IDENTIFICATION OF SULFATED POLYSACCHARIDES WITH ACTIVITY AGAINST *Plasmodium falciparum* CYTOADHERENCE PROTEINS USING TARGETED SEQUENCING AND *IN SILICO* CHARACTERISATION

JENNIFER MUMBE MUTISYA

MASTER OF SCIENCE

(Molecular Biology and Bioinformatics)

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

2022

Identification of Sulfated Polysaccharides with Activity against *Plasmodium falciparum* Cytoadherence Proteins using Targeted Sequencing and *In Silico* Characterisation

Jennifer Mumbe Mutisya

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Molecular Biology and Bioinformatics of the Jomo Kenyatta University of Agriculture and Technology

2022

DECLARATION

This thesis is my original work and has not been presented for a degree in any other university.

Signature: ……. Date: …….

Jennifer Mumbe Mutisya

This thesis has been presented for examination with our approval as University supervisors.

Signature: ……. Date: …….

Prof. Johnson Kinyua, PhD
JKUAT, Kenya

Signature: ……...……. Date: ……

Dr. Victor Mobegi, PhD
UoN, Kenya

Signature: ……. Date: …….

Dr. Ben Andagalu
USAMRD-MDR, Kenya

DEDICATION

Dedicated to my mother Beatrice and my father Titus for being present, supportive and endlessly loving. And to my son, may you read my creativity one day.

ACKNOWLEDGEMENTS

I would like to acknowledge Developing Excellence in Leadership and Genetics Training for Malaria Elimination in Sub-Saharan Africa (DELGEME) and Department of Emerging and Infectious Diseases (DEID) of the United States Army Medical Research Directorate-Africa (USAMRD-A) for funding this research project. I also acknowledge the support accorded by The Malaria Drug Resistance Laboratory-Kisumu (MDR) team through the leadership of Dr. Andagalu and Dr. Akala for helping me carry out all my molecular assays as well as training me thoroughly to be an experienced scientific researcher. I acknowledge the efforts and support offered by my supervisors; Dr. Andagalu for being very supportive in growing my career through the application of grants and molecular experiments guidance, Dr. Mobegi for guiding me through proposal development, experimental designs and research execution and Dr. Kinyua for helping me meet all academic requirements during the period of this research. To my friends with whom we have walked this path together, thank you for the gift of friendship and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

ACTs	Artemisinin-based combination therapies
AMA	Apical Membrane Antigen
CEBIB	Centre for Biotechnology and Bioinformatics
DBLMSP2	Duffy binding-like domain merozoite surface protein 2
DNA	Deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphate
EBA	Erythrocyte Binding antigens
ELM	Erythrocyte Linear Motif
ExPASy	Expert Protein Analysis System
JKUAT	Jomo Kenyatta University of Agriculture and Technology
KEMRI-USAMRD/A	Kenya Medical Research Institute-United Stated Army Medical Research Directorate- Africa
LAH	Low anticoagulant heparin
MAST	Motif Alignment Search Tool
MEME	Multiple EM for Motif Elicitation
$MgCL_2$	Magnesium Chloride

MSP	Merozoite Surface Protein
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PDBQT	Protein Data Bank Partial Charge and Atom Type
PEXEL	*Plasmodium* Export Element
***Pf*Rh**	*P. falciparum* reticulocyte-binding homologs
PHISTb/RLP1	*Plasmodium* helical interspersed Sub telomeric domain b with RESA like protein.
PRESAN	*Plasmodium* RESA N-terminal
RBCs	Red Blood Cells
RBL	Reticulocyte binding-like
RESA	Ring infected erythrocyte surface antigen
RMSD	Root Mean Square Deviation
SDF	Standard Database Format
SNP	Single Nucleotide Polymorphism
3D	Three Dimension
TAE	Tris Acetate

TOMTOM	Motif Comparison Tool
UoN	University of Nairobi
UV	Ultraviolet
WHO	World Health Organization

ABSTRACT

Malaria is a global health problem that remains a leading cause of death especially in Sub-Saharan Africa. *Plasmodium* parasites in the human host undergo different stages with survival tactics to evade immune system clearance. One crafty mechanism of survival is cytoadherence, through which the parasite attached to human erythrocytes thus avoids splenic clearance. Cytoadherence is mediated by Merozoite surface and exported proteins. Diversity exhibited by these proteins is one of the main challenges in developing drugs targeting intracellular parasites. This study aimed at identifying sulfated polysaccharides with ability to block these proteins by investigating two drug targets as representatives: Merozoite surface protein with Duffy Binding-Like Domain (DBLMSP2) and *Plasmodium*-Helical Interspersed Subtelomeric domain b containing Ring-infected Erythrocyte Antigen (RESA) Like Protein (PHISTb/RLP1). Sanger sequencing was used to obtain sequences of these drug targets from whole blood samples collected from malaria-endemic zones in Kenya. The sequencing employed novel primers designed for the target genes from Primer3 plus. Reference sequences were obtained from Plasmodb database from 3d7 strain and used for genome mapping as well as structure comparisons. Functional Motifs were predicted using MEME software with default settings for consensus output. Protein structures were modelled using I-TASSER and LOMETs tools and validated using Galaxy Refine and RAMPAGE tools. Sulfated polysaccharides were screened from the Pub-Chem Compounds database. The criteria of selection were guided by the Lipinski rule of five for a drug compound. Interaction simulations were achieved using auto-dock vina. The docking results were visualized in PyMOL. Sequencing results were successful for the PHISTb/RLP1 gene but not for DBLMSP2. 157-point mutations were identified in PHISTb/RLP1 sequences and 12 of these single out for protein structure analysis. Protein structure prediction yielded high quality models passing >90% ramachandran score. PHISTb/RLP1 mutant structure folded differently compared to wild type structure. Eleven sulfated polysaccharides fulfilled the criteria of drug compounds. Alpha carrageenan, Amylopectin sulfate, cyclodextrine sulfate, ghatti sulfate, fucoidan and galactal interacted optimally with both drug targets after docking. Protein-Ligand interactions were different in wildtype PHISTb/RLP1 compared to wildtype inferring effects of mutations on drug binding. The study identifies lead compounds that can be added to malaria drug molecules pipeline. *In vitro* testing is recommended to obtain more data on inhibition capacity of the compounds. The study design has proven effective for initial stages of drug discovery and is recommended for other diseases as well.

CHAPTER ONE

INTRODUCTION

1.1 Background information

Malaria remains a serious health problem in African countries (WHO, 2017). The pathogenesis of the disease occurs during blood stages when merozoites invade erythrocytes. Different proteins located on merozoites initiate invasion steps on the RBCs. An example of these ligands are the merozoite surface proteins (MSPs), which initiate the RBC invasion process by binding to their receptors on these host cells. Other proteins released to strengthen the merozoite-erythrocyte bridge are the Erythrocyte-binding antigens (EBAs) and *P. falciparum* reticulocyte-binding homologs (PfRh) (Dobaño et al., 2007). The merozoites movement into the newly formed RBCs at this stage is facilitated by an actin-myosin–like motor (Cowman *et al.*, 2006). The invasion process is mediated by specific protein-protein interactions. Moreover, there are many proteins found on the merozoite surface whose functions are not determined (Duffy et al., 2014). Many of such proteins are thought to be involved in pathogenesis. Lack of knowledge in these parasite development areas has slowed down therapeutic innovations to block these parasite pathways of survival.

Research has reported that sulfated polysaccharide compounds interfere with the mechanisms of merozoite to RBCs (Veena et al., 2007). These molecules inhibit erythrocyte invasion by interacting with essential proteins that facilitate this process, such as MSP1-42 and MSP1-33 (Boyle *et al.,* 2010). Further research findings on heparin inhibition mechanisms against merozoite proteins have suggested the use of modified heparin-like molecules as receptors for invasion ligands (Boyle *et al.,* 2010). Clinical trials of heparin-derived compounds containing anti-plasmodial activity have been carried out with time. For example, Sevuparin, a heparin drug derivative, has been used to treat severe malaria, although it had a risk of causing over bleeding (Leitgeb *et al.,* 2011). The drug activity disrupted rosette formation and inhibited cytoadherence to endothelial cells

(Saiwaew *et al.,* 2017). It is paramount to continue investigating modified heparin-derived compounds that have improved efficacy and identify more amenable drug targets involved in *P. falciparum* parasite's cytoadherence mechanism to develop chemical compounds that can inactivate this parasite's mechanism and terminate asexual stages proliferation.

Plasmodium falciparum parasite uses the mechanism of cytoadherence to enhance its virulence in the human host (Lee *et al.,* 2019). Through this mechanism, the parasite can attach to uninfected erythrocytes and the inner layers of blood vessels and internal organs as a survival tactic in the body (Saiwaew *et al.,* 2017). DBLMSP2 was reported to enhance invasion by binding to its receptors on RBCs during the initial stages of merozoite invasion. This protein was established to interact with antibodies, including human immunoglobulins (Crosnier *et al.,* 2016). The DBL domain present in the DBLMSP genes was previously reported to exhibit extreme protein polymorphism (Hodder *et al.,* 2012). The genetic variability exhibited by these genes was the motivation for this study to understand the mutation pattern of these proteins in Kenyan parasites. The exported proteins induce morphological changes of the host cell such as giving the cells abilities to attach to other cells and blood vessels (Lee et al., 2019). These changes enable the malaria parasite to thrive in the human body, evade clearance by the spleen and increase parasitemia (Boddey *et al.,* 2016). Exported proteins include *Plasmodium* Helical Interspersed Subtelomeric (PHIST) family, which contains 89 proteins. These proteins give the remodelled host RBC more cytoadherence to vasculature and uninfected erythrocytes (Warncke *et al.,* 2016). An essential exported protein, PHISTb domain-containing Ring-infected Erythrocyte Surface Antigen (RESA) Like protein, (PHISTb/RLP1), is another protein that is associated with cytoadherence of *P. falciparum* (Moreira *et al.,* 2016). The protein controls cytoadherence by binding to the domain of *P. falciparum's* EMP1, a major virulence factor of the parasite located on the cytoskeleton of erythrocytes (Oberli *et al.,* 2016; Goel *et al.,* 2014).

The current study reports findings from investigation of sulfated polysaccharides interactions with *Pf*DBLMSP2 and *Pf*PHISTb/RLP1. These include identified point mutations, protein structure analysis, database search results and docking energies from protein-ligand analysis.

1.2 Problem Statement

Malaria continues to claim human lives in Africa despite the various methods devised to control the disease. The treatment of the disease is equally expensive in Africa, with countries putting a lot of money into the purchase of antimalarial drugs. The use of certain drugs such as chloroquine and sulfadoxine-Pyrimethamine (Fansidar®) has been terminated in malaria control due to *P. falciparum* resistance to these therapeutic compounds. There is reported parasite resistance against the currently used first line drugs, Artemisinin Combined Therapies. This threatens control of malaria and calls for urgent studies on discovery of new molecules. Cytoadherence is enhanced by many protein-protein interactions of the parasite and human erythrocytes. The proteins enhancing this survival mechanism of the parasite exhibit high genetic variability and there is need of current sequence analysis to understand mutation patterns. In the *Plasmodium* database, the target proteins of this study; DBLMSP2 and PHISTb/RLP1 gene ontology information on crystal structure and current Single Nucleotide Polymorphisms track data is not available. This is a challenge in drug discovery studies of targeted therapies against *Plasmodium* survival pathways that calls for urgent intervention.

1.3 Justification

The onset of Artemisinin resistance calls for the identification of new compounds for malaria treatment to guarantee continued control of the disease. Sulfated polysaccharides are reported to have active antiplasmodial activity through biochemical studies on heparin modifications to generate compounds with pre-determined antimalarial properties. The prior studies have not recorded any drug resistance against these compounds. The identified compounds have not been tested on parasites in many parts of Africa, including Kenya, where malaria transmission is high. This study fills this gap by testing these

compounds on *P. falciparum*'s targeted virulent proteins from Kenyan samples. The use of the *in silico* approach in drug discovery has proven to be an efficient, affordable, and faster method than traditional drug discovery methods; which are expensive and take longer. This study adopted this methodology for efficiency in data collection.

1.4 Hypothesis

This study hypothesizes that sulfated polysaccharides contain antimalarial activity and interact with cytoadherent proteins of *P. falciparum* therefore they are possible drug compounds inhibiting cytoadherence. The null hypothesis is sulfated polysaccharides do not contain activity against *Pf*DBLMSP2 not *Pf*PHISTb/RLP1.

1.5 Objectives

1.5.1 General objective

To investigate the interactions of *Pf*DBLMSP2 and *Pf*PHISTb/RLP1 proteins with sulfated polysaccharides.

1.5.2 Specific objectives

i. To analyze *Pf*DBLMSP2 and *Pf*PHISTb/RLP1 genetic variations from whole blood samples of uncomplicated malaria patients.

ii. To predict functional motifs within *Pf*DBLMSP2 and *Pf*PHISTb/RLP1 proteins that interact with sulfated polysaccharide compounds.

iii. To determine the effect of mutations in DBLMSP2 and PHISTb/RLP1 proteins on 3D structures and interactions with the sulfated polysaccharide ligands.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Malaria ranks high on the list of life-threatening diseases in the world, according to WHO, The Road to Malaria Eradication report of 2018 (World Health Organization, 2018). The disease is common in tropical and subtropical regions including South America, Africa, Eastern Mediterranean region and oceanic regions (Autino et al., 2012). African countries have the highest burden of malaria with 91% of global deaths. Five *Plasmodium* species contribute to human malaria, these include *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale* and *P. knowlesi. P. falciparum* is the most common in Africa (Hay et al., 2008). *P. vivax* dominates in South America over other species. South Eastern Asia and the Mediterranean region are dominated by both *P. falciparum* and *P. vivax*. *P. malariae* is common in all regions but generally has low transmission rates. *P. Ovale* is common in tropical Africa and *P. knowlesi* is found mostly in forested regions of South Eastern Asia (Sullivan, 2010). Currently, there is no competent vaccine for malaria. Available vaccines provide partial immunity to infants according to the clinical trial studies carried out (Wykes, 2013). Intervention measures such as insect treated nets and residual spraying are employed to overcome transmission of malaria by blocking and killing the main vector, the female *Anopheles* mosquito (Maia *et al.,* 2019).

