Analyzing the Essential Proteins Set of Plasmodium Falciparum PF3D7 for Novel Drug Targets Identification Against Malaria

Fawad Ali
Abdul Wali Khan University Mardan

Hira Wall
Abdul Wali Khan University Mardan

Saadia Jan
Abdul Wali Khan University Mardan

Muneeba Aslam
Abdul Wali Khan University Mardan

Imtiaz Ahmad
Abdul Wali Khan University Mardan

Sahib Gul Afridi
Abdul Wali Khan University Mardan

Suliaman Shams
Abdul Wali Khan University Mardan

Asifullah Khan (✉ asifullah111@gmail.com)
Abdul Wali Khan University Mardan  https://orcid.org/0000-0002-1444-4249

Research

Keywords: Plasmodium falciparum, malaria, comparative proteomics, anti-malarial therapeutics

DOI: https://doi.org/10.21203/rs.3.rs-383876/v1

License: ☺ ① This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

**Background:** *Plasmodium falciparum* is an obligate intracellular parasite of humans that causes malaria. *P. falciparum* is a major public health threat to human life responsible for high mortality. Currently, the risk of multi-drug resistance of *P. falciparum* is rapidly increasing. There is a need to address new anti-malarial therapeutics strategies to combat the drug-resistance threat.

**Methods:** We retrieved the *P. falciparum* essential proteins from the recently published studies. Pathogen essential proteins were initially scanned against human host and its gut microbiome proteome sets by comparative proteomics analyses. The human host non-homologs essential proteins of *P. falciparum* were additionally analyzed for druggability potential via *in silico* methods to possibly identify novel therapeutic targets.

**Results:** The analyses identified six *P. falciparum* essential and human host non-homolog proteins that follow the key druggability features. These druggable targets have not catalogued so far in the Drugbank repository. These prioritized proteins seem novel and promising drug targets against *P. falciparum* due to their key protein-protein interactions features in pathogen-specific biological pathways and to hold appropriate drug-like molecule binding pockets.

**Conclusion:** The prioritized protein targets may worthy to test in malarial drug discovery program to overcome the anti-malarial resistance issues. The *in-vitro* and *in-vivo* studies might be promising for additional validation of these prioritized lists of drug targets against malaria.

**Background**

Malaria is a life-threatening infectious disease caused by parasitic protozoan plasmodium. It is a vector born disease, transmitted to humans through the bite of infected carrier female anopheles mosquitoes. Among five parasite species that cause malaria in humans, the two species, i.e. *P. falciparum* and *P. vivax* have the greatest threat to human life [1]. The *P. falciparum*, a unicellular protozoan, belongs to the family Plasmodiidae and lies in the phylum Apicomplexa [2]. *P. falciparum* is the most deadly parasite and the major cause of morbidities and mortalities. About 50% of all the malarial cases caused due to *P. falciparum* [3, 4]. The *P. falciparum* alone responsible for almost all malaria-inflicted deaths in Sub-Saharan Africa, with the continent bearing over 90% of the global *P. falciparum* burden [5]. Asia is second to Africa in terms of malaria prevalence. In 2019, the WHO estimated 229 million malaria cases and about 409,000 deaths due to malaria worldwide [6]. More than 85% of confirmed recorded cases and deaths in Asia occurred in India, Indonesia, Myanmar, and Pakistan [7].

Many approved drugs, including chloroquine and the most recent anti-malarial drug Artemisinin, were found to be ineffective against *P. falciparum* [8]. Resistance to chloroquine was first observed in Thailand in 1957 and the Colombian-Venezuelan border in 1959 [9]. Drug resistance had spread across Sub-Saharan Africa by 1988 and today chloroquine is no longer effective in almost all parts of the world [10]. Specific polymorphisms in the *P. falciparum* chloroquine resistance transporter (PfCRT) have been found to be strongly linked to chloroquine resistance [11]. Likewise, the artemisinin resistance is observed in many regions of the world and was first documented in 2008 in the Thailand-Cambodia border regions [12, 13]. The artemisinin resistance was first associated with delayed parasitic clearance after three days of artemisinin monotherapy. Many studies have reported that artemisinin resistance emerged due to polymorphism in the *pfk13* gene [14, 15].

