

The targets and role of palmitoylation in *Plasmodium* parasites

Chwen Ling Tay (BSc Hons)

**Christ's College, University of Cambridge
Wellcome Trust Sanger Institute**

**This dissertation is submitted for the degree of
Doctor of Philosophy in Biological Science**

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically stated in the text.

This dissertation does not exceed the word limit of 60000 words for the Degree Committee for the Faculty of Biology.

Abstract

Palmitoylation is the post-translational reversible addition of the acyl moiety, palmitate, to cysteine residues of proteins, and has been shown to be important in regulating protein trafficking, localisation, stability and function. Palmitoylation is wide-spread in all eukaryotes, and recent work revealed the presence of more than 400 palmitoylated proteins in the *Plasmodium falciparum* intraerythrocytic schizont stages, including proteins involved in key aspects of malaria biology and pathogenesis. The work described in this dissertation advances our understanding of protein palmitoylation in *Plasmodium* by developing a novel method to specifically identify palmitoylated cysteines within the *P. falciparum* palmitome, and characterising for the first time, the *Plasmodium* DHHC and MBOAT proteins, which are thought to mediate protein palmitoylation.

In the first section of work, a quantitative mass spectrometry based approach was developed to identify palmitoylation sites, resulting in the identification of over 100 putative palmitoylation sites in the *P. falciparum* schizont palmitome. These potential palmitoylation sites can be used to guide further experiments into the role of palmitoylation in the function of specific proteins. Pilot experiments were also carried out with related parasites, *P. berghei* and *Toxoplasma gondii*, and revealed palmitoylation sites that were conserved across Apicomplexan species.

The *Plasmodium* DHHC protein family was characterised in *P. falciparum* and *P. berghei*, establishing that individual DHHC proteins are localised to distinct organelles, including specialised parasite-specific organelles such as the rhoptries and the IMC. DHHC protein localisation may therefore play some role in substrate specificity. Knock-out studies identified individual DHHC proteins that were essential for blood stage growth, as well as proteins that could be successfully disrupted, suggesting that a subset of DHHCs is functionally redundant. Lastly, an assay was developed to demonstrate the palmitoyl transferase activity of the *Plasmodium* DHHC proteins, confirming for the first time that these proteins are responsible for protein palmitoylation in *Plasmodium* parasites. This assay further revealed that different *P. falciparum* DHHC proteins could palmitoylate the same target protein, further confirming the existence of overlapping functionality for these proteins in *Plasmodium*.

The occurrence of palmitoylation on so many *Plasmodium* proteins, as well as the existence of a repertoire of *Plasmodium* proteins shown to demonstrate palmitoyl transferase activity, indicate that this post-translational modification may have an important role in the normal cellular function of the parasite. Further study of palmitoylation in *Plasmodium* may thus result in the discovery of potential therapeutic drug targets, and the assays developed here could assist in achieving this goal.

Acknowledgements

First and foremost, sincere thanks to Julian Rayner for excellent guidance and supervision during this entire PhD, as well as for careful reading of this dissertation. Special thanks also to Matt Jones for teaching, advice, discussion and assistance in all things. Many thanks also to Mark Collins for all the hard work on the site-ID purification protocol and mass spectrometry. Thanks also to Michel Theron for kind assistance in designing the *P. falciparum* growth assay and for the flow cytometry analysis, as well as to Leyla Bustamente for qPCR analysis. Thanks also to Ellen Bushell for the generation of *P. berghei* transgenic strains and PFGE analysis. All your time and effort is greatly appreciated.

Table of Contents

Abstract	1
Acknowledgements	2
Table of Contents	3
List of Abbreviations	5
List of Figures	8
List of Tables	10
Chapter 1: Introduction	11
1.1. Protein palmitoylation: Definition	11
1.2. Protein acyltransferases (PATs)	12
1.3. Acyl-protein thioesterases	15
1.4. Functions of protein palmitoylation	16
1.5. Types of palmitoylated proteins	19
1.6. Purification of palmitoylated proteins	21
1.7. Malaria: A major global parasitic disease	25
1.8. The life cycle of <i>Plasmodium</i> parasites	26
1.9. The need for new drug targets	29
1.10. Palmitoylation in <i>Plasmodium</i>	29
1.11. Aims and objectives	32
Chapter 2: Materials and Methods	37
2.1. Generation of plasmid constructs	37
2.2. <i>Plasmodium in vitro</i> cell culture and transfection	43
2.3 Genotyping of transgenic parasite strains	47
2.4. Human Embryonic Kidney 293 (HEK293) <i>in vitro</i> cell culture and transfection	50
2.5. Protein detection and analysis	51
2.6. Palmitome purification in <i>P. falciparum</i> parasites and HEK293 cells	53
Chapter 3: The sites of palmitoylation in the <i>Plasmodium falciparum</i> palmitome	58
3.1. Development of the site-identification (ID) palmitome purification method	58
3.2. Analysis of the overlaps between the five trial site-ID palmitome purifications	82
3.3. Trial site-ID palmitome purification from <i>Plasmodium berghei</i> schizonts and <i>Toxoplasma gondii</i> tachyzoites	95
Chapter 4: Protein acyltransferases in <i>Plasmodium berghei</i>	103

4.1. Repertoire of DHHC-domain-containing proteins in <i>Plasmodium</i>	103
4.2. Expression and localisation of DHHC proteins in <i>Plasmodium berghei</i>	106
4.3. Essentiality of the DHHC proteins in <i>P. berghei</i>	114
Chapter 5: Protein acyltransferases in <i>Plasmodium falciparum</i>	117
5.1. Expression and localisation of DHHC and MBOAT proteins in <i>Plasmodium falciparum</i>	117
5.2. Essentiality of DHHC and MBOAT proteins in <i>P. falciparum</i>	128
Chapter 6: Protein acyltransferase activity of <i>Plasmodium falciparum</i> DHHC proteins	142
6.1. PAT activity assay – the method	142
6.2. Expression and localisation of <i>P. falciparum</i> proteins in HEK293 cells	145
6.3. PAT activity assay of PfDHHC proteins with PfSec22 and PfARO	153
6.4. Point mutations of potential palmitoylation sites in PfSec22 and PfARO	160
6.5. Site-directed mutagenesis of the cysteine in the DHHC domain of PfDHHCS	163
Chapter 7: Discussion	169

List of Abbreviations

2-BMP – 2-bromopalmitate

3-HA – triple-haemagglutinin

5FC – 5-fluorocytosine

6-FAM – 6-carboxyfluorescein

17-ODYA – 17-octadecynoic acid

ABE – Acyl-biotinyl exchange

AMPA – α -amino-3-hydroxy-5-methyl-4-isoxazole propionate

APT – Acyl-protein thioesterase

ARO – Armadillo repeats-only

β -ME – β -mercaptoethanol

BSA – Bovine serum albumin

CDPK1 – Calcium-dependent protein kinase 1

CRT – Chloroquine resistance transporter

DGAT – Diacylglycerol acyltransferase

DHHC – Aspartate-Histidine-Histidine-Cysteine

DMSO – Dimethyl sulphoxide

eNOS – Endothelial nitric oxide synthase

ER – Endoplasmic reticulum

FBS – Foetal bovine serum

GAP45 – Glideosome-associated protein 45

GO – Gene ontology

GPI – glycosylphosphatidylinositol

HEK293 – Human embryonic kidney 293

Hh – Hedgehog

Hhat – Hedgehog acyltransferase

HRP – Horse radish peroxidase

IAA - Iodoacetamide

IMC – Inner membrane complex

KAHRP – Knob-associated histidine-rich protein

MBOAT – Membrane-bound O-acyl transferase

MESA – Mature-parasite-infected erythrocyte surface antigen

MSP1 – Merozoite surface protein 1

MTIP – myosin A tail domain-interacting protein

MWCO – Molecular weight cut-off

NEM – N-ethylmaleimide

PAT – Protein acyl transferase

PBS – Phosphate buffered saline

PBST – PBS with 0.1% Tween-20

PCR – Polymerase chain reaction

PEI – Polyethylenimine

PfEMP1 – *P. falciparum* erythrocyte membrane protein 1

PFGE – Pulsed field gel electrophoresis

PPT – Protein palmitoylthioesterase

PSD-95 – Post synaptic density protein-95

PTM – Post-translational modification

RAP1 – Rhoptry associated protein 1

RT-PCR – Reverse transcription-PCR

SDS – Sodium dodecyl sulphate

SDS-PAGE – SDS-polyacrylamide gel electrophoresis

Shh – Sonic Hedgehog

SILAC – Stable isotope labelling with amino acids in cell culture

SNARE – Soluble NSF attachment protein receptor

TBTA – Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine

TCEP – Tris(2-carboxyethyl)phosphine

TCR – T cell receptor

TM – Transmembrane

List of Figures

Figure 1.1: Reversible S-palmitoylation compared to irreversible N-palmitoylation	12
Figure 1.2: Structure of DHHC-domain and MBOAT-domain-containing PATs	15
Figure 1.3: Function of palmitoylation on the subcellular trafficking of Ras proteins in mammalian cells	17
Figure 1.4: Schematic of the acyl-biotinyl exchange (ABE) method of purifying palmitoyl-proteins	23
Figure 1.5: Schematic of the metabolic labelling and click chemistry method of purifying palmitoyl-proteins	24
Figure 1.6: Schematic of the <i>Plasmodium</i> parasite's life cycle	28
Figure 2.1: Schematic of the <i>P. falciparum</i> and HEK293E expression plasmid constructs	41
Figure 3.1: Trial 1 site-ID palmitome purification by ABE	63
Figure 3.2: Trial 2 site-ID palmitome purification by ABE	68
Figure 3.3: Trial 3A site-ID palmitome purification by ABE	71
Figure 3.4: Trial 3B site-ID palmitome purification by ABE	74
Figure 3.5: Trial 4 site-ID palmitome purification by ABE	77
Figure 3.6: Overlaps between the different trial site-ID palmitome purifications and the total palmitome	83
Figure 3.7: Gene ontology (GO) term analysis of enriched proteins in the overlapping trial site-ID purification datasets	86
Figure 3.8: Over-represented sequence motifs within the Trial 4 site-ID palmitome purification dataset	93
Figure 3.9: Trial site-ID palmitome purification by ABE for <i>Plasmodium berghei</i> and <i>Toxoplasma gondii</i>	96
Figure 4.1: Repertoire of DHHC-containing proteins in <i>Plasmodium</i>	105
Figure 4.2: Generation of PbDHHC triple-HA-tagged and knock-out transgenic lines	107
Figure 4.3: Expression of PbDHHC proteins in purified <i>P. berghei</i> schizonts	108
Figure 4.4: Localisation of DHHC proteins in <i>P. berghei</i> schizonts	111
Figure 5.1: Generation of PfDHHC and PfMBOAT triple-HA-tagged transgenic lines	119
Figure 5.2: Expression of PfDHHC and PfMBOAT proteins in <i>P. falciparum</i> schizonts	121
Figure 5.3: Localisation of DHHC and MBOAT proteins in <i>P. falciparum</i>	124
Figure 5.4: Generation of PfDHHC and PfMBOAT knock-out transgenic lines in <i>P. falciparum</i>	133
Figure 5.5: Reverse transcription (RT)-PCR analysis of PfDHHC5-KO and PfDHHC9-KO transgenic lines	136
Figure 5.6: Growth assay comparing the growth of PfDHHC5-KO and PfDHHC9-KO with wild-type 3D7	138

Figure 6.1: PAT activity assay	144
Figure 6.2: Localisation of codon-optimised <i>P. falciparum</i> proteins in HEK293 cells	150
Figure 6.3: Immunoprecipitation of PfSec22 and PfARO proteins using antibodies against the c-Myc tag	153
Figure 6.4: PAT activity assay demonstrating the PAT activity of PfDHHC proteins on the target protein, PfSec22	155
Figure 6.5: PAT activity assay demonstrating the PAT activity of PfDHHC proteins on the target protein, PfARO	158
Figure 6.6: Immunoprecipitation of mutant PfSec22 and PfARO proteins using antibodies against the c-Myc tag	162
Figure 6.7: Immunoprecipitation of PfSec22 and PfARO proteins when co-expressed with mutant PfDHHC5 using antibodies against the c-Myc tag	165

List of Tables

Table 1.1: Common classes of palmitoylated proteins	20
Table 2.1: Primers used for the generation of <i>P. falciparum</i> triple-HA-tagged plasmid Constructs	38
Table 2.2: Primers used for the generation of <i>P. falciparum</i> double-crossover knock-out plasmid constructs	39
Table 2.3: PlasmogEM vectors used for the generation of 3HA-tagged and knock-out transgenic strains in <i>P. berghei</i>	40
Table 2.4: Primers used for the generation of point mutations in HEK293 expression plasmid constructs	43
Table 2.5: Primers used for the genotyping of <i>P. falciparum</i> triple-HA-tagged lines	47
Table 2.6: Primers used for the genotyping of <i>P. falciparum</i> knock-out lines	48
Table 2.7: Primers used for the PCR analysis of the cDNA of PfDHHC5-KO and PfDHHC9-KO transgenic clones	49
Table 2.8: Primers and probes used for the qPCR analysis of the cDNA of PfDHHC5-KO and PfDHHC9-KO transgenic clones	50
Table 2.9: All primary and secondary antibodies used in this work along with their appropriate working dilutions	51
Table 3.1: Summary of all five trial site-ID purifications and the analyses performed on the five datasets	81
Table 3.2: The 14 enriched peptides identified in all 4 trial site-ID palmitome purifications	84
Table 3.3: A selected set of enriched proteins found to be common between both label-free trial site-ID datasets (Trial 3B and Trial 4), as well as the total palmitome, along with a selected set of enriched peptides common between both the trial datasets	88
Table 3.4: A selected set of enriched proteins found to be common between the Trial 4 site-ID dataset and the total palmitome, along with a selected set of corresponding enriched peptides	91
Table 3.5: Summary of the analyses performed on the overlaps between the different trial site-ID palmitome purifications	94
Table 3.6: Summary of the trial site-ID purifications for <i>P. berghei</i> and <i>T. gondii</i> , and the analyses performed on the two datasets	97
Table 3.7: A selected set of enriched proteins purified in the <i>P. berghei</i> trial site-ID palmitome purification, along with their <i>P. falciparum</i> homologues, present in the <i>P. falciparum</i> Trial 4 site-ID palmitome purification	98
Table 3.8: A selected set of enriched proteins purified in the <i>T. gondii</i> trial site-ID palmitome purification, along with their <i>P. falciparum</i> homologues (homologues were determined by protein-protein BLAST search) found to be present in the <i>P. falciparum</i> Trial 4 site-ID palmitome purification	100
Table 4.1: DHHC-domain-containing proteins in <i>Plasmodium berghei</i> and <i>Plasmodium Falciparum</i>	104

Chapter 1

Introduction

The targets and role of palmitoylation in *Plasmodium* parasites

The post-translational modification (PTM) of proteins plays an essential role in modulating protein activity and function, and thus greatly influences the normal cell biology of most organisms, including *Plasmodium* parasites, the causative agents of malaria. The work described here explores palmitoylation, one of the lesser-known protein PTMs, which is used broadly to regulate protein function in other eukaryotes, but has not been well-studied in *Plasmodium*. This dissertation focuses on two broad aims: developing new approaches to identify palmitoylated cysteines in *Plasmodium falciparum* proteins, and characterising for the first time the enzymes responsible for palmitoylation in *Plasmodium* parasites. To set the background for this work, this introduction outlines what is currently known about protein palmitoylation in other eukaryotic organisms in general, and in *Plasmodium* parasites in particular.

1.1. Protein Palmitoylation: Definition

Palmitoylation is a PTM involving the addition of long-chain fatty acids to one or more internal cysteine residues via thioester (S-acyl) linkages (Figure 1.1). Proteins can be S-acylated with different fatty acid chains and thus this modification can also be referred to by the more general term, 'S-acylation'. However, the fatty acid typically used is palmitic acid, a 16-carbon saturated fatty acid, hence the term 'palmitoylation'. Critically, the thioester bond between palmitic acid and the cysteine residue is reversible, making palmitoylation a dynamically regulated protein modification, just like the much more widely studied protein PTM, phosphorylation [1].

The reversibility of the thioester bond makes palmitoylation distinct from other lipid modifications that commonly occur, such as myristoylation and prenylation. Myristoylation refers to the covalent attachment of myristate to the N-terminal glycine residue of the protein, which occurs co-translationally upon the removal of the initiator methionine residue [2]. Prenylation is the addition of a farnesyl or geranylgeranyl isoprenoid to a cysteine residue that is four amino acids from the C-terminus via a thioether linkage, which leads to subsequent proteolytic removal of the three terminal amino acids and the carboxyl-methylation of the prenylated cysteine [3]. Both myristoylation and prenylation are irreversible modifications.

While the majority of palmitoylation occurs in the form of the reversible S-acylation described above, palmitic acid can be irreversibly attached to cysteine residues if they are at the extreme N-terminus of proteins. In this case, it is thought that an intermediate thioester linkage is first formed and this is

then rearranged into a stable amide linkage (Figure 1.1). This irreversible version of palmitoylation is referred to as 'N-palmitoylation' [4].

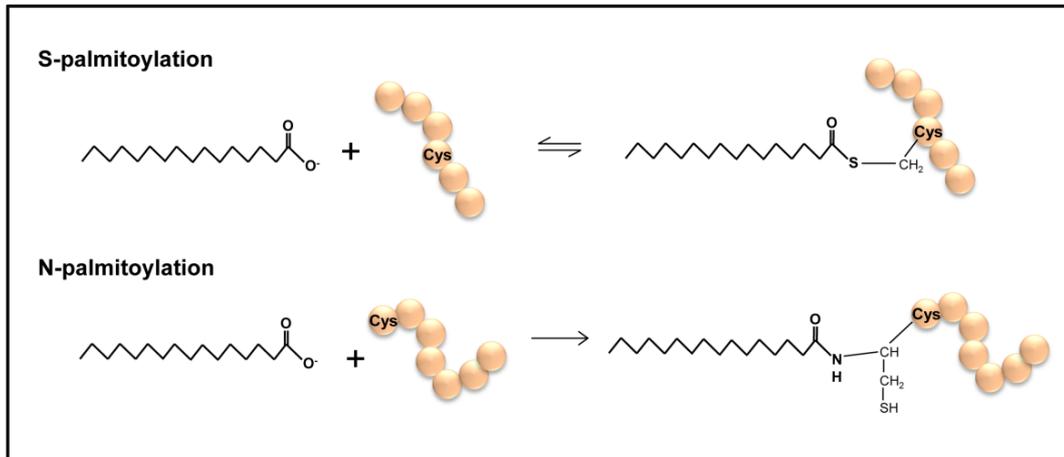


Figure 1.1: Reversible S-palmitoylation compared to irreversible N-palmitoylation. Chemical structures of palmitate attached to a protein via the reversible thioester linkage (S-palmitoylation) and the irreversible amide linkage (N-palmitoylation).

1.2. Protein acyltransferases (PATs)

The formation of the thioester bond between the palmitoyl moiety and the cysteine side chain can occur spontaneously *in vitro* in the presence of palmitoyl-CoA. It was therefore initially thought that palmitoylation took place via non-enzymatic mechanisms *in vivo* [5]. However, the recent discovery of proteins responsible for catalysing protein palmitoylation, first in *Saccharomyces cerevisiae* and later in mammalian cells, now indicates that palmitoylation is in fact an enzyme-mediated process [4, 6, 7]. These proteins are referred to as protein acyltransferases (PATs) and can be divided into two main families: DHHC-domain-containing PATs (DHHC-PATs) and MBOAT-domain-containing PATs (MBOAT-PATs) [5].

1.2.1. DHHC-PATs:

DHHC-PATs are transmembrane (TM) proteins and typically contain four to six TM-domains. The defining feature of these proteins is that they all contain a highly conserved Asp-His-His-Cys (DHHC) motif within a cysteine-rich domain [5] and this motif is commonly located on the loop between two TM-domains, and is exposed on the cytosolic side of the membrane [8, 9] (Figure 1.2A). The DHHC

signature motif, which is part of the larger approximately 50 residue cysteine-rich domain, is conserved across eukaryotes. This conserved domain allows the bioinformatics identification of DHHC-PATs, and each eukaryotic genome contains a variable number, ranging from 7 in *S. cerevisiae* to 23 in humans [5].

Besides the DHHC domain, two other regions display significant homology across the DHHC-PATs. An Asp-Pro-Gly (DPG) motif is typically found next to the second TM-domain and a Thr-Thr-x-Glu (TTxE) motif is usually found next to the fourth TM-domain, although the functional significance of these conserved areas are still unknown [5].

The DHHC-domain is thought to be directly involved in the palmitoyl transfer reaction, although not all details of the enzymatic process are currently known. All DHHC proteins so far have been found to undergo autopalmitoylation and when the cysteine residue in the DHHC motif is mutated, autopalmitoylation is abolished. This has led to the suggestion that palmitoylation of the substrate may occur via a two-step process where the PAT is first autopalmitoylated, resulting in the production of a palmitoyl-enzyme intermediate, and this is followed by transfer of the palmitoyl moiety to the substrate [5, 10]. Evidence suggests that the DHHC motif is essential for this two-step palmitoylation reaction [10]. However, this is by no means proven, and other models posit that instead of acting as an intermediate in the palmitoyl transferase reaction, the palmitoylated DHHC protein could be involved in allosteric regulation of substrate binding or in substrate specificity [5]. Indeed, other work suggests that the DHHC domain is not merely a catalytic unit and may in fact possess substrate specificity determinants [11].

The general question of DHHC substrate specificity –whether individual DHHC-PATs palmitoylate a specific set of targets- is a subject of particular interest for the field as a whole. For some palmitoyl-proteins, it appears that palmitoylation can be catalysed by several different PATs [11]. For example, deletion of the yeast PAT, Pfa3, reduced but did not abolish palmitoylation of the vacuolar protein, Vac8, which may indicate that the palmitoylation of Vac8 may be taken over by another PAT in the absence of Pfa3 [12, 13]. Similarly the palmitoylation of Ras in yeast was also only partially reduced in the absence of its cognate PAT, Erf [14]. Furthermore, it was also shown that multiple DHHC-PATs could be knocked-out in yeast without completely abolishing palmitoylation of substrates, suggesting possible overlap in DHHC-PAT functionality [14]. However, some palmitoyl-proteins do appear to have dedicated DHHC-PATs. For example, in yeast, palmitoylation of SNARE proteins appear to be the responsibility of the DHHC-PAT, Swf1, while another DHHC-PAT, Pfa4, appears to palmitoylate only the amino acid permeases [14], and the activities of both Swf1 and Pfa4 cannot be performed by other yeast PATs [11]. The question of whether DHHC-PATs are substrate specific therefore seems to

depend on the particular DHHC-PAT or the particular group of palmitoyl-protein substrates. However, it should be pointed out that only a handful of DHHC-PAT/substrate partnerships have been studied in detail, and there have been no studies on the global impact of DHHC-PAT deletion on palmitoylation sites, in large part because few global datasets of palmitoylation sites have so far been generated.

1.2.2. MBOAT-PATs

The membrane-bound O-acyl transferase (MBOAT) superfamily of proteins consists of membrane-bound proteins, which generally contain eight to ten TM-domains. Although they are less conserved compared to the DHHC family of proteins, MBOAT proteins possess an invariant histidine residue found within a long hydrophobic region which is thought to form part of the active site [15, 16] (Figure 1.2B). The majority of the members of the MBOAT superfamily act as enzymes for the transfer of fatty acids to the hydroxyl groups of membrane-bound target proteins [15]. However, several MBOAT proteins have also been found to catalyse N-palmitoylation, in particular, the N-palmitoylation of secreted proteins such as Hedgehog (Hh), Spitz and Wnt [17-20]. For example, in humans, the secreted morphogen Sonic hedgehog (Shh), the homologue of the *Drosophila* Hh protein, is N-palmitoylated, and the transfer of palmitate to the N-terminal cysteine is catalysed by Hedgehog acyltransferase (Hhat), a member of the MBOAT family [21]. Porcupine, another MBOAT protein, is also known to mediate palmitoylation of Wnt secreted proteins [20]. Thus, the MBOAT-PATs appear to be mainly responsible for the irreversible N-palmitoylation of secreted signalling proteins.

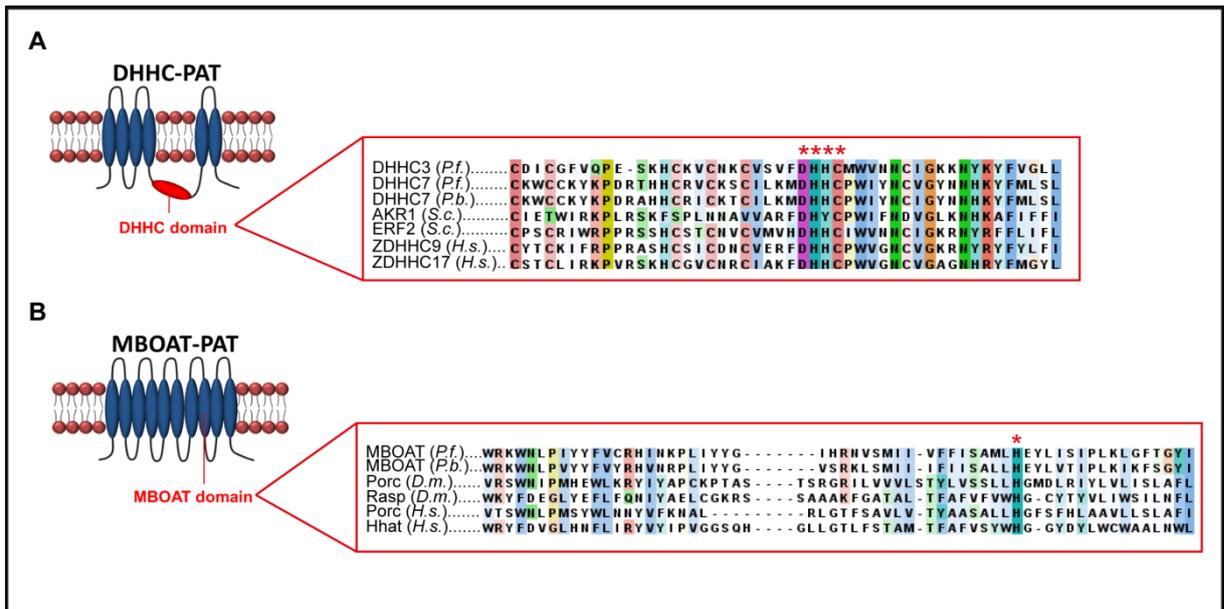


Figure 1.2: Structure of DHHC-domain and MBOAT-domain-containing PATs. (A) The generic structure of DHHC-PATs shown schematically. The conserved DHHC-domain is shown in red. Representative amino acid sequences of the DHHC domains of known DHHC-PATs from different organisms are shown aligned on the right. The characteristic DHHC sequence is highlighted with asterisks. **(B)** The generic structure of MBOAT-PATs shown schematically. The hydrophobic domain containing the conserved histidine residue is indicated as the MBOAT domain. Representative sequences of MBOAT proteins are shown aligned on the right. The conserved histidine residue which is thought to form part of the active site is highlighted with an asterisk. *P.f.* - *Plasmodium falciparum*, *P.b.* - *Plasmodium berghei*, *S.c.* - *Saccharomyces cerevisiae*, *D.m.* - *Drosophila melanogaster*, *H.s.* - *Homo sapiens*.

1.3. Acyl-protein thioesterases

As mentioned above, S-palmitoylation is a reversible reaction as the thioester linkage can be broken, allowing the removal of the palmitate group from proteins. A class of enzymes known as acyl-protein thioesterases have been found to catalyse the depalmitoylation of proteins [22]. There are two types of acyl-protein thioesterases. The first type, known as the protein palmitoylthioesterases (PPT), of which there are two - PPT1 and PPT2, are localised in lysosomes and are involved in depalmitoylating proteins during protein degradation [23]. The second type of acyl-protein thioesterase, known as the APTs, appears to be the enzymes responsible for regulating the depalmitoylation of S-palmitoylated proteins [22]. APT1 is found in the cytosol and has been shown to remove palmitate from signalling proteins such as G-proteins, Ras and endothelial nitric oxide synthase (eNOS) [24, 25]. Although APT1 was originally found to act as a lysophospholipase, later work indicates that it in fact prefers palmitoyl-thioesters as a substrate [24, 26]. Furthermore, the inhibition of APT1 appears to

interfere with Ras localisation and signalling, which is regulated by cycles of palmitoylation and depalmitoylation [27]. This indicates that the APTs may be important players in regulating protein palmitoylation, although much is still unknown about these proteins and further study is now being performed in order to determine how and where these enzymes function *in vivo*.

1.4. Functions of protein palmitoylation

Protein palmitoylation is a widespread PTM and occurs in a large and diverse repertoire of proteins. As a lipid modification, the basic function of palmitoylation is to regulate the interaction of proteins with the lipid bilayers of cellular membranes. It is however, more than just a simple lipid anchor, and a significant role of palmitoylation appears to be in the trafficking of proteins to specific cellular membranes [28]. A well-studied example of proteins which uses palmitoylation for specific intracellular targeting is the dually-acylated Ras family of small GTPases. H-Ras and N-Ras proteins are first farnesylated, which targets the Ras proteins to the endoplasmic reticulum where they are then palmitoylated. Palmitoylated Ras proteins are trafficked to the Golgi and then to the plasma membrane via the classical secretory pathway. At the plasma membrane, Ras proteins are depalmitoylated, allowing their trafficking back to the Golgi where they can be repalmitoylated and sent back to the membrane [4, 8, 29, 30] (Figure 1.3). This dynamic cycle of palmitoylation and depalmitoylation is required for regulating the subcellular distribution of Ras proteins and in turn, may be important in regulating the locations in the cell where Ras proteins are available for signalling [3, 31].

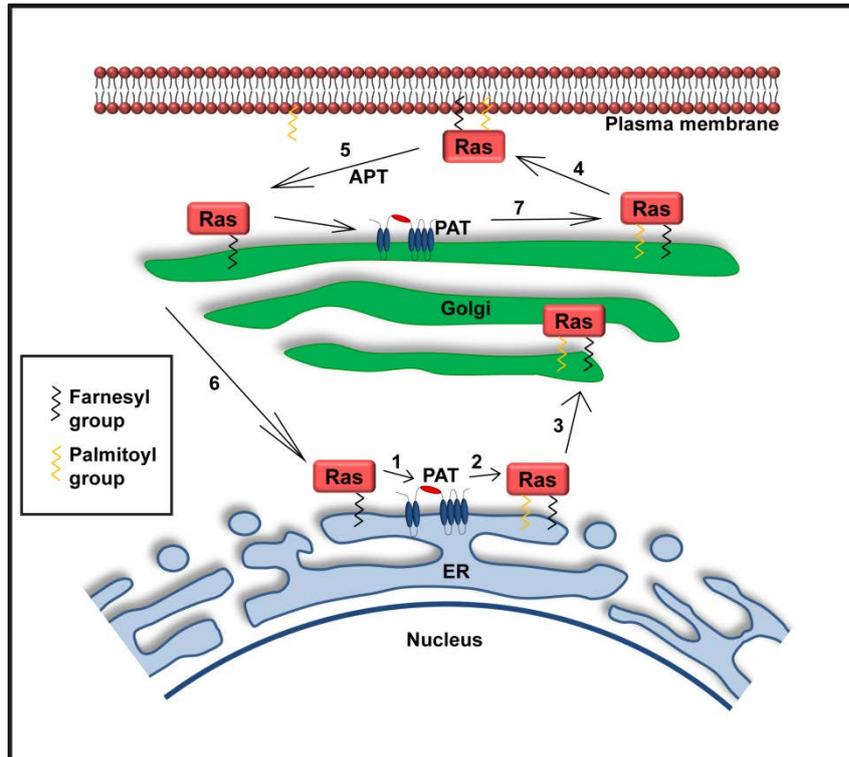


Figure 1.3: Function of palmitoylation on the subcellular trafficking of Ras proteins in mammalian cells. Ras proteins are dually-acylated proteins. The first acyl modification is the addition of a farnesyl moiety (black) and this farnesylation directs Ras proteins to the endoplasmic reticulum (ER). Farnesylation alone is not sufficient to anchor Ras proteins to the ER and a second lipid modification is required. Thus, farnesylated Ras proteins are then (1) palmitoylated with a palmitoyl group (yellow) by the putative protein acyltransferase (PAT) in the ER. This results in (2) dually-acylated Ras proteins which are either anchored at the ER membrane, or are (3) trafficked to the Golgi. From the Golgi, dually-acylated Ras can then be (4) trafficked to the plasma membrane, where they are (5) depalmitoylated by the putative acyl-protein thioesterase (APT) and trafficked back to the Golgi. At the Golgi, the farnesylated Ras proteins can either be (6) trafficked back to the ER, or (7) repalmitoylated by a putative PAT in the Golgi, to be once again delivered to the plasma membrane. This dynamic cycle of palmitoylation and depalmitoylation regulates the subcellular distribution of Ras proteins in mammalian cells.

As well as targeting proteins to specific intracellular membranes, palmitoylation is also used to regulate protein localisation between microdomains within a specific membrane [4]. Again, Ras family proteins are the most well-studied example, where H-Ras can be palmitoylated on one or both of two cysteine residues, Cys181 and Cys184. Monopalmitoylation on Cys184 mostly retains H-Ras at the Golgi, with some inefficient delivery to the plasma membrane. For the fraction that is delivered to the plasma membrane, monopalmitoylation on Cys184, as well as palmitoylation of both cysteine residues, promotes GTP-regulated lateral diffusion of H-Ras between cholesterol-dependent and cholesterol-independent microdomains. This means that GTP-loaded Cys184 monopalmitoylated Ras,

as well as GTP-loaded dually-palmitoylated Ras segregate to cholesterol-independent microdomains, while GDP-loaded Cys184 monopalmitoylated Ras and GDP-loaded dually-palmitoylated Ras segregate to cholesterol-dependent microdomains. Monopalmitoylation of H-Ras on Cys181 meanwhile directs H-Ras to the plasma membrane, where monopalmitoylation on Cys181 dramatically reverses the lateral segregation of Ras and promotes localisation of GTP-loaded Cys181 monopalmitoylated Ras into cholesterol-dependent microdomains [32]. Palmitoylation thus appears to also influence the fine-tuning of Ras localisation on cellular membranes.

In addition to soluble proteins such as Ras, many TM-domain-containing proteins, such as multi-TM-domain spanning receptors, are also palmitoylated, despite the fact that the presence of the TM domain is sufficient for membrane binding [1]. One of the reasons for the palmitoylation of TM-domain-containing proteins could again be the fine-tuning of membrane localisation. For example, CD4 and CD8 -the co-receptors of the T cell receptor (TCR) complex- are palmitoylated, and palmitoylation is required for the localisation of these co-receptors into lipid rafts, which in turn promotes lipid raft aggregation and is important for TCR-mediated downstream signalling events [33-35]. Another example of the use of palmitoylation is the α -amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA) receptor, which is palmitoylated on different cysteines in its second TM domain and also on cysteines in its C-terminal domain. Palmitoylation of these cysteines applies multiple levels of regulation on the localisation of the receptor. For example, palmitoylation of cysteines in the second TM domain causes accumulation of the AMPA receptor at the Golgi, and palmitoylation of cysteines in the C-terminal domain promotes regulated endocytosis of the AMPA receptors, removing the AMPA receptors from the cell surface [36]. In addition to that, the dynamic palmitoylation of the post synaptic density protein, PSD-95, controls the synaptic clustering of PSD-95 and this in turn controls the recruitment of AMPA receptors to the post synaptic density. This then regulates the cycling of AMPA receptors on and off the synaptic membrane and influences synaptic plasticity and strength [37].

Finally, palmitoylation can also regulate protein-protein interactions and protein stability. An example of this function of palmitoylation is demonstrated by the yeast t-SNARE, Tlg1. This single TM-domain-containing protein is palmitoylated at a cysteine residue close to its TM-domain. Inhibition of palmitoylation results in the interaction of Tlg1 with Tul1 -an E3 ubiquitin ligase- which causes the ubiquitination of Tlg1, its targeting to the vacuole, and its subsequent degradation by the proteasome [4, 38, 39]. Another extreme and well-known example of the effect of palmitoylation on protein stability is its impact on the stability and aggregation of the huntingtin protein. The intracellular aggregation of the mutant form of huntingtin is a characteristic feature of Huntington's

disease, a debilitating neurological disorder [40]. Huntingtin is palmitoylated and the mutant form of huntingtin shows decreased palmitoylation. Inhibition of the palmitoylation of wild-type huntingtin promotes the accumulation of huntingtin protein aggregates and this is even more pronounced when palmitoylation of mutant huntingtin is inhibited. Conversely, protein aggregation is decreased when the palmitoylation of huntingtin is increased [40]. Thus, in this case, palmitoylation appears to play a protective role by regulating protein aggregation, and may play a direct role in the pathogenesis of Huntington's disease.

Reversible S-palmitoylation can therefore be used as a protein regulatory tool through a variety of mechanisms, including specifying locations within or between membrane-bound organelles, and regulating protein-protein interactions and protein stability. Although N-palmitoylation is an irreversible process, this form of palmitoylation has also been found to have an important influence on protein activity, particularly that of secreted signalling ligands such as Hh and Wnt proteins [1]. The Hh and Wnt families of proteins are required for pattern formation during embryonic development. These proteins act as morphogens, forming the signalling gradients which specify growth patterning, proliferation and differentiation during the developmental process [17]. Both these families of proteins are N-palmitoylated on specific cysteine residues and the inhibition or removal of N-palmitoylation results in defective pattern formation and reduced signalling activity [1, 20, 41]. N-palmitoylation is thus thought to regulate the distribution and localisation, as well as the signalling of these secreted proteins, which in turn affect the protein gradients and pattern formation controlled by these proteins.

1.5. Types of palmitoylated proteins

From the examples described above, it is clear that palmitoylation is more than merely a membrane anchor and in fact appears to have several important regulatory functions within the cell. However, predicting which proteins might be regulated in this manner is not straightforward because there is no clear consensus sequence for palmitoylation, other than the requirement for a cysteine residue [3]. However, palmitoyl-proteins have been grouped into several classes based on the location of known palmitoylated cysteine residues and on the context of the sequences surrounding the cysteine residues [3, 8]. Examples of these classes of palmitoyl-proteins are shown in Table 1.1.

Lastly, Type IV palmitoyl-proteins are also dually-acylated proteins, which are modified with myristate and palmitate. These proteins possess the N-myristoylation consensus sequence, Met-Gly-X-X-X-Ser/Thr, at the N-terminal end of the protein, where the glycine residue is the N-myristoylated residue. For most of myristoylated and palmitoylated proteins, the amino acid in position three of the N-myristoylation consensus sequence is a cysteine residue, and myristoylation of glycine in position two appears to facilitate the palmitoylation of the cysteine residue [28]. However, for some proteins, palmitoylation can also occur on cysteines up to 20 amino acids away from the N-terminus [3]. Members of this class of palmitoyl-proteins include the Src family of protein tyrosine kinases [28] (Table 1.1).

While splitting palmitoylated proteins into these basic classes is helpful, it obscures the fact that there are palmitoyl-proteins that do not fall into any of these groups. The true scale at which palmitoylation is used in the cell, and the number of palmitoyl-proteins that fall outside the classical groupings has become apparent through the recent development of two independent biochemical approaches for purifying the total complement of palmitoylated proteins.

1.6. Purification of palmitoylated proteins

Early methods of identifying palmitoylated proteins involved the use of metabolic labelling with tritiated palmitic acid, which was time-consuming and mainly focused on individual proteins rather than global, whole proteome analyses [14, 43, 44]. Recently, advances in proteome-based technologies to purify and identify all palmitoylated proteins in an organism have allowed the systematic characterisation of the whole palmitoylated proteome, or 'palmitome', in different organisms. Most of these studies revealed that the size of most palmitomes numbered in the hundreds [8, 14, 45-47]. The palmitoyl-protein enrichment methods can be divided into two main techniques. The first is based on a chemical-exchange process known as acyl-biotinyl exchange (ABE) chemistry [43] and the second involves metabolic labelling followed by 'click chemistry' [45].

1.6.1 Acyl-biotinyl exchange (ABE)

Acyl-biotinyl exchange (ABE) is based on the exchange of the palmitate group with a detectable label, such as biotin, and relies on the susceptibility of the thioester linkage to cleavage by neutral hydroxylamine [43]. ABE consists of three sequential chemical steps. In the first step, the extracted and solubilised proteome sample is treated with N-ethylmaleimide (NEM), a chemical which

irreversibly binds to free cysteine thiols, causing the irreversible blockage of existing free thiols in the sample. This is followed by treatment with hydroxylamine, which specifically cleaves the thioester bonds between the palmitoyl-groups and the cysteine residues, releasing the palmitoyl-groups and exposing the previously palmitoylated thiol groups. The resultant newly-exposed free thiols are then biotinylated by a thiol-reactive biotinylation reagent, such as HPDP-biotin, and the biotinylated palmitoyl-proteins can then be affinity-purified using streptavidin agarose [14, 43, 48]. An important control in this method of palmitoyl-protein purification is the mock treatment of an equal quantity of protein sample – that is, parallel treatment of a sample under identical conditions but in the absence of hydroxylamine. In this control sample, palmitoyl-groups are not removed and thus are not biotinylated or purified, and therefore, only contaminant proteins which may occur due to inappropriate biotinylation, non-specific streptavidin binding or the non-specific purification of highly abundant proteins are expected to be purified [48]. This control sample is essential in order to assist in determining false-positive identifications, and purification of palmitoyl-proteins are measured by their enrichment in the hydroxylamine-treated ‘palmitome’ sample over the non-hydroxylamine-treated ‘control’ sample [48]. As ABE is based on the cleavage of the thioester linkage by hydroxylamine, only S-acylated proteins are purified, and proteins that are modified by the irreversible N-palmitoylation will not be purified by this method. A schematic showing the chemical exchange steps of ABE is shown in Figure 1.4.

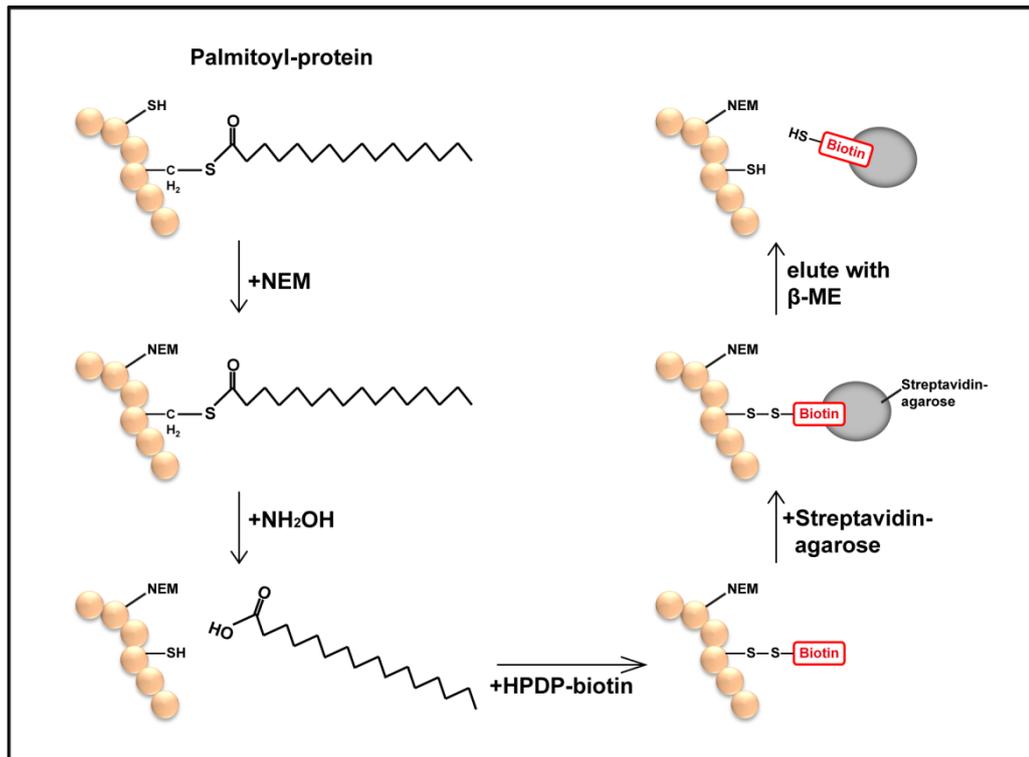


Figure 1.4: Schematic of the acyl-biotinyl exchange (ABE) method of purifying palmitoyl-proteins. This method relies on the exchange of palmitoyl groups with biotin. The extracted proteome is first treated with N-ethylmaleimide (NEM) which binds irreversibly to free thiols of unmodified cysteine residues. Palmitoyl groups are removed by treatment with hydroxylamine (NH_2OH), which cleaves the thioester linkage, resulting in the formation of newly-exposed free thiols. These free thiols can be biotinylated by a thiol-reactive biotinylation reagent, such as HPDP-biotin. The biotinylated proteins can then be purified by streptavidin-agarose affinity purification and eluted from the streptavidin-agarose with β -mercaptoethanol (β -ME). In the control sample, the protein extract is left untreated with hydroxylamine, the thioester linkages are thus not cleaved and palmitoyl groups are not exchanged with biotin. Proteins which are significantly enriched in the hydroxylamine-treated palmitome samples compared to the non-hydroxylamine-treated control samples can thus be classified as palmitoylated.

1.6.2 Metabolic labelling and click chemistry

This method involves the use of a palmitic acid analog that is incorporated into cellular proteins, then post protein extraction, can be conjugated to azide-linked labels using standard 'click chemistry' conditions for purification or detection [45]. Cells are first metabolically labelled with the palmitic acid analogue, 17-octadecynoic acid (17-ODYA), which contains an alkyne group. The proteome is then extracted from the metabolically labelled cells and reacted with biotin-azide under conditions conducive for the copper (I)-catalysed azide-alkyne cycloaddition (click chemistry) reaction. The click chemistry reaction essentially joins the azide and alkyne groups and thus results in the biotinylation of

the palmitoyl-proteins, which can then be affinity purified by streptavidin-agarose [8, 44, 45, 49]. As a control, cells are mock-labelled with the palmitic acid analogue and this control is important for the determination of false-positives and contaminant proteins [8, 45], just as the mock treatment control is critical in ABE purification. As this method is based on the metabolic incorporation of a palmitic acid analogue, both the thioester-linked S-palmitoylation and the irreversible N-palmitoylation can be purified by this method. A schematic showing the steps of this method of palmitoyl-protein purification is shown in Figure 1.5. Metabolic labelling with 17-ODYA followed by click chemistry methods can also be used in pulse-chase labelling experiments in order to study the dynamics of protein palmitoylation, by distinguishing between stably palmitoylated proteins and proteins which undergo rapid turnover of palmitoylation [44, 49]. Importantly, this metabolic labelling method can also be used for alternative purposes other than purification, simply by modifying the azide-linked label that is used in the click chemistry reaction. For example, replacement of biotin-azide with rhodamine-azide allows visualisation of all 17-ODYA-labelled proteins by in-gel fluorescence [45].

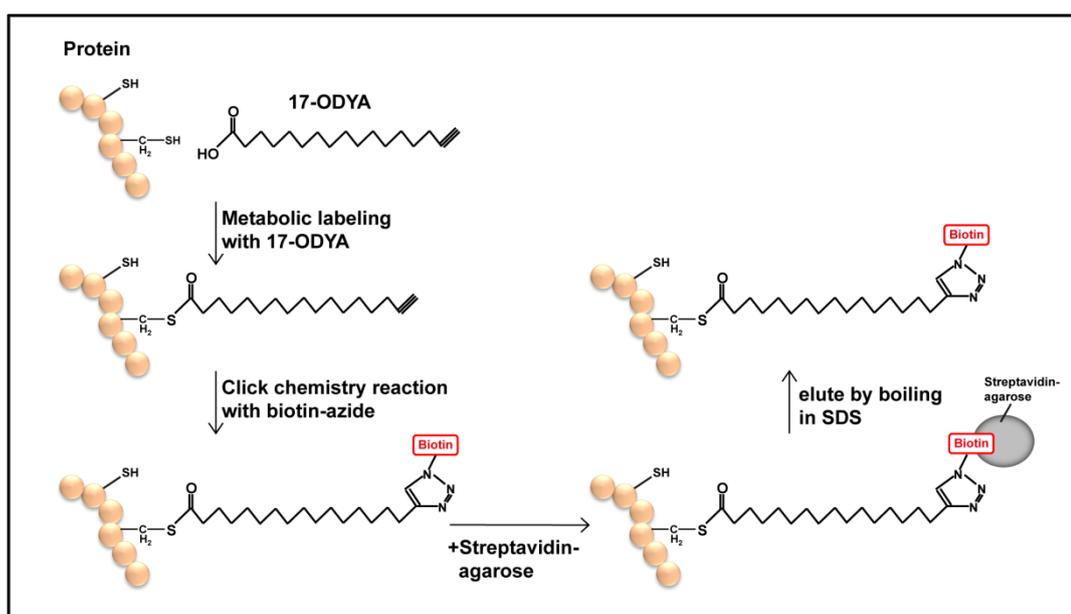


Figure 1.5: Schematic of the metabolic labelling and click chemistry method of purifying palmitoyl-proteins. This method relies on the *in vitro* metabolic labelling of proteins with a palmitic acid analog. Cells are metabolically labelled with the alkyne-containing palmitic acid analog, 17-octadecynoic acid (17-ODYA). Proteins are extracted from the cells and proteins labelled with 17-ODYA are reacted with biotin-azide under copper (I)-catalysed azide-alkyne cycloaddition (click chemistry) reaction conditions, resulting in the biotinylation of 17-ODYA-labelled proteins. The biotinylated proteins are purified by streptavidin-agarose affinity purification and eluted from the streptavidin-agarose beads by boiling in SDS. In the control sample, cells are mock-labelled by treatment with DMSO instead of 17-ODYA, and subjected to the same conditions as described above. Proteins which are significantly enriched in the 17-ODYA-treated samples compared to the mock DMSO-treated control samples can thus be classified as palmitoylated.

1.6.3. Comparison of palmitome purification methods

Both methods of palmitome purification described above bring with them their own set of false-positive identifications. ABE is based on the complete and irreversible blockage of free thiols with NEM in order to prevent the purification of non-palmitoylated free thiols. Insufficient blockage with NEM can thus give rise to false identifications. This can also result in the enrichment of highly abundant proteins whose free thiols may not have been completely blocked. Additionally, ABE will also enrich all other proteins which use thioester linkages, such as enzymes which use thioester intermediates as part of their reaction mechanisms [44, 50]. Meanwhile, metabolic labelling and click chemistry methods can result in the enrichment of proteins which have incorporated the 17-ODYA metabolic label, but are not in fact S-acylated. For example, glycosylphosphatidylinositol (GPI)-anchored proteins can incorporate the 17-ODYA label although they are not palmitoylated [50].

Although false-positive identifications are reduced in click chemistry purification methods, this method relies on the ability of the proteome of interest to be labelled with 17-ODYA, as well as on the efficiency of the metabolic labelling. Despite the larger background found with ABE, all the chemical exchange reactions of ABE take place post extraction of the proteome and thus does not have to rely on the *in vitro* or *in vivo* metabolic incorporation of an exogenous label.

Thus, both methods of palmitome purification have their strengths and weaknesses, and datasets from both methods, as with any large-scale profiling method, will require careful quantitation, analysis and validation. Nevertheless, these methods have still resulted in the identification of several hundreds of potential palmitoylated proteins in multiple organisms and can provide new insight into the extent of protein palmitoylation in biology.

1.7. Malaria: A major global parasitic disease

Malaria, a disease caused by the mosquito-borne Apicomplexan parasite, *Plasmodium*, still remains a wide-spread global health problem. According to recent reports from the World Health Organisation (WHO), in 2010, there were about 219 million cases of malaria and an estimated 660,000 deaths due to malaria, with approximately 90% of all malaria deaths occurring in the African continent [51]. Although this disease has been eradicated in some parts of the world, more than two billion of the world's population still live in malarial regions and thus are at risk of contracting malaria [52]. Across the years, there have been several combined efforts working towards the global eradication of malaria. These campaigns, which include the use of more efficacious drugs and anti-vector measures such as insecticide-treated bed nets, along with economic development and improved health

infrastructure in some countries, have helped to reduce the burden of malaria, but whether these methods alone will be able to completely eradicate the disease is contentious [53, 54], and the ongoing problem of *Plasmodium* drug resistance means that any gains that are being made currently are liable to erosion unless new drugs are continually developed.

1.8. The life cycle of *Plasmodium* parasites

Plasmodium parasites are members of the phylum Apicomplexa, which is characterised by the apical complex, a group of specialised organelles found at apical end of the invasive parasite stages, and the apicoplast, a rudimentary plastid descended from alga through secondary endosymbiosis [55]. Four species of *Plasmodium*, namely *Plasmodium falciparum*, *P. vivax*, *P. malariae* and *P. ovale*, are found to cause malaria in humans and are transmitted by female *Anopheles* mosquitoes. Of the four, *P. falciparum* is the most virulent and deadly, causing the most severe forms of malaria and responsible for most of the disease morbidity and mortality [53, 54]. More recently, another species of *Plasmodium*, *P. knowlesi*, which normally infects macaque monkeys, has also been found to infect humans in some areas [56].

Plasmodium parasites are obligate intracellular parasites and thus spend most of their life cycle within the cells of their host, evading the immune system and feeding on the host cell. The life cycle of these parasites in both the mosquito vector and the vertebrate host is complex, involving the invasion of a variety of cell types and many transitions from one developmental stage to another at different points in the life cycle.

The life cycle begins when the human host is bitten by an infected *Anopheles* mosquito, resulting in the transmission of sporozoites, a motile form of the parasite, into the bloodstream. The sporozoites migrate to the liver and invade hepatocytes, where they then develop into tens of thousands of merozoites. The merozoites then enter the bloodstream and rapidly invade erythrocytes, thus beginning the asexual intraerythrocytic stages. During the intraerythrocytic cycle, the parasites develop and replicate within the host erythrocytes, firstly from ring stages, to the highly metabolically active trophozoite stages, and lastly to the mature schizont stages where replication occurs to form daughter merozoites. Eventually, the infected erythrocytes burst, releasing the daughter merozoites into the bloodstream, where they are free to reinvade new erythrocytes. Each intraerythrocytic cycle takes approximately 48 hours and it is the synchronous rupture of infected erythrocytes every 48 hours which induces the periodic symptoms of malaria such as fever, chills and anaemia. Furthermore, a distinctive feature of *P. falciparum* is the ability of the intraerythrocytic

stages of the parasite to bind to endothelium and sequester in organs such as the brain, which can lead to the symptoms of cerebral malaria [53].

During the intraerythrocytic cycle, some parasites can differentiate into male and female sexual forms, known as gametocytes, in response to as yet unknown factors. These forms remain in the bloodstream until they are taken up by a mosquito during a blood meal. Once in the mosquito midgut, they transform into the activated male and female gametes and fertilization occurs, resulting in the production of the diploid ookinete. The motile ookinete crosses the mosquito midgut wall and forms an oocyst on the midgut basal lamina, where hundreds of haploid, genetically distinct sporozoites are formed. Rupture of the oocyst occurs, releasing the sporozoites, which migrate to the salivary glands and are transmitted to the human host during the next blood meal [53, 54]. A diagram adapted from [8], showing all the steps of the *Plasmodium* life cycle, in both the mosquito vector and the human host, is shown in Figure 1.6.

All of the steps taken by the parasite in order to complete its life cycle, even during just the intraerythrocytic cycle, from the invasion of erythrocytes [57, 58], to the remodelling of the infected erythrocyte to promote the survival of the parasite [59], and to the egress of the parasite in order to invade new erythrocytes [60], involve the use of many different proteins and multiple signalling cascades. This provides a potential source of targets which can be used for therapeutic interventions.

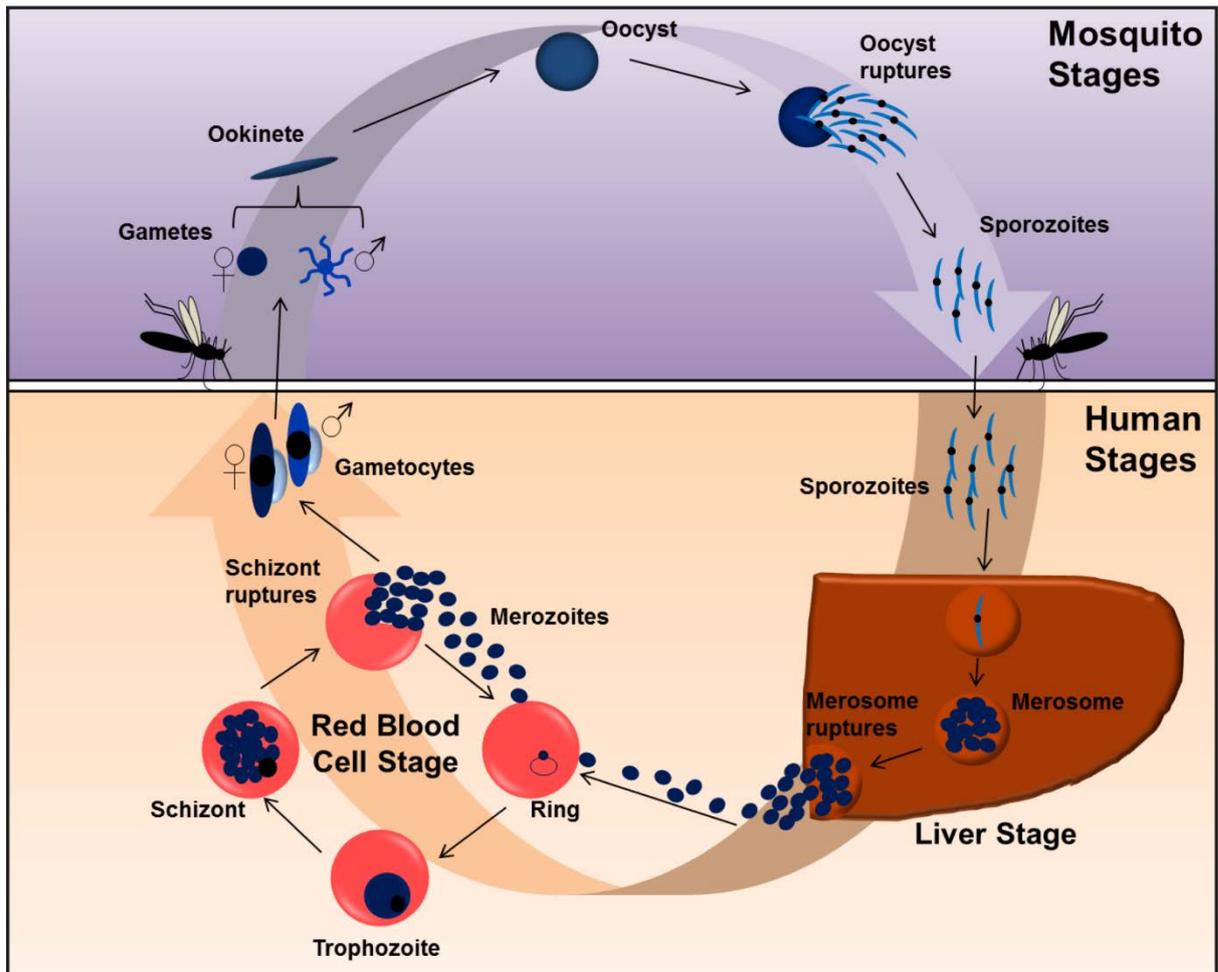


Figure 1.6: Schematic of the *Plasmodium* parasite's life cycle. The bite of an infected *Anopheles* mosquito transmits the motile sporozoite forms of the parasite into the host bloodstream. The sporozoites invade the hepatocyte cells of the liver, where they develop into tens of thousands of merozoites. The merozoites enter the bloodstream and invade erythrocytes. This is the beginning of the asexual intraerythrocytic stages of *Plasmodium*, where the parasites replicate and develop within host erythrocytes. The cycle begins at the ring stages, moves on to the highly metabolically active trophozoite stages, and ends at the mature schizont stages, where replication occurs to form daughter merozoites. The merozoites burst from the erythrocytes and re-enter the bloodstream, where they are free to invade new erythrocytes. During the intraerythrocytic stages, some parasites differentiate to form male and female gametocytes, the sexual forms of the parasites. The gametocytes are taken up by the mosquito during a blood-meal, and fertilisation occurs in the mosquito, producing the diploid ookinete. The ookinete crosses the mosquito midgut wall and forms the oocyst, where hundreds of haploid, genetically distinct sporozoites are formed. The sporozoites are released when the oocyst ruptures, and migrate to the mosquito salivary glands, where they can be transmitted to the human host during the next bloodmeal. This schematic was adapted from [8].