2.2 Life cycle of *P. falciparum*

Among the five listed species of *Plasmodium* parasite that transmit malaria to humans, *P. falciparum* and *P. vivax* are the most severe species. *P. vivax* is most dominant outside Sub-Saharan countries (Baird et al., 2012). *P. falciparum* accounts for the most malaria cases in African countries ("*Plasmodium falciparum* mortality in Africa between 1990 and 2015," 2017). Transmission to a human host occurs when the female *Anopheles* mosquito sucks a blood meal from a human (Lafferty, 2016,). During a blood meal, sporozoites from the midgut of the mosquito enter the circulation system of the victim and

are transferred to the liver. In the hepatocytes, they develop into schizoint that are released into blood circulation and invade RBCs in which they transform into merozoites (Joana *et al.*, 2016). The merozoites develop to trophozoites, which is the stage that leads to clinical manifestations of the disease. (Figure 2.1)

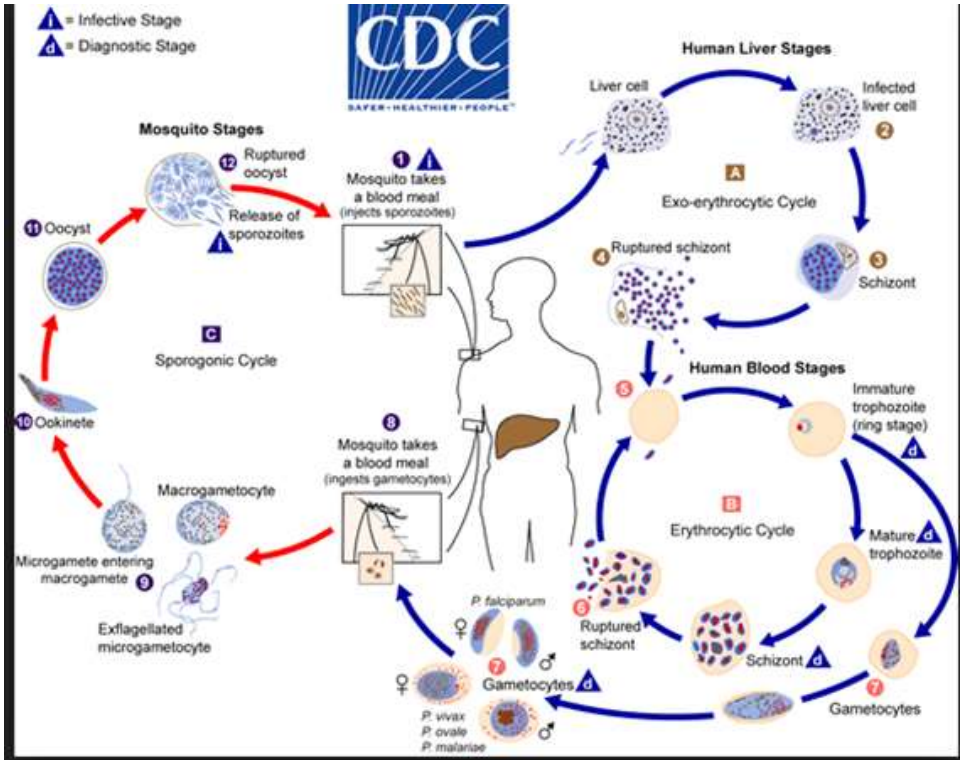

Figure 2.1: The figure shows the life cycle of the *P. falciparum* in mosquito and human hosts. The different sexual and asexual stages are highlighted. Image used is by the Centre for Disease Control and Prevention (CDC, 2020).

2.3 Epidemiology of Malaria

Malaria is reported to be more common in tropical and subtropical regions of Africa, Asia, Central and South America (Phillips et al., 2017). As previously mentioned, over 90% of clinical Malaria transmission occurs in Sub-Saharan Africa. Regions characterized by highland and desert climatic conditions generally have low Malaria transmission. As previously mentioned, *P. falciparum* is the most infectious species in Africa and

SouthEast Asia (Gething et al., 2016). Immunity against Malaria is naturally acquired in high transmission zones. Exposure is likely to occur in early childhood and as one grows up the severity of the disease lessens due to buildup of immunity (Fowkes et al., 2016,). In Africa, Malaria extreme cases and deaths largely affect children under five years and pregnant mothers (Maitland, 2016). In high endemic zones, Malaria manifests more as anemia however the incidence of cerebral Malaria increases as well. In Sub-Saharan Africa, Malaria and HIV/AIDs were reported to occur in high frequencies concurrently. Severity of Malaria was reported to be high in individuals infected with HIV compared to HIV negative individuals (Kwenti, 2018). Intervention measures to reduce the risk of infection such as insect treated nets and medicinal drugs have been put in place to curb the diseases. However, Malaria treatment has faced dire challenges due to mutations that enable the parasite to survive despite effective antimalarial drug administration (Hughes, 2004).

2.4 Treatment of Malaria

The treatment of malaria depends on age, severity of the disease, efficacy, cost and availability of the drugs. Available antimalarials include Chloroquine, ACTs, Quinine sulfate, Atovaquone-proquianil and Primaquine phosphate (Arrow *et al.*, 2004; Tang et al., 2020). Quinolines which include Chloroquine and structurally related drugs act by interfering with the crucial hemozoin formation process of the malaria parasites (Müller & Hyde, 2010). The usage of these drugs was stopped due to emergence and spread of parasite drug resistance (Sidhu et al., 2002). The current drug for Malaria treatment recommended by WHO is ACTs (WHO, 2018). They are used widely to control malaria in African countries. Artemisinin resistance was reported in *P. falciparum,* posing challenges to the control of malaria. The prevalence of the resistance has now spread across the mainland in Southeast Asia. The resistance was associated with development of mutations in the kelch13 gene of the *P. falciparum* (Uwimana *et al.*, 2020)

The cost of treating malaria is generally high for the communities in high transmission zones of Malaria in Africa who cannot afford to purchase the ACTs at market price (Alonso *et al.*, 2019). The high cost of malaria treatment has fallen back to the health

ministries in African countries, with governments incurring a lot of expenses to purchase the malaria drugs for its people to have access to them at their local clinics. The expensive cost of malaria treatment requires research on new Malaria drug compounds that can overcome resistance and reduce treatment costs (Arrow *et al.,* 2004). Research on new drug compounds is the only way to keep malaria under control. There is an urgent need to widen the scope of drug discovery in malaria through new methodologies and identification of antiplasmodial molecules.

2.5 Sulfated Polysaccharides

Heparin and other low anticoagulant sulfated polysaccharides were reported to have antimalarial activity (Boyle *et al.,* 2010). Research on heparin showed that the compound could enter erythrocytes just before merozoites egress and interfere with the attachment of infected erythrocytes to various host receptors and disrupt *P. falciparum* rosettes (Glushakova *et al.,* 2017). Inside the RBCs, the parasites perforate the enclosing erythrocyte membrane shortly before egress. The perforated membrane forms a pathway for entry of these hydrophilic compounds to trap the merozoites inside the infected erythrocyte and prevent invasion into other cells. Heparin-cell interaction revealed that the compound interacted with merozoite surface proteins interfering with parasite dispersion. Heparin was also reported to bind the inner erythrocyte membrane preventing it from rupturing (Boyle *et al.,* 2010). These heparin interactions with intracellular parasites convinced researchers that it could be used as a drug compound to treat severe malaria by the mechanism of interfering with parasitemia increase.

One reported source of sulfated polysaccharides was extracts from marine plants such as sea cucumbers, red algae and the marine sponge. Compounds from these marine plants were reported to have antimalarial activity by inhibiting the growth of *P. falciparum* (Marques *et al.,* 2016). Another source of these compounds was tropical plants. These plant species include *Gossipium hirsutum, Helianthus tuberosus,* and *Hedysarum polybotrys*. Chemical modifications such as addition of sulfate groups to sulfated polysaccharides extracts from these plants was reported to reduce anticoagulation

properties of these chemicals, increasing their safety as drug compounds (Oliveira *et al.*, 2016).

The pathways targeted by sulfated polysaccharides compounds in *Plasmodium* parasites are sialic dependent and independent (Kobayashi *et al.,* 2013). In *P. falciparum*, sialic independent pathway was reported to be utilized during RBCs invasion by the merozoites. These invasion processes were characterized by multiple interactions between ligands on the merozoites and their receptors on human RBCs (Rosalynn *et al*., 2012). Antimalarial compounds containing heparin, a sulfated polysaccharide, were developed to target these interactions and successfully block them (Svetlana *et al.,* 2017). In a study conducted to investigate the effect of these compounds on inhibiting merozoites from infecting human RBCs *in vitro,* fucoidans and dextran sulfates were recorded to bind to the Duffy antigens on the merozoites and preventing them from invading human erythrocytes (John *et al*., 1991).

The drug activity of sulfated polysaccharides was further investigated in clinical trials. Heparin, for instance, was used as a drug to treat severe malaria cases (Lantero et al., 2020). The compound, however, would cause excessive bleeding in patients due to its anticoagulant activity (Lantero et al., 2020). These resulted in the clinical trials' termination and encouragement of more research to further modify the compound by reducing its anticoagulant activity to make it safe for use. In Africa, the use of low anticoagulant heparin (LAH) was studied in Cameroon and found to be effective and safe in disrupting rosettes in fresh isolates from children suffering from malaria (Leitgeb *et al*., 2011).

Modified heparin derivatives were identified and successfully reported to have reduced anticoagulant activity and have high *P. falciparum* anti-proliferation properties (Boyle et al., 2017). One example of such a compound was curdlan sulfate, which was tested in *in vivo* studies and found to reduce the severity of malaria (Havlik *et al.,* 2005). Studies of improving the structure and invasion properties of heparin-like molecules suggested that minimizing the monosaccharide chain to six monomers and reducing the sulfate groups

was essential for optimal inhibitory activity. Other suggested modifications of polysaccharides were esterifications (Boyle *et al.,* 2017). There were several identified polysaccharide inhibitors of *P. falciparum* with low anticoagulant activities. These included carrageenan compounds, whose source was sea algae, and other compounds identified by modifying their chemical properties (Boyle *et al.,* 2010). Sulfated polysaccharides were investigated further and reported to have other therapeutic properties other than antimalarial properties. These included antiviral activity by preventing replication of enveloped viruses, for example, *Herpes simplex* virus, human immunodeficiency virus, Dengue virus, and respiratory syncytial viruses (Damonte *et al.,* 2004; Luescher-Mattli, 2003; Witvrouw *et al.*, 1997; Schaeffer *et al.,* 2000;). Carrageenans from red seaweed were reported in a research article to inhibit dengue virus proliferation *in vitro* (Talarico *et al.*, 2007). Sulfated polysaccharides were reported to have therapeutic activity by treating diseases attacking other animals, for example, the use of extracted seaweeds to protect chicken embryos from mumps and influenza B viruses (Wei et al., 2022). Other reported activities of these compounds include antithrombotic activities through activating thrombin inhibitors and antioxidant activities (Guangling *et al.,* 2011).

2.6 *P. falciparum* proteins interactions with heparin modified compounds

Heparin-like molecules were reported to interact with surface merozoite proteins involved in RBCs invasion processes such as MSP1 (Boyle *et al.,* 2010). The rhoptry and microneme protein families were investigated and reported to play key roles in reorientation and signaling steps of invasion that trigger strong deformation of the RBCs (Sherling *et al.*, 2019). Heparin was reported to interact with these proteins and block their functions in *P. falciparum*. Duffy Binding Like (DBL) and reticulocyte binding-like (RBL) family of proteins have a high affinity for heparin molecules. Heparin was also reported to inhibit PfRH2 and PfRH4 (Tham et al., 2012). All these ligands mediated the sialic-dependent and independent pathways (Tham *et al.*, 2012). DBL proteins localized in microneme were secreted onto the merozoite surface during merozoite egress interacting with their specific receptors during the invasion (Beeson et al., 2016). It was

previously reported that blocking DBL domains would interfere with the protein-protein interactions (Kobayashi *et al.,* 2013). In this research, I report interactions of sulfated polysaccharides with DBLMSP2 and RESA-like genes. DBLMSP2 was reported to bind to ligands of human RBCs and enhance merozoite invasion (Crosnier et al., 2016). The PHIST family of protein was shown to be expressed across the sexual stages of *P. falciparum,* and some protein members of this family were involved in controlling the cytoadherence process (Moreira et al., 2016). The PHIST domain clusterered into three subgroups across the *Plasmodium* genus, PHISTa, PHISTb, and PHISTc (Parish et al., 2013). The proteins containing these domains were reported to be expressed in all *Plasmodium* species with major parasite survival tactics associated with them. The PHISTb subgroup was reported to expand in *P. falciparum* with a specific subgroup that contains RESA and the DnaJ domains (Sargeant *et al.,* 2006a). PHISTb sub-class of PHIST localized to the periphery of the RBCs (Tarr et al., 2014). These proteins contained the PRESAN domain, which gives the RESA a specific function of binding the cytoskeleton of the erythrocytes (Moreira *et al.,* 2016). The PRESAN domain in PHISTb proteins' role was reported to be transfer the proteins to the periphery of the RBCs where they are needed. PHISTb protein export mechanism was enhanced by the cleaving of the *Plasmodium* export element (PEXEL) motif by Plasmepsin V (Boddey *et al.*, 2016). The PEXEL motif was shown to be a short sequence of amino acids conserved all through the *Plasmodium* genus and has allowed for the identification of these exported proteins. (Sargeant *et al.*, 2006a). The PHISTb and RESA-like genes were suggested to be involved in multiple functions that lead to erythrocyte remodeling, essential for the survival of intracellular parasites (Moreira *et al.,* 2016).

This study fills in the gap of research in finding more molecules with antimalarial properties that are added to the pipeline of screening therapeutics that can overcome drug resistance and improve drug efficacy. It reports analysis into PHISTb/RLP1 and DBLMSP2 genes which play key roles in cytoadherence and gives chemical inhibitors for these protein families. This is novel results of sequencing primers, mutations and protein-ligand interactions results reported in this work for the first time.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Study Settings, ethical approval and sample collection

This study was carried out in Kisumu Kenya at the United States Army Medical Research Directorate-Africa (USAMRD-A), The Malaria Drug Resistance (MDR) Laboratory. The laboratory is located at The Kenya Medical Research Institute-Center for Global Health Research (KEMRI-CGHR). The study analyzed samples that were collected for an ongoing project on epidemiology of malaria and drug resistance patterns in MDR laboratory and archived in the laboratory. The epidemiology project started in 2008 and recruited study participants from five different malaria hotspots in Kenya with distinct transmission patterns. The ethical guidelines of the study were assessed and approved by review committee members from KEMRI and Walter Reed Army Institute of Research (WRAIR). The approved protocols were KEMRI #1330/WRAIR #1384 and KEMRI #3628/WRAIR #2454. These were the ethical approval for the main epidemiology of Malaria project from which my study used a subset of collected samples. The sample sites were selected according to differential distribution of Malaria endemic zones in Kenya. These included lake region endemic zones, Kisumu District Hospital (KDH) and Kombewa (KOM). Kisii (KSI) and Kericho (KCH) were selected as highland epidemic areas as well as Marigat (MGT) as seasonal malaria transmission zones. Malindi (MDH) was also included in the study to represent coastal areas with declining endemicity. The umbrella epidemiology study sample collection guidelines were such that samples were collected at 4 time points. Day 0, 7, 14, 21 and 28 according to severity of the disease and response of the patient to treatment. The subset of the samples used in this current study were from one time point, which was day zero. Day zero meant the first sample collection from a study participant before any Malaria treatment.

