

# Identification of New Candidate Inhibitors Able to Prevent Erythrocyte Invasion in Malaria by Drug Screening

*Sıtma Hastalığında Eritrosit İstilasını Engelleyebilen Yeni Aday İnhibitorlerin İlaç Taraması ile Belirlenmesi*

✉ Ahmet Burak Doğanoglu<sup>1</sup>, ✉ Vildan Enisoğlu Atalay<sup>2</sup>

<sup>1</sup>Üsküdar University Faculty of Medicine, Department of Bioinformatics, İstanbul, Türkiye

<sup>2</sup>İstanbul Technical University, Graduate School of Science/Bioinformatics, İstanbul, Türkiye

Cite this article as: Doğanoglu AB, Enisoğlu Atalay V. Identification of new candidate inhibitors able to prevent erythrocyte invasion in malaria by drug screening. Türkiye Parazit Derg. [Epub Ahead of Print]

## ABSTRACT

**Objective:** In Nowadays malaria still remains the parasitic disease causing the highest number of deaths, accounting for 619,000 fatalities. The *Plasmodium* parasites that cause malaria have separate life cycles in both humans and female Anopheles mosquitoes, existing in various forms throughout this process. The main reason for observing the disease is that merozoites sustain their existence by invading erythrocytes. Existing drugs affect the parasite's ability to digest hemoglobin. Drug resistance is also involved in this process. In this study, have been focused to develop new drug candidate molecules for evade drug resistance. To evade drug resistance, the aim was to prevent merozoites from invading erythrocytes.

**Methods:** The invasion of merozoites into erythrocytes consists of several stages: Attachment, deformation, apical junction formation, and tight junction formation. For this purpose, the docking calculations have been done between the invasion proteins such as MSP1, pVDBP, pRHS, AMA1 and candidates. The candidates obtained from the malaria box set were subjected to conformational scanning and geometry optimization in the Spartan'14 program to determine their physicochemical properties. According to the obtained results from the AutoDock Vina and multiple regression analyses were conducted for each protein to examine the relationship between binding affinities and the calculated physicochemical parameters of the candidates.

**Results:** In the regression study of 200 molecules examined for 4 different proteins, 108 molecules were included for DBP, 96 for MSP1, 90 for AMA1 and 96 for RHS, and 21 common molecules were observed for all proteins.

**Conclusion:** Twenty-one molecules showed correlation with the proteins studied. Among these molecules, MMV019074, MMV019662 and MMV665881 were suggested as candidate drug leads in terms of their binding affinities, physicochemical properties and SwissADME values.

**Keywords:** Docking, erythrocytes, malaria, malaria box, merozoite

## ÖZ

**Amaç:** Günümüzde sıtma, 619.000 ölümle en çok ölüme neden olan paraziter hastalık olmaya devam etmektedir. Sıtmaya neden olan *Plasmodium* parazitlerinin hem insanlarda hem de dişi Anopheles sivrisineklerinde ayrı yaşam döngüleri vardır ve bu süreç boyunca çeşitli biçimlerde var olurlar. Hastalığın gözlemlenmesinin temel nedeni, merozoitlerin eritrositleri istila ederek varlıklarını sürdürmeleridir. Mevcut ilaçlar parazitin hemoglobini sindirme yeteneğini etkiler. İlaç direnci de bu süreçte yer alır. Bu çalışmada, ilaç direncinden kaçınmak için yeni ilaç adayları molekülleri geliştirmeye odaklanılmıştır. İlaç direncinden kaçınmak için amaç, merozoitlerin eritrositleri istila etmesini önlemektir.

**Yöntemler:** Merozoitlerin eritrositlere invazyonu birkaç aşamadan oluşur: Tutunma, deformasyon, apikal bağlantı oluşumu ve sıkı bağlantı oluşumu. Bu amaçla, MSP1, pVDBP, pRHS, AMA1 gibi istilacı proteinler ile adaylar arasında kenetleme hesaplamaları yapılmıştır. Malaria box setinden elde edilen adaylar, fizikokimyasal özelliklerini belirlemek için Spartan'14 programında konformasyonel tarama ve geometri optimizasyonuna tabi tutulmuştur. AutoDock Vina'dan elde edilen sonuçlara göre, her protein için bağlanma afiniteleri ile adayların hesaplanan fizikokimyasal parametreleri arasındaki ilişkiyi incelemek için çoklu regresyon analizleri yapıldı.

**Bulgular:** Dört farklı protein için incelenen 200 molekülün regresyon çalışmasında DBP için 108, MSP1 için 96, AMA1 için 90 ve RHS için 96 molekül dahil edilmiş olup tüm proteinler için 21 tane ortak molekül olduğu gözlemlenmiştir.

Received/Geliş Tarihi: 11.10.2023 Accepted/Kabul Tarihi: 24.03.2025 Epub: 29.04.2025

**Address for Correspondence/Yazar Adresi:** Vildan Enisoğlu Atalay, İstanbul Technical University, Department of Computational Science and Engineering, İstanbul, Türkiye

**E-mail/E-Posta:** enisogluatalayv@itu.edu.tr **ORCID ID:** orcid.org/0000-0002-9830-9158



**Sonuç:** Yirmi bir molekül çalışılan proteinlerle korelasyon gösterdi. Bu moleküller arasında, bağlanma afiniteleri, fizikokimyasal özellikleri ve SwissADME değerleri açısından MMV019074, MMV019662 ve MMV665881 aday ilaç öncülleri olarak önerildi.