1.9. The need for new drug targets

There is currently no effective vaccine for any *Plasmodium* species, and all past efforts to control malaria have been based on the use of anti-malarial drugs for the prevention and treatment of infection, and insecticides for the elimination of the mosquito vector. However, these efforts have been repeatedly thwarted by the development of resistance of the parasite against the available anti-malarial drugs as well as resistance of the mosquito against insecticides [54, 61]. Drug resistance is one of the major problems facing malaria control. Over the past several decades, parasites have become resistant to most of the commonly used anti-malarial drugs, such as chloroquine, and the antifolate drugs such as sulfadoxine-pyrimethamine, both of which were once frontline therapies for malaria but are now largely ineffective due to widespread resistance [62]. Artemisinin derivatives and artemisinin-based combination therapies are now the standard frontline therapeutic interventions for malaria [51, 61]. However, sobering reports have now revealed that parasites with increased tolerance to these new drugs have emerged in South East Asia [51, 63] and with it, the fear that the continued widespread use of the artemisinin-based drugs could potentially result in the global spread of artemisinin resistance [63].

Furthermore, the lack of a vaccine against the malaria-causing parasites hinders efforts to completely eradicate malaria. Currently, the most promising candidate vaccine is RTS,S, which is based on the *P. falciparum* circumsporozoite protein and targets the pre-erythrocytic stages of infection. However, although the RTS,S vaccine has been shown to prevent infection and reduce the occurrence of severe malaria, this vaccine is still only approximately 50% efficacious [53, 54, 64]. There is therefore an urgent and on-going need for the identification of new drug targets, and the signalling pathways and proteins involved in regulating parasite growth and development through PTMs are one potential source of novel targets. Given that palmitoylation plays a role in regulating cellular processes in other eukaryotes, palmitoylation is a process that warrants investigating as a potential drug target in *Plasmodium* parasites.

1.10. Palmitoylation in *Plasmodium*

As detailed above, protein palmitoylation appears to play diverse and important roles in the regulation of protein localisation and function in many organisms, and is widely used in many cellular processes, as indicated by the large sizes of the palmitomes of the various organisms studied. It is gradually being discovered that this is also the case in *Plasmodium* parasites. Previously, only three proteins were experimentally shown to be palmitoylated in *P. falciparum*. Glideosome-associated

protein 45 (GAP45) and calcium-dependent protein kinase 1 (CDPK1) are components of the invasion motor complex, a multi-protein complex found at the inner membrane complex (IMC) of the parasite invasive stages, and thought to be required for the generation of the force required to actively propel the parasite into the host cell during invasion [57, 58, 65]. Calpains are cysteine proteases with well-conserved catalytic domains that are thought to be involved in various cellular processes such as cell signalling and migration, cell differentiation and development, and cell-cycle regulation, although the exact role of the calpains is still poorly characterised [66]. These three proteins -GAP45, CDPK1 and calpain- are dually-acylated in *P. falciparum*, modified by both a myristate and a palmitate group, and these acyl modifications appear to be important for the membrane attachment, targeting and localisation of these proteins [67-69].

Recently however, using both ABE and metabolic labelling followed by click chemistry methods, coupled with quantitative mass spectrometry, it was discovered that protein palmitoylation in *Plasmodium* is more extensive than previously known, with more than 400 putative palmitoyl-proteins found in the asexual intraerythrocytic schizont stages of *P. falciparum* [50]. Even more intriguing was the discovery that treatment of schizont-stage parasites with 2-bromopalmitate (2-BMP), a small molecule compound which inhibits palmitoylation [70], for 6 hours appears to have different effects on different palmitoyl-proteins. Some proteins are unaffected and remain palmitoylated, while other proteins appear to be either moderately or severely affected by 2-BMP treatment. These different effects induced by the inhibition of palmitoylation point to the existence of both stable as well as dynamic palmitoylation in *Plasmodium* [50].

The proteins identified in the *P. falciparum* palmitome include those which fall into the general types of palmitoyl-proteins described above, as well as proteins which are part of processes known to involve protein palmitoylation in other organisms [8, 50]. However, more interestingly, a significant number of proteins identified as palmitoylated are proteins which are involved in parasite specific processes such as invasion, drug resistance, cytoadherence and parasite development [50].

Parasite invasion of the host erythrocyte is one of the parasite specific processes which appear to involve palmitoylation. Invasion is a complicated process involving the recognition and binding of surface receptors on the erythrocyte, and an actin-myosin invasion motor complex, which generates the force required to push the parasite into the host erythrocyte [57, 58]. All the steps of the invasion process involve many different proteins, and bioinformatics analysis suggest that more than 400 proteins play a role in the invasion of parasites into erythrocytes [71]. Of these invasion-associated proteins, 74 are identified as palmitoylated [50]. The fact that so many of the invasion-linked proteins appear to be palmitoylated suggests that palmitoylation may be critically involved in

regulating the invasion process of the parasite. This suggestion is supported by the fact that inhibition of palmitoylation with 2-BMP significantly inhibits erythrocyte invasion [50]. In fact, palmitoylation appears to be directly involved in maintaining the stability of components of the invasion motor complex such as GAP45 and myosin A tail domain-interacting protein (MTIP), as treatment of parasites with 2-BMP results in the degradation of both GAP45 and MTIP [50]. Although exactly how protein palmitoylation regulates invasion is still unknown, it is clear that this PTM appears to be required for the normal function of some of the proteins responsible for parasite invasion.

Perhaps more surprising was the discovery that a fraction of schizont proteins identified as palmitoylated are parasite exported proteins [50]. Following invasion, several hundred parasite proteins are exported across the parasite plasma membrane to specific sites in the erythrocyte where they modify and remodel the erythrocyte in order to support parasite survival and development. This remodelling results in structural and morphological changes within the erythrocyte, such as increased rigidity and adhesiveness, and decreased deformability of the plasma membrane, which in turn leads to cytoadherence and sequestering of infected erythrocytes, a pathological characteristic of *P. falciparum* infection [59].

Some of these parasite exported proteins, including the major virulence factor, *P. falciparum* erythrocyte membrane protein 1 (PfEMP1), the protein responsible for anchoring PfEMP1 to the surface of the erythrocyte, knob-associated histidine-rich protein (KAHRP), and mature-parasite-infected erythrocyte surface antigen (MESA) were identified as palmitoylated [50]. All of the parasite exported proteins require trafficking across several membranes in order to move from the parasite to its site in the host erythrocyte. Even within the erythrocyte, the exported proteins are either retained in the Maurer's clefts, parasite organelles which are tethered to the erythrocyte membrane, or transported to the erythrocyte membrane itself. The fact that some of these exported proteins are palmitoylated raises the question as to whether palmitoylation is involved in regulating the trafficking and localisation of these proteins from the parasite to the erythrocyte.

This initial work suggests for the first time that palmitoylation plays an important role in regulating multiple aspects of *Plasmodium* biology. To further our understanding of *Plasmodium* palmitoylation, two critical questions need to be addressed. Firstly, where and when the *Plasmodium* PATs are expressed, as well as the substrates for individual PATs, are still unknown. Secondly, although a list of palmitoylated proteins in *Plasmodium* now exists, the exact sites which are palmitoylated are still not known. The work described in this dissertation was performed in an attempt to answer these questions.

1.11. Aims and objectives

The work described in this dissertation can be divided into two main aims as detailed below:

1.11.1. Aim 1: To determine the sites of palmitoylation in Plasmodium.

The *P. falciparum* palmitome has now been purified from intraerythrocytic schizont stages, resulting in the identification of over 400 putative palmitoyl-proteins [50]. However, the specific cysteines which are palmitoylated are still unknown for most *P. falciparum* palmitoyl-proteins and due to the lack of a consensus sequence for palmitoylation, are impossible to predict in the majority of cases. As is the case for all eukaryotes, while some proteins identified in the *P. falciparum* palmitome fall into the defined classes of palmitoyl-proteins described above where palmitoylation sites are predictable (for example, 44% of the palmitome are TM-domain-containing proteins, and 8 proteins identified in the total palmitome contain the N-myristoylation motif followed by a cysteine residue within 10 amino acids of the motif [50]), in the majority of cases, palmitoylation sites are impossible to predict. Information on the sites of palmitoylation could be used in the generation of testable hypotheses about the function of palmitoylation for a particular protein of interest, and would also provide critical information for subsequent studies of DHHC-PAT substrate specificity.

The first goal of this project was therefore to expand further on the analysis of the *P. falciparum* palmitome by developing methods to allow the identification of specific palmitoylated cysteines on individual palmitoyl-proteins. This was achieved by adapting the ABE methodology to purify palmitoylated peptides, rather than intact palmitoyl-proteins. As such peptides should by definition contain a palmitoylated cysteine residue, a palmitome-wide list of palmitoylated cysteines could be generated.

1.11.2. Aim 2: To investigate the molecular basis of palmitoylation by characterising the Plasmodium PATs.

Little is currently known about PATs in *Plasmodium*. A protein-protein BLAST search of the *P. falciparum* genome using the 50 amino acid DHHC cysteine-rich domain [5] reveals 12 potential DHHC-PATs, all of which share some homology with DHHC-PATs from other eukaryotes. These DHHC-PATs also have homologues in other *Plasmodium* species, although one of these proteins, PFB0140w, has no homologue in rodent *Plasmodium* species, which only have 11 DHHC-PATs. Some *P. falciparum* DHHC-PATs appear to be expressed constitutively throughout the life cycle stages, while other DHHC-PATs appear to be more regulated, with higher expression in particular life cycle stages

[8]. Only one of the *P. falciparum* DHHC-PATs (PFC0160w) has been characterised, and is found to be localised to the Golgi [72], one of the common localisations of DHHC proteins in other organisms. The localisation of the other DHHC-domain-containing proteins in *Plasmodium* is currently unknown, and whether they actually act as PATs in *Plasmodium* has never been formally tested.

In contrast to the DHHC proteins, only one protein is identified by protein-protein BLAST search as a potential MBOAT-protein in *P. falciparum*. However, this protein (PFC0995c) has been characterised to act as a diacylglycerol O-acyltransferase [73], an activity common to MBOAT family proteins in other organisms [15]. It is unknown whether this protein also acts as an MBOAT-PAT in *Plasmodium*.

Given this significant gap in our understanding of palmitoylation in *Plasmodium*, the second aim of this project was to characterise for the first time the *Plasmodium* DHHC and MBOAT proteins. This was achieved by determining the different localisations of as many of these proteins as possible in the parasite, as well as by determining the effect of deleting DHHC-PAT encoding genes on parasite growth and development. In addition, in order to formally prove that *Plasmodium* DHHC and MBOAT proteins have palmitoyl transferase activity, a novel PAT activity assay was developed using a mammalian cell expression system. This assay was used to establish enzymatic activity, and to carry out investigations into the substrate-specificity of individual *P. falciparum* DHHC-PATs.

References:

1. Resh, M.D., *Palmitoylation of ligands, receptors, and intracellular signaling molecules*. Sci STKE, 2006. **2006**(359): p. re14.
2. Farazi, T.A., *The Biology and Enzymology of Protein N-Myristoylation*. Journal of Biological Chemistry, 2001. **276**(43): p. 39501-39504.
3. Smotrýs, J.E. and M.E. Linder, *Palmitoylation of intracellular signaling proteins: regulation and function*. Annu Rev Biochem, 2004. **73**: p. 559-587.
4. Linder, M.E. and R.J. Deschenes, *Palmitoylation: policing protein stability and traffic*. Nat Rev Mol Cell Biol, 2007. **8**(1): p. 74-84.
5. Mitchell, D.A., et al., *Protein palmitoylation by a family of DHHC protein S-acyltransferases*. J Lipid Res, 2006. **47**(6): p. 1118-1127.
6. Lobo, S., *Identification of a Ras Palmitoyltransferase in Saccharomyces cerevisiae*. Journal of Biological Chemistry, 2002. **277**(43): p. 41268-41273.
7. Roth, A.F., *The yeast DHHC cysteine-rich domain protein Akr1p is a palmitoyl transferase*. The Journal of Cell Biology, 2002. **159**(1): p. 23-28.
8. Jones, M.L., C.L. Tay, and J.C. Rayner, *Getting stuck in: protein palmitoylation in Plasmodium*. Trends in Parasitology, 2012. **28**(11): p. 496-503.
9. Politis, E.G., *Transmembrane Topology of the Protein Palmitoyl Transferase Akr1*. Journal of Biological Chemistry, 2005. **280**(11): p. 10156-10163.

10. Mitchell, D.A., et al., *Mutational Analysis of Saccharomyces cerevisiae Erf2 Reveals a Two-step Reaction Mechanism for Protein Palmitoylation by DHHC Enzymes*. Journal of Biological Chemistry, 2010. **285**(49): p. 38104-38114.
11. Gonzalez Montoro, A., et al., *Specificity of transmembrane protein palmitoylation in yeast*. PLoS One, 2011. **6**(2): p. e16969.
12. Smotrys, J.E., *The vacuolar DHHC-CRD protein Pfa3p is a protein acyltransferase for Vac8p*. The Journal of Cell Biology, 2005. **170**(7): p. 1091-1099.
13. Hou, H., et al., *Analysis of DHHC Acyltransferases Implies Overlapping Substrate Specificity and a Two-Step Reaction Mechanism*. Traffic, 2009. **10**(8): p. 1061-1073.
14. Roth, A.F., et al., *Global analysis of protein palmitoylation in yeast*. Cell, 2006. **125**(5): p. 1003-1013.
15. Hofmann, K., *A superfamily of membrane-bound O-acyltransferases with implications for wnt signaling*. Trends Biochem Sci, 2000. **25**(3): p. 111-112.
16. Guo, Z.Y., *The Active Site His-460 of Human Acyl-coenzyme A:Cholesterol Acyltransferase 1 Resides in a Hitherto Undisclosed Transmembrane Domain*. Journal of Biological Chemistry, 2005. **280**(45): p. 37814-37826.
17. Mann, R.K. and P.A. Beachy, *Novel lipid modifications of secreted protein signals*. Annu Rev Biochem, 2004. **73**: p. 891-923.
18. Micchelli, C.A., et al., *Rasp, a putative transmembrane acyltransferase, is required for Hedgehog signaling*. Development, 2002. **129**(4): p. 843-851.
19. Miura, G.I., et al., *Palmitoylation of the EGFR ligand Spitz by Rasp increases Spitz activity by restricting its diffusion*. Dev Cell, 2006. **10**(2): p. 167-176.
20. Galli, L.M., et al., *Porcupine-mediated lipid-modification regulates the activity and distribution of Wnt proteins in the chick neural tube*. Development, 2007. **134**(18): p. 3339-3348.
21. Buglino, J.A. and M.D. Resh, *Identification of conserved regions and residues within Hedgehog acyltransferase critical for palmitoylation of Sonic Hedgehog*. PLoS One. **5**(6): p. e11195.
22. Zeidman, R., C.S. Jackson, and A.I. Magee, *Protein acyl thioesterases (Review)*. Mol Membr Biol, 2009. **26**(1): p. 32-41.
23. Hofmann, S.L., *Lysosomal Targeting of Palmitoyl-protein Thioesterase*. Journal of Biological Chemistry, 1996. **271**(26): p. 15831-15836.
24. Duncan, J.A., *A Cytoplasmic Acyl-Protein Thioesterase That Removes Palmitate from G Protein alpha Subunits and p21RAS*. Journal of Biological Chemistry, 1998. **273**(25): p. 15830-15837.
25. Yeh, D.C., *Depalmitoylation of Endothelial Nitric-oxide Synthase by Acyl-protein Thioesterase 1 Is Potentiated by Ca²⁺-Calmodulin*. Journal of Biological Chemistry, 1999. **274**(46): p. 33148-33154.
26. Duncan, J.A., *Characterization of Saccharomyces cerevisiae Acyl-protein Thioesterase 1, the Enzyme Responsible for G Protein alpha Subunit Deacylation in Vivo*. Journal of Biological Chemistry, 2002. **277**(35): p. 31740-31752.
27. Dekker, F.J., et al., *Small-molecule inhibition of APT1 affects Ras localization and signaling*. Nat Chem Biol, 2010. **6**(6): p. 449-56.
28. Resh, M.D., *Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins*. Biochim Biophys Acta, 1999. **1451**(1): p. 1-16.
29. Rocks, O., et al., *An acylation cycle regulates localization and activity of palmitoylated Ras isoforms*. Science, 2005. **307**(5716): p. 1746-52.
30. Choy, E., et al., *Endomembrane Trafficking of Ras The CAAX Motif Targets Proteins to the ER and Golgi*. Cell, 1999. **98**(1): p. 69-80.
31. Greaves, J. and L.H. Chamberlain, *Palmitoylation-dependent protein sorting*. The Journal of Cell Biology, 2007. **176**(3): p. 249-254.
32. Roy, S., et al., *Individual Palmitoyl Residues Serve Distinct Roles in H-Ras Trafficking, Microlocalization, and Signaling*. Mol Cell Biol, 2005. **25**(15): p. 6722-6733.

33. Balamuth, F., J.L. Brogdon, and K. Bottomly, *CD4 raft association and signaling regulate molecular clustering at the immunological synapse site*. J Immunol, 2004. **172**(10): p. 5887-92.
34. Fragoso, R., et al., *Lipid raft distribution of CD4 depends on its palmitoylation and association with Lck, and evidence for CD4-induced lipid raft aggregation as an additional mechanism to enhance CD3 signaling*. J Immunol, 2003. **170**(2): p. 913-21.
35. Arcaro, A., et al., *Essential role of CD8 palmitoylation in CD8 coreceptor function*. J Immunol, 2000. **165**(4): p. 2068-76.
36. Hayashi, T., G. Rumbaugh, and R.L. Huganir, *Differential Regulation of AMPA Receptor Subunit Trafficking by Palmitoylation of Two Distinct Sites*. Neuron, 2005. **47**(5): p. 709-723.
37. El-Husseini Ael, D., et al., *Synaptic strength regulated by palmitate cycling on PSD-95*. Cell, 2002. **108**(6): p. 849-63.
38. Valdez-Taubas, J. and H. Pelham, *Swf1-dependent palmitoylation of the SNARE Tlg1 prevents its ubiquitination and degradation*. Embo J, 2005. **24**(14): p. 2524-32.
39. Reggiori, F. and H.R.B. Pelham, *A transmembrane ubiquitin ligase required to sort membrane proteins into multivesicular bodies*. Nature Cell Biology, 2002. **4**(2): p. 117-123.
40. Yanai, A., et al., *Palmitoylation of huntingtin by HIP14 is essential for its trafficking and function*. Nature Neuroscience, 2006. **9**(6): p. 824-831.
41. Chen, M.H., et al., *Palmitoylation is required for the production of a soluble multimeric Hedgehog protein complex and long-range signaling in vertebrates*. Genes Dev, 2004. **18**(6): p. 641-659.
42. Zhang, F.L. and P.J. Casey, *Protein prenylation: molecular mechanisms and functional consequences*. Annu Rev Biochem, 1996. **65**: p. 241-69.
43. Drisdell, R.C., et al., *Assays of protein palmitoylation*. Methods, 2006. **40**(2): p. 127-134.
44. Tom, C.T.M.B. and B.R. Martin, *Fat Chance! Getting a Grip on a Slippery Modification*. ACS Chemical Biology, 2013. **8**(1): p. 46-57.
45. Martin, B.R. and B.F. Cravatt, *Large-scale profiling of protein palmitoylation in mammalian cells*. Nat Methods, 2009. **6**(2): p. 135-138.
46. Yang, W., et al., *Proteome scale characterization of human S-acylated proteins in lipid raft-enriched and non-raft membranes*. Mol Cell Proteomics. **9**(1): p. 54-70.
47. Kang, R., et al., *Neural palmitoyl-proteomics reveals dynamic synaptic palmitoylation*. Nature, 2008. **456**(7224): p. 904-909.
48. Wan, J., et al., *Palmitoylated proteins: purification and identification*. Nat Protoc, 2007. **2**(7): p. 1573-1584.
49. Martin, B.R., et al., *Global profiling of dynamic protein palmitoylation*. Nat Methods, 2012. **9**(1): p. 84-9.
50. Jones, Matthew L., et al., *Analysis of Protein Palmitoylation Reveals a Pervasive Role in Plasmodium Development and Pathogenesis*. Cell Host & Microbe, 2012. **12**(2): p. 246-258.
51. WHO, *World Malaria Report 2012*, W.H. Organisation, Editor. 2012: Geneva, Switzerland
52. Snow, R.W., et al., *The global distribution of clinical episodes of Plasmodium falciparum malaria*. Nature, 2005. **434**(7030): p. 214-7.
53. Winzeler, E.A., *Malaria research in the post-genomic era*. Nature, 2008. **455**(7214): p. 751-756.
54. Greenwood, B.M., et al., *Malaria: progress, perils, and prospects for eradication*. J Clin Invest, 2008. **118**(4): p. 1266-1276.
55. Baumeister, S., et al., *The malaria parasite Plasmodium falciparum: cell biological peculiarities and nutritional consequences*. Protoplasma. **240**(1-4): p. 3-12.
56. Singh, B., et al., *A large focus of naturally acquired Plasmodium knowlesi infections in human beings*. The Lancet, 2004. **363**(9414): p. 1017-1024.
57. Cowman, A.F. and B.S. Crabb, *Invasion of red blood cells by malaria parasites*. Cell, 2006. **124**(4): p. 755-766.

58. Baum, J., et al., *Host-cell invasion by malaria parasites: insights from Plasmodium and Toxoplasma*. Trends in Parasitology, 2008. **24**(12): p. 557-563.
59. Maier, A.G., et al., *Malaria parasite proteins that remodel the host erythrocyte*. Nat Rev Microbiol, 2009. **7**(5): p. 341-354.
60. Blackman, M.J., *Malarial proteases and host cell egress: an 'emerging' cascade*. Cell Microbiol, 2008. **10**(10): p. 1925-1934.
61. Sanchez, C.P., et al., *Transporters as mediators of drug resistance in Plasmodium falciparum*. Int J Parasitol. **40**(10): p. 1109-1118.
62. Baird, J.K., *Effectiveness of antimalarial drugs*. N Engl J Med, 2005. **352**(15): p. 1565-77.
63. Noedl, H., et al., *Artemisinin Resistance in Cambodia: A Clinical Trial Designed to Address an Emerging Problem in Southeast Asia*. Clinical Infectious Diseases, 2010. **51**(11): p. e82-e89.
64. Greenwood, B.M. and G.A. Targett, *Malaria vaccines and the new malaria agenda*. Clin Microbiol Infect, 2011. **17**(11): p. 1600-7.
65. Green, J.L., et al., *The Motor Complex of Plasmodium falciparum: PHOSPHORYLATION BY A CALCIUM-DEPENDENT PROTEIN KINASE*. Journal of Biological Chemistry, 2008. **283**(45): p. 30980-30989.
66. Russo, I., et al., *A calpain unique to alveolates is essential in Plasmodium falciparum and its knockdown reveals an involvement in pre-S-phase development*. Proceedings of the National Academy of Sciences, 2009. **106**(5): p. 1554-1559.
67. Rees-Channer, R.R., et al., *Dual acylation of the 45 kDa gliding-associated protein (GAP45) in Plasmodium falciparum merozoites*. Mol Biochem Parasitol, 2006. **149**(1): p. 113-116.
68. Moskes, C., et al., *Export of Plasmodium falciparum calcium-dependent protein kinase 1 to the parasitophorous vacuole is dependent on three N-terminal membrane anchor motifs*. Mol Microbiol, 2004. **54**(3): p. 676-691.
69. Russo, I., A. Oksman, and D.E. Goldberg, *Fatty acid acylation regulates trafficking of the unusual Plasmodium falciparum calpain to the nucleolus*. Mol Microbiol, 2009. **72**(1): p. 229-45.
70. Jennings, B.C., et al., *2-Bromopalmitate and 2-(2-hydroxy-5-nitro-benzylidene)-benzo[b]thiophen-3-one inhibit DHHC-mediated palmitoylation in vitro*. J Lipid Res, 2009. **50**(2): p. 233-242.
71. Hu, G., et al., *Transcriptional profiling of growth perturbations of the human malaria parasite Plasmodium falciparum*. Nat Biotechnol, 2010. **28**(1): p. 91-8.
72. Seydel, K.B., et al., *Plasmodium falciparum: Characterization of a late asexual stage Golgi protein containing both ankyrin and DHHC domains*. Experimental Parasitology, 2005. **110**(4): p. 389-393.
73. Palacpac, N.M., et al., *Evidence that Plasmodium falciparum diacylglycerol acyltransferase is essential for intraerythrocytic proliferation*. Biochem Biophys Res Commun, 2004. **321**(4): p. 1062-1068.

Chapter 2

Materials and methods

2.1. Generation of plasmid constructs

2.1.1. Generation of *Plasmodium falciparum* plasmid constructs

Triple-HA (3-HA)-tagged plasmid constructs: An approximately 1000bp fragment from the 3' region of the open reading frame of the gene of interest, excluding the STOP codon, was amplified by polymerase chain reaction (PCR) from the genomic DNA of *Plasmodium falciparum* strain 3D7 using the PCR primers (containing the appropriate restriction enzyme sites) listed in Table 2.1. All PCR reactions were performed using KOD Hot Start DNA polymerase (Novagen) in a 50 μ L reaction according to manufacturer's instructions. The PCR cycling parameters were as follows:

Step	Temperature (°C)	Time (min)
1	95	5
2	95	1
3	45-58	0.5
4	65	2
5	Go to 2, repeat 35 times	
6	65	5
7	4	Forever

The PCR amplified DNA was digested with the appropriate restriction enzymes overnight at 37°C in a 60 μ L reaction. The pCAM-BSD-3HA vector, which allows the endogenous tagging of the gene of interest via single-crossover recombination (Figure 5.1A) [1], was also digested with the same restriction enzymes overnight at 37°C in a 40 μ L reaction. All restriction enzymes used were from New England Biolabs (NEB) and restriction digestions were performed according to manufacturer's instructions. The digested PCR amplified insert DNA was collected by ethanol precipitation. The digested pCAM-BSD-3HA vector was resolved by agarose gel electrophoresis using a 1% agarose gel, with electrophoresis performed at a constant voltage of 120V until the DNA was adequately separated, and the required DNA band was gel purified using a QIAquick gel extraction kit (QIAGEN) according to manufacturer's instructions. The digested PCR amplified insert DNA was introduced into the digested pCAM-BSD-3HA vector using T₄ DNA ligase (NEB), in a 10 μ L ligation reaction, and incubated for 2 hours at 24°C, according to manufacturer's instructions. The products of the ligation reaction were introduced into One Shot® TOP10 chemically competent *E. coli* (Life Technologies) by heat shock for 45 seconds at 42°C, according to manufacturer's instructions. Positive colonies were selected for using LB-agar plates containing 50 μ g/mL of ampicillin (Sigma-Aldrich). DNA was extracted from individual colonies using a QIAprep Spin Miniprep Kit (QIAGEN) according to manufacturer's instructions, and tested for the presence of the plasmid of interest by restriction

digestion with the appropriate restriction enzymes. DNA containing the desired plasmid was re-introduced into One Shot® TOP10 chemically competent *E.coli* as described above. Single colonies were inoculated into 100 mL of LB broth containing 50 µg/mL ampicillin and grown overnight at 37°C with shaking at 250 rpm. DNA was extracted from the overnight cultures using a GenElute™ HP Plasmid Midiprep Kit (Sigma-Aldrich) and the presence of the desired plasmid again tested by restriction digestion with the appropriate restriction enzymes. A schematic of the pCAM-BSD-3HA plasmid construct is shown in Figure 2.1A.

Triple-HA-tagged constructs	Primers used (5' - 3')		Restriction sites	Vector
PfiBOAT	Forward	GCctgagCATGAGGTGTGTTATGAAGCTAGG	PstI and HindIII	pCAM-BSD-2HA
	Reverse	CCCTCGACTCAaagctTTAAGCGTAATCTGG		
PfdHHC3	Forward	GCctgagGGATCGTATAATGATATATCC	PstI and XhoI	pCAM-BSD-3HA
	Reverse	GCGctgagATAATTTTTTAATGTAATTTCTCC		
PfdHHC5	Forward	GCctgagCTATACCCACTGATGTTGC		
	Reverse	GCGctgagTTCAATGTTTCATTCGTTTAATCGC		
PfdHHC7	Forward	GCctgagGGAGAACGAAGACATTGTAAATGG		
	Reverse	GCGctgagTATATTTGTTTTATTGGAATAATTTCC		
PfdHHC8	Forward	GCctgagGATTGGTAATTGTGTAGGAAATCG		
	Reverse	GCGctgagGAGCGTTTTGAAAATGTTAATATATTTTTGAC		
PfdHHC9	Forward	GCctgagCCGATTTTTATAAATGCTCTTTGTGCC		
	Reverse	GCGctgagATCTCCATCTTCTTTATGTTTTC		

Table 2.1: Primers used for the generation of *P. falciparum* triple-HA-tagged plasmid constructs. The restriction enzyme sites are highlighted in red and indicated, along with the vector used, in the columns on the right.

Double-crossover knock-out plasmid constructs: Generation of the double-crossover knock-out plasmid constructs were performed essentially as described above, except that an approximately 600-800bp fragment from both the N-terminal and C-terminal regions of the open reading frame of the gene of interest was amplified by PCR and introduced, one fragment at a time, into multiple cloning site 1 and multiple cloning site 2 of the pCC1 vector respectively, which allows the disruption of a gene of interest by double-crossover recombination (Figure 5.4A) [2]. PCR cycling parameters were as described above and the restriction enzymes used are as described in Table 2.2. In this case, ligation reactions were performed in 20 µL volumes and left overnight at 24°C, after which the products of the ligation reactions were collected by ethanol precipitation, before being introduced into competent cells. A schematic of the double-crossover knock-out plasmid construct is shown in Figure 2.1B.

Knock-out constructs	Fragment	Primers used (5' - 3')		Restriction sites	Vector
PfMBOAT	Nt	Forward	GC ccg cggaATAATAAAAAATGATGAGGGAC	SacII and SpeI	pCC-1
		Reverse	GCG actag tCATAGCTATAAATTGAGGGTGG		
	Ct	Forward	GC gaattc CCATTAATAAATTAGGTTTTACCGG	EcoRI and AvrII	
		Reverse	GCG cctagg CACTTTTGAGATATGTGTCCATGGTGG		
PfDHH3	Nt	Forward	GC ccg cgGCTTATAATATAATCCCAAATTTTCC	SacII and SpeI	
		Reverse	GCG actag tCACATTTATTACAGACTTTACAGTGC		
	Ct	Forward	GC gaattc CCTTAATACATCTCACCAC	EcoRI and NcoI	
		Reverse	GCG cctagg CCTCTTTCAATTCGTGTTTTTCC		
PfDHH5	Nt	Forward	GC ccg cgCAACATCATGTGTGTCCATTAACG	SacII and SpeI	
		Reverse	GCG actag tGCTGCCCAATCAATTACCGTAC		
	Ct	Forward	GC gaattc GTATGGGTATTGAAAATGCTAG	EcoRI and AvrII	
		Reverse	GCG cctagg CTTATAATTACACTTAATAAATTACCAC		
PfDHH7	Nt	Forward	GC ccg cgGACAAAGAATAAGAATGTTGAAG	SacII and SpeI	
		Reverse	GCG actag tCCTTATAAATTAAGGCAAACAGTG		
	Ct	Forward	GC gaattc CGTTCCTTATCCCTCATAGTAACATG	EcoRI and NcoI	
		Reverse	GCG cctagg GTTTTTATTGGAATAATTCCTCTACTG		
PfDHH8	Nt	Forward	GC ccg cgGGTTCAAATACATGATGCAAAATCTCC	SacII and SpeI	
		Reverse	GCG actag tCCAGGGTCACAAAATGCTGTAG		
	Ct	Forward	GC gaattc CCTTTCCATCCTCCCTTTTG	EcoRI and NcoI	
		Reverse	GCG cctagg GCGTTTTGAAAATGTTAATATATTTTTG		
PfDHH9	Nt	Forward	GC ccg cgGAATAATTATTTGGCATTATC	SacII and SpeI	
		Reverse	GCG actag tCCTGCAGTGATGAGCTCTCTG		
	Ct	Forward	GC cctagg cgctatctcttttagATGAACTGTC	EcoRI and AvrII	
		Reverse	GCG cctagg CTCCATCTTCTTTATGTTTTTC		

Table 2.2: Primers used for the generation of *P. falciparum* double-crossover knock-out plasmid constructs. The restriction enzyme sites are highlighted in red and indicated, along with the vector used, in the columns on the right.

2.1.2. Generation of *Plasmodium berghei* plasmid constructs

P. berghei triple-HA-tagging and knock-out targeting vectors were obtained from the open access *PlasmoGEM* resource hosted at the Wellcome Trust Sanger Institute (<http://plasmogem.sanger.ac.uk>). Construction of the *PlasmoGEM* vectors are as previously described [3], and the transfection of these vectors into *P. berghei* strain ANKA 2.34 purified schizonts by electroporation was performed as previously described [4]. All *P. berghei* transfection procedures and the extraction of blood from infected mice were kindly performed by Ellen Bushell, Billker and Rayner labs. The *PlasmoGEM* vectors used here are listed in Table 2.3.

PbDHHc	<i>Plasmo</i> DB Gene names	Vector Design	<i>Plasmo</i> GEM transfection vector names
PbDHHc3	PBANKA_092730	TAG	PbGEM-094114
		KO	PbGEM-111866
PbDHHc4	PBANKA_142090	TAG	PbGEM-065194
		KO	PbGEM-065186
PbDHHc5	PBANKA_133780	TAG	PbGEM-058319
		KO	PbGEM-072266
PbDHHc6	PBANKA_083330	TAG	PbGEM-112088
		KO	PbGEM-027807
PbDHHc7	PBANKA_124300	TAG	PbGEM-052430
		KO	PbGEM-104070
PbDHHc8	PBANKA_141970	TAG	PbGEM-065002
		KO	PbGEM-225987
PbDHHc9	PBANKA_093210	TAG	PbGEM-121234
		KO	PbGEM-121226
PbDHHc10	PBANKA_051200	TAG	PbGEM-112097
		KO	PbGEM-015165
PbDHHc11	PBANKA_031260	TAG	PbGEM-225995
		KO	PbGEM-121242

Table 2.3: PlasmoGEM vectors used for the generation of 3HA-tagged and knock-out transgenic strains in *P. berghei*.

2.1.3. Generation of Human Embryonic Kidney 293E (HEK293E) expression plasmid constructs

The sequences coding for the *P. falciparum* proteins of interest were codon-optimised for expression in Human embryonic kidney 293 (HEK293) cells using the *GeneART* gene synthesis service (Life Technologies). *Plasmodium* parasites lack the enzymatic machinery required for the N-linked glycosylation of proteins, thus all *P. falciparum* proteins lack N-linked glycosylation moieties. All potential N-linked glycosylation sites (N-X-S/T) were therefore removed from the *P. falciparum* proteins of interest by substituting the serine/threonine residues at these sites with alanine. The sequence coding for either the FLAG[®] tag or the c-Myc tag was added to the C-terminal region of the proteins of interest, followed by a STOP codon. The entire sequence of the protein of interest, including the FLAG/c-Myc tag and the STOP codon, was then flanked by unique NotI (5'-end) and AscI (3'-end) restriction sites in order to introduce the sequences into a pTT3-based expression vector, which also contains a region coding for the immunoglobulin-like domains 3 and 4 of rat CD4 [5], using the *GeneART* gene synthesis service. As CD4 was not required in this case, the STOP codon was introduced directly after the FLAG or c-Myc tag, upstream of the CD4 region. A schematic of the HEK293 expression construct containing the c-Myc/FLAG-tagged codon-optimised *P. falciparum* gene of interest is shown in Figure 2.1C.

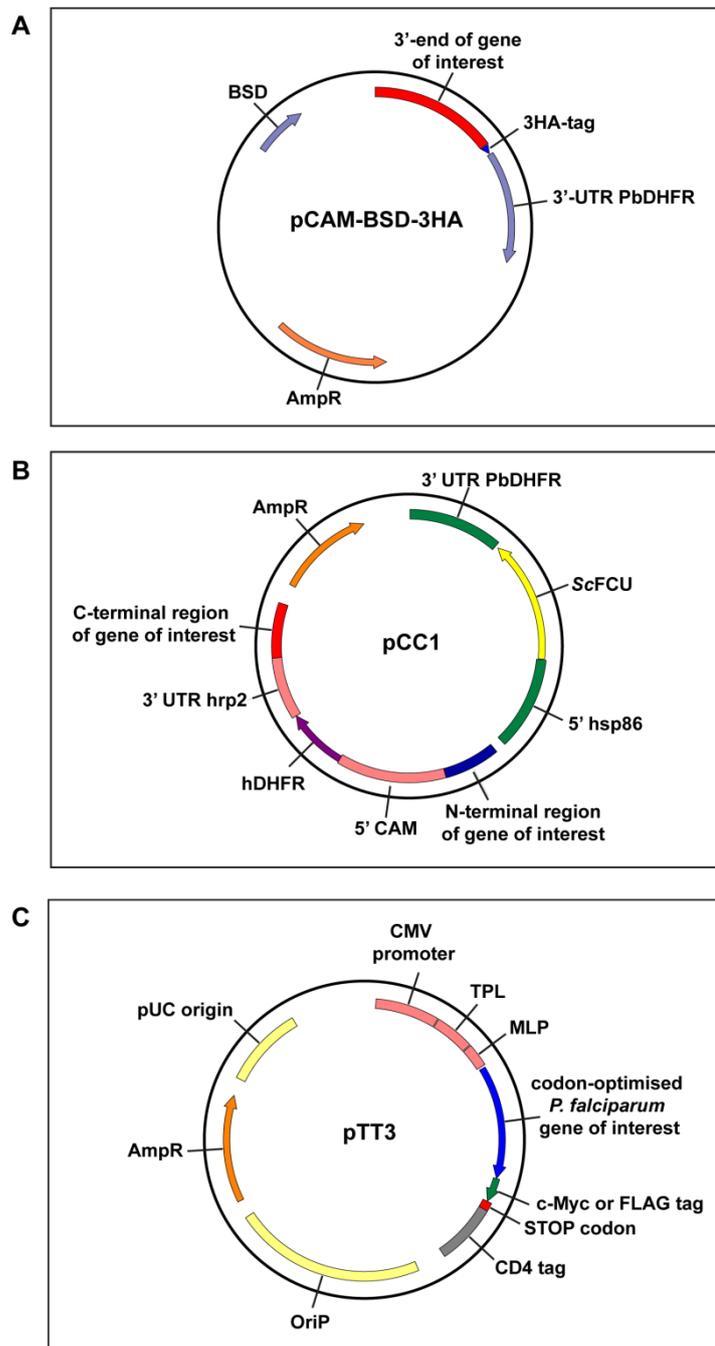


Figure 2.1: Schematic of the *P. falciparum* and HEK293E expression plasmid constructs. (A) Schematic of the pCAM-BSD-3HA plasmid construct used for the C-terminal 3-HA-tagging of *P. falciparum* genes of interest. 3'UTR PbDHFR - 3' untranslated region of the *P. berghei* dihydrofolate reductase, AmpR - ampicillin resistance gene. **(B)** Schematic of the pCC1 plasmid construct used to knock-out *P. falciparum* genes of interest via double-crossover homologous recombination. 5' CAM - calmodulin promoter, hDHFR - human dihydrofolate reductase gene, 3' UTR hrp2 - 3' untranslated region of *P. falciparum* hrp2 gene, 5' hsp86 - *P. falciparum* hsp86 promoter, ScFCU - chimaeric *S. cerevisiae* cytosine deaminase and uracil phosphoribosyl transferase, 3' UTR PbDHFR - 3' untranslated region of *P. berghei* dihydrofolate reductase, AmpR - ampicillin resistance gene. **(C)** Schematic of the pTT3-based HEK293 expression plasmid construct used for the expression of codon-optimised *P. falciparum* genes of interest in HEK293 cells. CMV promoter - cytomegalovirus promoter, TPL - tripartite leader sequence, MLP - adenovirus major late promoter enhancer, OriP - Epstein-Barr virus origin of replication, AmpR - ampicillin resistance gene, pUC origin - bacterial origin of replication.

2.1.4. Site-directed mutagenesis and point mutations of HEK293 expression plasmid constructs

PfDHHc5_CdA: In order to mutate the cysteine residue within the DHHc domain of PfDHHc5 into an alanine residue, the *GeneART*-synthesized expression plasmid coding for PfDHHc5 was subjected to site-directed mutagenesis using the primers listed in Table 2.4 and the QuikChange II XL Site Directed Mutagenesis Kit (Agilent Technologies), according to manufacturer's instructions. Primers were designed using the free online tool, QuikChange Primer Design Program (Agilent Technologies). DNA containing the desired mutation was introduced into One Shot® TOP10 chemically competent *E.coli* as described above. Single colonies were inoculated into 100 mL of LB broth containing 50 µg/mL ampicillin and grown overnight at 37°C with shaking at 250 rpm. DNA was extracted from the overnight cultures using a GenElute™ HP Plasmid Midiprep Kit (Sigma-Aldrich) according to manufacturer's instructions, and used for the transient transfection of HEK293 cells (described below).

PfSec22 and PfARO: In order to mutate the cysteine residues of interest in PfSec22 and PfARO, the primers listed in Table 2.4 (which contain the desired cysteine to alanine point mutations) were used to amplify the genes of interest by PCR, using the *GeneART*-synthesized expression plasmids as templates. PCR conditions and cycling parameters are as previously described above. The amplified DNA was restriction digested using the restriction enzymes listed in Table 2.4, and re-introduced into the *GeneART* expression plasmids using T₄ DNA ligase, following the methods previously described above. DNA containing the desired mutation was introduced into One Shot® TOP10 chemically competent *E.coli* as described above. Single colonies were inoculated into 100 mL of LB broth containing 50 µg/mL ampicillin and grown overnight at 37°C with shaking at 250 rpm. DNA was extracted from the overnight cultures using a GenElute™ HP Plasmid Midiprep Kit (Sigma-Aldrich) and used for the transient transfection of HEK293 cells (described below).

Site-directed mutagenesis construct	Primers (5'-3')		Restriction sites	Vector
PfDHH5-CdA	Sense	GAAATCTTCGACCACCAC <u>GCC</u> GACTTCACCCTGAACTG	-	pTT3-CD4 expression vector
	Antisense	CAGTTCAGGGTGAAGTC <u>GCC</u> GTGGTGGTCGAAGATTTTC		
PfSec22-C2dA	Forward	GC <u>gggcccgc</u> ATG <u>GCC</u> GATGTGGTGTGCTGTGCAG	NotI and Ascl	
	Reverse	GCG <u>ggcgcgcc</u> TCAGAGATCCTCTTCAGAAATCAGC		
PfARO-C5dA	Forward	GC <u>gggcccgc</u> ATGGGCAACAAT <u>GCC</u> TGCGCCGGC	NotI and Ascl	
	Reverse	GCG <u>ggcgcgcc</u> TCAGAGATCCTCTTCAGAAATCAGC		
PfARO-C6dA	Forward	GC <u>gggcccgc</u> ATGGGCAACAATTGC <u>GCC</u> GCCGGC	NotI and Ascl	
	Reverse	GCG <u>ggcgcgcc</u> TCAGAGATCCTCTTCAGAAATCAGC		

Table 2.4: Primers used for the generation of point mutations in HEK293 expression plasmid constructs. The codon targeted for the point mutation is underlined and highlighted in red. The restriction enzyme sites are highlighted in red and indicated along with the vector used in the columns on the right.

2.2. Plasmodium in vitro cell culture and transfection

2.2.1. Plasmodium falciparum culture

Plasmodium falciparum strain 3D7 was maintained in 'complete' media consisting of RPMI 1640 media (Life Technologies) supplemented with 30 mM HEPES (Sigma-Aldrich), 0.05 mg/mL hypoxanthine (Sigma-Aldrich), 0.025 mg/mL Gentamicin (Sigma-Aldrich), 2 mg/mL glucose (Sigma-Aldrich), 0.24% sodium bicarbonate (Sigma-Aldrich), and 10% O⁺ heat-inactivated human serum or 0.5% AlbuMAX[®] II (Life Technologies), along with O⁺ human erythrocytes at 5% haematocrit. All cultures were gassed with 5% CO₂ and 0.5% O₂ in N₂, according to established methods [6].

2.2.2. Synchronisation of P. falciparum parasites by sorbitol lysis

The pellet from *P. falciparum*-infected blood-stage culture was collected by centrifugation at 800xg for 5 minutes with low breaks and resuspended in 5 volumes of 5% sorbitol (Sigma-Aldrich) in water. The suspension was left standing at room temperature for 5 minutes and was subjected again to centrifugation in order to remove the sorbitol. The pellet was washed once in RPMI before being resuspended in complete media and put back into culture.

2.2.3. Transfection of *Plasmodium falciparum*

Transfection of ring-stage 3D7 parasites was performed according to published protocols [7]. Approximately 100-150 µg of DNA to be transfected was collected by ethanol precipitation and fully resuspended in 40 µL of sterile 1xTE buffer. Incomplete cytomix (120 mM KCl, 0.15 mM CaCl₂, 10 mM K₂HPO₄/KH₂PO₄ pH 7.6, 25 mM HEPES/2 mM EGTA pH 7.6, 5 mM MgCl₂) was added to the DNA, to a total volume of 400 µL, and the DNA/incomplete cytomix mixture placed on ice. The pellet from 5 mL of *P. falciparum* *in vitro* blood-stage culture (at 5% haematocrit with ring stage parasites at 8-12% parasitemia) was collected by centrifugation at 800xg for 5 minutes with low brakes and washed in 5 mL of incomplete cytomix. The pellet was then resuspended in the DNA/incomplete cytomix mixture and transferred into the bottom of a cold 2 mm electroporation cuvette (Bio-Rad). The entire mixture was electroporated using the Gene Pulser Xcell™ Electroporation System (Bio-Rad) with the following settings: 0.31 kV, 950 µF. The transfected cells were washed with 10 mL of RPMI media and then resuspended in 10 mL of pre-warmed complete media with 0.5 mL of freshly washed erythrocytes (50% haematocrit), and returned to culture. One day post-transfection, positive drug selection was started using either 2.5 µg/mL Blasticidin-S (Fisher Scientific) or 2.5 nM WR99210. Drug selection was maintained until stable parasite growth was obtained, which normally occurred 14 to 25 days post-selection. In order to select for parasites containing the construct integrated via homologous recombination, drug cycling was performed, where the transfected parasites were grown without drug pressure for 3 weeks, after which drug pressure was reapplied until stable parasite growth was once again attained. For knock-out strains transfected with the pCC1 plasmid, after drug cycling, negative selection was then performed by adding 1 µM 5-fluorocytosine (5FC) (Sigma-Aldrich), in the presence of 2.5 nM WR99210, in order to select for parasites which had integrated the construct via double cross-over recombination, and to select against parasites still harbouring the episomal plasmid. This negative selection was maintained until stable parasite growth was obtained [2].

2.2.4. Cloning by limiting dilution

P. falciparum blood-stage culture was diluted in complete media and added into the wells of a round-bottomed 96-well plate at concentrations of 0.5 parasites/well and 1 parasite/well. A concentration of 3 parasites/well was also included in two rows of the 96-well plate, as a positive control for the limiting dilution. Freshly washed erythrocytes were added into each well until a 2% haematocrit was achieved, and the volume in each well was then brought up to a total of 100 µL. Media was changed to selective media (complete media containing 2.5 µg/mL Blasticidin-S or complete media containing

2.5 nM WR99210) 10-12 days later. Smears were taken from each of the wells on day 14 up to day 21 until parasites were seen. Approximately 10-15 clones were selected, preferably from wells with higher dilutions, and gradually raised to 10 mL of complete media at 5% haematocrit.

2.2.5. Saponin lysis of *P. falciparum* infected erythrocytes

The pellet from *P. falciparum*-infected blood-stage culture was collected by centrifugation at 800xg for 5 minutes with low brakes. The pellets were resuspended in 10 volumes of 0.1% saponin in PBS and left at room temperature for 10 minutes to allow the lysis of erythrocytes to occur. Parasite material was collected by centrifugation at 3220xg for 10 minutes and the supernatant containing lysed erythrocyte material was discarded. The resultant parasite pellet was repeatedly washed with 0.1% saponin/PBS until the supernatant was clear. The supernatant from the final wash was removed and the saponin pellet frozen at -80°C until use. For SDS-PAGE and Western blot analysis (described in Section 2.5.1), the saponin pellet was lysed in 4% SDS/50 mM Tris-Cl pH 7.4/150 mM NaCl/5 mM EDTA in the appropriate volume which would result in a concentration of 5×10^8 parasites/mL. Lysis was allowed to occur for 30 minutes at 37°C with shaking. Any insoluble material was removed by centrifugation at 20238xg for 5 minutes and 10-20 μ L of the resultant supernatant was diluted in 2X Laemmli sample buffer (4% SDS, 120 mM Tris-Cl pH 6.8, 20% glycerol, 0.02% bromophenol blue) with 10% β -mercaptoethanol (Sigma-Aldrich) and analysed as described in Section 2.5.1.

2.2.6. *P. falciparum* growth assay

In vitro blood-stage culture infected with *P. falciparum* transgenic strains of interest, as well as with wild-type 3D7 strain as a control, was diluted appropriately to produce a suspension of 2% haematocrit and 1% parasitemia. A volume of 100 μ L of each suspension was added into each well of a round-bottomed 96-well plate (labelled 'Day 0') and the plate incubated for 2 days under the standard *P. falciparum* culture conditions described in Section 2.2.1. Additionally, 10 μ L of each parasite suspension was also fixed in 50 μ L of fixative (2% formaldehyde/0.2% glutaraldehyde/PBS) for 45 minutes at 4°C. The fixed cells were collected by centrifugation at 450xg for 2 minutes and washed with PBS. These fixed cells were named 'Day 0' and kept in 50 μ L PBS until further use. After the 2 day incubation period, 10 μ L aliquots from each well of the 'Day 0' 96-well plate were taken into a new 96-well plate (labelled 'Day 2') and fixed in 50 μ L per well of fixative as described above. These fixed cells, along with the previously fixed 'Day 0' cells, were then permeabilised in 50 μ L per well of 0.3% Triton X-100/PBS for 10 minutes at 24°C, followed by a wash with PBS. The cells were

then incubated in 50 μ L/well of ribonuclease A (MP Biomedicals), at a concentration of 0.5 mg/mL, for 45 minutes at 37°C, followed by another wash with PBS. Finally, the cells were stained with 50 μ L/well of SYBR® Green I (Invitrogen)/PBS, at a concentration of 1:5000, for 45 minutes at 37°C. The stained cells were washed twice with PBS and were then resuspended in 200 μ L/well of PBS. Parasitemia of the samples from 'Day 0' and 'Day 2' was then counted by flow cytometry as previously described [8], with kind assistance from Michel Theron, Rayner lab. The cultures in the original 'Day 0' plate were then diluted 1:5 or 1:10 as appropriate and incubated for a further 2 days under the standard *P. falciparum* culture conditions. Samples were again taken two days later (cells collected were then labelled 'Day 4') and the cells fixed and stained as described above. The cultures in the original 'Day 0' plate were again diluted 1:5 or 1:10 accordingly. This was repeated every two days until 'Day 10' samples were collected, fixed and stained. Each sample for every transgenic strain tested was performed in triplicate.

2.2.7. Purification of *P. berghei* schizonts

P. berghei-infected mouse blood was kindly provided by Ellen Bushell, Billker lab, and added into 50 mL of schizont culture media, which consisted of RPMI 1640 (supplemented with 25 mM HEPES and 2 mM L-Glutamine) (Sigma-Aldrich), 24 mM sodium bicarbonate pH 7.2 (Sigma-Aldrich), 25% heat-inactivated foetal bovine serum (FBS) and 1:100 Penicillin/Streptomycin (Life Technologies). The schizont culture was gassed (5% CO₂ and 0.5% O₂ in N₂) and incubated overnight at 37°C with shaking. The red blood cell pellet was then collected by centrifugation at 180xg for 15 minutes with low acceleration/brakes. The supernatant was removed and the pellet was layered on top of a Nycodenz gradient (2.75 mL of Nycodenz stock in 2.25 mL of PBS) and subjected to centrifugation at 300xg for 30 minutes with low acceleration/brakes. The schizont-containing interlayer was then removed and washed with schizont culture media. The purified schizont pellet was collected from the wash by centrifugation at 300xg for 8 minutes with low acceleration/brakes. For immunofluorescence assays, 1-2 μ L of the purified schizonts were resuspended in 500 μ L FBS, and 100 μ L of the schizont/FBS suspension was dropped onto poly-L-lysine coverslips (BD Biosciences) and was analysed as described in Section 2.5.2. For SDS-PAGE and Western blot analysis, the purified schizont pellet was resuspended in 30-50 μ L of 2X Laemmli sample buffer (4% SDS, 120 mM Tris-Cl pH 6.8, 20% glycerol, 0.02% bromophenol blue) with 10% β -mercaptoethanol (Sigma-Aldrich) and boiled for 5 minutes at 95°C, followed by analysis as described in Section 2.5.1.

2.3. Genotyping of transgenic parasite strains

2.3.1. Genotyping by PCR

Genomic DNA was extracted from *in vitro* blood-stage culture infected with *P. falciparum* transgenic triple-HA-tagged or knock-out strains of interest, or the wild-type 3D7 strain as a control, using the QIAamp DNA Blood Mini Kit (QIAGEN) according to manufacturer's instructions. The genomic DNA extracted for each transgenic strain was genotyped using the specific primers listed in Table 2.5 and Table 2.6. The PCR cycling parameters were as described in Section 2.1.1.

Triple-HA-tagged line	Primer	Sequence (5'-3')
PfDHH3	P1	GTAGGATTGTTATCTACCTTAACCC
	P2	CTCGATATTATTAACAAGCCGATC
	P3	GAACATATTTATTAAGTGCAG
	P4	GCTATTTACATGCATGTGCATGCAC
	P5	CACATTTGTGCATTTAGAGATGATAC
PfDHH5	P1	GTTGCCCATATTTCTTAGCATTTATAC
	P2	CAATGTTTCATTTCCGTTTAAATCGC
	P3	GAACATATTTATTAAGTGCAG
	P4	GCTATTTACATGCATGTGCATGCAC
	P5	GAAGCAAATGGTACTATAGGCC
PfDHH7	P1	CCAGAGGAATTATTGAGATGGGGG
	P2	CATATAAATGCATATTATTCAAGCAG
	P3	GAACATATTTATTAAGTGCAG
	P4	GCTATTTACATGCATGTGCATGCAC
	P5	GGATACGTCAAATAATGTGATGTTGG
PfDHH8	P1	GGAATTATGTTGATTTATCATTGCC
	P2	CGAGGTTTATGATCCCCTGCGG
	P3	GAACATATTTATTAAGTGCAG
	P4	GCTATTTACATGCATGTGCATGCAC
	P5	-
PfDHH9	P1	CCATGGATAGGCACCTTGTGATAGG
	P2	CTCCATCTTCCTTTATGTTTTTC
	P3	GAACATATTTATTAAGTGCAG
	P4	GCTATTTACATGCATGTGCATGCAC
	P5	CGATTCAGACTTGTAATAAGTGC
PfMBOAT	P1	GATACTGAACCCACACTTTCATC
	P2	CCATGGTGGTAAAATATATGCACGTG
	P3	GAACATATTTATTAAGTGCAG
	P4	GCTATTTACATGCATGTGCATGCAC
	P5	GAATAATAAAAATGATGAGGGAC

Table 2.5: Primers used for the genotyping of *P. falciparum* triple-HA-tagged lines.

Knock-out line	Primer	Sequence (5'-3')
PfDHC3	P6	GAACGTATAATCTTTATAAAGTTTTGCC
	P7	CCTCTTTCAATTTTCGTGTTTTTCC
	P8	GGAATACTAAATATATATCCAATGGCCCC
	P9	GCGACGATGCAGTTTAGCGAACC
	P10	GGGTTAAGGTAGATAACAATCCTAC
	P11	-
	P12	GTAGGATTGTTATCTACCTTAACCC
PfDHC5	P6	GTTTATCCTTTAACCTTTTTATAGTATG
	P7	CTTATAATTACACTTAATAAATTACCAC
	P8	GGAATACTAAATATATATCCAATGGCCCC
	P9	GCGACGATGCAGTTTAGCGAACC
	P10	GTCATCCGTTCCATGTCGTAAC
	P11	GTTTTACAGTATTGAAGTTAATTTGC
	P12	GTTTACGACATAGGAACGGATGAC
PfDHC7	P6	GATGTATAATAAAGACGAATGAAGTGTC
	P7	GTTTTATTGGAATAATTTCTCTACTG
	P8	GGAATACTAAATATATATCCAATGGCCCC
	P9	GCGACGATGCAGTTTAGCGAACC
	P10	CCATTTACAATGTCTTCGTTCTCC
	P11	-
	P12	GGAGAACGAAGACATTGTAATGG
PfDHC8	P6	GTTCTCAACAGCAGTAATACATAC
	P7	GCGTTTTGAAAATGTTAATATATTTTTG
	P8	GGAATACTAAATATATATCCAATGGCCCC
	P9	GCGACGATGCAGTTTAGCGAACC
	P10	GTATAACAATGTTTACTTCTAGGTTT
	P11	-
	P12	GAACCTAGAAGTAAACATTGTTATAC
PfDHC9	P6	CAATTATGCAATATGTTGTATAAATG
	P7	CTCCATCTTCTTTATGTTTTT
	P8	GGAATACTAAATATATATCCAATGGCCCC
	P9	GCGACGATGCAGTTTAGCGAACC
	P10	GGCACAAGAGCATTATATAAAATCGG
	P11	GAAAATAATGGAATGTTAAATTAGGAAATAC
	P12	CCGATTTTTATAAATGCTCTTTGTGCC
PfMBOAT	P6	GTGTGTTTCGTAATATTCACTTTTAG
	P7	CACTTTTGAGATATGTGTCCATGGTGG
	P8	GGAATACTAAATATATATCCAATGGCCCC
	P9	GCGACGATGCAGTTTAGCGAACC
	P10	GTGTATGAATATTATGAGCACACG
	P11	-
	P12	CGTGTGCTCATAATATTCATACAC
P13	CCATATGCGGTGTGAAATACCGC	

Table 2.6: Primers used for the genotyping of *P. falciparum* knock-out lines.

