Received: 17 December 2023 | Revised: 29 May 2024 | Accepted: 12 July 2024 | Published online: 24 July 2024

RESEARCH ARTICLE

Medinformatics
2025, Vol. 00(00) 1–9
DOI: 10.47852/bonviewMEDIN42022318

BON VIEW PUBLISHING

Machine Learning Predicts 1-[(4-Fluorophenyl) Methyl] Indole-2,3-Dione as Drug Lead for Peptide Deformylase in *Plasmodium falciparum*

Omiyale Olumakinde Charles[1,*], Zainab Hassan[2], Okunbi Favour Onasokhare[3], Aisha Bello[2], Olaiyapo F. Oluwafemi[4], Blessing Ojajuni[5], Sulaimon Olajuwon Abdul[6], Roseline Boboye[7], Martin Egbulefu[8], Chiziterem Jaachinma Ezeano[9], Ahmed Rufai Nurudeen[10], Rasaq Awosemo[11] and Marian Mark[12]

[1]*Department of Pharmacology, University of Lagos, Nigeria*

[2]*Department of Biochemistry, Lagos State University, Nigeria*

[3]*Department of Microbiology, Mountain Top University, Nigeria*

[4]*Department of Mathematics, Emory University, USA*

[5]*Department of Computer Science, Houdegbe North American University, Benin Republic*

[6]*Department of Pharmacy, Olabisi Onabanjo University, Nigeria*

[7]*Department of Biochemistry, Federal University of Technology Akure, Nigeria*

[8]*Department of Electronics and Electrical Engineering, Imo State University, Nigeria*

[9]*Department of Computer Science, Nnamdi Azikiwe University Awka, Nigeria*

[10]*Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology, Ghana*

[11]*Department of Biology and Environmental Science, University of New Haven, USA*

[12]*Department of Chemistry, University of Ibadan, Nigeria*

Abstract: Malaria continues to be a serious illness for world health. Deadly parasites that cause malaria infect female Anopheles mosquitoes, which then bite humans, acting as the disease's vector. Since *Plasmodium falciparum* is the deadliest of the five *Plasmodium* species and research indicates that medicines for malaria are increasingly revealing drug-resistance mechanisms, it is imperative that new and effective medications be developed. This project aims to identify and produce a novel therapeutic lead against a *P. falciparum* target that has been validated. Target2Scan, a programmed tool, was employed in this investigation to find potential therapeutic targets for *P. falciparum* peptide deformylase (PDF). After the tool received the target signature, it ran a Basic Local Alignment Search Tool (BLAST) methodology to find targets that resembled PDF. To ascertain the binding affinities of protein–ligand complexes, molecular docking was performed using PyRx and CBDock. Of the 12,561 compounds produced, 11,304 were used as the training set and 1,256 as the test set. Six ligands were produced by machine learning as potential therapeutic leads, and using molecular docking, 1-[(4-fluorophenyl) methyl] indole-2,3-dione showed the greatest binding effect on PDF when compared to 5-chloro-1-(2-phenylethyl) indole-2,3-dione. The higher binding affinity of 1-[(4-fluorophenyl) methyl] indole-2,3-dione over 5-chloro-1-(2-phenylethyl) indole-2,3-dione and other ligands on PDF suggests that it has the ability to suppress *P. falciparum* PDF activity.

Keywords: Target2Scan, molecular docking, malaria, peptide deformylase, *Plasmodium falciparum*, BLAST

1. Introduction

Approximately 6.6 billion people are still susceptible to malaria, which results in 200 million cases being reported in 97 countries, with nearly 600,000 deaths from the disease [1]. For this reason, malaria remains a major problem. Four distinct species of the unicellular parasite *Plasmodium* parasitize human red blood cells, resulting in malaria. As a result, the parasites multiply inside the blood cells and transfer the disease from person to person by Anopheles mosquitoes that feed on human blood. Furthermore, complicating matters is the fact that this "arms race" is being built step-by-step, making disease control the top priority [2].

*Corresponding author: Omiyale Olumakinde Charles, Department of Pharmacology, University of Lagos, Nigeria. Email: 199096019@live.unilag.edu.ng

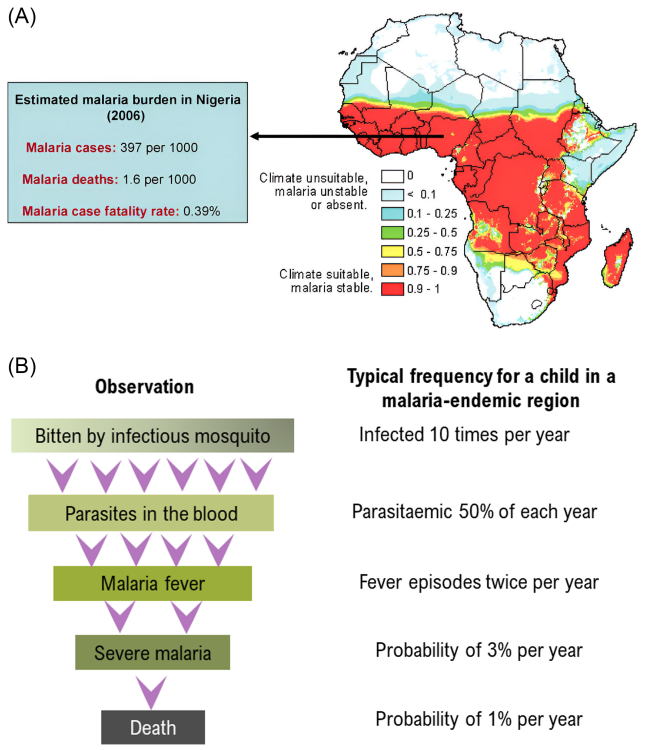

Figure 1. Malaria distribution and progression. (A) Human malaria distribution in Africa and its estimated burden in Nigeria. Adapted from WHO, 2008a. World Malaria Report 2008. World Health Organization, Geneva, pp. 7–15, 99–101. (B) Progression of malarial disease in a malaria-endemic region. Adapted from MARA/AMRA

P. falciparum is estimated to be responsible for 300–500 million clinical cases reported annually and 1–3 million fatalities annually, accounting for the majority of infections and deaths globally [3]. See Figure 1 for more information on malaria distribution and progression. The majority of mortality is caused by *Plasmodium falciparum*, with other *Plasmodium* species accounting for only 1% of deaths [4]. *Plasmodium* species are newly discovered parasites that infect female Anopheles mosquitoes and arthropod hosts alternately. Their life cycle is refilled by the interaction of various zoite cell types. In the ten years prior, many epidemiologists have voiced concerns about the increasing resistance of several major human pathogenic bacteria and parasites to a class of drugs often used in medicine. For instance, research seems to show that *Plasmodium* is not very responsive to aromatic drugs [5].

