Molecular Immunology 116 (2019) 131–139

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

In-silico structural modeling and epitope prediction of highly conserved *Plasmodium falciparum* protein AMR1

Bijara Devi Sanasam, Sanjeev Kumar*

Department of Life science & Bioinformatics, Assam University, Silchar, 788011, India

ARTICLE INFO

Keywords:
Malaria
Plasmodium falciparum
Vaccine
Epitopes
AMR1

ABSTRACT

Malaria caused by *Plasmodium falciparum* is the most deadly and a major health issue worldwide. In spite of several control programs, there hasn't been much improvement in keeping the disease under control. The appearance of drug resistant strains of Plasmodium in addition to insecticide resistance of the Anopheles vector has been a hurdle. Therefore, it is highly desirable to identify new potential candidates that can be targeted for therapeutic intervention. The present study identifies AMR1, a highly conserved essential protein of *Plasmodium falciparum*, as a potential candidate for vaccine development. AMR1 is an exposed surface protein with high antigenic property and conservancy among other species of the parasite. Reverse vaccinology approach (RV) is adopted to determine the best epitopes of AMR1 protein. The protein was further evaluated for several important physiochemical parameters. The study revealed the 3D structure of AMR1, as well as the best B cell and helper T-cell epitopes of the protein. These resulted epitopes might be of great importance in the development of an effective vaccine to combat the deadly disease.

1. Introduction

Malaria is caused by Plasmodium species and it is a major cause for millions of death worldwide (World Health organisation, 2019). *P. falciparum* infection can cause cerebral malaria which is a fatal case of neurological syndrome causing mental deficits or death (Maccormick et al., 2014). Plasmodium parasite completes its life cycle in two hosts i.e., the mosquito vector and the vertebrate host (Schmidt et al., 2015). The pre-erythrocytic stage occurring in the liver of the host is an asymptomatic stage, whereas, erythrocytic stage also known as blood stage or symptomatic stage is the main reason for clinical manifestation and complications of the disease (Yazdani et al., 2006). Plasmodium has a complex multi-staged life cycle employing various evasion mechanisms to subvert host immune system and establishing infection successfully within the host (Schmidt et al., 2015). The parasite expresses varying immune dominant antigens and multiple proteins to accomplish one or more similar functions which enables the parasite to evade the host immune system (Hafalla et al., 2011). Further, malaria parasites execute a range of genetic changes including allelic variation, expression of different proteins during different stages and intracellular replication thereby increases the probability of survival within the host (Gomes et al., 2016).

A number of drugs are available for the treatment of malaria, but the parasite has become resistant with time. The emergence of resistance for commonly used anti-malarial drugs and also for Artemisinin Combination Therapy (ACT) fuels the need of more effective therapeutic approaches (Mokgethi-Morule and N'Da, 2016; Turschner and Efferth, 2009). Understanding the interaction between the host and the pathogen plus the immune response evoked by the pathogen is necessary to develop an effective vaccine to control and possibly to eradicate malaria completely. Both B-cells and T-cells are involved in providing immune protection against malaria infection. However, the immune response during blood stage is mainly directed by humoral immunity since RBCs are devoid of MHC-1 needed for antigen processing mechanisms (Yazdani et al., 2006). Although the role of CD8 + T cells are limited during the blood stage, CD4 + T cells play significant role in the development of blood stage immunity by producing cytokines and regulating the immune responses (Perlmann and Troye-Blomberg, 2000). Studies have shown that vaccine construct that can induce both the B and T cell responses are more effective and likely to generate long lasting memory (Kaba et al., 2012; Wang et al., 2004). So far, there is no full proof vaccine available for malaria despite several report on malaria vaccine concepts and potential candidates for pre-erythrocytic, blood-stage and transmission-blocking vaccines (Coelho et al., 2017). Currently the most advanced vaccine is RTS,S/AS01, which is a pre-erythrocytic stage hybrid recombinant protein vaccine that received a

---

* Corresponding author.
E-mail address: skpuniya@gmail.com (S. Kumar).

https://doi.org/10.1016/j.molimm.2019.10.003
Received 19 July 2019; Received in revised form 29 September 2019; Accepted 1 October 2019

positive regulatory assessment in its Phase 3 trial (WHO, 2016).

Reverse vaccinology, the approach of using the genomic information for developing vaccines has become a popular strategy. RV has been successfully used to design meningococcus vaccine (Sette and Rappuoli, 2010). Traditional method of using the whole organisms for vaccine production has several disadvantages and might cause major problems. The development of fully synthetic peptide-based vaccines might be able to overcome the shortcomings (Skwarczynski and Toth, 2016). With the advanced technology and the availability of rich genomic data, finding of potential vaccine candidates are easier and effortless using the comparative genomic approach and immunoinformatics (Tomar and De, 2010). But, it is crucial to understand the complex process of invasion as well as major protein that may be targeted for development of newer therapeutic interventions (Sanasam and Kumar, 2019). Now several bio-informatical tools and softwares are available to find potential candidates for designing and subsequent development of new vaccine. Nowadays, researchers employ computational methods to identify potential targets prior to laboratory experimentation. Several proteins such as Schizont egress antigen 1 (PfSEA-1A) (Nixon et al., 2017), Merozoite adhesive erythrocytic binding protein (MAEBL) (Cravo et al., 2018), PRE-binding protein (Sanasam and Kumar, 2019) etc. were identified as potential targets by immunoinformatics approach. The present work deals with a highly conserved intra-erythrocytic protein AMR1of *P. falciparum* (Roberts et al., 2013), which is proven as a potential candidate that can be exploited for the development of an effective malaria vaccine. Although B-cell epitope identification is the fundamental for developing epitope based vaccines, inclusion of T cell epitopes in vaccine might be the key for an effective malaria vaccine since both the immune cells go hand-in-hand (Crotty, 2011). Therefore, the protein was studied to identify the potential B and T-cell epitopes for the development of epitope-based vaccine against malaria by employing a number of bio-informatical tools

2. Methods and materials

2.1. Sequence retrieval and conservancy determination

Sequence of *P. falciparum* (3D7) conserved protein, AMR1, was retrieved from PlasmoDB (http://plasmodb.org/plasmo/) and the conservancy of the protein among other Plasmodium species was determined by NCBI Blast with default settings. Conservancy is one of the most important criteria to be considered, since conservation of a gene indicates the essentiality of the genes which make it a better target for therapeutic intervention (Doyle et al., 2010).

2.2. Determination of homology with human host and gut microbiota

The protein sequence was subjected to NCBI Blast against human proteome and a set of selected gut microbiota proteome (Supplementary data 1) to check the presence of homologs so as to avoid any unwanted cross reaction with the human host or the beneficial gut flora population.