2.3.2. Reverse transcription (RT)- PCR

Total RNA was extracted from *in vitro* blood-stage cultures infected with *P. falciparum* transgenic strains of interest, or the wild-type 3D7 strain as a control, using a QIAamp RNA Blood Mini Kit (QIAGEN) according to manufacturer's instructions. The extracted RNA was DNase treated using the DNA-free™ Kit (Ambion by Life Technologies) according to manufacturer's instructions. Approximately 2 µg of the total RNA was then used to generate cDNA using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to manufacturer's instructions. The cDNA was tested by PCR amplification of regions specific to the gene of interest using the primers listed in Table 2.7.

Knock-out construct	Primers for testing cDNA (5'-3')	
	Forward	Reverse
PfDHC5-KO	CCATTCTTCAATTTGTTGAAG	GGTGTATGGGCCTATAGTACC
PfDHC9-KO	GAATAATTATTTGGCATTATC	CTCCATCTTCCTTTATGTTTTTC

Table 2.7: Primers used for the PCR analysis of the cDNA of PfDHC5-KO and PfDHC9-KO transgenic clones.

Quantitative PCR (qPCR) was also performed on the generated cDNA using gene-specific primers, along with gene-specific probes labelled with 6-carboxyfluorescein (6-FAM) on the 5'-end, and a non-fluorescent quencher (MGBNFQ, Applied Biosystems) on the 3'-end. A concentration of 300 nmol/L was used for both primers and probes in a 25 µL reaction (the primers and probes used are listed in Table 2.8). The cycling parameters for the qPCR reaction are as shown below:

Step	Temperature (°C)	Time (min)
1	50	2
2	95	10
3	95	0.25
4	60	1
5	Go to 3, repeat 39 times	

Each qPCR reaction also included a wild-type 3D7 cDNA control, as well as a standard internal control using the housekeeping gene, seryl-tRNA-synthase (primers and probes used for the seryl-tRNA-synthase housekeeping gene are also shown in Table 2.8). All samples were assayed in triplicate. The qPCR assay described here was kindly performed by Leyla Bustamante, Rayner lab.

Gene	qPCR primers		qPCR 6-FAM probe
	Forward	Reverse	
PfDHC5-KO	CCCCCGAATTTACCAGTT	TGCCTTCTTTTGACATATCCTA	CTAAATAGTAAGCAGGTAT
PfDHC9-KO	TCCTACAGTGACATGAATCCTT	AAAACCGCTTTCCAATTATT	TGACTTGGGAATATA
Seryl-tRNA-synthase	CCACACAAGGAGAAGATA	GGGAAAGACAATAGACA	TTAAAGTTTGTCTCGCTTGAGCCC

Table 2.8: Primers and probes used for the qPCR analysis of the cDNA of PfDHC5-KO and PfDHC9-KO transgenic clones.

2.4. Human Embryonic kidney 293 (HEK293) *in vitro* cell culture and transfection

2.4.1. Human embryonic kidney 293E (HEK293E) cell culture

Human embryonic kidney 293E (HEK293E) cells were maintained in ‘complete’ culture media, consisting of Gibco® FreeStyle™ 293 expression media (Life Technologies) supplemented with 1% heat-inactivated FBS under standard humidified conditions (37°C and 5% CO₂), essentially as previously described [9], except that the cells were grown in T75 tissue culture flasks (at 1x10⁶ cells/flask) with 25 mL of complete media, instead of in suspension.

2.4.2. Transfection of HEK293 cells

Transient transfection of HEK293 cells with *GeneArt* constructs coding for codon-optimized *P. falciparum* proteins were performed using polyethylenimine (PEI) as previously described [5, 9]. Briefly, cells were split the day before transfection to give a concentration of 1x10⁶ cells per transfection. On the day of transfection, 12.5 µg total DNA was added into 1 mL of unsupplemented FreeStyle media, and mixed with 25 µL of PEI. The transfection mixture was left for 15 minutes at 24°C before being added to the cells, and left overnight under the standard culture conditions described in Section 2.4.1. The transfected cells were collected by centrifugation at 3220xg for 15 minutes and culture supernatants were discarded. For immunofluorescence assays, the cells were split onto poly-l-lysine coverslips (BD Biosciences) in 12-well plates at a concentration of 30000 cells/well in 1 mL of complete culture media the day before transfection. The transfection mix was scaled down appropriately according to the volume of media, and the cells were fixed after overnight transfection as described in Section 2.5.2.

2.5. Protein detection and analysis

2.5.1. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting

Proteins were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using NuPAGE® Novex® 4-12% Bis-Tris pre-cast gels (Life Technologies), at a constant voltage of 200V for 50 minutes, according to manufacturer's instructions. After electrophoresis was complete, the proteins were then transferred to Immobilon®-P PVDF membranes (Sigma-Aldrich) using the XCell II™ blot module (Life Technologies) and 1X NuPAGE® transfer buffer (Life Technologies) (with 10% methanol), at a constant voltage of 30V for one hour, according to manufacturer's instructions. After the transfer was complete, the PVDF membranes were blocked overnight in 5% milk (Marvel)/PBS. The membranes were probed with primary antibodies, diluted in PBS-0.1% Tween-20 (Sigma-Aldrich) (PBST) or 2% milk/PBS, for 2 hours at 24°C with rotation, followed by 3 washes with PBST. The membranes were then probed with horse radish peroxidase (HRP)-conjugated secondary antibodies (Amersham ECL, GE Healthcare), diluted in PBST or 2% milk/PBS, for 1 hour at room temperature with rotation, followed by another 3 washes with PBST. The presence of the HRP-conjugated secondary antibodies on the membranes was detected using enhanced chemiluminescence (Amersham ECL, GE Healthcare). All primary and secondary antibodies used are listed in Table 2.9 with the appropriate dilutions.

Antibodies		Source	Dilution	
			IFA	Western blot
Primary	Anti-HA-tag mouse monoclonal (6E2)	Cell Signaling Technology	1:200	1:400
	Anti-HA-tag rabbit monoclonal (C29F4)		1:200	1:400
	Anti-HA-tag rabbit polyclonal	Life Technologies	-	1:400
	Anti- <i>P. falciparum</i> ERD2 rabbit polyclonal (MRA-1)	MR4, ATCC	1:2000	-
	Anti- <i>P. falciparum</i> GRP (BiP) rat polyclonal	MR4, ATCC	1:2000	-
	Anti- <i>P. falciparum</i> GAP45 rabbit	Matt Jones	1:1000	-
	Anti- <i>P. falciparum</i> MSP1 mouse monoclonal (MRA-94)	MR4, ATCC	1:1000	-
	Anti- <i>P. falciparum</i> RAP1 mouse monoclonal (MRA-79)	MR4, ATCC	1:1000	-
	Anti- <i>myc</i> mouse monoclonal (9E10)	Life Technologies	1:1000	1:500-1:1000
	Anti-c-myc rat monoclonal (JAC6)	Abcam	1:1000	1:500-1:1000
	Anti-FLAG® mouse monoclonal (M2)	Sigma Aldrich	1:1000	1:1000
	Anti-FLAG® rabbit polyclonal	Sigma Aldrich	1:1000	1:1000
	Anti-calnexin rabbit polyclonal	Abcam	1:1000	-
	Anti-pan-cadherin rabbit polyclonal	Abcam	1:1000	-
Secondary	Alexa Fluor® 488 Goat Anti-Mouse/Rabbit/Rat IgG (H+L)	Life Technologies	1:1000	-
	Alexa Fluor® 555 Goat Anti-Mouse/Rabbit/Rat IgG (H+L)	Life Technologies	1:500-1:1000	-
	Amersham ECL™ HRP-linked Anti-Mouse/Rabbit/Rat IgG	GE Healthcare	-	1:4000

Table 2.9: All primary and secondary antibodies used in this work along with their appropriate working dilutions.

2.5.2. Immunofluorescence Assay

In suspension: The pellet from 600 μL of *in vitro* blood-stage culture infected with *P. falciparum* transgenic strains of interest was collected by centrifugation at 450xg for 1 minute, and fixed in 4% formaldehyde/0.01% glutaraldehyde/PBS at for one hour at 24°C with rotation. The fixed pellet was then permeabilised with 0.1% Triton X-100 for 10 minutes at 24°C with rotation, followed by blocking in 3% bovine serum albumin (BSA) (Sigma-Aldrich)/PBS for one hour at 24°C with rotation. Immunodetection was performed by incubating the fixed cell pellets with primary antibodies, diluted in 1% BSA/PBS (all primary antibodies used and the corresponding dilutions are shown in Table 2.9), for 2 hours at 24°C with rotation. After 3 washes in 1% BSA/PBS, this was followed by incubation with secondary antibodies, diluted in 1% BSA/PBS (all secondary antibodies used and the corresponding dilutions are shown in Table 2.9), for 1 hour at 24°C with rotation. During the secondary antibody incubation, the nuclear DNA was also stained with DAPI at a concentration of 1:4000. After a final 3 washes, the pellet was resuspended in 250 μL FBS and 7 μL of the resuspended pellet was spread out on a glass slide and allowed to dry. This was then mounted in Prolong anti-fade mounting reagent (Life Technologies).

On poly-l-lysine coverslips: Poly-l-lysine coverslips (BD Biosciences) were placed into the wells of 12-well plates and cells were either grown on the poly-l-lysine coverslips (as described in Section 2.4.2), or dropped onto the coverslips (as in Section 2.2.7) and left for 15 minutes to attach. The cells were then fixed in 500 μL of 4% formaldehyde/0.01% glutaraldehyde/PBS for 15 minutes at 24°C. The permeabilisation, blocking, and primary and secondary antibody incubation steps were performed as described above, except that all incubations were done in the wells of the 12-well plate without rotation. After the final washes, the coverslips were removed from the 12-well plate and mounted onto glass slides using Prolong anti-fade mounting reagent (Life Technologies).

For both methods of immunofluorescence assay described above (in suspension and on coverslips), confocal images were acquired using a Zeiss LSM 510 Laser Scanning confocal microscope.

2.5.3. Immunoprecipitation of c-Myc-tagged proteins

HEK293 cells were either co-transfected with both the FLAG-tagged PfDHHC protein of interest and a c-Myc-tagged potential substrate, or co-transfected with the c-Myc-tagged potential substrate and the empty CD4 vector control, as described in Section 2.4.2. After 24 hours of transfection at 37°C under standard culture conditions, transfected cells were collected by centrifugation at 3220xg for 10 minutes. The cell pellets were lysed in 500 μL of IP buffer (1% Triton X-100, 50 mM Tris-Cl pH 7.4, 150

mM NaCl, 5 mM EDTA, protease inhibitor cocktail (Roche)) and incubated for 30 minutes at 37°C with shaking. Cell debris was removed by centrifugation at 20238xg for 5 minutes, and the resulting supernatants were pre-cleared by incubating with 15 µL of Protein G Sepharose® (Sigma-Aldrich) per sample for 1 hour at 4°C with rotation. The Protein G Sepharose resin was collected by centrifugation at 20238xg for 1 minute at 4°C. The resulting supernatants were then incubated with 2 µg of mouse α-c-Myc antibody (Life Technologies) per sample and incubated overnight at 4°C with rotation. This was followed by incubation with 30 µL of Protein G Sepharose per sample for 2 hours at 4°C with rotation. The Protein G Sepharose resin was collected by centrifugation at 20238xg for 5 minutes at 4°C and subjected to 3 washes with IP buffer. The immunoprecipitated proteins were then eluted from the Protein G Sepharose beads by incubating with 60 µL of 2% SDS/50 mM Tris-Cl pH 7.4/5 mM EDTA for 5 minutes at 95°C with shaking. For SDS-PAGE and Western blot analysis, 10 µL of each eluate was diluted in 2X Laemmli sample buffer (4% SDS, 120 mM Tris-Cl pH 6.8, 20% glycerol, 0.02% bromophenol blue) with 10% β-mercaptoethanol (Sigma-Aldrich) and analysed as described in Section 2.5.1.

2.6. Palmitome purification in *P. falciparum* parasites and HEK293 cells

2.6.1. Purification of palmitoylated peptides in *P. falciparum* schizonts using Acyl-biotinyl exchange (ABE) – Trial 1

Parasites were extracted from *in vitro* blood-stage cultures infected with *P. falciparum* strain 3D7 by saponin-lysis (described in 2.2.5). The saponin-extracted parasite pellet was first washed in ice cold ABE lysis buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 5 mM EDTA) and resuspended at a concentration of 1×10^9 parasites/mL in ice-cold ABE lysis buffer containing 1.7% Triton X-100, 10 mM N-ethylmaleimide (NEM) (Sigma-Aldrich), Protease inhibitor cocktail (Sigma-Aldrich) and Phosphatase inhibitor cocktail 2 (Calbiochem). The lysates were incubated for one hour at 4°C with rotation, followed by precipitation of proteins from the lysates by chloroform-methanol (C/M) precipitation. Precipitated proteins were resuspended in 4% SDS/50 mM Tris-Cl pH 7.4/5 mM EDTA/10 mM NEM and incubated at 37°C with shaking until fully solubilised. The solubilised proteins were diluted in ABE lysis buffer containing 1 mM NEM and 0.2% Triton X-100, and incubated overnight at 4°C with rotation. The samples were then subjected to 3 sequential C/M precipitations to ensure all NEM was removed. After the final C/M precipitation, the samples were split into two equal portions. One portion was treated with hydroxylamine by resuspension in (+)hydroxylamine buffer (0.7 M hydroxylamine (Sigma-Aldrich) pH 7.4, 1 mM EZ-link HPDP-biotin (Thermo Scientific),

0.2% Triton X-100), and the second portion was mock treated by resuspension in (-)hydroxylamine buffer (50 mM Tris-Cl pH 7.4, 1 mM EZ-link HPDP-biotin, 0.2% Triton X-100). The (+) and (-) hydroxylamine treatments were incubated for one hour at 24°C with rotation, followed by C/M precipitation. The precipitated samples were then further treated with HPDP-biotin by resuspension in HPDP-biotin buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 5mM EDTA, 0.2 mM EZ-link HPDP-biotin, 0.2% Triton X-100) for 2 hours at 24°C with rotation. The HPDP-biotin buffer was removed by 3 sequential C/M precipitations and the precipitated proteins were resuspended in 2% SDS/50 mM Tris-Cl pH 7.4/5 mM EDTA.

The samples were then diluted in ABE lysis buffer containing 0.2% Triton X-100, at a volume that resulted in the dilution of the SDS concentration in the samples to 0.1%. Streptavidin-agarose resin (Thermo Scientific), at a bed volume of 200 µL, was added into each sample and the samples were incubated for 2 hours at 24°C with rotation. The streptavidin-agarose resin was precipitated by centrifugation at 3200xg for 2 minutes and washed once with ABE lysis buffer for 10 minutes. This was followed by a further 3 washes with 2 M urea (Sigma-Aldrich)/100 mM ammonium bicarbonate (Sigma-Aldrich), at 10 minutes per wash. The resin was then treated with 20 µg of Trypsin Gold (Promega) in 1 mL total volume of 1 M urea/50 mM ammonium bicarbonate. Trypsin digestion was allowed to proceed for 2 hours at 37°C with shaking. The resin was then loaded into microcentrifuge spin columns (Thermo Scientific) and washed 3 times with 2 M urea/100 mM ammonium bicarbonate, using syringes. Peptides were then eluted from the resin by incubating with 50 µL of 10 mM Tris(2-carboxyethyl)phosphine(TCEP) for 10 minutes at 37°C with shaking. This was repeated with a further 50 µL of TCEP and both elutions were pooled. The purification methods described above were performed in collaboration with Mark Collins, Choudhary lab.

For quantitative analysis of the eluted peptides, samples were desalted and subjected to stable isotope dimethyl labelling as previously described [10]. Briefly, (-)hydroxylamine samples were labelled with 'medium' deuterated formaldehyde (D2) and sodium cyanoborohydride, and the (+)hydroxylamine samples were labelled with 'heavy' ¹³C-deuterated formaldehyde (13C-D2) and sodium cyanoborodeuteride. Labelled peptides were pooled and acidified using 10% formic acid for analysis by LC-MS/MS with an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific). The stable isotope dimethyl labelling and LC-MS/MS analysis was kindly performed by Mark Collins, Choudhary lab.

2.6.2. Purification of palmitoylated peptides in *P. falciparum* schizonts using Acyl-biotinyl exchange (ABE) – Trial 2, 3A, 3B and 4

Parasites were extracted from *in vitro* blood-stage cultures infected with *P. falciparum* strain 3D7 by saponin-lysis (described in 2.2.5). The saponin-extracted parasite pellet was lysed in 1 mL of extraction buffer (4% SDS, 0.1 M Tris-Cl pH 8.0, 2 µg/µL Aprotinin/Leupeptin, 0.5 mM PMSF, 20 µM ZnCl₂, 5 mM EDTA and 25 mM TCEP) and homogenised 25 times. The homogenised lysate was then heated for 10 minutes at 70°C and passed through a fine gauge needle 10 times to sheer the DNA. The lysates were subjected to centrifugation at 20238xg for 5 minutes and the supernatants transferred to a new tube. The remaining pellet was then further lysed in 400 µL 8 M urea/100 mM Tris-Cl pH 8.0, and vortexed for 2 minutes. This urea-treated lysate was subjected to centrifugation at 20238xg for 5 minutes. The resultant supernatant was pooled with the first supernatant, and a 10 µL aliquot of the pooled lysates was kept aside for measurement of protein concentration. Protein concentration was measured using the QuantiPro BCA Assay Kit (Sigma-Aldrich) according to manufacturer's instructions. Iodoacetamide (IAA) (Sigma-Aldrich) was added to the lysate at a final concentration of 50 mM. Urea was also added to a final concentration of 8 M. The samples were incubated for 2 hours at 24°C in the dark with rotation. Cysteine, at a final concentration of 50 mM, was then added to the sample and incubated for 30 minutes at 24°C with rotation, to quench the remaining IAA.

The lysate was then split into two equal portions and transferred into two pre-washed (wash buffer: 8 M urea/100 mM Tris-Cl pH 8.0) Amicon Ultra-15 Centrifugal Filter Units (30 kDa molecular weight cut-off) (Millipore). The samples were washed 4 times with 2 mL 8 M urea/100 mM Tris-Cl pH 8.0 buffer. To one sample – the (+)hydroxylamine sample- the following reagents were added to the upper chamber of the filter units: 210 µL 100 mM Tris-Cl pH 7.4, 20 µL 50 mM EZ-link HPDP-biotin and 350 µL 2 M hydroxylamine pH 7.4. To the second sample – the (-)hydroxylamine sample- the following reagents were added to the upper chamber of the filter units: 560 µL 100 mM Tris-Cl pH 7.4 and 20 µL 50 mM EZ-link HPDP-biotin. Samples were fully resuspended and incubated for one hour at 24°C with shaking. The ABE reagents were removed by centrifugation at 3220xg for 20 minutes, followed by 4 washes with 2 mL 8 M urea/100 mM ammonium bicarbonate buffer and a final wash with 1 mL 100 mM ammonium bicarbonate buffer. The samples were made up to 2 mL with 100 mM ammonium bicarbonate buffer and the appropriate amount of Trypsin Gold (Promega) was added to each sample at an enzyme:substrate ratio of 1:50. The samples were digested for 4 hours at 37°C with shaking. The digested peptides were collected by centrifugation at 3220xg for 15 minutes. The upper chambers of the spin columns were then washed with 2 mL wash buffer (100

mM Tris-Cl pH 7.4, 300 mM NaCl, 10mM EDTA pH 7.4, 0.2% SDS, 0.4% Triton X-100), and the wash buffer pooled with the previously collected peptides.

Streptavidin-agarose resin (at a bed volume of 200 μ L per sample) was washed with LB buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 5 mM EDTA pH 7.4, 0.1% SDS, 0.2% Triton X-100), and added to the peptide samples. The samples were incubated for one hour at 24°C with rotation. The resins were then washed twice with 10 mL of LB buffer (at 10 minutes per wash), and washed a further two times with 10 mL of 2 M urea/100 mM ammonium bicarbonate buffer (at 5 minutes per wash). The resins were loaded into microcentrifuge spin columns and washed a further two times with 10 mL of 2 M urea/100 mM Tris-Cl pH 7.4 buffer, followed by one wash with 2 mL water, using syringes. Peptides were then eluted from the resin by incubating with 50 μ L 10 mM TCEP for 10 minutes at 37°C with shaking. This was repeated with another 50 μ L of 10 mM TCEP, and both elutions were pooled. The eluted peptides were either subjected to stable isotope dimethyl labelling followed by acidification with 10% formic acid as described above (Section 2.6.1), or acidified with 10% formic acid without labelling (for label-free analysis). The labelled and label-free peptides were then analysed by LC-MS/MS with an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific). The purification methods described above were performed in collaboration with Mark Collins, Choudhary lab. Stable isotope dimethyl labelling and LC-MS/MS analysis was kindly performed by Mark Collins, Choudhary lab.

2.6.3. Purification of palmitoylated proteins in HEK293 cells by metabolic labelling and click chemistry

HEK293 cells –either co-transfected with both the FLAG-tagged PfDHHC protein of interest and a c-Myc-tagged potential substrate, or co-transfected with the c-Myc-tagged potential substrate and the empty CD4 vector control- were treated with 25 μ M of the palmitic acid analogue, 17-octadecynoic acid (17-ODYA) (Cayman Chemical), or mock-treated with an equal volume of DMSO (Sigma-Aldrich), 24 hours after transfection. The 17-ODYA treatment was allowed to continue for a further 6 hours at 37°C under the standard culture conditions described in Section 2.4.1. The transfected and 17ODYA-labelled cells were collected by centrifugation at 3220xg for 15 minutes and the supernatants were removed. The resulting cell pellets were each lysed in 500 μ L of buffer containing 1% Triton X-100/50 mM Tris-Cl pH 7.4/ 150 mM NaCl/EDTA-free protease inhibitor cocktail (Roche), and incubated for 30 minutes at 37°C with shaking. The lysates were subjected to centrifugation at 20238xg for 5 minutes and the supernatants were C/M precipitated and then resuspended in 2% SDS/PBS. Protein concentrations were measured for all conditions using the QuantiPro BCA assay kit (Sigma-Aldrich) according to manufacturer’s instructions. The click chemistry reaction was set up with approximately 2 mg of protein for each of the conditions using the following chemicals: 100 μ M biotin-azide

(Invitrogen), 1 mM TCEP (Sigma-Aldrich), 100 μ M Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, Sigma-Aldrich) dissolved in DMSO/*tert*-butanol (20%/80%), and 1 mM CuSO₄ (Sigma-Aldrich). The reaction was made up to a total volume of 500 μ L with PBS and allowed to proceed for 1.5 hours at 23°C. C/M precipitation was then performed to completely remove the reactants of the click chemistry reaction. The precipitated proteins were resuspended in 2% SDS/PBS and diluted 10X with 0.2% Triton X-100/PBS. Pre-washed streptavidin-agarose resin (wash buffer: 0.2% Triton X-100/PBS) was added to the samples, at 50 μ L bed volume of streptavidin-agarose per sample, and incubated for 2 hours at 24°C with rotation. The streptavidin-agarose resins were washed 4 times with 0.2% Triton X-100/PBS. Proteins were eluted from the resins by adding 100 μ L of 2% SDS/50 mM Tris-Cl pH 7.4, 5 mM EDTA pH 7.4 and incubating for 5 minutes at 95°C with shaking. For SDS-PAGE and Western blot analysis, 12 μ L of each eluate was diluted in 2X Laemmli sample buffer (4% SDS, 120 mM Tris-Cl pH 6.8, 20% glycerol, 0.02% bromophenol blue) with 10% β -mercaptoethanol (Sigma-Aldrich) and analysed as described in Section 2.5.1.

References:

1. Abdi, A., et al., *SAM domain-dependent activity of PfTKL3, an essential tyrosine kinase-like kinase of the human malaria parasite Plasmodium falciparum*. Cellular and Molecular Life Sciences, 2010. **67**(19): p. 3355-3369.
2. Maier, A.G., et al., *Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in Plasmodium falciparum for targeted gene deletion by double crossover recombination*. Mol Biochem Parasitol, 2006. **150**(1): p. 118-121.
3. Pfander, C., et al., *A scalable pipeline for highly effective genetic modification of a malaria parasite*. Nat Methods, 2011. **8**(12): p. 1078-1082.
4. Janse, C.J., J. Ramesar, and A.P. Waters, *High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite Plasmodium berghei*. Nat Protoc, 2006. **1**(1): p. 346-356.
5. Durocher, Y., S. Perret, and A. Kamen, *High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells*. Nucleic Acids Res, 2002. **30**(2): p. E9.
6. Trager, W. and J.B. Jensen, *Human malaria parasites in continuous culture*. Science, 1976. **193**(4254): p. 673-5.
7. Crabb, B.S., et al., *Transfection of the human malaria parasite Plasmodium falciparum*. Methods Mol Biol, 2004. **270**: p. 263-76.
8. Theron, M., et al., *An adaptable two-color flow cytometric assay to quantitate the invasion of erythrocytes by Plasmodium falciparum parasites*. Cytometry Part A, 2010. **77A**(11): p. 1067-1074.
9. Bushell, K.M., et al., *Large-scale screening for novel low-affinity extracellular protein interactions*. Genome Research, 2008. **18**(4): p. 622-630.
10. Boersema, P.J., et al., *Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics*. Nat Protoc, 2009. **4**(4): p. 484-494.

Chapter 3

**The sites of palmitoylation
in the *Plasmodium
falciparum* palmitome**

The *Plasmodium falciparum* palmitome

Recent advances in proteome-based technologies have allowed the characterisation of palmitoylated proteins to shift from studies of individual palmitoyl-proteins, to the global characterisation of whole palmitomes in particular organisms [1-3]. The two main techniques which have been used to successfully purify whole palmitomes are acyl-biotinyl exchange (ABE) [2] and metabolic labelling with click chemistry [3](described in detail in Section 1.6). Both these techniques were recently used to purify the palmitome of the asexual intraerythrocytic schizont stages of *Plasmodium falciparum*, and coupled with quantitative mass spectrometry, these purification methods revealed that there are at least 494 putative palmitoyl-proteins in blood-stage *P. falciparum* parasites [4].

However, these palmitome purification techniques do not provide site-identifying information, and as noted in Section 1.5, there is no consensus sequence for palmitoylation. Therefore, although our understanding of palmitoylation in *P. falciparum* has increased with the purification of the *P. falciparum* palmitome, it is still unknown which cysteine residue(s) within these 494 palmitoyl-proteins are actually palmitoylated. Identification of the actual palmitoylated cysteine residues is important, as a global study of palmitoylation sites could potentially assist in our understanding of how individual PATs recognise their substrates. Furthermore, information on the sites of palmitoylation could be used to experimentally investigate the function of palmitoylation for a particular protein. Thus, the first aim of this project was to adapt current methods of palmitome purification to incorporate the identification of specific sites of palmitoylation on individual palmitoyl-proteins, in order to allow a global study of palmitoylation sites within the *P. falciparum* schizont palmitome.

3.1. Development of the site-identification (ID) palmitome purification method

3.1.1. Method used for the purification of the *P. falciparum* total schizont palmitome

In an attempt to increase the number of known palmitoylation sites in *P. falciparum*, a series of trials were performed in order to develop a site-identification (ID) palmitome purification method. The site-ID palmitome purification method was adapted from the ABE method of palmitoyl-protein purification (described in detail in Section 1.6.1), which was one of the techniques previously used to purify the total palmitome from *P. falciparum* schizonts [4]. ABE was the method of choice to adapt the site-ID purification method as all the chemical steps of ABE take place post-extraction of the proteome and does not rely on *in vitro* metabolic labelling. This would then allow the site-ID

purification method to be used to identify palmitoylation sites in other stages of the *Plasmodium* life cycle (such as the mosquito stages) or in other *Plasmodium* species (such as *P. berghei*), which are not amenable to the metabolic labelling required in the metabolic labelling and click chemistry method of palmitoyl-protein purification (described in Section 1.6.2).

Two key differences were present between the site-ID purification method and the original ABE purification method used in the purification of the total schizont palmitome. Firstly, the ABE purification of the total palmitome was coupled with stable isotope labelling with amino acids in cell culture (SILAC), where parasites were metabolically labelled with 'light' or 'heavy' versions of isoleucine in culture, in order to allow accurate protein quantification by mass spectrometry [4-6]. However, as mentioned above, it was intended that the site-ID purification method should be compatible for use in other stages of the *Plasmodium* life cycle, or in other Apicomplexan species, which would not be amenable to *in vitro* metabolic labelling. Thus, it was decided that SILAC labelling would not be used in the site-ID purifications and quantification by mass spectrometry would be performed using stable isotope dimethyl labelling [7], which is performed after the ABE purification is complete, or using label-free methods [8] (as will be described in detail below).

Secondly, for the purification of the total schizont palmitome, the proteome was extracted from SILAC labelled schizont stages and subjected to the different chemical steps of ABE, followed by streptavidin affinity purification and elution (as described in Section 1.6.1). The eluted proteins were then separated by SDS-PAGE and digested in-gel with trypsin, before analysis by high-resolution tandem mass spectrometry (LC-MS/MS) [4]. In order to identify palmitoylation sites, digestion with trypsin would have to take place before the elution of the samples from the streptavidin resin (instead of in-gel digestion after elution) so that only palmitoylated peptides are purified, and this was what was performed in the site-ID purification trials (as will be described in detail below).

During the development of the site-ID palmitome purification method, several different conditions were trialled in order to determine the optimum conditions required, and each of these trial purifications are described in the sections below. All trials of the site-ID purification method were performed on *P. falciparum* schizont stages, both because schizonts contain multiple individual merozoites, and therefore allow the generation of larger amounts of material than earlier stages of development where only a single parasite is present, and also to allow comparison of the site-ID datasets with the previously published schizont total palmitome dataset [4]. Although it is unlikely that the site-ID purification method would produce a dataset which completely overlapped with the total palmitome, comparison between the site-ID datasets and the total palmitome could still act as a rough guide for the efficiency and accuracy of the trial purifications.

The development and optimisation of the site-ID palmitome purification method described below was performed together with Mark Collins, Choudhary lab, and all quantitative mass spectrometry analysis was kindly performed by Mark Collins, Choudhary lab.

3.1.2. Development of the site-ID palmitome purification method – Trial 1

Briefly, ABE consists of three main sequential chemical steps which take place post-extraction of the proteome: (1) treatment of samples with NEM, which irreversibly blocks any free cysteine thiols present, (2) cleavage of the thioester bonds between palmitoyl groups and cysteine residues using hydroxylamine and (3) biotinylation of newly-exposed free thiols using the thiol-reactive biotinylation reagent, HPDP-biotin, which allows the streptavidin affinity purification of biotinylated proteins, that can be then eluted from streptavidin-agarose beads by reduction with a reducing agent. Importantly, the control treatment of the proteome, where an equal amount of the proteome is processed in parallel under identical conditions, but in the absence of hydroxylamine cleavage (during step 2), must always be present, as the quantitative measurement of the enrichment of a protein in the hydroxylamine-treated 'palmitome' samples over the non-hydroxylamine-treated 'control' samples determines whether the protein can be regarded as palmitoylated.

For the first trial of the site-ID palmitome purification, approximately 10^9 *P. falciparum* strain 3D7 schizont stage parasites (equivalent to 200 mL of parasite *in vitro* blood-stage culture at 5% haematocrit and 8-10% parasitemia) were extracted from infected erythrocytes by saponin lysis, as described in Section 2.2.5 of the Materials and Methods. The proteome was extracted from saponin-treated parasite pellets by lysis in a Triton X-100-based buffer and subjected to the three chemical steps of ABE, followed by streptavidin affinity purification, as described above and in Section 2.6.1 of the Materials and Methods. As before, an equal amount of the proteome was subjected to identical steps of ABE, but without hydroxylamine treatment (instead of hydroxylamine, an equal volume of 100 mM Tris-Cl pH 7.4 was added into the control samples). Before elution of the biotinylated palmitome and control samples from streptavidin-agarose beads, both samples were digested with trypsin while bound to the beads in order to digest the proteins into peptide fragments. The samples were then stringently washed so that unbound peptide fragments were washed away and only the biotinylated peptide fragments remained bound to streptavidin-agarose. The peptide fragments were then eluted from the beads using the reducing agent Tris(2-carboxyethyl)phosphine (TCEP). The trypsin digestion of the samples before elution ensured that only biotinylated peptide fragments

(which by definition contained the palmitoylated cysteine) were purified. The steps of Trial 1 of the site-ID palmitome purification are as shown in the schematic in Figure 3.1A.

The eluted peptide fragments were then subjected to stable isotope dimethyl labelling in order to allow quantitative mass spectrometry analysis of the peptide fragments. Stable isotope dimethyl labelling uses formaldehyde and cyanoborohydride to modify all primary amines (N-termini and the side chain of lysine residues). By using different isotopomers of formaldehyde and cyanoborohydride, the mass added to each sample can be varied. The labelled samples are then pooled and analysed by LC-MS/MS, where peptides from the different isotope-labelled samples can be identified due to their mass difference. Quantification is then performed by measuring the ratio of the peak intensities of the isotope-labelled peptide pairs [7, 9]. For Trial 1 of the site-ID purification method, non-hydroxylamine-treated control samples were treated with 'medium' formaldehyde (deuterated formaldehyde) and sodium cyanoborohydride, and hydroxylamine-treated palmitome samples were treated with 'heavy' formaldehyde (^{13}C -deuterated formaldehyde) and sodium cyanoborodeuteride. The samples were analysed by LC-MS/MS and MaxQuant [10] was used to generate specific enrichment ratios for each of the peptide identifications, where the enrichment ratio was the intensity of the peptide in the palmitome (hydroxylamine-treated) sample compared to the intensity of the peptide in the control (non-hydroxylamine-treated) sample.

In total, 158 peptides, which had enrichment ratios and contained a free cysteine, were identified. This corresponded to a total of 103 protein identifications, indicating that some proteins had multiple peptides which contained a palmitoylated cysteine, suggesting that some proteins may be multiply-palmitoylated. Alternatively, the same peptide fragment could also be identified more than once, perhaps due to the abundance of the protein, and this also contributed to the higher number of peptide identifications compared to protein identifications. The 158 peptides identified were then grouped into enriched or highly enriched classes based on the same cut-off criteria which was used previously in the analysis of the total schizont palmitome [4]:

- Enriched peptides: All peptides with an enrichment ratio above the median enrichment ratio plus the median of the absolute deviation of each individual enrichment ratio from the median enrichment ratio.
- Highly enriched peptides: All peptides with an enrichment ratio above the mean enrichment ratio plus the mean of the absolute deviation of each individual enrichment ratio from the mean enrichment ratio.

Using these cut-off criteria, 41 peptides (corresponding to 32 protein identifications) were defined as enriched, and out of these 41 peptides, 9 peptides (corresponding to 6 protein identifications) could be further defined as highly enriched. A scatter plot showing the median peptide intensities against the MaxQuant generated enrichment ratios for all 158 identified peptides are shown in Figure 3.1B, with peptides classified as enriched and highly enriched highlighted in blue and red respectively.

Of the 32 proteins which had peptide identifications classified as enriched, 26 proteins were also present in the total schizont palmitome [4]. This indicated that 81% of the proteins classified as enriched in this dataset (26 out of 32 proteins) were also present in the total palmitome (Figure 3.1C). As mentioned above, a complete overlap between the site-ID purifications and the total palmitome is not expected. This is firstly due to the fact that the site-ID purification was based solely on ABE, whereas the total palmitome was compiled from multiple replicates of both ABE and metabolic labelling with click chemistry purifications, both of which purify overlapping as well as unique sets of proteins. Secondly, in the site-ID purifications, individual peptide identifications were analysed, whereas in the total palmitome purification, protein identifications were analysed, and each protein was identified by contributions from multiple peptide fragments. Nevertheless, the presence of 81% of the enriched proteins of this site-ID dataset in the total palmitome was encouraging.

Included in the overlap list were highly studied palmitoylated proteins such as glideosome-associated protein 45 (GAP45), chloroquine resistance transporter (CRT) and myosin A tail domain interacting protein (MTIP) [4]. This, along with the good overlap between enriched proteins in this dataset and the total palmitome, indicated that in this first trial of the site-ID palmitome purification, palmitoylated proteins were indeed being purified and the use of this method resulted in the identification of specific palmitoylated peptide fragments, containing the putative palmitoylated cysteines, for each of the enriched proteins purified.

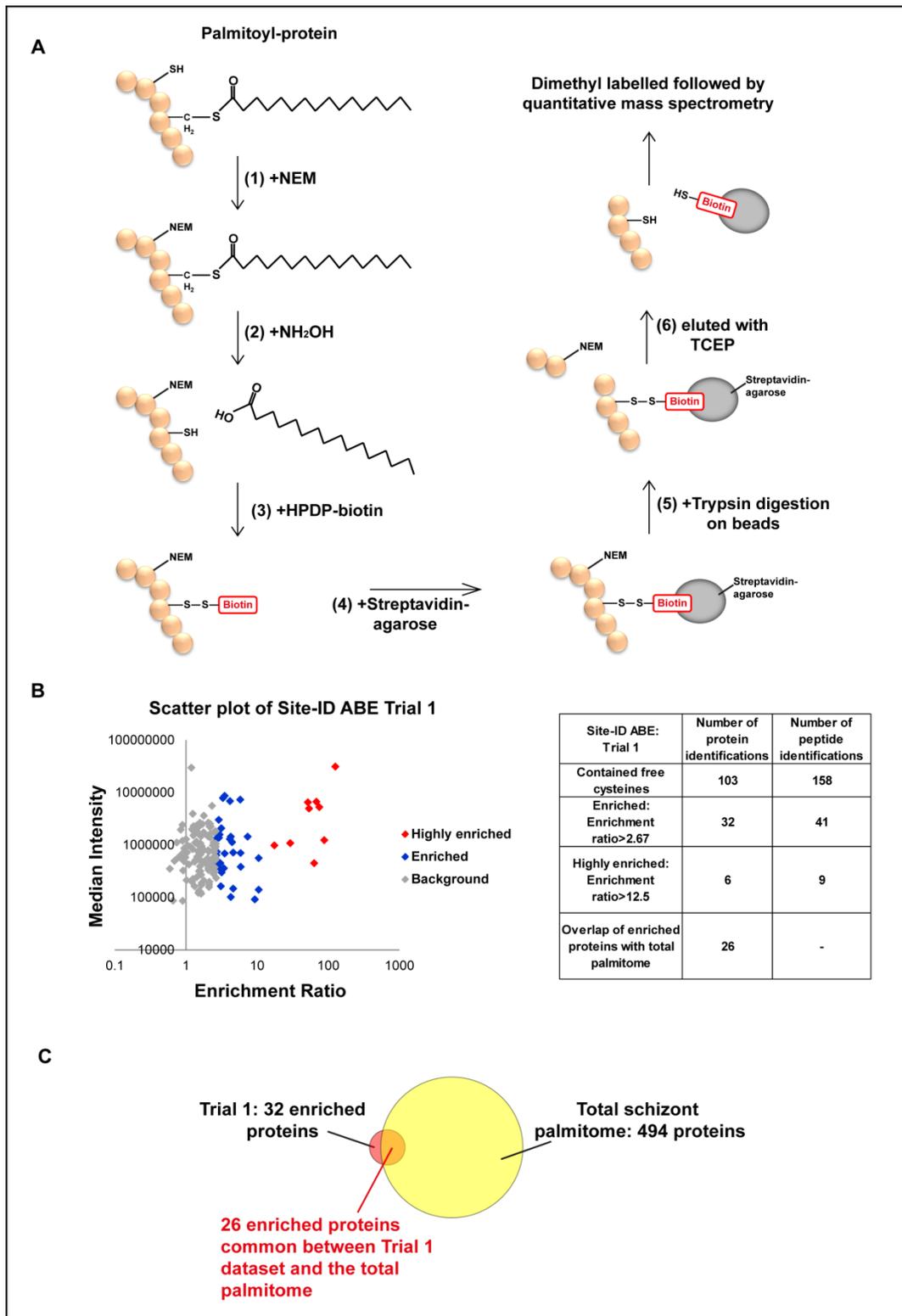


Figure 3.1: Trial 1 site-ID palmitome purification by ABE. (A) The site-ID palmitome purification was based on the ABE method of palmitoyl-protein purification. In the Trial 1 site-ID purification, the proteome extracted from saponin-treated parasite pellets was (1) first treated with NEM to irreversibly block free thiols. This was followed by (2) cleavage of palmitate groups using hydroxylamine (NH₂OH) and the (3) biotinylation of newly-exposed free thiols with HPDP-biotin. As always, a 'control' sample, which was subjected to identical conditions as the 'palmitome' sample, but without hydroxylamine treatment, was also prepared. Biotinylated proteins were purified by (4) streptavidin affinity purification, and (5) trypsin digestion was performed while the

samples were bound to streptavidin-agarose beads, in order to digest the proteins into peptide fragments. After stringent washes, only biotinylated peptide fragments remained bound to the beads and were (6) eluted by reduction with TCEP, followed by dimethyl labelling and quantitative mass spectrometry analysis. **(B)** Scatter plot displaying the median protein intensities against the MaxQuant generated enrichment ratios for all the peptides purified by the Trial 1 site-ID purification. Peptides classified as enriched and highly enriched (according to the cut-off criteria described in the text) are highlighted in blue and red respectively. The numbers of protein and peptide identifications found are listed in the table on the right. **(C)** Venn diagram representing the overlap between the enriched proteins in the Trial 1 dataset and the total schizont palmitome.

However, as described in Section 1.6.3, a number of false positive identifications are unavoidable, and were probably present in this dataset, even in the proteins defined as enriched. This is due firstly to the ABE method itself, which enriches other proteins that also use thioester linkages, and is highly dependent on the complete and irreversible blockage of free thiols before hydroxylamine cleavage (detailed in Section 1.6.3). Secondly, false positive identifications can also occur due to the inherent background that occurs in large-scale protein purifications and in streptavidin bead enrichments.

Additionally, cysteine-containing peptides were also directly identified by MaxQuant in the control (non-hydroxylamine-treated) samples. A low level of cysteine-containing peptides are expected to be present in control samples, and in fact, relative proteomic quantitation and the generation of the enrichment ratios is based on the presence of these cysteine-containing-peptides in the control samples. Typically, these peptides should usually be at a very low abundance in control samples, and can be regarded as 'noise'. The identification of a particular peptide ion by MaxQuant in the palmitome sample will then result in a search by MaxQuant for the corresponding peptide ion in the noise of the control sample, based on matching the mass and retention time of the peptide ion, in order to generate an enrichment ratio. The peptide intensities in the noise of the control samples are usually too low to be directly identified by MaxQuant. The fact that some cysteine-containing-peptides were being directly sequenced and identified by MaxQuant in the control samples of this purification however indicates that there is a background of cysteine-containing peptides present in the control samples, which is higher than the normal noise usually found in control samples. This suggests that the initial treatment with NEM to block free thiols was not efficient or may not have gone to completion. The presence of this high background of cysteine-containing peptides in the control samples could have an effect on the enrichment ratios generated, and this could subsequently have an effect on the definition of the cut-off criteria and the classification of proteins as enriched, potentially increasing the number of false-positive identifications.

In summary, although palmitoylated proteins with corresponding cysteine-containing peptides were identified in this first trial, the number of enriched protein identifications was relatively small compared to the total number of proteins in the *P. falciparum* proteome (there are over 5000 proteins coded for in the *P. falciparum* genome [11] with over 1000 proteins up-regulated in schizont stages [12]). Additionally, although 81% of the enriched proteins (26 out of 32 enriched proteins) were also present in the total schizont palmitome, this was equivalent to only 5% of the total schizont palmitome (26 out of 494 proteins in the total palmitome) (Figure 3.1C), indicating that many palmitoylated proteins were potentially not purified, or did not pass the cut-off criteria in this trial purification. Taken together, this suggested that the total complement of palmitoylated proteins was not being purified by this first trial purification and further optimisation was still required.

3.1.3. Development of the site-ID palmitome purification method – Trial 2

For Trial 2 of the site-ID palmitome purification, approximately 10^9 *P. falciparum* strain 3D7 schizont stage parasites (200 mL *in vitro* culture at 5% haematocrit and 8-10% parasitemia) was extracted from infected erythrocytes by saponin-lysis, as was done for Trial 1. However, in this trial purification, the extraction of the proteome from saponin-treated parasite pellets was performed more extensively (as described in Section 2.6.2 of the Materials and Methods), in order to ensure that the maximum amount of the parasite proteome was extracted. Briefly, lysis of the saponin-treated parasite pellets was performed using an SDS-based buffer instead of a Triton X-100-based buffer, and the lysate was also homogenised, and then sheered by passing through a fine gauge needle. Additionally, the insoluble material remaining from the initial lysis with SDS buffer was further treated with buffer containing 8 M urea in order to solubilise and denature the more hydrophobic proteins.

The initial irreversible blockage of free thiol groups in the proteome sample was performed using iodoacetamide (IAA), another sulphhydryl-reactive alkylating agent, instead of NEM, in order to determine whether the blockage of free cysteine thiols occurred more efficiently and completely using IAA. The IAA-blocked proteome sample was then subjected to the remaining two chemical steps of ABE – cleavage of palmitoyl groups using hydroxylamine and biotinylation of newly-exposed free thiols with HPDP-biotin. In this trial purification however, the hydroxylamine-cleavage and biotinylation steps, as well as all the wash steps in between, were performed using 30 kDa molecular weight cut-off (MWCO) spin columns. For the purification of the total palmitome [4] and in Trial 1 of the site-ID purification, buffer exchange during washes and when changing to the different ABE

chemical steps, occurred by performing multiple chloroform-methanol (C/M) precipitations, which was time-consuming and increased the chances of sample loss between each C/M precipitation. The use of the 30 kDa MWCO spin columns here allowed buffer exchange to occur more easily and efficiently, and reduced the loss of samples between each wash and each chemical step. As before, the non-hydroxylamine-treated control was also prepared from an equal amount of the proteome, and this control sample was also treated in the 30 kDa MWCO spin columns. Additionally, all wash steps were performed in buffer containing 8 M urea, in order to ensure that the proteins were fully denatured and unfolded so that all palmitate groups would be exposed for hydroxylamine cleavage.

After biotinylation of the free thiols produced from hydroxylamine cleavage of palmitate groups, the samples were treated with trypsin in order to digest the proteins into peptide fragments. Trypsin digestion was also performed in the 30 kDa MWCO spin columns. Biotinylated peptide fragments were then purified by streptavidin affinity purification. The trypsin digestion of the protein samples before streptavidin affinity purification ensured that only biotinylated peptide fragments (which by definition contained the palmitoylated cysteine) would be bound by the streptavidin-agarose beads. Extensive and stringent wash steps were then performed in order to remove any unbound peptide fragments and the remaining bound palmitoylated peptide fragments were then eluted from the streptavidin-agarose beads using TCEP. The steps of Trial 2 of the site-ID palmitome purification are as shown in the schematic in Figure 3.2A.

The eluted peptide fragments were then subjected to stable isotope dimethyl labelling as before (described in Section 3.1.2) and analysed by LC-MS/MS. MaxQuant was used to generate specific enrichment ratios (palmitome over control) for each of the peptide identifications.

In total, 431 peptides containing enrichment ratios and a free cysteine were identified, and this corresponded to a total of 291 proteins. This set of 431 peptide identifications were then grouped into enriched and highly enriched classes based on the same cut-off criteria described in Section 3.1.2 above. Using these cut-off criteria, 139 peptides (corresponding to 117 protein identifications) were defined as enriched, and out of these 139 peptides, 28 peptides (corresponding to 24 protein identifications) could be further defined as highly enriched. A scatter plot showing the median peptide intensities against the MaxQuant generated enrichment ratios for all 431 peptide identifications are shown in Figure 3.2B. Of the 117 proteins which had peptide identifications classified as enriched, 58 proteins were also present in the total palmitome [4]. This indicated that 50% of the enriched proteins in this dataset (58 out of 117 proteins) were present in the total palmitome (Figure 3.2C). Again, proteins previously shown to be palmitoylated, such as GAP45 and CRT, were also present and classified as enriched in this trial purification.

Although only 50% of enriched proteins in this dataset overlapped with the total schizont palmitome, this was equivalent to 12% of the total palmitome (58 proteins out of 494 proteins in the total palmitome) (Figure 3.2C). This implied that more of the total palmitome was being purified in this trial compared to Trial 1, suggesting that more palmitoylated proteins were being purified or managed to pass the cut-off criteria in this purification. However, it also appeared a larger portion of the Trial 2 enriched dataset did not overlap with the total palmitome (50% of the Trial 2 enriched dataset overlapped with the total palmitome compared to 81% of the Trial 1 enriched dataset). As explained in Section 3.1.2 above, a complete overlap between the site-ID purifications and the total palmitome is not expected. This is further exacerbated for the Trial 2 purification where there were significant differences in the methodology (for example, different lysis conditions and treatment with IAA instead of NEM), and this could be the reason why only half of the Trial 2 enriched dataset overlapped with the total palmitome.

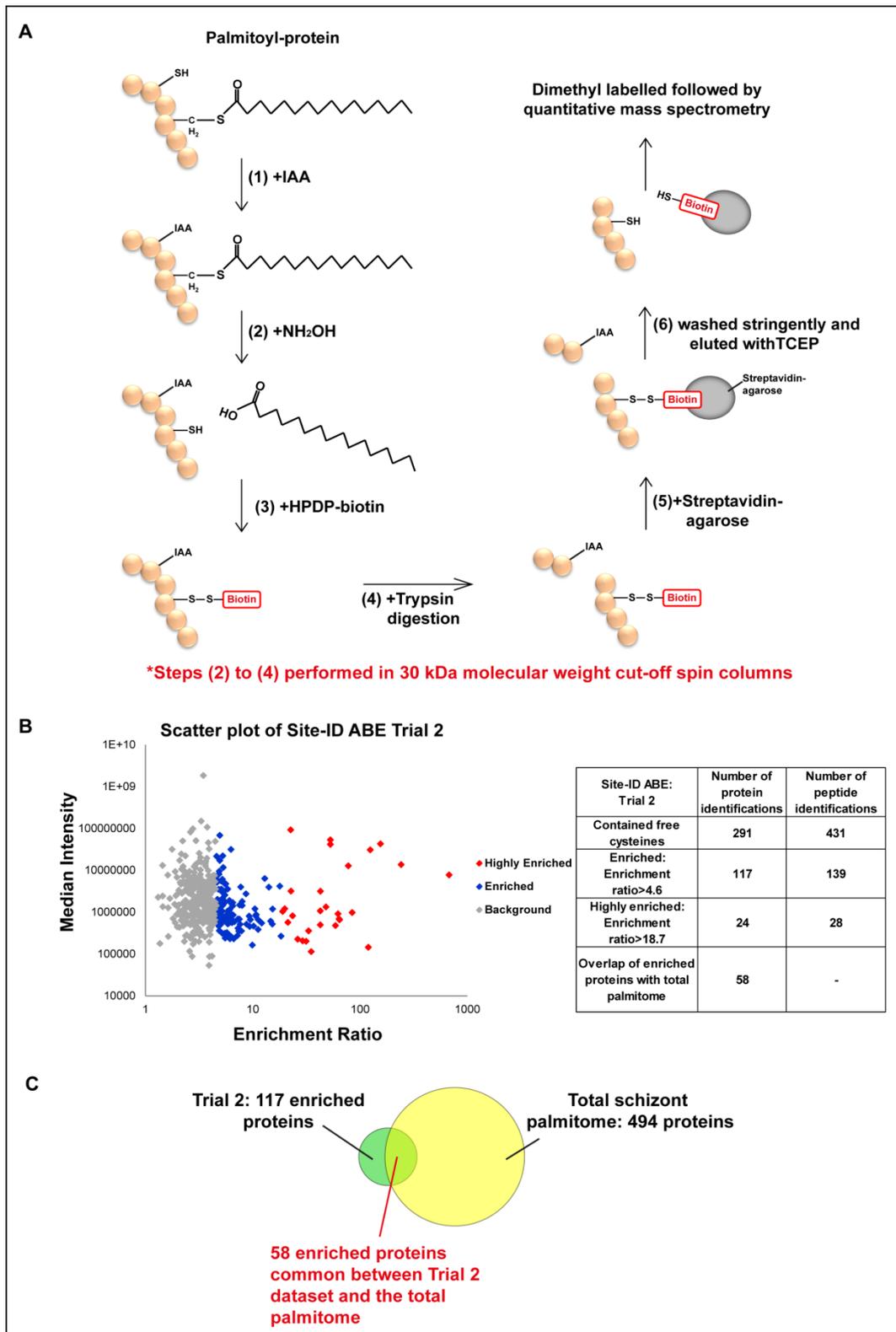


Figure 3.2: Trial 2 site-ID palmitome purification by ABE. (A) In the Trial 2 site-ID purification, the proteome was extensively extracted from saponin-treated parasite pellets (as described in the text) and was (1) treated with IAA to irreversibly block free thiols. This was followed by (2) cleavage of palmitate groups using hydroxylamine (NH₂OH) and the (3) biotinylation of newly-exposed free thiols with HPDP-biotin. As always, a non-hydroxylamine-treated control sample was processed in parallel. (4) Trypsin digestion was then performed in order to digest the proteins into peptide fragments. Steps (2) to (4) were performed in 30 kDa molecular weight cut off (MWCO) spin columns in order to ease buffer exchange during washes and changes to different

chemical steps. Biotinylated peptides were purified by (5) streptavidin affinity purification, and unbound peptides removed by stringent washes. The biotinylated peptide fragments which remained bound to the beads were (6) eluted by reduction with TCEP, followed by dimethyl labelling and quantitative mass spectrometry analysis. **(B)** Scatter plot displaying the median protein intensities against the MaxQuant generated enrichment ratios for all the peptides purified by the Trial 2 site-ID purification. Peptides classified as enriched and highly enriched (according to the cut-off criteria described in the text) are highlighted in blue and red respectively. The numbers of protein and peptide identifications found are listed in the table on the right. **(C)** Venn diagram representing the overlap between the enriched proteins in the Trial 2 dataset and the total schizont palmitome.

In summary, the site-ID palmitome purification method used in this Trial 2 purification appeared to be an improved purification compared to that of Trial 1, as evidenced by the increased total number of peptides identified, the increased number of peptides classified as enriched, as well as the greater spread of enrichment ratios, and the identification of more proteins which were also present in the total schizont palmitome. However, cysteine-containing peptides were still directly identified by MaxQuant in the control samples, indicating that the background of cysteine-containing peptides, with intensities above the normal 'noise' usually found in control samples, was still present. This means that the use of IAA instead of NEM did not improve the blockage of free thiols. Additionally, the dataset of protein identifications was again smaller than total palmitome dataset, indicating that further improvements were still required.

3.1.4. Development of the site-ID palmitome purification method – Trial 3A

For the next trial of the site-ID palmitome purification, approximately 10^9 *P. falciparum* strain 3D7 schizont stage parasites (200 mL *in vitro* culture at 5% haematocrit and 8-10% parasitemia) was extracted from infected erythrocytes by saponin-lysis as performed for Trial 1 and 2. Extraction of the proteome from saponin-treated parasite pellets was then performed exactly the same as in Trial 2 (described in Section 3.1.3). Initial blockage of free thiols in the proteome was performed using IAA and the hydroxylamine treatment and biotinylation steps of ABE performed in the 30 kDa MWCO spin columns, as in Trial 2 (Section 3.1.3). As always, the non-hydroxylamine-treated control sample was also prepared in parallel. In this trial purification however, during the hydroxylamine cleavage and biotinylation steps, NEM was added only into control samples, in an attempt to block any free thiols that may become exposed due to palmitate groups falling off during the ABE procedure. This was performed in the hopes that the background of cysteine-containing peptides observed in control

samples in the previous trial purifications could be reduced. Trypsin digestion of the samples in order to produce peptide fragments, streptavidin affinity purification of only biotinylated peptide fragments and elution of the peptides were all performed similar to Trial 2 (Section 3.1.2). The steps of Trial 3A of the site-ID palmitome purification are as shown in the schematic in Figure 3.3A.

Aliquots of the palmitome and control elutions from this Trial 3A purification were kept aside for label-free analysis in order to determine whether a label-free analysis would produce a better dataset (as will be described in the section below). The rest of the palmitome and control elutions were subjected to stable isotope dimethyl labelling (as for the previous trial purifications), and were analysed by LC-MS/MS. MaxQuant was used to generate the specific enrichment ratios (palmitome over control) for each of the peptide identifications.

In total, 700 peptides, which had enrichment ratios and contained a free cysteine, were identified, and this corresponded to a total of 440 protein identifications. This set of 700 peptides was grouped into enriched and highly enriched classes using the same cut-off criteria as above (detailed in Section 3.1.2). Using these cut-off criteria, 214 peptides (corresponding to 179 protein identifications) were defined as enriched, and out of these 214 peptides, 62 peptides (corresponding to 53 protein identifications) could be further defined as highly enriched. A scatter plot showing the median peptide intensities against the MaxQuant generated enrichment ratios for all 431 peptide identifications are shown in Figure 3.3B. Of the 179 proteins which had peptide identifications classified as enriched, 69 proteins were also present in the total palmitome. This indicated that 39% of the proteins classified as enriched in this dataset (69 out of 179 proteins) were present in the total palmitome, and this was equivalent to 13% of the total palmitome (69 proteins out of 494 proteins in the total palmitome) (Figure 3.3C).

