUCTION

Plasmodium species, the causative agents of malaria, have a complex life cycle in vertebrate host and mosquito vector. Parasite-infected erythrocytes multiply via an asexual replication cycle called schizogony to release merozoites that infect new red blood cells. A small percentage of these develop into sexual stage precursor cells (gametocytes) which, after uptake with the blood meal of a feeding mosquito, begin a rapid process of gamete formation and fertilization inside the mosquito midgut.

The resultant zygotes undergo meiosis and transform into elongated forms termed ookinetes, which traverse the midgut epithelium and then round up to form the oocysts. In the following weeks, young oocysts grow and divide by a process known as sporogony to generate hundreds of daughter cells named sporozoites. After egress from the oocyst, sporozoites invade and inhabit the salivary glands, and are transmitted to new hosts by mosquito bite to first infect liver cells from which new malaria blood stage infections are initiated to complete the life cycle.

The merozoite, ookinete, and sporozoite constitute the three motile and invasive stages in the *Plasmodium* life cycle. These so-called "zoite" stages possess a characteristic cortical structure termed the pellicle. The pellicle is defined by a double membrane structure termed the inner membrane complex (IMC) situated directly underneath the plasma membrane, which is equivalent to a sutured system of flattened sacs or alveoli (Bannister et al., 2000; Morrissette and Sibley, 2002; Santos et al., 2009). On the cytoplasmic face of the IMC, and tightly associated with it, sits a network of intermediate filaments termed the subpellicular network (SPN), a viscoelastic membrane skeleton that supports the IMC and provides tensile strength to the cell (Mann and Beckers, 2001). A family of IMC1 proteins, now called alveolins, have been identified as major components of the SPN (Mann and Beckers, 2001; Khater et al., 2004). Members of the alveolin family are found in apicomplexans and chromerids, as well as in ciliates and dinoflagellate algae, which together with the apicomplexans form the Alveolata superphylum (Gould et al., 2008). The alveolins are part of a larger class of proteins called epiplastins that aside alveolates have also been identified in euglenids, glaucophytes and cryptophytes (Goodenough et al., 2018). In the genus *Plasmodium*, 13 conserved and syntenic alveolin family members have thus far been identified that are differentially expressed among the three different zoites stages (Al-Khattaf et al., 2015; Kaneko et al., 2015). In addition, two PhIL1 interacting proteins: PIP2 and PIP3, show structural homology with alveolins (Kono et al., 2013; Parkyn Schneider et al., 2017).

It has been shown in the rodent malaria species *P. berghei* that disruption of the alveolins IMC1a, IMC1b, or IMC1h gives rise to morphological aberrations that are accompanied by reduced tensile strength of the zoite stages in which they are expressed. The same null mutant parasites also display motility defects, indicating that these alveolins also participate in parasite locomotion through an as yet unknown mechanism (Khater et al., 2004; Tremp et al., 2008; Tremp and Dessens, 2011; Volkmann et al., 2012). The SPN effectively separates the main cytosol from a smaller cortical cytoplasm that contains the molecular machinery that drives apicomplexan zoite motility, invasion and egress. Motility of apicomplexan zoites relies on an actinomyosin motor system that is situated in the space between the plasma membrane and the IMC. The conventional model is that the molecular motor and its auxiliary proteins is linked to cell surface adhesins via actin filaments and bridging proteins, and is internally anchored into the IMC. Motor force drives the actin filaments and adhesins rearward, thereby creating a traction force that propels the cell in the opposite direction against a substrate (Frenal et al., 2010). The IMC is underlain by the rigid yet flexible SPN, and this is most likely how alveolins assert their role in motility, either indirectly by providing mechanical support to the IMC, or through interactions with components of the motility apparatus.

The structural homologies between alveolin proteins are largely confined to conserved domains containing tandem repeat sequences (Al-Khattaf et al., 2015), herein referred to as "alveolin" domains. A subset of alveolins possess additional conserved modules that are structurally unrelated to the archetypal "alveolin" module. The roles of these "non-alveolin" domains in protein function are poorly understood. In this study, we investigated the functional contributions of the "alveolin" vs. "non-alveolin" modules of IMC1h, using the rodent malaria parasite species *P. berghei* and a strategy of allelic replacement and GFP tagging. IMC1h is expressed in both the ookinete and sporozoite life cycle stages of the parasite, where it carries out equivalent roles (Tremp and Dessens, 2011; Volkmann et al., 2012), thus allowing our investigations to be conducted across two distinct zoite stages. The results obtained indicate that the two IMC1h modules play distinct parts in facilitating morphogenesis, tensile strength and motility. The implications of these results are discussed in the context of parasite infectivity.

MATERIALS AND METHODS

Animal Use

Experiments were conducted in 6–8 weeks old female CD1 mice, specific pathogen free and maintained in filter cages. Animal welfare was assessed daily and animals were humanely killed upon reaching experimental or humane endpoints. Mice were infected with parasites by intraperitoneal injection, or by infected mosquito bite on anesthetized animals. Parasitemia was monitored regularly by collecting of a small drop of blood from a superficial tail vein. Drugs were administered by intraperitoneal injection or where possible were supplied in drinking water. Parasitized blood was harvested by cardiac bleed under general anesthesia without recovery.

Gene Targeting Vectors

To delete domain 1 of IMC1h (amino acids 103 to 306) primers IMC1hdeltadomain1-F (GAGGTTCACAATATTTGAA TAACAATCAAGCACA) and IMC1hdeltadomain1-R (AAAT ATTGTGAACCTCCATACAAAGTGTGTT) were used to PCR amplify plasmid pLP-IMC1h/GFP (Tremp and Dessens, 2011). Template plasmid was removed after the PCR by *Dpn*I digestion, and the PCR product was circularized by in-fusion, to give plasmid pLP-IMC1hΔ1. This mutation deletes 205 amino acids from the IMC1h::GFP fusion protein (**Figure 1**). The same approach was used to delete domain 2 (amino acids 397 to 504), using primers IMC1hdeltadomain2-F (ATAATTCGGT TAAAGCTATCCAGAAAAACAT) and IMC1hdeltadomain2-R (GCTTTAACCGAATTATTTTGTCTATAATCCATATTTGA) to give plasmid pLP-IMC1hΔ2. This mutation deletes 109 amino acids from the IMC1h::GFP fusion protein (**Figure 1**).

FIGURE 1 | Generation and genotyping of IMC1h mutants. (A) Schematic diagram of the parental and modified *imc1h* alleles. The *imc1h* gene is indicated with coding sequence (wide bars) and 5′ and 3′ untranslated regions (narrow bars). The relative positions of domains 1 and 2 are indicated (gray). Also shown are the *gfp* module, positions of primers P1–P3 used for diagnostic PCR amplification, and the predicted lengths of the amplicons (arrows). (B) PCR with primers P1 and P3 diagnostic for integration of the modified, GFP-tagged *imc1h* alleles into the *imc1h* locus, or with primers P1 and P2 diagnostic for absence/presence of the unmodified parental *imc1h* allele. Amplicon sizes (kb) are indicated on the right-hand side and correspond to those depicted in (A). (C) Western blot analysis of purified ookinetes (∼ 50,000 per lane) of IMC1h mutant parasite lines using anti-GFP antibodies, showing successful expression of IMC1h::GFP fusion proteins. Identified band sizes reflect the introduced domain deletions in IMC1h, resulting in proteins of 86 kDa (IMC1h/GFP), 62 kDa (IMC1hΔ1), and 74 kDa (IMC1hΔ2). The ookinete surface protein Pbs21 is shown as a loading control.

Generation and Genotyping of Genetically Modified Parasites

Parasite transfection, pyrimethamine selection and dilution cloning were performed as previously described (Waters et al., 1997). Prior to performing transfections, plasmid DNA was double-digested with KpnI and SacII to remove the plasmid backbone. Genomic DNA extraction was performed as previously described (Dessens et al., 1999). After transfection, drug resistant parasites were subjected to limiting dilution cloning. Integration of the GFP-tagged IMC1h-encoding sequence into the *imc1h* locus was confirmed by diagnostic PCR across the 5' integration site with primers P1 (CCATTTTTA TGTTGAGCTTGAG) and P3 (GTGCCCATTAACCATCACC) (**Figure 1**). The absence of the unmodified *imc1h* allele in the clonal parasite lines was confirmed by diagnostic PCR with primers P1 and P2 (TTCTATAATCTTTAATTGTTCAGAA ATGTG) (**Figure 1**).

Sporozoite Size Measurements

Images of individual midgut sporozoites were captured by microscopy on Zeiss LSM510 inverted laser scanning confocal microscope. Using Zeiss LSM image browser software the circumference was measured from which the occupied surface area ("footprint") was calculated as a measure of cell size. Statistical analysis was carried out using Student's *t*-test.

Zoite Motility

Ookinete motility was assessed essentially as previously described (Moon et al., 2009). Aliquots of neat ookinete cultures were added to equal volumes of Matrigel (BD Biosciences) on ice, mixed thoroughly, spotted onto a microscope slide, and covered with a cover slip. After sealing with nail varnish, the Matrigel was

allowed to set at room temperature for 30 min before analysis. Sporozoites were gently released from ∼ 20 salivary glands in 200 µl RPMI medium in a Dounce homogenizer on ice. Following addition of an equal volume of RPMI supplemented with 20% fetal bovine serum, the sporozoites were transferred to an Eppendorf tube and collected by centrifugation in a swing-out rotor for 10 min at 1,000 × g at 4°C, followed by removal of excess supernatant. Aliquots of resuspended sporozoites were mixed with an equal volume of Matrigel on ice before transfer to microscope slides. After sealing the cover slip with nail varnish, the Matrigel was allowed to set at room temperature for 30 min before analysis. Cells were examined and time-lapse images taken on a Zeiss Axioplan II microscope. Movies were analyzed with ImageJ using the Manual Tracking plugin. Statistical analysis was carried out using ANOVA and Tukey's multiple comparison.

Osmotic Shock and Viability Assays

Ookinetes in neat culture were subjected to hypo-osmotic shock of 0.5 × normal osmotic strength by adding an equal volume of water. Sporozoites were released from oocyst-infected midguts at 15 days post-infection and were subjected to 0.33 × normal osmotic strength by adding two equal volumes of water. After 5 min, normal osmotic conditions were restored by adding an appropriate amount of 10 × PBS. Cell viability was scored by fluorescence microscopy in the presence of 0.5% propidium iodide and 1% Hoechst 33258. Ookinetes whose nuclei stained positive for both propidium iodide and Hoechst were scored as non-viable, whereas ookinetes whose nuclei stained positive only for Hoechst were scored as viable. Values were normalized to 100% viability in untreated cells.

Mosquito Infection

At 6 days before infecting mosquitoes, mice were injected intraperitoneally with phenylhydrazine (6 mg/ml in PBS, 10 μl/g body weight) to induce reticulocytosis. At 3 days before mosquito feeding, mice were infected intraperitoneally with 10^7 parasitized red blood cells. The day of the feed, parasitemia and gametocytemia were checked by using a Giemsa-stained blood film. Mice were anesthetized and placed on a cage containing up to 50 starved female mosquitoes. Insects were allowed to blood feed in a draft-free, darkened environment at room temperature for 15 min. The day after feeding, unfed or partially fed mosquitoes (i.e., those unlikely to be infected) were removed if desired. For sporozoite transmission, the prevalence of infection of mosquito batches was determined and combined with the number of blood meal-positive insects after feeding to estimate the number of sporozoite-infected mosquitoes that fed.

RESULTS

Generation and Genotyping of IMC1h Mutants

Multiple alignment of amino acid sequences of IMC1h orthologs from different *Plasmodium* species clearly reveals the presence of its single conserved "alveolin" module (domain 1), as well as a conserved carboxy-terminal module that is structurally unrelated (domain 2) (Tremp and Dessens, 2011). We previously generated and characterized parasites lines stably expressing IMC1h::GFP from the native *imc1h* promoter (named IMC1h/GFP) (Tremp and Dessens, 2011). To study the role of domains 1 and 2 we used the same allelic replacement strategy, generating transgenic parasites lines that express IMC1h::GFP without domain 1 (named IMC1hΔ1) or IMC1h::GFP lacking domain 2 (named IMC1hΔ2) (**Figure 1A**). Diagnostic PCR with primers P1 and P3 (**Figure 1A**) of clonal parasite lines amplified expected products of ∼ 2.2, 1.6, and 1.9 kb from parasite lines IMC1h/GFP, IMC1hΔ1, and IMC1hΔ2, respectively, confirming integration of the modified *imc1h* alleles into the *imc1h* locus (**Figure 1B**). Moreover, amplification with primers P1 and P2 (**Figure 1A**) amplified an ∼ 3 kb product only from the parental parasite, confirming absence of the unmodified *imc1h* allele in the transgenic lines (**Figure 1B**). Western blot analysis of purified ookinete samples using anti-GFP antibodies showed comparable expression levels of IMC1h::GFP fusion proteins in all three parasite lines, with protein sizes of the IMC1h::GFP fusion proteins corresponding to the introduced amino acid deletions in IMC1h (**Figure 1C**). This indicated that the truncated IMC1h proteins are stably expressed in ookinetes and confirms the successful introduction of the domain 1 and domain 2 deletions.

Ookinete-Specific Subcellular Localization of Mutant IMC1h::GFP and Cell Shape

To study the subcellular localization of the truncated IMC1h::GFP fusion proteins in parasite lines IMC1hΔ1 and IMC1hΔ2, live ookinetes were examined for GFP fluorescence. As described previously (Tremp and Dessens, 2011), ookinetes expressing full-length IMC1h::GFP had normal
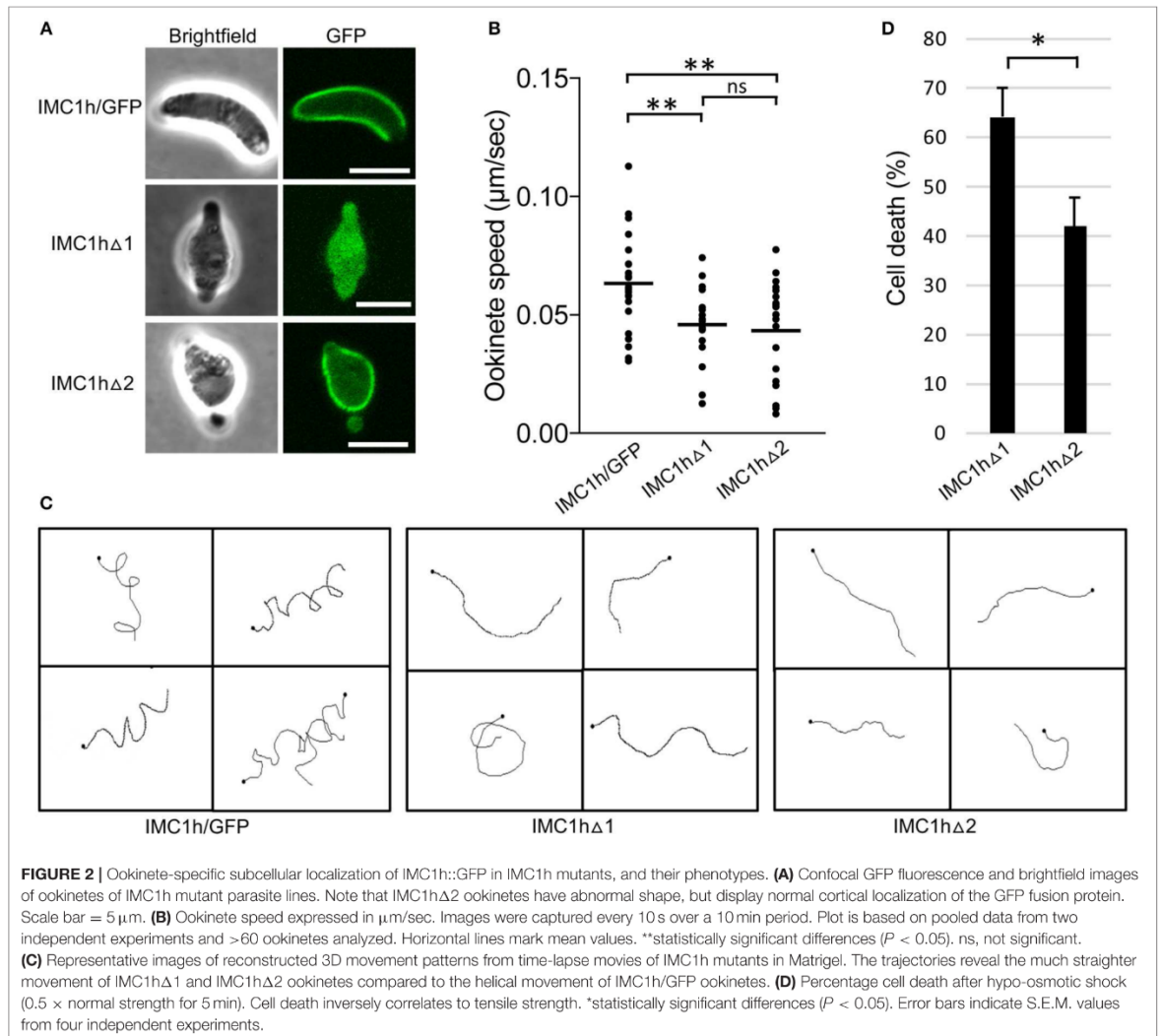
shape and displayed a predominantly cortical localization of GFP fluorescence (**Figure 2A**), consistent with the recruitment of IMC1h to the SPN. Both IMC1hΔ1 and IMC1hΔ2 ookinetes were misshapen, lacking the typical crescent shape and possessing a bulging area in the center (**Figure 2A**) similar to the shape of IMC1h-KO ookinetes (Tremp and Dessens, 2011). This indicates that domain 1 and domain 2 are both required for IMC1h to facilitate normal ookinete morphogenesis. Subcellular localization of GFP fluorescence was markedly different between IMC1hΔ1 and IMC1hΔ2 ookinetes: while the GFP signal in IMC1hΔ2 ookinetes was predominantly found at the cell cortex like in IMC1h/GFP ookinetes, IMC1hΔ1 ookinetes displayed only cytoplasmic fluorescence (**Figure 2A**). These observations show that domain 1 is required for recruitment of IMC1h to the SPN.

