In Silico Pharmacology (2021) 9:38
https://doi.org/10.1007/s40203-021-00097-8

ORIGINAL RESEARCH

Structure based design, stability study and synthesis of the dinitrophenylhydrazone derivative of the oxidation product of lanosterol as a potential *P. falciparum* transketolase inhibitor and in-vivo antimalarial study

Olatomide A. Fadare[1] · Nusrat O. Omisore[2] · Oluwaseun B. Adegbite[1] · Oladoja A. Awofisayo[3] · Frank A. Ogundolie[4] · Julius K. Adesanwo[1] · Craig A. Obafemi[1]

Received: 30 April 2021 / Accepted: 31 May 2021

Abstract

The growing resistance to the current antimalarial drugs in the absence of a vaccine can be effectively tackled by identifying new metabolic pathways that are essential to the survival of the malaria parasite and developing new drugs against them. Triterpenes and steroids are the most abundant group of natural products with a great variety of biological activities. However, lanosterol is not known to possess any significant biological activity. In this study the binding and interactions of a dinitrophenyl hydrazine (DNP) derivative of lanosterol, LAN (a derivative that incorporates a substantially polar moiety into the steroid) with *P. falciparum* transketolase was studied by molecular docking and MD simulation with the view to exploit the DNP derivative as a lead in antimalarial chemotherapy development considering that the *P. falciparum* transketolase (PfTk) is a novel target in antimalarial chemotherapy. The enzyme catalyses the production of ribose sugars needed for nucleic acid synthesis; it lacks a three-dimensional (3D) structure necessary for docking because it is difficult to obtain a crystalline form. A homology model of PfTk was constructed using *Saccharomyces cerevisiae* transketolase (protein data bank ID of 1TRK) as the template. The compound was observed to have Free Energy of Binding higher than that of the cofactor of the protein (Thiamine Pyrophosphate, TPP) and a synthetic analog (SUBTPP) used as reference compounds after MD Simulation. The compound was synthesized in a two-step, one-pot reaction, utilizing a non-acidic and mild oxidant to oxidize the lanosterol in order to avoid the rearrangement that accompanies the oxidation of sterols using acidic oxidants. The LAN was characterized using IR spectroscopy and NMR experiments and tested in-vivo for its antimalarial chemo suppression using a murine model with Chloroquine as a standard. The LAN at a concentration of 25 mg/kg was found to have a comparable activity with Chloroquine at 10 mg/kg and no mortality was observed among the test animals 24 days post drug administration showing that the compound indeed has potential as an antimalarial agent and a likely inhibitor of PfTk considering that there is a strong agreement between the in-silico results and biological study.

Keywords Transketolase · Molecular dynamics · Homology modelling · Synthesis · Chemo suppression

✉ Olatomide A. Fadare
tofadare@oauife.edu.ng

[1] Department of Chemistry, Obafemi Awolowo University, Ile-Ife, Nigeria

[2] Department of Pharmacology, Faculty of Pharmacy, Obafemi Awolowo University, Ile-Ife, Nigeria

[3] Department of Pharmaceutical and Medicinal Chemistry, University of Uyo, Uyo, Nigeria

[4] Enzymology and Enzyme Biotechnology Unit, Department of Biochemistry, Federal University of Technology, Akure, Nigeria

Introduction

Malaria is a mosquito-borne infectious disease caused by a parasitic protozoan belonging to the *plasmodium* type. This parasite gets transmitted via a bite of an infected female Anopheles mosquito. Malaria is one of the diseases that pose a significant threat to human lives. Malaria is prominent in most parts of Africa, Asia, and South America due to the weather condition of these areas. According to the 2019 World Malaria Report (WHO 2010), Nigeria had the highest number of global malaria cases (25% of global

Published online: 18 June 2021

malaria cases) in 2018 and accounted for the highest number of deaths (24% of global malaria deaths). Malaria is a major public health problem in Nigeria where it accounts for more cases and deaths than any other country in the world. Malaria is a risk for 97% of Nigeria's population. The remaining 3% of the population live in the malaria free highlands. There are an estimated (average) 100 million malaria cases with over 300,000 deaths per year in Nigeria. This compares with 215,000 deaths (average) per year in Nigeria from HIV/AIDS. It is worthy to note that malaria contributes to an estimated 11% of maternal mortality as well in Nigeria (Nigeria Malaria Fact sheet—us embassy in Nigeria). By virtue of the fact that Malaria is a leading cause of death and disease in many developing countries especially in Nigeria, where young children and pregnant women are the groups most affected, and the propensity of the malaria parasite to develop resistance to known antimalarial drugs, especially, recently, that there have been reported cases of resistance to the most dependable and commonly used antimalarial drug, Artemisinin and derivatives (WHO 2017; Amato et al. 2018), it has become expedient that new antimalaria drugs are developed fast to counter the effect of artemisinin resistance, hence, this study. A more effective way to resolve the antimalarial resistance issue is to identify new metabolic pathways in the parasite that are essential for its survival and developing new drugs against them (Olliaro and Yuthavong 1999; Ridley 2002; Rosenthal 2001). The Pentose phosphate pathway (PPP) is an essential pathway in the parasite, this pathway is responsible for the replication of the parasite, PPP produces ribose sugar used in the synthesis of nucleotide and nucleic acid. Pentose phosphate pathway is an enzyme driven pathway and one of the critical enzyme in this pathway is transketolase, it is a novel target in the anti-malarial drug discovery effort even though it lacks an experimentally determined 3D structure. However, homology modelling has helped in bridging the gap between sequence and structure space, enabling scientists to construct protein models that are hard to crystallize or for which NMR spectroscopy structure determination is not tractable. It helps to construct a three-dimensional (3D) model of an unknown "target" protein from its amino acid sequence and an experimental 3D structure of a related homologous protein which is referred to as the "template".

Triterpenoids and steroids are the most abundant group of natural products with a great variety of biological activities. Their diverse use in the pharmaceutical industry have distinguished them as economically important substances that generates annual sales of 12.4 billion USD (Dzubak et al. 2006; Nazaruk and Borym-Kluczyk 2015). Triterpenoids have been shown to possess various activities like anti-parasitic, anti-fungal, anti-bacterial, anti-inflammatory, anti-proliferative, anti-diabetics etc. Nogueira and Lopes (2011) and Nasomjai et al. (2014) reported that a class of triterpenoid "tetracyclic triterpenoid" isolated from *Momordica balsamina L.* (Cucurbitaceae) having the lanostane skeleton (steroids) exhibit in-vitro anti-plasmodial activity against blood stages of parasites, hence this study of exploiting (lanosterol also a steroid and endogenous) as an antimalarial agent. It is thought that lanosterol, being lipohilic and endogenous would be harmless to the human body and if derivatized in such a way that it posses polar functional groups, it might have some medicinal properties. This present research envisaged to use computational techniques to investigate the interactions of the proposed compound with the *P. falciparum* transketolase enzyme and synthesize dinitrophenylhydrazone derivative of oxidized lanosterol as a potential inhibitor of the enzyme by occupying the cavity that the substrates would normally occupy and binding with critical polar residues within the active site of the enzyme such that the metabolic function of the enzyme is diminished and ultimately leading to the inhibition of the plasmodium causing malaria knowing that the *P. falciparum transketolase* enzyme has least homology with the human transketolase enzyme.

Materials and methods

A model of *Plasmodium falciparum* transketolase was constructed by homology modelling using the software package, Modeller v9.1. The homology model was optimized by molecular dynamic simulation using the GROMACS v5.0 software package and the optimized model was visualized and analysed using Visual Molecular Dynamics (VMD) v1.9. The docking study was carried out using Autodock vina v12.0. The structures of the compounds were drawn in Chemdraw v12.0 and energy minimization of the structures done by Chem 3D, visualization of the protein and structures was done in Pymol v1.7.4.5. The molecular dynamics simulation of the modelled protein with the proposed lanosterol derivative (LAN) was done in a box of water using the GROMACS software and grommos 54A7 forcefield to establish the trajectory after 50 ns from which information on protein–ligand complex stability was determined, root mean square deviation (rmsd) and fluctuation (rmsf), radius of gyration (ROG) and solvent accessible surface area (SASA), after which the g_MMPBSA module was used to estimate the free energy of binding. The MD simulation (as well as free energy estimate) was replicated for two reference compounds, thiamine pyrophosphate (TPP) and a synthetic analog (SUBTPP) that are cofactors of the transketolase enzyme, for comparison (Fig. 1). The lanosterol derivative was synthesized and characterized using spectroscopic techniques (NMR experiments and UV–Vis spectrocopy). The synthesized compound was then tested for antimalaria chemosuppression using a murine model.

