

3D Assessment of *Dhfr* and *Dhps* Genes Obtained from *Plasmodium vivax* in Anambara State, Nigeria

Isaac Okezie Godwin¹,
Ifeoma Mercy Ekejindu¹,
George Uchenna
Eleje², Chukwuemeka
Chukwubuikem Okoro² and
George Oche Ambrose^{3*}

Abstract

Plasmodium vivax is the most widespread human malaria, putting 2.5 billion people at risk of infection. Its unique biological and epidemiological characteristics pose challenges to control strategies that have been principally targeted against *Plasmodium falciparum*. It occurs across the widest geographic area of the human malaria, extending well beyond the limits of *P. falciparum* into temperate climates. Outside of Africa, *P. vivax* is the dominant species, with relatively high prevalence of infection in the South Asian and Western Pacific regions. *Plasmodium vivax* resistance to antifolates is prevalent throughout Asian countries such as India, which is caused by point mutations within the parasite Dihydrofolate Reductase (DHFR)-thymidylate synthase. Hence, the need to understudy the 3D structure and pattern of mutations associated with both *dhfr* and *dhps* in *Plasmodium vivax* and their closest neighbors. A total number of 390 pregnant women were recruited, 336 of them were taking SP while 54 of the pregnant women were using other anti-malaria drugs for treatment of malaria. The Polymerase Chain Reaction (PCR) technique was used to characterize the specie of the isolated *Plasmodium* while Sanger sequencing method of molecular genotyping was adopted for the subjection of the *Plasmodium* to resistance studies using *dhfr* and *dhps* genes to identify possible mutations. The 3D structure of both *dhfr* and *dhps* was determined using Swiss Model. Quality assessment of the model indicated that the model is reliable. Understanding the 3D structures of *dhfr* and *dhps* will enhance the designs and development of more potent antimalarial..

Keywords: Plasmodium vivax; dhfr; dhps; 3D structure; Swiss Model

- 1 Department of Medical Laboratory Science, Faculty of Health Sciences and Technology, Nnamdi Azikiwe University, Nnewi Campus, Nnewi, Anambra State, Nigeria
- 2 Department of Obstetrics and Gynaecology, Nnamdi Azikiwe University Teaching Hospital, Nnewi, Anambra State, Nigeria
- 3 Center for Malaria and Other Tropical Diseases, University of Ilorin Teaching Hospital, Kwara State, Nigeria

*Corresponding author:

George Oche Ambrose

✉ Ocheab1@gmail.com

Center for Malaria and Other Tropical Diseases, University of Ilorin Teaching Hospital, Kwara State, Nigeria

Received: August 25, 2021; **Accepted:** September 02, 2021; **Published:** September 26, 2021

Introduction

Malaria is a parasitic infectious disease caused by parasites of the genus *Plasmodium* and is transmitted by mosquitoes. Drug resistance is one of the greatest challenges of malaria control programs. *Sulphadoxine*-Pyrimethamine (SP) resistance is linked to substitutions of amino acids in the enzymes Dihydropteroate Synthetase (DHPS) and Dihydrofolate Reductase (DHFR) in the folate biosynthetic pathway [1,2].

Folate metabolism is important for malaria parasite survival, and several enzymes in the pathway have been well characterized to be the targets for several classes of antimalarial drugs. A combination of pyrimethamine and sulfadoxine (Fansidar) was widely used to treat malaria until resistance emerged 10 years

after its introduction [3,4]. Previous studies revealed that point mutations in the Dihydrofolate Reductase (DHFR) of *Plasmodium falciparum* (PfDHFR) and the Dihydropteroate Synthase (DHPS) of *P. falciparum* (PfDHPS) contributed to antifolate and sulfa drug resistance, respectively [5-8]. Due to the conserved nature of the enzymes in the folate metabolic pathway, similar polymorphisms in the *P. vivax* DHFR (PvDHFR) and *P. vivax* DHPS (PvDHPS) have also been suggested to reduce the efficacy of antifolates and sulfa drug treatment in *P. vivax* infection [9].