The indispensable proteins of *P. falciparum* recently explored from well experimental approaches in some studies. These repositories are promising to identify suitable targets to overcome the drug-resistant *P. falciparum* infection [16, 17]. The Zhang et al., (2018) experimentally analyzed 5399 genes and identified 2680 as essential for optimal growth of *P. falciparum* during asexual blood stages. These essential genes coding the *P. falciparum* vital proteins, including drug targets and potential vaccine candidates. Besides, there are over 1,000 of *plasmodium*-conserved essential genes with unknown biological functions so far. In the current study, these essential proteins were retrieved from the two recently published studies and assessed for druggable potential based on comparative proteomics, protein-protein interactions and drug-like molecules binding potential.

**Methodology**

The methodological layout of the current study is depicted in Fig. 1.

**Retrieval of non-paralogous essential proteins**
The essential proteins of *P. falciparum* strain 3D7 were retrieved from two recently published studies [i.e. 16, 17]. The paralogous proteins were removed by CD-HIT clustering analysis with 80% sequence similarity cutoff [18].

**Human host non-homologous and virulent proteins identification**

The *P. falciparum* essential proteins non-homolog to human host were identified by comparative sequence analyses via BLASTp tool [19]. The threshold values of 35% query coverage and sequence identity were set during this analysis [20]. The proteins having significant similarity with human proteome were discarded and the remaining non-homologs were shortlisted for further analysis. The non-homology search against human gut microbiota proteins sequences was also carried out with a threshold cutoff, i.e. *E* value 0.001 [21, 22]. The Vectors database was screened for *P. falciparum* 3D7 virulent proteins annotation. The Vectors database contains 5304 virulent proteins data from various parasites including *P. falciparum* [23].

**Drugbank database scanning**

The shortlisted essential proteins of *P. falciparum* from above analyses were scanned against the Drugbank database to identify novel drug targets with 60% query coverage and percent identity threshold of BLASTp [24].

**Structure homologs search**

The proteins data bank (PDB) was screened to identify the homologous 3D structures of pathogenic proteins [25]. The pathogen proteins were BLAST against the entire PDB database entries with 60% percent sequence identity and query coverage [26, 22]. The pathogen sequences having < 60% homology were modeled with Swiss Model [27] and verified by ERRAT [28] and RAMPAGE [29].

**Druggability analysis**

The prioritized list of essential proteins shortlisted from above analyses were tested for druggability potential. The drug-like molecules binding pockets of the targets were identified by PockDrug-server [30]. The subcellular localization was performed with CELLO v.2.5 [31]. The protein-protein interaction (PPI) and molecular weight analysis were performed [32]. The PPI analysis was performed by STRING database and the Hub proteins were identified based on node degree \((K \geq 5)\) that represent the number of interactions [33].

**Results**

**Subtractive proteomic analyses**

Essential genes perform key cellular functions for the survival of pathogens [34]. The *P. falciparum* strain 3D7 essential genes information were obtained from the recently published articles [16, 17] and total 3380 essential protein were identified. These proteins sequences were retrieved from Uniprot. Seven paralogous protein sequences were excluded by CD-HIT analysis and the remaining 3373 sequences were considered for downstream analysis (Supplementary Table S1). The non-paralogous protein sequences were subjected to BLASTp against human proteome as well as human gut microbiome proteome data with threshold parameters. The gut flora is helpful to the host in many ways like vitamins biosynthesis and absorption of short chain-fatty acids [35]. The unintended inhibition of gut microbe leads to a decrease in gut flora and colonization of pathogenic bacteria in the host gut [36]. The comparative sequences analyses identified total 183 *P. falciparum* strain 3D7 essential proteins non-homologous to human as well as human's gut flora proteins (Supplementary Table S2). The homology screening against Drugbank repository for these 183 proteins inferred no homology with already reported drug targets deposited in drugbank database. The five among these 183 were annotated as *P. falciparum* 3D7 virulent proteins during screening of Vectors database (Table 1).