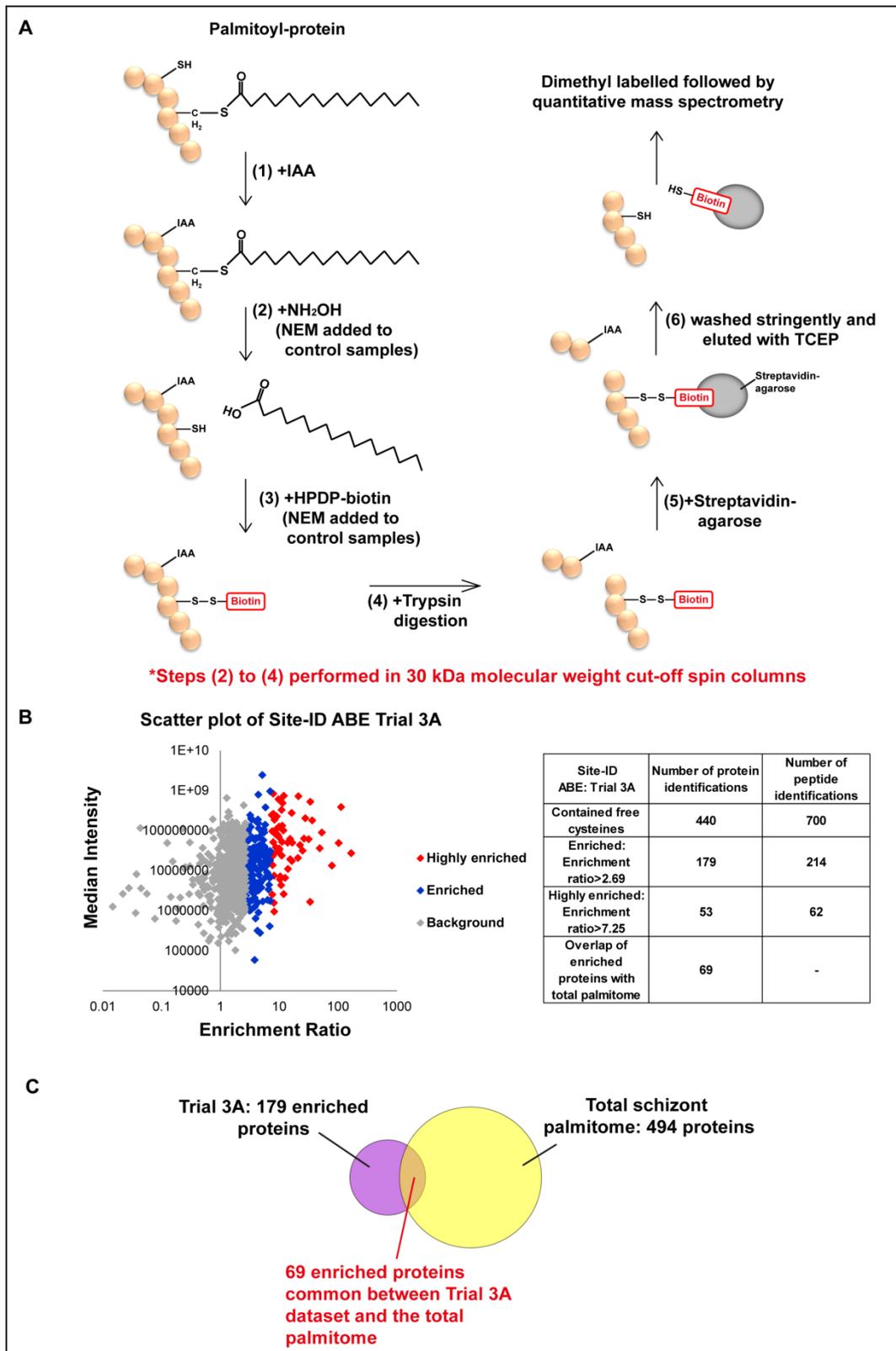


Figure 3.3: Trial 3A site-ID palmitome purification by ABE. (A) The Trial 3A site-ID purification was similar to that of Trial 2. The proteome was extensively extracted from saponin-treated parasite pellets as in Trial 2, and protein samples were (1) blocked with IAA, (2) treated with hydroxylamine (NH₂OH) (except for control samples) and (3) biotinylated using HPDP-biotin. In this Trial 3A purification however, NEM was added only into control samples during steps (2) and (3) in an attempt to reduce the high background of cysteine-containing peptides seen in the control samples. (4) Trypsin digestion was then performed to digest the proteins into

peptide fragments. Steps (2) to (4) were performed in 30 kDa MWCO spin columns as before. Biotinylated peptides were purified by (5) streptavidin affinity purification, and unbound peptides removed by stringent washes. The biotinylated peptide fragments which remained bound to the beads were (6) eluted by reduction with TCEP. Aliquots of palmitome and control samples were kept aside for label-free analysis. The remaining palmitome and control samples were dimethyl labelled and analysed by quantitative mass spectrometry as before. **(B)** Scatter plot displaying the median protein intensities against the MaxQuant generated enrichment ratios for all the peptides purified by the Trial 3A site-ID purification. Peptides classified as enriched and highly enriched (according to the cut-off criteria described in the text) are highlighted in blue and red respectively. The numbers of protein and peptide identifications found are listed in the table on the right. **(C)** Venn diagram representing the overlap between the enriched proteins in the Trial 3A dataset and the total schizont palmitome.

Again, proteins previously shown to be palmitoylated, such as CRT, were enriched in this dataset. Strangely however, GAP45, a protein known to be palmitoylated and usually identified as enriched or highly enriched in the previous site-ID purifications, was identified in this dataset but did not pass the cut-off criteria to be defined as enriched.

Although the number of enriched proteins identified in this dataset had increased, the overlap with the total palmitome was not improved from the Trial 2 purification (the Trial 2 dataset purified 12% of the total palmitome while the Trial 3A dataset purified 13% of the total palmitome), and the palmitoylated protein GAP45, which had previously been consistently purified and classified as enriched, was not classified as enriched in this trial. This implied that Trial 3A was not an improvement from the previous purification, although the addition of NEM into the control samples did indeed reduce the direct identification of cysteine-containing peptides in the control samples by MaxQuant.

3.1.5. Development of the site-ID palmitome purification method – Trial 3B

As mentioned in Section 3.1.4 above, aliquots of the palmitome and control elutions from the Trial 3A site-ID purification were kept aside for label-free analysis in order to determine whether quantification using a label-free approach would be more appropriate for the site-ID palmitome purification compared to stable isotope dimethyl labelling.

In the dimethyl labelling approach, the palmitome and control samples –each labelled with a different isotopomer of formaldehyde- were pooled together and analysed by LC-MS/MS in a single run. In contrast, for label-free analysis, the palmitome and control samples were kept separate and

were subjected to individual LC-MS/MS runs. Protein quantification was then measured by comparing the peak intensities from each individual run for a particular peptide [8]. Due to the run-to-run experimental variations of LC-MS/MS, label-free quantification is thus less accurate and less reproducible, and is more variable compared to quantification with dimethyl labelling, resulting in the need for more rigorous analysis of the data acquired. However, despite the lower accuracy and lower reliability when measuring small quantitative differences, label-free analysis may provide greater depth of proteome coverage as well as a higher dynamic range of quantification [13]. Additionally, dimethyl labelling of samples requires that no buffers containing primary amines be present during the labelling reaction. Thus, samples need to be desalted prior to dimethyl labelling, and this can lead to some sample losses. No desalting is required for label-free analysis, thus reducing the loss of samples. Lastly, although dimethyl labelling is not expected to change the chemical properties of labelled peptides, it still involves modification of the peptides, and as the peptides being labelled here contain free cysteines, may result in unwanted modification at the side chains of the free cysteines by formaldehyde, which could lead to reduced identifications. With the label-free approach however, the eluted peptides are not further modified prior to mass spectrometry analysis. Thus, due to these differences, the label-free approach was tested to determine its suitability for the site-ID palmitome purifications.

The aliquots of the palmitome and control elutions from the Trial 3A site-ID purification (method described in Section 3.1.4) was analysed by LC-MS/MS without any dimethyl labelling. MaxQuant was again used to generate specific enrichment ratios (palmitome over control) for each of the peptide identifications.

In total, 2295 peptides, which had enrichment ratios and contained a free cysteine, were identified, and this corresponded to a total of 1173 proteins. This dataset of peptide identifications obtained using the label-free approach was clearly larger than previously obtained using dimethyl labelling, suggesting that a label-free analysis might provide better coverage for the site-ID palmitome purifications. The 2295 peptides was then grouped into enriched and highly enriched classes following the same cut-off criteria used for the previous site-ID trial purifications (detailed in Section 3.1.2). Using these cut-off criteria, 713 peptides (corresponding to 521 protein identifications) were defined as enriched, and out of these 713 peptides, 165 peptides (corresponding to 138 protein identifications) could be further defined as highly enriched. A scatter plot showing median peptide intensities against the MaxQuant generated enrichment ratios for all 2295 peptide identifications found in this Trial 3B site-ID palmitome purification is shown in Figure 3.4A.

Of the 521 proteins which had peptide identifications classified as enriched, 181 proteins were also present in the total schizont palmitome. This indicated that 37% of the total palmitome (181 out of 494 proteins in the total palmitome) was identified in this dataset (Figure 3.4B), suggesting that potentially more putative palmitoyl-proteins were being identified when using a label-free quantitative approach with the site-ID purification. As for the previous trial site-ID purifications, proteins shown previously to be palmitoylated, such as GAP45 and CRT, were identified as enriched in this dataset.

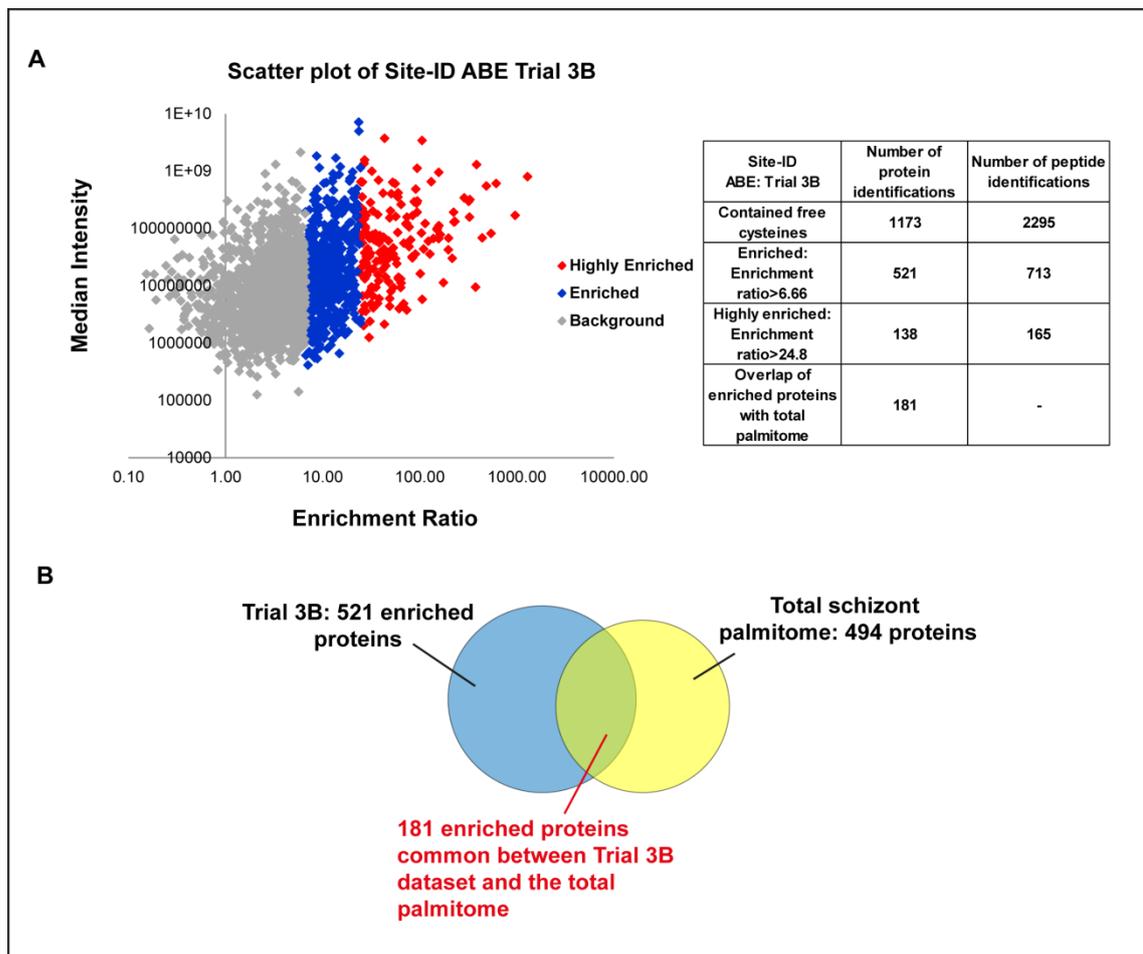


Figure 3.4: Trial 3B site-ID palmitome purification by ABE. The aliquots kept aside from the Trial 3A site-ID purification were analysed using a label-free approach by quantitative mass spectrometry, and this analysis was named Trial 3B. The scatter plot displays the median protein intensities against the MaxQuant generated enrichment ratios for all the peptides purified by the Trial 3B site-ID purification. Peptides classified as enriched and highly enriched (according to the cut-off criteria described in the text) are highlighted in blue and red respectively. The numbers of protein and peptide identifications found are listed in the table on the right. **(B)** Venn diagram representing the overlap between the enriched proteins in the Trial 3B dataset and the total schizont palmitome.

Although 37% of the total palmitome was found to be enriched in this dataset, this was equivalent to only 35% of the enriched proteins in this dataset (181 out of 521 proteins) (Figure 3.4B), which was less than previously observed in the other trial purifications. This indicated that more than half of the proteins in this dataset were not found in the total palmitome. This could be due to the differences between the purification methods as detailed in the sections above (for example, the use of ABE purification alone in these site-ID purifications and differences in methodology such as the use of IAA instead of NEM), which would reduce the overlap between the two datasets. Additionally, another reason could be due to the fact that the analysis of the total palmitome was performed with SILAC labelling, whereas Trial 3B was performed using label-free analysis, which has been shown to provide greater proteome coverage compared to labelling approaches, including metabolic labelling [14, 15]. However, as described above, due to the nature of label-free analysis, quantification is generally more inaccurate, and although the dataset is larger, it must be more cautiously and rigorously analysed. For example, the accuracy of this label-free approach could be improved by performing further replicates and accepting only those proteins that were repeatedly identified.

In addition to the dataset of peptide identifications possessing enrichment ratios, a dataset of peptides with peak intensities found only in the palmitome (hydroxylamine-treated) sample and not in the control (non-hydroxylamine-treated) sample was also identified in this label-free trial analysis. It is possible that these peptides were truly only present in the palmitome sample, indicating that these peptides truly contained the palmitoylated cysteine. Indeed, peptides from known palmitoylated proteins, such as GAP45, were present in this 'palmitome-only' dataset. However, it is also equally possible that the corresponding peak for a particular peptide was not matched in control samples, due to run-to-run variation or differences in the retention time, which are the errors associated with label-free analysis. This was not observed in the previous site-ID trial purifications because in dimethyl labelling approaches, analysis is based on the co-elution of the differentially-labelled versions of a particular peptide. As the peptides in this palmitome-only dataset were not detected in the control samples, an enrichment ratio could not be assigned and cut-off criteria could not be applied to this dataset, making it even more difficult to distinguish between truly palmitoylated peptides and false-positive identifications. Thus, the peptide identifications in this 'palmitome-only' dataset were not further considered at this time.

Despite the lower accuracy of the label-free approach, this Trial 3B site-ID label-free palmitome purification appeared to be improved compared to the previous purifications, with increased proteome coverage and more enriched protein identifications found to be also present in the total

palmitome, suggesting that the use of label-free quantification may provide more useful information for these site-ID purifications.

3.1.6. Development of the site-ID palmitome purification method – Trial 4

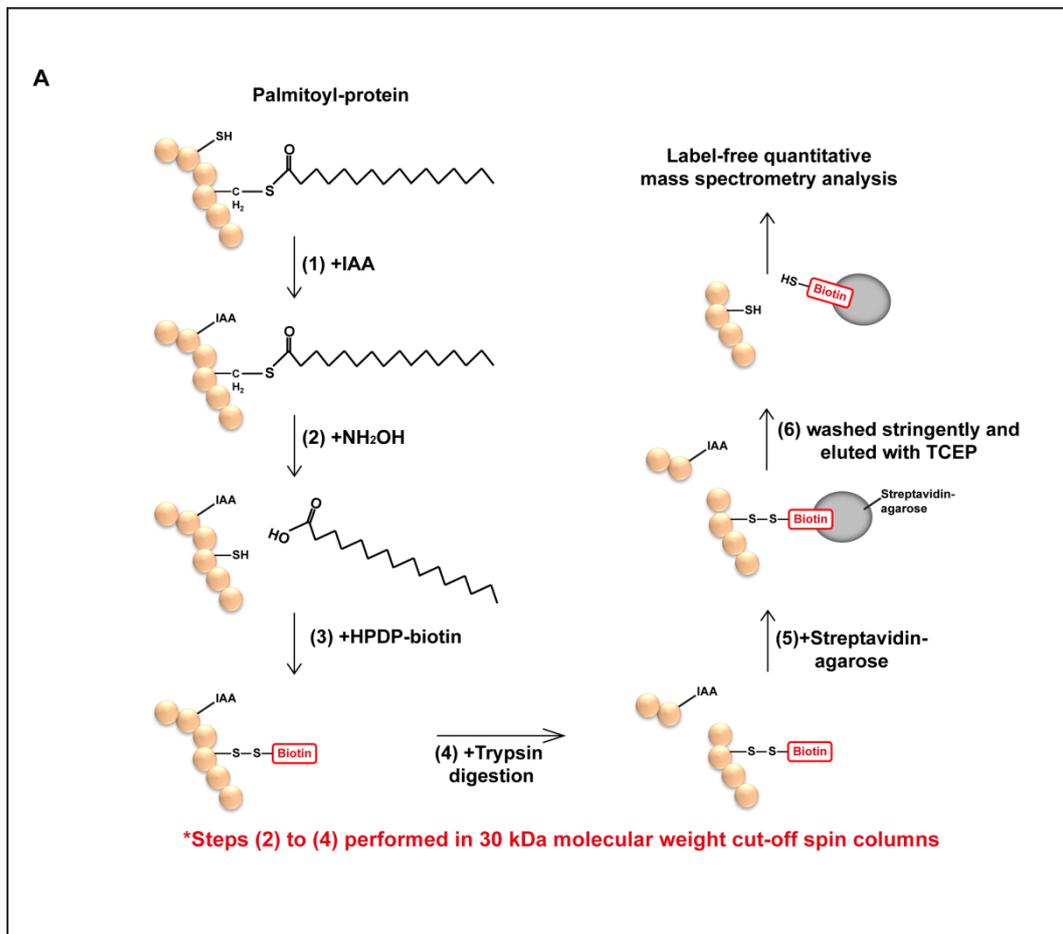
The previous trial purifications revealed that using a label-free approach appeared to be better suited for the site-ID palmitome purifications. Thus, another site-ID purification was next performed based on the Trial 3A purification method (shown in the schematic in Figure 3.3A) coupled with the Trial 3B label-free analysis (described in Section 3.1.5 above).

For this large-scale purification, two biological replicates and two technical replicates were performed. Each biological replicate consisted of approximately 10^9 *P. falciparum* strain 3D7 schizont stage parasites (200 mL of *in vitro* culture at 5% haematocrit and 8-10% parasitemia). Parasites were extracted from infected erythrocytes by saponin-lysis, and extraction of the proteome from saponin-treated parasite pellets was performed exactly the same as in Trial 3A. Initial blockage of free thiols was performed using IAA, and the hydroxylamine-cleavage and biotinylation steps were performed in 30 kDa MWCO spin columns as in Trial 3A. As always, the non-hydroxylamine-treated control sample was prepared in parallel using an equal amount of proteome. Although the addition of NEM into control samples during the hydroxylamine-cleavage and biotinylation steps of the Trial 3A purification appeared to reduce the high background of cysteine-containing peptides in the control sample, there was the possibility that the addition of NEM only to the control sample could end up artificially biasing the results, and might result in a dataset that did not truly represent the palmitoylated peptides in schizont stage parasites. Thus, for this Trial 4 purification, NEM was not added into control samples during the hydroxylamine-cleavage and biotinylation steps. Trypsin digestion in order to produce peptide fragments, streptavidin affinity purification of only biotinylated peptide fragments and elution of the peptides were all performed similar to Trial 3A (Section 3.1.4). The steps of this Trial 4 site-ID palmitome purification are shown in the schematic in Figure 3.5A.

The palmitome and control elutions were then analysed label-free by LC-MS/MS. MaxQuant was used to generate specific enrichment ratios (palmitome over control) from the median intensity of each peptide identification across the different replicates analysed.

In total, 5088 peptides containing enrichment ratios and a free cysteine were identified, and this corresponded to a total of 1801 proteins. This set of 5088 peptides was then grouped into enriched and highly enriched classes according to the same cut-off criteria used for the previous site-ID trial purifications (detailed in Section 3.1.2). Using these cut-off criteria, 1615 peptides (corresponding to

940 protein identifications) were defined as enriched, and out of these 1615 peptides, 377 peptides (corresponding to 295 protein identifications) could be further defined as highly enriched. A scatter plot showing the median peptide intensities against the MaxQuant generated enrichment ratios for all 5088 peptides identified in this Trial 4 site-ID palmitome purification is shown in Figure 3.5B. Of the 940 proteins which had peptide identifications classified as enriched, 208 proteins were also present in the total schizont palmitome, indicating that 42% of the total palmitome (208 out of 494 proteins in the total palmitome) was identified in this trial dataset (Figure 3.5C), which was higher than previously observed in the Trial 3B purification, suggesting that even more putative palmitoyl-proteins were being purified in this trial purification. As in the previous site-ID trial purifications, proteins known to be palmitoylated such as GAP45 and CRT, were identified and classified as enriched in this Trial 4 dataset. The presence of known palmitoyl-proteins, the larger enriched dataset obtained, and the increased number of proteins enriched in this trial that were also found in the total palmitome, indicated that this trial purification was the most improved compared to the previous trial purifications.



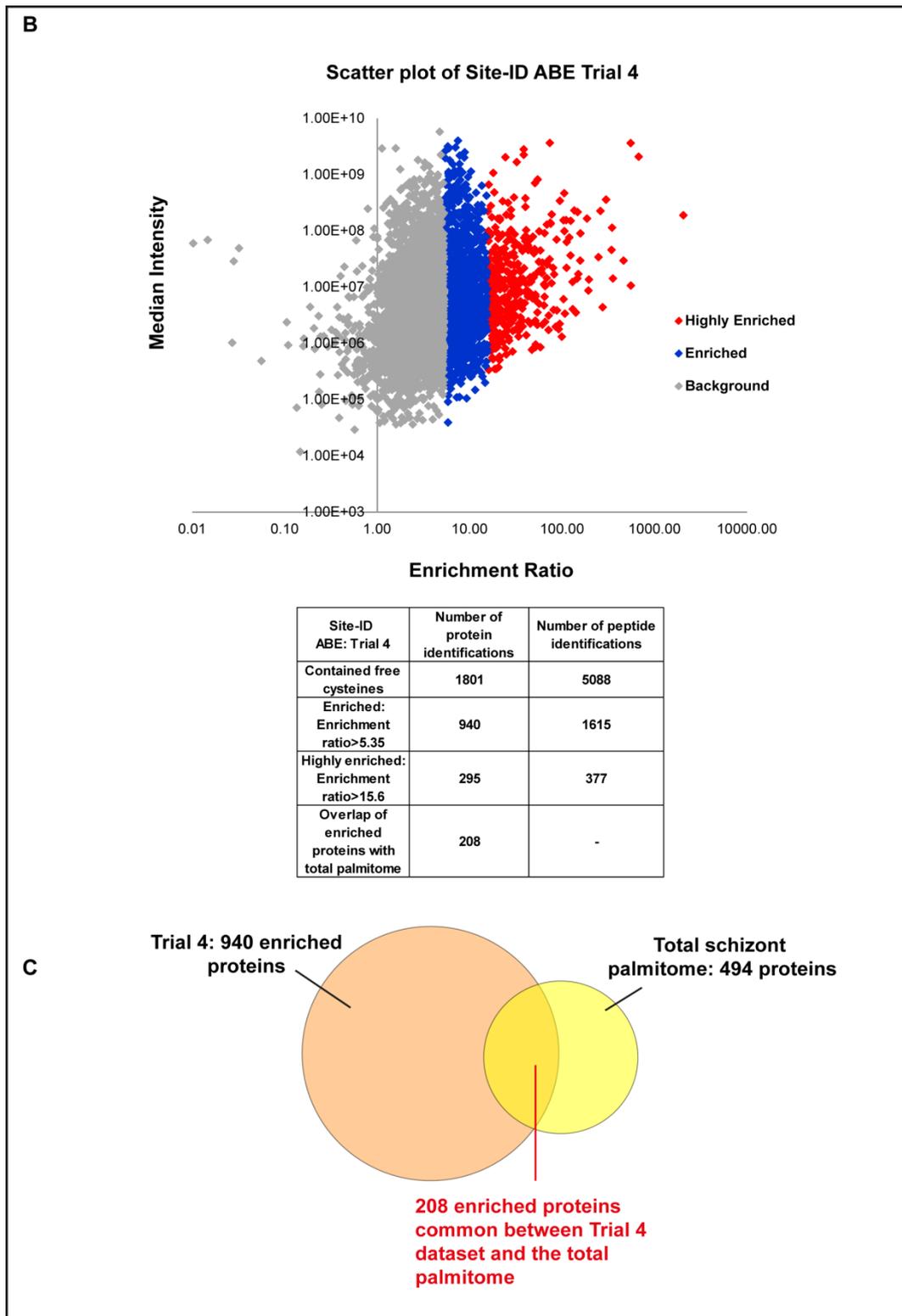


Figure 3.5: Trial 4 site-ID palmitome purification by ABE. (A) The Trial 4 site-ID purification was similar to that of Trial 3A and B. The proteome was extensively extracted from saponin-treated parasite pellets as in Trial 3A, and protein samples were (1) blocked with IAA, (2) treated with hydroxylamine (NH₂OH) (except for control samples) and (3) biotinylated using HPDP-biotin. In this Trial 4 purification however, NEM was not added into control samples during steps (2) and (3), as the addition of NEM into control samples only may skew the results obtained. (4) Trypsin digestion was then performed to digest the proteins into peptide fragments. Steps (2) to (4) were performed in 30 kDa MWCO spin columns as before. Biotinylated peptides were purified by (5)

streptavidin affinity purification, and unbound peptides removed by stringent washes. The biotinylated peptide fragments which remained bound to the beads were (6) eluted by reduction with TCEP and analysed label-free by quantitative mass spectrometry. **(B)** Scatter plot displaying the median protein intensities against the MaxQuant generated enrichment ratios for all the peptides purified by the Trial 4 site-ID purification. Peptides classified as enriched and highly enriched (according to the cut-off criteria described in the text) are highlighted in blue and red respectively. The numbers of protein and peptide identifications found are listed in the table below the plot. **(C)** Venn diagram representing the overlap between the enriched proteins in the Trial 4 dataset and the total schizont palmitome.

Although 42% of the total palmitome was enriched in the Trial 4 dataset, this was equivalent to only 22% of the Trial 4 enriched dataset (208 out of 940 proteins) (Figure 3.5C), indicating that approximately 80% of the enriched proteins in this purification were not found in the total palmitome. Additionally, this dataset was much larger than those of the previous purifications, and the classification of 940 proteins as enriched (and therefore potentially palmitoylated) may be inaccurate, considering the fact that just over 1000 proteins are upregulated in schizont stages [12].

These issues could be due in part to the nature of the label-free analysis, which is more prone to inaccurate quantification and run-to-run variability (as described in detail in Section 3.1.5). Another reason could be the high background presence of cysteine-containing peptides in the control samples, which was observed in the previous trial purifications and was also present in this Trial 4 dataset. This background, which could be due to incomplete blockage of free thiols during the initial treatment with IAA, or due to palmitate groups falling off during the other chemical steps of ABE, was not reduced in this Trial 4 purification, as NEM was not added to the control samples during the purification. The presence of this background in the control samples could thus affect the enrichment ratios and subsequently influence the cut-off criteria for enrichment, especially when the cut-off criteria were based on the median of the enrichment ratios. This indicates that more stringent analysis and more rigorous cut-off criteria may be needed in order to obtain useful information from this dataset.

Similar to what was observed in the Trial 3B site-ID purification (Section 3.1.5), a dataset of peptides that had peak intensities identified only in palmitome samples and not in control samples were also observed in this Trial 4 dataset. As for the Trial 3B purification, due to the lack of enrichment ratios for these peptide identifications, this dataset was not further analysed, although it must be noted that the presence of truly palmitoylated peptides within this dataset cannot be discounted.

In summary, five different trials of the site-ID palmitome purification were attempted in order to determine the optimum conditions required. All five trials purified peptides from a number of proteins which had been previously shown to be palmitoylated, indicating that the basic methodology was correct. All five trial purifications also produced enriched datasets which overlapped with the total schizont palmitome to different degrees. Analysis of the datasets from all trial purifications indicated that the use of a label-free approach improved the dataset of enriched proteins obtained. However, due to the inherent lower accuracy of label-free analysis, more stringent analysis of the datasets obtained is required. The best purification obtained out of the five trial purifications was considered to be the Trial 4 site-ID purification, as multiple biological and technical replicates were performed in this purification, resulting in an extensive dataset of peptide identifications and high overlap with the total schizont palmitome, although due to the large size of the dataset and the fact that the a label-free approach was used, more rigorous analysis of this dataset is still required. A summary of all five trial purifications and the main analyses performed on the five datasets is shown in Table 3.1.

Site-ID ABE purification	Method	Proteins Identified	Overlap with total palmitome
Trial 1	Full ABE protocol, pulled down with streptavidin-agarose, trypsin digestion on beads, eluted and dimethyl labelled	158 peptides (103 proteins) with ratios. 41 peptides (32 proteins) defined as enriched, 9 of these peptides (6 proteins) could be further defined as highly enriched.	26 enriched protein IDs also found in total palmitome (494 proteins) - 81% of the enriched dataset was equivalent to 5% of the total palmitome.
Trial 2	Blocked with IAA, ABE and trypsin digestion performed in spin columns, pulled down with streptavidin-agarose, eluted and dimethyl labelled.	431 peptides (291 proteins) with ratios. 139 peptides (117 proteins) defined as enriched, 28 of these peptides (24 proteins) could be further defined as highly enriched.	58 enriched protein IDs also found in total palmitome (494 proteins) - 50% of the enriched dataset was equivalent to 12% of the total palmitome.
Trial 3A	Same method as Trial 2, but initial block was with IAA, followed by NEM added only to control during ABE steps, dimethyl labelled	700 peptides (440 proteins) with ratios. 214 peptides (179 proteins) defined as enriched, 62 peptides (53 proteins) could be further defined as highly enriched.	69 enriched protein IDs also found in total palmitome (494 proteins) - 39% of the enriched dataset was equivalent to 13% of total palmitome.
Trial 3B	Same samples from 3A but not dimethyl labelled - label free analysis	2295 peptides (1173 proteins) with ratios. 713 peptides (521 proteins) defined as enriched, 165 peptides (138 proteins) could be further defined as highly enriched.	181 enriched protein IDs also found in total palmitome (494 proteins) - 35% of the enriched dataset was equivalent to 37% of total palmitome.
Trial 4	Same method as 3B: Initial block with IAA, but no NEM further added into control during ABE steps, not dimethyl labelled - label free analysis. 2 biological replicates and 2 technical replicates	5088 peptides (1801 proteins) with ratios. 1615 peptides (940 proteins) defined as enriched, 377 peptides (295 proteins) could be further defined as highly enriched.	208 enriched protein IDs also found in total palmitome (494 proteins) - 22% of the enriched dataset was equivalent to 42% of total palmitome.

Table 3.1: Summary of all five trial site-ID purifications and the analyses performed on the five datasets.

3.2. Analysis of the overlaps between the five trial site-ID palmitome purifications

As described in the sections above, label-free analysis of the site-ID palmitome purifications appeared to provide the most peptide identifications classified as enriched and the largest overlap with the total schizont palmitome. However, due to the relatively high background of cysteine-containing peptides found in the control samples, as well as the reduced accuracy of label-free quantification, more stringent analysis of the label-free datasets was required. One way to tackle this would be to investigate the overlaps between all of the trial purifications performed during the development of the site-ID purification method, in order to determine which peptides were consistently purified in all the trial purifications. If a particular peptide was consistently identified in the different trial purifications and was consistently classified as enriched, that peptide could be more confidently considered a putative palmitoylated peptide.

3.2.1. Overlap between all five trial site-ID palmitome purifications (Trial 1, 2, 3A, 3B and 4)

The first overlap analysed was the overlap between proteins identified in all five trial site-ID purification datasets. In this analysis, only proteins classified as enriched in all five datasets were considered. It is possible that proteins classified as enriched in one trial purification could be classified as not enriched in another purification, due to the differences in methodology and to whether the analysis was performed label-free or with dimethyl labelling (as described in Section 3.1 above). However, to reduce the complexity of this analysis, only the enriched protein datasets were considered at this time.

Only 6 enriched proteins were found to be common between all five trial site-ID purifications, and all of these proteins were also found in the total schizont palmitome. The small size of this overlap is not unexpected given that the Trial 1 dataset is an outlier with respect to size, with only 32 proteins classified as enriched (Section 3.1.2). Repeating the overlap analysis with proteins classified as enriched in all the later datasets (Trial 2, 3A, 3B and 4) revealed 33 enriched protein identifications in common. Of these 33 common enriched proteins, 22 proteins were also found to be present in the total schizont palmitome (Figure 3.6A).

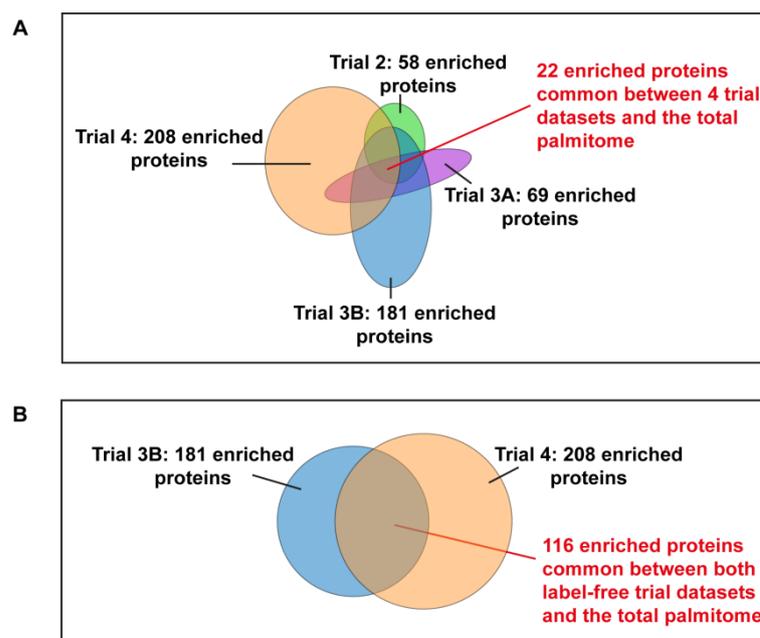


Figure 3.6: Overlaps between the different trial site-ID palmitome purifications and the total palmitome. Venn diagrams representing the overlaps between proteins classified as enriched in each of the different trial site-ID palmitome purifications. All the enriched proteins listed here for each trial dataset were also found to be present in the total schizont palmitome. The numbers of enriched proteins common between the trial site-ID datasets and the total schizont palmitome are indicated in red. (A) Overlap between 4 trial site-ID palmitome purifications (excluding Trial 1) and the total palmitome. (B) Overlap between the two label-free trial site-ID palmitome purifications and the total palmitome.

To assess whether the same peptides were consistently enriched for these 22 proteins, the peptides for all 22 enriched proteins common between the 4 site-ID trial purifications, as well as the total palmitome, were compared. For these proteins, 94 unique peptides passed the cut-off criteria for enrichment in at least one of the 4 trial datasets (for peptides which were identified multiple times, only one peptide was considered, and the peptide with the higher enrichment ratio was chosen), but only 14 unique enriched peptides were found to be present in all 4 trial datasets. As these peptides appeared to be consistently purified by all 4 trial site-ID purifications, and passed the cut-off criteria for enrichment used for each dataset, these peptides should be considered as high confidence palmitoylated peptides, and the cysteine residue present in each of these 14 common enriched peptides considered as high confidence sites of palmitoylation. For those peptides with more than one cysteine residue present, the cysteine which is palmitoylated is unknown and will have to be determined through experimental means. Table 3.2 shows all 14 enriched peptides common between the 4 datasets, along with their enrichment ratios in each dataset. These 14 peptides also include 2 peptides which were also classified as enriched in the Trial 1 dataset, and the enrichment

ratios of these 2 peptides in the Trial 1 purification is also shown in Table 3.2. It must be noted however that as mentioned above, ABE can result in the false enrichment of highly abundant proteins. Thus, the presence of peptides from highly abundant proteins, such as heat shock 70 kDa protein (PF08_0054) and glyceraldehyde-3-phosphate dehydrogenase (PF14_0598), within the dataset of 14 enriched common peptides may also be due to the high abundance of these proteins.

Although 14 common enriched peptides were found in this analysis, this degree of overlap might appear rather small. However, it should be noted that the different trial purifications had very different methodologies, as they were part of a methods development process. These relatively large differences in sample preparation, extraction and labelling will have a major impact on the subset of the proteome that is captured, making the identification of only 14 high confidence palmitoylation sites not completely surprising.

Protein	Protein name	Sequence	Enrichment Ratio				
			Trial 1	Trial 2	Trial 3A	Trial 3B	Trial 4
PF08_0035	conserved Plasmodium protein, unknown function	NYNTM V NDHN C TK	74.86	53.12	11.07	52.68	9.09
		NV N CNYEL N IK	51.87	155.19	33.88	32.27	6.06
PF08_0054	heat shock 70 kDa protein	LV N F C VED F KR	-	4.86	3.89	13.29	15.08
PF11_0351	heat shock protein hsp70 homologue	AKLEEL C H D LLK	-	6.33	12.00	614.39	10.41
PF13_0338	cysteine-rich surface protein	TIG F V C PVK	-	10.50	2.93	8.47	7.97
PF14_0578	conserved Plasmodium protein, unknown function	NS C ETLLESE Q NLLK	-	12.97	3.20	59.40	7.71
PF14_0598	glyceraldehyde-3-phosphate dehydrogenase	QLIVSNAS C TT N CLAP L AK	-	5.86	3.21	12.84	44.76
PFD1165w	Serine/Threonine protein kinase, FIKK family	LIT C SYNSDK	-	10.70	31.82	94.82	158.42
PFF0435w	ornithine aminotransferase	EQLDE C TEI V K	-	4.95	7.51	10.32	26.17
PFF0690c	organic anion transporter	Y V CF S GLDK	-	242.35	10.47	26.70	38.75
		IDEN I D G CGK	-	53.01	11.45	27.12	9.68
PFI0875w	Heat shock protein 70 (HSP70) homologue	LKDLEAV C QPI V K	-	6.28	9.47	23.68	38.68
PFI1445w	High molecular weight rhoptry protein-2	N CTLLTDF M K	-	6.52	5.20	53.43	6.45
PFL1300c	conserved Plasmodium protein, unknown function	NIETQDLN L N C PC W NR	-	5.63	4.68	198.62	43.24

Table 3.2: The 14 enriched peptides identified in all 4 trial site-ID palmitome purifications. These peptides were found to be present in either all 5 trial purifications, or in 4 of the trial purifications (excluding Trial 1). The enrichment ratios in each trial dataset for each of the peptides is shown here, with enrichment ratios above the cut-off criteria for highly enriched peptides shown in red, and enrichment ratios above the cut-off criteria for enriched peptides shown in blue. The putative palmitoylated cysteines are highlighted in bold and in green.

3.2.2. Overlap between the two label-free trial site-ID palmitome purifications (Trial 3B and 4)

A large part of the divergence between the trial datasets may be because two of the datasets were dimethyl labelled, while the remaining two were analysed using label-free approaches, which produces far larger datasets. Therefore, in the next analysis, only the overlap between the two label-free trial datasets (Trial 3B and 4) were analysed. As before, only proteins classified as enriched in each of the datasets were considered.

In this analysis, 324 enriched protein identifications were found to be common between the 2 label-free trial purifications considered. Of these 324 enriched proteins, 116 proteins were also found to be present in the total palmitome (Figure 3.6B). Gene ontology (GO) term analysis of the 116 common enriched proteins revealed an enrichment for proteins involved in localisation and transport (Figure 3.7A), which was previously observed for the total palmitome and for palmitomes in other systems [4, 16]. Thus, the degree of overlap between the 2 label-free trial datasets appeared to be better than the previous overlaps analysed, and may represent a more extensive dataset of putative palmitoylated proteins, as evidenced by GO term analysis of this overlap, which was similar to GO term analyses of other palmitomes.

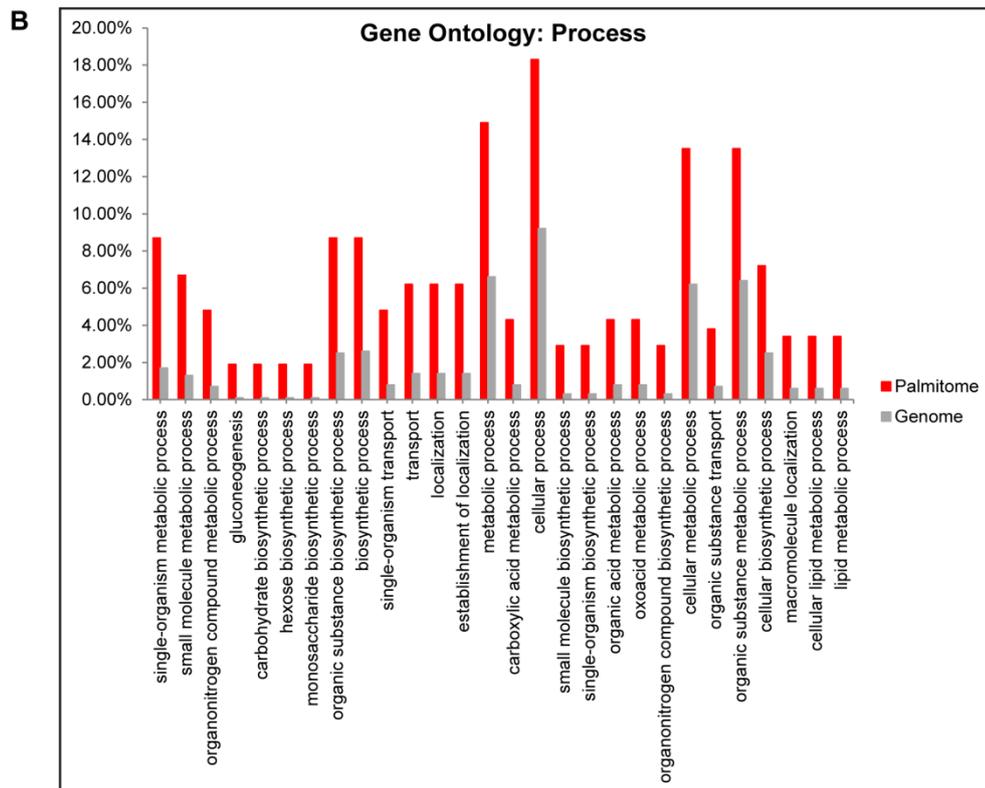
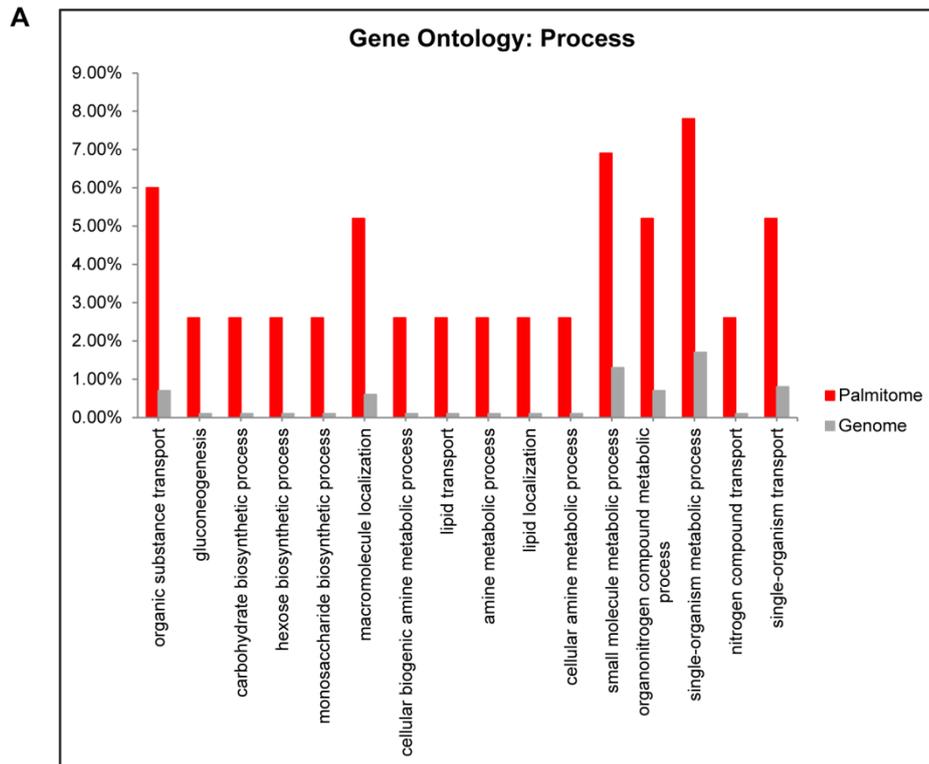


Figure 3.7: Gene ontology (GO) term analysis of enriched proteins in the overlapping trial site-ID purification datasets. GO analysis is of biological process annotations and is presented in comparison to corresponding percent genome values. p values are ≤ 0.05 for all displayed terms. **(A)** Enriched proteins present in both label-free datasets (Trial 3B and 4) and the total palmitome. **(B)** Enriched proteins present in both the Trial 4 dataset and the total palmitome. GO term analysis was performed using the following online tool: Generic Gene Ontology (GO)Term Finder (<http://go.princeton.edu/cgi-bin/GOTermFinder>).

The 116 enriched proteins which were common between the 2 label-free site-ID purifications, as well as the total palmitome, corresponded to a total of 323 unique peptides, which were classified as enriched in at least one of the 2 label-free trial datasets. Of these 323 unique peptides, a total of 142 unique enriched peptides were found to be present in both trial datasets. Although these peptides were present in both trial datasets considered here, and managed to pass the cut-off criteria for enrichment in both datasets, it must be noted that both the datasets analysed here were quantified through label-free methods, which is less accurate and more prone to error (as described in detail in Section 3.1 above), and were also affected by issues with the relatively high background of cysteine-containing peptides in control samples (Section 3.1.6). Thus, the peptides found to be common between both these datasets can be considered medium confidence palmitoylated peptides, and the cysteine residues present within these peptides regarded as medium confidence putative palmitoylation sites. Table 3.3 displays a selected set of enriched peptides common between the 2 trial datasets, with their corresponding protein identifications, which were also present in the total palmitome, along with the enrichment ratios of each peptide in each trial dataset.

Protein	Protein name	Peptide Sequence	Enrichment Ratios	
			Trial 3B	Trial 4
MAL13P1.117	DHHC-type zinc finger protein, putative	VYESNNIFF C K	329.46	186.76
MAL13P1.214	phosphoethanolamine N-methyltransferase	VLDIGSGLGGG C MYINEK	16.24	22.80
MAL7P1.27	chloroquine resistance transporter	IGGGS C IGK	24.62	7.13
MAL8P1.32	nucleoside transporter, putative	VGLLTGYG C STLYK	42.95	14.88
PF07_0033	Cg4 protein	LTAYYENTPDLPSN C IK	9.64	25.33
		EMLSYPPTFTN C QK	6.93	7.97
PF08_0035	conserved Plasmodium protein, unknown function	NYNTMVDHNIN C TK	52.68	9.09
		NVNCNYELINK	32.27	6.06
PF10_0039	membrane skeletal protein IMC1-related	ISSEEIKPAG CC SAACT	441.70	248.58
PF10_0220	phospholipid scramblase 1, putative	I CIPMNMK	157.02	54.60
		CC SYNFNLFDPSSNK	39.14	72.16
		SILSPMQT C K	14.53	9.75
PF11_0168	moving junction protein	E I CEPQNGLIDET L TK	31.13	9.65
PF13_0338	cysteine-rich surface protein	TIGFV C PVK	8.47	7.97
PF14_0102	rhoptry-associated protein 1, RAP1	FGT C IGSFGEHHLR	122.68	27.78
PFC0120w	Cytoadherence linked asexual protein 3.1	VLDMM C DHESVYSEK	326.77	30.83
		LFESLIQ C IEK	12.01	8.65
PFE0785c	metabolite/drug transporter, putative	FIG C GEEYMP L R	14.82	75.68
PFD1165w	Serine/Threonine protein kinase, FIKK family	LIT C SYNSDKDAK	94.82	158.42
		EICDNEN C VNLEDIK	8.94	7.41
		LCDFANSAPIYTYNNR	7.59	12.41
PFF0675c	myosin E	AYEN C FFENR	50.93	7.22
		IGNMNEIS C YK	91.40	6.02
PFI1730w	cytoadherence linked asexual protein 9(CLAG9)	KE C NIYESDR	6.74	81.45
PFL1090w	glideosome-associated protein 45	SVTP C DMNK	43.53	18.14

Table 3.3: A selected set of enriched proteins found to be common between both label-free trial site-ID datasets (Trial 3B and Trial 4), as well as the total palmitome, along with a selected set of enriched peptides common between both the trial datasets. The enrichment ratios in each trial dataset for each of the peptides is shown here, with enrichment ratios above the cut-off criteria for highly enriched proteins shown in red and enrichment ratios above the cut-off criteria for enriched proteins shown in blue. The putative palmitoylated cysteines are shown in bold and in green.

3.2.3. Overlap between Trial 4 site-ID palmitome purification and the total palmitome

As described in Section 3.1.6, the Trial 4 site-ID purification was considered the best purification, with multiple biological and technical replicates performed, a large dataset of enriched peptides obtained and the greatest coverage of the total palmitome. However, it was also noted that this dataset had to be treated with caution due to the label-free quantification used, as well as the high background of cysteine-containing peptides in the control samples, thus indicating that the results of this dataset cannot be accepted by itself. Overlap of the Trial 4 dataset with the Trial 3B dataset (described in Section 3.2.2 above) yielded some useful information, with a total of 142 medium confidence palmitoylation sites determined. Nevertheless, as the Trial 4 dataset was larger than the Trial 3B dataset, and overlapped more with the total palmitome, some palmitoylated peptides may have been missed simply due to the lower proteome coverage of the Trial 3B purification. Furthermore, in the Trial 3B purification, NEM was added into control samples in an attempt to reduce the background in the control, which as explained in Section 3.1.6, may instead inadvertently bias the results. Thus, an analysis based on just the Trial 4 dataset may provide further useful information on the sites of palmitoylation.

As an analysis of the peptides purified by the Trial 4 purification based solely on the cut-off criteria for enrichment could possibly have many false positive identifications due to the effect of the high background in control samples, only those peptides corresponding to proteins which were also found in the total palmitome were considered here.

As described in Section 3.1.6, 208 proteins classified as enriched in the Trial 4 dataset were found to be also present in the total schizont palmitome (Figure 3.5C). GO term analysis of the 208 common enriched proteins revealed an enrichment of proteins involved in localisation, transport and metabolic processes (Figure 3.7B). The GO term analysis of all proteins classified as enriched in just the Trial 4 purification resulted in the enrichment of proteins involved mainly in metabolic processes (data not shown). Although enrichment for proteins involved in metabolic processes was still observed in the GO term analysis of the 208 proteins common between the Trial 4 dataset and the total palmitome, the enrichment of proteins involved in localisation and transport seen here, which have been previously observed in other palmitomes, implied that this set of 208 common proteins may be a more accurate representation of palmitoylated proteins.

The 208 enriched proteins present in both the Trial 4 purification and the total palmitome corresponded to a total of 383 unique enriched peptides (which included the 142 unique enriched peptides already determined from the analysis of the overlap between the 2 label-free datasets and

the total palmitome described in Section 3.2.2 above). Although the dataset of 208 enriched proteins did appear to be a more representative list of palmitoyl-proteins, the peptide data was still only from one round of purification (albeit with two biological and two technical replicates). Thus, the 383 unique enriched peptides described here can also be regarded as medium confidence palmitoylated peptides, with the cysteine residues present in each peptide regarded as medium confidence palmitoylated sites. Table 3.4 displays a selected set of enriched proteins common between the Trial 4 dataset and the total palmitome, along with a selected set of the enriched peptides, and the enrichment ratios of each peptide in the Trial 4 dataset.

Protein	Protein Names	Peptide Sequence	Enrichment Ratio
MAL13P1.117	DHHC-type zinc finger protein, putative	VYESNNIFF C K	186.76
MAL13P1.310	calpain	KKEE C NIIENVEGNNVGNK	12.69
MAL7P1.229	Cytoadherence linked asexual protein	TVTT C YWFPSPIK	6.61
MAL7P1.27	chloroquine resistance transporter	IGGG S IGK	7.13
PF07_0033	Cg4 protein	ITAYENTPDIPS N CIK	25.33
		C EII C DIADSNIGGR	17.53
PF10_0039	membrane skeletal protein IMC1-related	ISSEEIKPAG C CSAACT	248.58
PF10_0220	phospholipid scramblase 1, putative	C CSYFNIFDPSNNK	72.16
		I CIPMNMK	54.60
		SEF C YR	44.16
		SIISPMQ T C K	9.75
PF11_0168	moving junction protein	E I CEPQNGIIDETITK	9.65
PF13_0338	cysteine-rich surface protein	C MISFDEGNNWK	27.03
		SIIII C SNIHDWK	9.86
		TIGF V CPVK	7.97
PF14_0102	rhoptry-associated protein 1, RAP1	FGT C IGSFGEHHR	27.78
PF14_0164	NADP-specific glutamate dehydrogenase	F CQAFMNEIYR	106.37
		QTAVVSGSGNVAI V C V QK	35.78
		HIGP C TDVPAGDIGVGGR	10.96
		SNNI I CPSK	7.68
		VAESYIEQ G C F	6.67
PF14_0455	multidrug resistance protein 2 (heavy metal transport family)	C NGEYTEMWNNMQSK	14.84
PFB0815w	Calcium-dependent protein kinase 1	QIISG I CYIHK	6.27
PFA0130c	Serine/Threonine protein kinase, FIKK family, putative	IS C DP I CIEK	349.65
PFC0120w	Cytoadherence linked asexual protein 3.1	VIDMM C DHESVYYSEK	30.83
		IFESIQ C IEK	8.65
PFC0185w	membrane skeletal protein IMC1-related	N C F IFNINR	19.88
PFD0895c	Bet3 transport protein, putative	SDIS F CEDFEETVNVIK	29.27
PFE0785c	metabolite/drug transporter, putative	FIG C GEEYMPIR	75.68
PF11730w	cytoadherence linked asexual protein 9(CLAG9)	E CNIYESDR	81.45
PFL1090w	glideosome-associated protein 45	SVTP C DMNK	18.14

Table 3.4: A selected set of enriched proteins found to be common between the Trial 4 site-ID dataset and the total palmitome, along with a selected set of corresponding enriched peptides. The enrichment ratios from the Trial 4 purification is shown here for each of the peptides, with enrichment ratios above the cut-off criteria for highly enriched proteins shown in red and enrichment ratios above the cut-off criteria for enriched proteins shown in blue. The putative palmitoylated cysteines are highlighted in bold and in green.

The sequences of all these 383 unique enriched peptides were also analysed by motif-x, an online software tool which looks for over-represented sequence motifs within sequence data (<http://motif-x.med.harvard.edu>) [17], in order to determine whether any conserved motifs existed around the palmitoylated cysteines. No significant consensus motif for palmitoylation was determined from this analysis. However, there did appear to be a significant prevalence for the presence of hydrophobic, branched amino acids, such as isoleucine and valine, one or two amino acids upstream or downstream of the putative palmitoylated cysteine. As one of the classes of proteins commonly palmitoylated are proteins that are palmitoylated close to TM domains, it is not surprising that the palmitoylated cysteine appears to be often close to hydrophobic amino acids. Indeed, analysis of the 208 proteins corresponding to the 383 unique enriched peptides revealed that 27% of the enriched proteins contained TM-domains (58 out of 208 proteins). The sequence motifs found to be overrepresented by the motif-x software is as shown in Figure 3.8.

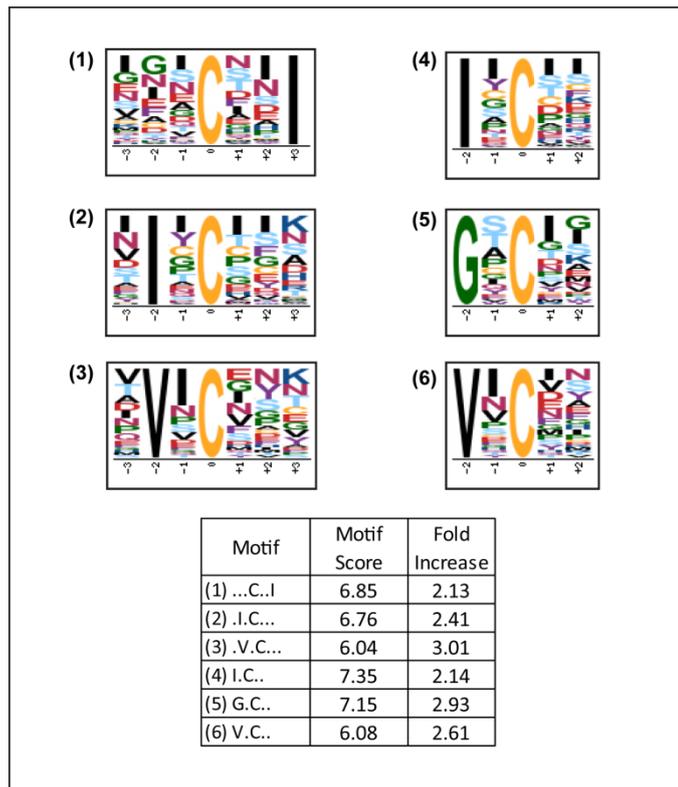


Figure 3.8: Over-represented sequence motifs within the Trial 4 site-ID palmitome purification dataset. The sequences of the 383 unique enriched peptides of the Trial 4 dataset, which corresponded to protein identifications also present within the total schizont palmitome, were analysed by motif-x software (<http://motif-x.med.harvard.edu>) and 6 over-represented sequence motifs were found. The motif score is calculated by the software and is the sum of the $-\log(\text{probability})$ of each fixed position in the motif. The motif score roughly correlates with statistical significance, where motifs with higher motif scores are more statistically significant (although all motifs found by the motif-x software is statistically significant) [17]. The fold increase measures the enrichment of the particular motif in the sequence dataset of interest compared to the background [17] (which in this case was sequences of the entire *P. falciparum* proteome obtained from *PlasmoDB*). The motif-x software was run using the following standard parameters: significance threshold - 0.000001, occurrences threshold - 20, width - 7 (for motifs 1 to 3) or 5 (for motifs 4 to 6).