The 3D structure information of *dhfr* and *dhps* will help us to

understand the mechanisms underlying antifolate and sulfa drug resistance and interaction of their domains with their ligands. The gap between the available sequence of protein and its solved structure was created because 3D structure prediction of proteins requires X-ray crystallography and NMR spectroscopy which consumes a lot of time, tedious approach and generate a large amount of data. *In silico* method of 3D structure prediction has bridged this gap. Computational study of biological sequences has become a very informative field of modern science which is highly interdisciplinary, where statistical and algorithmic methods play a vital role [10]. In this present study we performed sequence analysis on *dhfr* and *dhps* with their secondary and tertiary structures analysis. We finally ensured the quality of the predicted model.

Materials and Methods

DNA extraction and purification from dried blood spots

DNA extracted from bloodspots on filter papers using QIA amp DNA mini kit [11].

Polymerase Chain Reaction (PCR)

Quick load, one Taq, one step polymerase chain reaction was used. Quick load one step PCR master with catalog number NEB MO486S was purchased from Inqaba Biotech., Hart field, South Africa incorporated and was used according to manufacturer's instructions. The malaria diagnosis was established with PCR only and *P. falciparum* and *P. vivax* were distinguished with molecular diagnosis (PCR) only.

Preparation of agarose gel

One-point zero percent Agarose gel (1%) was prepared by dissolving 1.0 g in 100 ml Tris EDTA Buffer. The mixture was then heated in a microwave for 5 minutes to dissolve completely. It was then allowed to cool at 56°C and 6 µl of Ethidium bromide was added to it. The Agarose gel was poured into the electrophoresis chambers with gel comb, and allowed to solidify.

Electrophoresis

Five micro liters of the amplified PCR products was analyzed on 1.0% Agarose gel containing Ethidium bromide in Tris EDTA buffer. Electrophoresis was performed at 90 V for 60 minutes. After electrophoresis the PCR products were visualized by Wealth Dolphin Doc UV transilluminator and photographed. Molecular weights were calculated using molecular weight standard of the maker.

Polymerase chain reaction product cleaning and purification

The PCR products were cleaned using exonuclease/shrimp alkaline phosphatase. Purification was done with ABI V.3.1 Big dye kit according to manufacturer's instructions. The labeled products were then cleaned with ZR DNA Sequencing Clean-Up Kits Sequencing.

The Ultra-pure DNA was sequenced with ABI3500XL analyzer with a 50 cm array, using POP 7 at Inqaba Biotechnical Industries Ltd. (Hatfield, South Africa). Sequences data generated were analyzed with Geneious version 9.0.5 and phylogenetic trees were constructed using neighbor joining. The sequences were subsequently deposited in the National Center for Biotechnology Information (NCBI) database with their corresponding accession IDs (**Table 1**).

The gene sequences were later translated using Expsy to their corresponding amino acids sequences.

Table 1: *Dhfr* and *Dhps* genes with their corresponding accession IDs.

S. No.	Gene	NCBI Accession ID
1	<i>dhfr</i>	MT577725
2	<i>dhps</i>	MT577726

Secondary structure prediction

Secondary structures of *dhfr* and *dhps* sequences were predicted using GOR IV. The GOR IV method is based on information theory and was developed by J. Garnier, D. Osguthorpe and B. Robson [12]. Predict Protein was also employed for calculating and analyzing the secondary structure features of *dhfr* and *dhps* sequences [13].

3D structure prediction using homology approach

In silico prediction of 3D structure of protein is based on Threading, Ab-Initio and Homology modeling. Homology modeling can only be useful for 3D structure prediction of the target protein if sequence similarity search of target sequence is more than 60%. Template was searched against Swiss-model (an online tool for 3D prediction) template library using Blast and HHBlits algorithms. Based on maximum identity and GMQE values, the best templates with PDB IDs: 4yuu.A and 4i31.1.A were selected automatically for modeling. Homology modeling was done using the template sequences whose structures were solved by X-ray diffraction.

Quality and reliability assessments

After the 3D model structure of *dhfr* and *dhps* were generated, energy minimization was performed by GROMOS96 force field in a Swiss-PDB Viewer. Stereochemical analyses and structural evaluation were carried out using ProSA-web [14]. Also, Procheck Ramachandran plot and Z-scores were carried out. Visualization of the generated model was carried out using Discovery studio visualizer.

Results and Discussion

The secondary structure of *dhfr* and *dhps* was predicted using GOR IV and the result revealed that the sequences are mainly composed of Alpha helix and Beta sheets. It is predicted from the result that *dhfr* is chiefly made up of random coil (47.67%),

extended strand (28.49%) and alpha helix (23.84%) while *dhps* constitute mainly of random coil (59.36%) and extended strand (36.90%).