2.3. Determination of protein localization and antigenicity

Surface proteins are the major target for vaccine development since they are exposed and accessible to the immune system. The subcellular localization of the protein was determined by SubCons webserver (Salvatore et al., 2018). Then, antigenicity of the protein was predicted by using VaxiJen with a threshold value of 0.4. VaxiJen is an alignment-independent tool for the prediction of protective antigens, which classified the antigen exclusively based on the physicochemical properties of the proteins (Doytchinova and Flower, 2007).

2.4. Domain analysis and tertiary structure prediction and validation

The AMR1protein sequence was subjected for domain analysis in Pfam database (Finn et al., 2016) followed by Robetta tool and the 3D structure model of the protein domains were modeled domain-wise using Robetta (Kim et al., 2004). Robetta uses ROSETTA software offering both *ab-initio* and comparative modeling of protein domains, in which comparative modeling are done using the RosettaCM protocol whereas *ab-initio* modeling are done with the Rosetta de novo protocol (Song et al., 2013). The quality of the protein structure generated was analyzed and validated using PROCHECK program in SAVES server (Laskowski et al., 1993).

2.5. Conformational B-cell epitope prediction

The discontinuous B-cell epitopes of the protein was predicted using Ellipro tool which predicts discontinuous B-cell epitopes based on the protein's 3D structure. The generated 3D structure model of the protein in PDB format was used for the prediction with default setting, which is 0.5 for minimum score and maximum distance 6A^{O}. Residues with greater score shows greater solvent accessibility whereas larger value of distance R indicates larger discontinuous epitopes prediction (Ponomarenko et al., 2008).

2.6. Continuous linear B-cell epitopes prediction

B-cell epitopes is important for vaccine design, immunodiagnostic tests, and antibody production that make its identification and characterization a significant task. Continuous linear B-cell epitopes were predicted using three different web based tools namely- AAP (Chen et al., 2007), BCPred (EL-Manzalawy et al., 2008) and FBCPred (El-Manzalawy et al., 2008) available in web-based server BCPREDS. AAP predicts epitope based on the finding that B-cell epitopes favor particular amino acid pairs whereas FBCPred and BCPred predict linear B-cell epitopes using the subsequence kernel.

2.6.1. Surface accessibility and beta-turn prediction

The surface accessibility of AMR1 protein was predicted by the method of Emini taking a threshold value of 1.00 (Emini et al., 1985). β-turn regions are usually of high accessibility and found to be correlated with the location of immunogenic and antigenic peptides. Therefore, β-turn of the protein was predicted using Chou and Fasman beta-turn prediction with a threshold value of 1.00 (Chou and Fasman, 1977).

2.6.2. Flexibility and hydrophilicity

Karplus and Schulz’s flexibility scale was used for the prediction of flexibility with a threshold value of 1.00 (Karplus and Schulz, 1985). Evaluation of the hydrophilicity was done by using Parker hydrophilicity scale with threshold value of 2.00 (Parker et al., 1986). These tools are based on amino acid property scales. Each scale consists of 20 values assigned to each of the amino acid residues on the basis of their relative propensity to possess the property described by the scale.

2.6.3. Similarity search

Epitopes were subjected to Smith-Waterman similarity search tool to identify experimentally validated B-cell epitopes recorded in IEDB, if any. Experimentally validated epitope found in IEDB are verified epitopes which are proven potential candidates. Nonetheless, epitopes not available in IEDB can be a new potential vaccine candidate.

2.7. Helper T-cell epitope

It is a well-known fact that B-cell activation by protein antigen is helper T-cell dependent and epitope specific helper T-cell binding is necessary for that. Hence, it is expected that the inclusion of epitopes

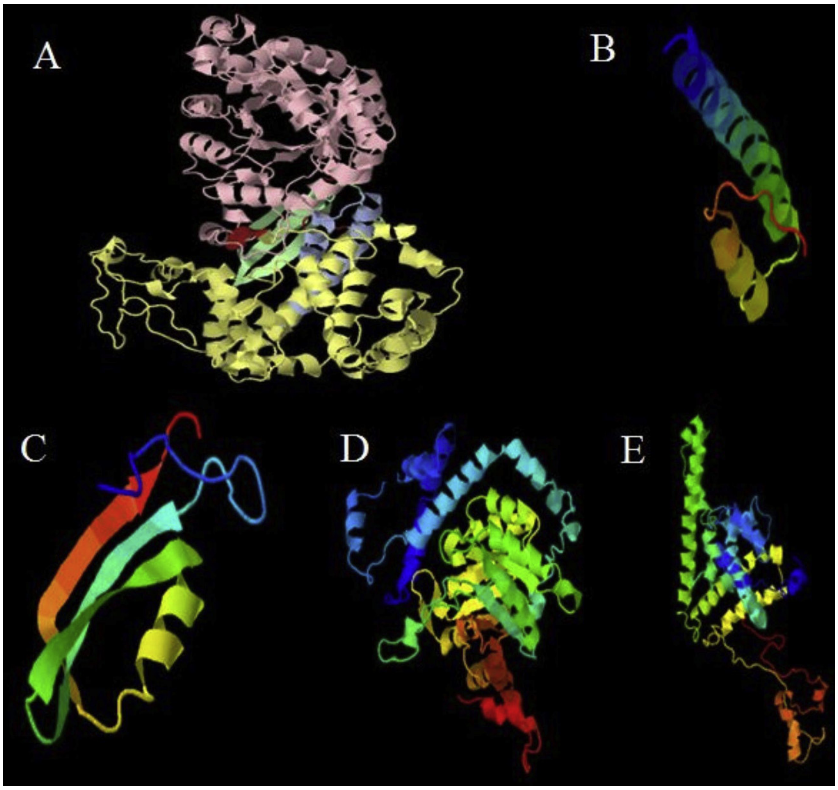

Fig. 1. Modeled structures of conserved protein AMR1 and its domains. (**A**) Whole protein structure (**B**) Domain 1 (**C**) Domain 2 (**D**) Domain 3 and (**E**) Domain 4.

with overlapping B and T cell epitopes in the vaccine will be more effective and desirable since it can activate both the response. Likewise, the addition of T-cell epitopes capable of binding to numerous HLA alleles (promiscuous epitopes) into the vaccine may enhance their effectiveness for general population. Therefore, promiscuous binding CD4$^+$ T cell epitopes of AMR1 protein for the 26 most frequently occurring HLA alleles was predicted using Tepitool (Paul et al., 2016) by following IEDB recommended prediction method. Moreover, a comparison was carried out between the complete result of predicted CD4$^+$T cell epitopes and the selected B-cell epitopes for overlapping epitopes.

2.7.1. Helper T-cell epitope immunogenicity prediction

The relative ability of a peptide-MHC complex to elicit an immune response represents the immunogenicity of the peptide. CD4$^+$T cell immunogenicity prediction tool which is freely available in IEDB was used for the prediction of immunogenicity of the epitopes by following IEDB recommended method (Dhanda et al., 2018).

