



RESEARCH ARTICLE

REVISED **Characterization of sulfated polysaccharide activity against virulent *Plasmodium falciparum* PHISTb/RLP1 protein [version 2; peer review: 2 approved]**

Jennifer M. Mutisya ^{1,2}, Victor A. Mobegi ³, Johnson K. Kinyua², Martha N. Kivecu ¹, Raphael O. Okoth¹, Gladys C. Chemwor¹, Edwin W. Mwakio¹, Agnes C. Cheruiyot¹, Redempta A. Yeda¹, Charles O. Okello¹, Jackline A. Juma¹, Benjamin H. Opot¹, Dennis W. Juma¹, Amanda L. Roth¹, Hosea M. Akala¹, Ben M. Andagalu¹

¹Department of Emerging and Infectious Diseases (DEID), United States Army Medical Research Directorate-Africa (USAMRD-A), Kenya Medical Research Institute (KEMRI)/Walter Reed Project, Kisumu, Kenya

²Department of Biochemistry, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

³Department of Biochemistry, University of Nairobi, Nairobi, Kenya

V2 **First published:** 23 Oct 2020, 9:1268
<https://doi.org/10.12688/f1000research.26756.1>
Latest published: 29 Apr 2022, 9:1268
<https://doi.org/10.12688/f1000research.26756.2>

Abstract

Background: The emergence of artemisinin resistance in South East Asia calls for urgent discovery of new drug compounds that have antiplasmodial activity. Unlike the classical compound screening drug discovery methods, the rational approach involving targeted drug discovery is less cumbersome and therefore key for innovation of new antiplasmodial compounds. *Plasmodium falciparum* (Pf) utilizes the process of host erythrocyte remodeling using Plasmodium-helical interspersed sub-telomeric domain (PHIST) containing proteins, which are amenable drug targets. The aim of this study is to identify inhibitors of PHIST from sulfated polysaccharides as new antimalarials.

Methods: 251 samples from an ongoing study of epidemiology of malaria and drug resistance sensitivity patterns in Kenya were sequenced for PHISTb/RLP1 gene using Sanger sequencing. The sequenced reads were mapped to the reference Pf3D7 protein sequence of PHISTb/RLP1 using CLC Main Workbench. Homology modeling of both reference and mutant protein structures was achieved using the LOMETs tool. The models were refined using ModRefiner for energy minimization. Ramachandran plot was generated by ProCheck to assess the conformation of amino acids in the protein model. Protein binding sites predictions were assessed using FT SITE software. We searched for prospective antimalarials from PubChem. Docking experiments were achieved using AutoDock Vina and analysis results visualized in PyMOL.

Open Peer Review**Approval Status**

	1	2
version 2		
(revision)		
29 Apr 2022		
version 1		
23 Oct 2020		

1. **Eusébio Macete**, Centro de Investigação em Saúde de Manhiça (CISM), Maputo, Mozambique
2. **Abraham Madariaga** , National Autonomous University of Mexico, Mexico City, Mexico

Any reports and responses or comments on the article can be found at the end of the article.

Results: Sanger sequencing generated 86 complete sequences. Upon mapping of the sequences to the reference, 12 non-synonymous single nucleotide polymorphisms were considered for mutant protein structure analysis. Eleven drug compounds with antiplasmodial activity were identified. Both modeled PHISTb/RLP1 reference and mutant structures had a Ramachandran score of >90% of the amino acids in the favored region. Ten of the drug compounds interacted with amino acid residues in PHISTb and RESA domains, showing potential activity against these proteins.

Conclusion: This research identifies inhibitors of exported proteins that can be used in *in vitro* tests against the *Plasmodium* parasite.

Keywords

Exported proteins, Sulfated polysaccharides, PHISTb/RLP1, Interactions, *P. falciparum*, Anti-malarials



This article is included in the [Pathogens gateway](#).



This article is included in the [Chemical Information Science gateway](#).

Corresponding authors: Victor A. Mobegi (vmobegi@gmail.com), Amanda L. Roth (Amanda.Roth@usamru-k.org)

Author roles: **Mutisya JM:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Mobegi VA:** Conceptualization, Supervision, Validation, Writing – Review & Editing; **Kinyua JK:** Supervision, Writing – Review & Editing; **Kivecu MN:** Writing – Review & Editing; **Okoth RO:** Resources, Writing – Review & Editing; **Chemwor GC:** Methodology, Writing – Review & Editing; **Mwakio EW:** Writing – Review & Editing; **Cheruiyot AC:** Writing – Review & Editing; **Yeda RA:** Writing – Review & Editing; **Okello CO:** Writing – Review & Editing; **Juma JA:** Writing – Review & Editing; **Opot BH:** Methodology, Resources, Writing – Review & Editing; **Juma DW:** Project Administration, Resources, Writing – Review & Editing; **Roth AL:** Funding Acquisition, Project Administration, Resources, Writing – Review & Editing; **Akala HM:** Conceptualization, Funding Acquisition, Methodology, Project Administration, Resources, Validation, Writing – Review & Editing; **Andagalu BM:** Funding Acquisition, Investigation, Project Administration, Resources, Supervision, Validation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: Funding for this study was provided by the Armed Forces Health Surveillance Branch (AFHSB) - Global Emerging Infections Surveillance and Response System. This work was also supported through the DELTAS Africa Initiative [DELGEME grant 107740/Z/15/Z]. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [DELGEME grant 107740/Z/15/Z] and the UK government. The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government. Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the author, and are not to be construed as official, or as reflecting true views of the Department of the Army or the Department of Defense. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70-25.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2022 Mutisya JM *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Mutisya JM, Mobegi VA, Kinyua JK *et al.* **Characterization of sulfated polysaccharide activity against virulent *Plasmodium falciparum* PHISTb/RLP1 protein [version 2; peer review: 2 approved]** F1000Research 2022, 9:1268 <https://doi.org/10.12688/f1000research.26756.2>

First published: 23 Oct 2020, 9:1268 <https://doi.org/10.12688/f1000research.26756.1>

REVISED Amendments from Version 1

The newly revised article contains additional information and analysis of the data as suggested by the reviewers. The included analysis is Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) of all the reported drug molecules. A positive control, Low molecular weight heparin was included in the analysis for the results comparison. The control was selected on the basis that it belongs to the sulphated-polysaccharides group and has gone through clinical trials as reported in previous studies. In addition to these, figures elaborating hydrophobic and hydrogen bonding interactions between the ligands and protein targets were included in the discussion in line with reviewer's comments. All other missing information regarding explanation of the protocol as reviewers raised were also addressed and included in the new version. The specific sections are: explanation of the docking protocol, grid box parameters, energy minimization and the choice of modelled structure in LOMETs. An explanation of how samples were selected for this study and the treatment of patients was included in the methods.

Any further responses from the reviewers can be found at the end of the article

Introduction

African countries have 94% of malaria cases and the highest malaria-related death rates according to the [2019 World Health Organization \(WHO\) Malaria Report \(WHO, 2019\)](#). Despite its prioritization in the Millennium Development Goals and other large scale global health initiatives, efforts and strategies to reduce the burden of malaria by 40% have stalled over the years due to different challenges ([WHO, 2018](#)). These challenges include emergence of *Plasmodium falciparum* resistance to first-line treatment, the lack of an efficacious vaccine, vector resistance to insecticides, the great diversity of malaria parasite, and insufficient funding towards the control of the disease. These challenges have created a gap that calls for quick intervention to reduce the burden of the disease ([Dondorp et al., 2012](#)). The latest drug resistance development of *P. falciparum* to artemisinin combination therapies in Southeast Asia ([Takala-harrison et al., 2015](#)), calls for the development of new drug compounds that can overcome parasite resistance to existing drugs.

Studies have shown that sulfated polysaccharides have anti-malarial activities ([Mourão, 2015](#)). Unlike most drugs that target the blood stages of the parasite, these compounds exploit newly identified pathways to interact with intracellular parasites. It has been established that merozoites perforate the erythrocyte membrane before egress ([Boyle et al., 2010](#)). Heparin enters the infected erythrocyte through these pores and prevents merozoite egress ([Glushakova et al., 2017](#)). These findings have led to therapies using heparin to treat severe malaria ([Vogt et al., 2006](#)). Moreover, heparin has been shown to reduce the cytoadherence of infected red blood cells and thus reduce parasitemia. However, heparin use was discontinued due to the induction of severe bleeding in patients ([Leitgeb et al., 2011](#)). Low concentrations of heparin were later used successfully in safety and efficacy trials to disrupt cytoadherence and rosette formation and additional research has been conducted to identify low anticoagulant sulfated polysaccharides with anti-malarial activity. Modified heparin compounds and polysaccharide

inhibitors were successfully profiled for activity against intracellular parasites ([Boyle et al., 2017](#)). These compounds have been identified in marine organisms and plants ([Marques et al., 2016](#)).

The interaction of sulfated polysaccharide molecules with merozoite proteins and the *P. falciparum* erythrocyte membrane protein family have been studied intensively. Members of the Duffy binding-like domain and reticulocyte binding-like domain were shown to have interaction with heparin at different affinities ([Saiwaew et al., 2017](#)). In addition, sulfated polysaccharides interacts with all *P. falciparum* reticulocyte binding homologues, including PfRH2, PfRH4 and PfRH5 ([Somner et al., 2000](#)). The *P. falciparum* parasite uses not only the merozoite surface proteins for invasion but also exported proteins that are found on the surface of the erythrocyte to form cross-linkers with erythrocytes ([Tarr et al., 2014](#)). The exported proteins have been classified to be essential for parasite survival and virulence, hence contribute to the pathogenesis of malaria ([Maier et al., 2008](#)). The Plasmodium-helical interspersed sub-telomeric domain (PHIST) family of exported proteins are expressed in most *Plasmodium* species but they are largely expanded in *P. falciparum*. The PHIST domain clusters into three subgroups across the *Plasmodium* genus, PHISTa, PHISTb, and PHISTc. PHISTa and PHISTb localize specifically in *P. falciparum*. PHISTb is found in *P. vivax* and *P. knowlesi*. One major virulent protein of the PHIST family is PHISTb containing ring-infected erythrocyte surface antigen (RESA)-like protein, abbreviated to PHISTb/RLP1 (PlasmoDB accession number, [PF3D7_0201600](#)). This protein has been involved in enabling attachment of infected erythrocytes to organs and blood vessels, as well as erythrocyte remodeling mechanism. Studies have associated this protein with controlling expression mechanism of PfEMP1 (*P. falciparum* erythrocyte membrane protein) ([Moreira et al., 2016](#); [Oberli et al., 2016](#); [Warncke et al., 2016](#)). As a drug target in the *P. falciparum* genome, PHISTb/RLP1 is identified in Tropical Disease Research Targets database (TDR Targets ID: [3288](#))

In this study, we profile specific protein-ligand interactions that give the first insight into drug compounds targeting exported proteins. Furthermore, we outline old and novel mutations found within PHISTb/protein sequences from whole blood samples collected from various malaria endemic sites in Kenya. We link the effects of these mutations on protein structure, active sites and interaction with the drug molecules.