<table>
<thead>
<tr>
<th>Protein Accession number</th>
<th>Protein Name</th>
<th>Nature of protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q9TY99_PLAF7</td>
<td>knob-associated histidine-rich protein</td>
<td>Virulent</td>
</tr>
<tr>
<td>Q8I5P1_PLAF7</td>
<td>Cell traversal protein for ookinetes and sporozoites (CelTOS), putative</td>
<td>Virulent</td>
</tr>
<tr>
<td>C6KTB1_PLAF7</td>
<td>4-methyl-5(B-hydroxyethyl)-thiazol monophosphate biosynthesis enzyme</td>
<td>Virulent</td>
</tr>
<tr>
<td>Q8I6Z5_PLAF7</td>
<td>plasmepsin V</td>
<td>Virulent</td>
</tr>
<tr>
<td>Q8IDM6_PLAF7</td>
<td>nucleoside transporter 1</td>
<td>Virulent</td>
</tr>
</tbody>
</table>
Druggability analyses

The Drugbank non-homologous proteins were prioritized for downstream druggability analyses. The subcellular localization is one of the key aspect of druggability and the cytoplasmic proteins are considered as suitable drug targets [37, 38]. The 15 *P. falciparum* proteins among shortlisted prioritized targets were annotated as cytoplasmic proteins. The proteins 3D structures identified by Swiss model were validated with the ERRAT tool with quality factor score of > 50, which is accepted as high quality model [39]. Ramachandran plot identified 80–90% of modeled proteins residues in the allowed region assuring good quality structure modeling of the target proteins (Table 2). Finally, six (06) proteins were prioritized on the basis of (i) pockdrug probability score ≥0.5 [30], (ii) ERRAT quality factor ≥ 90 [28], and (iii) protein-protein interaction node degree i.e. K ≥ 5 [33] (Table 2) (Fig. 2). These six prioritized targets are speculating to hold promising druggable pockets to anchor small drug-like molecules and act as indispensable hub proteins in *P. falciparum* metabolic network.
Table 2
Druggability analysis of shortlisted 15 cytoplasmic proteins of *P. falciparum* 3D7 essential and human host non-homolog proteins. The ‘steric’ indicates shortlisted 06 best druggable targets.