Motility, Tensile Strength, and Infectivity of IMC1h Mutant Ookinetes

As the truncated IMC1h::GFP fusion protein expressed in IMC1hΔ1 ookinetes fails to reach its site of action: the SPN (**Figure 2A**), this parasite is effectively an IMC1h null mutant. Indeed, assessment of ookinete motility in Matrigel revealed that IMC1hΔ1 ookinetes had a significantly lower average speed to that of IMC1h/GFP ookinetes (**Figure 2B**), as was reported for IMC1h null mutants (Tremp and Dessens, 2011). IMC1hΔ2 ookinetes also had a significantly lower average speed than IMC1h/GFP ookinetes, that was comparable to that of IMC1hΔ1 ookinetes (**Figure 2B**). The manner of ookinete movement as judged by their 3D trajectories was also examined: this showed that IMC1h/GFP ookinetes moved in a clearly helical fashion, while the trajectories of both IMC1hΔ1 and IMC1hΔ2 ookinetes were much more linear (**Figure 2C**). Similar differences were reported between wildtype and IMC1h null mutant ookinetes (Volkmann et al., 2012; Kan et al., 2014). Examination of videos of 3D ookinete movement showed that the abnormally shaped ookinetes of parasite lines IMC1hΔ1 and IMC1hΔ2 still possessed mildly helical movement albeit with a smaller radius, explaining why their trajectories appeared more linear (**Videos S1–S3**). These collective findings indicate that deletion of either domain 1 or domain 2 from IMC1h results in a loss-of-function phenotype with respect to ookinete motility.

We also assessed tensile strength of ookinetes using a hypo-osmotic shock assay. Hypo-osmotic conditions cause the cells to draw in water and swell, and the degree of hypo-osmotic stress a cell can tolerate is a measure of its tensile strength (Menke and Jockusch, 1991). In this assay, IMC1hΔ2 ookinetes were significantly more resistant to hypo-osmotic shock than their IMC1hΔ1 counterparts (**Figure 2D**), indicating that IMC1hΔ2 ookinetes have superior tensile strength to IMC1hΔ1 ookinetes. Thus, IMC1h lacking domain 2 is able to facilitate tensile strength above null mutant levels.

To assess the effects of the IMC1h module deletions on parasite infectivity, we infected *Anopheles stephensi* mosquitoes and recorded oocyst numbers as a measure of ookinete infectivity (**Table 1**). Reproducibly, both IMC1hΔ1 and IMC1hΔ2 ookinetes gave rise to statistically significantly

FIGURE 2 | Ookinete-specific subcellular localization of IMC1h::GFP in IMC1h mutants, and their phenotypes. **(A)** Confocal GFP fluorescence and brightfield images of ookinetes of IMC1h mutant parasite lines. Note that IMC1hΔ2 ookinetes have abnormal shape, but display normal cortical localization of the GFP fusion protein. Scale bar = 5 μm. **(B)** Ookinete speed expressed in μm/sec. Images were captured every 10 s over a 10 min period. Plot is based on pooled data from two independent experiments and >60 ookinetes analyzed. Horizontal lines mark mean values. **statistically significant differences (P < 0.05). ns, not significant. **(C)** Representative images of reconstructed 3D movement patterns from time-lapse movies of IMC1h mutants in Matrigel. The trajectories reveal the much straighter movement of IMC1hΔ1 and IMC1hΔ2 ookinetes compared to the helical movement of IMC1h/GFP ookinetes. **(D)** Percentage cell death after hypo-osmotic shock (0.5 × normal strength for 5 min). Cell death inversely correlates to tensile strength. *statistically significant differences (P < 0.05). Error bars indicate S.E.M. values from four independent experiments.

(P < 0.05, t-test) reduced oocyst numbers compared to IMC1h/GFP control parasites (**Table 1**), indicating that they are less infective than their wildtype counterparts. IMC1hΔ2 ookinetes produced significantly higher oocyst numbers than IMC1hΔ1 ookinetes in one experiment (Experiment II, **Table 1**), indicating that they are more infective than their IMC1hΔ1 counterparts.

Sporozoite-Specific Subcellular Localization of Mutant IMC1h::GFP and Cell Shape

Sporozoites of parasite line IMC1hΔ1 were misshapen possessing a bulging area (**Figure 3A**), similar to IMC1h null mutant sporozoites (Tremp and Dessens, 2011). Like IMC1hΔ1 ookinetes (**Figure 2A**), these sporozoites displayed

only cytoplasmic fluorescence (**Figure 3A**) reflecting absence of SPN targeting. By contrast, the large majority (95%, n = 100) of IMC1hΔ2 midgut sporozoites had a normal crescent shape and displayed cortical localization of GFP fluorescence (**Figure 3A**) indicative of normal SPN targeting as observed in IMC1hΔ2 ookinetes (**Figure 2A**). Assessment of sporozoite sizes revealed that IMC1hΔ1 sporozoites had a significantly smaller average size than IMC1hΔ2 sporozoites (P < 0.001), which in turn were significantly smaller than IMC1h/GFP sporozoites (P < 0.005) (IMC1hΔ1 footprint: 6.9 ± 0.16 μm²; IMC1hΔ2 footprint: 8.5 ± 0.14 μm²; IMC1h/GFP footprint; 10.0 ± 0.42 μm²; n = 30). Interestingly, IMC1hΔ2 sporozoites lost their normal shape during transition from the midgut to the salivary glands, resulting in salivary gland sporozoites possessing a bulging area (**Figure 3A**). These combined observations indicate that IMC1hΔ2

TABLE 1 | Development of IMC1h mutant parasite lines in *Anopheles stephensi*.

Experiment	Parasite line	Infection prevalence[a]	Mean ± SEM oocyst number[b]	Median oocyst number	Mean salivary gland sporozoite number[c]	Salivary gland sporozoites per oocyst[d]
I	IMC1h/GFP	80 (15)	90 ± 19	82	n/a	n/a
	IMC1hΔ1	27 (15)	8.8 ± 5	5	n/a	n/a
	IMC1hΔ2	60 (15)	4.7 ± 1	4	n/a	n/a
II	IMC1h/GFP	100 (15)	102 ± 22	62	7,230 (13)	71
	IMC1hΔ1	100 (15)	7.9 ± 2.3	5	300 (20)	38
	IMC1hΔ2	100 (15)	32 ± 7.3	25	2,300 (20)	72
III	IMC1h/GFP	100 (21)	104 ± 25	48	n/a	n/a
	IMC1hΔ1	100 (21)	27 ± 6.5	20	625 (20)	23
	IMC1hΔ2	100 (22)	39 ± 10	26	2,760 (20)	70
IV	IMC1h/GFP	83 (24)	132 ± 26	85	7,968 (15)	60
	IMC1hΔ1	88 (24)	37 ± 7.3	19	528 (15)	14
	IMC1hΔ2	78 (27)	35 ± 6.8	20	2,496 (15)	71

[a]*Percentage of mosquitoes with at least one oocyst. (n) denotes the total number of mosquitoes analyzed.*
[b]*Oocysts were counted between 9 and 11 days post infection. Only infected insects were included.*
[c]*Average number of sporozoites per mosquito was calculated from (n) pooled salivary glands.*
[d]*Average number of sporozoites per mosquito divided by mean oocyst number.*

parasites possess an intermediate phenotype with respect to sporozoite morphogenesis.

Tensile Strength, Motility, and Infectivity IMC1h Mutant Sporozoites

Tensile strength assessment of midgut sporozoites indicated that IMC1hΔ2 sporozoites had higher tensile strength than their IMC1hΔ1 counterparts, but lower tensile strength than IMC1h/GFP sporozoites (**Figure 3B**). Both IMC1hΔ1 and IMC1hΔ2 sporozoites had similar average speed that was significantly reduced compared to that of their IMC1h/GFP counterparts (**Figure 3C**). The manner of sporozoite 3D movement as judged by their trajectories in Matrigel was also different between the parasite lines examined: while IMC1h/GFP sporozoites moved in a predominantly circular fashion in Matrigel, the trajectories of IMC1hΔ1 and IMC1hΔ2 sporozoites were more meandering and less circular (**Figure 3D**), as indeed was reported for IMC1h null mutants (Volkmann et al., 2012). These collective observations indicate that deletion of either domain 1 or domain 2 from IMC1h results in a loss-of-function phenotype with respect to sporozoite motility.
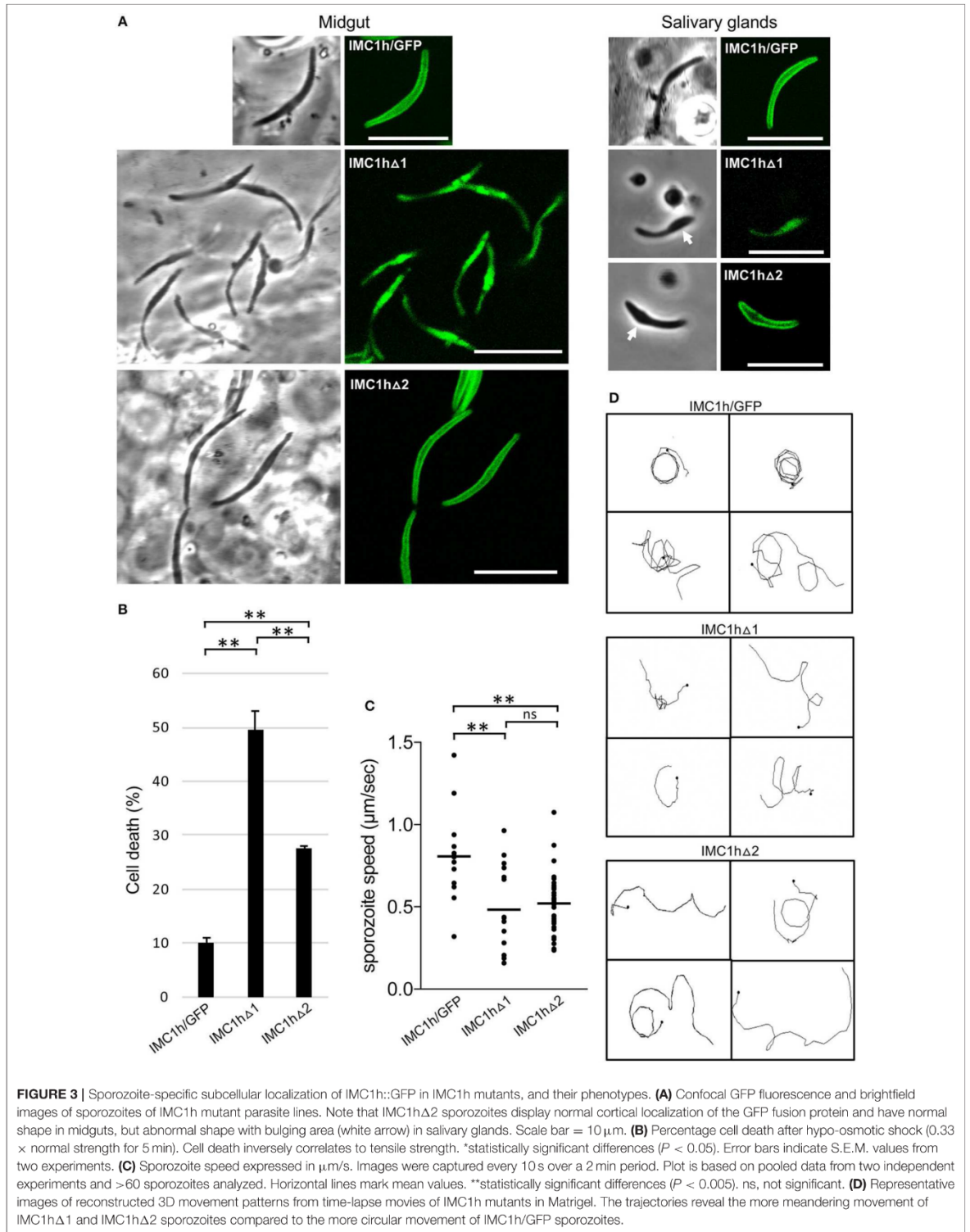
Sporozoite infectivity to the mosquito was assessed by counting salivary gland sporozoite numbers 3 weeks post-infection. Assuming a linear correlation between the number of oocyst and the number of sporozoites produced, these data indicated that IMC1hΔ2 sporozoites were almost three times more infective to salivary glands than IMC1hΔ1 sporozoites ($P < 0.05$, t-test), and had similar efficacy to IMC1h/GFP sporozoites in colonizing the salivary glands (**Table 1**). This is consistent with previous findings that IMC1h null mutant sporozoites are less invasive than their wildtype counterparts (Tremp and Dessens, 2011; Volkmann et al., 2012). Despite their presence in salivary glands, we failed to transmit IMC1hΔ1 and IMC1hΔ2 sporozoites by mosquito bite to naive mice, indicating

that their infectivity to the mammalian host by natural route of transmission is compromised.

DISCUSSION

Plasmodium alveolins have important roles in morphogenesis, tensile strength and motility of the zoite stages, and in many cases their disruption leads to loss of parasite fitness, infectivity and transmission (Khater et al., 2004; Tremp et al., 2008, 2014; Tremp and Dessens, 2011). Here, we present data from a structure-function analysis of the alveolin IMC1h, aimed to determine the contributions of its "alveolin" and "non-alveolin" modules to protein function. This was based on a research strategy of allelic replacement of *imc1h* with modified versions of the gene fused to a GFP module for localization purposes. Our results demonstrate that the alveolin module of IMC1h is necessary for targeting the protein to the cortical membrane skeleton. This is in full agreement with a report showing that the alveolin domain of TgIMC3 can target YFP to the SPN of *Toxoplasma gondii* in a wildtype parasite background (Anderson-White et al., 2011). Our data show furthermore that the alveolin domain of IMC1h is sufficient to increase tensile strength above null mutant levels. Given the filamentous nature of the SPN (Mann and Beckers, 2001), the assembly or incorporation of the alveolins into such structures is likely to be a prerequisite for the provision of mechanical strength. Hence, our findings strongly suggest that the alveolin module contains the properties for intermediate filament formation, consistent with the fact that it is found in all alveolins and indeed is the distinguishing feature of this protein family. The processes of filament formation and SPN recruitment could be mechanistically linked as proposed (Tremp et al., 2017).