3.2 Criteria of inclusion and exclusion

The age of study participants was from six months and above. Minors were represented by their parent or guardian. Individuals showing symptoms of malaria and/or testing positive for uncomplicated malaria by rapid diagnostic test (mRDT; Parascreen® (Pan/Pf), Zephyr Biomedicals, Verna Goa, India) were recruited into the study after providing written informed consent or assent. This was done at the out-patient departments of general hospitals in each study site.

Volunteers unable to complete the questionnaire or participate in the blood draw due to language barrier, religious beliefs, unavailable consenting guardian or other individual reasons were excluded from the study.

3.3 Sample size calculation

The sample size for this study was determined using Cochran's formula 1963 (Duru *et al*., 2016).

The formula states that $n = Z^2pq/e^2$

Where:

- Z is the critical value for a 95% confidence level (1.96)
- e is the desired level of precision (margin of error) given as 5%
- p is the (estimated) proportion of the population which has the attribute in question (20%)
- q is 1 – p.

Sample size of 300 was used in this study; distribution was evenly, 50 samples per site. After removing sample vials which were depleted, a total of 251 samples were fit to be used for analysis.

3.3 Study Design

This study was experimental based research. The aim was to study how sulfated polysaccharides would interact with the targeted proteins *Pf*DBLMSP2 and *Pf*PHISTbRLP1 from *P. falciparum* parasite sequences from Kenyan population. The sequences would represent Kenyan *Plasmodium falciparum* variability and individual site variation was not a key objective in this present work, even though it is an important measure of genetic diversity study designs.

3.3 Molecular biology assays

3.3.1 Extraction of *Plasmodium* parasite DNA using Qiagen Kit

The parasite DNA from whole blood samples was extracted using the Qiagen DNA Mini extraction spin protocol (Qiagen, Valencia, CA) according to the given protocol. In the first step, microcentrifuge tubes were taken in which 20µl of proteinase K was transferred followed by 200µl of whole blood into each tube. To break down the cells, the next step was addition of 200µl of lysis buffer to the tubes which were then vortexed for about fifteen seconds for complete mixing of the contents. The tubes were incubated for ten minutes on a heat block at a temperature of 56°C. The tubes were then centrifuged at 8000 revolutions per minute (RPM) to remove any droplets inside the lids. Afterwards, 200µl of absolute ethanol (96-100%) was added for DNA precipitation purposes. The tubes were vortexed and centrifuged. After this step, the mixture was carefully transferred to QIAamp mini spin columns without wetting the rim. The columns were centrifuged for one minute at 8000 RPM. The DNA was trapped inside the mini columns silica membrane while the filtrate was discarded. The QIAamp spin columns were transferred into fresh collecting tubes. 500µl wash one buffer was added and the tubes centrifuged for one minute at 8000 RPM. This step was repeated again by adding 500µl of wash two buffer and centrifuged at maximum speed 14,000 RPM for three minutes. Following the washing steps, the last step was eluting the purified DNA. The columns were transferred into clean 2 ml collecting tubes, 200µl of elution buffer was added into the columns and allowed to

incubate for five minutes at room temperature. The tubes were centrifuged for one minute at 8000 RPM. The extracted DNA was stored in the freezer at -20^{o}C.

3.3.2 *Plasmodium falciparum* speciation by Realtime PCR

According to the MDR laboratory workflow, after DNA extraction the samples were first tested for the *Plasmodium* genus using primers for amplification of parasite 18S rRNA genes in an assay developed by (Kamau *et al., 2011*). This step was to eliminate false positive samples from RDT test kits. The samples that turned out positive for *Plasmodium* were tested for *P. falciparum* species of interest. The primers and probes used in the qPCR assay were from previous studies (Kamau *et al.,* 2013; Veron *et al.,* 2009). The *Plasmodium* genus test was called PLU assay and *P. falciparum speciation* experiment was referred to as the FAL assay. The human RnaseP gene was used as control for negative samples. The probes used were labelled with 6-carboxyfluorescein (FAM) in 5' and at the 3' carboxytetramethyl-rhodamine. The reaction mixture amounts were calculated according to the number of samples run per assay. Quantifast master mix was used in the assay and 2µl of the DNA template was added to each well with the reaction mixture. The real-time PCR assay was carried out in Applied Biosystem Quantstudio 6 Flex (Quantstudio^{Tm} real-time PCR applied Biosystems by Thermos Fisher Scientific). The cycling condition were optimized in a previous study carried out at the laboratory (Kamau *et al.,* 2013). A cycle threshold of at least 32 in each sample were considered positive for *P. falciparum* species.

3.3.3 *Pf*DBLMSP2 and *Pf*PHISTb/RLP1 Gene-specific Primer design

Gene-specific primers were designed using Primer3 plus (Untergasser *et al*. 2012). The primers were designed using reference sequences from *P. falciparum* 3D7 strain PlasmoDB sequence entry; *Pf*DBLMSP2 Gene ID; PF3D7_1036300 2289 base pairs (Bahl, 2003). Three pairs of overlapping primers were designed for this gene to amplify the gene in three fragments, as shown in table 3.1. The designing was repeated for *Pf*PHISTb/RLP1 gene using *P. falciparum* reference sequence; Gene ID PF3D7_0201600

1558 base pairs long (Bahl, 2003). Two pairs of primers were designed to amplify the gene in two different fragments.

Table 3.1: DBLMSP2 and PHISTb/RLP1 gene-specific primers designed in Primer3 plus tool.

Gene Name	Target Region	Primers	Product Size (bp)
*Pf*DBLMSP2	80-996	>SP2_F1 5'-TTGTTGTAAATGAGGGGAAC-3' >SP2_R1 5'-GTTATCTTTCTGCGCACTCT-3'	917
*Pf*DBLMSP2	927-1659	>SP2_F2 5'-GTGGACAGAAAACAGGCATC-3' >SP2_R2 5'-GCTCACGCTACCTTGGTTTA-3'	733
*Pf*DBLMSP2	1495-2249	>SP2_F3 5'-ACACAACAGGAAAATCAACC-3' >SP2_R3 5'-TTCAAGGTAGAATCACTTCCA-3'	755
*Pf*PHISTb/RLP1	80-998	>RLP1_F1 5'-CAATGACCATGTGTTCAGA-3' >RLP1_R1 5'-TCTGGAGATAAAGCATACCC-3'	919
*Pf*PHISTb/RLP1	791-1450	>RLP1_F2 5'-GCAAAACCAACGTGTAATC-3' >RLP1_R2 5'-TTCAACATCTCTTTCGACTG-3'	660

3.3.4 Mapping primers to sequence

Primer mapping on *Pf*3D7 DBLMSP2 and PHISTb/RLP1 gene reference sequences was achieved using the online Sequence Manipulation Suite tool (Stothard, 2000). The tool accepts the DNA sequence and maps the primer annealing positions. The output revealed that the designed primers aligned onto the specific part of the sequence that was targeted.

3.3.5 Conventional PCR assays for gene amplification

Conventional PCR conditions for the DBLMSP2 gene were similar for all the primers; denaturation, 94^{o}C- 5 minutes, 30 seconds, annealing 61 ^{o}C - 30 seconds, and extension 72 ^{o}C - 1 minute, 5 minutes. PHISTb/RLP1 PCR conditions for the first pair of primers amplifying the first fragment were 94 ^{o}C - 5 minutes, 30 seconds, annealing 59 ^{o}C - 30 seconds, and extension 72 ^{o}C - 1 minute, 5 minutes. For the second pair of primers, the annealing temperature was 58^{o}C. All the other conditions were similar to the primer pair one. GE health beads were used as a master mix ("PuReTaq ready-to-Go (RTG) PCR beads"). Each sample was amplified three times for DBLMSP2 and two times for PHISTb/RLP1. Amplification was carried out in different plates according to the primer pairs. The thermocycler was set for 40 cycles in each amplification. The primer annealing temperatures were calculated using a temperature calculator and adjusted according to the PCR product output after running agarose gel electrophoresis ("NEB Tm calculator"). The conditions are novel for the set of primers used to amplify the target genes.

3.3.6 Gel electrophoresis

Following PCR, gel electrophoresis was conducted to show whether the samples amplified successfully for each gene (PfDBLMSP2 and PfPHISTb/RLP1 samples from each plate as carried out in primary PCR). 2% Agarose gel was prepared using 10% Tris Acetate EDTA (TAE). To prepare the gel, 2.25g of agarose powder was weighed into a conical flask and 150 ml of 10% TAE buffer was added. The gel was then heated in the microwave for 3 minutes to allow mixing and cooled until it was lukewarm, then added 15 µl of gel red. After that, the gel was poured into the gel tank in which gel combs had been inserted. The gel was incubated at room temperature for 30 to 60 minutes to allow the gel to form (Lee et al., 2012).

To load the samples, 4 µl of each sample template was mixed with with 4 µl of loading dye and loaded to the gel wells. Additionally, 1.5 µl each of 1 kb hyper ladder 1(Bioline) was loaded into the first and final wells on the gel. After loading all the samples, running buffer (10% TAE) was then added, and the gel was run at 230 volts for 50 minutes. The gel was then read on UVI save HD5 - Gel documentation (Uvitec Limited).

3.3.7 Sanger sequencing

The primary PCR amplicons were purified thoroughly before the actual sequencing. The first step was ExoSAP clean-up procedure which employed an enzymatic cleaning procedure using the Alkaline Phosphatase enzyme supplied by Affymetrix (ThermoFisher Scientific). The enzymes act by deactivating the dNTPs that were unused during conventional PCR assay. Using the recommended 96 well-plates, 2µl of the ExoSAP were pippeted into each well followed by the addition of 8 µl of PCR amplicons. The plates were placed into the thermocycler and the cycling conditions were; 37°C for twenty minutes and 80°C for addition twenty minutes. The following step was a secondary PCR where dideoxynucleotides (ddNTPs) were used to enable the chain termination process (Maraka et al., 2020). The samples were divided such that forward and reverse strand sequencing was carried out separately by separating the forward and reverse primers. Upon completion of this process, the secondary PCR products were purified using Sephadex powder supplied by Sigma-Aldrich. The powder would be put in special 96 well plates followed by addition of double distilled water. The powder is hygroscopic and would form pores upon addition of water. The plates were incubated for twenty-four hours. The secondary PCR sequencing products were then pipetted into each specific well of the sephadex plates followed by centrifugation at 2300 RPM for 5 minutes. This process would trap any impurities and allow pure DNA-amplicons to pass through into collecting plates. 10 µl of HiDi formamide was added into the sequencing products for preservation. The sequence reading was achieved using capillary electrophoresis ABI 3130/3500xL Genetic analyzer machine (ThermoFisher Scientific) according to guidelines of Sanger sequencing (Heather & Chain, 2016).

3.4 Data analysis methods

3.4.1 Processing and analysis of sequenced data

Sequencing electropherogram data were imported into CLC Main Workbench version 8.6.1 by QIAGEN for cleaning and analysis. The electropherogram data was organized into three folders for the DBLMSP2. Each sample had six different files because of the

three different primer pairs sequencing different regions. Each fragment was sequenced twice using both forward and reverse primers. The PHISTb/RLP1 chromatogram data was arranged into two folders. Since sequencing of this gene was done using two pairs of primers both forward and reverse, each sample had four files. The data was imported into two different folders named according to the gene name. Each reference sequence was also uploaded to respective folders. Each sample's electropherogram files was assembled against the 3D7 reference sequences for corrections on base calling and creating the consensus sequence. The cleaned consensus for each sample was copied from CLC Main Workbench workspace and saved in FASTA format to notepad. Two different text files were generated for each gene with each sample sequence in FASTA format.

3.4.2 SNPs identification in PHISTb/RLP1 protein sequences

Multiple sequence alignment was done using Multiple Sequence Comparison by Log-Expectation (MUSCLE) (Edgar, 2004) in Jalview version 2.11.0 (Alzohairy, 2014; Waterhouse *et al.,* 2009). Bioedit version 7.2.5 (Hall, 1999) was used to enable intron and gap deletions as well as translation of the nucleotides to protein sequences. SNPs were identified from the alignment and compared against SNPs recorded in PlasmoDB (www.plasmodb.org) genetic variation tracks for PHISTb/RLP1 3D7 reference (PF3D7_0201600). SNPs not recorded in the database were reported as novel mutations. The alignment was copied to Microsoft Excel and frequencies of the observed SNPs counted, as well as generate the frequency bar graph.

3.4.3 Motif search

Motifs within the translated sequences were searched using MEME (Multiple Em for Motif Elicitation**)** Suite version 5.1.0 (Bailey *et al.,* 2009). The Motif search was conducted according to default parameters which include prediction on either side of the strand, minimum motif width of 8, and maximum motif width of 100, no limit on E value and zero or one motif site per sequence. Within MEME Suite, TOMTOM version 5.1.0 tool for motif comparison was used to search the motif results against a known database

(Gupta *et al.,* 2007). Discovered Motifs were also searched against sequence databases for similarities using MAST (Motif Alignment & Search Tool) (Bailey *et al.,* 1998).

3.4.4 Homology modelling, model verification and binding site predictions

Protein structure modelling was achieved by *ab initio* method. DBLMSP2 protein modelling was achieved by ITASSER (Iterative Threading ASSEmbly Refinement) tool version 5.1 (Yang *et al.,* 2015). The tool enabled a hierarchical approach to proteins structure. The PHISTb/RLP1 reference and mutant protein modelling were achieved through the LOMETS Tool, which gives more accurate models for multi-domain proteins. PHISTb/RLP1 is a multi-domain protein and thus LOMETs was preferred for its structure prediction to ITASSER) (Zheng et al., 2019). The tool has an internal selection of templates through inbuilt multiple sequence alignment and ranking the templates in descending order according to a normalized Z score. A Z score greater than or equal to one is considered a good alignment. The PDB hits used as templates for the reference structure were PDB IDs 4jleA, 5ez3A and 6d03E all with a normalized z-score ≥1. The templates selected for the PHISTb/RLP1 mutant structure modelling were 4jleA, 2ziqB and 6d03E which had a normalized z-score ≥ 1. Two templates were similar for the reference and mutant protein structure threading because of sequence similarities at the starting and ending domains of the two protein sequences. The middle template was different for both protein structure threading of mutant and reference structures due to the presence of point mutations in this region of the mutant PHISTb/RLP1 protein sequence. The models were submitted to ModRefiner (Xu & Zhang, 2011) for energy minimization. The energy minimized structure of PHISTb/RLP1 reference had a Template Model (TM) score of 0.97 to the initial model. The TM score is calculated between 0 and 1, whereby the higher the TM score, the higher the perfect match between the two structures. The models were verified using Galaxy Refine web server tool (Ko *et al.,* 2012) for correction of wrong rotamers. The protein models stability were validated through the parameter of percentage residues lying within the favoured and allowed regions using RAMPAGE tool (Lovell *et al.,* 2003) and also overall stability confirmed using ProCheck web tool (Laskowski *et al.,* 1993). Following models validations, the functional sites within the

protein structures were predicted in the web server tool FT Site (Ngan *et al.,* 2012) (Brenke *et al.,* 2009). FT Site algorithm was reported to achieve near experimental accuracy of predicting druggable hotspots in 94% of apo-proteins used in the evaluation of binding sites methods. The outcome of FT Site prediction is a Photoshop Element (.pse) visualized in PyMOL (Schrondiger LLC). For confirmation of the binding Site Clusters, COACH metaserver tool, which uses comparisons of other servers' prediction to profile highly accurate protein-ligand binding sites was used (Yang *et al.,* 2013). The software uses Confidence score (C-score) to rank the binding sites clusters. The C-score ranges between 0-1 where a higher C-score gives a more reliable prediction.