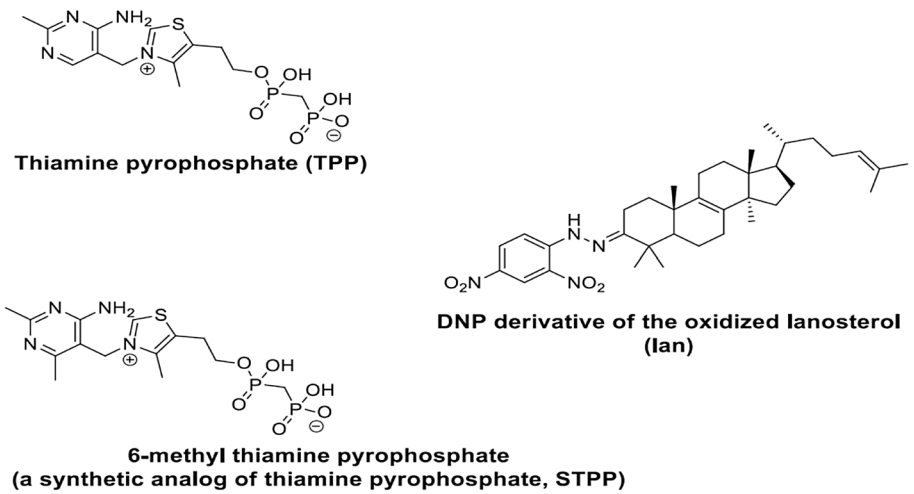

Fig. 1 The structure of the compounds studied as ligands in the molecular docking and MD simulation

In-silico study of the interaction of study compound with transketolase enzyme

This was done prior to wet lab synthesis using available software, this intends to gain insight into the possible interaction of the DNP derivative of the oxidized lanosterol with the transketolase enzyme to see if it will have any potential as an inhibitor (in-silico) before going ahead to synthesize and carry out any biological test.

Model construction (homology modelling)

The amino acid sequence of *Plasmodium falciparum* transketolase (target protein) was obtained in FASTA format from National Centre for Biotechnology Information (NCBI), a similarity search was performed in the Protein Data Bank (PDB) using protein-basic local alignment search tool (BLASTp) in order to get a homologous protein which will serve as the template protein. The percentage homology with the target and the resolution of the 3D structure as well as the statistical significance (E-value) of the alignment between the target and the template protein were taken into consideration when selecting a template that was used in the construction of the target protein. The amino acid sequence of the target protein (*P. falciparum* transketolase) was aligned with the amino acid sequence of the template protein (*Saccharomyces cerevisiae* transketolase) using the python script *align2d.py* of the software, MODELLER, in order to obtain optimal alignment. *P. falciparum* transketolase model was constructed using the python script *Build_profile.py* of the same MODELLER software (Šali and Blundell 1993). This software was used to construct five similar models of which the one with the lowest Discreet Optimized Potential Energy (DOPE) and highest GA341 score was chosen as the best model. The model built was evaluated by the python script, *evaluate_model.py*; it was also validated by constructing the Ramachandran plot for the model using the Zlab server for Ramachandran plot generation (University of Massachusetts Medical School). The RMSD of the modelled protein structure was also estimated as a measure of the validity of the model.

Protein structure optimization

The *P. falciparum* transketolase homology model was subjected to molecular dynamic simulation using the software GROMACS (Abraham et al. 2015). A topology was first generated: topology contains all the information necessary to define the molecule within a simulation, the force field, OPLS AA was used and a box (rhombic-dodecahedron) filled with water molecules was used for the simulation, ions were added to the system to balance the charges of the solvated protein, the homology model (3D structure) was relaxed by energy minimization in order to ensure that the system has no steric clashes or inappropriate geometry, then the solvent and ions around the protein were equilibrated to a desired temperature, pressure and density. After the system was well equilibrated, the production molecular dynamics was run for data collection and the simulated protein was analysed for root mean square deviation (rmsd), stability by assessment of radius of gyration and solvent accessible surface area.

Molecular docking

The docking studies were carried out using Autodock tools v1.5.4 (Morris et al. 2009) and AutodockVina v1.1.2. (Trott and Olson 2010). The chemical structures were drawn using Chemdraw v12.0 and energy minimization was done using Chem 3D v12.0. The lanosterol derivative (LAN), thiamine pyrophosphate (TPP) and 6-methyl thiamine pyrophosphate (SUBTTP) were subjected to molecular docking

using Autodock tools. TPP (a cofactor of transketolase) and 6-methyl TPP (SUBTPP) served as reference compounds. The protein was converted from.pdb to.pdbqt format, after which its binding site was analyzed and assigned a grid box (search space area) using AutoDOCK Tool. The ligands were prepared for docking using Chem3D and AutoDock Tools. After the preparation of the ligands and protein, each ligand was docked into the protein using AutoDock Vina to estimate the binding affinity and best binding pose.

MD simulation and free energy estimation

Molecular dynamics study of the protein–ligand complexes were done using GROMACS 2018.3 (Abraham et al. 2015) software which was installed in ubuntu 18.04 LTS. Docked structures of the protein–ligand complex were used in the simulation study. The protein was processed and the topology file was prepared by using pdb2gmx and GROMOS54a7atb.ff (imported from the automated topology builder website) force field while the ligand topology file was prepared by using the Automated Topology Builder (ATB) version 3.0. The solvent addition was done in a cubic box by using a box distance 1.0 nm from closest atom in the protein. The addition of Cl^- ions was used to neutralize the system. Energy minimization was done by using steepest descent algorithm taking 50,000 steps and 5 kJ/mol maximum force and Verlet cut-off scheme taking Particle Mesh Ewald (PME) columbic interactions. Position restraints were applied in the equilibration step. After that, NVT equilibration was done in 300 K and 100 ps of steps and NPT equilibration taking Parrinello-Rahman (pressure coupling), 1 bar reference pressure and 100 ps of steps. LINCS algorithm was applied to constraint all the bonds length. For long-range electrostatics, Particle-mesh Ewald (PME) algorithm was used. And finally the production MD of the protein–ligand complex was run for 50 ns.

After successful completion of molecular dynamic simulation for 50 ns, Root mean square deviation (RMSD) of backbone residues, number of Hydrogen bonds, Root mean square fluctuations (RMSF), Radius of gyration (RG), Solvent accessible surface area (SASA) (Marsh and Teichmann 2011) were calculated. Free energy of binding calculations were done following the g_mmpbsa protocol by Kumari and Kumar (2014), developed for GROMOS96 43a1 force field. Molecular Mechanics Poisson-Boltzmann surface area (MM-PBSA) is a widely used method to calculate the free binding energy and to predict the stability of the complex. Modified Molecular Mechanics–Poisson Boltzmann Surface Area (MM-PBSA) is open-source software, and employed to calculate the free energy of binding between two defined groups. Recently, MM-PBSA algorithm has been utilized as a scoring function in computational drug design (Pant et al. 2020). In this study, MM-PBSA method was employed to calculate the free energy of binding between LAN, TPP and SUBTPP and the modelled transketolase. The free energy of binding calculation is based on the following theory:

$$G_{binding} = G_{complex} - \left(G_{protein} + G_{ligand}\right)$$

where $G_{complex}$ is the total free energy of the protein–inhibitor complex and $G_{protein}$ and G_{ligand} are total free energies of the separated form of protein and inhibitor in solvent, respectively (Egan et al. 2000). The average binding energy calculations were done by a python script provided in g_mmpbsa stand-alone program.

Synthesis

All the chemicals and the reagents used were of synthetic and analytical grades. The synthesis of the dinitrophenylhydrazone of oxidized lanosterol was carried out by first oxidizing lanosterol with freshly prepared quinoliniumchlorochromate (QCC), stirring at room temperature for 18 h, after which the oxidized product was reacted with 2,4-dinitrophenylhydrazine using an acid catalyst (5 drops of conc H_2SO_4) by stirring at room temperature for 24 h. After the synthesis, the product was purified using column chromatography and then recrystallized using appropriate solvent. The dinitrophenylhydrazone derivative of oxidized lanosterol (Compound A) was characterised by analysing the Retention factor (Rf), Melting point, Ultraviolet–Visible spectrum, Infra-red Spectrum, ^1H and ^{13}C NMR Spectra. The melting point of the DNP derivative of the oxidized lanosterol was determined in an open capillary tube, the compound was also confirmed for its purity on TLC plate using suitable solvent system. IR Spectrum of the synthesized compound was recorded using KBR pellets in the range of 400–4000 cm^{-1} on Agilent Cary 630 FT-IR spectrometer. ^1H NMR and ^{13}C of the synthesized compound were recorded in $CDCl_3$ by Bruker 400 MHz FT-NMR Spectrometer**.**

Procedure for the formation of 2,4-dinitrophenylhydrazone derivative of oxidized lanosterol (LAN)

This synthesis involves a two-step, one-pot reaction (intermediate ketone was not isolated).

Step 1. Oxidation of lanosterol

One molar equivalent of Lanosterol (10 g; 23.67 mmols) was added to the freshly prepared quinolinium chlorochromate (QCC) (6.22 g; 23.67 mmols) in dichloromethane (DCM) and stirred at room temperature for 18 h. The reaction mixture was worked up by extraction with dichloromethane and concentrated in-vacuo to get the crude product which was dried and weighed

(70% yield), melting point of 116–120 °C. IR (KBr): 1709 cm^{-1} (C=O), 2926 cm^{-1} (sp3-C–H stretch), 1464 and 1377 cm^{-1} (C–H bend of hydrocarbon).