In summary, the site-ID palmitome purification method developed here has successfully allowed the purification of palmitoylated proteins and can localise the putative palmitoylated cysteine. However, due to the reasons discussed in detail in the sections above, the whole enriched datasets from each of the different trial purifications might not be accepted as completely accurate when considered individually. Additionally, the excessive background of cysteine-containing peptides in control samples may disappointingly have a negative effect on the definition of enrichment in some of the datasets. Nevertheless, analysis of the overlaps between these different datasets still produced valuable information, providing sets of high confidence and medium confidence palmitoylation sites.

Although it cannot be stated that these sites are definitely palmitoylated, and the definition of the palmitoylation sites identified as high confidence or medium confidence is mostly arbitrary, the data presented here can still act as a guide to which cysteines may be possibly palmitoylated, and these can then be further validated experimentally. Indeed, the site-ID data from the Trial 4 purification was used to guide experiments in other parts of this project (described in Chapter 6). A summary of the analysis of the overlaps between the different trial purifications is shown in Table 3.5, and is represented visually by the Venn diagrams in Figure 3.6.

Site-ID Datasets	Overlaps between datasets	Overlap with total palmitome	Unique enriched peptides
All 5 datasets (Trial 1, 2, 3A, 3B and 4)	6 enriched protein IDs common	6 enriched protein IDs also found in total palmitome (494 proteins) - 100% of overlapping protein IDs equivalent to 1.2% of total palmitome	Total of 34 unique enriched peptides - only 2 peptides found in all 5 datasets with corresponding protein IDs found in the total palmitome.
4 datasets (Trial 2, 3A, 3B and 4)	33 enriched protein IDs common	22 enriched protein IDs also found in total palmitome (494 proteins) - 67% of overlapping protein IDs equivalent to 4.5% of total palmitome.	Total of 94 unique enriched peptides - only 14 peptides found in all 4 datasets with corresponding protein IDs also found in the total palmitome.
2 datasets (Trial 3B and 4) - label free datasets	324 enriched protein IDs common	116 enriched protein IDs also found in total palmitome (494 proteins) - 35% of overlapping protein IDs equivalent to 24% of total palmitome	Total of 323 unique enriched peptides - 142 peptides found in both datasets with corresponding protein IDs also found in the total palmitome.
1 label-free dataset (Trial 4)	-	208 enriched protein IDs also found in total palmitome (494 proteins).	Total of 383 unique enriched peptides with corresponding protein IDs also found in the total palmitome.

Table 3.5: Summary of the analyses performed on the overlaps between the different trial site-ID palmitome purifications.

3.3. Trial site-ID palmitome purification from *Plasmodium berghei* schizonts and *Toxoplasma gondii* tachyzoites

As mentioned in Section 3.1 above, the site-ID palmitome purification method was developed based on ABE, and without SILAC metabolic labelling, in order to allow the use of this site-localising method on other stages of the parasite life cycle, or on other Apicomplexan species, which are not amenable to metabolic labelling. As detailed in the sections above, the site-ID purification method did successfully purify palmitoyl-proteins and identified some putative palmitoylated cysteines in *P. falciparum* schizonts. In order to determine whether this site-ID purification could be used to determine sites of palmitoylation in other related Apicomplexan parasites, the final site-ID purification method (as performed in the *P. falciparum* Trial 4 site-ID purification) was applied to both the schizont stages of the rodent malaria parasite, *Plasmodium berghei*, and on the tachyzoite stages of the related Apicomplexan species, *Toxoplasma gondii*.

3.3.1. Trial site-ID palmitome purification from *P. berghei* schizonts

Approximately 10^9 *P. berghei* strain ANKA 2.33 schizont stage parasites was extracted from infected mouse blood by saponin-lysis as described in the Materials and Methods (infection of mice with *P. berghei* strain ANKA 2.33 and isolation of blood from infected mice was kindly performed by Katarzyna Modrynska, Billker and Rayner lab). Extraction of the proteome from saponin-treated parasite pellets and all other steps of the site-ID purification were performed exactly as for the *P. falciparum* Trial 4 site-ID palmitome purification (described in Section 3.1.6 and shown in the schematic in Figure 3.5A). Palmitome and control elutions were analysed using the label-free approach by LC-MS/MS as before. MaxQuant was used to generate specific enrichment ratios (palmitome over control) for each of the peptide identifications.

In total, 805 peptides, which had enrichment ratios and contained a free cysteine, were identified in this *P. berghei* trial purification, and this corresponded to a total of 498 proteins. This set of 805 peptides was then grouped into enriched and highly enriched classes following the same cut-off criteria used for the *P. falciparum* trial site-ID datasets (detailed in Section 3.1.2). Using these cut-off criteria, 238 peptides (corresponding to 193 protein identifications) were defined as enriched, and out of these 238 peptides, 53 peptides (corresponding to 38 protein identifications) could be further defined as highly enriched. A scatter plot showing the median peptide intensities against the MaxQuant generated enrichment ratios for all the 805 peptides identified in this trial purification is shown in Figure 3.9A.

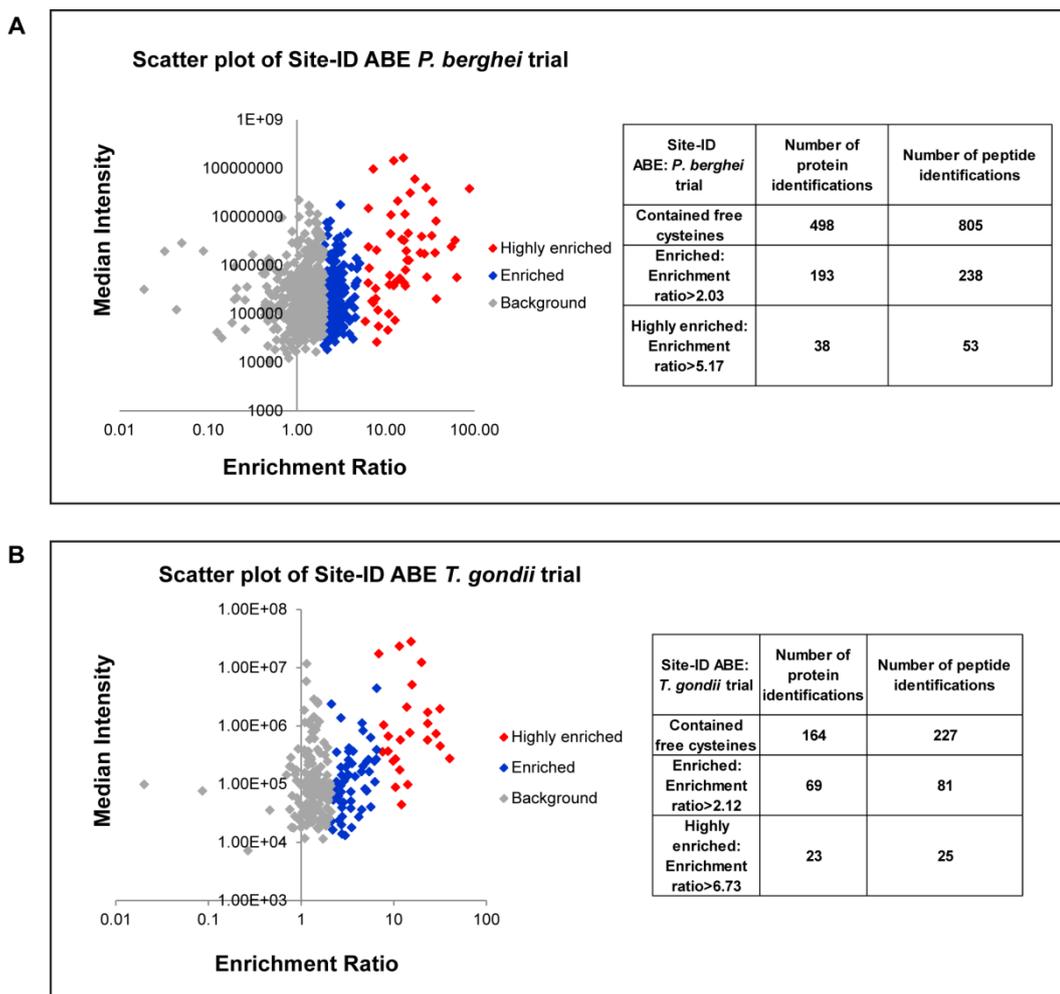


Figure 3.9: Trial site-ID palmitome purification by ABE for *Plasmodium berghei* and *Toxoplasma gondii*. The site-ID trial purification for *P. berghei* and *T. gondii* was performed identical to that of the *P. falciparum* Trial 4 site-ID palmitome purification, and purified peptides were analysed label-free by quantitative mass spectrometry. **(A)** Scatter plot displaying the median protein intensities against the MaxQuant generated enrichment ratios for all the peptides purified by the *P. berghei* site-ID purification. Peptides classified as enriched and highly enriched (according to the cut-off criteria described in the text) are highlighted in blue and red respectively. The numbers of protein and peptide identifications found are listed in the table on the right. **(B)** Scatter plot displaying median protein intensities against MaxQuant generated enrichment ratios for all the peptides purified by the *T. gondii* site-ID purification. Peptides classified as enriched and highly enriched (according to the cut-off criteria described in the text) are highlighted in blue and red respectively. The numbers of protein and peptide identifications found are listed in the table on the right.

The dataset from the *P. berghei* trial site-ID purification was smaller than previously observed for *P. falciparum*. In fact, the median protein intensities appeared to be in general lower than that of the previous *P. falciparum* purifications (Figure 3.9A compared to Figure 3.5B), suggesting that the overall amount of protein present was lower. This indicates that further optimisation of the amount of initial parasite material required for the purification must still be performed.

Despite the smaller size of the *P. berghei* trial dataset, known palmitoylated proteins such as GAP45 and MTIP, were found to be present and was classified as enriched in this dataset. The presence of known palmitoyl-proteins was encouraging, as this suggested that putative palmitoyl-proteins were indeed being purified in this trial purification. A summary of the analysis performed on the *P. berghei* trial dataset is shown in Table 3.6.

Site-ID ABE purification	Method	Proteins Identified
<i>P. berghei</i> trial	Similar to <i>P. falciparum</i> Trial 4 - label free	804 peptides (498 proteins) with ratios. 237 peptides (193 proteins) defined as enriched, 52 of these peptides (38 proteins) further defined as highly enriched.
<i>T. gondii</i> trial	Similar to <i>P. falciparum</i> Trial 4 - label free	227 peptides (164 proteins) with ratios. 81 peptides (69 proteins) defined as enriched, 25 of these peptides (23 proteins) further defined as highly enriched.

Table 3.6: Summary of the trial site-ID purifications for *P. berghei* and *T. gondii*, and the analyses performed on the two datasets.

Of the 193 enriched *P. berghei* proteins identified in this dataset, 177 enriched *P. berghei* proteins had homologues in *P. falciparum* (the list of *P. berghei* and corresponding *P. falciparum* homologues was kindly provided by Lia Chappell, Berriman and Rayner lab). This list of 177 *P. falciparum* homologues of the *P. berghei* enriched proteins was compared with the *P. falciparum* Trial 4 site-ID enriched dataset, as this purification was the most similar to the *P. berghei* trial purification in terms of methodology and was also analysed using label-free quantification. Of the 177 homologues, 100 homologues were found to be present in the *P. falciparum* Trial 4 dataset, indicating that 56% of the *P. berghei* trial dataset overlapped with the *P. falciparum* Trial 4 dataset.

A selected set of 8 *P. berghei* proteins, and their peptides classified as enriched in this trial purification is shown in Table 3.7, along with the corresponding *P. falciparum* homologues, and their peptides which were classified as enriched in the *P. falciparum* Trial 4 purification. In this set of homologous pairs, the same homologous peptide was purified from both site-ID purifications for four of the homologous pairs. For the other four homologous pairs however, different peptides appeared to be enriched in the *P. berghei* dataset compared to their homologues in the *P. falciparum* dataset. However, as the *P. berghei* dataset was from only a single purification, it is unknown at this time whether different peptides are palmitoylated in some of the homologues of the different species, or the enrichment of different peptides was due to false positive identification or the inaccurate quantitation of the label-free approach. Further replicates are still required before it can be determined whether the same or different peptides are palmitoylated in the homologous pairs of these two *Plasmodium* species. Nevertheless, the purification of the same peptides for some of the enriched proteins, the good overlap with the *P. falciparum* Trial 4 dataset, and the presence of known palmitoyl-proteins in this *P. berghei* dataset indicates that this site-ID purification can be used for site-ID palmitome purification in *P. berghei*, although further optimisation is still required.

<i>P. berghei</i>				<i>P. falciparum</i> homologue			
Protein	Protein Name	Sequence	Enrichment Ratio	Protein	Protein Name	Sequence	Enrichment Ratio
PBANKA_141970	zinc finger protein, putative	VYESNNIF C QGK	33.13	MAL13P1.117	DHHC-type zinc finger protein, putative	VYESNNIFF C K	186.76
PBANKA_111530	glideosome-associated protein 40, putative (GAP40)	FIG C GEEYMPIR	25.53	PFE0785c	metabolite/drug transporter, putative	FIG C GEEYMPIR	75.68
PBANKA_124060	membrane skeletal protein, putative	I ACC SGDNVFD	18.95	PFE1285w	membrane skeletal protein IMC1-related	NI YCS YA	9.01
		EN NYC TTV SCK	6.42			ESIMEVPEV NC PHIDISK	8.14
PBANKA_011220	myosin-like protein, putative	GYEN C YASFAK	88.27	PFF0675c	myosin E	Q NFIC EFPSNYK	10.77
		F CTI HIENE EIK	2.23			AY ENC FFENR	7.22
						IG NMNEIS CYK	6.02
PBANKA_143760	glideosome-associated protein 45, putative	STTP C DIDKMD ETA K	12.32	PFL1090w	glideosome-associated protein 45	SVTP C DM NK	18.14
PBANKA_130290	copper transporter, putative	SG NDEC KPIIDIN HIG SE GK	4.77	PF14_0369	copper transporter putative	IS FM Y NC W K	10.52
		C DVDPHKE Y GD PACCGC	2.29			C SPDDID VHK	8.60
PBANKA_132640	glyceraldehyde-3-phosphate dehydrogenase, putative	F PC EVTPTEGGIM VGNK	2.34	PF14_0598	glyceraldehyde-3-phosphate dehydrogenase	C QVD WC ESTGV FI T K	56.39
						Q IIV SN ASC TT NC I API A K	44.76
						C AIS NI PAST GAAK	30.32
PBANKA_070770	conserved Plasmodium protein, unknown function	N FISGM C GG A K	33.93	PF08_0035	conserved Plasmodium protein, unknown function	NY NTM MND H NI N C T K	9.09
		H EDA IES N VT C FS K	21.37			NV NC Y E I IN K	6.06
		YD CT N PT NY I Q NY NC M K	17.50				
		CC N IP N F C R	17.08				

Table 3.7: A selected set of enriched proteins purified in the *P. berghei* trial site-ID palmitome purification, along with their *P. falciparum* homologues, present in the *P. falciparum* Trial 4 site-ID palmitome purification. The enriched peptides are shown for each enriched protein in both *P. berghei* and *P. falciparum* datasets, along with the enrichment ratios of each peptide. Enrichment ratios above the cut-off criteria for highly enriched proteins are shown in red and enrichment ratios above the cut-off criteria for enriched proteins are shown in blue. The putative palmitoylated cysteines are highlighted in bold and in green.

3.3.2. Trial site-ID palmitome purification from *T. gondii* tachyzoites

Next, the site-ID palmitome purification method was attempted on the tachyzoite stages of *T. gondii* parasites (a stage in the *T. gondii* life cycle which is similar to the schizont stages of *Plasmodium*). The proteome was extracted from approximately 10^8 *T. gondii* RH strain tachyzoite stage parasites (kindly provided by Karine Frenal, Soldati-Favre lab) using the same extensive extraction methods as before (described in the Materials and Methods and in Section 3.1.3). All other steps of the site-ID purification were performed exactly as for the *P. falciparum* Trial 4 site-ID palmitome purification (described in Section 3.1.6 and shown in the schematic in Figure 3.5A). Palmitome and control elutions were analysed label-free by LC-MS/MS as before, and MaxQuant was used to generate specific enrichment ratios (palmitome over control) for each of the peptide identifications.

In total, 227 peptides with enrichment ratios and containing a free cysteine were identified in this *T. gondii* trial purification, and this corresponded to a total of 164 proteins. The 227 peptides were grouped into enriched and highly enriched classes according to the same cut-off criteria used for the *P. falciparum* trial purifications (detailed in Section 3.1.2). Using these cut-off criteria, 81 peptides (corresponding to 69 protein identifications) were defined as enriched, and out of these 81 peptides, 25 peptides (corresponding to 23 protein identifications) could be further defined as highly enriched. A scatter plot showing the median protein intensities against the enrichment ratios for all the 227 peptides identified in this trial purification is shown in Figure 3.9B.

The *T. gondii* trial dataset was smaller than the previous datasets, and the average peptide intensity was again lower than previous purifications. In fact, the average peptide intensity was even lower than that of the *P. berghei* trial purification, indicating that an overall low amount of protein was present and further optimisation of the starting parasite material used in the purification is still required. Despite the small size of this dataset, proteins known to be palmitoylated in other systems, such as GAP45 and Bet3 transporter, were identified in this dataset. A summary of the analysis performed in the *T. gondii* trial dataset is shown in Table 3.6.

Additionally, the amino acid sequences of a selected set of seven *T. gondii* proteins, classified as enriched in this dataset, was used in a protein-protein BLAST search in order to determine potential *P. falciparum* homologues for these proteins. These *T. gondii* proteins, and their peptides enriched in this trial purification, are shown along with their corresponding potential *P. falciparum* homologues, and their peptides enriched in the *P. falciparum* Trial 4 purification, in Table 3.8. The same homologous peptide appeared to be enriched in both datasets for one of the homologous pairs. Conversely, different peptides were enriched in each dataset for two of the homologous pairs. For

the remaining 5 homologous pairs, both the same homologous peptides, as well as additional different peptides, were enriched in the two datasets. However, as mentioned in Section 3.3.1 above, further replicates are still required before any conclusions can be made.

<i>T. gondii</i>				<i>P. falciparum</i> homologue			
Protein	Protein Name	Sequence	Enrichment Ratio	Protein	Protein Name	Sequence	Enrichment Ratio
TGGT1_006420	transport protein particle component Bet3, putative	TGMGG CDC FR	5.65	PFD0895c	Bet3 transport protein, putative	SDISF C EDFEETVNVIAK VNCYFVK	29.27 10.26
TGGT1_033520	glyceraldehyde-3-phosphate dehydrogenase, putative	SSDVVSNAS CTTNC IAPIAK	3.30	PF14_0598	glyceraldehyde-3-phosphate dehydrogenase	CQVDV C ESTGVFITK	56.39
		VPVPDVSWDIT C K	2.70			QIVSNAS CTTNC IAPIAK	44.76
						VPIGTVSWD I VC	6.44
TGGT1_078320	conserved hypothetical protein	CGCDIGDQHDENE C PICR SVVGYTVTP C DMAAIDETAK	15.39 11.52	PFL1090w	glideosome-associated protein 45	SVTP C DMNK	18.14
TGGT1_101540	conserved hypothetical protein	FIG C GEEYMPMR	28.75	PFE0785c	metabolite/drug transporter, putative	FIG C GEEYMPIR	75.68
TGGT1_107040	enolase, putative	IEESIGSD C QYAGAGFR	3.55	PF10_0155	enolase	SDQMVPV P CINVINGGSHAGNK	94.67
						NACNA I IK	16.24
						IIGM N CTEQK	10.41
TGGT1_112260	cell division protein, putative	AVASE C SANFVSIK IDEVGYDDIG G CR	2.23 2.20	PFF0940c	cell division cycle protein 48 homologue, putative	KCPVAENVPIDF I AQK	111.19
						HST C I ^{II} INDNDIDEGK	35.85
						TAGFSGAD I A E ICQR	16.98
TGGT1_112840	heat shock protein, putative	MDF C VQDFK	2.93	PF08_0054	heat shock 70 kDa protein	IVN F CVEDFKR	15.08
						NSIEN C YGVK	9.84
						EKIQA E I T CMK	7.28

Table 3.8: A selected set of enriched proteins purified in the *T. gondii* trial site-ID palmitome purification, along with their *P. falciparum* homologues (homologues were determined by protein-protein BLAST search) found to be present in the *P. falciparum* Trial 4 site-ID palmitome purification. The enriched peptides are shown for each enriched protein in both *T. gondii* and *P. falciparum* datasets, along with the enrichment ratios of each peptide. Enrichment ratios above the cut-off criteria for highly enriched proteins are shown in red and enrichment ratios above the cut-off criteria for enriched proteins are shown in blue. The putative palmitoylated cysteines are highlighted in bold and in green.

Although further optimisation is still required for the *T. gondii* site-ID palmitome purification, the presence in this dataset of known palmitoylated proteins, as well as the presence of *P. falciparum* homologues enriched in previous datasets, indicate that this purification method can be used to determine sites of palmitoylation in *T. gondii* parasites.

For both the *P. berghei* and *T. gondii* trial purifications, further optimisation and certainly more replicates are still required before any truly relevant site-ID data can be obtained. Unfortunately, this was unable to be achieved before the writing of this dissertation, due to time constraints and due to the availability of mass spectrometry facilities. Nevertheless, these two trial purifications still provide valuable preliminary information on the sites of palmitoylation in *P. berghei* schizonts and *T. gondii* tachyzoites.

Conclusion

In this chapter, a method for the identification of palmitoylation sites within the palmitome of an organism was developed based on the ABE method of palmitoyl-protein purification. This site-ID palmitome purification method was successful in purifying palmitoylated peptides (containing the putative palmitoylated cysteine) in *P. falciparum*, *P. berghei* and *T. gondii* parasites. However, the continued presence of a relatively high background of cysteine-containing peptides in control samples, and the use of the less accurate label-free quantification method, have made analysis of the data more difficult, and require the treatment of datasets to be more stringent and rigorous. Although this method still requires some optimisation and refinement (for example, the reduction of the background of cysteine-containing peptides in control samples would be beneficial), some high and medium confidence palmitoylation sites have been identified in *P. falciparum* schizonts, and these data can now be validated experimentally and used to guide future experiments, such as those carried out in Chapter 6.

References:

1. Drisdell, R.C., et al., *Assays of protein palmitoylation*. Methods, 2006. **40**(2): p. 127-134.
2. Roth, A.F., et al., *Global analysis of protein palmitoylation in yeast*. Cell, 2006. **125**(5): p. 1003-1013.
3. Martin, B.R. and B.F. Cravatt, *Large-scale profiling of protein palmitoylation in mammalian cells*. Nat Methods, 2009. **6**(2): p. 135-138.
4. Jones, Matthew L., et al., *Analysis of Protein Palmitoylation Reveals a Pervasive Role in Plasmodium Development and Pathogenesis*. Cell Host & Microbe, 2012. **12**(2): p. 246-258.
5. Nirmalan, N., P.F. Sims, and J.E. Hyde, *Quantitative proteomics of the human malaria parasite Plasmodium falciparum and its application to studies of development and inhibition*. Mol Microbiol, 2004. **52**(4): p. 1187-1199.
6. Prieto, J.H., et al., *Large-scale differential proteome analysis in Plasmodium falciparum under drug treatment*. PLoS One, 2008. **3**(12): p. e4098.
7. Kovanich, D., et al., *Applications of stable isotope dimethyl labeling in quantitative proteomics*. Analytical and Bioanalytical Chemistry, 2012. **404**(4): p. 991-1009.
8. Zhu, W., J.W. Smith, and C.-M. Huang, *Mass Spectrometry-Based Label-Free Quantitative Proteomics*. Journal of Biomedicine and Biotechnology, 2010. **2010**: p. 1-6.
9. Boersema, P.J., et al., *Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics*. Nat Protoc, 2009. **4**(4): p. 484-494.
10. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification*. Nat Biotechnol, 2008. **26**(12): p. 1367-72.
11. Gardner, M.J., et al., *Genome sequence of the human malaria parasite Plasmodium falciparum*. Nature, 2002. **419**(6906): p. 498-511.

12. López-Barragán, M.J., et al., *Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum*. BMC Genomics, 2011. **12**(1): p. 587.
13. Bantscheff, M., et al., *Quantitative mass spectrometry in proteomics: a critical review*. Analytical and Bioanalytical Chemistry, 2007. **389**(4): p. 1017-1031.
14. Li, Z., et al., *Systematic Comparison of Label-Free, Metabolic Labeling, and Isobaric Chemical Labeling for Quantitative Proteomics on LTQ Orbitrap Velos*. Journal of Proteome Research, 2012. **11**(3): p. 1582-1590.
15. Patel, V.J., et al., *A comparison of labeling and label-free mass spectrometry-based proteomics approaches*. J Proteome Res, 2009. **8**(7): p. 3752-9.
16. Wilson, J.P., et al., *Proteomic Analysis of Fatty-acylated Proteins in Mammalian Cells with Chemical Reporters Reveals S-Acylation of Histone H3 Variants*. Molecular & Cellular Proteomics, 2010. **10**(3): p. M110.001198-M110.001198.
17. Schwartz, D. and S.P. Gygi, *An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets*. Nature Biotechnology, 2005. **23**(11): p. 1391-1398.

Chapter 4

Protein acyltransferases in *Plasmodium berghei*

4.1. Repertoire of DHHC-domain-containing proteins in *Plasmodium*

In order to achieve the aim of characterising the putative PATs in *Plasmodium*, it was first necessary to identify the repertoire of DHHC-domain-containing proteins in both *Plasmodium* species of interest here, *P. falciparum* and *P. berghei*. Search by protein-protein BLAST (blastp), using the 50 amino acid characteristic conserved DHHC cysteine-rich domain of the yeast DHHC protein, Erf2, revealed 12 proteins containing the conserved DHHC-domain in *P. falciparum*, all except one possessing a homologue in *P. berghei*, which has 11 proteins with the DHHC-domain. All of these putative DHHC proteins in both *P. falciparum* and *P. berghei* adhered to the generic structure of DHHC-PATs (shown in Figure 1.2), containing at least four TM-domains and possessing the DHHC-cysteine rich domain within an inter-TM-domain loop. For both *Plasmodium* species, each had one protein which contained a DHYC rather than the characteristic DHHC motif. This is similar to one of the yeast DHHC proteins, Akr1, which also contains the DHYC motif but is known to function as a DHHC-PAT [1]. A multiple sequence alignment of the amino acid sequences of the DHHC-cysteine rich regions of all these proteins across *P. falciparum* and *P. berghei* confirmed that this signature motif was indeed highly conserved in all the DHHC-containing proteins in both *Plasmodium* species, as shown in Figure 4.1A. The two other motifs usually conserved in DHHC proteins, the DPG and TTxE motifs, are also present in most of the *Plasmodium* DHHC proteins, as indicated in Table 4.1.

Phylogenetic analysis using neighbour-joining (NJ) [2] and maximum likelihood (ML) [3] methods on the conserved DHHC domains of *P. falciparum* and *P. berghei*, and the related Apicomplexan parasites, *Toxoplasma gondii* and *Neospora caninum*, revealed that an evolutionary relationship does exist between the DHHC proteins of these Apicomplexan parasites. For *P. falciparum* and *P. berghei*, all the DHHC proteins appeared to be grouped in orthologous pairs, except for the one *P. falciparum* DHHC protein (PFB0140w) which did not have a homologue in *P. berghei*. Although homologues of PFB0140w were present in other human *Plasmodium* species, there were no homologues of this protein in all rodent *Plasmodium* species, and PFB0140w did not group with the other Apicomplexan DHHC-domain-containing proteins, suggesting that this protein may have been gained in human and closely related simian *Plasmodium* species. Five of the DHHC proteins in *P. falciparum* and *P. berghei* appeared to have direct one-to-one orthologues in *T. gondii*. The rest of the *Plasmodium* DHHC proteins had related DHHC proteins in *T. gondii*, but clear one-to-one orthologues could not be identified as *T. gondii* has an expanded DHHC protein family of 18 proteins. A phylogenetic tree was generated from this data using both the NJ distance analysis and the ML analysis, and is shown in Figure 4.1B, where only nodes with a bootstrap value more than 80 are indicated and values of more than 95 are considered significant. The phylogenetic analysis described here was performed by Arnault Graindorge, Soldati-Favre lab as part of a collaborative project on *P. berghei* and *T. gondii*

DHHCs which was recently published in Traffic [4]. Based on these homology relationships, the *Plasmodium* DHHC proteins have now been named according to their closest *T. gondii* homologue, while preserving the orthology between *P. falciparum* and *P. berghei* pairs. The naming of the *Plasmodium* DHHC proteins along with their accession numbers is as shown in Table 4.1.

<i>Plasmodium berghei</i>			<i>Plasmodium falciparum</i>		
Name	Plasmodb accession	C-terminal motif	Name	Plasmodb accession	C-terminal motif
PbDHHC1	PBANKA_040200	NPG; TFxE	PfDHHC1	PFC0160w	SPG; TFxE
PbDHHC2	PBANKA_010830	DPG; TTxE	PfDHHC2	PFF0485c	DPG; TTxE
PbDHHC3	PBANKA_092730	DPL; TTxE	PfDHHC3	PF11_0217	DPL; TTxE
PbDHHC4	PBANKA_142090	DPG; TTxE	PfDHHC4	MAL7P1.68	DPG; TTxE
PbDHHC5	PBANKA_133780	NPG; TLxE	PfDHHC5	MAL13P1.126	NPG; TSxE
PbDHHC6	PBANKA_083330	NPG; TTxE	PfDHHC6	PFI1580c	NPG; -
PbDHHC7	PBANKA_124300	PPG; TTxE	PfDHHC7	PFE1415w	SPG; TTxE
PbDHHC8	PBANKA_141970	DPG; TTxE	PfDHHC8	MAL13P1.117	DPG; TTxE
PbDHHC9	PBANKA_093210	NPG; TTxE	PfDHHC9	PF11_0167	NPG; -
PbDHHC10	PBANKA_051200	SPG; TTxE	PfDHHC10	PF10_0273	NPG; TTxE
PbDHHC11	PBANKA_031260	NPG; -	PfDHHC11	PFB0725c	NPG; TTxxE
-	-	-	PfDHHC12	PFB0140w	DPG; TxxE

Table 4.1: DHHC-domain-containing proteins in *Plasmodium berghei* and *Plasmodium falciparum*. The nomenclature of the DHHC proteins in both parasite species are shown along with their *PlasmoDB* accession numbers. The two additional conserved motifs (besides the characteristic DHHC domain) are also indicated.

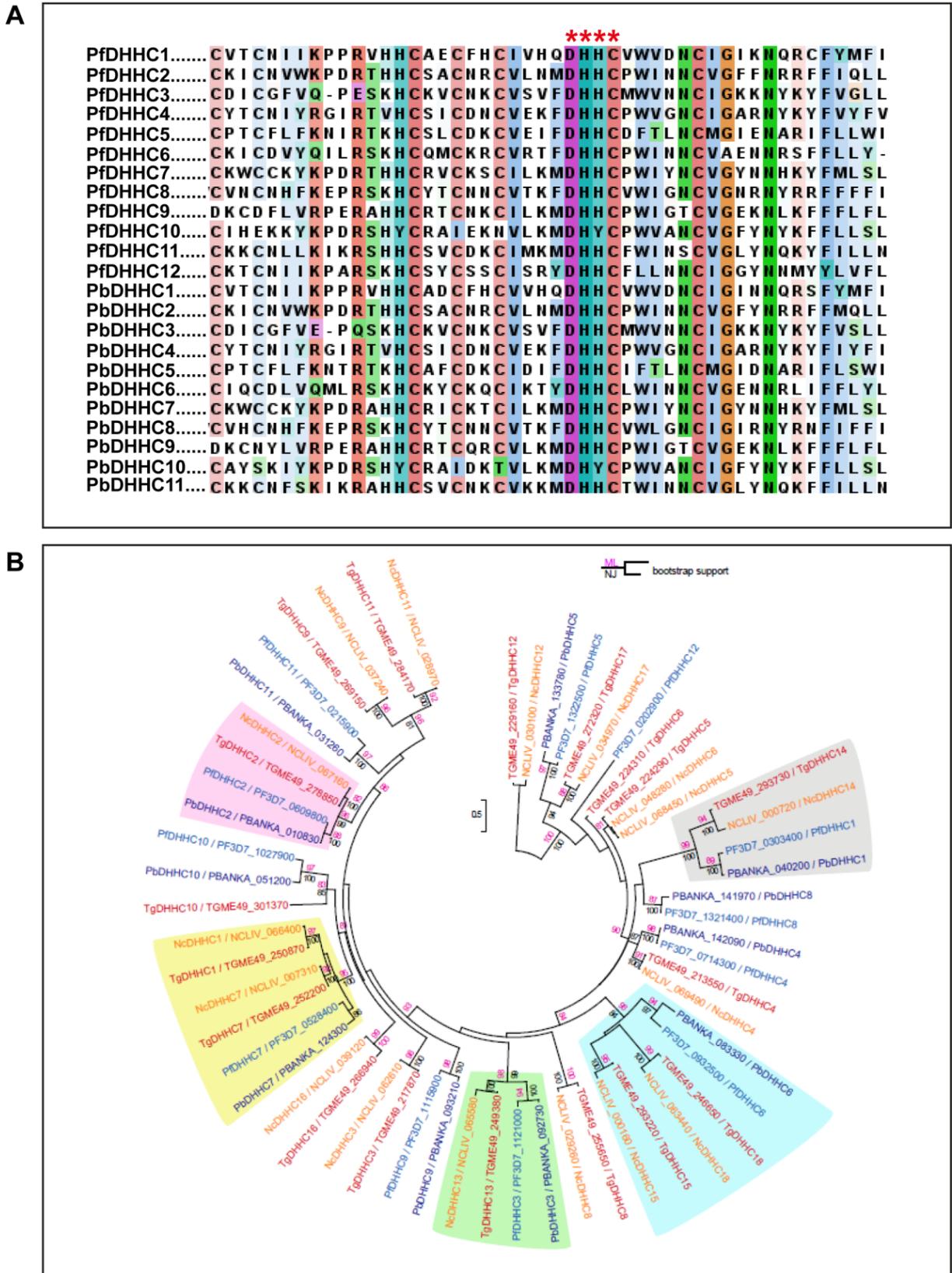


Figure 4.1: Repertoire of DHC-containing proteins in *Plasmodium*. (A) Multiple sequence alignment of DHC-domain containing proteins in *P. falciparum* and *P. berghei*. The amino acid sequences of the DHC-cysteine rich domains of all the DHC proteins in *P. falciparum* and *P. berghei* are shown aligned using ClustalW2 and Jalview. The characteristic DHC motif is highlighted with asterisks. (B) Phylogenetic tree of the DHC protein family in *P. falciparum*, *P. berghei*, *T. gondii* and *N. caninum* based on NJ distance analysis and ML analysis [4].

Advances in *P. berghei* genetic modification technology, developed by the Billker and Rayner labs, now allows quicker, large-scale and much more efficient means of genetic manipulation in *P. berghei* [5]. Thus, the study of the *Plasmodium* DHHC proteins was chosen to be performed first in *P. berghei*, where analysis of the full repertoire of DHHC proteins was feasible, before following up specific DHHC proteins of interest in *P. falciparum*. In order to analyse all the *Plasmodium* DHHC proteins in *P. berghei*, vectors from the *PlasmoGEM* resource (<http://plasmogem.sanger.acuk>) [5] were used to introduce a C-terminal triple-HA (3-HA) epitope tag to as many of the PbDHHC proteins as possible, as well as to perform knock-out studies on these same PbDHHC genes.

4.2. Expression and localisation of DHHC proteins in *Plasmodium berghei*

4.2.1. Generation of PbDHHC triple-HA (3-HA)-tagged transgenic parasite lines

As described above, there are 11 DHHC-containing proteins in *P. berghei*, now referred to as PbDHHC1 to 11. Tagging vectors from the *PlasmoGEM* resource [5] were available for 9 of the PbDHHC genes (PbDHHC3-11), and these vectors were introduced into *P. berghei* strain ANKA 2.34 as described in the Materials and Methods. Integration of the tagging constructs was expected to occur by double-crossover recombination of the homologous regions, resulting in the removal of the STOP codon and the insertion of the 3-HA tag, as well as the insertion of the *hDHFR/yfcu* selection cassette, which confers resistance to the drug pyrimethamine [6] (Figure 4.2A). The successful integration of each of the nine PbDHHC tagging constructs into the expected chromosomes was demonstrated by Southern blotting of chromosomes separated by pulsed field gel electrophoresis (PFGE) (Figure 4.2B), indicating the generation of 3-HA-tagged transgenic lines for 9 of the PbDHHCs (PbDHHC3 to 11).

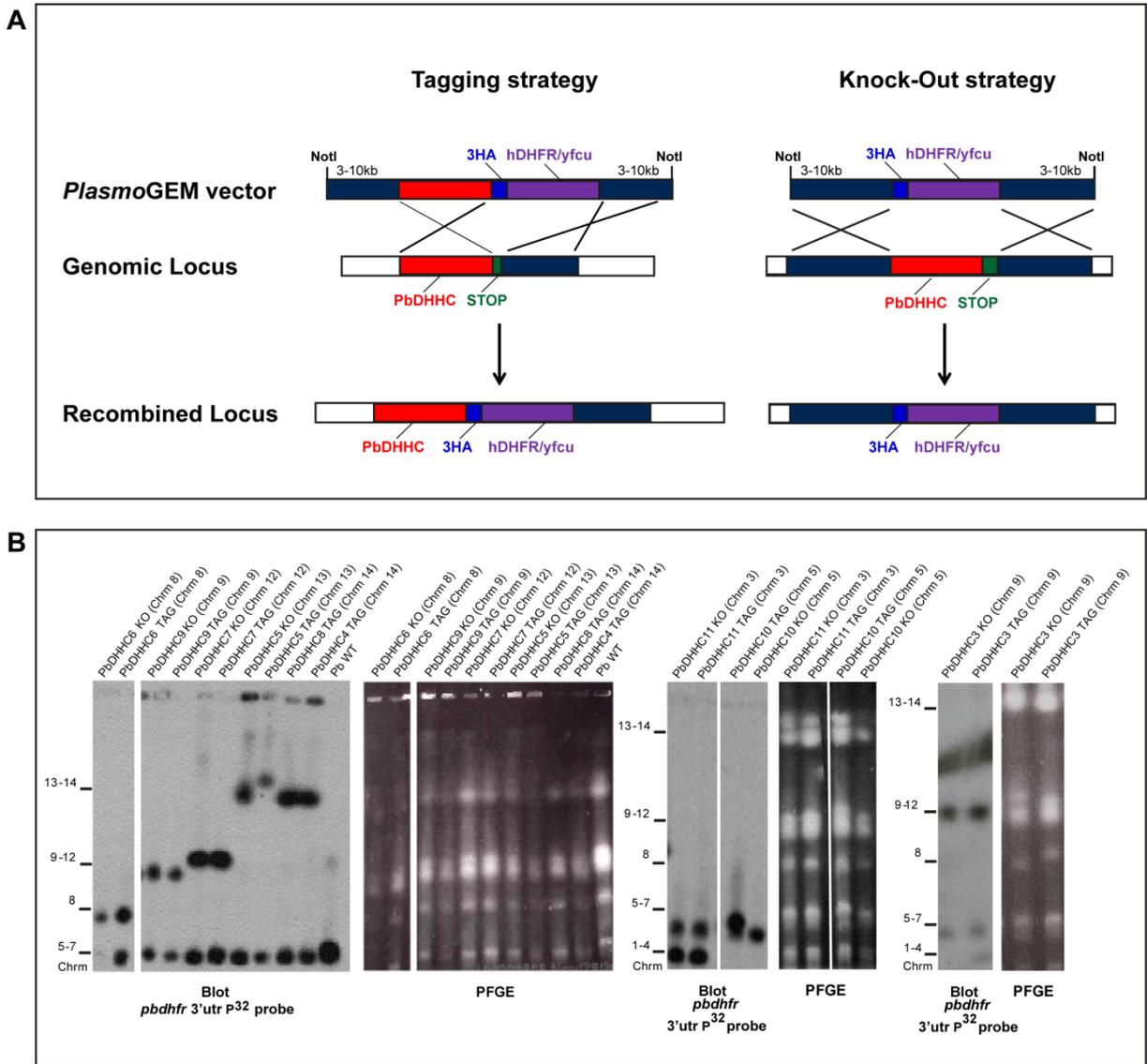


Figure 4.2: Generation of PbDHHC triple-HA-tagged and knock-out transgenic lines. (A) Scheme of the strategy used to C-terminally-tag and knock-out the endogenous locus of the PbDHHC genes of interest. **(B)** Pulsed field gel electrophoresis (PFGE) and southern blot analysis of size-separated *P. berghei* chromosomes using a probe specific to the PbDHFR 3' UTR, indicating the integration of the tagging and knock-out vectors into the expected chromosomes.

4.2.2. Five DHHC proteins are expressed in *P. berghei* schizont stages

In order to confirm that the 3-HA-tagged PbDHHC proteins from the transgenic parasite lines were expressed in parasite blood-stages, immunoblots of purified schizont preparations, obtained from all of the nine 3-HA-tagged transgenic *P. berghei* lines, were performed using antibodies against the 3-HA epitope tag. It is important to note that the tagging strategy results in integration of the 3-HA epitope at the endogenous locus, meaning that the PbDHHC protein is still expressed under its native promoter.

The predicted sizes of the PbDHHC proteins listed on *PlasmoDB* (<http://plasmodb.org>) ranges from 35 kDa to 160 kDa. As all the PbDHHC proteins have 4 TM-domains, the difference in their sizes is mainly due to the presence or absence of ankyrin repeats in their N-terminal domains, or in the case of PbDHHC4, a very long C-terminal domain. Immunoblot of the PbDHHC schizont extracts revealed single bands that ran at the expected sizes for five of the nine PbDHHC schizont preparations (PbDHHC3, 5, 7, 8 and 9) as shown in Figure 4.3.

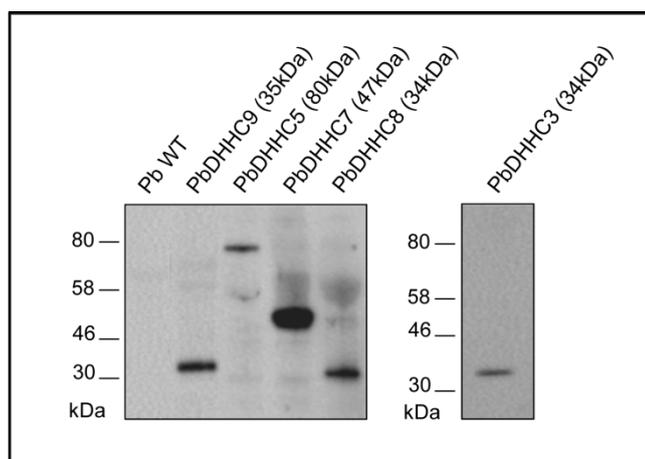


Figure 4.3: Expression of PbDHHC proteins in purified *P. berghei* schizonts. Immunoblot analysis was performed on total protein extracts from purified *P. berghei* schizonts, expressing the 3-HA-tagged PbDHHC proteins. Membranes were probed with α -HA antibodies and the expected protein sizes are shown in brackets.

Bands were not detected for the remaining tagged DHHC proteins (PbDHHC4, 6, 10 and 11). This could be due to low expression levels of the tagged protein. Alternatively, the inability to detect the protein could be due to stage-specific expression. For the purposes of this work, only the intraerythrocytic schizont stages were studied and it is possible that the DHHC proteins not detected in schizonts could perhaps be expressed at different stages, either of the intraerythrocytic cycle (for

example, ring or trophozoite stages), or of the whole life cycle (for example, the liver or mosquito stages). Some transcription data supports such an explanation for some of the DHHC genes – RNAseq data indicates that the expression of *P. falciparum* homologues of PbDHHC6 and 10 appear to be more up-regulated in gametocyte stages [7], and neither PbDHHC6 or 10 were detected in schizonts here. This suggests that while *Plasmodium* parasites have a similar number of DHHCs as other single-celled eukaryotes, such as *S. cerevisiae*, because of their complex life cycle, they may rely on a more restricted subset of these genes at specific life cycle stages.

4.2.3. Subcellular localisation of DHHC proteins in *P. berghei*

In other eukaryotes, DHHCs are predominantly found in the ER and the Golgi. For example, the yeast DHHC-PAT, Erf2 localises to the ER [8], while another DHHC-PAT, Akr1, localises to the Golgi [1]. However, *Plasmodium* parasites possess several unique organelles not found in other species. In order to establish the subcellular localisation of *Plasmodium* DHHCs, and specifically to determine whether any were localised to the unique parasite organelles, immunofluorescence microscopy was performed on the parasite lines containing 3-HA-tagged versions of the five PbDHHC proteins (PbDHHC3, 5, 7, 8 and 9) that could be detected by immunoblot. All five of these PbDHHCs appeared to distribute to discrete foci in *P. berghei* schizonts (Figure 4.4). The possible localisations of these proteins were determined by comparing the immunofluorescence staining of all five PbDHHC proteins to that of two proteins with well-established localisations, for which high quality antibodies were available: ERD2 (Golgi marker) [9] and merozoite surface protein 1 (MSP1) (plasma membrane marker) [10].

PbDHHC8 displayed an intracellular staining which did not co-localise to either ERD2 or MSP1 (Figure 4.4D). Further investigation using the endoplasmic reticulum (ER) marker, BIP [11], revealed that PbDHHC8 also did not co-localise with BIP (Figure 4.4D). However, this intracellular discrete staining, that was not Golgi or ER staining, could suggest localisation to vesicles. Unfortunately, the number of available localisation markers in *P. berghei* is limited, and thus, the specific location of PbDHHC8 is as yet unknown.

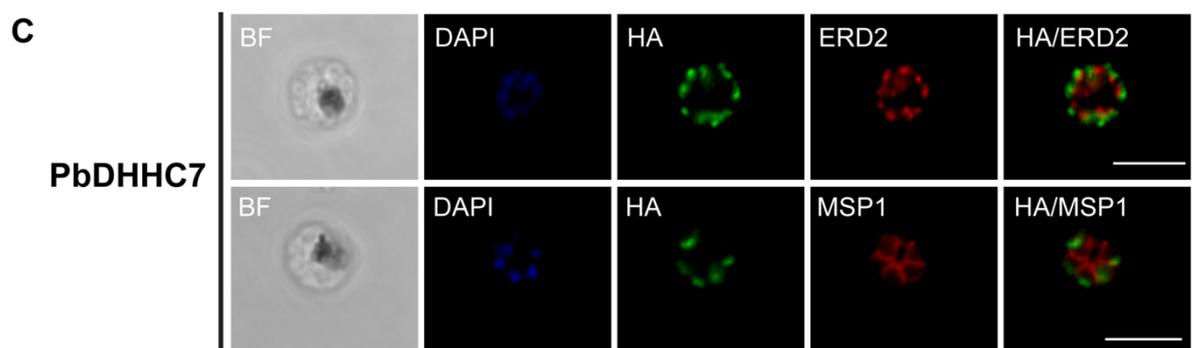
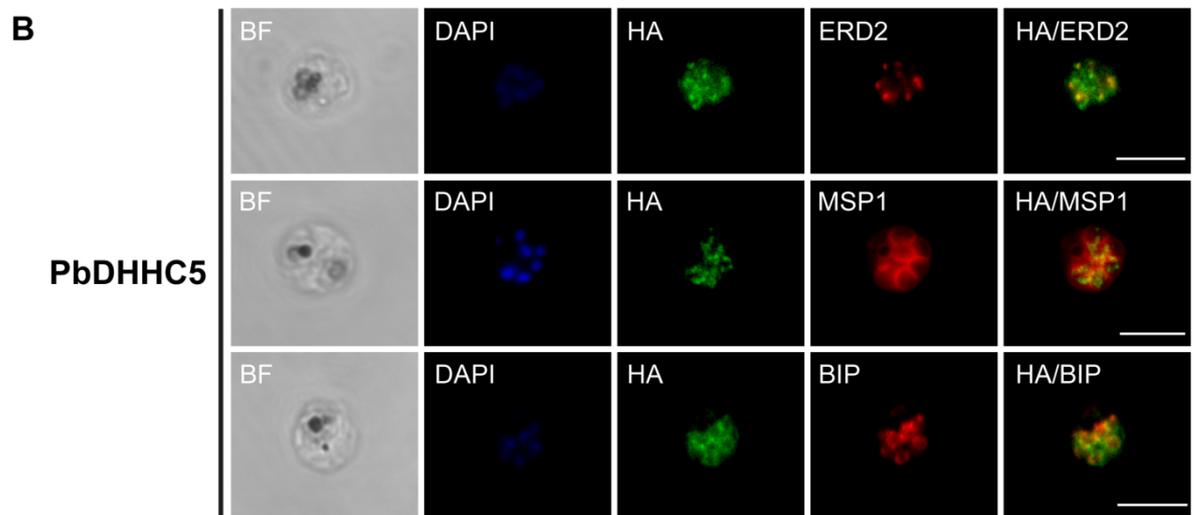
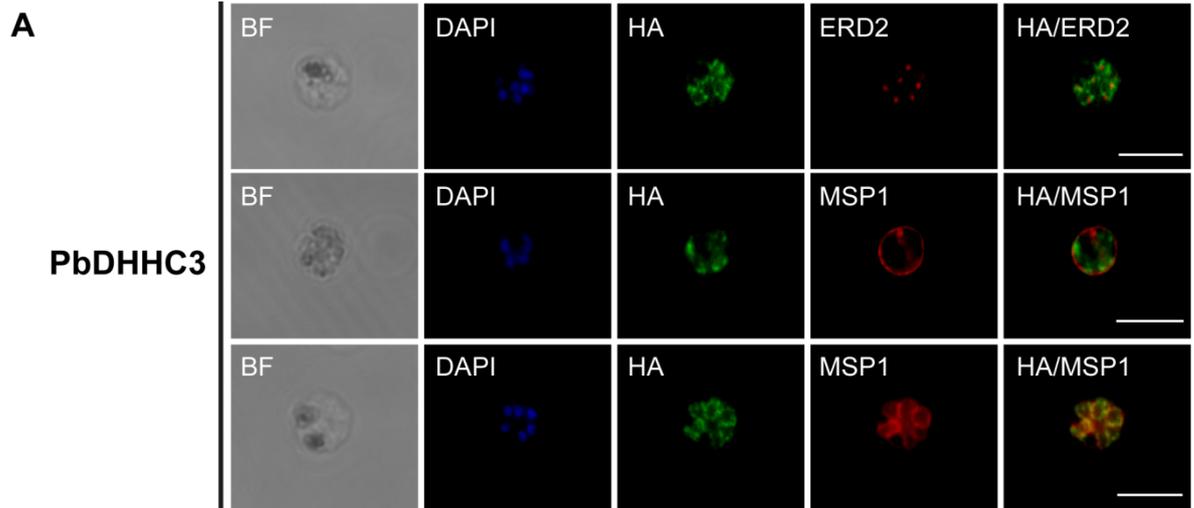
Both PbDHHC3 and PbDHHC9 appeared to co-localise with MSP1 in late schizonts (Figure 4.4A and E). However, this co-localisation was not observed in earlier stages of schizogony (data shown for PbDHHC3, Figure 4.4A), suggesting an inner membrane complex (IMC) localisation rather than a plasma membrane localisation. The IMC is a distinct morphological feature of alveolate organisms such as *Plasmodium*, and consist of flattened, membranous vesicles that underlie the plasma

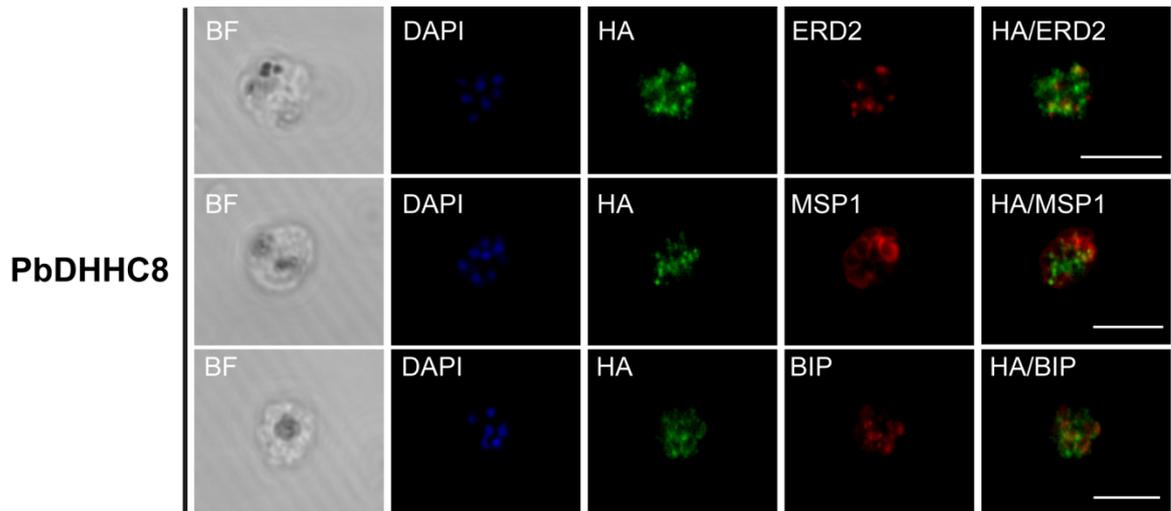
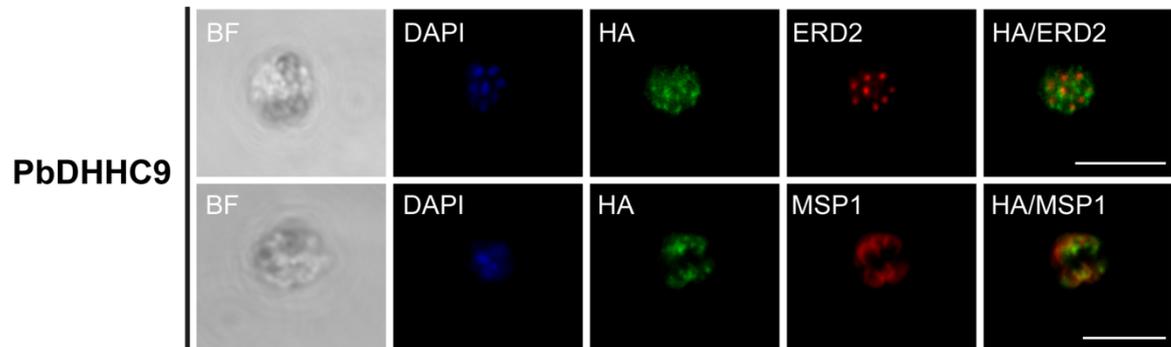
membrane. In mature schizonts where the IMC is fully developed, it lies closely beneath the plasma membrane, and at this stage, staining of the IMC is closely similar to staining of the plasma membrane [12]. However, in the earlier stages of the life cycle, such as the ring and trophozoite stages, as well as during the early stages of schizogony, the IMC is still in the process of developing, and does not yet lie closely against the plasma membrane. At these stages, the staining of the IMC appears more punctate, and does not appear similar to the plasma membrane [12]. Thus, as this was observed for both PbDHC3 and 9, it appears that these PbDHC proteins are localised to the IMC rather than the plasma membrane.

PbDHC5 did not co-localise with MSP1 but displayed some partial co-localisation with ERD2 (Figure 4.4B). Further investigation using the ER marker, BiP, also revealed partial co-localisation of PbDHC5 with BiP (Figure 4.4B), suggesting both a Golgi and ER localisation.

Lastly, PbDHC7 appeared to have a punctate distribution, different from that of ERD2 (Figure 4.4C). Co-staining with MSP1 suggested that PbDHC7 was apically located (Figure 4.4C). This suggests possible rhoptry localisation. Unfortunately, localisation to the rhoptries was unable to be confirmed due to the lack of antibodies against rhoptry markers in *P. berghei*.

The remaining five PbDHC proteins (PbDHC4, 6, 10 and 11) were unable to be detected by immunofluorescence assay, again perhaps due to low expression levels or stage-specific expression.