3.4.5 Sulfated polysaccharides search

The Sulfated polysaccharides containing antimalarial properties were searched from Pub-Chem Compounds database, which stores chemical structures of identified chemical compounds and their biochemical activities (Kim *et al.,* 2016). The drug-likeness of the screened compounds were analyzed for Lipinski Rule of Five: Molecular mass less than 500 Dalton; High lipophilicity (expressed as LogP less than 5), Less than five hydrogen bond donors, less than ten hydrogen bond acceptors and Molar refractivity should be between 40-130. This was achieved in the Lipinski Rule of Five webserver (Gimenez *et al.,* 2010)

3.4.6 Preparation of proteins and ligands, molecular docking and output visualization

The Standard Database Format (.SDF) files of the drug compounds obtained were converted to Protein Data Bank (PDB) format files using Openbabel tool version 2.3.1 on Linux command line (O'Boyle *et al.,* 2011). Ligand files, receptor files (protein models), and grid parameters were prepared using MGL tools version 1.5.6 (Morris *et al.,* 2010). Proteins (receptors) and the ligands (drug compounds) were converted from PDB format to PDBQT file format required by autodock tools. Docking simulations were achieved using auto-dock vina version 1.1.2 (Anil, 2019). The output of autodock vina results was visualized using PyMOL version 2.3.5 by Schrodinger LLC.

CHAPTER FOUR

RESULTS

4. 1 Quantitative PCR results for *Plasmodium* genus specificity testing

A total of 175 out of 251 (70%) samples tested positive for *Plasmodium* parasite by quantitative PCR. The samples had at least 32 cycles and were considered to be positive for *Plasmodium* genus as shown in figure 4.1.

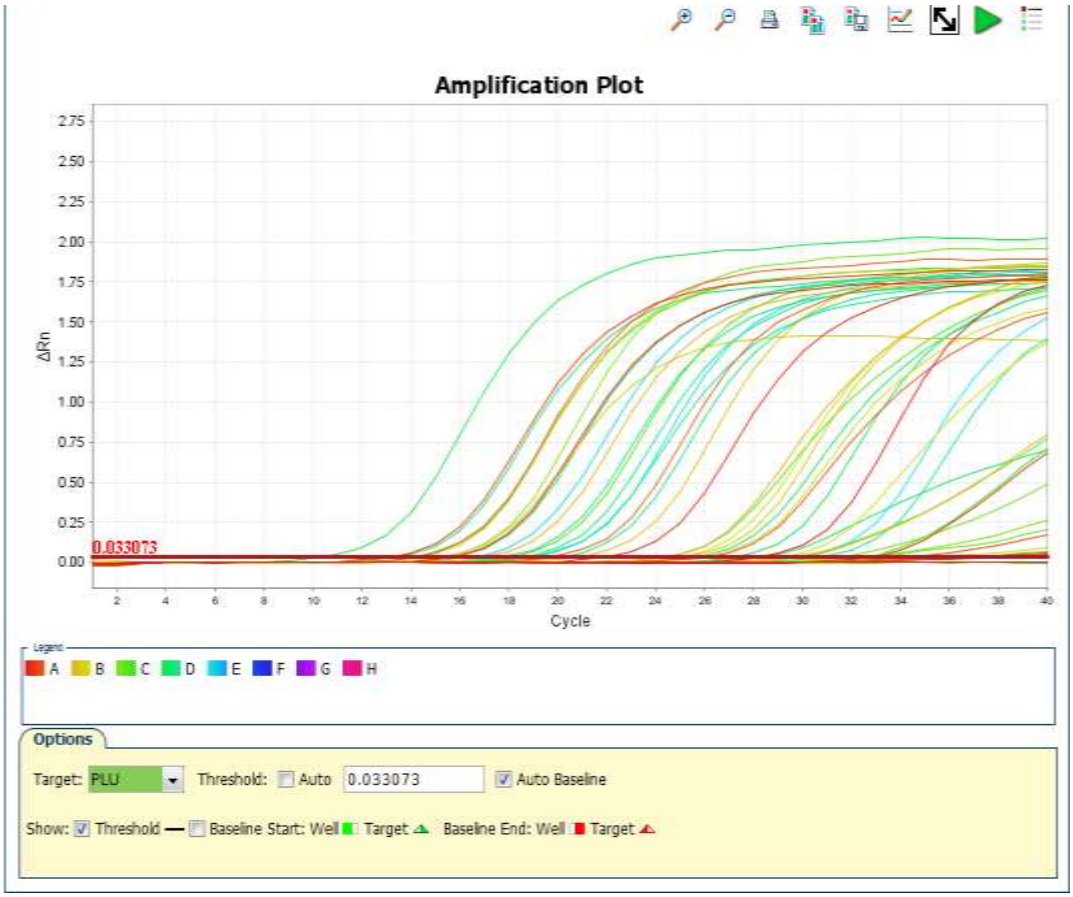

Figure 4.1: A graph showing positive samples of *Plasmodium* genus. Samples with high parasitemia were captured by the shown threshold of 0.033073. The sigmoid curves represent different samples which were colour-coded and visualized in quantstudio.

4.2 *P. falciparum* speciation results

The 70% positive samples were further tested for *P. falciparum* species, and 63% were positive. The 63% (n=110) positive samples were used for the downstream experiments. The positive samples were used for sequencing of DBLMSP2 and PHISTb/RLP1 genes. These samples were tested for other *Plasmodium* species in other studies carried out in the MDR laboratory. Other species documented to be present were *Plasmodium ovale walkeri, Plasmodium ovale curtisi* and *Plasmodium malariae* (Akala *et al.,* 2021). This would account for the difference not accounted for between the total samples containing the *Plasmodium* parasite genome and the *Plasmodium falciparum* positive samples. 37% samples were negative for *Plasmodium falciparum*, could be other species that were not tested in the assay.

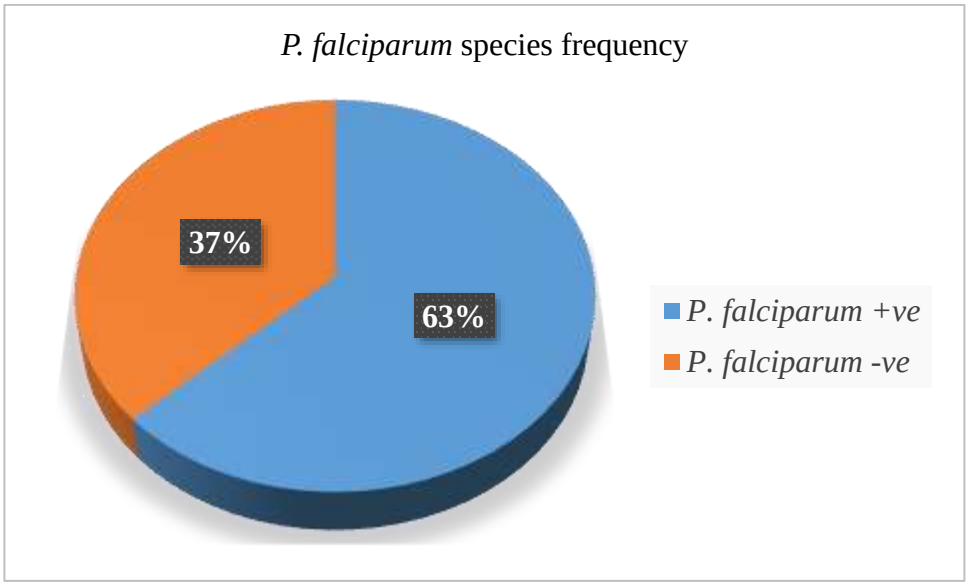

Figure 4.2: The pie chart showing the proportion of the *Plasmodium* positive samples that tested positive for the *P. falciparum* species (63%).

4.3 Gel electrophoresis results

4.3.1 Primer optimization agarose gel electrophoresis results

The gel bands seen in figure 4.3 below show the size of different fragments of DBLMSP2 that were successfully amplified by different primers. The samples were different

Plasmodium falciparum 3d7 strains (nibsc, hb3, dd2).The first bands correspond to different gene fragments: 917, 733 and the third of 755 base pairs respectively.

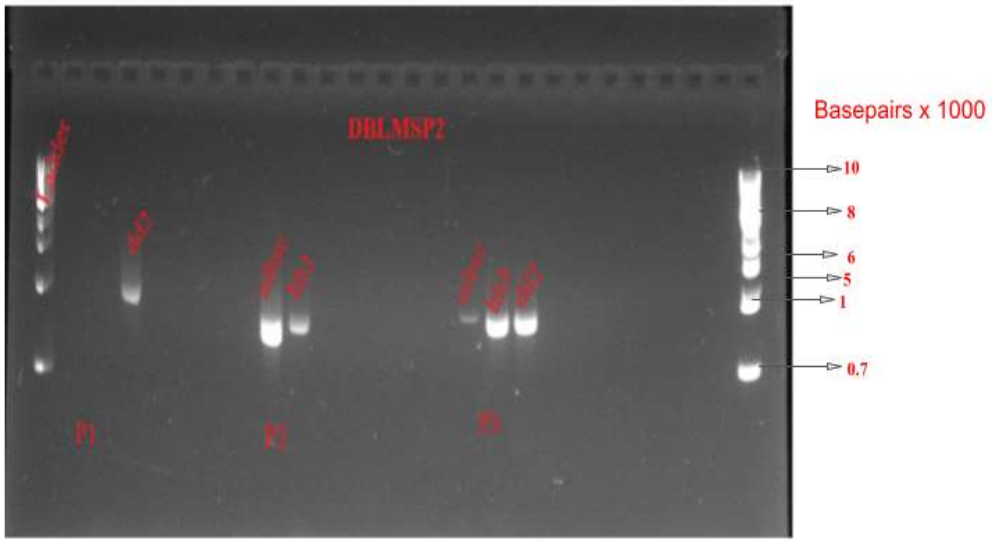

Figure 4.3: Gel electrophoresis image showing amplified *P. falciparum* controls using all the primers for DBLMSP2. The presence of the bands showed successful optimization of the cycling conditions. The bands correspond to the different sizes of the fragmented gene 917,733 and 755 base pairs since they lie between 1000 and 700 base pair bands of the DNA ladder.

4.3.2 Agarose gel electrophoresis results of PHISTb/RLP1

The PCR amplification of PHISTb/RLP1 gene was successful as seen in the pictures after running an agarose gel electrophoresis. The gene was amplified in two fragments. The gel below figure 4.4 shows successful bands seen after running PCR on fragment one whose size was 919 base pairs.

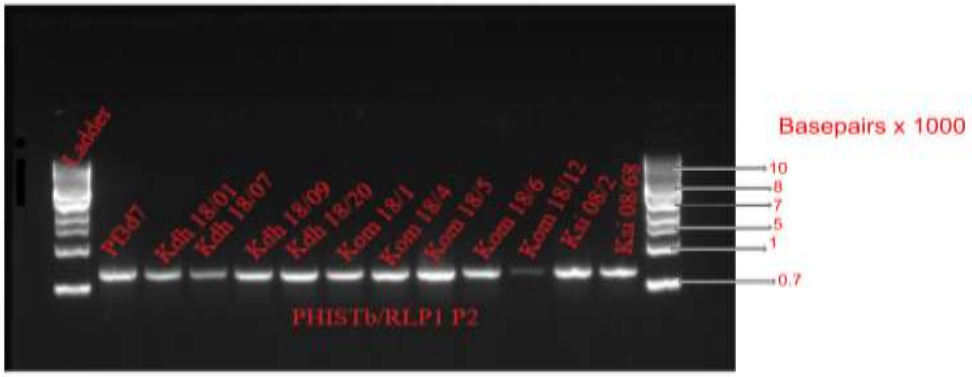

Figure 4.4: **Picture of gel showing bands of the first fragment amplification in PHISTb/RLP1 gene. The band is between 1000 and 700 basepairs ladder band hence corresponding to the right size of 919 base pairs of the actual fragment.**

4.4 Sanger sequencing and Multiple Sequence Alignment of PHISTb/RLP1 results

The sequencing of the PHISTb/RLP1 gene was successful in 102 samples. The sequences were aligned against the reference gene (PHISTb/RLP1 sequence from 3D7 strain) as shown in figure 4.5.