2.7.2. Population coverage

MHC molecules are extremely polymorphic and in case of human over a thousand different MHC (HLA) alleles are known. Different HLA types are expressed at dramatically different frequencies in different ethnicities. Therefore, careful consideration should be made so that the developed vaccine has large population coverage and not ethnically biased. The fraction of population that respond to the given epitope set on the basis of HLA genotypic frequencies and MHC-II binding T cell restriction data was calculated using the Population Coverage tool (Bui et al., 2006).

2.8. Epitope conservancy determination

Five homologs of AM1 protein in *P. falciparum* HB3 and the other 4 human Plasmodium species were selected by employing Blast. The conservancy of the epitopes were determined by using Epitope Conservancy Analysis tool with identity match level set at ≥70% (Bui et al., 2007). These conserved epitopes are expected to provide wider protection across different strains or species, than the epitopes with variable genome regions.

3. Results

3.1. Sequence retrieval and conservancy

The protein sequence of the protein AMR1 was retrieved from PlasmoDB (Aurrecoechea et al., 2009), which is 790 amino acids long with a molecular weight of 92.79 kDa. After running Blast against the other species of Haemosporida, the protein was found to be highly conserved among the other Plasmodium species. The conservancy of the protein and the presence of homologs in other species show the essentiality of the protein in the survival of the organism. The protein was further checked for homologs in human by running Blast and the result showed that the protein doesn't have any homolog in human indicating its uniqueness to the parasite.

3.2. Determination of homology with human host and gut microbiota

The cross-species analysis of AMR1with the human proteome and gut micro-biota by running NCBI Blast showed the absence of any homolog of the protein in human as well as gut micro-biota population. This showed the uniqueness of AMR1 protein as a Plasmodium protein that might be essential for parasite survival.

3.3. Protein localization and antigenicity

The localization of the protein is an important factor in finding potential vaccine candidates since it determines the accessibility of the protein by the host immune cells. SubCons predicted AMR1 as a surface membrane protein, which indicates the possible accessibility by the immune system. In addition, a vaccine should be immunogenic and capable to trigger immune response to develop memory cells against the pathogenic epitopes. VexiJen determined the protein to have antigenicity probability of 0.6529. The antigenicity of the protein is higher than the threshold value (0.4), representing the antigenic nature of the protein (Doytchinova and Flower, 2007).

3.4. Domain detection and protein structure modeling

Domain analysis of AMR1 protein showed that the query protein contains four domains. The third domain of the protein belongs to Indole-3-Glycerol Phosphate Synthase like class. The structures of the domains present in the protein were modeled domain-wise using Robetta program, which offers both *ab-initio* as well as comparative method of protein structure modeling (Kim et al., 2004). The full protein 3D structure of AMR1 protein was then generated which is shown in Fig. 1.

3.5. Validation of the protein model

The quality of the modeled protein structure was evaluated and validated using PROCHECK in SAVES server (http://servicesn.mbi.ucla.edu/PROCHECK/). Ramachandran plot generated by the PROCHECK described the amino acid positions in the plot as well as the overall quality of the protein model (Laskowski et al., 1993). The plot showed that 91.3% amino acids were arranged in most favored core regions with 7.8% in allowed region, 0.9% generously allowed region and 0.0% (none) in disallowed region (Fig. 2). The results showed that the generated structure of AMR1 protein was good considering the overall geometry and can be used for further study.

3.6. Conformational B-cell epitope prediction

Conformational B-cell epitopes are discontinuous epitopes which are generated due to the position of the amino acids in the protein structure resulted from protein folding. Ellipro (Ponomarenko et al., 2008) predicted a total of 14 epitopes with maximum score of 0.921 and minimum of 0.542. The minimum threshold was taken as 0.7 and those epitope with scores greater than 0.7 were selected. The four selected discontinuous epitopes, their number of the residues and the scores obtained by each epitopes are provided in Table 1 and the structure of the epitopes in relation with protein structure are given in Fig. 3.

3.7. Continuous linear B-cell epitopes

The linear B-cell epitopes of AMR1 protein was predicted using BCPred, AAP and FBCpred programs (Chen et al., 2007; El-Manzalawy et al., 2008; EL-Manzalawy et al., 2008). The epitope length was set at 16 amino acids (aa) for all the three predictions. Each of the three programs generates a wide variety of potential B-cells epitopes: BCPred predicted 11 epitopes, whereas AAP and FBCpred predicted 20 epitopes each. All the predicted epitopes with prediction score ≥0.9 were selected not to miss out any potential epitopes and a manual comparison was done between the predictions to merge all the overlapping epitopes. The merging of overlapping epitopes gives rise to a total of 37 epitopes for further analysis.

3.8. Epitopes screening based on amino acid propensity scale

The locations of continuous epitopes are correlated with various parameters such as hydrophilicity, flexibility, beta turns and surface accessibility propensity of polypeptide chain. To toughen the reliability of the predicted B-cell epitopes, predictions were done based on these aforementioned physiochemical parameters and comparisons were carried out with the results from the previous prediction. The resulted epitopes with their peptide lengths, sequences and the position of the peptides along the sequence length are summarized in Table 2. The graphical representation of the epitope prediction of the protein on the basis of surface accessibility, surface flexibility, hydrophilicity and beta-turn prediction are shown in Fig. 4.

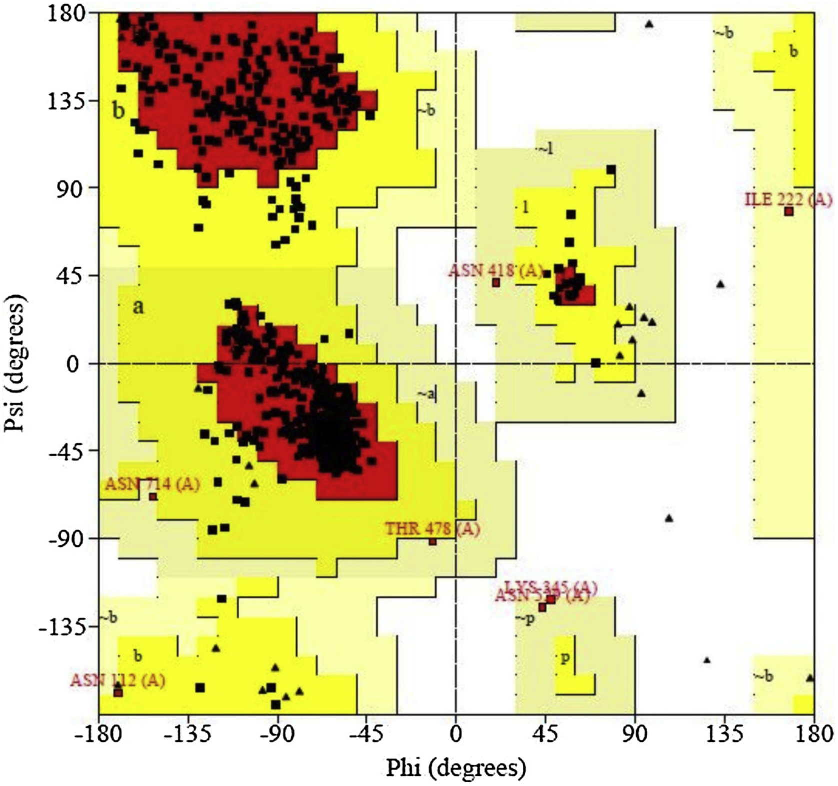

Fig. 2. Ramachandran plot generated to validate the modeled 3D structure of AMR1protein showing 91.3% residues in the favored region.