Methods

Study setting

The study was conducted at the Malaria Drug Resistance laboratory in Kisumu, Kenya, under the United States Army Medical Research Directorate-Africa and Kenya Medical Research Institute. The study analyzed a subset of archived *P. falciparum* samples from ongoing approved research, studying the epidemiology of malaria drug resistance patterns in Kenya. Samples were collected from the following sites: Kisumu and Kombewa (endemic regions), Kericho and Kisii (highland epidemic areas), Marigat (seasonal transmission zone), and Malindi (coastal area with declining endemicity). The proportion of samples collected was roughly the same

between the six collection sites. Samples were selected using non-random method due to time and limited resources. Fifty samples were selected from each site from the laboratory inventory which contained samples from 2008 to 2018. Factors considered were availability of the sample collected at day zero of study participant recruitment.

Sample collection

Samples were collected from volunteers consenting to participate in the study, who were aged 6 months. 2.5 ml of venous blood was collected from participants presenting signs of uncomplicated Malaria. 0.5ml of the blood was transferred to the Acid Citrate Dextrose tubes for sterile culture. Participants that tested positive for Malaria were given Artemether lumefantrine and follow up samples taken at day 7, 14 and 28 following treatment. This study analyzed samples collected at day zero before treatment. At least 1ml of the blood was transferred into an ethylene glycol-bis(2-aminoethylether)-N,N,N',N'tetraacetic acid (EDTA) tube for molecular assays which included DNA extraction. From the remaining blood sample, 3–5 drops of the syringe blood sample was transferred onto Whatman filter paper by blotting on three separate spots and archived at room temperature in a pouch.

Ethical clearance

Written informed consent was provided by participants and/or their legal guardians. The study was carried out in accordance to approved guidelines by the Ethical Review Committee of the Kenya Medical Research Institute (KEMRI), Nairobi, Kenya and Walter Reed Army Institute of Research (WRAIR) Institutional Review Board, Silver Spring, MD. The study was conducted under the approved study protocols KEMRI #1330/WRAIR #1384 and KEMRI #3628/WRAIR #2454.

Sequencing of PHISTb/RLP1 gene using molecular assays

A total of 251 viable samples were used in this study. *P. falciparum* parasite DNA extraction was carried out using Qiagen DNA Mini extraction spin protocol (Qiagen, Valencia, CA) as per manufacturer's instructions. Extracted DNA was stored at -20°C for preservation.

Plasmodium testing and *P. falciparum* speciation using quantitative PCR

First the samples were tested for the presence of *Plasmodium* parasite. *Plasmodium* genus was tested with primers F1 and R1 shown in Table 1, the probe was labeled with FAM (6-carboxyfluorescein) in 5' and at the 3' TAMRA (6-carboxytetramethyl-rhodamine). All the *Plasmodium* positive samples were then tested for *P. falciparum* species. *P. falciparum* species-specific assay was carried out with primers and probes designed with FAM reporter, as reported in previous work (Kamau *et al.*, 2013; Veron *et al.*, 2009). The samples were amplified in 0.1 milliliters 96-well plates. The kit used for this assay was Quantifast (QIAGEN). The components of each well were, 2µl DNA template, recommended amount of Quantifast master mix and recommended amount of primers and probes. The reaction mixture amounts were calculated according to the number of samples run per assay.

The real-time PCR assay was the carried out in Applied Biosystem Quantstudio 6 Flex (Quantstudio™ real-time PCR applied Biosystems by Thermo Fisher Scientific) with the following conditions: 95°C for 10 minutes, 40 cycles at 95°C for 15 minutes, and then 60°C for 1 minute, as previously published (Kamau *et al.*, 2013). RNase-P was used the housekeeping gene and Ct value was awarded by the equipment as the value for which the amplification was high enough to pass the detection threshold The results were exported into Excel worksheet showing the cycle threshold (CT) value for each sample. Only samples positive by *P. falciparum* confirmatory PCR with a cycle threshold value of at least 32 and below were considered for downstream assays (Table 1).

Conventional PCR assays for *P. falciparum* PHISTb/RLP1 gene amplification and gel electrophoresis

Gene-specific primers were designed using Primer3Plus, as described in the literature (Untergasser *et al.*, 2012). Mapping of the primers to sequences to test the correct annealing on sequence was achieved using Sequence Manipulation Suite version 2. Primers are included in Table 2.

PCR was carried out in Applied Biosystems thermocycler machine, the reaction conditions were as follows: the first

Table 1. Primers and probes used to diagnose *Plasmodium* parasite and detect *Plasmodium falciparum* species in real time PCR. These are primers and probes were published in previous work by Kamau *et al.* (2013).

	Primer/probe name	Primer sequence
<i>Plasmodium</i> primers	F1	5'-GCTCTTTCTTGATTCTTGGATG-3'
	R1	5'-AGCAGGTTAAGATCTCGTTCG-3'
	Probe (FAM-TAMRA)	5'-ATGGCCGTTTTTAGTTCGTG-3',
<i>Plasmodium falciparum</i>	<i>Pf</i> -1	5'ATTGCTTTTGAGAGTTTTGTACTTT3'
	<i>Pf</i> -2	5'GCTGTAGTATTCAAACACAATGAACCAA3'
	Probe (FAM-MGB)	5'CATAACAGACGGGTAGTCAT3'

Table 2. PHISTb/RLP1 specific gene primers for *Plasmodium falciparum*, as designed in Primer3 plus tool. The table shows the regions on the gene where each primer should anneal as well as the expected PCR product. These are novel primers for this gene, designed and used in this study for the first time.

Gene	Region	Primers Forward & Reverse	PCR Product
<i>Pf</i> PHISTb/RLP1	80-998	>RLP1_F1: 5'-CAATGACCATGTGTTTCAGA-3'	919
		>RLP1_R1: 5'-TCTGGAGATAAAGCATACCC-3'	
<i>Pf</i> PHISTb/RLP1	791-1450	>RLP1_F2: 5'-GCAAAACCAACGTGTAATC-3'	660
		>RLP1_R2: 5'-TTCAACATCTCTTCGACTG-3'	

pair of primers amplifying the first fragment were 94°C for 5 minutes during initial denaturation phase and 30 seconds in the normal denaturation, annealing 59°C for 30 seconds, and extension 72°C for 1 minute and 5 minutes in the final elongation stage. The second primer pair annealing temperature was 58°C for 30 seconds while the denaturation and extension phases were the same as primer pair one. We used the puRE Taq Ready-To-Go PCR beads standard master mix (GE Healthcare, CITY, STATE), as per manufacturer's instructions.

Following successful amplification, PCR amplicons were visualized in 2% agarose gel electrophoresis for confirmation. The 2% agarose gel was prepared using 2.25gm of agarose powder in 150 ml of 10% Tris Acetate EDTA (TAE) buffer. The gel was then heated in the microwave for 3 min to allow mixing. The gel was then cooled until it was lukewarm, then 15µl of gel red was added. Thereafter 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 set. Upon setting, the gel was flooded with TAE running buffer just above the wells for loading of samples. Samples were then loaded into the wells on the gel wells by mixing 4µl of the sample with 4µl of loading dye. Additionally, 1.5µl each of 1kb hyper ladder 1 (Bioline) was loaded into the first and final wells on the gel. After loading all the samples, the gel was completely flooded with running buffer (10% TAE) and was run at 230 voltage for 50 min. The gel was then read on UVI save HD5 - Gel documentation (Uvitec Limited, United Kingdom).

Sanger sequencing of *P. falciparum* PHISTb/RLP1 gene

Amplification products were purified using ExoSAP-IT (Affymetrix, CITY, STATE) enzymatic cleanup procedure. The enzymatic reaction involves an exonuclease enzyme Shrimp Alkaline Phosphatase (SAP) enzyme, which cleaves phosphate group from unincorporated dNTPs. Using 0.2 ml 96 well plates, each well contained 2µl of ExoSAP-IT enzyme mixed with 8µl of PCR product. The plates were incubated at 37°C for 20 minutes to allow the enzymes to work then for a further 20 minutes at 80°C to inactivate the ExoSAP-IT enzymes prior to sequencing. Big dye termination PCR was carried out using primary PCR primers and amplification conditions. Big dye terminator sequencing is a modification of

Sanger sequencing where dideoxynucleotides (ddNTPs) labeled with a specific fluorescent dye corresponding to each nucleotide base are added to the reaction. The sequencing products were cleaned using Sephadex supplied by Sigma. 10µl of HiDi formamide (Applied Biosystems) was added into the purified sequence products. The plates were then sealed and heated at 96°C for 3 minutes to denature the DNA and then analyzed using capillary electrophoresis ABI 3130/3500xL Genetic analyzer. Read assembling was done on Qiagen CLC main Workbench version 8.0.1 (CLC Bio, 2014).

Single nucleotide polymorphisms analysis in PHISTb/RLP1 protein sequences

Multiple sequence alignment was achieved 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. The alignment was copied to Microsoft Excel and frequencies of the observed single nucleotide polymorphisms (SNPs) counted, as well as generation of frequency bar graphs.

Protein 3D structure modelling, model verification and binding site predictions

Protein structures of the PHIST/RLP1 reference (Pf3D7) and mutant proteins structure were predicted via *ab initio* multithreading tool LOMETS which is verified to model multi-domain protein structures (Wu & Zhang, 2007; Zheng *et al.*, 2019). The LOMETS tool has an internal selection of templates through in-built 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.

Protein Data Bank (PDB) hits used as templates for the PHISTb/RLP1 reference structure were 4jle, 5ez3 and 6d03 all with a normalized Z-score ≥ 1 . The templates selected for the PHISTb/RLP1 mutant structure modelling were 4jle, 2ziq and 6d03, 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 the protein sequences. The middle

template was different for both protein structure threading of mutant and reference structures due to 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 model stability were validated through the parameter of percentage residues lying within the favored and allowed regions using [Rampage tool](#) (Lovell *et al.*, 2003) and overall stability confirmed using [PROCHECK web tool](#) (Laskowski *et al.*, 1993). Following model validations, the functional sites within the protein structure 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 evaluation of binding sites methods. For confirmation of our binding site clusters, [COACH](#), a metaserver tool that uses comparisons of other servers prediction to profile highly accurate protein-ligand binding sites, was used (Yang *et al.*, 2013).

