Distribution Agreement

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my dissertation in whole or in part in all forms of media, now or hereafter known, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this dissertation. I also retain the right to use in future works; all or part of this dissertation.

_____   _____
            Sheila Akinyi   Date

Proteomic Identification and Biochemical Characterization of Modifications Produced on

Plasmodium cynomolgi and *Plasmodium knowlesi* Infected Erythrocyte Membranes

By

Sheila Akinyi

Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences

Immunology and Molecular Pathogenesis

_____

Mary R. Galinski, Ph.D.

Advisor

_____   _____

Rustom Antia, PhD   Jan Mead, PhD

Committee Member   Committee Member

_____   _____

Patrick J Lammie, PhD   Alberto Moreno, M.D.

Committee Member   Committee Member

_____

Venkatachalam Udhayakumar, Ph.D.

Committee Member

Accepted

_____

Lisa A. Tedesco, Ph.D.

Dean of The James T. Laney School of Graduate Studies

_____

Date

Proteomic Identification and Biochemical Characterization of Modifications Produced on
Plasmodium cynomolgi and *Plasmodium knowlesi* Infected Erythrocyte Membranes

By

Sheila Akinyi

B.S., State University of New York, College at Cortland, 2003

Advisor

Mary R. Galinski, Ph.D.

An abstract of

A dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies
of Emory University in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Immunology and Molecular Pathogenesis Program
2010

Abstract

Proteomic Identification and Biochemical Characterization of Modifications Produced on

Plasmodium cynomolgi and *Plasmodium knowlesi* Infected Erythrocyte Membranes

By Sheila Akinyi

Plasmodium parasites modify their host red blood cell (RBC) antigenically and structurally, enabling parasite protein export, nutrient import, and RBC adherence. While knobs and Maurer's clefts are observed in *P. falciparum*-infected erythrocytes, RBCs infected with other *Plasmodium* species display alterations such as caveolae and caveola-vesicle complexes (CVCs). Information on the antigenic makeup and function of these ultrastructures is scarce. This dissertation utilizes genomics, proteomics and biochemical tools as well as the simian malaria species *P. knowlesi* and *P. cynomolgi* to characterize and gain functional insights into these structures.

This dissertation first focuses on investigations of glyceraldehyde-3-phosphate dehydrogenase, an enzyme involved in glycolysis for ATP production. We identified *Plasmodium*-specific amino acid substitutions within the functional domains of this protein that may prove to be of importance, thus providing insights for antimalarial drug-targeting studies.

Secondly, we identified the gene encoding a protein known to be part of the CVCs of *P. vivax*, and believed to have an ortholog in the related parasite, *P. cynomolgi*. Immunoprecipitation of detergent extracts followed by trypsin peptide cleavage and mass spectrometric analysis identified this protein as an 81 kDa member of the *Plasmodium* helical interspersed sub-telomeric (PHIST) superfamily, termed PcyPHIST-81, with homologs in five other *Plasmodium* species. All PHIST-81 homologs have a *Plasmodium* export motif, conserved tryptophans, and four consecutive alpha helices in their amino acid sequences. Immunoelectron tomography studies on *P. cynomolgi*-infected RBCs using rabbit antisera to recombinant *P. vivax* PHIST-81 revealed that PcyPHIST-81 localizes to the tubular extensions of CVCs.

Finally, we utilized a global approach to identify proteins that are localized to the membranes of *P. cynomolgi* and *P. knowlesi*-infected erythrocytes. Mass spectrometric analysis of parasite-infected RBC membrane ghosts identified 109 *P. cynomolgi* proteins and 129 *P. knowlesi* proteins, which included members of the PHIST and Pk-fam-c protein families as well as hypothetical proteins.

Proteomic Identification and Biochemical Characterization of Modifications Produced on
Plasmodium cynomolgi and *Plasmodium knowlesi* Infected Erythrocyte Membranes

By

Sheila Akinyi

B.S., State University of New York, College at Cortland, 2003

Advisor

Mary R. Galinski, Ph.D.

A dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies
of Emory University in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Immunology and Molecular Pathogenesis Program
2010

Table of Contents

List of Tables and Figures

CHAPTER FIVE

CHAPTER ONE

MALARIA AND THE RED BLOOD CELL MEMBRANE:

ULTRASTRUCTURAL MODIFICATIONS, SIMIAN MODELS AND NOVEL

MEMBRANE PROTEINS

1.1. Malaria

Malaria is estimated to cause about 247 million cases of infection, leading to nearly one million deaths every year (1). Most malaria-related deaths occur in children under the age of five who live in sub-Saharan Africa. More than three billion people are at risk of infection from this disease, which is endemic in 109 countries (1). It should be noted that the estimates of malaria burden may be inaccurate due to their reliance on clinical statistics which, in many cases, are not confirmed by laboratory diagnoses such as positive blood smears or rapid diagnostic testing (RDT) (Reviewed in (2)). Moreover, light microscopy and RDTs may not detect parasites in individuals who have gametocytes, and are capable of transmitting these to a mosquito (3).

The asexual form of *Plasmodium* is responsible for malaria-related pathology. There are over 120 known *Plasmodium* species, five of which cause malaria in humans: *Plasmodium falciparum*, *P. vivax*, *P. malariae*, *P. ovale* and the simian malaria species *P. knowlesi*. The first four are considered human malaria species, but *P. knowlesi* has also been shown to be responsible for hundreds of malaria cases in humans (4, 5). While *P. falciparum* causes the most virulent and deadly infection, *P. vivax* is also a significant contributor to malaria-associated morbidity and mortality (Reviewed in (6-10)). The other *Plasmodium* species infect several different vertebrates, some of which are utilized as laboratory models to enhance our understanding of human malaria pathogenesis and parasite biology; they include monkey, ape and rodent species (11).

Interventions proposed by the World Health Organization (WHO) for the control of malaria include indoor residual spraying of insecticides to kill or deter the anopheline mosquito vector, the use of long-lasting insecticide-treated bed nets to reduce the

incidences of mosquito bites, intermittent preventive malaria treatment for women during pregnancy in regions with high malaria transmission, prompt and accurate malaria diagnosis, and access to appropriate antimalarial drug therapy (1).

Efforts to control malaria in South America, Southeast Asia and other regions have been complicated by increased *P. falciparum* resistance to former first-line antimalarial drugs chloroquine and sulfadoxine-pyrimethamine (S/P) (12, 13). This problem has fueled renewed efforts to create alternative therapeutics for malarial treatment that are effective against blood stage parasites and are relatively non-toxic and affordable. Currently, the most effective drugs available are artemisinin-based combination (ACT) therapies, which combine an artemisinin derivative with a longer-lasting partner drug. Artemisinin is a very potent but short-lived antimalarial endoperoxide purified from the *Artemisia annua* plant. Artemisinin derivatives rapidly clear young, circulating ring-stage parasites, prevent parasite maturation and sequestration (14, 15), and destroy gametocytes, thereby reducing parasite transmission. The partner drug functions to reduce parasitemia in the long term. If malaria is to be eliminated, and eventually eradicated (16), developing antimalarials and vaccines with transmission-blocking characteristics should become a high priority. Examples of ACTs currently in use include artemether-lumefantrine, artesunate-amodiaquine, artesunate-mefloquine and artesunate-sulfadoxine-pyrimethamine (1). A fifth ACT, dihydroartemisinin-piperaquine, was recently added to the treatment options (17).

The WHO currently recommends ACTs for the treatment of *P. falciparum* infections, and a 14-day ACT regimen in combination with primaquine for the treatment of *P. vivax* infections, taking into account the risk of primaquine-induced hemolysis in

glucose-6-phosphate dehydrogenase-deficient individuals (1). ACT treatment combined with distribution of insecticide-treated bednets produced a ten-fold decrease in the prevalence of parasitemia in Zanzibar within one year (18). The administration of artemisinin derivatives in combination with vector control efforts has also had great success in Eritrea (19) and KwaZulu Natal, South Africa (20). Recently, however, resistance to artemisinin has emerged on the Thai-Cambodian border and in southern Cambodia, findings that have grave implications for malaria control and elimination efforts (21, 22). In light of these recent developments, it is imperative that the discovery and development of novel antimalarial drugs that are safe and fast-acting takes precedence.

The emergence of vector resistance to insecticides and parasite resistance to drugs needs to be closely monitored. The evolution of parasite- and vector-resistant strains is almost inevitable since there are currently very few effective compounds utilized for widespread antimalarial treatment and mosquito control, creating an environment where large-scale selection of parasite- and vector-resistant strains occurs. For example, both dichlorodiphenyltrichloroethane (DDT) and pyrethroids, which are currently utilized for indoor-residual spraying and insecticide-treated nets respectively, target the same voltage-gated sodium channel protein responsible for neuronal signal transmission (Reviewed in (2)), and research shows that mutations that confer resistance to DDT also confer resistance to pyrethroids (23).

P. vivax research has lagged behind that of *P. falciparum* despite its relevance as a parasite of global importance (6-8). This disparity is made evident in the number of vaccine candidates undergoing preclinical or clinical trials for either species – 70 for *P.*

falciparum versus only nine for *P. vivax*. For *P. vivax*, vaccine candidates include: two pre-erythrocytic stage antigens (Thrombospondin-Related Anonymous Protein and Circumsporozoite Protein), six erythrocytic stage antigens (Duffy Binding Protein, Merozoite Surface Proteins-1, -3 and -9, Apical Membrane Antigen-1 and Reticulocyte-Binding Proteins) and the sexual stage antigen, Pvs25 (8, 24-26). Unfortunately, *P. falciparum* vaccines will not be cross-protective against *P. vivax* infections due to the phylogenetic and antigenic differences between the two species. Therefore, there is a need for more rigorous research into *P. vivax* candidate vaccine development and drug discovery.

Of all vaccine candidates, the most advanced in clinical trials is the pre-erythrocytic stage vaccine RTS,S; it is a fusion protein consisting of the central tandem repeat and C-terminal regions of the *P. falciparum* circumsporozoite protein (CSP) fused with the N-terminal of the Hepatitis B virus surface (S) antigen, combined with the adjuvant AS02 (27). Three doses of RTS,S delayed the onset of parasitemia by nine weeks in 74% of Gambian adults, but this efficacy waned with time; a booster dose the following season was efficacious for nine weeks in 47% of the immunized adults (28). When three doses of the vaccine were administered to infants and children aged 1-4, the vaccine was reported to be efficacious against clinical malaria in 35% of the subjects and effective against severe malaria in 49% of those studied for a period of 18 months (29, 30). Moreover, it was found that RTS,S maintained its efficacy when lyophilized ASO2 was reconstituted at the time of immunization rather than being coformulated during vaccine manufacture (31).

For a long time, malaria drug development and vaccine efforts were almost exclusively focused on *P. falciparum*, the main cause of malaria-related deaths. However, with a shift of focus from malaria control (the situation where malaria causes significantly fewer illnesses and deaths) to malaria elimination (the situation where malaria transmission in a certain region has ceased, and can only be introduced from another area) and eventual eradication (when the *Plasmodium* parasite completely disappears), it has become clear that further research is needed in gaining a deeper understanding of the biology and pathogenesis of the other three human malaria species, especially *P. vivax* (Reviewed in (8)). In 2007, Bill and Melinda Gates challenged the scientific community to make malaria eradication its goal (32). Eradicating *P. vivax* and *P. ovale* will be more challenging than *P. falciparum* because the former two species can form hypnozoites (dormant forms of the parasite, which can cause a relapse weeks, months or even years after the primary infection). Large-scale treatment will require the administration of a hypnozoite-killing drug that is much safer than primaquine, which is known to cause hemolysis in glyceraldehyde-6-phosphate dehydrogenase-deficient individuals (Reviewed in (2)).

Although current interventions may not be adequate to meet the ambitious goal of malaria eradication, scaling up current control efforts such as the use of insecticide-treated bed nets and the correct administration of effective drugs will go a long way in achieving malaria elimination in many areas. Additionally, the concerted efforts of programs such as the President's Malaria Initiative and the Global Fund to Fight AIDS, Tuberculosis and Malaria, coordinated with a push to increase funding from the World

Bank and the governments of affected countries, is vital to achieving the goals of malaria elimination and eventual eradication (32).

The following sections of this chapter cover the life cycle of *Plasmodium*, symptoms of malaria infection and an overview of the simian malaria species used as models for *P. vivax* infection: *P. cynomolgi* and *P. knowlesi*. It should be noted that because *P. knowlesi* merozoites are very stable, the species is an ideal model for studying erythrocyte invasion mechanisms, which can be applied to both *P. falciparum* and *P. vivax* studies (33). Additionally, the antigenic and structural modifications produced in the erythrocyte after parasite invasion are discussed. In particular, the significant progress that has been made in understanding *P. falciparum*-induced alterations to infected RBCs is reviewed. Furthermore, this chapter presents current knowledge of the characteristics of exported parasite proteins, and introduces a recently identified novel family of proteins, the *Plasmodium* helical interspersed subtelomeric (PHIST) family.

1.2. The *Plasmodium* Life Cycle

The *Plasmodium* life cycle is complex, involving both vertebrate and invertebrate hosts. It consists of three main phases: the sporogonic, pre-erythrocytic and erythrocytic stages.

The sporogonic stage occurs inside the female *Anopheles* mosquito. The mosquito injects sporozoites from its salivary glands into the skin of a human host when it takes a blood meal. Then, the sporozoites migrate to the liver via the bloodstream and invade hepatocytes, initiating the pre-erythrocytic stage.

Inside hepatocytes, the sporozoites mature into liver schizonts and multiply

asexually for about 2-16 days, depending on the *Plasmodium* species, to produce

thousands of merozoites (Reviewed in (34, 35)). These merozoites rupture out of the

hepatocytes and migrate through the bloodstream within membrane-bound vesicles

known as merosomes, which enable them to avoid immune detection (36). Merosomes

were detected by studying the rodent species *P. berghei* (36), and have also been detected

in *P. yoelii* (37) but are presumed to be present in other species. While in the liver, certain

species including *P. vivax*, *P. ovale* and *P. cynomolgi* may produce hypnozoites (38).

Once in the bloodstream, the merozoites invade red blood cells (RBCs) in a

process that involves recognition, attachment, reorientation and entry (Reviewed in (39)).

This signals the start of the erythrocytic stage. When a merozoite binds to a RBC, it

reorients using actin filaments and myosin-based motors to bring its apical end in contact

with the erythrocyte; this process is irreversible (Reviewed in (40)). *Plasmodium*

rhoptries and micronemes, which are specialized organelles located at the apical end of

the merozoite, secrete proteins that structurally alter the RBC membrane, causing

membrane invagination, RBC cytoskeletal stripping at the point of attachment, and band

3 degradation (41). Other changes to the host cell include the phosphorylation of RBC

membrane skeletal components and a transient decrease in RBC membrane rigidity (42,

43). Then, by the action of actin and myosin motors, the merozoite enters the RBC

through this prepared region (44).

Once RBC invasion is complete, the *Plasmodium* parasite is enclosed in a

parasitophorous vacuole (PV) that enlarges as the parasite grows and matures. The PV

membrane is formed from a combination of RBC and parasite material, and undergoes

extensive modification after *Plasmodium* invasion and during development (45, 46). The intraerythrocytic parasite digests cytoplasmic hemoglobin as a nutrient source, and polymerizes the undigested heme residue to produce a pigment known as hemozoin.

Within RBCs, merozoites begin cycles of multiplication known as schizogony, which involve the development of rings that then differentiate into trophozoites. The ring and trophozoite stages are marked by increased RNA, DNA and protein synthesis, and the beginning of hemoglobin digestion. *P. falciparum* digests more than 80% of its host's hemoglobin to support growth and replication (47, 48). The parasite then matures to first form multinucleated, then segmented, schizonts by asexual reproduction, resulting in the production of daughter merozoites. The merozoites then rupture out of erythrocytes to invade uninfected RBCs and initiate the next round of the erythrocytic stage (34, 35). *P. vivax* merozoites preferentially invade reticulocytes and the reticulocyte-binding proteins PvRBP-1 and PvRBP-2 have been proposed as critical proteins that may target these host cells (49).

During the erythrocytic stage, some of the merozoites will differentiate into sexual forms known as the male and female gametocytes. When a mosquito takes up these gametocytes during a blood meal, they migrate to the mosquito gut where the male gametocyte (microgamete) fertilizes the female gametocyte (macrogamete) forming a zygote that develops into an ookinete. The ookinete migrates through the midgut epithelium and basal lamina where it differentiates into an oocyst. Within two weeks of differentiation, the oocyst produces daughter sporozoites by asexual multiplication, and the sporozoites migrate to the mosquito's salivary glands where they can then be transmitted to susceptible hosts, beginning a new life cycle for the *Plasmodium* parasite.

P. vivax gametocyte development can occur at lower temperatures than *P. falciparum* parasites require in mosquitoes, and this is the reason why *P. vivax* infections are predominant outside of tropical regions (Reviewed in (2)).

The development of novel interventions to break the cycle of infection will require a better understanding of the molecular mechanisms the parasite utilizes to recognize and invade erythrocytes, form gametocytes and then succeed in thriving in the mosquito host.

1.3. Clinical Symptoms of Malaria

All of the clinical symptoms associated with malaria are caused by the asexual erythrocytic stage of *Plasmodium*. Malaria infections are classified as uncomplicated or complicated (severe) depending on whether they affect the function of vital organs. Clinical symptoms of uncomplicated malaria include fever, sweats, nausea and headaches.

Severe malaria manifests itself through such conditions as cerebral malaria, severe malarial anemia, hyperparasitemia, hemoglobinuria, metabolic acidosis, cardiovascular collapse, pulmonary edema, severe thrombocytopenia, splenic rupture and acute respiratory distress syndrome (50). Although *P. falciparum* has been thought to cause the majority of severe malaria cases in humans, *P. vivax* is increasingly being identified as a significant cause of severe malaria (Reviewed in (9, 51)).

1.4. *Plasmodium vivax*, *P. knowlesi* and *P. cynomolgi*

Long term *P. vivax in vitro* culture is limited by a number of factors, mainly that this species is reticulocyte-specific, leaving the generation of parasite materials to either field isolates or experimental infections in New World monkeys. *P. vivax*, *P. cynomolgi* and *P. knowlesi* are phylogenetically closely related (52), therefore the latter two species have provided good *in vitro* and *in vivo* simian models for *P. vivax* infections because of the ease in obtaining large quantities of parasite samples from infection of rhesus macaques.

P. knowlesi is recognized as an important model for studying RBC invasion and antigenic variation of proteins inserted into the erythrocyte membrane (Reviewed in (53, 54)). The direct relevance of this species to human malaria research is underscored by the documented cases of *P. knowlesi* natural infections in humans in Malaysia, Thailand and other parts of Southeast Asia (4, 5, 55, 56). Furthermore, the recent completion of *P. knowlesi* (57) and *P. vivax* (58) genomes contributes to databases from which crucial genes can be determined. In contrast to *P. vivax*, *P. falciparum* and *P. cynomolgi*, all of which have 48-hour cycles of schizogony, intraerythrocytic *P. knowlesi* has a 24-hour cycle.

In addition to *P. falciparum* (59), *P. vivax* (58) and *P. knowlesi* (57), the genome of the simian malaria parasite *P. cynomolgi* is being completed and will be published soon. Ongoing genome sequencing projects for several isolates of *P. falciparum* and *P. vivax* will contribute to databases from which genes that are crucial to the survival of *Plasmodium* within its host can be determined. While identifying antigens to analyze as potential drug and vaccine targets remains a challenge, newly available proteomics and

genomics techniques make it possible to selectively analyze proteins of potential importance.

1.5. The Red Blood Cell Membrane

The mature human RBC has been referred to as a 'floating corpse' (60) because it is denucleated and lacks protein synthesis and trafficking capabilities present in other eukaryotic cells. A multinucleated reticulocyte originates from the bone marrow, travels through the circulation and matures within about 48 hours. The average life span of a human RBC is 120 days, during which time it continuously undergoes massive deformation in order to traverse the narrow capillaries of the microvasculature. The erythrocyte is unique in that its plasma membrane accounts for all of its antigenic and mechanical characteristics (Reviewed in (61)). Unfortunately, current knowledge of erythrocyte proteins accounts for only 15% of the 300 proteins recently identified by proteomics analysis (62).

The RBC membrane is composed of equal proportions (by weight) of cholesterol and anionic phospholipids, anchored to a two-dimensional network of skeletal proteins (63). The RBC cytoskeleton is tethered to the plasma membrane through sites on the cytoplasmic domains of transmembrane proteins embedded in the lipid bilayer (64). In addition, many erythrocyte cytoskeletal proteins interact with the anionic phospholipids facilitating additional cytoskeletal attachments to the lipid bilayer (65, 66). Cholesterol is distributed evenly between the inner and outer leaflets of the RBC membrane, while the phospholipids are asymmetrically distributed: phosphatidylcholine and sphingomyelin are mostly located on the outer leaflet, while most of the phosphatidylethanolamine and

all the phosphatidylserine are concentrated on the inner leaflet (67, 68). The cytoskeleton

provides stability and support to the plasma membrane, and regulates erythrocyte shape

and deformability while maintaining the cell's surface area (69, 70). The major

components of the RBC skeletal network are spectrin, actin, protein 4.1, adducin,

dematin, tropomyosin and tropomodulin (63, 71-73).

1.6. *Plasmodium* and the Red Blood Cell Membrane

When malaria parasites invade RBCs, they impose greater metabolic demands on

the cells than is the case for uninfected, quiescent erythrocytes, in order to support

parasite growth and development. The glucose uptake in a *Plasmodium*-infected RBC is

10-40 fold greater than that in an uninfected erythrocyte (74). In addition, the parasite

induces ultrastructural, antigenic, biochemical and functional alterations in the plasma

membrane and erythrocyte cytoplasm. The ultrastructural alterations assume different

morphological forms, depending upon the *Plasmodium* species infecting the erythrocyte.

The *Plasmodium* parasite remodels the host RBC by establishing a complex

membrane network that connects the PV to the host plasma membrane and is involved in

protein transport and trafficking, allows for import of nutrients, export of waste products,

host cell lysis, and affects the adherence properties of infected RBCs. The requirement

for an extraparasitic secretory system poses a challenge that most intracellular organisms

do not have to face, and has led to increased interest in investigating these parasite-

induced membrane-bound compartments.

Once parasite proteins are inserted into the RBC membrane, the properties of the

membrane change as well (Reviewed in (75-79)):

- Species-specific structures such as knobs, clefts, caveolae and caveola-vesicle complexes are produced.

- The infected RBC loses its normal discoid shape.

- There are changes to the phosphorylation states of RBC membrane skeletal proteins.

- The fluidity of RBC membrane phospholipids changes.

- There is a shift in the lateral distribution of host RBC integral membrane proteins.

- The deformability of parasitized RBCs decreases.

- RBC membrane permeability to a number of small molecular weight solutes is altered.

- Hidden epitopes of membrane proteins on uninfected RBCs are exposed on the surface of *Plasmodium*-infected RBCs.

- *Plasmodium* antigens are inserted into the RBC membrane.

- Membrane-localized parasite proteins interact with RBC membrane proteins. These *Plasmodium* protein-RBC protein interactions are important for parasite biology and malaria pathogenesis.

Parasite proteins that associate with the RBC cytoskeleton are insoluble in nonionic detergents such as Triton X-100 (Reviewed in (80)). Known interactions among *P. falciparum* and RBC proteins include the mature parasite-infected erythrocyte surface antigen (MESA) with protein 4.1 (81), ring-infected erythrocyte surface antigen (RESA) with spectrin (82), and *P. falciparum* erythrocyte membrane protein-1 (PfEMP-1) with actin and spectrin (83). These interactions are discussed further in section **1.7.b**.

1.7. *Plasmodium falciparum*-infected RBC Ultrastructures and Proteins

1.7. a. RBC remodeling by *P. falciparum*

The complete sequencing and annotation of the *P. falciparum* genome (59), development of transfection systems for asexual stage parasites (84), increasingly sophisticated morphological analyses, and the identification of novel parasite proteins exported to the RBC have all advanced knowledge of host RBC modifications by this species.

In the absence of endogenous trafficking machinery, the *Plasmodium* parasite faces the challenge of transporting proteins across multiple membranes: the parasite's plasma membrane, the PV membrane and, for some proteins, the RBC membrane. Active areas of research pertaining to protein trafficking in *P. falciparum*-infected RBCs include:

1. The identification of machinery/molecules that facilitate protein transport across the PVM, the RBC cytoplasm and insertion into the host RBC membrane

2. The identification and mechanistic action of signals that direct proteins to particular extracellular destinations, and

3. Determining whether trafficking pathways represent novel mechanisms and molecules that have evolved in adaptation to life inside the RBC, or if the parasite exports proteins that are already involved in intraparasitic protein trafficking, thus utilizing biologically conserved mechanisms and molecules (85).

All *P. falciparum* proteins analyzed to date that are detected in the infected RBC cytoplasm traverse the parasitophorous vacuole (86-89). Soluble proteins are moved into the PV when transport vesicles fuse with the parasite plasma membrane (89, 90). Some

proteins remain in this compartment whereas others, such as the knob-associated histidine-rich protein (KAHRP) and *P. falciparum* erythrocyte membrane protein 3 (PfEMP-3), are directed outwards across the PVM. There is evidence to suggest that PV-resident and forward-destined proteins might be separated into different sub-compartments within the PV (88, 91). Retrograde transport of membrane vesicles from the RBC membrane has not been observed in *P. falciparum* (92), although a novel pathway seems to allow uptake of raft-associated proteins from the RBC membrane (93).

1.7. b. *P. falciparum* trafficking beyond the parasitophorous vacuole in infected RBCs

As *P. falciparum* matures, structures known as Maurer's clefts (MCs) are formed in the RBC cytoplasm (94, 95). By electron microscopy (EM), MCs appear to be flattened, membranous and disk-like with translucent lumens and electron-dense coats (96). MCs are visible at the early trophozoite stage of parasite development and persist throughout the remainder of the erythrocytic cycle. Initially, the twisted and branched MC structures are located near the PV (97, 98), but they gradually relocate closer to the RBC membrane skeleton and are linked to the RBC membrane by tubular tethers (99). The degree of elaboration of MCs appears to vary between different *P. falciparum* strains, where observations range from singular to a few tubular or flat cisternae surrounded by vesicles of various sizes (100-102).

The origin of MCs is not known, and researchers have debated over the years whether peripheral MCs are independent structures or subdomains of the tubulovesicular network (TVN) (94, 103-105). Earlier data from fluorescence microscopy on RBCs

labeled with lipid probes, as well as from serial EM sectioning, suggested that MCs

formed a continuous network (98, 105, 106). However, live cell imaging of green

fluorescent protein (GFP) chimeras of MC-associated cargo and the TVN indicated that

MCs were distinct, independent entities (96, 107). Consistent with this view was the

observation that reporter proteins transported to the PV did not diffuse into the lumen of

MCs (88, 89). Perhaps the most compelling data thus far has been from recent whole cell

imaging studies using three-dimensional structured illumination microscopy and electron

tomography, which support the hypothesis that MCs are not connected to any extensions

of the PV membrane, but rather occur as independent entities (99). This study also

provided further evidence that MCs in *P. falciparum* 3D7 parasites form tethers that

facilitate their attachment to the RBC membrane (99).

In three-dimensional EM reconstructions of *P. falciparum*-infected erythrocytes

25 nm vesicles were observed in the vicinity of MCs and close to the RBC membrane

(106, 108, 109). They were distinct in size and appearance from the coated vesicles

involved in ER-to-Golgi transport in the parasite cytoplasm, and might be involved in

protein trafficking between MCs and the RBC membrane.

1.7. c. Proteins associated with *P. falciparum* ultrastructures

Most *P. falciparum* proteins destined for the RBC membrane, cytoskeleton and

surface traffic through MCs (105, 107), with proteins such as the Maurer's cleft-

associated histidine rich protein (MAHRP), ring stage exported protein (REX) and the *P.*

falciparum skeleton binding protein 1 (PfSBP1) residing in these structures (110-112).

PfSBP1is a 48 kDa integral membrane protein that spans the MC membrane (110), with

its N-terminal domain located within the cleft and its C-terminal domain exposed to the

infected RBC cytoplasm and interacting with a RBC membrane skeleton protein (110).

This interaction probably serves to anchor the MCs to the RBC membrane skeleton (110).

PfSBP1lacks a classical N-terminal signal peptide (110), yet it is targeted outside the

parasite. MAHRP is also an MC-resident protein that lacks an N-terminal secretory

signal, has a long phenylalanine-rich transmembrane domain and might be inserted into

MCs with the C-terminal region facing the RBC cytosol (112).

Other proteins that associate transiently associate with MCs *en route* to their final

destinations include Pf332, KAHRP, PfEMP-1 and PfEMP-3 (Reviewed in (113)); the

subtelomeric variable open reading frame (STEVOR) proteins (114) and the Maurer's

cleft two-transmembrane domain proteins (Pfmc2-tm) (115).

Several other novel proteins identified by proteomics analysis are also thought to

associate with MCs (116). Since MCs interact with the parasitized RBC membranes

throughout the intraerythrocytic stage, they can be recovered in ghost preparations, which

consist of the RBC membrane and submembrane skeleton (116).

P. falciparum proteins associated with the infected RBC membrane include

PfEMP-1, the knob-associated histidine-rich protein (KAHRP), glyceraldehyde-3-

phosphate dehydrogenase (G3PDH) and rhoptry-associated proteins (116). Membrane

associated proteins like PfEMP-3, RESA, MESA, *falciparum*-interspersed repeat antigen

(FIRA) and *falciparum*-exported serine-threonine kinase (FEST) appear to be distributed

evenly across the host RBC skeleton (Reviewed in (113)).

The unit membranes surrounding the TVN and the PVM lack the electron-dense coat associated with MCs, and have distinct antigen profiles e.g. exported proteins-1 and –2 are present in the PVM and the TVN network but not in MCs (117).

The functions of some of these exported proteins have been elucidated, while most remain unknown. For example, deleting the *kahrp* gene blocks knob formation on the RBC membrane, abrogating RBC adherence to the vascular endothelium (118). Moreover, disruption of the *pfemp-3* gene hinders PfEMP-1 trafficking to the RBC surface, also decreasing cytoadherence (119).

A number of recent studies have identified and characterized the binding domains on parasite proteins that interact with the RBC membrane, with the goal of developing novel therapeutics that would interfere with those interactions that are found to be essential for parasite growth and/or virulence (Reviewed in (113)). For example, MESA binds to protein 4.1 through a 19-amino acid sequence (81), while RESA binds to spectrin via a 48-residue domain (120). The N-terminal of KAHRP contains a spectrin-binding domain and it also binds the cytoplasmic tail of PfEMP-1(121). Some malaria protein interactions disrupt associations between RBC proteins, causing changes to the morphological or mechanical characteristics of the host cell. For example, MESA competes for the same protein 4.1 domain that glycophorin C and p55 bind to, interfering with the binding of the latter proteins (122).

1.8. RBC Ultrastructures and Proteins in other *Plasmodium* species

Morphological changes in other *Plasmodium*-infected RBCs were first observed as stippling of the erythrocyte membrane by light microscopy on Romanowsky-stained

blood films (Reviewed in (123)). Depending on the *Plasmodium* species, the stippling was referred to as Schuffner's dots (*P. vivax*, *P. ovale*), Sinten and Mulligan's stippling (*P. cynomolgi, P. simium*), or Ziemann's stippling (Reviewed in (123)). Schuffner's stippling appears as multiple, small dots in Giemsa-stained thin films of *P. vivax*, *P. ovale*, *P. cynomolgi* and *P. simium* infected erythrocytes, and it is important for malaria diagnosis (124). Immunofluorescence microscopy studies using *P. vivax* and *P. cynomolgi* immune serum have shown that the stippling observed on infected RBCs corresponds to the presence of parasite antigens (125). Electron microscopic studies later confirmed the presence of ultrastructural modifications induced by various human and simian *Plasmodium* species when they invaded RBCs. While *P. vivax*, *P. cynomolgi* and *P. simium* form caveola-vesicle complexes (126-128), *P. knowlesi* produce caveolae on infected erythrocyte membranes. And similar to *P. falciparum*, the human malaria species *P. malariae* as well as simian species *P. coatneyi* and *P. brasilianum* form knobs on RBC membranes.

1.8. a. Caveolae

Plasmodium caveolae are flask-shaped invaginations of the membrane skeleton that are thought to play a role in importing host plasma proteins by pinocytosis or exporting parasite-derived proteins (Reviewed in (80)). Electron microscopy studies indicate that these structures are about 90 nm in diameter with flattened bases, and their inner surface is coated with an electron-dense fuzzy material. Caveolae resemble the coated pits of nucleated eukaryotic cells (97) and they increase in number as the intraerythrocytic parasite matures.

Caveolae are found in almost all eukaryotic cell types, but are in greater abundance in endothelial cells, adipocytes, muscle cells and fibroblasts. In these cells, caveolae help incorporate membrane components (e.g. glycosphingolipids and glycosylphosphatidylinositol-anchored proteins), extracellular ligands (e.g. folic acid, albumin and autocrine motility factors), bacterial toxins (e.g. cholera and tetanus toxins), and several non-enveloped viruses (e.g. SV40 and polyoma virus; (129, 130). Caveolae may also play a role in endogenous processes such as calcium and cholesterol homeostasis, cellular signaling, recycling of GPI-anchored proteins, glycosphingolipid transport and transcytosis of serum components (129-133).

1.8.b. Caveola-vesicle complexes (CVCs)

CVCs are 50 nm spherical or tube-like membrane-bound vesicles that cluster around the base of caveolae (124, 128, 134). These ultrastructures have been observed in *P. vivax*, *P. cynomolgi*, *P. simium* and *P. ovale* infected RBCs, and their numbers increase as the parasite matures from the late ring stage to a mature trophozoite (97). It is not known how *Plasmodium* caveolae and CVCs are inserted into the RBC cytoskeleton, although it has been suggested that parasite proteins accumulating at specific sites may initiate membrane invagination leading to their formation (97). The antigenic composition and function of CVCs is not known, although it has been proposed that they may function in endocytosis (128). These structures have only been observed in RBCs infected by primate *Plasmodium* species; none to date have been reported in other mammalian, reptilian or avian *Plasmodium* species.

1.9. Proteins Associated With Caveolae and Caveola-vesicle Complexes

Immunofluorescence microscopy provided the first evidence that malarial antigens, which had originally been thought to be parasite-altered host antigens, associated with *Plasmodium* ultrastructures (125, 135, 136). Later, immunoelectron microscopic analyses determined that while some parasite antigens localized to the CVCs of *P. vivax*-infected RBCs, others were detected on the 50-60 nm cytoplasmic vesicles (124, 134, 137). It has been proposed that these freely floating vesicles transport CVC-destined proteins across the infected RBC cytoplasm and then fuse with the caveolar membrane at the erythrocyte surface to release their contents to the extracellular environment (Reviewed in (80)). Since immuno-EM did not detect these CVC-localized parasite proteins in association with the intracellular parasite, the parasite membrane, PV, or PV membrane, it is likely that upon their production, they are rapidly translocated across these membranes to their final destination (124, 134).

To begin analyzing the trafficking of novel parasite proteins, determine if vesicle-mediated endocytosis occurs, and identify potential protein targeting signals, protein characteristics such as the presence of N-terminal signal sequences, hydrophobic transmembrane domains, and parasite protein export motifs need to be assessed. It is likely that there are as yet unknown proteins located in the various membrane compartments that function as carrier molecules for exported parasite proteins.

1.10. Algorithms to Detect *Plasmodium* Exported/Secreted Proteins

1.10. a. Signal sequences

In most eukaryotic proteins, a hydrophobic sequence of approximately 15 residues located 3-17 amino acids downstream of the N-terminus directs proteins to the secretory pathway. In *Plasmodium*, this N-terminal signal sequence (SS) is frequently encoded on a short first exon and is responsible for co-translational translocation of *Plasmodium* proteins into the ER and export into the PV. Proteins destined for sites in the ER, parasite plasma membrane, PV, PV membrane and apical organelles appear to have this classical SS (138-140).

Alternatively, several *Plasmodium* proteins that are directed past the PVM to the RBC cytosol may contain a longer (up to 30 amino acids) stretch of hydrophobic amino acids that is recessed (up to 80 residues) from the N terminus, and termed the recessed SS (138, 140). It should be noted that this recessed hydrophobic region may not always serve as a signal sequence; for example, proteins destined for the parasite food vacuole have a recessed hydrophobic region which functions as a membrane anchor, and is cleaved only after transit to the parasite plasma membrane (141). Because secretion of proteins containing classical and recessed signal sequences can be inhibited by brefeldin A (BFA) (89, 142, 143), both types of signal sequences can direct translocation across the ER membrane. BFA specifically blocks classical vesicle-mediated trafficking by interfering with the *Plasmodium* ADP ribosylation factor, GDP-GTP exchange protein (a key molecule required for the assembly of coat protein that covers the transport vesicles (144)).

Algorithms designed to predict classical SS, such as SignalP (http://www.cbs.dtu.dk/services/SignalP/)(145), do not easily recognize the recessed hydrophobic signal sequences of exported *Plasmodium* proteins because these sequences are preceded by a 50-amino acid hydrophilic spacer. Therefore, fewer proteins are predicted to be exported than is actually the case. To resolve this, an algorithm targeted to *Plasmodium* proteins known as MalSig (http://bioserve.latrobe.edu.au/cgi-bin/pfsigseq.py) was created, and it recognizes both classical and non-canonical sequences (146). It is not known whether recessed SS are cleaved in the lumen of the ER by signal peptidases or whether the extended N-terminal domains serve additional, as yet unknown, functions.

1.10. b. *Plasmodium* export element/ Host targeting signal (RxLxQ/E/D)

A recently identified *Plasmodium*-specific motif targets parasite proteins for onward transport into the RBC cytoplasm and beyond. Two studies independently identified this second conserved motif, termed either the *Plasmodium* export element (PEXEL)(147) or host targeting signal (HTS)(148), following bioinformatics analyses of aligned N-terminal sequences from proteins known to be exported from the PV into the erythrocyte. Both reports described a short stretch of alternating charged and hydrophobic amino acids separated by uncharged amino acids (RxLxE/Q/D) located a short distance downstream of the signal sequence and provided an *in silico* prediction of proteins exported into the host RBC (an 'exportome'). While the HTS motif (148) utilized a matrix based pattern that was used to search a database of predicted secretory *P*.

falciparum proteins independent of their gene structure, Marti *et al* used the linear

PEXEL signal (147) to search a *P. falciparum* database for a specific two-exon structure.

Many of the proteins identified in these two studies were members of large

variant-antigen families known or predicted to be exported to the host RBC (147, 148).

However, the list also includes a number of proteins for which export out of the parasite

had not been previously shown, including heat-shock proteins, kinases, phosphatases and

putative transporters. Data from both studies also indicated that several proteins for

which strong experimental evidence existed supporting their export out of the parasite,

lacked the PEXEL motif; for example, PfSBP1 and MAHRP that are associated with

cytoplasmic MCs.

Analysis of the transcription profiles of genes encoding proteins bearing a PEXEL

motif indicate that the majority share a temporal profile that is characteristic of proteins

synthesized in the ring stage and exported into the erythrocyte from mid-ring to late ring

stages (147). Reporter constructs in which gene fragments encoding the N-terminal

regions of putative exported proteins are appended to reporter proteins like GFP or

luciferase have been introduced into *P. falciparum* parasites using either transient or

stable transfection systems; the locations of the proteins are then assessed and confirmed

by fluorescence microscopy or biochemical assays (87-89, 149, 150). However, it should

be noted that varying results might be obtained depending on the reporter protein

employed. Using GFP technology, Marti *et al.* demonstrated that non-*P. falciparum*

PEXEL motifs shared the same export function as that in *P. falciparum* by replacing the

P. falciparum PEXEL sequence in GFP fusion constructs (147). It was also confirmed

that non-*P. falciparum* proteins lacking a PEXEL motif could still be exported into the RBC, similar to their *P. falciparum* counterparts (147).

The PEXEL motif was demonstrated to play a role in the export of both soluble and integral membrane proteins out of the parasite (151). PEXEL is cleaved and N-acetylated upstream of the BFA-sensitive trafficking step in the *P. falciparum* secretory pathway, indicating that this step occurs in the parasite's ER, after which the processed N-terminus is recognized within the PV (151). Observations suggested the direct export of soluble proteins from the PV into the RBC cytosol (89, 152) although trafficking them through the PV membrane via a putative translocon also seemed feasible. Recent models have suggested that *P. falciparum* integral membrane proteins are either trafficked by vesicles budding from the PV membrane or by traversing the TVN to their final distribution in MCs or the RBC plasma membrane (90, 104, 106).

While the *Plasmodium* exportome data collected by Marti *et al.* (147) and Hiller *et al.* (148) have aided in annotation efforts and provided a foundation for experimental studies, the two predictive algorithms identify incompletely overlapping sets of proteins. Recently, a new algorithm named ExportPred was developed to predict the *Plasmodium* exportome and determine its conservation across species, from *P. falciparum*, *P. vivax* and *P. knowlesi* to rodent malaria species (153). ExportPred works by utilizing the available genome sequences from different *Plasmodium* species to identify genus-wide exported proteins and determine specific conserved domains (153). This involves the use of a generalized hidden Markov model (GHMM)(154) to model both the SS and the PEXEL motifs required for protein export.

Information obtained from the above studies has contributed to a growing body of literature that describes the multistep processes underlying the trafficking of *P. falciparum*-encoded proteins to the cytoplasm and the surface of infected erythrocytes. The current model proposes protein recruitment to the parasite ER (87, 89), secretion into the PV lumen (89, 107, 149, 155), and transport across the PV membrane mediated by the PEXEL sequence (107, 147, 148, 155). Recent studies have determined that the aspartic acid protease plasmepsin V is responsible for cleaving the PEXEL motif in *P. falciparum* exported proteins, revealing the xE/Q/D portion at the N-terminal (156, 157). This is followed by the transport of cleaved proteins across the PVM into the host RBC by an ATP-dependent translocator, recently identified in *P. falciparum* as the *Plasmodium* translocon of exported proteins (PTEX) complex (158). The PTEX complex consists of the ATPase heat shock protein 101, two novel proteins named PTEX150 and PTEX88, thioredoxin 2 and exported protein 2 (158).