Table 1
Discontinuous B-cell epitopes predicted by Ellipro.

S/N	Residues	No. of residues	Scores	3D Structure
1.	A:H725, A:K726, A:N727, A:L728, A:T729, A:G730, A:Y731, A:P732, A:M733, A:G734, A:T735, A:L736, A:Q737, A:N738, A:S739, A:P740, A:M741, A:N742, A:N743, A:L744, A:D745, A:I746, A:Y747, A:N748, A:L749, A:T750, A:K751, A:E752, A:Y753, A:F754, A:G755, A:E756, A:T757, A:Q758, A:N759, A:S760, A:D761, A:A762, A:H763, A:T764, A:E765, A:N766, A:I767, A:F768, A:D769, A:S770, A:D771, A:Q772, A:N773, A:K774, A:L775, A:L776, A:D777, A:E778, A:S779, A:T780, A:E781	57	0.921	Fig. 3(A)
2.	A:D502, A:K503, A:N504, A:I505, A:S506, A:N507, A:N508, A:N579, A:Y583, A:T585, A:E586, A:G587, A:N588, A:K589, A:N590, A:N591, A:Y592, A:A593, A:K596, A:S598, A:F599, A:E600, A:N601, A:K602, A:C603, A:I604, A:D605, A:N606, A:K607, A:N608, A:I609, A:K610, A:H611, A:T612, A:N613, A:M614, A:N615, A:K616, A:K617, A:E618, A:E619, A:I620, A:L621, A:S622, A:K623, A:E624, A:E625, A:K626, A:Q627, A:I628, A:V629, A:Q630, A:N631, A:F632, A:K633, A:N634, A:E635, A:K636, A:K637, A:R638, A:E639, A:L640, A:M641, A:L642, A:L643, A:S644, A:Q645, A:M646, A:K647, A:E648, A:I649, A:I650, A:K651, A:E652, A:V653, A:D654, A:D655, A:Q656, A:C657, A:K658, A:N659	81	0.791	Fig. 3(B)
3.	A:D782, A:S783, A:S784, A:R785, A:G786, A:L787, A:D788, A:S789, A:S790	9	0.758	Fig. 3(C)
4.	A:M1, A:I2, A:A3, A:R4, A:I5, A:K7, A:I72, A:K73, A:N74, A:D75, A:G76, A:E77, A:Y79, A:K94, A:N95, A:D96, A:I97, A:G98, A:G99, A:R100	20	0.737	Fig. 3(D)

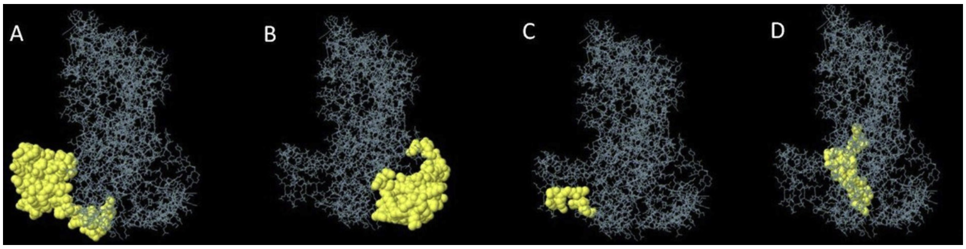

Fig. 3. 3D structures of the conformational B-cell epitopes of AMR1 protein.

3.8.1. *Assessment of solvent accessibility regions*

Emini surface accessibility propensity scale (Emini et al., 1985) predicted a total of 18 peptides as potential epitopes when employed with a threshold value of 1.0 and gave an average of 1.0 with a minimum of 0.043 and maximum of 6.860 (Fig. 4A). A comparison was made with these predicted epitopes and the epitopes resulted from the three prediction tools in the previous step. The analysis of solvent accessibility resulted in 14 peptides which can be carried forward for other analysis.

3.8.2. *Beta-turn prediction by Chou and Fasman beta-turn prediction*

Sequence regions with beta-turn are known to be antigenic and highly accessible. Chou and Fasman method was used for beta-turn prediction with a threshold of 1.00. The protein was predicted to have an average value of 1.005 with minimum value of 0.619 and maximum of 1.397 (Fig. 4B). Epitopes were compared with the resulted residues score manually and residues with below threshold were removed. One epitope 710′VEQKNK′715 was found to score lower than the threshold, thus removed. Remaining 13 epitopes proceeded for further analysis.

3.8.3. *Epitope flexibility and hydrophilicity prediction*

The amino acids in the epitope are usually charged and hydrophilic in nature as well as flexible. Karplus and Schulz scale was used to assess the flexibility prediction (Karplus and Schulz, 1985). Flexibility of an average value 1.006 with minimum and maximum value of 0.880 and 1.126 respectively were predicted (Fig. 4C). Whereas, Parker hydrophilicity scale (Parker et al., 1986) predicted hydrophilicity score to an average value of 1.957 with minimum value −7.814 and maximum value 7.686 (Fig. 4D). All the 13 epitopes were compared with the resulted residue score and only those with significant scores were selected. A minimum length of 6aa was set for selecting the epitope after removing the low-ranking residues. Two epitopes 215 ‘NLQDEKYI ‘222 and 499‘LYDDKNI‘505 didn't fulfill the required criteria and thus removed. The selected epitopes with their length and position are provided in Table 2. The selected epitopes might be of greater importance as vaccine subunits for epitope based vaccine design.

3.8.4. *Epitope similarity search*

The search of similar epitopes in IEDB for the predicted B-cell

Table 2
Predicted antigenic B-cell epitopes and overlapping helper T cell epitopes of conserved Plasmodium protein AMR1.