**Anahtar Kelimeler:** Kenetleme, eritrosit, sıtma, malaria box, merozoit

## INTRODUCTION

Malaria is an acute febrile disease caused by *Plasmodium* parasites transmitted by female Anopheles mosquitoes. The life cycle of the parasites starts when sporozoites are injected into the skin by female Anopheles mosquitoes. Due to environmental changes and immune factors, sporozoites need to quickly reach the liver (1). During this process, sporozoites utilize surface proteins like circumsporozoite protein and thrombospondin-related anonymous protein to pass through Kupffer macrophage cells, which are the first obstacle they encounter, and invade hepatocytes (2,3). After sporozoites transform into schizonts within hepatocytes, the schizonts release merozoites into the bloodstream by rupturing the hepatocyte.

After merozoites invade erythrocytes, it leads to the appearance of the disease's symptoms. The invasion process begins with the merozoite surface protein 1 (MSP1) attachment to the erythrocyte surface. Moreover, the MSP1 proteins abundantly present on the surface are highly polymorphic due to being exposed to a large number of antibodies (4). In the next stage, erythrocyte binding antigen (EBA) proteins and reticulocyte binding protein homologs (RH) move the merozoite towards erythrocytes. Thus, facilitating contact between the merozoite apex and the erythrocyte. The apex of the merozoite then attaches to the erythrocyte by the pRHS protein and becoming ready for invasion. Rhoptry neck protein 2 (RON2) is released from the rhoptry secretory organelles and binds to the erythrocyte surface. This allows Apical Membrane Antigen 1 (AMA1) protein to attach to the erythrocyte. The merozoite establishes a mobile tight junction with the AMA1-RON2 complex, to infect the erythrocyte (5). The stages of erythrocyte invasion are shown in Figure 1.

Merozoites differentiate into male and female gametocytes by invading erythrocytes. During erythrocyte invasion, not all merozoites differentiate into gametocytes; some of them revert back to merozoites after erythrocyte invasion. Once gametocytes are formed, the human life cycle comes to an end. When an infected individual is bitten by a mosquito again, the gametocytes are transferred to the mosquito (6). After erythrocyte invasion, the parasites sustain their lives by utilizing hemoglobin digestion within the erythrocyte. Trophozoites take up hemoglobin into the cytoplasm through endocytosis, creating a digestive vacuole with a pH of 5.2. Heme molecules produced from hemoglobin breakdown are detoxified inside this vacuole because of their toxic nature (7). All drugs that affect the current blood stage are

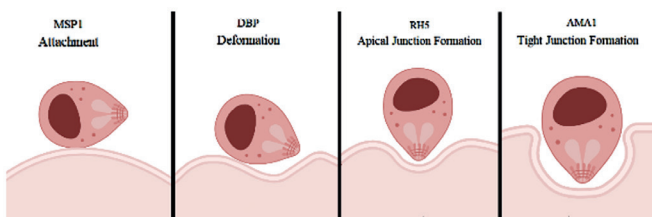
used to prevent the detoxification of heme molecules, and all resistance mechanisms develop along this pathway. Merozoites can compensate for any issues that may arise during erythrocyte invasion by establishing alternative pathways (8).

Therefore, this study, to target merozoite erythrocyte invasion by inhibiting pvDBP, MSP1, AMA1, and pRHS proteins to evade resistance mechanisms. It has been demonstrated that the C-terminal region of the MSP1 protein (MSP1 19) binds to and inhibits the protein by targeting amino acids 29-36 and 51-56 through a ligand called NIC (9,10). In this context, the same region has been selected as the target binding sequence. The EBL protein family, which only has one member called the DBP protein in *P. vivax*, utilizes Duffy binding-like domains to attach to erythrocyte receptors (11). To determine the compatibility of the binding region of pvDBP with other proteins, sequence alignments were performed between erythrocyte membrane protein 1 (EMP1), EBA175 (12), EBA140 (13), and pvDBP. The DBP-Duffy Antigen Receptor (DARC) complex was found to be the preferred option (the aligned proteins are 4NUU\_A, 1ZRO\_A, AUF72533, 4GF2\_A, and SBT76215). According to Batchelor et al. (11), the binding site of the PhRH5 protein involved in the RH5-basigin interaction, which is responsible for binding, has been identified as the active region. The protein has a pdb code of 4WAT (14). The active region amino acids for the AMA1 protein have been selected as the ones that interact with the AMA1-R1 inhibitor (15) and the AMA1-pfRON2 (16) complex in the 3SRI structure (17). Docking was done with these proteins and molecules in the "Malaria box" molecule set published by the Medicines for Malaria Venture Foundation. Thus, the correlation between the physicochemical properties of molecules and their binding affinities was studied. This novel approach provides a new drug and treatment pathway.

Additionally, the MMV020291 molecule from the study of Dans et al. (18) was preferred to compare binding affinities with experimental data. MMV020291, a molecule from the "Pathogen box" set published by the Medicines for Malaria Venture Foundation. It was observed that MMV020291 was in the deformation and tight junction stages specified in Figure 1. For this reason, the binding affinities of MMV020291 were accepted as the threshold value for the molecules to be recommended.