The characterization of newly-identified proteins will likely reveal novel roles for those located in the PVM, within the RBC cytoplasm or expressed on the surface of infected RBCs, including the much sought-after ion/nutrient channels (159-161). Understanding the pathways and molecular mechanisms by which *Plasmodium* parasites export proteins to the RBC membrane is critical to the development of novel therapeutic approaches.

1.11. Monoclonal Antibodies Raised Against *P. vivax* Schizonts That Localize Specific Proteins to Parasitized RBC Ultrastructures

About twenty years ago, Barnwell *et al.* raised monoclonal antibodies (mAbs) against *P. vivax* schizonts and began to investigate parasite-encoded proteins that localized to the infected RBC membrane (124, 134). Mice were immunized with 2 x 10^8 *P. vivax* schizonts intraperitoneally at 4-week intervals. Spleen cells from the immunized mice were then fused with FOX-NY myeloma cells by standard techniques. Hybridoma supernatants were screened by indirect immunofluorescence assays on unfixed smears of *P. vivax* infected blood from *Saimiri boliviensis* monkeys and those secreting the mAbs were inoculated intraperitoneally into pristane-primed CD.1F2 mice. The resulting ascites fluid was purified on Protein A MAPS columns after absorption with uninfected squirrel monkey erythrocytes. The mAbs from this earlier research that have been utilized in this dissertation are listed below:

Table 1.1: *P. vivax* monoclonal antibody characteristics (124, 134)

Monoclonal antibody	Molecular weight of recognized *P. vivax* antigen	Localization by immunoEM
1H4.B6	95 kDa	CVCs
2H12.B4	95 kDa	CVC vesicles and cytoplasmic clefts
2H8.E10	95 kDa	CVCs
4C12.B4	95 kDa	CVCs

Immunofluorescence assays carried out with mAbs 2H8.E10, 4C12.B4 and 2H12.B4 on *P. vivax* trophozoite-infected erythrocytes produced a pattern of discrete, irregular dots distributed throughout the RBC membranes (162). Fewer dots were observed on ring-infected RBCs, with the number increasing as the parasite matured to middle and late stage trophozoites. The observed staining pattern was similar to the number, size and distribution of Schuffner's dots seen in Giemsa-stained *P. vivax*-infected RBCs by light microscopy (134). All four monoclonal antibodies are of the IgG1 subclass, cross react with *P. cynomolgi*, *P. ovale* and *P. knowlesi*-infected RBCs and recognize a protein that localizes to membrane-bound CVCs and cytoplasmic vesicles by immunoEM (124, 134). The mAbs immunoprecipitated a 95 kDa antigen from [^{35}S] methionine-labeled detergent extracts of *P. vivax*-infected RBCs. This protein may be associated with cytoskeletal matrix proteins since it is solubilized with sodium dodecyl sulfate (SDS) but not with nonionic detergents such as Triton X-100 (134).The identity and function(s) of the 95 kDa CVC protein, as well as the function(s) of the CVC, have not been previously determined (134).

1.12. *Plasmodium* Helical Interspersed Subtelomeric (PHIST) Family of Proteins

The so-called PHIST family was identified recently in post genomic bioinformatics investigations (153). The PHIST genes were shown to be located in the subtelomeric regions of all *Plasmodium* chromosomes with 71 paralogs in *P. falciparum*, 39 in *P. vivax* and 27 in *P. knowlesi* (153). The defining feature of members of this protein family is the conserved and unique PHIST domain of approximately 150 amino acids, consisting of four consecutive alpha helices and conserved tryptophan residues

(153). PHIST domains cluster into three subgroups based on the presence and position of the tryptophan residues – PHISTa, PHISTb and PHISTc. PHISTa is a *P. falciparum*-specific subfamily, while PHISTc is the most diverse subgroup (153). To date, the structure, function and localization of PHIST proteins as well as *phist* gene expression characteristics remain unknown.

1.13. Overview of *Plasmodium*-infected Red Blood Cell Studies

The purpose of this dissertation is to investigate the structural and antigenic modifications produced by *P. vivax* in their host erythrocyte membranes utilizing the simian models *P. cynomolgi* and *P. knowlesi*. This research involves the identification of specific genes and the respective proteins that are associated with infected RBC membrane structures, using a targeted monoclonal antibody proteomic approach as well as global RBC membrane proteomic methodologies. This research has identified proteins that are either unique to one species or conserved across species, analyzed the caveola-vesicle complexes in greater detail, and provided information on proteins that may serve as novel targets for therapeutic intervention.

REFERENCES

1. Organization, W. H. 2008. World Malaria Report 2008. WHO.
2. Greenwood, B. M., D. A. Fidock, D. E. Kyle, S. H. Kappe, P. L. Alonso, F. H. Collins, and P. E. Duffy. 2008. Malaria: progress, perils, and prospects for eradication. *J Clin Invest 118:1266*.
3. Bousema, J. T., L. C. Gouagna, C. J. Drakeley, A. M. Meutstege, B. A. Okech, I. N. Akim, J. C. Beier, J. I. Githure, and R. W. Sauerwein. 2004. Plasmodium falciparum gametocyte carriage in asymptomatic children in western Kenya. *Malar J 3:18*.
4. Jongwutiwes, S., C. Putaporntip, T. Iwasaki, T. Sata, and H. Kanbara. 2004. Naturally acquired Plasmodium knowlesi malaria in human, Thailand. *Emerg Infect Dis 10:2211*.
5. Singh, B., L. Kim Sung, A. Matusop, A. Radhakrishnan, S. S. Shamsul, J. Cox-Singh, A. Thomas, and D. J. Conway. 2004. A large focus of naturally acquired Plasmodium knowlesi infections in human beings. *Lancet 363:1017*.
6. Mendis, K., B. J. Sina, P. Marchesini, and R. Carter. 2001. The neglected burden of Plasmodium vivax malaria. *Am J Trop Med Hyg 64:97*.
7. Baird, J. K. 2008. Real-world therapies and the problem of vivax malaria. *N Engl J Med 359:2601*.
8. Galinski, M. R., and J. W. Barnwell. 2008. Plasmodium vivax: who cares? *Malar J 7 Suppl 1:S9*.
9. Price, R. N., N. M. Douglas, and N. M. Anstey. 2009. New developments in Plasmodium vivax malaria: severe disease and the rise of chloroquine resistance. *Curr Opin Infect Dis*.
10. Mueller, I., M. R. Galinski, J. K. Baird, J. M. Carlton, D. K. Kochar, P. L. Alonso, and H. A. del Portillo. 2009. Key gaps in the knowledge of Plasmodium vivax, a neglected human malaria parasite. *Lancet Infect Dis 9:555*.
11. Sherman, I. W. 1998. *Malaria: Parasite biology, pathogenesis and protection*. ASM Press, Washington DC.
12. Gregson, A., and C. V. Plowe. 2005. Mechanisms of resistance of malaria parasites to antifolates. *Pharmacol Rev 57:117*.
13. Wellems, T. E., and C. V. Plowe. 2001. Chloroquine-resistant malaria. *J Infect Dis 184:770*.
14. ter Kuile, F., N. J. White, P. Holloway, G. Pasvol, and S. Krishna. 1993. Plasmodium falciparum: in vitro studies of the pharmacodynamic properties of drugs used for the treatment of severe malaria. *Exp Parasitol 76:85*.
15. Udomsangpetch, R., B. Pipitaporn, S. Krishna, B. Angus, S. Pukrittayakamee, I. Bates, Y. Suputtamongkol, D. E. Kyle, and N. J. White. 1996. Antimalarial drugs reduce cytoadherence and rosetting Plasmodium falciparum. *J Infect Dis 173:691*.
16. Roberts, L., and M. Enserink. 2007. Malaria. Did they really say. eradication? *Science 318:1544*.
17. Health, O. W. 2009. World Malaria Report 2009. WHO.
18. Bhattarai, A., A. S. Ali, S. P. Kachur, A. Martensson, A. K. Abbas, R. Khatib, A. W. Al-Mafazy, M. Ramsan, G. Rotllant, J. F. Gerstenmaier, F. Molteni, S. Abdulla, S. M. Montgomery, A. Kaneko, and A. Bjorkman. 2007. Impact of

artemisinin-based combination therapy and insecticide-treated nets on malaria burden in Zanzibar. *PLoS Med 4:e309*.

19. Nyarango, P. M., T. Gebremeskel, G. Mebrahtu, J. Mufunda, U. Abdulmumini, A. Ogbamariam, A. Kosia, A. Gebremichael, D. Gunawardena, Y. Ghebrat, and Y. Okbaldet. 2006. A steep decline of malaria morbidity and mortality trends in Eritrea between 2000 and 2004: the effect of combination of control methods. *Malar J 5:33*.

20. Barnes, K. I., D. N. Durrheim, F. Little, A. Jackson, U. Mehta, E. Allen, S. S. Dlamini, J. Tsoka, B. Bredenkamp, D. J. Mthembu, N. J. White, and B. L. Sharp. 2005. Effect of artemether-lumefantrine policy and improved vector control on malaria burden in KwaZulu-Natal, South Africa. *PLoS Med 2:e330*.

21. Dondorp, A. M., F. Nosten, P. Yi, D. Das, A. P. Phyo, J. Tarning, K. M. Lwin, F. Ariey, W. Hanpithakpong, S. J. Lee, P. Ringwald, K. Silamut, M. Imwong, K. Chotivanich, P. Lim, T. Herdman, S. S. An, S. Yeung, P. Singhasivanon, N. P. Day, N. Lindegardh, D. Socheat, and N. J. White. 2009. Artemisinin resistance in Plasmodium falciparum malaria. *N Engl J Med 361:455*.

22. Rogers, W. O., R. Sem, T. Tero, P. Chim, P. Lim, S. Muth, D. Socheat, F. Ariey, and C. Wongsrichanalai. 2009. Failure of artesunate-mefloquine combination therapy for uncomplicated Plasmodium falciparum malaria in southern Cambodia. *Malar J 8:10*.

23. Martinez-Torres, D., F. Chandre, M. S. Williamson, F. Darriet, J. B. Berge, A. L. Devonshire, P. Guillet, N. Pasteur, and D. Pauron. 1998. Molecular characterization of pyrethroid knockdown resistance (kdr) in the major malaria vector Anopheles gambiae s.s. *Insect Mol Biol 7:179*.

24. Ballou, W. R., M. Arevalo-Herrera, D. Carucci, T. L. Richie, G. Corradin, C. Diggs, P. Druilhe, B. K. Giersing, A. Saul, D. G. Heppner, K. E. Kester, D. E. Lanar, J. Lyon, A. V. Hill, W. Pan, and J. D. Cohen. 2004. Update on the clinical development of candidate malaria vaccines. *Am J Trop Med Hyg 71:239*.

25. Moorthy, V. S., M. F. Good, and A. V. Hill. 2004. Malaria vaccine developments. *Lancet 363:150*.

26. Malkin, E., F. Dubovsky, and M. Moree. 2006. Progress towards the development of malaria vaccines. *Trends Parasitol 22:292*.

27. Stoute, J. A., M. Slaoui, D. G. Heppner, P. Momin, K. E. Kester, P. Desmons, B. T. Wellde, N. Garcon, U. Krzych, and M. Marchand. 1997. A preliminary evaluation of a recombinant circumsporozoite protein vaccine against Plasmodium falciparum malaria. RTS,S Malaria Vaccine Evaluation Group. *N Engl J Med 336:86*.

28. Bojang, K. A., P. J. Milligan, M. Pinder, L. Vigneron, A. Alloueche, K. E. Kester, W. R. Ballou, D. J. Conway, W. H. Reece, P. Gothard, L. Yamuah, M. Delchambre, G. Voss, B. M. Greenwood, A. Hill, K. P. McAdam, N. Tornieporth, J. D. Cohen, and T. Doherty. 2001. Efficacy of RTS,S/AS02 malaria vaccine against Plasmodium falciparum infection in semi-immune adult men in The Gambia: a randomised trial. *Lancet 358:1927*.

29. Alonso, P. L., J. Sacarlal, J. J. Aponte, A. Leach, E. Macete, P. Aide, B. Sigauque, J. Milman, I. Mandomando, Q. Bassat, C. Guinovart, M. Espasa, S. Corachan, M. Lievens, M. M. Navia, M. C. Dubois, C. Menendez, F. Dubovsky, J. Cohen, R.

Thompson, and W. R. Ballou. 2005. Duration of protection with RTS,S/AS02A malaria vaccine in prevention of Plasmodium falciparum disease in Mozambican children: single-blind extended follow-up of a randomised controlled trial. *Lancet 366:2012*.

30. Aponte, J. J., P. Aide, M. Renom, I. Mandomando, Q. Bassat, J. Sacarlal, M. N. Manaca, S. Lafuente, A. Barbosa, A. Leach, M. Lievens, J. Vekemans, B. Sigauque, M. C. Dubois, M. A. Demoitie, M. Sillman, B. Savarese, J. G. McNeil, E. Macete, W. R. Ballou, J. Cohen, and P. L. Alonso. 2007. Safety of the RTS,S/AS02D candidate malaria vaccine in infants living in a highly endemic area of Mozambique: a double blind randomised controlled phase I/IIb trial. *Lancet 370:1543*.

31. Kester, K. E., D. A. McKinney, N. Tornieporth, C. F. Ockenhouse, D. G. Heppner, Jr., T. Hall, B. T. Wellde, K. White, P. Sun, R. Schwenk, U. Krzych, M. Delchambre, G. Voss, M. C. Dubois, R. A. Gasser, Jr., M. G. Dowler, M. O'Brien, J. Wittes, R. Wirtz, J. Cohen, and W. R. Ballou. 2007. A phase I/IIa safety, immunogenicity, and efficacy bridging randomized study of a two-dose regimen of liquid and lyophilized formulations of the candidate malaria vaccine RTS,S/AS02A in malaria-naive adults. *Vaccine 25:5359*.

32. Grabowsky, M. 2008. The billion-dollar malaria moment. *Nature 451:1051*.

33. Johnson, J. G., N. Epstein, T. Shiroishi, and L. H. Miller. 1980. Factors affecting the ability of isolated Plasmodium knowlesi merozoites to attach to and invade erythrocytes. *Parasitology 80:539*.

34. Miller, L. H., M. F. Good, and G. Milon. 1994. Malaria pathogenesis. *Science 264:1878*.

35. Schofield, L., and G. E. Grau. 2005. Immunological processes in malaria pathogenesis. *Nat Rev Immunol 5:722*.

36. Sturm, A., R. Amino, C. van de Sand, T. Regen, S. Retzlaff, A. Rennenberg, A. Krueger, J. M. Pollok, R. Menard, and V. T. Heussler. 2006. Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science 313:1287*.

37. Tarun, A. S., K. Baer, R. F. Dumpit, S. Gray, N. Lejarcegui, U. Frevert, and S. H. Kappe. 2006. Quantitative isolation and in vivo imaging of malaria parasite liver stages. *Int J Parasitol 36:1283*.

38. Krotoski, W. A. 1985. Discovery of the hypnozoite and a new theory of malarial relapse. *Trans R Soc Trop Med Hyg 79:1*.

39. Bannister, L., and G. Mitchell. 2003. The ins, outs and roundabouts of malaria. *Trends Parasitol 19:209*.

40. Galinski, M. R., A. R. Dluzewski, and J. W. Barnwell. 2005. Malaria Merozoites and Invasion of Erythrocytes. In *Molecular Approaches to Malaria*. I. W. Sherman, ed. ASM Press, New York, p. 113.

41. Roggwiller, E., M. E. Betoulle, T. Blisnick, and C. Braun Breton. 1996. A role for erythrocyte band 3 degradation by the parasite gp76 serine protease in the formation of the parasitophorous vacuole during invasion of erythrocytes by Plasmodium falciparum. *Mol Biochem Parasitol 82:13*.

42. Wilson, R. J. 1990. Biochemistry of red cell invasion. *Blood Cells 16:237*.

43.   Rangachari, K., A. Dluzewski, R. J. Wilson, and W. B. Gratzer. 1986. Control of malarial invasion by phosphorylation of the host cell membrane cytoskeleton. *Nature 324:364.*

44.   Pinder, J. C., R. E. Fowler, A. R. Dluzewski, L. H. Bannister, F. M. Lavin, G. H. Mitchell, R. J. Wilson, and W. B. Gratzer. 1998. Actomyosin motor in the merozoite of the malaria parasite, Plasmodium falciparum: implications for red cell invasion. *J Cell Sci 111 (Pt 13):1831.*

45.   Dluzewski, A. R., G. H. Mitchell, P. R. Fryer, S. Griffiths, R. J. Wilson, and W. B. Gratzer. 1992. Origins of the parasitophorous vacuole membrane of the malaria parasite, Plasmodium falciparum, in human red blood cells. *J Cell Sci 102 (Pt 3):527.*

46.   Dluzewski, A. R., D. Zicha, G. A. Dunn, and W. B. Gratzer. 1995. Origins of the parasitophorous vacuole membrane of the malaria parasite: surface area of the parasitized red cell. *Eur J Cell Biol 68:446.*

47.   Ball, E. G., K. R. Mc, and et al. 1948. Studies on malarial parasites; chemical and metabolic changes during growth and multiplication in vivo and in vitro. *J Biol Chem 175:547.*

48.   Morrison, D. B., and H. A. Jeskey. 1948. Alterations in some constituents of the monkey erythrocyte infected with Plasmodium knowlesi as related to pigment formation. *J Natl Malar Soc 7:259.*

49.   Galinski, M. R., C. C. Medina, P. Ingravallo, and J. W. Barnwell. 1992. A reticulocyte-binding protein complex of Plasmodium vivax merozoites. *Cell 69:1213.*

50.   CDC. 2006. Malaria.

51.   Baird, J. K. 2007. Neglect of Plasmodium vivax malaria. *Trends Parasitol 23:533.*

52.   Qari, S. H., Y. P. Shi, N. J. Pieniazek, W. E. Collins, and A. A. Lal. 1996. Phylogenetic relationship among the malaria parasites based on small subunit rRNA gene sequences: monophyletic nature of the human malaria parasite, Plasmodium falciparum. *Mol Phylogenet Evol 6:157.*

53.   Barnwell, J. W., and M. R. Galinski. 1991. The adhesion of malaria merozoite proteins to erythrocytes: a reflection of function? *Res Immunol 142:666.*

54.   Galinski, M. R., and V. Corredor. 2004. Variant antigen expression in malaria infections: posttranscriptional gene silencing, virulence and severe pathology. *Mol Biochem Parasitol 134:17.*

55.   Lee, K. S., J. Cox-Singh, and B. Singh. 2009. Morphological features and differential counts of Plasmodium knowlesi parasites in naturally acquired human infections. *Malar J 8:73.*

56.   van Hellemond, J. J., M. Rutten, R. Koelewijn, A. M. Zeeman, J. J. Verweij, P. J. Wismans, C. H. Kocken, and P. J. van Genderen. 2009. Human Plasmodium knowlesi infection detected by rapid diagnostic tests for malaria. *Emerg Infect Dis 15:1478.*

57.   Pain, A., U. Bohme, A. E. Berry, K. Mungall, R. D. Finn, A. P. Jackson, T. Mourier, J. Mistry, E. M. Pasini, M. A. Aslett, S. Balasubrammaniam, K. Borgwardt, K. Brooks, C. Carret, T. J. Carver, I. Cherevach, T. Chillingworth, T. G. Clark, M. R. Galinski, N. Hall, D. Harper, D. Harris, H. Hauser, A. Ivens, C. S.

Janssen, T. Keane, N. Larke, S. Lapp, M. Marti, S. Moule, I. M. Meyer, D. Ormond, N. Peters, M. Sanders, S. Sanders, T. J. Sargeant, M. Simmonds, F. Smith, R. Squares, S. Thurston, A. R. Tivey, D. Walker, B. White, E. Zuiderwijk, C. Churcher, M. A. Quail, A. F. Cowman, C. M. Turner, M. A. Rajandream, C. H. Kocken, A. W. Thomas, C. I. Newbold, B. G. Barrell, and M. Berriman. 2008. The genome of the simian and human malaria parasite Plasmodium knowlesi. *Nature 455:799*.

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

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

60. Gratzer, W. 1984. Cell biology. More red than dead. *Nature 310:368*.

61. Mohandas, N., and P. G. Gallagher. 2008. Red cell membrane: past, present, and future. *Blood 112:3939*.

62. Pasini, E. M., M. Kirkegaard, P. Mortensen, H. U. Lutz, A. W. Thomas, and M. Mann. 2006. In-depth analysis of the membrane and cytosolic proteome of red blood cells. *Blood 108:791*.

63. Walensky, L. D., N. Mohandas, and S. E. Lux. 2003. *Blood, Principles and Practice of Haematology*. Lippincott Williams & Wilkins, Philadelphia.

64. Mohandas, N., and E. Evans. 1994. Mechanical properties of the red cell membrane in relation to molecular structure and genetic defects. *Annual Review of Biophysics and Biomolecular Structure 23:787*.

65. Rybicki, A. C., R. Heath, B. Lubin, and R. S. Schwartz. 1988. Human erythrocyte protein 4.1 is a phosphatidylserine binding protein. *J Clin Invest 81:255*.

66. An, X., X. Guo, H. Sum, J. Morrow, W. Gratzer, and N. Mohandas. 2004. Phosphatidylserine binding sites in erythroid spectrin: location and implications for membrane stability. *Biochemistry 43:310*.

67. Verkleij, A. J., R. F. Zwaal, B. Roelofsen, P. Comfurius, D. Kastelijn, and L. L. van Deenen. 1973. The asymmetric distribution of phospholipids in the human red cell membrane. A combined study using phospholipases and freeze-etch electron microscopy. *Biochim Biophys Acta 323:178*.

68. Zwaal, R. F., and A. J. Schroit. 1997. Pathophysiologic implications of membrane phospholipid asymmetry in blood cells. *Blood 89:1121*.

69. Bennett, V., and S. Lambert. 1991. The spectrin skeleton: from red cells to brain. *J Clin Invest 87:1483*.

70. Bennett, V. 1985. The membrane skeleton of human erythrocytes and its implications for more complex cells. *Annu Rev Biochem 54:273*.

71. Bennett, V. 1989. The spectrin-actin junction of erythrocyte membrane skeletons. *Biochim Biophys Acta 988:107*.

72. Bennett, V., and A. J. Baines. 2001. Spectrin and ankyrin-based pathways: metazoan inventions for integrating cells into tissues. *Physiol Rev 81:1353*.

73. Mohandas, N., and X. An. 2006. New insights into function of red cell membrane proteins and their interaction with spectrin-based membrane skeleton. *Transfus Clin Biol 13:29*.

74. Sherman, I. W. 1979. Biochemistry of Plasmodium (malarial parasites). *Microbiol Rev 43:453*.

75. Allred, D. R., J. E. Gruenberg, and I. W. Sherman. 1986. Dynamic rearrangements of erythrocyte membrane internal architecture induced by infection with Plasmodium falciparum. *J Cell Sci 81:1*.

76. Ginsburg, H., S. Kutner, M. Krugliak, and Z. I. Cabantchik. 1985. Characterization of permeation pathways appearing in the host membrane of Plasmodium falciparum infected red blood cells. *Mol Biochem Parasitol 14:313*.

77. Howard, R. J. 1988. Malarial proteins at the membrane of Plasmodium falciparum-infected erythrocytes and their involvement in cytoadherence to endothelial cells. *Prog Allergy 41:98*.

78. Sherman, I. W. 1985. Membrane structure and function of malaria parasites and the infected erythrocyte. *Parasitology 91 (Pt 3):609*.

79. Sherman, I. W., and J. R. Greenan. 1986. Plasmodium falciparum: regional differences in lectin and cationized ferritin binding to the surface of the malaria-infected human erythrocyte. *Parasitology 93 (Pt 1):17*.

80. Barnwell, J. W. 1990. Vesicle-mediated transport of membrane and proteins in malaria-infected erythrocytes. *Blood Cells 16:379*.

81. Lustigman, S., R. F. Anders, G. V. Brown, and R. L. Coppel. 1990. The mature-parasite-infected erythrocyte surface antigen (MESA) of Plasmodium falciparum associates with the erythrocyte membrane skeletal protein, band 4.1. *Mol Biochem Parasitol 38:261*.

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

83. Oh, S. S., S. Voigt, D. Fisher, S. J. Yi, P. J. LeRoy, L. H. Derick, S. Liu, and A. H. Chishti. 2000. Plasmodium falciparum erythrocyte membrane protein 1 is anchored to the actin-spectrin junction and knob-associated histidine-rich protein in the erythrocyte skeleton. *Mol Biochem Parasitol 108:237*.

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

85. Cooke, B. M., K. Lingelbach, L. H. Bannister, and L. Tilley. 2004. Protein trafficking in Plasmodium falciparum-infected red blood cells. *Trends Parasitol 20:581*.

86. Ansorge, I., J. Benting, S. Bhakdi, and K. Lingelbach. 1996. Protein sorting in Plasmodium falciparum-infected red blood cells permeabilized with the pore-forming protein streptolysin O. *Biochem J 315 (Pt 1):307*.

87. Waller, R. F., M. B. Reed, A. F. Cowman, and G. I. McFadden. 2000. Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. *Embo J 19:1794*.

88. Adisa, A., M. Rug, N. Klonis, M. Foley, A. F. Cowman, and L. Tilley. 2003. The signal sequence of exported protein-1 directs the green fluorescent protein to the parasitophorous vacuole of transfected malaria parasites. *J Biol Chem 278:6532*.

89. Wickham, M. E., M. Rug, S. A. Ralph, N. Klonis, G. I. McFadden, L. Tilley, and A. F. Cowman. 2001. Trafficking and assembly of the cytoadherence complex in Plasmodium falciparum-infected human erythrocytes. *Embo J 20:5636*.

90. Taraschi, T. F., D. Trelka, S. Martinez, T. Schneider, and M. E. O'Donnell. 2001. Vesicle-mediated trafficking of parasite proteins to the host cell cytosol and erythrocyte surface membrane in Plasmodium falciparum infected erythrocytes. *Int J Parasitol 31:1381*.

91. Spielmann, T., D. J. Fergusen, and H. P. Beck. 2003. etramps, a new Plasmodium falciparum gene family coding for developmentally regulated and highly charged membrane proteins located at the parasite-host cell interface. *Mol Biol Cell 14:1529*.

92. Haldar, K., A. F. de Amorim, and G. A. Cross. 1989. Transport of fluorescent phospholipid analogues from the erythrocyte membrane to the parasite in Plasmodium falciparum-infected cells. *J Cell Biol 108:2183*.

93. Adisa, A., M. Rug, M. Foley, and L. Tilley. 2002. Characterisation of a delta-COP homologue in the malaria parasite, Plasmodium falciparum. *Mol Biochem Parasitol 123:11*.

94. Lanzer, M., H. Wickert, G. Krohne, L. Vincensini, and C. Braun Breton. 2006. Maurer's clefts: a novel multi-functional organelle in the cytoplasm of Plasmodium falciparum-infected erythrocytes. *Int J Parasitol 36:23*.

95. Wickert, H., and G. Krohne. 2007. The complex morphology of Maurer's clefts: from discovery to three-dimensional reconstructions. *Trends Parasitol 23:502*.

96. Spycher, C., M. Rug, N. Klonis, D. J. Ferguson, A. F. Cowman, H. P. Beck, and L. Tilley. 2006. Genesis of and trafficking to the Maurer's clefts of Plasmodium falciparum-infected erythrocytes. *Mol Cell Biol 26:4074*.

97. Atkinson, C. T., and M. Aikawa. 1990. Ultrastructure of malaria-infected erythrocytes. *Blood Cells 16:351*.

98. Wickert, H., W. Gottler, G. Krohne, and M. Lanzer. 2004. Maurer's cleft organization in the cytoplasm of plasmodium falciparum-infected erythrocytes: new insights from three-dimensional reconstruction of serial ultrathin sections. *Eur J Cell Biol 83:567*.

99. Hanssen, E., P. Carlton, S. Deed, N. Klonis, J. Sedat, J. Derisi, and L. Tilley. 2009. Whole cell imaging reveals novel modular features of the exomembrane system of the malaria parasite, Plasmodium falciparum. *Int J Parasitol*.

100. Bannister, L. H., J. M. Hopkins, G. Margos, A. R. Dluzewski, and G. H. Mitchell. 2004. Three-dimensional ultrastructure of the ring stage of Plasmodium falciparum: evidence for export pathways. *Microsc Microanal 10:551*.
101. Slomianny, C., and G. Prensier. 1990. A cytochemical ultrastructural study of the lysosomal system of different species of malaria parasites. *J Protozool 37:465*.
102. Van Wye, J., N. Ghori, P. Webster, R. R. Mitschler, H. G. Elmendorf, and K. Haldar. 1996. Identification and localization of rab6, separation of rab6 from ERD2 and implications for an 'unstacked' Golgi, in Plasmodium falciparum. *Mol Biochem Parasitol 83:107*.
103. Behari, R., and K. Haldar. 1994. Plasmodium falciparum: protein localization along a novel, lipid-rich tubovesicular membrane network in infected erythrocytes. *Exp Parasitol 79:250*.
104. Haldar, K., B. U. Samuel, N. Mohandas, T. Harrison, and N. L. Hiller. 2001. Transport mechanisms in Plasmodium-infected erythrocytes: lipid rafts and a tubovesicular network. *Int J Parasitol 31:1393*.
105. Przyborski, J. M., H. Wickert, G. Krohne, and M. Lanzer. 2003. Maurer's clefts--a novel secretory organelle? *Mol Biochem Parasitol 132:17*.
106. Wickert, H., F. Wissing, K. T. Andrews, A. Stich, G. Krohne, and M. Lanzer. 2003. Evidence for trafficking of PfEMP1 to the surface of P. falciparum-infected erythrocytes via a complex membrane network. *Eur J Cell Biol 82:271*.
107. Knuepfer, E., M. Rug, N. Klonis, L. Tilley, and A. F. Cowman. 2005. Trafficking determinants for PfEMP3 export and assembly under the Plasmodium falciparum-infected red blood cell membrane. *Mol Microbiol 58:1039*.
108. Hanssen, E., R. Sougrat, S. Frankland, S. Deed, N. Klonis, J. Lippincott-Schwartz, and L. Tilley. 2008. Electron tomography of the Maurer's cleft organelles of Plasmodium falciparum-infected erythrocytes reveals novel structural features. *Mol Microbiol 67:703*.
109. Kriek, N., L. Tilley, P. Horrocks, R. Pinches, B. C. Elford, D. J. Ferguson, K. Lingelbach, and C. I. Newbold. 2003. Characterization of the pathway for transport of the cytoadherence-mediating protein, PfEMP1, to the host cell surface in malaria parasite-infected erythrocytes. *Mol Microbiol 50:1215*.
110. Blisnick, T., M. E. Morales Betoulle, J. C. Barale, P. Uzureau, L. Berry, S. Desroses, H. Fujioka, D. Mattei, and C. Braun Breton. 2000. Pfsbp1, a Maurer's cleft Plasmodium falciparum protein, is associated with the erythrocyte skeleton. *Mol Biochem Parasitol 111:107*.
111. Hawthorne, P. L., K. R. Trenholme, T. S. Skinner-Adams, T. Spielmann, K. Fischer, M. W. Dixon, M. R. Ortega, K. L. Anderson, D. J. Kemp, and D. L. Gardiner. 2004. A novel Plasmodium falciparum ring stage protein, REX, is located in Maurer's clefts. *Mol Biochem Parasitol 136:181*.
112. Spycher, C., N. Klonis, T. Spielmann, E. Kump, S. Steiger, L. Tilley, and H. P. Beck. 2003. MAHRP-1, a novel Plasmodium falciparum histidine-rich protein, binds ferriprotoporphyrin IX and localizes to the Maurer's clefts. *J Biol Chem 278:35373*.
113. Cooke, B. M., N. Mohandas, and R. L. Coppel. 2004. Malaria and the red blood cell membrane. *Semin Hematol 41:173*.

114. Kaviratne, M., S. M. Khan, W. Jarra, and P. R. Preiser. 2002. Small variant STEVOR antigen is uniquely located within Maurer's clefts in Plasmodium falciparum-infected red blood cells. *Eukaryot Cell 1:926*.

115. Sam-Yellowe, T. Y., L. Florens, J. R. Johnson, T. Wang, J. A. Drazba, K. G. Le Roch, Y. Zhou, S. Batalov, D. J. Carucci, E. A. Winzeler, and J. R. Yates, 3rd. 2004. A Plasmodium gene family encoding Maurer's cleft membrane proteins: structural properties and expression profiling. *Genome Res 14:1052*.

116. Vincensini, L., S. Richert, T. Blisnick, A. Van Dorsselaer, E. Leize-Wagner, T. Rabilloud, and C. Braun Breton. 2005. Proteomic analysis identifies novel proteins of the Maurer's clefts, a secretory compartment delivering Plasmodium falciparum proteins to the surface of its host cell. *Mol Cell Proteomics 4:582*.

117. Johnson, D., K. Gunther, I. Ansorge, J. Benting, A. Kent, L. Bannister, R. Ridley, and K. Lingelbach. 1994. Characterization of membrane proteins exported from Plasmodium falciparum into the host erythrocyte. *Parasitology 109 (Pt 1):1*.

118. Crabb, B. S., B. M. Cooke, J. C. Reeder, R. F. Waller, S. R. Caruana, K. M. Davern, M. E. Wickham, G. V. Brown, R. L. Coppel, and A. F. Cowman. 1997. Targeted gene disruption shows that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress. *Cell 89:287*.

119. Waterkeyn, J. G., M. E. Wickham, K. M. Davern, B. M. Cooke, R. L. Coppel, J. C. Reeder, J. G. Culvenor, R. F. Waller, and A. F. Cowman. 2000. Targeted mutagenesis of Plasmodium falciparum erythrocyte membrane protein 3 (PfEMP3) disrupts cytoadherence of malaria-infected red blood cells. *Embo J 19:2813*.

120. Foley, M., L. Corcoran, L. Tilley, and R. Anders. 1994. Plasmodium falciparum: mapping the membrane-binding domain in the ring-infected erythrocyte surface antigen. *Exp Parasitol 79:340*.

121. Kilejian, A., M. A. Rashid, M. Aikawa, T. Aji, and Y. F. Yang. 1991. Selective association of a fragment of the knob protein with spectrin, actin and the red cell membrane. *Mol Biochem Parasitol 44:175*.

122. Chishti, A. H., G. J. Maalouf, S. Marfatia, J. Palek, W. Wang, D. Fisher, and S. C. Liu. 1994. Phosphorylation of protein 4.1 in Plasmodium falciparum-infected human red blood cells. *Blood 83:3339*.

123. Garnham, P. C. C. 1966. *Malaria parasites and other Haemosporidia*. Blackwell Scientific Publications, Oxford.

124. Matsumoto, Y., M. Aikawa, and J. W. Barnwell. 1988. Immunoelectron microscopic localization of vivax malaria antigens to the clefts and caveola-vesicle complexes of infected erythrocytes. *Am J Trop Med Hyg 39:317*.

125. Tobie, J. E., and G. R. Coatney. 1961. Fluorescent antibody staining of human malaria parasites. *Exp Parasitol 11:128*.

126. Aikawa, M., Y. Uni, A. T. Andrutis, and R. J. Howard. 1986. Membrane-associated electron-dense material of the asexual stages of Plasmodium falciparum: evidence for movement from the intracellular parasite to the erythrocyte membrane. *Am J Trop Med Hyg 35:30*.

127. Sterling, C. R., T. M. Seed, M. Aikawa, and J. Rabbege. 1975. Erythrocyte membrane alterations induced by Plasmodium simium infection in Saimiri sciureus: relation to Schuffner's dots. *J Parasitol 61:177*.

128. Aikawa, M., L. H. Miller, and J. Rabbege. 1975. Caveola--vesicle complexes in the plasmalemma of erythrocytes infected by Plasmodium vivax and P cynomolgi. Unique structures related to Schuffner's dots. *Am J Pathol 79:285*.
129. Parton, R. G. 1996. Caveolae and caveolins. *Curr Opin Cell Biol 8:542*.
130. Pelkmans, L., and A. Helenius. 2002. Endocytosis via caveolae. *Traffic 3:311*.
131. Bender, F., M. Montoya, V. Monardes, L. Leyton, and A. F. Quest. 2002. Caveolae and caveolae-like membrane domains in cellular signaling and disease: identification of downstream targets for the tumor suppressor protein caveolin-1. *Biol Res 35:151*.
132. Brazer, S. C., B. B. Singh, X. Liu, W. Swaim, and I. S. Ambudkar. 2003. Caveolin-1 contributes to assembly of store-operated Ca2+ influx channels by regulating plasma membrane localization of TRPC1. *J Biol Chem 278:27208*.
133. Sharma, D. K., A. Choudhury, R. D. Singh, C. L. Wheatley, D. L. Marks, and R. E. Pagano. 2003. Glycosphingolipids internalized via caveolar-related endocytosis rapidly merge with the clathrin pathway in early endosomes and form microdomains for recycling. *J Biol Chem 278:7564*.
134. Barnwell, J. W., P. Ingravallo, M. R. Galinski, Y. Matsumoto, and M. Aikawa. 1990. Plasmodium vivax: malarial proteins associated with the membrane-bound caveola-vesicle complexes and cytoplasmic cleft structures of infected erythrocytes. *Exp Parasitol 70:85*.
135. el Nahal, H. M. 1969. Immunofluorescence studies on the erythrocytic and sporogonic stages of the malaria parasite: stippling in infected erythrocytes. *Bull World Health Organ 40:158*.
136. Voller, A., and R. S. Bray. 1962. Fluorescent antibody staining as a measure of malarial antibody. *Proc Soc Exp Biol Med 110:907*.
137. Udagama, P. V., C. T. Atkinson, J. S. Peiris, P. H. David, K. N. Mendis, and M. Aikawa. 1988. Immunoelectron microscopy of Schuffner's dots in Plasmodium vivax-infected human erythrocytes. *Am J Pathol 131:48*.
138. Albano, F. R., M. Foley, and L. Tilley. 1999. Export of parasite proteins to the erythrocyte cytoplasm: secretory machinery and traffic signals. *Novartis Found Symp 226:157*.
139. Von Heijne, G. 1985. Signal sequences. The limits of variation. *Journal of Molecular Biology 184:99*.
140. Lingelbach, K. R. 1993. Plasmodium falciparum: a molecular view of protein transport from the parasite into the host erythrocyte. *Exp Parasitol 76:318*.
141. Klemba, M., W. Beatty, I. Gluzman, and D. E. Goldberg. 2004. Trafficking of plasmepsin II to the food vacuole of the malaria parasite Plasmodium falciparum. *J Cell Biol 164:47*.
142. Adisa, A., F. R. Albano, J. Reeder, M. Foley, and L. Tilley. 2001. Evidence for a role for a Plasmodium falciparum homologue of Sec31p in the export of proteins to the surface of malaria parasite-infected erythrocytes. *J Cell Sci 114:3377*.
143. Benting, J., D. Mattei, and K. Lingelbach. 1994. Brefeldin A inhibits transport of the glycophorin-binding protein from Plasmodium falciparum into the host erythrocyte. *Biochem J 300 (Pt 3):821*.

144. Baumgartner, F., S. Wiek, K. Paprotka, S. Zauner, and K. Lingelbach. 2001. A point mutation in an unusual Sec7 domain is linked to brefeldin A resistance in a Plasmodium falciparum line generated by drug selection. *Mol Microbiol 41:1151*.

145. Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak. 2004. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol 340:783*.

146. Logan, E. 2004. Lecture Notes in Artificial Intelligence. In *Knowledge-based Intelligent Information and Engineering Systems Proceedings*. N. M. e. al, ed. Springer-Verlag, Wellington, New Zealand, p. 1023.

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

148. Hiller, N. L., S. Bhattacharjee, C. van Ooij, K. Liolios, T. Harrison, C. Lopez-Estrano, and K. Haldar. 2004. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science 306:1934*.

149. Lopez-Estrano, C., S. Bhattacharjee, T. Harrison, and K. Haldar. 2003. Cooperative domains define a unique host cell-targeting signal in Plasmodium falciparum-infected erythrocytes. *Proc Natl Acad Sci U S A 100:12402*.

150. Burghaus, P. A., and K. Lingelbach. 2001. Luciferase, when fused to an N-terminal signal peptide, is secreted from transfected Plasmodium falciparum and transported to the cytosol of infected erythrocytes. *J Biol Chem 276:26838*.

151. Chang, H. H., A. M. Falick, P. M. Carlton, J. W. Sedat, J. L. DeRisi, and M. A. Marletta. 2008. N-terminal processing of proteins exported by malaria parasites. *Mol Biochem Parasitol 160:107*.

152. Knuepfer, E., M. Rug, N. Klonis, L. Tilley, and A. F. Cowman. 2005. Trafficking of the major virulence factor to the surface of transfected P. falciparum-infected erythrocytes. *Blood 105:4078*.

153. Sargeant, T. J., M. Marti, E. Caler, J. M. Carlton, K. Simpson, T. P. Speed, and A. F. Cowman. 2006. Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol 7:R12*.

154. Rabiner, L. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc IEEE 77:257*.

155. Przyborski, J. M., S. K. Miller, J. M. Pfahler, P. P. Henrich, P. Rohrbach, B. S. Crabb, and M. Lanzer. 2005. Trafficking of STEVOR to the Maurer's clefts in Plasmodium falciparum-infected erythrocytes. *Embo J 24:2306*.