S/N	Start	B-cell Epitopes	End	Length	T-cell epitopes with overlapping residues (15aa)	Immunogenicity
1.	90	VRKRKN	95	6	KYIGGKKEKVRKRKN	93.9997
2.	105	YGKDSRN	112	7	GGRVRFRRYGKDSRN	91.0673
3.	164	LKRSNNDE	171	8	NKLSESLKRSNNDEK	87.4549
4.	193	KKKEKDHKTK	202	10	IFCSIEKEKKKEKDH	97.6621
					RKIFCSIEKEKKKEK	95.0038
5.	241	NNEEKKK	247	7	NNEEKKKNINNNYIY	79.3128
6.	305	TLPRNK	310	6	KDIIKSTRTLPRNKR	89.9552
7.	577	KDNPMK	582	6	KDNPMKYKTEGNKNN	97.4333
					KRYNVLNKDNPMKYK	66.344
					YNVLNKDNPMKYKTE	71.3009
					RYNVLNKDNPMKYKT	70.9217
					VLNKDNPMKYKTEGN	91.6356
8.	584	KTEGNKNNYA	593	10	MKYKTEGNKNNYAEK	98.0223
9.	661	DTNTDEQNQKK	671	11	EQNQKKEEKLESLLE	98.8169
10.	757	TQNSDAH	763	7	LTKEYFGETQNSDAH	99.3072
11.	778	ESTEDS	783	6	FDSDQNKLLDESTED	99.117

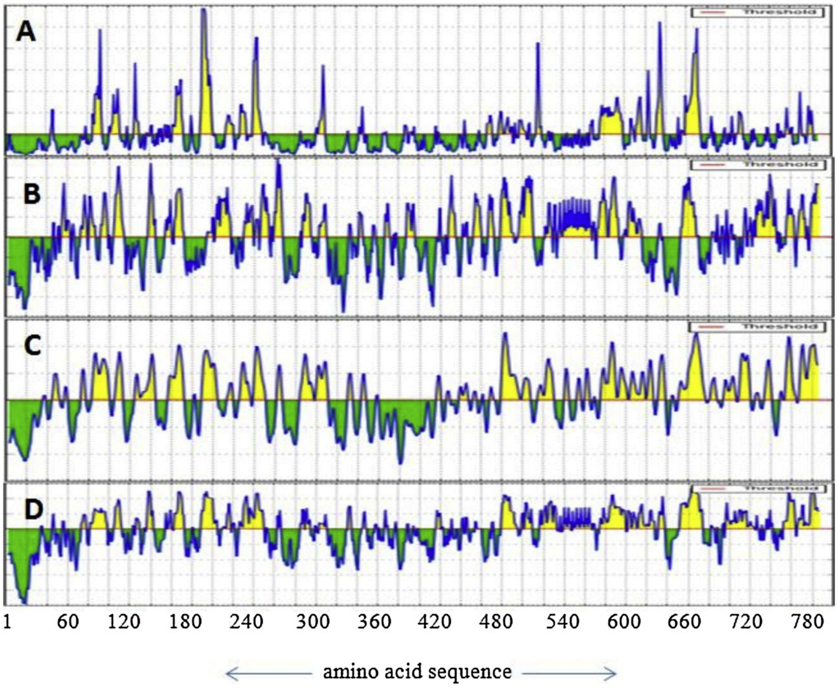

Fig. 4. Graphical representation of Epitope prediction of AMR1 protein by (A) Emini surface accessibility prediction (threshold: 1.000), (B) Chou and Fasman beta-turn prediction (threshold: 1.00), (C) Karplus and Schulz flexibility prediction (threshold: 1.000) and (D) Parker hydrophilicity prediction (threshold: 2.00).

epitopes gave rise to many epitopes with varying matching residues. 100% identity matching was found for epitope 90′VRKRKN’95, whereas for the remaining 10 epitopes although not exact match but similarity at some level were found with the validated epitopes. The details of the similarity search of the predicted epitopes with their matching validated epitopes are provided as Supplementary data 2.

3.9. Helper T-cell epitope prediction

Promiscuous MHC-II Binding epitopes for the most frequently occurring HLA-II alleles set as predicted by the prediction tool were of 15mer. The predicted T-cell epitopes were crosschecked with the 12 B-cell epitope for overlapping epitopes. The T cell epitopes with overlapping residues are provided against the corresponding B-cell epitopes in Table 2. Furthermore, best promiscuous helper T-cell epitopes were also selected. The best T-cell epitopes along with their binding alleles count, peptide position and scores are provided in Table 3.

Table 3
Promiscuous helper T-cell epitopes for the most frequently occurring HLA-II alleles set.

S/N	Start	Peptide	End	No. of binding alleles (alleles with consensus percentile ≤20.0)	Immunogenicity score
1.	3	ARIYKYIITLFYFLL	17	19	88.5875
2.	8	YIITLFYFLLYIFIL	22	15	92.1979
3.	13	FYFLLYIFILFHCLN	27	17	70.8064
4.	20	FILFHCLNSLIIKYN	34	16	87.5475
5.	29	LIIKYNVKVEGLRLV	43	13	81.632
6.	60	GILHNHYISTLFIKN	74	13	89.991
7.	249	INNNYIYQHNFLNLS	263	13	83.8985
8.	315	VVDDIIIHPIQIALA	329	13	95.2567
9.	335	DGVILNLSYLKNDME	349	13	88.2721
10.	462	KTLLYLNRNNYLKEF	476	13	54.4328

3.9.1. T-cell epitopes immunogenicity analysis

Both set of T-cell epitopes were subjected to immunogenicity analysis separately and the score assigned to each epitope signifies the immunogenicity of the epitope. The higher the predicted score of a peptide, the greater is the possibility of stimulating an immune response. The immunogenicity scores of the epitopes were provided alongside its respective T-cell epitope in Tables 2 and 3.

3.9.2. Population coverage

Among different ethnicities the expression of the different HLA alleles are at extremely different frequencies. On the basis of HLA genotypic frequencies as well as MHC binding data, the fraction of individuals responding to the overlapping T cell epitopes were calculated using the Population Coverage tool. The world coverage chart of the epitopes is shown in Fig. 5. The world coverage percent of each epitope is provided as Supplementary data 3. The promiscuous T-cell epitopes are supposed to bind a number of HLA covering a large population, therefore population analysis was not performed.

3.10. Epitope conservancy determination

Conservancy of an epitope is the presence of the fraction of sequence (epitope) in close or distant relatives. Conservancy of the 11 B-cell epitopes and 10 promiscuous T-cell epitopes were determined by the tool Epitope Conservancy Analysis hosted by IEDB (Bui et al., 2007). The epitopes were subjected against orthologues of AMR1 protein in other human Plasmodium species and strain. In the present case, AMR1 orthologs in *P. falciparum* (HB3), P.vivax, P. ovale, P.knowlesi and P.malarae were selected. The similarity percentage was set at ≥50% to identify conserved and semi-conserved epitopes of 50% and above. The selected epitopes of the protein exhibited varying conservancy whose details are provided in Table 4. Epitopes showing 100% conservancy are the best epitopes that can be used to target the different species of Plasmodium.