Step 2. Formation of the DNP derivative of oxidized lanosterol

One molar equivalent of oxidized lanosterol was added together with 2,4-dinitrophenylhydrazine in ethanol using an acid (1 mL of conc. H_2SO4) as a catalyst. The mixture was stirred at room temperature for 24 h. Being monitored with thin layer chromatography (TLC), the reaction mixture was filtered, dried and packed on a column to purify. The product of the reaction was concentrated, dried and weighed, melting point of 198-200 °C. IR (KBr): 3372 cm^{-1} (N–H), 3103 cm^{-1} (Sp^2-C–H) 2951 cm^{-1} (Sp^3-C–H), 1618 cm^{-1} (C=N), 1518 and 1334 cm^{-1} (NO_2), 1591 cm^{-1} (C=C Ar), 1471 cm^{-1} (C=C–H bend), 1132 cm^{-1} (C–N). The ^{1}H NMR data is presented as follows; 1H, d, δ11.20 (J=4.3 Hz); 1H,s, δ11.10; 1H,d, δ9.13 (J=2.6 Hz); 1H,dd, δ8.30 (J=8.9, 5.6, 2.6 Hz); 1H,dd, δ7.95 (J=8.4, 5.9, 2.4 Hz); 1H,s, δ5.30; 1H,t, δ5.10; 2H,m, δ 2.46–2.85; 1H,t δ 2.32 (J=6.8 Hz); 2H, s, δ 2.18; 4H,m, δ 1.98–2.13; 2H, m, δ1.79–1.98; 7H,m, 1.52–1.76; 3H,m, 1.36–1.51;4H, m, 1.28–1.35; 4H,d δ 1.24 (J=6.5 Hz); 4H, m, δ1.10–1.20; 2H,dd, δ 1.03 (J=26.8, 6.9 Hz); 7H, pd δ 0.88 (J=7.9, 7.0, 3.9 Hz), 2H,d, δ0.70 (J=9.5 Hz). And the ^{13}C NMR data is also presented as follows (δ values, ppm); 35.55, 27.81, 38.85, 50.35, 18.22, 26.46, 134.35, 134.35, 36.97, 20.97, 30.94, 44.43, 49.76, 30.82, 28.18, 50.35, 15.71, 19.11, 36.24, 18.61, 36.33, 24.89, 125.22, 130.91, 25.72, 17.62, 27.93, 15.40, 24.43.

Biological study

Lethal dose determination

The acute toxicity of Compound A was determined using Lorke's method (Lorke 1983). Twelve mice were divided into four groups (A, B, C and D). The four groups were administered orally with graded concentrations (10, 100, 1000, 5000 mg kg^{-1} respectively) of the compound. These animals were watched for 24 h.

Antimalarial chemo suppressive test

Swiss white mice were divided into groups of five and inoculated with 1.0 *10^7 *P. berghei* parasites intraperitoneal. Different doses of Compound A were dissolved in 1:1 mixture of DMSO and H_2O. Three of the groups were given oral doses of 25, 50 and 100 mg/kg of the compound respectively. Two other groups were administered with 10 mg/kg chloroquine (reference drug) and 0.2 ml of vehicle of preparation (negative control) respectively. The test agents were administered for four (4) consecutive days. On the fifth day, thin blood smear was prepared and the blood films were stained with Giemsa solution and then microscopically examined with 100 × magnification (oil immersion). The percentage chemo suppression of parasitaemia was calculated for each dose level by comparing the % parasitaemia of the controls with those of treated mice (Knight and Peters 1980). The percentage parasitaemia of the experimental animals after the four days of drug administration was estimated as count of parasitized red blood cells divided by the total red blood cells counted within 20 microscopic fields (having an average of 2000 total red blood cells counted).

The percentage chemo suppression (Chart 2) was calculated as a relative value taking into account the % parasitaemia of the negative control. The formula used is as described below

$$\%\ \text{Parasitaemia} = \frac{\text{Parasitized red blood cells}}{\text{Total red blood cells}} \times 100$$

$$\%\ Chemosuppression = \frac{(X - Y)}{X} \times 100$$

X = % Parasitaemia of the negative control Y = % Parasitaemia of the test compound.

Result and discussion

Homology modelling

The target protein, *P. falciparum* transketolase, has 672 amino acid residues in its sequence. Homology modelling was used to construct a 3D structure for the target protein. *Saccharomyces cerevisea* transketolase having 49% homology with the target protein was used as the template. The alignment of the amino acid sequence of the target and the template done by Clustal Omega is shown in Fig. 2. Five different models of this target protein were successfully constructed by the MODELLER software, the model with the lowest DOPE score of − 80,725.75 was selected. The model constructed was evaluated by plotting the energy profile, Fig. 3, the energy profile also shows the alignment between

Table 1 Binding affinities of the ligands

S/N	Compounds docked	Binding affinity (kcal/mol)
1	DNP derivative of oxidized lanosterol (compound A)	− 8.0
2	TPP	− 7.0
3	Synthetic Analog of TPP	− 7.0

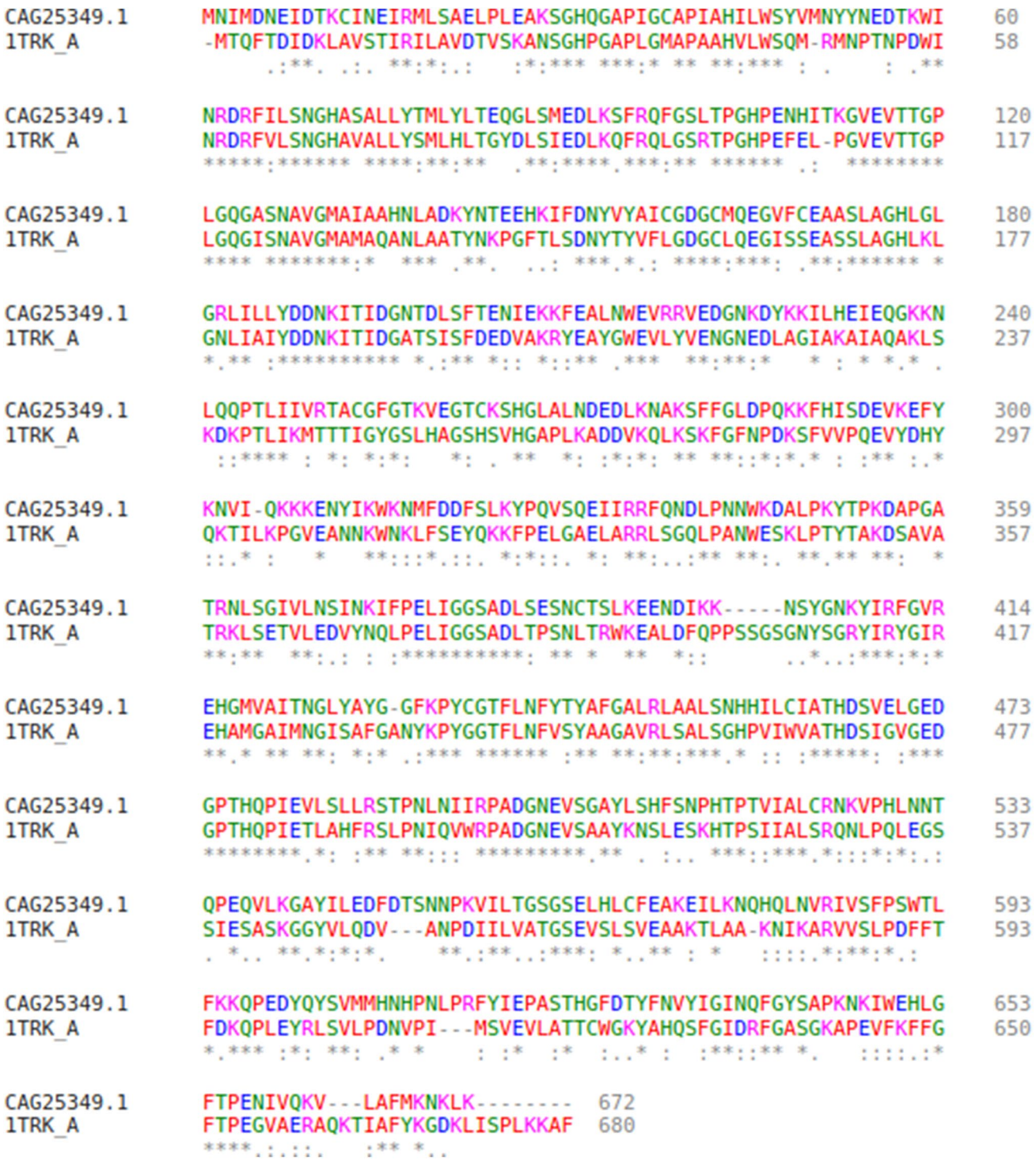

Fig. 2 A screenshot showing alignment between the amino acid sequence of *P. falciparum* transketolase and *Saccharomyces cerevisea* transketolase. The target is represented as CAG25349.1 and template as 1TRK-A (PDB code)

the template protein and the target model in a graphical form, the energy profile in Fig. 3 shows close alignment for most regions of the proteins. A Ramanchandran plot that shows the stereo chemical assessment of the model constructed is shown in Fig. 4. The Ramanchandran plot, Fig. 4, shows that 95.1% and 4.2% of the amino acid residues are in favoured and allowed regions respectively, only 0.7% was found to be in the outlier region, this also confirmed that the model

