Molecular Diversity
https://doi.org/10.1007/s11030-022-10594-3

ORIGINAL ARTICLE

In silico investigation of falcipain-2 inhibition by hybrid benzimidazole-thiosemicarbazone antiplasmodial agents: A molecular docking, molecular dynamics simulation, and kinetics study

Nyiang Kennet Nkungli[1] · Aymard Didier Tamafo Fouegue[2] · Stanley Numbonui Tasheh[1,3] · Fritzgerald Kogge Bine[3] · Abrar Ul Hassan[4] · Julius Numbonui Ghogomu[1,3]

Received: 2 October 2022 / Accepted: 20 December 2022

Abstract

The emergence of artemisinin-resistant variants of *Plasmodium falciparum* necessitates the urgent search for novel antimalarial drugs. In this regard, an in silico study to screen antimalarial drug candidates from a series of benzimidazole-thiosemicarbazone hybrid molecules with interesting antiplasmodial properties and explore their falcipain-2 (FP2) inhibitory potentials has been undertaken herein. FP2 is a key cysteine protease that degrades hemoglobin in *Plasmodium falciparum* and is an important biomolecular target in the development of antimalarial drugs. Pharmacokinetic properties, ADMET profiles, MM/GBSA-based binding free energies, reaction mechanisms, and associated barrier heights have been investigated. DFT, molecular dynamics simulation, molecular docking, and ONIOM methods were used. From the results obtained, four ^{4}N-substituted derivatives of the hybrid molecule (*E*)-2-(1-(5-chloro-1*H*-benzo[*d*]imidazol-2-yl)ethylidene)hydrazine-1-carbothioamide (1A) denoted 1B, 1C, 1D, and 1E are drug-like and promising inhibitors of FP2, exhibiting remarkably small inhibitory constants ($5.94\times10^{-14}-2.59\times10^{-04}$ *n*M) and favorable binding free energies (−30.32 to −17.17 kcal/mol). Moreover, the ONIOM results have revealed that 1B and possibly 1C and 1D may act as covalent inhibitors of FP2. The rate-determining step of the thermodynamically favorable covalent binding mechanism occurs across a surmountable barrier height of 24.18 kcal/mol in water and 28.42 kcal/mol in diethyl ether. Our findings are useful for further experimental investigations on the antimalarial activities of the hybrid molecules studied.

✉ Nyiang Kennet Nkungli
nyiangken@gmail.com

[1] Department of Chemistry, Faculty of Science, The University of Bamenda, Bambili, P. O. Box 39, Bamenda, Cameroon

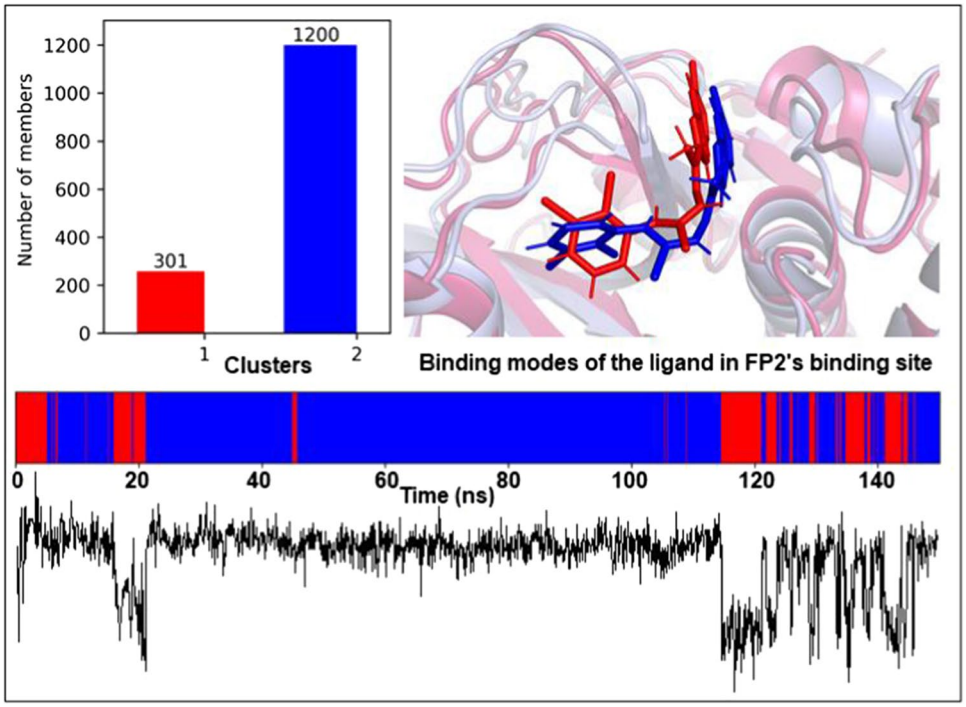
[2] Department of Chemistry, Higher Teacher Training College Bertoua, University of Bertoua, P.O. Box 652, Bertoua, Cameroon