D**E**

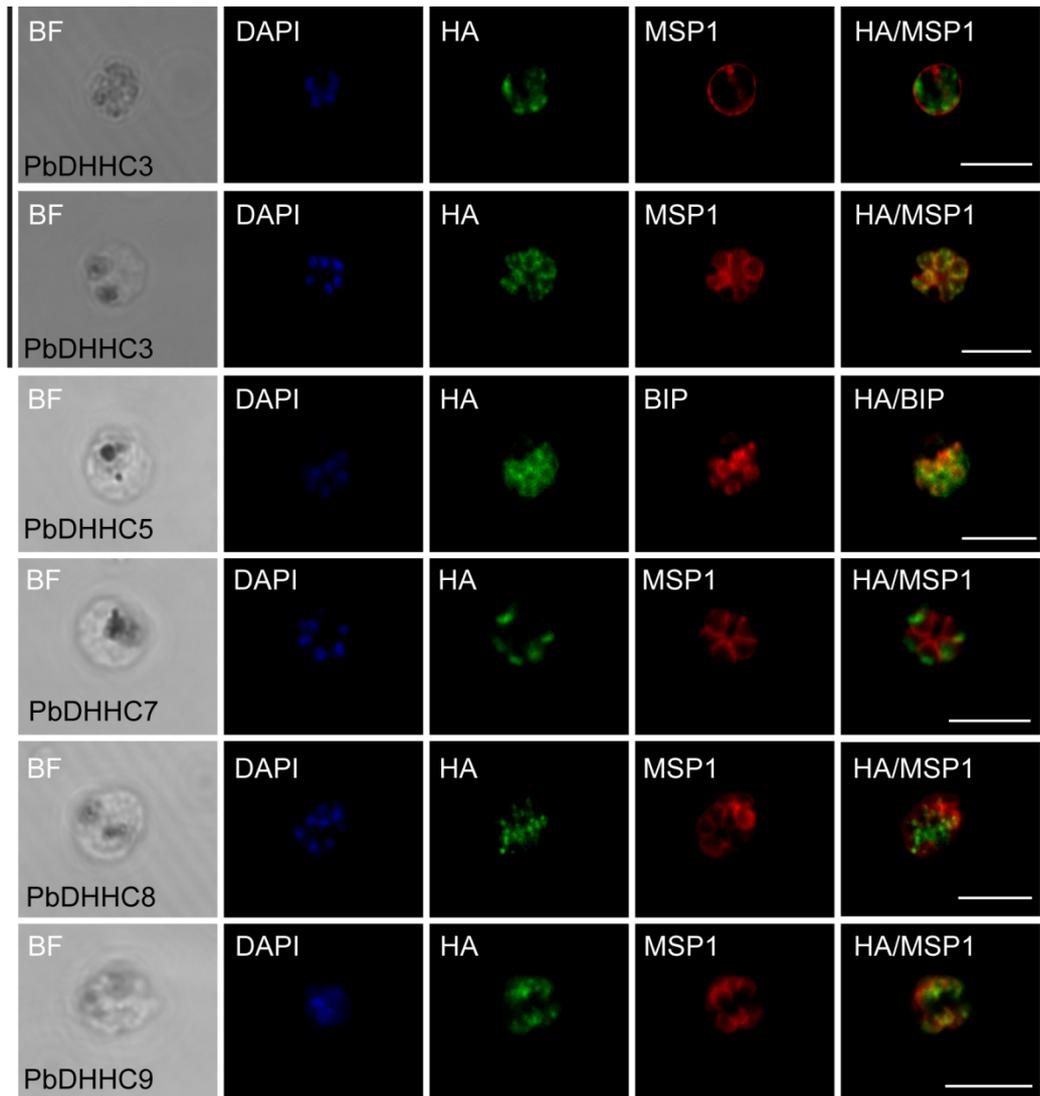
F

Figure 4.4: Localisation of DHHC proteins in *P. berghei* schizonts. Triple-HA-tagged PbDHHC proteins were localised by immunofluorescence using antibodies against the 3-HA tag (green) and comparing their immunofluorescence staining against that of the following known localisation markers (red): ERD2 (Golgi marker), MSP1 (plasma membrane marker), BIP (endoplasmic reticulum marker). Nuclear staining by DAPI is shown in blue. **(A)** Localisation of PbDHHC3. Staining is shown for both an early schizont and a late schizont in order to differentiate IMC from plasma membrane, as IMC staining is similar to plasma membrane staining in late stage schizonts but not early schizonts. **(B)** Localisation of PbDHHC5. **(C)** Localisation of PbDHHC7. **(D)** Localisation of PbDHHC8. **(E)** Localisation of PbDHHC9. **(F)** Summary of the localisations of the five 3-HA-tagged PbDHHC proteins. Scale bar: 5 μ m.

The localisation data described above has revealed that the five PbDHC proteins expressed in *P. berghei* schizonts appear to be localised to four different sites within the parasite. This includes typical DHC localisations, such as the ER and Golgi, as well as parasite specific organelles, such as the IMC and potentially the rhoptries (Figure 4.4F). The fact that the DHC proteins appear to be distributed to different locations in the parasite can imply two things. Firstly, the substrate specificity of the DHC proteins could be governed by their location, in that the DHC proteins only come in contact with a certain subset of target proteins at their specific location, thus regulating the specificity of the enzyme-substrate binding of these DHC proteins. Secondly, palmitoylation of target proteins could be regulated to occur only at a specific organelle, that is, the target protein is only palmitoylated when transported to the particular organelle where its DHC-PAT is located, which may be important in regulating protein activity.

The generation of the tagged transgenic parasite lines, PFGE analysis of the tagged lines, and isolation of infected mouse blood described in this section were performed by Ellen Bushell, Billker and Rayner lab, while I performed the *P. berghei* schizont purification, parasite protein extraction, immunoblotting, immunofluorescence assay, and the acquiring of images by confocal microscopy.

4.3. Essentiality of the DHC proteins in *P. berghei*

Knock-out vectors from the *PlasmoGEM* resource [5] were available for 9 out of the 11 PbDHC genes (PbDHC3-11), and were introduced into *P. berghei* strain ANKA 2.34 as described in the Materials and Methods. Integration of the knock-out vectors was expected to occur via double-crossover recombination of the homologous regions, resulting in the removal of the PbDHC gene of interest, and its replacement with the *hDHFR/yfcu* gene selection cassette, as shown in Figure 4.2A.

PFGE confirmed the integration of the knock-out vectors into the expected chromosomes for 7 of the 11 PbDHC genes: PbDHC3, 5-7 and 9-11, indicating the successful generation of transgenic knock-out lines for these seven PbDHC proteins (Figure 4.2B). This indicates that these seven PbDHCs are not essential for *P. berghei* asexual growth, and thus suggests that these genes could be functionally redundant in the asexual intraerythrocytic stages. Additionally, two of the PbDHC proteins successfully knocked-out here -PbDHC6 and 10- had no detectable blood-stage expression by epitope-tagging, as described above. This further supports the possibility that these proteins could be functionally important in another life cycle stage besides the intraerythrocytic stages.

Although knock-out vectors were available for PbDHC4 and 8, attempts to generate transgenic knock-out lines for these genes were unsuccessful. This could be an indication that these genes are not amenable to disruption, suggesting that these genes could be essential for blood-stage growth. However, confirmation of this hypothesis, by providing an episomally-expressed gene copy whilst simultaneously disrupting the gene locus, was not within the scope of this project.

Nevertheless, the results described here suggest that there may be two subsets of DHC proteins in *P. berghei*: one subset appears to be functionally redundant in the blood-stages, whereas the other subset of DHC proteins could be essential for blood-stage growth.

The generation of the knock-out transgenic lines and PFGE analysis of the knock-out lines described in this section were performed by Ellen Bushell, Billker and Rayner labs, and this data is included here due to its relevance to work described in other later chapters.

The work described in this chapter has since been published in [4].

Conclusion

This analysis of the repertoire of DHC-containing-proteins in *P. berghei* has provided several important insights. Firstly, only a subset of DHC proteins appears to be expressed in schizonts, implying that the other DHC proteins not detected in schizonts could be expressed in other life cycle stages of the parasite. Additionally, the fact that some DHC proteins appear to be essential for blood-stage growth, while other DHC proteins seem to be redundant, further supports the suggestion that stage-specific expression exists for the *Plasmodium* DHCs. Secondly, the PbDHC proteins appear to have different locations within the parasite, which may have implications on how substrate specificity of the DHCs could be regulated. However, the main focus of this project was to understand the palmitoylation site specificity of the *Plasmodium* DHC proteins. One of the ways this could be achieved was to combine the generation of DHC knock-out parasite lines with the site-ID palmitome purification method (described in Chapter 3), in an attempt to determine the specific cysteines palmitoylated by individual DHC-PATs. The palmitome is more well-established in *P. falciparum*, and large amounts of material for biochemical approaches can be more easily generated using *P. falciparum*. Thus, the next step in this study was to investigate the DHC-domain-containing proteins in *P. falciparum*, focusing on those that are more highly expressed in schizonts.

References:

1. Roth, A.F., *The yeast DHHC cysteine-rich domain protein Akr1p is a palmitoyl transferase*. The Journal of Cell Biology, 2002. **159**(1): p. 23-28.
2. Saitou, N. and M. Nei, *The neighbor-joining method: a new method for reconstructing phylogenetic trees*. Mol Biol Evol, 1987. **4**(4): p. 406-25.
3. Guindon, S. and O. Gascuel, *A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood*. Syst Biol, 2003. **52**(5): p. 696-704.
4. Fréchal, K., et al., *Global Analysis of Apicomplexan Protein S-Acyl Transferases Reveals an Enzyme Essential for Invasion*. Traffic, 2013. **14**(8): p. 895-911.
5. Pfander, C., et al., *A scalable pipeline for highly effective genetic modification of a malaria parasite*. Nat Methods, 2011. **8**(12): p. 1078-1082.
6. Janse, C.J., J. Ramesar, and A.P. Waters, *High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite Plasmodium berghei*. Nat Protoc, 2006. **1**(1): p. 346-356.
7. López-Barragán, M.J., et al., *Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum*. BMC Genomics, 2011. **12**(1): p. 587.
8. Lobo, S., *Identification of a Ras Palmitoyltransferase in Saccharomyces cerevisiae*. Journal of Biological Chemistry, 2002. **277**(43): p. 41268-41273.
9. Elmendorf, H.G. and K. Haldar, *Identification and localization of ERD2 in the malaria parasite Plasmodium falciparum: separation from sites of sphingomyelin synthesis and implications for organization of the Golgi*. Embo J, 1993. **12**(12): p. 4763-73.
10. Holder, A.A., et al., *A malaria merozoite surface protein (MSP1)-structure, processing and function*. Mem Inst Oswaldo Cruz, 1992. **87 Suppl 3**: p. 37-42.
11. Van Dooren, G.G., et al., *Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of Plasmodium falciparum*. Mol Microbiol, 2005. **57**(2): p. 405-419.
12. Kono, M., et al., *Evolution and Architecture of the Inner Membrane Complex in Asexual and Sexual Stages of the Malaria Parasite*. Molecular Biology and Evolution, 2012. **29**(9): p. 2113-2132.

Chapter 5

Protein acyltransferases in *Plasmodium falciparum*

Protein acyltransferases in *Plasmodium falciparum*

The results described in the previous chapter (Chapter 4) provided a first look at the DHHC protein family in *Plasmodium* parasites. However, in order to allow for further downstream follow-up studies (for example, comparative site-ID palmitome purification of DHHC knock-out lines) the focus was now shifted to the DHHC protein family in *P. falciparum*.

As described in Section 4.1, there are 12 proteins containing the conserved DHHC domain in *P. falciparum*, all except one (PFB0140w/PfDHHC12) possessing a homologue in *P. berghei*. These DHHC proteins have subsequently been named PfDHHC1 to 12, corresponding to their homologues in *P. berghei* and to their closest homologues in *T. gondii* (Table 4.1). BLAST search also revealed that *P. falciparum* appears to have only one open reading frame annotated to be a putative MBOAT family member [1], the second family of proteins thought to act as PATs. Due to the technical difficulties of genetic manipulation in *P. falciparum*, 5 out of the 12 PfDHHC proteins which appear to be more highly expressed in schizonts (PfDHHC3, 5, 7, 8 and 9) [2], were chosen to be C-terminally tagged with a 3-HA epitope tag, in order to determine their subcellular localisation. Additionally, the MBOAT protein in *P. falciparum* was also chosen to be C-terminally tagged with the 3-HA epitope tag. Gene disruption strategies were also performed on the same five PfDHHC proteins and one PfMBOAT protein in an attempt to investigate the essentiality and function of these PATs in *P. falciparum*.

5.1. Expression and localisation of DHHC and MBOAT proteins in *Plasmodium falciparum*

5.1.1. Generation of PfDHHC and PfMBOAT triple-HA (3-HA)-tagged transgenic parasite lines

In order to insert the 3-HA epitope tag into the C-terminus of the PAT genes of interest (PfDHHC3, 5, 7, 8 and 9, and PfMBOAT) at the endogenous loci, approximately 1000 bp of the C-terminal end of each gene was PCR amplified and introduced into a *P. falciparum* tagging vector (pCAM-BSD-3HA) containing the 3-HA epitope tag and the Blastocidin S deaminase gene (*BSD*), which confers resistance to the drug Blastocidin S [3], as described in the Materials and Methods. These PfDHHC and PfMBOAT tagging vectors were then introduced into *P. falciparum* strain 3D7 using the transfection methods described in Materials and Methods, and parasites that had taken up the tagging vector were selected for using resistance to Blastocidin S.

After transfection and drug selection, transgenic parasites resistant to Blastocidin S were obtained for all six PAT proteins studied here. The insertion of the 3-HA epitope tag into the genome was expected to occur via single-crossover recombination at the homologous C-terminal region, as

described in Figure 5.1A, meaning that the tagged DHHC gene would be under the control of its native promoter, which is critical as it removes the possibility of expression-related mis-localisation artefacts which can occur when an episomal strategy is used.

The genomic DNA of all the transgenic parasites were tested for the presence of sequences specific to the integrated construct, the wild-type locus and the episome, by PCR amplification using the primer pairs shown in the schematic in Figure 5.1A, and listed in the table in Figure 5.1B and in Table 2.5. For each of the genes of interest, successful PCR amplification of sequences specific to the integration of the tagging construct into the genomic DNA (primer pair: P5+P4 or P1+P4) indicated that the 3-HA tag was successfully inserted into the C-terminal end of the genes of interest (Figure 5.1B). However, in each of the tagged PfDHHC lines, the continued PCR amplification of sequences specific to the wild-type locus (primer pair: P1+P2) and to the episome (primer pair: P3+P4) indicated that for each transgenic parasite line, a proportion of the parasite population appeared to retain the DHHC tagging vector as an episome, and did not integrate the tagging vectors into the genomic DNA. Nevertheless, transgenic parasites containing the integrated construct could be identified using the expression of the 3-HA epitope tag, as expression of the 3-HA tag was not expected to be detectable when the vector was retained as an episome. For the PfMBOAT transgenic line, sequence specific to the wild-type locus was not amplified at the correct size (Figure 5.1B). This could mean that the wild-type locus is no longer present due to integration of the construct and the bands observed were just background amplification. However, the clear amplification of the sequence specific to integration of the construct indicated that there was indeed integration, and parasites with the integrated construct could be identified by the expression of the 3-HA tag.

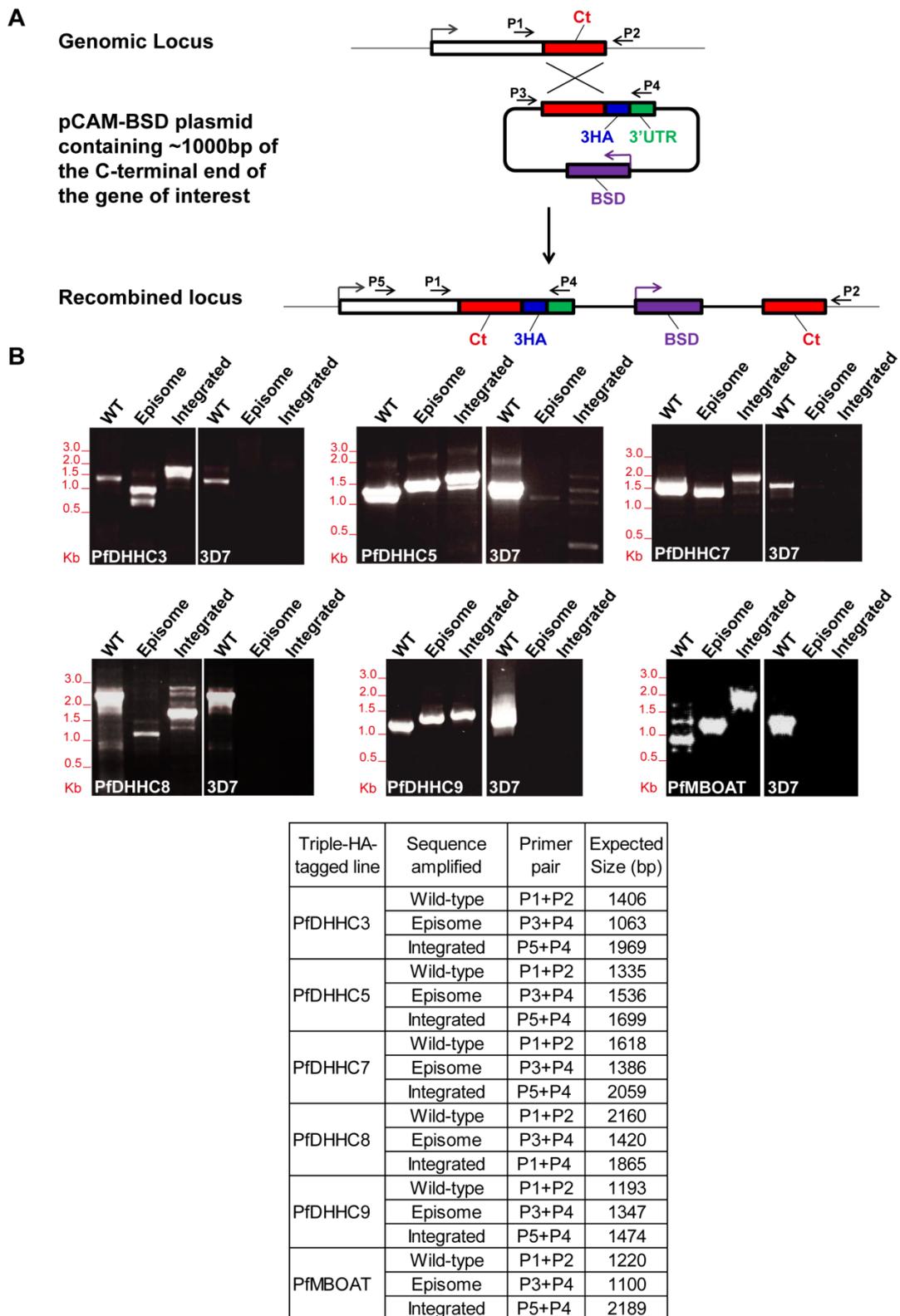


Figure 5.1: Generation of PfDHC and PfMBOAT triple-HA-tagged transgenic lines. (A) Scheme of the strategy used to C-terminally-tag the endogenous locus of the PfDHC and PfMBOAT genes of interest. The primer positions illustrated here indicate the primers used for genotyping of the 3-HA-tagged transgenic lines. Flags indicate promoter regions. **(B)** Genotyping of the PfDHC and PfMBOAT tagged transgenic lines by genomic PCR analysis. Primer pairs used for the amplification of sequences specific to the wild-type locus, episome and integrated construct are as listed in the table, along with the expected sizes of the fragments. For all six

transgenic lines, sequences specific to the integrated construct is amplified. For PfDHC3, 5, 7, 8 and 9, sequences specific to the wild-type locus and the episome are still amplified. For PfMBOAT, sequence specific to the wild-type locus is not amplified at the right size, which may indicate that the wild-type locus is no longer present and the bands observed may be due to background amplification.

5.1.2. Expression of DHC and MBOAT proteins in *P. falciparum*

In order to establish whether these tagged proteins were expressed in schizonts as expected, immunoblots of saponin-lysed schizont pellets, collected from all tagged transgenic *P. falciparum* lines, were performed using antibodies against the 3-HA epitope tag. As in *P. berghei*, the expected sizes of the proteins listed on PlasmoDB (<http://plasmodb.org>) ranged from 35 kDa to 160 kDa, depending on the presence or absence of N-terminal ankyrin repeats and the length of the C-terminal region. Immunoblot of the schizont pellets resulted in single bands which ran at the expected sizes for three of the PfDHC proteins (PfDHC5, 7 and 9), as well as for PfMBOAT protein, as shown in Figure 5.2. This indicated that PfDHC 5, 7, and 9, as well as PfMBOAT, were expressed in *P. falciparum* schizonts. Bands were not detected for PfDHC3 and 8. This could perhaps indicate that these DHC proteins are expressed at levels that are too low to be detected by immunoblot or by tagging at the endogenous locus. Interestingly, both homologues of these proteins in *P. berghei* (PbDHC3 and 8, Section 4.2.2), were successfully detected by immunoblotting. This could imply that the expression levels of these proteins differ between the species, perhaps due to the fact that there is one less DHC protein in *P. berghei* and another DHC protein may be needed to compensate.

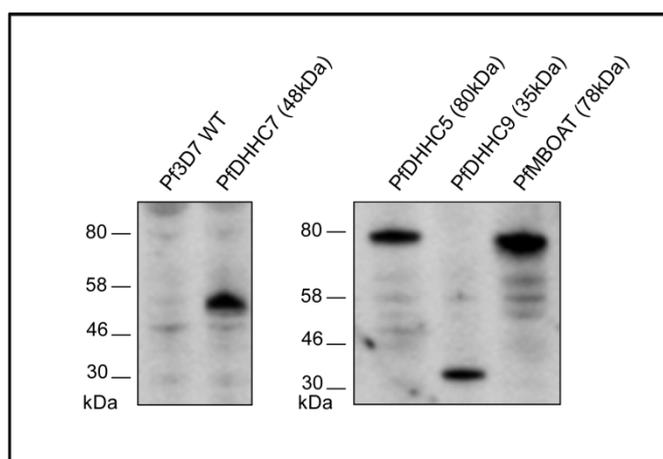


Figure 5.2: Expression of PfDHC and PfMBOAT proteins in *P. falciparum* schizonts. Immunoblot analysis was performed on total protein lysates from saponin-lysed *P. falciparum* schizonts, expressing triple-HA-tagged DHC and MBOAT proteins. Membranes were probed with α -HA antibodies and the expected protein sizes are shown in brackets.

5.1.3 Subcellular localisation of DHC and MBOAT proteins in *P. falciparum*

In order to establish the subcellular localisation of these DHC proteins in *P. falciparum*, immunofluorescence microscopy was performed on the parasite lines containing the 3-HA-tagged versions of the five PfDHC proteins, as well as PfMBOAT, using antibodies against the 3-HA tag. Immunofluorescence signal was successfully detected for four of the PfDHCs (PfDHC3, 5, 7 and 9) as well as for PfMBOAT (Figure 5.3). Although the expression of PfDHC3 was not detected by immunoblot, it was successfully detected by immunofluorescence. This suggests that the lack of detection by immunoblot (Section 5.1.2) might not be due to low expression levels, but could instead be due to solubility of the protein in the detergents used for protein extraction. Membrane proteins as a whole are difficult to extract due to their high hydrophobicity, and this is even truer for multi-TM-domain proteins, such as the DHCs. Nevertheless, the successful detection of PfDHC3 by immunofluorescence microscopy indicates that this protein is expressed in schizonts. Immunofluorescence signal was not detected for PfDHC8, suggesting that the expression of PfDHC8 is probably at a level that is too low to be detected by tagging at the endogenous locus. As the homologue of this protein in *P. berghei* (PbDHC8, Section 4.2.3) was successfully detected by immunofluorescence, this further suggests that the expression levels of this DHC might differ between the species.

Similar to what was observed in *P. berghei* (Section 4.2.3), the PfDHC proteins, as well as PfMBOAT, appeared to be distributed to discrete foci in *P. falciparum* schizonts (Figure 5.3). The potential

subcellular localisations of all these proteins were determined by comparing their immunofluorescence signals to that of the following proteins with well-established localisations, for which high quality antibodies were available: ERD2 (Golgi marker) [4], BiP (ER marker) [5], rhoptry associated protein 1 (RAP1) (rhoptry marker) [6] and GAP45 (IMC marker) [7] (Figure 5.3).

The immunofluorescence staining of PfDHH3 was compared to the staining of all five localisation markers described above. PfDHH3 staining did not co-localise with BiP, GAP45, MSP1 or RAP1, indicating that PfDHH3 was not localised at the ER, IMC, plasma membrane or the rhoptries. However, PfDHH3 staining co-localised with ERD2 staining, indicating a Golgi localisation (Figure 5.3A). This was in contrast to its *P. berghei* homologue, PbDHH3, which was localised to the IMC (Section 4.2.3). The difference in location of this DHH protein between the two *Plasmodium* species could again be due to the fact that *P. berghei* is missing one DHH protein, and thus may require another DHH protein to compensate.

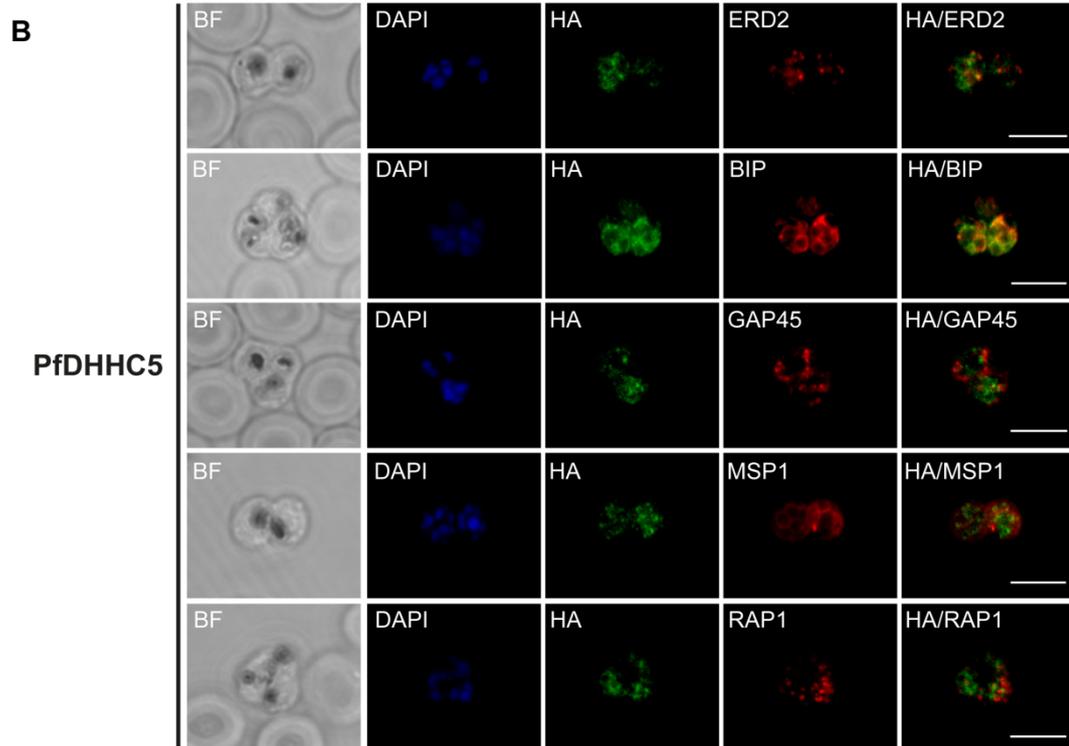
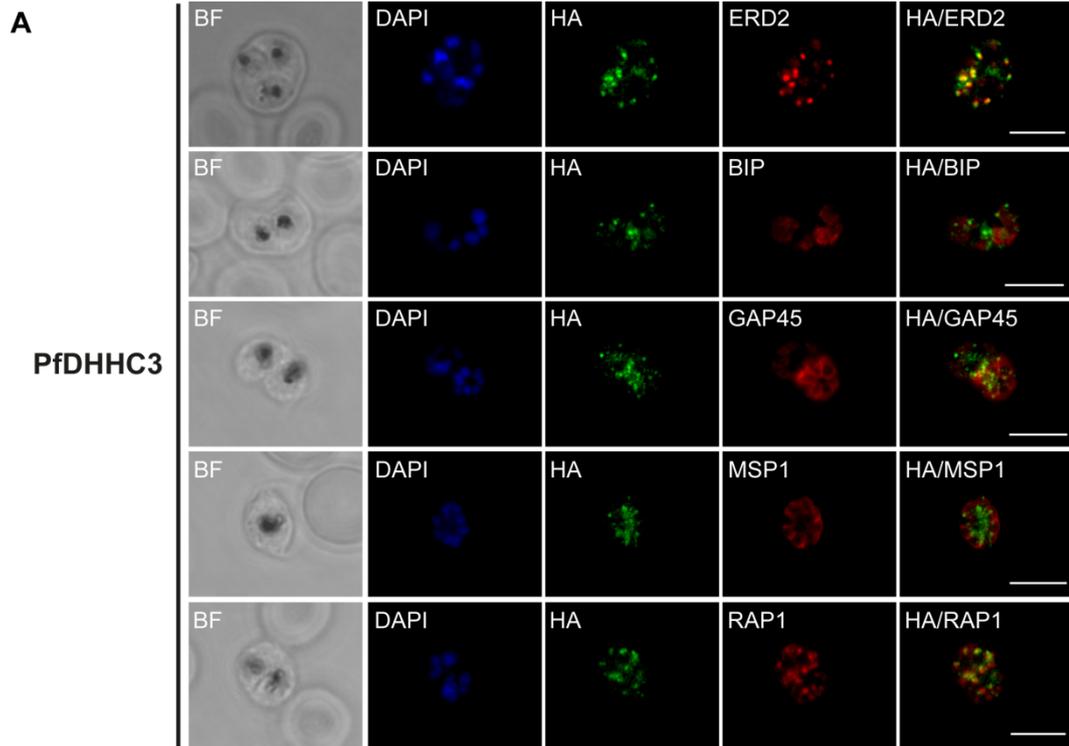
PfDHH5 staining did not co-localise with ERD2, GAP45, MSP1 or RAP1 staining, but co-localised with BiP staining (Figure 5.3B), suggesting an ER localisation, and this was similar to the localisation of its *P. berghei* homologue, PbDHH5 (Section 4.2.3).

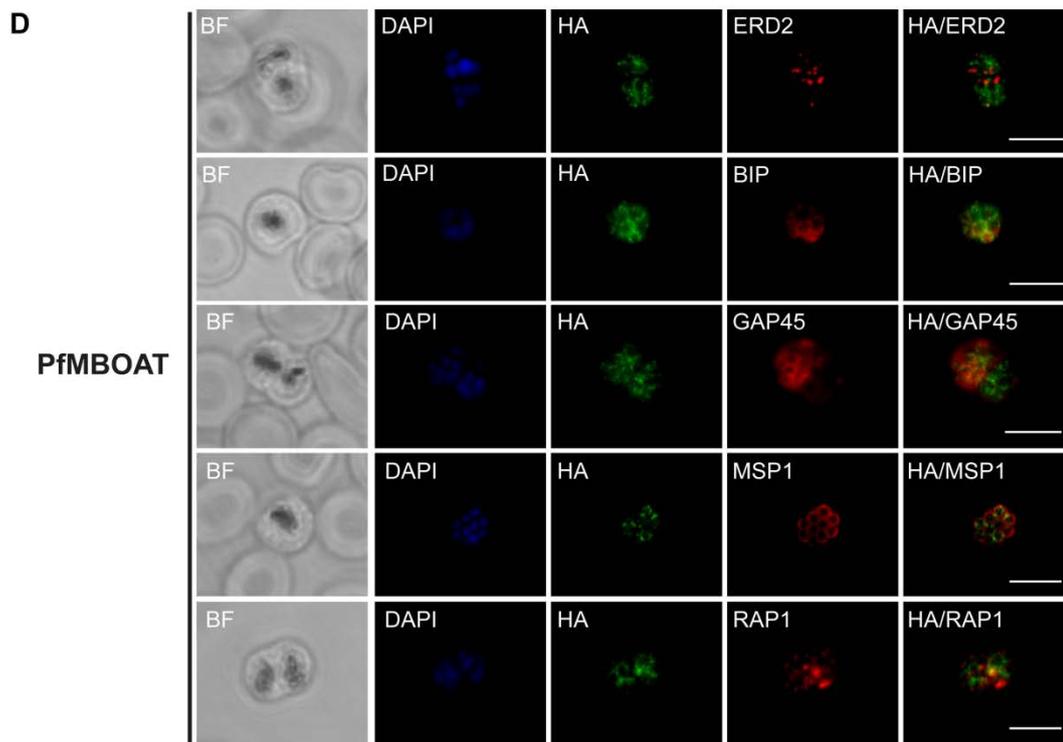
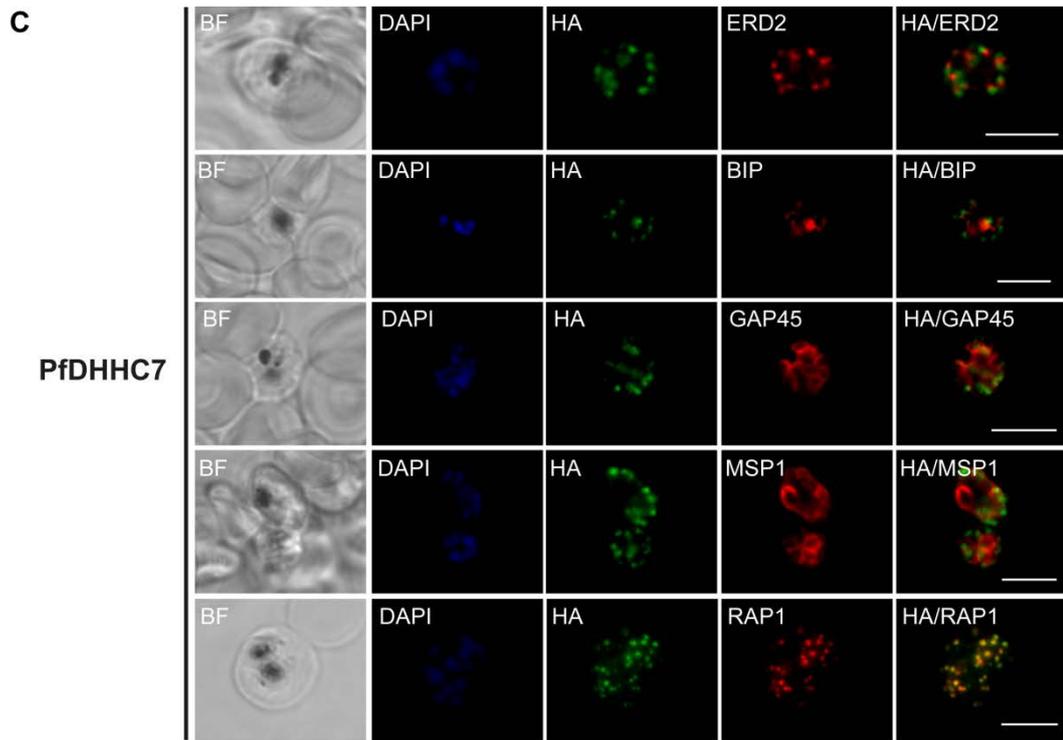
PfDHH7 displayed a punctate distribution that was not similar to that of ERD2 staining, indicating that it was not localised to the Golgi. PfDHH7 also did not co-localise with BiP, GAP45 or MSP1, although staining with MSP1 indicated that PfDHH7 appeared to be distributed apically. However, PfDHH7 co-localised with RAP1 staining (Figure 5.3C), indicating a rhoptry localisation. PbDHH7 staining was also punctate and was apically located in *P. berghei* schizonts (Section 4.2.3). That, and the fact that PfDHH7 co-localises with RAP1, suggests that PbDHH7 may also localise to the rhoptries.

PfDHH9 did not co-localise with ERD2, BiP or RAP1 staining. However, PfDHH9 appeared to co-localise with both GAP45 and MSP1 staining in late schizonts (Figure 5.3E). Further investigation by comparing PfDHH9 staining to GAP45 and MSP1 staining in early schizonts, revealed that PfDHH9 co-localised with GAP45 but not with MSP1 staining in the earlier stages of schizogony (Figure 5.3E). As described in the previous chapter (Section 4.2.3), IMC and plasma membrane staining is similar in mature schizonts but is distinctly different in early schizonts. Thus, this indicates an IMC localisation for PfDHH9, which is in keeping with the localisation of its *P. berghei* homologue, PbDHH9 (Section 4.2.3).

Lastly, PfMBOAT did not co-localise with ERD2, GAP45 or MSP1, indicating that it was not localised to the Golgi, IMC or plasma membrane. PfMBOAT did however co-localise with BIP staining (Figure 5.3D), indicating an ER localisation.

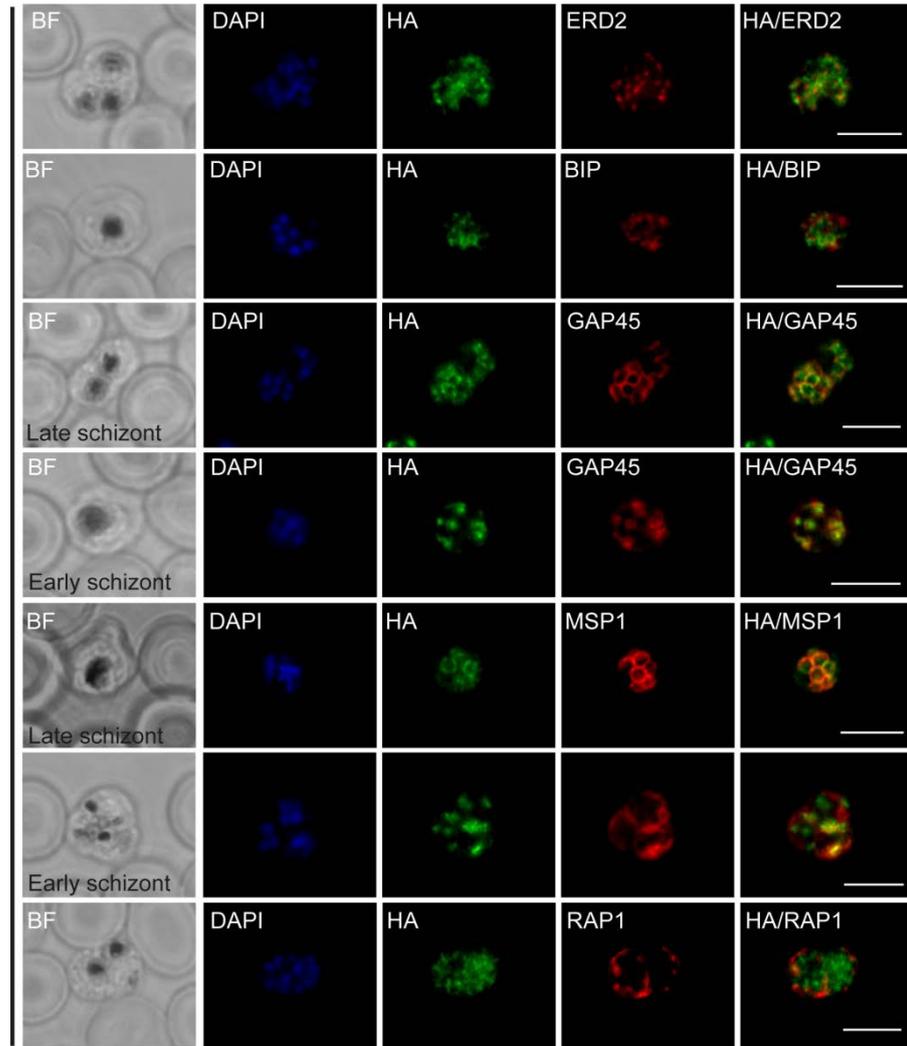
In summary, the five *Plasmodium* PATs studied here appear to be localised to different membrane compartments within the parasite. PfDHHC3 is localised to the Golgi, PfDHHC5 localises to the ER, PfDHHC7 localises to the rhoptries, PfDHHC9 localises to the IMC, and PfMBOAT also localises to the ER (Figure 5.3F). Similar to *P. berghei*, the *P. falciparum* PATs appear to be distributed both to organelles where PATs in other eukaryotic organisms are typically localised to, such as the ER and Golgi, as well as unique parasite organelles, such as the IMC and rhoptries. The results described here also indicate that although most of the homologues of the two *Plasmodium* species appear to be localised to the same place, this is not always the case, as PfDHHC3 and PbDHHC3 localised to different sites in the two *Plasmodium* species.





E

PfDHH9



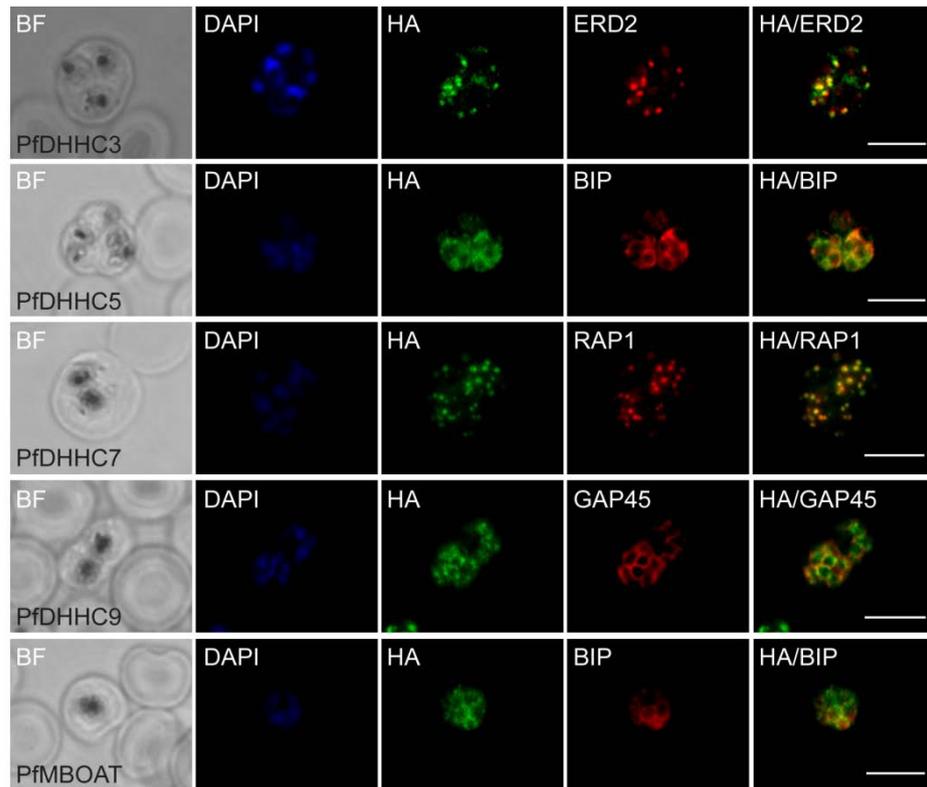
F

Figure 5.3: Localisation of DHHC and MBOAT proteins in *P. falciparum*. Triple-HA-tagged PfDHHC and PfMBOAT proteins were localised by immunofluorescence using antibodies against the 3-HA tag (green). Immunofluorescence staining of each of the tagged PfDHHC and PfMBOAT proteins were compared against that of the following known localisation markers (red): ERD2 (Golgi marker), BIP (endoplasmic reticulum marker), RAP1 (rhoptry marker) and GAP45 (inner membrane complex marker). Nuclear staining by DAPI is shown in blue. **(A)** Localisation of PfDHHC3. **(B)** Localisation of PfDHHC5. **(C)** Localisation of PfDHHC7. **(D)** Localisation of PfMBOAT. **(E)** Localisation of PfDHHC9. For the staining of PfDHHC9 with GAP45 and MSP1, both a late schizont, as well as an early schizont, is shown in order to differentiate between IMC and plasma membrane localisation. **(F)** Summary of the localisations of all PfDHHC and PfMBOAT proteins. Scale bar: 5 μ m.

5.2. Essentiality of DHHC and MBOAT proteins in *P. falciparum*

5.2.1. Generation of PfDHHC and PfMBOAT knock-out transgenic parasite lines

In order to knock-out the PAT genes of interest (PfDHHC3, 5, 7, 8, and 9, and PfMBOAT), 600-800 bp of the N-terminal and C-terminal regions of each gene of interest were PCR amplified and introduced into the *P. falciparum* negative-selectable knock-out vector (pCC-1), which contains two selection cassettes: the human dihydrofolate reductase gene (*hDHFR*), which confers resistance to the antifolate inhibitor WR99210 [8], and the *Saccharomyces cerevisiae* cytosine deaminase/uracil phosphoribosyltransferase chimeric gene (*ScFCU*), the 'suicide gene' which converts 5-fluorocytosine (5-FC) into the toxic 5-fluorouracil when 5-FC is present [9].

The two selection cassettes present in the pCC-1 knock-out vector allows two types of drug selection to be performed on transgenic parasites which have taken up the vector. Firstly, transgenic parasites are selected for using WR99210 (positive drug selection). The drug resistant parasite population, which consists of parasites which have integrated the construct into the genome, and parasites which are harbouring the construct episomally, can then be put through several rounds of drug cycling, where drug pressure is removed for 3 weeks before being reapplied. This is because episomal plasmids are unstable, and can be lost in the absence of drug pressure, whereas the stable integrant will be maintained [10]. Thus, the use of drug cycling should enrich for parasite populations which have integrated the construct and are not harbouring the construct episomally. Secondly, transgenic parasites containing the integrated construct can be further selected for using 5-FC (negative drug selection). In the presence of the episome, *ScFCU* is expressed and converts 5-FC to the toxic 5-fluorouracil. When the knock-out vector is integrated by double-crossover recombination, the episomal plasmid is linearised and *ScFCU* is not expressed, resulting in transgenic parasites which survive in the presence of 5-FC [9]. Unfortunately, the use of 5-FC negative selection does not select against single-crossover recombination, as in this form of recombination, the entire plasmid is inserted into the genome and *ScFCU* is inserted into the endogenous locus in a 3' to 5' configuration, with its promoter also in a 3' to 5' configuration, and is thus also not expressed.

The *P. falciparum* PAT knock-out constructs were designed such that the N-terminal and C-terminal targeting regions of each gene were inserted into the 5' and 3' flanking regions of the *hDHFR* selection cassette respectively, as shown in Figure 5.4A and described in Materials and Methods. These PfDHHC and PfMBOAT knock-out vectors were then introduced into *P. falciparum* strain 3D7 using the transfection methods described in Materials and Methods. After transfection and positive drug selection using WR99210, as well as after drug cycling, transgenic parasites that had taken up

the knock-out constructs and thus gained resistance to WR99210 were obtained for all six PAT genes studied here (PfDHHC3, 5, 7, 8 and 9, and PfMBOAT). The genomic DNA of all six transgenic parasite lines were then genotyped using PCR amplification in order to determine whether the knock-out construct had been integrated into the genomic DNA or was being maintained episomally.

5.2.2. Genotyping of PfDHHC and PfMBOAT knock-out transgenic parasite lines

The integration of the PfDHHC and PfMBOAT knock-out constructs into the genomic DNA was expected to occur via double-crossover recombination at the homologous N-terminal and C-terminal targeting regions. This would result in the replacement of a section of the genes of interest –which includes the conserved DHHC domain (for the PfDHHC proteins) or the conserved histidine residue (for PfMBOAT protein) - with the *hDHFR* selection cassette, thus causing the disruption of the genes of interest, as shown in Figure 5.4A.

Integration of these knock-out constructs could also potentially occur via single-crossover recombination at either the N-terminal or the C-terminal homologous targeting regions (Figure 5.4A). In this case, instead of just the *hDHFR* selection cassette, the entire plasmid is inserted into the endogenous locus. As mentioned above, treatment with the negative selection drug, 5-FC, would not select against this form of integration, as in both cases, *ScFCU* is inserted into the genome in a 3' to 5' configuration and is thus not expressed. With single-crossover recombination, although the gene of interest is still disrupted, a complete copy of the gene is still present. However, if the single-crossover recombination were to occur at the N-terminal region, the complete copy of the gene is moved further downstream, away from its endogenous promoter (Figure 5.4A), and expression of the wild-type gene could potentially be reduced.

As mentioned above, after transfection and positive drug selection, transgenic parasites which had taken up the knock-out constructs were obtained for all six PAT genes studied here. The genomic DNA of all six transgenic parasites lines were analysed by PCR amplification for the presence of sequences specific to the integrated construct –both double-crossover and single crossover (N-terminal and C-terminal) - the wild-type locus and the episomal plasmid, using the primer pairs shown in the schematic in Figure 5.4A and listed in the table in Figure 5.4B and in Table 2.6.

(1) Genotyping of PfDHHC3 and 7, and PfMBOAT knock-out transgenic parasite lines

For these 3 transgenic strains, analysis by PCR amplification revealed the presence of sequences specific to the wild-type locus (primer pair: P6+P10) and the episome (primer pair: P8+P9). No sequences specific to the double-crossover integrated construct (primer pair:

P6+P9), N-terminal single-crossover integrated construct (primer pair: P8+P10) or C-terminal single-crossover integrated construct (primer pair: P12+P13) were observed (Figure 5.4B). Furthermore, amplification of the full-length gene locus (primer pair: P6+P7) resulted in a fragment with a size equal to the expected size of the wild-type locus (Figure 5.4B). This was still the case even after drug cycling. The three transgenic lines were then treated with the negative selection drug, 5-FC, in an attempt to remove parasites harbouring the plasmid episomally, with the intention to select for the parasite population which contained the integrated construct, if such a population existed. Strangely, sequence for the episome was still amplified even after 5-FC treatment, indicating that the episome still remained in the presence of 5-FC. This was observed even in increased concentrations of 5-FC. The reason for this is unknown. However, a potential explanation could be that the PfDHH3 and 7 and PfMBOAT genes are not amenable to disruption, but the transgenic parasites have somehow gained resistance to 5-FC in an attempt to survive. A similar phenotype, where 5-FC resistance was gained without loss of the *ScFCU*-containing plasmid has also been previously observed to occur in *P. falciparum* [11], and indeed, spontaneous resistance to 5-FC has been documented in other species [12, 13]. Given that there was no evidence of integration for any of these 3 strains, they were not analysed any further.

(2) Genotyping of PfDHH8 knock-out transgenic parasite line

PCR analysis of this transgenic line resulted in the amplification of sequences specific to the wild-type locus (primer pair: P6+P10) and the episome (primer pair: P8+P9). No sequence was amplified for the double-crossover integrated construct (primer pair: P6+P9), N-terminal single-crossover integrated construct (primer pair: P8+P10), or the C-terminal single-crossover integrated construct (primer pair: P12+P13) (Figure 5.4B). Amplification of the full-length gene locus (primer pair P6+P7) resulted in a fragment with a size equal to that expected for the wild-type locus (Figure 5.4B). This was still observed after drug cycling. However, in this case, subsequent negative selection with 5-FC resulted in parasite death and no parasites were observed even after three weeks of culturing. This could be an indication that the PfDHH8 gene cannot be disrupted, and suggests that PfDHH8 could thus be an essential gene. As with PfDHH3 and 7, and PfMBOAT transgenic lines, the PfDHH8 transgenic line was not analysed any further.

(3) Genotyping of PfDHHc5 knock-out transgenic parasite line

After positive drug selection, drug cycling, as well as negative selection with 5FC, PCR analysis of this transgenic line revealed the presence of sequences specific to the double-crossover integrated construct (primer pair P6+P9). Sequences specific to the wild-type locus (primer pair: P6+P10) and the episome (primer pair P8+P9) were no longer amplified (Figure 5.4B), and no amplification was observed for the N-terminal single-crossover integrated construct (primer pair P8+P10) or the C-terminal single-crossover integrated construct (primer pair: P12+P13). Amplification of the full-length gene locus (primer pair: P6+P7) resulted in a fragment with a size equal to the size expected if double-crossover integration had occurred (Figure 5.4B). This was further confirmed using an additional primer (primer P11) located in the 3' UTR of the PfDHHc5 gene (Figure 5.4A). Amplification of the full-length gene locus using primer P6 with primer P11 also resulted in a fragment with a size equal to the size expected if double-crossover integration had occurred (Figure 5.4B). This indicated that for this transgenic line, double-crossover recombination had occurred, and the PfDHHc5 gene had been disrupted, resulting in a PfDHHc5 knock-out transgenic line. The PfDHHc5 knock-out line was then cloned by limiting dilution, as described in Materials and Methods, in order to generate parasite clones derived from single parasites. All clones were also analysed by PCR in order to ensure that each clonal population displayed double-crossover integration (data not shown).

(4) Genotyping of PfDHHc9 knock-out transgenic parasite line

After positive drug selection, drug cycling and negative selection with 5FC, PCR analysis revealed the presence of sequence specific to single-crossover integration at the N-terminus (primer pair: P8+P10) (Figure 5.4B). Sequence specific to the episome (primer pair: P8+P9) and the unmodified wild-type locus (primer pair: P6+P10) were no longer amplified, indicating that the episome was not retained and that integration had indeed occurred (Figure 5.4B). Sequence specific to the C-terminal single-crossover integrated construct (primer pair: P12+P13) was also not amplified (Figure 5.4B). Amplification of sequence indicating double-crossover integration (primer pair: P6+P9) was observed, and amplification of the full-length gene locus (primer pair: P6+P7) resulted in a fragment with a size equal to the size expected if double-crossover integration had occurred. However, the use of both these double-crossover integration primer pairs described above would result in amplification of fragments which also exist in an N-terminal single-crossover integrated construct, as shown in the schematic in Figure 5.4A. Therefore, further analysis by PCR was

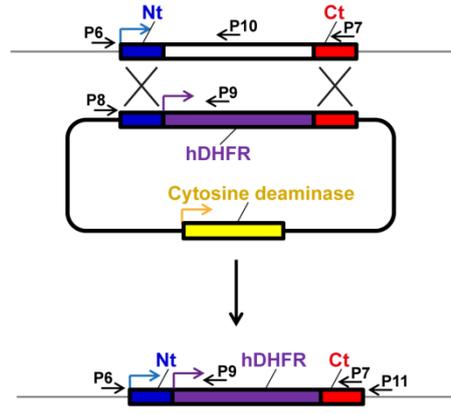
performed using an additional primer located in the 3' UTR of the PfDHHC9 gene (primer P11, Figure 5.4A and B). Amplification of the full-length gene locus using primer P6 with primer P11 did not amplify either the fragment expected if double-crossover recombination had occurred, or the fragment expected for the wild-type locus (Figure 5.4B). However, sequence specific to the presence of the wild-type locus located downstream of the inserted knock-out construct (primer pair: P8+P11) was amplified instead (Figure 5.4B), confirming that the knock-out vector had indeed integrated via single-crossover at the N-terminus. This means that although the PfDHHC9 gene has been disrupted, a full-length copy of the wild-type gene is still present. However, as the wild-type copy has now moved away from its endogenous promoter, expression of the gene could potentially be reduced, perhaps resulting in a knock-down rather than a knock-out line. Again, the PfDHHC9 transgenic line was cloned by limiting dilution and all clones analysed by PCR (data not shown).

The genotyping of all the transgenic parasite strains described above indicated that although transfection of the knock-out constructs was successful and parasites resistant to WR99210 were obtained for all six PATs, successful integration was detected for only two of the proteins of interest (PfDHHC5 and 9). It must be noted however, that all the genotyping described here was performed by PCR analysis and further confirmation of the integration of these knock-out constructs should be also performed by Southern blot analysis. Unfortunately, due to time constraints, data from Southern blot analysis was not yet available at the time of writing this dissertation.

A

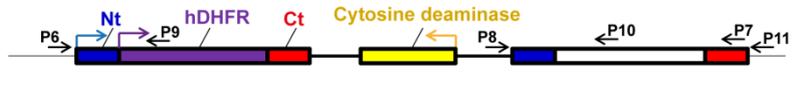
Genomic Locus

pCC-1 plasmid containing ~600-800bp of the N-terminal and C-terminal end of the gene of interest

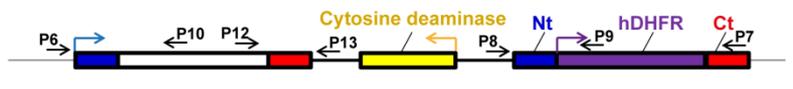


Recombined locus (double cross-over)

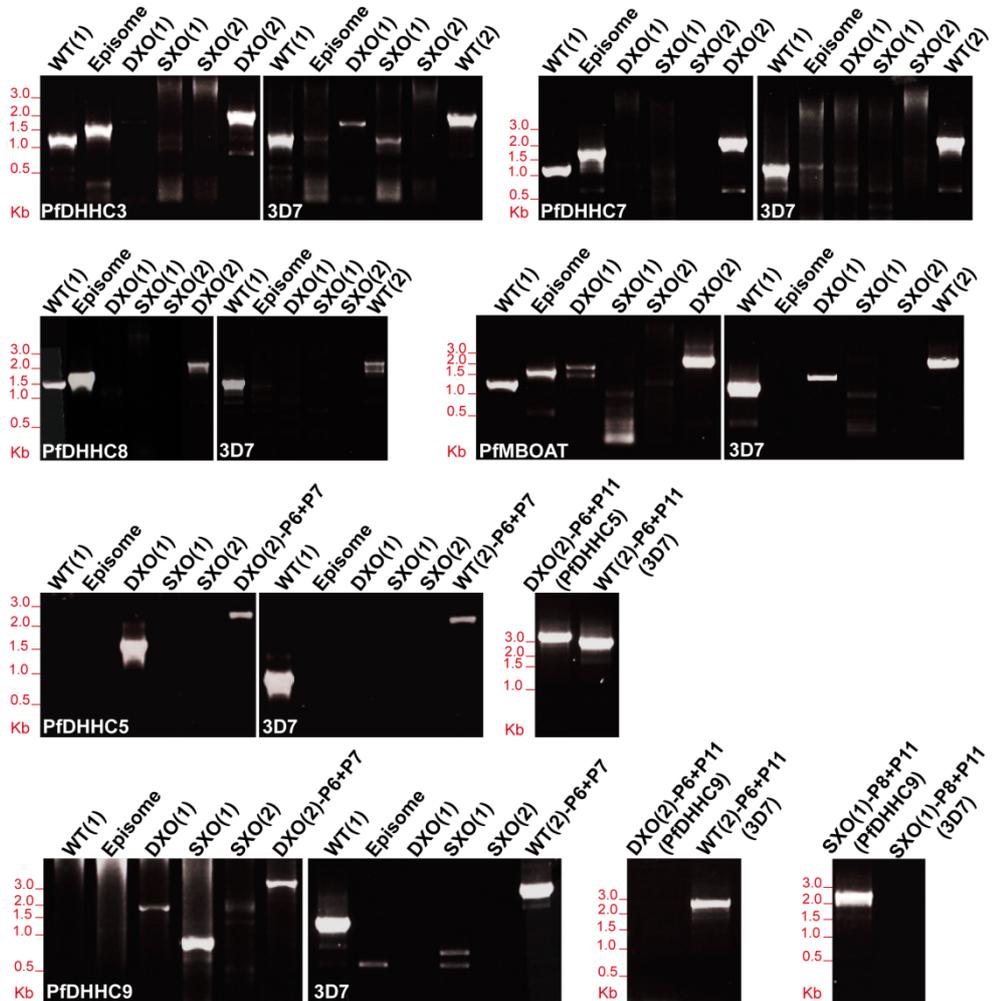
Recombined locus (single-crossover at N-terminus)



Recombined locus (single-crossover at C-terminus)



B



Knock-out line	Sequence amplified	Primer pair	Expected Size (bp)
PfDHHHC3	WT(1) - Wild-type (Fragment)	P6+P10	1294
	WT(2) - Wild-type (Full-length)	P6+P7	2291
	Episome	P8+P9	1769
	DXO(1) - Double-crossover integration (Fragment)	P6+P9	1978
	DXO(2) - Double-crossover integration (Full-length)	P6+P7	3727
	SXO(1) - Single-crossover integration (Nt)	P8+P10	1050
	SXO(2) - Single-crossover integration (Ct)	P12+P13	1075
PfDHHHC5	WT(1) - Wild-type (Fragment)	P6+P10	923
	WT(2) - Wild-type (Full-length)	P6+P7 and P6+P11	2753 and 2908
	Episome	P8+P9	1794
	DXO(1) - Double-crossover integration (Fragment)	P6+P9	1780
	DXO(2) - Double-crossover integration (Full-length)	P6+P7 and P6+P11	3462 and 3617
	SXO(1) - Single-crossover integration (Nt)	P8+P10	920
	SXO(2) - Single-crossover integration (Ct)	P12+P13	1907
PfDHHHC7	WT(1) - Wild-type (Fragment)	P6+P10	1268
	WT(2) - Wild-type (Full-length)	P6+P7	2369
	Episome	P8+P9	1801
	DXO(1) - Double-crossover integration (Fragment)	P6+P9	1918
	DXO(2) - Double-crossover integration (Full-length)	P6+P7	3596
	SXO(1) - Single-crossover integration (Nt)	P8+P10	1151
	SXO(2) - Single-crossover integration (Ct)	P12+P13	1903
PfDHHHC8	WT(1) - Wild-type (Fragment)	P6+P10	1468
	WT(2) - Wild-type (Full-length)	P6+P7	2769
	Episome	P8+P9	1755
	DXO(1) - Double-crossover integration (Fragment)	P6+P9	2036
	DXO(2) - Double-crossover integration (Full-length)	P6+P7	3765
	SXO(1) - Single-crossover integration (Nt)	P8+P10	1153
	SXO(2) - Single-crossover integration (Ct)	P12+P13	1380
PfDHHHC9	WT(1) - Wild-type (Fragment)	P6+P10	1342
	WT(2) - Wild-type (Full-length)	P6+P7 and P6+P11	2406 and 2539
	Episome	P8+P9	1651
	DXO(1) - Double-crossover integration (Fragment)	P6+P9	1992
	DXO(2) - Double-crossover integration (Full-length)	P6+P7 and P6+P11	3711 and 3844
	SXO(1) - Single-crossover integration (Nt)	P8+P10 and P8+P11	1001 and 2168
	SXO(2) - Single-crossover integration (Ct)	P12+P13	1144
PfmBOAT	WT(1) - Wild-type (Fragment)	P6+P10	1190
	WT(2) - Wild-type (Full-length)	P6+P7	2414
	Episome	P8+P9	1684
	DXO(1) - Double-crossover integration (Fragment)	P6+P9	1979
	DXO(2) - Double-crossover integration (Full-length)	P6+P7	3755
	SXO(1) - Single-crossover integration (Nt)	P8+P10	895
	SXO(2) - Single-crossover integration (Ct)	P12+P13	1473

Figure 5.4: Generation of PfDHHHC and PfmBOAT knock-out transgenic lines in *P. falciparum*. (A) Scheme of the strategy used to knock-out the endogenous locus of the PfDHHHC and PfmBOAT genes of interest. The primer positions illustrated here indicate the primers used for genotyping of the PfDHHHC and PfmBOAT knock-out transgenic lines. Flags indicate promoter regions. (B) Genotyping of the PfDHHHC and PfmBOAT knock-out transgenic lines by genomic PCR analysis. Primer pairs used for the amplification of sequences specific to the wild-type locus, episome and integrated (double crossover and single crossover) constructs are as listed in the table, along with the expected sizes of the fragments. For PfDHHHC5 and PfDHHHC9 knock-out transgenic lines, additional PCR analysis using the primer P11 was used to confirm integration of the construct.