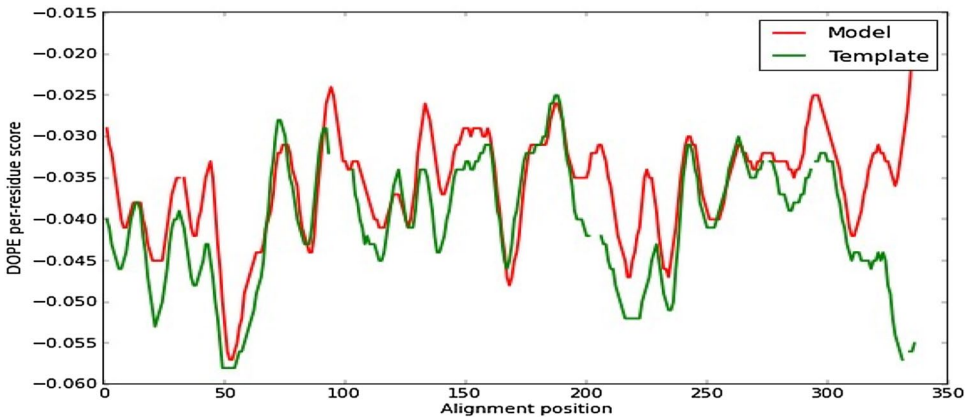

Fig. 3 Energy profile of the alignment between the target protein model and the template

is a very good one. The superimposition of the model and the template gives a RMSD value of 1.03 Å which further buttress the fact that the model is a good model since an RMSD less than 7 is said to signify a good model. The protein model was optimized by molecular dynamic simulation in a box of water with OPLS force field. The model, after being subjected to molecular dynamic simulation was also assessed using a Ramanchandran plot, Fig. 5, and the number of amino acid residues in the favoured and allowed regions are 87.6% and 10.6% respectively. Only 1.8% are found in the outlier region. The superimposition of the homology model and the 3D structure obtained after optimization with molecular dynamic simulation gives a RMSD value of 2.65 Å.

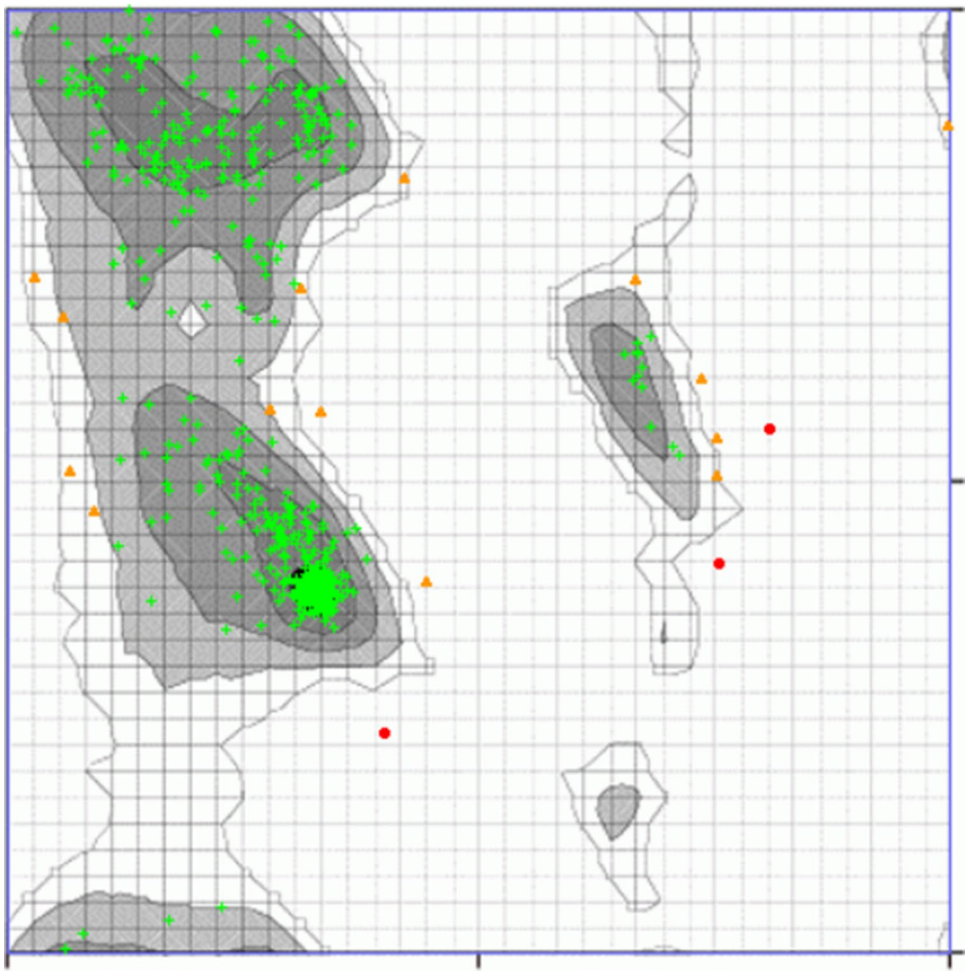

Fig. 4 Ramachandran plot of the homology model of the target protein (Green crosses - highly preferred (favoured); Brown triangles - preferred (allowed) and Red circles - questionable observations (outlier region))

Table 2 Infrared spectral data of lanosterol, oxidized lanosterol and compound A

Compound	Lanosterol (starting material)	Oxidized lanosterol (intermediate)	Compound A
Formula	$C_{30}H_{50}O$	$C_{30}H_{48}O$	$C_{35}H_{52}N_4O_4$
Frequency (cm^{-1})	3306 (O–H)	1709 (C=O)	1618 (C=N)
	2953 (sp^3-C–H)	2926 (sp^3-C–H)	2951 (sp^3-C–H)
	1371 and1458 (C–H bend of hydrocarbon)	1464 and 1377 (C–H bend of hydrocarbon)	1471 (C=C–H bend)
			3372 (N–H)
			1518 and 1334 (NO_2)
			1132 (C–N)
			3103 (sp^2-C–H)
			1591 (C=C Ar)

Docking studies

Molecular docking provided insight into the interactions and the binding modes of the protein and compound A. The active site of the target protein model was determined by visual inspection and superimposition of the templates 3D structure on that of the protein model. The proteins and the compound were prepared as described earlier. The 3D structure of compound A was docked into the target model with the gridbox (dimension of the grid box) centred around the active site. The cofactor (TPP) and its synthetic analog (6-methyl TPP) were also docked into the protein to serve as reference. Compound A has the highest binding affinity of − 8.0 kcal/mol relative to that of the cofactor (TPP, − 7.0 kcal/mol) and synthetic analog (6-methyl TPP, − 7.0 kcal/mol). The higher the binding affinity, the stronger the interaction between the protein and the ligand. The resultant binding affinities from all the interactions for the docked compounds are shown in Table 1. This implies that the enzyme has higher affinity for compound A relative to that of cofactor, which suggest that compound A has the capacity to displace the cofactor and eventually disrupting the activity of the enzyme, thereby effectively inhibiting the enzyme, and preventing the enzyme from carrying out its catalytic function and subsequently preventing the parasite from carrying out its metabolic activity which will eventually lead to the death of the parasite. This observation was corroborated from a follow-up in-vivo study using animal model.

MD simulation and free energy estimation

The conformational changes that took place during the simulation were estimated by determining the RMSD, RMSF, ROG, SASA and H-bonds. The extent of conformational change was used as a measure of the stability of the system and extent of interaction between the protein and the ligands. The rmsd plot of the protein backbone fitted to the ligands (is shown in Fig. 3) shows that the simulations reached equilibrium at around 25 ns for TPP and SUBTPP and at 35 ns for LAN. The RMSF plots for the protein (C_alpha) fitted to the ligands shows that the ligands fluctuated around the same set of residues within the protein except for the residues between 250 and 350 that showed marked differences in the fluctuations for the three ligands as observed in the plots for RMSF (Fig. 6). The radius of gyration (ROG) plots (Fig. 6) which indicates the compactness of the protein–ligand complex decreased steadily showing that the protein–ligand complex for TPP and the lanosterol derivative (LAN) was getting more compact as the simulation progressed but the complex of SUBTPP with the protein did not achieve the kind of compactness that TPP and lanosterol had over the course of the simulation.