[3] Department of Chemistry, Faculty of Science, University of Dschang, P. O. Box 67, Dschang, Cameroon

[4] Department of Chemistry, University of Gujrat, Gujrat 54400, PK, Pakistan

Published online: 09 January 2023

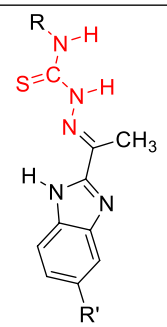
Graphical abstract



Keywords Falcipain-2 · Benzimidazole-thiosemicarbazone · Kinetics · Docking · ADMET

Introduction

Malaria is a life-threatening parasitic disease caused by intracellular protozoans of the genus *Plasmodium.* Among the etiological agents of the disease, *Plasmodium falciparum* is the most prevalent [1, 2]. The disease is transmitted to humans through the bites of infected female *Anopheles* mosquitoes. Its pathogenesis is characterized by high morbidity and mortality, most especially in children under five years of age [3–5]. According to the 2021 world malaria report [6], there were an estimated 241 million malaria cases and 627,000 malaria deaths across the globe in 2020, and about 47,000 of the death cases were linked to service disruptions during the COVID-19 pandemic. According to that report, most of the malaria cases and deaths occurred in Africa. In fact, the severity and lethality of malaria are almost entirely restricted to the disease-endemic regions, including Sub-Saharan Africa, South-East Asia, Eastern Mediterranean, Western Pacific, and Latin America [6, 7]. During the intraerythrocytic stage of their life cycle, malaria parasites digest the host hemoglobin to obtain the amino acids needed for protein synthesis [1, 8]. Unfortunately for the parasites, free heme that is highly cytotoxic is released as a by-product of hemoglobin metabolism. However, the parasites evade the toxicity of free heme by converting it into a nontoxic insoluble crystalline substance called hemozoin, which is structurally similar to β-hematin [1, 4, 9]. In an effort to eradicate this devastating disease, several antimalarial drugs comprising quinolines, atovaquones, antifolates, and artemisinin-based combination therapies (ACTs) have been developed in the past years [2]. Their modes of action include the inhibition of free heme biocrystallization into hemozoin, disruption of mitochondrial function, inhibition of folate metabolism, and the prevention of hemoglobin degradation in plasmodia [10, 11]. However, the emergence of antimalarial drug-resistant variants of plasmodium species with reduced sensitivity toward the ACTs (first-line



Hybrid	R	R'
1A	4-FC_6H_4	Cl
1B	2,3-ClC_6H_3	Cl
1C	4-IC_6H_4	Cl
1D	COC_6H_5	Cl
1E	4-MeC_6H_4	Cl
2A	C_6H_{11}	H
2B	iso-Propyl	H
2C	4-$MeOC_6H_4$	H

Fig. 1 Molecular structures of the benzimidazole-thiosemicarbazone hybrids studied. The TSC moiety is shown in red

antimalarial drug regimens) has been reported, necessitating the development of new and more effective antimalarial drugs [2, 4, 5, 7].