<table>
<thead>
<tr>
<th>Uniprot accession numbers</th>
<th>PDB homolog ID's</th>
<th>Protein Names</th>
<th>ERRAT Quality Factor</th>
<th>Ramachandran plot (residues in favored region %)</th>
<th>QMEAN &gt; -4</th>
<th>Pock Drug score &gt; 0.5 (Residues in pocket)</th>
<th>Molecular weight (Dalton)</th>
<th>Node degree (K) ≥ 5 (STRING analysis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O97259_PLAF7</td>
<td>3fga.1.C</td>
<td>Serine/threonine-protein phosphatase</td>
<td>99.3103</td>
<td>94.75%</td>
<td>-2.11</td>
<td>0.24 (15)</td>
<td>35765.00</td>
<td>7.45</td>
</tr>
<tr>
<td>C0H4F3_PLAF7*</td>
<td>5cfl.1.A</td>
<td>Bis(5'-nucleosyl)tetraphosphatase [asymmetrical]</td>
<td>96.1832</td>
<td>97.20%</td>
<td>-0.37</td>
<td>0.98 (8)</td>
<td>17750.19</td>
<td>6.36</td>
</tr>
<tr>
<td>Q8I3A9_PLAF7</td>
<td>4nc6.1.A</td>
<td>GTPase-activating protein, putative</td>
<td>90.2778</td>
<td>90.71%</td>
<td>-3.37</td>
<td>0.99 (20)</td>
<td>50301.05</td>
<td>4.36</td>
</tr>
<tr>
<td>Q8I2N2_PLAF7*</td>
<td>4lac.1.C</td>
<td>Serine/threonine-protein phosphatase</td>
<td>93.5484</td>
<td>97.54%</td>
<td>-0.72</td>
<td>0.91 (14)</td>
<td>36386.98</td>
<td>6.18</td>
</tr>
<tr>
<td>Q8I082_PLAF7</td>
<td>1ryt.1.A</td>
<td>Rifin</td>
<td>69.7368</td>
<td>93.41%</td>
<td>-2.95</td>
<td>0.32 (10)</td>
<td>40648.33</td>
<td>3.82</td>
</tr>
<tr>
<td>Q8I1U3_PLAF7</td>
<td>2c9o.1.A</td>
<td>RuvB-like helicase</td>
<td>88.7505</td>
<td>92.14%</td>
<td>-2.41</td>
<td>53405.94</td>
<td>18244.08</td>
<td>6.55</td>
</tr>
<tr>
<td>Q8I599_PLAF7*</td>
<td>2ibj.1.A</td>
<td>Cytochrome b5, putative</td>
<td>94.5205</td>
<td>95.00%</td>
<td>-0.14</td>
<td>1 (19)</td>
<td>14234.31</td>
<td>8.18</td>
</tr>
<tr>
<td>Q8IDG3_PLAF7</td>
<td>6muw.1.C</td>
<td>Proteasome subunit alpha type</td>
<td>92.2727</td>
<td>93.72%</td>
<td>-1.63</td>
<td>27947.73</td>
<td>34952.82</td>
<td>10</td>
</tr>
<tr>
<td>Q8IM27_PLAF7*</td>
<td>2a22.2.A</td>
<td>Vacuolar protein sorting-associated protein 29</td>
<td>92.3977</td>
<td>95.63%</td>
<td>0.23</td>
<td>0.96 (16)</td>
<td>21774.02</td>
<td>5.09</td>
</tr>
<tr>
<td>Q8IM19_PLAF7*</td>
<td>6kax.1.B</td>
<td>Multifunctional methyltransferase subunit TRM112, putative</td>
<td>92.233</td>
<td>91.80%</td>
<td>-1.87</td>
<td>1 (10)</td>
<td>14234.31</td>
<td>8.18</td>
</tr>
<tr>
<td>Q8ILZ5_PLAF7</td>
<td>6y4i.1.A</td>
<td>ER membrane protein complex subunit 2</td>
<td>100</td>
<td>95.92%</td>
<td>-2.03</td>
<td>0.97 (16)</td>
<td>34952.82</td>
<td>4.36</td>
</tr>
<tr>
<td>C6S3I2_PLAF7</td>
<td>6t69.1.A</td>
<td>MORN repeat-containing protein 5</td>
<td>68.3333</td>
<td>87.14%</td>
<td>-0.84</td>
<td>0.98 (12)</td>
<td>17938.95</td>
<td>1.33</td>
</tr>
<tr>
<td>Q8IKK5_PLAF7</td>
<td>1w98.1.B</td>
<td>Mediator of RNA polymerase II transcription subunit 20, putative</td>
<td>88.6179</td>
<td>93.80%</td>
<td>-3.17</td>
<td>0.99 (11)</td>
<td>36709.76</td>
<td></td>
</tr>
<tr>
<td>Q8IKC5_PLAF7</td>
<td>41o2.1.A</td>
<td>Diacylglycerol kinase</td>
<td>85.8025</td>
<td>81.58%</td>
<td>-8.09</td>
<td>0.98 (32)</td>
<td>57642.42</td>
<td>7.09</td>
</tr>
<tr>
<td>Q8ID85_PLAF7*</td>
<td>4ww4.1.B</td>
<td>RuvB-like helicase</td>
<td>91.3242</td>
<td>94.46%</td>
<td>-1.49</td>
<td>1 (14)</td>
<td>54629.36</td>
<td>7.27</td>
</tr>
</tbody>
</table>