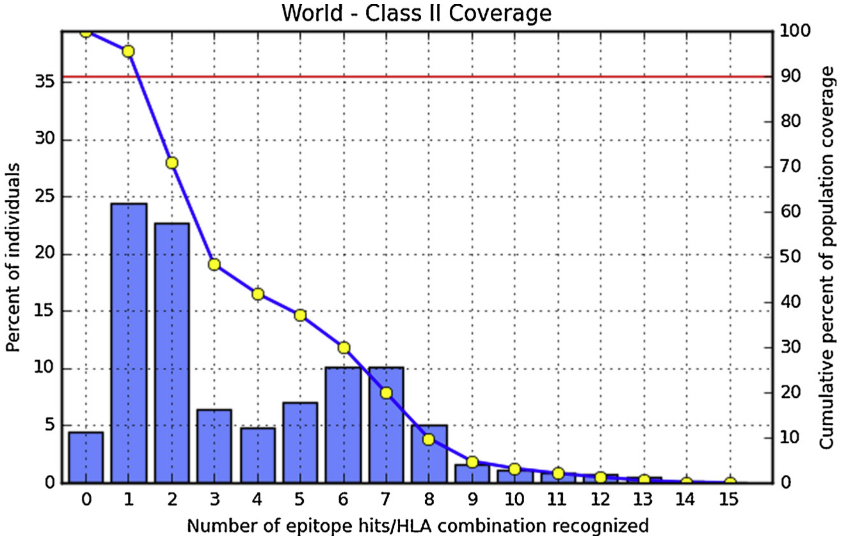

Fig. 5. Chart showing the population coverage by the epitopes.

4. Discussion

In the present genomic era bioinformatics play significant role in the field of vaccinology (Tomar and De, 2010). Conventional methods are time consuming, costly and require intensive facilities with high labor demand. Finding all the required information using the available genome data resources beforehand saves a significant fraction of the time, labor and resources (Delany et al., 2013).Therefore, use of bioinformatical tools for the screening of the best candidate followed by laboratory experiment for the validation is a good strategy. Vaccines based on the whole sporozoite of the parasite (Hollingdale and Sedegah, 2017), MSP1 protein (Takala et al., 2007), EBA175 protein (Peek et al., 2006) etc. were developed and experimentally tested for its efficacy and further trials are still under consideration. Even after the generous efforts put by the researchers to tackle the disease, there hasn't been much improvement over the last decades in developing an effective vaccine against malaria. But continuous effort is required to control this persistent disease and identification of new potential targets is extremely necessary.

Parasite proteins which are protective antigen with high immunogenicity and good surface accessibility are potential targets for vaccine development (Doytchinova and Flower, 2007; Ghosh et al., 2012; Mahanty et al., 2015). Both the innate and acquired immune systems are required for the generation of an effective immune response to an immunization leading to the induction of lasting stimulation of the adaptive system (the humoral and cell-mediated immunity) by the production of effector cells and memory cells (Clem, 2011). Epitope based vaccine can have either B-cell epitopes that are targeted to activate the humoral immunity or T-cell epitopes that are targeted to activate cell mediated immunity or both as the subunits. But stimulating both the branches of adaptive immunity will be more effective. The current study identifies both the linear and the conformational B-cell epitopes along with helper T-cell epitopes. The conformational B-cell epitopes are the one which are formed due to the position of the residues caused by the folding of the protein and determining these conformational B-cell epitopes is very important in vaccine development (Ahmad et al., 2016). 3D structure of the AMR1 protein was generated for conformational B-cell epitope prediction and quality analysis of the generated structure was performed which showed that the overall geometry of the structure was good for carrying out the study. The protein was found to contain IGPS domain which belongs to the family lyases. Erythrocytic stage of malarial parasite is a key

Table 4
Epitope conservancy analysis of the linear B and T cell epitopes by percent of protein sequence matches at identity ≥50%.

S/N	Epitope type	Epitope sequence	Epitope length	Percent matches at identity ≥50%	Minimum identity	Maximum identity
1.	**B-cell**	VRKRKN	6	100.00% (5/5)	100.00%	100.00%
2.		YGKDSRN	7	100.00% (5/5)	71.43%	100.00%
3.		LKRSNNDE	8	100.00% (5/5)	62.50%	100.00%
4.		KKKEKDHKTK	10	100% (5/5)	50.0%	100.00%
5.		NNEEKKK	7	100.00% (5/5)	57.14%	100.00%
6.		TLPRNK	6	100.00% (5/5)	50.00%	100.00%
7.		KDNPMK	6	100% (5/5)	50.00%	100.00%
8.		KTEGNKNNYA	10	20.00% (1/5)	40.00%	100.00%
9.		DTNTDEQNQKK	11	20.00% (1/5)	36.36%	100.00%
10.		TQNSDAH	7	20.00% (1/5)	28.57%	100.00%
11.		ESTEDS	6	100.00% (5/5)	50.00%	100.00%
1.	**T-cell**	ARIYKYIITLFYFLL	15	20.00% (1/5)	26.67%	100.00%
2.		YIITLFYFLLYIFIL	15	20.00% (1/5)	26.67%	100.00%
3.		FYFLLYIFILFHCLN	15	20.00% (1/5)	26.67%	100.00%
4.		FILFHCLNSLIIKYN	15	20.00% (1/5)	26.67%	100.00%
5.		LIIKYNVKVEGLRLV	15	20.00% (1/5)	33.33%	100.00%
6.		GILHNHYISTLFIKN	15	20.00% (1/5)	26.67%	100.00%
7.		INNNYIYQHNFLNLS	15	100.00% (5/5)	60.00%	100.00%
8.		VVDDIIIHPIQIALA	15	100.00% (5/5)	93.33%	100.00%
9.		DGVILNLSYLKNDME	15	100.00% (5/5)	80.00%	100.00%
10.		KTLLYLNRNNYLKEF	15	100.00% (5/5)	80.00%	100.00%

research area (Satchwell, 2016). AMR1 is an intra-erythrocytic protein and efficiency of vaccines based on blood stage proteins in parasite clearance is strongly supported by experimental data (Mahanty et al., 2003). The study revealed the best B-cell and helper T-cell epitopes with high conservancy and population coverage that might be used to target the different species of the parasite whilst covering majority of the world population. Epitopes thus resulted after the various analyses are expected to be the best subunit candidate of the protein for peptide based vaccine. These epitopes can be used to develop recombinant vaccine that target the blood stage or can be used in compliments with subunits from other stage to form multistage vaccine. The present study is an encouraging step towards the application of reverse vaccinology in finding potential vaccine target and vaccine subunits for peptide-based vaccine development against malaria.