Previous studies have shown that knockout of IMC1a, IMC1b, or IMC1h reduced motility and tensile strength at the same time (Khater et al., 2004; Tremp et al., 2008; Tremp and Dessens,

FIGURE 3 | Sporozoite-specific subcellular localization of IMC1h::GFP in IMC1h mutants, and their phenotypes. **(A)** Confocal GFP fluorescence and brightfield images of sporozoites of IMC1h mutant parasite lines. Note that IMC1hΔ2 sporozoites display normal cortical localization of the GFP fusion protein and have normal shape in midguts, but abnormal shape with bulging area (white arrow) in salivary glands. Scale bar = 10 μm. **(B)** Percentage cell death after hypo-osmotic shock (0.33 × normal strength for 5 min). Cell death inversely correlates to tensile strength. *statistically significant differences ($P < 0.05$). Error bars indicate S.E.M. values from two experiments. **(C)** Sporozoite speed expressed in μm/s. Images were captured every 10 s over a 2 min period. Plot is based on pooled data from two independent experiments and >60 sporozoites analyzed. Horizontal lines mark mean values. **statistically significant differences ($P < 0.005$). ns, not significant. **(D)** Representative images of reconstructed 3D movement patterns from time-lapse movies of IMC1h mutants in Matrigel. The trajectories reveal the more meandering movement of IMC1hΔ1 and IMC1hΔ2 sporozoites compared to the more circular movement of IMC1h/GFP sporozoites.

2011). For this reason, it has not been previously possible to dissect the individual contributions of tensile strength and motility to zoite infectivity. This study shows that zoites of parasite line IMC1hΔ2 are better at infecting mosquito tissues than the equivalent IMC1hΔ1 parasites (**Table 1**). As IMC1hΔ1 and IMC1hΔ2 zoites do not differ discernibly in their motilities, their differences in tensile strength are likely to be the cause of their distinct invasive capacities. These experiments thus show for the first time that tensile strength is an important contributor to zoite infectivity in the mosquito. Cell rigidity and flexibility are likely to be important when ookinetes escape the blood meal in the midgut lumen and cross the peritrophic matrix and midgut epithelium, a process that has been described to cause major cell constrictions (Vernick et al., 1999; Han et al., 2000), and the same can be envisaged in sporozoites when entering mosquito salivary glands.

IMC1hΔ2 sporozoites are larger than IMC1hΔ1 sporozoites and smaller than wildtype sporozoites, and they display normal shape when first formed. These observations indicate that IMC1h without its non-alveolin module causes an intermediate phenotype with regards to sporozoite morphogenesis. Interestingly, mutations of putative amino- or carboxy-terminal palmitoylation sites of the sporozoite-specific alveolin IMC1a have also been shown to affect sporozoite shape and size, being in between that of null mutant and wildtype sporozoites (Al-Khattaf et al., 2017). Collectively, these findings point to a fundamental role of the alveolins in determining sporozoite size and shape. The mechanisms by which alveolins participate in zoite morphogenesis remain poorly understood and, interestingly, knockouts of other SPN proteins that are structurally unrelated to alveolins such as G2 and PhIL1 cause similar cell shape abnormalities (Barkhuff et al., 2011; Tremp et al., 2013). Zoite morphogenesis is concurrent with the formation of the IMC and SPN structures (Hu et al., 2002; Tremp et al., 2008), and one possibility is that this is a highly constrained process that is very sensitive to structural disruption of any SPN component.

The apparent lack of correlation between motility and tensile strength (**Figures 2**, **3**) indicates that these two processes are in fact mechanistically uncoupled. This corroborates previous findings that IMC1h null mutant ookinetes have comparably reduced speed to similarly misshapen G2 null mutant ookinetes, despite these two parasite lines possessing different tensile strengths (Tremp et al., 2013). Thus, these observations support a role of IMC1h in motility through interactions with the motility apparatus, possibly via IMC-resident bridging proteins, rather than by simply contributing mechanical support to the IMC. Furthermore, our finding that IMC1hΔ2 parasites fail to rescue the IMC1h null mutant phenotype with respect to motility points to an involvement of the non-alveolin domain of IMC1h in the motility process. We therefore postulate that domain 2 is the part of IMC1h that engages with the motility apparatus, while the alveolin module is primarily involved in filament formation, SPN recruitment and viscoelasticity. Comparative alveolin interactome studies using IMC1hΔ2, IMC1hΔ1, and IMC1h/GFP ookinetes are underway to test this hypothesis and identify candidate IMC proteins interacting with domain 2. In this context, it is also worth noting that the other two

alveolins with demonstrated roles in motility: IMC1a and IMC1b, also possess a conserved carboxy-terminal domain that has no structural relationship to the alveolin module (Khater et al., 2004; Tremp et al., 2008), and which may fulfill a similar role in motility to domain 2 of IMC1h.

Both ookinetes and sporozoites possess chirality, which is thought to be responsible for the circular nature of their directional movement (Kudryashev et al., 2012; Kan et al., 2014). During circular forward movement in Matrigel, ookinetes also rotate on their axis resulting in a helical, corkscrew-like 3D trajectory. By contrast, sporozoites possess dorsoventral polarity (Kudryashev et al., 2012) and helical movement is largely missing, explaining their predominantly circular directional movement in both 2D and 3D environments. Previous studies showed that the simultaneous absence of the alveolins IMC1h and IMC1b did not further affect ookinete cell shape, but further reduced ookinete speed and tensile strength compared with the respective single mutants, indicating that these two alveolins operate autonomously and contribute to motility independently of cell shape (Tremp and Dessens, 2011). We show here that the deletion of domains 1 or 2 from IMC1h not only reduces ookinete and sporozoite speed, but also alters their 3D movement to a markedly more linear fashion compared to that of their counterparts expressing full-length IMC1h. Nonetheless, examination of IMC1h mutant ookinetes reveals that the underlying motion remains mildly helical (**Videos S1–S3**), indicating that the biomechanics of movement have remained fundamentally the same as those of normal-shaped ookinetes. Changes in ookinete shape resulting from IMC1h knockout reduce the level of chirality and it was proposed that this determines the way in which these cells move in a 3D environment (Kan et al., 2014). The same could apply to the IMC1h mutant zoites described here. Consolidating the entire spectrum of motility observations, a model is emerging whereby zoite speed is largely cell shape-independent, whilst 3D movement is affected by the shape of the cells.

Salivary gland sporozoite numbers indicate that the infectivity of IMC1hΔ2 sporozoites to the mosquito is higher than that of IMC1hΔ1 sporozoites (**Table 1**), despite both mutant sporozoite populations possessing similar motility defects. These observations suggest that motility makes a relatively minor contribution to the sporozoites' invasive power. This is consistent with findings that IMC1h null mutant sporozoites have similar *in vitro* hepatoma cell transmigration and infection rates to their wildtype counterparts, despite having significantly reduced speed (∼ 2-fold) and more meandering movement patterns (Volkmann et al., 2012). Heat shock protein 20 null mutant sporozoites, which can move only very slowly, are also able to invade hepatocytes at the same efficiency as wild-type sporozoites, indicating that at least rapid gliding is not essential for efficient invasion (Montagna et al., 2012). The invasive power of IMC1hΔ2 sporozoites with regards to salivary glands was however not reflected in a greater infectivity to the mouse, as we were repeatedly unable to transmit these parasites by the normal route of mosquito bite. IMC1h null mutants are not naturally transmissible, but are infective to

mice when injected intravenously (Volkmann et al., 2012). These collective observations thus identify traversal of the dermis as a major bottleneck for sporozoite infection of the mammalian host.

DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

ETHICS STATEMENT

This study was carried out in accordance with the Laboratory Animal Science Association guidelines. All laboratory animal work was approved by the Animal Welfare and Ethical Review Board of the London School of Hygiene and Tropical Medicine, and by the United Kingdom Home Office. Work was carried out in accordance with the United Kingdom Animals (Scientific Procedures) Act 1986 implementing European Directive 2010/63 for the protection of animals used for experimental purposes.

AUTHOR CONTRIBUTIONS

JD and CV contributed conception and design of the study. MC, AT, SS, and JD performed experiments and interpreted results. MC wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb. 2019.00266/full#supplementary-material

Video S1 | Time lapse movie of a representative IMC1h/GFP ookinete in Matrigel.

Video S2 | Time lapse movie of a representative IMC1hΔ1 ookinete in Matrigel.

Video S3 | Time lapse movie of a representative IMC1hΔ2 ookinete in Matrigel.

REFERENCES

Al-Khattaf, F. S., Tremp, A. Z., and Dessens, J. T. (2015). *Plasmodium alveolins* possess distinct but structurally and functionally related multi-repeat domains. *Parasitol. Res.* 115, 631–639. doi: 10.1007/s00436-014-4226-9

Al-Khattaf, F. S., Tremp, A. Z., El-Houderi, A., and Dessens, J. T. (2017). The *Plasmodium alveolin* IMC1a is stabilised by its terminal cysteine motifs and facilitates sporozoite morphogenesis and infectivity in a dose-dependent manner. *Mol. Biochem. Parasitol.* 211, 48–56. doi: 10.1016/j.molbiopara.2016.09.004

Anderson-White, B. R., Ivey, F. D., Cheng, K., Szatanek, T., Lorestani, A., Beckers, C. J., et al. (2011). A family of intermediate filament-like proteins is sequentially assembled into the cytoskeleton of *Toxoplasma gondii*. *Cell Microbiol.* 13, 18–31. doi: 10.1111/j.1462-5822.2010.01514.x

Bannister, L. H., Hopkins, J. M., Fowler, R. E., Krishna, S., and Mitchell, G. H. (2000). A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages. *Parasitol. Today* 16, 427–433. doi: 10.1016/S0169-4758(00)01755-5

Barkhuff, W. D., Gilk, S. D., Whitmarsh, R., Tilley, L. D., Hunter, C., and Ward, G. E. (2011). Targeted disruption of TgPhIL1 in *Toxoplasma gondii* results in altered parasite morphology and fitness. *PLoS ONE* 6:e23977. doi: 10.1371/journal.pone.0023977

Dessens, J. T., Beetsma, A. L., Dimopoulos, G., Wengelnik, K., Crisanti, A., Kafatos, F. C., et al. (1999). CTRP is essential for mosquito infection by malaria ookinetes. *EMBO J.* 18, 6221–6227. doi: 10.1093/emboj/18.22.6221

Frenal, K., Polonais, V., Marq, J. B., Stratmann, R., Limenitakis, J., and Soldati-Favre, D. (2010). Functional dissection of the apicomplexan glideosome molecular architecture. *Cell Host Microbe* 8, 343–357. doi: 10.1016/j.chom.2010.09.002

Goodenough, U., Roth, R., Kariyawasam, T., He, A., and Lee, J. H. (2018). Epiplasts: membrane skeletons and epiplastin proteins in euglenids, glaucophytes, cryptophytes, ciliates, dinoflagellates, and apicomplexans. *mBio* 9:18. doi: 10.1128/mBio.02020-18

Gould, S. B., Tham, W. H., Cowman, A. F., Mcfadden, G. I., and Waller, R. F. (2008). Alveolins, a new family of cortical proteins that define the protist infrakingdom Alveolata. *Mol. Biol. Evol.* 25, 1219–1230. doi: 10.1093/molbev/msn070

Han, Y. S., Thompson, J., Kafatos, F. C., and Barillas-Mury, C. (2000). Molecular interactions between *Anopheles stephensi* midgut cells and *Plasmodium berghei*: the time bomb theory of ookinete invasion of mosquitoes. *EMBO J.* 19, 6030–6040. doi: 10.1093/emboj/19.22.6030

Hu, K., Mann, T., Striepen, B., Beckers, C. J., Roos, D. S., and Murray, J. M. (2002). Daughter cell assembly in the protozoan parasite *Toxoplasma gondii*. *Mol. Biol. Cell* 13, 593–606. doi: 10.1091/mbc.01-06-0309

Kan, A., Tan, Y. H., Angrisano, F., Hanssen, E., Rogers, K. L., Whitehead, L., et al. (2014). Quantitative analysis of *Plasmodium* ookinete motion in three dimensions suggests a critical role for cell shape in the biomechanics of malaria parasite gliding motility. *Cell Microbiol.* 16, 734–750. doi: 10.1111/cmi. 12283

Kaneko, I., Iwanaga, S., Kato, T., Kobayashi, I., and Yuda, M. (2015). Genome-wide identification of the target genes of AP2-O, a *Plasmodium* AP2-family transcription factor. *PLoS Pathog.* 11:e1004905. doi: 10.1371/journal.ppat.1004905

Khater, E. I., Sinden, R. E., and Dessens, J. T. (2004). A malaria membrane skeletal protein is essential for normal morphogenesis, motility, and infectivity of sporozoites. *J. Cell Biol.* 167, 425–432. doi: 10.1083/jcb.200406068

Kono, M., Prusty, D., Parkinson, J., and Gilberger, T. W. (2013). The apicomplexan inner membrane complex. *Front. Biosci.* 18, 982–992. doi: 10.2741/4157

Kudryashev, M., Munter, S., Lemgruber, L., Montagna, G., Stahlberg, H., Matuschewski, K., et al. (2012). Structural basis for chirality and directional motility of *Plasmodium* sporozoites. *Cell Microbiol.* 14, 1757–1768. doi: 10.1111/j.1462-5822.2012.01836.x

Mann, T., and Beckers, C. (2001). Characterization of the subpellicular network, a filamentous membrane skeletal component in the parasite *Toxoplasma gondii*. *Mol. Biochem. Parasitol.* 115, 257–268. doi: 10.1016/S0166-6851(01)00289-4

Menke, A., and Jockusch, H. (1991). Decreased osmotic stability of dystrophin-less muscle cells from the mdx mouse. *Nature* 349, 69–71. doi: 10.1038/349069a0

Montagna, G. N., Buscaglia, C. A., Munter, S., Goosmann, C., Frischknecht, F., Brinkmann, V., et al. (2012). Critical role for heat shock protein 20 (HSP20) in migration of malarial sporozoites. *J. Biol. Chem.* 287, 2410–2422. doi: 10.1074/jbc.M111.302109

Moon, R. W., Taylor, C. J., Bex, C., Schepers, R., Goulding, D., Janse, C. J., et al. (2009). A cyclic GMP signalling module that regulates gliding motility in a malaria parasite. *PLoS Pathog.* 5:e1000599. doi: 10.1371/journal.ppat.1000599

Morrissette, N. S., and Sibley, L. D. (2002). Cytoskeleton of apicomplexan parasites. *Microbiol. Mol. Biol. Rev.* 66, 21–38. doi: 10.1128/MMBR.66.1.21-38.2002

Parkyn Schneider, M., Liu, B., Glock, P., Suttie, A., Mchugh, E., Andrew, D., et al. (2017). Disrupting assembly of the inner membrane complex blocks *Plasmodium falciparum* sexual stage development. *PLoS Pathog.* 13:e1006659. doi: 10.1371/journal.ppat.1006659

Santos, J. M., Lebrun, M., Daher, W., Soldati, D., and Dubremetz, J. F. (2009). Apicomplexan cytoskeleton and motors: key regulators in morphogenesis, cell division, transport and motility. *Int. J. Parasitol.* 39, 153–162. doi: 10.1016/j.ijpara.2008.10.007

Tremp, A. Z., Al-Khattaf, F. S., and Dessens, J. T. (2014). Distinct temporal recruitment of *Plasmodium alveolins* to the subpellicular network. *Parasitol. Res.* 113, 4177–4188. doi: 10.1007/s00436-014-4093-4

Tremp, A. Z., Al-Khattaf, F. S., and Dessens, J. T. (2017). Palmitoylation of *Plasmodium alveolins* promotes cytoskeletal function. *Mol. Biochem. Parasitol.* 213, 16–21. doi: 10.1016/j.molbiopara.2017.02.003

Tremp, A. Z., Carter, V., Saeed, S., and Dessens, J. T. (2013). Morphogenesis of *Plasmodium zoites* is uncoupled from tensile strength. *Mol. Microbiol.* 89, 552–564. doi: 10.1111/mmi.12297

Tremp, A. Z., and Dessens, J. T. (2011). Malaria IMC1 membrane skeleton proteins operate autonomously and participate in motility independently of cell shape. *J. Biol. Chem.* 286, 5383–5391. doi: 10.1074/jbc.M110.187195

Tremp, A. Z., Khater, E. I., and Dessens, J. T. (2008). IMC1b is a putative membrane skeleton protein involved in cell shape, mechanical strength, motility, and infectivity of malaria ookinetes. *J. Biol. Chem.* 283, 27604–27611. doi: 10.1074/jbc.M801302200

Vernick, K. D., Fujioka, H., and Aikawa, M. (1999). *Plasmodium gallinaceum*: a novel morphology of malaria ookinetes in the midgut of the mosquito vector. *Exp. Parasitol.* 91, 362–366. doi: 10.1006/expr.1998.4388

Volkmann, K., Pfander, C., Burstroem, C., Ahras, M., Goulding, D., Rayner, J. C., et al. (2012). The alveolin IMC1h is required for normal ookinete and sporozoite motility behaviour and host colonisation in *Plasmodium berghei*. *PLoS ONE* 7:e41409. doi: 10.1371/journal.pone.0041409

Waters, A. P., Thomas, A. W., Van Dijk, M. R., and Janse, C. J. (1997). Transfection of malaria parasites. *Methods* 13, 134–147. doi: 10.1006/meth.1997.0506

Chapter 5

Studies towards resolving the alveolin structure

5.1 Introduction

Thirteen distinct alveolin genes have to date been identified in *Plasmodium* species (Khater *et al.*, 2004, Al-Khattaf *et al.*, 2015, Kaneko *et al.*, 2015). In addition, two PhIL1 interacting proteins: PIP2 and PIP3, show structural homology with alveolins (Parkyn Schneider *et al.*, 2017, Kono *et al.*, 2013), and these could represent additional family members. Sequence homology between these proteins is confined to conserved modules named 'alveolin' modules that are composed of tandem repeat sequences of 12 amino acid periodicity (Al-Khattaf *et al.*, 2015). Their conservation and structural similarity makes it likely that these 'alveolin' modules hold the physical properties that facilitate the putative intermediate filament (IF)-forming nature of the alveolins.