Discussion

In the current study, the 3380 essentially reported proteins of *P. falciparum strain* 3D7 were analyzed to address potent novel druggable targets. These proteins were analyzed based on their non-homology with the human host as well as human gut microbiome proteome. The targets were additionally shortlisted based on strict threshold criteria of basic druggability features. Among the shortlisted targets...
The proliferation of bacteria, viruses and plasmodium, and inhibiting the DNA unwinding activity reduces the replication of these diseases. The PMTs are highly attractive among the histone-modifying enzymes and act as drug targets [50]. The protein methyltransferases (PMTs) have been linked to the pathogenesis of a variety of diseases, including human cancers, inflammatory diseases, metabolic diseases, and neurodegenerative diseases. The PMTs are highly attractive among the histone-modifying enzymes and act as drug targets [55, 56]. However, to date no study has been conducted about the targeting or inhibition of *P. falciparum* methyltransferase. This also verified by transcriptomic analysis, where the PfPP1 transcript levels reported higher than PP2A after 24 hours of RBC infection [48]. The okadaic acid (OA), a toxin initially isolated from a marine sponge, i.e. *Halichondria okadai* has been identified as a selective inhibitor of serine/threonine protein phosphatases (PPPs) and reported to strongly inhibits PP1, 2A, and 2B in-vitro [49]. Out of 30 examined protein phosphatase, the 16 protein phosphatases along with PP1 and putative phosphatases seem to be important for blood-stage parasites [47]. Moreover, some studies also showed that PfPP1 is indispensable for blood-stage parasite survival [50]. Many phosphatases play key roles in the pathological pathways, and their inactivation may help to prevent or postpone the emergence of human diseases. Therefore, the potent inhibitors for such phosphatases might be of great therapeutic benefit.

The enzyme cytochrome b5 Reductase (cb5r) (Q8I599_PLAF7) plays a role in fatty acid elongation, cholesterol biosynthesis, and cytochrome P450-mediated detoxification of xenobiotics [51]. This protein has been thoroughly studied in mammals, but still need to be characterized in microorganisms, such as fungi and parasites, including *P. falciparum*. There is a close phylogenetic relationship between the plant and *P. falciparum* cb5r proteins. The plant cb5r has been identified as a novel herbicidal target [51]. This protein reported essential for *P. falciparum* survival and was found human host non-homolog and possibly a potent therapeutic target, thereby might be a worthy candidate for drug development against malaria.

The vacuolar protein sorting-associated protein 29 (VPS29) (Q8IM27_PLAF7) is involved in the essential metabolic process of proteins translocation to the subcellular organelles. The *P. falciparum* sort and traffic newly synthesized proteins to target intracellular organelles as well as beyond the plasma membrane into the host cell in some cases [52]. The *P. falciparum* VPS29 (i.e. PfVPS29) is the functional component in the assembly of the retromer complex [53]. During the PPI analysis, the PfVPS29 shows direct interactions with other retromer complex components i.e. PfVPS26, VPS9, VPS10 as shown in Fig. 2. The PfVPS29 is located in the cytosol and highly expressed in early trophozoite and schizont stages [54]. Inhibiting the activity of PfVPS29 may lead to the disassembling of the retromer complex and possibly halt the protein sorting function of the *P. falciparum*.

The multifunctional methyltransferase subunit (Q8IM19_PLAF7) have methyltransferase activity during post-translational modifications, chromatin remodeling and protein heterodimerization activity [40]. The protein methyltransferases (PMTs) have been linked to the pathogenesis of a variety of diseases, including human cancers, inflammatory diseases, metabolic diseases, and neurodegenerative diseases. The PMTs are highly attractive among the histone-modifying enzymes and act as drug targets [55, 56]. However, to date no study has been conducted about the targeting or inhibition of *P. falciparum* methyltransferase.