As part of the ongoing search for more reliable antimalarial therapeutics, medicinal chemists have directed considerable research attention toward thiosemicarbazones (TSCs) due to their wide range of biological activities, including antimalarial properties. Accordingly, several TSCs have been synthesized in the past decades and found to exhibit interesting antiplasmodial properties [12–16]. However, there are hardly any literature reports of TSC-based antimalarial lead compounds with known biomolecular targets and precise mechanisms of action [17]. Against this backdrop, Chellan and co-workers [18] investigated the ability of 3,4-dichloroacetophenone thiosemicarbazone to inhibit hemozoin formation in *Plasmodium falciparum*. Their results revealed that the molecule's antimalarial properties cannot be attributed to the prevention of free heme detoxification in plasmodia, which is somewhat in disagreement with our previous findings on a similar TSC, 4-methoxyacetophenone thiosemicarbazone [19], which demonstrated the thermochemical viability of its interaction with free heme. Chellan and co-workers further speculated that the antimalarial mechanism of TSCs may be dependent on some TSC-induced processes that disrupt hemoglobin digestion by plasmodia, such as inhibition of key cysteine proteases in the food vacuole of the parasites [18]. To aid the development of antimalarial drugs containing the TSC pharmacophore, lead compounds with a clear mechanism of action are imperative. The development of such compounds has been immensely facilitated by in silico methods, which do not necessitate the synthesis of every compound studied. Moreover, many sophisticated computational chemistry programs that can be used to determine reaction mechanisms, pharmacokinetic profiles, and pharmacodynamic properties of drug candidates have been developed in the recent decades.

Among the TSC-based compounds exhibiting significant antiplasmodial properties are those synthesized by Divatia and co-workers in 2019 [14] and earlier in 2013 [20]. They demonstrated experimentally that ^{4}N-substitution in the TSC moiety of two benzimidazole-thiosemicarbazone hybrid molecules, namely (*E*)-2-(1-(5-chloro-1*H*-benzo[*d*]imidazol-2-yl)ethylidene)hydrazine-1-carbothioamide (1A) and (*E*)-2-(1-(1*H*-benzo[*d*]imidazol-2-yl)ethylidene)hydrazine-1-carbothioamide (2A) has remarkable effects on their antiplasmodial activities [14]. According to their results, certain ^{4}N-substituents significantly improved the antiplasmodial activities of these hybrid molecules, whereas other substituents produced an opposite effect. Unfortunately, there is no literature report on the antiplasmodial mode of action, pharmacokinetic properties, and toxicity of these hybrid molecules and their ^{4}N-substituted derivatives, as far as we know. Therefore, in silico studies to complement the experimental findings on these molecules are imperative. A plausible explanation of the impact of ^{4}N-substitution on the antiplasmodial activities of the molecules is needed, as this may constitute a means of modulating their antiplasmodial potencies and/or efficacies. It is noteworthy that undesirable pharmacokinetic properties and toxicity of drug candidates are among the main causes of clinical trial failures [21–23]. It is therefore worthwhile to carry out early toxicological and pharmacokinetic analyses of drug candidates in the drug discovery and development process.

The main objective of this work was to investigate the falcipain-2 inhibitory potentials and mechanisms, pharmacokinetic/ADMET profiles, and toxicity of some ^{4}N-substituted derivatives of the benzimidazole-thiosemicarbazone hybrid molecules 1A and 2A as depicted in Fig. 1 using computational chemistry methods such as the density functional theory (DFT), molecular dynamics (MD) simulation, molecular docking and Our own N-Layered Integrated Molecular Orbital and Molecular Mechanics (ONIOM). Among several known antimalarial drug targets [17], falcipain-2 (FP2) was chosen because the cysteine residue in its active site may react covalently with the TSC moiety in the hybrid molecules. Moreover, it is known that FP2 can be inactivated by TSCs either reversibly or irreversibly [24]. In addition, FP2 is one of the key proteases involved in different processes of the erythrocytic life cycle of plasmodium parasites, including the degradation of the host hemoglobin and the invasion/rupture of the erythrocytes [5, 9, 17, 25]. Indeed, FP2 has become an important biomolecular target for the development of novel antimalarials [2, 8, 17, 26]. To achieve the above goal, the benzimidazole-thiosemicarbazone hybrids studied were first subjected to pharmacokinetic and toxicological screening of the most drug-like molecules. This was followed by molecular docking of the adequately drug-like molecules (ligands) in the binding site of FP2 (receptor) to generate ligand−FP2 complexes with the most probable binding modes of the ligands. Then, MD simulations to assess the stability of the ligand−FP2 complexes and to generate conformational structures of the complexes useful for binding free energy calculation through the Molecular Mechanics Generalized Born Surface Area (MM/GBSA) method were next performed. Finally, mechanistic and kinetics studies on the covalent reaction between FP2 and the most biologically active molecule among the drug-like benzimidazole-thiosemicarbazone hybrids was performed using the ONIOM method.