## METHODS

As a target proteins pvDBP, MSP1, AMA1, and pRHS were selected, which play a crucial role in the erythrocyte invasion mechanism mentioned above, have been identified for the binding procedures. The crystal structures of the these proteins were selected from the RCSB Protein Data Bank (19) as follows: pdb id: 4NUU, 1B9W, 3SRI and 4WAT. The central coordinates of the active region of the selected protein structures for the binding procedures are specified in Table 1, and the grid box has been scaled to 35x35x35Å<sup>3</sup>. The binding studies were performed using the AutoDock Vina program (20), and the interaction maps and visualization processes were carried out using the BIOVIA Discovery Studio Visualizer program (21).



**Figure 1.** Erythrocyte invasion stages of merozoite (created with BioRender.com)

**Table 1.** The grid box coordinates

PDB ID	X	Y	Z
4NUU	47.655875	-44.833875	98.381875
1B9W	-3.335750	20.720000	53.636500
3SRI	9.933917	-1.080167	3.548250
4WAT	46.476750	38.165625	51.009750

The determine the binding and inhibition constants of new candidate drug molecules for treating malaria are crucial. To achieve this, 200 molecules were selected from the “Malaria box” molecules that target the blood stage (22). The chosen molecules went through a process of conformational searching using the molecular mechanics force field (MMFF) method, with single bonds being rotated in 60° increments in the Spartan’14 program (23) to obtain their most stable conformer structures. The molecules with stable conformers then underwent geometry optimizations using the semi-empirical PM6 method. Additionally, the drug-likeness potential of the candidate molecules was investigated using the SwissADME program (24). Their compliance with Lipinski’s rules (25) was also assessed in the study.

### Statistical Analysis

Linear correlations between binding affinities against physicochemical parameter values such as lipophilicity (LogP), dipole moment, molecular weight ( $M_w$ ), volume (V), number of rotatable bonds (NRB) and topological polar surface area (TPSA) calculated as a result of geometry optimizations were determined using the Excel program (26).

## RESULT

In this context, new candidate drug molecules targeting the 4WAT, 4NUU, 3SRI, and 1B9W pdb coded structures responsible for erythrocyte invasion in malaria treatment’s blood stage were identified using computational methods. Table 2 displays the molecular weight ( $M_w$ ), TPSA, NRB, LogP, dipole moment ( $\mu$ ), and hardness [ $\eta = (E_{\text{HOMO}} - E_{\text{LUMO}})/2$ ] values, and SwissADME maps for 21 molecules that displayed correlation with all the proteins among the chosen 200 molecules (calculated parameters and binding affinity values for 200 molecules are shown in Supplementary Tables 1-3). The table also includes values for the MMV020291 structure. In the SwissADME maps that provide the initial criteria for drug-likeness in a collective manner, it was observed that only the compound with the number MMV665881 exceeded the limit value at the unsaturation point. However, all the other suggested candidate molecules meet Lipinski’s drug-likeness criteria.

In this study, different affinity values were observed based on the QSAR parameters of the examined molecules. The binding affinity values for the reference structure MMV020291 were obtained as -7.1 and -7.7 kcal.mol<sup>-1</sup> for the protein structures with pdb codes 4NUU and 3SRI, respectively. On the other hand, experimental data such as toxicity and inhibition were examined for the 200 molecules. In the binding calculations between the 4NUU pdb coded protein structure and the molecules, the binding affinities of MMV665881, MMV019074, and MMV019662 were calculated as -7.8, -8.2, and -8.4 kcal.mol<sup>-1</sup>, respectively. For the 3SRI protein structure, the binding affinities of these molecules were calculated as -9.0, -8.2, and -8.9 kcal.mol<sup>-1</sup>, respectively. Based on the obtained binding energies, it can be observed

that the increase in affinity compared to the reference molecule MMV020291 is approximately 3 to 20 times for the 3SRI protein structure and between 5 to 20 times for the 4NUU structure.

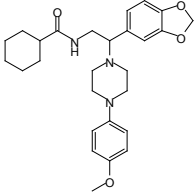
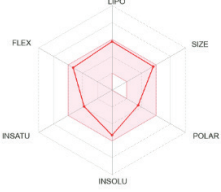
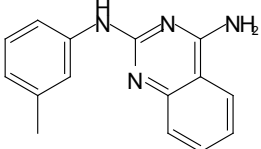
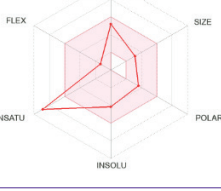
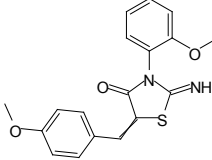
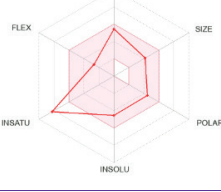
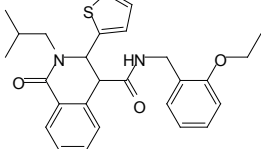
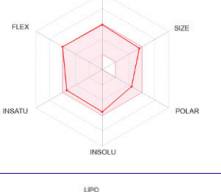
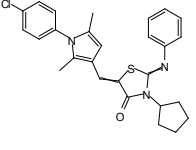
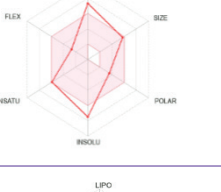
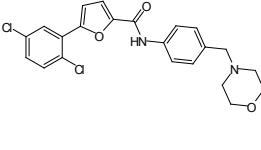
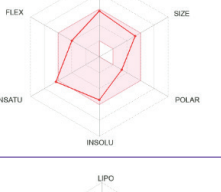
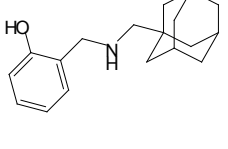
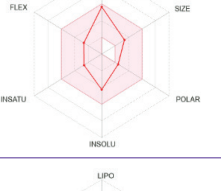
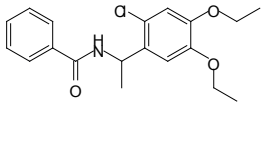
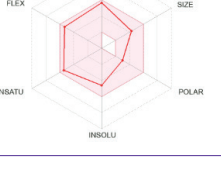
In the conducted experimental studies, inhibition could not be achieved with the 1B9W and 4WAT pdb-coded structures, which are defined as the first and third stages of erythrocyte invasion stages, respectively (attachment and apical junction formation). However, in the binding studies, affinity values of -6.4 and -7.0 kcal.mol<sup>-1</sup> were obtained for these structures with the MMV020291 reference molecule. On the other hand, it was determined that the increase in binding affinity for the prominent candidate structures MMV665881, MMV019074, and MMV019662 from the examined molecule set ranged from 2 to 8 times.