Sulfated polysaccharides search and toxicity analysis

The sulfated polysaccharides containing anti-malarial properties were identified through chemical modifications, as per previous research (Boyle *et al.*, 2017). We searched for these compounds from the [PubChem database](#), which stores chemical structures of identified chemical compounds and their biochemical activities (Kim *et al.*, 2016). Low molecular weight heparin was used a positive control. The compound belongs to sulfated polysaccharides group and was reported to have inhibition concentration of 0.4 mg•mL⁻¹ against *Plasmodium* parasites (Skidmore *et al.*, 2008). The drug likeness of the screened compounds were analyzed for the [Lipinski Rule of Five](#), as follows: molecular mass <500 Dalton; high lipophilicity (expressed as LogP <5); <5 hydrogen bond donors; <10 hydrogen bond acceptors; and molar refractivity should be between 40–130. This was achieved using the [Lipinski Rule of Five](#) webserver (Gimenez *et al.*, 2010). The absorption, distribution, metabolism, excretion, and toxicity (ADMET) analysis of the screened compound was carried out using SwissADME. The tool was certified to conduct pharmacokinetics and their drug-likeness parameters of chemical molecules (Daina *et al.*, 2017). The algorithm contains intergrated tools such as iLOG for testing water solubility and BOILED-egg for absorption, metabolism and excretion of the screened compounds (Daina *et al.*, 2014) (Daina & Zoete, 2016)

Preparation of proteins and ligands, molecular docking and output visualization

The Standard Database Format (.SDF) files of the drug compounds obtained were converted to PDB format files using [Open Babel 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 Molecular Graphics Laboratory (MGL), [MGL tools](#) version 1.5.6 (Morris *et al.*, 2009). The suite contains intergrated the Python Molecular Viewer (PMV), a general purpose molecular viewer [AutoDockTools](#) (ADT) a set of PMV commands specifically developed to support AutoDock users. These parameters enabled the preparation of the ligand and receptor filed to autodock file format in preparation for docking. The ligands underwent geometric minimization and editing to remove contrary ions. Grid parameters used for mutant protein: center_x = 127.746, center_y = 109.332, center_z = 115.398 size_x = 60, size_y = 60, size_z = 60,reference protein grid parameters used: center_x = 1, center_y = 44, center_z = 58, size_x = 60, size_y = 60, size_z = 60 Proteins (receptors) and the ligands (drug compounds) were converted from PDB format to PDBQT file format required by autodock tools using mgl tools. Docking simulations were achieved using [AutoDock Vina](#) version 1.1.2 (Trott & Olson, 2010). The output of AutoDock Vina results was visualized using [PyMOL](#) version 2.3.5 (Schrodinger LLC). Specific protein ligand interactions including hydrogen bonding, vander waals forces as well as hydrophobic interactions were analyzed using [LigPlot](#) version 4.5.3 (Wallace *et al.*, 1995).

Results

Diagnosis for *Plasmodium* genus and *Plasmodium falciparum* species detection

A total of 175 out of 251 (70%) samples tested positive for *Plasmodium* parasite. The 70% positive samples were further tested for *P. falciparum* species and 63% were positive. The positive samples (n=110) were used for the downstream experiments, i.e. they were used to sequence PHISTb/RLP1 gene.

Genetic mutations in PHISTb/RLP1 sequences

102 of the 110 samples processed for Sanger sequencing yielded sequences. Of these, 86 samples generated complete sequences; 16 samples had poor sequences.

Non-synonymous SNPs identified in the PHISTb/RLP1 sequenced data

A total of 157 non-synonymous SNPs were observed across the full length of the protein sequences (485 amino acid in length). The SNPs occurred at different frequencies in the total sequenced samples. Only SNPs occurring at a frequency >50% were considered for protein structure analysis (n=20) (Table 3).

SNPs considered for mutant protein structure analysis

Within the 20 SNPs occurring at high frequency, codons F145L, D146R, Y147D, S156H, S208L and L219H were all novel mutations identified in our samples across all different sites where our samples were collected. We compared our data against SNPs recorded in PlasmoDB genetic variation tracks for the PHISTb/RLP1 3D7 reference (PF3D7_0201600). These six mutations were not present in the database; therefore, they could be newly identified mutations for PHISTb/RLP1 gene as observed in our samples. Analyzing the amino acids substitutions further, we observed that mutations at codons N108T, F145L, W209F, Y210N, V269A, V274I and M277L were all substitutions of amino acids within

the same group. The F145L, W209F, V269A, V274I and M277L mutations were all substitutions within the non-polar group, while N108T and W209F were substitutions in the polar group. Amino acids within the same group play similar functional roles in the protein sequence and thus these SNPs were not considered for protein function analysis. On the other hand, I129T, T142V, Y147D, E154Q, S156H, T167I, S208L, M211T, L219H, D387N, D390N, and E403K were substitutions of amino acids across different functional groups. We postulated that these point mutations are most likely to change

the folding of the protein structure and have an effect on the function of the protein. These codons were modelled in the mutant protein sequence for PHISTb/RLP1 structure. Their effect on the tertiary protein structure as well as effect on interaction with a drug compound were further investigated (Figure 1).

Homology modeling, structure refinement and binding sites prediction

PHISTb/RLP1 reference and mutant structures were modelled successfully. The LOMETs modelling results revealed five structures with significant close scores. All structures were assessed for correct positioning of the amino acid side chains using RAMPAGE tool. The model with high Ramachandran score above 90% which is the recommended value for a significant model was selected to be used for further analysis. Following this, the fourth model in wildtype and mutant structure prediction were both fit to be used for downstream docking experiments. The energy minimized structure of PHISTb/RLP1 reference had a 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 PHISTb/RLP1 mutant protein had a TM score of 0.98 following energy minimization. The models were submitted to Galaxy Refine for further model refinement and then we chose the first models with best scores. The Galaxy refined models were submitted to RAMPAGE software for Ramachandran plot assessment. The reference structure had 95% and the mutant structure 94% of the residues in the favored region. The PROCHECK results displayed 92% of the residues in the reference structure to be in the most favored region. The mutant structure had 90% of the residues to be in the favored region. Ramachandran plot analysis required a good model to have at least 90% of the amino acids in the favored region. These Ramachandran plot scores qualified the two models to be used for the docking experiments and were submitted to binding sites prediction tools.

Table 3. Total observed non-synonymous SNPs in the sequenced data after read cleaning, multiple sequence alignment, intron deletion and conversion to protein sequences.

The total number of SNPs was 157 across the entire protein length, in all the 102 sequenced samples. 107 SNPs were only overserved in 1-10% of the samples and were excluded. 30 of the remaining SNPs occurred in 10-50% samples and this frequency was not high enough to be considered for downstream analysis, they were excluded as well. The final 20 codons of the total observed non-synonymous SNPs had a frequency >50% of the total samples sequenced. These were considered for protein structure analysis. The frequency was high thus the implications of these SNPs was analyzed further.

Non-synonymous SNPs (n=157)	Frequency in 102 samples
107	10% and below
30	between 10% to 50%
20	Above 50%

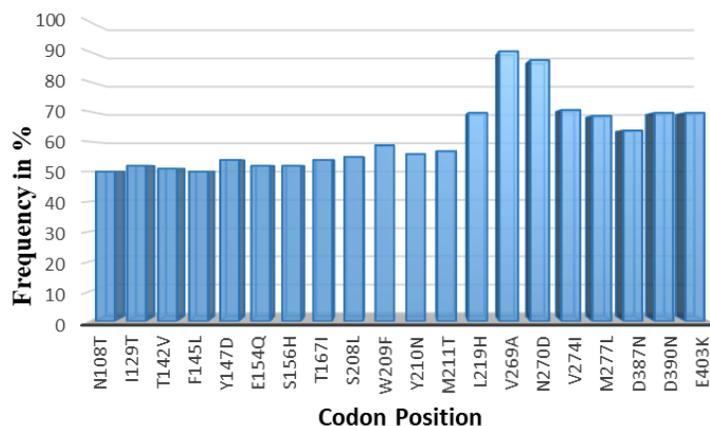


Figure 1. High frequency non-synonymous SNPs observed in sequenced data for PHISTb/RLP1 protein. Non-synonymous SNPs occurring in 50% and above of the total sample size. Codons I129T, T142V, Y147D, E154Q, S156H, T167I, S208L, M211T, L219H, D387N, D390N, and E403K were substitution of amino acids across different functional groups. These were modelled and their implications on protein structure and interactions with a drug compound investigated.

FT Site prediction outcome included Photoshop Element (.pse) files that were visualized in PyMol to visualize the active amino acids in the binding sites clusters for both reference and mutant PHISTb/RLP1 protein structures, as shown in [Table 4](#). We also considered residues that were predicted by COACH software and had the highest confidence score (C-score) of 0.05 for PHISTb/RLP1 reference protein structure. COACH predicted residues in PHISTb/RLP1 protein structure had a C-score between 0.07 and 0.08. The C-score ranges between 0–1 where a higher C-score gives a more reliable prediction. In both the mutant and the reference structures, there were active residues predicted by both tools. This double prediction emphasized

the accuracy of the binding sites as functional regions of the proteins ([Figure 2](#) and [Table 5](#)).

Sulfated polysaccharides search and Absorption, Distribution, Metabolism, Excretion, and Toxicity analysis

The PubChem database search yielded the following compounds: beta carrageenan, alpha carrageenan, dextrin sulfate, amylopectin sulfate, zinc sulfate, ghatti sulfate, 2,4-diaminoanisoole sulfate, cyclodextrin sulfate, fucoidan, 3-aminophenylboronic acid and 3,6-Di-O-benzoyl-D-galactal. In this group of compounds, dextrin sulfate deviated from the Lipinski Rule of

Table 4. Galaxy refine scores for the best model of the protein structures for PHISTb/RLP1 reference and mutant proteins after structure refinement. The score considered included the poor rotamers scores in which they are very low in both inferring to correction in R chain conformations. The Rama favored scores were also used to consider the models. The scores above 90 for Rama favored shows that the models are good to be used for further analysis.