The TSC moiety in the benzimidazole-thiosemicarbazone hybrids is a reactive warhead endowed with an electrophilic center (the thiocarbonyl carbon), which is a probable site of attack by the active site cysteine thiol/thiolate group in FP2 [27]. Such a nucleophilic attack on the electrophilic center can lead to reversible or irreversible covalent modification of the cysteine residue in the enzyme's binding site [18, 28].

Generally, covalent inhibitors of falcipains rely on a chemical warhead to accomplish their inhibitory activities [8, 29]. The catalytic/inhibitory activities of most cysteine proteases begin with the activation of the active site cysteine residue by a proximal histidine residue leading to the formation of a thiolate–imidazolium ion pair. This is usually followed by covalent bond formation between the thiolate sulfur of the activated cysteine residue and an electrophilic center in the substrate/inhibitor molecule [28, 30].

Computational details and relevant theory

Preparation of the initial structures of the receptor and the ligands

The crystallographic structure of FP2 (PDB code: 3bpf) was downloaded from the RCSB protein data bank and chain A of the protein was extracted. The native ligand, E64, was eliminated along with all crystallographic solvent species. The missing amino acid residues were added with the aid of the CHARMM-GUI pdb manipulator [31]. The protein was next protonated at pH 7.4 using the H++ webserver [32], taking account of the pKa of the titratable residues. The guess geometries of the ligands were constructed with Avogadro 1.2 [33] and optimized using the composite DFT-based r^2SCAN-3c [34] method implemented in the ORCA 5.0.3 program package [35]. Harmonic frequency calculations at the same level of theory confirmed all optimized structures as minima on the potential energy surface.

Drug-likeness and ADMET screening of the ligands

It is well known that the pharmacokinetic (PK) properties of drug candidates, most especially absorption, distribution, metabolism, excretion, and toxicity (ADMET), are crucial during the drug discovery and development process [36]. Accordingly, the benzimidazole-thiosemicarbazone hybrid molecules under study were first subjected to critical drug-likeness screening based on computed physicochemical and ADMET-related properties via the SwissADME [37] and ADMETlab 2.0 [22] webservers. The SMILES strings of the ligands used as molecular data input for both servers are provided in supplementary Table S1. Indeed, undesirable PK properties of drug candidates are a major cause of failure during clinical trials. Therefore, preclinical evaluation of the PK properties of potential drugs is necessary [38].

Molecular docking

The docking abilities of two programs from the Autodock suite, namely Autodock vina 1.2 (ADVina) [39, 40] and AutodockFR (ADFR) [41], with FP2 as receptor were first assessed through self-docking (re-docking) of the native ligand E64. The best program was then used for molecular docking of the sufficiently drug-like and safer ligands. The PDBQT files of both the ligands and the receptor were prepared using the python scripts *prepare_ligand.py* and *prepare_receptor.py*, respectively, provided in the ADFR program suite. To prepare the PDBQT files, polar hydrogen atoms were added and the non-polar ones were merged with the heavy atoms. Moreover, Gasteiger charges were computed and the ligand torsional tree root and rotatable bonds were defined. In all ADFR docking calculations, 60 Lamarckian genetic algorithm (LGA) runs were performed alongside 2,500,000 energy evaluations and 27,000 generations per LGA run, to search for ligand binding poses and predict their binding affinities. In the case of ADVina, an exhaustiveness of 32 was used together with 2,500,000 energy evaluations. In both cases, a search grid box with center $x = -57.042$ Å, $y = -1.174$ Å, $z = -15.913$ Å (center of the native ligand) and size 22.5 Å × 22.5 Å × 22.5 Å was used in all docking computations. All post-docking processing was performed with the programs: Biovia Discovery Studio v.21.1.0 [42], LigPlot$^+$ v2.2 [43] and PyMol v2.6 [44, 45].