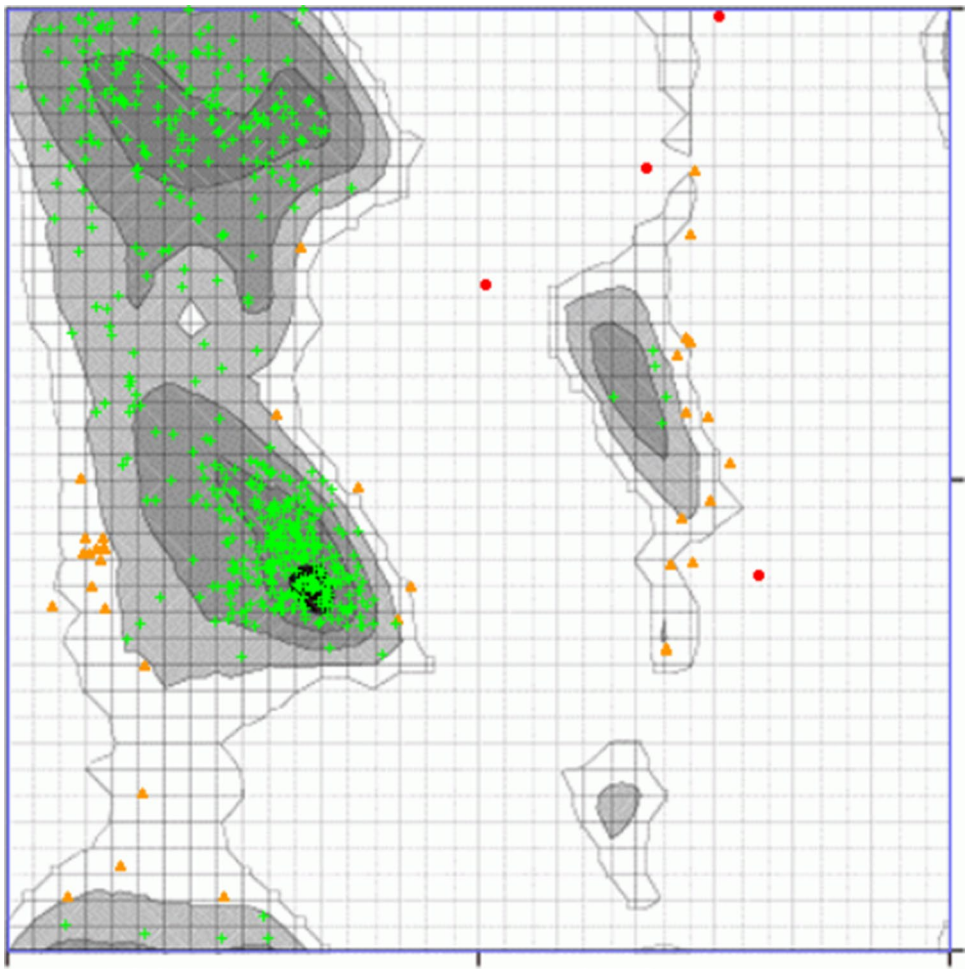

Fig. 5 The Ramachandran plot of the simulated protein in a box of water, for optimization (Green crosses - highly preferred (favoured); Brown triangles - preferred (allowed) and Red circles - questionable observations (outlier region))

The gradual decrease in the solvent accessible area of the complexes as the simulation progressed show that the binding of the ligands with the protein caused the protein to constrict in such a way that the surface area available for interaction of water molecules reduced which is typical for proteins having strong interactions with ligands. The plots for hydrogen bonding show that TPP and SUBTPP have plenty of hydrogen bond interactions with the protein throughout the course of the simulation which was expected because of the high number of polar functional groups (considering the heterocyclic core of the TPP and SUBTPP is also charged) which facilitated strong electrostatic interactions Fig. 7. The plot of resdiues contributing to free energy (Fig. 8) shows some strong repulsive interactions between some residues and ligands (LYS191 and ASP195 for TPP; ASP189 and LYS191 for SUBTPP)—these repulsive interactions must have been offset by some other strong attractive interactions (though smaller in magnitude but more) which culminated in an overall strong electrostatic interaction unlike the lanosterol derivative which did not have much of the electrostatic interactions, observed for TPP and SUBTPP, due to being largely hydrophobic. The Van der Waals forces of attraction between the lanosterol derivative and protein appears to be more pronounced than its electrostatic interactions, however, TPP and SUBTPP also have significant VDW interactions (Fig. 7c). The Overall free energy of binding of the three ligands shows that the lanosterol derivative has the highest free energy of binding relative to TPP and SUBTPP which had markedly reduced Free energy of binding despite their high electrostatic interaction contributions. The overall free energy must have been offset by the high cost of desolvation of the ligands. Their charged/polar nature must have caused them to associate strongly with the water molecules such that the penalty for stripping them of water molecules at the binding site must have been very high as observed in the breakdown shown in Fig. 9. It is a well-known fact that the traditional, enthalpy-dominated way of regarding ligand–protein interactions essentially ignores solvation effects, which has now been establihsed to strongly affect the thermodynamic profile of a binding event (Bronowska 2011). Several detailed binding free energy studies have suggested that differences in solvation may play a critical role in the observed differences in the free energy of binding between relatively similar compounds (Jiao et al.

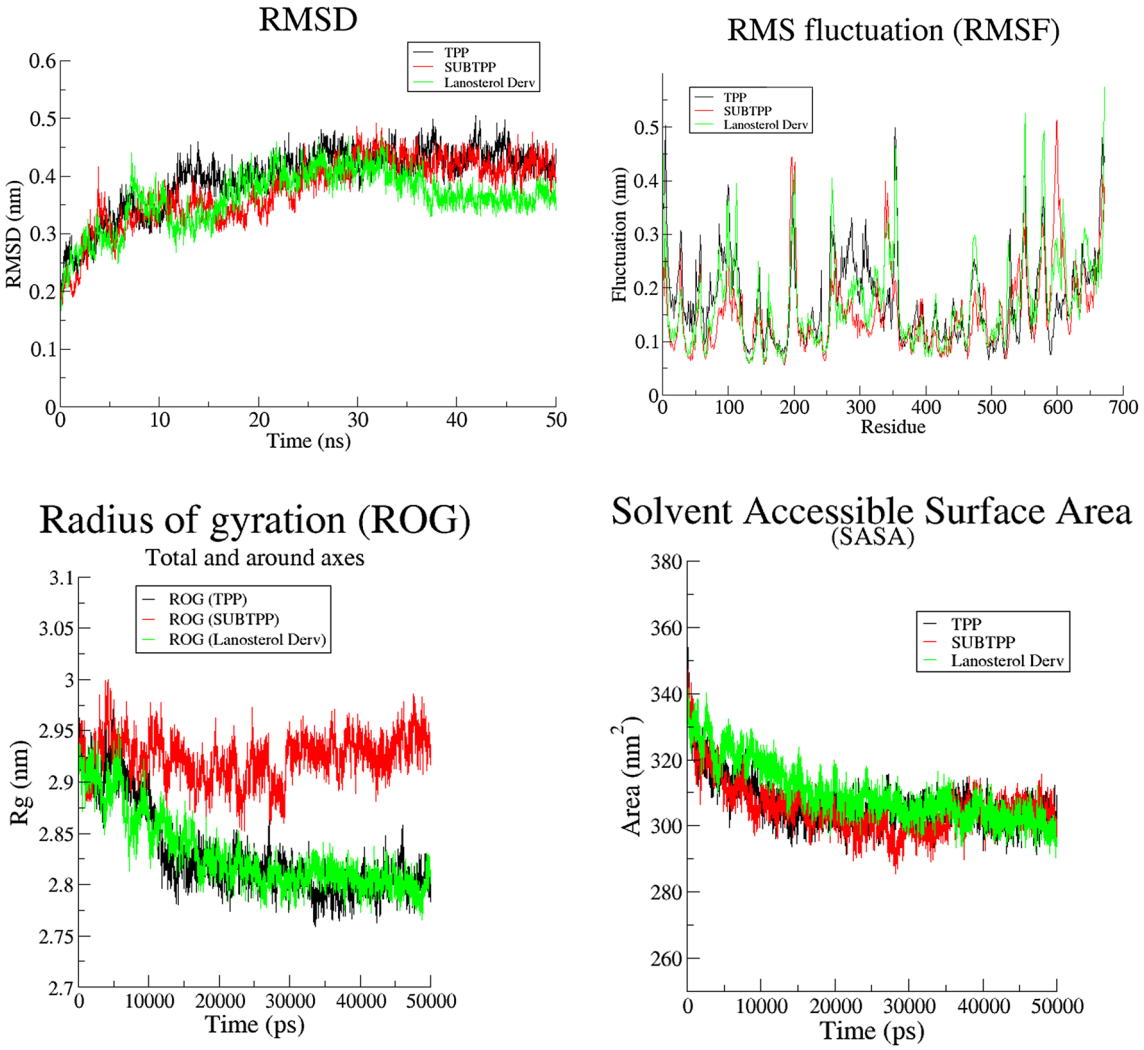

Fig. 6 The plots of RMSD, RMSF, ROG and SASA

2008; Reddy and Erion 2001). Two molecules might have similar interactions with a protein, similar strain energies, etc., but have different solvation properties in water, leading to solvation-driven differences in binding free energies (Bronowska 2011). Recently it became clear that studying the hydration state of a protein binding pocket in the apo (unbound) state should be a routine procedure in rational drug design, as the role of solvation in tuning binding affinity is critical (Bronowska 2011). Solvation costs appear to be a valid reason why some ligands, despite fitting into a binding site, fail during experimental tests as inhibitors. Young and co-workers showed that an optimised inhibitor of factor Xa turns virtually inactive when the isopropyl group interacting in the S4 pocket of factor Xa is replaced by hydrogen. Substitution of this group by hydrogen, apart from trimming down the number of favourable hydrophobic interactions, leads to unfavourable solvation of the binding pocket (Young et al. 2007).