Model	GDT-HA	RMSD	MolProbity	Clash Score	Poor Rotamers	Rama Favored
PHISTb/RLP1 Reference	0.9454	0.433	2.346	22.9	1.1	94.8
PHISTb/RLP1 Mutant	0.9438	0.436	2.348	25.7	1.5	93.6

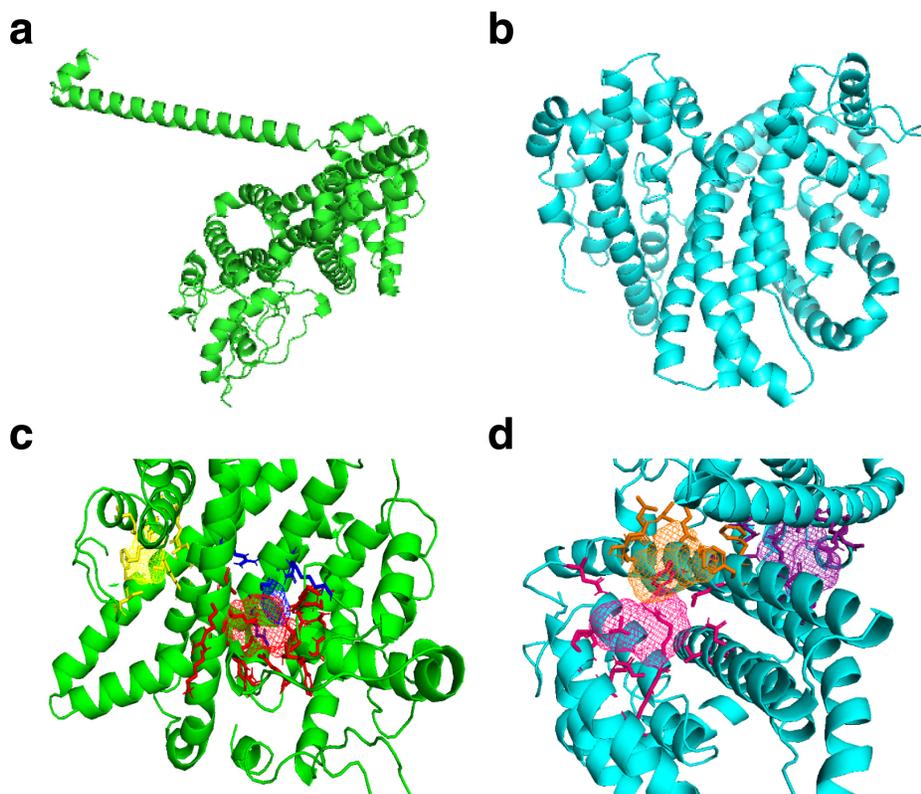


Figure 2. PHISTb/RLP1 reference and mutant homology modelled structures and their predicted binding sites clusters. (a) PHISTb/RLP1 reference protein structure shown in green; (b) structure of PHISTb/RLP1 mutant protein structure shown in cyan; (c) PHISTb/RLP1 reference protein binding site clusters, site 1 in red, site 2 in blue and site 3 in yellow; (d) PHISTb/RLP1 mutant protein binding sites clusters, site 1 in pink, site 2 in orange and site 3 in dark purple. All models visualized in PyMOL.

Table 5. Amino-acids found within binding sites clusters in both reference and mutant PHISTb/RLP1 protein structures as predicted by FT-Site and COACH softwares. The highlighted residues are predicted by both softwares. This puts emphasis on the accuracy of the predictions.

Protein Structure	FT-Site Binding Predictions	COACH Meta-server redictions	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

Five. Beta carrageenan has hydrogen bond acceptors of 12, the compound was considered for interaction experiments because the Rule of Five allows a window of one deviation of the physical properties. 3-aminophenylboronic acid PDB file contained atoms that could not be recognized by AutoDock Vina. The compound was left out for docking experiments. All the other drug compounds adhered to Lipinski Rule of Five hence passing the toxicity test of drug likeness and were tested for activity against the exported protein PHISTb/RLP1. Most compounds contain high gastrointestinal absorption while alpha and Beta carrageenan show low absorption rates. The distribution of the compounds was shown by significant score of bioavailability score of 0.55 which is the recommended value for medicinal drugs. The compounds do not inhibit cytochrome isoenzymes and they do not cross the blood brain barrier (BBB). The skin permeability score were negative showing less permeability of the molecules (Table 6, Table 7).

PHISTb/RLP1 protein-ligand interactions

The identified compounds interacted with the target protein. Alpha carrageenan, amylopectin sulfate, cyclodextrin sulfate and Fucoidan exhibited optimum interactions with the PHISTb/RLP1 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. These interactions significantly show the potential of interfering with the function of the exported protein. Beta carrageenan and dextrin sulfate compounds had specific interactions with the target protein. Amino acids K376, E402 and E403 were identified in the binding site clusters. 2, 4-Diaminoanisole sulfate and zinc sulfate interactions with the protein were not specific with the identified binding site clusters. However, we noted that the amino acids interacting with these compounds are within the PHIST domain which spans amino acids 1 to 167 of the protein sequence. Ghatti sulfate showed interaction with the reference protein as well, although not specific.

The same drug compounds were tested on the modelled PHISTb/RLP1 mutant protein structure with the identified non-synonymous SNPs from the sequenced data. The aim was to test whether mutations found within the protein interfere with the interactions or not. The interactions of the drug

compounds with mutant protein were not as strong as those of the reference protein. The interactions in the mutant protein occurred at different binding residues because the mutations within the proteins affected the folding of the protein structure. Due to these mutations, there was a shift of the binding sites clusters. Alpha carrageenan, ghatti sulfate and 3,6-Di-O-benzoyl-D-galactal depicted specific interactions with the mutant protein. Residues K35, N90, D17 and N18 were within the PHIST domain. The other compounds were interacting with different residues in the protein domains. The interactions were not optimized to the predicted binding site clusters. However, we noted that some of the interacting residues were within the PHISTb domain, which is the functional region of the protein. The interactions were compared to the positive control low molecular weight heparin which contains predetermined activity against *Plasmodium falciparum*. The interaction visualization of low molecular weight heparin with PHISTb/RLP1 reference protein reveal specific interactions between the compound and N197, K205 and S161 all predicted with high confidence score in the target binding site as shown in Figure 3. These interactions, therefore, gave insight to the action of the drug compounds against the mutant protein.

When selecting the poses after docking experiments, we selected the first pose of the AutoDock Vina pdbqt output file. The docking energies of the screened reference and mutant targets were very low. The energies released therefore supported the drug likeness of these compounds. The resulting Root Mean Square Deviation (RMSD) was zero (first pose RMSD are always zero). There was a difference in docking energies of the drug compounds interacting with PHISTb/RLP1 reference and mutant proteins because of the difference in binding site clusters. The interactions with different residues caused by the shifting of binding sites due to mutations resulted in the slight difference in energy released during docking (Figure 4, Table 8–Table 10)

Discussion

Point mutations, protein structure analysis and binding sites

In *P. falciparum*, SNPs have been reported to cluster in subtelomeric regions of the chromosomes. A study comparing synteny of exported proteins in *P. vivax* and *P. falciparum*

Table 6. Screened sulfated polysaccharides searched from PubChem database. Accession numbers, physical properties and Lipinski Rule of Five properties of each compound are curated. The Lipinski rule of five requires that a drug compound meets all the properties or deviates with at most one property. All the screened compound met the rules of a drug compound, except dextrine sulfate. Low molecular weight heparin included as a positive control.

Compound	CID Number	Molecular weight g/mol	Hydrogen bond donor count	Hydrogen bond acceptor count	High lipophilicity (Log P)	Molar Refractivity
Low molecular weight heparin	118989588	241.26	2	2	1.2	6
Alpha carrageenan	102199625	416.394	4	12	1.7	85
Beta carrageenan	102199626	336.337	4	9	1.9	75
Dextrin sulfate	129722329	598.478	11	20	1.2	113
Amylopectin sulfate	23675774	190.189	0	4	0.8	40
Zinc sulfate	24424	161.436	0	4	0.2	15
Ghatti sulfate	3423265	332.431	0	5	3	78
2,4-Diaminoanisoole sulfate	38221	236.242	4	7	-0.1	49
Cyclodextrin 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

Table 7. ADMET analysis data for screened drug compounds. The compounds are highly soluble in water with logS score greater than -4 which is the range of most market drugs. The compounds have high gastrointestinal absorption except for carrageenan and 2,4-Diaminoanisoole sulfate. The bioavailability value is highly significant for most of the compounds. The drug compounds do not inhibit cytochromes hence no formation of toxic compounds. Skin permeation values are negative depicting less permeation. Lipophilicity greater than 1 in 3,6-Di-O-benzoyl-D-galactal and ghatti sulfate shows the ability of these compounds to reach target cells.

	Water solubility (LogS)	Lipophilicity Log P_{ow}	Bioavailability score	BBB Permeant	Cytochrome inhibitor	Skin permeation	GI absorption
Low molecular weight heparin	0.57	-3.94	0.55	No	No	-8.94 cm/s	Low
Ghatti sulfate	-3.82	0.76	0.55	No	No	-4.69cm/s	High
Cyclodextrine sulfate	-0.11	-2.69	0.55	No	No	-8.62 cm/s	High
Beta Carrageenan	1.82	-1.27	0.55	No	No	-9.62 cm/s	low
Amylopectin sulfate	-0.6	-1.82	0.55	No	No	-6.86 cm/s	High
Alpha carrageenan	2.22	-1.62	0.11	No	No	-10.38 cm/s	Low
3-Aminophenylboronic acid	-0.90	-0.54	0.55	No	No	-7.03 cm/s	High
3,6-Di-O-benzoyl-D-galactal	-4.05	2.57	0.55	No	No	-6.39 cm/s	High
2,4-Diaminoanisoole sulfate	0.28	-1.01	0.55	No	No	-8.33 cm/s	Low
Fucoidan	1.42	-1.06	0.56	No	No	-8.86 cm/s	High