The implementation of a law-and-order system in the healthcare industry will undoubtedly have unfavorable effects, but that is not the focus of this study; instead, the development of a new antimicrobial medication may be the result, and resistant microbes will most likely survive the storm. This is not feasible, though, as lab-based conventional screening is limited to identifying compounds as potential pathogen candidates based on their capacity to combat the pathogen in the lab. It also requires little rigorous work to confirm whether the candidates are successful or unsuccessful in target cells. However, as one component emerges, another portion of the entire also manifests itself, and this is the process of developing alternate techniques as well as envisioning something new. It concerns me, and for this reason, the first I would face is the molecular target that will succeed the whole panel of the libraries, starting from the drugs to the final targets comprising thousands of products designed, and I would test the items that offer working chances one by one. I think that the hardest task is picking the right target audience as it can have a tremendous impact on the success of my campaign. It is generally agreed [6] that such a target should (a) be part of the genetic structure of several human infectious pathogens that essentially harbour, (b) plainly not located in human cells, (c) serve a contribution to the virulence pathway, and (d) not be restricted by the common antimicrobials (drug-resistant).

It clearly points out that peptide deformylase (PDF) could be the most suitable creature since it holds a certain amount of contribution for research purposes and will definitely serve as a good example. Besides accepting the whole substrate moiety, the multiple binding sites of the PDF enzyme govern the reaction process delaying fMet deformylation product formation by proteins synthesized in bacteria cells and organelles such as chloroplasts and mitochondria [7]. The necessity of deformylation makes PDF an attractive target in the development of new antimalarial drugs. The identification of PDF as an attractive target for new drugs to treat infectious diseases has led to an extensive search for PDF inhibitors.

Yet, as the global big data and artificial intelligence (AI) revolution progresses, it will lead to the impact of those technologies in bioinformatics, genetic science, and drug exploration even in such a narrow field of knowledge. The field of bioinformatics and computational biology along with AI has earned increased attention as data-driven methods and AI technologies are leveraged. It is necessary to point out that for genomics/proteomics and drug discovery, one can use machine learning within these areas [8–10]. Machine learning, a division of AI, provides a set of tools that can improve discovery and decision-making for well-specified questions with abundant, high-quality data. In small drug molecule discovery, the recent surge in the development of machine learning-based protein–ligand interaction tools has helped researchers in identifying small drug molecules that can interact with a particular drug target [11–13]. Mechanical learning (ML) methods have been developed in this context to identify compound-target occurrences and to establish an inverse relationship between compounds and their corresponding targets. Our tools' workflow is built on the integration and thoughtful selection of the key components that together form the robust bioinformatics, computational biology, and AI-driven drug discovery revolutions, all of which are combined into a single, straightforward workflow. The tool user may supply dependent signatures of the molecule in the target's amino acid sequence format or the target's nucleotide sequence when asked to find drugs against a new drug target. In the end, the programmatic workflow for locating compounds that bind and target the given molecule begins as follows: The instrument performs a Basic Local Alignment Search Tool (BLAST) program against the new target signatures provided by the user. From the output, many known protein drug targets are recognized, which are somehow similar to the new target provided by the user. Since they are accessible with data available, some of the high-interest protein drug targets that are screened using the tool's BLAST technique as well as the drug lead generation based on quantitative structure-activity relationship (QSAR)-scaffolds study could be used in PubChem. The tool will obtain experimental data regarding the inhibitory activity of PubChem compounds on the target as well as the molecular descriptors of the active compounds in order to generate QSAR models during the lead drug generation process based on the QSAR technique. In this instance, the 6–7 million ligand library of PubChem was used in the construction, validation, and prediction stages of the ML-based AutoQSAR technique for drug lead generation. The interactions