Desolvation of the ligand itself may sometimes control the binding free energy. Several detailed binding free energy studies have suggested that differences in solvation may play an important role in differences in binding free energy between relatively similar compounds (Jiao et al. 2008; Reddy and Erion 2001; Mobley and Dill 2009). Two molecules might have similar interactions with a protein, similar strain energies, etc., but have different solvation properties in water, leading to solvation-driven differences in binding free energies (Mobley and Dill 2009).

For highly hydrophilic ligands, the desolvation costs may be very high and make unfavourable contributions to the binding (Daranas et al. 2004; MacRaild et al. 2007; Syme et al. 2010). Essentially, a charged ligand may make favorable electrostatic interactions in a polar binding site, but it may also cost a huge amount of energy to remove it from water (Brenk et al. 2006; Gilson and Zhou 2007; Shoichet et al. 1999). In other cases, a small modification to a ligand

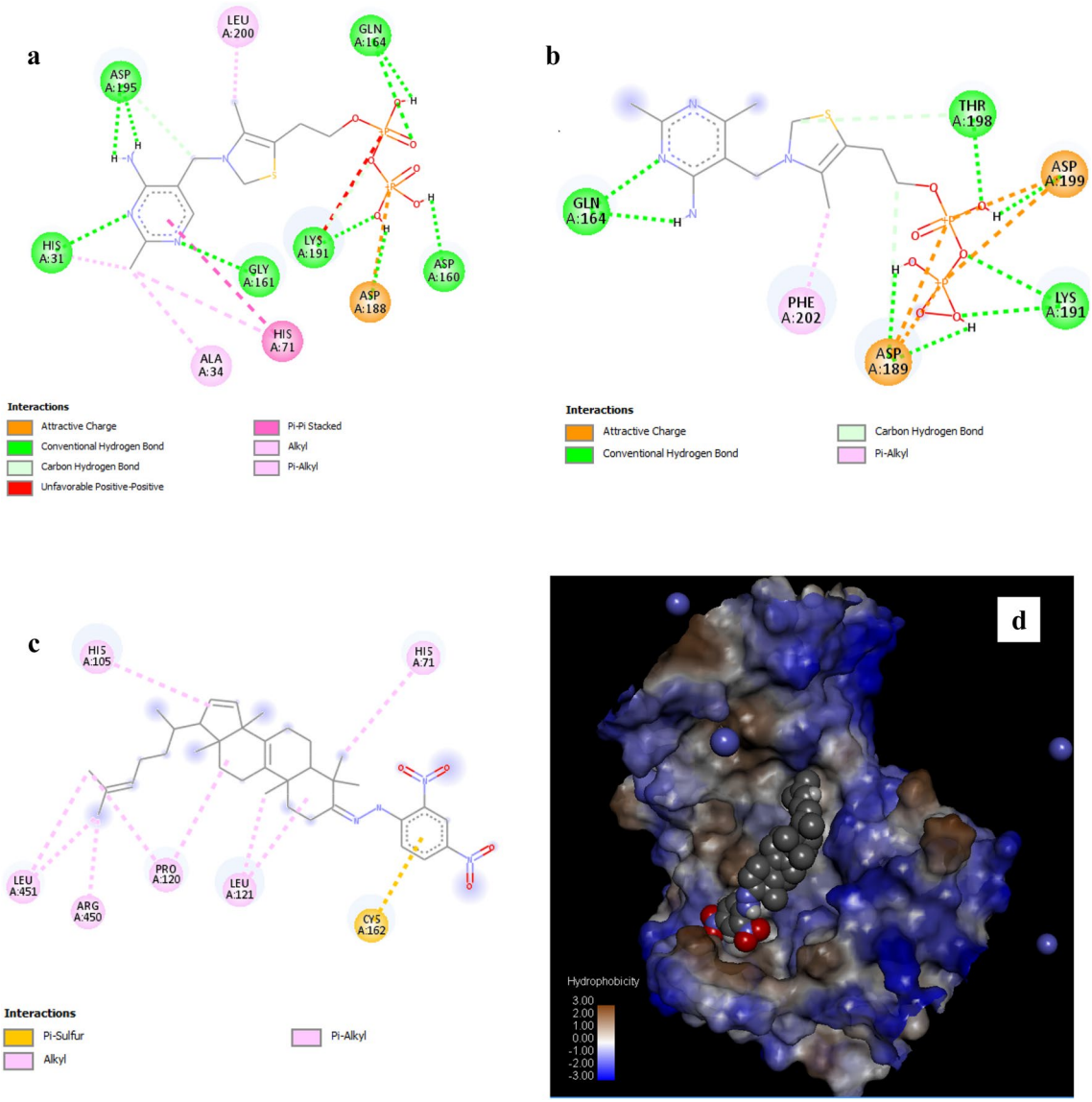

Fig. 7 **a** TPP showing residues within close proximity in the binding pocket with the different types of interactions. Particularly the unfavourable positive-positive interaction between Lysine 191 and a positively charged phosphorus (enhanced electrophilicity) in the phosphate group. **b** SUBTPP showing residues within close proximity in the binding pocket showing more electrostatic interactions around the phosphate groups. **c** The lanosterol derivative within the binding site of the transketolase, showing more of Van der Waals interactions. **d** The protein rendered as an hydrophobic surface, showing the lanosterol derivative (CPK rendering) with the hydrocarbon part hovering over the entrance of the active site with the polar groups occupying positions that are tending towards hyrophilic on the surface—noting that the TPP and SUBTPP have preference for deeper in the binding site. The images are obtained from a midpoint of a cluster from geometric clustering performed using the gmx cluster module in the GROMACS program (2018.3) with the trajectory from the 50 ns simulation of the ligand with the protein

can potentially lead to affinity gains due to a change in the desolvation cost (Kangas and Tidor 2001). Therefore, in this study, it would have been expected that the strong electrostatic interactions between the reference compounds, TPP and SUBTPP within the deep binding pocket would amount to high Free Energy of Bininding but it turns out that the cost of desolvation played a significant role in the reduction of the binding energy. Furthermore, the SUBTPP being a synthetic derivative of TPP with a methyl substituent at the 6-position hard a markedly reduced reduction in the free energy of binding relative to the TPP (TPP is -33 ± 1.529 while SUBTPP is -6.152 ± 1.529 kJ/mol) which confirms the observations and conclusions made by Jiao et al. (2008) as well as Reddy and Erion (2001), for polar/charged ligands within a deep binding pocket filled with water. The overall binding energy for the lanosterol derivative being 5 times

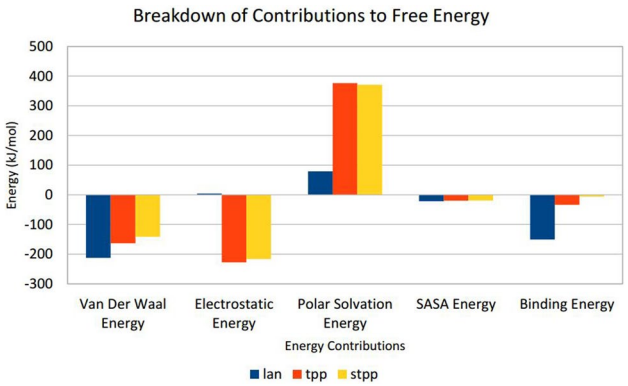

Fig. 8 Chart of amino residues that made significant contributions to the Free Energy of Binding

more than that of TPP (a cofactor) shows that the lanosterol derivative would be a potent inhibitor of the *P. falciparum* transketolase and a possible antimalaria agent whit high efficacy if it is bioavailable at the site of action.