In higher eukaryotes, IF proteins such as lamins, vimentins and α-keratins have an underlying architecture that includes a helical rod domain that forms coiled-coils, which is thought to be fundamental to their filament-forming properties. Coiled-coils are typically formed by tandem repeat sequences of seven amino acid periodicity that adopt an α-helical fold. The interesting parallel of having tandem repeat sequences in both alveolins and metazoan IF proteins raises the possibility that these molecules could be variations on the same theme, and that alveolins adopt similar α-helical configurations to allow filament formation, despite the apparent differences in periodicity. Arguing against this concept is the fact that the 'alveolin' modules are relatively rich in proline residues (Al-Khattaf *et al.*, 2015). Proline is a cyclic and bulky residue known to disrupt α helices and for this reason this amino acid is largely absent from the α-helical domains of metazoan IF proteins. Finally, despite strong circumstantial evidence (Mann and Beckers, 2001) there is to date no direct experimental evidence that alveolins actually form filaments.

Whilst alveolins have been characterised through affinity purification assays (Bullen *et al.*, 2009), they have yet to be characterised for their structures using biophysical techniques such as X-ray crystallography or electron microscopy. As alveolins are relatively uncharacterised in this regard, there is scope to understand how the protein behaves in an environment where it is overexpressed. Naturally aggregating proteins have been well characterised before for their structures, being expressed in highly denaturing conditions that prevent the protein aggregation and becoming insoluble (Nicolet *et al.*, 2010, Herrmann and Aebi, 2016).

To address some of these questions, structural data of the alveolins was set out to be obtained through a programme of recombinant protein expression, purification and ultimately structural analysis, the results of which so far are compiled in this chapter.

5.2 Materials and Methods

5.2.2 Plasmid constructs

To generate pGEX4T-1 constructs, alveolin sequences were PCR amplified from *P. berghei* genomic DNA with specific primer pairs (Table 5.1) and introduced into *Bam*HI-digested pGEX4T-1 by in-fusion cloning.

To generate pETM6T1 constructs, alveolin sequences were PCR amplified from *P. berghei* genomic DNA with specific primer pairs (Table 5.2) and introduced into *Bam*HI/*Xho*I-digested pETM6T1 by in-fusion cloning.

Table 5.1. PCR primer sequences used to amplify alveolin-specific sequences for cloning into pGEX4T-1.

Region amplified	Forward Primer	Reverse Primer
IMC1c	GGTTCCGCGTGGATCAATGGA TAATTCACATTTAAGCAGAAA	GGAATTCCGGGGATCCTTATC TGCATGTACCTGTACAGCA
IMC1c-N	GGTTCCGCGTGGATCAATGGA TAATTCACATTTAAGCAGAAA	GGAATTCCGGGGATCTTACCA ATTTATGTTTTTGGAGTCTCT T
IMC1g	GGTTCCGCGTGGATCAATGTG TTCTACACCCAATAAATTAGC	GGAATTCCGGGGATCCTAGTT ACAGTAAAAACCTCTATGAAC TACATTT
IMC1g-N	GGTTCCGCGTGGATCAATGTG TTCTACACCCAATAAATTAGC	GGAATTCCGGGGATCTTATAT TTCAGGAATTTCTTTAATACT TTCTTTT

Table 5.2. PCR primer sequences used to amplify alveolin-specific sequences for cloning into pETM6T1.

Plasmid Construct	Forward Primer	Reverse Primer
IMC1c	GGGCGCCATGGGATCCATGGA TAATTCACATTTAAGCAGAAA C	GGTGGTGGTGCTCGATTATCT AGCTGTACCTGTAGCAGCTCC AACTGGTTTAGCTTC
IMC1c-N	GGGCGCCATGGGATCCATGGA TAATTCACATTTAAGCAGAAA C	GGTGGTGGTGCTCGATTATCC TGCTTTCCCTTCTGTATATT
IMC1g	GGGCGCCATGGGATCCATGGC CTCTACACCCAATAAATTAGC CGCTGCTTCTGGAGA	GGTGGTGGTGCTCGATTAGTT ACAGTAAAAACCTCTATGAAC TACATTTA
IMC1g-N	GGGCGCCATGGGATCCATGGC CTCTACACCCAATAAATTAGC CGCTGCTTCTGGAGA	GGTGGTGGTGCTCGATTAATT TGGGAAAATTATTTCAGGAAT TT
IMC1e	GGGCGCCATGGGATCCGACAA AGTCCCAGAAAATGC	GGTGGTGGTGCTCGATTATGT GTTAAAATTGTCTTCTTGATA ATATAATG
IMC1h	GGGCGCCATGGGATCCATGCA TCCAAATACTAGCGAGG	GGTGGTGGTGCTCGATTACAA ATATTGTGTCGGTTTTAATGG

A glycine-serine linker between the TEV cleavage site and the alveolin sequence was introduced by PCR-based site directed mutagenesis (see Chapter 2). pETM6T1-IMC1e plasmids were PCR amplified with primers SDM glycine linker-F (AATCTTTATTTTCAGGGTTCGGGTTCGGGTTCGAGCGGTTCGGGCGCCATGGGATCC) and SDM glycine linker-R (CTGAAAATAAAGATTCTCACCCGC), and the amplicon circularized by In-Fusion, to give plasmid pETM6T1-NusA-GS-IMC1e. This changes the amino acid sequence from …NLYFQ/GAMGS… to …NLYFQ/GSGSGSSGSGAMGS… (/ denotes TEV protease cleavage junction).

5.2.5 Chromatography calibration curve

A size exclusion chromatography gel calibration curve was made by running using standard molecular weight markers blue dextran (M_r 2,000,000), bovine serum

albumin (M_r 66,463), chymotrypsinogen (M_r 25,000) and ribonuclease A (M_r 13,700) through the size exclusion column. The *Kav* value was calculated using the equation: $Kav = (Ve - Vo)/(Vt - Vo)$, where *Ve* is equal to the elution volume, *Vo* is equal the void volume, and *Vt* is equal the total volume held by the column. The void volume is the volume at which blue dextran runs off the column.

5.3 Results

5.3.1 Structure predictions

Bioinformatics tools were used to make predictions about the structure of alveolins. The programme PrDOS was used to look at the disorder probability of the alveolins, a measure of the conformational flexibility of proteins and their domains. This revealed that the conserved alveolin domains are predicted to have an overall ordered structure, whilst the non-conserved regions upstream and downstream have disordered structures, as is illustrated for IMC1b (Fig. 5.1A). Similar predictions were made for other alveolin family members (data not shown). This indicates that the conserved modules are conformationally stable, which is consistent with the notion that they constitute the functional domains of the protein.

The programme Quick2D (MPI Bioinformatics Toolkit) uses different algorithms to make secondary structure predictions of the alveolins. As illustrated for IMC1b, alveolin domains 1 and 2 are predicted to have a predominantly β-strand structure, whilst its non-alveolin domain 3 is predicted to be mostly α-helical (Fig. 5.1B). Similar secondary structure predictions were also done at a late stage for IMC1e, IMC1c and IMC1g, but due to a changed algorithm repertoire used by Quick2D, only PSIPRED consistently predicted β-strand fold for the conserved domains within these alveolins (Fig. 5B-E). This highlights the unpredictable nature of secondary structure prediction algorithms. Nonetheless, these findings could suggest that the alveolin modules adopt a β-pleated fold rather than the α-helical fold.

IMC1e, IMC1c and IMC1g were all chosen to be recombinantly expressed due to their different properties. IMC1e is the most representative alveolin of the family, IMC1c is a

small alveolin that is found in all three motile zoite stages of the malaria life cycle, and

IMC1g is similar in both being a small alveolin being expressed in all three zoite stages.
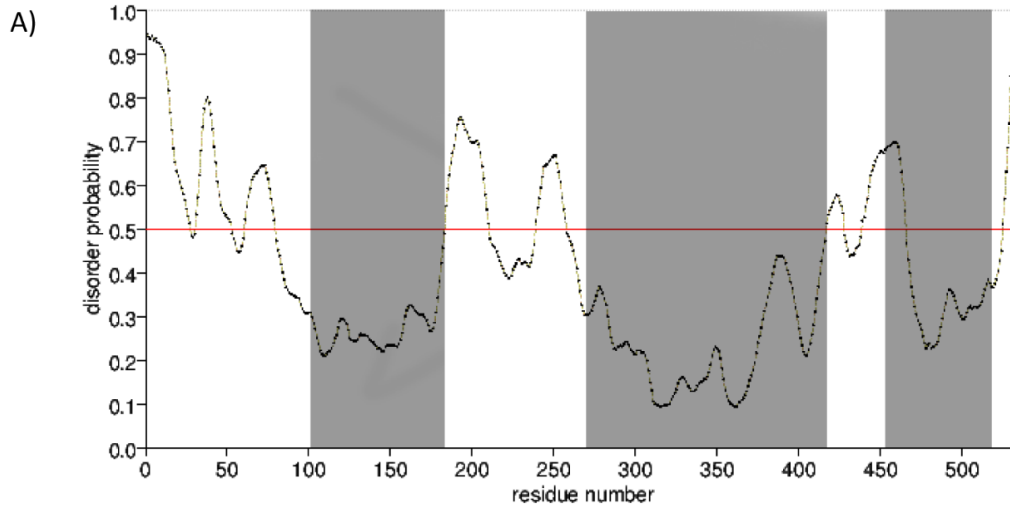
C)

```
AA_QUERY      1    MKNNCGEGSTYNNQGYKFCETSNCGTKYVDQNGKTKSQGDQQYNIMDNIFLKESNNGNDNNYLDKVPENARMLKPLVQEKIIEIM    85
SS_PSIPRED                                             EEE     EEEEE EEEE            EE    EE   EEEEE EEEE
SS_SPIDER3
SS_PSSPRED4                     EEEE        E          HHHHHHHEEEE                HHHHHHHHHHHHHHHHHHHH

AA_QUERY     86    KPEIEEKIIEVPQIQYVEKLVEVPHVILQEKLIHVPKPVIHERIKKCPKTIFQEKIVEVPQIKIVDKIIEVPQYVYQEKIIQVPK   170
SS_PSIPRED          EEEEEEE       EEE   EE      EEE         EEEEEEEE   EEEEEE EE   EEEEEEEEEE   EEEE
SS_SPIDER3                                                                     E      EEEEEEEE
SS_PSSPRED4        HHHHHHHH  HHHHHHHHHHH    HHHHHHHHHH    HHHHHHHHH  HHHHHHHHHH  EHHHHHHHHHHHHHHHHHHHHH  H

AA_QUERY    171    VMVQERIIPVPKKVIKEKIVEIPQIELKNINIEKIEEIPEYIPEVVKRDVPYTQIIDRPFHVEKIVEVPHVQHIYRNIVSPQYRH   255
SS_PSIPRED        EEEEEEE     EEEEEEEE    EEE EEEEEEE      EEE      EEEE     EEEE  EEEEEEE  EE    EE
SS_SPIDER3        EEEE          EEEEEEE  EEEEEEE EEEEEE      EEEE          EEEEE
SS_PSSPRED4       HHHHH       HHHHHHHHH    EEE    HHHHHHHHHHHHHHHHH              EEEEE    HHHHHHHH  HHH

AA_QUERY    256    IPKPVEVPMAHYRTFPVEKLVDRNVPVPVEIQIVQEFLCPKIEARYKEIPVPVHVQRIIEHPIPKDAMNNPFLLPLYYQEDNFNT   340
SS_PSIPRED        E          EEE   EEE            EEEEE   EEE EEE        EEE                     EEE
SS_SPIDER3
SS_PSSPRED4                  HHHHH        EEEEEHHHH HHHHHHH      HHHHHHHH          EEEEEE
SS_NETSURFP2      E     EE   EEEEEEEEEEEEEEEEE  EEEEEEEEEEEE  EEEEEEEEE   EEEEEEEEEEEEEHHH       EEEE

AA_QUERY    341    SKNSTVNGSNTPQNNKGCFMFNWNKNHNQKTMKSRNNSVELLLHNDNKNNEFFNINEANINPNINLKTSINQYSGNTFRDINNIP   425
SS_PSIPRED            EE                                  HHHHH
SS_SPIDER3                                                EEEE
SS_PSSPRED4               EEEE      H HHH     EEEEEE                    EE

AA_QUERY    426    MDANIMNMHALGQCQSINPIDNSNNQYNFIHNPAAMNYSSPNFNSEIPYNPNASLKISNSMSYPMESIPLKKPYPIVSHIVPPVL   510
SS_PSIPRED
SS_SPIDER3
SS_PSSPRED4            HHHH                                          EEEE

AA_QUERY    511    NE   512
SS_PSIPRED
SS_SPIDER3
SS_PSSPRED4
```