156. Boddey, J. A., A. N. Hodder, S. Gunther, P. R. Gilson, H. Patsiouras, E. A. Kapp, J. A. Pearce, T. F. de Koning-Ward, R. J. Simpson, B. S. Crabb, and A. F. Cowman. An aspartyl protease directs malaria effector proteins to the host cell. *Nature 463:627*.

157. Russo, I., S. Babbitt, V. Muralidharan, T. Butler, A. Oksman, and D. E. Goldberg. Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte. *Nature 463:632*.

158. de Koning-Ward, T. F., P. R. Gilson, J. A. Boddey, M. Rug, B. J. Smith, A. T. Papenfuss, P. R. Sanders, R. J. Lundie, A. G. Maier, A. F. Cowman, and B. S. Crabb. 2009. A newly discovered protein export machine in malaria parasites. *Nature 459:945*.

159. Desai, S. A., S. M. Bezrukov, and J. Zimmerberg. 2000. A voltage-dependent channel involved in nutrient uptake by red blood cells infected with the malaria parasite. *Nature 406:1001*.

160. Kirk, K., H. A. Horner, B. C. Elford, J. C. Ellory, and C. I. Newbold. 1994. Transport of diverse substrates into malaria-infected erythrocytes via a pathway showing functional characteristics of a chloride channel. *J Biol Chem 269:3339*.

161. Elliott, J. L., K. J. Saliba, and K. Kirk. 2001. Transport of lactate and pyruvate in the intraerythrocytic malaria parasite, Plasmodium falciparum. *Biochem J 355:733*.

162. Barnwell, J. W. 1986. Antigens of Plasmodium vivax blood stage parasites identified by monoclonal antibodies. *Memorias do Instituto Oswaldo Cruz, Rio de Janeiro 81:59*.

CHAPTER TWO

PHYLOGENETIC AND STRUCTURAL INFORMATION ON GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE IN *PLASMODIUM* PROVIDES FUNCTIONAL INSIGHTS

Sheila Akinyi [a], Jenny Gaona [b], Esmeralda V-S. Meyer [a], John W. Barnwell [c], Mary R. Galinski [a,d], Vladimir Corredor [a,b]

[a] Emory Vaccine Center, Yerkes National Primate Research Center, Emory University, 954 Gatewood Road, Atlanta, GA 30329, USA

[b] Unidad de Parasitolgı´a, Departamento de Salud Pu´blica, Facultad de Medicina, Universidad Nacional de Colombia, Bogotá´, Colombia

[c] Division of Parasitic Diseases, National Center for Infectious Diseases, Centers for Disease Control, Atlanta, USA

[d] Department of Medicine, Division of Infectious Diseases, Emory University, 954 Gatewood Road, Atlanta, GA 30329, USA

2.1. Abstract

Plasmodium is dependent on glycolysis for ATP production. The glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (G3PDH) plays an important role in glycolysis and is, therefore, a potential target for antimalarial drug development. The *g3pdh* gene of nine *Plasmodium* species was sequenced from genomic DNA and the type and origin determined by phylogenetic analysis. Substitutions were analyzed over a wide phylogenetic spectrum in relation to the known three-dimensional structures of the *P. falciparum* and human proteins. Substitutions were found within the functional domains (Rossman NAD+-binding and catalytic domains). A number of replacements within the adenosyl-binding surfaces were found to be conserved within the Chromalveolates, others in the Apicomplexa, and still others within the genus *Plasmodium*, all of which were different from the human sequence. These sites may prove to be of functional importance and provide insights for drug-targeting studies, as have other regions examined in *Leishmania* and *Toxoplasma* G3PDH research.

2.2. Introduction

   Plasmodium parasites cause more than one million deaths and sicken 300–500

million people with malaria infections each year (1). The discovery of new drug targets is

a global priority to help combat the spread of drug-resistant parasites (2, 3).

Glyceraldehyde-3-phosphate dehydrogenase, G3PDH) is an essential enzyme in

Plasmodium and recent crystallographic studies suggest it can be a target for therapeutic

intervention (4, 5). When *Plasmodium* parasites invade red blood cells (RBCs), they must

provide their own molecular machinery to support growth and development, and in

essence remodel the infected host cell (6). Mature RBCs have been referred to as

''floating corpses'' due to their lack of nuclei, protein synthesis capabilities and

trafficking machinery (7). To achieve its needs, the intraerythrocytic *Plasmodium* is

heavily dependent on glycolysis for the production of ATP, as illustrated by the much

higher glycolytic activity of parasitized RBCs over uninfected ones (6, 8). This is of

particular importance given that the parasite seems to lack a complete tricarboxylic acid

cycle in its asexual stages due to the absence of a pyruvate dehydrogenase complex (9,

10).

   G3PDH is a homotetrameric enzyme with dihedral symmetry consisting of two

major functional domains: the Rossman N-terminal NAD+ cofactor-binding domain and

the C-terminal catalytic domain (11, 12). The crystal structure of the *P. falciparum*

G3PDH has been solved at 2.25 Å (4) and 2.6 Å (5) resolution revealing that NAD+

binds to one of each of the four subunits on the G3PDH tetramer and that with the

exception of a bulge, created by the so-called S-loop, that separates the NAD+-binding

cavities of adjacent subunits, the structure is very similar to the human G3PDH. The S-

loop is structurally different from the human sequence and in close proximity to the catalytic site, and hence it is considered as a potential drug target site. Sequence similarity between *P. falciparum* G3PDH and human G3PDH is about 63.5% (11).

There are greater than 5000 known species within the phylum Apicomplexa, which includes the clinically relevant parasites *Plasmodium*, *Toxoplasma*, *Cryptosporidium*, *Eimeria* and *Babesia* (13, 14). In addition to the glycolytic function, there is evidence in some species including humans and fungi that cytosolic G3PDH may have non-glycolytic functions associated with membrane fusion, microtubule bundling, nuclear RNA export, DNA repair, apoptosis, and cell adhesion (15-17). Evidence also exists showing that *P. falciparum* G3PDH may play a role in apical complex biogenesis (18) and this activity is inhibited by the hemoglobin degradation product, ferriprotoporphyrin IX (19). The N-terminus of *P. falciparum* G3PDH has been shown to interact with the GTPase Rab2 and mediate its recruitment to microsomal membranes in a HeLa cell experimental system (18). These data support the proposition that *Plasmodium* G3PDH may have acquired multiple functions during its course of evolution, and underscores the importance of G3PDH in the survival of malaria parasites.

In *P. falciparum* the enzyme is encoded by a single copy, two-exon, 1.3 kb gene that is transcribed in ring, trophozoite and schizont stage parasites. It is maximally expressed in early developing schizonts and the RNA levels are reduced in late stage schizonts (11, 20) and our unpublished data).

The conserved albeit evolutionarily distinct G3PDH enzymes present in humans, *Plasmodium* and other Apicomplexa have slightly differing structures that could be targeted for the development of specific inhibitors and may provide a means for

therapeutic intervention for malaria infections (21-23). This strategy has had success in drug-targeting studies involving *Trypanosoma* and *Leishmania* G3PDH where adenosine analogs have been developed and function as selective inhibitors of *Trypanosoma* G3PDH (see (24, 25). The adenosine analogs were designed to act as tight binding inhibitors that occupied the adenosyl binding pocket on *Toxoplasma brucei* and *Leishmania mexicana* G3PDH (26). Just as is the case with *Plasmodium*, glycolysis provides virtually all the energy for the bloodstream form of *T. brucei*, as well as being an important source of energy for *T. cruzi* and *L. mexicana* (27, 28). The availability of the crystal structure of *P. falciparum* G3PDH (4, 5) allows for a detailed comparison of G3PDH sequences and structures in *Plasmodium*, with other Apicomplexa, and humans.

In this study, the *g3pdh* gene from various *Plasmodium* species was sequenced and a comparative analysis performed. The aim was to identify amino acid substitutions in the *Plasmodium* G3PDH protein that may reveal residues of potential functional importance (including functions other than glycolysis), distinct from the human host and thus provide insights for future experimentation and the development of therapeutic interventions.

2.3. Materials and Methods

In addition to including several available published *Plasmodium g3pdh* sequences, the *g3pdh* gene was cloned and sequenced from a number of additional *Plasmodium* species and additional strains: the human species *P. falciparum* (FVO), *P. vivax* (Salvador I), and *P. malariae* (Uganda 1), the chimpanzee species *P. reichenowi* (CDC1), and the simian species *P. knowlesi* (H strain)*, P. coatneyi* (Type strain), *P. cynomolgi* (Berok)*, P. brasilianum* (Peru II)*, and P. fragile* (Nilgiri)*,* using purified genomic DNA that had been preserved from primate blood-stage infections. Standard PCR amplifications were performed using *g3pdh*-specific primers (Table 1), the Roche High Fidelity System (Roche, Indianapolis, IN) and KOD HotStart DNA Polymerase (Novagen, San Diego, CA) with the following conditions: 30 cycles of denaturation (94 °C, 1 min), annealing (48 °C, 1 min) and extension (68 °C, 2.5 min). Two independent rounds of PCR amplification and sequencing were performed. All PCR products were purified using the QIAquick PCR purification kit (Qiagen, Valencia, CA) and cloned into the pCR2.1-TOPO vector. Purified plasmid DNA from 10 positive clones was sequenced using BigDye Terminator v. 3.1 Cycle sequencing kit (Applied Biosystems, Foster City, CA). The sequences were assembled and aligned using the MacVector software package CLUSTALW and refined manually.

Table 2.1: List of oligonucleotides used in this study

Name	Sequence
VC25F	5′-ATT AAT GGA TTT GGT CGT ATC-3′
VC24R	5′-CAA GTA CAC GGT TTG AGT ATC-3'
PfGP44F	5′-GAC GTT TAG TAT TTA GAG C-3′
PfGP366F	5′-TGC TAT CTC TTG AAA TAC-3′
PfGP507F	5′-CAA GTA GAT GTT GTA TGT-3′
PfGP534R	5′-CAG TTG ATT CAC ATA CAA-3′
PfGP1230R	5′-TAG TTG TTA GTA ATG TGT AC-3′
PvGP450F	5′-ACT TGT GCT ACT TGC TCA-3′
PmGP566F	5′-CTA CTA GCG TAC GAC TC-3′
PbrGP484F	5′-CTA CTA GCG TAC GAC TCA-3′
PbrGP361F	5′-GAG TAT GTT CAT GCA GTT-3′
PcoGP454F	5′-AGC ACT TGT GTT ACT TGC-3′
PcyGP467R	5′-TTG CAA TTG GAG CTA AAC-3′

A *g3pdh* sequence database was built using new and published *Plasmodium* sequences as well as *g3pdh* sequences from other eukaryotes (Table 2). Protein sequences were aligned according to G3PDH type groups with 3Dcoffee (29), using available G3PDH crystal structures as a reference. Secondary structure predictions (obtained by DSSP) were drawn over the alignments using the Structural Alignment Package, STRAP (30). Position assignment to protein surfaces was achieved using the Computed Atlas of Surface Topography of proteins (31). The designation of ambiguous positions and alignment refinements were carried out by eye using Bioedit 7.1.2 (32) and the outputs from 3DCoffee and STRAP as a guide. A phylogenetic tree was reconstructed using maximum likelihood methods with the PHYML V2.4.3 program (33). The best substitution model and its parameter values were obtained using ProtTest (34). A phylogenetic tree was reconstructed under the WAG + I evolutionary model; gamma distribution was calculated using four rate categories, pinv 0.05, alfa 0.86 and homogeneous rates across the tree. Bootstrap values over 500 replicates are indicated on each branch in percentage values.

2.4. Results

To advance research in this direction, the type and origin of *g3pdh* in *Plasmodium* was determined through a phylogenetic analysis that included 12 *Plasmodium* species and sequences from 110 species along a wide phylogenetic spectrum. Then, the pattern of substitutions that differed between the human and *Plasmodium* sequences was analyzed within the framework of an alignment covering *g3pdh* sequences from diverse species within the "supergroup" Chromalveolates, including members of the Apicomplexa (Figs. 2.1 & 2.2).

The phylogenetic tree in Fig. 2.1 was built to determine the origin and identity of the *Plasmodium g3pdh* gene. Fig. 2.1 shows that both cytosol- and plastid-targeted G3PDH of Chromalveolates (Heterokonts and Alveolates) are of eukaryotic origin, as previously noted (35). Similar results with slightly different topologies are obtained with the Bayesian and NJ methods (data not shown). Positions homologous to D32, L187, P188 of the *S. stearothermophilus* G3PDH sequence (M24493), which are involved in the specific binding of NAD+ to the enzyme, allow for the classification of G3PDH sequences as plastid-targeted (NADP+) or cytosol-targeted enzymes (NAD+). The corresponding positions were checked in the G3PDH sequences of all taxa in this study and all *Plasmodium* G3PDH sequences analyzed were found to be compatible with a cytosolic classification. Fig. 2.1 also shows that the origin of *Plasmodium g3pdh* sequences is monophyletic.
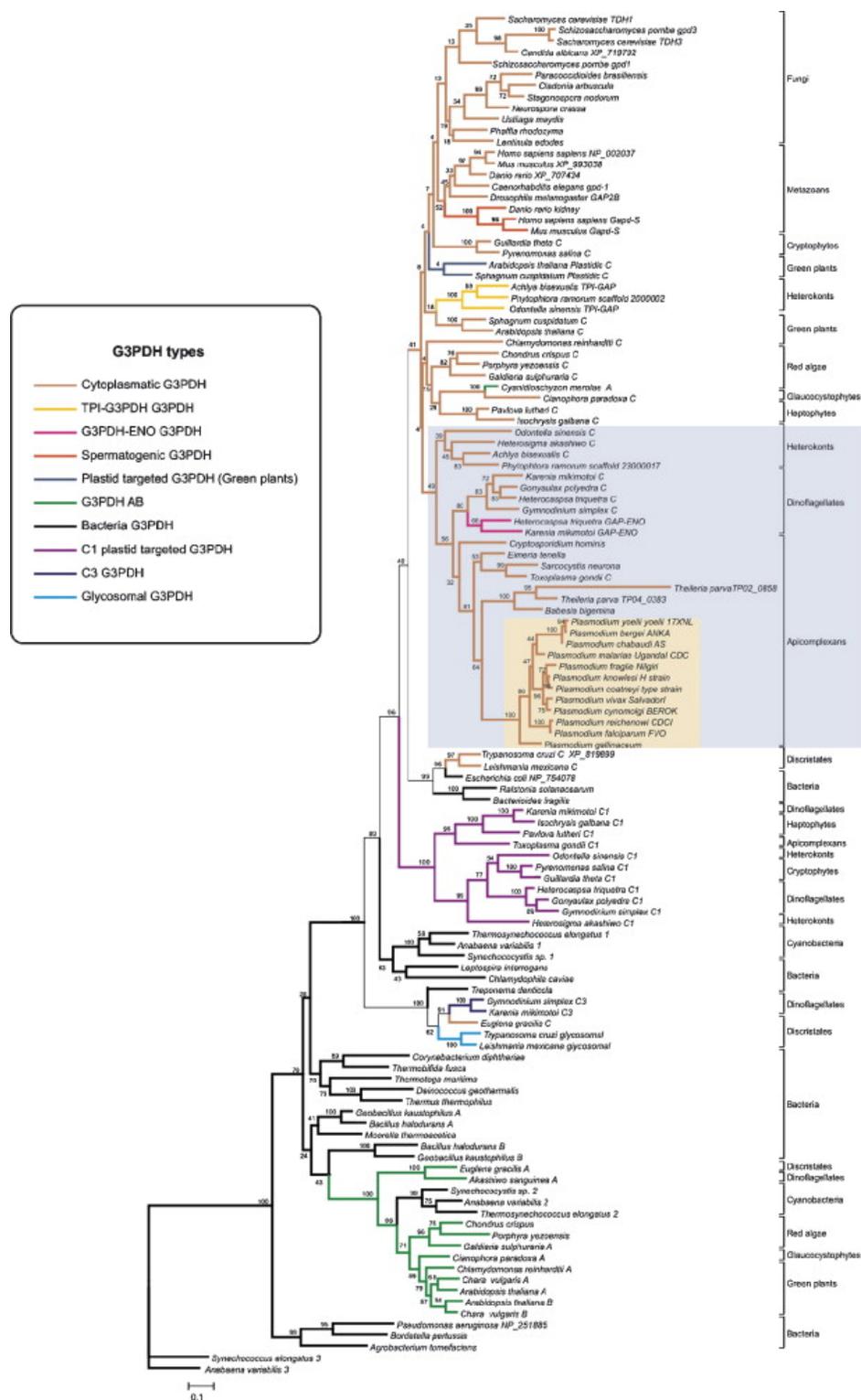
Figure 2.1. Maximum likelihood tree of G3PDH. Bootstrap values are indicated above the branches. Branch lengths are the number of substitutions per site according to the scale. G3PDH types are indicated by color branches, Chromalveolates are highlighted in grey and the *Plasmodium* genus is indicated in brown.
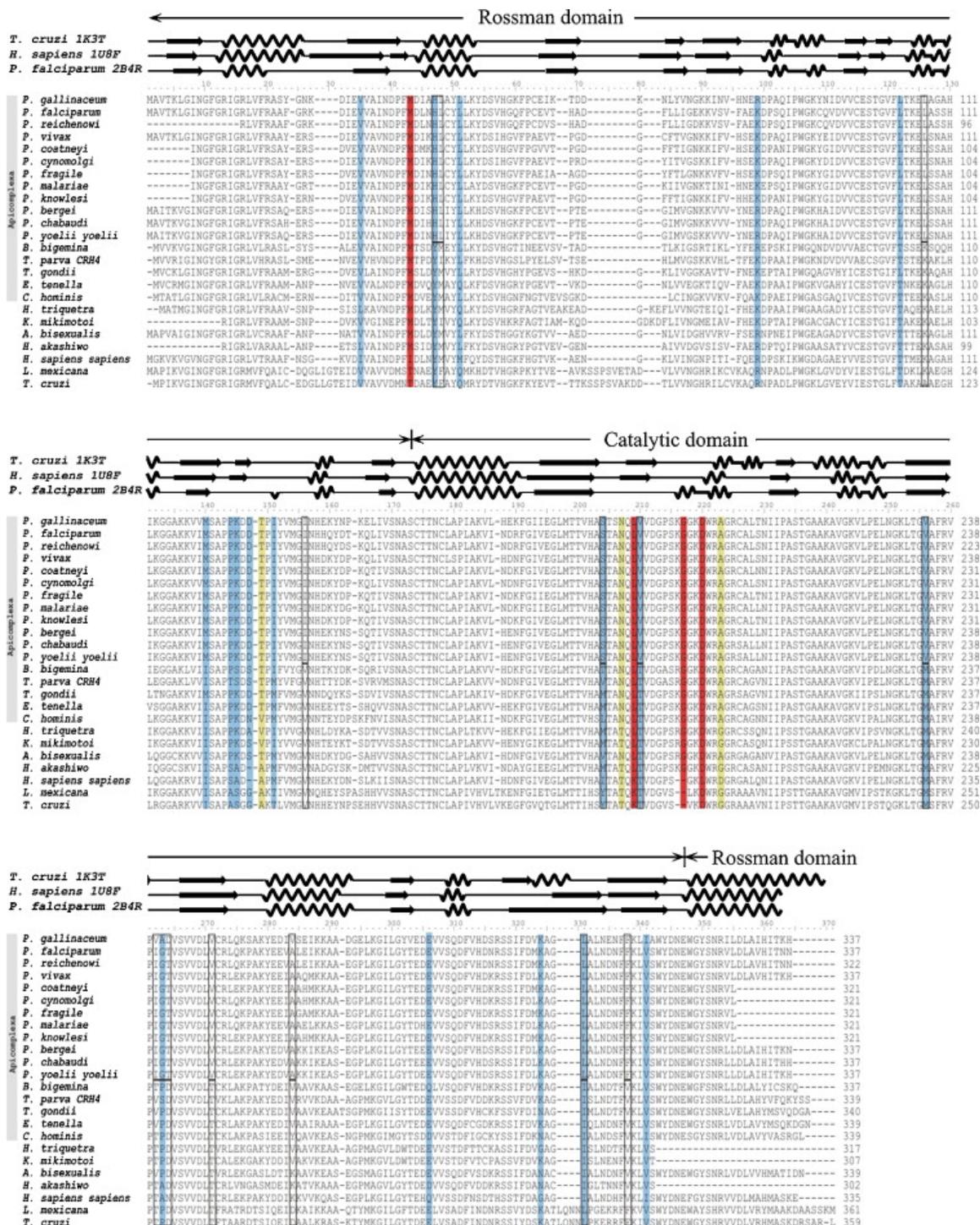
Figure 2.2. Amino acid sequence alignment of the functional domains of G3PDH—the Rossman NAD+-binding and catalytic domains. The alpha helix (〰〰) and beta sheet (→) secondary structures are indicated on top of the alignment. Substitutions that are conserved in Chromalveolates (red box), Apicomplexa (yellow) or *Plasmodium* (blue) but different from the human sequence and part of the NAD-binding pocket surface are shown. In addition, the figure shows replacements that are unique to *Plasmodium* species (boxed residues).

Table 2.2. List of species, identification numbers and database of origin for the sequences used in this study

Species/Strain	Sequence ID	Database	Species/Strain	Sequence ID	Database
P. falciparum 3D7	PF14_0598	PlasmoDB 3.5	*C. parvum*	AAEE01000002	CryptoDB 3.4
P. falciparum FVO	EU045404 (This study)	GenBank	*C. hominis*	AAEL01000124	CryptoDB 3.4
P. reichenowi CDC1	EU045408 (This study)	GenBank	*N. caninum*	TC3080 TC1626	USDA-WashU Neospora EST Project
P. knowlesi H strain	EU045406 (This study)	GenBank	*S. neurona*	TC2288 TC1094 TC1379 TC479	USDA-WashU gene index Project
P. coatneyi Type strain	EU045403 (This study)	GenBank	*T. gondii*	80.m00003 IX-4-3473658-3472567	ToxoDB 4.1
P. cynomolgi Berok	EU045409 (This study)	GenBank		59.m00091 VIII-4-5999894-5998908	
P. vivax Salvador I	EU045410 (This study)	GenBank	*H. triquetra*	AB195834	GenBank
P. fragile Nilgiri	EU045405 (This study)	GenBank		AB106701	
P. brasilianum Peru II	EU045402 (This study)	GenBank		AB106700	
P. malariae Uganda I	EU045407 (This study)	GenBank	*G. polyedra*	AF028560	GenBank
P. y. yoelii	PY03280	PlasmoDB 3.5		AF028562	
P. berghei	PB000084.03.0	PlasmoDB 3.5	*G. simplex*	AB106693	GenBank
P. chabaudi	PC000143.05.0	PlasmoDB 3.5		AB106694	
P. gallinaceum	Pgal1137d09.q1k	Sanger		AB106695	
	Pgal1137b01.q1k	*P. gallinaceum*	*A. sanguinea*	AB106697	GenBank
	Pgal0998b08.p1k	Genome		AB106696	
	Pgal1137d09.p1k	Project	*K. mikimotoi*	AB164183	GenBank
	Pgal0362b01.p1k			AB164184	
	Pgal0373g08.q1k			AB164185	
	Pgal1128h02.p1k			AB164186	
	Pgal0647a02.p1k		*H. akashiwo*	AF319448	GenBank
	Pgal0547d07.q1k			AF319449	
B. bigemina	Contig4141.0		*A. bisexualis*	AF063107	GenBank
B. bovis	chr*2*924*18			AF063106	
T. parva	AAGK01000004	GenBank	*G. theta*	U40032	GenBank
	AAGK01000002			U39873	

Species/Strain	Sequence ID	Database	Species/Strain	Sequence ID	Database
T. annulata	TA15530_chr2	Sanger	*A. thaliana*	NM_103456	GenBank
	TA08145_chr4	Theileria		NM_101161	
		annulata		NM_101214	
		Genome		NM_101496	
		project		NM_106601	
O. sinensis	AF063800	GenBank		NM_111283	
	AF063801			NM_113576	
	AF063802		*S. cuspidatum*	AJ246022	GenBank
P. ramorum	scaffold_23000017	DOE Joint		AJ246021	
	scaffold_2000002	Genome		AJ246031	
		Institute.		AJ246032	
		Phytophthora	*C. vulgaris*	AJ246015	GenBank
		ramorum V1.1		AJ246016	
P. lutheri	AY292376	GenBank		DQ270262	
	Cluster Id	TBestDB		AJ246014	
	PLL00000015		*C. reinhardtii*	L27669	GenBank
L. galbana	Cluster Id	TBestDB		L27668	
	ISL00000338		*C. crispus*	X73033	GenBank
	Cluster Id			X73036	
	ISL00001421		*P. yezoensis*	AY273819	GenBank
P. salina	U40033	GenBank		AY273820	
	U39897		*G. sulphuraria*	AJ012286	GenBank
C. albicans	XM_714816	GenBank		contig951	The *Galdieria sulphuraria* Genome Project
	XM_714699		*C. merolae*	c10f0002	*Cyanidioschyzon merolae* Genome Project
S. cerevisiae	S000003588	Saccharomyces		c13f0002	
	S000003769	genome		c10f0010	
	S000003424	database	*C. paradoxa*	AJ313316	GenBank
S. pombe	NM_001021142	GenBank		DQ270258	
	NM_001022073		*E. coli*	NP_753744	GenBank
P. brasiliensis	AF396657	GenBank		NP_754078	
C. arbuscula	AY170750	GenBank	*R. solanacearum*	NP_520870	GenBank
S. nodorum	AJ271155	GenBank	*B. fragilis*	YP_098251	GenBank
N. crassa	XM_951884	GenBank	*T. elongatus*	NP_680834	GenBank
U. maydis	X07879	GenBank		NP_682256	
P. rhodozyma	AF006483	GenBank		NP_682731	
L. edodes	AB012862	GenBank	*A. variabilis*	YP_321014	GenBank
D. melanogaster	NM_076134	GenBank		YP_324215	
	NM_001038847			YP_322831	
C. elegans	NM_076134	GenBank	*Synechocystis sp. PCC 6803*	NP_440929	GenBank
				NP_442821	
	NM_063836		*L. interrogans*	NP_711885	GenBank
	NM_076133		*C. caviae*	NP_828990	GenBank
	NM_063791		*T. denticola*	NP_972094	GenBank
D. rerio	BC066528	GenBank	*C. diphtheriae*	NP_939663	GenBank
	XM_702342		*T. fusca*	YP_290073	GenBank
M. musculus	XM_987944	GenBank	*T. maritima*	NP_228497	GenBank
	NM_008085		*D. geothermalis*	YP_604599	GenBank
H. sapiens	NM_002046	GenBank	*T. thermophilus*	YP_004524	GenBank
	NM_014364		*G. kaustophilus*	YP_148579	GenBank

Species/Strain	Sequence ID	Database	Species/Strain	Sequence ID	Database
B. pertussis	NP_879794	GenBank		YP_148911	GenBank
L. mexicana	X65226	GenBank	P. aeruginosa	NP_249242	GenBank
	X65220			NP_251691	
E. gracilis	L21904	GenBank		NP_251885	
	L39772		A. tumefaciens	AAL44547	GenBank
T. cruzi	XM_814806	GenBank	M. thermoacetica	YP_429140	GenBank
	XM_810475		B. halodurans	NP_244015	GenBank
				NP_244427	

A representative alignment of the Chromalveolates (Fig. 2.2) reveals 117 amino

acids within the functional Rossman and catalytic domains that are conserved.

Constraints in G3PDH within the *Plasmodium* lineage follow the same general pattern of

G3PDH sequences compared across different phylogenetic lineages; i.e., the catalytic

domain is more conserved than the Rossman domain. When the human and *Plasmodium*

sequences are compared, most variable sites are located within the Rossman domain

(68%). In contrast the catalytic domain contains fewer variable sites (30%). Constraints

in the catalytic domain are mainly due to the way the substrate interacts with the

holoenzyme: glyceraldehyde-3-phosphate and cysteine 174 interact to form a covalent

hemithioacetal intermediate which is then oxidized to a thioester. NAD+ bound to the

enzyme acts as the receptor of a hydride ion and is reduced to NADH. This reaction is

facilitated by the conserved histidine 202 residue in the active site followed by a

phosphorolytic attack on the thioester by inorganic phosphate that releases the final 1, 3-

bisphosphoglycerate product. Therefore, the adenosyl-binding pocket surface may

constitute an obvious target for drug development.

Twenty-five sites at the NAD-binding pocket surface, as determined using the

CASTp program (31), display variation between the human and *Plasmodium* sequences.

Sites conserved along different phylogenetic spectra (i.e. Chromalveolates, Alveolates,

Apicomplexa or the *Plasmodium* lineages) and which are therefore assumed to represent

different functional constraints, may represent attractive target sites. Four sites (Fig. 2.2;

shaded in red) are conserved in the Chromalveolates. Three (L209, G217 and D220) are

located in the so-called S-loop near the catalytic site and residue 216, also within the S-

loop, has only two morphs (a conservative substitution R → K). Of those, site 220 is only

variable in the human sequence (a hydrophobic small amino acid versus a hydrophilic

small amino acid) and therefore seems to be highly constrained in the Chromalveolate

lineage. Two positions (Fig. 2.2; shaded in yellow) are conserved in the Apicomplexa

lineage (N207 and A223) and are close or within the S-loop and represent attractive

targets for functional studies.

Fifteen substitutions (or 17 if excluding *P. gallinaceum*) are conserved within the

Plasmodium lineage (Fig. 2.2, shaded in blue) within the NAD-binding surface.

Substitutions Y47, R/K99, P144, K145 and G263 are of particular interest. In the *P.

falciparum* protein Y47 makes a hydrogen bond with S215, R99 binds NAD+ via a

carbonyl oxygen, P144 and K145 are close to the active site, and G263 may be part of a

surface thought to possibly bind the small molecule CGP-3466 that inhibits the pro-

apoptotic G3PDH activity in human cells (36). While many sites conserved within

Plasmodium may allow for a variety of substitutions in other phylogenetic groups, a few

sites are only dimorphic (H47, R/K99, V210, L122, K216, M256, V341) suggesting

greater constraints. These dimorphic sites could be considered for their potential as

additional inhibitory targets, but also as sites worth probing experimentally from a

functional perspective. It is of value to note that a KG insertion within the S-loop

previously observed in *P. falciparum* is conserved in all *Plasmodium* species (K216,

G217; Fig. 2.2) and the Chromalveolates, further increasing the potential importance of this site as a possible drug target.

Fourteen replacements are unique to mammalian-infecting *Plasmodium* species—H47, L48, L126, I156, S204, V210, V256, I262, G263, T264, V271, A284, L331 and F338. Of those, H47, S204, V210, V256, G263 and L331 are part of the NAD-binding surface. Residues I262 and T264 are part of the central channel formed in the quaternary structure by the assembly of all four monomers.

One would expect that the wider the phylogenetic spectrum being analyzed, the fewer the number of conserved substitutions. Interestingly, when analyzing conserved G3PDH residues that differ from the human sequence within large taxonomic divisions, the majority of them fall into the adenosyl-binding pocket. For example, there are 113 conserved sites between the human sequence and the Chromalveolates and 117 conserved sites in the Chromalveolates, of the four variable residues between the human sequence and the Chromalveolates three are part of the adenosyl pocket, further increasing the interest of those residues from a functional point of view.

2.5. Discussion

This analysis shows that G3PDH in *Plasmodium* is monophyletic and of the cytosolic type. It contains residues with potential functional importance that can be experimentally probed and may eventually constitute appropriate targets for drug development. One hundred and thirty-three positions were identified within the two functional (Rossman NAD+-binding and catalytic) domains that are conserved within the Apicomplexa. Some of these residues are part of the NAD+-binding pocket surface, and as a consequence may be of functional importance: four are conserved in the Chromalveolates, and two within Apicomplexa, while 15 are conserved among *Plasmodium* species. Also of interest are five amino acid substitutions within this surface that are uniquely conserved among *Plasmodium* species. Drug design for malaria and other pathogens has the challenge of designing specific inhibitors that do not affect the function of counterpart proteins present in humans. This study reveals unique amino acid substitutions within functionally important sites that provide attractive targets for therapeutic intervention.

REFERENCES

1. Roll Back Malaria (RBM), W. H. O., and a. U. (WHO). 2005. World Malaria Report, Geneva, p. xvii.

2. Biot, C., and K. Chibale. 2006. Novel approaches to antimalarial drug discovery. *Infect Disord Drug Targets 6:173*.

3. Newton, P., and N. White. 1999. Malaria: New developments in treatment and prevention. *Annual Reviews in Medicine 50:179*.

4. Robien, M. A., J. Bosch, F. S. Buckner, W. C. van Voorhis, E. A. Worthey, p. Myler, C. Mehlin, E. E. Boni, O. Kalyuzhniy, L. Anderson, A. Lauricalla, M. Soltis, F. Zucker, C. L. Verlinde, E. A. Merritt, L. W. Schoenfeld, and W. G. Hol. 2006. Crystal structure of Glyceraldehyde-3-Phosphate Dehydrogenase from Plasmodium falciparum at 2.25A resolution reveals intriguing extra electron density in the active site. *PROTEINS: Structure, Function and Bioinformatics 62:570*.

5. Satchell, J. F., R. L. Malby, C. S. Luo, A. Adisa, A. E. Alpyurek, N. Klonis, B. J. Smith, and P. M. Colman. 2005. Structure of glyceraldehyde-3-phosphate dehydrogenase from *Plasmodium falciparum. Acta Crystallographica Section D 61:1213*.

6. Roth, E. F., Jr., M. C. Calvin, I. Max-Audit, J. Rosa, and R. Rosa. 1988. The enzymes of the glycolytic pathway in erythrocytes infected with Plasmodium falciparum malaria parasites. *Blood 72:1922*.

7. Gratzer, W. 1984. Cell biology. More red than dead. *Nature 310:368*.

8. Roth, E., Jr. 1990. Plasmodium falciparum carbohydrate metabolism: a connection between host cell and parasite. *Blood Cells 16:453*.

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

10. Sherman, I. W. 1998. *Malaria: Parasite Biology, Pathogenesis, and Protection*. ASM Press, Washington, DC.

11. Daubenberger, C. A., F. Poltl-Frank, G. Jiang, J. Lipp, U. Certa, and G. Pluschke. 2000. Identification and recombinant expression of glyceraldehyde-3-phosphate dehydrogenase of *Plasmodium falciparum. Gene 246:255*.

12. Nagradova, N. K. 2001. *Biochemistry 66:1067*.

13. Rich, S. M., and F. J. Ayala. 2003. Progress in Malaria Research: the Case for Phylogenetics. *Advances in Parasitology 54:255*.

14. Waller, R. F., P. J. Keeling, R. G. Donald, B. Striepen, E. Handman, N. Lang-Unnasch, A. F. Cowman, G. S. Besra, D. S. Roos, and G. I. McFadden. 1998. Nuclear-encoded proteins target to the plastid in Toxoplasma gondii and Plasmodium falciparum. *Proc Natl Acad Sci U S A 95:12352*.

15. Barbosa, M. S., S. N. Bao, P. F. Andreotti, F. P. de Faria, M. S. Felipe, L. dos Santos Feitosa, M. J. Mendes-Giannini, and C. M. Soares. 2006. Glyceraldehyde-3-phosphate dehydrogenase of Paracoccidioides brasiliensis is a cell surface protein involved in fungal adhesion to extracellular matrix proteins and interaction with cells. *Infect Immun 74:382*.

16. Colell, A., J. E. Ricci, S. Tait, S. Milasta, U. Maurer, L. Bouchier-Hayes, P. Fitzgerald, A. Guio-Carrion, N. J. Waterhouse, C. W. Li, B. Mari, P. Barbry, D.

D. Newmeyer, H. M. Beere, and D. R. Green. 2007. GAPDH and autophagy preserve survival after apoptotic cytochrome c release in the absence of caspase activation. *Cell 129:983*.

17. Sirover, M. A. 1999. New insights into an old protein: the functional diversity of mammalian glyceraldehyde-3-phosphate dehydrogenase. *Biochim Biophys Acta 1432:159*.

18. Daubenberger, C. A., E. J. Tisdale, M. Curcic, D. Diaz, O. Silvie, D. Mazier, W. Eling, B. Bohrmann, H. Matile, and G. Pluschke. 2003. The N'-terminal domain of glyceraldehyde-3-phosphate dehydrogenase of the Apicomplexan *Plasmodium falciparum* mediates GTPase Rab2-dependent recruitment to membranes. *Biological Chemistry 384:1227*.

19. Campanale, N., C. Nickel, C. A. Daubenberger, D. A. Wehlan, J. J. Gorman, N. Klonis, K. Becker, and L. Tilley. 2003. Identification and characterization of heme-interacting proteins in the malaria parasite, *Plasmodium falciparum*. *Journal of Biological Chemistry 278:27354*.

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

21. Brady, R. L., and A. Cameron. 2004. Structure-based approaches to the development of novel anti-malarials. *Curr Drug Targets 5:137*.

22. Dobeli, H., A. Trzeciak, D. Gillessen, H. Matile, I. K. Srivastava, L. H. Perrin, P. E. Jakob, and U. Certa. 1990. Expression, purification, biochemical characterization and inhibition of recombinant Plasmodium falciparum aldolase. *Mol Biochem Parasitol 41:259*.

23. Wanidworanun, C., R. L. Nagel, and H. L. Shear. 1999. Antisense oligonucleotides targeting malarial aldolase inhibit the asexual erythrocytic stages of Plasmodium falciparum. *Mol Biochem Parasitol 102:91*.

24. Lakhdar-Ghazal, F., C. Blonski, M. Willson, P. Michels, and J. Perie. 2002. Glycolysis and proteases as targets for the design of new anti-trypanosome drugs. *Curr Top Med Chem 2:439*.

25. Opperdoes, F. R., and P. A. Michels. 2001. Enzymes of carbohydrate metabolism as potential drug targets. *Int J Parasitol 31:482*.

26. Aronov, A. M., S. Suresh, F. S. Buckner, W. C. Van Voorhis, C. L. Verlinde, F. R. Opperdoes, W. G. Hol, and M. H. Gelb. 1999. Structure-based design of submicromolar, biologically active inhibitors of trypanosomatid glyceraldehyde-3-phosphate dehydrogenase. *Proc Natl Acad Sci U S A 96:4273*.

27. Bressi, J. C., C. L. Verlinde, A. M. Aronov, M. L. Shaw, S. S. Shin, L. N. Nguyen, S. Suresh, F. S. Buckner, W. C. Van Voorhis, I. D. Kuntz, W. G. Hol, and M. H. Gelb. 2001. Adenosine analogues as selective inhibitors of glyceraldehyde-3-phosphate dehydrogenase of Trypanosomatidae via structure-based drug design. *J Med Chem 44:2080*.

28. Suresh, S., J. C. Bressi, K. J. Kennedy, C. L. Verlinde, M. H. Gelb, and W. G. Hol. 2001. Conformational changes in Leishmania mexicana glyceraldehyde-3-phosphate dehydrogenase induced by designed inhibitors. *Journal of Molecular Biology 309:423*.

29. O'Sullivan, O., K. Suhre, C. Abergel, D. G. Higgins, and C. Notredame. 2004. 3DCoffee: combining protein sequences and structures within multiple sequence alignments. *J Mol Biol 340:385*.

30. Gille, C., and C. Frommel. 2001. STRAP: editor for STRuctural Alignments of Proteins. *Bioinformatics 17:377*.

31. Binkowski, T. A., S. Naghibzadeh, and J. Liang. 2003. CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic Acids Res 31:3352*.

32. Hall, T. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series [0261-3166] 41:95*.

33. Guindon, S., F. Lethiec, P. Duroux, and O. Gascuel. 2005. PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res 33:W557*.

34. Abascal, F., R. Zardoya, and D. Posada. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics 21:2104*.

35. Fast, N. M., J. C. Kissinger, D. S. Roos, and P. J. Keeling. 2001. Nuclear-encoded, plastid-targeted genes suggest a single common origin for Apicomplexan and Dinoflagellate plastids. *Molecular Biology and Evolution 18:418*.

36. Jenkins, J. L., and J. J. Tanner. 2006. High-resolution structure of human D-glyceraldehyde-3-phosphate dehydrogenase. *Acta Crystallogr D Biol Crystallogr 62:290*.

CHAPTER THREE

***PLASMODIUM VIVAX* PHIST-81 IS IDENTIFIED AND VISUALIZED BY 3-D IMAGING IN THE CAVEOLA-VESICLE COMPLEXES OF *P. CYNOMOLGI*-INFECTED ERYTHROCYTES**

Sheila Akinyi[a], Eric Hanssen[b], Cindy C. Korir[a], Balwan Singh[a], Esmeralda VS Meyer[a], John W. Barnwell[c], Leann Tilley[b] and Mary R. Galinski[a,d].

[a]Emory Vaccine Center, Yerkes National Primate Research Center, Emory University, Atlanta, GA 30329, USA.

[b]Department of Biochemistry and Centre of Excellence for Coherent X-ray Science, La Trobe University, VIC 3086, Australia.

[c]Laboratory Research & Development Unit, Malaria Branch, Division of Parasitic Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30341, USA

[d]Department of Medicine, Division of Infectious Diseases, Emory University, Atlanta, GA 30329, USA.

This chapter will be submitted for publication.