5.2.3. Analysis of gene expression in PfdHHC5 and 9 knock-out transgenic parasite lines

In order to confirm that the expression of PfdHHC5 and 9 proteins was abolished or at least reduced in the PfdHHC5 and 9 knock-out lines, RNA was extracted from blood-stage cultures infected with the transgenic lines of interest and reverse transcribed to make cDNA, as described in Materials and Methods. This was done for three of the limiting dilution clones of each knock-out/knock-down line: PfdHHC5 clones 6, 8, and 19, and PfdHHC9 clones 1, 2, and 3, as well as for wild-type 3D7 as a control. The presence of sequences specific to PfdHHC5 and 9 within the cDNA of all parasite lines was tested by PCR amplification, using the primer pairs listed in Table 2.7. No amplification was observed in the cDNA of all three PfdHHC5 transgenic clones when using primers specific for sequences within the PfdHHC5 gene (Figure 5.5A), indicating that PfdHHC5 is not expressed and the PfdHHC5 clones are indeed knock-out clones. For the cDNA of all three PfdHHC9 clones, reduced PCR amplification was observed when using primers specific for sequences within the PfdHHC9 gene (Figure 5.5B), indicating a reduced expression of PfdHHC9 protein. This suggests that the PfdHHC9 clones, although not truly knocked-out, are at least knock-down lines.

Further analysis by quantitative PCR (qPCR) was performed using the cDNA from one clone of each PfdHHC knock-out line (PfdHHC5 clone 19 and PfdHHC9 clone 3, along with wild-type 3D7 as a control). Primers and 6-carboxyfluorescein (FAM)-labelled probes that were specific for each PfdHHC gene were used in the qPCR reaction, as described in the Materials and Methods (the qPCR experiment described here was kindly performed by Leyla Bustamante, Rayner lab). The qPCR analysis revealed that the expression of the PfdHHC5 gene was completely abolished in the PfdHHC5 clone 19 transgenic line (Figure 5.5C). Meanwhile, although some expression of the PfdHHC9 gene could still be detected, expression was greatly reduced in the PfdHHC9 clone 3 transgenic line (Figure 5.5C). Taken together, this indicated that PfdHHC5 was successfully knocked-out, while PfdHHC9 was knocked down, in *P. falciparum*.

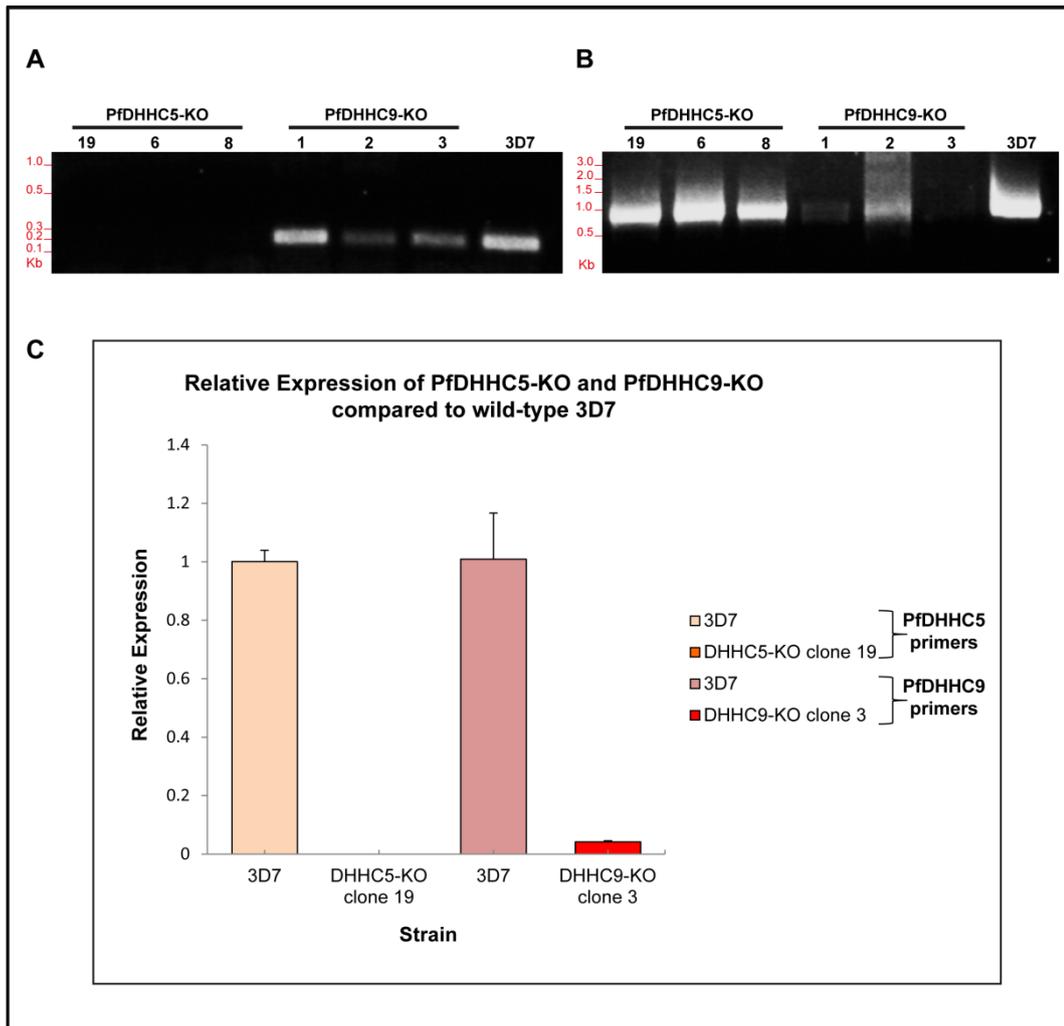


Figure 5.5: Reverse transcription (RT)-PCR analysis of PfDhHc5-KO and PfDhHc9-KO transgenic lines. cDNA was made by RT-PCR for PfDhHc5-KO clones (19, 6, and 8) and PfDhHc9-KO clones (1, 2, and 3), as well as for wild-type 3D7. Sequence specific to each PfDhHc gene was amplified by PCR using the cDNA of all clones, along with wild-type 3D7. **(A)** The primers used amplify a sequence specific to the PfDhHc5 gene. **(B)** The primers used amplify a sequence specific to the PfDhHc9 gene. **(C)** Quantitative PCR (qPCR) analysis of PfDhHc5-KO clone 19 and PfDhHc9-KO clone 3 relative to wild-type 3D7, using primers and 6-FAM-labelled probes specific to each gene.

5.2.3. Effect of the knock-out/knock-down of PfDHHC5 and 9 on parasite growth

The results described above indicate that of the six transgenic lines made, only PfDHHC5 and 9 transgenic parasite lines had integrated the knock-out constructs, and were disrupted at the genes of interest, with completely abolished expression for PfDHHC5 and reduced expression for PfDHHC9. For ease of description, both PfDHHC5 and 9 transgenic lines are now referred to as PfDHHC5-KO and PfDHHC9-KO. In order to determine the effect of disrupting these DHHC proteins on parasite growth, a growth assay was set up for both PfDHHC5-KO and PfDHHC9-KO as described in Materials and Methods. The growth assay was performed on three PfDHHC5-KO clones (6, 8, and 19) and three PfDHHC9-KO clones (1, 2, and 3), as well as on wild-type 3D7 as a control.

Briefly, the parasitemia of all knock-out parasite lines, including the wild-type 3D7 control, was measured by flow cytometry (with the kind assistance of Michel Theron, Rayner lab) every 2 days, over a period of 10 days. The growth of the PfDHHC5-KO and PfDHHC9-KO clones was then compared to that of wild-type 3D7. For both PfDHHC5-KO and PfDHHC9-KO, there was no significant difference in growth compared to wild-type 3D7 (Figure 5.6), indicating that the transgenic parasites appeared to be able to survive and grow normally in the absence of PfDHHC5 and with reduced PfDHHC9.

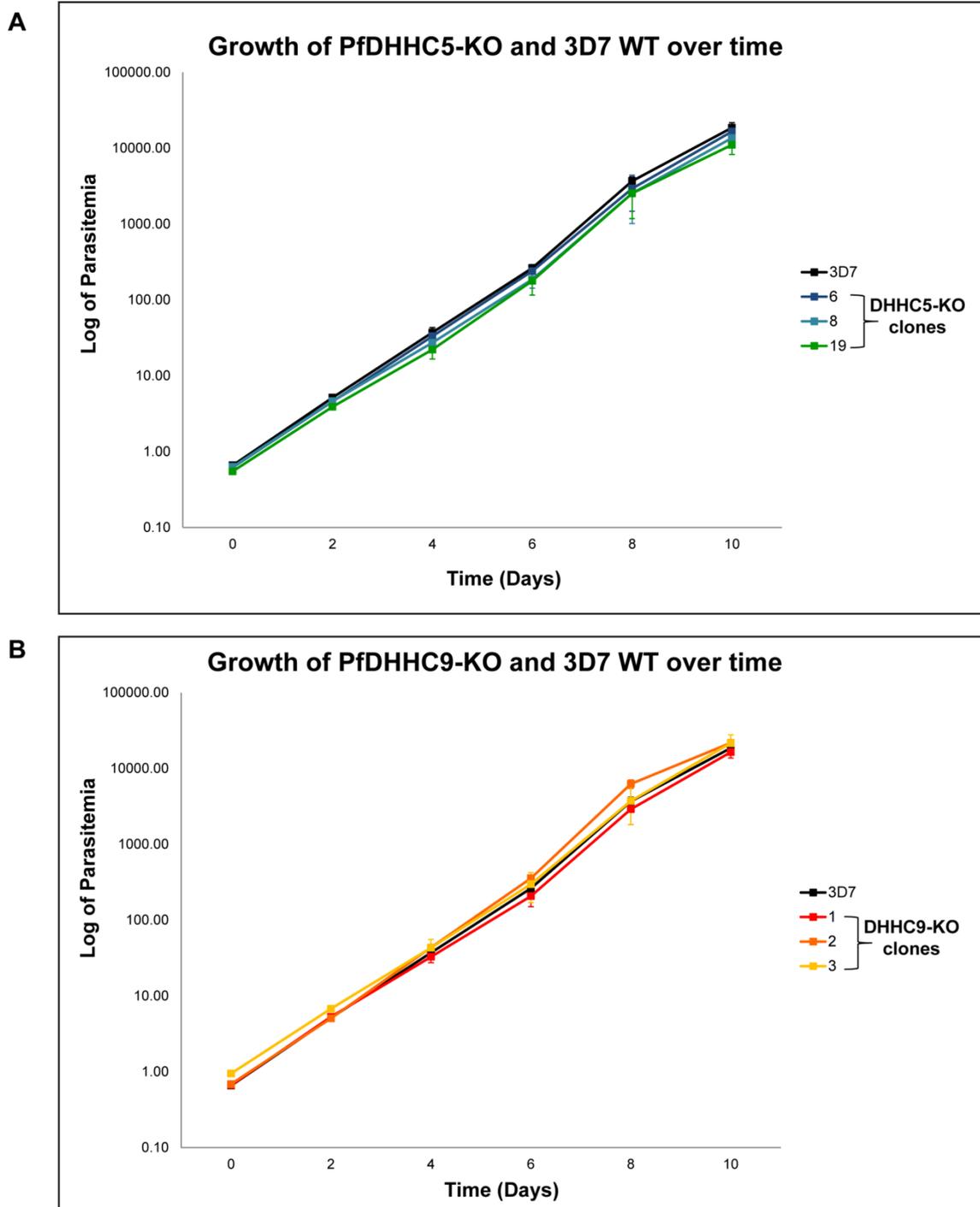


Figure 5.6: Growth assay comparing the growth of PfdHHC5-KO and PfdHHC9-KO with wild-type 3D7. (A) Parasitemia of PfdHHC5-KO clones (6, 8, and 19), along with wild-type 3D7, measured every 2 days for a total of 10 days. **(B)** Parasitemia of PfdHHC9-KO clones (1, 2, and 3) along with wild-type 3D7 measured every 2 days for a total of 10 days.

In summary, of the five *P. falciparum* DHHC-domain-containing proteins analysed here, only two proteins, namely PfDHHC5 and 9, were successfully disrupted, and disruption of these proteins appeared to have no detrimental effect on parasite growth and survival in parasite blood stages. However, it must be noted that further phenotyping of the knock-out strains generated here was unable to be performed due to time constraints. The remaining DHHC proteins were unable to be disrupted, at least in the intraerythrocytic stages studied here.

This implies that a subset of the *P. falciparum* DHHC proteins appear to be amenable to disruption in the intraerythrocytic stages, suggesting a possible functional redundancy for some of the PfDHHC proteins. This was similar to *P. berghei* DHHC proteins, where PbDHHC5 and 9 were also successfully disrupted in the blood stages (Section 4.3), just as PfDHHC5 and PfDHHC9 could be disrupted in *P. falciparum* blood stages. Conversely, some of the *P. falciparum* DHHC proteins were unable to be disrupted in the intraerythrocytic stages, suggesting that some DHHC proteins could be essential for *P. falciparum* blood-stage growth. This was again similar to *P. berghei*, where PbDHHC8 was unable to be disrupted in the blood stages (Section 4.3), just as PfDHHC8 could not be disrupted in *P. falciparum*.

Interestingly, PfDHHC3 and 7 were unable to be disrupted in *P. falciparum*, despite successful disruption in *P. berghei* (Section 4.3). This could possibly be due to the more efficient genetic manipulation techniques available in *P. berghei*, as well as superior gene targeting vectors [14], which greatly improve the efficiency of genetic modification in this parasite. On the other hand, this could also be truly a difference in gene function between the two *Plasmodium* species, again because of the difference in the total number of DHHC proteins between the two species, as well as the difference in localisation of some of the DHHC proteins.

In the case of PfMBOAT, the PfMBOAT coding gene was unsuccessfully disrupted here. Other groups have reported similar findings [1], suggesting that PfMBOAT may be essential in the intraerythrocytic stages. PfMBOAT is actually annotated on *PlasmoDB* to act as a diacylglycerol acyltransferase (DGAT), a common activity performed by members of the MBOAT family, and other work has shown that PfMBOAT does in fact exhibit DGAT activity [1]. Whether or not PfMBOAT also acts as a PAT is still unknown. However, the fact that it does play a role as a DGAT, which is an important enzyme in the biosynthetic pathway of triacylglycerol, may be the reason why this gene does not appear to be amenable to disruption.

Conclusion

This analysis of PAT proteins in *P. falciparum* has revealed similar insights to that of the study of these proteins in *P. berghei*. Firstly, as in *P. berghei*, *P. falciparum* DHHC proteins are localised to different membrane-bound sites in the parasite, including specialised parasite organelles. Secondly, some of the DHHC proteins in *P. falciparum* could possibly be essential for blood-stage growth, while other DHHC proteins appear to have functional redundancy. However, there appears to be some differences in the localisation and essentiality between some of the homologues of these DHHC proteins in the two *Plasmodium* species. Nevertheless, the fact that these DHHC proteins may play an important role in *Plasmodium* biology seems apparent. However, whether these proteins do in fact exhibit PAT activity is still unknown, and is addressed in the next chapter.

References:

1. Palacpac, N.M., et al., *Evidence that Plasmodium falciparum diacylglycerol acyltransferase is essential for intraerythrocytic proliferation*. *Biochem Biophys Res Commun*, 2004. **321**(4): p. 1062-1068.
2. Le Roch, K.G., *Discovery of Gene Function by Expression Profiling of the Malaria Parasite Life Cycle*. *Science*, 2003. **301**(5639): p. 1503-1508.
3. Abdi, A., et al., *SAM domain-dependent activity of PfTKL3, an essential tyrosine kinase-like kinase of the human malaria parasite Plasmodium falciparum*. *Cellular and Molecular Life Sciences*, 2010. **67**(19): p. 3355-3369.
4. Elmendorf, H.G. and K. Haldar, *Identification and localization of ERD2 in the malaria parasite Plasmodium falciparum: separation from sites of sphingomyelin synthesis and implications for organization of the Golgi*. *Embo J*, 1993. **12**(12): p. 4763-73.
5. Van Dooren, G.G., et al., *Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of Plasmodium falciparum*. *Mol Microbiol*, 2005. **57**(2): p. 405-419.
6. Baldi, D.L., et al., *RAP1 controls rhoptry targeting of RAP2 in the malaria parasite Plasmodium falciparum*. *Embo J*, 2000. **19**(11): p. 2435-43.
7. Langsley, G., et al., *Subcellular Location, Phosphorylation and Assembly into the Motor Complex of GAP45 during Plasmodium falciparum Schizont Development*. *PLoS One*, 2012. **7**(3): p. e33845.
8. Fidock, D.A. and T.E. Wellems, *Transformation with human dihydrofolate reductase renders malaria parasites insensitive to WR99210 but does not affect the intrinsic activity of proguanil*. *Proc Natl Acad Sci U S A*, 1997. **94**(20): p. 10931-6.
9. Maier, A.G., et al., *Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in Plasmodium falciparum for targeted gene deletion by double crossover recombination*. *Mol Biochem Parasitol*, 2006. **150**(1): p. 118-121.
10. Crabb, B.S., et al., *Transfection of the human malaria parasite Plasmodium falciparum*. *Methods Mol Biol*, 2004. **270**: p. 263-76.

11. Patzewitz, E.-M., E.H. Wong, and S. Müller, *Dissecting the role of glutathione biosynthesis in Plasmodium falciparum*. Mol Microbiol, 2012. **83**(2): p. 304-318.
12. Normark, S. and J. Schonebeck, *In vitro studies of 5-fluorocytosine resistance in Candida albicans and Torulopsis glabrata*. Antimicrob Agents Chemother, 1972. **2**(3): p. 114-21.
13. Block, E.R., A.E. Jennings, and J.E. Bennett, *5-fluorocytosine resistance in Cryptococcus neoformans*. Antimicrob Agents Chemother, 1973. **3**(6): p. 649-56.
14. Pfander, C., et al., *A scalable pipeline for highly effective genetic modification of a malaria parasite*. Nat Methods, 2011. **8**(12): p. 1078-1082.

Chapter 6:

**Protein acyltransferase
activity of *Plasmodium
falciparum* DHHC proteins**

Protein acyl transferase (PAT) activity of *P. falciparum* DHHC proteins

The results described in the previous chapters characterised the localisation and essentiality of *P. falciparum* and *P. berghei* homologues of the family of proteins –namely the DHHC protein family– known to catalyse protein palmitoylation in other eukaryotic organisms. However, whether any of these *Plasmodium* DHHC-domain-containing proteins actually have palmitoyl transferase activity has never been formally shown. In this chapter, a novel assay was developed to directly test the PAT activity of *P. falciparum* DHHC proteins characterised in Chapter 5, and was used to determine whether individual PfDHHC proteins are substrate specific.

6.1. PAT activity assay – the method

The PAT activity of several of the DHHC proteins in other eukaryotic organisms, such as yeast and humans, have been previously shown using tritiated palmitoyl-CoA in *in vitro* radiolabeling assays, or using tritiated palmitic acid in *in vivo* metabolic labelling assays. For example, in the *in vitro* radiolabeling assays, the DHHC protein of interest is mixed with the putative target protein in the presence of tritiated palmitoyl-CoA, and the PAT activity of the DHHC protein determined by measuring the incorporation of tritiated palmitate. These assays successfully demonstrated the PAT activity of DHHC proteins, such as Erf2 and Akr1 in yeast, and DHHC9 in humans [1-3].

The use of radiolabeling however is time-consuming, with blots containing tritiated proteins sometimes requiring months of exposure. Thus, for the purposes of this project, a PAT activity assay incorporating metabolic labelling and click chemistry was developed to increase the throughput and to allow for specific experimental questions to be followed in the time available. The principle of the PAT activity assay was to express a potential palmitoylated target protein in the presence and absence of a PfDHHC protein, and then to purify only palmitoylated versions of the target protein using metabolic labelling and click chemistry methods (described in Section 1.6.2). The activity of the PfDHHC protein of interest could then be assessed by comparing the level of target protein palmitoylation in the presence and absence of the PfDHHC protein. If, palmitoylation of the target protein only occurred in the presence of the PfDHHC protein, this would demonstrate the PAT activity of said PfDHHC protein. Comparison of palmitoylation levels in the presence of different PfDHHCs could then be used to test whether the PAT activity of a particular PfDHHC protein is substrate specific.

The PAT activity assay was developed using the Human Embryonic Kidney 293 (HEK293) cell expression system, which has been used previously for the successful production of recombinant *P. falciparum* proteins [4]. Potential *P. falciparum* target proteins were codon-optimised for expression in mammalian cells and C-terminally-tagged with a c-Myc tag, to allow their detection. These constructs were then expressed in HEK293 cells by transient transfection, and co-expressed with either a putative PfDHHC –also codon-optimised for expression in mammalian cells but C-terminally-tagged with a FLAG tag for detection- or with an empty CD4 control vector which does not express any PfDHHC. After 18 hours post-transfection, the cells were then metabolically labelled for 6 hours with the palmitic acid analogue, 17-ODYA, at a concentration of 25 μ M, or mock-treated with DMSO as a control. After 6 hours of labelling with 17-ODYA, proteins were extracted from the HEK293 cells (using a Triton X-100-based buffer) and then reacted with biotin-azide under standard click chemistry reactions (described in Section 1.6.2 and in the Materials and Methods), resulting in the biotinylation of all 17-ODYA-labelled proteins. Biotinylated proteins were then streptavidin affinity purified, and the ‘palmitoylation status’ of the proteins was determined by their enrichment in the 17-ODYA-treated sample compared to the DMSO-treated control sample. The presence of the target protein in the 17-ODYA-treated sample was specifically determined using the C-terminal c-Myc tag. As the PfDHHC is tagged with a different tag, autopalmitoylation of the PfDHHC could also potentially be determined by looking for the presence of the PfDHHC in the 17-ODYA-treated sample using its C-terminal FLAG tag. Expression of both *P. falciparum* target proteins and PfDHHC proteins in HEK293 cells could also be confirmed by immunoblotting of whole protein extracts and by immunofluorescence assay using the c-Myc and FLAG tags. A schematic detailing the steps of this assay is shown in Figure 6.1.

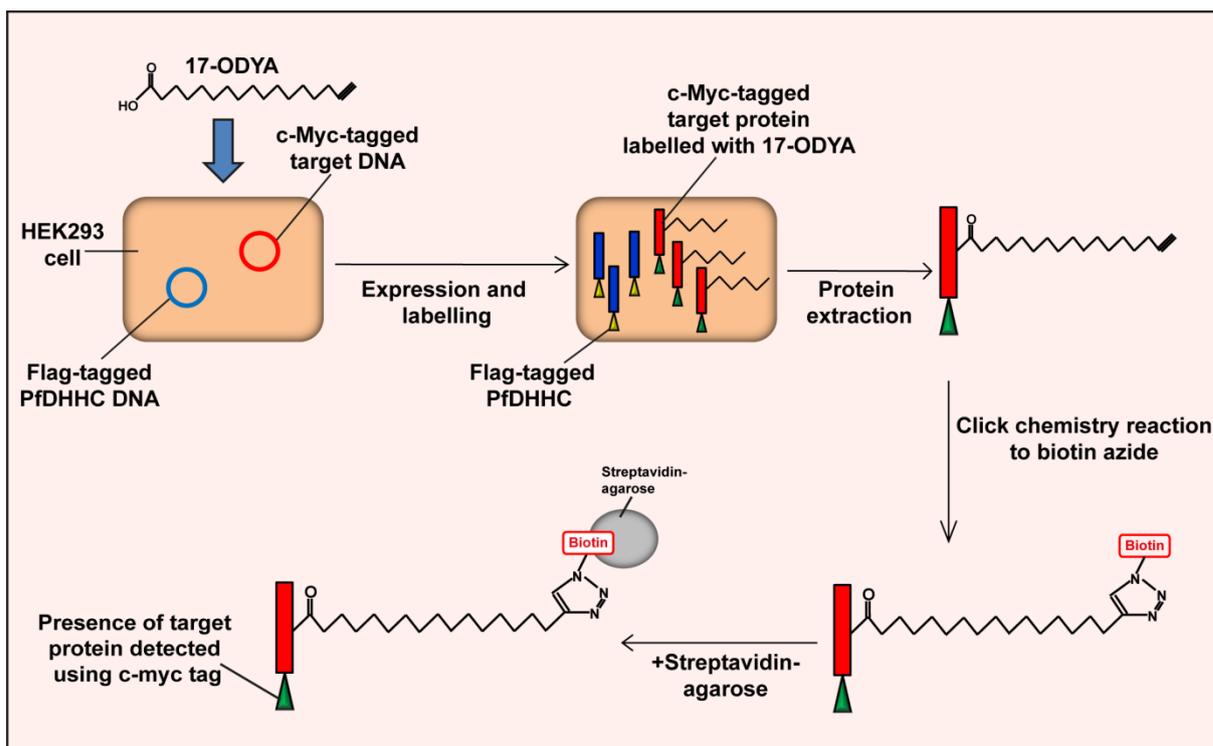


Figure 6.1: PAT activity assay. This assay incorporates the expression of codon-optimised *P. falciparum* proteins in a mammalian cell expression system along with metabolic labelling and click chemistry methods of palmitoyl-protein purification. HEK293 cells were co-transfected with FLAG-tagged PfDdHC DNA along with the c-Myc-tagged DNA of a potential target palmitoyl-protein, both of which were codon-optimised for expression in human cells. The HEK293 cells were then treated with the metabolic label, 17-ODYA. Proteins were extracted and reacted to biotin-azide under standard click chemistry conditions, resulting in the biotinylation of 17-ODYA-labelled proteins. Biotinylated proteins were affinity-purified using streptavidin-agarose and the presence of the target protein detected using antibodies against the c-Myc tag. As a standard control, transfected cells were also mock-labelled with DMSO instead of 17-ODYA. As a control for the background palmitoylation of target proteins by endogenous HEK293 PAT proteins, cells were also co-transfected with the c-Myc-tagged target protein and an empty CD4 control vector, in the absence of the PfDdHC. The level of target protein palmitoylation when in the presence of the PfDdHC compared to when in the absence of the PfDdHC was expected to be an indication of whether the particular PfDdHC was responsible for the palmitoylation of the target protein.

This assay was adapted from work described in [5], where tagged versions of human DdHC proteins of interest were co-expressed in mammalian cells, along with putative target proteins tagged with a different epitope tag, and metabolically labelled with 17-ODYA. In the previous assay described in [5], the target proteins were immunoprecipitated using their epitope tags and the click chemistry reaction was performed on the immunoprecipitates using rhodamine azide, which allows the visualisation of labelled proteins by in-gel fluorescence. Due to the lack of facilities for in-gel fluorescence, the PAT activity assay developed here was performed using biotin-azide in the click

chemistry reaction instead. The labelled proteins were then purified by streptavidin affinity purification and the presence of the proteins visualised using their C-terminal epitope tags. Additionally, click chemistry reactions were performed on total protein extracts rather than immunoprecipitates.

The effect of the presence of the putative PfDHHHC protein on the palmitoylation of the target protein was determined by comparing the level of target protein palmitoylation in the presence, as well as in the absence of a given PfDHHHC (that is, when the target protein was co-expressed with only the empty CD4 control vector). Due to the potential overlapping functionality of the DHHHC-PATs, and how little is currently known about their substrate specificity in any eukaryotic species, some background palmitoylation of the target protein could occur due to endogenous PATs present in HEK293 cells. However, it was expected that if the PfDHHHC of interest was responsible for palmitoylating the target protein, the amount of palmitoylated target protein would be greater when in presence of the PfDHHHC, compared to when in the absence of the PfDHHHC.

6.2. Expression and localisation of *P. falciparum* proteins in HEK293 cells

6.2.1. Expression of *P. falciparum* proteins in HEK293 cells

All five PfDHHHC proteins which were studied in Chapter 5 –PfDHHHC3, 5, 7, 8 and 9- were codon-optimised for expression in HEK293 cells and introduced into HEK293 expression vectors [6], along with sequence coding for a C-terminal FLAG tag in order to allow the detection of the PfDHHHC proteins, as described in the Materials and Methods. The HEK293 expression vectors used here also contained immunoglobulin-like domains 3 and 4 of rat CD4 (thought to aid folding of heterologous proteins and usually also used as an epitope tag), which would normally be present following the open reading frame of the PfDHHHC proteins [7]. However, as CD4 was not required for the purposes of this assay, a STOP codon was introduced downstream of the C-terminal FLAG tag and upstream of CD4. Thus the recombinant FLAG-tagged PfDHHHC proteins were not expected to contain CD4.

A number of potential *P. falciparum* target proteins were considered based on their localisation in relation to the localisation of the PfDHHHC proteins (established in Chapter 5) and on the identification of putative palmitoylated cysteines within the target proteins (described in Chapter 3). Due to time constraints, only two potential target proteins –PfSec22 (PFC0890w) and Armadillo Repeats-Only (ARO) (PFD0720w) - were studied here. PfSec22 is annotated as a SNARE protein containing a single TM domain in *PlasmoDB* (<http://plasmodb.org>). The site-ID palmitome

purification experiments (described in Chapter 3) identified a peptide from the N-terminal region of PfSec22, which contained two cysteine residues, Cys2 and Cys8, either or both of which could potentially be palmitoylated. PfSec22 has been found to be involved in ER to Golgi transport in *P. falciparum* and appears to be localised predominantly at the ER [8]. PfSec22 could therefore potentially be palmitoylated by PfDHHC5, which also localises to the ER, or PfDHHC3, which localises to the Golgi. PfARO is a dually-acylated protein which localises to the rhoptries, where it is attached to the membrane of the rhoptries by its myristate and palmitate anchors [9]. Although PfARO was not identified in the site-ID palmitome purification experiments (Chapter 3), this protein has been experimentally shown to be palmitoylated in *P. falciparum*, and either or both of two cysteines located within the first 20 N-terminal amino acids, Cys5 and Cys6, are thought to be the potential palmitoylation sites [9]. As PfARO appears to be targeted to the rhoptries, it could thus potentially be palmitoylated by PfDHHC7, which is also located at the rhoptries, or by PfDHHC3, which PfARO could potentially encounter in the Golgi during transport to the rhoptries. These two potential target proteins were also codon-optimised for expression in HEK293 cells and introduced, along with a C-terminal c-Myc tag, followed by a STOP codon, into the same HEK293 CD4-containing expression vector. An 'empty' expression vector containing only CD4 was used as the negative vector control for the target proteins.

6.2.2. Subcellular localisation of *P. falciparum* proteins expressed in HEK293 cells

In order to confirm that the *P. falciparum* proteins were expressed properly in HEK293 cells, as well as to determine where the parasite proteins localised to in these heterologous mammalian cells, immunofluorescence microscopy was performed on HEK293 cells transiently transfected with each of the proteins of interest, using antibodies against the FLAG tag and c-Myc tag respectively.

(1) Subcellular localisation of FLAG-tagged PfDHHC proteins expressed in HEK293 cells

All five PfDHHC proteins of interest (PfDHHC3, 5, 7, 8 and 9) were transfected into HEK293 cells using polyethylenimine (PEI). After 24 hours with the transfection mixture, cells were fixed in 4% formaldehyde and immunofluorescence staining was performed using antibodies against the FLAG tag. Fluorescence staining was detected for four of the PfDHHCs (PfDHHC3, 5, 7, and 9) (Figure 6.2A), indicating that these four parasite proteins were expressed in mammalian cells. Immunofluorescence signal was not detected for PfDHHC8. However, PfDHHC8 was also not detectable by both immunoblot and immunofluorescence assay when tagged in *P. falciparum* (Chapter 5). This, along with the lack of detection in HEK293 cells,

suggests that this protein may be expressed at levels that are too low for detection, even in mammalian cells, or else issues with protein solubility preclude its detection using the current extraction procedures.

The subcellular localisations of the four PfDHC proteins were determined by comparing their immunofluorescence signals to that of two mammalian proteins with established localisations: calnexin (ER marker) [10] and cadherin (plasma membrane marker) [11] (Figure 6.2A). It must be noted however, that although the cadherins are integral membrane proteins found at the plasma membrane, they are mainly required for cell-cell adhesion by forming adherens junctions between cells. In these experiments, mainly single cells with no cell-cell junctions were studied, and thus membrane localisation of the cadherins may not be as clear. Nevertheless, staining of the HEK293 cells with antibodies against cadherin still displayed a mainly peripheral, and therefore presumably plasma membrane signal. A low level of cytoplasmic staining was also detected, which may be due to the presence of soluble cadherins, and such staining has been described by the antibody manufacturer (Abcam).

In HEK293 cells, PfDHC3 staining co-localised with that of calnexin, suggesting an ER localisation (Figure 6.2A). In *P. falciparum*, this DHC protein is targeted to the Golgi (Section 5.1.3). However, it appeared that when expressed in mammalian cells, this protein localised instead to the ER, perhaps because as a parasite protein, it may lack the mammalian exit signal that would direct it to the Golgi in mammalian cells.

PfDHC5 staining did not co-localise with either calnexin or cadherin, indicating that this protein was not localised to the ER or the plasma membrane in HEK293 cells (Figure 6.2A). PfDHC5 was localised to the ER in *P. falciparum* (Section 5.1.3). However, PfDHC5 contains ankyrin repeats in its N-terminus, and other N-terminal ankyrin repeat-containing DHC proteins, such as DHC17 in humans [12] and Akr1 in yeast [2], have been localised to the Golgi. Thus, it would perhaps not be unexpected if PfDHC5 were to be targeted to the Golgi when expressed in a mammalian cell. Although immunofluorescence co-staining was performed with antibodies against two different Golgi-localised proteins (syntaxin-6 and 58K Golgi protein) both these antibodies did not produce good Golgi immunofluorescence staining and unfortunately, the localisation of PfDHC5 to the Golgi cannot be proven at this time.

PfDHH7 co-localised with cadherin staining in HEK293 cells, indicating a plasma membrane localisation (Figure 6.2A). In *P. falciparum*, PfDHH7 was localised to the rhoptries (Section 5.1.3), specialised organelles at the apical end of the parasite, close to the cell periphery. As no such organelle exists in mammalian cells, localisation to the plasma membrane instead is not unexpected.

Lastly, PfDHH9 staining co-localised with that of calnexin (Figure 6.2), indicating that PfDHH9 was localised to the ER in mammalian cells. In *P. falciparum* (Section 5.1.3) PfDHH9 was localised to the IMC, another specialised parasite organelle consisting of flattened membrane-bound vesicles close to the plasma membrane. Although not much is known about the formation of the IMC, some proteins which associate with the IMC in *P. falciparum* have been shown to be trafficked to the IMC via the ER [13]. As the IMC does not exist in mammalian cells, it could be that in HEK293 cells, PfDHH9 is retained in the ER instead.

(2) Subcellular localisation of c-Myc-tagged PfARO and PfSec22 in HEK293 cells

The two potential target proteins, PfARO and PfSec22, were transfected into HEK293 cells respectively using polyethylenimine (PEI). After 24 hours, cells were fixed in 4% formaldehyde and immunofluorescence staining was performed using antibodies against the c-Myc tag. Immunofluorescence staining was successfully detected for both proteins (Figure 6.2B), indicating that both these proteins were expressed in mammalian cells. The subcellular localisations of these two proteins were determined by comparing their immunofluorescence signals to that of two mammalian proteins with established localisations: calnexin (ER marker) [10] and cadherin (plasma membrane marker) [11].

PfARO staining appeared to be mainly cytoplasmic, with some partial co-localisation with cadherin at the cell periphery (Figure 6.2B), indicating that although mainly localised to the cytoplasm, some PfARO proteins were also localising to the plasma membrane. In *P. falciparum*, ARO is localised to the rhoptries, where it associates with the cytosolic face of the rhoptry membrane, most likely via the myristate and palmitate attachments on its N-terminal domain. Removal of either of these acylation motifs causes ARO localisation to become cytoplasmic [9]. As this protein has no TM-domains, and is being expressed in mammalian cells without its endogenous DHH protein, the mainly cytoplasmic localisation suggests that the majority of PfARO is not palmitoylated, while the fraction which is localised to the plasma membrane could be due to background palmitoylation by endogenous HEK293 PATs.

PfSec22 co-localised with calnexin staining, indicating that this protein was localised to the ER in HEK293 cells (Figure 6.2B). PfSec22 is annotated on *PlasmoDB* as a SNARE protein, and has been found to be predominantly localised in the ER in *P. falciparum* [8]. Both yeast and human versions of Sec22 are involved in ER to Golgi transport, and are also mainly found in the ER membrane [14]. A protein-protein BLAST (blastp) search using the PfSec22 amino acid sequence revealed some homology to both yeast and human Sec22 (identities of approximately 30%). Thus, the localisation of PfSec22 to the ER membrane of HEK293 cells is expected.

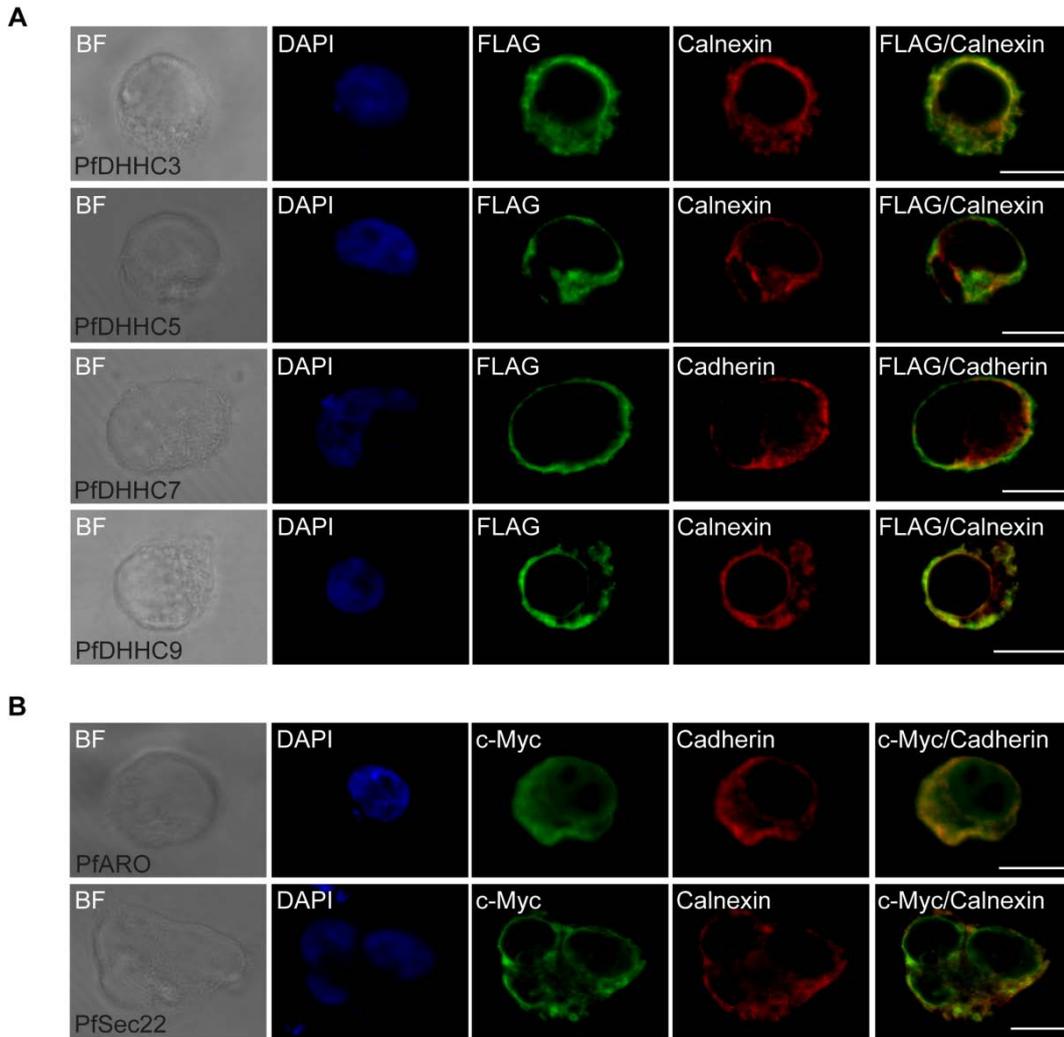


Figure 6.2: Localisation of codon-optimised *P. falciparum* proteins in HEK293 cells. (A) Codon-optimised *P. falciparum* DHHC proteins (PfDHHC3, 5, 7, and 9) were expressed in HEK293 cells and expression of the proteins determined by immunofluorescence assay using antibodies against the FLAG tag. Localisation of the PfDHHC proteins (green) was determined by comparing the immunofluorescence signal to that of the following known localisation markers (red): calnexin (endoplasmic reticulum marker) and cadherin (plasma membrane marker). Nuclear staining by DAPI is shown in blue. **(B)** Codon-optimised *P. falciparum* proteins PfARO and Pfsec22 were expressed in HEK293 cells. Expression was determined by immunofluorescence assay using antibodies against the c-Myc tag. Both PfARO and PfSec22 (green) were localised by immunofluorescence against the following mammalian localisation markers (red): calnexin (endoplasmic reticulum marker) and cadherin (plasma membrane marker). Nuclear staining by DAPI is shown in blue. Scale bar: 10 μ m.

In summary, all the *P. falciparum* proteins of interest, with the exception of PfDHHC8, were successfully expressed using the HEK293 mammalian cell expression system, albeit with quite different intracellular localisations compared to their endogenous localisations in *P. falciparum*, highlighting a major confounder of heterologous expression systems. PfDHHC3 and 9 localised to the ER, PfDHHC7 localised to the plasma membrane and PfDHHC5 may be potentially localised to the Golgi. Meanwhile, the two potential target proteins, PfARO and PfSec22, were also successfully expressed using the HEK293 cell expression system. PfARO displayed a mostly cytoplasmic localisation, with some localisation at the plasma membrane, while PfSec22 was localised to the ER.

6.2.3. Immunoprecipitation of c-Myc-tagged PfSec22 and PfARO in the presence of PfDHHC proteins

In order to determine whether the *P. falciparum* potential target proteins could be visualised by immunoblot using antibodies against the c-Myc tag, PfSec22 and PfARO were expressed in HEK293 cells, each either in the presence of one of the PfDHHC proteins (PfDHHC3, 5, 7, 8 and 9), or in the presence of the empty CD4 control vector. PfSec22 and PfARO were then immunoprecipitated from the protein extracts using antibodies against the c-Myc tag. The immunoprecipitates were separated by SDS-PAGE and the presence of PfSec22 and PfARO was determined by immunoblot, using antibodies against the c-Myc tag from a different species.

PfSec22 is predicted to be 26 kDa, and previously published immunoblots of *P. falciparum* protein extracts with polyclonal antibodies raised against recombinant PfSec22 produced bands running at approximately this size [8]. However, this analysis of recombinant PfSec22 immunoprecipitated from HEK293 cells, using α -c-Myc tag antibodies, resulted in two clear bands: a stronger lower molecular weight band at approximately 10 kDa and a weaker higher molecular weight band at approximately 17 kDa (Figure 6.3). This discrepancy in apparent molecular weight could be due to the fact that this parasite protein is being expressed recombinantly in mammalian cells, where the protein folding and processing may differ from that of parasite cells, thus resulting in the aberrant migration of the protein. Additionally, post-translational modifications of the protein which occur in the parasite might not occur when the protein is expressed in mammalian cells and *vice versa*, thus causing a difference in the expected molecular weight. Alternatively, the difference in apparent molecular weight could be due to aberrant migration because of the presence of coiled-coil structure in the PfSec22 SNARE protein.

The presence of two clear bands raises the question as to whether the higher molecular weight band corresponds to the palmitoylated version of PfSec22. Although protein palmitoylation does not

consistently result in a shift in molecular weight, this has been shown to occur in some cases. For example, immunoprecipitation of the palmitoylated *P. falciparum* invasion motor complex protein, GAP45, from *P. falciparum* schizonts, produces two bands, and only the higher molecular weight band of the doublet corresponds to palmitoylated GAP45 [15]. Interestingly, the intensity of the higher molecular weight band of PfSec22 immunoprecipitated from HEK293 cells was strongest when PfSec22 was co-expressed with PfDHH5, and the higher molecular weight band was readily detected when PfSec22 was co-expressed with PfDHH3 and 7. In contrast, the higher molecular weight band was only weakly detected when PfSec22 was co-expressed with PfDHH8 and 9, and the empty CD4 control vector. If indeed the higher molecular weight band is palmitoylated PfSec22, this difference in intensity of the band may be an indication of which PfDHH is responsible for palmitoylating PfSec22.

PfARO is predicted to be 30 kDa, and previous work involving the immunoblot of *P. falciparum* schizonts, using antibodies against PfARO, has produced bands migrating at the expected size [9]. Here, PfARO immunoprecipitated from HEK293 cells using α -c-Myc antibodies revealed two clear bands: a lower molecular weight band at approximately 28 kDa and a higher molecular weight band at approximately 45 kDa. Additionally a faint third band is observed above the 46 kDa marker (Figure 6.3). The discrepancy in the molecular weight of the 28 kDa band from the predicted molecular weight of PfARO is much smaller than the discrepancy of PfSec22 above, but could also be due to differences in protein folding or processing when expressed recombinantly in a mammalian cell line. As with PfSec22, the higher molecular weight band (at approximately 45 kDa) could be due to palmitoylation of PfARO. The intensity of the 45 kDa band was strongest when PfARO was co-expressed with PfDHH3 and 5, and the intensity of the band was weaker when co-expressed with PfDHH7 and 9. The 45 kDa band was weakest when PfARO was co-expressed with PfDHH8 and with the empty vector.

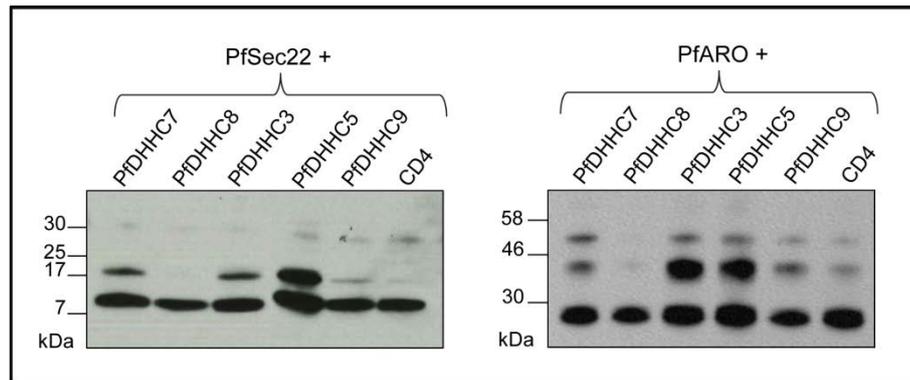


Figure 6.3: Immunoprecipitation of PfSec22 and PfARO proteins using antibodies against the c-Myc tag. HEK293 cells were co-transfected with plasmids coding for the expression of c-Myc-tagged PfSec22 or PfARO, along with the indicated FLAG-tagged PfDHC proteins (PfDHC3, 5, 7, 8 and 9) or the empty control vector (CD4). Pfsec22 and PfARO were immunoprecipitated from cell lysates using an α -c-Myc antibody. The proteins were separated by SDS-PAGE and visualised by immunoblot, using α -c-Myc antibody from a different species.

In summary, both PfSec22 and PfARO could be successfully expressed in HEK293 cells and immunoprecipitation of both these proteins resulted in two distinct bands. For both these proteins, the higher molecular weight band differed in intensity depending on which PfDHC it was co-expressed with, which could suggest specificity in their palmitoylation. In order to directly test whether these higher molecular weight bands were actually the palmitoylated versions of the target proteins, 17-ODYA metabolic labelling and click chemistry reactions were carried out on co-transfected cells in order to perform the PAT activity assay described above.

6.3. PAT activity assay of PfDHC proteins with PfSec22 and PfARO

The PAT activity assay (as described above) was performed for both PfSec22 and PfARO, co-transfected with each of the four PfDHC proteins that could be expressed in HEK293 cells (PfDHC3, 5, 7 and 9), or a CD4 empty vector control. As the expression of PfDHC8 was unable to be detected in HEK293 cells by immunofluorescence microscopy, suggesting low expression of the protein, any difference observed in the palmitoylation of the target protein could thus also be due to the lack of expression of PfDHC8, rather than the fact that PfDHC8 did not palmitoylate the particular target. Thus, PfDHC8 was not used for any further analysis.

6.3.1. PAT activity assay of PfDHHC proteins with PfSec22

PfSec22 was co-expressed in HEK293 cells in the presence of each of the four PfDHHCs (PfDHHC3, 5, 7 and 9), or with only the CD4 empty vector control. Eighteen hours post-transfection, cells were metabolically labelled with 17-ODYA, or treated with DMSO as a control, for a further six hours, after which cells were lysed for protein extraction. An aliquot of each whole cell lysate was kept for confirmation of protein expression. The PAT activity assay was performed on the remaining cell lysates as described above and in the Materials and Methods. Both the initial cell lysates and the PAT activity assay click chemistry elutions were separated by SDS-PAGE and the presence of PfSec22 was specifically determined by immunoblot using antibodies against the c-Myc tag.

Immunoblot of the initial cell lysates indicated the presence of the two bands observed previously (Section 6.2.3): a lower molecular weight band at approximately 10 kDa and a higher molecular weight band at approximately 17kDa (Figure 6.4). However, immunoblot of the elutions after the click chemistry reaction was performed resulted in a single band running at approximately 17 kDa, and this band was present only in samples treated with 17-ODYA and not in the samples treated with DMSO (Figure 6.4). The detection of this band only in 17-ODYA-treated samples and not in DMSO-treated samples confirmed that the 17 kDa band corresponds to palmitoylated PfSec22. The molecular weight of the palmitoylated PfSec22 band corresponded exactly with the higher molecular weight band observed in the initial cell lysate and in the immunoprecipitation analysis of PfSec22 described above (Section 6.2.3 and Figure 6.3). Therefore, this indicated that the higher molecular weight band visible in immunoblots and immunoprecipitates was the palmitoylated version of PfSec22, and the absence of any lower molecular weight band in the 17-ODYA-labelled samples confirms that the lower molecular weight 10 kDa band was non-palmitoylated PfSec22.

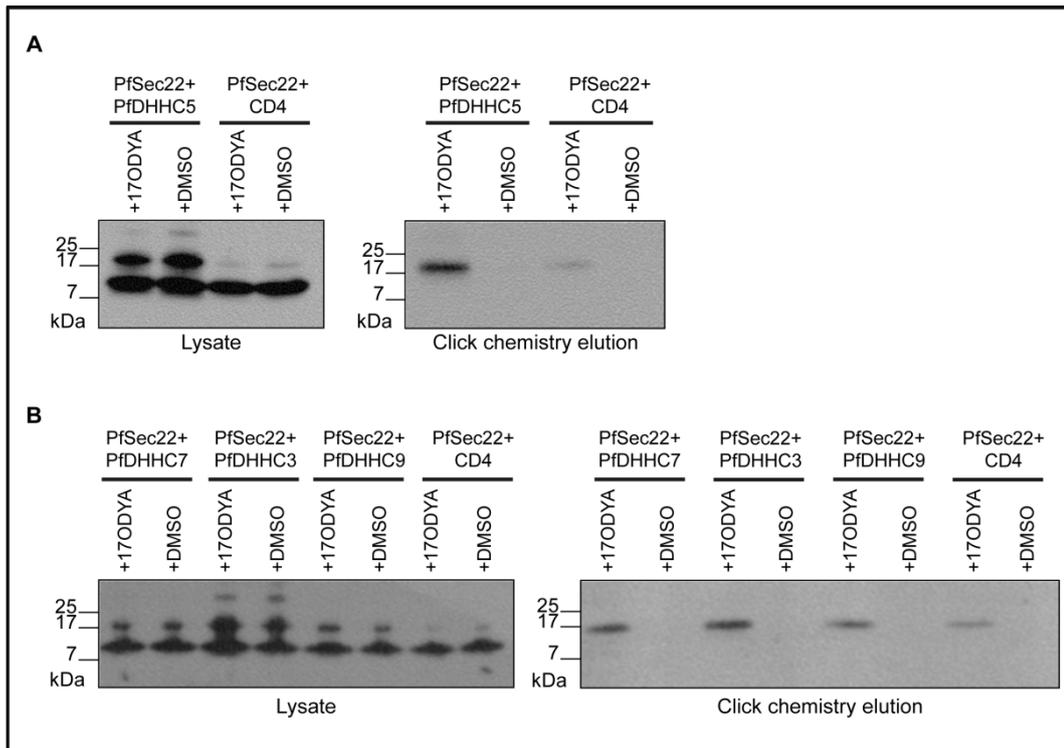


Figure 6.4: PAT activity assay demonstrating the PAT activity of PfDHC proteins on the target protein, PfSec22. HEK293 cells were co-transfected with plasmids expressing c-Myc-tagged PfSec22, along with the indicated FLAG-tagged PfDHC protein or the empty vector control (CD4). Cells were either treated with the metabolic label, 17-ODYA, or mock-treated with DMSO. Proteins were extracted and an aliquot of each lysate kept aside to confirm protein expression. The remaining lysates were put through click chemistry reactions to biotin-azide, which resulted in the biotinylation of 17-ODYA-labelled proteins. These biotinylated proteins were subsequently affinity purified by streptavidin-agarose and eluted by boiling in SDS. Samples from the initial whole cell lysates and the click chemistry elutions were separated by SDS-PAGE and the presence of c-Myc-tagged PfSec22 in each of the samples was observed by immunoblot using antibodies against the c-Myc tag. **(A)** HEK293 cells co-transfected with c-Myc-PfSec22 + FLAG-PfDHC5 or c-Myc-PfSec22 + CD4 empty vector. **(B)** HEK293 cells co-transfected with c-Myc-PfSec22 + FLAG-PfDHC7, c-Myc-PfSec22 + FLAG-PfDHC3, c-Myc-PfSec22 + FLAG-PfDHC9 or c-Myc-PfSec22 + CD4 empty vector.

The intensity of the PfSec22 palmitoylated band in the click chemistry elutions was clearly greater when PfSec22 was co-expressed with PfDHC5 than when PfSec22 was co-expressed with the empty vector (Figure 6.4A). This indicates that while there may be some background level of PfSec22 palmitoylation by endogenous HEK293 PATs, the co-expression of PfSec22 with PfDHC5 significantly increased the palmitoylation of PfSec22. This is the first direct proof that a *P. falciparum* DHC protein has PAT activity.

PfDHC3, 7 and 9 all similarly increased the intensity of the palmitoylated PfSec22 band, compared to when PfSec22 was co-transfected with the empty control vector (Figure 6.4B). This suggests that

all four PfDHHCs tested can palmitoylate PfSec22 and demonstrates the PAT activity of these 4 PfDHHC proteins.

While the ability to palmitoylate PfSec22 was clearly not restricted to a single PfDHHC, the increase in the intensity of the PfSec22 palmitoylated band in the click chemistry elutions above the intensity of the band due to background palmitoylation appeared to be greatest when PfSec22 was co-expressed with PfDHHC5 (Figure 6.4A). However, it must be noted that these click chemistry elutions are not strictly quantitative, as the intensity of the band is also affected by the efficiency of the metabolic labelling. Additionally, each lane is protein material taken from different transfections, and each transfection was individually streptavidin affinity purified, which may lead to differences in the amount of protein present. These differences cannot be completely discounted when a difference in band intensity is observed. Furthermore, the metabolic label 17-ODYA is actually an inhibitor of cytochrome P450 metabolism [16]. Cells treated with 17-ODYA are thus generally less healthy, and treatment with 17-ODYA appears have a slight effect on cell growth and protein expression, especially if treatment is excessive. These effects on cell growth and protein expression can affect the amount of protein produced for the PAT assay and can thus also contribute to differences in band intensity.

Nevertheless, even in the click chemistry elutions, the intensity of the PfSec22 palmitoylated band was consistently greater when PfSec22 was in the presence of the PfDHHC proteins, indicating that these DHHC proteins were indeed palmitoylating PfSec22. Additionally the difference in the intensities of the higher molecular weight palmitoylated band of PfSec22, in the presence of the different DHHC proteins, was also previously observed when PfSec22 was immunoprecipitated and analysed by immunoblot (Figure 6.3), as described in Section 6.2.3. In this case, as the cells were not treated with either 17-ODYA or DMSO, the difference in band intensity can be assumed to be mainly due to the presence of the particular PfDHHC protein.