The 3D structure of protein is very important in understanding Proteins functions, their localization and interactions. The most common structure prediction method is Homology modeling. The prediction of the tertiary structure or 3D structures of *dhfr* and *dhps* were done using Swiss Model with an initial search for solved templates with similar sequences. Templates having the best E-value, maximum number of query sequence covered and percentage similarity were selected for homology modeling.

4yuu.A and 4i31.1.A were selected for homology modeling which are X-ray diffraction models for *dhfr* and *dhps* respectively. The 3D structures of the modeled *dhfr* and *dhps* are shown in (Figure 1). The Z-score and ramachandran plot, among several assessment methods, were used to check the quality and reliability structure of the model. The Z-score is indicative of overall model quality and is used to check whether the input structure is within the

range of scores typically found for native proteins of similar size [15]. The PROSA web was used to find the value of Z-score of both template and query. The Z-score of *dhfr* and *dhps* were -0.93 and -0.81 respectively (Figure 2). The stereochemical quality of a protein was checked using procheck server. This is done by analyzing residue-by-residue geometry and overall structure geometry. The quality of this model was assured by using this tool to determine ramachandran plot. The result of the ramachandran plots showed 84.6% and 100% of residues in the most favorable region for *dhfr* and *dhps* respectively, indicating that *dhps* is most reliable and good quality model (Figure 3). A model having more than 90% residues in the most favorable region is considered as a good quality model. Furthermore, the reliability of the model was assessed by ERRAT that analyzes the statistics of non-bonded interactions between different atom types and plots the value of error function versus position of a 9-residue sliding window [15]. The ERRAT result shows the overall quality for both *dhfr* and *dhps* as a factor of 100 (Figure 4).

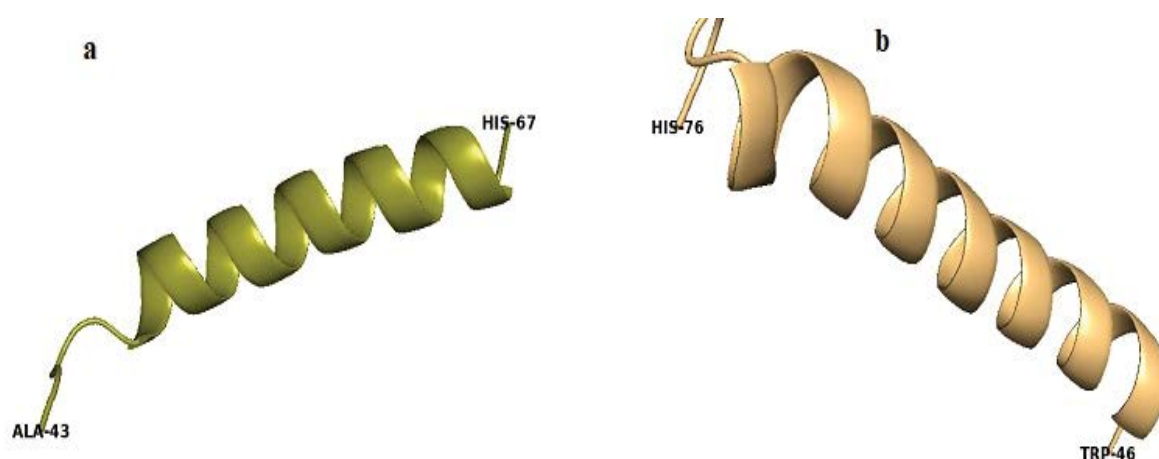


Figure 1 3D structure; (a): *dhfr*; (b): *dhps*.