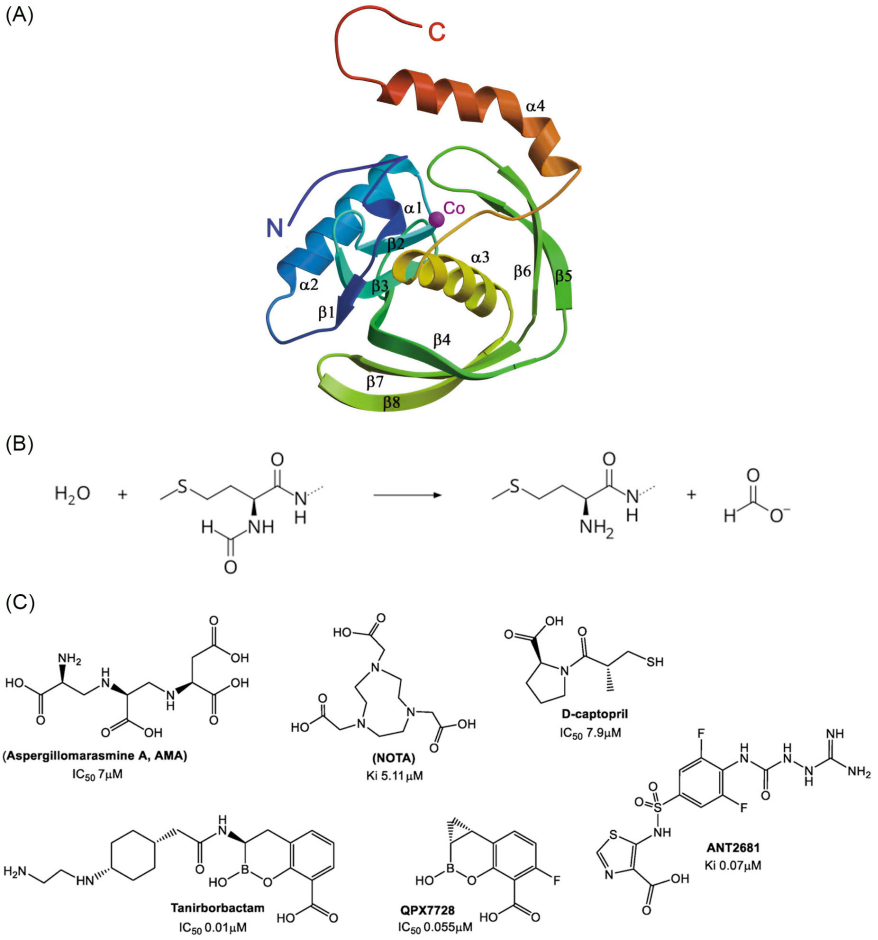

Figure 2. ***Plasmodium falciparum*** **peptide deformylase (PfPDB) structure and function. (A) Secondary structure of PfPDB (ribbon representation) showing α-helices, β-sheets, and loops. (B) Catalytic reaction mechanism of PfPDB. (C) Structures of six drug leads against PfPDB: (1) indole-2,3-dione, 5-chloro-1-(2-phenylethyl), (2) [2,3-f] 5-ethylthieno indole-6,7-dione, (3) 1-[methyl (4-fluorophenyl)] 2,3-dione of indole, (4) 1-[methyl 5-bromothiophen-2-yl] 2,3-dione of indole, (5) indole-2,3-dione 7-bromo-1-(thiophen-3-ylmethyl), and (6) indole-2,3-dione 7-bromo-1-(thiophen-2-ylmethyl)**

between compounds, identified as drug leads using the tool, and the protein targets were also predicted through in silico modelling. See Figure 2 for more information on Plasmodium falciparum peptide deformylase (PfPDB) structure and function. AutoDock was a popular high throughput virtual screening package that was used programmatically through the tool. Then the calculators can be built, and the results should be put into the temporary storage folder.

See Figure 1, malaria continues to be one of the most devastating infectious diseases worldwide, especially in low-resource settings. This life-threatening disease, caused by the *Plasmodium* parasites, particularly "*Plasmodium falciparum*," is transmitted to humans through the bites of infected Anopheles mosquitoes. Malaria is responsible for approximately 200 million cases and nearly 600,000 deaths annually, predominantly in sub-Saharan Africa, where it is endemic [1, 14]. The high morbidity and mortality rates associated with **P. falciparum** highlight the urgent need for effective control and treatment strategies [15, 16].

Socioeconomic factors and limited access to healthcare contribute significantly to the global burden of malaria. The disease is particularly prevalent in regions where climatic conditions, such as consistent temperatures and high humidity, favor mosquito breeding and transmission [17]. In Nigeria, which accounts for a quarter of all malaria cases in Africa, the disease exhibits year-round transmission in the south and seasonal peaks in the north, exacerbating the public health challenge [18].

Malaria control efforts have focused on vector management, including the widespread use of insecticide-treated bed nets and indoor residual spraying (IRS), which have successfully reduced mosquito populations and transmission rates [19]. The introduction of the RTS,S/AS01 malaria vaccine represents a significant advancement, showing promise in reducing the incidence of clinical malaria, particularly in young children [20].

However, the emergence and spread of drug-resistant strains of *P. falciparum* pose a formidable challenge to existing malaria control measures. The rapid evolution of resistance to frontline antimalarial drugs, such as artemisinin, necessitates the development of new therapeutic agents [21, 22]. Traditional drug discovery methods, which are often time-consuming and costly, are proving inadequate in keeping pace with the emergence of drug-resistant strains.

Recent advances in computational biology and bioinformatics offer promising alternatives for accelerating drug discovery processes. Machine learning and AI have revolutionized the field by enabling the analysis of vast biological datasets to predict potential drug candidates with high accuracy and efficiency. These

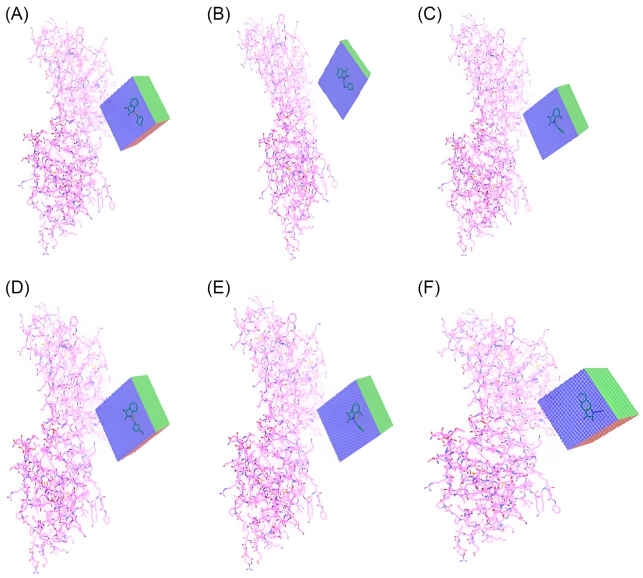

Figure 3. Grid boxes of ligands (figures courtesy of the authors). (A) 1-[(4-fluorophenyl) methyl] indole-2,3-dione. (B) 5-chloro-1-(2-phenylethyl) indole-2,3-dione. (C) 7-bromo-1-(thiophen-3-ylmethyl) indole-2,3-dione. (D) 1-[(5-bromothiophen-2-yl) methyl] indole-2,3-dione. (E) 7-bromo-1-(thiophen-2-ylmethyl) indole-2,3-dione. (F) 5-ethylthieno[2,3-f] indole-6,7-dione

technologies facilitate the identification of novel compounds, prediction of their biological activities, and optimization of their pharmacokinetic and pharmacodynamic properties [11–13].