The RuvB-like helicase (Q8ID85_PLAF7) also shortlisted as therapeutic target in the current study. The RuvB-like helicase function like ATP- dependent helicases. It has a vital role in the cell cycle and transcription [57, 58, 59]. The RUVBL proteins (RUVBL1 & 2) are known to regulate various essential cellular processes in different organisms like *S. cerevisiae*, *drosophila* and *C. elegans* [60, 61, 62]. Three types of RuvB, i.e., PfRuvB1, PfRuvB2, and PfRuvB3 are present in the *P. falciparum*. The PfRuvB1 possesses ssDNA-stimulated ATPase activity and function as a helicase that unwind the DNA in a 5’ to 3’ direction [58]. The PfRuvB2 works similar to PfRuvB1, however, its helicase activity is comparatively weak. The PfRuvB3 function only as ATPase with no helicase activity during schizont/merozoits or intraerythrocytic mitosis [63]. During the developmental stages of the parasite, the PfRuvB1 and PfRuvB2 are expressed in the asexual phase, while the PfRuvB3 expresses only during the schizont stage, where intraerythrocytic mitosis of *P. falciparum* occurs [64]. The PfRuvBL3 protein is a true homolog of yeast RuvBL2. Since in yeast, RuvBL proteins are shown to be extremely essential for survival and known to regulate the transcription of almost 5% of yeast genes [60]. The RuvB-like helicases are suitable drug targets to control malaria due to their essentiality for pathogen and non-homology with human host proteome. It is reported that helicases are needed for the proliferation of bacteria, viruses and plasmodium, and inhibiting the DNA unwinding activity reduces the replication of these
pathogens in cell cultures and animal models [65, 66, 67]. The PfRuvB1 ATPase activity is formerly reported to be inhibited by actinomycin, novobiocin, and ethidium bromide [68].

**Conclusion**

We took advantage from the recently published essential proteins of *P. falciparum* and employed the comparative subtractive proteomics in couple with *in silico* druggability approaches to identify novel and suitable drug targets against *P. falciparum*. The study based on comparative sequence analysis, updated biological databases scanning and multi-direction druggability analyses. This ultimately prioritized and addressed several novel druggable targets against *P. falciparum* infection not highlighted before. Additional consideration of these shortlisted targets in future drug discovery projects may worthy to combat the anti-malarial drug-resistant issues.

**Abbreviations**

Pf: *Plasmodium falciparum*

WHO: World Health Organization

PfCRT: *Plasmodium falciparum* Chloroquine resistant transporter

Pfk13 gene: *Plasmodium falciparum* Kelch-13 gene

BLAST: Basic Local Alignment Search Tool

PDB: Protein Data Bank

PfPP: *Plasmodium falciparum* protein phosphatase

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

All authors read and agreed to publish the study.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

Not applicable.

**Authors’ contributions**

FA, HW and S.J performed the analysis and prepared the initial draft. MA, IA, SS and SGA reviewed the critical analysis and helped in draft final version preparation. AK finalized the draft and supervised the overall study. All authors have read and agreed to the published version of the manuscript.

**Acknowledgements**

Not applicable.

**References**


28. Dym O, Eisenberg D, Yeates TO. ERRAT. Wiley Online Library; 2012;
51. Malvisi L. Functional characterization of cytochrome b5 reductase and its electron acceptor cytochrome b5 in Plasmodium falciparum. 2009;

Figures
Figure 1

The stepwise workflow adopted for novel anti-malarial drug targets identification.

Figure 2

Protein-protein interaction plots of best interacting proteins. The red color indicates query protein. (A) Bis(5'-nucleosyl)-tetraphosphatase. (B) Serine/threonine-protein phosphatase. (C) Cytochrome b5, putative. (D) Vacuolar protein sorting-associated protein 29. (E)
Multifunctional methyltransferase subunit TRM112, putative. (F) RuvB-like helicase. All these proteins showed vital interaction with other Pfalciparum proteins.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- SupplimentaryTableS1.xlsx
- SupplimentaryTableS2.xlsx