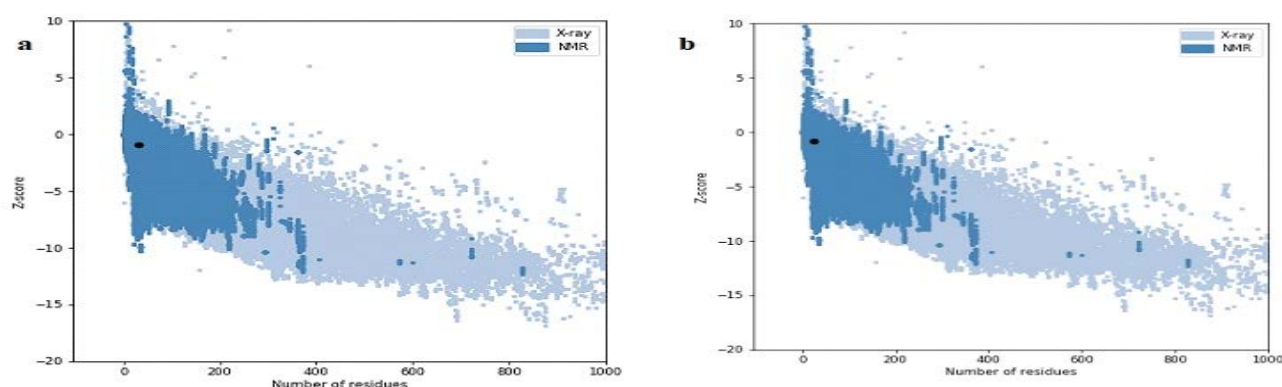


Figure 2 Z-score of (a): *dhfr*; (b) *dhps* using PROSA web.

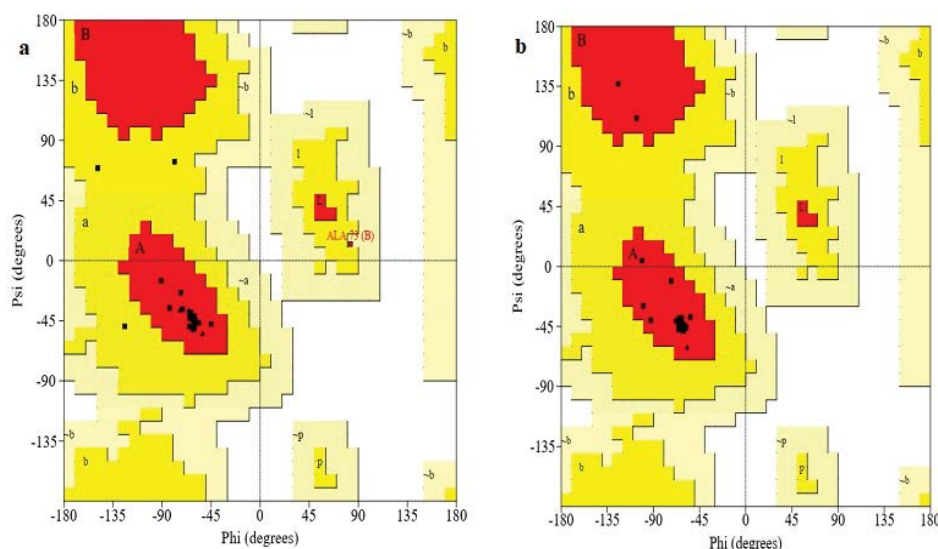


Figure 3 Ramachandran plot of (a): *dhfr* (b): *dhps* using Procheck

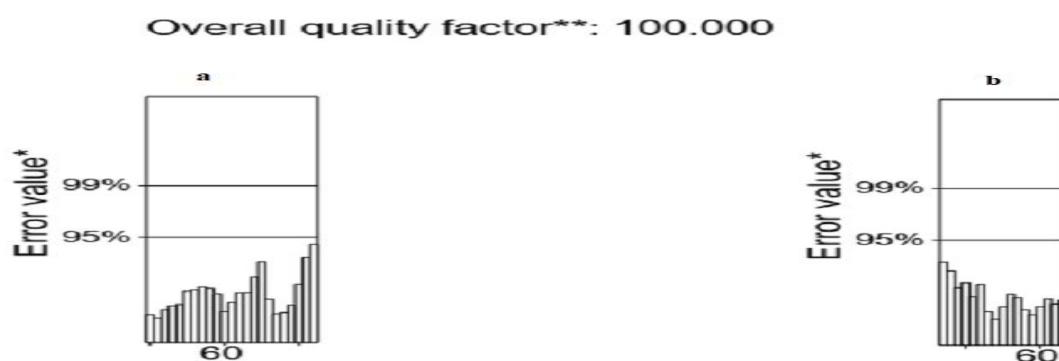


Figure 4 Overall quality factor checked by ERRAT; (a): *dhfr* (b): *dhps*

Conclusion

The major aim of this study was to perform sequence analysis, structure analysis and homology modeling on *dhfr* and *dhps* resistance genes. We have adopted various sequences and structure analysis tools which were useful in understanding the sequence and its structure. Finally, as part of our study, we considered homology modeling approach to propose the first 3D structure of the *dhfr* and *dhps* resistance genes obtained from Anambra State. The predicted 3D structure will give more insight in understanding the function and structure of the proteins. In addition, this structure can be used in drug design or understanding interactions between proteins.