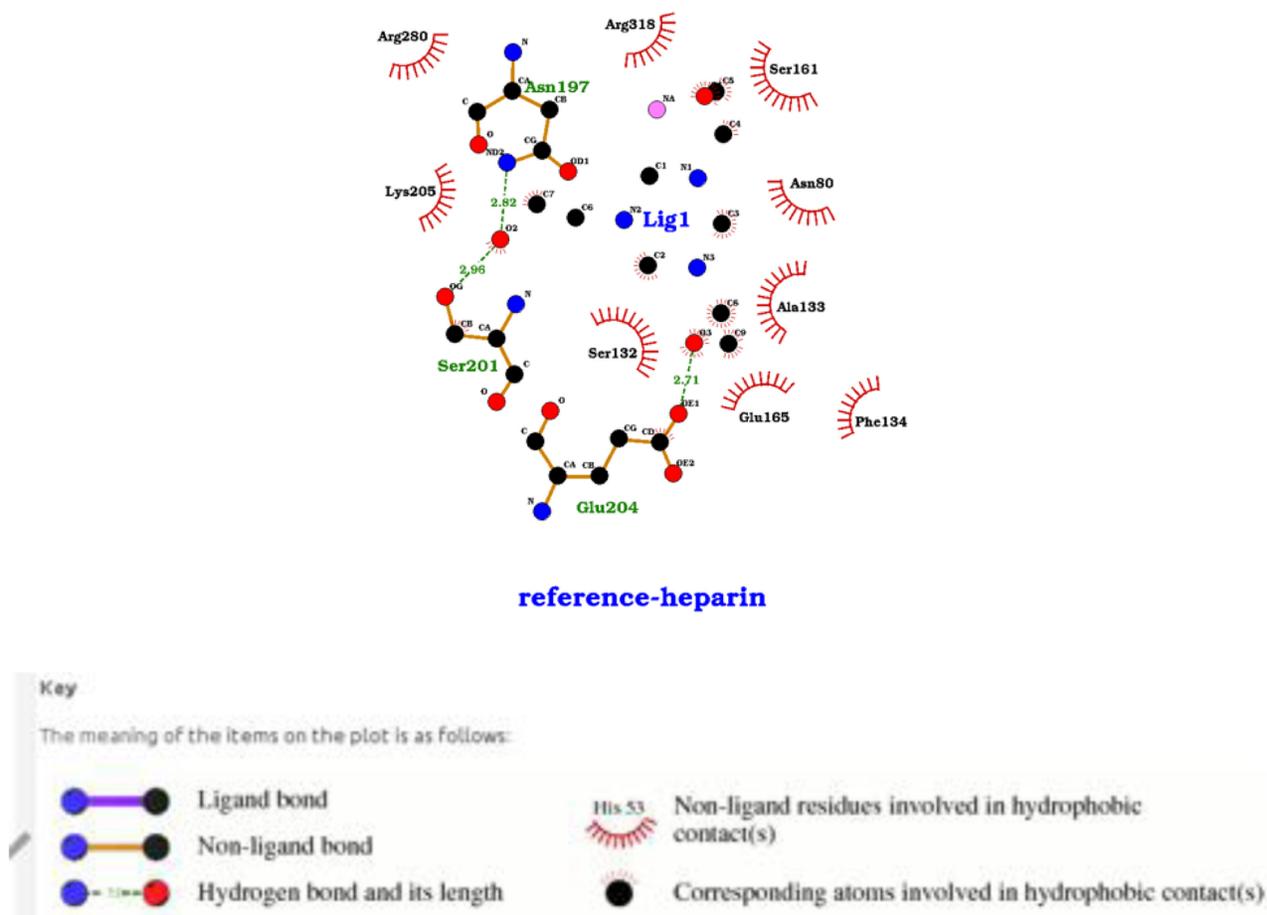


Figure 3. A LigPlot showing interactions of the low molecular weight heparin compound with residues in the active site of the PHISTb/RLP1 reference protein. The compound forms hydrophobic interactions with S161 and K205 amino acids. Other interactions shown include hydrogen bonding N197 S201 and E204 all within the PHISTb/RLP1 binding pocket.

reported a large number of SNPs in chromosome 2 and 10 subtelomeric regions (Sargeant *et al.*, 2006). The subtelomeric regions of the *P. falciparum* genome have been reported to be highly variable. The genes found in these regions have shown sequence diversity within and across different *P. falciparum* isolates. The drug target PHISTb/RLP1 is located in chromosome 2. Our genetic diversity analysis of this protein in Kenyan isolates depicted many SNPs, with many of the polymorphisms having been identified before. Novel SNPs were reported with high frequency across our sample size. Novel point mutations were identified in our sequences across different groups of amino acids. These included I129T, T142V, Y147D, E154Q, S156H, T167I, S208L, M211T, L219H, D387N, D390N, and E403K. These substitutions are hypothesized to affect the function of the protein and we considered them to analyze their effect on the structure of the protein, as well as interactions with drug compounds.

The homology modelling of full length PHISTb/RLP1 protein revealed that the 3D structure of this drug target is

characterized by alpha helices. This was confirmed by the solved PHIST domain crystal structure PDB ID 4JLE. Upon comparing the reference and mutant structures, there was a difference in the folding of the alpha helices. This difference was clearly depicted by the difference in clustering of the binding sites. This can be accounted for by the point mutations, which changed the conformation of the protein structure. The tool used to model the two proteins selected one different template for threading, which explains the difference in the overall writing of the pdb file for the two structures. The PHIST domain spans amino acid residues 1 to 167 of the protein sequence. Both structures had a cluster of functional sites towards the C-terminal. This emphasizes crucial function of this part of the protein. Other clusters were distributed in the middle and towards the N-terminal of the structures. The PHISTb/RLP1 is a multi-domain protein, it contains the PRESAN domain (Plasmodium-RESA N-terminal) and the RESA domain, which is a fusion of other domains, namely the DnaJ domain and the DnaJ-associated X domain Pfam accession number PF0987. The distribution of these domains across the protein sequence

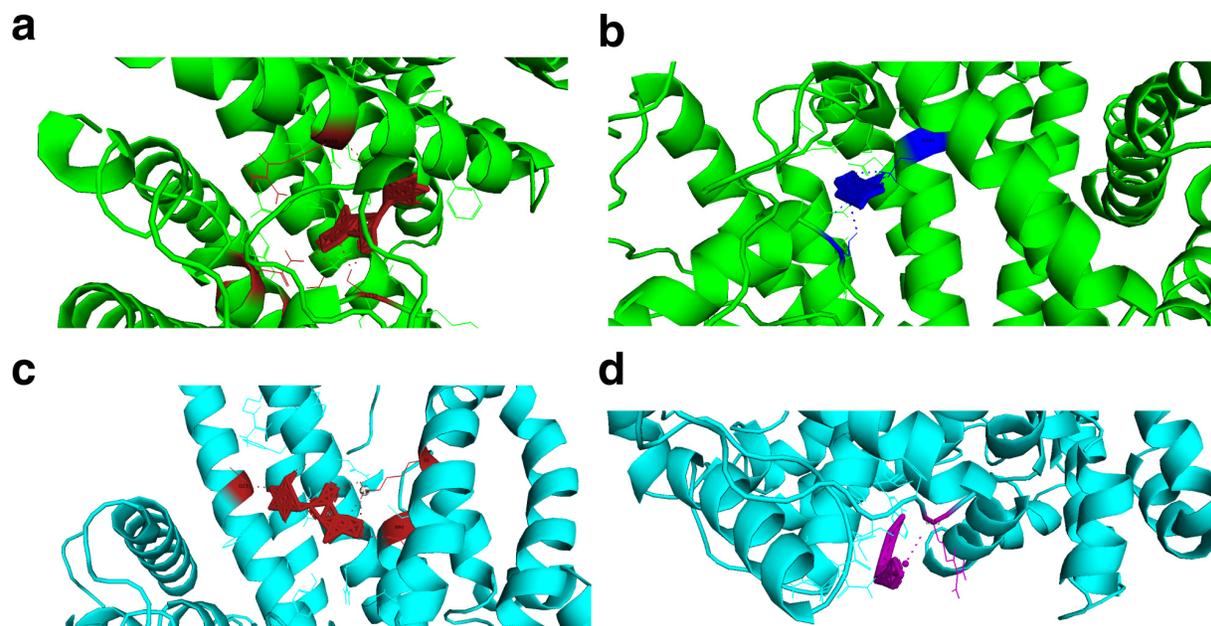


Figure 4. Visualization of protein-ligand interactions for both PHISTb/RLP1 reference and mutant proteins. (a and b) visualizations of alpha carrageenan (firebrick) and fucoidan (blue) drug compounds interacting with PHISTb/RLP1 reference protein structure (green) at specific binding sites. The interacting residues and polar contacts are shown in firebrick and blue PyMOL colours respectively. (c and d) are visualizations of alpha carrageenan (firebrick) and ghatti sulfate (purple) compounds interacting with PHISTb/RLP1 mutant protein structure binding pockets. The interacting residues and polar contacts are shown in firebrick and purple PyMOL colours, respectively. These specific interactions reveal potential inhibitory activity of these compounds against PHISTb/RLP1 proteins.

Table 8. Specific residue interactions of the drug compounds and the PHISTb/RLP1 reference protein structure as visualized in PyMOL by showing polar contacts.

The highlighted residues show optimized interactions with amino acids that were predicted by the binding site predicting tools. Alpha carrageenan, beta carrageenan, amylopectin sulfate, cyclodextrin sulfate and fucoidan compounds interacted specifically with amino acids predicted in binding sites of the reference protein structure. They showed potential ability to clock functional domains of PHISTb/RLP1 protein.

Compound	Interacting residues PHISTb/RLP1 Reference Protein
Low molecular weight heparin	N197 S161 K205
Alpha Carrageenan	S132, R315, N80
Beta Carrageenan	K376, N408, E402, E403
Dextrin sulfate	K376, E377, Q404, D406, E402
Amylopectin sulfate	S201, N197, S132, K84, K205
Phenoxyacetyl cellulose sulfate/zinc sulfate	S113, Y137, T29, K111
Ghatti sulfate	K375, G317
2,4-Diaminoanisoole sulfate	N7, S12
Cyclodextrin sulfate	S132, A133, F134
Fucoidan	K205, S132, A133
3,6-Di-O-benzoyl-D-galactal	N406, K371

Table 9. Specific amino acids of the PHISTb/RLP1 mutant protein interacting with different drug compounds at different clusters of the active sites.

The highlighted residues are specific amino acids predicted by the binding sites prediction tool. Alpha carrageenan, ghatti sulfate and 3,6-Di-O-benzoyl-D-galactal compounds interacted optimally with predicted amino acids in the mutant protein binding sites. The other compounds interacted with amino acids within the PHIST and RESA domains of the protein. These interactions revealed potential activity of these compounds as inhibitors of PHISTb/RLP1 protein from carrying out its functions.

Compound	Interacting residue PHISTb/RLP1 Mutant Protein
Low molecular weight heparin	Q237, S245, N90, T86
Alpha carrageenan	K35, N90, N214, Q237
Beta carrageenan	D392, E402, K347
Dextrin sulfate	R315, E95
Amylopectin sulfate	D356, E353
Phenoxyacetyl cellulose sulfate/zinc sulfate	S113, K261, N69, N121
Ghatti sulfate	R152
2,4-Diaminoanisoole sulfate	N121, K125, L71
Cyclodextrin sulfate	E95, K94
Fucoidan	R318, N215, K94
3,6-Di-O-benzoyl-D-galactal	K19, N18, D17

Table 10. Docking energies of each drug compound on interacting with PHISTb/ RLP1 reference and mutant proteins. The energies released were all very low, supporting the activity of the sulfated polysaccharides as drug compounds against the exported proteins. The low binding energies show high binding affinities. Generally, the docking energies in PHISTb/RLP1 mutant protein was slightly higher than in reference protein. As seen in the interaction results, the compounds are interacting with different residues in the two proteins hence the difference in the released energy.