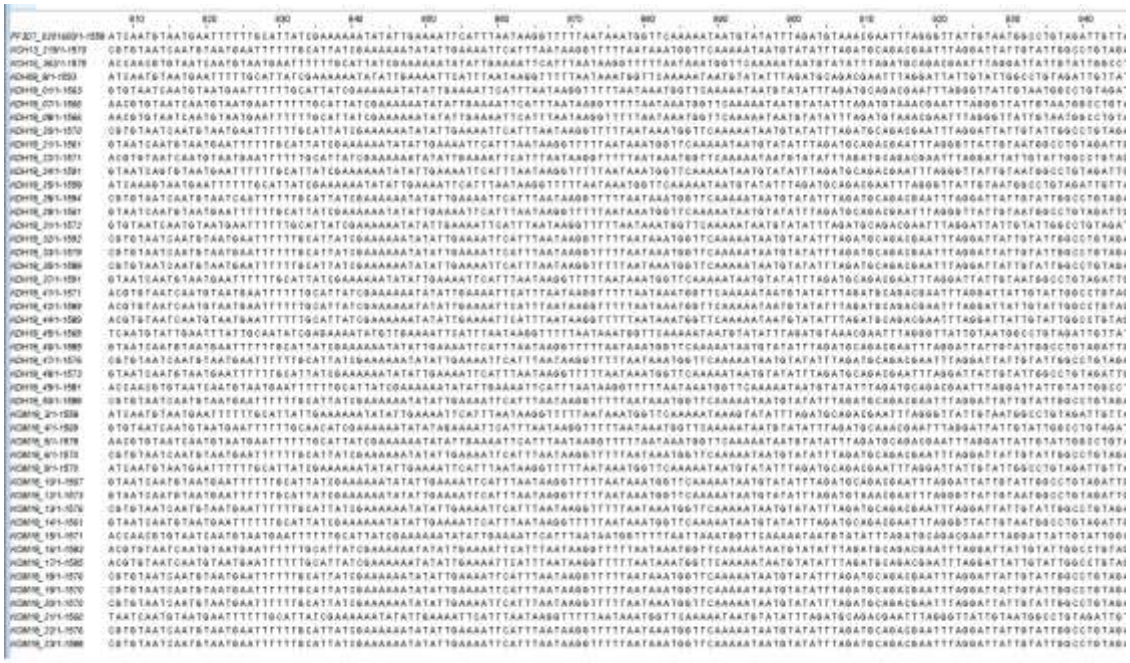

Figure 4.5: **A picture showing a section of multiple sequence alignment of PHISTb/RLP1 sequences. From the figure the sequencing was successful as shown by the presence of nucleotides.**

4.5 PHISTbRLP1 observed Single Nucleotide Polymorphisms (SNPs) results

4.5.1 Total observed Single Nucleotide Polymorphism results

A total of 157 non-synonymous SNPs were observed across the full length of the protein sequences (485 amino acids) as outlined in table 4.1. The SNPs occurred at different frequencies in the total sequenced samples. Only SNPs occurring at a frequency above 50% were considered for protein structure analysis. This is because they were occurring in a high frequency within the sample size hence justified to be true mutations.

Table 4.1: **The frequency distribution of Non-synonymous SNPs identified in PHISTb/RLP1 gene. The table shows the frequency distribution of the number of SNPs against the total sample size. Only 20 SNPs occurred 50% and above in the sample size and were warranted for further analysis.**

Non-Synonymous SNPs total 157	Frequency in N= 102
107	10% and below
30	between 10% to 50%
20	Above 50%

The total number of SNPs was 157 across the entire protein length, in all the 102 sequenced samples. All SNPs occurring in the samples with a frequency below 50% were excluded from further analysis. 20 codons of the total observed Non-Synonymous SNPs had a frequency above 50% of the total samples sequenced. These were considered for protein structure analysis. The frequency was high; thus, the implications of these SNPs were analyzed further.

4.5.2 Non-Synonymous SNPs analysis in PHISTb/RLP1 protein sequences

The graph in figure 4.6 shows all codons with SNPs occurring in 50% frequency with the sample size. F145L, D146R, Y147D, S156H, S208L, and L219H were all novel mutations identified in the collected samples. Codons N108T, F145L, W209F, Y210N, V269A, V274I and M277L omitted since they were substitutions within same functional groups. On the other hand, I129T, T142V, Y147D, E154Q, S156H, T167I, S208L, M211T, L219H, D387N, D390N, and E403K were the substitution of amino acids across different functional groups and were taken further for protein structure analysis.

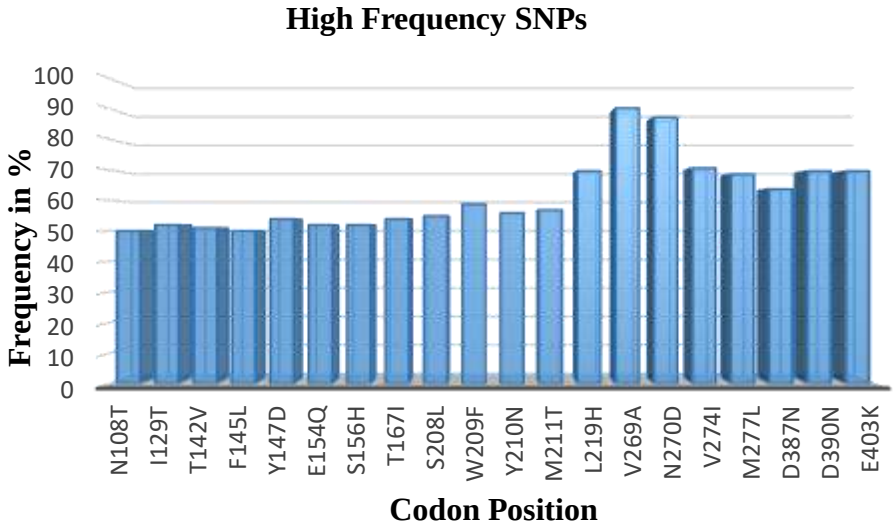

Figure 4.6: **A chart showing non-synonymous SNPs occurring in more than 50% frequency of the sample size.**

4.6 Sequence analysis of DBLMSP2 gene results

The sequencing of this gene was unsuccessful following three different attempts. The N denoted in the sequences shows that no correct nucleotide could be included. No mutation results could be acquired for the gene from the sequenced data.

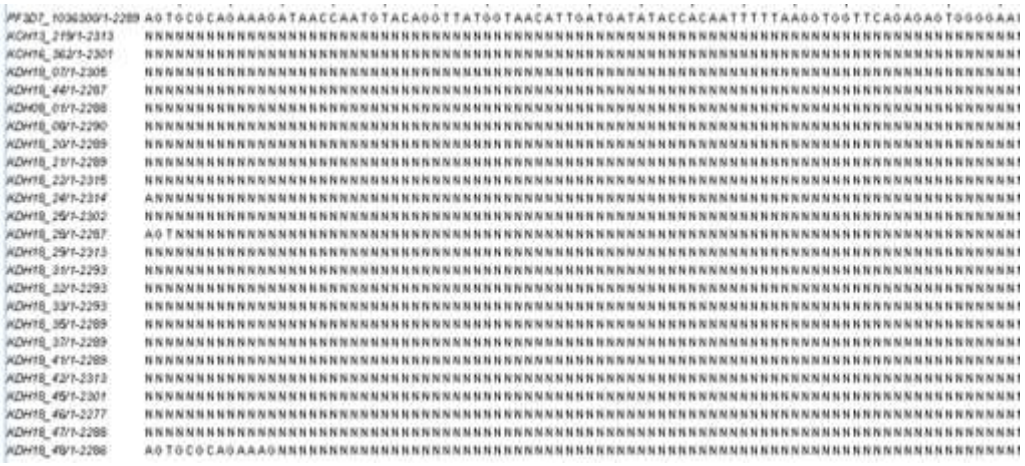

Figure 4.7: **A Jalview image showing a section of Multiple Sequence Alignment of DBLMSP2 sequences. The 'N' residue represents the unsequenced region of complexity in the gene after three attempts of sequencing.**

4.7 Motif search results within PHISTb/RLP1 sequences

MEME suite identified the following motifs in the PHISTb/RLP1 complete sequences (n=86) including Pf3D7_0201600 reference sequence as shown in table 4.2.

Table 4.2: **Results from MeMe software for motif search in PHISTb/RLP1 sequenced data. The table shows the identified motifs and supporting probability values supporting the significance of the results. ELM database comparison reveals that the motifs are involved in signal transduction.**

Motif number	Consensus sequence	Location in sequence	E-value	Frequency in sample size	MAST p value <0.001	TOMTOM ELM match
1	ACRLLWRKTLATLKEEGMLYLQKPFEALHHERKKIHKRGGR	280-320	3.7e-2174	60%	11	12, top hit ELME000187 (Endosome-Lysosome-Basolateral sorting signals)
2	DYDDYDEKQYDKNGEVIVGANEDPSYEYNYHYEEPFILTPELIEAIERAV	396 -445	3.3e-2389	59%	13	21 top hits (ELME000145 human MYND binding motif)
3	RIHTIWHNVMKSEKEKFNLLNTYFNNNFYHLRKKYKTPFNYAKPTCNQCN	190 - 239	5.9e-2668	63%	5	12, top hit ELME000150 (Endosome-Lysosome-Basolateral sorting signals)

4.8 Homology modeling and structure refinement for PHISTb/RLP1 reference protein results

The energy minimized structure of the PHISTb/RLP1 reference had a Template Model (TM) score of 0.97 to the initial model. The Ramachandran plot assessment in RAMPAGE tool (Lovell *et al.,* 2003) revealed 95% of amino acids in the right region. The ProCheck (Laskowski *et al.,* 1993) results displayed 92% of the residues in the reference structure to be in the most favoured region. These Ramachandran plot scores qualified the to be used for the docking experiments and were submitted to binding sites prediction tools. The structure was composed of alpha helical regions as shown in figure 4.8 below in green.

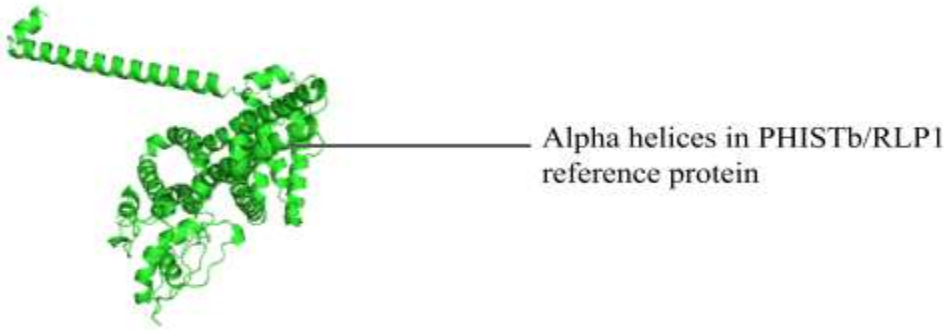

Figure 4.8: **Model of PHISTb/RLP1 in green showing the alpha helices folding of the tertiary protein structure.**

4.9 Homology modelling and structure refinement for PHISTb/RLP1 mutant protein results

The PHISTb/RLP1 mutant protein had a TM score of 0.98 following energy minimization. The structure was characterised by alpha helices as shown in figure 4.8 below coloured cyan. Rampage score of the mutant model prediction yielded 94% of the residues in the favoured region while procheck score was 90%. Ramachandran plot analysis required a good model to have at least 90% of the amino acids in the favoured region therefore the model was fit for use in analysis of drug interactions.

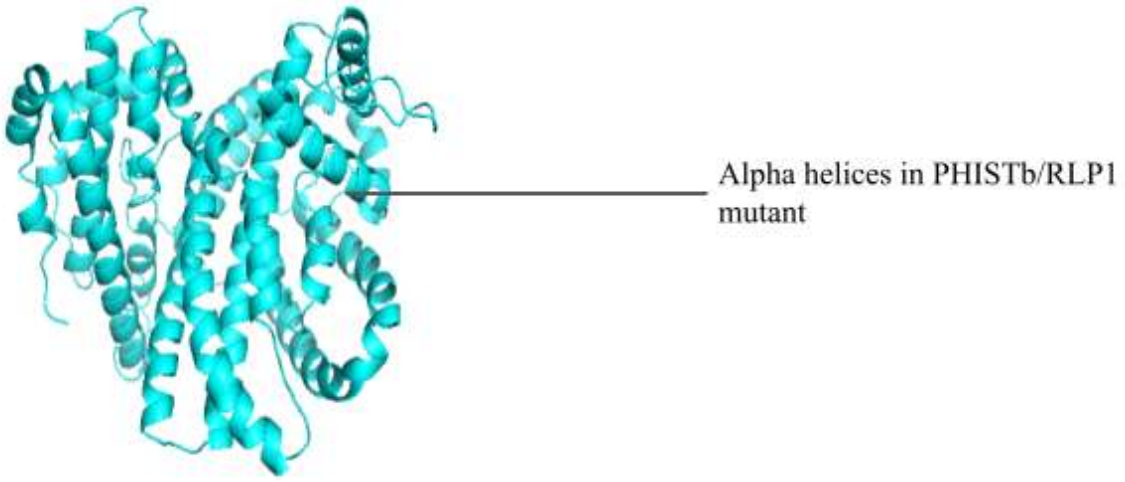

Figure 4.9: **Structure of PHISTb/RLP1 mutant protein structure shown in cyan. All images visualized in PyMOL. Alpha helices fold together and show distinct difference from PHISTb/RLP1 reference structure.**

4.10 Binding sites prediction results in PHISTb/RLP1 reference and mutant structures

FT Site results for both reference and mutant PHISTb/RLP1 protein structures revealed the specific amino acids in the binding sites clusters as shown in table 4.3. The COACH metaserver (Yang *et al.,* 2013) predicted residues in PHISTb/RLP1 protein structure had a C-score between 0.07 and 0.08.

Table 4.3: Amino-acids found within binding sites clusters in both reference and mutant PHISTb/RLP1 protein structures as predicted by FT-Site and COACH softwares

Protein Structure	FT-Site Binding Predictions	COACH Meta-server Predictions	Residues Predicted in Both Tools
PHISTb/RLP1 Reference	K80, K84, K131, S132, A133, F134, N135, S161, E165, G162, T167, F241, S245, K205, K314, R318, A319, K128, N197, S201, E204, K205, R280, L282, K285, K376, E379, D380, E403 and I412	F83, T86, F87, N90, G162, I194, N197, K200, A319, E422 and Y423	N80, F241, K205 and E403
PHISTb/RLP1 Mutant	I15, L16, D17, N18, N20, P27, M28, C31, K35, T86, R89, N90, T211, N214, L244, Y248, F241, F242, S245, N233, K246, T289, E292, E293, I412, V413, G414, A415, N416, E422, D448, V449, E452 and R456	N210, L244, K247, K285, T286, F102, K107, V198, S201, E204, K205, L244, Y328 and E329	K35, R89, N90, L244 and Y248

4.11 DBLMSP2 reference protein structure modelling results

The refined structure had a Ramachandran score of 91.4% of the amino acids in the favored region. Therefore, this qualified the structure to use for docking experiments. The structure is characterized by many alpha helices and two beta sheets as seen in PyMOL (Figure 4.10) coloured orange. The refined structure for the DBLMSP2 was submitted to FT Site and COACH Metaserver web tools for binding sites prediction.