References

1. Cowman AF, Mony MJ, Biggs BS, Cross GA, Foote SJ (1998) Amino acid changes linked to pyrimethamine resistance in the dihydrofolate reductase-thymidylate synthetase gene of *Plasmodium falciparum* dihydropterin pyrophosphokinase-dihydropterote synthetase. *J of Bioorganic and Medi Chem* 14: 4433-4443.
2. Brooks DR, Wang P, Read M, Watkins WM, Sims PF (1994) Sequence variation of the hydroxymethyldihydropterin in pyrophosphokinase: Dihydropterote synthetase genes in lines of the human malaria parasite, *Plasmodium falciparum* with differing resistance to sulfadoxine. *Eur J of Biochem* 224: 397-405.
3. Thaithong S, Chan SW, Songsomboon S, Wilairat P, Seesod N, et al. (1992) Pyrimethamine resistant mutations in *Plasmodium falciparum*. *Mol Biochem Parasitol* 52: 149-157.
4. Pinichpongse S, Doberstyn EB, Cullen JR, Yisunsri L, Thongsombun Y, et al. (1982) An evaluation of five regimens for the outpatient therapy of *falciparum* malaria in Thailand 1980-81. *Bull World Health Organ* 60: 907-912.
5. Cowman AF, Morry MJ, Biggs BA, Cross GA, Foote SJ. (1988) Amino acid changes linked to pyrimethamine resistance in the dihydrofolate reductase-thymidylate synthase gene of *Plasmodium falciparum*. *Proc Natl Acad Sci USA* 85: 9109-9113.
6. Wang P, Read M, Sims PF, Hyde JE. (1997) Sulfadoxine resistance in the human malaria parasite *Plasmodium falciparum* is determined by mutations in dihydropterote synthetase and an additional factor associated with folate utilization. *Mol Microbiol* 23: 979-986.
7. Brooks DR, Wang P, Read M, Watkins WM, Sims PF, et al. (1994) Sequence variation of the hydroxymethyldihydropterin

pyrophosphokinase: Dihydropteroate synthase gene in lines of the human malaria parasite, *Plasmodium falciparum*, with differing resistance to sulfadoxine. *Eur J Biochem* 224: 397-405.

8. Triglia T, Menting JG, Wilson C, Cowman AF (1997) Mutations in dihydropteroate synthase are responsible for sulfone and sulfonamide resistance in *Plasmodium falciparum*. *Proc Natl Acad Sci USA* 94: 13944-13949.
9. Korsinczyk M, Fischer K, Chen N, Baker J, Rieckmann K, et al. (2004) Sulfadoxine resistance in *Plasmodium vivax* is associated with a specific amino acid in dihydropteroate synthase at the putative sulfadoxine-binding site. *Antimicrob Agents Chemother* 48: 2214-2222.
10. Giancarlo R, Siragusa A, Siragusa E, Utro F (2007) A basic analysis toolkit for biological sequences. *Algorithms for Molecular Biology* 2: 1-10.
11. Chandrasekhar BN, Jagannath M, Mulakkapurath NM, Sukriti M, Phani K (2016) Protocol for DNA Purification from Dried Blood Spots. Quality Improvement Agency amplification DNA mini and Blood mini Handbook, fifth edition, Public Library of Science. San Francisco California United States of America 42-43.
12. Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ (1998) JPred: A consensus secondary structure prediction server. *Bioinformatics (Oxford, England)* 14: 892-893.
13. Rost B, Yachdav G, Liu J (2004) The predict protein server. *Nucleic acids research* 32: 321-326.
14. Wiederstein M, Sippl MJ (2007) ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research* 35: 407-410.
15. Idrees S, Nadeem S, Kanwal S, Ehsan B, Yousaf A, et al. (2012) In silico sequence analysis, homology modeling and function annotation of *Ocimum basilicum* hypothetical protein G1CT28_OCIBA. *Int J Bioautomat* 16: 111-118.