In our study, we focus on the identification and production of a novel therapeutic lead targeting PDF in *P. falciparum*. PDF is an essential enzyme in bacterial protein synthesis, absent in human cells, making it an attractive target for antimalarial drug development [7]. Using a combination of molecular docking, QSAR models, and machine learning techniques, we identified 1-[(4-fluorophenyl) methyl] indole-23-dione as a potential lead compound with high binding affinity to PDF. Shown in Figure 3.

The integration of machine learning with traditional drug discovery methods in our research underscores the potential of these advanced computational techniques to streamline the identification of effective antimalarial compounds. By addressing the pressing need for new treatments, our study contributes to the global effort to control and ultimately eradicate malaria.

2. Materials and Methods

2.1. Identification of target with data availability on PubChem

Target2Scan is a software tool that the scientists used in their work. It incorporates several features from the early days of bioinformatics, computational biology, and AI-driven drug launches, all of which come together in a single process. This strategy is sometimes called "hybridizing" the conventional drug development procedure for the contemporary day. They employ extensive protein, genomic, and other data collection methods to thoroughly train their machine learning algorithms. The algorithm will then be applied to obtain drug targets. Using modeling and training data from pharmacological targets with known ligands, Target2Scan is tasked with identifying the ligands of novel targets. Finding new treatments, understanding how proteins bind to different drugs, and determining if a medicine is successful for a particular patient are all made possible by it. Additionally, it ensures optimum ligand bioactivity and supports perimeter markers. Although the core of Target2Scan is data mining, which offers the tool greater leverage in making educated decisions to recommend cost-effective preventive actions, AI and machine learning reduce the tool's error-proneness and time-consuming processes. The Target2Scan in general is an ultra-potent weapon in the arsenal of drug discovery.

Using the target's signature, the tool was configured to identify possible medications or ligands for the target. The chemicals are ligands for the target that the tool retrieved from PubChem. The molecular descriptors of the ligands are summarized in Table 1 below. We looked for druggable or well-known protein targets that were comparable to the new target using the BLAST. Here, the software limited the findings to only those targets for which PubChem, an open public database of chemical compounds and biological functions, provided background information. The list of possible medications or ligands for the targeted receptor could now be reduced thanks to the drug discovery tool.

2.2. Ligand retrieval for the target and its characteristics

The chemicals are ligands for the target that the tool retrieved from PubChem. The tool not only identifies new drug leads but also ensures their optimum bioactivity and supports perimeter markers. The possible medications and ligands can be further examined using these descriptions.

Table 1. Molecular descriptors of ligands

Parameters	Ligand 3	Ligand 4
CID	13603431	28492614
IUPAC name	5-ethylthieno[2,3-f] indole-6,7-dione	1-[(5-bromothiophen-2-yl) methyl] indole-2,3-dione
Molecular weight (G/Ml)	231.27	322.18
Heavy atom count	16	18
XLogP	2.3	3.1
Complexity	344	376
Hydrogen bond acceptor counts	3	3
Monoisotopic mass (Da)	231.0354	320.9459
Rotatable bond counts	1	2
TA/SA (A^2)	65.6	65.6

Table 2. Ligands and their binding affinities

S/N	IUPAC name	Binding affinity
1	1-[(4-fluorophenyl) methyl] indole-2,3-dione	−4.3
2	5-chloro-1-(2-phenylethyl) indole-2,3-dione	−4.2
3	7-bromo-1-(thiophen-3-ylmethyl) indole-2,3-dione	−4.1
4	1-[(5-bromothiophen-2-yl) methyl] indole-2,3-dione	−4.0
5	7-bromo-1-(thiophen-2-ylmethyl) indole-2,3-dione	−3.9
6	5-ethylthieno[2,3-f] indole-6,7-dione	−3.8

2.3. Retrieval of compounds that share structural characteristics with recognized ligands for the target

Compounds with structures similar to the known ligands for the target were retrieved by the tool along with their molecular characteristics. These compounds recovered molecular characteristics including rotatable, monoisotope, XLogP, complexity, heavy atom, molecular weight, hydrogen bond acceptor, and TPSA (topological polar surface area).

2.4. Generation of drugs leads by machine learning

Drug target interaction prediction studies can be analyzed and resolved with the aid of computer technology called machine learning. This method was developed to forecast how prospective medications or ligands would interact with the target protein. The target protein's details were obtained using BLAST search results, which served as the foundation for the employed technique. The results of the BLAST searches were then used to retrieve information about potential drugs/ligands from PubChem, a public database of chemicals and their properties. The final dataset was generated by an AutoQSAR script, which combined the information from both sources. Get PubChem Property (which was used to overwrite the downloaded file to final_data_frame) and AutoQSAR script (generated TOP_CID_123, which retrieved data from PubChem that are related to the protein target given).

The machine learning algorithms that comprise AutoQSAR are AI software that builds, validates, and uses QSAR models. A software program called AutoQSAR integrates machine learning methods with statistics related to the creation and development of QSAR models. A vast number of models are first created, and then an iterative process that involves choosing the best models and optimizing the selected model is implemented. This is completed by testing the models against a series of validation compounds to ensure their correctness and robustness. The model works well for predicting the characteristics of new chemicals that will be applied to environmental chemistry, toxicological research, and medicinal development.

In order to ascertain the biological qualities (such as solubility or reactivity) of chemical molecules, it filtered the enormous libraries of bioactive compounds (PubChem) based on their structural composition.

Two sets of compounds were identified based on their properties that resembled well-known ligands: a training set and a holdout set. The machine learning model was built using the training component of the data, and its accuracy was verified using the test portion of the data. Processing in this manner is required since it makes it easier to examine the model and identify its weak points.