3.1. Abstract

In each *Plasmodium* species, the intracellular blood-stage parasite forms membranous networks within the red blood cell (RBC) cytoplasm, allowing communication between the parasite's plasma membrane and the erythrocyte surface, and enabling processes like nutrient import, protein export and RBC cytoadherence. While structures known as Maurer's clefts and electron dense knobs are characteristic of *Plasmodium falciparum*-infected erythrocytes, extensive cytoplasmic cleft membranes and caveola-vesicle complexes (CVCs) are typical for *Plasmodium vivax*. Information on the makeup and function of these *P. vivax* features is scarce. Building upon earlier published data showing that four monoclonal antibodies raised against *P. vivax* schizonts recognize a 95 kDa antigen by SDS-PAGE that localizes to CVCs and cytoplasmic vesicles by immunoelectron microscopy, in this study we used proteomic methods to identify the genes encoding the corresponding protein. LC-MS/MS analysis of the immunoprecipitated 95 kDa protein from *P. cynomolgi* infected RBCs identified this protein as a homolog of an 80.73 kDa *P. vivax* member of the *Plasmodium* helical interspersed sub-telomeric (PHIST) protein superfamily. We have named the respective proteins PcyPHIST-81 and PvPHIST-81 and have identified and compared the most closely related homologs present in *P. knowlesi*, *P. falciparum*, *P. berghei* and *P. yoelii* databases. Each contains a characteristic PEXEL motif near the N-terminus, and four consecutive alpha helices and four conserved tryptophans near the C-terminus. In this report we also highlight a diverse central repetitive domain, evident at the amino acid or nucleotide levels, which has not been reported previously as a feature of PHIST family members. PcyPHIST-81 is expressed throughout the erythrocytic stages, with notably

high expression in trophozoites, when CVCs are being produced at the surface of the infected host cell. Moreover, immunoelectron tomography studies using a cross reactive rabbit antiserum raised against the recombinant PvPHIST-81 protein indicates that PcyPHIST-81 localizes to the cytoplasmic side of the tubular extensions of the CVCs. This study represents the first confirmed biological localization of a member of the PHIST superfamily, and a step towards understanding the biogenesis, structure and function(s) of the CVCs, and revealing possible novel targets of intervention.

3.2. Introduction

Plasmodium vivax research, facilitated by simian models, particularly the closely

related *P. cynomolgi* species, continues to be critical for developing much needed

vaccines and drugs. Such tools are undeniably essential in light of *Plasmodium vivax*

causing an estimated 80 million to 350 million clinical cases of malaria illness each year,

with severe disease and lethal cases also being revealed with increasing regularity (1-4).

Efforts to address *P. vivax* research priorities have concluded that the molecular and cell

biological basis underlying how *P. vivax* parasites invade and remodel their host red

blood cells, the reticulocytes, represents an important gap in knowledge, relevant for

revealing drug or vaccine targets. Furthermore, since *P. vivax* cannot be cultured

continuously *in vitro*, and the quantities of *P. vivax* parasites from *in vivo* sources are

limited, *P. cynomolgi* has been recognized as a valuable model to address such questions

experimentally (5).

When a *Plasmodium* merozoite invades a host red blood cell (RBC), it creates and

then resides within a parasitophorous vacuole (4). Although the enucleated RBC lacks

protein synthesis and trafficking machinery, the parasite developing within has the

immediate capability to produce and transport proteins to the parasitophorous vacuole

membrane, onwards to newly formed membranous structures in the cytoplasm called

clefts, and to the surface of the RBC membrane. Investigations of nascent parasite-

induced cytoplasmic networks and RBC membrane structures created as the parasite

remodels the host cell have been most advanced for *P. falciparum,* resulting in many

breakthroughs in the scientific understanding of this species' transport biology (6-8).

This research has been aided by the availability of well established long-term *in vitro P.*

falciparum blood-stage culture systems, a *P. falciparum* genome database (9), and functional genomics technologies (10). In contrast, reliable *P. vivax* culture systems have not been established, and the first *P. vivax* genome was just published recently (11). Investigations relating to the unique species-specific modifications of *P. vivax*-infected RBCs have therefore lagged behind the research momentum established for studies on *P. falciparum*. However, today, with the availability of a *P. vivax* genome database and phylogenetically closely related simian experimental models (12, 13), there is much promise for moving this field of research forward. This aim is in line with the global goal to focus increased attention on all species of *Plasmodium* infecting humans, to work towards the elimination and ultimate eradication of malaria (14).

Parasite-derived structures within the infected RBC cytoplasm and at the surface of the host cell differ for each species of *Plasmodium*. While *P. falciparum* produces electron-dense protrusions called knobs at the surface of the infected host cell, *P. vivax* induces the formation of caveola-vesicle complexes (CVCs) (15). In *P. falciparum*, knobs are associated with the accumulation of erythrocyte membrane protein 1 (PfEMP1), the main antigenic ligand in this species that is responsible for both cytoadherence and antigenic variation (16-18). How this protein and others are trafficked from the parasite to their final destination at or near the surface of the infected RBC has been of extreme interest in regards to identifying critical functions of the parasite proteins and developing associated interventions (7, 19). There is evidence that the *P. vivax* CVCs also contain parasite-derived proteins (20, 21) (22-24) and knowledge on these molecules, and their associated functions and metabolic pathways may lead to the development of inhibitors that can disrupt their function.

The CVCs in *P. vivax* and *P. cynomolgi* infected RBCs were initially observed by electron microscopy in 1975, at which time accompanying immunochemical experiments supported the view that these structures were open to the exterior of the host cell (15). The frequency of the CVCs along the plasmalemma also supported the idea that these structures corresponded to the numerous Schuffner's dots (i.e., a speckling appearance, (25)) resulting from the accumulation of Giemsa stain at the surface of trophozoite and schizont-infected RBCs on a blood smear. The CVCs had been described as 90-100 nm caveolae with 40-50 nm wide electron-dense vesicles extending from the caveolar base; some of the vesicles appeared alveolar in shape, while others, uniquely so far observed in *P. cynomolgi*-infected RBCs, appeared more elongated and tubular (15). How they are formed and the functional significance of CVCs is not known, although it has been suggested that they may function to incorporate RBC plasma proteins, or to transport and release *Plasmodium* antigens from infected erythrocytes (15, 24, 26, 27). CVCs are also a characteristic feature of the human malaria species *P. ovale* and the simian species *P. simium* (15, 28-30). Each of these species can be identified in blood smears in part based upon the profuse speckling appearance caused by Giemsa staining of their CVCs. *Plasmodium knowlesi*, on the other hand, only has caveolae, without the associated vesicle complexes, and the caveloae in these species are few in number at the surface of the host cell (15, 24, 26, 27).

We set out to investigate the composition and structure of the CVCs of *P. vivax* by capitalizing on the routine availability of *P. cynomolgi* infected RBCs. *P. cynomolgi* is phylogenetically closely related to *P. vivax* (12, 13) and rhesus macaques can be experimentally infected with this species, providing a greater amount of parasite material

for molecular and cell biological studies than is possible from experimental *P. vivax* infections of small New World monkeys (31). Importantly, the close phylogenetic and biological relatedness between *P. vivax* and *P. cynomolgi* with genes and proteins generally showing 85% to 95% identity permits the development of cross reactive reagents, database mining using the *P. vivax* genome database (11), and inferences with regards to structural and functional insights.

In this study, a 95 kDa protein was immunoprecipitated from *P. cynomolgi* trophozoite detergent extracts with a panel of cross-reacting monoclonal antibodies raised against *P. vivax* schizonts. The mAbs had previously localized the homologous 95 kDa protein by immunoelectron microscopy to the CVCs of *P. vivax* (22-24) (22, 27). Proteomic analysis of the immunoprecipitated 95 kDa protein revealed its identity as a member of the *Plasmodium* Helical Interspersed SubTelomeric (PHIST) family of proteins with a calculated molecular weight of 80.73 kDa; thus, termed here PHIST-81. In light of recent advances using electron tomography to study the secretory system of *P. falciparum* infected RBCs in 3-dimensions (32), tomograms were similarly created to study the CVCs of *P. cynomolgi*. Here we show the elaborate multi-vesicular/tubular nature of the CVC structures for the first time via 3-D imaging and define the location of a specific resident protein, PcyPHIST-81, on the cytoplasmic side of the tubules. These studies mark the beginning of post-genomic investigations focused on the stage-specific development of these intricate structures, the trafficking and functioning of the PHIST-81 among other particular proteins, and the search for specific targets of intervention against *P. vivax.*

3.3. Materials and Methods

Parasites

P. cynomolgi (Berok)-infected RBCs were inoculated into a rhesus monkey

following approvals and protocols from the Institutional Animal Care and Use

Committee at Emory University. At approximately 5% parasitemia, blood was collected

from the monkey and passed through glass beads and cellulose CF11 columns to remove

platelets and white blood cells, respectively (33). The blood was then enriched for

trophozoite or schizont-infected RBCs by centrifugation over a 48% or 52% Percoll

cushion, respectively, as described (22).

***P. vivax* Monoclonal Antibodies**

P. vivax monoclonal antibodies 2H8.E10, 2H12.B4, 4C12.B4 and 1H4.B6 were

produced and characterized as previously described (22, 27). Total IgG was then purified

using the Protein A MAPS affinity isolation system (BioRad) following the

manufacturer's protocol.

Production of Recombinant Protein and Antisera

Pvphist-81 gene-specific primers PvPHIST-81F (5'tat gga tcc ATG AGT CCC

TGC AAC ATC) and PvPHIST-81R (5'ata ctc gag TTA GAG TTT GCT GTG TTT CT)

were used to amplify the full length of the gene under standard PCR conditions following

the manufacturer's protocol (Calbiochem). The amplicon was cloned into the expression

vector pGEX 4T-2 (GE Healthcare) using the *Bam*-HI and *Xho*-I restriction sites. Positive

clones were confirmed by sequencing using the ABI 3100 DNA sequencer. The clones

were then re-transformed into *E. coli* BL21 Star™ (DE3) cells (Invitrogen) for protein expression. Soluble protein was purified using Glutathione Sepharose™ 4B (GE Healthcare) slurry according to the manufacturer's protocol. Recombinant PvPHIST-81 (rPvPHIST-81) was inoculated into a New Zealand White Rabbit (Covance) for production of polyclonal antiserum, rabbit anti-rPvPHIST-81.

Indirect immunofluorescence assays (IFA) and Western blots

The cross-reactivity of *P. vivax* mAbs was tested by IFA on air-dried, acetone-fixed thin films of RBCs infected with late trophozoite or schizont *P. cynomolgi* parasites. Following incubation with the primary antibodies, expression was detected using affinity-purified goat IgG anti-mouse conjugated to Alexa Fluor 488 (Invitrogen) as secondary antibodies. The mAbs were tested at 1:100, 1:200, 1:400 and 1:800 dilutions in phosphate-buffered saline (PBS; Lonza) containing 0.2% bovine serum albumin (BSA; Sigma).

IFA studies were carried out with rabbit anti-rPvPHIST-81 at dilutions of 1:100, 1:200, 1:400 and 1:800. This was followed by anti-rabbit antibodies conjugated to Alexa Fluor 488 (Invitrogen) at a dilution of 1:200. The parasite nuclei were visualized with DAPI, contained in ProLong Gold Antifade Reagent (Invitrogen). The slides were examined with an Axioscope Z.1 microscope with filters appropriate for the fluorescent dyes, and the images merged.

For immunoblot analyses, *P. cynomolgi* trophozoite extracts were electrophoretically separated on SDS-polyacrylamide gels and then transferred to a 0.2 μm nitrocellulose membrane (Schleicher & Schuell) and probed with the mAbs diluted

1:1000. Membranes were incubated with the corresponding alkaline phosphatase-conjugate as a secondary antibody (Promega) and immunoreactivity was detected by incubating with NBT/BCIP substrate (Promega).

Immunoprecipitation of *P. cynomolgi* Extracts with Monoclonal Antibodies

    P. cynomolgi-infected RBC extracts were prepared as follows: Ice-cold 1X NET/1% NP-40 containing protease inhibitors (10 mM EDTA-Na$_2$; 1 mM PMSF; 0.1 mM each of TPCK, TLCK, Leupeptin, Chymostatin, Antipain and 3,4-DCI; 10 µM EP-64 and 1 µM Pepstatin A; Sigma) was added to *P. cynomolgi*-infected erythrocyte pellet, and extracted for 30 minutes on ice with occasional vortexing. The samples were then transferred to pre-chilled microcentrifuge tubes and spun at 14,000rpm for 20 minutes at 4ºC. The resulting NP-40 extracts were transferred and stored, while the pellet was extracted in 1% SDS by occasional vortexing for 10 minutes at RT. The sample was then centrifuged at 14,000rpm for 20 minutes at RT. The SDS extracts were combined with the NP-40 extracts for use in immunoprecipitations.

    160 µl of rProtein G agarose suspension (Invitrogen) was incubated with 200 µg of each *P. vivax* mAb overnight at 4°C, rotating. Meanwhile, extracts were preclarified with rProtein G agarose overnight at 4°C, on a rotational shaker. The extracts were then separated from the preclarifying beads. Dimethyl pimelimidate (Sigma) was added to the mAb/rProtein G mix to a final concentration of 20 mM and incubated for 6 h, 4°C, rotating. The uncoupled mAb was then removed by pulse spin, and the mixture washed with NETT (150 mM NaCl/ 5 mM EDTA/ 50 mM Tris/ 0.5% Triton X-100). One ml of the combined extract was added to the coupled mAb/rProtein G mix and incubated

overnight. The beads were then washed twice with NETT, twice with NETT/0.5 M NaCl and once with NETT/0.05% SDS. The samples were resuspended in 2X SDS-PAGE loading buffer and resolved on 4-20% gradient SDS-PAGE gels (Bio-Rad).

Mass Spectrometric Analysis of Immunoprecipitated Protein

After resolving the extracts and membrane samples on 4-15% SDS-PAGE gradient gels, the gels were stained with colloidal Coomassie blue (Imperial Protein Stain; Thermo Scientific). Gel slices were then excised, destained, dried, and processed as described previously (34). Briefly, the gel pieces were digested with trypsin (Sigma) and the resulting peptides extracted with trifluoroacetic acid (Sigma). The samples were then desalted and concentrated using ZipTip pipette tips (Millipore).

Cleaned peptides were analyzed by reverse-phase liquid chromatography coupled with tandem mass spectrometry (35) using an LTQ-Orbitrap mass spectrometer (Thermo Finnigan). A reverse database strategy using the SEQUEST algorithm was implemented to evaluate false discovery rate; the matched peptides were filtered according to matching scores to remove all false matches from the reverse database (36). Only proteins that were matched by at least two peptides were accepted to further improve the confidence of identification. The peptides were then searched against the NCBI database, with searches being limited to *Plasmodium* results.

Bioinformatics for peptide and gene analysis

Acquired protein matches from LC-MS/MS were searched against *Plasmodium* (PlasmoDB; www.plasmodb.org) and general (National Center for Biotechnology and

Information; http://www.ncbi.nlm.nih.gov/sites/entrez?db=Protein&itool=toolbar)

databases to determine the homologous genes.

Signal peptide cleavage sites and transmembrane domains were predicted with

SignalP V3.0 (http://www.cbs.dtu.dk/services/SignalP/) and TMpred

(www.ch.embnet.org /software/TMPRED_form.html) software, respectively. Multiple

alignments of PHIST protein sequences were generated using ClustalW2

(http://www.ebi.ac.uk/Tools/clustalw2/index.html) with all default parameters.

The GOR4 program (http://npsa-pbil.ibcp.fr/cgi-

bin/npsa_automat.pl?page=npsa_gor4.html; (37) was used to predict secondary structure

using the amino acid sequences of the PHIST family members.

Isolation of *P. cynomolgi* DNA and RNA

P. cynomolgi genomic DNA (gDNA) was prepared from schizont-stage parasites

using the QIAamp DNA Blood Extraction kit (Qiagen), as described previously (33).

Polymerase chain reaction (PCR) amplification and cloning of *pcyphist-81*

Because the *P. cynomolgi* PHIST-81 gene sequence was unavailable, PCR

amplification of *P. cynomolgi* gDNA was performed using a combination of *pvphist-81*

and *pkphist-81*gene-specific primers and the Expand High Fidelity System (Roche) kit as

per the manufacturer's protocol. PCR products were then purified using the Qiaquick

purification system (Qiagen), cloned into the pCR2.1 vector (Invitrogen) and sequenced

using the ABI Prism BigDye Terminator v3.0 cycle sequencing kit (Applied Biosystems).

By gene walking, and the subsequent design of *pcyphist-81*-specific primers, the first

2200 bases of the *pcyphist-81* gene were sequenced and verified.

In order to sequence the 3' end of *pcyphist-81* up to the stop codon, degenerate

reverse primers were designed based on the 3' UTR sequences of *P. vivax* and *P.

knowlesi phist-81*. The primer pair PcyPHIST.2023.F (5'GAT GCA AGA GTA CAT

TAT GC) and PcyPHIST.3'UTR.R (5'CAA AA(A/C) GTT CTC CTA TGA CG)

amplified the sequence up to the stop codon. This result was confirmed by sequencing the

entire gene using *pcyphist81*-specific primers:

PcyPHIST.1.F (5'ATG AGT CCC TGC AAC ATC)

PcyPHIST.263.R (5'CTC AGA GAG ATA TGC TCA AA)

PcyPHIST.566.R (5'CAT CTC CTC CTC TTG CCA)

PcyPHIST.762R (5'TCA GAG GGA TCG GTA TCG)

PcyPHIST.1229.R (5'CAC CTC TTC CGT GGT ATT)

PcyPHIST.2023.F (5'GAT GCA AGA GTA CAT TAT GC)

PcyPHIST.2160.R (5'GAG TAT TGC ATA ATG TAC TC)

PcyPHIST.2433.R (5'TAC AAT TTA CTG TGT TTC TTC).

Immunoelectron Microscopy

Permeabilization of *P. cynomolgi*-infected RBCs with Equinatoxin II (EqtII) was

performed as previously described (32, 38). Briefly, infected RBCs were fixed in

incomplete RPMI containing 2% paraformaldehyde (PFA) and then permeabilized with

30 μg EqtII (39). The samples were refixed in PBS containing 2% paraformaldehyde,

blocked with 3% BSA in PBS and then incubated with rabbit anti-rPvPhist-81 as a

primary antibody at a 1:10 dilution in PBS/3% BSA. After washing, cells were incubated with gold-conjugated Protein A beads (Aurion) at a 1:10 dilution in PBS/3% BSA.

After washing, cells were fixed with 1% glutaraldehyde/0.5% PFA/0.1 M cacodylate buffer overnight. The samples were then embedded in 3% agarose and rinsed with 0.175 M cacodylate buffer. After post-fixation with 1% osmium tetroxide, the cells were stained 'en-bloc' with 1% uranyl acetate. The cells were then serially dehydrated and then embedded in LR White resin. The samples were sectioned to 70nm thickness and after staining with lead citrate and uranyl acetate, observed on a 2010HC microscope (Jeol, Japan; La Trobe University EM Facility) at 120kV.

Electron Tomography

Tomography was performed as described previously (32). 300nm serial sections were cut and collected on a grid then incubated with fiducial gold particles. The sections were then contrasted with lead citrate and uranyl acetate and observed on a tilt series from -69 degrees to 69 degrees at every 1.5 degrees between captured images for the first axis and every 3 degrees between captured images for the second axis using a Tecnai G2 TF30 (FEI Company; Bio21 Institute, Melbourne) transmission electron microscope at 200kV. The tilt images were aligned using IMOD software (40) and tomograms were then generated.

3.4. Results

Proteomics identify PcyPHIST-81 as a member of the PHIST protein superfamily in *P. cynomolgi* infected RBCs, corresponding to PvPHIST-81, a *P. vivax* CVC protein

Barnwell *et al.* had previously shown by immuno-EM that *P. vivax* mAbs 2H12.B4, 2H8.E10, 4C12.B4, and 1H4.B6 localized to the CVCs of *P. vivax* blood-stage parasites and immunoprecipitated an antigen from ^{S}35-metabolically-labeled *P. vivax* blood-stage parasite SDS detergent extracts, which migrated at 95 kDa on SDS-PAGE gels (22, 23). Given the restricted availability of *P. vivax* infected RBCs for routine in-depth biological experiments, this set of *P. vivax* mAbs were tested by IFA on blood smears containing *P. cynomolgi* blood-stage parasites as a surrogate to advance studies on the CVCs. Antibody cross reactivity was observed with *P. cynomolgi* infected RBCs and a clear fluorescence pattern with numerous speckles was produced with three of the antibody reagents (Figure 3.1a). This heavily dotted pattern resembles Schuffner's dots on Giemsa-stained thin blood smears.

(a)

P. vivax mAbs	IFA
4C12.B4	
2H8.E10	
2H12.B4	
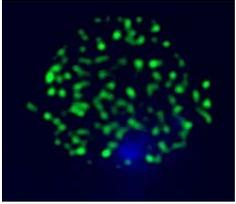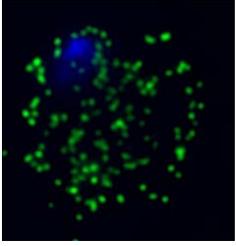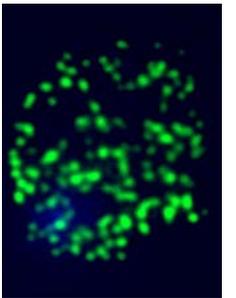
(b)
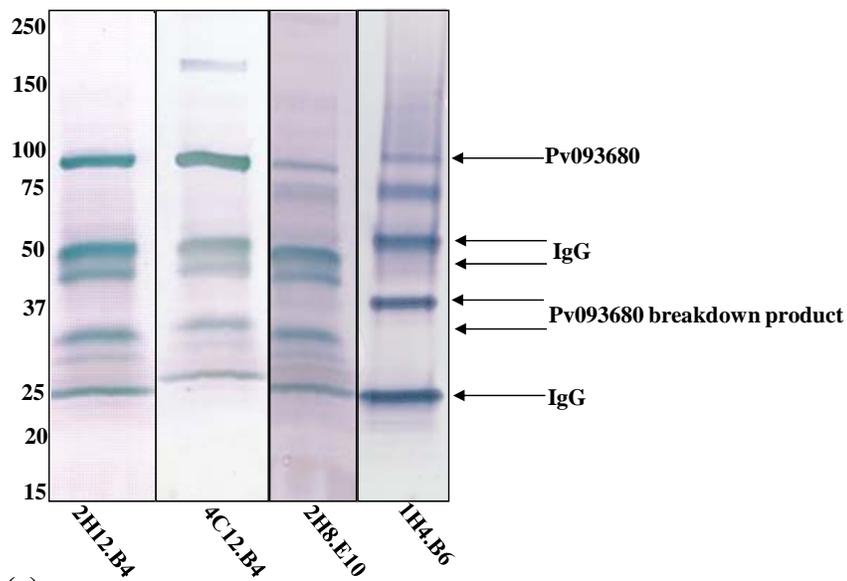
(c)
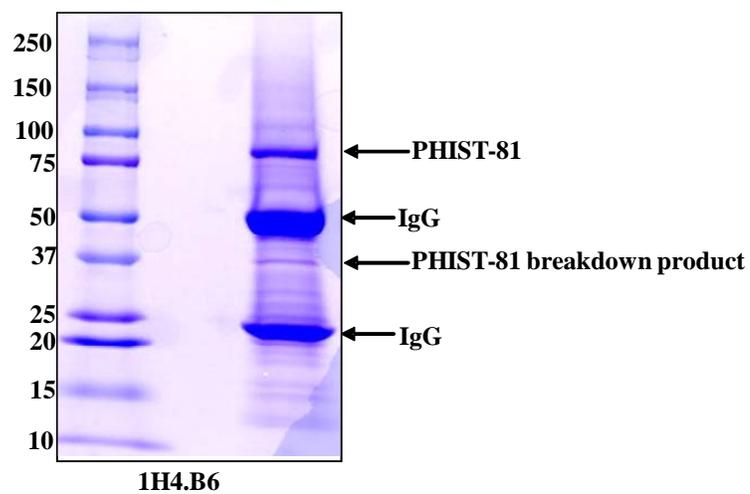
Figure 3.1: Proteomics identify PcyPHIST-81 as a member of the PHIST protein superfamily in *P. cynomolgi* infected RBCs.

(a) Fixed smears of *P. cynomolgi*-infected trophozoite-stage erythrocytes were incubated with *P. vivax* mAbs 4C12.B4, 2H8.E10 and 2H12.B4. Cross-reactivity was detected using Alexafluor-488-conjugated anti-mouse IgG (Invitrogen) and an Axioscope 4.1 microscope.

(b) Western immunoblot analysis of *P. cynomolgi* trophozoite-infected RBCs subjected to SDS-PAGE under reducing conditions on a 4-20% gradient gel followed by probing with *P. vivax* mAbs 4C12.B4, 2H8.E10 and 2H12.B4. The identities of the specific bands at 95 kDa, 50 kDa, ~37 kDa and 20 kDa were determined by LC-MS/MS.

(c) 4-20% SDS-PAGE gel showing protein bands immunoprecipitated with mAb 1H4.B6. Similar results were obtained with mAbs 4C12.B4, 2H8.E10 and 2H12.B4 (not shown).

We confirmed by immunoblot analysis of *P. cynomolgi* extracts that all four mAbs initially tested with *P. vivax* showed positive reactivity with an antigen migrating at 95 kDa (Figure 3.1b). Other protein bands detected on the immunoblot were alternatively confirmed by mass spectrometry to correspond to either IgG or PcyPHIST protein breakdown products (Fig. 3.1b). All four *P. vivax* mAbs also immunoprecipitated a protein from *P. cynomolgi* extracts that migrated at 95 kDa on SDS-PAGE gels; this result is exemplified by mAb 1H4.B6 in Figure 3.1c. The immunoprecipitated *P. cynomolgi* 95 kDa protein was excised from the SDS-PAGE gel slices and analyzed by reverse-phase liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). When the resulting peptide sequences were searched against the *P. vivax* genome database a member of the *Plasmodium* helical interspersed subtelomeric (PHIST)

superfamily of proteins was identified with the gene ID: Pvx_093680 and a predicted

mass of 80.73 kDa (henceforth referred to as PvPHIST-81).

The *pvphist-81* gene represented by Pvx_093680 is 2,133 nucleotides long in a

single exon (11). The encoded 710 amino acid PvPHIST-81 protein's calculated mass of

80.73 kDa is slightly lower than the electrophoretic mobility of 95 kDa. Bioinformatic

analysis with the MalSig algorithm (http://bioserve.latrobe.edu.au/cgi-bin/pfsigseq.py)

indicates that PvPHIST-81 has a predicted recessed signal sequence or signal anchor

starting 52 amino acids downstream of the N-terminus, and spanning 23 amino acids

(Table 3.1). A PEXEL motif, which predictably is cleaved and N-acetylated in the

parasite's endoplasmic reticulum (41), is located almost immediately downstream of this

sequence. PvPHIST-81 can be classified as aspartate (11.7%) and glycine (11%) rich

(http://ca.expasy.org/tools/protparam.html; Table 3.1), though this may simply be due to

the presence of degenerated repeated nucleotide sequence motifs (discussed further

below) by chance encoding a correspondingly higher content of these amino acids.

Table 3.1: Characteristics of *P. vivax* PHIST-81

Accession Number[a]	Pvx_093680
Calculated Mass	80.73 kDa
Approx. Size of Gel Slice Excised[b]	95kDa
P. vivax Database Annotation[c]	PHIST protein (Pf-fam-b)
IFA Localization[d]	Stippling pattern, suggestive of CVCs
IEM Localization	CVCs (22)
Predicted Signal Peptide[e]	Yes
Predicted Transmembrane Domains[f]	None
Pexel/Host Targeting Domain[g]	Yes
Size of Coding DNA Sequence[h]	2,133 nucleotides
Protein Expression Profile (Blood Stages)[i]	R, T, S
P. cynomolgi Homolog (% identity)[j]	PcyPHIST-81 (100%)
P. knowlesi Homolog (% identity)[j]	PkPHIST-105 PkH_011720 (89%)
P. falciparum Homolog (% identity)[j]	PfPHIST-147 PF08_0137 (55%)
P. yoelii Homolog (% identity)[j]	PyPHIST-110 PY01786 (32%)
P. berghei Homolog (% identity)[j]	PbPHIST-122 PB108348.00.0 and PB000848.03.0 (31%)

[a]NCBI accession number of PvPHIST-81 identified by mass spectrometry analysis.

[b]Approximate size of the gel slice that was excised from the SDS-PAGE gel.

[c]Annotation of PvPHIST-81.

[d]Localization of PvPHIST-81 determined by the fluorescence pattern in IFAs using rabbit antisera against rPvPHIST-81. CVC=caveola-vesicle complex.

[e]Recessed signal peptides were predicted by MalSig (42, 43); malaria secretary signal predictions; http://bioserve.latrobe.edu.au/cgi-bin/pfsigseq.py).

[f]Transmembrane domains were predicted by TMHMM program (44);http://www.cbs.dtu.dk/services/TMHMM/).

[g]*Plasmodium* export element/ host targeting domain (45, 46).

^hPredicted size of spliced DNA sequence.

ⁱProtein expression profile determined by IFA and immunoblot

^j The NCBI *Plasmodium* database was blasted using the PvPHIST sequence, and separately with the PHIST domain from this protein. The most closely related *Plasmodium* homolog protein sequences were aligned with the *P. vivax* sequence using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html) and the percent amino acid similarity to the C-terminus of PvPHIST-81 determined.

Bioinformatic analyses of *pvphist-81, pcyphist-81*and other predicted orthologs reveal degenerate repeated motifs between the PEXEL motif and PHIST domain.

To support future investigations on the biological and functional characteristics of the CVCs using the *P. cynomolgi* model, we proceeded to identify and characterize the *pcyphist-81* gene. Degenerate gene-specific and 3'UTR degenerate primer pairs were designed (see materials and methods) based on *pvphist-81* coding and non coding 3'UTR regions, complemented with primers based on the corresponding gene sequence from *P. knowlesi* (PkH_011720) (47) to successfully amplify by PCR and sequence the *pcyphist-81* gene (Berok strain). RT-PCR and sequencing verified the *pcyphist-81* gene to be 2,433 nucleotides, encoding a 720 amino acid protein, in two exons (Figure 3.2). A 270 nucleotide intron separates a short initial exon of 213 nucleotides from the second exon of 1,950 nucleotides. This *pcyphist-81*gene structure differs from the reported *pvphist-81* gene structure with one exon (11), which we verified by RT-PCR and sequencing. The coding region of these genes is otherwise 83% identical and the amino acid identity is 75%.
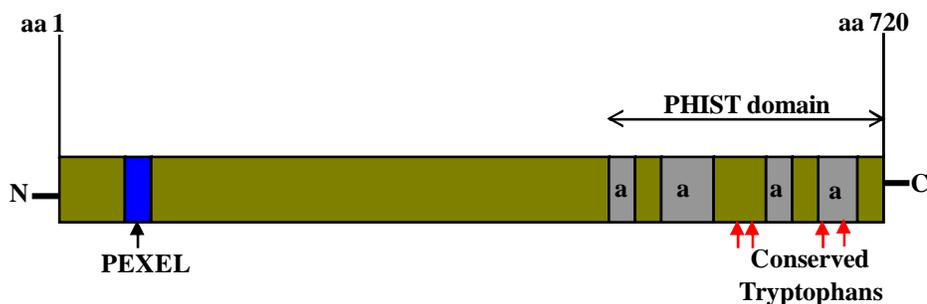
Figure 3.2: Protein structure of PcyPHIST-81. The schematic represents the full-length amino acid sequence of *P. cynomolgi* PHIST-81. The protein is 720 amino acids long, with a PEXEL motif located 86 amino acids downstream of the start methionine (blue). Two characteristic features of PHIST proteins are highlighted: the predicted PHIST domain spanning 150 amino acids and characterized by four consecutive alpha helices (grey), and the conserved tryptophan residues (red arrows).

Additional predicted *pvphist-81* orthologs were detected in *P. falciparum* and *P. knowlesi* through BLAST (Basic Linear Alignment Search Tool) analyses with the most closely related sequences identified from the PlasmoDB database as Gene IDs Pf08_0137 and PkH_011720, respectively. The *P. falciparum* protein encoded by Pf08_0137 is annotated as a *Plasmodium* exported protein of unknown function. It has a calculated molecular mass of 147 kDa and shares 20% overall amino acid sequence identity with PvPHIST-81 (Table 3.1). The *P. knowlesi* protein encoded by PkH_011720 on the other hand is annotated as a conserved, hypothetical protein, and it has a calculated molecular mass of 105 kDa and has 56% overall sequence identity with PvPHIST-81 at the amino acid level (Table 3.1). When the C-terminus region alone is compared, containing the conserved PHIST domain, much higher amino acid identities are reflected; 55% and 89%, for *P. falciparum* and *P. knowlesi*, respectively (Table 3.1). Given these

relationships at the gene and protein levels, we suggest the designation of these proteins as PfPHIST-147 and PkPHIST-105. Clustal alignments highlighting the relationship of PvPHIST-81, PcyPHIST-81, PfPHIST-147 and PkPHIST-105 proteins are shown in Figure 3a. Each of these proteins can be classified as members of the PHISTc subfamily, based on the position of their four conserved tryptophans, a defining characteristic of PHIST proteins (48).

```
PvPHIST-81   MSPCNIPIIVLPATSSPADIKKAIDANNQTVAKSDGRRNKCKSRKSFADFNLASLPYKVL 60
PcyPHIST-81  MSPCNIPIIVLPATSSPADIKRAIDANNQAVAKSDGGRNKCKSKKSFADFNLASLPYKVL 60
PkPHIST-105  MNPCNIPIIVLPGTSSPADIRRAIDTNNQTVATSDGKKSKCKSRKSFADFNLASLPYKVL 60
PfPHIST-147  MIFVKSKILYFLKWPSVAIEENFSGSFKCLFKNKRNKYVRNNLSRFILSVSLILFFFHFV 60

PvPHIST-81   FIFGVIVVLLQ----NNN-SSSLEVSRNGRSLGEYYRDDHYDYDPRGRRG---------- 105
PcyPHIST-81  FIFGVIVVLLQ----NNTTSSSLEVSRNGRSLGEYYQDDLYEYEARGRRGG--------- 107
PkPHIST-105  FIFGVIVVLLQVSEGKNNTSSNLEVSRNGRSLGAYSHDDSYEYYGSGR----------- 108
PfPHIST-147  LCSTIENVEILKNDYNTLTESHNIINRRSRNLGANPESISLGYELSEKDEGNKNDLINSA 120

PvPHIST-81   -----------SRWQEEEDMYNPRMRGGN------------------------------ 123
PcyPHIST-81  -----------NRWQEEE-MYNPRMRGGN------------------------------ 124
PkPHIST-105  ------------KWEDEEDMYNLRMRGRDENDGDEKNRE-------------------- 135
PfPHIST-147  TDVSTELENLKERLFPELELYTNDQNSRNNTPNLRKGSLGFDSFKKLELGTLNQFDKDKM 180

PvPHIST-81   ------------------------------------------------------------
PcyPHIST-81  ------------------------------------------------------------
PkPHIST-105  ------------------------------------------------------------
PfPHIST-147  INLKDETNMNEFEGFLGRNSMASNVVTSELFDEPVDDSSSTTTSTGTKLQNVPSNDNNGE 240

PvPHIST-81   ------------------PDYYDER----------------------------------- 130
PcyPHIST-81  ------------------HDYYDDR----------------------------------- 131
PkPHIST-105  ---HRSDADDKRTFHDMKPAPAYDDEKDDKRTFHDMKPAPAYDDEKDDKSKIIHDNKPVP 192
PfPHIST-147  LLKDEPIDDYINNNSKVESEDNYYAQQNMQSQSKDNYASEQNVADQSTDNYPTQHDVPVQ 300

PvPHIST-81   ------------------------------------------------------------
PcyPHIST-81  ------------------------------------------------------------
PkPHIST-105  PKP---------DGDEKNKEHRSDADDKRTFHDMKPAPAYDDEKDDKSKIIHDNKPVPPK 243
PfPHIST-147  LRDNYASEQEYFDRGEQLNDVSADNNTSNKLKDEPVDNNTSNKLKDEPVDNNTSNKLKDE 360

PvPHIST-81   ------SGGYRSGPDDRNTYHNVHPIPAYDDQR--------------------DDKDNV 163
PcyPHIST-81  ------SMGYRSGPDDRNTFHRVEPLPAYDDER--------------------DDKGNV 164
PkPHIST-105  PDGDEKNKEHRSDADDKRTFHDMKPAPAYDDEK--------------------DDKSKI 282
PfPHIST-147  PVDDNTSNKLKDEPVDNNTINKLKDEPVDDNTSNILKDEPVDDHAGKHLKDEPVDDHAGK 420

PvPHIST-81   IRPDQPAPVKPDGDDTGKGDDSSVTPSPENPD-DPNNPPSTTETPGNSDGE--------- 213
PcyPHIST-81  INPAKPAPAKPDDDDQSKVDDHKVAPNPEAPDTDPSDPP----KPDNPDSE--------- 211
PkPHIST-105  IHDNKPVPPKPDGDDPGKGDGHNVVPSPDVPDDDPTNPPATTAAPRTPGGDKH------- 335
PfPHIST-147  HMKDEPVDIDRTNIKKGLNEQHVNPWTTTLADLKNINNSMKIEKNNKSNEQVKNTSVSKS 480

PvPHIST-81   -------------------------------------HKDDEGNVIRPAGKHVVKPDDE 235
PcyPHIST-81  -------------------------------------HKDDKNDEIHIDDNHEIKPDDD 233
PkPHIST-105  ------------------------------------DDRHDDKDNVIKIIKTHKIEPDDE 359
PfPHIST-147  CDIIKPSKFNKKNLFEQRLQSVEGKNFFEGRSQNLEGRSNFDERSQIVEQRRNFDDRDQN 540

PvPHIST-81   KDDKS--DNDPIKPNVPS-------EHHHDGSHDGSDDGSHDGSHPDQPGHHDGLKGTMS 286
PcyPHIST-81  KDDKGKIISDPTQPIEP-----------------SDDPHHDGSQPDKPGPDGGLKGVVG 275
PkPHIST-105  HDDDRRYDRHDDRHYDDRRFDRHDDRHDDDRRFDRHDDDRRYDKHDDRHYDDRRFDRHDD 419
PfPHIST-147  IMDRKNFDERNQQVNDRRNFDERNQQVNDRRNFDDRDQNVMDRRNFDERNQQVNDRRNFD 600

PvPHIST-81   -QGPYGPDPR--GDRSGDEREFR--HNQHG----------------------------- 311
PcyPHIST-81  GEGPYGPDGR--GDRSGDEREFRNYDNQRG----------------------------- 303
PkPHIST-105  RRFDRHDDRR--YDRHDDDRHYDKHDDRHDDKDNVIKIIKTHKIEPDDEHDDDRRYDRHD 477
PfPHIST-147  ERNQQVNDRRNFDDRDQNVMDRRNFDERNQQVNDRRNFDERNQQVNDRRNFDDRDQNVMD 660

PvPHIST-81   -RVFDDRRGRGGFDDYNGPDGYGSDYGRRYGDDGREIHYSRSEKSFDDEYHG---RGG-- 365
PcyPHIST-81  -RHMDDRRGRGGYDDYNGPDGYGSDYGRRYGDDSREIHYSNTERMYDEEYHG---RGGRG 359
PkPHIST-105  DRHYDDRRFDRHDDRHDDDRHYDDRRFDRHDDGSHPNHRGRRVGVTGAMSQGPDPRGDRR 537
PfPHIST-147  RRNFDERNQQVNDRRNFDERNQQVNDRRNFDERNQHVNDRRNFDERNQNVNDRRNFDERN 720

PvPHIST-81   -DDRSFH-FSKTNRVIDENMPYP------------------------------PNGPFR 392
PcyPHIST-81  GDDRSIH-YAENHRMIDEDMPYP------------------------------PNGRFR 387
PkPHIST-105  GDEREFRNYDERGRNLDDRRGRNRYDD----------------FNGLDEFGGYYNRRFR 580
PfPHIST-147  QNVNDRRNFDERNQQVNDRRNFDERYQNVNERRNFDERNQQVNDRRNFDERNQHVNERYQ 780
```

```
PvPHIST-81      GGDNRSIRSEQIAAMNYEEQFHQGPRGG--------RMGSANPFNVPPRGGRDDDHTFHS 444
PcyPHIST-81     --DDRSANYDRYNSMRFEDDYNQGPRGG--------RMGSANPFNVPPRGGRDDDHMIHT 437
PkPHIST-105     RGDDSSAHYDRFNSMRYEGDFSQGSRGG--------RMRNENPFSVP-RDRRDD------ 625
PfPHIST-147     NVNDRRNFDERNQQVNDRRNFDERNQHVNERRNFDERNQHVNERYQNVNDRRNFDERNQH 840

PvPHIST-81      ATPHRSVDDVND---------PRNRGRRDAGDDPRSFN-----------------PARGN 478
PcyPHIST-81     AHANKMIDDMND--------PDGRNRPMPGGPQGNLGG-------------RDQPGRGG 475
PkPHIST-105     VNGPRKSHDSRP---------PRG----NPEDDPRQFP---------------------- 650
PfPHIST-147     VNERRNFDQRAPNVEERRYMDPRNPNIPYVRFPHHQWGQGMMYGRPYYPWVPFMGDGRGY 900

PvPHIST-81      --QRHDDQFFDEGRFG--PPGGPGKFPDGRRP---PVPGPHGGPGGRGQPGRGGR----- 526
PcyPHIST-81     RLDMEEDAYGSRGDFMRRPPQDGRKFHHSRNPHAKPEDDPRQIHSSRVNPMKPDD----- 530
PkPHIST-105     ---------NSRG-----PMEREDEFEDGRNR--RMRAGPQGDEGRRDERGRDRR----- 689
PfPHIST-147     NFYNPHQHMVYGRPYYWVPPPPALEYTKGFNPMEQRREEDRGHMGGRGSRYPEEERYNYN 960

PvPHIST-81      ----------LEMEDDAFGSRG---DFMRR----SHD--GRGRHG-PHGDN--------- 557
PcyPHIST-81     ----------HFFDDDRFGPPGGPGKFQRG----RPGPVGRDMPGRPQGGN--------- 567
PkPHIST-105     ----------LDMEEDRFNSRE---DFMRRSADGRRGQVGRDVHGRPEGDD--------- 727
PfPHIST-147     NKRSNSIPEGRNYEENAYERGGGNNKWDFRNMYDRLRDEDENDYDQPPSTSSSNRGRGNE 1020

PvPHIST-81      -----------------------------------------------EQLPFGCTRAELQE 571
PcyPHIST-81     -----------------------------------------------EQLPFGCTRAELQE 581
PkPHIST-105     -----------------------------------------------ERLPFGCTQAELEE 741
PfPHIST-147     RYSQSRDRREERNNYNSDYYTRGNERTYNNSNVTSSSNRELIPYKKEILPFGVSNSELED 1080
```

```
                     Helix 1                              Helix 2
PvPHIST-81      QMTEEELNSKIKNLRPNATVKEMFVLFNQILSFERKKFVKMQEYIMQYSQYLQKTLLLPT 631
PcyPHIST-81     QMTEEELNSKIKNLRPNATVKEMFVLFNQILSFERKKFVKMQEYIMQYSQYLQKTLLLPT 641
PkPHIST-105     QMTEEELNRRIQNLRPNATVKEMFVLFNQILSFERKKFVKTQEYIMRYSQYLQKTLLLPT 801
PfPHIST-147     KLTEEELNERIRRLDYTVSVKDMFILWNHILAHERKKYTKMQEYLMYYSQYLEKTYLVPT 1140
```

```
                                   Helix 3
PvPHIST-81      PIRMKYWWRAHYNMTDELIKKERGDFQDFYAFVSKGQCQRWDFLYFANAKRKSWDELRDL 691
PcyPHIST-81     PIRMKYWWRAHYNMTDELIKKERGDFQDFYAFVSKGQCQRWDFLYFANAKRKSWDELRDL 701
PkPHIST-105     PIRMKYWWRAHYNMTEELIKKERGDFQDFYAFVNRGPCQKWDFLYFVNAKRKSWEELSDL 861
PfPHIST-147     AFRKKYWWRVHYMLTEEVVKRERTDNLDFHQFLRKGSCEKREFLYFINSKRKGWADLTET 1200
```

```
                     Helix 4
PvPHIST-81      MKSIWMEILTYKMKKHSKL 710
PcyPHIST-81     MKSIWMEILTYKMKKHSKL 720
PkPHIST-105     MKSVWMEILTYKMKKHSKL 880
PfPHIST-147     MKNIWMERLTYKMRKYSGA 1219
```

Figure 3.3a: Amino acid sequence alignment of PcyPHIST-81 with primate *Plasmodium* homologs. The alignments were performed using ClustalW. Residue numbering for each sequence is shown on the right. Positions of identity are highlighted in pink, the putative PEXEL motif is boxed and highlighted in blue, while conserved tryptophans, a distinguishing feature of PHIST proteins, are highlighted in red. The regions predicted to form alpha helices using the GOR4 algorithm (37) are shaded in grey.