Compound	PHISTb/RLP1 WILD TYPE		PHISTb/RLP1 MUTANT	
	affinity (kcal/mol)	RMSD upper bound	affinity (kcal/mol)	RMSD upper bound
Low molecular weight heparin	-7.9	0	-7.5	0
Alpha carrageenan	-9.4	0	-7.9	0
Beta carrageenan	-8	0	-7.2	0
Dextrin sulfate	-11.9	0	-10.3	0
Amylopectin sulfate	-6.2	0	-5.9	0
Phenoxyacetyl cellulose sulfate	-4.7	0	-5.1	0
Ghatti sulfate	-6.6	0	-6.1	0
2,4-Diaminoanisoole sulfate	-4.4	0	-4.7	0
Cyclodextrin sulfate	-7	0	-6.2	0
Fuoidan	-7.6	0	-7.2	0
3,6-Di-O-benzoyl-D-galactal	-9.4	0	-8.4	0

explains the clustering of the binding sites in the protein structure of the PHISTb/RLP1 protein.

Sulfated polysaccharides drug compounds

Sulfated polysaccharides are a wide group of biochemical molecules with therapeutic properties including anti-thrombotic, anti-viral and anti-plasmodial activities (Mourão, 2015). We narrowed down our search to sulfated polysaccharides with potential anti-malarial properties (Boyle *et al.*, 2017). The database search yielded ten compounds that are curated in PubChem compounds: Beta carrageenan, alpha carrageenan, dextrin sulfate, amylopectin sulfate, zinc sulfate, ghatti sulfate, 2,4-Diaminoanisoole sulfate, cyclodextrin sulfate, Fuoidan, 3-Aminophenylboronic acid and 3,6-Di-O-benzoyl-D-galactal. The positive control, low molecular weight heparin, was previously reported to contain activity against Plasmodium parasites (Leitgeb *et al.*, 2011). The compound was reported to contain a significant inhibition concentration (IC 50 = 1.97×10^{-2} and 3.05×10^{-3} mg/mL (-1)) against Plasmodium falciparum (Skidmore *et al.*, 2008). The compound is a glycosaminoglycan and therefore gives structural and functional comparison with the screened sulfated polysaccharides (Sun *et al.*, 2016). The compounds met most of the ADMET parameters with minor deviations. The water solubility score for most of the compounds was below six and above -4 which is the reported range for market drugs (Ritchie *et al.*, 2013). The

compounds do not cross the Blood Brain Barrier not inhibit cytochrome enzymes which supports their abilities form toxic compounds in the body (Daina & Zoete, 2016). The more negative the score of skin permeability is the less the permeability. The bioavailability score 0.55 is well correlated with most market drugs (Riyadi *et al.*, 2021). The compounds have lipophilicity score near one which is the recommended score for a drug. This depicted possibility of these molecules to reach target cells through hydrophobic interactions with lipids (Daina & Zoete, 2016). The lipophilicity scores are less than four which is recommended for optimal absorption and less toxicity of the compounds. 3,6-Di-O-benzoyl-D-galactal and Ghatti sulfate contain optimal score of lipophilicity meaning that they the physicochemical properties qualify them as drug molecules (Gao *et al.*, 2017). This is significance in showing the drugs can be metabolized and excreted safely in the body. The compounds generally have a low skin permeation ability and can hence be retained in the body longer (Potts & Guy, 1992). The solubility score showed that these compounds could be easily handled in drug testing procedures such as invitro assays of Plasmodium cultures. The compounds were tested for activity against exported protein PHISTb/RLP1 protein except for 3-Aminophenylboronic acid. Despite dextrin sulfate deviating from Lipinski's Rule of Five, it was still tested. Dextrin sulfate has been supported by previous research to contain antimalarial activity (Boyle *et al.*, 2017). The drug activity of

this compound supports its investigation as an antimalarial with further chemical modifications to suit rules of a drug compound. Carrageenan compounds antimalarial activity had previously been reported, however, the compound had to be further modified to reduce its toxicity (Adams *et al.*, 2005; Recuenco *et al.*, 2014). The compound interacted optimally with the reference and mutant protein through different hydrogen bonds as well as van der Waals forces as shown in Figure 4 below. All these compounds interacted with the PHISTb/RLP1 at different affinities and this inferred their inhibitory activity against the PHIST family of proteins.

Interactions of sulfated polysaccharides with PHISTb/RLP1

The identified drug compounds interacted with the exported protein PHISTb/RLP1. Alpha carrageenan compound interacted with both the reference and the mutant proteins. The compound shows potential inhibitory activity against exported proteins. Amylopectin sulfate, cyclodextrin sulfate, ghatti sulfate, Fucoidan and 3,6-Di-O-benzoyl-D-galactal compounds have specific interactions with the protein. The interactions of the drug compounds with specific amino acids found in binding site clusters depicted that these compounds have the potential to block the protein domains used to invade red blood cells. 2,4-Diaminoanisole sulfate and zinc sulfate showed weak

interactions with the proteins. The interactions with the mutant protein are generally weaker. The mutations found in the PHISTb/RLP1 changed the tertiary folding of the protein thus interfering with the active sites. The identified mutations I129T, T142V, Y147D, E154Q, S156H and T167I are within the PHIST domain (PDB ID: 4JLE) (Oberli *et al.*, 2014). I129T, T142V and T167I represent polar to non-polar and non-polar to polar substitutions. Y147D shows substitution of a polar amino acid to negatively charged aspartic acid residue. E154Q is a substitution from negatively charged to polar uncharged glutamine and S156H is a substitution from polar uncharged group to a positively charged amino acid. These changes affected the protein structure as well as enhancing the protein function. The binding affinity was different among the amino acids found in different structures and was shown by the difference in docking energies. The interactions of these compounds with exported proteins support the wide activity of sulfated polysaccharides against the *P. falciparum* parasite.

P. falciparum parasite uses the mechanism of exporting proteins to the host erythrocyte to enhance its virulence. These changes induced by the exported proteins include changing the physical properties of the cell and giving the cell adhesive properties. These changes enhance the pathogenesis of malaria in humans (Boddey *et al.*, 2016). Among the exported

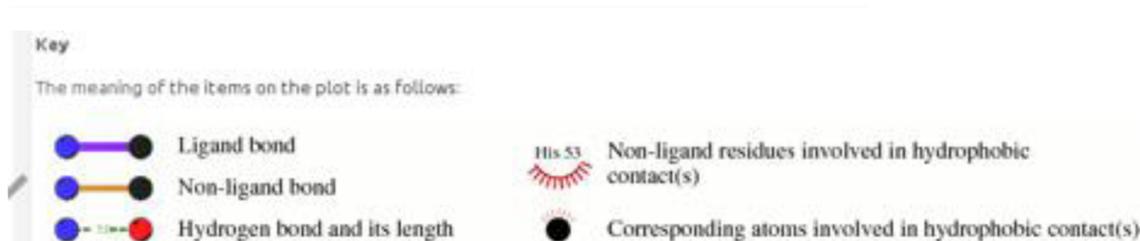
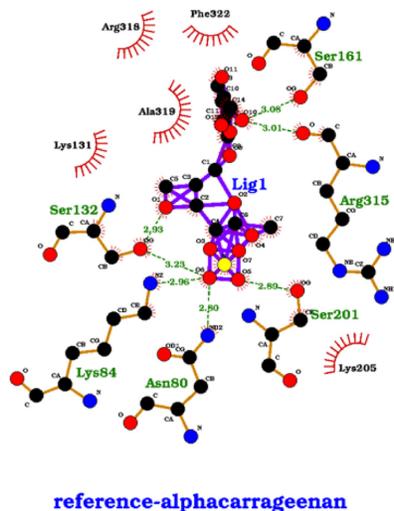


Figure 5. Ligplot showing hydrogen bonding, van der Waals forces and hydrophobic interactions between carrageenan and PHISTb/RLP1 reference protein. The compound interacts with S132, R315 and N80 amino acids located in the protein model binding site. This reveals the high affinity of interaction of this compound to the protein target.

proteins, the PHIST family, which contains 89 proteins, has been identified to play a role in making the remodeled host cell more cytoadherent (Warncke *et al.*, 2016). The key target we have studied in this research, PHISTb/RLP1 (Pf3D7_0201600) plays a key role in remodeling of the host erythrocyte to enhance malaria virulence through cytoadherence mechanism (Warncke *et al.*, 2016). The PRESAN domain present in this protein contains the *Plasmodium* protein export element (PEXEL) motif that enables the export mechanism of the protein (Boddey *et al.*, 2016; Moreira *et al.*, 2016). The interactions of the screened drug compounds with amino acids found in the functional domains of this protein reveal novel chemical inhibitors targeting exported proteins. The compounds have the ability to inhibit the PRESAN and PHIST domains from carrying out the export functions of the proteins. In *P. falciparum*, the PHIST family of proteins are expressed in the early and late ring stages, as well as trophozoites. Using PHISTb/RLP1 as a representative of this protein family, the interactions of the protein with sulfated polysaccharides infers that these compounds can deactivate exported proteins. The PHISTb proteins have been linked with controlling the *P. falciparum* Erythrocyte Membrane Protein (PfEMP1) major virulent factor of the parasite (Tarr *et al.*, 2014). The interactions of the screened drug molecules with the mutant protein of PHISTb/RLP1 show that even though the mutations change the protein folding, the functional domains are still being blocked by the chemicals. The screened sulfated polysaccharides adhere to the Lipinski Rule of Five. Their toxicity levels are very low and can be investigated further in *in vitro* analysis and clinical trials.

Conclusion

The interactions of specific sulfated polysaccharide compounds with PHISTb/RLP1 protein are the first findings showing compounds that can act against exported proteins of the *Plasmodium* parasite. These findings support further drug discovery downstream processes with these compounds as lead compounds for developing the next class of antimalarial agents. Crystal structures solving PHISTb/RLP1 and other exported proteins is recommended to enable more insight into the implications of structural variants on the protein structure and functions.

Data availability

Open Science Framework: Novel Characterization of Sulfated Polysaccharides Activity against Virulent *Plasmodium falciparum* PHIST Proteins, <https://doi.org/10.17605/OSF.IO/YFQAZ> (Mutisya, 2020).

This project contains the following underlying data:

- Sequenced reads of PHISTb/RLP1 acquired through Sanger sequencing
- Multiple Sequence Alignment files
- Protein .pdb files (of both PHISTb/RLP1 reference and mutant protein structures)
- Sulfated polysaccharides .sdf and .pdbqt files.

Data are available under the terms of the [Creative Commons Zero “No rights reserved” data waiver](#) (CC0 1.0 Public domain dedication).