In the literature, experimental inhibition tests were conducted by Voorhis et al. (22). When Table 3 is examined, it can be observed that the structures MMV019662 and MMV019074 provided inhibition in both early and late ring stages for the HB3 strain. On the other hand, the structure MMV665881 inhibited subsequent steps instead of the targeted inhibition steps. The selected molecules exhibited nearly complete inhibition in the *Plasmodium* 3D7 strain, where no drug resistance was observed. However, achieving our main goal, these molecules particularly MMV019662, MMV665881, and MMV019074, demonstrated over 70% inhibition in the *Plasmodium* DD2 strain, which is known for chloroquine resistance, indicating their potential to evade resistance mechanisms.

When the physicochemical parameters of the candidate molecules listed in Table 2 were examined, a clear trend between calculated logP, hardness, and TPSA values and binding affinities could not be observed. However, a more distinct trend was obtained with the dipole moment ( $\mu$ ) and molecular weight (MW) properties. It is observed that the logP value, which has an effect on binding affinity, varies between 0.88 for the reference MMV020291 structure and in the range of (-1.41) - (+1.86) for the other prominent molecules. According to this, a very sharp change between logP values and binding affinity has not been observed. On the other hand, decreased dipole moment ( $\mu$ ) has been shown to impact binding affinity. The structure MMV019662, which showed the highest affinity, has a dipole moment of  $\mu=3.92$ , while the reference molecule MMV020291 has a dipole moment of  $\mu=6.17$ . The decrease in the dipole moment ( $\mu$ ) of the other molecules given in Table 2 has resulted in higher binding affinities. It’s important to mention that there is a clear correlation between higher molecular weight ( $M_w$ ) values and a stronger binding affinity.

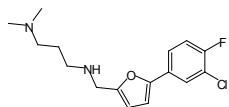
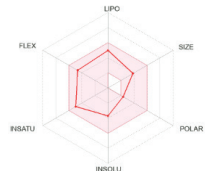
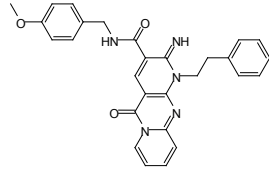
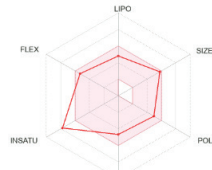
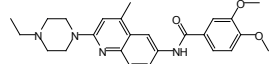
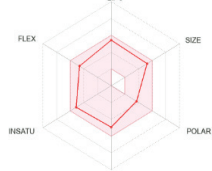
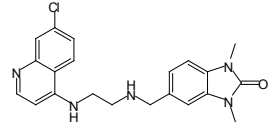
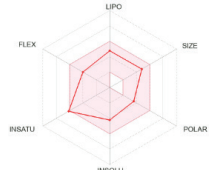
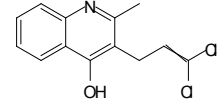
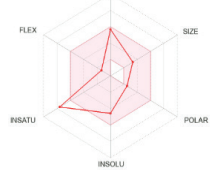
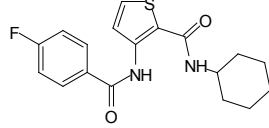
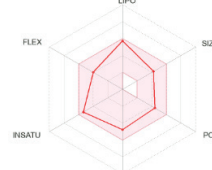
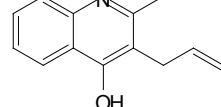
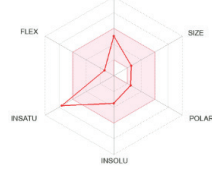
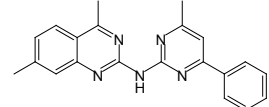
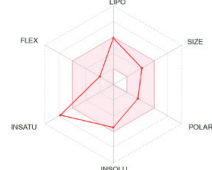
In the significant regression analyses obtained from the correlation curves performed for 6 different parameters with the calculated values given in Table 2 (using Microsoft Excel), a total of 107, 96, 90, and 96 candidate molecules have been included in the regression curves for the 4NUU, 1B9W, 3SRI, and 4WAT pdb-coded structures, respectively (molecules included in the regressions are shown in Supplementary Tables 4-7). The results are provided in Table 4 below. Furthermore, when the reference structure MMV020291 was included in the multiple regression analysis, no correlation was observed. In order to explain the binding affinities and identify important amino acids, interaction maps with the active site amino acids were generated for the candidate molecules MMV019662, MMV665881, MMV019074, and the reference structure MMV020291. The Table 5 provides