BLAST analyses were also performed using the PvPHIST-81 amino acid

sequence to reveal novel orthologs in the NCBI database from *P. yoelii* (Py01786) and *P.*

berghei (PB108348.00.0 and PB000848.03.0) (Figure 3.3b). Py01786 has been annotated

as dentin phosphoryn, while the *P. berghei* fragments PB108348.00.0/PB000848.03.0 are

annotated as hypothetical proteins in PlasmoDB. These two sequences also share the four

conserved tryptophan residues and consecutive alpha-helical domains discovered in the

C-termini of PvPHIST-81, PcyPHIST-81, PfPHIST-147 and PkPHIST-105 (Figure 3.3b).

```
PvPHIST-81   MSPCNIPIIVLPATSSPADIKKAIDANNQTVAKSDGRRNKCKSRKSFADF 50
PcyPHIST-81  MSPCNIPIIVLPATSSPADIKRAIDANNQAVAKSDGGRNKCKSKKSFADF 50
PkPHIST-105  MNPCNIPIIVLPGTSSPADIRRAIDTNNQTVATSDGKKSKCKSRKSFADF 50
PyPHIST-110  MDSGKLPIIILPSSGSSVSLIKKNVND----RKLKKGNRRAYRKNSFERF 46
PbPHIST-122  MDSGKLPIIILPSSGPSVSLINKNVND----RKLKKGNRRGYRKNSFERF 46
PfPHIST-147  MIFVKSKILYFLKWPSVAIEENFSGSFKCLFKNKRNKYVRNNLSRFILSV 50

PvPHIST-81   NLASLPYKVLFIFGVIVVLLQ----NNN-SSSLEVSRNGRSLGEYYRDDH 95
PcyPHIST-81  NLASLPYKVLFIFGVIVVLLQ----NNTTSSSLEVSRNGRSLGEYYQDDL 96
PkPHIST-105  NLASLPYKVLFIFGVIVVLLQVSEGKNNTSSNLEVSRNGRSLGAYSHDDS 100
PyPHIST-110  NLASTFYKIILAFGIIIVVFLDNDSVHSNSEFSRIAKEGRNLSEAVAQDE 96
PbPHIST-122  NLASAFYKVILAFGIIIVVFLDNDSIYSNSGFSRITKEGRNLSEAIVPDE 96
PfPHIST-147  SLILFFFHFVLCSTIENVEILKNDYNTLTESHNIINRSRNLGANPESIS 100

PvPHIST-81   YDYDPRGRRG---------------------------------------- 105
PcyPHIST-81  YEYEARGRRGG--------------------------------------- 107
PkPHIST-105  YEYYGSGR------------------------------------------ 108
PyPHIST-110  ATNVGFTGTEVVDNDEKHETEQINEKHE---------------------- 124
PbPHIST-122  TPSTGFVETGVSENDENHG------------------------------- 115
PfPHIST-147  LGYELSEKDEGNKNDLINSATDVSTELENLKERLFPELELYTNDQNSRNN 150

PvPHIST-81   ----------------------------SRWQEEEDMYNPR-MRGGN- 123
PcyPHIST-81  ----------------------------NRWQEEE-MYNPR-MRGGN- 124
PkPHIST-105  ----------------------------KWEDEEDMYNLR-MRGRDE 126
PyPHIST-110  ----------TEQNDEKNDEQNDEQNVEQNVETNDSDSTHAAS-SEDSSK 163
PbPHIST-122  ----------------------------EKNDNDSTHATS-FEDSNK 133
PfPHIST-147  TPNLRKGSLGFDSFKKLELGTLNQFDKDKMINLKDETNMNEFEGFLGRNS 200

PvPHIST-81   -------------------------PDYYDER----------------- 130
PcyPHIST-81  -------------------------HDYYDDR----------------- 131
PkPHIST-105  NDGDEKNREHRSDADDKRTFHDMKPAPAYDDEKDDKRTFHDMKPAPAYDD 176
PyPHIST-110  SESSHTDLDVTKDKENEQTNDSNENGAQYFQESSLSSNQAEYTHDAKLTE 213
PbPHIST-122  SESSHTYLDTTKEKENEENNDSKGNVAQNFQKSNISSNQVECTDDVKLTE 183
PfPHIST-147  MASNVVTSELFDEPVDDSSSTTTSTGTKLQNVPSNDNNGELLKDEPIDDY 250

PvPHIST-81   --------------------------------------------------
PcyPHIST-81  --------------------------------------------------
PkPHIST-105  EKDDKSKIIHDNKPVPPKPDGDEKNKEHRSDADDKRTFHDMKPAPAYDDE 226
PyPHIST-110  TTN------YEHGYEQ-GDEHYYVESQYQQESYGQQISNNPDIQYGISVI 256
PbPHIST-122  TDN------YDKEYEQNGDEHYYAEKKYKQESYDQNILHNPDIQHALSMI 227
PfPHIST-147  INNNSKVESEDNYYAQQNMQSQSKDNYASEQNVADQSTDNYPTQHDVPVQ 300

PvPHIST-81   ----------------------SGGYRSG-PDDRNTYHNVHPIP----- 151
PcyPHIST-81  ----------------------SMGYRSG-PDDRNTFHRVEPLP----- 152
PkPHIST-105  KDDKSKIIHDNKPVPPKPDGDEKNKEHRSD-ADDKRTFHDMKPAP----- 270
PyPHIST-147  PSDSVIRSFLHTSLTISPEDDELDNPYREKTQDDVNTIHFTEWVG----- 301
PbPHIST-110  PSDSIIKSFLHTSLTISPEDDEIDNPYREKTQHDIDSMHFTEPLKPEQKY 277
PfPHIST-147  LRDNYASEQEYFDRGEQLNDVSADNNTSNKLKDEPVDNNTSNKLKDEPVD 350

PvPHIST-81   --------------------------------------------------
PcyPHIST-81  --------------------------------------------------
PkPHIST-105  --------------------------------------------------
PyPHIST-110  YMKHAIQEVDQGHE------------------------------------ 315
PbPHIST-122  YIEPSKPEQEENIEPLKPEQEENVDPLKPEQEENIKPLKPEQKENIKPLK 327
PfPHIST-147  NNTSNKLKDEPVDDNTSNKLKDEPVDNNTINKLKDEPVDDNTSNILKDEP 400
PvPHIST-81   --------------------------------------------------
PcyPHIST-81  --------------------------------------------------
PkPHIST-105  --------------------------------------------------
PyPHIST-110  --------------------------------------------------
PbPHIST-122  PEQEENIKPLKPEQEENVDPLKPEQEENVDPLKPEQEENVDPLKPEQEEN 377
PfPHIST-147  VDDHAGKHLKDEPVDDHAGKHMKDEPVDIDRTNIKKGLNEQHVNPWTTTL 450
```

```
PvPHIST-81   ------------------------------------------------AYDD 155
PcyPHIST-81  ------------------------------------------------AYDD 156
PkPHIST-105  ------------------------------------------------AYDD 274
PyPHIST-110  ---------------------------------------QPSTSEVPRP 325
PbPHIST-122  VDPLKPEQEENIKPLKPEQKENIKPLKPEQEENVDPLKPEQEENVDPLKP 427
PfPHIST-147  ADLKNINNSMKIEKNNKSNEQVKNTSVSKSCDIIKPSKFNKKNLFEQRLQ 500


PvPHIST-81   QRDDKDNVIRPDQPAPVKPDGDDTGKGDDSSV-----------TPSPENP 194
PcyPHIST-81  ERDDKGNVINPAKPAPAKPDDDDQSKVDDHKV-----------APNPEAP 195
PkPHIST-105  EKDDKSKIIHDNKPVPPKPDGDDPGKGDHNV-----------VPSPDVP 313
PyPHIST-110  EQPSTSEVPRPEQPSTSEIPRPEQPSTSEVPR-----------PEQPSTS 364
PbPHIST-122  EQEENIKPLKPEQKENIKPLKPEQEENVDPLK-----------PEQDENI 466
PfPHIST-147  SVEGKNFFEGRSQNLEGRSNFDERSQIVEQRRNFDDRDQNIMDRKNFDER 550


PvPHIST-81   D-DPNNPPSTTETPGNSDGE----HKDDEGNVIRPAGKHVVKPDDEKDDK 239
PcyPHIST-81  DTDPSDPP----KPDNPDSE----HKDDKNDEIHIDDNHEIKPDDDKDDK 237
PkPHIST-105  DDDPTNPPATTAAPRTPGGDKHDDRHDDKDNVIKIIKTHKIEPDDEHDDD 363
PyPHIST-110  EIPRPEQPSTSEIPRPEQPSTSEIPRPEQEAEYATINTKLPMGSENHQTN 414
PbPHIST-122  KPLKPEQEENIKPLKPEQEENIKPLKPEQEVEYAPLNNKFIMESKNNQK- 515
PfPHIST-147  NQQVNDRRNFDERNQQVNDRRNFDDRDQNVMDRRNFDERNQQVNDRRNFD 600


PvPHIST-81   S--DNDPIKPNVPS-------EHHHDGSHDGSDDGSHDGSHPDQPGHHDG 280
PcyPHIST-81  GKIISDPTQPIEP------------------SDDPHHDGSQPDKPGPDGG 269
PkPHIST-105  RRYDRHDDRHYDDRRFDRHDDRHDDDRRFDRHDDDRRYDKHDDRHYDDRR 413
PyPHIST-110  QQSGQGASNYNYNLDIIDENEINMLFGDPKSYGGAKRKTNFGSHHTEKES 464
PbPHIST-122  KQLGHFDSSYDYNLDIIDETEMNLLFEDSKSYSDIQKTTNNGSHYSEKEP 565
PfPHIST-147  ERNQQVNDRRNFDDRDQNVMDRRNFDERNQQVNDRRNFDERNQQVNDRRN 650


PvPHIST-81   LKGTMS-QGPYG-PDPRGDRSGDEREFR--HNQHG--------------- 311
PcyPHIST-81  LKGVVGGEGPYG-PDGRGDRSGDEREFRNYDNQRG--------------- 303
PkPHIST-105  FDRHDDRRFDRH-DDRRYDRHDDDRHYDKHDDRHDDKDNVIKIIKTHKIE 462
PyPHIST-110  MNMWNEPTDLFK-NKKSSPPKSSNSNVRGNEEMSMG---MWDGLRPSSIS 510
PbPHIST-122  LSMRDEPSGLYK-NKTSGNSKGSNPNVRGNEEVSMG---IWDGLKPSSMS 611
PfPHIST-147  FDDRDQNVMDRRNFDERNQQVNDRRNFDERNQQVND---RRNFDERNQHV 697


PvPHIST-81   ----------------RVFDDRRG----RGGFDDYNGPDGYGSDYGRRYG 341
PcyPHIST-81  ----------------RHMDDRRG----RGGYDDYNGPDGYGSDYGRRYG 333
PkPHIST-105  PDDEHDDDRRYDRHDDRHYDDRRF----DRHDDRHDDDRHYDDRRFDRHD 508
PyPHIST-110  RTHGMNSRQGMRPDLYYPSSSRNQ----GMRPDQYPPSSSRQSMIPEQYP 556
PbPHIST-122  RSQGMTINQ-------YPSTSGTQ----GITINQYPSTSGTQGMTRNQYP 650
PfPHIST-147  NDRRNFDERNQNVNDRRNFDERNQNVNDRRNFDERNQQVNDRRNFDERYQ 747


PvPHIST-81   DDGREIHYSRSEKSFDDEYHG------RGG---DDRSFH-FSKTNRVIDE 381
PcyPHIST-81  DDSREIHYSNTERMYDEEYHG------RGGRGGDDRSIH-YAENHRMIDE 376
PkPHIST-105  DGSHPNHRGRRVGVTGAMSQGPDP---RGDRRGDEREFRNYDERGRNLDD 555
PyPHIST-110  PSSSRQGMRHDLYSPSSSRQGMRPDQYPPPNSRQSMRPDQYPPSSSRQGM 606
PbPHIST-122  STSGTQGMTRNQYPSTSGTQGMTRNQYPSTSGTQGMTRNQYPSTSGTQGM 700
PfPHIST-147  NVNERRNFDERNQQVNDRRNFDERNQHVNERYQNVNDRRNFDERNQQVND 797


PvPHIST-81   NMPYP-----------------PNGPFRGGDNRSIRSEQIAAMN-YEEQ 412
PcyPHIST-81  DMPYP-----------------PNGRFR--DDRSANYDRYNSMR-FEDD 405
PkPHIST-105  RRGRNRY----DDFNGLDEFGGYYNRRFRRGDDSSAHYDRFNSMR-YEGD 600
PyPHIST-110  RPDQYPP----SSSRNQGMSRNKLIGSLDSLEDYKPEDSMWSGLKSSYQP 652
PbPHIST-122  TRNQYP-----STSGTQGMTRNQLIGEMYNIDERKIEEPMWLGLKSSYKP 745
PfPHIST-147  RRNFDERNQHVNERRNFDERNQHVNERYQNVNDRRNFDERNQHVNERRNF 847


PvPHIST-81   FHQGPRGGRMGSANP---------------------------------- 427
PcyPHIST-81  YNQGPRGGRMGSANP---------------------------------- 420
PkPHIST-105  FSQGSRGGRMRNENP---------------------------------- 615
PyPHIST-110  ISRKPSVSNIPGMDA---------------------------------- 667
PbPHIST-122  TPRKPYASNFSDMET---------------------------------- 760
PfPHIST-147  DQRAPNVEERRYMDPRNPNIPYVRFPHHQWGQGMMYGRPYYPWVPFMGDG 897


PvPHIST-81   ------------------FNVPPRGGRDDDHTFHSATPHRSVDDVNDP 457
PcyPHIST-81  ------------------FNVPPRGGRDDDHMIHTAHANKMIDDMNDP 450
PkPHIST-105  ------------------FSVP-RDRRDD---------------VNGP 629
PyPHIST-110  ------------------FQREGYSRNEHEMFDRFDSNSMSRDYSSQY 697
PbPHIST-122  ------------------FPRENYSRNKHEAFDRFDSHNMPRGHSTKY 790
PfPHIST-147  RGYNFYNPHQHMVYGRPYYWVPPPPALEYTKGFNPMEQRREEDRGHMGGR 947
```

```
PvPHIST-81   R--NRGRR---DAGDDPRSFNPARGN-QRHDDQFFDEGRFGPPGGPGKFP  501
PcyPHIST-81  RKFHHSRNPHAKPEDDPRQIHSSRVNPMKPDDHFFDDDRFGPPGGPGKFQ  500
PkPHIST-105  RKSHDSRPPRGNPEDDPRQFPNSRGPMERED----------------EFE  663
PyPHIST-110  PTPTSSKSRNPMNSYNYRDSSRSGSKLNSPTTSLRGNDPYRSDN------  741
PbPHIST-122  PTRTSSEFHNPMNRYNSRGTSNSGSILNS-MKSLRGNDSHRLDNGQNYMS  839
PfPHIST-147  GSRYPEEERYNYNNKRSNSIPEGRNYEENAYERGGGNNKWDFRNMYDRLR  997
```

```
PvPHIST-81   DGRRPPVPG-PHGGPGGRG-QPGRGGRL-EMEDDAFGSRGDFMRR-SHDG  547
PcyPHIST-81  DGRNRPMPGGPQGNLGGRD-QPGRGGRL-DMEEDAYGSRGDFMRRPPQDG  548
PkPHIST-105  DGRNRRMRAGPQGDEGRRD-ERGRDRRL-DMEEDRFNSREDFMRRSADGR  711
PyPHIST-110  ESASTLSIEEDKPCFRCVNYQSTKENKS-VSHQSKSNDDEDLAQQEQLII  790
PbPHIST-122  EPISTTSMGNHKLYFGSKN---TQETIS-TSFKSKSNDYKGLDQQ-----  880
PfPHIST-147  DEDENDYDQPPSTSSSNRGRGNERYSQSRDRRREERNNYNSDYYTRGNERT 1047
```

```
                                             Helix 1
PvPHIST-81   RGRHG---------PHGDNEQLPFG-CTRAELQEQMTEEELNSKIKNLRP  587
PcyPHIST-81  RGRPGPVGRDMPGRPQGGNEQLPFG-CTRAELQEQMTEEELNSKIKNLRP  597
PkPHIST-105  RGQVG---RDVHGRPEGDDERLPFG-CTQAELEEQMTEEELNRRIQNLRP  757
PyPHIST-110  QTINKEKPGTVAVRGSREVSDASFSPKKSTDVQIKEMDDHINSKISALDV  840
PbPHIST-122  NKASNYNSDTDAVIDSKEVSETNFSPRKSPDAQIKEMNDHINEKIDMLDE  930
PfPHIST-147  YNNSNVTSSSNRELIPYKKEILPFG-VSNSELEDKLTEEELNERIRRLDY 1096
```

```
                       Helix 2
PvPHIST-81   NATVKEMFVLFNQILSFERKKFVKMQEYIMQYSQYLQKTLLLPTPIRMKY  637
PcyPHIST-81  NATVKEMFVLFNQILSFERKKFVKMQEYIMQYSQYLQKTLLLPTPIRMKY  647
PkPHIST-105  NATVKEMFVLFNQILSFERKKFVKTQEYIMRYSQYLQKTLLLPTPIRMKY  807
PyPHIST-110  NTTPEALYDIWTDFIIAETKKFMLTQEYVLQYSIYLQKRNKLNPELRTKA  890
PbPHIST-122  NATPDVFRNIWTIFISAETKKLMMTQEYVLQYSIYLQKRNKLYPEMRTAA  980
PfPHIST-147  TVSVKDMFILWNHILAHERKKYTKMQEYLMYYSQYLEKTYLVPTAFRKKY 1146
```

```
                     Helix 3
PvPHIST-81   WWRAHYNMTDELIKKERGDFQDFYAFVSKGQCQRWDFLYFANAKRKSWDE  687
PcyPHIST-81  WWRAHYNMTDELIKKERGDFQDFYAFVSKGQCQRWDFLYFANAKRKSWDE  697
PkPHIST-105  WWRAHYNMTEELIKKERGDFQDFYAFVNRGPCQKWDFLYFVNAKRKSWEE  857
PyPHIST-110  WWKVFYSLVNKFVKHEKEDVSEFKELLKSSTNVSFNFVEFIKNKKEAWRE  940
PbPHIST-122  WWKVHYSMINKFIKREKEEVSELKELLKSSTNTYSNFVTFIRNKMESWKE 1030
PfPHIST-147  WWRVHYMLTEEVVKRERTDNLDFHQFLRKGSCEKREFLYFINSKRKGWAD 1196
```

```
              Helix 4
PvPHIST-81   LRDLMKSIWMEILTYKMKKHSKL---  710
PcyPHIST-81  LRDLMKSIWMEILTYKMKKHSKL---  720
PkPHIST-105  LSDLMKSVWMEILTYKMKKHSKL---  880
PyPHIST-147  LQSEIKDTWMAALTYKMNKYSYQSD-  965
PbPHIST-122  FQKEIKDTWLATLTYKINKYSSQDNY 1056
PfPHIST-147  LTETMKNIWMERLTYKMRKYSGA--- 1219
```

Figure 3.3b: Amino acid sequence alignment of PcyPHIST-81 with all identified *Plasmodium* **homologs**. The alignments were performed using ClustalW. Residue numbering for each sequence is shown on the right. Positions of identity are highlighted in pink, the putative PEXEL motif is boxed and highlighted in blue, while conserved tryptophans, a distinguishing feature of PHIST proteins, are highlighted in red. The regions predicted to form alpha helices using the GOR4 algorithm (37) are shaded in grey.

When comparing the PvPHIST-81 homologs from all of the species evaluated to date, each has the *Plasmodium* export element/Host targeting sequence (PEXEL/HTS; RxLxE/Q/D; (45, 46) in the N-terminal domain, a characteristic of many *Plasmodium* exported proteins (Figures 3.3a & 3.3b). A second feature of PHIST family members is

the PHIST domain, characterized by the presence of four consecutive alpha helices spanning approximately 150 amino acid residues (48). Using the GOR4 program, we identified four putative alpha helical domains in *P. vivax* PHIST-81, which are conserved among the three predicted orthologous PHIST proteins, and propose that these regions form their respective PHIST domains (Figures 3.3a & 3.3b).

An unexpected additional attribute of the *P. knowlesi* and *P. falciparum* PHIST-81 sequences is the presence of a repeated amino acid motifs and degenerated repeats positioned between the N-terminal PEXEL and the C-terminal PHIST domains (Fig. 3.4), with these sequences predominating as the majority of the protein's central domain. The specific repeats differ between homologs, and are of varying amino acid lengths within each protein sequence. In contrast, PvPHIST-81 and PcyPHIST-81, lack such repetitive amino acid sequences (Fig. 3.4), but have discernible repeats at the nuclei acid level. According to analyses using the Tandem Repeats Finder program, the *pvphist-81* and *pcyphist-81* genes each had 4 repeats at the nucleotide level.

P. vivax PHIST-81

MSPCNIPIIVLPATSSPADIKKAIDANNQTVAKSDGRRNKCKSRKSFA
DFNLASLPYKVLFIFGVIVVLLQNNNSSSLEVSRNGRSLGEYYRDDHY
DYDPRGRRGSRWQEEEDMYNPRMRGGNPDYYDERSGGYRSGPDDRNTY
HNVHPIPAYDDQRDDKDNVIRPDQPAPVKPDGDDTGKGDDSSVTPSPE
NPDDPNNPPSTTETPGNSDGEHKDDEGNVIRPAGKHVVKPDDEKDDKS
DNDPIKPNVPSEHHHDGSHDGSDDGSHDGSHPDQPGHHDGLKGTMSQG
PYGPDPRGDRSGDEREFRHNQHGRVFDDRRGRGGFDDYNGPDGYGSDY
GRRYGDDGREIHYSRSEKSFDDEYHGRGGDDRSFHFSKTNRVIDENMP
YPPNGPFRGGDNRSIRSEQIAAMNYEEQFHQGPRGGRMGSANPFNVPP
RGGRDDDHTFHSATPHRSVDDVNDPRNRGRRDAGDDPRSFNPARGNQR
HDDQFFDEGRFGPPGGPGKFPDGRRPPVPGPHGGPGGRGQPGRGGRLE
MEDDAFGSRGDFMRRSHDGRGRHGPHGDNEQLPFGCTRAELQEQMTEE
ELNSKIKNLRPNATVKEMFVLFNQILSFERKKFVKMQEYIMQYSQYLQ
KTLLLPTPIRMKYWWRAHYNMTDELIKKERGDFQDFYAFVSKGQCQRW
DFLYFANAKRKSWDELRDLMKSIWMEILTYKMKKHSKL

P. cynomolgi PHIST-81

MSPCNIPIIVLPATSSPADIKRAIDANNQAVAKSDGGRNKCKSKKSFA
DFNLASLPYKVLFIFGVIVVLLQNNTTSSSLEVSRNGRSLGEYYQDDL
YEYEARGRRGGNRWQEEEMYNPRMRGGNHDYYDDRSMGYRSGPDDRNT
FHRVEPLPAYDDERDDKGNVINPAKPAPAKPDDDDQSKVDDHKVAPNP
EAPDTDPSDPPKPDNPDSEHKDDKNDEIHIDDNHEIKPDDDKDDKGKI
ISDPTQPIEPSDDPHHDGSQPDKPGPDGGLKGVVGGEGPYGPDGRGDR
SGDEREFRNYDNQRGRHMDDRRGRGGYDDYNGPDGYGSDYGRRYGDDS
REIHYSNTERMYDEEYHGRGGRGGDDRSIHYAENHRMIDEDMPYPPNG
RFRDDRSANYDRYNSMRFEDDYNQGPRGGRMGSANPFNVPPRGGRDDD
HMIHTAHANKMIDDMNDPDGRNRPMPGGPQGNLGGRDQPGRGGRLDME
EDAYGSRGDFMRRPPQDGRKFHHSRNPHAKPEDDPRQIHSSRVNPMKP
DDHFFDDDRFGPPGGPGKFQRGRPGPVGRDMPGRPQGGNEQLPFGCTR
AELQEQMTEEELNSKIKNLRPNATVKEMFVLFNQILSFERKKFVKMQE
YIMQYSQYLQKTLLLLPTPIRMKYWWRAHYNMTDELIKKERGDFQDFYA
FVSKGQCQRWDFLYFANAKRKSWDELRDLMKSIWMEILTYKMKKHSKL

P. knowlesi PHIST-105

MNPCNIPIIVLPGTSSPADIRRAIDTNNQTVATSDGKKSKCKSRKSFA
DFNLASLPYKVLFIFGVIVVLLQVSEGKNNTSSNLEVSRNGRSLGAYS
HDDSYEYYGSGRKWEDEEDMYNLRMRGRDENDGDEKNREHRSDADDKR
TFHDMKPAPAYDDEKDDKRTFHDMKPAPAYDDEKDDKSKIIHDNKPVP
PKPDGDEKNKEHRSDADDKRTFHDMKPAPAYDDEKDDKSKIIHDNKPV
PPKPDGDEKNKEHRSDADDKRTFHDMKPAPAYDDEKDDKSKIIHDNKP
VPPKPDGDDPGKGDGHNVVPSPDVPDDDPTNPPATTAAPRTPGGDKHD
DRHDDKDNVIKIIKTHKIEPDDEHDDDRRYDRHDDRHYDDRRFDRHDD
RHDDDRRFDRHDDDRRYDKHDDRHYDDRRFDRHDDRRFDRHDDRRYDR
HDDDRHYDKHDDRHDDKDNVIKIIKTHKIEPDDEHDDDRRYDRHDDRH
YDDRRFDRHDDRHDDDRHYDDRRFDRHDDGSHPNHRGRRVGVTGAMSQ
GPDPRGDRRGDEREFRNYDERGRNLDDRRGRNRYDDFNGLDEFGGYYN
RRFRRGDDSSAHYDRFNSMRYEGDFSQGSRGGRMRNENPFSVPRDRRD
DVNGPRKSHDSRPPRGNPEDDPRQFPNSRGPMEREDEFEDGRNRRMRA
GPQGDEGRRDERGRDRRLDMEEDRFNSREDFMRRSADGRRGQVGRDVH
GRPEGDDERLPFGCTQAELEEQMTEEEELNRRIQNLRPNATVKEMFVLF
NQILSFERKKFVKTQEYIMRYSQYLQKTLLLPTPIRMKYWWRAHYNMT
EELIKKERGDFQDFYAFVNRGPCQKWDFLYFVNAKRKSWEELSDLMKS
VWMEILTYKMKKHSKL

P. falciparum PHIST-147 (Pf08_0137)

MIFVKSKILYFLKWPSVAIEENFSGSFKCLFKNKRNKYVRNNLSRFIL
SVSLILFFFHFVLCSTIENVEILKNDYNTLTESHNIINRRSRNLGANP
ESISLGYELSEKDEGNKNDLINSATDVSTELENLKERLFPELELYTND
QNSRNNTPNLRKGSLGFDSFKKLELGTLNQFDKDKMINLKDETNMNEF
EGFLGRNSMASNVVTSELFDEPVDDSSSTTTSTGTKLQNVPSNDNNGE
LLKDEPIDDYINNNSKVESEDNYYAQQNMQSQSKDNYASEQNVADQST
DNYPTQHDVPVQLRDNYASEQEYFDRGEQLNDVSADNNTSNKLKDEPV
DNNTSNKLKDEPVDNNTSNKLKDEPVDDNTSNKLKDEPVDNNTINKLK
DEPVDDNTSNILKDEPVDDHAGKHLKDEPVDDHAGKHMKDEPVDIDRT
NIKKGLNEQHVNPWTTTLADLKNINNSMKIEKNNKSNEQVKNTSVSKS
CDIIKPSKFNKKNLFEQRLQSVEGKNFFEGRSQNLEGRSNFDERSQIV
EQRRNFDDRDQNIMDRKNFDERNQQVNDRRNFDERNQQVNDRRNFDDR
DQNVMDRRNFDERNQQVNDRRNFDERNQQVNDRRNFDDRDQNVMDRRN
FDERNQQVNDRRNFDERNQQVNDRRNFDDRDQNVMDRRNFDERNQQVN
DRRNFDERNQQVNDRRNFDERNQHVNDRRNFDERNQNVNDRRNFDERN
QNVNDRRNFDERNQQVNDRRNFDERYQNVNERRNFDERNQQVNDRRNF
DERNQHVNERYQNVNDRRNFDERNQQVNDRRNFDERNQHVNERRNFDE
RNQHVNERYQNVNDRRNFDERNQHVNERRNFDQRAPNVEERRYMDPRN
PNIPYVRFPHHQWGQGMMYGRPYYPWVPFMGDGRGYNFYNPHQHMVYG
RPYYWVPPPPALEYTKGFNPMEQRREEDRGHMGGRGSRYPEEERYNYN
NKRSNSIPEGRNYEENAYERGGGNNKWDFRNMYDRLRDEDENDYDQPP
STSSSNRGRGNERYSQSRDRREERNNYNSDYYTRGNERTYNNSNVTSS
SNRELIPYKKEILPFGVSNSELEDKLTEEELNERIRRLDYTVSVKDMF
ILWNHILAHERKKYTKMQEYLMYYSQYLEKTYLVPTAFRKKYWWRVHY
MLTEEVVKRERTDNLDFHQFLRKGSCEKREFLYFINSKRKGWADLTET
MKNIWMERLTYKMRKYSGA

P. yoelii PHIST-110

MDSGKLPIIILPSSGSSVSLIKKNVNDRKLKKGNRRAYRKNSFERFNL
ASTFYKIILAFGIIIVVFLDNDSVHSNSEFSRIAKEGRNLSEAVAQDE
ATNVGFTGTEVVDNDEKHETEQINEKHETEQNDEKNDEQNDEQNVEQN
VETNDSDSTHAASSEDSSKSESSHTDLDVTKDKENEQTNDSNENGAQY
FQESSLSSNQAEYTHDAKLTETTNYEHGYEQGDEHYYVESQYQQESYG
QQISNNPDIQYGISVIPSDSVIRSFLHTSLTISPEDDELDNPYREKTQ
DDVNTIHFTEWVGYMKHAIQEVDQGHEQPSTSEVPRPEQPSTSEVPRP
EQPSTSEIPRPEQPSTSEVPRPEQPSTSEIPRPEQPSTSEIPRPEQPS
TSEIPRPEQEAEYATINTKLPMGSENHQTNQQSGQGASNYNYNLDIID
ENEINMLFGDPKSYGGAKRKTNFGSHHTEKESMNMWNEPTDLFKNKKS
SPPKSSNSNVRGNEEMSMGMWDGLRPSSISRTHGMNSRQGMRPDLYYP
SSSRNQGMRPDQYPPSSSRQSMIPEQYPPSSSRQGMRHDLYSPSSSRQ
GMRPDQYPPPNSRQSMRPDQYPPSSSRQGMRPDQYPPSSSRNQGMSRN
KLIGSLDSLEDYKPEDSMWSGLKSSYQPISRKPSVSNIPGMDAFQREG
YSRNEHEMFDRFDSNSMSRDYSSQYPTPTSSKSRNPMNSYNYRDSSRS
GSKLNSPTTSLRGNDPYRSDNESASTLSIEEDKPCFRCVNYQSTKENK
SVSHQSKSNDDEDLAQQEQLIIQTINKEKPGTVAVRGSREVSDASFSP
KKSTDVQIKEMDDHINSKISALDVNTTPEALYDIWTDFIIAETKKFML
TQEYVLQYSIYLQKRNKLNPELRTKAWWKVFYSLVNKFVKHEKEDVSE
FKELLKSSTNVSFNFVEFIKNKKEAWRELQSEIKDTWMAALTYKMNKY
SYQSD

P. berghei PHIST-122

MDSGKLPIIILPSSGPSVSLINKNVNDRKLKKGNRRGYRKNSFERFNL
ASAFYKVILAFGIIIVVFLDNDSIYSNSGFSRITKEGRNLSEAIVPDE
TPSTGFVETGVSENDENHGEKNDNDSTHATSFEDSNKSESSHTYLDTT
KEKENEENNDSKGNVAQNFQKSNISSNQVECTDDVKLTETDNYDKEYE
QNGDEHYYAEKKYKQESYDQNILHNPDIQHALSMIPSDSIIKSFLHTS
LTISPEDDEIDNPYREKTQHDIDSMHFTEPLKPEQKYYIEPSKPEQEE
NIEPLKPEQEENVDPLKPEQEENIKPLKPEQKENIKPLKPEQEENIKP
LKPEQEENVDPLKPEQEENVDPLKPEQEENVDPLKPEQEENVDPLKPE
QEENIKPLKPEQKENIKPLKPEQEENVDPLKPEQEENVDPLKPEQEEN
IKPLKPEQKENIKPLKPEQEENVDPLKPEQDENIKPLKPEQEENIKPL
KPEQEENIKPLKPEQEVEYAPLNNKFIMESKNNQKKQLGHFDSSYDYN
LDIIDETEMNLLFEDSKSYSDIQKTTNNGSHYSEKEPLSMRDEPSGLY
KNKTSGNSKGSNPNVRGNEEVSMGIWDGLKPSSMSRSQGMTINQYPST
SGTQGITINQYPSTSGTQGMTRNQYPSTSGTQGMTRNQYPSTSGTQGM
TRNQYPSTSGTQGMTRNQYPSTSGTQGMTRNQYPSTSGTQGMTRNQLI
GEMYNIDERKIEEPMWLGLKSSYKPTPRKPYASNFSDMETFPRENYSR
NKHEAFDRFDSHNMPRGHSTKYPTRTSSEFHNPMNRYNSRGTSNSGSI
LNSMKSLRGNDSHRLDNGQNYMSEPISTTSMGNHKLYFGSKNTQETIS
TSFKSKSNDYKGLDQQNKASNYNSDTDAVIDSKEVSETNFSPRKSPDA
QIKEMNDHINEKIDMLDENATPDVFRNIWTIFISAETKKLMMTQEYVL
QYSIYLQKRNKLYPEMRTAAWWKVHYSMINKFIKREKEEVSELKELLK
SSTNTYSNFVTFIRNKMESWKEFQKEIKDTWLATLTYKINKYSSQDNY

Figure 3.4: PHIST-81 orthologs contain varied repeat motifs. Amino acid sequences of the PHIST-81 homologs, highlighting amino acid repeat domains, which were determined by eye. The various repeat sequences are shown in different colors, while the PEXEL and PHIST domains are illustrated in pink.

PcyPHIST-81 is expressed throughout the erythrocytic stages in *P. cynomolgi*, predominantly in trophozoites.

After identifying the *pvphist-81* gene via the proteomic experiments outlined above, we produced a specific *P. vivax* recombinant protein representing 710 amino acids (rPvPHIST-81) as described in the experimental procedures, and raised a polyclonal rabbit antiserum to this protein (anti-rPvPHIST-81). This antiserum was then evaluated in IFA assays on thin blood smears containing mixed stages of *P. cynomolgi*-infected erythrocytes, and positive reactivity was evident at the ring, trophozoite and schizont stages (Fig. 3.5a). The anticipated fluorescent speckling pattern was evident at the ring stage, and it increased in abundance and intensity in the trophozoite and schizont stages, while keeping the time of exposure constant. In contrast, the control rabbit pre-immune serum was negative (data not shown). The PcyPHIST-81 protein was detected by Western immunoblot in all stages of development, although maximal expression was apparent in repeated experiments at the trophozoite stage (Fig. 3.5b).

(a)

Stage	Rings	Trophozoites	Mixed schizonts
IFA using anti-rPvPHIST81 antiserum			

(b)

Rings   Trophozoites   Schizonts
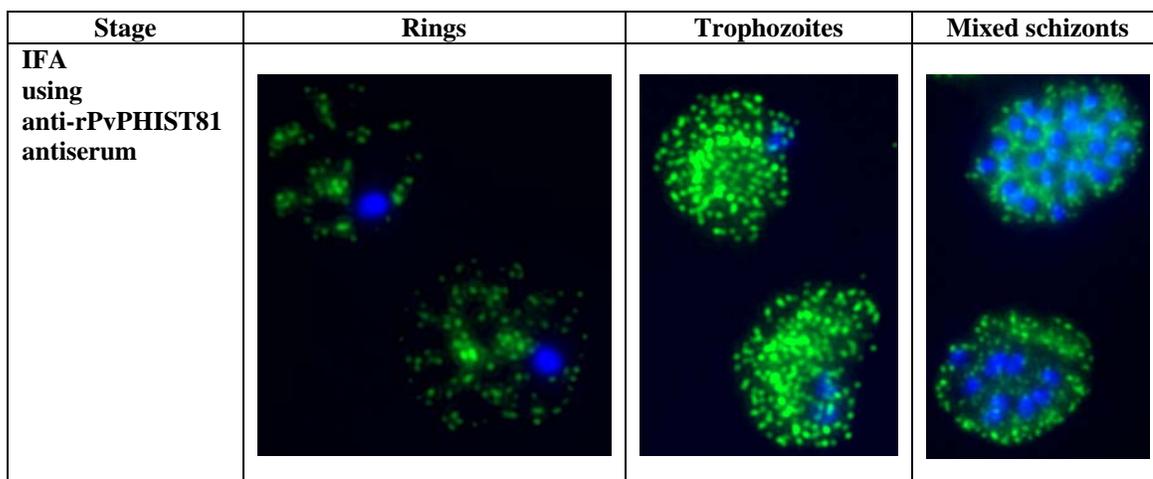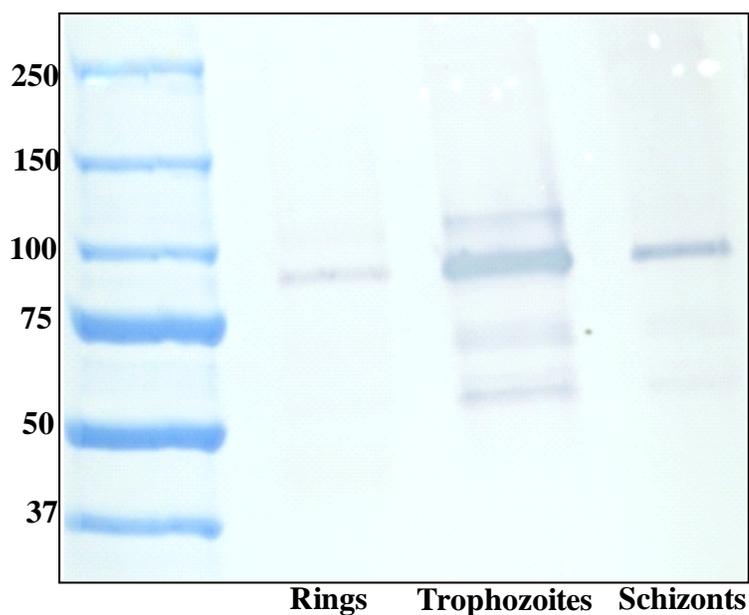
Figure 3.5: PcyPHIST-81 is expressed in the ring, trophozoite and schizont stages.