5. Conclusion

The present study was designed to find the best epitopes of the highly conserved protein *P. falciparum*, AMR1, for the development of an effective vaccine against malaria. AMR1 protein is an exposed surface protein with high antigenicity and high conservancy which makes it a potential target. The study revealed the best T cell and B cell epitopes of AMR1 protein as well as the 3D structure of the protein. Crosscheck of the predicted B-cell epitopes in IEDB revealed some degree of identity with validated epitopes. The identified potential epitopes might be of great use as vaccine subunits, developing therapeutics and other clinical applications.

Declaration of Competing Interest

The authors declare no conflicting interest in publishing the article.

Acknowledgements

The authors would like to acknowledge Assam University, Silchar for providing research and infrastructure facilities for carrying out the present work.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.molimm.2019.10.003.

References

Ahmad, T.A., Eweida, A.E., Sheweita, S.A., 2016. B-cell epitope mapping for the design of vaccines and effective diagnostics. Trials Vaccinol. 5, 71–83. https://doi.org/10.1016/j.trivac.2016.04.003.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Miller, J.A., Nayak, V., Pennington, C., Pinney, D.F., Roos, D.S., Ross, C., Stoeckert, C.J., Treatman, C., Wang, H., 2009. PlasmoDB: a functional genomic database for malaria parasites. Nucleic Acids Res. 37, D539–D543. https://doi.org/10.1093/nar/gkn814.

Bui, H., Sidney, J., Li, W., Fusseder, N., Sette, A., 2007. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinf. 8, 361. https://doi.org/10.1186/1471-2105-8-361.

Bui, H.-H., Sidney, J., Dinh, K., Southwood, S., Newman, M.J., Sette, A., 2006. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinf. 7, 153. https://doi.org/10.1186/1471-2105-7-153.

Chen, J., Liu, H., Yang, J., Chou, K., 2007. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale Ã. Amino Acids 33, 423–428. https://doi.org/10.1007/s00726-006-0485-9.

Chou, P.Y., Fasman, G.D., 1977. Secondary structural prediction of proteins from their amino acid sequence. Trends Biochem. Sci. 2, 128–131.

Clem, A.S., 2011. Fundamentals of vaccine immunology. J. Glob. Infect. Dis. 3, 73–78. https://doi.org/10.4103/0974-777X.77299.

Coelho, C.H., Doritchamou, J.Y.A., Zaidi, I., Duffy, P.E., 2017. Advances in malaria vaccine development: report from the 2017 malaria vaccine symposium. npj Vaccines 2, 1–5. https://doi.org/10.1038/s41541-017-0035-3.

Cravo, P., Machado, R.B., Leite, J.A., Leda, T., Suwanarusk, R., Bittencourt, N., Albrecht, L., Judice, C., Lopes, S.C.P., Lacerda, M.V.G., Ferreira, M.U., Soares, I.S., Goh, Y.S., Bargieri, D.Y., Nosten, F., Russell, B., Rénia, L., Costa, F.T.M., 2018. In silico epitope mapping and experimental evaluation of the merozoite adhesive erythrocytic binding protein (MAEBL) as a malaria vaccine candidate. Malar. J. 17, 1–9. https://doi.org/10.1186/s12936-017-2144-x.

Crotty, S., 2011. Follicular helper CD4 T cells (TFH). Annu. Rev. Immunol. 29, 621–663. https://doi.org/10.1146/annurev-immunol-031210-101400.

Delany, I., Rappuoli, R., Seib, K.L., 2013. Vaccines, reverse vaccinology, and bacterial pathogenesis. Cold Spring Harbor Perspect. Med. 3https://doi.org/10.1101/cshperspect.a012476. a012476–a012476.

Dhanda, S.K., Karosiene, E., Edwards, L., Grifoni, A., Paul, S., Andreatta, M., Weiskopf, D., Sidney, J., Nielsen, M., Peters, B., Sette, A., 2018. Predicting HLA CD4 immunogenicity in human populations. Front. Immunol. 9, 1–14. https://doi.org/10.3389/fimmu.2018.01369.

Doyle, M.A., Gasser, R.B., Woodcroft, B.J., Hall, R.S., Ralph, S.A., 2010. Drug target prediction and prioritization: using orthology to predict essentiality in parasite genomes. BMC Genomics 11, 222. https://doi.org/10.1186/1471-2164-11-222.

Doytchinova, I.A., Flower, D.R., 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinf. 8, 4. https://doi.org/10.1186/1471-2105-8-4.

El-Manzalawy, Y., Dobbs, D., Honavar, V., 2008. Predicting flexible length linear B-cell epitopes. Comput. Syst. Bioinf. Conf. 7, 121–132.

EL-Manzalawy, Y., Dobbs, D., Honavar, V., 2008. Predicting linear B-cell epitopes using string kernels. J. Mol. Recognit. 21, 243–255. https://doi.org/10.1002/jmr.893.

Emini, E.A., Hughes, J.V., Perlow, D.S., Boger, J., 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. 55, 836–839.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44, D279–D285. https://doi.org/10.1093/nar/gkv1344.

Ghosh, A., Chattopadhyay, S., Chawla-Sarkar, M., Nandy, P., Nandy, A., 2012. In Silico study of rotavirus VP7 surface accessible conserved regions for antiviral drug/vaccine design. PLoS One 7. https://doi.org/10.1371/journal.pone.0040749.

Gomes, P.S., Bhardwaj, J., Rivera-correa, J., Freire-de-lima, C.G., Morrot, A., 2016. Immune escape strategies of malaria parasites. Front. Microbiol. 7, 1–7. https://doi.org/10.3389/fmicb.2016.01617.

Hafalla, J.C., Silvie, O., Matuschewski, K., 2011. Cell biology and immunology of malaria. Immunol. Rev. 240, 297–316. https://doi.org/10.1111/j.1600-065X.2010.00988.x.

Hollingdale, M.R., Sedegah, M., 2017. Development of whole sporozoite malaria vaccines. Expert Rev. Vaccines 16, 45–54. https://doi.org/10.1080/14760584.2016.1203784.

Kaba, S.A., Mccoy, M.E., Doll, T.A.P.F., Brando, C., Guo, Q., Dasgupta, D., Yang, Y., Mittelholzer, C., Spaccapelo, R., Crisanti, A., Burkhard, P., Lanar, D.E., 2012. Protective antibody and CD8+ T-Cell responses to the *Plasmodium falciparum* circumsporozoite protein induced by a nanoparticle vaccine. PLoS Med. 7, 1–10. https://doi.org/10.1371/journal.pone.0048304.

Karplus, P.A., Schulz, G.E., 1985. Prediction of chain flexibility in proteins. Naturwissenschaften 72, 212–213. https://doi.org/10.1007/BF01195768.