**Table 2.** 2D plots, SwissAdme maps, calculated logP,  $\mu$ ,  $\eta$ , MW, NRB, TPSA values and binding affinities (kcal.mol<sup>-1</sup>) of examined structures for 4WAT, 4NUU, 3SRI and 1B9W protein structures

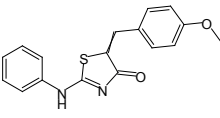
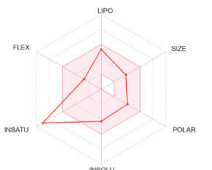
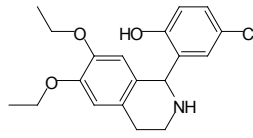
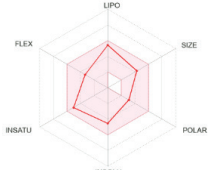
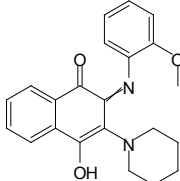
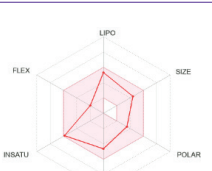
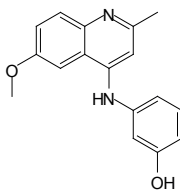
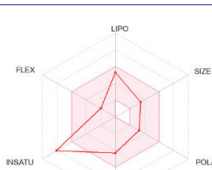
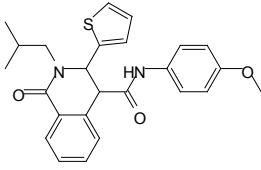
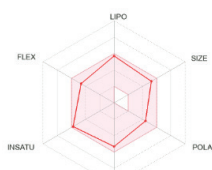
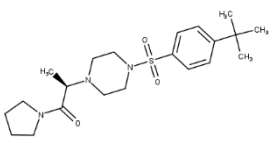
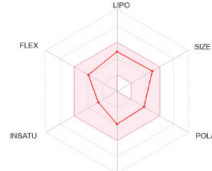
MMV019662			logP	$\mu$	$\eta$	MW	NRB
			-0.49	3.92	-3.87	465.58	8
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			63.27	-8.4	-7.3	-8.9	-7.1
MMV019110			logP	$\mu$	$\eta$	MW	NRB
			0.59	1.85	-3.77	286.76	2
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			63.83	-6.8	-7.1	-6.9	-6.0
MMV666691			logP	$\mu$	$\eta$	MW	NRB
			-0.32	4.22	-3.98	340.4	4
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			87.92	-6.8	-6.3	-7.8	-6.6
MMV000653			logP	$\mu$	$\eta$	MW	NRB
			1.09	1.49	-4.21	462.6	9
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			86.88	-7.1	-7.1	-8.6	-7.0
MMV666688			logP	$\mu$	$\eta$	MW	NRB
			1.86	5.09	-3.81	476.03	4
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			62.9	-8.2	-7.4	-8.6	-7.3
MMV020651			logP	$\mu$	$\eta$	MW	NRB
			-1.19	4.92	-3.53	431.31	6
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			54.71	-7.6	-7.5	-7.8	-6.4
MMV020788			logP	$\mu$	$\eta$	MW	NRB
			2.23	3.36	-4.49	307.86	4
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			32.26	-6.7	-6.5	-7.8	-5.8
MMV665800			logP	$\mu$	$\eta$	MW	NRB
			0.22	0.96	-4.19	347.84	8
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			47.56	-6.8	-6.7	-7.0	-6.6



**Table 2.** Continued

MMV020505			logP	$\mu$	$\eta$	MW	NRB
			-0.6	1.17	-3.83	347.26	7
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			28.41	-6.6	-7.0	-6.9	-6.1
MMV665881			logP	$\mu$	$\eta$	MW	NRB
			1.2	5.82	-3.66	479.53	8
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			101.48	-7.8	-7.9	-9.0	-7.3
MMV019074			logP	$\mu$	$\eta$	MW	NRB
			-1.41	3.06	-3.61	434.53	7
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			66.93	-8.2	-7.3	-8.2	-6.8
MMV665875			logP	$\mu$	$\eta$	MW	NRB
			-1.16	7.52	-3.99	432.35	6
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			63.88	-7.1	-7.3	-7.8	-6.6
MMV665994			logP	$\mu$	$\eta$	MW	NRB
			1.47	6.37	-4.18	268.14	2
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			33.12	-6.4	-6.7	-6.4	-5.7
MMV665939			logP	$\mu$	$\eta$	MW	NRB
			0.19	3.24	-3.87	346.42	6
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			86.44	-7	-6.6	-7.7	-6.8
MMV006913			logP	$\mu$	$\eta$	MW	NRB
			0.8	6.05	-4.16	199.25	2
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			33.12	-5.2	-6.4	-6.2	-5.6
MMV007041			logP	$\mu$	$\eta$	MW	NRB
			3.97	1.99	-3.98	341.41	3
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			63.59	-7.6	-7.3	-7.9	-6.5