Under the moniker AutoQSAR, QSAR created two linear and nonlinear regression models based on a variety of chemical descriptor types. The models with the highest R2 value or the models' R2 values closest to 1 were chosen from among those QSAR models, and the prediction model was the one with the R2 value closest to 1. The screening was done, yielding several chemicals from the PubChem database. The Lipinski's drug-likeness criteria (which relate molecular descriptors of the retrieved compounds/ligands with known drug activity) were used to eliminate undesirable compounds; the top 50 compounds of the prediction that satisfied the criteria were printed out as drug leads against the target. The speed, efficiency, robustness, flexibility, and accuracy make AutoQSAR a good choice for machine learning.

2.5. Ligand and receptor (target model) preparation for docking

We used the AutoDock-Vina 1.5.6 tool to get the receptor and ligands ready for docking. The Protein Data Bank in Europe (PDBe: https://www.ebi.ac.uk/pdbe/entry/pdb/2VGB) is a centralized source of structural data for biological macromolecules. This is where the receptor model was obtained in PDB format. After that, AutoDock-Vina was used to analyze the receptor model and convert it to PDBQT format. Water molecules were eliminated, polar hydrogens were added, and Kollman charges were incorporated during this conversion. Moreover, the AutoDock-Vina tool was used to convert the ligands into PDBQT format.

2.6. Molecular docking

AutoDock-Vina (version 1.5.6) was used to optimize known drugs and identify novel binders by predicting their binding mode and affinity of each ligand given. The available open-source program has a powerful heuristic search algorithm that allows for fast and accurate docking of molecules to a target protein. It has a user-friendly interface that sets docking simulations and results analyses very easy.

In order to forecast the intensity of the binding between the ligand and the target and to generate possible postures that are subsequently assessed using scoring functions, autodocking was utilized to position the ligands within the target's binding region. Furthermore, this method produced many poses, which were then evaluated using scoring systems to determine their feasibility.

The binding affinity and strength of the intermolecular interactions between the ligand and the target are roughly predicted by the mathematical functions known as scoring functions.

2.7. Visualization and determination of protein–ligand interaction using PyMOL program

Using autodocking, the ligands were positioned inside the target's binding site to anticipate the strength of the binding between the ligand and the target. This allowed for the production of possible poses, which were subsequently assessed using scoring systems. This method also produced other positions, which were then evaluated using scoring systems to see how feasible they were.

The scoring functions are mathematical functions that are used to predict, to some extent, the intensity of the intermolecular interactions and the binding affinity between the ligand and target.

To determine the chemical and geometric elements that influence the binding behavior of the tested ligands, the chemical groups involved in specific interactions were discovered, and the interaction's geometry (distance and angle), or the preferred groups, was examined. The target molecules and the molecule with the highest affinity toward the negatively bound ligand, which were investigated thereafter for visibility and additional analysis, are in poses that indicate the molecular interaction.

The protein–ligand complex's interactions, or bonds, were manually identified using PyMOL version 1.7.4.5's molecular graphics tool. The parameters for determining the bonds are the angle, the atoms (carbon atoms in the case of hydrophobic contacts), and the distance limit (4.0^{-} for hydrophobic bonds, 4.1^{-} for hydrogen bonds, and 5.5^{-} for salt bridge bonds).

3. Results and Discussion

3.1. Peptide deformylase as the protein of interest

In bacteria and in organelles found in some eukaryotes, such as mitochondria and chloroplasts, PDF is in charge of eliminating the formyl group from the N-terminal fMet amino acid of freshly synthesized proteins [23, 24]. Since bacterial PDF activity is essential to their survival [25], researchers are looking at using bacterial PDF as a possible target for novel therapeutics. Currently, a number of pharmaceutical companies are focused on creating antibacterial medicines that specifically target PDF [25–27]. In eukaryotic cells, such as *P. falciparum*, where they are thought to be found in the apicoplast based on a potential leader peptide sequence, the role of bacterial PDFs is different [28]. The Apicomplexa phylum, which includes malaria parasites, has the apicoplast, a crucial organelle encircled by many membranes. The catalytic activity of the recombinant PfPDF indicates that the malaria parasite's apicoplast is where the formylation/deformylation process takes place. The growth of malaria parasites was moderately inhibited by two *E. coli* PDF (EcPDF) inhibitors. Thus, PfPDF offers new opportunities to design novel antimalarial drugs [28].

3.1.1. P. falciparum peptide deformylase structure

With a root mean square deviation (rmsd) of 2.2 Å for 161 equivalent Cα atoms, PfPDF's structure is comparable to that of the *E. coli* enzyme (Protein Data Bank code 1DFF). Nonetheless, the two architectures differ in a few significant ways. For instance, in PfPDF, three insertions are made in the loop that joins β strands β3, and the presence of Pro 220 causes the C-terminal helix (residues 213–230) to become kinked. The *P. falciparum* enzyme's active site and helix α2's general structure are altered by the three insertions in the loop that connects β strands β3 and β4. The distinct substrate selectivity and activity of the *P. falciparum* enzyme are probably influenced by these variations.

PyMOL, a system software, was used to view the 3D structure of the protein target (PDF), which was obtained from the Protein Data Bank (PDB) database. The resolution of the structure is 2.18 Å. Two subunits totaling more than a thousand amino acid residues make up the target. Fourteen β-sheets and two ten α-helices encircle the central α-helix, which houses the metal ion and the active site (bound by a conserved HEXXH motif).

3.1.2. Reaction catalyzed by peptide deformylase

Peptide deformylase (EC 3.5.1.88) is a type of enzyme in enzymology that catalyzes the following chemical reaction:

$$\text{Formate} + \text{methionyl peptide} + \text{formyl-L-methionyl peptide} + H_2O$$

The hydrolysis of formyl-L-methionyl peptide into formate and methionyl peptide is catalyzed by this enzyme. It belongs to the category of hydrolases, more precisely amidohydrolases, which break down links between carbon and nitrogen apart from peptide bonds. This family of enzymes is known by its systematic name, formyl-L-methionyl peptide amidohydrolase.