Taken together, the results from the immunoprecipitation of PfSec22 and the metabolic labelling and click chemistry reaction suggest that of the PfDHHCs tested, the palmitoylation of PfSec22 appears to be greater in the presence of PfDHHC5, followed by PfDHHC3 and 7. Although PfDHHC9 also appears to be able to palmitoylate PfSec22, the amount of palmitoylation seems very low, and PfDHHC9 may not normally palmitoylate PfSec22 in the parasite. The palmitoylation of PfSec22 by 3 different PfDHHC proteins is not surprising given that DHHC-PATs are known to have overlapping functionalities in other organisms and may act redundantly by palmitoylating similar substrates [17].

The palmitoylation of PfSec22 by PfdDHC5 (which was shown to localise to the ER in *P. falciparum*, as described in Chapter 5) and PfdDHC3 (which was shown to localise to the Golgi in *P. falciparum*, as described in Chapter 5) is consistent with the fact that PfSec22 is a SNARE protein involved in transport between the ER and Golgi, and is mainly localised in the ER. Palmitoylation of PfSec22 by PfdDHC7 is more unexpected as PfdDHC7 is localised to the rhoptries in *P. falciparum*. However, due to their different localisations, PfSec22 would probably not come in contact with PfdDHC7. Thus, this may be an example of how the particular localisation of the DHHC protein might in part regulate the target proteins that it palmitoylates.

6.3.2. PAT activity assay of PfdDHC proteins with PfARO

PfARO was co-expressed in HEK293 cells in the presence of each of the four PfdDHCs (PfdDHC3, 5, 7, and 9), or with only the CD4 empty vector control. Eighteen hours post-transfection, cells were metabolically labelled with 17-ODYA or treated with DMSO as a control, for a further 6 hours, after which cells were lysed for protein extraction. An aliquot of each whole cell lysate was kept for confirmation of protein expression. The PAT activity assay was performed on the remaining cell lysates, as described above and in the Materials and Methods. Both the initial cell lysates and the PAT assay click chemistry elutions were separated by SDS-PAGE, and the presence of PfARO was specifically determined by immunoblot using antibodies against the c-Myc tag.

Immunoblot of the initial cell lysates indicated the presence of the same bands observed previously for PfARO (Section 6.2.3): a lower molecular weight band at approximately 28 kDa and a higher molecular weight band at approximately 45 kDa, along with a faint third band just above the 46 kDa marker (Figure 6.5). However, immunoblot of the elutions after click chemistry was performed resulted in a single band running at approximately 45 kDa, and this band was present only in samples treated with 17-ODYA and not in the samples treated with DMSO (Figure 6.5). The detection of this band only in 17-ODYA-treated samples and not in DMSO-treated samples indicated that this band corresponded to palmitoylated PfARO. As the apparent molecular weight of this band was approximately 45 kDa, this band corresponded to the 45 kDa molecular weight band observed in the initial cell lysate and in the immunoprecipitation analysis of PfARO (Figure 6.3). Therefore, this indicated that this 45 kDa band was the palmitoylated version of PfARO, and the lower molecular weight band at approximately 28 kDa must then correspond to non-palmitoylated PfARO. The faint third band above the 46 kDa marker was not observed in the click chemistry elutions, indicating that it is not likely to be corresponding to a doubly- or multiply-palmitoylated version of PfARO. This third

band could instead perhaps correspond to the formation of a dimer, or perhaps to a version of PfARO which was modified by a different PTM, such as phosphorylation, which could also cause a shift in the apparent molecular weight.

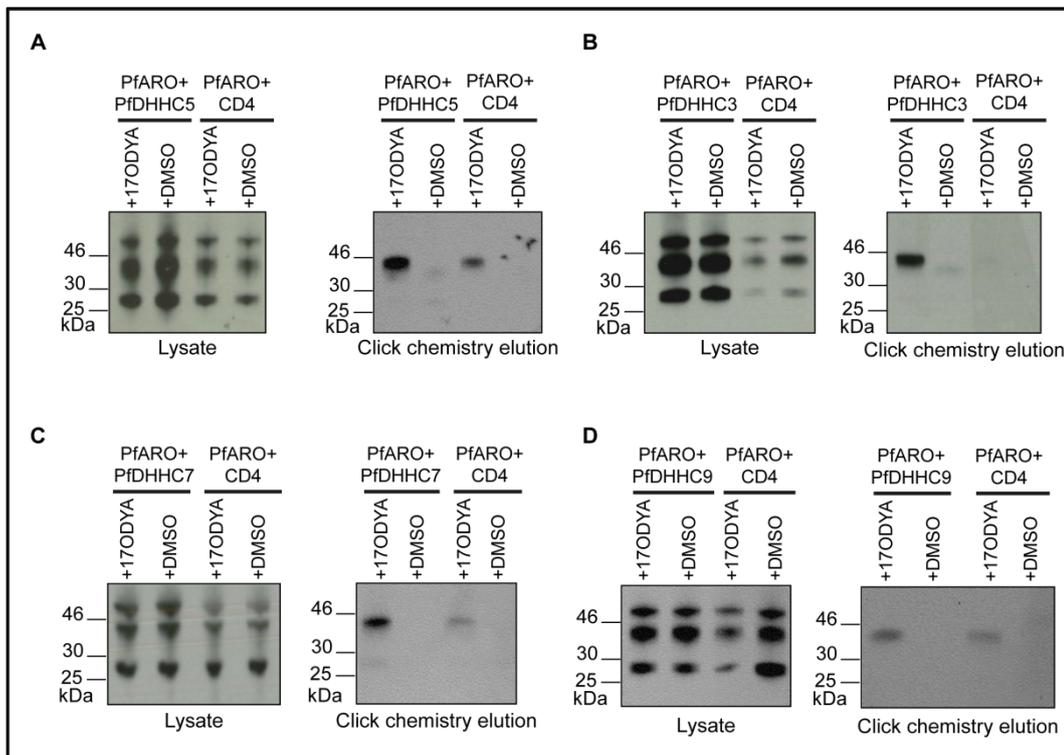


Figure 6.5: PAT activity assay demonstrating the PAT activity of PfDHC proteins on the target protein, PfARO. HEK293 cells were co-transfected with plasmids expressing c-Myc-tagged PfARO, along with the indicated FLAG-tagged PfDHC protein or the empty vector control (CD4). Cells were either treated with the metabolic label, 17-ODYA, or mock-treated with DMSO. Proteins were extracted and an aliquot of each lysate kept aside to confirm protein expression. The remaining lysates were put through click chemistry reactions to biotin-azide, which resulted in the biotinylation of 17-ODYA-labelled proteins. Biotinylated proteins were subsequently affinity purified by streptavidin-agarose and eluted by boiling in SDS. Samples from the initial lysates and the click chemistry elutions were separated by SDS-PAGE and the presence of c-Myc-tagged PfARO in each of the samples was observed by Western blot using antibodies against the c-Myc-tag. **(A)** HEK293 cells co-transfected with c-Myc-PfARO + FLAG-PfDHC5 or c-Myc-PfARO + CD4 empty vector. **(B)** HEK293 cells co-transfected with c-Myc-PfARO + FLAG-PfDHC3 or c-Myc-PfARO + CD4 empty vector. **(C)** HEK293 cells co-transfected with c-Myc-PfARO + FLAG-PfDHC7 or c-Myc-PfARO + CD4 empty vector. **(D)** HEK293 cells co-transfected with c-Myc-PfARO + FLAG-PfDHC9 or c-Myc-PfARO + CD4 empty vector.

The intensity of the PfARO palmitoylated band in the click chemistry elutions was clearly greater when PfARO was co-expressed with PfDHH5 compared to when PfARO was co-expressed with the empty vector (Figure 6.5A). The same thing was also observed for PfDHH3, 7 and 9, where the intensity of the PfARO palmitoylated band was greater when PfARO was co-expressed with each of these DHHC proteins, compared to when co-expressed with the empty vector (Figure 6.5B, C and D). Unfortunately, in these experiments with PfARO, immunoblot of the initial cell lysates indicated that protein expression in each of the different transfection and labelling conditions appeared to vary, perhaps due to differences in cell numbers or in the health of the cells. This difference in protein expression could contribute to the difference seen in the intensity of the PfARO palmitoylated band in the click chemistry elutions and this means that the increased intensity of the PfARO palmitoylated band in the presence of the PfDHHs over that of the bands due to background palmitoylation may not be completely due to palmitoylation by the PfDHHs, but due in part also to the differences in protein expression.

However, as in the case of PfSec22, the difference in the intensities of the higher molecular weight (45 kDa) palmitoylated band of PfARO, in the presence and absence of the different PfDHH proteins, was also observed when PfARO was immunoprecipitated and analysed by immunoblot (Figure 6.3), as described in Section 6.2.3. In the immunoblot of immunoprecipitated PfARO, protein expression in each transfection condition appeared to be more similar, and thus should not contribute as much to the differences in the intensity of the PfARO palmitoylated band. In this case, the intensity of the PfARO palmitoylated band appeared to be greater in the presence of PfDHH3, 5, 7 and 9, compared to the intensity of the palmitoylated band in the absence of any PfDHH (Figure 6.3). This suggests that like PfSec22, PfARO was also able to be palmitoylated by at least 4 different PfDHH proteins, further demonstrating the potential overlapping functionality of the DHHC proteins.

In summary, the use of the PAT assay with PfARO as a target has again demonstrated the PAT activity of the four *P. falciparum* DHHC proteins studied here – PfDHH3, 5, 7 and 9. Additionally, the intensity of the PfARO palmitoylated band appeared to be the greatest in the presence of PfDHH3 and 5 (Figure 6.3), implying that of the four PfDHHs tested in this PAT activity assay, PfDHH3 and 5 appeared to have the most PAT activity towards PfARO. As PfARO localises to the rhoptries in *P. falciparum*, it was expected that palmitoylation of PfARO might mainly be performed by PfDHH7, which also localises to the rhoptries in the parasite. However, the results of the PAT assay show that PfDHH5 (which localises to the ER) and PfDHH3 (which localises to the Golgi) appeared to have strong PAT activity towards PfARO, implying that PfARO might be palmitoylated when in the

secretory pathway instead, before attachment to the rhoptry membrane. PfDHHC7 meanwhile may be required for the palmitoylation of other rhoptry proteins. For example, RAP1, another rhoptry-localising protein, was identified as palmitoylated by the site-ID palmitome purifications (described in Chapter 3), and could potentially be a target for PfDHHC7. Alternatively, another possibility could be that some of these parasite DHHC proteins could perhaps be non-functional when expressed in a mammalian cell expression system. It is interesting that PfDHHC5 was clearly able to palmitoylate both PfARO as well as PfSec22, perhaps indicating better expression, or increased activity, of this particular PfDHHC protein in mammalian cells compared to others.

Lastly, PfDHHC9, which localises to the IMC in *P. falciparum*, did not appear to have strong PAT activity towards either PfARO or PfSec22 (at least in these PAT activity assays), despite exhibiting an ER localisation in HEK293 cells. This indicates that despite the change in localisation, which allows PfDHHC9 to come in contact with potential targets in the ER, palmitoylation of these targets does not occur as much as when in the presence of the other PfDHHCs, suggesting that PfDHHC9 may possibly only palmitoylate a specific set of target proteins, which are likely to be IMC-localised proteins. This could perhaps be an example of the potential substrate specificity of this DHHC protein, although this is by no means proven at this time.

6.4. Point mutations of potential palmitoylation sites in PfSec22 and PfARO

The results described above have shown that PfSec22 and PfARO are indeed palmitoylated proteins and are palmitoylated mainly by PfDHHC3 and 5. For both these palmitoyl-proteins, evidence exists indicating that cysteine residues in the N-terminal regions of these proteins may potentially be the sites of palmitoylation, as described in Section 6.2.1. In order to determine which cysteine residues in both these target proteins were the palmitoylated cysteines, point mutations, which mutated the potentially palmitoylated cysteine residue into an alanine residue, were performed on the PfSec22 and PfARO HEK293 expression constructs, as described in the Materials and Methods. These mutated target proteins were then expressed in HEK293 cells, along with the appropriate PfDHHC protein, in order to determine whether any difference in the level of target protein palmitoylation was observed. If the particular cysteine residue was indeed the site of palmitoylation, it was expected that mutation of that cysteine residue into an alanine residue would result in a decrease in the level of target protein palmitoylation.

6.4.1. Point mutation of Cys2 in PfSec22 does not affect PfSec22 palmitoylation by PfDHHc5

As mentioned in Section 6.2.1, two cysteines in the N-terminal region of PfSec22 (Cys2 and Cys8) were identified as potentially palmitoylated by the site-ID palmitome purification experiments described in Chapter 3. In order to determine whether Cys2 was the site of PfSec22 palmitoylation, a point mutant was made, where Cys2 was mutated into an alanine residue, and co-expressed in HEK293 cells along with PfDHHc5, the DHHc protein which produced the greatest increase in PfSec22 palmitoylation over background palmitoylation (Section 6.3.1). The PfSec22 Cys2 point mutant (named PfSec22-C2dA) was then immunoprecipitated from protein extracts using antibodies against the c-Myc tag, and visualised by immunoblot using antibodies against the c-Myc tag from a different species.

Immunoblot analysis of PfSec22-C2dA immunoprecipitates resulted in the same two bands observed previously (Section 6.2.3): the lower molecular weight band running at approximately 10 kDa and the higher molecular weight band running at approximately 17 kDa, which was determined above to correspond to the palmitoylated version of the protein (Figure 6.6A). In the presence of PfDHHc5, there was no difference in the intensity of the palmitoylated band of PfSec22-C2dA compared to the intensity of the palmitoylated band of wild-type PfSec22 (Figure 6.6A). Additionally the intensity of the PfSec22-C2dA palmitoylated band (when co-expressed with PfDHHc5) was still greater than that of the band due to background palmitoylation (when wild-type PfSec22 was co-expressed with the empty vector) (Figure 6.6A). This indicates that the mutation of Cys2 into an alanine residue did not have an effect on the level of PfSec22 palmitoylation by PfDHHc5. It can thus be inferred that Cys2 of PfSec22 does not appear to be the site of palmitoylation. This suggests that Cys8 is likely to be the palmitoylated cysteine instead. Although it was intended to perform a point mutation of Cys8 as well, due to time constraints, the data was not yet available during the writing of this dissertation.

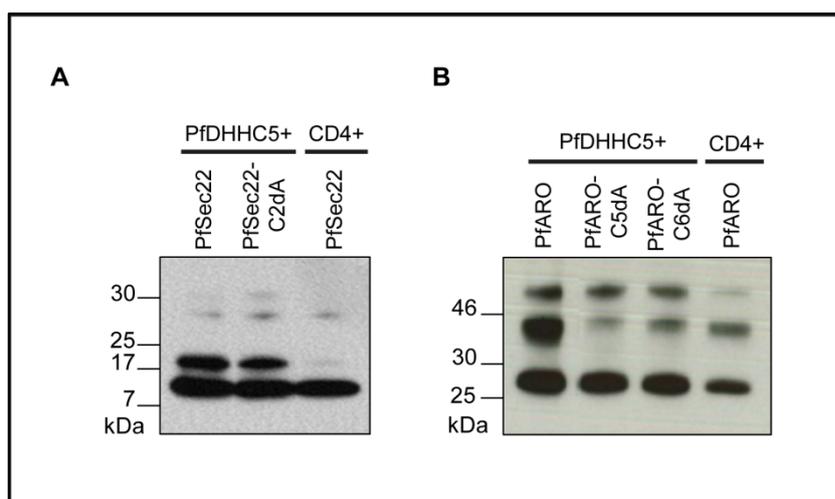


Figure 6.6: Immunoprecipitation of mutant PfSec22 and PfARO proteins using antibodies against the c-Myc tag. HEK293 cells were co-transfected with plasmids coding for the expression of mutant c-Myc-tagged PfSec22 or PfARO (which had the predicted palmitoylated cysteine residues point mutated to alanine residues), along with FLAG-tagged PfDHHHC5 or the empty vector (CD4). The mutant PfSec22 and PfARO proteins were immunoprecipitated from cell lysates using an α -c-Myc antibody. The proteins were separated by SDS-PAGE and visualised by immunoblot, using α -c-Myc antibody from a different species. **(A)** HEK293 cells co-transfected with c-Myc-PfSec22 + FLAG-PfDHHHC5, c-Myc-PfSec22-C2dA + FLAG-PfDHHHC5 or c-Myc-PfSec22 + CD4 empty vector. **(B)** HEK293 cells co-transfected with c-Myc-PfARO + FLAG-PfDHHHC5, c-Myc-PfARO-C5dA + FLAG-PfDHHHC5, c-Myc-PfARO-C6dA + FLAG-PfDHHHC5 or c-Myc-PfARO + CD4 empty vector.

6.4.2. Point mutation of Cys5 and Cys6 in PfARO reduces palmitoylation by PfDHHHC5

PfARO has been previously shown to be palmitoylated, and two cysteine residues in the N-terminal region of PfARO (Cys5 and Cys6) are thought to be the potential sites of palmitoylation [9]. In order to determine whether Cys5 or Cys6 was the palmitoylated cysteine, point mutants were made for both Cys5 and Cys6, where each cysteine residue was mutated into an alanine residue individually. Each of these point mutants were co-expressed in HEK293 cell, again along with PfDHHHC5. The PfARO Cys5 point mutant (name PfARO-C5dA) and the PfARO Cys6 point mutant (named PfARO-C6dA) were each immunoprecipitated from protein extracts using antibodies against the c-Myc tag, and visualised by immunoblot using antibodies against the c-Myc tag from another species, as described above.

Immunoblot of the immunoprecipitates from both PfARO point mutants resulted in the same bands as previously observed (Section 6.2.3): the lower molecular weight band running at approximately 28 kDa and the higher molecular weight band running at approximately 45 kDa, which was determined above to correspond to the palmitoylated version of the protein (Section 6.3.2) (Figure 6.6B).

Additionally, the third band running above the 46 kDa marker was also present in the immunoprecipitates from both point mutants (Figure 6.6B).

In the presence of PfDHC5, there was a decrease in the intensity of the palmitoylated band of PfARO-C5dA compared to the intensity of the palmitoylated band of wild-type PfARO (Figure 6.6B). In fact, the intensity of the PfARO-C5dA palmitoylated band did not appear to be greater than that of the band due to background palmitoylation (when wild-type PfARO was co-expressed with only the empty vector control, not with PfDHCs) (Figure 6.6B). This indicated that the mutation of Cys5 to an alanine residue appeared to affect the level of PfARO palmitoylation by PfDHC5, thus suggesting that Cys5 may be a palmitoylated cysteine.

The same result was observed for the PfARO-C6dA point mutant, where in the presence of PfDHC5, there was a decrease in the intensity of the palmitoylated band compared to that of wild-type PfARO (Figure 6.6B). In this case as well, the intensity of the PfARO-C6dA palmitoylated band did not appear to be greater than that of the band due to background palmitoylation (Figure 6.6B). This indicated that the mutation of Cys6 into an alanine residue also appeared to affect the level of palmitoylation by PfDHC5, suggesting therefore that Cys6 of PfARO may also be a palmitoylated cysteine.

The results described here suggest that both cysteine residues (Cys5 and Cys6) in the N-terminal region of PfARO appear to be the sites of palmitoylation, and mutation of either of these cysteine residues into alanine residues results in a decrease in the level of PfARO palmitoylation. Presumably, the continued presence of the palmitoylated band when either of the cysteine residues was individually mutated could be due to the palmitoylation of the non-mutated cysteine residue. It would be interesting to determine whether, if both cysteine residues were to be simultaneously mutated, the palmitoylated band would no longer be present, indicating that palmitoylation had been abolished. This would further confirm that both Cys5 and Cys6 of PfARO were in fact the sites of palmitoylation. Unfortunately, due to time constraints, this data was not yet available at the time of writing this dissertation.

6.7. Site-directed mutagenesis of the cysteine in the DHHC domain of PfDHC5

Not much is known about the exact mechanism of palmitate transfer by the DHHC proteins, and the role played by the highly conserved DHHC signature motif is as yet unknown [18]. Most of the DHHC proteins studied so far have been found to be palmitoylated when incubated individually with

palmitoyl-CoA *in vitro*, indicating that these DHHC proteins are autopalmitoylated [1-3, 19]. This led to the hypothesis that palmitate transfer occurs via a palmitoyl-enzyme intermediate. As autopalmitoylation is abolished when the cysteine residue of the DHHC domain is mutated, that cysteine residue is thought to be a possible candidate for the site of modification during the formation of the palmitoyl-enzyme intermediate, although autopalmitoylation of another cysteine residue is also possible [18].

In order to determine whether the cysteine residue in the DHHC domain of the PfDHHC proteins was important for the palmitoylation of target proteins, site-directed mutagenesis was performed in order to mutate that cysteine residue into an alanine residue in the PfDHHC5 HEK293 expression construct, as described in the Materials and Methods. The PfDHHC5 point mutant (named PfDHHC5-CdA) was then co-expressed with PfSec22 or PfARO in HEK293 cells. The target proteins were immunoprecipitated from protein extracts using antibodies against the c-Myc-tag, and visualised by immunoblot using α -c-Myc antibodies from a different species, in order to determine if any difference in target protein palmitoylation was observed.

Immunoblot of PfSec22 resulted in the same two bands as observed previously: one band running at approximately 10 kDa and a second band running at approximately 17 kDa, which was shown above to correspond to palmitoylated PfSec22. In the presence of PfDHHC5-CdA, the intensity of the palmitoylated band of PfSec22 appeared to be the same as the intensity of the palmitoylated band of PfSec22 when in the presence of wild-type PfDHHC5 (Figure 6.7A). The palmitoylated band of PfSec22 due to background palmitoylation (when PfSec22 was co-expressed with the empty CD4 vector) was not clearly seen here, although background palmitoylation of PfSec22 was observed to occur in other experiments. However, in this experiment, protein expression as a whole appeared to be lower than normal, and this was probably why this lower intensity background palmitoylated band was not seen in this case. Nevertheless, the palmitoylated band due to palmitoylation by PfDHHC5-CdA was clearly present, and was greater than when PfSec22 was co-expressed with the empty CD4 vector (Figure 6.7A). This indicates that mutation of the cysteine residue of the DHHC domain did not appear to affect the level of PfSec22 palmitoylation, which argues against a role for this residue in the formation of the palmitoyl-enzyme intermediate, or at least a role that is absolutely required for enzyme activity.

Immunoblot of PfARO again resulted in the same bands as observed previously: one band running at approximately 28 kDa and one band running at approximately 45 kDa, which was shown previously to correspond to the palmitoylated version of PfARO. In this case, the third band running above the 46 kDa marker that was observed in previous experiments was not observed here. However, as

mentioned above, protein expression in this experiment was relatively lower than normal, and as a consequence, this band, which was usually of a lower intensity, could not be seen here. In the presence of the PfDHHc5-CdA, the intensity of the palmitoylated band of PfARO appeared to be the same as the intensity of the palmitoylated band when in the presence of wild-type PfDHHc5 (Figure 6.7B). The intensity of the palmitoylated band was also greater than that of the band due to background palmitoylation (when PfARO was co-expressed with the CD4 empty vector) (Figure 6.7B). Again, this indicates that the mutation of the cysteine residue in the DHHc domain did not appear to affect the level of PfARO palmitoylation.

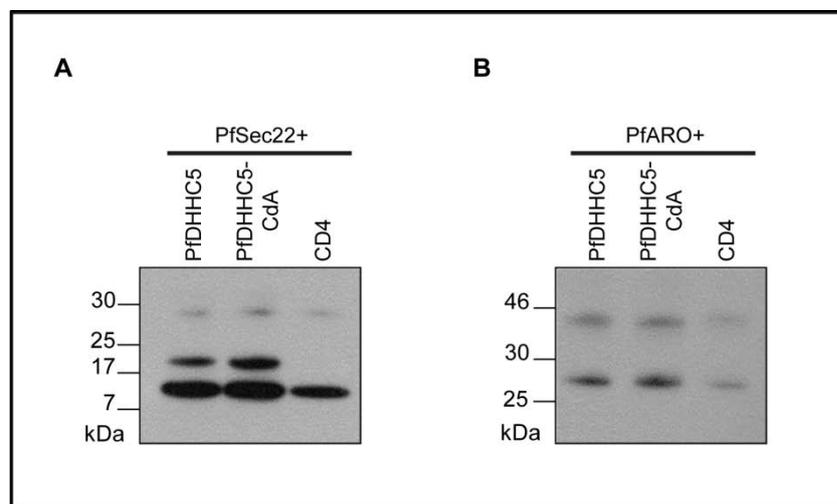


Figure 6.7: Immunoprecipitation of PfSec22 and PfARO proteins when co-expressed with mutant PfDHHc5 using antibodies against the c-Myc tag. HEK293 cells were co-transfected with plasmids coding for the expression of c-Myc-tagged PfSec22 or PfARO along with FLAG-tagged mutant PfDHHc5 (which had the cysteine residue of the DHHc domain mutated into an alanine residue) or the empty vector (CD4). PfSec22 and PfARO proteins were immunoprecipitated from cell lysates using an α -c-Myc antibody. The proteins were separated by SDS-PAGE and visualised by immunoblot, using α -c-Myc antibody from a different species. **(A)** HEK293 cells co-transfected with c-Myc-PfSec22 + FLAG-PfDHHc5, c-Myc-PfSec22 + FLAG-PfDHHc5-CdA or c-Myc-PfSec22 + CD4 empty vector. **(B)** HEK293 cells co-transfected with c-Myc-PfARO + FLAG-PfDHHc5, c-Myc-PfARO + FLAG-PfDHHc5-CdA, or c-Myc-PfARO + CD4 empty vector.

The results described above suggest that, at least in the case of PfDHHc5, the mutation of the cysteine residue of the DHHc domain into an alanine residue did not have an effect on PfSec22 and PfARO palmitoylation. This therefore implies that the cysteine residue of the DHHc domain of PfDHHc5 does not appear to be involved in the transfer of palmitate to target proteins.

For the yeast DHHC proteins, Erf2 and Akr1, mutation of the first histidine residue of the DHHC domain has been found to reduce palmitoylation of their substrate [1, 2]. Interestingly, for both Erf2 and Akr1, mutation of the cysteine residue of the DHHC domain, besides abolishing autopalmitylation, also resulted in the loss of PAT activity [1, 2]. This was not observed here for PfDHHC5. However, as other residues of the DHHC domain (such as the first histidine residue) appear to also be involved in palmitoylation, it could be that in the case of PfDHHC5, other residues within the DHHC domain are important for palmitate transfer to the substrate, whilst the cysteine residue is involved in the autopalmitylation of the DHHC protein itself. Unfortunately, due to the size of the PfDHHC5 protein (80.8 kDa) and to its solubility in detergents, PfDHHC5 expressed in HEK293 cells was unable to be immunoprecipitated or visualised by immunoblot. Thus, it is unknown whether the mutation of the cysteine residue in the DHHC domain affects autopalmitylation of PfDHHC5 at this time.

Alternatively, other models of protein palmitoylation suggest that instead of acting as the palmitoyl-enzyme intermediate, the DHHC domain may be involved in the allosteric regulation of the DHHC protein, controlling substrate specificity and binding, whilst palmitate transfer occurs on other cysteine residues [18]. As the results described here indicate that mutation of the cysteine residue did not affect target protein palmitoylation, this could imply that the cysteine residue of the DHHC domain is not involved in substrate specificity or binding, if that is indeed the role of the DHHC domain.

It would be interesting to determine whether the same results are observed for the other *P. falciparum* DHHC proteins studied here, as well as to investigate which residue within the DHHC domain is actually important for the palmitoylation of target proteins, in order to shed some light on the exact mechanism of palmitate transfer performed by the DHHC proteins.

Conclusions

The PAT activity assay developed here has successfully shown that four PfDHHC proteins of interest (PfDHHC3, 5, 7 and 9) exhibit PAT activity and can palmitoylate *P. falciparum* proteins expressed in a

heterologous system. This assay also confirms that two potential palmitoyl-proteins, PfSec22 and PfARO, are indeed palmitoylated. Furthermore, this assay has revealed that like the DHHC proteins of other eukaryotic organisms, the *Plasmodium* DHHC proteins have overlapping functionalities, where several different PfDHHC proteins appear to be able to palmitoylate the same substrate protein, although factors such as the localisation of the DHHC protein in relation to the localisation of the target protein may affect target protein palmitoylation. Lastly, at least in the case of PfDHHC5, the cysteine residue of the DHHC domain does not appear to be directly involved in palmitate transfer, implying that the other residues of this highly conserved domain may be the residues involved in target protein palmitoylation.

References:

1. Lobo, S., *Identification of a Ras Palmitoyltransferase in Saccharomyces cerevisiae*. Journal of Biological Chemistry, 2002. **277**(43): p. 41268-41273.
2. Roth, A.F., *The yeast DHHC cysteine-rich domain protein Akr1p is a palmitoyl transferase*. The Journal of Cell Biology, 2002. **159**(1): p. 23-28.
3. Swarthout, J.T., *DHHC9 and GCP16 Constitute a Human Protein Fatty Acyltransferase with Specificity for H- and N-Ras*. Journal of Biological Chemistry, 2005. **280**(35): p. 31141-31148.
4. Crosnier, C., et al., *Basigin is a receptor essential for erythrocyte invasion by Plasmodium falciparum*. Nature, 2011.
5. Hicks, Stuart W., et al., *Subcellular Targeting of Salmonella Virulence Proteins by Host-Mediated S-Palmitoylation*. Cell Host & Microbe, 2011. **10**(1): p. 9-20.
6. Durocher, Y., S. Perret, and A. Kamen, *High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells*. Nucleic Acids Res, 2002. **30**(2): p. E9.
7. Bushell, K.M., et al., *Large-scale screening for novel low-affinity extracellular protein interactions*. Genome Research, 2008. **18**(4): p. 622-630.
8. Ayong, L., et al., *The Longin Domain Regulates the Steady-State Dynamics of Sec22 in Plasmodium falciparum*. Eukaryotic Cell, 2009. **8**(9): p. 1330-1340.
9. Cabrera, A., et al., *Dissection of Minimal Sequence Requirements for Rhoptry Membrane Targeting in the Malaria Parasite*. Traffic, 2012. **13**(10): p. 1335-1350.
10. Kleizen, B. and I. Braakman, *Protein folding and quality control in the endoplasmic reticulum*. Curr Opin Cell Biol, 2004. **16**(4): p. 343-9.
11. Geiger, B., et al., *Broad spectrum pan-cadherin antibodies, reactive with the C-terminal 24 amino acid residues of N-cadherin*. J Cell Sci, 1990. **97 (Pt 4)**: p. 607-14.
12. Yanai, A., et al., *Palmitoylation of huntingtin by HIP14 is essential for its trafficking and function*. Nature Neuroscience, 2006. **9**(6): p. 824-831.
13. Yeoman, J.A., et al., *Tracking Glideosome-Associated Protein 50 Reveals the Development and Organization of the Inner Membrane Complex of Plasmodium falciparum*. Eukaryotic Cell, 2011. **10**(4): p. 556-564.
14. Tang, B.L., D.Y. Low, and W. Hong, *Hsec22c: a homolog of yeast Sec22p and mammalian rsec22a and msec22b/ERS-24*. Biochem Biophys Res Commun, 1998. **243**(3): p. 885-91.

15. Rees-Channer, R.R., et al., *Dual acylation of the 45 kDa gliding-associated protein (GAP45) in Plasmodium falciparum merozoites*. Mol Biochem Parasitol, 2006. **149**(1): p. 113-116.
16. Martin, B.R. and B.F. Cravatt, *Large-scale profiling of protein palmitoylation in mammalian cells*. Nat Methods, 2009. **6**(2): p. 135-138.
17. Roth, A.F., et al., *Global analysis of protein palmitoylation in yeast*. Cell, 2006. **125**(5): p. 1003-1013.
18. Mitchell, D.A., et al., *Protein palmitoylation by a family of DHHC protein S-acyltransferases*. J Lipid Res, 2006. **47**(6): p. 1118-1127.
19. Smotrys, J.E., *The vacuolar DHHC-CRD protein Pfa3p is a protein acyltransferase for Vac8p*. The Journal of Cell Biology, 2005. **170**(7): p. 1091-1099.

Chapter 7

Discussion

Palmitoylation, one of the lesser known protein PTMs, was initially thought to be merely a simple membrane anchor. Recently however, the development of new methods to purify whole 'palmitomes' from individual organisms has revealed the broad scope of this PTM in many different eukaryotic organisms, and has raised questions about the potentially diverse roles played by this PTM in protein function. The fact that palmitoylation is the only known reversible lipid modification, as well as evidence that palmitoylation can have active roles in many aspects of protein localisation, trafficking and function, has brought about the view that this protein PTM is important for normal cellular function [1, 2]. Furthermore, the recent discovery of protein families capable of mediating the addition of palmitate to proteins (the DHHC and MBOAT protein families), as well as protein families capable of removing palmitate from proteins (the APT protein family), raises interesting questions regarding the dynamic regulation of palmitoylation, and the effect and role of this dynamic regulation on the function of the palmitoylated protein [2, 3]. Until recently, only three proteins were confirmed to be palmitoylated in *Plasmodium* parasites (GAP45 [4], CDPK1 [5] and calpain [6]). Application of the new palmitome purification technologies however has revealed that more than 400 putative palmitoyl-proteins are found in the intraerythrocytic schizont stages of *P. falciparum*, which is a similar percentage of the total proteome as found in other eukaryotic species [7]. In the work described in this dissertation, a more in-depth study of protein palmitoylation in *P. falciparum* was performed, and can be divided into two main aims – a global analysis of the sites of palmitoylation within the palmitome of *P. falciparum* schizonts, and the characterisation of the DHHC and MBOAT proteins in *Plasmodium* parasites.

In order to fulfil the first aim, a method for the purification of palmitoylated peptide fragments was developed based on previously used purification techniques. This method was applied to the schizont stages of *P. falciparum* and resulted in the successful purification of putative palmitoylated peptide fragments, which contained the putative palmitoylated cysteine residues, resulting in the identification of potential sites of palmitoylation within the *P. falciparum* schizont palmitome (Chapter 3).

Although the application of this site-ID palmitome purification method resulted in the identification a number of putative palmitoylated peptides, further development is clearly needed. Two main technical issues remain to be overcome. Firstly, the quantification method that resulted in the best proteome coverage and produced the largest dataset of enriched peptides was label-free quantification. However, this method of quantification is less accurate and more prone to error and variability (due to the inherent run-to-run experimental variations which occur during label-free mass spectrometry analysis) compared to other methods of quantification, such as SILAC metabolic

labelling and stable isotope dimethyl labelling [8, 9]. This means that the datasets produced, although certainly containing actual palmitoylated peptides, may possibly contain more false-positive identifications due to inaccurate quantification, and subsequently, generation of inaccurate enrichment ratios. These datasets must therefore be treated with higher stringency when deciding the cut-off point for enrichment. One way to combat this would be to perform many replicates of different purifications and considering only the peptides which are consistently identified as enriched across the replicates, as consistent peptide identifications across multiple replicates are less likely to be false-positive contaminants. For the site-ID palmitome purifications from *P. falciparum* schizonts described here, multiple label-free replicates were unable to be performed due to time constraints. However, by considering the overlaps between the datasets produced from all the different purifications performed during the development of the purification method, and accepting only the peptides that were consistently present in the different datasets, palmitoylation sites could still be identified, which will be valuable for follow-up experiments.

The second issue observed during the development of the site-ID palmitome purification method was the relatively high background of cysteine-containing peptides identified in control samples. A 'noise' level of cysteine-containing peptides is always expected to be present in control samples, and in fact, is necessary for relative quantification of abundance by mass spectrometry. In these analyses, enrichment ratios were used as a measure of the enrichment of a particular peptide in the palmitome samples compared to the control samples, and thus as an indication of whether the peptide was palmitoylated or not. These enrichment ratios were generated by comparing the intensity of the particular cysteine-containing peptide in palmitome samples with its intensity in the control samples. This means that a level of cysteine-containing peptides in control samples is required for the generation of the enrichment ratios. The abundance of these cysteine-containing peptides in control samples are usually very low and are not usually high enough to be directly sequenced and detected by MaxQuant. These peptides are instead usually found by MaxQuant in the control samples using the 'match between runs' feature, which matches peptide ions by their mass and retention time. However, in these site-ID purifications, some cysteine-containing peptides in control samples was directly identified by MaxQuant, indicating that the background of cysteine-containing-peptides in the control samples was unfortunately higher than expected.

This high background in control samples would not have a great effect on peptides which had much higher intensities in the palmitome samples, thus resulting in high enrichment ratios. However, this background could certainly affect those peptides with lower intensities, and could possibly in fact mask the fold enrichment of the peptide. This is further complicated by the fact that label-free

quantification is less sensitive to small quantitative differences [9], and thus might not be able to accurately measure the lower fold enrichment due to increased background in the control samples. This background of cysteine-containing peptides could thus have an effect on the enrichment ratios, and this would ultimately affect the classification of peptides as enriched, as the cut-off criteria used here to define enrichment is based on the enrichment ratios (described in Section 3.1.2). This is especially true when the cut-off criteria is based on the median of the enrichment ratios, as the generation of relatively low enrichment ratios as a result of the background in the control would result in a lower median, and subsequently, a lowered cut-off point.

The presence of this high background of cysteine-containing peptides in the control samples points to either the incomplete blockage of free thiols by the initial IAA (or NEM) treatment, or the possibility that palmitate groups were somehow falling off during the later steps of purification. Treatment with both IAA and NEM is sensitive to pH, with specific blockage of free thiols taking place at a specific pH. The specificity of the blockage and the rate of the reaction are thus dependent on the pH of the reaction. Although the pH was measured and adjusted to the optimum pH during the initial blocking of free thiols, minor changes in the pH could have occurred during the course of the blockage step (the samples were usually blocked for two hours), especially since a reducing agent, TCEP, was present in the lysis buffer for the extraction of the proteome, in order to reduce any disulphide bonds present. This could potentially contribute to the background observed in control samples, although it is unlikely to be the only cause of the high background in the control. Unfortunately, due to time constraints, this issue was unable to be solved before the writing of this dissertation.

However, despite the presence of this background in the datasets already obtained, analysis of the overlaps between the datasets, as described above, can still provide useful information on the potential sites of palmitoylation. By considering just the peptides that were consistently purified in the two label-free datasets (Trial 3B and 4), and which corresponded to proteins previously identified in the total schizont palmitome [7], 142 putative palmitoylation sites were identified. Furthermore, of these 142 common enriched palmitoylation sites, 14 palmitoylation sites were found to be common across four of the trial purification datasets (Trial 2, 3A, 3B and 4). Although false positive identifications are undoubtedly still present within these datasets, not only due to the issues reported above, but also due to the contaminant proteins that are inherently purified by ABE methods [7], this list of potential palmitoylation sites can still be used as a guide for the design of further experiments and the generation of testable hypotheses.

It was interesting to note that in all the datasets collected, not only the *P. falciparum* schizont dataset, but also the preliminary *P. berghei* schizont dataset, as well as the *T. gondii* tachyzoite dataset, members of the DHHC protein family were found to be present and were classified as enriched. The DHHC protein family have been shown to be the enzymes responsible for catalysing protein palmitoylation in other eukaryotic organisms, and all the DHHC proteins studied thus far have been shown to also be palmitoylated themselves [3]. The identification of these proteins in the site-ID palmitome purification datasets is therefore reassuring, not only as additional validation for the site-ID palmitome purification method, but also as possible confirmation that these proteins are also palmitoylated in Apicomplexan organisms. Each of the three Apicomplexan species mentioned above possess a relatively large repertoire of DHHC-domain-containing proteins, with 12 DHHC proteins in *P. falciparum*, 11 DHHC proteins in *P. berghei* and 18 DHHC proteins in *T. gondii*. However, not all of the DHHC proteins were identified in the different site-ID trial purifications, and in fact only one of the PfDHHC proteins was consistently identified in more than one of the *P. falciparum* trial purifications. This may yet again be the result of the issues faced with the site-ID protocol (as described above), or may instead reflect the difficulties of sufficiently extracting these multi-pass TM-domain-containing proteins, some of which have relatively high molecular weights. Nevertheless, the identification of some members of the DHHC protein family in these site-ID purification datasets was a first step in the characterisation of this protein family in Apicomplexan parasites.

Importantly, in all the site-ID trial datasets, the cysteine residue identified as palmitoylated for the DHHC proteins was not the cysteine residue of the conserved DHHC motif. One hypothesis for the transfer of palmitate to cysteine residues of target proteins involves the formation of a palmitoyl-enzyme intermediate, and the cysteine of the DHHC motif has long been thought to be a possible candidate for the site of modification on DHHC proteins, although modification at another cysteine residue is also thought to be possible [3]. The data obtained from the site-ID palmitome purifications however indicate that palmitoylation of the DHHC proteins appeared to occur on a cysteine residue which was not the cysteine of the DHHC motif, at least for the DHHC proteins identified in the site-ID purifications. It is possible that the cysteine of the DHHC motif is also palmitoylated but was not identified in the site-ID purifications, due either to chance or to the issues with the site-ID protocol mentioned above, and the cysteine identified as palmitoylated by the site-ID purifications could perhaps have other roles besides palmitate transfer, such as in the localisation of the DHHC protein itself or in regulating substrate specificity. Conversely, another possibility is that the cysteine residue involved in palmitate transfer might not be the cysteine of the DHHC motif, if palmitate transfer does indeed occur via a palmitoyl-enzyme intermediate. Whether these cysteine residues are truly palmitoylated, and whether palmitoylation at these sites are actually required for the palmitoylation

of the target proteins by the DHCs can now be validated and tested using the PAT activity assay also developed in the work for this dissertation (Chapter 6), and this may then provide some insight into the mechanisms of palmitate transfer by the DHC proteins.

In order to further characterise these putative PATs in *Plasmodium* parasites, the repertoire of DHC-domain-containing proteins were localised and knocked-out in both species of *Plasmodium*: *P. berghei* and *P. falciparum* (Chapter 4 and 5). In both *Plasmodium* species, the DHC proteins appeared to localise to different membrane-bound compartments within the cell. These membrane compartments consisted of organelles such as the Golgi [10] and the ER [11], to which the DHC proteins of other eukaryotic organisms are commonly localised to, as well as specialised, parasite-specific organelles, such as the rhoptries and the IMC. The distribution of the DHC proteins to various different membranous compartments suggests that at least one aspect of the regulation of DHC protein substrate specificity may be reliant on the localisation of the DHC protein, which dictates the potential target proteins that come into contact with the particular DHC protein. Additionally, the expression of the DHC proteins appeared to be stage-specific, with only a subset of proteins highly expressed in the schizont stages studied here, which further implied that the activity of different DHC proteins may be required at different stages of the parasite life cycle. The localisation of these proteins to different sites in the cell, and the differential expression of these proteins at different stages of the life cycle, may be the reason why so many of the DHC proteins are present in each species, and may be how substrate specificity of these proteins is regulated, especially considering the fact that these proteins have been found to be relatively promiscuous in terms of substrate specificity in other organisms [12].

Studies on the essentiality of the DHC protein family in both *Plasmodium* species revealed the existence of two subsets of DHC proteins. One subset consisted of proteins that were unable to be knocked-out in the intraerythrocytic stages of the parasite, implying that these proteins were essential for blood stage growth, and the second subset consisted of proteins that were successfully knocked-out in the blood stages without any detrimental effect on blood stage growth, implying a possible redundancy in the function of these proteins, at least in the blood stages studied here. Overlapping functionality of the DHC proteins has been demonstrated before in yeast, where individual deletions of yeast DHC proteins had minimal effect on substrate palmitoylation. In fact, a significant decrease in substrate palmitoylation was only observed when multiple DHC proteins were simultaneously deleted [12]. The ability to delete some of the DHC proteins in both *P. berghei* and *P. falciparum* imply that the same overlapping functionality may exist in *Plasmodium*. However, whereas all the yeast DHC proteins were found to be non-essential when individually deleted [12],

the results shown here indicate that at least a few of the DHHC proteins in *P. falciparum*, for example PfDHHC3 (localised to the Golgi), PfDHHC7 (localised to the rhoptries) and PfDHHC8 (localisation undetermined due to low expression), and in *P. berghei*, for example, PbDHHC4 (localisation unknown) and PbDHHC8 (localisation undetermined due to low expression), appeared to be essential to blood stage growth, perhaps reflecting the more complicated life cycle of the parasite and possible life cycle stage-specific expression and function of the DHHC proteins.

Due to time constraints, only the intraerythrocytic stages were studied here for both *P. falciparum* and *P. berghei*. Recent work by others however, revealed that the disruption of PfDHHC9 (which was shown in the work presented in this dissertation to be non-essential in the blood-stages of both *P. falciparum* and *P. berghei*) prevented the formation of mature gametocytes in *P. falciparum*, implying that this DHHC protein is essential for gametocytogenesis [13], a process which is necessary for the transmission of the parasite to the mosquito vector. This indicates that further work on the localisation and essentiality of the DHHC proteins in other stages of the parasite life cycle besides the asexual intraerythrocytic stages, such as the sexual gametocyte stages, would certainly provide further insight on the function and importance of the DHHC protein family in *Plasmodium* biology. These studies could be more easily done using *P. berghei* instead of *P. falciparum*, as the phenotype of a particular DHHC knock-out strain could be observed in all stages of the *Plasmodium* life-cycle, and the effect of the disruption of the DHHC protein on parasite transmission to the mosquito could also be studied.

The next step in understanding the specificity of palmitoylation in *Plasmodium* parasites would be to combine the DHHC knock-out transgenic parasite lines generated here with the newly-developed site-ID palmitome purification protocol, in order to determine which specific or overlapping sites were palmitoylated by individual DHHCs. This would be achieved by performing the site-ID palmitome purification on each individual DHHC knock-out transgenic strain, and comparing the enriched palmitoylation sites obtained, with that of the wild-type strain, in order to determine whether any palmitoylation sites were no longer present when a particular DHHC protein was deleted. This would then be an indication of which sites were palmitoylated by a specific DHHC protein. Similar studies in yeast revealed that the yeast DHHC proteins palmitoylate overlapping sets of substrate proteins [12]. However, given that unlike in yeast, some of the DHHC proteins in *Plasmodium* are not redundant and are not amenable to disruption in the blood stages, this might not be the case for this set of non-redundant DHHC proteins, which may possibly be responsible for palmitoylating specific, non-overlapping substrates. In order to perform the comparative site-ID studies on these essential DHHCs however, conditionally regulated versions of the essential DHHCs

would be required. While there are currently no robust regulatable systems in *Plasmodium*, recent developments, such as the DiCre conditional recombinase system [14] may offer the way forward.

Given the issues with the site-ID palmitome purification method (described above), this sort of comparative palmitoylation site analysis may potentially be difficult and complex. However, if a sufficient number of replicates are performed, and stringent and rigorous analysis applied to the datasets, valuable information can still be gleaned. Also, as mentioned above, label-free analysis is known to be less sensitive to small quantitative differences [9], and in the case of a comparative palmitoylation site analysis, might not be sensitive enough to detect any small changes in enrichment for a particular palmitoylation site. Thus, although the results presented in this work has determined that label-free quantitative analysis provided better coverage and larger datasets, in the case of the comparative site analysis, it might be more beneficial to use a chemical or metabolic labelling approach rather than a label-free approach, for the sake of higher quantitative accuracy. As stable isotope dimethyl labelling methods appeared to be provide the least information, SILAC metabolic labelling could be used instead, at least in the case of the *P. falciparum* knock-out strains, as this form of labelling worked well in the purification of the total schizont palmitome [7], and can be used for the efficient metabolic labelling of *P. falciparum* parasites in culture. Again due to time constraints, these comparative site analyses have not yet been performed, but are certainly the logical next steps in future experiments.

The repertoire of DHHC proteins in *Plasmodium* is relatively large, and the differences in the localisation and essentiality of these proteins suggested that these proteins may have important roles in parasite biology. Nevertheless, whether the role of these proteins in *Plasmodium* involved catalysing protein palmitoylation was still unknown. The development of the PAT activity assay in HEK293 cells described here however, revealed that the *Plasmodium* DHHC proteins did indeed exhibit PAT activity, at least for the four *P. falciparum* DHHC proteins studied (Chapter 6), indicating that the function of these proteins is indeed to mediate protein palmitoylation. Moreover, analysis using the PAT activity assay revealed that different PfDHHC proteins were able to palmitoylate the same target protein. This acted as further confirmation that, like in yeast, *Plasmodium* DHHC proteins may also possess overlapping functionalities.

Initial PAT activity experiments involving the mutation of the cysteine residue in the highly conserved DHHC domain suggested that for the *Plasmodium* DHHC proteins, the cysteine residue might not be the residue involved in palmitate transfer to target proteins. This contrasted with experiments performed in yeast, where mutation of the cysteine residue in the DHHC domain abolished the palmitoylation of substrate proteins [10, 11], but was consistent with the fact that the cysteine

residue identified as palmitoylated in some PfDHC proteins by the site-ID palmitome purification method was not the cysteine of the DHC motif, raising the question as to whether the transfer of the palmitoyl group to the cysteine residues of target proteins could occur via different mechanisms in *Plasmodium*. Further analysis using site-directed mutagenesis and the PAT activity assay, in order to determine which residues are responsible for palmitate transfer, may provide valuable insights into the mechanism of palmitate transfer by the *Plasmodium* DHC-PATs, and as noted above, can be used to test the function of the palmitoylation sites identified within the DHCs themselves. The PAT activity assay developed here can thus be used not only for the demonstration of PAT activity, but also for determining substrate specificity, and for exploring the mechanisms of palmitate transfer.

In order to perform the PAT activity assay, one of the first steps was to choose target proteins that could potentially be palmitoylated by the DHC proteins of interest. This was done by considering both the data obtained from the site-ID palmitome purifications, as well as the localisation data of the DHC proteins being studied. The target proteins were chosen based on evidence of palmitoylation in the site-ID datasets and their localisation in relation to the localisation of the DHC proteins. For example, one of the target proteins chosen, PfSec22, was identified in the Trial 4 site-ID palmitome purification dataset, with a high enrichment ratio that passed the cut-off criteria for highly enriched proteins. Additionally, PfSec22 was annotated as a SNARE protein, found to be localised to the ER, which made it a candidate for palmitoylation by the DHC proteins localised to the ER and Golgi. Subsequent experiments using the PAT activity assay revealed that PfSec22 was indeed palmitoylated, providing further validation for the ability of the site-ID palmitome purification method to truly purify palmitoylated proteins. Furthermore, it was shown that PfSec22 appeared to be mainly palmitoylated by the PfDHC proteins located at the ER (PfDHC5) or Golgi (PfDHC3), contributing to the possibility that substrate specificity may be affected by the localisation of the DHC proteins in relation to the target proteins. The cysteine residues identified as potentially palmitoylated by the site-ID purification can now be validated experimentally using the PAT activity assay, in order to confirm the site of palmitoylation for PfSec22.

It also must be noted that PfSec22 was identified and classified as enriched only in the Trial 4 site-ID purification, and not in the other trial purifications. This indicates that there is the possibility that a protein which was identified in one dataset, but not in other datasets, could actually still be palmitoylated, despite the lack of overlap between datasets. Thus, when using the site-ID datasets to decide whether a protein may be palmitoylated or not, besides the overlap between the different datasets, it is also important to consider the enrichment ratio. PfSec22 for example, had a relatively

high enrichment ratio which was consistent across the two biological and two technical replicates that were performed in the Trial 4 site-ID purification. Additionally, due to the lack of overlap between datasets for a particular protein, experimental validation of palmitoylation should also be performed, which was done here in the case of PfSec22, using the PAT activity assay.

The second target chosen for the PAT activity assay was PfARO. PfARO was chosen mainly due to its localisation to the rhoptries, which made it a good candidate for palmitoylation by either PfDHHC7, which also localised at the rhoptries, or PfDHHC3 and 5 located at the Golgi and ER respectively. Additionally, PfARO was shown experimentally to be a palmitoylated protein [15], although PfARO was not identified in any of the site-ID trial palmitome purifications performed here. This thus presents a contrasting example of how these site-ID palmitome purifications can also possibly miss some truly palmitoylated proteins, perhaps due to the various issues described previously or perhaps merely by chance. Thus, although the list of putative palmitoyl-proteins and their potential palmitoylation sites is extensive, some palmitoylated proteins still may not be captured by this method of purification. Nevertheless, the datasets generated by the site-ID purifications can still act as an initial guide when looking for potential palmitoyl-proteins, and with the development of the PAT activity assay, there is now the means to easily validate the 'palmitoylation status' of these proteins without having to produce transgenic parasite strains, which particularly in *P. falciparum*, is more difficult and more time consuming.

Conclusions

In the work described here, potential palmitoylation sites were identified for a range of putative palmitoyl-proteins through the development of a site-identification palmitome purification method coupled with quantitative mass spectrometry approaches. The DHHC protein family was characterised in *Plasmodium* parasites, and found to be distributed in different cellular compartments, raising the question as to whether the localisation of the DHHC protein may play a part in regulating substrate specificity. The DHHC family was also found to consist of genes that were both essential and non-essential for blood stage growth, implying a functional redundancy for some of these DHHC proteins. It would be interesting to determine whether any of these DHHC proteins are essential for other stages of parasite development, which would imply a specific functional requirement for that DHHC protein at a particular stage of development.

Lastly, the development of the PAT activity assay allowed the PAT activity of the *Plasmodium* DHHC proteins to be demonstrated, showing for the first time that these parasite proteins did indeed act as

PATs. Studies using this PAT activity assay also confirmed that *Plasmodium* DHHC proteins exhibited overlapping functionalities and were capable of palmitoylating the same substrate, although for some target proteins, there appeared to be a preference for particular DHHC proteins. This assay can now be used to validate the palmitoylation sites identified by the site-ID palmitome purifications, as well as to try and tease out the mechanisms involved in palmitate transfer by the DHHC proteins.

Given the wide-spread occurrence of palmitoylation in *Plasmodium*, the reversibility of this lipid PTM, the presence of *Plasmodium* proteins possessing palmitoyl transferase activity, and the negative effect of the inhibition of palmitoylation on parasite growth and invasion, further pursuing the role of palmitoylation in *Plasmodium* parasites may provide novel drug targets that could be used for therapeutic interventions, especially considering the need for new therapies in the fight against malaria. For example, the set of DHHC proteins which were found to be essential for *Plasmodium* blood-stage growth are logically potential targets for future drug development. The PAT activity assay developed here could then also potentially be used to screen for compounds which inhibit the PAT activity of the essential *Plasmodium* DHHC proteins. Furthermore, the fact that one of the DHHC proteins appears to be essential for gametocytogenesis [13] raises the question as to whether this DHHC protein could be a potential transmission blocking target, and whether there are other DHHC proteins which act at similar stages of the life cycle and could thus also be used as transmission blocking drug targets. However, it must be noted that DHHC proteins are also present in humans (there are 23 DHHC proteins in humans), and the human DHHCs play essential roles in normal cellular function [16, 17]. Targeting the *Plasmodium* DHHC proteins could possibly also result in the cross-inhibition of the human DHHC proteins. An ideal drug would thus be one which can specifically target the *Plasmodium* DHHC proteins without affecting the DHHC proteins and palmitoylation in the human host. Clearly, a better understanding of palmitoylation and the DHHC proteins in *Plasmodium* parasites is still required before this can be successfully achieved. Nevertheless, the work described in this dissertation has provided new information about this lesser known protein modification in *Plasmodium*, and also provided useful assays that can hopefully contribute to achieving this goal.

References:

1. Resh, M.D., *Palmitoylation of ligands, receptors, and intracellular signaling molecules*. Sci STKE, 2006. **2006**(359): p. re14.
2. Linder, M.E. and R.J. Deschenes, *Palmitoylation: policing protein stability and traffic*. Nat Rev Mol Cell Biol, 2007. **8**(1): p. 74-84.
3. Mitchell, D.A., et al., *Protein palmitoylation by a family of DHHC protein S-acyltransferases*. J Lipid Res, 2006. **47**(6): p. 1118-1127.
4. Rees-Channer, R.R., et al., *Dual acylation of the 45 kDa gliding-associated protein (GAP45) in Plasmodium falciparum merozoites*. Mol Biochem Parasitol, 2006. **149**(1): p. 113-116.
5. Moskes, C., et al., *Export of Plasmodium falciparum calcium-dependent protein kinase 1 to the parasitophorous vacuole is dependent on three N-terminal membrane anchor motifs*. Mol Microbiol, 2004. **54**(3): p. 676-691.
6. Russo, I., A. Oksman, and D.E. Goldberg, *Fatty acid acylation regulates trafficking of the unusual Plasmodium falciparum calpain to the nucleolus*. Mol Microbiol, 2009. **72**(1): p. 229-45.
7. Jones, Matthew L., et al., *Analysis of Protein Palmitoylation Reveals a Pervasive Role in Plasmodium Development and Pathogenesis*. Cell Host & Microbe, 2012. **12**(2): p. 246-258.
8. Zhu, W., J.W. Smith, and C.-M. Huang, *Mass Spectrometry-Based Label-Free Quantitative Proteomics*. Journal of Biomedicine and Biotechnology, 2010. **2010**: p. 1-6.
9. Bantscheff, M., et al., *Quantitative mass spectrometry in proteomics: a critical review*. Analytical and Bioanalytical Chemistry, 2007. **389**(4): p. 1017-1031.
10. Roth, A.F., *The yeast DHHC cysteine-rich domain protein Akr1p is a palmitoyl transferase*. The Journal of Cell Biology, 2002. **159**(1): p. 23-28.
11. Lobo, S., *Identification of a Ras Palmitoyltransferase in Saccharomyces cerevisiae*. Journal of Biological Chemistry, 2002. **277**(43): p. 41268-41273.
12. Roth, A.F., et al., *Global analysis of protein palmitoylation in yeast*. Cell, 2006. **125**(5): p. 1003-1013.
13. Ikadai, H., et al., *Transposon mutagenesis identifies genes essential for Plasmodium falciparum gametocytogenesis*. Proceedings of the National Academy of Sciences, 2013. **110**(18): p. E1676-E1684.
14. Collins, C.R., et al., *Robust inducible Cre recombinase activity in the human malaria parasite Plasmodium falciparum enables efficient gene deletion within a single asexual erythrocytic growth cycle*. Mol Microbiol, 2013. **88**(4): p. 687-701.
15. Cabrera, A., et al., *Dissection of Minimal Sequence Requirements for Rhoptry Membrane Targeting in the Malaria Parasite*. Traffic, 2012. **13**(10): p. 1335-1350.
16. Swarthout, J.T., *DHHC9 and GCP16 Constitute a Human Protein Fatty Acyltransferase with Specificity for H- and N-Ras*. Journal of Biological Chemistry, 2005. **280**(35): p. 31141-31148.
17. Yanai, A., et al., *Palmitoylation of huntingtin by HIP14 is essential for its trafficking and function*. Nature Neuroscience, 2006. **9**(6): p. 824-831.