**Table 2.** Continued

MMV665836			logP	$\mu$	$\eta$	MW	NRB
			0.66	3	-3.85	310.37	3
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			75.99	-6.6	-6.6	-8.3	-6.1
MMV000478			logP	$\mu$	$\eta$	MW	NRB
			-1.11	3.26	-4.2	347.84	5
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			50.72	-6.2	-6.6	-6.4	-6.0
MMV085203			logP	$\mu$	$\eta$	MW	NRB
			-0.12	4.81	-3.57	362.42	4
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			58.64	-6.9	-7.7	-7.6	-6.6
MMV008212			logP	$\mu$	$\eta$	MW	NRB
			-1.06	2.42	-3.82	317.78	3
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			54.38	-6.5	-7.1	-7.2	-6.2
MMV000662			logP	$\mu$	$\eta$	MW	NRB
			0.3	1.58	-3.86	434.55	7
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			86.88	-7.2	-6.8	-8.6	-6.8
MMV020291			logP	$\mu$	$\eta$	MW	NRB
			0.88	6.17	-4.085	407.57	6
			<b>TPSA</b>	<b>4NUU</b>	<b>1B9W</b>	<b>3SRI</b>	<b>4WAT</b>
			69.31	-7.1	-7	-7.7	-6.4

NRB: Number of rotatable bonds, TPSA: Topological polar surface area

information on which amino acids are involved in the closest and strongest interactions. When the interactions between the molecules and the 4NUU protein were examined, a relationship was observed between the binding energies, the number of amino acids involved in the interaction, and the distances of the interactions. Particularly, the hydrogen bond distance with amino acid R368 was obtained at a very good level with a distance of 2.49 Å. In addition, for the 4NUU protein, it has been observed that the active site amino acids A360 and Y363 have  $\pi$ -alkyl, alkyl, and hydrophobic interactions with all the molecules studied at distances ranging from 4.79 to 5.36 Å and 4.30 to 5.36 Å, respectively. Upon examination of the interactions obtained for the 1B9W structure, it was found that only amino acid R29 was repeated with all the selected molecules. Conversely, it has been reported in the literature that the reference structure MMV020291 molecule does

not experimentally interact with 1B9W. In this regard, the lengths of the interactions and the affinity value of  $-7.0 \text{ kcal.mol}^{-1}$  support the experimental data. It has been observed that the MMV665881 candidate structure has the best binding affinity with 1B9W with a value of  $-7.9 \text{ kcal.mol}^{-1}$ , the closest hydrogen bond distance of 3.09 Å, and the highest number of amino acid interactions. Upon analyzing the outcomes, for the 3SRI protein, it was determined as the structure with the highest number of amino acid interactions and consequently the best binding affinities. Despite the lower number of amino acids involved in the interaction for the MMV665881 structure, the obtained binding affinity value of  $-9.0 \text{ kcal.mol}^{-1}$ , which is the lowest, was interpreted as having two hydrogen bonds at distances of 2.95 and 3.04 Å. However, despite MMV019074 structure having three hydrogen bonds, its binding affinity is approximately reduced by 6 times compared to the MMV665881 structure due to the fact that

**Table 3.** Staged inhibition assay

	Molecules		
	MMV019662	MMV665881	MMV019074
DD2 strain, asexual blood stage inhibition (10 $\mu$ M)	71%	70%	80%
3D7 strain, asexual blood stage inhibition (5 $\mu$ M)	97%	91%	99%
HB3 strain, early ring (EH) stage	+	GH	+
HB3 strain, late ring (GH) stage	+	T	+
HB3 strain, trophozoite (T) stage	+	S	-
HB3 strain, schizont (S) stage	-	-	-
HB3 strain, merozoite (M) stage	-	-	-
HB3 strain, viability <50% (10 $\mu$ M)	+	+	+

**Table 4.** Multiple regression analysis results

PDB ID	Multiple R	R <sup>2</sup>	Adjustable R <sup>2</sup>	Observation
4NUU	0.936118	0.87632	0.86890	107
1B9W	0.904718	0.81851	0.80628	96
3SRI	0.917513	0.84183	0.83039	90
4WAT	0.905308	0.81958	0.80742	96

**Table 5.** Molecule-amino acid interaction map (blue, pink, purple, green, brown and yellow colors indicate Hydrophobic, Alkyl,  $\pi$ -Alkyl, H-Bond,  $\pi$ -Cation and  $\pi$ -Anion interactions respectively)

Molecule	4NUU							1B9W				
	R274	I277	Y278	A281	A360	Y363	R368	R29	C30	E36	K52	G55
MMV019662	5.10	4.24	5.21	4.08	4.94	4.30	2.49	5.00		4.74	4.72	
MMV665881		4.25		3.51	5.36	4.76	3.00	3.09	4.60	4.32	4.93	
MMV019074		5.15	5.43	4.93	4.79	4.76	2.86	4.46				3.40
MMV020291	2.75		3.74		5.16	5.36		4.69	5.10			3.60
Molecule	3SRI											
	L131	V137	Y142	P226	Y234	Y236	Y251	A254	P350			
MMV019662		4.95		4.91	5.41		2.79	4.02	4.03			
MMV665881		4.36	4.42		2.95			3.04				
MMV019074	4.28	3.75	3.61	3.04	3.73	3.41	3.21	3.68				
MMV020291	5.43						5.18		5.45			
Molecule	4WAT											
	S197	Y200	F350	N352	N354	R357	W447	T449				
MMV019662					3.23		4.81					
MMV665881			5.42				4.19	3.4				
MMV019074	3.65	3.8			3.34	4.57						
MMV020291				2.81								