Molecular dynamics simulation

All molecular dynamics (MD) simulations of the ligand-FP2 complexes in explicit water solvent were started from the atomic coordinates of the complexes obtained from molecular docking. The GROMACS-2022 software package [46] was used together with the Amber ff99SBildn force-field [47] for the protein, the generalized Amber force-field (GAFF2) [48] for non-standard residues, and the TIP3P solvent model [49] for the water molecules. The ff99SBildn force-field was preferred over ff99SB due to its improved amino acid side-chain torsion potentials (for Ile, Leu, Asp and Asn residues) compared to the latter. The second generation of GAFF (GAFF2) was chosen because of its updated bonded and non-bonded parameters, and its high-quality charge model [50]. The classical 3-point rigid water model TIP3P was used due to its popularity and computational efficiency. The GAFF2-based ligand topologies and parameters were generated with the aid of the Antechamber module [51] of AmberTools2021 [52]. The newly generated topologies and parameters were converted into GROMACS format using the ACPYPE software [53]. The partial charges of the ligand atoms were derived using the AM1-BCC method of the Antechamber program. The default parameters of the ff99SBildn force-field were used at physiological pH 7.4 for the termini, disulfide bridges and protonation states of the amino acid residues.

In the MD simulations, each ligand-FP2 complex was placed and centered in a periodic dodecahedral box at least 1.2 nm from the box edge. The system was neutralized and a salt concentration of 0.15 mol/L was obtained by adding Na^+ and Cl^- ions. The structure was next subjected to steepest descent energy minimization in 5000 steps in order to remove steric clashes and irregular geometry. Thereafter, a two-step equilibration run was performed to equilibrate the solvent around the protein. A short 250 ps position-restrained equilibration under the canonical ensemble (NVT) in which the system was coupled to a temperature of 310.15 K was followed by a 500 ps run under the isothermal-isobaric (NPT) ensemble at a reference pressure of 1 bar. The temperature and pressure couplings were accomplished by means of the Nosé-Hoover thermostat and the Parrinello-Rahman barostat, respectively. A force constant of 1000 kJ/mol in the LINCS algorithm was used to restrain all heavy-atom positions so as to preserve the original protein structure during equilibration. The LINCS algorithm was also used to constrain all covalent bonds to hydrogen atoms. After successful minimization and equilibration, the position restrain potential was removed and 5, 10, and 15 ns independent production runs per ligand-FP2 complex were performed under the NPT ensemble maintained at 310.15 K temperature and 1 bar pressure. Coordinate sampling was performed every 10, 20, and 30 ps, respectively. Furthermore, a 150-ns production run was performed on the ligand-FP2 complex of the most biologically active drug-like ligand. In the latter case, coordinate sampling was performed every 100 ps. All simulations were performed at 2 fs time-step using periodic boundary conditions (PBC) to eliminate the effects of walls. The long-range electrostatic interactions were treated by the Particle-Mesh Ewald (PME) algorithm. The verlet cut-off distance of 1.0 nm was chosen for both van der Waals and short-term electrostatic interactions.

The 150 ns trajectory was analyzed using various GROMACS tools to obtain the Root-Mean-Square Deviation (RMSD), Root-Mean-Square Fluctuation (RMSF), radius of gyration, hydrogen bond, and protein secondary structure plots of the ligand-FP2 complex concerned. All trajectory analyses were performed after aligning the protein with respect to the backbone atoms to eliminate translations and rotations due to diffusion during the simulation.

MM/GBSA free energy calculations

Docking programs use highly simplified scoring functions to predict protein–ligand binding affinities, hence their ability to determine sufficiently accurate binding strengths is limited [54–57]. To obtain more accurate FP2-binding affinities of the ligands studied, their binding free energies were calculated using the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method as implemented in the gmx_MMPBSA-1.5.1 program [58, 59]. Here, end-state free energy calculations were performed based on the single trajectory (ST) approach, using the Amber ff99SBildn and GAFF2 force fields for the protein and ligands, respectively. 500 snapshots obtained from the trajectory of each of the 5, 10 and 15 ns MD simulations discussed earlier were used in the binding free energy calculations.