D)

```
AA_QUERY      1    MDNSHLSRNSYQKIDNIDTKETSTVDRKWVALTAYQPVDVVTKTVEVPVIKTVEKFVPKTIIQEKIIHVPKNVTHIVEKIVEVPE    85
SS_PSIPRED                     EEEEE     EEE EE    EEEEEEEE   EEEEE EE    EEE EEEE    EEE    EE
SS_SPIDER3                                                           EE
SS_PSSPRED4              HHH           EEEEEEE    EEEEEEEEEEE EEEHHHH   HHHHHHHHHH    HHHHHHHHHHH

AA_QUERY     86    VKYIEKVVEVPHIHYKNKYVPKIEVVEKIVERQKIIEKWHDKIVEVPQIKEVVRYKEIEDTEEVIKYVPRDSKNINWEEEYKKYT   170
SS_PSIPRED        EE   EEE    EEE EE   EEEEEEEEEEEEEE           EEEEEEEE     EEEEEE            EE
SS_SPIDER3                              EEEEE                                EEEEEE        EEEEEEE
SS_PSSPRED4       HHHHHHHHH    EE       HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH    HHHHHH       HHHHHHHHH

AA_QUERY    171    EGKAGKFSLNNTYQQQMNSYNQINENAYARNSHSRSINVMNSKTANDQTNMRSEDFSQINFFNQYNGESMDQHIMMPSSSFNMNG   255
SS_PSIPRED
SS_SPIDER3
SS_PSSPRED4            EE    HHHHHHHHHHHHHHHHHHH      EEEEE             HHHHHHH      EE

AA_QUERY    256    SFQFKRLPSEEAKPVGCCTGTCR    278
SS_PSIPRED
SS_SPIDER3           H
SS_PSSPRED4       EEEEE
```
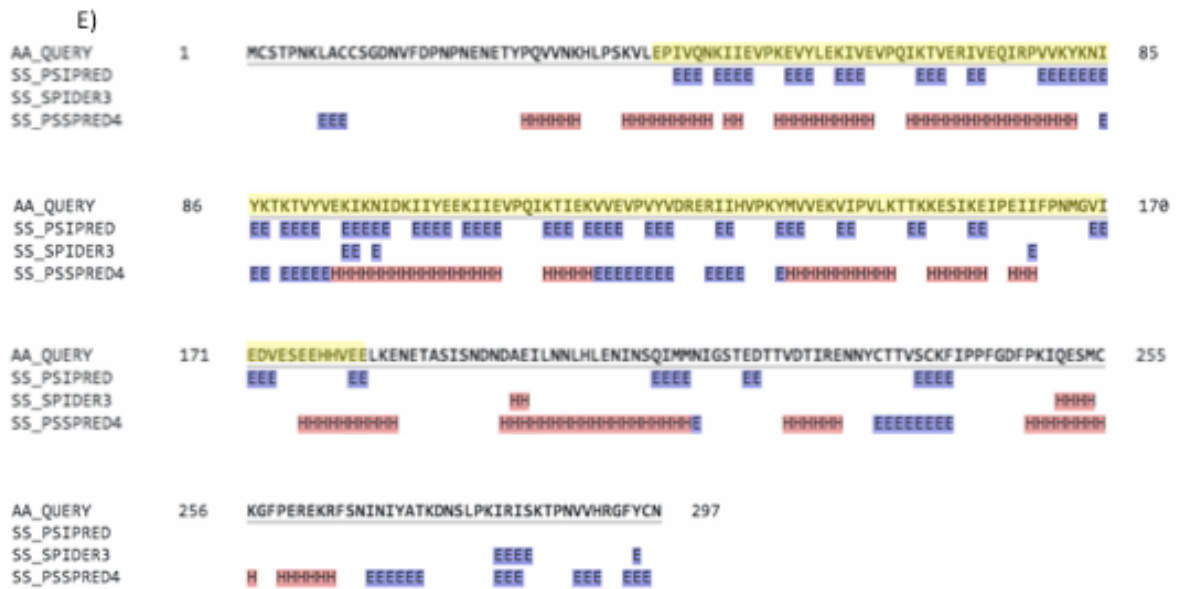
Fig. 5.1 (Pages 80 - 82). Predictions of the structure of the *P. berghei* alveolin IMC1b. A) Protein disorder prediction using PrDOS, indicating that its three conserved domains (grey blocks) (Tremp *et al.*, 2008) positively correlate with conformational stability. B) Secondary structure predictions of IMC1b using Quick2D, indicating a predominantly β-strand structure (blue E) of domains 1 and 2, and a chiefly α-helical structure (red H) of domain 3. C) Secondary structure predictions of IMC1e. D) Secondary structure predictions of IMC1c. E) Secondary structure predictions of IMC1g. The domains are defined either by sequence conservation (yellow) (Tremp *et al.*, 2008), or by the presence of tandem repeat structures (green) (Al-Khattaf *et al.*, 2015).

5.3.2 Recombinant protein expression

To bacterially express recombinant alveolins, or alveolin domains, we started using the pGEX4T system. This system expresses recombinant proteins in *E. coli* as fusions with glutathione S-transferase (GST) at the N-terminus. Expression of recombinant proteins is driven by the IPTG-inducible *tac* promoter. Subsequently, GST fusion proteins can be purified by immobilization on agarose beads bound to reduced glutathione, and the GST portion can be cleaved with thrombin. We started out making constructs for GST fusions with the alveolins IMC1g and IMC1c, as well as the amino terminal halves of IMC1c (IMC1c-N) and IMC1g (IMC1g-N) that contain the predicted 'alveolin' modules. These two alveolins were chosen as they are the smallest members of the *Plasmodium* alveolin family (Al-Khattaf *et al.*, 2015).

PCR amplification of the various parts of *imc1c* and *imc1g* with the specific primer pairs (Table 5.1) gave rise to DNA fragments of the expected sizes (Fig. 5.2A). In-Fusion cloning of these fragments into *Bam*HI-digested pGEX4T-1 was verified by digestion with *Hinc*II, showing that all fragments were successfully cloned into pGEX4T-1 (Fig. 5.2B). Plasmids were then transformed into *E. coli* BL21 and protein expression in small scale (10ml) cultures was induced with 100µM IPTG for 3 hours at 37°C. SDS-PAGE analysis of total cell lysates revealed that GST fusions with the N-terminal halves of IMC1g or IMC1c were expressed above background levels, giving rise to proteins of the expected sizes (45kDa GST::IMC1g-N; 46kDa GST::IMC1c-N) (Fig. 5.2C). By contrast, full-length IMC1g fused to GST was not expressed at discernible levels (Fig. 5.2C) and the same was true for full-length GST::IMC1c (data not shown). However, comparing soluble and insoluble cell fractions revealed that solubility of the GST fusion proteins was poor, with the majority of recombinant protein present in insoluble form (data not shown).
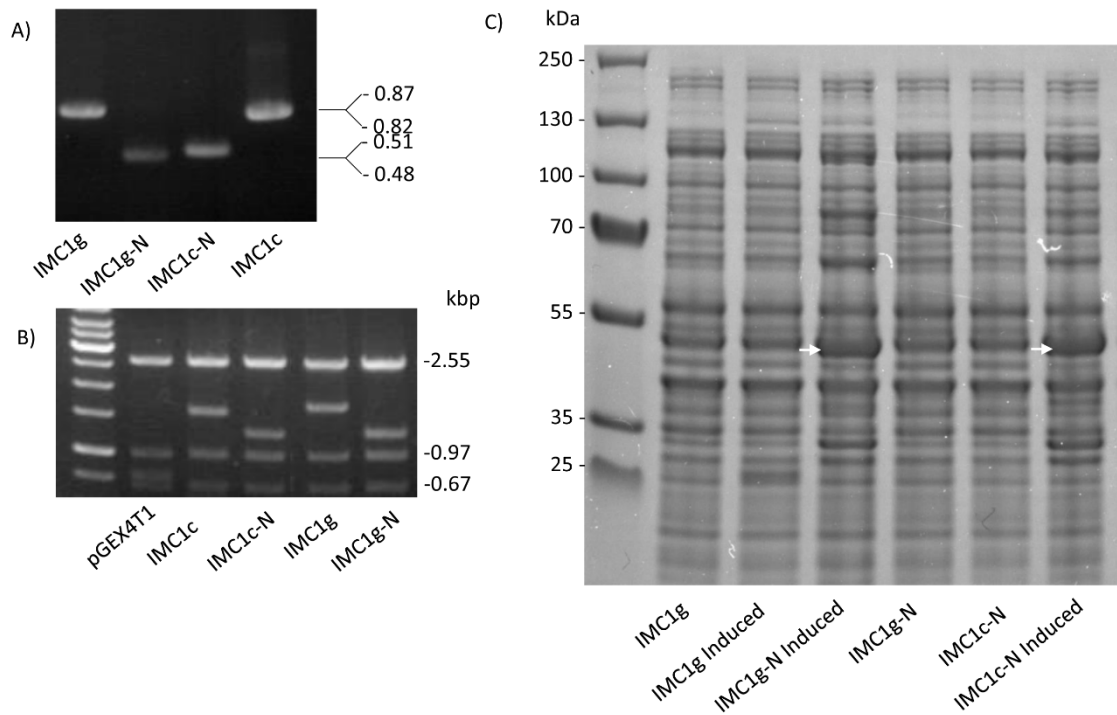
Fig. 5.2. Plasmid construction and alveolin protein expression using the pGEX4T system. A) PCR amplification of the DNA sequences encoding *P. berghei* IMC1c, IMC1g, and the N-terminal portions of IMC1c (IMC1c-N) and IMC1g (IMC1g-N) using specific primers as shown in Table 5.1. Amplicon sizes are indicated on the right hand side. B) Diagnostic digest with *Hinc*II shows successful introduction of the PCR fragments into *Bam*HI-digested pGEX4T-1. This digest gives rise to three fragments of invariable length (sizes indicated on right hand side) and one fragment of variable size corresponding to the insert. 'Empty' pGEX4T-1 is shown as the reference. C) Protein expression of the GST::alveolin proteins. Total bacterial cell lysates before and after induction with 100mM IPTG for 3h at 37°C were fractionated by SDS-PAGE and gels stained with Coomassie brilliant blue. The position of the recombinant fusion proteins are indicated with white arrows. Protein molecular weight markers (kDa) are shown and sizes indicated on the left hand side.

As an alternative to the pGEX system, the pETM6T1-based bacterial expression system was used. The latter allows expression of recombinant proteins as carboxy-terminal fusions to N-utilisation substance protein A (NusA), a fusion tag that is well documented to confer stability and increases solubility of its bound protein (De Marco *et al.*, 2004, Dummler *et al.*, 2005, Turner *et al.*, 2005). The pET system drives expression from the bacteriophage T7 promoter by an IPTG-inducible T7 RNA polymerase that is encoded by the bacterial host. The NusA tag contains a 6×His tag and a TEV protease cleavage site, to release and remove the tag from the fusion protein after expression. For this expression system, we also included the central alveolin domain of IMC1e, as this alveolin is the most similar to the *Plasmodium*

alveolin population as a whole and thus arguably most representative of the family (Al-Khattaf *et al*., 2015). The conserved alveolin domains from IMC1e, IMC1c and IMC1g, as well as the entire IMC1c and IMC1g proteins, were successfully PCR amplified from *P. berghei* genomic DNA (Fig. 5.3A, not shown for IMC1e) using specific primer pairs (Table 5.2). The resulting fragments were successfully introduced into *Bam*HI/*Xho*I digested pETM6T1 as determined by diagnostic digests with *Eco*RV (Fig. 5.3B). The resulting plasmids were transformed into *E. coli* BL21(DE3). Initial expression studies were carried out on small scale cultures induced with 100μM IPTG for 3 hours at 37°C. This resulted in considerably higher expression levels of the expected recombinant fusion proteins than was observed in the pGEX system (Fig. 5.3C). Furthermore, comparing soluble and insoluble cell fractions revealed that solubility of the NusA fusion proteins was much high under the expression conditions used (at least 50%) (Fig. 5.3D). Due to this superior performance of the pETM6T1 system, all ensuing experiments were carried out with this setup.
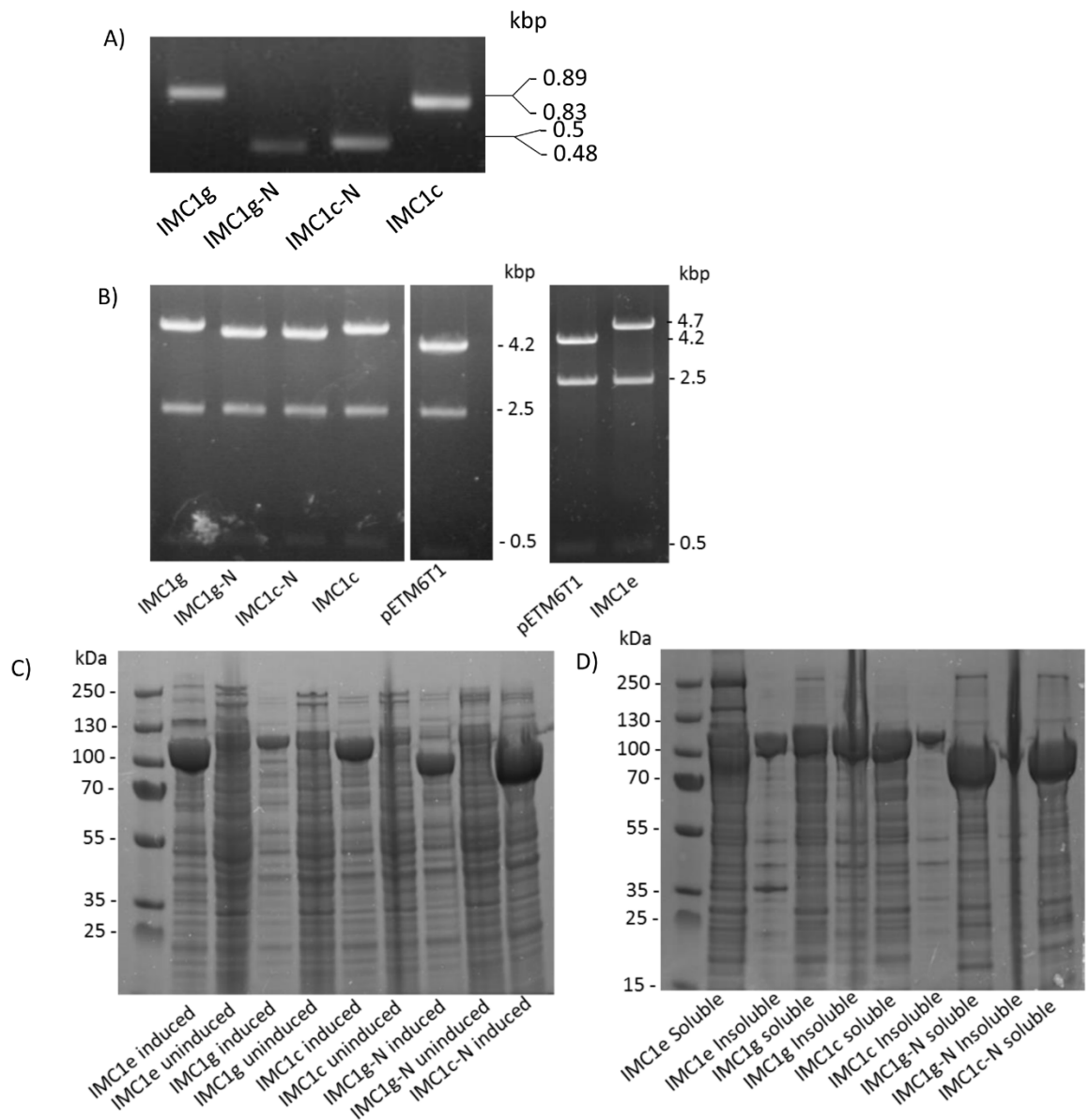
Fig. 5.3. Plasmid construction and alveolin protein expression using the pETM6T1 system. A) PCR amplification of the DNA sequences encoding *P. berghei* IMC1c, IMC1g, and the N-terminal portions of IMC1c (IMC1c-N) and IMC1g (IMC1g-N) using specific primers as shown in Table 5.2. Amplicon sizes are indicated on the right hand side. B) Diagnostic digest with *Eco*RV shows successful introduction of the PCR fragments into *Bam*HI/*Xho*I-digested pET6T1. This digest gives rise to two fragments of invariable length (2.5 and 0.5kb) and one fragment of variable size corresponding to the insert (sizes indicated on right hand side). 'Empty' pETM6T1 is shown as the reference. C) Protein expression of the NusA::alveolin proteins. Total bacterial cell lysates before and after induction with 100mM IPTG for 4h at 37°C were fractionated by SDS-PAGE and gels stained with Coomassie brilliant blue. Protein molecular weight markers (kDa) are shown and sizes indicated on the left hand side. D) Equivalent soluble vs insoluble fractions of bacterial lysates after IPTG induction. Protein molecular weight markers (kDa) are shown and sizes indicated on the left hand side.

5.3.2 Purification of IMC1c-N

After small scale cultures determined that the NusA expression system was efficient, large scale cultures were used to increase recombinant protein yields. This was first carried out with IMC1c-N, which corresponds to one of the smaller alveolin domains. After recombinant protein expression, bacterial cells were lysed and ultracentrifuged to remove insoluble material. The lysate was then incubated with TALON metal affinity resin (i.e. cobalt ions immobilized on agarose beads) binding the recombinant NusA::alveolin fusion protein (via its 6×His-tag) and allowing non-target bacterial proteins to be washed away. The part-purified NusA::alveolin fusion protein was then eluted off the TALON batch column with imidazole-containing buffer, and incubated with TEV protease to cleave off the NusA tag. The latter step was carried out at 4°C overnight, whilst dialysing to remove excess imidazole from the solution. In this experiment, the TEV protease efficiently cleaved the fusion protein into its individual NusA and IMC1c-N components (Fig. 5.4A).