(a) IFAs. Fixed smears of ring, trophozoite and schizont-stage *P. cynomolgi*-infected RBCs were incubated with rabbit anti-rPvPHIST-81 serum and mounted in Prolong Gold Antifade Reagent containing DAPI (Invitrogen). Positive antibody reactivity was detected using Alexa Fluor 488-conjugated anti-rabbit IgG (Invitrogen).

(b) Western immunoblot. SDS detergent extracts from 1 x 10^8 *P. cynomolgi* ring, trophozoite and schizont stage parasite were diluted 1:100 in sample buffer and subjected to SDS-PAGE under non-reducing conditions on a 4-20% gradient gel. Rabbit anti-rPvPHIST-81 was used at a 1:1000 dilution.

Electron tomography reveals the intricate 3-dimensional signatures of the CVCs from *P. cynomolgi*-infected RBCs and localizes PcyPHIST-81 to the tubular extensions.

To gain an enhanced perspective of CVC structures and to precisely localize PcyPHIST-81, electron tomography and immunoelectron microscopy were performed using mAb 4C12.B4 (Figure 3.6 a-d) and anti-rPvPHIST-81 serum (Figure 3.6 e-h). These reagents were examined on sections of *P. cynomolgi* trophozoite-infected RBCs, previously treated with equinatoxin II. The antibodies localized to the CVCs (Figure 3.6), as previously observed by immuno-EM with the mAb and *P. vivax*-infected RBCs (15, 24, 26, 27). More specifically, we show here that by immuno-electron tomography the gold beads were observed to concentrate on the cytoplasmic side of the CVC tubular extensions (Figure 3.6 d & h). These and other images presented show, for the first time, the capability of high resolution electron tomography to illustrate the elaborate nature of individual CVCs in *P. cynomolgi.* Figures 3.6 d and 3.6 h depict the 3-dimensional reconstruction of *P. cynomolgi* serial electron microscopic sections. In these images, the interconnected nature of the caveolar opening of the CVCs and the multiple tubular extensions can also be better appreciated than by transmission EM alone.
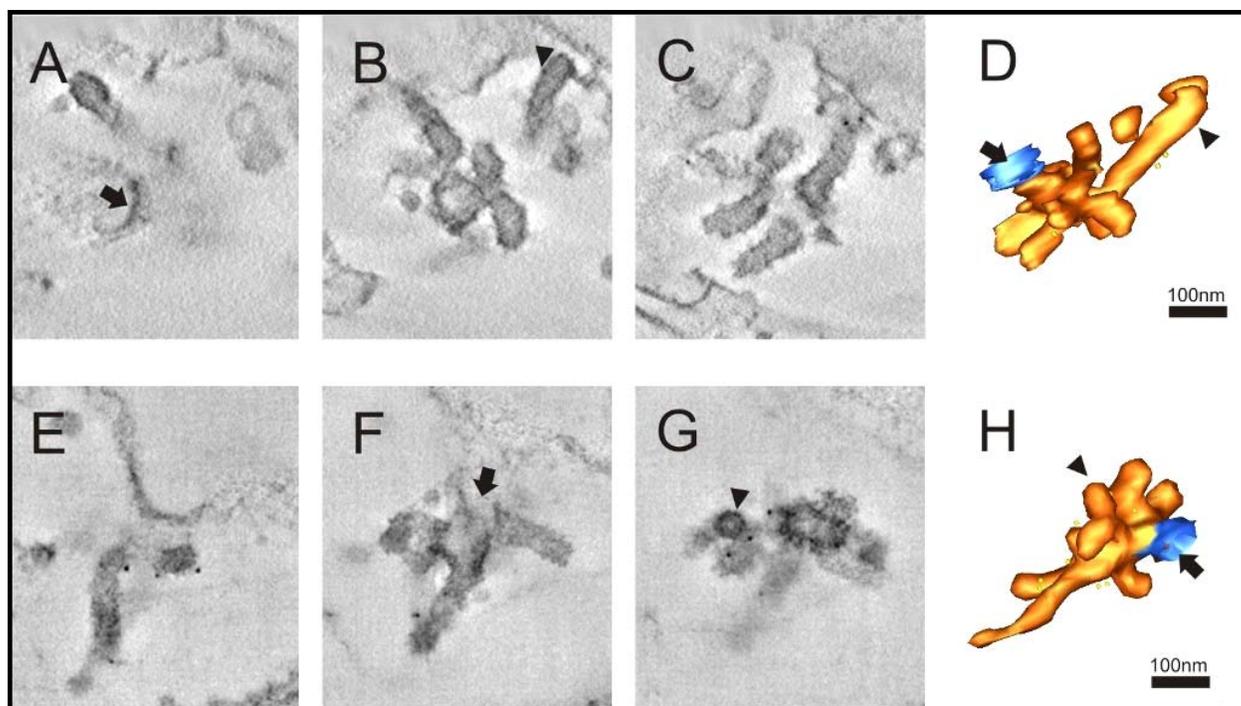
Figure 3.6: Electron tomography of caveola-vesicle complexes in *P. cynomolgi*-infected erythrocytes. EqtII-permeabilized *P. cynomolgi*-infected RBCs were labeled with two different antibodies (a – d, mAb 4C12.B4 and e – h, rabbit anti-rPvPHIST-81), fixed, en-bloc stained, embedded in LR White resin, sectioned (300 nm) and stained with lead citrate. The sections were observed with a Tecnai G2 TF30 at an accelerating voltage of 200 kV. A dual tilt series was collected from -69 degrees to 69 degrees every 1.5 degrees for the first axis and every 3 degrees for the second axis. The tomograms were generated using the IMOD package. (a-c) Selected virtual sections (20 nm) showing caveola-vesicle complexes (CVC). (d) The segmentation model was created with IMOD. The RBC membrane is rendered in blue and the CVC in orange. The opening on the surface of the RBC is indicated with thick arrows (a,d,f,h). A section through one of the tubular extensions is indicated with an arrowhead (g) and the same extension is indicated with an arrowhead in (h). The lumen of the CVC opens onto the RBC surface (b, d, thick arrows). The gold particles are concentrated on the tubular extensions. Bars: 100 nm.

3.5. Discussion

With the ultimate goal of malaria eradication in mind (14), all species of *Plasmodium* infecting humans must be investigated, and an increasing arsenal of tools to prevent and treat malaria infections must be generated. With this goal, it is important to gain knowledge on the composition and functions of the various membranous structures created as *Plasmodium* invades and remodels its host RBC.

This study is the first in about two decades to investigate the morphology and molecular make-up of caveolae and CVC structures in *Plasmodium vivax* and *P. cynomolgi*. Advanced microscopy and post-genomic tools have greatly expanded the potential of such investigations, which were hitherto hindered and developing slowly. Today, the recent publication of *P. vivax* and *P. knowlesi* genome databases and proteomic technologies has helped once again to jump-start this field. The future availability of additional or refined *P. vivax* and *P. knowlesi* genomic databases and, importantly, the inclusion of *P. cynomolgi* genome sequencing data will become a major plus for advancing this area of research at an expedited pace. The recent introduction of rapid and cost-effective "next generation" genome sequencing technologies (49) in fact make it likely that the first *P. cynomolgi* genome sequence will be initiated and rapidly completed in the near future.

Here we have used proteomics to identify the first gene encoding a protein that localizes to the CVCs of *P. vivax* and the simian malaria *P. cynomolgi,* denoted here as *phist-81*. *P. cynomolgi* infected RBCs became the main focus of these exploratory studies, instead of *P. vivax,* because they are readily accessible for pilot experimental manipulations as well as subsequent in-depth investigations. And, being very closely

related phylogenetically, with shared morphological and biological features, there is the reasonable expectation that the protein composite and biology of each of these species will be basically similar, with a limited degree of parasite-specific nuances. Similar direct studies on *P. vivax* are planned to explore this premise and define and compare the infected RBC membrane proteome for both species. Meanwhile, as anticipated, hypotheses and ongoing experimental work can proceed using the *P. cynomolgi* model to build upon the findings presented in this report.

3-D imaging by electron tomography has corroborated previous transmission EM reports suggesting that the caveolae of CVCs are open to the surface of *P. cynomolgi*-infected RBCs (22, 27, 50). In addition, the tomograms definitively show that the CVC tubules are of different sizes and that they can become quite elongated compared to the vesicles of the CVCs. While this information could be inferred from transmission EMs, the 3-D images showing different visual angles and perspectives provide a more complete view and appreciation of the elaborate nature of these structures, each one seemingly with its "unique morphological signature". As one scans the RBC plasmalemma, frequently marked with these intricate structures, haunting questions remain regarding how these structures are formed, and for what purpose(s). Why would the parasites expend the energy and resources to produce many of these structures in the course of their approximate 48 hour cycle in the blood? While some limited cytoadherent characteristics may be revealed in future investigations of *P. vivax* and *P. cynomolgi* infected RBCs, to date there is no evidence of any dominant cytoadherent or sequestration phenotypes at all akin to that generally observed for the trophozoite and schizont stages of *P. falciparum* and attributed to the adhesive domains of the variant antigens expressed at the surface of

electron dense knobby protrusions. In fact, although notorious for these characteristics and their predicted associations with severe disease, *P. falciparum* is quite unique in this regard, as most other primate species do not cytoadhere and sequester in the strict sense: all blood stage forms circulate for the other three human malaria species *P. vivax*, *P. malariae* and *P. ovale* and for all the simian malaria species, including among others *P. cynomolgi* and *P. knowlesi*. Necessary metabolic functions seem more likely for the CVCs, thus making them potential promising targets of intervention.

The *P. cynomolgi* model is now poised for advanced exploration of the many open questions relating to the biogenesis and function of the CVCs, along with comparative analyses of *P. vivax* (and *P. knowlesi*), to determine if and how caveolae (as in *P. knowlesi*) and CVCs (as in *P. cynomolgi* and *P. vivax*) could result from the invagination of the infected RBC membrane at numerous junctions, or strictly through the biogenesis of the composite proteins, lipids and cytoplasmic vesicles and their progressive or pre-formed docking at the surface of the infected erythrocyte. Transmission EM images of *P. vivax* trophozoite and schizont stage parasites showing clusters of cytoplasmic vesicles (Dluzewski and Galinski, unpublished data) and images generated with *P. cynomolgi* in this study showing individual vesicles near the surface of the infected RBC membrane would suggest the latter is occurring, though these observations require further investigation.

Through the current study, a previously reported 95 kDa protein – extractable only in SDS and thus predicted in the past to be associated with the cytoskeleton – is now confirmed to correspond to a PHIST protein with a calculated MW of 81 kDa and localized to the cytoplasmic face of the CVC tubules. While these data are clear, future

corroboration with advancing technologies and other antibody reagents will be welcomed, particularly knowing that certain fixation methods or defined antibody reagents may be preferable to best visualize specific antigens in immuno-EM studies. By transmission immuno-EM, mAbs 2H12.B4, 1H4.B6 and 4C12.C6 appeared to also recognize the caveolae of the *P. vivax* CVCs (22, 27, 50), but gold staining of the *P. cynomolgi* caveolae was not observed in this current study using immuno-ET. Whether the PHIST-81 protein is also present as a component of the caveolae is an open question that requires further clarification.

The relevance of the main defining characteristic of PHIST superfamily members, the conserved PHIST domain, is unknown. Particularly in the case of this protein, with degenerated nucleic acid repeated motifs dominating the central region between the PEXEL motif and the PHIST domain, it would seem that the PHIST domain may be the most critical part of the protein. Functional studies focused on the selective mutation or disruption of this domain are now in progress to see if the formation of the CVCs will be altered or blocked.

It also remains to be determined whether all PHIST family members are expressed and if they have different or similar functions associated with the infected RBC membrane. There are 71 predicted PHIST paralogs in *P. falciparum*, 39 in *P. vivax* and 27 in *P. knowlesi* (48). Here we have aligned the sequences of apparent PvPHIST-81 and PcyPHIST-81 homologs. BLAST searches using these full sequences, or restricted to the PHIST domain, also identified closely related PHIST members in *P. knowlesi, P. falciparum, P yoelii* and *P. berghei,* albeit with different MWs. These may be viewed as

orthologs; however, in the strict sense based on a confirmed ancestral relationship there is not enough evidence to make that call.

The PHIST homologs reported here from *P. falciparum*, *P. knowlesi*, *P. yoelii* and *P. berghei* all have strongly conserved C-terminal PHIST domains, but evident degenerated amino acid repeated motifs of varied lengths and sequences located between the N-terminal PEXEL and C-terminal PHIST domains. Because repeated motifs have not been noted previously as a characteristic of PHIST family members, it was surprising to see them predominate in the related gene/protein family members focused on in this report. These repeated motifs became evident first by eye, because the central domain of each PHIST protein when compared across the species seemed to have drastically unique sequences, yet with species-specific common amino acids occurring. While such amino acid repeated motifs are not apparent in the PvPHIST-81 and PcyPHIST-81 sequences, repeated sequences are evident at the nucleotide level in the corresponding genes from *P. vivax* and *P. cynomolgi*. This suggests that the repeated motifs per say are not functionally important for these proteins, but they may represent the traces of an evolutionary relationship of the associated genes, perhaps akin to the widely diverse repeated motifs maintained in the central regions of the otherwise conserved *P. cynomolgi* circumsporozoite genes from different strains (51), a phenomenon found to be common in various *Plasmodium* proteins (52-54).

Furthermore, because *P. falciparum*, *P. knowlesi*, *P. yoelii* and *P. berghei* do not produce CVCs in their host RBC membranes, their respective PHIST-81 homologs may not be functionally similar to PvPHIST-81 or PcyPHIST-81. Indeed, it was recently determined that the most closely related PHIST in *P. falciparum,* which we have called

PHIST-147 based on its deduced MW (gene ID:Pf08_0137), has been detected in association with the newly-described *P. falciparum* translocon of exported proteins (PTEX) complex, which has been predicted to transport proteins from the parasite across the parasitophorous vacuole membrane to the erythrocyte cytosol (55). Whether the PHIST-147 protein was simply caught in transit or if it is a functional component of the PTEX remains to be determined. In any event, the protein is expressed in *P. falciparum*, but presumably with a different specific functional role than the related PHIST-81 proteins in the CVCs of *P. cynomolgi* and *P. vivax*.

Further studies are obviously needed to determine what purpose the CVCs and their associated proteins perform in *P. vivax* and *P. ovale* parasite-infected RBCs, and if they are required for parasite survival. To this end, gene disruption studies are in progress for the PcyPHIST-81 family member. Overall, it will be important to determine how the CVC and caveolae architecture develops during the growth and multiplication of the intraerythrocytic parasite, and to determine if unique biological and biochemical processes can be identified and disrupted through specific drug intervention(s). With today's potential in medicinal chemistry to design novel drug compounds based on structure associated relationships (SAR), further explorations in this direction seem highly warranted.

REFERENCES

1. Price, R. N., N. M. Douglas, and N. M. Anstey. 2009. New developments in Plasmodium vivax malaria: severe disease and the rise of chloroquine resistance. *Curr Opin Infect Dis*.

2. Baird, J. K. 2008. Real-world therapies and the problem of vivax malaria. *N Engl J Med 359:2601*.

3. Galinski, M. R., and J. W. Barnwell. 2008. Plasmodium vivax: who cares? *Malar J 7 Suppl 1:S9*.

4. Galinski, M. R., A. R. Dluzewski, and J. W. Barnwell. 2005. Malaria Merozoites and Invasion of Erythrocytes. In *Molecular Approaches to Malaria*. I. W. Sherman, ed. ASM Press, New York, p. 113.

5. Mueller, I., M. R. Galinski, J. K. Baird, J. M. Carlton, D. K. Kochar, P. L. Alonso, and H. A. del Portillo. 2009. Key gaps in the knowledge of Plasmodium vivax, a neglected human malaria parasite. *Lancet Infect Dis 9:555*.

6. Haldar, K., and N. Mohandas. 2007. Erythrocyte remodeling by malaria parasites. *Curr Opin Hematol 14:203*.

7. Tilley, L., R. Sougrat, T. Lithgow, and E. Hanssen. 2008. The twists and turns of Maurer's cleft trafficking in P. falciparum-infected erythrocytes. *Traffic 9:187*.

8. Maier, A. G., B. M. Cooke, A. F. Cowman, and L. Tilley. 2009. Malaria parasite proteins that remodel the host erythrocyte. *Nat Rev Microbiol 7:341*.

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

10. Di Girolamo, F., C. Raggi, E. Bultrini, A. Lanfrancotti, F. Silvestrini, M. Sargiacomo, C. Birago, E. Pizzi, P. Alano, and M. Ponzi. 2005. Functional genomics, new tools in malaria research. *Ann Ist Super Sanita 41:469*.

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

12. Waters, A. P., D. G. Higgins, and T. F. McCutchan. 1993. Evolutionary relatedness of some primate models of Plasmodium. *Mol Biol Evol 10:914*.

13. Qari, S. H., Y. P. Shi, N. J. Pieniazek, W. E. Collins, and A. A. Lal. 1996. Phylogenetic relationship among the malaria parasites based on small subunit

rRNA gene sequences: monophyletic nature of the human malaria parasite, Plasmodium falciparum. *Mol Phylogenet Evol 6:157.*

14. Mendis, K., A. Rietveld, M. Warsame, A. Bosman, B. Greenwood, and W. H. Wernsdorfer. 2009. From malaria control to eradication: The WHO perspective. *Trop Med Int Health 14:802.*

15. Aikawa, M., L. H. Miller, and J. Rabbege. 1975. Caveola--vesicle complexes in the plasmalemma of erythrocytes infected by Plasmodium vivax and P cynomolgi. Unique structures related to Schuffner's dots. *Am J Pathol 79:285.*

16. Pasternak, N. D., and R. Dzikowski. 2009. PfEMP1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite Plasmodium falciparum. *Int J Biochem Cell Biol 41:1463.*

17. Kraemer, S. M., and J. D. Smith. 2006. A family affair: var genes, PfEMP1 binding, and malaria disease. *Curr Opin Microbiol 9:374.*

18. Kyes, S., P. Horrocks, and C. Newbold. 2001. Antigenic variation at the infected red cell surface in malaria. *Annu Rev Microbiol 55:673.*

19. Marti, M., J. Baum, M. Rug, L. Tilley, and A. F. Cowman. 2005. Signal-mediated export of proteins from the malaria parasite to the host erythrocyte. *J Cell Biol 171:587.*

20. Barnwell, J. W. 1990. Vesicle-mediated transport of membrane and proteins in malaria-infected erythrocytes. *Blood Cells 16:379.*

21. Tobie, J. E., and G. R. Coatney. 1961. Fluorescent antibody staining of human malaria parasites. *Exp Parasitol 11:128.*

22. Barnwell, J. W., P. Ingravallo, M. R. Galinski, Y. Matsumoto, and M. Aikawa. 1990. Plasmodium vivax: malarial proteins associated with the membrane-bound caveola-vesicle complexes and cytoplasmic cleft structures of infected erythrocytes. *Exp Parasitol 70:85.*

23. Barnwell, J. W. 1986. Antigens of Plasmodium vivax blood stage parasites identified by monoclonal antibodies. *Memorias do Instituto Oswaldo Cruz, Rio de Janeiro 81:59.*

24. Udagama, P. V., C. T. Atkinson, J. S. Peiris, P. H. David, K. N. Mendis, and M. Aikawa. 1988. Immunoelectron microscopy of Schuffner's dots in Plasmodium vivax-infected human erythrocytes. *Am J Pathol 131:48.*

25. Schuffner, W. A. P. 1899. Beitrag zur kenntnis der malaria. *Deutsch Arch. f. Klin. Med. 64:428.*

26. Wunderlich, F., H. Stubig, and E. Konigk. 1982. Intraerythrocytic development of Plasmodium knowlesi: structure, temperature- and Ca2+-response of the host and parasite membranes. *J Protozool 29:49.*

27. Matsumoto, Y., M. Aikawa, and J. W. Barnwell. 1988. Immunoelectron microscopic localization of vivax malaria antigens to the clefts and caveola-vesicle complexes of infected erythrocytes. *Am J Trop Med Hyg 39:317.*

28. Aikawa, M., Y. Uni, A. T. Andrutis, and R. J. Howard. 1986. Membrane-associated electron-dense material of the asexual stages of Plasmodium falciparum: evidence for movement from the intracellular parasite to the erythrocyte membrane. *Am J Trop Med Hyg 35:30.*

29. Matsumoto, Y., S. Matsuda, and Y. Yoshida. 1986. Ultrastructure of human erythrocytes infected with Plasmodium ovale. *Am J Trop Med Hyg 35:697.*

30. Aikawa, M., and L. H. Miller. 1983. Structural alteration of the erythrocyte membrane during malarial parasite invasion and intraerythrocytic development. *Ciba Found Symp 94:45*.

31. Herrera, S., B. L. Perlaza, A. Bonelo, and M. Arevalo-Herrera. 2002. Aotus monkeys: their great value for anti-malaria vaccines and drug testing. *Int J Parasitol 32:1625*.

32. Hanssen, E., R. Sougrat, S. Frankland, S. Deed, N. Klonis, J. Lippincott-Schwartz, and L. Tilley. 2008. Electron tomography of the Maurer's cleft organelles of Plasmodium falciparum-infected erythrocytes reveals novel structural features. *Mol Microbiol 67:703*.

33. Galinski, M. R., C. C. Medina, P. Ingravallo, and J. W. Barnwell. 1992. A reticulocyte-binding protein complex of Plasmodium vivax merozoites. *Cell 69:1213*.

34. Korir, C. C., and M. R. Galinski. 2006. Proteomic studies of Plasmodium knowlesi SICA variant antigens demonstrate their relationship with P. falciparum EMP1. *Infect Genet Evol 6:75*.

35. Peng, J., and S. P. Gygi. 2001. Proteomics: the move to mixtures. *J Mass Spectrom 36:1083*.

36. Peng, J., J. E. Elias, C. C. Thoreen, L. J. Licklider, and S. P. Gygi. 2003. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res 2:43*.

37. Combet, C., C. Blanchet, C. Geourjon, and G. Deleage. 2000. NPS@: network protein sequence analysis. *Trends Biochem Sci 25:147*.

38. Jackson, K. E., T. Spielmann, E. Hanssen, A. Adisa, F. Separovic, M. W. Dixon, K. R. Trenholme, P. L. Hawthorne, D. L. Gardiner, T. Gilberger, and L. Tilley. 2007. Selective permeabilization of the host cell membrane of Plasmodium falciparum-infected red blood cells with streptolysin O and equinatoxin II. *Biochem J 403:167*.

39. Anderluh, G., J. Pungercar, B. Strukelj, P. Macek, and F. Gubensek. 1996. Cloning, sequencing, and expression of equinatoxin II. *Biochem Biophys Res Commun 220:437*.

40. Kremer, J. R., D. N. Mastronarde, and J. R. McIntosh. 1996. Computer visualization of three-dimensional image data using IMOD. *J Struct Biol 116:71*.

41. Chang, H. H., A. M. Falick, P. M. Carlton, J. W. Sedat, J. L. DeRisi, and M. A. Marletta. 2008. N-terminal processing of proteins exported by malaria parasites. *Mol Biochem Parasitol 160:107*.

42. Albano, F. R., M. Foley, and L. Tilley. 1999. Export of parasite proteins to the erythrocyte cytoplasm: secretory machinery and traffic signals. *Novartis Found Symp 226:157*.

43. Lingelbach, K. R. 1993. Plasmodium falciparum: a molecular view of protein transport from the parasite into the host erythrocyte. *Exp Parasitol 76:318*.

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

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

46. Hiller, N. L., S. Bhattacharjee, C. van Ooij, K. Liolios, T. Harrison, C. Lopez-Estrano, and K. Haldar. 2004. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science 306:1934.*

47. Pain, A., U. Bohme, A. E. Berry, K. Mungall, R. D. Finn, A. P. Jackson, T. Mourier, J. Mistry, E. M. Pasini, M. A. Aslett, S. Balasubrammaniam, K. Borgwardt, K. Brooks, C. Carret, T. J. Carver, I. Cherevach, T. Chillingworth, T. G. Clark, M. R. Galinski, N. Hall, D. Harper, D. Harris, H. Hauser, A. Ivens, C. S. Janssen, T. Keane, N. Larke, S. Lapp, M. Marti, S. Moule, I. M. Meyer, D. Ormond, N. Peters, M. Sanders, S. Sanders, T. J. Sargeant, M. Simmonds, F. Smith, R. Squares, S. Thurston, A. R. Tivey, D. Walker, B. White, E. Zuiderwijk, C. Churcher, M. A. Quail, A. F. Cowman, C. M. Turner, M. A. Rajandream, C. H. Kocken, A. W. Thomas, C. I. Newbold, B. G. Barrell, and M. Berriman. 2008. The genome of the simian and human malaria parasite Plasmodium knowlesi. *Nature 455:799.*

48. Sargeant, T. J., M. Marti, E. Caler, J. M. Carlton, K. Simpson, T. P. Speed, and A. F. Cowman. 2006. Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol 7:R12.*

49. Schuster, S. C. 2008. Next-generation sequencing transforms today's biology. *Nat Methods 5:16.*

50. Bracho, C., I. Dunia, M. Romano, G. Raposo, M. De La Rosa, E. L. Benedetti, and H. A. Perez. 2006. Caveolins and flotillin-2 are present in the blood stages of Plasmodium vivax. *Parasitol Res 99:153.*

51. Galinski, M. R., D. E. Arnot, A. H. Cochrane, J. W. Barnwell, R. S. Nussenzweig, and V. Enea. 1987. The circumsporozoite gene of the Plasmodium cynomolgi complex. *Cell 48:311.*

52. Kemp, D. J., R. L. Coppel, and R. F. Anders. 1987. Repetitive proteins and genes of malaria. *Annu Rev Microbiol 41:181.*

53. Anders, R. F., G. V. Brown, R. L. Coppel, and D. J. Kemp. 1986. Induction of protective immunity to malaria. *Lepr Rev 57 Suppl 2:245.*

54. Ramasamy, R. 1991. Repeat regions of malaria parasite proteins: a review of structure and possible role in the biology of the parasite. *Indian J Malariol 28:73.*

55. de Koning-Ward, T. F., P. R. Gilson, J. A. Boddey, M. Rug, B. J. Smith, A. T. Papenfuss, P. R. Sanders, R. J. Lundie, A. G. Maier, A. F. Cowman, and B. S. Crabb. 2009. A newly discovered protein export machine in malaria parasites. *Nature 459:945.*

CHAPTER FOUR

MASS SPECTROMETRIC ANALYSIS OF *PLASMODIUM*-INFECTED RHESUS ERYTHROCYTE MEMBRANES IDENTIFIES NOVEL PARASITE AND HOST PROTEINS

Cindy C. Korir [a*], Sheila Akinyi [a*], Mary R. Galinski [a,b]

[a] Emory Vaccine Center, Yerkes National Primate Research Center, Emory University, 954 Gatewood Road, Atlanta, GA 30329, USA

[b] Department of Medicine, Division of Infectious Diseases, Emory University, 954 Gatewood Road, Atlanta, GA 30329, USA

[*] Authors contributed equally to this work

This chapter will be submitted for publication.

4.1. Abstract

Plasmodium parasites alter their host red blood cells (RBCs) antigenically and structurally. Changes to the RBC membrane have been extensively studied in *P. falciparum*, but much less so in other *Plasmodium* species of importance in human malaria, such as *P. vivax*. In this study, we utilize a global proteomic approach to identify proteins associated with the membranes of *P. cynomolgi* and *P. knowlesi*-infected erythrocytes. We report the identification of 109 *P. cynomolgi* orthologs and 129 *P. knowlesi* proteins on infected erythrocyte membrane ghosts. Of the *P. knowlesi* proteins identified, 12 are members of the *Plasmodium* helical interspersed subtelomeric (PHIST) protein family while three belong to the recently identified Pk-fam-c protein family. In addition, this is the first presentation of Rhesus macaque erythrocyte membrane proteomic data, where reverse-phase liquid chromatography followed by tandem mass spectrometric (LC-MS/MS) analyses yielded 41 host cell membrane hits from *P. cynomolgi*-infected RBC membranes and 81 hits from *P. knowlesi*-infected RBC membranes.

The identification of RBC membrane-localized proteins using a sensitive proteomics technique like LC-MS/MS provides an initial characterization of the more than 3000 hypothetical proteins in the *Plasmodium* genome. These novel proteins could be critical for parasite protein export, nutrient import, antigenic variation and membrane ultrastructural modifications leading to the creation of knobs, caveolae and caveola-vesicle complexes. Importantly, cross-species membrane analyses will facilitate the identification of proteins that may be functional across species, providing potential therapeutic targets.

4.2. Introduction

Malaria continues to be a significant global health burden, causing one to three million deaths annually, with an estimated 300-500 million cases of morbidity each year (1). There is also an increase in documented cases of malaria drug resistance, therefore necessitating research to identify novel therapeutic targets and develop effective vaccines. The *Plasmodium* genome consists of over 5000 genes, of which 65% remain uncharacterized (2). This creates a challenge to determining which proteins are the best candidates to be studied as potential drug and vaccine targets. Even so, newly available proteomics and genomics techniques make it easier to selectively analyze proteins of potential importance.

While *Plasmodium falciparum* is well known to cause the most severe cases of human malaria, *P. vivax* has been historically neglected and is lagging far behind in basic biological research. The study of *P. vivax* is hampered by the lack of a long-term *in vitro* culture system and thus heavily relies on its closely related simian malaria model species, *P. cynomolgi* and *P. knowlesi* (3-5). *P. cynomolgi* and *P. knowlesi* provide good *in vivo* and *in vitro* simian models for *P. vivax* infections because of the ease in obtaining sufficient quantities of parasite material from rhesus macaque infections. The recent completion of *P. vivax* (6) and *P. knowlesi* (7) genomes provides valuable data from which crucial genes can be determined.. The very close biological relationship between *P. vivax* and *P. cynomolgi*, with 85-95% sequence identity between them, allows for a reasonable estimate of the *P. vivax*/*P. cynomolgi*- infected RBC membrane proteome. *P. knowlesi* is also recognized as an important model for studying RBC invasion and

antigenic variation (reviewed in (8, 9)) and its ability to cause natural infections in humans (10, 11) emphasizes its direct relevance to human malaria research.

The intraerythrocytic *Plasmodium* parasite is enclosed within a parasitophorous vacuole (PV) from which it exports proteins that facilitate such functions as the import of nutrients, cytoadherence and the production of cytoplasmic and membrane ultrastructures (12-14). The functions of a large majority of these exported *Plasmodium* proteins remain unknown. Therefore, identifying and characterizing proteins with potential functional significance initially requires the recognition of certain characteristic motifs, which act as general predictors of a protein's export to the host RBC. For example, in most eukaryotes, a sequence of approximately fifteen hydrophobic amino acids located three to seventeen residues downstream of the N-terminus directs proteins to the secretory pathway (15-17). This N-terminal signal sequence (SS) is typically encoded on a short first exon and is responsible for co-translational translocation of proteins into the endoplasmic reticulum (ER) and export into the PV (18).

A second, *Plasmodium*-specific motif targets many proteins for transport across the PV membrane into the RBC cytoplasm and beyond. This motif is known as the *Plasmodium* export element (PEXEL; (19) or host targeting sequence (HTS; (20)). PEXEL/HTS consists of a short stretch of alternating charged and hydrophobic amino acids separated by uncharged amino acids (RxLxE/Q/D) and is located a few residues downstream of the classic signal sequence. While a protein is in the ER, the N-terminal portion of the PEXEL motif is cleaved between the arginine and leucine residues, N-acetylated and the protein exported across the PV directed by the remaining PEXEL residue (21).

The objectives of this study were to identify *P. cynomolgi-* and *P. knowlesi-*specific proteins associated with infected RBC membranes, and to perform cross-species comparisons among their membrane proteomes that could reveal common entities, which may serve as targets for therapeutic intervention. *P. cynomolgi* and *P. knowlesi-*infected RBC membrane ghosts were prepared and the extracted proteins analyzed by reverse-phase liquid chromatography-tandem mass spectrometry (LC-MS/MS). The data was searched against *Plasmodium* (PlasmoDB) and general (National Center for Biotechnology and Information; NCBI) databases and over 100 protein identifications each for *P. cynomolgi* and *P. knowlesi* were obtained. These included members of two recently-discovered gene families, *Plasmodium* helical interspersed subtelomeric (PHIST) and Pk-fam-c (6, 7, 22). Moreover, our data provided an initial description of the rhesus macaque RBC membrane proteome, with over 40 host protein hits obtained for each species studied.

In addition to facilitating the identification of novel therapeutic targets, this contribution to the building of *Plasmodium-*infected RBC membrane proteomes will provide insights into the structural alterations produced by the parasite on its host cell.

4.3. Materials and Methods

Preparation of *Plasmodium*-infected RBC Membrane Ghosts

Parasite-infected RBC membrane ghosts were prepared as described previously (23). Briefly, percoll-purified trophozoite-infected RBCs were washed and resuspended in PBS containing protease inhibitors. The cells were then centrifuged on a 5%-7.5%-

15% Ficoll separation gradient and the fractions containing the membranes collected and stored at -80ºC (Figure 4.1).

(a)

(b)

(d)

(c)
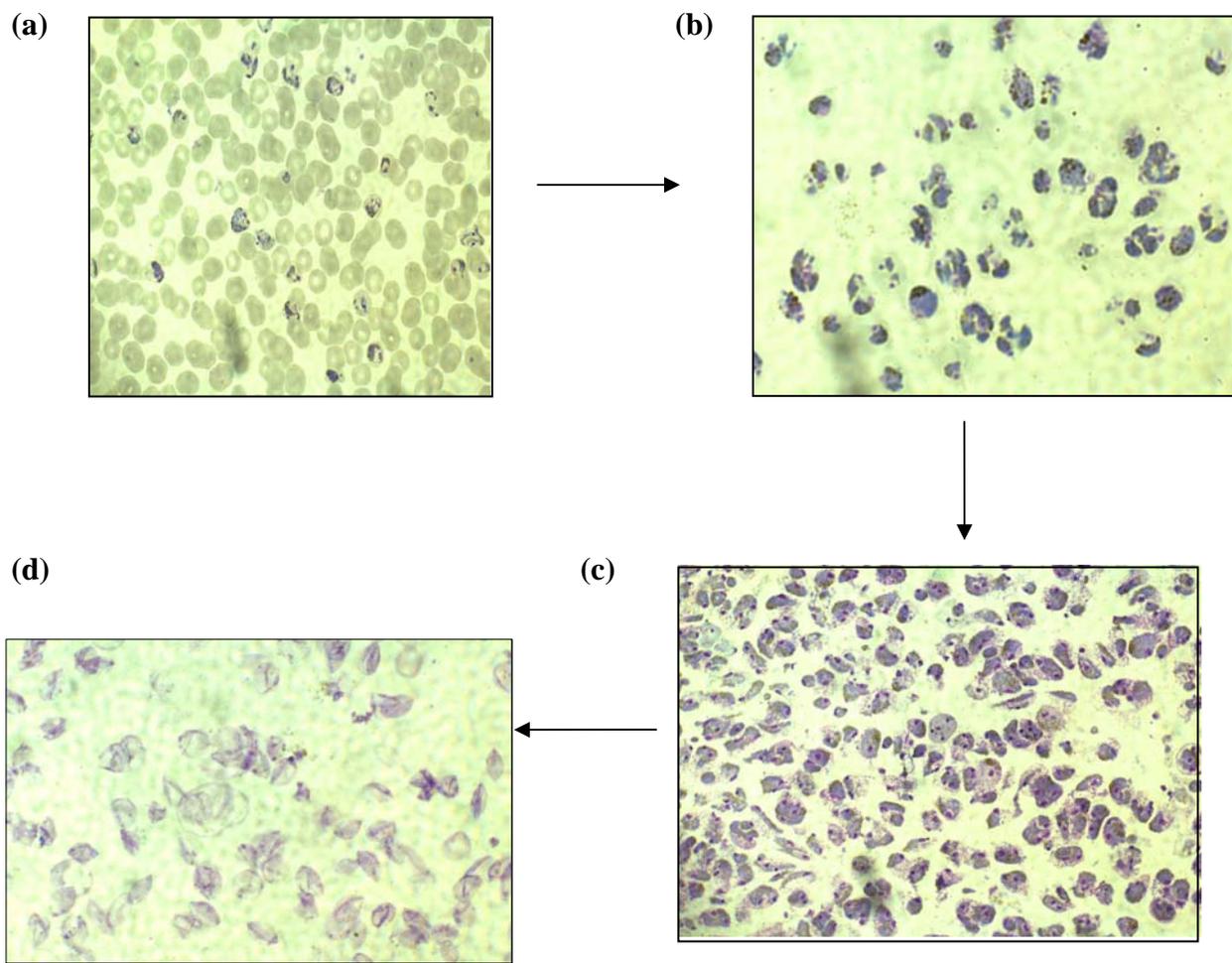
Figure 4.1: Preparation of *P. cynomolgi*-and *P. knowlesi*-infected erythrocyte membrane ghosts. *Plasmodium*-infected RBCs (a) were passed through a Percoll gradient to enrich for trophozoite-stage parasites (b). Parasites (c) were then separated from RBC membranes (d) using a Ficoll density gradient. (d) *Plasmodium*-infected RBC membrane ghosts.

Ethanol Solubilization and Precipitation of Membrane Ghosts

Parasitized RBC membrane ghost proteins were delipidized as described previously (24). Briefly, membrane pellets were diluted with four volumes of absolute ethanol and the solution brought to 50mM sodium acetate, using 2.5M sodium acetate, pH 5.0. 20μg of glycogen (Invitrogen) was then added per ml of original sample. The suspension was mixed at room temperature for 90 minutes then centrifuged for 10 minutes at room temperature. The supernatant was then discarded and the pellet prepared for mass spectrometric analysis.

LC-MS/MS Analysis of Immunoprecipitated Extracts and Infected Membrane Ghosts

After resolving the extract and membrane samples on 4-15% SDS-PAGE gradient gels, the gels were stained with colloidal Coomassie blue (Figure 4.2). Gel slices were then excised, destained, dried, and processed as described previously (25). Briefly, the gel pieces were digested with trypsin (Sigma; St. Louis MO) and the resulting peptides extracted with trifluoroacetic acid (Sigma; St. Louis, MO). The samples were then desalted and concentrated using ZipTip pipette tips (Millipore; Billerica, MA). Cleaned peptides were analyzed by reverse-phase liquid chromatography coupled with tandem mass spectrometry (26) using an LTQ-Orbitrap mass spectrometer (Thermo Finnigan, San Jose CA). A reverse database strategy using the SEQUEST algorithm was implemented to evaluate false discovery rate; the matched peptides were filtered according to matching scores to remove all false matches from the reverse database (27). Only proteins that were matched by at least two peptides were accepted to further

improve the confidence of identification. The peptides were then searched against the

NCBI database, with searches being limited to *Plasmodium* and *Macaca mulatta* results.
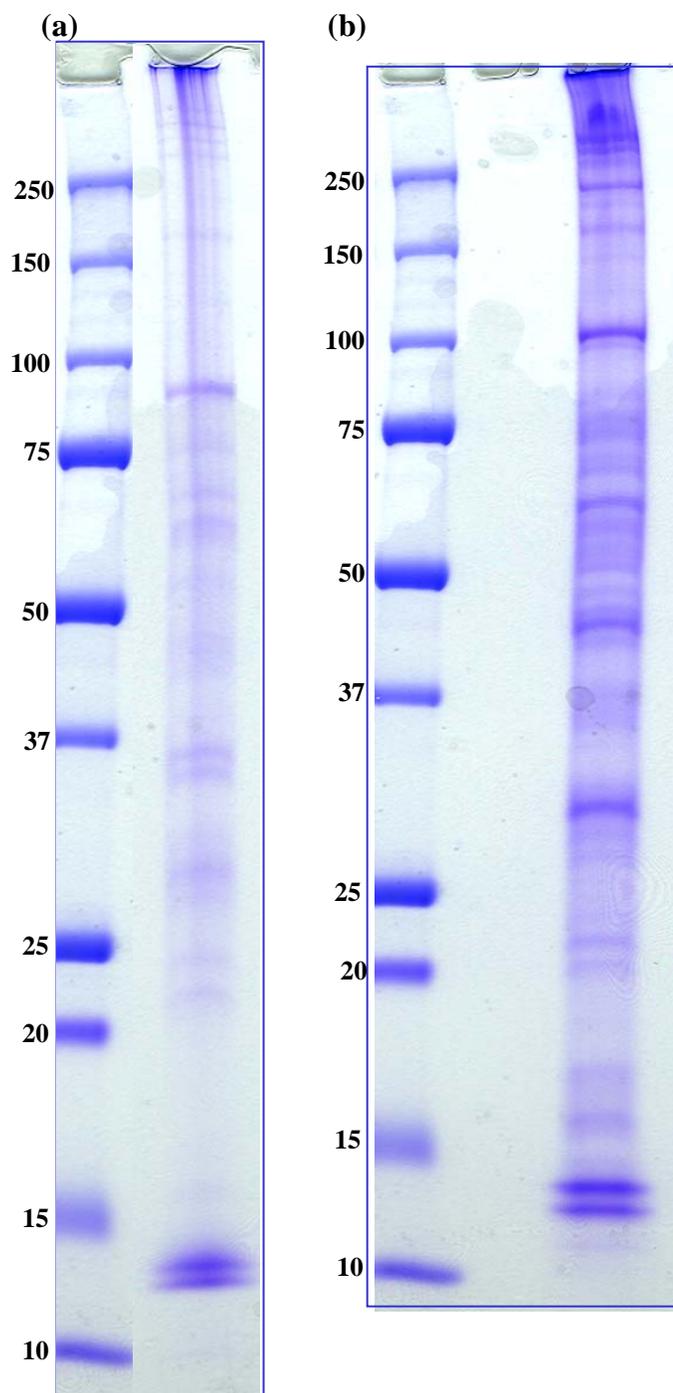
Figure 4.2: Gradient SDS PAGE gels of (a) *P. cynomolgi-* **and (b)** *P. knowlesi-***infected**

erythrocyte membrane ghosts.