In the MM/GBSA approach, the free energy change (ΔG_{bind}) arising from receptor/ligand binding to form a receptor–ligand complex is given by:

$$\Delta G_{bind} = G_{complex} - \left(G_{receptor} + G_{ligand}\right) \quad (1)$$

where $G_{complex}$ is the free energy of the complex, $G_{receptor}$ is the free energy of the receptor, and G_{ligand} is the free energy of the ligand. ΔG_{bind} is also calculated as;

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (2)$$

where

$$\Delta E_{MM} = \Delta E_{covalent} + \Delta E_{ele} + \Delta E_{vdW} \quad (3)$$

$$\Delta E_{covalent} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{torsion} \quad (4)$$

$$\Delta G_{solv} = \Delta G_{polar} + \Delta G_{non-polar} = \Delta G_{GB} + \Delta G_{SA} \quad (5)$$

In the above equations, ΔE_{MM}, ΔG_{solv} and $-T\Delta S$ represent changes in gas-phase molecular mechanical energy, solvation free energy, and conformational entropy, respectively, upon protein–ligand binding. In this work, these changes were computed by means of ensemble averaging over the 500 snapshots sampled from the various MD trajectories. ΔE_{MM} comprises the bond stretching energy (E_{bond}), angle bending energy (E_{angle}), torsional energy ($E_{torsion}$), electrostatic energy (E_{ele}), and van der Waals energy (E_{vdW}). ΔG_{solv} is the sum of two energy contributions namely electrostatic or polar solvation (ΔG_{polar}) and non-electrostatic or non-polar solvation ($\Delta G_{non-polar}$). Herein, ΔG_{polar} was obtained based on the generalized born continuum solvation model and $\Delta G_{non-polar}$ was estimated from the solvent accessible surface area (SASA). The entropic contribution $-T\Delta S$ is traditionally estimated by normal modes analysis (NM) on a few conformational snapshots generated by MD simulation, but the method is computationally expensive [58, 60]. Recently, cheaper alternatives such as the simplified versions of NM and interaction entropy (IE) [61] have been introduced [58, 62]. The IE method was used in this work, although the entropic contribution is often neglected due to the high uncertainty and computational cost associated with its calculation [54, 62, 63].

Table 1 Physicochemical properties and drug-likeness assessment of the hybrids based on SwissADME predictions

Molecule	MW (g/mol)	HBA	HBD	NRB	TPSA (Å^2)	LogP	LogS (ESOL)	LogS (Ali)	LogS (Silicos-IT)	Lipinski's rule	Ghose's rule	Veber's rule	Egan's rule	Muegge's rule
1A	361.82	3	3	5	97.19	3.87	− 4.74	− 5.75	− 7.01	Yes	Yes	Yes	Yes	Yes
1B	412.72	2	3	5	97.19	4.62	− 5.76	− 6.95	− 7.92	Yes	Yes	Yes	Yes	No
1C	469.73	2	3	5	97.19	4.21	− 5.76	− 6.32	− 7.58	Yes	Yes	Yes	Yes	Yes
1D	371.84	3	3	6	114.26	3.31	− 4.93	− 6.45	− 6.67	Yes	Yes	Yes	Yes	Yes
1E	357.86	2	3	5	97.19	3.89	− 4.88	− 6.02	− 7.12	Yes	Yes	Yes	Yes	Yes
2A	315.44	2	3	5	97.19	3.16	− 3.99	− 5.26	− 5.07	Yes	Yes	Yes	Yes	Yes
2B	275.37	2	3	5	97.19	2.39	− 3.16	− 4.21	− 4.47	Yes	Yes	Yes	Yes	Yes
2C	339.41	3	3	6	106.42	3.01	− 4.06	− 5.16	− 6.26	Yes	Yes	Yes	Yes	Yes

LogP: The consensus value of iLogP, XLogP3, WLogP, MLogP and LogP (SILICOS-IT) is reported