The cleaved solution was further fractionated on a size exclusion column, with the majority of the alveolin protein eluting in fractions 15-17 (Fig. 5.4A). Unexpectedly, the latter protein partially co-eluted with the much larger NusA protein (Fig. 5.4A). This observation suggested that the alveolin domain was not behaving as a globular protein and therefore eluted earlier than expected, together with the globular NusA protein.

Fractions 15-17 were pooled, and the alveolin protein was successfully purified from the NusA protein using a HiTrap Q ion exchange column: a pure alveolin fraction was present in the HiTrap Q wash through, while NusA and other proteins eluted as conductivity of the column increased (Fig. 5.4B).

To further purify and concentrate the IMC1c-N protein, another size exclusion column was used. Unexpectedly, the protein did not elute in a fraction expected for a 20kDa protein, instead eluting in fractions 16-18, much closer to the void (Fig. 5.4C). A size exclusion chromatography gel calibration curve was made to determine what molecular weight protein would be eluted in fractions 16-18 (Fig. 5.4D). This indicated that these fractions should contain protein of molecular weight closer to 80kDa, suggesting that the IMC1c-N protein could be forming tetramers.
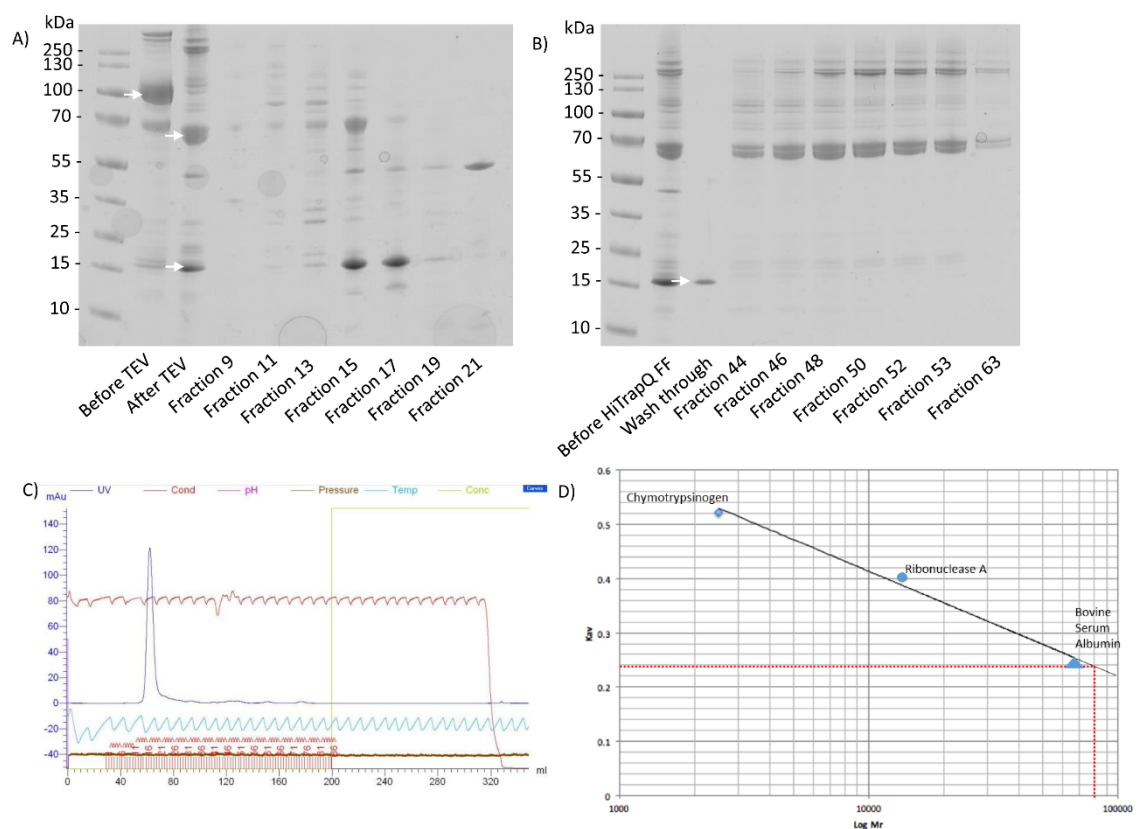

Fig. 5.4. Purification of bacterially expressed IMC1c-N. A) Coomassie-stained protein gel showing efficient TEV protease cleavage of NusA::IMC1c-N fusion protein (lane marked Before TEV, white arrow) into NusA and IMC1c-N (lane marked After TEV, white arrows). Other lanes marked Fraction 9 - Fraction 21 show elution of the alveolin portion in Fractions 15-17 from a size exclusion column. Protein molecular weight markers (kDa) are shown and sizes indicated on the left hand side. B) Coomassie-stained protein gel showing the presence of the IMC1c-N in the wash through of a HiTrap Q anion exchange column (lane marked Wash through). NusA elutes as conductivity of the column increases (lanes marked Fraction 44 – Fraction 63). Protein molecular weight markers (kDa) are shown and sizes indicated on the left hand side. C) Chromatogram of a size exclusion column showing that the IMC1c-N protein elutes in fractions 16-18. D) Size exclusion chromatography calibration curve made using Chymotrypsinogen, Ribonuclease A and Bovine aerum albumin, showing that fractions 16-18 contain protein with M_r close to 80kDa.

To further concentrate the protein for downstream analysis, the protein solution was centrifuged through a protein concentrator with a 10kDa molecular weight cut-off. During this process, the protein became insoluble and precipitated at a concentration of approximately 2.5mg/mL. The amount of protein left in solution was approximately 0.22mg/mL, but this was too low to conduct further downstream structural analyses. Subsequent repeats of these protocols using construct IMC1c-N failed to reproduce the results. Specifically, in these repeat experiments the target protein already precipitated during the TEV protease cleavage step.

5.3.4 Purification of IMC1e

Given the problems encountered with IMC1c-N, we switched our attention to construct pETM6T1-IMC1e, expressing NusA fused to the alveolin domain of IMC1e. IMC1e is structurally most similar to the *Plasmodium* alveolin family, but its alveolin domain is larger than that of IMC1c (Al-Khattaf *et al.*, 2015). The NusA::IMC1e fusion protein was successfully expressed in large scale culture and partly purified using the TALON metal affinity resin (Fig. 5.5A). However, in contrast to NusA::IMC1c-N, the NusA::IMC1e fusion protein was poorly cleaved with the TEV protease and no discernible cleavage product corresponding to the alveolin portion was detected (Fig. 5.5A).

TEV protease cleavage is generally very efficient. One possible explanation for the poor cleavage efficiency on this occasion was that the IMC1e domain was too close to the TEV cleavage site and its specific folding might prevent access to the cleavage site by the TEV protease. In this context, it is noteworthy that the NusA::IMC1c-N construct contains some 20 amino acids corresponding to the IMC1c amino terminus in between the TEV cleavage site and the predicted start of the IMC1c alveolin domain (Tremp *et*

al., 2014). For this reason a 10 amino acid glycine-serine linker was introduced immediately downstream of the TEV cleavage site in the NusA::IMC1e-encoding pETM6T1 plasmid (resulting in plasmid pETM6T1-NusA-GS-IMC1e).

The modified NusA::IMC1e fusion protein was again successfully expressed and part purified on TALON resin (Fig. 5.5B). Overnight TEV protease treatment in solution resulted in complete loss of the fusion protein and the concurrent appearance of its cleavage products corresponding to NusA and the IMC1e alveolin portion. This indicated that the addition of the glycine-serine linker between the TEV protease cleavage site and the alveolin domain had indeed made the cleavage site better accessible to the protease. The IMC1e portion migrated at a position smaller than expected of its calculated mass (32kDa). In contrast to off-column cleavage, TEV protease cleavage on the column (*i.e.* in the presence of the TALON resin) was much less efficient and only resulted in partial cleavage (Fig. 5.5B). In fact, the latter was also observed for on the column cleavage of the NusA::IMC1c-N protein (data not shown). Efforts to further purify and concentrate the recombinant IMC1e alveolin domain are ongoing. However, these efforts are being hindered by protein precipitation during the TEV protease cleavage step, similar to what was observed with the NusA::IMC1c-N fusion protein.
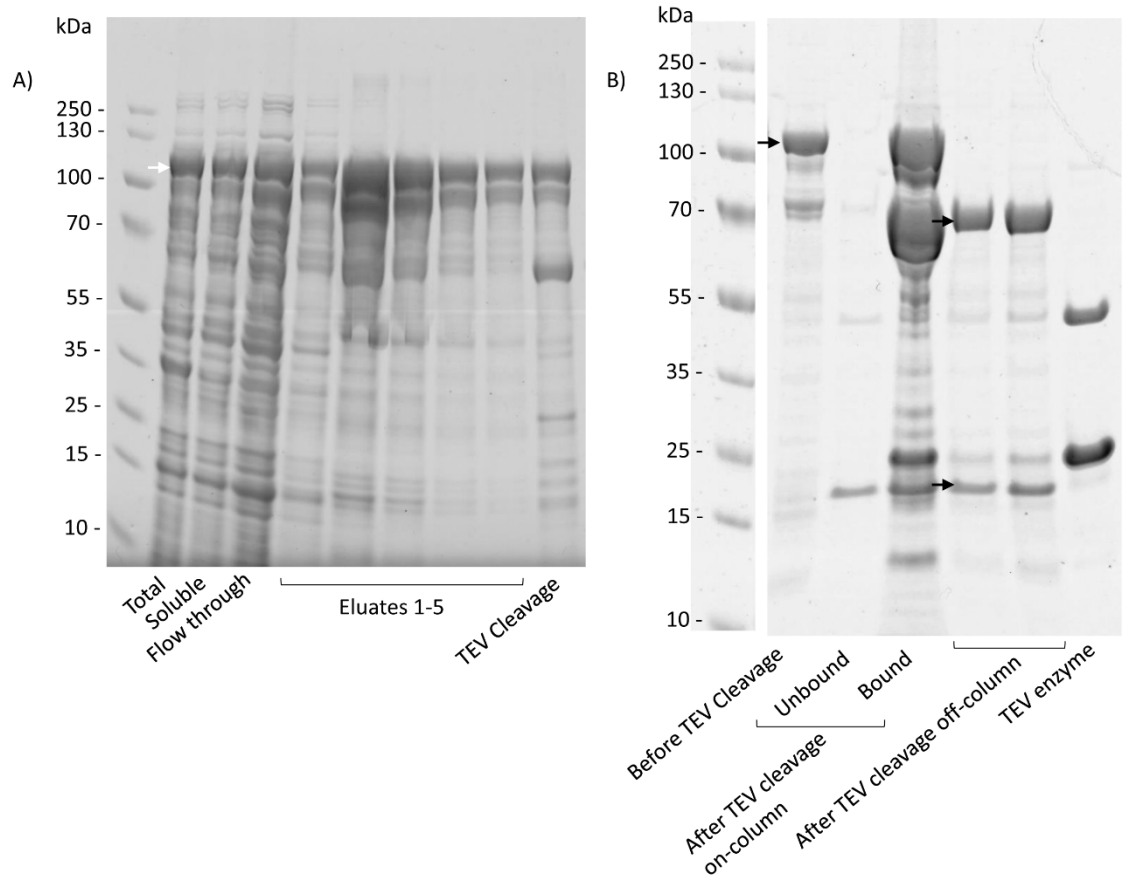
Fig. 5.5. Purification of bacterially expressed IMC1e. A) Coomassie-stained protein gel showing bacterial NusA::IMC1e expression (lane marked Total, white arrow), solubility (lane marked Soluble), part purification on TALON resin (lanes marked Eluates 1-5), but inefficient TEV protease cleavage (lane marked TEV cleavage). Protein molecular weight markers (kDa) are shown and sizes indicated on the left hand side. B) Coomassie-stained protein gel showing efficient expression and TALON purification of NusA::IMC1e with glycine-serine linker (lane marked Before TEV cleavage), and efficient TEV protease cleavage in solution (lanes marked After TEV cleavage off-column). Black arrows mark positions of the fusion protein and cleavage products. The TEV protease cleavage is much less efficient on the column (lanes marked After TEV cleavage on-column). A sample of the TEV protease is also shown (lane marked TEV enzyme). Protein molecular weight markers (kDa) are shown and sizes indicated on the left hand side.

5.4 Discussion

Secondary structure predictions (Fig. 5.1) suggest that the *Plasmodium* alveolin domains are conformationally stable and adopt a β-pleated structure. Indeed, the latter is corroborated by a recent study predicting that alveolins and related epiplastin proteins across different radiations adopt an overall β-strand fold (Goodenough *et al*., 2018). Although not common, there are examples where β-strands are involved in filament formation, most notable members of the β-keratin family in birds and reptiles (Fraser and Parry, 2009). In these proteins, dimers are assembled via interactions between short β-sheet domains within monomers that are rich in valine, isoleucine and proline residues. These protofilaments polymerase to form long filaments 3-4nm in diameter that possess viscoelastic properties and are detergent-insoluble (Fraser and Parry, 2009). It is compelling that many of these features are shared with alveolins and it is tempting to speculate that alveolins could form filaments through a similar mechanism. However, this can only be confirmed using biophysical techniques. For example, size exclusion chromatography with multiple angle light scattering (SEC-MALS) will generate data on the exact molecular weight of the protein in question and whether it forms oligomers of any kind with itself. Circular dichroism spectroscopy (CDS) will give an insight into whether the protein solution contains mainly α-helices or β-strands. X-ray crystallography or electron microscopy can be applied to shed light on the 3D atomic structure of the protein under investigation.

To facilitate such structural biological studies, a program of recombinant alveolin expression and purification was embarked on. Expression and purification of recombinant proteins are two important and vital aspects of protein structural biology and each presents its own challenges. For example, expression of some proteins within

bacteria can have detrimental effects and lead to toxic build-up of proteins within the cell that may interrupt vital cellular processes. Some proteins need very specific pH and temperatures for expression and can only be expressed in low amounts before aggregating within the cell. Some proteins need extensive post-translational modifications that make bacterial cell cultures unsuitable, and there are many other conditions that may need to be met in order to ensure a protein is stable enough to extract from a cell. Purification presents other challenges, for example with respect to protein stability and solubility.

The results presented in this chapter show that the bacterial expression system, as well as the conditions for TEV protease cleavage, were successfully optimised. Expression of the alveolins with an N-terminal NusA tag clearly works much better than a GST tag both in terms as expression level and solubility. Moreover, TEV protease cleavage works much better in solution than it does on the column. The results furthermore show that the alveolin domain must not be too close to the TEV protease recognition sequence as it could prevent protease access to the cleavage site. Despite these considerable advances, the project has so far fallen short of purifying sufficiently large amounts of recombinant alveolin protein to conduct the planned downstream structural analyses. One of the main problems encountered was the protein precipitating during the TEV protease cleavage step. The removal of the imidazole during TEV protease treatment is itself unlikely to initiate precipitation of the protein, which indicates that it is the release of the NusA tag from the alveolin portion that is causing loss of solubility. On other occasions, the protein precipitated during concentration, suggesting that alveolin solubility may also be linked to its concentration. Other observations provided evidence suggesting that the alveolin

module is not globular, and has a tendency to oligomerise. The latter may be an underlying feature of the alveolins, and perhaps the degree of polymerisation is linked to their solubility. Such properties are not incompatible with IF forming proteins, and in fact could be fundamentally linked to the ability of alveolins in the cell to assemble into the filamentous SPN structure, which is notoriously insoluble and detergent-resistant (Mann *et al.*, 2002).