4.4. Results

P. cynomolgi- and P. knowlesi -infected erythrocyte membrane ghosts were gel-purified and trypsin-digested for LC-MS/MS analysis. The resulting data was searched against the NCBI database for identification of P. vivax and P. knowlesi proteins. Proteomic analysis revealed 109 P. cynomolgi orthologs and 129 P. knowlesi proteins in the respective membrane ghosts. Among the proteins identified, four P. cynomolgi orthologs are members of the Plasmodium helical interspersed subtelomeric (PHIST) protein family (22) and six were hypothetical proteins (Table 4.1a; Supplementary table 4.1).

On the other hand, P. knowlesi membrane proteomic data identified 23 PEXEL-containing and 20 non-PEXEL motif-containing hypothetical proteins (Table 4.1b; Supplementary table 4.2). Of the 23 PEXEL-containing proteins, 12 either contain PHIST domains or have P. vivax orthologs that are members of the PHIST family, but have not been annotated as PHIST family members themselves (Table 4.1b; Supplementary table 4.2). Importantly, homologs of three of the identified P. knowlesi PHIST family membrane proteins were detected in P. cynomolgi-infected RBCs, indicating between-species conservation of function (Tables 4.1a & 4.1b; highlighted in blue). All of the identified P. cynomolgi PHIST homologs and P. knowlesi PHIST proteins have characteristics of exported proteins, including predicted signal sequences and the PEXEL motif. There are likely more PHIST proteins that are unique to P. cynomolgi, but those could not be detected due to the lack of a species-specific database.

Table 4.1a: PHIST and hypothetical proteins identified in *P. cynomolgi*-infected RBC membranes

Protein Description	MW (kD)	pI	Spectral count	Total peptides	Classic SS	Recessed SS	Transmembrane domain	Structural features	Orthologs
Phist protein (Pf-fam-b) Pvx_093680 (PvPHIST-81)	81	6.55	18	9	N	Y	None	PEXEL, PHIST domain, HØ region	Pk, Pf, Py, Pb, Pcy
Phist protein (Pf-fam-b) Pvx_088830	65	9.17	5	2	N	Y	None	PEXEL, PHIST domain, HØ region	Pk, Pf
Phist protein (Pf-fam-b) Pvx_112110	59	8.78	2	2	N	N	Three	PHIST domain	Pk
Tryptophan-rich antigen (Pv-fam-a) Pvx_101515	40	9.52	2	2	N	Y	One	PvTRAG family, PHIST domain	None
Hypothetical protein Pvx_083560	34	8.27	3	2	N	Y	None	12% lysine	Pk, Pf
Hypothetical protein Pvx_083270	42	7.85	2	2	N	Y	None		Pk, Pf, Pb, Pc, Py
Hypothetical protein Pvx_003555	119	3.82	1	2	N	N	None		Pk, Pf
Hypothetical protein Pvx_081830	57	4.64	5	4	N	Y	None	16% lysine	Pk, Pc
Hypothetical protein Pvx_123060	27	11.12	4	3	N	N	None	21% glycine	Pk, Pf, Pb, Pc, Py
Hypothetical protein Pvx_092070	48	4.48	4	2	Y	Classic	None	Pv-fam-d, PvTRAG family	Pk, Pf, Pb, Pc, Py

Table 4.1b: : PHIST and hypothetical proteins identified in *P. knowlesi*-infected RBC membranes

Protein Description	MW (kD)	pI	Spectral count	Total peptides	Classic SS	Recessed SS	Transmembrane domain	Structural features	Orthologs
Hypothetical protein PkH_090690	56	7.04	25	16	Y	N	None	PEXEL PHISTc 2 exons	Pv, Pf
Hypothetical protein PkH_011720	105	6.26	30	22	N	Y	None	PEXEL PHISTc 2 exons	Pvx_093680, Pf, Pk, Pb, Py
Hypothetical protein PkH_100030	97	5.1	23	18	N	N	None	PEXEL, 2 exons	Pvx_112110
Hypothetical Protein PkH_060050	157	4.11	17	14	Y	N	None	PEXEL PHISTc 2 exons	Pvx_088830
Hypothetical protein PkH_050060	72	8.18	7	6	N	Y	None	PEXEL, PHISTc, 2 exons	Pv
Hypothetical protein PkH_143340	58	6.75	7	4	N	Y	None	PEXEL, PHISTb, 2 exons	Pv, Pf
Hypothetical protein PkH_103240	40	8.19	6	5	Y	N	None	PEXEL, PHISTd, 2 exons	Pv
Hypothetical protein PkH_020250	54	6.73	5	3	N	N	None	PEXEL, PHISTc, 1 exon, fragment	None
Hypothetical protein PkH_051470	51	6.62	4	4	N	N	None	PEXEL, PHISTc, 1 exon, fragment	None
Hypothetical protein PkH_120030	58	4.95	4	3	N	Y	None	PEXEL, PHISTc, 2 exons	Pv, Pf
Hypothetical protein PkH_060020	66	8.61	3	3	Y	N	Three	PEXEL, PHIST, 2 exons	Pv, Pk
Hypothetical protein PkH_103230	47	9.27	2	2	N	Y	None	PEXEL, PHISTc, 2 exons	Pv, Pf
Hypothetical protein PKH_131790	59	4.39	19	14	N	Y	None	PEXEL, 2 exons	Pk
Hypothetical protein PKH_127040	92	10.03	14	14	N	N	None	16% lysine	None
Hypothetical protein PKH_031220	59	4.66	8	5	Y	N	None	16% glutamate 12% lysine	Pv, Pf
Tryptophan/ threonine- rich antigen, putative PKH_052760	67	6.58	8	8	N	N	None	11% serine	Pv, Pb, Pc, Py
Hypothetical protein PKH_110070	22	4.42	8	5	N	N	One	12% glutamate	Pv
Hypothetical protein PKH_114830	21	9.52	6	5	Y	N	One	PEXEL 15% asparagine	Pv
Hypothetical protein PKH_140050	33	8.96	6	5	N	Y	Two	PEXEL, Sh 1TM family,	Pv, Pf

Hypothetical protein PKH_031240	50	3.85	5	5	N	Y	None	PEXEL, Pk-fam-c, 2 exons	Pk only
Hypothetical protein PKH_125070	57	6.77	5	4	N	N	None	13% glutamate 12% lysine	Pv
Hypothetical protein PKH_072760	161	8.37	4	4	Y	N	None	11% lysine 10% leucine	Pk, Pf, Pb, Pc, Py
Hypothetical protein PKH_052790	41	4.76	4	4	N	Y	None	14% glutamate 12% lysine	Pv, Pb, Pc, Py
Hypothetical protein PKH_021540	56	4.47	4	4	N	Y	None	PEXEL, 2 exons	Pv, Pc
KIR-like protein PKH_020060	56	8.57	4	4	N	N	One	15% serine 12% glycine	Pv
Hypothetical protein PKH_073420	39	4.84	4	3	Y	N	None	PEXEL, Pk-fam-c, 2 exons	Pk only
Hypothetical protein PKH_103200	28	4.72	4	2	N	N	One	10% alanine	Pf (SBP1)
Hypothetical protein PKH_120640	42	7.86	3	3	N	Y	None		Pv, Pf, Pb, Pc, Py
Hypothetical protein PKH_120070	34	7.62	3	3	N	Y	None	PEXEL, 2 exons	Pv, Pf
Hypothetical protein PKH_131780	37	4.73	3	2	N	Y	None	PEXEL, 2 exons	Pv, Pb, Pc, Py
Early transcribed membrane protein PKH_052730	23	7.47	2	2	Y	N	One	13% threonine 12% serine 10% lysine	Pv, Pf
Hypothetical protein PKH_020090	76	4.97	2	2	N	Y	None	16% glutamate 11% lysine	Pv
Hypothetical protein PKH_134550	74	8.59	2	2	N	Y	Three	PEXEL, 1 exon	Pv, Pk
Hypothetical protein PKH_147000	31	5.34	2	2	N	N	None	15% glutamate 14% lysine 10% tryptophan	None
Hypothetical protein PKH_114880	38	4.39	2	2	N	N	None	17% glutamate	None
Hypothetical protein PKH_072300	24	4.73	2	2	Y	N	None	11% isoleucine 10% leucine	Pv, Pf, Pb, Pc, Py
Tryptophan/threonine-rich antigen, putative PKH_050090	52	4.83	2	2	N	Y	None	14% lysine 14% glutamate 10% asparagine	Pv, Pf, Pb, Pc, Py
Hypothetical protein PKH_010270	58	7.79	2	2	N	N	None	14% lysine 13% glutamate	Pv
Hypothetical protein (fragment) PKH_010030	24	9.73	2	2	N	N	None	PEXEL, Pk-fam-c, Fragment due to contig end,	Pk only

								1 exon	
Hypothetical protein PKH_092690	48	4.44	2	2	Y	N	None	12% asparagine 10% serine	Pv, Pf, Pb, Pc, Py
Hypothetical protein PKH_020070	34	10.34	2	2	N	Y	One	18% serine 11% asparagine	None
Hypothetical protein PKH_132260	33	3.96	2	2	N	N	None	18% gluatamate 13% aspartate 12% lysine 11% leucine	Pv, Pf, Py
Hypothetical protein PKH_020110	30	4.03	2	2	N	Y	One	14% glutamate	Pv, Pf

Table 4.1: Putative PHIST family members and hypothetical proteins identified by LC-MS/MS analysis of (a) *P. cynomolgi* and (b) *P. knowlesi*-infected RBC membrane ghosts. PHIST orthologs common to *P. cynomolgi* and *P. knowlesi* are highlighted in blue, while Pk-fam-c family members are highlighted in pink.

An alignment with six of the twelve identified *P. knowlesi* PHIST domain-containing proteins revealed conserved residues (47-81% sequence identity) and three conserved tryptophans, a characteristic feature of this family ((22); Figure 4.3a). In addition, we were able to identify and highlight the ~150 amino acid regions containing consecutive alpha helices within the C-termini of these six proteins, which are indicative of the PHIST domain ((22); Figure 4.3a). Whereas PkH_020250 and PkH_051470 are fragments with no starting methionine, they, in addition to PkH_090690, PkH_120030 and PkH_103230 are within-species paralogs, with only one known ortholog outside of *P. knowlesi*: the *P. vivax* Pvx_003555 protein that is also identified in our *P. cynomolgi*-infected RBC membrane proteome analysis (Table 4.1a; Supplementary table 4.1). When Pvx_003555 is aligned with the six *P. knowlesi* paralogs, its C-terminus is highly conserved with the others while its N-terminus is not, due to the presence of various tandem repeats in Pvx_003555 (Figure 4.3b).

Figure 4.3a

```
PkH_020250    ------------------------------------------------------------
PkH_051470    ------------------------------------------------------------
PkH_090690    --------------------------------MDEKKSTQRSSFFTKKTAMLFFVLAC 26
PkH_120030    --------------------------------MDEKKSTQRSSFFTKRTAMLFFVVAC 26
PkH_143340    MAHTNEIIINGASSRSMQSQCGNYQIVTARQERKMDERKSTQRSSFFTKRTAMLFFVVAC 60
PkH_103230    ---------------MQSQCGSCQIVGAKQERKMDEKKSTQRSSFFTKRTAMLFFVVAC 44


PkH_020250    ----HQDDARRQGVNAPLQVDNIVCRNLASKE--EATSSGQNDYEVGQTEVPSEGGRQAS 54
PkH_051470    ----HQDDARRQGVNASLQVHNKVGRNLASKE--EATPSGQNDYEVGQTEVPSEGGRQAS 54
PkH_090690    IFLKHQDDARRQGVNAPLQVDNIVGRNLASKE--EATSSGPYDYEVGQTEDPSEGGRQAS 84
PkH_120030    IFLKHQDDFSRQGVNAPLQVDNNVGRNLSSTESEDEAPSGPYDYEVGQTEDPSEGGRQAS 86
PkH_143340    IFLKHQDDARRQGVNASLQVHNKVGRNLSSKE--EATPSGPNDHRLGQTEDHSEGGRQAS 118
PkH_103230    IFLKHQDDTPRQGVNAPLQVDNIVCRNLGSTETEEEAPS-KNNSGFGGNKKVYLG--KGA 101


PkH_020250    PAVGENTWGMEYPNVRQDLQKTAEQETPLFGEKTNYVVDADLCVYAIFRTFYDNAFNDNK 114
PkH_051470    PAVGENTWGMEYPNVRQDLQKSAEQEASLFGERTNFVIEADFTPYADFVHTCMNAYYDDS 114
PkH_090690    PAAGEMTWGMEYPNVRQDVQKSAEQEAPLFGERTNYVAEGDLCVYAYFAHLCKNAFNDNK 144
PkH_120030    PAAGEMTLEMEYPNARQDVQKSAEQEAPLFGERTNYVAEGDLCPYAYFVHTCMNAFNDNK 146
PkH_143340    PAVGE----MEYPNVRQDVQKPAEQEAPLFGERTNYITDGDSGNYAIFESLCKNSFNDNK 174
PkH_103230    PKLNE-----------KKLQGKEDYVGNARMDDKTYLN--------TLQESYDHMDNDDK 142


PkH_020250    KKKKGKKEIEEEKKKEEALLKEKEKSQKKLNNWNQKVYRGPLAALEKYEDELEEERKRSL 174
PkH_051470    KKTKKQREMEE-KKKEEALLKEKEKSQKELSKGKKKVYSGPLAALEKYEDELEEERKRSL 173
PkH_090690    KKKKGKKEIEEEKKKEEALLKEKEKSQKELSKGKKKVYSGPLAAIEKYEDEQEEERKRSL 204
PkH_120030    KKKKGEKEIEEEKKKEEALLKEKEKSQNKLSIGNTKIYSGPLAALEKYEDELEEERKRSL 206
PkH_143340    KKKKGKKEIEEEKKKEEALLKEKEKSQNKLSIGNTKVYSGPLAALEKYEDELEEERKRSL 234
PkH_103230    KEQKRK----EEKKKE------KKKCQKKFS------------TWEKRKDKMEECK---- 176


PkH_020250    RYDKEERKRKMKGKGKGY-------SPIYTNGAHCDDIDAYMAKNMNGKNVEIEDWELRK 227
PkH_051470    RYDKEKRRKERD-------------------------ERVILNSMDKEKDIIADHRER- 206
PkH_090690    RYDKEERKRKMKEMEK---------------------ERSFLNKFHSPLYTNGAHSDDI 242
PkH_120030    RYDKEERKRKMKGKGKGQSLLDKWQSIVYEVGAYRDETEVNKRKCMNGEYKEIEDWEISK 266
PkH_143340    RYDKEERRKNKERE----------------------EYEKIEKWGMDKKVVVENDDN- 269
PkH_103230    -----EKK---------------------------------------------------- 179


PkH_020250    LKRAGYDVSDYGKHIPKDNTRSSGQDEEEHIYDEEGMNELLFGEFKKWEDSGAHTIPSSH 287
PkH_051470    ---------EYKEELEEERKRSLRRGKEELIDDEEGMHELLLGEFKKWENSGTHTIPSSH 257
PkH_090690    RMYRAED-RGDDEYEEKNNTRSLRYDEEEHIYDEEGMNDLLFGEFKKWEDSGTHTIPSSH 301
PkH_120030    LKKGGHDESDYSKYIRKDNTRSLRYDEEEHIDDEEEMNELLFGEFKKWEDSGAHTIPSSH 326
PkH_143340    ----------YLEQISEKRKRSLRYDEEEHIDDEEEMNELLFGEFKKWENYGTHTIPSSH 319
PkH_103230    ----------YKDKISKNRKKCLRHDEQKHIDEAGELDKLLFGELELLEKYGTHKVSTFH 229

                                      Helix 1                        Helix 2
PkH_020250    YEVLCFEDRLNDYEINKELNEMEEMPKKCELVSLYWQSFMNEKRKYINANKRLFKKFLEL 347
PkH_051470    YEVLCFEDRLNDYEINKELKKMEEMPKKCELVSLYWQSYINEKSKYINANKRLFKKFLEL 317
PkH_090690    YEVLCFEDRLNDYEINKELNEMEEMPKKCELVSLYWQSFMNEKRKYINANKRLFKKFLEL 361
PkH_120030    YEVLCFEDRLNDYEINKELNEMEEMPKKCELVSLYWQSFMNEESKYINASKRLFEKFLEL 386
PkH_143340    YEVLCFEDRLNDYEINKELNEMEEMPKKCELVSLYWQSFMNEKRKYINANKRLFKKFLEL 379
PkH_103230    YEMIPMDKLLTDDEINKELNEMEEMPKKCELVALYWKSYVNEIRNYVDTIRYLFEKFLEL 289

                                                          Helix 3
PkH_020250    KKKQNFETIAKYNNKWKKCSKIVGTNFKEQREYVNDIFYTHMTKENLSKDEFKEILGHVR 407
PkH_051470    KKKQNFGTMGKYNNKWKKCSKIVGTNFKEQREYVNDIFYTHMTKENLSKDEFKGILGHVR 377
PkH_090690    KKKQNFETLGKCNNKWKKCSKIVGTNFKEQREYVNDIFYTHMTKENLSKDEFKGILGHVR 421
PkH_120030    KEKQNFETMGKYNNKWKKCSKIVGTNFKEQREYVNDIFYTHMTKDNLSKDEFKEILGHVR 446
PkH_143340    KKKQNFDTMGKCNNKWKKCSKIVGTNFKEQREYVNDVFYSHMTKENLSKDEFKEILGHVR 439
PkH_103230    KKKQNFETLGKCNNKWKKCSKIVSTNFKDQRDYVNHIFYAHMTKENLSRDEFKEILGHLR 349




                      Helix 4
PkH_020250    DSWKEVTLKVTKECAAILEKPIVPDVKILDYNPYSGESYFKVSRMSSPQVSS 459
PkH_051470    DSWKEVTLKVTQECVALLEEPIVPDVKILDYDPYDGIAYFKVSRITSPQVNS 429
PkH_090690    ESWKEVTFKVTKECAAILEKPIVPDVKILDYHPYDGVAYFKVSRMSSPQVNS 473
PkH_120030    DSWKELNLKVTQECVALLEEPVVPDVKILDYDPHHGYAYFKVTRMSSPEVSS 498
PkH_143340    DSWKELTLKVTQECAALLEEPIVPDVKILDYDPYYGHGYFKVTRMSSPEVSS 491
PkH_103230    DSWKELTLKVTQECVALLEEPIVPDVKILDYKPYSGKAFFKVTRMSSPQVSS 401
```

Figure 4.3a: ClustalW alignment of six paralogous *P. knowlesi* PHIST protein sequences identified by LC-MS/MS analysis on infected RBC membranes. Conserved sequences are highlighted in pink, the PEXEL motif in blue, conserved tryptophans in red and the PHIST domain consisting of four consecutive alpha helices is shaded in grey.

Figure 4.3b

```
PkH_020250     ------------------------------------------------------------
PkH_120030     ------------------------------------------------------------
PkH_090690     ------------------------------------------------------------
PkH_051470     ------------------------------------------------------------
PkH_143340     ------------------------------------------------------------
PkH_103230     ------------------------------------------------------------
Pvx_003555     MHDKYRCRGLKLRALRHKDDFERQNAAAKLQVSNKVVRNLATEKEGAQSEDELELEDMEL 60

PkH_020250     ------------------------------------------------------------
PkH_120030     ------------------------------------------------------------
PkH_090690     ------------------------------------------------------------
PkH_051470     ------------------------------------------------------------
PkH_143340     ------------------------------------------------------------
PkH_103230     ------------------------------------------------------------
Pvx_003555     TEGGEEGEETYEEAEEEANADEELEVELEGGEEEVVEGEGEEEAVEGEGEGEEEAVEGEG 120

PkH_020250     ------------------------------------------------------------
PkH_120030     ------------------------------------------------------------
PkH_090690     ------------------------------------------------------------
PkH_051470     ------------------------------------------------------------
PkH_143340     ------------------------------------------------------------
PkH_103230     ------------------------------------------------------------
Pvx_003555     EGEEEAVEGEEAVEGEEAVEGEEAVEGEEAVEGEEAAEGEEAAEGEEAAEGEEAAEGEEA 180

PkH_020250     ------------------------------------------------------------
PkH_120030     ------------------------------------------------------------
PkH_090690     ------------------------------------------------------------
PkH_051470     ------------------------------------------------------------
PkH_143340     ---------------------------------------------------MAHTNEII 8
PkH_103230     ------------------------------------------------------------
Pvx_003555     AEGEEAAEGEEAAEGEEAAEGEEAAEGEEAAEGEEAAEGEEAAEGEEAAEGEEAAEGEAA 240

PkH_020250     ------------------------------------------------------------
PkH_120030     -----------------------MDEKKSTQRSSFFTKRT----------------- 17
PkH_090690     -----------------------MDEKKSTQRSSFFTKKT----------------- 17
PkH_051470     ------------------------------------------------------------
PkH_143340     INGASSRSMQSQCGNYQIVTARQERKMDERKSTQRSSFFTKRT---------------- 51
PkH_103230     --------MQSQCGSCQIVGAKQERKMDEKKSTQRSSFFTKRT---------------- 35
Pvx_003555     EGEAAEGEEAAEGKAAEGEEAAEGEAAEEEAAAESGGLLRGKTPTKPVIEDGENVTLTVD 300

PkH_020250     -----------------------HQDDARRQ-------------------------- 8
PkH_120030     ----------AMLFFVVACIFLKHQDDFSRQ-------------------------- 38
PkH_090690     ----------AMLFFVLACIFLKHQDDARRQ-------------------------- 38
PkH_051470     -----------------------HQDDARRQ-------------------------- 8
PkH_143340     ----------AMLFFVVACIFLKHQDDARRQ-------------------------- 72
PkH_103230     ----------AMLFFVVACIFLKHQDDTPRQ-------------------------- 56
Pvx_003555     DGEELYTDVNSGEYFGDATVEYEKLLDDVPAESADELAKNPYIRSFLEKALKQASRTDLY 360

PkH_020250     ----GVNAPLQVDN------------------------------------------- 18
PkH_120030     ----GVNAPLQVDN------------------------------------------- 48
PkH_090690     ----GVNAPLQVDN------------------------------------------- 48
PkH_051470     ----GVNASLQVHN------------------------------------------- 18
PkH_143340     ----GVNASLQVHN------------------------------------------- 82
PkH_103230     ----GVNAPLQVDN------------------------------------------- 66
Pvx_003555     EYEYDYNVDLSVDQSQESKDDDVEVDDNGRRLPRKAAPPVDKAKQDVMKDIVNYLSKNML 420

PkH_020250     ---IVCRNLASKE--EATSSGQNDYEVG----------------------------- 41
PkH_120030     ---NVGRNLSSTESEDEAPSGPYDYEVG----------------------------- 73
PkH_090690     ---IVGRNLASKE--EATSSGPYDYEVG----------------------------- 71
PkH_051470     ---KVGRNLASKE--EATPSGQNDYEVG----------------------------- 41
PkH_143340     ---KVGRNLSSKE--EATPSGPNDHRLG----------------------------- 105
PkH_103230     ---IVCRNLGSTETEEEAPS-KNNSGFG----------------------------- 90
Pvx_003555     AFVRQKRNVSGKE--GEAPTGPSGAQGGDSSQYASKFTFTDHSVDFSKYNKLDKEKFAAK 478
```

```
PkH_020250     --------------------------------------------------------QTEVP 46
PkH_120030     --------------------------------------------------------QTEDP 78
PkH_090690     --------------------------------------------------------QTEDP 76
PkH_051470     --------------------------------------------------------QTEVP 46
PkH_143340     --------------------------------------------------------QTEDH 110
PkH_103230     --------------------------------------------------------GNKKV 95
Pvx_003555     DDLKSRLKNEVVASMLDTEGDILTEEFGYLLRNYFDKVKLEEKKSQEAESAKPAEQEEEA 538


PkH_020250     SEGGRQASPAVGENTWGMEYPNVRQDLQKTAEQETPLFGEKTNYVVDADLCVYAIFRTFY 106
PkH_120030     SEGGRQASPAAGEMTLEMEYPNARQDVQKSAEQEAPLFGERTNYVAEGDLCPYAYFVHTC 138
PkH_090690     SEGGRQASPAAGEMTWGMEYPNVRQDVQKSAEQEAPLFGERTNYVAEGDLCVYAYFAHLC 136
PkH_051470     SEGGRQASPAVGENTWGMEYPNVRQDLQKSAEQEASLFGERTNFVIEADFTPYADFVHTC 106
PkH_143340     SEGGRQASPAVGE----MEYPNVRQDVQKPAEQEAPLFGERTNYITDGDSGNYAIFESLC 166
PkH_103230     YLG--KGAPKLNE-----------KKLQGKEDYVGNARMDDKTYLN--------TLQESY 134
Pvx_003555     EEAPEQKEEATAEKATEETTEAATEETTEAATEETTEAATEETTEAATEETTEAATEETT 598


PkH_020250     DNAFND-----------------------------------------NKKKKKGK 120
PkH_120030     MNAFND-----------------------------------------NKKKKKGE 152
PkH_090690     KNAFND-----------------------------------------NKKKKKGK 150
PkH_051470     MNAYYD-----------------------------------------DSKKTKKQ 120
PkH_143340     KNSFND-----------------------------------------NKKKKKGK 180
PkH_103230     DHMDND-----------------------------------------DKKEQKRK 148
Pvx_003555     EAATEETTEAATEETTEAATEEATEGATEEGAEETTEEATEEGAEEATEEGAEEATEEGA 658


PkH_020250     KEIEEKKKEEALLKEKEKSQK----------------------------------- 142
PkH_120030     KEIEEKKKEEALLKEKEKSQN----------------------------------- 174
PkH_090690     KEIEEKKKEEALLKEKEKSQK----------------------------------- 172
PkH_051470     REMEE-KKKEEALLKEKEKSQK---------------------------------- 141
PkH_143340     KEIEEKKKEEALLKEKEKSQN----------------------------------- 202
PkH_103230     ----EEKKKE------KKKCQK--------------------------------- 160
Pvx_003555     EETTEEATEEGAEETTEETTEEGAEEEATEEGAEETTEEGAEEAAEEGAEEGAEAATEEA 718


PkH_020250     ----------------KLNNWNQKVYRGPLAALEKYEDELEEERKRSLRYDKEERKRKM 185
PkH_120030     ----------------KLSIGNTKIYSGPLAALEKYEDELEEERKRSLRYDKEERKRKM 217
PkH_090690     ----------------ELSKGKKKVYSGPLAAIEKYEDEQEEERKRSLRYDKEERKRKM 215
PkH_051470     ----------------ELSKGKKKVYSGPLAALEKYEDELEEERKRSLRYDKEKRRKER 184
PkH_143340     ----------------KLSIGNTKVYSGPLAALEKYEDELEEERKRSLRYDKEERRKNK 245
PkH_103230     ----------------KFS-----------TWEKRKDKMEECKEK--KYKDK------ 183
Pvx_003555     TEEATEEATEEATEEATEEATEEATEEATAEVAEAATPEKVTEEATEEATEEGDNEPAEQAAEK 778


PkH_020250     KGKGKG-----------------------YSPIYTNGAHCDDIDAYMAKNMNGKNVEIE 221
PkH_120030     KGKGKGQSLL----------------DKWQSIVYEVGAYRDETEVNKRKCMNGEYKEIE 260
PkH_090690     KEMEKERSFL----------------NKFHSPLYTNGAHSDDIRMYRAEDR-------- 250
PkH_051470     DE--------------------------RVILNSMDKEKDIIADHR------------ 204
PkH_143340     EREE---------------------YEKIEKWGMDKKVVVEND------------ 267
PkH_103230     ------------------------------------------------------
Pvx_003555     EEDVKGGLMDNETYYNTLQELYEEIENDDKKEKEKIQKAKEQEELEKKLFKESKKGKKKE 838


PkH_020250     DWELRKLKRAGYDVSDYGKHIPKDNTRSSGQDEEEHIYDEEGMNELLFGEFKKWEDSGAH 281
PkH_120030     DWEISKLKKGGHDESDYSKYIRKDNTRSLRYDEEEHIDDEEEMNELLFGEFKKWEDSGAH 320
PkH_090690     ----------GDDE-----YEEKNNTRSLRYDEEEHIYDEEGMNDLLFGEFKKWEDSGTH 295
PkH_051470     -------------EREYKEELEEERKRSLRRGKEELIDDEEGMHELLFGEFKKWENSGTH 251
PkH_143340     --------------DNYLEQISEKRKRSLRYDEEEHIDDEEEMNELLFGEFKKWENYGTH 313
PkH_103230     -------------------ISKNRKKCLRHDEQKHIDEAGELDKLLFGELELLEKYGTH 223
Pvx_003555     KKRRKKLCKMAKIVEKYAEEIPKDSERSLRYDKEEHIDDPDEMDDLLFGEFKTLEKYGTH 898


PkH_020250     TIPSSHYEVLCFEDRLNDYEINKELNEMEEMPKKCELVSLYWQSFMNEKRKYINANKRLF 341
PkH_120030     TIPSSHYEVLCFEDRLNDYEINKELNEMEEMPKKCELVSLYWQSFMNEESKYINASKRLF 380
PkH_090690     TIPSSHYEVLCFEDRLNDYEINKELNEMEEMPKKCELVSLYWQSFMNEKRKYINANKRLF 355
PkH_051470     TIPSSHYEVLCFEDRLNDYEINKELNEMEEMPKKCELVSLYWQSYINEKSKYINANKRLF 311
PkH_143340     TIPSSHYEVLCFEDRLNDYEINKELNEMEEMPKKCELVSLYWQSFMNEKRKYINANKRLF 373
PkH_103230     KVSTFHYEMIPMDKLLTDDEINKELNEMEEMPKKCELVALYWKSYVNEIRNYVDTIRYLF 283
Pvx_003555     KTSTFYYEMTCFDERLRDFEINTKLKEMEEVPEKWELLSLYWQSYRNERHKYLAVKKYLL 958


PkH_020250     KKFLELKKKQNFETIAKYNNKWKKCSKIVGTNFKEQREYVNDIFYTHMTKENLSKDEFKE 401
PkH_120030     EKFLELKEKQNFETMGKYNNKWKKCSKIVGTNFKEQREYVNDIFYTHMTKDNLSKDEFKE 440
PkH_090690     KKFLELKKKQNFETLGKCNNKWKKCSKIVGTNFKEQREYVNDIFYTHMTKENLSKDEFKG 415
PkH_051470     KKFLELKKKQNFGTMGKYNNKWKKCSKIVGTNFKEQREYVNDIFYTHMTKENLSKDEFKG 371
PkH_143340     KKFLELKKKQNFDTMGKCNNKWKKCSKIVGTNFKEQREYVNDVFYSHMTKENLSKDEFKE 433
PkH_103230     EKFLELKKKQNFETLGKCNNKWKKCSKIVSTNFKDQRDYVNHIFYAHMTKENLSRDEFKE 343
Pvx_003555     EKFLELKTNQSTEALPKYNKKWKQCEEIVDNNFTKQHEHVNDVFYTFVAKENLSRDEFKE 1018
```

```
PkH_020250    ILGHVRDSWKEVTLKVTKECAAILEKPIVPDVKILDYNPYSGESYFKVSRMSSPQVSS 459
PkH_120030    ILGHVRDSWKELNLKVTQECVALLEEPVVPDVKILDYDPHHGYAYFKVTRMSSPEVSS 498
PkH_090690    ILGHVRESWKEVTFKVTKECAAILEKPIVPDVKILDYHPYDGVAYFKVSRMSSPQVNS 473
PkH_051470    ILGHVRDSWKEVTLKVTQECVALLEEPIVPDVKILDYDPYDGIAYFKVSRITSPQVNS 429
PkH_143340    ILGHVRDSWKELTLKVTQECAALLEEPIVPDVKILDYDPYYGHGYFKVTRMSSPEVSS 491
PkH_103230    ILGHLRDSWKELTLKVTQECVALLEEPIVPDVKILDYKPYSGKAFFKVTRMSSPQVSS 401
Pvx_003555    ILNDVRASWKKVTLKTRDECVALFEEPIVELEVKTPTQPPKGLKYWKKFKFKGFFV-- 1074
```

Figure 4.3b: ClustalW alignment of the six *P. knowlesi* PHIST paralogs with their known ortholog in *P. vivax*. Conserved sequences are highlighted in pink, the PEXEL motif in blue and conserved tryptophans in red. In addition, the N-terminal repeat motifs are highlighted in color, and tandem motifs indicated by underlining.

Fourteen of the identified hypothetical proteins were found to be either unique or shared among very few species. Pvx_101515 (Table 4.1a; Supplementary table 4.1) is unique to the *P. vivax/P. cynomolgi* lineage while Pvx_112110 and Pvx_123750 have orthologs in *P. cynomolgi* and *P. knowlesi* only (Table 4.1a; Supplementary table 4.1). Interestingly, 12 of the identified *P. knowlesi* proteins either have within-species homologs only or are unique (Table 4.1b; Supplementary table 4.2). Three of these 12 proteins, PkH_010030, PkH_073420 and PkH_031240, are known members of the Pk-fam-c gene family, which is *P. knowlesi*-specific (Table 4.1b, highlighted in pink(6, 7)). These Pk-fam-c family members contain the PEXEL motif and have an amino acid sequence identity between 45-68% among them (Figure 4.4), but their function is unknown. The above unique proteins indicate possible specific roles within the *P. knowlesi* lineage, the *P.vivax*/*P. cynomolgi* lineage or the *P. vivax/P. cynomolgi/ P. knowlesi* lineage.

Figure 4.4a

```
PkH_031240    MAVPTKQNKKARKVSSAFPFCSKISLYTLLIWIVNCSNSCQYDGNSYSVMNNSLGKAFDS 60
PkH_010030    ------------------------------------------------------------
PkH_073420    MAVPTKQNKKARKAG-AFPFCSKVSLYTLLIWIVNCSNSCQYDGNSYSVMNNSLGKAFDS 59

PkH_031240    RALRLLAEAVVYEDEVEQDGPVVYEDEIEQDGPVIYEEDIEQDGPVVYEEDIEQDGPVVY 120
PkH_010030    ------------------------------------------------------------
PkH_073420    RTLRLLAEAVE------------------------------------------------- 70

PkH_031240    EEDIEQDGPVVYEEDIEQDGPVVYEEDIEQDGPVVYEEDIEQDGAVESEDEIEQDGAVES 180
PkH_010030    ---------------------HEDDFSRQG---------VSTLLQVHNKVGRNLASKE 28
PkH_073420    ---------------------SEDEIEQDGPVVYEEDIEQDGPVVYEEDIEQDGPVVH 107

PkH_031240    EDEIEQDGPVVYEDEVEQDGPVVYEDEIEQDGPVIYEEDIEQDGPVVYEDEIEQDGPVVY 240
PkH_010030    EKKEEKK---------------------------EEKKEEK----------------- 42
PkH_073420    EEDIEQDGP-----------------------VVHEEDIEQDGP------------VVH 131

PkH_031240    EEEIEQDGPVVYEEDIEQDGPVVYEEVAQQVPKQLGKKGKKEYKDGRSTFVGARDINKEK 300
PkH_010030    KEEKK------------------EEKKKTNKTLLGKKGKKEYKDGRSTFVGARDINKEK 83
PkH_073420    EEDIE------------QDGPVVHEEVAQQVPQHSGKKGKKEHKDGKSTSVGLRGANKE- 178

PkH_031240    SEMDSYRQRKLDFWEHFEPTITANFEEVLKRCISRKAGEQDVDEYYSAGLPKLGWNADPY 360
PkH_010030    SEMDSYRQRKLDFWEHFEPTMTANFEKVLKRCFARKAGEEDMDEDYSVVLPKLGWNADPY 143
PkH_073420    NETDSYDKGEFCLREHCEKKITADLEEALKRCAVQEA-QEDMDGQLSAGMPKLGGDGDPY 237

PkH_031240    GVLKKTKKGTVQEGYRKMLENNFRNVPYIDDTQNDHKSNEPYLEREYGRVELDADAKTST 420
PkH_010030    GVLKKTKKGTVQEGYRKMLENNFRNVPYVHDSQNAHKSNEPYLEREYGRVELGADAKT-- 201
PkH_073420    GLLKKEKKGGAQEGYKKMLESNFSNIPYVDDSQNAHKSNKTHLEREYGRVDLTPPTKTTK 297

PkH_031240    NLRKGAEKKMNKKM------------------------------------- 434
PkH_010030    --------------------------------------------------
PkH_073420    NVTKNAETKVDKEKVEEGNGKEKVKEGMKEKEGKKEKEKEKKKKKKLLFCF 348
```

Figure 4.4b
PkH_031240

```
MAVPTKQNKKARKVSSAFPFCSKISLYTLLIWIVNCSNSCQYDGNSYSVMNNSLGKAFDSRALRLLAEAVV
YEDEVEQDGPVVYEDEIEQDGPVIYEEDIEQDGPVVYEEDIEQDGPVVYEEDIEQDGPVVYEEDIEQDGPV
VYEEDIEQDGPVVYEDEIEQDGAVESEDEIEQDGAVESEDEIEQDGPVVYEDEVEQDGPVVYEDEIEQDGP
VIYEEDIEQDGPVVYEDEIEQDGPVVYEEEIEQDGPVVYEEDIEQDGPVVYEEVAQQVPKQLGKKGKKEYK
DGRSTFVGARDINKEKSEMDSYRQRKLDFWEHFEPTITANFEEVLKRCISRKAGEQDVDEYYSAGLPKLGW
NADPYGVLKKTKKGTVQEGYRKMLENNFRNVPYIDDTQNDHKSNEPYLEREYGRVELDADAKTSTNLRKGA
EKKMNKKM
```

PkH_010030 (fragment)

```
HEDDFSRQGVSTLLQVHNKVGRNLASKEEKKEEKKEEKKEEKKEEKKEEKKKTNKTLLGKKGKKEYKDGRS
TFVGARDINKEKSEMDSYRQRKLDFWEHFEPTMTANFEKVLKRCFARKAGEEDMDEDYSVVLPKLGWNADP
YGVLKKTKKGTVQEGYRKMLENNFRNVPYVHDSQNAHKSNEPYLEREYGRVELGADAKT
```

PkH_073420

```
MAVPTKQNKKARKAGAFPFCSKVSLYTLLIWIVNCSNSCQYDGNSYSVMNNSLGKAFDSRTLRLLAEAVES
EDEIEQDGPVVYEDEIEQDGPVVYEEDIEQDGPVVHEEDIEQDGPVVHEEDIEQDGPVVHEEDIEQDGPVV
HEEVAQQVPQHSGKKGKKEHKDGKSTSVGLRGANKENETDSYDKGEFCLREHCEKKITADLEEALKRCAVQ
EAQEDMDGQLSAGMPKLGGDGDPYGLLKKEKKGGAQEGYKKMLESNFSNIPYVDDSQNAHKSNKTHLEREY
GRVDLTPPTKTTKNVTKNAETKVDKEKVEEGNGKEKVKEGMKEKEGKKEKEKEKKKKKKLLFCF
```

Figure 4.4: Pk-fam-c proteins identified by LC-MS/MS analysis of *P. knowlesi*-infected

erythrocyte membranes. (a) ClustalW alignment showing conserved residues (pink), PEXEL

motif (blue) and conserved lysines (green).

(b) Repeat motifs identified by eye in the three membrane-localized Pk-fam-c family members.

Repeat motifs are highlighted in color, and tandem motifs indicated by underlining.

We also identified a 70kDa member of the heat shock protein (HSP) family,

annotated in *P. vivax* as a putative 78kDa glucose-regulated protein precursor (GRP78;

Pvx_099315; supplementary table 4.1) and in *P. knowlesi* as a putative heat shock protein

(PkH_071520; supplementary table 4.2). Because this protein has an endoplasmic

reticulum (ER) retention signal (C-terminal sequence SDEL), one would expect that it

would exclusively be localized to the ER, but we confirmed its presence on *P. knowlesi*-

infected RBC membranes by Western blot analysis (Figure 4.5). In this experiment,

PkH_071520 was detected in infected RBC membrane ghosts but absent in uninfected

membrane ghosts, ruling out the possibility of cross-reactivity with host HSPs. The *P.*

falciparum homolog of this protein has previously been detected in Maurer's clefts (28),

and is predicted to interact with the exported protein PfEMP1 (29).
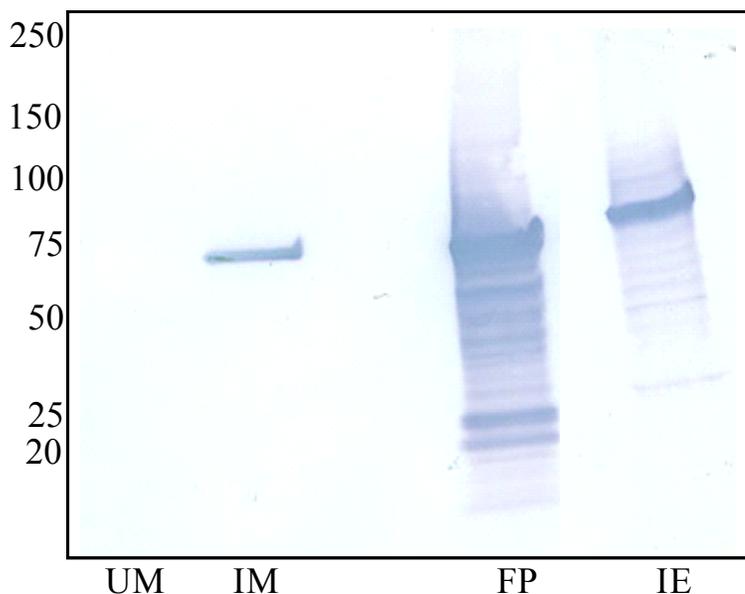
Figure 4.5: Western blot of *P. knowlesi*-infected RBC membrane ghosts using rabbit antisera

against recombinant heat shock protein-2 (anti-rPkHSP2). Uninfected membrane ghosts (UM), *P.*

knowlesi-infected membrane ghosts (IM), Ficoll pellet (FP) and whole *P. knowlesi*-infected RBCs

(IE) are shown.