LogS scale: Insoluble < − 10 < Poorly soluble < − 6 < Moderately soluble < − 4 < Soluble < − 2 < Very soluble < 0 < Highly soluble

Muegge's rule: 200 ≤ MW ≤ 600, − 2 ≤ XLogP ≤ 5, TPSA ≤ 150; # of rings ≤ 7, # of carbons > 4, # of heteroatoms > 1, NRB ≤ 15, HBA ≤ 10 and HBA ≤ 5

Ghose's rule: 160 ≤ MW ≤ 480, − 0.4 ≤ WLogP ≤ 5.6, 40 ≤ Molar Refractivity (MR) ≤ 130 and 20 ≤ Atoms ≤ 70

Veber's rule: NRB ≤ 10 and TPSA ≤ 140

Egan's rule: WLogP ≤ 5.88 and TPSA ≤ 131.6

Trajectory clustering and ONIOM calculations

Because covalent interactions are best treated quantum mechanically, a multiscale quantum chemical method comprising the DFT and the Grimme's extended Tight-Binding (xTB) method was used to study the covalent interaction between the active site cysteine residue (Cys42) of FP2 and the benzimidazole-thiosemicarbazone hybrids under study. The hybrids were represented by 1B because it possesses a favorable docking score and is the most biologically active molecule among the drug-like and safer hybrids. It must be pointed out that the covalent binding study was preceded by noncovalent docking and MD simulation because the initial contacts or collisions between FP2 and each ligand must initially form an encounter complex before a covalent bond is formed [64, 65]. The most representative configuration of the 1B−FP2 complex was first generated by clustering its 150 ns MD simulation trajectory with the TTClust-4.10.0 [66] program using the RMSD approach. While the backbone atoms of the protein were used for alignment, the ligand atoms were used for the computation of RMSD values. Apart from the atom selection for RMSD calculation, the default TTClust parameters were used.

The geometry of the dominant configuration of the 1B−PF2 complex was first energy-minimized with the L-BFGS minimizer in ORCA 5.0.3 using parameters of the Amber force-field earlier utilized in the MD simulations. Thereafter, a quantum mechanical cluster model comprising the ligand 1B, the catalytic dyad of FP2 (Cys42 and His174), and other amino acid residues with atoms within 4.0 Å of 1B (Gln36, Gly40, Trp43, Cys80, Asn81, Gly82, Gly83, Leu84, Ser149, Val152, Ser153, Ala157, Leu172, Asn173, Ala175, and Trp206) in the 1B−PF2 complex was constructed. The model was used to study the covalent reaction between 1B and FP2 by means of the two-layered ONIOM method in ORCA 5.0.3. The ONIOM calculations were aimed at determining the minimum energy path (MEP) and the associated stationary points of all reaction pathways studied. In these calculations, the atoms of the directly reacting moieties (fragments of the catalytic dyad and the chemical warhead) were represented with the DFT-based composite electronic structure method r^{2}SCAN-3c, while the rest of the system was described by the xTB method via the GFN2-xTB Hamiltonian [67]. The initial structures of the reactant, products, and the thiolate-imidazolium ion pair, gotten from relaxed geometry scans along the various reaction coordinates, were further optimized without constraints of any kind. Vibrational frequency calculations to confirm the above-mentioned states as minima on the potential energy surface (PES) were also performed. The MEP of each reaction pathway was next determined using the climbing image Nudged Elastic Band method followed by Transition State optimization (NEB-TS) [68]. Each transition state had one and only

Table 2 Physicochemical properties and drug-likeness assessment of the hybrids based on ADMETLab 2.0 results