Synthesis

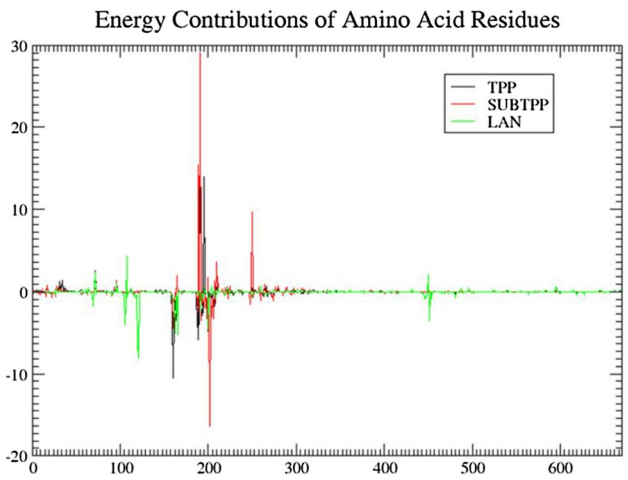

Fig. 9 Breakdown of components of Free energy that make up the total Free energy of binding of the ligands

The synthetic strategy of the 2,4-dinitrophenylhydrazone derivative of oxidized lanosterol is depicted in Fig. 7. Lanosterol was oxidized by freshly prepared QCC which is a mild oxidizing agent to minimize the incidence of rearrangement or prevent it completely (Dhar and Singh 1977). Attempts to oxidize sterols in literature, using acidic oxidants such as Jones reagent or chromic acid usually led to a mixture of products due to the possibility of rearrangements of double bond position that the acidic medium affords. This rearrangement will normally accompany the oxidation of the sterol such that, starting with one sterol will yield an expected ketone and other ketones borne from rearrangement of double bond positions. It is expected that using QCC which is not acidic will minimize the incidence of rearrangement or prevent it completely (Dhar and Singh 1977). The observation from the oxidation of lanosterol with QCC in this study shows that the incidence of rearrangement is greatly minimized, since the TLC of the oxidized product shows one major ketone and another present in trace amount. The oxidized lanosterol was reacted with 2,4-dinitrophenylhydrazine to produce the lanosterol derivative (LAN). The lanosterol derivative was purified by packing on a column for chromatography (eluted with a mixture of n-hexane and ethyl acetate at different ratios) which was also recrystallized. The LAN was characterised by UV, FTIR, ^{1}H and ^{13}C NMR.

The infrared spectra of the lanosterol derivative (LAN), intermediate ketone and starting sterol were compared. The absorption profiles confirm that the expected reactions from sterol to ketone to DNP derivative occurred (Table 2). A sharp band at 3372 cm^{-1} characteristic of N–H stretching vibration (which must be due to the hydrazinyl moiety) was observed in the IR spectrum of LAN, a band at 1618 cm^{-1} (sharp with medium intensity) was also observed which is attributed to the C=N bond stretching vibration. Two characteristics absorption for $–NO_2$ vibration at 1518 cm^{-1} and 1334 cm^{-1} were also observed as well as bands at 1591 cm^{-1} and 1471 cm^{-1} for C=C of benzene which show that the DNP moiety has been incorporated into the target molecule. The ^{1}H NMR spectrum shows three signals in the aromatic region which is expected for the DNP moiety: 1H, d at δ9.13 ppm; 1H, dd at δ8.30 ppm (J = 8.9 Hz) and 1H, dd at 7.95 ppm (8.4 Hz). The number of protons in the hydrocarbon region is characteristic of a steroid (lanosterol in particular), this represents the steroidal part of the molecule. The region further downfield, beyond the aromatic region is expected to have one signal corresponding to the hydrazinyl proton but shows two signals, close to each other, and their integrations are fractions of a whole which suggest a form of tautomerism occurring in solution such that the compound switches between an imine form and eneamine form (Fig. 10). This suspected tautomerism is confirmed by the second observed signal in the alkene region of the spectrum, where only one signal is expected but two are observed. The ^{13}C NMR spectrum also shows the same trend (suggesting a form of tautomerism). The hydrocarbon region matches with that of lanosterol from literature (Emmons et al. 1989) (Table 3).

Table 3 Comparison of ^{13}C spectroscopic data of lanosterol from literature with that of the synthesized compound (2, 4-dinitrophenylhydrazone derivative of oxidized lanosterol)

S/N	Carbons	Dinitrophenylhydrazone derivative of lanosterol (pure, experimental)	Lanosterol (Emmons et al. 1989)
1	C1	40.15	35.55
2	C2	28.90	27.81
3	C3 (C=N)	167.98	Absent
4	C4	45.45	38.85
5	C5	51.86	50.35
6	C6	19.19	18.22
7	C7	28.57	26.46
8	C8	146.18	134.35
9	C9	146.10	134.35
10	C10	45.01	36.97
11	C11	24.46	20.97
12	C12	40.07	30.94
13	C13	48.36	44.43
14	C14	50.48	49.76
15	C15	37.58	30.82
16	C16	36.81	28.18
17	C17	50.93	50.35
18	C18	18.35	15.71
19	C19	23.40	19.11
20	C20	42.80	36.24
21	C21	21.66	18.61
22	C22	42.94	36.33
23	C23	25.51	24.89
24	C24	124.15	125.22
25	C25	131.55	130.91
26	C26	26.34	25.72
27	C27	19.09	17.62
28	C28	31.46	27.93
29	C29	16.43	15.40
30	C30	24.85	24.43
31	DNP Carbon	117.03	Absent
32	DNP Carbon	129.45	Absent
33	DNP Carbon	129.69	Absent
34	DNP Carbon	133.54	Absent
35	DNP Carbon	164.98	Absent
36	DNP Carbon	136.13	Absent
37	DNP Carbon	137.99	Absent
38	DNP Carbon	155.77	Absent
39	DNP Carbon	145.71	Absent
40	C=C (transient)	140.29	Absent

In-vivo study

Lethal dose determination

The oral safety of the test compound (LAN) was ascertained by the outcome of the LD_{50} assay, none of the animals died even at the stipulated highest dose of 5000 mg/kg (OECD guidelines). This gives the benefit for dosing in the use of the substance since any LD_{50} beyond 5000 mg/kg is of no practical interest.

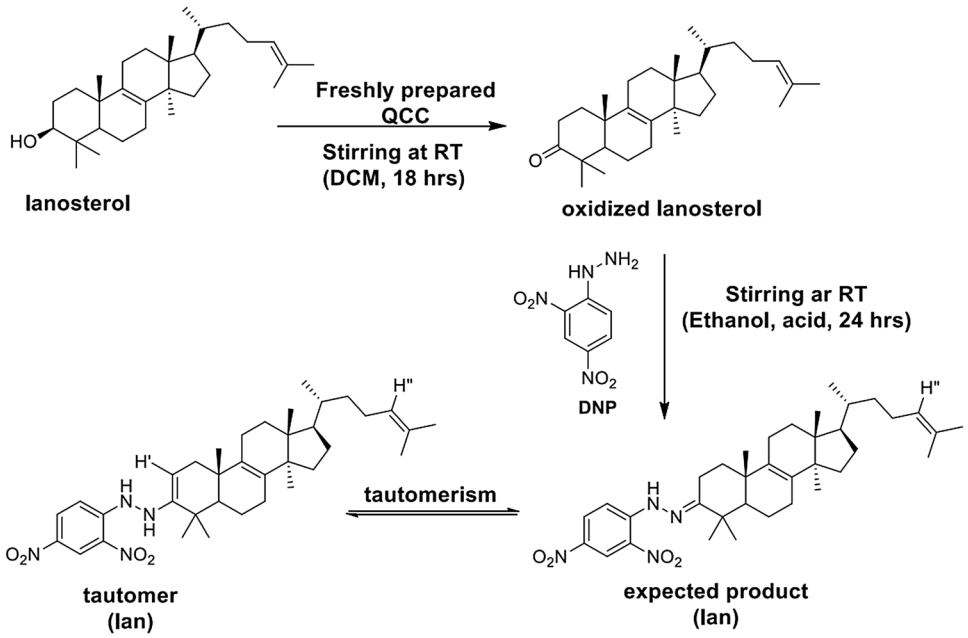

Fig. 10 Imine-enamine tautomerism of Compound A to show the occurrence of the second N–H group and the olefin hydrogen

Antimalarial chemo suppressive test

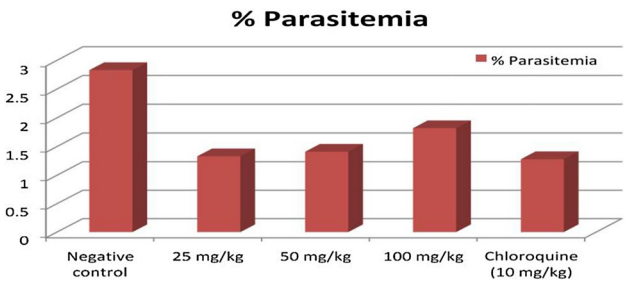

Fig. 11 Graphical representation of % Parasitemia

Figure 11 shows the percentage parasitemia of the compound at different doses, the lowest percentage parasitemia for LAN, 1.33% is observed at 25 mg/kg, which is comparable with that of the positive control chloroquine which is 1.28%. As it can be seen from Fig. 12, the survival rate observed in 25 mg/kg which has the optimal activity compared favourably with that of the reference drug (chloroquine) too. This shows the ability of the compound to sustain the host (murine model) after the dosing is stopped. The % chemo suppression, Fig. 12, of the test compound (LAN) at 25 mg/kg (53.13%) is comparable with that of the positive control (CQ) having % chemo suppression of 54.93%, while at 50 mg/kg, it is 50.41% and at 100 mg/kg, it is 35.67%. This shows that the compound compares favourably with the