41 rhesus RBC proteins were detected in *P. cynomolgi*-infected RBC membranes (Supplementary table 3) while 81 were detected in *P. knowlesi*-infected RBC membranes (Supplementary table 4), with different isoforms of these proteins being identified. Some isoforms, such as isoform 2 of spectrin-alpha, were more abundant than others. The isoforms of proteins presented in this paper are those with high spectral scores and/or with unique peptides (Supplementary tables 4.3 and 4.4).

4.5. Discussion

We sought to produce and analyze the *P. cynomolgi* and *P. knowlesi*-infected RBC membrane proteomes with the aim of undertaking cross-species comparisons that would highlight unique proteins and common entities between the species that could possibly lead to the identification of therapeutic targets. Since direct proteome analysis of *P. vivax* membranes remains a challenge, we used the evolutionarily close species *P. cynomolgi* to identify membrane proteins with orthologs in *P. vivax*. We first separated *P. cynomolgi*- and *P. knowlesi*-infected RBCs membrane ghosts from parasites using a Ficoll gradient, separated the membrane proteins on SDS-PAGE gradient gels and digested them with trypsin. The extracted peptides were then subjected LC-MS/MS analysis. Our results yielded 109 *P. cynomolgi* homologs and 129 *P. knowlesi* proteins.

Perhaps the most striking group of proteins repeatedly identified in high abundance belongs to the PHIST protein family. There are 71 known PHIST paralogs in *P. falciparum*, 39 in *P. vivax* and 27 in *P. knowlesi* and the PHIST protein structure is unique (22). We have localized, for the first time, twelve *P. knowlesi* PHIST family members and three *P. cynomolgi* PHIST orthologs to the infected RBC membranes. *P.*

knowlesi orthologs of the three *P. cynomolgi* PHIST proteins identified were also present in *P. knowlesi* membrane ghosts (Table 4.1a & 4.1b). These common entities provide evidence for possible similarities in functions for these PHIST proteins in the host RBC membranes of their respective *Plasmodium* species. Thus far, the function(s) of PHIST proteins remain unknown (22). More research is needed to more precisely determine the localization of these proteins, in addition to determining their function(s). PHIST domains cluster into three subgroups based on the presence and position of conserved tryptophan residues – PHISTa, PHISTb and PHISTc. PHISTc is the most diverse subgroup and is represented by eight proteins in our mass spectrometric analysis of *P. knowlesi*-infected RBC membranes (Supplementary table 4.2; Table 4.1b).

We have utilized a second approach to identifying novel *Plasmodium*-infected erythrocyte membrane proteins using a panel of monoclonal antibodies raised against *P. vivax* schizonts. This has allowed us to localize the *P. cynomolgi* ortholog of an 81 kDa *P. vivax* PHIST protein (Pvx_093680; Table 4.1a; Supplementary table 4.1) to the caveola-vesicle complexes of *P. cynomolgi*-infected RBC membranes (manuscript in preparation). We determined that the closely related proteins to Pvx_093680, including PkH_011720 (Table 4.1b; Supplementary table 4.2) are conserved in the functional C-terminus as well as in the N-terminus, so we expected that the same trend would be seen across other PHIST proteins. However, we discovered that of the twelve PHIST domain-containing hypothetical proteins identified by mass spectrometric analysis of *P. knowlesi*-infected RBC membranes, six of these were within-species paralogs with significant sequence homology among them when aligned by ClustalW (Figure 4.3a). Their only

known ortholog, Pvx_003555, contains tandem repeats in the C-terminus, but the significance of this is unknown.

Because of the degree of sequence identity between *P. vivax* and *P. cynomolgi*, the *P. vivax* database was utilized for the proteomic analysis of *P. cynomolgi*-infected RBC membranes. Nevertheless, it is likely that the absence of a *P. cynomolgi* genome database may have reduced the chances of detecting more *P. cynomolgi*-specific proteins. We made this observation due to the significantly greater number of proteins, hypothetical and otherwise, identified from *P. knowlesi*-infected RBC membranes compared to the *P. vivax* homologs of *P. cynomolgi* proteins.

In this study, we have provided some insights into the host (rhesus) RBC proteomic make-up when infected with *P. cynomolgi* and *P. knowlesi*. Using animal models to study human disease requires a more complete database of the model species, and the relative lack of protein annotations for rhesus models is a major obstacle in utilizing them for the study of human disease states, including malaria. There are approximately 42,000 protein annotations for rhesus compared to over 330,000 for humans (NCBI.org). Rhesus monkeys are routinely infected with *P. cynomolgi* and *P. knowlesi* to obtain samples for various malaria studies. Thus, the relative lack of protein information limits the use of rhesus monkeys for proteomics studies (30). Comparing the rhesus and human RBC proteome in the presence and absence of malaria could highlight differences of potential importance between malarial species, especially when analyzing species that are biologically and evolutionarily similar such as *P. vivax* (human) and *P. cynomolgi* (non-human primate). In addition, proteomics on diseased RBCs versus normal RBCs offers the ability to identify markers that are altered or absent in the

diseased state (reviewed in (31). It could also lead to the identification of targets for therapeutics. It is likely that we may not be identifying proteins that are differentially regulated in malaria-infected versus normal RBCs, or other proteins that may be post-translationally modified.

This study, with its unique methodology and high throughput potential, serves to efficiently advance *Plasmodium* research related to the function of the infected RBC membrane. It is likely that these proteins may play crucial roles in the survival and functioning of the intracellular parasite. The methods used here represent a global approach to identifying proteins that will establish *Plasmodium*-infected RBC membrane proteomes. The number of signal peptide and PEXEL-containing hypothetical proteins suggests that these and other hitherto unidentified proteins may have potentially important roles within the membrane of infected RBCs. For example, proteins predicted to have multiple transmembrane domains, such as (Pvx_112110 (three), PkH_060020 (three), PkH_134550 (two) and PkH_140050 (two) may form channels through the infected RBC membrane. Future studies include determining the proteomic profile of *P. vivax*-infected erythrocyte membrane proteins and comparing/contrasting the structural and antigenic composition among the RBC membranes infected by this species to those of *P. knowlesi* and *P. cynomolgi.*

4.6. Supplementary Tables

Supplementary Table 4.1: Putative *P. vivax* orthologous proteins detected after LC-MS/MS analysis of *P. cynomolgi*-infected membrane ghosts.

MW- protein molecular weight. pI (32-34) – protein isoelectric point. PEXEL– *Plasmodium* export element.

Signal peptide is predicted using SignalP algorithm

(https://www.cbs.dtu.dk/services/SignalP/; (35).

Recessed signal sequence is detected using *Plasmodium*-specific MalSig algorithm

(http://bioserve.latrobe.edu.au/cgi-bin/pfsigseq.py; (36).

Transmembrane domain predicted by TMHMM

(http://www.cbs.dtu.dk/services/TMHMM-2.0/; (37).

Supplementary Table 4.2: Putative *P. knowlesi* – infected erythrocyte membrane proteins detected after LC-MS/MS analysis of *P. knowlesi*-infected membrane ghosts.

MW- protein molecular weight. pI – protein isoelectric point. *Pexel* – Plasmodium export element.

Signal peptide is predicted using SignalP algorithm

(https://www.cbs.dtu.dk/services/SignalP/; (35).

Recessed signal sequence is detected using *Plasmodium*-specific MalSig algorithm

(http://bioserve.latrobe.edu.au/cgi-bin/pfsigseq.py; (36)

Transmembrane domain predicted by TMHMM

(http://www.cbs.dtu.dk/services/TMHMM-2.0/; (37)

Supplementary Table 4.3: Rhesus proteins identified in membranes of *P. cynomolgi-*

infected RBCs by LC-MS/MS.

Supplementary Table 4.4: Rhesus proteins identified in membranes of *P. knowlesi-*

infected RBCs by LC-MS/MS.

REFERENCES

1. WHO. 2008. World Malaria Report 2008, Vol. 2009.

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

3. Waters, A. P., D. G. Higgins, and T. F. McCutchan. 1993. Evolutionary relatedness of some primate models of Plasmodium. *Mol Biol Evol 10:914*.

4. Qari, S. H., Y. P. Shi, N. J. Pieniazek, W. E. Collins, and A. A. Lal. 1996. Phylogenetic relationship among the malaria parasites based on small subunit rRNA gene sequences: monophyletic nature of the human malaria parasite, Plasmodium falciparum. *Mol Phylogenet Evol 6:157*.

5. Vargas-Serrato, E., V. Corredor, and M. R. Galinski. 2003. Phylogenetic analysis of CSP and MSP-9 gene sequences demonstrates the close relationship of Plasmodium coatneyi to Plasmodium knowlesi. *Infect Genet Evol 3:67*.

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

7. Pain, A., U. Bohme, A. E. Berry, K. Mungall, R. D. Finn, A. P. Jackson, T. Mourier, J. Mistry, E. M. Pasini, M. A. Aslett, S. Balasubrammaniam, K. Borgwardt, K. Brooks, C. Carret, T. J. Carver, I. Cherevach, T. Chillingworth, T. G. Clark, M. R. Galinski, N. Hall, D. Harper, D. Harris, H. Hauser, A. Ivens, C. S. Janssen, T. Keane, N. Larke, S. Lapp, M. Marti, S. Moule, I. M. Meyer, D. Ormond, N. Peters, M. Sanders, S. Sanders, T. J. Sargeant, M. Simmonds, F. Smith, R. Squares, S. Thurston, A. R. Tivey, D. Walker, B. White, E. Zuiderwijk, C. Churcher, M. A. Quail, A. F. Cowman, C. M. Turner, M. A. Rajandream, C. H. Kocken, A. W. Thomas, C. I. Newbold, B. G. Barrell, and M. Berriman. 2008. The genome of the simian and human malaria parasite Plasmodium knowlesi. *Nature 455:799*.

8. Barnwell, J. W., and M. R. Galinski. 1991. The adhesion of malaria merozoite proteins to erythrocytes: a reflection of function? *Res Immunol 142:666*.

9. Galinski, M. R., and V. Corredor. 2004. Variant antigen expression in malaria infections: posttranscriptional gene silencing, virulence and severe pathology. *Mol Biochem Parasitol 134:17*.

10. Singh, B., L. Kim Sung, A. Matusop, A. Radhakrishnan, S. S. Shamsul, J. Cox-Singh, A. Thomas, and D. J. Conway. 2004. A large focus of naturally acquired Plasmodium knowlesi infections in human beings. *Lancet 363:1017*.
11. Jongwutiwes, S., C. Putaporntip, T. Iwasaki, T. Sata, and H. Kanbara. 2004. Naturally acquired Plasmodium knowlesi malaria in human, Thailand. *Emerg Infect Dis 10:2211*.
12. Udeinya, I. J., J. A. Schmidt, M. Aikawa, L. H. Miller, and I. Green. 1981. Falciparum malaria-infected erythrocytes specifically bind to cultured human endothelial cells. *Science 213:555*.
13. Deitsch, K. W., and T. E. Wellems. 1996. Membrane modifications in erythrocytes parasitized by Plasmodium falciparum. *Mol Biochem Parasitol 76:1*.
14. Barnwell, J. W. 1990. Vesicle-mediated transport of membrane and proteins in malaria-infected erythrocytes. *Blood Cells 16:379*.
15. von Heijne, G. 1985. Ribosome-SRP-signal sequence interactions. The relay helix hypothesis. *FEBS letters 190:1*.
16. Lingelbach, K. R. 1993. Plasmodium falciparum: a molecular view of protein transport from the parasite into the host erythrocyte. *Exp Parasitol 76:318*.
17. Albano, F. R., M. Foley, and L. Tilley. 1999. Export of parasite proteins to the erythrocyte cytoplasm: secretory machinery and traffic signals. *Novartis Found Symp 226:157*.
18. Wickham, M. E., M. Rug, S. A. Ralph, N. Klonis, G. I. McFadden, L. Tilley, and A. F. Cowman. 2001. Trafficking and assembly of the cytoadherence complex in Plasmodium falciparum-infected human erythrocytes. *Embo J 20:5636*.
19. Marti, M., R. T. Good, M. Rug, E. Knuepfer, and A. F. Cowman. 2004. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science 306:1930*.
20. Hiller, N. L., S. Bhattacharjee, C. van Ooij, K. Liolios, T. Harrison, C. Lopez-Estrano, and K. Haldar. 2004. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science 306:1934*.
21. Boddey, J. A., R. L. Moritz, R. J. Simpson, and A. F. Cowman. 2009. Role of the Plasmodium export element in trafficking parasite proteins to the infected erythrocyte. *Traffic 10:285*.
22. Sargeant, T. J., M. Marti, E. Caler, J. M. Carlton, K. Simpson, T. P. Speed, and A. F. Cowman. 2006. Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol 7:R12*.
23. Aley, S. B., J. W. Barnwell, W. Daniel, and R. J. Howard. 1984. Identification of parasite proteins in a membrane preparation enriched for the surface membrane of erythrocytes infected with Plasmodium knowlesi. *Mol Biochem Parasitol 12:69*.
24. Pasini, E. M., M. Kirkegaard, P. Mortensen, H. U. Lutz, A. W. Thomas, and M. Mann. 2006. In-depth analysis of the membrane and cytosolic proteome of red blood cells. *Blood 108:791*.
25. Korir, C. C., and M. R. Galinski. 2006. Proteomic studies of Plasmodium knowlesi SICA variant antigens demonstrate their relationship with P. falciparum EMP1. *Infect Genet Evol 6:75*.
26. Peng, J., and S. P. Gygi. 2001. Proteomics: the move to mixtures. *J Mass Spectrom 36:1083*.

27. Peng, J., J. E. Elias, C. C. Thoreen, L. J. Licklider, and S. P. Gygi. 2003. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res 2:43*.

28. Vincensini, L., S. Richert, T. Blisnick, A. Van Dorsselaer, E. Leize-Wagner, T. Rabilloud, and C. Braun Breton. 2005. Proteomic analysis identifies novel proteins of the Maurer's clefts, a secretory compartment delivering Plasmodium falciparum proteins to the surface of its host cell. *Mol Cell Proteomics 4:582*.

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

30. Tannu, N. S., and S. E. Hemby. 2007. De novo protein sequence analysis of Macaca mulatta. *BMC Genomics 8:270*.

31. Pasini, E. M., H. U. Lutz, M. Mann, and A. W. Thomas. 2009. Red Blood Cell (RBC) membrane proteomics - Part II: Comparative proteomics and RBC patho-physiology. *J Proteomics*.

32. Baird, J. K. 2008. Real-world therapies and the problem of vivax malaria. *N Engl J Med 359:2601*.

33. Galinski, M. R., and J. W. Barnwell. 2008. Plasmodium vivax: who cares? *Malar J 7 Suppl 1:S9*.

34. Price, R. N., N. M. Douglas, and N. M. Anstey. 2009. New developments in Plasmodium vivax malaria: severe disease and the rise of chloroquine resistance. *Curr Opin Infect Dis*.

35. Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak. 2004. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol 340:783*.

36. Logan E., H. R., Klonis N., Herd S., Tilley L. 2004. Fuzzy classification of secretory signals in proteins encoded by the Plasmodium falciparum genome. In *Lecture Notes in Computer Science*, Vol. 3213/2004. N. M. G. e. a. (Eds.), ed. Springer-Verlag Berlin, Heidelberg, p. 1023.

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

CHAPTER FIVE

CONCLUSION

5.1. Introduction

The *Plasmodium* genome consists of more than 5000 genes (1-4). Despite the availability of complete genome sequence data for *P. falciparum* (3), *P. vivax* (1), *P. knowlesi* (4) and *P. yoelii* (2), and the application of advanced sequencing and post-genomic technologies (Reviewed in (5)), many of the encoded proteins have not yet been annotated. A large number of these hypothetical proteins are exported beyond the intraerythrocytic parasite to various organelles, cytoplasmic clefts and vesicles, onward to the RBC membrane, and still others are exported out of the host RBC. In this dissertation, we sought to elucidate the antigenic composition of *P. cynomolgi* and *P. knowlesi* infected erythrocyte membranes with the aim of gaining insights into the structural and functional modifications created by each species.

In chapter 2, we identified residues of potential functional importance on *Plasmodium* glyceraldehyde-3-phosphate dehydrogenase (G3PDH), which could serve as potential drug targets. We also In chapters 3 and 4, we utilized two proteomic approaches to identify proteins localized to the membranes of *Plasmodium*-infected erythrocytes – a targeted approach (Chapter 3) where monoclonal antibodies raised against *P. vivax* blood-stage schizonts were used to immunoprecipitate the *P. cynomolgi* ortholog of an 81 kDa *P. vivax* PHIST family protein; and a global proteomic approach (Chapter 4) where membrane ghosts of *P. cynomolgi* and *P. knowlesi* infected RBCs were prepared for analysis by liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). The global approach identified 109 *P. cynomolgi* orthologs and 129 *P. knowlesi* membrane associated proteins, including PHIST family members, Fam-c proteins, G3PDH and heat shock protein 2 (HSP2).

The predicted localization of PkG3PDH and PkHSP2 was further assessed by western immunoblot analysis of *P. knowlesi* infected RBC membrane ghosts using rabbit antiserum against recombinant PkG3PDH and PkHSP2; the proteins were present in *P. knowlesi* infected RBC membrane preparations but absent in uninfected RBC membrane preparations.

The identification of PkG3PDH in our proteomic analysis of *P. knowlesi* infected RBC membranes as a protein with a potentially relevant biological function at this location is supported by studies that detected G3PDH on the cell membranes of other organisms. For example, G3PDH of the fungus *Paracoccidiodes brasiliensis* is a cell surface protein involved in fungal adhesion and invasion, and it binds extracellular matrix components fibronectin, laminin and type-1 collagen (6). Also, membrane-bound G3PDH of group A streptococci was reported to bind cell membrane proteins fibronectin, lysozyme, myosin and actin, indicating a function in bacterial colonization (7). Lastly, the cell wall-associated G3PDH of *Candida albicans* is a fibronectin and laminin-binding protein (8). It is therefore likely that, in addition to its role in the nucleus, PkG3PDH may have a function in the membranes of infected erythrocytes.

Interestingly, over the years, a number of novel functions have been assigned to G3PDH in different species: in addition to being implicated in cell signaling, microtubule bundling and DNA replication and repair (9-12), G3PDH has been shown to accumulate in mitochondria during apoptosis where it induces mitochondrial membrane permeabilization (13), leading to downstream pro-apoptotic events. But in the absence of caspase activation, G3PDH was shown to prevent cell death despite mitochondrial membrane permeabilization and cytochrome c release; G3PDH-mediated protection from

caspase-independent cell death involved upregulation of glycolysis and nuclear function (14).

5.2. Protein-protein Associations May Predict Potential Functions For Membrane-localized Proteins

PkHSP2 contains a putative four-residue endoplasmic reticulum (ER)-retentive motif (SDEL) on the C-terminus of its amino acid sequence. Its *P. vivax* ortholog is a 78 kDa glucose-regulated protein (PvGRP78). Two studies have demonstrated that the ER-retentive XDEL motif may not confine a protein, such as PkHSP2, to the ER. In the first study, when NG108-15 cell surface proteins were cross-linked using the membrane-impermeable compound 3,3'-dithiobis(sulfo succinimidyl propionate)(DTSSP), a 78 kDa glucose-regulated protein (GRP78) containing a KDEL motif in the C-terminus of its amino acid sequence was detected in association with other cell surface proteins (15). GRP78 formed high molecular weight complexes with these proteins, which were stable when electrophoresed under non-reduced conditions but disappeared under reduced conditions (15). In the second study, ER-resident proteins GRP78, calreticulin and protein disulfide isomerase were identified in RBC membrane ghosts obtained from patients suffering from congenital dyserythropoietic anemia, a disorder that causes ineffective erythropoiesis (16). These observations provide indirect evidence that PkHSP2 may indeed be a resident of *Plasmodium* infected RBC membranes. Determining the interacting partners of proteins with unknown functions or those detected in unusual locations, such as membrane PkHSP-2, can help define their potential functions.

In chapter 3, we demonstrated that *P. cynomolgi* PHIST-81 (PcyPHIST-81)

localizes to the cytoplasmic face of CVC tubules. It is possible that at this location,

PcyPHIST-81 may associate with RBC cytoskeletal proteins such as actin, spectrin or

ankyrin. Performing co-localization immunofluorescence assays and immunoelectron

microscopy using the available rabbit antisera against recombinant *P. vivax* PHIST-81 in

combination with monoclonal antibodies raised against each cytoskeletal protein may

provide evidence for or against this proposed association. If it is determined that

PcyPHIST-81 interacts with cytoskeletal proteins, this finding would have implications

for its function. Specifically, it would provide clues as to whether the protein is simply

involved in remodeling the RBC cytoskeleton resulting in CVC formation, or if it is a

structural component of the CVCs. While experimental evidence is required to answer

this question, the latter function seems more likely due to this protein's consistent and

abundant localization in the CVCs.

Other techniques that have previously been utilized to identify protein interactors

– and can be used to determine the potential association of membrane PkG3PDH,

PkHSP2 and PcyPHIST-81 with other proteins on the RBC membrane – include *in vitro*

binding assays, blue native gel electrophoresis (BN PAGE), co-immunoprecipitation and

protein cross-linking (17). *In vitro* binding assays would involve generating multiple

recombinant proteins that consist of a number of amino acid residues spanning the entire

protein sequence and testing these for binding to inside out RBC vesicles. Such

experiments have the potential to identify the specific domain(s) within the protein of

interest that may interact with RBC cytoskeletal proteins. This approach was used to

determine the spectrin- and actin-binding domains of *P. falciparum* membrane protein

PfEMP3 (18).

Functions assigned to heat shock protein family members include translocating

proteins from the cytoplasm to the ER, sequestering proteins in the ER until they are

glycosylated and assembled into complexes, and forming tight complexes with misfolded

or improperly glycosylated proteins until they are reassembled properly (19, 20).

Chemical cross-linking in combination with BN PAGE, in-gel digestion and LC-MS/MS

can be used in attempts to identify proteins that interact with membrane-associated

PkHSP2, as has previously been done for merozoite surface protein-1 localized to

detergent-resistant membranes (17). The chemicals used for cross-linking included

membrane-permeable compounds disuccinimidyl suberate (DSS), disuccinimidyl

gluterate (DSG) and dithiobis(succinimidyl propionate; DSP) and the membrane-

impermeable DTSSP (17). Transcriptional arrays performed in parallel with protein

cross-linking determined whether associations identified at the proteomic level were

present at the transcript level by investigating the transcriptional regulation profile of

potential interactors (17). If it is discovered that PkHSP2 associates with membrane

cytoskeletal proteins, one could speculate that the KDEL motif did not limit PkHSP-2 to

the ER for a couple of reasons: First, while performing its function of binding to and

assisting other proteins in their folding and maturation in the ER, the protein's KDEL

motif may be concealed from its respective receptor, resulting in PkHSP2 being

transported to the membrane in association with the membrane-targeted proteins to which

it was bound. Secondly, in relation to the heat-shock response of the HSP family, it is

possible that PkHSP2 is activated in response to stress induced by *P. knowlesi* malarial infection, leading to its export out of the ER in order to stabilize the RBC membrane.

5.3. PcyPHIST-81 Trafficking In *P. cynomolgi* Infected RBCs

The ability to visualize the trafficking of soluble PcyPHIST-81 will contribute to knowledge about the route this protein follows from its production in the parasite through the parasitophorous vacuole membrane and onward to the RBC membrane. I hypothesize that once the protein is translated in the ER, its PEXEL motif is cleaved and N-acetylated. Then, the protein's recessed signal sequence directs translocation across the ER membrane where its processed N-terminus is recognized in the PV by an ATP-powered translocon, such as the *Plasmodium* translocon of exported proteins (PTEX) complex (21), which then actively transports PcyPHIST-81 across the PV membrane. At this point, one of two pathways is followed:

1. Cytoplasmic clefts that dock at the PV membrane envelope PcyPHIST-81, transport and release the protein onto CVC tubules on the RBC membrane, or

2. PcyPHIST-81 associates with cytoplasmic vesicles which bud from the PVM, travel through the RBC cytoplasm and fuse with caveolae on the RBC membrane, forming CVCs with the protein already attached.

Fluorescent exporter protein technology, utilizing fluorophores such as Green Fluorescent Protein (GFP), has been widely employed in the *P. falciparum* field to map the path exported proteins follow in real time to their final destinations (22). This

technology has a number of advantages: the fluorophores are biologically stable and non-toxic to *P. falciparum* parasites and they are easy to use with widely available fluorescein filters (22). Thus, it can be envisioned that these experimental biological tools can be utilized to study the transport of PcyPHIST81.

Gene knock-out or disruption studies can also be carried out to determine whether PcyPHIST-81 is required for *P. cynomolgi* survival at the intraerythrocytic stage. Gene disruption, knock-out and transfection protocols are well-established for *P. falciparum* (23) and *P. knowlesi* (24), but much less so for *P. vivax* (25) and *P. cynomolgi* (26). An important follow-up to this dissertation is to determine if PcyPHIST-81 is essential for the formation of CVCs, and for parasite survival, *in vivo*. To this end, I started to develop *pcyphist-81* and the control circumsporozoite protein *pcycsp* gene constructs using a pBluescript plasmid backbone containing a mutated *Toxoplasma gondii* dihydrofolate reductase (TgDHFR) gene that confers pyrimethamine resistance. TgDHFR is flanked by 5' and 3' *P. berghei* untranslated regions (UTR) that provide a promoter and stop signal, respectively, for the TgDHFR construct. 5' and 3' sequences for *pcyphist-81* and *pcycsp* genes were amplified with the primers below:

Table 5.1: Primers used for PCR amplification of 5' and 3' constructs

Gene	Location	Enzyme	Primer Sequence
pcycsp	5' Forward	BamHI	5' tcgggatccATGAAGAACTTCATTCTCTTG 3'
pcycsp	5' Reverse	NdeI	5' ggcgtgcatatgCCTGCCTGATTTCCTCC 3'
pcycsp	3' Forward	NdeI	5'gccgtgacatatgGGAGGAGCAAATGCGGG 3'
pcycsp	3' Reverse	NotI	5'cgacgtctgcggccgTTAATTGAATAATGCTAG GACTA3'
pcyphist-81	5' Forward	BamHI	5'tcgggatcc ATGAGTCCCTGCAACATC 3'
pcyphist-81	5' Reverse	NdeI	5'ggcgtgcatatgTGGTCGTCTACCTTACTC 3'
pcyphist-81	3' Forward	NdeI	5'gccgtgacatatgGTAGGCAGAGACATGCC 3'
pcyphist-81	3' Reverse	NotI	5'cgacgtctgcggccgcTACAATTTACTGTGTTTC TTC 3'

PcyCSP 5'	Pb 5' UTR	TgDHFR	Pb 3' UTR	PcyCSP 3'

TOPO_XL

PcyPHIST-81 5'	Pb 5' UTR	TgDHFR	Pb 3' UTR	PcyPHIST-81 3'

TOPO_XL

Figure 5.1: Schematic of the *P. cynomolgi* protein knockout constructs

The TOPO_XL vector is a backbone for these constructs. The circumsporozoite protein (CSP) is a pre-erythrocytic antigen whose expression has previously been determined to be unnecessary for propagation of the blood-stage parasite.

Prior to performing the transfections, *P. cynomolgi* parasites will be reactivated from cryopreserved stocks and inoculated into a donor *Macaca mulatta* monkey as previously described (27). Parasitized blood will be monitored and collected for transfections as previously described (28). The transfection constructs will be linearized by restriction enzyme digest and the plasmids purified by DNA ethanol precipitation. Transfections will be performed on fresh *P. cynomolgi* infected RBCs from the donor animal using the AMAXA or BioRad transfection systems, and each transfectant will be intravenously injected into a rhesus monkey. The parasitemia will be monitored by

Giemsa-stained blood smears, and selection for transfected parasites will be performed by administering pyrimethamine intramuscularly.

We anticipate that the parasites with *pcycsp* transgenes will survive and the parasitemia will persist until the monkey is treated. We predict that the parasites with *pcyphist81* transgenes, and disruption of the *pcyphist81* locus, will die off, consistent with the PHIST-81 protein being required for parasite survival. If the protein is non-essential, the parasites containing the transgenes will survive and flourish.

5.4. Exploring Caveola-vesicle Complexes as Possible Antimalarial Targets

If PcyPHIST-81 is strictly localized to the cytoplasmic face of CVC tubules, it is unlikely to be an antibody target. However, the CVC structures themselves may be suitable drug targets if it is determined that they are essential for parasite survival *in vivo*.

Caveolae are found in almost all mammalian cell types, most abundantly in endothelial cells, adipocytes, muscle cells and fibroblasts (29). Cholesterol is a necessary constituent of eukaryotic caveolar vesicles, and its depletion blocks caveola-mediated events (29, 30). Administering cholesterol-depleting drugs such as the sterol-binding antibiotics nystatin and filipin III (30-32) and the cholesterol-sequestering agent methyl-β-cyclodextrin (33, 34) to cells resulted in the flattening and disappearance of caveolar structures.

I propose that nystatin, filipin III and methyl-β-cyclodextrin may present an innovative approach to eliminating CVC morphology and disrupting the functional connections among CVC proteins in *P. cynomolgi* and *P. vivax*-infected RBCs. The effects of nystatin, filipin III and methyl-β-cyclodextrin on CVC formation, if any, can be

tested by culturing *P. cynomolgi*-infected RBCs with different concentrations of the drugs and:

1. Monitoring the parasitemia by light microscopy and recording possible reduction in Giemsa staining of CVCs, that is, changes in their typical speckling appearance, and

2. Evaluating any presumptive changes in the infected RBC membrane morphology through electron microscopy.

Prior to such experiments, these drugs could be evaluated for their possible inhibitory effect on the growth and development of the *P. cynomolgi* infected RBCs and the IC_{50} determined in this laboratory's standard drug testing assays based on ^{3H}hypoxanthine uptake (35). It would also be possible to determine if a compensatory mechanism for CVC production and/or parasite survival emerges *in vitro* by LC-MS/MS proteomic analysis to test the antigen expression profiles of drug treated versus non-treated *P. cynomolgi*-infected erythrocytes. If it is determined that CVC development is hindered by these drugs, and that this inhibition in CVC development results in parasite death, follow-up studies could be designed to investigate the possible multiple effects of these drugs on infected RBCs. Ultimately, a therapeutic intervention would need to specifically target the CVCs in parasite-infected RBCs without affecting the caveolae present on host cells.

5.5. Concluding Remarks

The data generated in this dissertation contributes to our understanding of the antigenic makeup of *P. cynomolgi* and *P. knowlesi* infected erythrocyte membranes, and

sheds light on counterpart molecules and structures in *P. vivax*. Importantly, we

confirmed, for the first time, that a *P. cynomolgi* CVC-localized protein is a member of a

recently discovered *Plasmodium*-specific family of proteins, the PHIST family (Chapter

3). The fact that a number of unique PHIST proteins, in addition to the most predominant

member PHIST-81, were localized to the membrane of infected erythrocytes suggests

that this protein family may have multiple important roles in the infected erythrocyte

membranes (Chapter 4). Each protein would have to be analyzed individually for

potential functions. Antigenic characterization coupled with the selection of interesting

proteins for further functional analyses will contribute to the pool of potential

antimalarial therapeutic targets.

The data included in this dissertation has contributed to information on the

antigenic and structural modifications produced on *P. cynomolgi* and *P. knowlesi*-infected

erythrocyte membranes by applying advanced proteomic and microscopic techniques and

utilizing the available *Plasmodium* genomes. Nevertheless, a deeper understanding of

parasite biology is still required to more fully explore the pathogenic implications of such

alterations, in addition to their necessity for parasite survival and function. Also critical

for furthering the current research is the need for a *P. cynomolgi* genome database;

sequencing efforts to achieve this goal are currently underway. It is through the study and

characterization of functions and processes unique to each *Plasmodium* species, in

addition to those shared among species, that we will be better equipped to undertake the

development of better suited drug and vaccine candidates, with the ultimate goal of

malaria eradication.

REFERENCES

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

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

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

4. Pain, A., U. Bohme, A. E. Berry, K. Mungall, R. D. Finn, A. P. Jackson, T. Mourier, J. Mistry, E. M. Pasini, M. A. Aslett, S. Balasubrammaniam, K. Borgwardt, K. Brooks, C. Carret, T. J. Carver, I. Cherevach, T. Chillingworth, T. G. Clark, M. R. Galinski, N. Hall, D. Harper, D. Harris, H. Hauser, A. Ivens, C. S. Janssen, T. Keane, N. Larke, S. Lapp, M. Marti, S. Moule, I. M. Meyer, D. Ormond, N. Peters, M. Sanders, S. Sanders, T. J. Sargeant, M. Simmonds, F. Smith, R. Squares, S. Thurston, A. R. Tivey, D. Walker, B. White, E. Zuiderwijk, C. Churcher, M. A. Quail, A. F. Cowman, C. M. Turner, M. A. Rajandream, C. H. Kocken, A. W. Thomas, C. I. Newbold, B. G. Barrell, and M. Berriman. 2008. The genome of the simian and human malaria parasite Plasmodium knowlesi. *Nature 455:799*.

5. Winzeler, E. A. 2008. Malaria research in the post-genomic era. *Nature 455:751*.

6. Barbosa, M. S., S. N. Bao, P. F. Andreotti, F. P. de Faria, M. S. Felipe, L. dos Santos Feitosa, M. J. Mendes-Giannini, and C. M. Soares. 2006. Glyceraldehyde-3-phosphate dehydrogenase of Paracoccidioides brasiliensis is a cell surface protein involved in fungal adhesion to extracellular matrix proteins and interaction with cells. *Infect Immun 74:382*.

7.   Pancholi, V., and V. A. Fischetti. 1992. A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate-dehydrogenase with multiple binding activity. *J Exp Med 176:415*.

8.   Gozalbo, D., I. Gil-Navarro, I. Azorin, J. Renau-Piqueras, J. P. Martinez, and M. L. Gil. 1998. The cell wall-associated glyceraldehyde-3-phosphate dehydrogenase of Candida albicans is also a fibronectin and laminin binding protein. *Infect Immun 66:2052*.

9.   Sirover, M. A. 1999. New insights into an old protein: the functional diversity of mammalian glyceraldehyde-3-phosphate dehydrogenase. *Biochim Biophys Acta 1432:159*.

10.   Sirover, M. A. 2005. New nuclear functions of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in mammalian cells. *J Cell Biochem 95:45*.

11.   Hara, M. R., M. B. Cascio, and A. Sawa. 2006. GAPDH as a sensor of NO stress. *Biochim Biophys Acta 1762:502*.

12.   Chuang, D. M., C. Hough, and V. V. Senatorov. 2005. Glyceraldehyde-3-phosphate dehydrogenase, apoptosis, and neurodegenerative diseases. *Annu Rev Pharmacol Toxicol 45:269*.

13.   Tarze, A., A. Deniaud, M. Le Bras, E. Maillier, D. Molle, N. Larochette, N. Zamzami, G. Jan, G. Kroemer, and C. Brenner. 2007. GAPDH, a novel regulator of the pro-apoptotic mitochondrial membrane permeabilization. *Oncogene 26:2606*.

14.   Colell, A., J. E. Ricci, S. Tait, S. Milasta, U. Maurer, L. Bouchier-Hayes, P. Fitzgerald, A. Guio-Carrion, N. J. Waterhouse, C. W. Li, B. Mari, P. Barbry, D. D. Newmeyer, H. M. Beere, and D. R. Green. 2007. GAPDH and autophagy preserve survival after apoptotic cytochrome c release in the absence of caspase activation. *Cell 129:983*.

15.   Xiao, G., T. F. Chung, H. Y. Pyun, R. E. Fine, and R. J. Johnson. 1999. KDEL proteins are found on the surface of NG108-15 cells. *Brain Res Mol Brain Res 72:121*.

16.   Alloisio, N., P. Texier, L. Denoroy, C. Berger, E. Miraglia del Giudice, S. Perrotta, A. Iolascon, F. Gilsanz, G. Berger, and J. Guichard. 1996. The cisternae decorating the red blood cell membrane in congenital dyserythropoietic anemia (type II) originate from the endoplasmic reticulum. *Blood 87:4433*.

17.   Sanders, P. R., G. T. Cantin, D. C. Greenbaum, P. R. Gilson, T. Nebl, R. L. Moritz, J. R. Yates, 3rd, A. N. Hodder, and B. S. Crabb. 2007. Identification of protein complexes in detergent-resistant membranes of Plasmodium falciparum schizonts. *Mol Biochem Parasitol 154:148*.

18.   Waller, K. L., L. M. Stubberfield, V. Dubljevic, W. Nunomura, X. An, A. J. Mason, N. Mohandas, B. M. Cooke, and R. L. Coppel. 2007. Interactions of Plasmodium falciparum erythrocyte membrane protein 3 with the red blood cell membrane skeleton. *Biochim Biophys Acta 1768:2145*.

19.   Kassenbrock, C. K., P. D. Garcia, P. Walter, and R. B. Kelly. 1988. Heavy-chain binding protein recognizes aberrant polypeptides translocated in vitro. *Nature 333:90*.

20. Tatu, U., and A. Helenius. 1997. Interactions between newly synthesized glycoproteins, calnexin and a network of resident chaperones in the endoplasmic reticulum. *J Cell Biol 136:555*.

21. de Koning-Ward, T. F., P. R. Gilson, J. A. Boddey, M. Rug, B. J. Smith, A. T. Papenfuss, P. R. Sanders, R. J. Lundie, A. G. Maier, A. F. Cowman, and B. S. Crabb. 2009. A newly discovered protein export machine in malaria parasites. *Nature 459:945*.

22. Tilley, L., G. McFadden, A. Cowman, and N. Klonis. 2007. Illuminating Plasmodium falciparum-infected red blood cells. *Trends Parasitol 23:268*.

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

24. Kocken, C. H., H. Ozwara, A. van der Wel, A. L. Beetsma, J. M. Mwenda, and A. W. Thomas. 2002. Plasmodium knowlesi provides a rapid in vitro and in vivo transfection system that enables double-crossover gene knockout studies. *Infect Immun 70:655*.

25. Pfahler, J. M., M. R. Galinski, J. W. Barnwell, and M. Lanzer. 2006. Transient transfection of Plasmodium vivax blood stage parasites. *Mol Biochem Parasitol 149:99*.

26. Kocken, C. H., A. van der Wel, and A. W. Thomas. 1999. Plasmodium cynomolgi: transfection of blood-stage parasites using heterologous DNA constructs. *Exp Parasitol 93:58*.

27. Galinski, M. R., C. C. Medina, P. Ingravallo, and J. W. Barnwell. 1992. A reticulocyte-binding protein complex of Plasmodium vivax merozoites. *Cell 69:1213*.

28. Galinski, M. R., P. Ingravallo, C. Corredor-Medina, B. Al-Khedery, M. Povoa, and J. W. Barnwell. 2001. Plasmodium vivax merozoite surface proteins-3beta and-3gamma share structural similarities with P. vivax merozoite surface protein-3alpha and define a new gene family. *Mol Biochem Parasitol 115:41*.

29. Parton, R. G., M. Hanzal-Bayer, and J. F. Hancock. 2006. Biogenesis of caveolae: a structural model for caveolin-induced domain formation. *J Cell Sci 119:787*.

30. Smart, E. J., and R. G. Anderson. 2002. Alterations in membrane cholesterol that affect structure and function of caveolae. *Methods Enzymol 353:131*.

31. Lisanti, M. P., Z. L. Tang, and M. Sargiacomo. 1993. Caveolin forms a hetero-oligomeric protein complex that interacts with an apical GPI-linked protein: implications for the biogenesis of caveolae. *J Cell Biol 123:595*.

32. Rothberg, K. G., Y. S. Ying, B. A. Kamen, and R. G. Anderson. 1990. Cholesterol controls the clustering of the glycophospholipid-anchored membrane receptor for 5-methyltetrahydrofolate. *J Cell Biol 111:2931*.

33. Rodal, S. K., G. Skretting, O. Garred, F. Vilhardt, B. van Deurs, and K. Sandvig. 1999. Extraction of cholesterol with methyl-beta-cyclodextrin perturbs formation of clathrin-coated endocytic vesicles. *Mol Biol Cell 10:961*.

34. Subtil, A., I. Gaidarov, K. Kobylarz, M. A. Lampson, J. H. Keen, and T. E. McGraw. 1999. Acute cholesterol depletion inhibits clathrin-coated pit budding. *Proc Natl Acad Sci U S A 96:6775*.

35.   Webster, H. K., E. F. Boudreau, K. Pavanand, K. Yongvanitchit, and L. W. Pang. 1985. Antimalarial drug susceptibility testing of Plasmodium falciparum in Thailand using a microdilution radioisotope method. *Am J Trop Med Hyg 34:228*.

Acknowledgements

My sincerest appreciation to Dr. Mary R. Galinski for supporting my graduate research, to members of the Galinski and Moreno labs for encouragement and scientific support and to our collaborators in Colombia and Australia who have been influential in my scientific development.

Special thanks to members of my dissertation committee.
I would also like to acknowledge my family and friends for their love and support.
I dedicate this dissertation to the memory of my mother Ms. Grace A. Obare.