References

- Adams Y, Smith SL, Schwartz-Albiez R, *et al.*: **Carrageenans Inhibit the in Vitro Growth of *Plasmodium Falciparum* and Cytoadhesion to CD36.** *Parasitol Res.* 2005; **97**(4): 290–94.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Alzohairy A: **(3) Building a Multiple Sequence Alignment (8) Jalview Basics in More Details.** 2014.
[Publisher Full Text](#)
- AutoDock Vina.**
[Reference Source](#)
- Boddey JA, O'Neill MT, Lopaticki S, *et al.*: **Export of Malaria Proteins Requires Co-Translational Processing of the PEXEL Motif Independent of Phosphatidylinositol-3-Phosphate Binding.** *Nat Commun.* 2016; **7**: 10470.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Boyle MJ, Skidmore M, Dickerman B, *et al.*: **Identification of Heparin Modifications and Polysaccharide Inhibitors of *Plasmodium Falciparum* Merozoite Invasion That Have Potential for Novel Drug Development.** *Antimicrob Agents Chemother.* 2017; **61**(11): e00709–17.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Boyle MJ, Richards JS, Gilson PR, *et al.*: **Interactions with Heparin-like molecules during Erythrocyte Invasion by *Plasmodium Falciparum* Merozoites.** *Blood.* 2010; **115**(22): 4559–4568.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Brenke R, Kozakov D, Chuang GY, *et al.*: **Fragment-Based Identification of Druggable ‘hot Spots’ of Proteins Using Fourier Domain Correlation Techniques.** *Bioinformatics.* 2009; **25**(5): 621–27.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- CLC Bio: **CLC Main Workbench Manual.** *Software Manual.* 2014.
[Reference Source](#)
- Daina A, Michielin O, Zoete V: **ILOGP: A simple, robust, and efficient description of N-Octanol/Water partition coefficient for drug design using the GB/SA approach.** *J Chem Inf Model.* 2014; **54**(12): 3284–3301.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Daina A, Michielin O, Zoete V: **SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules.** *Sci Rep.* 2017; **7**(1): 42717.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Daina A, Zoete V: **A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules.** *ChemMedChem.* 2016; **11**(11): 1117–21.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dondorp AM, Nosten F, Yi P, *et al.*: **Artemisinin Resistance in *Plasmodium Falciparum* Malaria.** *N Engl J Med.* 2012; **361**(5): 455–67.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Edgar RC: **MUSCLE: A Multiple Sequence Alignment Method with Reduced Time and Space Complexity.** *BMC Bioinformatics.* 2004; **5**: 113.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Gao Y, Gesenberg C, Zheng W: **Oral formulations for preclinical studies.** *Developing Solid Oral Dosage Forms.* 2017; 455–495.
[Publisher Full Text](#)
- Gimenez BG, Santos MS, Ferrarini M, *et al.*: **Evaluation of Blockbuster Drugs under the Rule-of-Five.** *Pharmazie.* 2010; **65**(2): 148–52.
[PubMed Abstract](#)
- Glushakova S, Busse BL, Garten M, *et al.*: **Exploitation of a Newly-Identified Entry Pathway into the Malaria Parasite-Infected Erythrocyte to Inhibit Parasite Egress.** *Sci Rep.* 2017; **7**(1): 12250.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hall TA: **BioEdit: A User-Friendly Biological Sequence Alignment Editor and**

Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series.* 1999; **41**: 95–98.

[Reference Source](#)

Kamau E, Alemayehu S, Feghali KC, *et al.*: **Multiplex QPCR for Detection and Absolute Quantification of Malaria.** *PLoS One.* 2013; **8**(8): e71539.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Kim S, Thiessen PA, Bolton EE, *et al.*: **PubChem Substance and Compound Databases.** *Nucleic Acids Res.* 2016; **44**(D1): D1202–13.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Ko J, Park HH, Heo L, *et al.*: **GalaxyWEB Server for Protein Structure Prediction and Refinement.** *Nucleic Acids Res.* 2012; **40**(Web Server issue): W294–97.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Laskowski RA, MacArthur MW, Moss DS, *et al.*: **PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures.** *J Appl Cryst.* 1993; **26**(2): 283–91.

[Publisher Full Text](#)

Leitgeb AM, Blomqvist K, Cho-Ngwa F, *et al.*: **Low Anticoagulant Heparin Disrupts *Plasmodium falciparum* Rosettes in Fresh Clinical Isolates.** *Am J Trop Med Hyg.* 2011; **84**(3): 390–96.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Lovell SC, Davis LW, Arendall WB, *et al.*: **Structure Validation by Calpha Geometry: Phi, Psi and Cbeta Deviation.** *Proteins.* 2003; **50**(3): 437–50.

[PubMed Abstract](#) | [Publisher Full Text](#)

Maier AG, Rug M, O'Neill MT, *et al.*: **Exported Proteins Required for Virulence and Rigidity of *Plasmodium falciparum*-Infected Human Erythrocytes.** *Cell.* 2008; **134**(1): 48–61.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Marques J, Vilanova E, Mourão PAS, *et al.*: **Marine Organism Sulfated Polysaccharides Exhibiting Significant Antimalarial Activity and Inhibition of Red Blood Cell Invasion by *Plasmodium*.** *Sci Rep.* 2016; **6**: 24368.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Moreira CK, Naissant B, Coppi A, *et al.*: **The *Plasmodium* PHIST and RESA-Like Protein Families of Human and Rodent Malaria Parasites.** *PLoS One.* 2016; **11**(3): e0152510.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Morris GM, Huey R, Lindstrom W, *et al.*: **AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.** *J Comput Chem.* 2009; **30**(16): 2785–91.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Mourão PAS: **Perspective on the Use of Sulfated Polysaccharides from Marine Organisms as a Source of New Antithrombotic Drugs.** *Mar Drugs.* 2015; **13**(5): 2770–84.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Mutisya J: **Novel Characterization of Sulfated Polysaccharides Activity against Virulent *Plasmodium falciparum* PHISTb/RLP1 Protein.** 2020. <http://www.doi.org/10.17605/OSF.IO/YFQAZ>

Ngan CH, Hall DR, Zerbe B: **FTSite: high accuracy detection of ligand binding sites on unbound protein structures.** *Bioinformatics.* 2012; **28**(2): 286–87.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

O'Boyle NM, Banck M, James CA, *et al.*: **Open Babel: An open chemical toolbox.** *J Cheminform.* 2011; **3**(10): 33.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Oberli A, Slater LM, Cutts E, *et al.*: **A *Plasmodium falciparum* PHIST protein binds the virulence factor PfEMP1 and comigrates to knobs on the host cell surface.** *FASEB J.* 2014; **28**(10): 4420–33.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Oberli A, Zurbrugg L, Rusch S, *et al.*: ***Plasmodium falciparum* Plasmodium helical interspersed subtelomeric proteins contribute to cytoadherence and anchor *P. falciparum* erythrocyte membrane protein 1 to the host cell cytoskeleton.** *Cell Microbiol.* 2016; **18**(10): 1415–28.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Potts RO, Guy RH: **Predicting skin permeability.** *Pharm Res.* 1992; **9**(5): 663–669.

[PubMed Abstract](#) | [Publisher Full Text](#)

Recuenco FC, Takano R, Chiba S, *et al.*: **Lambda-carrageenan treatment exacerbates the severity of cerebral malaria caused by *Plasmodium berghei* ANKA in BALB/c mice.** *Malar J.* 2014; **13**(1): 487.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Ritchie TJ, Macdonald SJ, Peace S, *et al.*: **Increasing small molecule drug developability in sub-optimal chemical space.** *MedChemComm.* 2013; **4**(4): 673.

[Publisher Full Text](#)

Riyadi PH, Romadhon SID, Kurniasih RA, *et al.*: **SwissADME predictions of pharmacokinetics and drug-likeness properties of small molecules present in *spirulina platensis*.** *IOP Conference Series: Earth and Environmental Science.* 2021; **890**(1): 012021.

[Publisher Full Text](#)

Saiwaew S, Sritabal J, Piaraksa N, *et al.*: **Effects of sevuparin on rosette formation and cytoadherence of *Plasmodium falciparum* infected erythrocytes.** *PLoS One.* 2017; **12**(3): e0172718.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Sargeant TJ, Marti M, Caler E, *et al.*: **Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites.** *Genome Biol.* 2006; **7**(2): R12.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Somner EA, Black J, Pasvol G: **Multiple human serum components act as bridging molecules in rosette formation by *Plasmodium falciparum*-infected erythrocytes.** *Blood.* 2000; **95**(2): 674–82.

[PubMed Abstract](#)

Skidmore MA, Dumax-Vorzet AF, Guimond SE, *et al.*: **Disruption of rosetting in plasmodium falciparum malaria with chemically modified heparin and low molecular weight derivatives possessing reduced anticoagulant and other serine protease inhibition activities.** *J Med Chem.* 2008; **51**(5): 1453–1458.

[PubMed Abstract](#) | [Publisher Full Text](#)

Sun W, Saldaña MD, Fan L, *et al.*: **Sulfated polysaccharide heparin used as carrier to load hydrophobic lappaconitine.** *Int J Biol Macromol.* 2016; **84**: 275–280.

[PubMed Abstract](#) | [Publisher Full Text](#)

Takala-harrison S, Jacob CG, Arze C, *et al.*: **Independent emergence of artemisinin resistance mutations among *Plasmodium falciparum* in Southeast Asia.** *J Infect Dis.* 2015; **211**(5): 670–9.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Tarr SJ, Moon RW, Hardege I, *et al.*: **A conserved domain targets exported PHISTb family proteins to the periphery of Plasmodium infected erythrocytes.** *Mol Biochem Parasitol.* 2014; **196**(1): 29–40.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Trott O, Olson AJ: **Autodock Vina: Improving the Speed and Accuracy of Docking with a new scoring function, efficient optimization, and multithreading.** *J Comput Chem.* 2010; **31**(2): 455–61.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Untergasser A, Cutcutache I, Koressaar T, *et al.*: **Primer3—new capabilities and interfaces.** *Nucleic Acids Res.* 2012; **40**(15): e115.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Veron V, Simon S, Carme B: **Multiplex real-time PCR detection of *P. falciparum*, *P. vivax* and *P. malariae* in human blood samples.** *Exp Parasitol.* 2009; **121**(4): 346–51.

[PubMed Abstract](#) | [Publisher Full Text](#)

Vogt AM, Pettersson F, Moll K, *et al.*: **Release of sequestered malaria parasites upon injection of a glycosaminoglycan.** *PLoS Pathog.* 2006; **2**(9): e100.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Wallace AC, Laskowski RA, Thornton JM: **Ligplot: A program to generate schematic diagrams of protein-ligand interactions.** *Protein Eng.* 1995; **8**(2): 127–134.