3.2. Artificial intelligence-generated medication leads

Six distinct ligands that may be used as prospective drug leads were identified by using the machine learning AutoQSAR technique. Out of the top 50 compounds that satisfied Lipinski's requirements for drug similarity, these particular compounds were identified as prospective therapeutic leads. The machine learning system identified a set of pharmacological leads that include chemicals like 1-[(4-fluorophenyl) methyl], 5-ethylthieno[2,3-f] indole-6,7-dione, and 5-chloro-1-(2-phenylethyl) indole-2,3-dione. Additionally, it identified three different compounds: 1-[(5-bromothiophen-2-yl) methyl] indole-2,3-dione, 7-bromo-1-(thiophen-3-ylmethyl) indole- 2,3-dione, and 7-bromo-1-(thiophen-2-ylmethyl) indole-2,3-dione. Structures available in Figure 3. The tables below describe the drug leads and their molecular descriptors.

3.3. Molecular docking

Scoring functions produced by molecular docking are exhaustiveness = 8, and energy level = 4, respectively, for all drug leads (ligands). All six ligands were examined to determine the ligand with the highest binding affinity (lowest binding energy) for PDF as a target. The size of the resulting grid box was:

$$\text{Size\_x} = 40;\ \text{Center\_x} = 0$$

$$\text{Size\_y} = 40;\ \text{Center\_y} = 0$$

$$\text{Size\_z} = 40;\ \text{Center\_z} = 0$$

3.4. Peptide deformylase-1-[(4-fluorophenyl) methyl] indole-2,3-dione interaction

The maximum binding affinity was obtained when 1-[(4-fluorophenyl) methyl] indole-2,3-dione bound to PDF of −4.3. Interactions between this ligand and the protein target were determined and visualized with PyMOL. Distances of interactions were measured as follows:

Hydrogen bond interaction = 3.6 (ILE-186 and hydrogen atom of the ligand)
Hydrophobic interactions = 3.9 (HIS-191 and hydrogen atom of the ligand), 4.3 (LYS-186 and hydrogen atom of the ligand)
Salt bridge = 5.5 (LEU-187 and oxygen atom of the ligand)

The orientation of the ligand bound to the receptor is shown in Figure 4a and b below.

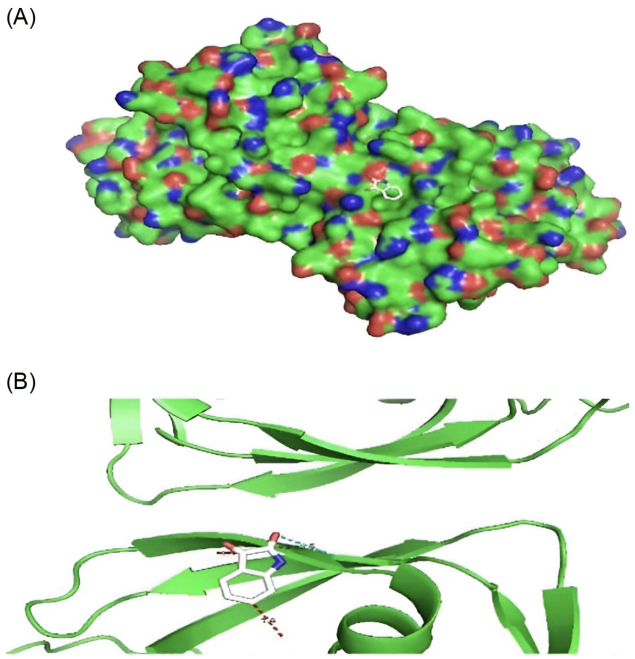

Figure 4. Peptide deformylase-1-[(4-fluorophenyl) methyl] indole-2,3-dione interaction. (A) Surface representation of peptide deformylase-1-[(4-fluorophenyl) methyl] indole-2,3-dione interaction. (B) 3D representation of peptide deformylase-1-[(4-fluorophenyl) methyl] indole-2,3-dione interaction.

4. Conclusion

Since the discovery of quinine from the cinchona tree, antimalarial medications have been the main tool used to treat and prevent malaria. However, the emergence of resistance in *P. falciparum* has posed significant obstacles to the creation of new medications and a vaccine. The ability of *Plasmodium* to elude antimalarial medications and the human immune system is a key hurdle to the development of effective treatments. Although there are now effective medications available to treat *P. falciparum* malaria, further research is necessary due to the possibility of resistance. Nonetheless, the risk of resistance necessitates continued investigation into novel therapeutic targets and enhanced approaches to malaria therapy and prevention. The fight against malaria requires not only the development of drugs but also the implementation of vector control measures such as bed nets, IRS, and mosquito control. Gaining insight into the biology and genetics of *Plasmodium* could lead to the development of innovative methods for managing this debilitating illness.

The search for novel chemicals that function as ligands to attach to the target gene of *P. falciparum* is crucial in reducing the complexity associated with antimalarial drug resistance, given the high mortality rate of malaria. The interaction between 1-[(4-fluorophenyl) methyl] indole-2,3-dione and *P. falciparum* PDF produced the lowest binding affinity of −4.3 out of the six ligands (drug leads) identified by machine learning in this study. A metric known as the binding free energy is used to quantify the binding affinity between a ligand and a target protein; a lesser binding affinity is indicative of a stronger bond with the lowest binding affinity, and 1-[(4-fluorophenyl) methyl] indole-2,3-dione is simply more competitive than other drug leads, *P. falciparum* PDF function more successfully. See Table 2 for the binding affinities of the ligands.

In summary, as PDF is a desirable target for novel medications, more study is required to develop 1-[(4-fluorophenyl) methyl] indole-2,3-dione as a possible antimalarial therapeutic candidate. In order to prevent resistance, it's also critical to safeguard novel compounds and their use in conjunction with other antimalarial medications. Research should also be funded and conducted to discover new drug leads for the development of antimalarial drugs.

5. Recommendations

Over the past ten years, the global impact of malaria has considerably decreased, partly as a result of the widespread use of IRS, bed nets treated with insecticides, and artemisinin combination therapy for both prevention and treatment. However, the rise in resistance to antimalarial medications—both old and new—highlights the necessity of ongoing study and advancement. It is evident that the battle against malaria is far from being won.