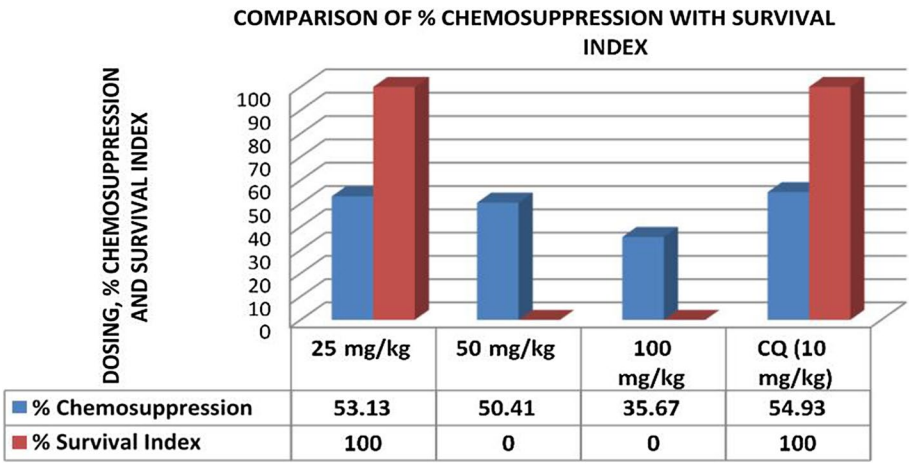

Fig. 12 Graphical Representation of % Chemosuppression with Survival Index

reference drug, chloroquine, at both 25 mg/kg and 50 mg/kg. The % Chemosuppression may be declining as the concentration of the dosage increases due to a form of toxicity (at high doses) which must have led to a reduction of survival index (eventual death of the test animals).

Conclusion

In this study, the 3D model of *P. falciparum* transketolase was constructed and made available for docking which gave insight into its interaction with the synthesized compound. The MD simulation also gave insight into the stability of the complex with the study ligands and revealed the advantage of the lanosterol derivative (LAN) binding Free Energy in terms of solvation energy contribution to the Total Free Energy of Binding. The oxidation of lanosterol with a mild oxidizing agent, QCC, also reduces the extent of rearrangement of pi-bond rearrangement (shift) and degradation, this was confirmed by the TLC and the spectroscopic data obtained for the DNP derivative (LAN). The in-vivo experiments corroborated the result from in-silico study, it showed that the dinitrophenylhydrazone derivative of oxidized lanosterol possesses antimalarial activity comparable to that of chloroquin at the lowest dose administered for LAN, 25 mg/kg and all the animals tested survived throughout the course of the experiment (at the 25 mg/kg dose for LAN) which shows that the compound is active as well as less toxic at low dosage, which is one of the goals set in drug design. 2,4-dinitrophenylhydrazone derivative of oxidized lanosterol can be exploited as lead in the antimalarial chemotherapy, the compound can be modified to get a better activity by replacing the DNP group with other groups and creating a variety of derivatives. The compound may also be subjected to an in-vitro enzyme inhibition assay for transketolase in a further study—which would require the cloning and overexpression of the transketolase to obtain sufficient amount for in-vitro study.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40203-021-00097-8.

Acknowledgements The authors would like to acknowledge the Berkman's Lab, Washington State University, USA for assisiting with all the NMR experiments in this study.

Funding The entire project was self-funded.

References

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1–2:19–25. https://doi.org/10.1016/j.softx.2015.06.001

Amato R, Pearson RD, Almagro-Garcia J, Amaratunga C, Lim P, Suon S, Sreng S, Drury E, Stalker J, Miotto O, Fairhurst RM, Kwiatkowski DP (2018) Origins of the current outbreak of multidrug-resistant malaria in southeast Asia: a retrospective genetic study. Lancet Infect Dis 18:337

Brenk R, Vetter SW, Boyce SE, Goodin DB, Shoichet BK (2006) Probing molecular docking at a charged model binding site. J Mol Biol 357:1449–1470

Bronowska AK (2011) Thermodynamics of ligand-protein interactions: implications for molecular design. Thermodyn Interact Stud Solids Liquids Gases. https://doi.org/10.5772/19447

Daranas AH, Shimizu H, Homans SW (2004) Thermodynamics of binding of Dgalactose and deoxy derivatives thereof to the L-arabinose-binding protein. J Am Chem Soc 126(38):11870–11876

Dhar DN, Singh AK (1977) Pyridinium Chlorochromate oxidation of some steroidal system. Zeitschrift Fur Naturforschung 32B:1476–1477

Dzubak P, Haajduch M, Vydra D, Hustova A, Kvasnica M, Biedermann D, Markova L, Urban M, Sarek J (2006) Pharmacological activities of natural triterpenoids and their therapeutic implications. Nat Prod Rep 23(3):394–411

Egan WJ, Merz KM, Baldwin JJ (2000) Prediction of drug absorption using multivariate statistics. J Med Chem 43:3867–3877. https://doi.org/10.1021/jm000292e

Emmons GT, Wilson WK, Shroepfer GJ (1989) 1H and 13C NMR assignments for lanostan-3B-ol Derivatives: revised assignments for lanosterol. Magn Reson Chem 27:1012–1024

Gilson MK, Zhou H-X (2007) Calculation of protein-ligand binding affinities. Annu Rev Biophys Biomol Struct 36:21–42

Jiao D, Golubkov PA, Darden TA, Ren P (2008) Calculation of protein-ligand binding free energy using a polarizable potential. Proc Nat Acad Sci USA 105:6290–6295

Kangas E, Tidor B (2001) Electrostatic complementarity at ligand binding sites: application to chorismate mutase. J Phys Chem B 105:880–888

Knight DJ, Peters W (1980) The antimalarial activity of N-benzyloxydihydrotriazines. I. The activity of clociguanil (BRL 50216) against rodent malaria, and studies on its mode of action. Ann Trop Med Parasitol 74:393–404

Kumari R, Kumar R (2014) Open Source Drug Discovery Consortium, Lynn A. g_mmpbsa–a GROMACS tool for high-throughput MM-PBSA calculations. J Chem Inf Model 54(7):1951–1962. https://doi.org/10.1021/ci500020m

Lorke D (1983) A new approach to tropical acute toxicity testing. Arch Toxicol 53:275–285

MacRaild CA, Daranas AH, Bronowska A, Homans SW (2007) Global changes in local protein dynamics reduce the entropic cost of carbohydrate binding in the arabinose-binding protein. J Mol Biol 368(3):822–832

Marsh J, Teichmann S (2011) Relative solvent accessible surface area predicts protein conformational changes upon binding. Structure 19:859–867. https://doi.org/10.1016/j.str.2011.03.010

Mobley DL, Dill KA (2009) Binding of Small-Molecule Ligands to Proteins: "What You See" Is Not Always "What You Get." Structure 17(4):489–498. https://doi.org/10.1016/j.str.2009.02.010

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 30(16):2785–2791. https://doi.org/10.1002/jcc.21256

Nasomjai P, Arpha K, Sodngam S, Brandt SD (2014) Potential antimalarial derivatives from astraodorol. Arch Pharmacal Res 37(12):1538–1545

Nazaruk J, Borym-Kluczyk M (2015) The role of triterpenes in the management of diabetes mellitus and its complications. Phytochem Rev 14(4):675–690

Nogueira CR, Lopes LMX (2011) Antiplasmodial natural products. Molecules 16(3):2146–2190

Olliaro PL, Yuthavong Y (1999) An overview of chemotherapeutic targets for antimalarial drug discovery. Pharmacol Ther 81:91–110

Pant R, Joshi A, Maiti P, Nand M, Pande V, Chandra S (2020) Identification of potential mycolyltransferase Ag85C inhibitors of Mycobacterium tuberculosis H37Rv via virtual High Throughput Screening and binding free energy studies. J Mol Graph Model. https://doi.org/10.1016/j.jmgm.2020.107584

Reddy MR, Erion MD (2001) Calculation of relative binding free energy differences for fructose 1,6-bisphosphatase inhibitors using the thermodynamic cycle perturbation approach. J Am Chem Soc 123:6246–6252

Ridley RG (2002) Medical need, scientific opportunity and the drive for antimalarial drugs. Nature 415:686–693

Rosenthal PJ (2001) Antimalarial chemotherapy: mechanisms of action, resistance, and new directions in drug discovery. Humana Press. Google Scholar, Totowa

Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815

Shoichet BK, Leach AR, Kuntz ID (1999) Ligand solvation in molecular docking. Proteins 34:4–16

Syme NR, Dennis C, Bronowska A, Paesen GC, Homans SW (2010) Comparison of entropic contributions to binding in a "hydrophilic" versus "hydrophobic" ligand−protein interaction. J Am Chem Soc 132(25):8682–8689. https://doi.org/10.1021/ja101362u

Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem 31:455–461

WHO Status report on artemisinin and ACT resistance (2017) WHO/HTM/GMP/2017.9

World Health Organization (2010) World Malaria Report. http://www.who.int/malaria/publications/worldmalariareport/en/

Young T, Abel R, Kim B, Berne BJ, Friesner RA (2007) Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. Proc Natl Acad Sci 104(3):808–813. https://doi.org/10.1073/pnas.0610202104