[PubMed Abstract](#) | [Publisher Full Text](#)

Warncke JD, Vakonakis I, Beck HP: **Plasmodium Helical Interspersed Subtelomeric (PHIST) Proteins, at the Center of Host Cell Remodeling.** *Microbiol Mol Biol Rev.* 2016; **80**(4): 905–27.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Waterhouse AM, Procter JB, Martin DMA, *et al.*: **Jalview Version 2—a multiple sequence alignment editor and analysis workbench.** *Bioinformatics.* 2009; **25**(9): 1189–91.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

WHO: **World Malaria Report 2019.** World Health Organization.

[Reference Source](#)

World Health Organization (WHO): **World Malaria Day 2018 'Ready to Beat Malaria'.** Malaria, no. March: 1–3. 2018.

[Reference Source](#)

Wu S, Zhang Y: **LOMETS: a local meta-threading-server for protein structure prediction.** *Nucleic Acids Res.* 2007; **35**(10): 3375–82.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Xu D, Zhang Y: **Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization.** *Biophys J.* 2011; **101**(10): 2525–34.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Yang J, Roy A, Zhang Y: **Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment.** *Bioinformatics.* 2013; **29**(20): 2588–95.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Zheng W, Zhang C, Wuyun Q, *et al.*: **LOMETS2: improved meta-threading server for fold-recognition and structure-based function annotation for distant-homology proteins.** *Nucleic Acids Res.* 2019; **47**(W1): W429–36.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status:  

Version 2

Reviewer Report 11 May 2022

<https://doi.org/10.5256/f1000research.122945.r136370>

© 2022 Macete E. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Eusébio Macete

Centro de Investigação em Saúde de Manhiça (CISM), Maputo, Mozambique

Thanks for the responses, from my point of view the article can be indexed.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 03 May 2022

<https://doi.org/10.5256/f1000research.122945.r136371>

© 2022 Madariaga A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Abraham Madariaga 

Institute of Chemistry, National Autonomous University of Mexico, Mexico City, Mexico

I agree with the new version of the manuscript, the authors have addressed most of the reviewers' comments. I have no further comments.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular modeling including docking and molecular dynamics simulations; QSAR studies, homology modeling

I confirm that I have read this submission and believe that I have an appropriate level of

expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 01 March 2022

<https://doi.org/10.5256/f1000research.29544.r121686>

© 2022 Madariaga A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Abraham Madariaga 

Institute of Chemistry, National Autonomous University of Mexico, Mexico City, Mexico

The manuscript describes the modeling of a protein and its mutant from *Plasmodium falciparum*, and the assessment of some ligands with these models by molecular docking analysis. The mutant was built based on experimental data from several samples. Both protein models were validated using online tools. Although the authors used several online tools for assessing the structural performance of the proteins, the mutant model is lacking confidence in its tridimensional structure. Furthermore, the docking protocol is not fully described. The results are too preliminary to suggest a potential activity of the selected ligands against PHISTb/RLP1 proteins. I would recommend major revision prior to indexing.

The modeled mutant protein resulted with a very different tridimensional configuration to the reference protein. For better prediction and assessment of a homology model, it is suggested to perform a molecular dynamics simulation of at least 50 ns in order to let the structure correctly minimize. Otherwise, the predicted protein may have a non-biological relevance.

It is not clear why the authors select the fourth model of each structure (reference and mutant) for energy minimization. Does LOMETs give a score? Furthermore, the TM score is significantly high for both structures, meaning a “perfect match” between the initial model and the generated model. It is surprising that the mutant model gets a TM as similar as the reference model, so the mutant has no apparent structural changes despite having a dozen different amino acids. Nevertheless we can see in Fig2a-b big differences among those structures. How do you explain this?

Rule of five is not synonymous with toxicity. Rule of five gives an empirical approximation of the pharmacokinetics of a molecule, in terms of absorption. Ro5 is a simple approach intended to select molecules with druggability properties. To examine toxicity of molecules, many other methods can be used (Ames mutagenicity prediction, ADME-Tox predictions, LD50, etc) in different online platforms.

Docking protocol. Were the ligands used in their salt forms? For better results, the ligands should undergo a “washing” procedure that includes removing contra ions, adjusting charges, and geometric minimizing process. It is not described in the methodology.

Regarding the gridbox, what size does it have? Depending on the size, the ligands (in this case the ligands are large molecules) are allowed to freely rotate or not. This may bias the docking results.

On the other hand, if a positive control is not used as a part of the docking experiments, it is difficult to argue that a molecule will have a potent activity, such in the case of the manuscript. It is preferable to suggest only a possible activity.

For completeness of the discussion, a figure depicting the ligand-protein interaction of the ligands with better affinity would be great.

Minor

Page 6, Section "Preparation of proteins and ligands, ...", line 8: Please change "autodock tolls" for "autodock tools".

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

No

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular modeling including docking and molecular dynamics simulations; QSAR studies, homology modeling

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 18 Apr 2022

Jennifer Mutisya, United States Army Medical Research Directorate-Africa (USAMRD-A), Kenya Medical Research Institute (KEMRI) / Walter Reed Project, Kisumu, Kenya

Comment: The manuscript describes the modeling of a protein and its mutant from *Plasmodium falciparum*, and the assessment of some ligands with these models by molecular docking analysis. The mutant was built based on experimental data from several samples. Both protein models were validated using online tools. Although the authors used several online tools for assessing the structural performance of the proteins, the mutant model is lacking confidence in its tridimensional structure. Furthermore, the docking protocol is not fully described. The results are too preliminary to suggest a potential activity of the selected ligands against PHISTb/RLP1 proteins. I would recommend major revision prior to indexing.

Response: All the tools used in modeling and assessment of the protein structures are open source tools with algorithms that are assessed and published. Furthermore, they have been used and cited in other scientific research prior to this work hence able to provide significant scientific data as documented in the provided links and cited publications for each tool used. The docking protocol details are included; that is preparation of ligand files in MGL tools, grid box parameters and python script used.

Comment: The modeled mutant protein resulted with a very different tridimensional configuration to the reference protein. For better prediction and assessment of a homology model, it is suggested to perform a molecular dynamics simulation of at least 50 ns in order to let the structure correctly minimize. Otherwise, the predicted protein may have a non-biological relevance.

Response: Energy minimization carried out using ModRefiner an algorithm for atomic-level, high-resolution protein structure refinement, which can start from either C-alpha trace. The models underwent energy minimization using this tool prior to docking experiments. Molecular dynamics simulation of at least 50ns were not possible to carry out due to minimal expertise in a tool that can do that. However, after energy minimization using ModRefiner the models quality improved and were fit for docking simulations.

Comment: It is not clear why the authors select the fourth model of each structure (reference and mutant) for energy minimization. Does LOMETs give a score? Furthermore, the TM score is significantly high for both structures, meaning a "perfect match" between the initial model and the generated model. It is surprising that the mutant model gets a TM as similar as the reference model, so the mutant has no apparent structural changes despite having a dozen different amino acids. Nevertheless, we can see in Fig2a-b big differences among those structures. How do you explain this?

Response: The fourth model was chosen after screening the Ramachandran score for the five suggested models hence the fourth model had the highest scores. Lomets does not provide scores for models since they are ranked from number 1 to 5, but for the template selected. Structures match because the middle template selected by the algorithm was similar in both reference and mutant modelling jobs. The PHIST domain is the only present crystal structure for this protein family in PDB. This is why the TM score matches closely. The difference between the two structures can be seen in the alpha-helix folding which is significantly distinct and postulated to be an effect of the

SNPs identified.

Comment: Rule of five is not synonymous with toxicity. Rule of five gives an empirical approximation of the pharmacokinetics of a molecule, in terms of absorption. Ro5 is a simple approach intended to select molecules with druggability properties. To examine the toxicity of molecules, many other methods can be used (Ames mutagenicity prediction, ADME-Tox predictions, LD50, etc) in different online platforms.

Response: ADMET analysis data was included using SwissADME tool as suggested.

Comment: Docking protocol. Were the ligands used in their salt forms? For better results, the ligands should undergo a "washing" procedure that includes removing contra ions, adjusting charges, and geometric minimizing process. It is not described in the methodology.

Response: The ligand files were converted using the Open Babel tool and prepared using MGL tools as described in the text. This editing enabled the removal of contradicting s ions prior to docking.

Comment: Regarding the grid box, what size does it have? Depending on the size, the ligands (in this case the ligands are large molecules) are allowed to freely rotate or not. This may bias the docking results.

Response: grid box dimensions were added in the main text.

Comment: On the other hand, if positive control is not used as a part of the docking experiments, it is difficult to argue that a molecule will have a potent activity, such in the case of the manuscript. It is preferable to suggest only a possible activity.

Response: Included low molecular weight heparin as the positive control compound since it is a well-researched sulfated polysaccharide with anti-malarial properties.

Comment: For completeness of the discussion, a figure depicting the ligand-protein interaction of the ligands with better affinity would be great.

Response: Included a figure in the discussion

Competing Interests: No competing interests were disclosed.

Reviewer Report 31 January 2022

<https://doi.org/10.5256/f1000research.29544.r100719>

© 2022 Macete E. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Eusébio Macete**

Centro de Investigação em Saúde de Manhiça (CISM), Maputo, Mozambique

My first comment is regarding the selection of the sites, the authors did not mention how the sites were selected.

Were the samples collected for this specific study? Or is this study a secondary analysis? If yes, what was the primary study?

Second, in the article the authors mentioned that the samples were tested for plasmodium falciparum, so, how were patients that had positive results treated? If that was the case?

In the method paragraph, the study design is not clear. How were the samples selected etc.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am specialist in public health working mostly in malaria drug and vaccines.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 07 Feb 2022

Jennifer Mutisya, United States Army Medical Research Directorate-Africa (USAMRD-A), Kenya Medical Research Institute (KEMRI) / Walter Reed Project, Kisumu, Kenya

Concerning the comments on sample sites selection and collection:
The samples used in this study were not collected specifically for this secondary study but

were used on basis of availability. The samples were collected for a bigger project that studies drug resistance patterns from the identified regions. The sampling for this study was hence non-randomly conducted from sample inventory of archived whole blood samples. This study was a sub-study in the Malaria Drug Resistance Laboratory and not the primary study of the institution. The sample sites were selected according to different Malaria endemic zones in Kenya and were guided by the proposal of the main study which studies epidemiology patterns of the disease in Kenya.

The patients who tested positive were treated with recommended Artemesinin combined therapies at day zero of testing and follow up samples were collected after 7, 14 and 28 days. These were done according to the objectives of a different study. The current study did not need the treatment data.

Competing Interests: No competing interest

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research