Kim, D.E., Chivian, D., Baker, D., 2004. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. 32, W526–W531. https://doi.org/10.1093/nar/gkh468.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291. https://doi.org/10.1107/S0021889892009944.

Maccormick, I.J.C., Beare, N.A.V., Taylor, T.E., Barrera, V., White, V.A., Hiscott, P., Molyneux, M.E., Dhillon, B., Harding, S.P., 2014. Cerebral malaria in children: using the retina to study the brain. Brain 137, 2119–2142. https://doi.org/10.1093/brain/awu001.

Mahanty, S., Prigent, A., Garraud, O., 2015. Immunogenicity of infectious pathogens and vaccine antigens. BMC Immunol. 16, 1–6. https://doi.org/10.1186/s12865-015-0095-y.

Mahanty, S., Saul, A., Miller, L.H., 2003. Review progress in the development of recombinant and synthetic blood-stage malaria vaccines. J. Exp. Biol. 206, 3781–3788. https://doi.org/10.1242/jeb.00646.

Mokgethi-Morule, T., N'Da, D.D., 2016. Cell based assays for anti-Plasmodium activity evaluation. Eur. J. Pharm. Sci. 84, 26–36. https://doi.org/10.1016/j.ejps.2016.01.001.

Nixon, C.E., Park, S., Pond-tor, S., Raj, D., Lambert, L.E., Orr-Gonzalez, S., Barnafo, E.K., Rausch, K.M., Friedman, J.F., Fried, M., Duffy, P.E., Kurtis, J.D., 2017. Identification of protective B-cell epitopes within the novel malaria vaccine candidate *Plasmodium falciparum* schizont egress antigen 1. Clin. Vaccine Immunol. 24, 1–6.

Parker, J.M., Guo, D., Hodges, R.S., 1986. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data : correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry 25, 1–2.

Paul, S., Sidney, J., Sette, A., Peters, B., 2016. TepiTool: a pipeline for computational prediction of T cell epitope candidates. Curr. Protoc. Immunol. 114, 18.19.1–18.19.24. https://doi.org/10.1002/cpim.12.

Peek, L.J., Brandau, D.T., Jones, L.T.S., Joshi, S.B., Middaugh, C.R., 2006. A systematic approach to stabilizing EBA-175 RII-NG for use as a malaria vaccine. Vaccine 24, 5839–5851. https://doi.org/10.1016/j.vaccine.2006.04.067.

Perlmann, P., Troye-Blomberg, M., 2000. Malaria blood-stage infection and its control by the immune system. Folia Biol (Praha) 46, 210–218. https://doi.org/10.1117/12.647092.

Ponomarenko, J., Bui, H.-H., Li, W., Fusseder, N., Bourne, P.E., Sette, A., Peters, B., 2008. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinf. 9, 514. https://doi.org/10.1186/1471-2105-9-514.

Roberts, R.N., Schlarman, M.S., Kariuki, M.M., Lacrue, A.N., Ou, R., Beerntsen, B.T., 2013. Expression profile of the *Plasmodium falciparum* intra-erythrocytic stage protein, PF3D7-1363700. Malar. J. 12, 1–14. https://doi.org/10.1186/1475-2875-12-66.

Salvatore, M., Shu, N., Elofsson, A., 2018. The SubCons webserver: a user friendly web interface for state-of-the-art subcellular localization prediction. Protein Sci. 27, 195–201. https://doi.org/10.1002/pro.3297.

Sanasam, B.D., Kumar, S., 2019. PRE-binding protein of *Plasmodium falciparum* is a potential candidate for vaccine design and development : an in silico evaluation of the hypothesis. Med. Hypotheses 125, 119–123. https://doi.org/10.1016/j.mehy.2019.01.006.

Satchwell, T.J., 2016. Erythrocyte invasion receptors for *Plasmodium falciparum*: new and old. Transfus. Med. 26, 77–88. https://doi.org/10.1111/tme.12280.

Schmidt, C.Q., Kennedy, A.T., Tham, W., 2015. More than just immune evasion: hijacking complement by *Plasmodium falciparum*. Mol. Immunol. 67, 71–84. https://doi.org/10.1016/j.molimm.2015.03.006.

Sette, A., Rappuoli, R., 2010. Reverse vaccinology: developing vaccines in the era of genomics. Immunity 33, 530–541. https://doi.org/10.1016/j.immuni.2010.09.017.

Skwarczynski, M., Toth, I., 2016. Peptide-based synthetic vaccines. Chem. Sci. 7, 842–854. https://doi.org/10.1039/C5SC03892H.

Song, Y., DiMaio, F., Wang, R.Y.-R., Kim, D., Miles, C., Brunette, T., Thompson, J., Baker, D., 2013. High-resolution comparative modeling with RosettaCM. Structure 21, 1735–1742. https://doi.org/10.1016/j.str.2013.08.005.

Takala, S.L., Coulibaly, D., Thera, M.A., Dicko, A., Smith, D.L., Guindo, A.B., Kone, A.K., Traore, K., Ouattara, A., Djimde, A.A., Sehdev, P.S., Lyke, K.E., Diallo, D.A., Doumbo, O.K., Plowe, C.V., 2007. Dynamics of polymorphism in a malaria vaccine antigen at a vaccine-testing site in Mali. PLoS Med. 4, 523–534. https://doi.org/10.1371/journal.pmed.0040093.

Tomar, N., De, R.K., 2010. Immunoinformatics: an integrated scenario. Immunology 131, 153–168. https://doi.org/10.1111/j.1365-2567.2010.03330.x.

Turschner, S., Efferth, T., 2009. Drug resistance in Plasmodium: natural products in the fight against malaria. Mini Rev. Med. Chem. 9, 206–2124. https://doi.org/10.2174/138955709787316074.

Wang, R., Epstein, J., Charoenvit, Y., Maria, F., Rahardjo, N., Gay, T., Banania, J., Chattopadhyay, R., Vega, P.D., Richie, T.L., Tornieporth, N., Doolan, D.L., Kester, K.E., Heppner, G., Norman, J., Carucci, D.J., Cohen, J.D., Hoffman, S.L., 2004. Induction in humans of CD8 + and CD4 + T cell and antibody responses by sequential immunization with malaria DNA and recombinant protein. J. Immunol. 172, 5561–5569. https://doi.org/10.4049/jimmunol.172.9.5561.

World Health organisation, 2019. Malaria. Retrived from. https://www.who.int/news-room/fact-sheets/detail/malaria.

WHO, 2016. Malaria vaccine: WHO position paper-January 2016. Weekly Epidemiol. Rec. 4, 33–52. https://doi.org/10.1371/jour.

Yazdani, S.S., Mukherjee, P., Chauhan, V.S., Chitnis, C.E., 2006. Immune responses to asexual blood-stages of malaria parasites. Curr. Mol. Med. 6, 187–203. https://doi.org/10.2174/156652406776055212.