these hydrogen bonds are at more distant distances. On the other hand, the fact that amino acid Y251 forms 2 hydrogen bonds indicates its effectiveness in the active site. When examining the interaction types and bond distances identified for 4WAT, it is observed that the MMV665881 molecule forms a hydrogen bond with amino acid T449 at a distance of 3.4 Å, and  $\pi$ -alkyl interactions with amino acids W447 and F350, resulting in the most stable binding affinity for the respective protein with a value

of -7.3 kcal.mol<sup>-1</sup>. The MMV019662 molecule, on the other hand, forms one hydrogen bond and one  $\pi$ -alkyl interaction, resulting in a binding affinity of -7.1 kcal.mol<sup>-1</sup>. For the 4WAT structure, the experimentally inactive MMV020291 molecule, which was selected as the reference structure, forms only one hydrogen bond with amino acid N352, resulting in the weakest binding energy with a value of -6.4 kcal.mol<sup>-1</sup>.

## DISCUSSION

In this study, the values of physicochemical parameters, radar plots, and binding affinities obtained from the docking processes with the 4WAT, 4NUU, 3SRI, and 1B9W PDB coded structures, which are responsible for erythrocyte invasion in malaria disease, are given in Table 2 above. In the regression analysis conducted for the 200 molecules on the 4 different proteins, 107 molecules were included for DBP (PDB id: 4NUU), 96 molecules for MSP1 (PDB id: 1B9W), 90 molecules for AMA1 (PDB id: 3SRI), and 96 molecules for RH5 (PDB id: 4WAT). It was observed that there are 21 common molecules (from Table 2) for all proteins. The results obtained from multiple regression analyses for each protein are provided in Table 6 below. The findings in Table 6 can serve as guiding factors for ligand design. These results have the potential to offer valuable insights for the design of molecules in future studies.

In the study conducted by Dans et al. (18), it was observed that the molecule with the code MMV020291 found in the “Pathogen box” inhibits the deformation and tight junction stages of erythrocyte invasion. Therefore, the binding affinities of MMV020291 with the 4NUU and 3SRI proteins involved in these stages were considered as the ideal binding affinities. At the same time, since it did not bind experimentally in the stages where 1B9W and 4WAT were included, the binding affinities obtained here were determined as the values that needed to be exceeded. MMV019074, MMV019662, and MMV665881 exceeded these threshold values and were recommended based on their physicochemical parameters and SwissADME values. On

the other hand, Voorhis et al. (22) experimentally observed that the structures MMV019662 and MMV019074 provide early and late ring stage inhibitions for the HB3 strain, while the structure MMV665881 inhibits the next steps. After merozoites invade the erythrocytes, ring, trophozoite, schizont and merozoite formation occurs, respectively. Since the aim of this study is to prevent erythrocyte invasion, it is very important to observe inhibition especially in the ring stages in the tests mentioned here. Therefore, MMV019662 and MMV019074 stand out slightly over MMV665881 according to inhibition tests.

## CONCLUSION

In conclusion, 200 molecules were screened for their potential to inhibit erythrocyte invasion. It was found that 21 of them could be effective in 4 different stages of erythrocyte invasion. When the threshold values determined for the reference molecule known to work in the experiment and the inhibition tests of the screened molecules are examined, MMV019074, MMV019662 and MMV665881 stand out. After analyzing the docking study results of the pertinent molecules, it is apparent that the modeled proteins, 4NUU and 3SRI, exhibit excellent binding affinities that align with the experimental data. These results have provided reliability to the obtained theoretical data, and the molecules MMV019074, MMV019662, and MMV665881 are suggested as potential drug candidates for future experimental stages in malaria treatment research.

**Table 6.** Affinity coefficient formulas

PDB ID	Affinity
4NUU	$-7,416 + (-0.010 * M_w) + (0.103 * NRB) + (0.041 * \mu) + (0.008 * TPSA) + (-0.066 * \log P) + (-0.752 * \eta)$
1B9W	$-9.685 + (-0.004 * M_w) + (-0.035 * NRB) + (-0.025 * \mu) + (0.009 * TPSA) + (-0.086 * \log P) + (-1.029 * \eta)$
3SRI	$-5.797 + (-0.010 * M_w) + (0.081 * NRB) + (0.039 * \mu) + (-0.004 * TPSA) + (-0.154 * \log P) + (-0.438 * \eta)$
4WAT	$-6.170 + (-0.005 * M_w) + (0.034 * NRB) + (0.051 * \mu) + (-0.004 * TPSA) + (0.042 * \log P) + (-0.323 * \eta)$



## \*Ethics

**Ethics Committee Approval:** This study did not involve human participants or animal subjects, and was entirely conducted using publicly available data and computer-based analysis; therefore, ethical approval was not required.

**Informed Consent:** Not applicable.

## Acknowledgments

Thank you to the Medicines for Malaria Venture foundation, which has made open source sets of molecules to find new drugs against malaria.

## Footnotes

## \*Authorship Contributions

Concept: A.B.D., V.E.A., Design: A.B.D., V.E.A., Data Collection or Processing: A.B.D., V.E.A., Analysis or Interpretation: A.B.D., V.E.A., Literature Search: A.B.D., V.E.A., Writing: A.B.D., V.E.A.