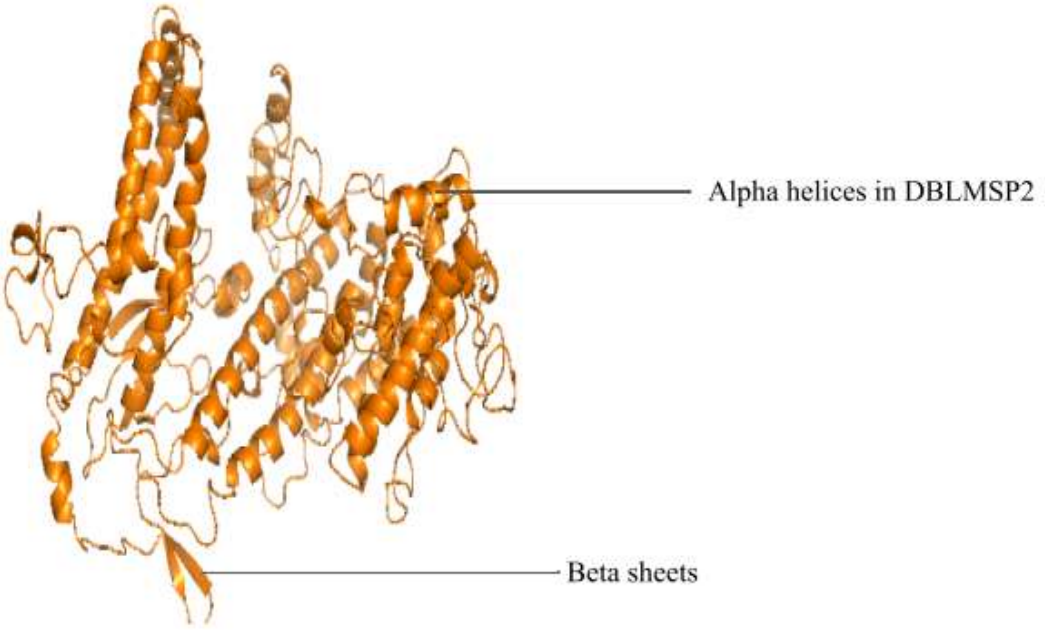

Figure 4.10: **The structure of DBLMSP2 reference protein structure (orange) as visualized in PyMOL. The structure is characterized by alpha helices and beta sheets. The Ramachandran plot score of 91.4% was scored by RAMPAGE.**

4.12: Binding sites prediction of DBLMP2 protein results

The FT site predicted three clusters as visualized in PyMOL. Coach Metaserver predicted active residues in the structure with high C-score between 0.05 and 0.06. Residues predicted by the two softwares, as shown in table 4.4 confirm the accuracy of the predictions and the functional sites of the DBLMSP2 protein. The DBL domain spans residues 156 to 460 (PDB ID: 3VUU) (Hodder *et al.*, 2012) of the entire protein model structure. Most of the predicted clusters were within the DBL domain, which is the functional region of the protein.

Table 4.4: **Active amino acids visualized in PyMOL present in the binding sites clusters of the DBLMSP2 protein structures.**

Protein Structure	FT-Site Binding Predictions	COACH Meta-server Predictions	Residues Predicted in Both Tools
DBLMSP2 Reference	K408, W411, T412, S416, F455, S476, N477, V479, L482, P483, S543, I544, I167, C168, S172, N173, R208, Q 209, S210, L211, R214, L217, K244, M277, T279, P280, T285, V8, F9,I16,Y51,N52,N53,T54, N199, T412, I416, V479, S543, I544, E546, V547 and Q545	H249, Y356, E231, I234, Y235, S240, R243, H316, A320, C323, C168, I167, D492, P377, C372, T379, H388, C393 and S394	K408, W411, T412, S476, S543, I167, C168, Q209 and L211

4.13 Database search of sulfated polysaccharides results

The following 11 compounds shown in table 4.5 resulted from Pub-Chem database search. The physical properties adhered to the likeness of drug compound with slight deviation in Dextrine sulfate and Alpha carrageenan.

Table 4.5: **Drug compounds searched from Pub-Chem Compounds database. Physical characteristics according to Lipinsky rule of five shown.**

Compound	CID number	Molecular weight g/mol	Hydrogen bond donor count	Hydrogen bond acceptor count	High lipophilicity (Log P)	Molar refractivity
Alpha carrageenan	102199625	416.394	4	12	1.7	85
Beta carrageenan	102199626	336.337	4	9	1.9	75
Dextrine sulfate	129722329	598.478	11	20	1.2	113
Amylopectin sulfate	23675774	190.189	0	4	0.8	40
Penoxyacetyl cellulos sulfate	24424	161.436	0	4	0.2	15
Ghatti sulfate	3423265	332.431	0	5	3	78
2,4-Diaminoanisole sulfate	38221	236.242	4	7	-0.1	49
Cyclodextrine sulfate	71317197	322.304	0	8	1.7	68
Fucoidan	92023653	242.242	3	7	1	50
3-Aminophenylboronic acid	92269	136.945	3	3	1	38
3,6-Di-O-benzoyl-D-galactal	11348785	354.348	1	6	3	83

4.14 Protein-Ligand interactions results of PHISTb/RLP1 protein and screened sulfated polysaccharides

The screened sulfated polysaccharides interacted with amino acids in the PHISTb and RESA-Like Protein domains of the PHISTb/RLP1 target. Table 4.6 shows each compound interaction with the amino acids in the binding sites clusters of PHISTb/RLP1 reference and mutant proteins.

Table 4.6: **Specific residue interactions of the drug compounds and the PHISTb/RLP1 reference and mutant protein structures as visualized in PyMOL**

Compound	Interacting residues in PHISTb/RLP1 reference protein	Interacting residues in PHISTb/RLP1 mutant protein
Alpha carrageenan	S132, R315, N80	K35, N90, N214, Q237
Beta carrageenan	K376, N408, E402, E403	D392,E402,K347
Dextrine sulfate	K376, E377, Q404, D406, E402	R315, E95
Amylopectin sulfate	S201, N197, S132, K84, K205	D356, E353
Penoxyacetyl cellulos sulfate	S113, Y137, T29, K111	S113, K261, N69, N121
Ghatti sulfate	K375, G317	R152
2,4-Diaminoanisole sulfate	N7, S12	N121, K125, L71
Cyclodextrin sulfate	S132, A133, F134	E95, K94
Fucoidan	K205, S132, A133	R318,N215,K94
3,6-Di-O-benzoyl-D-galactal	N406, K371	K19, N18, D17

4.14.1 Protein-Ligand interaction results in PHISTb/RLP1 reference protein

Alpha carrageenan, amylopectin sulfate, cyclodextrine sulfate and fucoidan exhibited optimum interactions with the PHISTb/RLP1 reference protein. The amino acids interacting with these compounds were identified in the binding sites. Of these interactions, amino acids S132, K84 and N80 were found within the PHISTb domain as seen in figure 4.11 below.

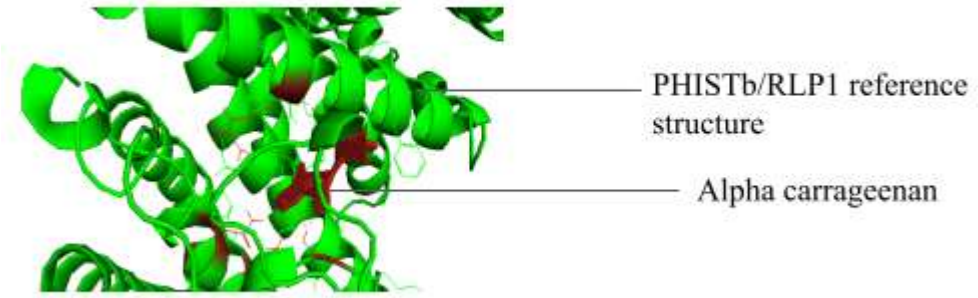

Figure 4.11: **An image showing optimum interaction of alpha carrageenan (brickred) and 3d7 strain PHISTb/RLP1 reference protein structure (green). The compound is deeply buried in the binding site and interacts with specific amino acids through hydrogen bonding.**

4.14.2 Protein-Ligand interaction results in PHISTb/RLP1 reference protein

The screened drug compounds were tested for interaction with PHISTb/RLP1 mutant protein structure. Alpha carrageenan, ghatti sulfate and 3, 6-Di-O-benzoyl-D-galactal depicted specific interactions with the mutant protein. Alpha carrageenan interacted optimally with amino acids in the binding sites of the mutant protein structure as shown in the image below.

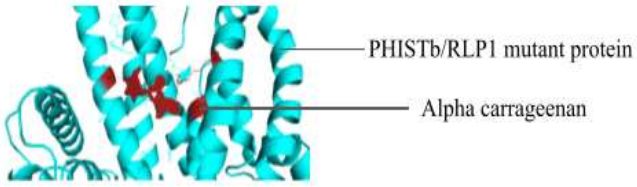

Figure 4.12: **Image showing interaction between PHISTb/RLP1 mutant protein structures (cyan) with alpha carrageenan compound brick brickred as visualized in PyMOL.**

4.15 Interactions of sulfated polysaccharides compounds with DBLMSP2 protein

The screened drug compounds interacted with the DBLMSP2 protein, with most of these interactions within the DBL domain. Fucoidan compound, a sulfated polysaccharide from marine organisms interacts optimally with active residues within the binding site clusters of the protein (figure 4.6). Alpha carrageenan, beta carrageenan, amylopectin sulfate, ghatti sulfate and 3, 6-Di-O-benzoyl-D-galactal interact with specific residues predicted in the binding sites. Table 4.7 outlines all amino acids involved in specific interactions with the screened drug compounds.

Table 4.7: **Specific interactions of the screened sulfated polysaccharides with DBLMSP2 protein.**

Compound	Interacting residues with the DBLMSP2 reference protein
Alpha carrageenan	Q376, T379, P377, D492
Beta carrageenan	Q376, D492
Dextrine sulfate	E484, S473, E471
Amylopectin sulfate	S415
Penoxyacetyl cellulos sulfate	S236,P149,Y153
Ghatti sulfate	S415,T412,K408
2,4-Diaminoanisole sulfate	S236,Y153,T154,P149,K232
Cyclodextrin sulfate	H388, D492
Fucoidan	T412, K408, E410, S415
3,6-Di-O-benzoyl-D-galactal	S543

These are amino acids making polar contacts with the drug compounds in PyMOL. Fucoidan and alpha carrageenan show optimum activity with specific residues in the binding sites as predicted earlier.

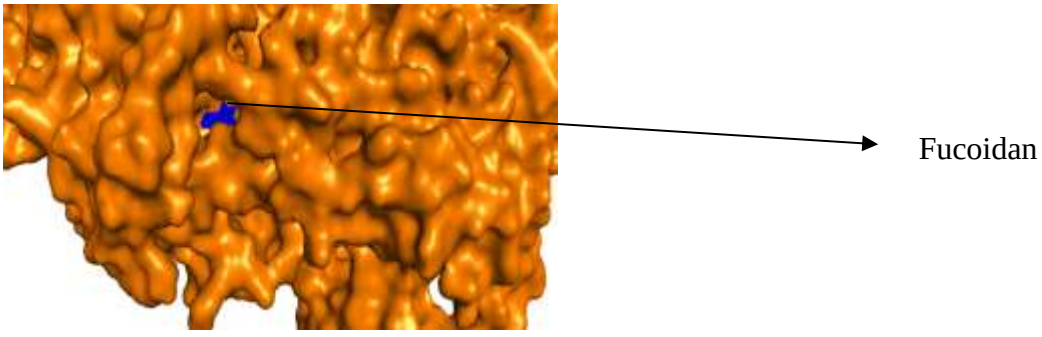

Figure 4.13: **Images showing interactions of DBLMSP2 (Orange) with fucoidan (blue coloured) as seen in PyMOL.**

4. 16 Docking energies comparison

4.16.1 Docking energies results of PHISTb/RLP1 protein interaction with sulfated polysacharides

The resulting docking energy of the first optimal pose shown in kcal/mol was generally low in all the tested compounds as shown in the table below (table 4.8). The resulting Root Mean Square Deviation (RMSD) was zero (First pose RMSD are always zero). This means that the resulting docking poses were accurately correlated with the predicted poses during binding sites prediction.

Table 4.8: **Table showing the docking energies released during interactions of the drug compounds with PHISTb/RLP1 protein.**

Compound	PHISTb/RLP1 reference protein target affinity (kcal/mol)	rmsd	PHISTb/RLP1 mutant protein target affinity (kcal/mol)	rmsd
Alpha Carrageenan	-9.4	0	-7.9	0
Beta Carrageenan	-8	0	-7.2	0
Dextrine sulfate	-11.9	0	-10.3	0
Amylopectin sulfate	-6.2	0	-5.9	0
Penoxyacetyl cellulos sulfate	-4.7	0	-5.1	0
Ghatti sulfatie	-6.6	0	-6.1	0
2,4-Diaminoanisole sulfate	-4.4	0	-4.7	0
Cyclodextrin sulfate	-7	0	-6.2	0
Fucoidan	-7.6	0	-7.2	0
3,6-Di-O-benzoyl-D-galactal	-9.4	0	-8.4	0

4.16.2 Docking energies results in DBLMSP2 protein and sulfated polysacharides

The binding energies released between protein-ligand interactions in DBLMSP2 were very low as seen in the output files after docking. Cyclodextrine sulfate and 3,6-Di-O-benzoyl-D-galactal first poses had no polar contacts with the target leading to the selection of second poses which had direct polar contacts with the protein hence the shift in RMSD (table 4.9).

Table 4.9: **Docking energies released during interactions of the drug compounds with DBLMSP2 protein**

	DBLMSP2 reference protein	
Compound	**affinity (kcal/mol)**	**rmsd**
Alpha Carrageenan	-9.5	0
Beta Carrageenan	-7.9	0
Dextrine sulfate	-11.6	0
Amylopectin sulfate	-6.7	0
Penoxyacetyl cellulos sulfate	-4.7	0
Ghatti sulfate	-6.5	0
2,4-Diaminoanisole sulfate	-4.4	0
Cyclodextrin sulfate	-6.9	27.92
Fucoidan	-8.1	0
3,6-Di-O-benzoyl-D-galactal	-8.7	2.864

CHAPTER FIVE

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

5.1 Discussion

In this study genetic diversity in the PHISTb/RLP1 gene was reported giving a picture of evolution of these proteins as drug targets. This gene was not sequenced before in Kenya and research of exported proteins was new in Kisumu where this study was conducted yet the site is one of the hotspots of Malaria in Kenya. This protein belongs to the chromosome two of the *P. falciparum* genome which was named a hypervariable zone by previous studies (Sargeant *et al.,* 2006b). The presence of numerous SNPs characterizes the subtelomeric regions of the *P. falciparum* genome since genes found in these regions have shown sequence diversity within and across different *P. falciparum* isolates (Figueiredo, 2002; Blythe *et al.,* 2009). The biology of *P. falciparum* survival mechanism in the human host was uncovered in many studies, cytoadherence being one of them (Craig *et al.,* 2012). The process is mediated by many ligands on the merozoite whose receptors are on the RBCs. Cytoadherence is present in other parasite species for instance in *P. knowlesi* infections, the parasite attaches to specific endothelial receptors (Fatih et al., 2012). Progression of infections through cytoadherence was reported to lead to malaria coma (Fatih *et al.,* 2012). The evolution of *Plasmodium* parasite genome exhibits high level gene variability in line with the environmental pressure such as population exposure to vaccines and medicine (Jeffares *et al.,* 2006,). The high genome variability enhances proliferation of the parasite by evading splenic clearance through cytoadherence (Oyebola *et al.,* 2014).