Studying 1-[(4-fluorophenyl) methyl] indole-2,3-dione as a possible therapeutic target is essential, considering the allure of PDF as a target in the fight against *P. falciparum*, the parasite that causes malaria. It is critical to stress the importance of ongoing study as well as this compound's potential for further investigation.

In order to strengthen our defenses against malaria, funds and resources must be set aside for further research projects that seek to identify novel drug candidates for the creation of antimalarial medications. Future studies developing more potent tactics to counteract this ongoing health danger will build on the current research.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support this work are available upon reasonable request to the corresponding author.

Author Contribution Statement

Omiyale Olumakinde Charles: Conceptualization, Investigation, Supervision. **Zainab Hassan:** Methodology, Investigation. **Okunbi Favour Onasokhare:** Validation, Investigation. **Aisha Bello:** Investigation, Writing – original draft, Writing – review & editing. **Oluwafemi F. Olaiyapo:** Investigation, Resources. **Blessing Ojajuni:** Software, Investigation, Data curation. **Sulaimon Olajuwon Abdul:** Investigation, Writing – review & editing. **Roseline Boboye:** Investigation, Writing – review & editing. **Martin Egbulefu:** Formal analysis, Investigation, Data curation. **Chiziterem Jaachinma Ezeano:** Formal analysis, Investigation. **Ahmed Rufai Nurudeen:** Investigation, Visualization. **Rasaq Awosemo:** Investigation, Project administration. **Marian Mark:** Investigation, Resources.

References

[1] Cowman, A. F., Healer, J., Marapana, D., & Marsh, K. (2016). Malaria: Biology and disease. *Cell*, *167*(3), 610–624. https://doi.org/10.1016/j.cell.2016.07.055

[2] Sato, S. (2021). *Plasmodium*—A brief introduction to the parasites causing human malaria and their basic biology.

Journal of Physiological Anthropology, *40*(1), 1. https://doi.org/10.1186/s40101-020-00251-9

[3] Hadiza, I. O. (2019). *Molecular surveillance of Plasmodium falciparum chloroquine resistant transporter genes among selected secondary school students in Ilorin East LGA, Kwara State*. Master's Thesis, Kwara State University.

[4] Excelence, E. A., Farouk, J. M., Munir, M. M., Ogbodo, E. C., Shaibu, G. E., & Adeniyi, K. A. (2019). Role of antibody diversity in the management of malaria parasite invasion. *GSC Biological and Pharmaceutical Sciences*, *9*(1), 1–4. https://doi.org/10.30574/gscbps.2019.9.1.0167

[5] Spoorthi, G. (2018). *To study clinical profile of malaria with reference to hepatopathy*. Doctoral Dissertation, Rajiv Gandhi University of Health Sciences.

[6] Murugaiyan, J., Kumar, P. A., Rao, G. S., Iskandar, K., Hawser, S., Hays, J. P., . . . , & van Dongen, M. B. M. (2022). Progress in alternative strategies to combat antimicrobial resistance: Focus on antibiotics. *Antibiotics*, *11*(2), 200. https://doi.org/10.3390/antibiotics11020200

[7] Fieulaine, S., Alves de Sousa, R., Maigre, L., Hamiche, K., Alimi, M., Bolla, J. M., . . . , & Giglione, C. (2016). A unique peptide deformylase platform to rationally design and challenge novel active compounds. *Scientific Reports*, *6*(1), 35429. https://doi.org/10.1038/srep35429

[8] UniProt Consortium. (2015). UniProt: A hub for protein information. *Nucleic Acids Research*, *43*(D1), D204–D212. https://doi.org/10.1093/nar/gku989

[9] Bhardwaj, A., Kishore, S., & Pandey, D. K. (2022). Artificial intelligence in biological sciences. *Life*, *12*(9), 1430. https://doi.org/10.3390/life12091430

[10] Dhudum, R., Ganeshpurkar, A., & Pawar, A. (2024). Revolutionizing drug discovery: A comprehensive review of AI applications. *Drugs and Drug Candidates*, *3*(1), 148–171. https://doi.org/10.3390/ddc3010009

[11] Stepniewska-Dziubinska, M. M., Zielenkiewicz, P., & Siedlecki, P. (2018). Development and evaluation of a deep learning model for protein–ligand binding affinity prediction. *Bioinformatics*, *34*(21), 3666–3674. https://doi.org/10.1093/bioinformatics/bty374

[12] Colwell, L. J. (2018). Statistical and machine learning approaches to predicting protein–ligand interactions. *Current Opinion in Structural Biology*, *49*, 123–128. https://doi.org/10.1016/j.sbi.2018.01.006

[13] Li, Y., & Yang, J. (2017). Structural and sequence similarity makes a significant impact on machine-learning-based scoring functions for protein–ligand interactions. *Journal of Chemical Information and Modeling*, *57*(4), 1007–1012. https://doi.org/10.1021/acs.jcim.7b00049

[14] World Health Organization. (2023). *WHO guidelines for malaria*. Retrieved from: https://www.who.int/publications/i/item/guidelines-for-malaria

[15] Tripathi, H., Bhalerao, P., Singh, S., Arya, H., Alotaibi, B. S., Rashid, S., . . . , & Bhatt, T. K. (2023). Malaria therapeutics: Are we close enough? *Parasites & Vectors*, *16*(1), 130. https://doi.org/10.1186/s13071-023-05755-8

[16] Bonam, S. R., Rénia, L., Tadepalli, G., Bayry, J., & Kumar, H. M. S. (2021). *Plasmodium falciparum* malaria vaccines and vaccine adjuvants. *Vaccines*, *9*(10), 1072. https://doi.org/10.3390/vaccines9101072

[17] Anyanwu, P. E., Fulton, J., Evans, E., & Paget, T. (2017). Exploring the role of socioeconomic factors in the development and spread of anti-malarial drug resistance: A qualitative study. *Malaria Journal*, *16*, 1–15. https://doi.org/10.1186/s12936-017-1849-1