Chapter 6

General Discussion

6.1 Discussion

Alongside microtubules and actin filaments, intermediate filaments (IFs) constitute one of the three main cytoskeletal systems in eukaryotic cells. Often referred to as 'membrane skeletons', IF networks provide viscoelastic support to membranes and as such are essential in maintaining cell integrity, shape and function. In contrast to animals and amoeba that have actin/spectrin-based plasma membrane skeletons, apicomplexan parasites possess a unique actin/spectrin-free membrane skeleton known as the SPN. The SPN is part of a unique structure known as the pellicle, which is found predominantly in the invasive and motile life cycle stages (zoites) of these parasites. The pellicle is defined by a double membrane layer named the IMC that is located directly underneath the plasma membrane. This IMC is equivalent to a system of single-membrane, flattened sacs known as alveoli, which are a unifying morphological feature that link the Apicomplexa with dinoflagellates and ciliates within the Alveolata superphylum. The SPN is situated on the cytoplasmic side of the IMC, with which it is tightly associated, forming an internal cytoskeletal basket supporting the IMC membranes and providing both rigidity and flexibility to the cell (Mann and Beckers, 2001).

Over a decade ago, a novel family of IMC1 proteins, now named alveolins, were reported to be major components of the SPN in *Plasmodium* and related apicomplexan parasites (Khater *et al*., 2004). Members of the alveolin protein family are also found in other alveolates (Gould *et al*., 2008) and are part of a larger class of proteins called epiplastins that have also been identified in euglenids, glaucophytes and cryptophytes (Goodenough *et al*., 2018). In the genus *Plasmodium*, 13 conserved and syntenic alveolin members have been identified that are differentially expressed among the

three different zoite stages of malaria parasites: the merozoite, ookinete and sporozoite (Al-Khattaf *et al.*, 2015; Kaneko *et al.*, 2015). It has been shown in the malaria species *P. berghei* that disruption of the alveolins IMC1a, IMC1b, IMC1h and IMC1i gives rise to morphological aberrations accompanied by reduced tensile strength of the zoite stages in which they are expressed (Kaneko *et al.*, 2015; Khater *et al.*, 2004; Tremp and Dessens, 2011; Tremp *et al.*, 2008; Volkmann *et al.*, 2012), identifying roles in morphogenesis and viscoelasticity. In addition, null mutant parasites of IMC1a, IMC1b and IMC1h were shown to display reduced and abnormal motility, demonstrating that at least a subset of alveolins also participate in parasite locomotion via a hitherto poorly understood mechanism (Khater *et al.*, 2004; Tremp and Dessens, 2011; Tremp *et al.*, 2008; Volkmann *et al.*, 2012). Simultaneous knockout of IMC1h and IMC1b severely reduces ookinete motility (Tremp and Dessens, 2011), indicating that these two alveolins in particular are main players in alveolin-facilitated locomotion. The combined impacts on zoite shape, tensile strength and motility result in marked reductions in infectivity of alveolin null mutant parasites, leading to inability to complete the life cycle (Kaneko *et al.*, 2015; Khater *et al.*, 2004; Tremp and Dessens, 2011; Tremp *et al.*, 2008; Volkmann *et al.*, 2012). Studies in the related parasite *Toxoplasma gondii* have identified 14 distinct alveolin species also localising predominantly in the SPN, with some having additional roles in cell division (Anderson-White *et al.*, 2011; Dubey *et al.*, 2017). This collective work indicates that alveolins occupy stage- and development-specific functional niches across the entire life cycle of apicomplexan parasites.

In this thesis, a body of work is presented that significantly advances our understanding of the structure and function of the alveolins. For example, evidence is

provided that the 'alveolin' module is the domain responsible for targeting the alveolin IMC1h to the SPN (Chapter 4). This is likely to also be the case for other alveolins, because the 'alveolin' domain is shared between all alveolin species. Evidence is also provided that the IMC1h-specific 'non-alveolin' domain (domain 2) of IMC1h has a likely role in ookinete and sporozoite motility (Chapter 4). This data supports the hypothesis that IMC1h facilitates motility by interacting, through its carboxy-terminal domain, with components of the motility apparatus directly, or indirectly via bridging proteins in the IMC. Indeed, the same could apply to the ookinete-expressed alveolin IMC1b, and to the sporozoite-expressed alveolin IMC1a, each of which also have roles in motility and possess a carboxy-terminal conserved 'non-alveolin' domain (domain 3) (Khater *et al*., 2004; Tremp and Dessens, 2011; Tremp *et al*., 2008). Indeed, in Chapter 3, attempts were made to address this question for IMC1b. Unexpectedly, many problems were encountered with achieving stable integration of the recombinant modified *imc1b* allelles into the target locus. The most likely reason for these problems is that the *imc1b* locus is more difficult to target by homologous recombination than was previously thought. However, it may be worth revisiting this topic, especially via the 'domain knockout' strategy that worked so well for IMC1h, in an approach designed to maximize stable integration, for example by using higher amounts of DNA and/or schizonts during transfections. If successful, it would be interesting to see if IMC1b behaves like IMC1h in such a structure-function analysis in live ookinetes, and the results could add further insights into the question of how alveolins contribute to zoite locomotion. Other future studies to shed light on this question would be to carry out comparative IMC1h interactome studies using parasite lines IMC1hΔ2, IMC1hΔ1, IMC1h/GFP and IMC1h-KO ookinetes, which would allow the identification of candidate IMC proteins including those interacting specifically with domain 2. The

availability of the GFP-tagged parasite lines generated would allow GFP pulldown of IMC1b protein complexes followed by quantitative mass spectrometry to identify individual protein constituents. This approach has been successfully used in the Dessens lab to characterize protein complexes of the ookinete-expressed LCCL lectin adhesive-like proteins (LAPs) (Tremp *et al.*, 2017).

Our understanding of alveolin function and mode of action in the parasite would be greatly enhanced if we possessed information on their structure, in particular that of the core 'alveolin' module that is a shared feature of the protein family. Structural information on these molecules could also be helpful to inform which compound might be able to interfere with alveolin assembly and/or function, for example through *in silico* screening platforms. For this reason, studies presented in Chapter 5 were focussed on the recombinant expression and purification of these molecules. A combination of different downstream biophysical analyses would then allow us to determine if the proteins oligomerise or form filaments, as well as their core secondary structures (e.g. α-helix, β-strand) and even high resolution 3D structures.

Our initial bioinformatic analyses predicted that the conserved alveolin domains are structured, and further pointed to a predominantly β-strand rather than α-helical secondary structure associated with the core 'alveolin' modules. This is supported by the fact that these sequences are abundant in proline residues, which is poorly compatible with α-helices. A suitable bacterial expression system was then identified based on the N-terminal fusion of the alveolin proteins with the NusA protein tag. This system allows high expression levels and solubility of the fusion proteins. The NusA portion can be removed from the target alveolin portions by cleavage with TEV protease, and we determined that this was best achieved in solution rather than on

the column. Furthermore, it is important that the TEV protease cleavage site is not too close to the alveolin domain, for risk of the site being prevented from protease access.

Unfortunately, we were so far unable to purify sufficiently large amounts of the recombinant alveolin proteins to carry out structural analyses. The main problem appears to be poor solubility of the alveolin once the NusA tag is removed, resulting in the protein precipitating out of solution. The same happens when the protein is concentrated to levels suitable for biophysical analyses. Whilst this has prevented meaningful structural analyses, these observations may nonetheless be giving some important clues about the physical properties of alveolins. For example, a concentration-dependent solubility could reflect the way by which these molecules assemble in the cell into the insoluble SPN structure, possibly via the intermittent formation of shorter oligomers (protofilaments).

Due to the uncharacterised alveolins ability to become insoluble and precipitate during these experiments, harsher denaturing conditions could circumvent this. Experiments using guanidine and urea in protein purification buffers denature the protein in question. The denatured protein is then able to be purified using, for example, a 6×His tag, whereby high concentrations of protein can be achieved. The protein is then able to be refolded using protein folding co-solvents, providing a solution to the problems encountered during this project (Wingfield, 2001).

Optimisation of buffers and protocols provide promising next steps for the purification of the alveolins, as well as using different protein tags that are perhaps small enough not to have to be cleaved, whilst at the same time maintaining solubility of the fusion proteins. We could also try to express different alveolin domains, a strategy which assumes that this insolubility is not intrinsically linked to all alveolins. It is encouraging

that in some experiments NusA cleavage was achieved without precipitation, and this would be a useful starting point for optimising steps to further purify and concentrate the cleaved alveolin proteins.

6.2 Future Work

The experiments conducted throughout this project allowed for further understanding of how the individual alveolin domains contribute to the proteins function, and how the protein behaves when expressed and purified from bacterial expression systems. IMC1h was characterised during this project and provided an insight in to how its alveolin domain contributes to the localisation of the protein to the subpellicular network. However, the question of how alveolin type 1 and type 2 domains contribute to the proteins function in the subpellicular network remains unanswered. It also remains unanswered if alveolins are able to form autofilaments with one another, or if filament-forming properties are restricted to each individual alveolin. Future experiments to investigate functional differences between between alveolin type 1 and type 2 domains, for example by structure-function analysis of IMC1b, would address these questions.

Substantial work has been done to study bacterial expression of the alveolins IMC1e, IMC1c and IMC1g. One way forward is to investigate the use of strongly denaturing conditions to purify the protein, and re-fold it using protein folding co-solvents. Another strategy going forward would be to utilise the *Plasmodium* itself to purify protein from. As *Plasmodium berghei* allows for large scale ookinete culturing, a large amount of alveolin protein could potentially be obtained from transgenic parasite lines by affinity purification. Alternatively, the mosquito stage-specific alveolin genes could be expressed under an asexual parasite promoter, allowing for overexpression in blood

stage parasites and potentially large amounts of material to be harvested. Using, for example, GFP pulldown assays, alveolins could be purified directly from the parasite, avoiding the need to express and purify protein from bacterial expression systems, and harvesting the material from a more physiologically relevant environment. This would potentially circumvent the problem of the protein precipitating during the purification process.

References

AL-KHATTAF, F. S., TREMP, A. Z. & DESSENS, J. T. 2015. Plasmodium alveolins possess distinct but structurally and functionally related multi-repeat domains. *Parasitol Res,* 114**,** 631-9.

AL-KHATTAF, F. S., TREMP, A. Z., EL-HOUDERI, A. & DESSENS, J. T. 2017. The Plasmodium alveolin IMC1a is stabilised by its terminal cysteine motifs and facilitates sporozoite morphogenesis and infectivity in a dose-dependent manner. *Mol Biochem Parasitol,* 211**,** 48-56.

ALAVI, Y., ARAI, M., MENDOZA, J., TUFET-BAYONA, M., SINHA, R., FOWLER, K., BILLKER, O., FRANKE-FAYARD, B., JANSE, C. J., WATERS, A. & SINDEN, R. E. 2003. The dynamics of interactions between Plasmodium and the mosquito: a study of the infectivity of Plasmodium berghei and Plasmodium gallinaceum, and their transmission by Anopheles stephensi, Anopheles gambiae and Aedes aegypti. *Int J Parasitol,* 33**,** 933-43.

ALY, A. S., VAUGHAN, A. M. & KAPPE, S. H. 2009. Malaria parasite development in the mosquito and infection of the mammalian host. *Annu Rev Microbiol,* 63**,** 195-221.

AMATO, R., PEARSON, R. D., ALMAGRO-GARCIA, J., AMARATUNGA, C., LIM, P., SUON, S., SRENG, S., DRURY, E., STALKER, J., MIOTTO, O., FAIRHURST, R. M. & KWIATKOWSKI, D. P. 2018. Origins of the current outbreak of multidrug-resistant malaria in southeast Asia: a retrospective genetic study. *Lancet Infect Dis,* 18**,** 337-345.

ANDERSON-WHITE, B. R., IVEY, F. D., CHENG, K., SZATANEK, T., LORESTANI, A., BECKERS, C. J., FERGUSON, D. J., SAHOO, N. & GUBBELS, M. J. 2011. A family of intermediate filament-like proteins is sequentially assembled into the cytoskeleton of Toxoplasma gondii. *Cell Microbiol,* 13**,** 18-31.

ANTINORI, S., GALIMBERTI, L., MILAZZO, L. & CORBELLINO, M. 2012. Biology of human malaria plasmodia including Plasmodium knowlesi. *Mediterr J Hematol Infect Dis,* 4**,** e2012013.

ARAI, M., BILLKER, O., MORRIS, H. R., PANICO, M., DELCROIX, M., DIXON, D., LEY, S. V. & SINDEN, R. E. 2001. Both mosquito-derived xanthurenic acid and a host blood-derived factor regulate gametogenesis of Plasmodium in the midgut of the mosquito. *Mol Biochem Parasitol,* 116**,** 17-24.

BANNISTER, L. H., MITCHELL, G. H., BUTCHER, G. A., DENNIS, E. D. & COHEN, S. 1986. Structure and development of the surface coat of erythrocytic merozoites of Plasmodium knowlesi. *Cell Tissue Res,* 245**,** 281-90.

BARKHUFF, W. D., GILK, S. D., WHITMARSH, R., TILLEY, L. D., HUNTER, C. & WARD, G. E. 2011. Targeted disruption of TgPhIL1 in Toxoplasma gondii results in altered parasite morphology and fitness. *PLoS One,* 6**,** e23977.

BAUM, J., GILBERGER, T. W., FRISCHKNECHT, F. & MEISSNER, M. 2008. Host-cell invasion by malaria parasites: insights from Plasmodium and Toxoplasma. *Trends Parasitol,* 24**,** 557-63.

BAUM, J., PAPENFUSS, A. T., BAUM, B., SPEED, T. P. & COWMAN, A. F. 2006. Regulation of apicomplexan actin-based motility. *Nature Reviews Microbiology,* 4**,** 621-628.

BELACHEW, E. B. 2018. Immune Response and Evasion Mechanisms of Plasmodium falciparum Parasites. *J Immunol Res,* 2018**,** 6529681.

BILLINGSLEY, P. F. 1990. Blood digestion in the mosquito, Anopheles stephensi liston (diptera: Culicidae): Partial characterization and post-feeding activity of midgut aminopeptidases. *Archives of Insect Biochemistry and Physiology,* 15**,** 149-163.

BULLEN, H. E., TONKIN, C. J., O'DONNELL, R. A., THAM, W. H., PAPENFUSS, A. T., GOULD, S., COWMAN, A. F., CRABB, B. S. & GILSON, P. R. 2009. A novel family of Apicomplexan glideosome-associated proteins with an inner membrane-anchoring role. *J Biol Chem,* 284**,** 25353-63.

CALDERARO, A., PICCOLO, G., GORRINI, C., ROSSI, S., MONTECCHINI, S., DELL'ANNA, M. L., DE CONTO, F., MEDICI, M. C., CHEZZI, C. & ARCANGELETTI, M. C. 2013. Accurate identification of the six human Plasmodium spp. causing imported malaria, including Plasmodium ovale wallikeri and Plasmodium knowlesi. *Malar J,* 12**,** 321.

CARTER, V*.,* SHIMIZU, S., ARAI, M. & DESSENS, J. T. 2008. PbSR is synthesized in macrogametocytes and involved in formation of the malaria crystalloids. *Mol Microbiol,* 68**,** 1560-9.

COWMAN, A. F., BERRY, D. & BAUM, J. 2012. The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *J Cell Biol,* 198**,** 961-71.

COWMAN, A. F. & CRABB, B. S. 2006. Invasion of red blood cells by malaria parasites. *Cell,* 124**,** 755-66.

CRAIG, A. G., GRAU, G. E., JANSE, C., KAZURA, J. W., MILNER, D., BARNWELL, J. W., TURNER, G. & LANGHORNE, J. 2012. The role of animal models for research on severe malaria. *PLoS Pathog,* 8**,** e1002401.

DE MARCO, V., STIER, G., BLANDIN, S. & DE MARCO, A. 2004. The solubility and stability of recombinant proteins are increased by their fusion to NusA. *Biochem Biophys Res Commun,* 322**,** 766-71.

DELVES, M. J., RUECKER, A., STRASCHIL, U., LELIEVRE, J., MARQUES, S., LOPEZ-BARRAGAN, M. J., HERREROS, E. & SINDEN, R. E. 2013. Male and female Plasmodium falciparum mature gametocytes show different responses to antimalarial drugs. *Antimicrob Agents Chemother,* 57**,** 3268-74.

DUMMLER, A., LAWRENCE, A. M. & DE MARCO, A. 2005. Simplified screening for the detection of soluble fusion constructs expressed in E. coli using a modular set of vectors. *Microb Cell Fact,* 4**,** 34.