Previous studies observed the *P. falciparum* PHISTb/RLP1 protein which uses the PEXEL motif for export from parasitophorous vacuole into the host cell (De Koning-Wa *et al.,* 2016; Warncke *et al.,* 2016). The Motif results of PHISTb/RLP1 generated top hit motifs involved in signal transduction in humans and other eukaryotes; Endosome-Lysosome-Basolateral sorting signals and human MYND binding motif (Kumar *et al.,* 2019). The locations of these motif consensus are within the fused Dnaj domains in the

RESA regions which are located towards the C-terminal in correlation with Endosome-Lysosome-Basolateral sorting signals (Goel *et al.,* 2014). These Motif consensuses inferred the functional sites of the protein.

Mapping mutations effect on protein structure reveals a lot of intermolecular interactions that when interfered with leads to disease progression (Glusman et al., 2017). The results of this study reveal how mutations within PHISTb/RLP1 protein change the folding of the 3D structure. The implication of the identified mutations within the *Plasmodium falciparum* sequence data was clearly depicted by the differences in 3D folding of the wild type and the mutant structures. The difference in clustering of the binding sites within the wild type and mutant structures support the effect of single point mutations. The clustering of the binding sites within the motifs confirmed these regions as functional sites of the protein. The distribution of the binding sites revealed the multi-domain nature of the PHISTb/RLP1 protein, which contains the PHIST and the fused domains of the RESA region (Tarr *et al.,* 2014; Goel *et al.,* 2014). The multidomain functional regions were further confirmed by the location of the predicted motifs within the sequenced PHISTb/RLP1 data. The challenge of sequencing DBLMSP2 showed the extent of the genomic variability of this region from the *Plasmodium falciparum* 3d7 reference genome. This evolution is a true indication that the presence of many mutations within the *Plasmodium* genome completely changes the genes and swift mutation analysis should be carried out (Moser *et al.,* 2020). The structure analysis of reference DBLMSP2 protein revealed folded domains which could be of functional importance in the survival mechanism of the parasite. The prediction of binding sites showed important amino acids within the DBL domain of the protein emphasizing the importance of finding chemical blockers in this region. Other binding sites located towards the C-terminal and the N-terminal of the protein structure, inferred these regions to be functional sites of the protein. Proteins containing the DBL domain in *P. falciparum* (DBLMSP1 and DBLMSP2) have been linked to play a role in the binding of the parasite to the human immunoglobulin (Crosnier *et al.,* 2016).

The identified sulfated polysaccharides from Pub-Chem had physical and chemical qualities if drug molecules (Kim *et al.,* 2016); A previous study identifying heparin-modified compounds with anti-plasmodial activity profiled a broad group of these sulfated compounds with an inhibition concentration of IC50 <10µg/ml (Boyle *et al.,* 2017). The compounds identified in this previous study targeted invasions of RBCs by the merozoite. The identified sulfated polysaccharides from PubChem Compounds database were from this heparin-modification and were fit for the novel drug discovery studies. Dextrine sulfate deviated from the Rule of five of a drug compound. Still, it was included in further interaction analysis due to its pre-determined antimalarial property, with recommendations of chemical properties modification. The identified drug compounds interacted with the exported protein PHISTb/RLP1 and DBLMSP2 drug targets at the predicted binding site clusters. Compounds such as Alpha carrageenan, Amylopectin sulfate, cyclodextrine sulfate, ghatti sulfate, fucoidan, 3, 6-di-o-benzoyl-d-galactal and beta carrageenan interacted specifically with amino acids located in binding sites of the drug targets. These compounds should further be tested for inhibitory activity against live *Plasmodium falciparum* parasites. Further confirmation of drug potential of these compounds was revealed by the low docking energies released during interactions. The low docking energies was inversely proportion to high affinity of the compounds to the proteins (Pantsar & Poso, 2018). This would mean that they could bind tightly to the functional domains and preventing protein-protein interactions.

The interactions of the screened sulfated polysaccharide compounds with the DBLMSP2 and PHISTb/RLP1 reveal these drug compounds to have the potential chemical inhibitory activity to terminate the cytoadherence survival mechanism of *P. falciparum*. These interactions of the screened drug compounds with the specific active sites found on the proteins show that they can block the parasite's ligands involved in cytoadherence from interacting with their receptors on the RBCs. The termination of cytoadherence mechanism by sulfated polysaccharides interferes with the life cycle of the parasite at the late schizont and merozoite stages. These compounds can prevent progression of

parasitemia to severe forms of malaria. The findings of this study support the antimalarial properties of sulfated polysaccharides, and yields lead compounds that can be taken up for in vitro drug testing and other down-stream drug testing procedures to refine molecules that can be used as the next class of antimalarial drugs.

5.2 Conclusion

The main objective of the study was met since the docking results revealed specific interactions between sulfated polysaccharides and *Pf*PHISTb/RLP1 as well as *Pf*DBLMSP2. specific objective one was met partly since only sequencing of PHISTb/RLP1 gene was successfully sequenced. Objective two was fulfilled by highling motif search results in the sequenced data on PHISTb/RLP1. Objective three was fulfilled by showing that mutations in sequenced data affected tertiary structure folding in the PHISTb/RLP1 reference protein. The failure of sequencing DBLMSP2 gene affected fulfillment of specific objectives two and three due to limited time and resources. The study majorly contributed a new approach for targeted drug discovery in which gives specific details of interactions and how they can be utilized for maximum drug efficacy. In conclusion, the study results support rejection of the null hypothesis which stated that sulfated polysaccharides do not contain antimalarial properties and hence cannot be chemical inhibitors of the cytoadherence proteins. The results lead to acceptance of the alternative hypothesis that sulfated polysaccharides contain antimalarial activity and interact optimally with cytoadherent proteins of *P. falciparum* therefore they are possible drug compounds inhibiting cytoadherence.

5.3 Recommendations

1. Sequencing of DBLMSP2 was a challenge for this current study. I recommend future studies to sequence this gene and investigate how the mutations affect interaction with these screened compounds or other antimalarial drugs.

2. I recommend future studies to identify signaling motifs in the merozoite and other exported proteins and other compounds that can block these proteins.-protein interactions.

3. I recommend future studies to solve crystal structures of exported proteins, since there is a limitation in obtaining templates for homology modeling due to insufficient data.

4. I recommend future studies to do in vitro testing of the screened compound on *Plasmodium* cultures. This data would be helpful to the drug discovery community to screen compounds which can be taken for further drug testing.

REFERENCES

Akala, H. M., Watson, O. J., Mitei, K. K., Juma, D. W., Verity, R., Ingasia, L. A., Opot, B. H., Okoth, R. O., Chemwor, G. C., Juma, J. A., Mwakio, E. W., Brazeau, N., Cheruiyot, A. C., Yeda, R. A., Maraka, M. N., Okello, C. O., Kateete, D. P., Managbanag, J. R., Andagalu, B., … Kamau, E. (2021). Plasmodium interspecies interactions during a period of increasing prevalence of *Plasmodium ovale* in symptomatic individuals seeking treatment: An observational study. *The Lancet Microbe*, *2*(4), e141-e150.

Alonso, S., Chaccour, C. J., Elobolobo, E., Nacima, A., Candrinho, B., Saifodine, A., … Zulliger, R. (2019). The economic burden of malaria on households and the health system in a high transmission district of Mozambique. *Malaria Journal*, *18*(1), 1–10.

Alzohairy, A. (2014). Building a Multiple Sequence Alignment (8) Jalview Basics in More details. *Conference: Genetics,* Genetics Departments, Faculty of Agriculture, Zagazig University.

Anil, K. T. J. W. (2019). Autodock vina: improving the speed and accuracy of docking. *Journal of Computational Chemistry*, *31*(2), 455–461.

Autino, B., Noris, A., Russo, R., & Castelli, F. (2012). Epidemiology of malaria in endemic areas. *Mediterranean Journal of Hematology and Infectious Diseases*, *4*(1), e2012060.

B. Duru, C., C. Nnebue, C., A. Uwakwe, K., C. Diwe, K., C. Agunwa, C., I. Achigbu, K., A. Iwu, C., & A. Merenu, I. (2016). Prevalence and pattern of herbal medicine use in pregnancy among women attending clinics in a tertiary hospital in IMO state, south east Nigeria. *International Journal of Current Research in Biosciences and Plant Biology*, *3*(2), 5-14.

Bahl, A. (2003). PlasmoDB: The *Plasmodium* genome resource. A database integrating experimental and computational data. *Nucleic Acids Research*, *31*(1), 212-215.

Baird, J. K., Maguire, J. D., & Price, R. N. (2012). Diagnosis and treatment of *Plasmodium vivax* malaria. *Advances in Parasitology*, 203-270.

Beeson, J. G., Drew, D. R., Boyle, M. J., Feng, G., Fowkes, F. J., & Richards, J. S. (2016). Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria. *FEMS Microbiology Reviews*, *40*(3), 343-372.

Blythe, J. E., Niang, M., Marsh, K., Holder, A. A., Langhorne, J., & Preiser, P. R. (2009). Characterization of the repertoire diversity of the *Plasmodium falciparum* stevor multigene family in laboratory and field isolates. *Malaria Journal*, *13*, 1–13.

Boddey, J. A., O'Neill, M. T., Lopaticki, S., Carvalho, T. G., Hodder, A. N., Nebl, T., … Cowman, A. F. (2016). Export of malaria proteins requires co-translational processing of the PEXEL motif independent of phosphatidylinositol-3-phosphate binding. *Nature Communications*, *7*(May 2015).

Boyle, M. J., Skidmore, M., Dickerman, B., Cooper, L., Devlin, A., Yates, E., … Beeson, J. G. (2017). Identification of heparin modifications and polysaccharide inhibitors of *Plasmodium falciparum* merozoite invasion that have potential for novel drug development. *Antimicrobial Agents and Chemotherapy*, *61*(11).

Boyle, M. J., Richards, J. S., Gilson, P. R., Chai, W., & Beeson, J. G. (2010). Interactions with heparin-like molecules during erythrocyte invasion by *Plasmodium falciparum* merozoites. *Blood*, *115*(22), 4559-4568.

Brenke, R., Kozakov, D., Chuang, G. Y., Beglov, D., Hall, D., Landon, M. R., Vajda, S. (2009). Fragment-based identification of druggable "hot spots" of proteins using Fourier domain correlation techniques. *Bioinformatics*, *25*(5), 621–627. Bio. (2014).

Craig, A. G., Khairul, M. F. M., & Patil, P. R. (2012). Cytoadherence and severe malaria. *Malaysian Journal of Medical Sciences*, *19*(2), 5–18.

Crosnier, C., Iqbal, Z., Knuepfer, E., Maciuca, S., Perrin, A. J., Kamuyu, G., … Wright, G. J. (2016). Binding of *Plasmodium falciparum* merozoite surface proteins DBLMSP and DBLMSP2 to human immunoglobulin M Is conserved among broadly

diverged sequence variants. *Journal of Biological Chemistry*, *291*(27), 14285–14299.

Damonte, E., Matulewicz, M., & Cerezo, A. (2004). Sulfated seaweed polysaccharides as antiviral agents. *Current Medicinal Chemistry*, *11*(18), 2399-2419.

De Koning-Ward, T. F., Dixon, M. W. A., Tilley, L., & Gilson, P. R. (2016). *Plasmodium* species: Master renovators of their host cells. *Nature Reviews Microbiology*, Vol. 14.

Duffy, P. E., Acharya, P., & Oleinikov, A. V. (2014). Cytoadherence. *Encyclopedia of Malaria*, 1-13.

Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*, 1–19.

Fatih, F. A., Siner, A., Ahmed, A., Woon, L. C., Craig, A. G., Singh, B., Krishna, S., & Cox-Singh, J. (2012). Cytoadherence and virulence - the case of *Plasmodium knowlesi* malaria. *Malaria Journal*, *11*(1).

Ferreira, L., Dos Santos, R., Oliva, G., & Andricopulo, A. (2015). Molecular docking and structure-based drug design strategies. *Molecules*, *20*(7), 13384-13421.

Figueiredo, L. M. (2002). A central role for *Plasmodium falciparum* subtelomeric regions in spatial positioning and telomere length regulation. *The EMBO Journal*, *21*(4), 815-824.

Fowkes, F. J., Boeuf, P., & Beeson, J. G. (2016). Immunity to malaria in an era of declining malaria transmission. *Parasitology*, *143*(2), 139-153.

Gething, P. W., Casey, D. C., Weiss, D. J., Bisanzio, D., Bhatt, S., Cameron, E., Battle, K. E., Dalrymple, U., Rozier, J., Rao, P. C., Kutz, M. J., Barber, R. M., Huynh, C., Shackelford, K. A., Coates, M. M., Nguyen, G., Fraser, M. S., Kulikoff, R., Wang, H., … Lim, S. S. (2016). Mapping *Plasmodium falciparum* Mortality in Africa between 1990 and 2015. *New England Journal of Medicine*, *375*(25), 2435-2445.

Gimenez, B. G., Santos, M. S., Ferrarini, M., & Fernandes, J. P. S. (2010). Evaluation of

blockbuster drugs under the rule-of-five. *Die Pharmazie*, *65*(2), 148–152.

Githinji, S., Oyando, R., Malinga, J., Ejersa, W., Soti, D., Rono, J., Snow, R. W., Buff, A. M., & Noor, A. M. (2017). Completeness of malaria indicator data reporting via the district health information software 2 in Kenya, 2011–2015. *Malaria Journal*, *16*(1).

Glusman, G., Rose, P. W., Prlić, A., Dougherty, J., Duarte, J. M., Hoffman, A. S., Barton, G. J., Bendixen, E., Bergquist, T., Bock, C., Brunk, E., Buljan, M., Burley, S. K., Cai, B., Carter, H., Gao, J., Godzik, A., Heuer, M., Hicks, M., … Deutsch, E. W. (2017). Mapping genetic variations to three-dimensional protein structures to enhance variant interpretation: A proposed framework. *Genome Medicine*, *9*(1).

Goel, S., Muthusamy, A., Miao, J., Cui, L., Salanti, A., Winzeler, E. A., & Gowda, D. C. (2014). Targeted Disruption of a Ring-infected Erythrocyte Surface Antigen (RESA) -like Export Protein Gene in *Plasmodium falciparum* Confers Stable Chondroitin 4-Sulfate Cytoadherence Capacity. *289*(49), 34408–34421.