**Conflict of Interest:** No conflict of interest was declared by the authors.

**Financial Disclosure:** The authors declared that this study received no financial support.

## REFERENCES

1. Somer PDA. Sıtma (Malarya) ve tedavisi. *Pediatrici*. 2018; 10: 35-40.
2. Wilder BK, Vigdorovich V, Carbonetti S, Minkah N, Hertoghs N, Raappana A, et al. Anti-TRAP/SSP2 monoclonal antibodies can inhibit sporozoite infection and may enhance protection of anti-CSP monoclonal antibodies. *NPJ Vaccines*. 2022; 7: 58.
3. Lu C, Song G, Beale K, Yan J, Garst E, Feng J, et al. Design and assessment of TRAP-CSP fusion antigens as effective malaria vaccines. *PLoS One*. 2020; 15: e0216260.
4. Putaporntip C, Kuamsab N, Rojrung R, Seethamchai S, Jongwutiwes S. Structural organization and sequence diversity of the complete nucleotide sequence encoding the *Plasmodium malariae* merozoite surface protein-1. *Sci Rep*. 2022; 12: 15591.
5. Wright GJ, Rayner JC. *Plasmodium falciparum* erythrocyte invasion: combining function with immune evasion. *PLoS Pathog*. 2014; 10: e1003943.
6. Garrido-Cardenas JA, González-Cerón L, Manzano-Agugliaro F, Mesa-Valle C. *Plasmodium* genomics: an approach for learning about and ending human malaria. *Parasitol Res*. 2019; 118: 1-27.
7. Wicht KJ, Mok S, Fidock DA. Molecular mechanisms of drug resistance in *Plasmodium falciparum* malaria. *Annu Rev Microbiol*. 2020; 74: 431-54.
8. Yahata K, Hart MN, Davies H, Asada M, Wassmer SC, Templeton TJ, et al. Gliding motility of *Plasmodium* merozoites. *Proc Natl Acad Sci U S A*. 2021; 118: e2114442118.
9. Chandramohanadas R, Basappa, Russell B, Liew K, Yau YH, Chong A, et al. Small molecule targeting Malaria merozoite surface protein-1 (MSP-1) prevents host invasion of divergent plasmodial species. *The J Infect Dis*. 2014; 210: 1616-26.
10. Chitarra V, Holm I, Bentley GA, Pêtres S, Longacre S. The crystal structure of C-terminal merozoite surface protein 1 at 1.8 Å resolution, a highly protective malaria vaccine candidate. *Molecular Cell*. 1999; 3: 457-64.
11. Batchelor JD, Malpede BM, Omattage NS, DeKoster GT, Henzler-Wildman KA, Tolia NH. Red blood cell invasion by *Plasmodium vivax*: structural basis for DBP engagement of DARC. *PLoS Pathog*. 2014; 10: e1003869.
12. Tolia NH, Enemark EJ, Sim BKL, Joshua-Tor L. Structural basis for the EBA-175 erythrocyte invasion pathway of the Malaria parasite *Plasmodium falciparum*. *Cell*. 2005; 122: 183-93.
13. Malpede BM, Lin DH, Tolia NH. Molecular basis for sialic acid-dependent receptor recognition by the *Plasmodium falciparum* invasion protein erythrocyte-binding antigen-140/BAEBL. *J Biol Chem*. 2013; 288: 12406-15.
14. Chen L, Xu Y, Healer J, Thompson JK, Smith BJ, Lawrence MC, et al. Crystal structure of Pfrh5, an essential P. falciparum ligand for invasion of human erythrocytes. *Elife*. 2014; 3: e04187.
15. Wang G, MacRaild CA, Mohanty B, Mobli M, Cowieson NP, Anders RF, et al. Molecular insights into the Interaction between *Plasmodium falciparum* apical membrane antigen 1 and an invasion-inhibitory peptide. *PLoS One*. 2014; 9: e109674.
16. Richard D, MacRaild CA, Riglar DT, Chan JA, Foley M, Baum J, et al. Interaction between *Plasmodium falciparum* apical membrane antigen 1 and the rophtry neck protein complex defines a key step in the erythrocyte invasion process of malaria parasites. *J Biol Chem*. 2010; 285: 14815-22.
17. Normand BVL, Tonkin ML, Lamarque MH, Langer S, Hoos S, Roques M, et al. Structural and functional insights into the Malaria parasite moving junction complex. *PLoS Pathog*. 2012; 8: e1002755.
18. Dans MG, Weiss GE, Wilson DW, Sleebs BE, Crabb BS, de Koning-Ward TF, et al. screening the medicines for Malaria venture pathogen box for invasion and egress inhibitors of the blood stage of *Plasmodium falciparum* reveals several inhibitory compounds. *Int J Parasitol*. 2020; 50: 235-52.
19. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. *Nucleic Acids Res*. 2000; 28: 235-42.
20. Trott O, Olson AJ. AutoDock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010; 31: 455-61.
21. Discovery studio modeling environment. San Diego: Dassault Systèmes: Dassault Syst'emes BIOVIA; 2016.
22. Voorhis WCV, Adams JH, Adelfio R, Ah Yong V, Akabas MH, Alano P, et al. Open source drug discovery with the malaria box compound collection for neglected diseases and beyond. *PLoS Pathogens*. 2016; 12: e1005763.
23. Spartan'14, Wavefunction, Inc., Irvine, CA. Available from: <https://www.wavefun.com/>
24. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep*. 2017; 7: 42717.
25. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. 2001; 46: 3-26.
26. Microsoft Corporation. (2018). Microsoft Excel. Available from: <https://office.microsoft.com/excel>