[18] Weiss, D. J., Lucas, T. C., Nguyen, M., Nandi, A. K., Bisanzio, D., Battle, K. E., . . . , & Gething, P. W. (2019). Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum*, 2000–17: A spatial and temporal modelling study. *The Lancet*, *394*(10195), 322–331. https://doi.org/10.1016/S0140-6736(19)31097-9

[19] World Health Organization. (2019). *World malaria report 2019*. Retrieved from: https://www.who.int/publications/i/item/9789241565721

[20] RTS,S Clinical Trials Partnership (2015). Efficacy and safety of RTS, S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: Final results of a phase 3, individually randomised, controlled trial. *The Lancet*, *386*(9988), 31–45. https://doi.org/10.1016/S0140-6736(15)60721-8

[21] Heller, L. E., & Roepe, P. D. (2019). Artemisinin-based antimalarial drug therapy: Molecular pharmacology and evolving resistance. *Tropical Medicine and Infectious Disease*, *4*(2), 89. https://doi.org/10.3390/tropicalmed4020089

[22] Siddiqui, F. A., Liang, X., & Cui, L. (2021). *Plasmodium falciparum* resistance to ACTs: Emergence, mechanisms, and outlook. *International Journal for Parasitology: Drugs and Drug Resistance*, *16*, 102–118. https://doi.org/10.1016/j.ijpddr.2021.05.007

[23] Meinnel, T., & Giglione, C. (2022). N-terminal modifications, the associated processing machinery, and their evolution in plastid-containing organisms. *Journal of Experimental Botany*, *73*(18), 6013–6033. https://doi.org/10.1093/jxb/erac290

[24] Pei, D. (2001). Peptide deformylase: A target for novel antibiotics? *Emerging Therapeutic Targets*, *5*(1), 23–40. https://doi.org/10.1517/14728222.5.1.23

[25] Chen, D. Z., Patel, D. V., Hackbarth, C. J., Wang, W., Dreyer, G., Young, D. C., . . . , & Yuan, Z. (2000). Actinonin, a naturally occurring antibacterial agent, is a potent deformylase inhibitor. *Biochemistry*, *39*(6), 1256–1262. https://doi.org/10.1021/bi992245y

[26] Campanini-Salinas, J., Andrades-Lagos, J., Mella-Raipan, J., & Vasquez-Velasquez, D. (2018). Novel classes of antibacterial drugs in clinical development, a hope in a post-antibiotic era. *Current Topics in Medicinal Chemistry*, *18*(14), 1188–1202. https://doi.org/10.2174/1568026618666180816162846

[27] Clements, J. M., Beckett, R. P., Brown, A., Catlin, G., Lobell, M., Palan, S., . . . , & Hunter, M. G. (2001). Antibiotic activity and characterization of BB-3497, a novel peptide deformylase inhibitor. *Antimicrobial Agents and Chemotherapy*, *45*(2), 563–570. https://doi.org/10.1128/aac.45.2.563-570.2001

[28] Bracchi-Ricard, V., Nguyen, K. T., Zhou, Y., Rajagopalan, P. R., Chakrabarti, D., & Pei, D. (2001). Characterization of an eukaryotic peptide deformylase from *Plasmodium falciparum*. *Archives of Biochemistry and Biophysics*, *396*(2), 162–170. https://doi.org/10.1006/abbi.2001.2631

How to Cite: Charles, O. O., Hassan, Z., Onasokhare, O. F., Bello, A., Olaiyapo, O. F., Ojajuni, B., . . . , & Mark, M. (2025). Machine Learning Predicts 1-[(4-Fluorophenyl) Methyl] Indole-2,3-Dione as Drug Lead for Peptide Deformylase in *Plasmodium falciparum*. *Medinformatics*, *2*(1), 1–9. https://doi.org/10.47852/bonviewMEDIN42022318

Supplementary Information

The supplementary tables provided here offer an extensive overview of the molecular descriptors of the ligands examined in this study. These supplementary data are intended to support the main findings by offering additional context and details that enhance the understanding of the ligand properties and their implications in the study. Researchers and readers are encouraged to refer to these supplementary tables for a comprehensive analysis and to gain deeper insights into the molecular characteristics that underpin the main results presented in the manuscript.

Table S1. Molecular descriptors of ligands

Parameters	Ligand 1	Ligand 2
CID	1645097	2930041
IUPAC name	1-[(4-fluorophenyl) methyl] indole-2,3-dione	5-chloro-1-(2-phenylethyl) indole-2,3-dione
Molecular weight (G/Ml)	255.24	285.72
Heavy atom count	19	20
XLogP	2.2	3.3
Complexity	376	392
Hydrogen bond acceptor counts	3	2
Monoisotopic mass (Da)	255.0696	285.0557
Rotatable bond counts	2	3
TA/SA (A^2)	37.4	37.4

This table lists various molecular descriptors of the ligands studied. These descriptors include properties such as molecular weight, logP, hydrogen bond donors, and acceptors, among others. The data provided here support the main findings presented in the Results section, offering additional details that are crucial for in-depth analysis but are supplementary to the main text.

Table S2. Molecular descriptors of ligands

Parameters	Ligand 5	Ligand 6
CID	61356610	61356832
IUPAC name	7-bromo-1-(thiophen-3-ylmethyl) indole-2,3-dione	7-bromo-1-(thiophen-2-ylmethyl) indole-2,3-dione
Molecular weight (G/Ml)	322.18	322.18
Heavy atom count	18	18
XLogP	2.8	2.8
Complexity	376	376
Acceptors of hydrogen bond counts	3	3
Mass of monoisotopes (Da)	320.9459	320.9459
Rotatable bond counts	2	2
TA/SA (A^2)	65.6	65.6

This table is a continuation of Table S1, providing further molecular descriptors of the ligands. The extended dataset includes additional ligands and their properties, allowing for a comprehensive understanding of the ligand characteristics. These details complement the primary data and are essential for a thorough interpretation of the study's results.