Gouw, M., Michael, S., Hugo, S., Kumar, M., Lang, B., Bely, B., … Gibson, T. J. (2018). The eukaryotic linear motif resource – 2018 update. *46*(November 2017), 428–434.

Havlik, I., Looareesuwan, S., Vannaphan, S., Wilairatana, P., Krudsood, S., Thuma, P., Kozbor, D., Watanabe, N., & Kaneko, Y. (2005). Curdlan sulphate in human severe/cerebral *Plasmodium falciparum* malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *99*(5), 333-340.

Hay, S. I., Smith, D. L., & Snow, R. W. (2008). Measuring malaria endemicity from intense to interrupted transmission. *The Lancet Infectious Diseases*, *8*(6), 369-378.

Heather, J. M., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*, *107*(1), 1-8.

Hodder, A. N., Czabotar, P. E., Uboldi, A. D., Clarke, O. B., Lin, C. S., Healer, J., … Cowman, A. F. (2012). Insights into Duffy Binding-like Domains through the Crystal Structure and Function of the Merozoite Surface Protein MSPDBL2 from *Plasmodium falciparum* *. *287*(39), 32922–32939.

Hughes, A. (2004). Faculty opinions recommendation of the malaria parasite's chloroquine resistance transporter is a member of the drug/metabolite transporter superfamily. *Faculty Opinions – Post-Publication Peer Review of the Biomedical Literature*.

Institute of Medicine (US) Committee on the Economics of Antimalarial Drugs, Arrow, K. J., Panosian, C., & Gelband, H. (2004). Antimalarial Drugs and Drug Resistance. In *Saving lives, buying time: Economics of malaria drugs in an age of resistance*. National Academies Press.

Jeffares, D. C., Pain, A., Berry, A., Cox, A. V., Stalker, J., Ingle, C. E., Thomas, A., Quail, M. A., Siebenthall, K., Uhlemann, A., Kyes, S., Krishna, S., Newbold, C., Dermitzakis, E. T., & Berriman, M. (2006). Genome variation and evolution of the malaria parasite *Plasmodium falciparum*. *Nature Genetics*, *39*(1), 120-125.

Kamau, E., Tolbert, L. S., Kortepeter, L., Pratt, M., Nyakoe, N., Muringo, L., Ogutu, B., Waitumbi, J. N., & Ockenhouse, C. F. (2011). Development of a highly sensitive genus-specific quantitative reverse transcriptase real-time PCR assay for detection and quantitation of *Plasmodium* by amplifying RNA and DNA of the 18S rRNA genes. *Journal of Clinical Microbiology*, *49*(8), 2946-2953.

Kamau, E., Alemayehu, S., Feghali, K. C., Saunders, D., & Ockenhouse, C. F. (2013). Multiplex qPCR for Detection and Absolute Quantification of Malaria. *8*(8).

Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., … Bryant, S. H. (2016). *PubChem Substance and Compound databases*. *44*(September 2015), 1202–1213.

Ko, J., Park, H., Heo, L., & Seok, C. (2012). GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Research*, *40*(W1), 294–297.

Kobayashi, K., Takano, R., Takemae, H., Sugi, T., Ishiwa, A., Gong, H., Recuenco, F. C., Iwanaga, T., Horimoto, T., Akashi, H., & Kato, K. (2013).

Analyses of interactions between heparin and the apical surface proteins of *Plasmodium falciparum*. *Scientific Reports*, *3*(1).

Kumar, M., Gouw, M., Michael, S., Sámano-Sánchez, H., Pancsa, R., Glavina, J., Diakogianni, A., Valverde, J. A., Bukirova, D., Čalyševa, J., Palopoli, N., Davey, N. E., Chemes, L. B., & Gibson, T. J. (2019). ELM—the eukaryotic linear motif resource in 2020. *Nucleic Acids Research*.

Kwenti, T. E. (2018). Malaria and HIV coinfection in sub-Saharan Africa: Prevalence, impact, and treatment strategies. *Research and Reports in Tropical Medicine, 9*, 123-136.

Lantero, E., Aláez-Versón, C. R., Romero, P., Sierra, T., & Fernàndez-Busquets, X. (2020). Repurposing heparin as antimalarial: Evaluation of multiple modifications toward in vivo application. *Pharmaceutics*, *12*(9), 825.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(2), 283–291.

Lee, P. Y., Costumbrado, J., Hsu, C., & Kim, Y. H. (2012). Agarose gel electrophoresis for the separation of DNA fragments. *Journal of Visualized Experiments*, (62).

Lee, W., Russell, B., & Rénia, L. (2019). Sticking for a cause: The Falciparum malaria parasites Cytoadherence paradigm. *Frontiers in Immunology*, *10*.

Leitgeb, A. M., Nde, P., Cho-Ngwa, F., Titanji, V., Samje, M., Blomqvist, K., & Wahlgren, M. (2011). Low anticoagulant heparin disrupts *Plasmodium falciparum* rosettes in fresh clinical isolates. *The American Journal of Tropical Medicine and Hygiene, 84*(3), 390-396.

Lovell, S. C., Davis, I. W., Arendall, W. B. 3rd, de Bakker, P. I. W., Word, J. M., Prisant, M. G., … Richardson, D. C. (2003). Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins, 50*(3), 437–450.

Luescher-Mattli, M. (2003). Algae, a possible source for new drugs in the treatment of HIV and other viral diseases. *Current Medicinal Chemistry -Anti-Infective Agents, 2*(3), 219-225

Maia, M. F., Kapulu, M., Muthui, M., Wagah, M. G., Ferguson, H. M., Dowell, F. E., … Cartwright, L. R. (2019). Detection of *Plasmodium falciparum* infected *Anopheles gambiae* using near - infrared spectroscopy. *Malaria Journal*, 1–11.

Maitland, K. (2016). Severe malaria in African children — The need for continuing investment. *New England Journal of Medicine*, *375*(25), 2416-2417.

Maraka, M., Akala, H. M., Amolo, A. S., Juma, D., Omariba, D., Cheruiyot, A., Opot, B., Okello Okudo, C., Mwakio, E., Chemwor, G., Juma, J. A., Okoth, R., Yeda, R., & Andagalu, B. (2020). A seven-year surveillance of epidemiology of malaria reveals travel and gender are the key drivers of dispersion of drug resistant genotypes in Kenya. *PeerJ, 8*, e8082.

Moreira, C. K., Naissant, B., Coppi, A., & Bennett, B. L. (2016). The *Plasmodium* PHIST and RESA-Like Protein Families of Human and Rodent Malaria Parasites. 1–27.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2010). Reference-36 docking simulation.pdf. *Journal of …*, *30*(16), 2785–2791.

Moser, K. A., Drábek, E. F., Dwivedi, A., Stucke, E. M., Crabtree, J., Dara, A., Shah, Z., Adams, M., Li, T., Rodrigues, P. T., Koren, S., Phillippy, A. M., Munro, J. B., Ouattara, A., Sparklin, B. C., Dunning Hotopp, J. C., Lyke, K. E., Sadzewicz, L., Tallon, L. J., … Silva, J. C. (2020). Strains used in whole organism *Plasmodium falciparum* vaccine trials differ in genome structure, sequence, and immunogenic potential. *Genome Medicine*, *12*(1).

Müller, I. B., & Hyde, J. E. (2010). Antimalarial drugs: Modes of action and mechanisms of parasite resistance. *Future Microbiology*, *5*(12), 1857-1873.

NEB Tm calculator. (n.d.). NEB Tm Calculator. https://tmcalculator.neb.com/#!/main

Ngan, C. H., Hall, D. R., Zerbe, B., Grove, L. E., Kozakov, D., & Vajda, S. (2012). FtSite: High accuracy detection of ligand binding sites on unbound protein structures. *Bioinformatics*, *28*(2), 286–287.

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An Open chemical toolbox. *Journal of Cheminformatics*, *3*(10), 2946.

Oberli, A., Zurbrügg, L., Rusch, S., Brand, F., Butler, M. E., Day, J. L., … Beck, H. P. (2016). *Plasmodium falciparum Plasmodium* helical interspersed subtelomeric proteins contribute to cytoadherence and anchor *P. falciparum* erythrocyte membrane protein 1 to the host cell cytoskeleton. *Cellular Microbiology*, *18*(10), 1415–1428.

Oligo analysis tool. (n.d.). Eurofins Genomics - Genomic services by experts. https://eurofinsgenomics.eu/en/dna-rna-oligonucleotides/oligo-tools/oligo-analysis-tool/?gclid=CjwKCAjw-

Oyebola, M. K., Idowu, E. T., Nyang, H., Olukosi, Y. A., Otubanjo, O. A., Nwakanma, D. C., Awolola, S. T., & Amambua-Ngwa, A. (2014). Microsatellite markers reveal low levels of population sub-structuring of *Plasmodium falciparum* in southwestern Nigeria. *Malaria Journal*, *13*(1).

Parish, L. A., Mai, D. W., Jones, M. L., Kitson, E. L., & Rayner, J. C. (2013). A member of the *Plasmodium falciparum* PHIST family binds to the erythrocyte cytoskeleton component band 4.1. *Malaria Journal*, *12*(1).

Pantsar, T., & Poso, A. (2018). Binding affinity via docking: Fact and fiction. *Molecules*, *23*(8), 1899.

Phillips, M. A., Burrows, J. N., Manyando, C., Van Huijsduijnen, R. H., Van Voorhis, W. C., & Wells, T. N. (2017). Malaria. *Nature Reviews Disease Primers*, *3*(1).

Plasmodium falciparum mortality in Africa between 1990 and 2015. (2017). *New*

England Journal of Medicine, 376(25), 2493-2494.

Saiwaew, S., Sritabal, J., Piaraksa, N., Keayarsa, S., Ruengweerayut, R., Utaisin, C., … Chotivanich, K. (2017). Effects of sevuparin on rosette formation and cytoadherence of *Plasmodium falciparum* infected erythrocytes. *5*, 1–15.

Sargeant, T. J., Marti, M., Caler, E., Carlton, J. M., Simpson, K., Speed, T. P., & Cowman, A. F. (2006b). Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biology*, *7*(2).

Schaeffer, D. J., & Krylov, V. S. (2000). Anti-HIV activity of extracts and compounds from algae and Cyanobacteria. *Ecotoxicology and Environmental Safety*, *45*(3), 208-227.

Sherling, E. S., Perrin, A. J., Knuepfer, E., Russell, M. R., Collinson, L. M., Miller, L. H., & Blackman, M. J. (2019). The *Plasmodium falciparum* rhoptry bulb protein RAMA plays an essential role in rhoptry neck morphogenesis and host red blood cell invasion. *PLOS Pathogens*, *15*(9), e1008049.

Sidhu, A. B., Verdier-Pinard, D., & Fidock, D. A. (2002). Chloroquine resistance in *Plasmodium falciparum* malaria parasites conferred by pfcrt mutations. *Science*, *298*(5591), 210-213.

Stothard, P. (2000). The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques*, 28:1102-1104

Sullivan, D. (2010). Uncertainty in mapping malaria epidemiology: Implications for control. *Epidemiologic Reviews*, *32*(1), 175-187.

Talarico, L. B., & Damonte, E. B. (2007). Interference in dengue virus adsorption and uncoating by carrageenans. *Virology*, *363*(2), 473-485.

Tang, Y., Ye, Q., Huang, H., & Zheng, W. (2020). An overview of available antimalarials: Discovery, mode of action and drug resistance. *Current Molecular Medicine*, *20*(8), 583-592.

Tarr, S. J., Moon, R. W., Hardege, I., & Osborne, A. R. (2014). A conserved domain targets exported PHISTb family proteins to the periphery of *Plasmodium* infected erythrocytes. *Molecular and Biochemical Parasitology*, *196*(1), 29-40.

Tham, W., Healer, J., & Cowman, A. F. (2012). Erythrocyte and reticulocyte binding-like proteins of *Plasmodium falciparum*. *Trends in Parasitology*, *28*(1), 23-30.

ThermoFisher Scientific. *ExoSAP-IT PCR product cleanup reagent*. Thermo Fisher Scientific - US. https://www.thermofisher.com/ke/en/home/life-science/sequencing/sanger-sequencing/sanger-sequencing-kits-reagents/exosap-it-pcr-product-cleanup.html

Untergasser, Andreas, Ioana Cutcutache, Triinu Koressaar, Jian Ye, Brant C. Faircloth, Maido Remm, and Steven G. Rozen. 2012. "Primer3-New Capabilities and Interfaces." *Nucleic Acids Research* 40 (15): 1–12.

Uwimana, A., Legrand, E., Stokes, B. H., Ndikumana, J. M., Warsame, M., Umulisa, N., Ngamije, D., Munyaneza, T., Mazarati, J., Munguti, K., Campagne, P., Criscuolo, A., Ariey, F., Murindahabi, M., Ringwald, P., Fidock, D. A., Mbituyumuremyi, A., & Menard, D. (2020). Emergence and clonal expansion of in vitro artemisinin-resistant *Plasmodium falciparum* kelch13 R561H mutant parasites in Rwanda. *Nature Medicine*, *26*(10), 1602-1608.

Veron, V., Simon, S., & Carme, B. (2009). Multiplex real-time PCR detection of *P. falciparum, P. vivax* and *P. malariae* in human blood samples. *Experimental Parasitology*, *121*(4), 346–351.

Warncke, J. D., Vakonakis, I., & Beck, H.-P. (2016). *Plasmodium* Helical Interspersed Subtelomeric (PHIST) Proteins, at the Center of Host Cell Remodeling. *Microbiology and Molecular Biology Reviews*, *80*(4), 905–927.

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

Wei, Q., Fu, G., Wang, K., Yang, Q., Zhao, J., Wang, Y., Ji, K., & Song, S. (2022). Advances in research on antiviral activities of Sulfated polysaccharides from seaweeds. *Pharmaceuticals*, *15*(5), 581.

Witvrouw, M., & De Clercq, E. (1997). Sulfated polysaccharides extracted from sea algae as potential antiviral drugs. *General Pharmacology: The Vascular System*, *29*(4), 497-511.

Wu, S., & Zhang, Y. (2007). LOMETS: A local meta-threading-server for protein structure prediction. *Nucleic Acids Research*, *35*(10), 3375–3382.

Wykes, M. N. (2013). Why haven't we made an efficacious vaccine for malaria ? *Nature Publishing Group*, *14*(8), 661.

Xu, D., & Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical Journal*, *101*(10), 2525–2534.

Yang, J., Roy, A., & Zhang, Y. (2013). Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics*, *29*(20), 2588–2595.

Zheng, W., Zhang, C., Wuyun, Q., Pearce, R., Li, Y., & Zhang, Y. (2019). LOMETS2: improved meta-threading server for fold-recognition and structure-based function annotation for distant-homology proteins. *Nucleic Acids Research*, *47*(W1), W429–W436.