FRASER, R. D. B., MACRAE, T. P., PARRY, D. A. D. & SUZUKI, E. 1971. The structure of feather keratin. *Polymer,* 12**,** 35-56.

FRASER, R. D. B. & PARRY, D. A. D. 2009. The role of beta-sheets in the structure and assembly of keratins. *Biophys Rev,* 1**,** 27.

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**,** 343-57.

FUEHRER, H. P. & NOEDL, H. 2014. Recent advances in detection of Plasmodium ovale: implications of separation into the two species Plasmodium ovale wallikeri and Plasmodium ovale curtisi. *J Clin Microbiol,* 52**,** 387-91.

GARNHAM, P. C. 1951. Patterns of exoerythrocytic schizogony. *Br Med Bull,* 8**,** 10-5.

GOODENOUGH, U., ROTH, R., KARIYAWASAM, T., HE, A. & LEE, J. H. 2018. Epiplasts: Membrane Skeletons and Epiplastin Proteins in Euglenids, Glaucophytes, Cryptophytes, Ciliates, Dinoflagellates, and Apicomplexans. *MBio,* 9.

GOULD, S. B., THAM, W. H., COWMAN, A. F., MCFADDEN, G. I. & WALLER, R. F. 2008. Alveolins, a new family of cortical proteins that define the protist infrakingdom Alveolata. *Mol Biol Evol,* 25**,** 1219-30.

GREGG, K., WILTON, S. D., PARRY, D. A. & ROGERS, G. E. 1984. A comparison of genomic coding sequences for feather and scale keratins: structural and evolutionary implications. *Embo j,* 3**,** 175-8.

GUEVARA PATINO, J. A., HOLDER, A. A., MCBRIDE, J. S. & BLACKMAN, M. J. 1997. Antibodies that inhibit malaria merozoite surface protein-1 processing and erythrocyte invasion are blocked by naturally acquired human antibodies. *J Exp Med,* 186**,** 1689-99.

HAN, Y. S., THOMPSON, J., KAFATOS, F. C. & BARILLAS-MURY, C. 2000. Molecular interactions between Anopheles stephensi midgut cells and Plasmodium berghei: the time bomb theory of ookinete invasion of mosquitoes. *Embo j,* 19**,** 6030-40.

HARDING, C. R. & MEISSNER, M. 2014. The inner membrane complex through development of Toxoplasma gondii and Plasmodium. *Cell Microbiol,* 16**,** 632-41.

HEINTZELMAN, M. B. 2006. Cellular and molecular mechanics of gliding locomotion in eukaryotes. *Int Rev Cytol,* 251**,** 79-129.

HERRMANN, H. & AEBI, U. 2016. Intermediate Filaments: Structure and Assembly. *Cold Spring Harb Perspect Biol,* 8.

HERRMANN, H., STRELKOV, S. V., BURKHARD, P. & AEBI, U. 2009. Intermediate filaments: primary determinants of cell architecture and plasticity. *J Clin Invest,* 119**,** 1772-83.

HICKS, M. R., HOLBERTON, D. V., KOWALCZYK, C. & WOOLFSON, D. N. 1997. Coiled-coil assembly by peptides with non-heptad sequence motifs. *Fold Des,* 2**,** 149-58.

HU, K., MANN, T., STRIEPEN, B., BECKERS, C. J., ROOS, D. S. & MURRAY, J. M. 2002. Daughter cell assembly in the protozoan parasite Toxoplasma gondii. *Mol Biol Cell,* 13**,** 593-606.

HYDER, C. L., ISONIEMI, K. O., TORVALDSON, E. S. & ERIKSSON, J. E. 2011. Insights into intermediate filament regulation from development to ageing. *J Cell Sci,* 124**,** 1363-72.

ISHIDA, T. & KINOSHITA, K. 2007. PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res,* 35**,** W460-4.

KAN, A., TAN, Y. H., ANGRISANO, F., HANSSEN, E., ROGERS, K. L., WHITEHEAD, L., MOLLARD, V. P., COZIJNSEN, A., DELVES, M. J., CRAWFORD, S., SINDEN, R. E., MCFADDEN, G. I., LECKIE, C., BAILEY, J. & BAUM, J. 2014. Quantitative analysis of Plasmodium ookinete motion in three dimensions suggests a critical role for cell shape in the biomechanics of malaria parasite gliding motility. *Cell Microbiol,* 16**,** 734-50.

KANEKO, I., IWANAGA, S., KATO, T., KOBAYASHI, I. & YUDA, M. 2015. Genome-Wide Identification of the Target Genes of AP2-O, a Plasmodium AP2-Family Transcription Factor. *PLoS Pathog,* 11**,** e1004905.

KANTELE, A. & JOKIRANTA, T. S. 2011. Review of cases with the emerging fifth human malaria parasite, Plasmodium knowlesi. *Clin Infect Dis,* 52**,** 1356-62.

KHATER, E. I., SINDEN, R. E. & DESSENS, J. T. 2004. A malaria membrane skeletal protein is essential for normal morphogenesis, motility, and infectivity of sporozoites. *J Cell Biol,* 167**,** 425-32.

KONO, M., PRUSTY, D., PARKINSON, J. & GILBERGER, T. W. 2013. The apicomplexan inner membrane complex. *Front Biosci (Landmark Ed),* 18**,** 982-92.

KUDRYASHEV, M., MUNTER, S., LEMGRUBER, L., MONTAGNA, G., STAHLBERG, H., MATUSCHEWSKI, K., MEISSNER, M., CYRKLAFF, M. & FRISCHKNECHT, F. 2012. Structural basis for chirality and directional motility of Plasmodium sporozoites. *Cell Microbiol,* 14**,** 1757-68.

MADEIRA, F., PARK, Y. M., LEE, J., BUSO, N., GUR, T., MADHUSOODANAN, N., BASUTKAR, P., TIVEY, A. R. N., POTTER, S. C., FINN, R. D. & LOPEZ, R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res,* 47**,** W636-w641.

MANN, T. & BECKERS, C. 2001. Characterization of the subpellicular network, a filamentous membrane skeletal component in the parasite Toxoplasma gondii. *Mol Biochem Parasitol,* 115**,** 257-68.

MANN, T., GASKINS, E. & BECKERS, C. 2002. Proteolytic processing of TgIMC1 during maturation of the membrane skeleton of Toxoplasma gondii. *J Biol Chem,* 277**,** 41240-6.

MATUSCHEWSKI, K. 2006. Getting infectious: formation and maturation of Plasmodium sporozoites in the Anopheles vector. *Cellular Microbiology,* 8**,** 1547-1556.

MENKE, A. & JOCKUSCH, H. 1991. Decreased osmotic stability of dystrophin-less muscle cells from the mdx mouse. *Nature,* 349**,** 69-71.

MONTAGNA, G. N., BUSCAGLIA, C. A., MUNTER, S., GOOSMANN, C., FRISCHKNECHT, F., BRINKMANN, V. & MATUSCHEWSKI, K. 2012. Critical role for heat shock protein 20 (HSP20) in migration of malarial sporozoites. *J Biol Chem,* 287**,** 2410-22.

MOON, R. W., TAYLOR, C. J., BEX, C., SCHEPERS, R., GOULDING, D., JANSE, C. J., WATERS, A. P., BAKER, D. A. & BILLKER, O. 2009. A cyclic GMP signalling module that regulates gliding motility in a malaria parasite. *PLoS Pathog,* 5**,** e1000599.

MORRISSETTE, N. S. & SIBLEY, L. D. 2002. Cytoskeleton of apicomplexan parasites. *Microbiol Mol Biol Rev,* 66**,** 21-38; table of contents.

MOTA, M. M., PRADEL, G., VANDERBERG, J. P., HAFALLA, J. C., FREVERT, U., NUSSENZWEIG, R. S., NUSSENZWEIG, V. & RODRIGUEZ, A. 2001. Migration of Plasmodium sporozoites through cells before infection. *Science,* 291**,** 141-4.

NICOLET, S., HERRMANN, H., AEBI, U. & STRELKOV, S. V. 2010. Atomic structure of vimentin coil 2. *Journal of Structural Biology,* 170**,** 369-376.

OTTO, T. D., BOHME, U., JACKSON, A. P., HUNT, M., FRANKE-FAYARD, B., HOEIJMAKERS, W. A., RELIGA, A. A., ROBERTSON, L., SANDERS, M., OGUN, S. A., CUNNINGHAM, D., ERHART, A., BILLKER, O., KHAN, S. M., STUNNENBERG, H. G., LANGHORNE, J., HOLDER, A. A., WATERS, A. P., NEWBOLD, C. I., PAIN, A., BERRIMAN, M. & JANSE, C. J. 2014. A comprehensive evaluation of rodent malaria parasite genomes and gene expression. *BMC Biol,* 12**,** 86.

PARASKEVOPOULOU, V. & FALCONE, F. H. 2018. Polyionic Tags as Enhancers of Protein Solubility in Recombinant Protein Expression. *Microorganisms,* 6.

PARKYN SCHNEIDER, M., LIU, B., GLOCK, P., SUTTIE, A., MCHUGH, E., ANDREW, D., BATINOVIC, S., WILLIAMSON, N., HANSSEN, E., MCMILLAN, P., HLISCS, M., TILLEY, L. & DIXON, M. W. A. 2017. Disrupting assembly of the inner membrane complex blocks Plasmodium falciparum sexual stage development. *PLoS Pathog,* 13**,** e1006659.

PRESTA, L. G. & ROSE, G. D. 1988. Helix signals in proteins. *Science,* 240**,** 1632-41.

PRUDENCIO, M., RODRIGUEZ, A. & MOTA, M. M. 2006. The silent path to thousands of merozoites: the Plasmodium liver stage. *Nat Rev Microbiol,* 4**,** 849-56.

ROBERT, A., HOOKWAY, C. & GELFAND, V. I. 2016. Intermediate filament dynamics: What we can see now and why it matters. *BioEssays : news and reviews in molecular, cellular and developmental biology,* 38**,** 232-243.

ROSALES-RONQUILLO, M. C. & SILVERMAN, P. H. 1974. In vitro ookinete development of the rodent malarial parasite, Plasmodium berghei. *J Parasitol,* 60**,** 819-24.

ROSENBERG, R. & RUNGSIWONGSE, J. 1991. The number of sporozoites produced by individual malaria oocysts. *Am J Trop Med Hyg,* 45**,** 574-7.

ROSENTHAL, P. J. 2018. Artemisinin Resistance Outside of Southeast Asia. *Am J Trop Med Hyg,* 99**,** 1357-1359.

SAEED, S., TREMP, A. Z. & DESSENS, J. T. 2015. Biogenesis of the crystalloid organelle in Plasmodium involves microtubule-dependent vesicle transport and assembly. *Int J Parasitol,* 45**,** 537-47.

SANTOS, J. M., LEBRUN, M., DAHER, W., SOLDATI, D. & DUBREMETZ, J. F. 2009. Apicomplexan cytoskeleton and motors: key regulators in morphogenesis, cell division, transport and motility. *Int J Parasitol,* 39**,** 153-62.

SHAHABUDDIN, M., TOYOSHIMA, T., AIKAWA, M. & KASLOW, D. C. 1993. Transmission-blocking activity of a chitinase inhibitor and activation of malarial parasite chitinase by mosquito protease. *Proc Natl Acad Sci U S A,* 90**,** 4266-70.

SIDEN-KIAMOS, I., ECKER, A., NYBÄCK, S., LOUIS, C., SINDEN, R. E. & BILLKER, O. 2006. Plasmodium berghei calcium-dependent protein kinase 3 is required for ookinete gliding motility and mosquito midgut invasion. *Mol Microbiol,* 60**,** 1355-63.

SINDEN, R. E. 1978. 3 - Cell Biology. *In:* KILLICK-KENDRICK, R. & PETERS, W. (eds.) *Rodent Malaria.* Academic Press.

SINDEN, R. E., HARTLEY, R. H. & WINGER, L. 1985. The development of Plasmodium ookinetes in vitro: an ultrastructural study including a description of meiotic division. *Parasitology,* 91 (Pt 2)**,** 227-44.

SINNIS, P. & COPPI, A. 2007. A long and winding road: the Plasmodium sporozoite's journey in the mammalian host. *Parasitol Int,* 56**,** 171-8.

SMITH, R. C. & JACOBS-LORENA, M. 2010. Plasmodium-Mosquito Interactions: A Tale of Roadblocks and Detours. *Adv In Insect Phys,* 39**,** 119-149.

TIRAWANCHAI, N. & SINDEN, R. E. 1990. Three non-repeated transmission blocking epitopes recognized in the 21 kD surface antigen of zygotes-ookinetes of Plasmodium berghei. *Parasite Immunol.* 12, 4, 435-46.

TOMAS, A. M., MARGOS, G., DIMOPOULOS, G., VAN LIN, L. H., DE KONING-WARD, T. F., SINHA, R., LUPETTI, P., BEETSMA, A. L., RODRIGUEZ, M. C., KARRAS, M., HAGER, A., MENDOZA, J., BUTCHER, G. A., KAFATOS, F., JANSE, C. J., WATERS, A. P. & SINDEN, R. E. 2001. P25 and P28 proteins of the malaria ookinete surface have multiple and partially redundant functions. *Embo J,* 20**,** 3975-83.

TREMP, A. Z., AL-KHATTAF, F. S. & DESSENS, J. T. 2014. Distinct temporal recruitment of Plasmodium alveolins to the subpellicular network. *Parasitol Res,* 113**,** 4177-88.

TREMP, A. Z., AL-KHATTAF, F. S. & DESSENS, J. T. 2017. Palmitoylation of Plasmodium alveolins promotes cytoskeletal function. *Mol Biochem Parasitol,* 213**,** 16-21.

TREMP, A. Z., CARTER, V., SAEED, S. & DESSENS, J. T. 2013. Morphogenesis of Plasmodium zoites is uncoupled from tensile strength. *Mol Microbiol,* 89**,** 552-64.

TREMP, A. Z. & DESSENS, J. T. 2011. Malaria IMC1 membrane skeleton proteins operate autonomously and participate in motility independently of cell shape. *J Biol Chem,* 286**,** 5383-91.

TREMP, A. Z., KHATER, E. I. & DESSENS, J. T. 2008. IMC1b is a putative membrane skeleton protein involved in cell shape, mechanical strength, motility, and infectivity of malaria ookinetes. *J Biol Chem,* 283**,** 27604-11.

TURNER, P., HOLST, O. & KARLSSON, E. N. 2005. Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in Escherichia coli by using a soluble fusion-tag and by tuning of inducer concentration. *Protein Expr Purif,* 39**,** 54-60.

UHLEMANN, A. C. & KRISHNA, S. 2005. Antimalarial multi-drug resistance in Asia: mechanisms and assessment. *Curr Top Microbiol Immunol,* 295**,** 39-53.

VERNICK, K. D., FUJIOKA, H. & AIKAWA, M. 1999. Plasmodium gallinaceum: a novel morphology of malaria ookinetes in the midgut of the mosquito vector. *Exp Parasitol,* 91**,** 362-6.

VOLKMANN, K., PFANDER, C., BURSTROEM, C., AHRAS, M., GOULDING, D., RAYNER, J. C., FRISCHKNECHT, F., BILLKER, O. & BROCHET, M. 2012. The alveolin IMC1h is required for normal ookinete and sporozoite motility behaviour and host colonisation in Plasmodium berghei. *PLoS One,* 7**,** e41409.

WATERS, A. P., THOMAS, A. W., VAN DIJK, M. R. & JANSE, C. J. 1997. Transfection of malaria parasites. *Methods,* 13**,** 134-47.

WEISS, M. M. & VANDERBERG, J. P. 1977. Studies on Plasmodium ookinetes: II. In vitro formation of Plasmodium berghei ookinetes. *J Parasitol,* 63**,** 932-4.

WILLIAMS, M., ALONSO, H., ENCISO, M., EGARTER, S., SHEINER, L., MEISSNER, M., STRIEPEN, B., SMITH, B. & TONKIN, C. 2015. Two Essential Light Chains Regulate the MyoA Lever Arm To Promote Toxoplasma Gliding Motility. *mBio,* 6.

WINGFIELD, P. T. 2001. Use of protein folding reagents. *Current protocols in protein science,* Appendix 3**,** Appendix-3A.