Molecule	MW (g/mol)	HBA	HBD	NRB	TPSA (Å^2)	LogP	LogS	LogD	Lipinski's rule	Pfizer's rule	GSK's rule	GT rule
1A	361.06	5	3	5	65.1	4.64	− 6.001	4.202	Yes	No	No	Yes
1B	410.99	5	3	5	65.1	5.52	− 6.392	4.563	Yes	No	No	Yes
1C	468.96	5	3	5	65.1	5.51	− 6.365	4.084	Yes	No	No	Yes
1D	371.06	6	3	6	82.17	4.12	− 5.342	4.018	Yes	Yes	No	Yes
1E	357.08	5	3	5	65.1	4.92	− 6.141	4.420	Yes	No	No	Yes
2A	315.15	5	3	5	65.1	4.09	− 4.505	4.067	Yes	No	No	Yes
2B	275.12	5	3	5	65.1	3.17	− 4.001	3.577	Yes	No	Yes	Yes
2C	339.12	6	3	6	74.33	3.95	− 5.407	3.854	Yes	No	Yes	Yes

LogS scale: Optimal values range from − 4 to ~ 0.5

Pfizer's rule: LogP > 3 and TPSA < 75

GlaxoSmithKline's (GSK's) rule: MW ≤ 400 and LogP ≤ 4

Golden Triangle (GT) rule: 200 ≤ MW ≤ 50 and LogD ≤ 5

one imaginary frequency, while the other stationary points had none. More accurate electronic energies of the stationary points were calculated at two levels of theory obtained by replacing the r^{2}SCAN-3c method with the functional and basis set combinations: M06-2X/def2-TZVPPD [69, 70] and DSD-PBEB95/def2-TZVPPD [70, 71] in conjunction with the Grimme D3 dispersion correction [72] with zero and Becke-Johnson damping [73, 74], respectively. Finally, the Shermo code [75] was used to perform thermochemical analyses at 310.15 K and 1 atm using the more accurate electronic energies, but with vibrational frequencies calculated using the r^{2}SCAN-3c method. Further details on the ONIOM and NEB-TS methods are provided in Supplementary Information.

Results and discussion

Drug-likeness and ADMET profiles of the benzimidazole-thiosemicarbazone hybrids

Based on the fact that a drug-like molecule is that which possesses sufficiently acceptable pharmacokinetic (PK) properties and a good ADMET profile [76, 77], the physicochemical properties of the benzimidazole-thiosemicarbazone hybrids under study (listed in Tables 1 and 2) were computed. The PK parameters important for good in vivo activity of a drug include its systemic exposure (dependent on absorption, distribution, metabolism, and excretion), its bioavailability (dependent on absorption and metabolism), and its elimination from the body (dependent on metabolism, distribution, and excretion) [78]. Some drug-likeness rules/filters based on physicochemical properties are widely used to filter out undesirable drug candidates [22, 36, 38]. Among these, the Lipinski's rule-of-five that is based on physicochemical properties used to predict the oral bioavailability of drugs is the most famous. According to the Lipinski's rule, human oral absorption is unlikely if a molecule violates two or more of the following guidelines: molecular weight (MW) ≤ 500 Da, logarithm of *n*-octanol/water partition coefficient (LogP) < 5, number of hydrogen bond acceptors (HBA) ≤ 10 , and number of hydrogen bond donors (HBD) ≤ 5 [79]. An additional guideline for number of rotatable bonds (NRB ≤ 10) is often added. Other drug-likeness rules include the Veber's rule, Ghose's rule, Egan's rule, Muegge's rule, Pfizer's rule, GlaxoSmithKline's (GSK's) rule, and the Golden Triangle (GT) rule (details are provided in the footnotes of Tables 1 and 2).

Inspection of Tables 1 and 2 reveals that the hybrids satisfy the Lipinski's, Ghose's, Veber's, Egan's, Muegge's, and the golden triangle rules, which is indicative of their likelihood of good oral bioavailability, especially good oral absorption and gastrointestinal permeability. This is supported by their moderate aqueous solubility (expressed by the LogS values), which can aid their transport across the intestinal epithelial cell lining via paracellular absorption. Accordingly, the physicochemical properties of these compounds are suitable enough to ensure desirable PK properties in the human body.

From the foregoing results, the benzimidazole-thiosemicarbazone hybrids studied can be regarded as drug-like molecules that may suffer minimal attrition rates during clinical trials. However, nearly all of the ^{4}*N*-substituted hybrids fail to satisfy the Pfizer's and GSK's rules, suggesting some deficiencies in their drug-likeness behavior.

Besides the physicochemical properties, some ADMET-related properties were computed. These include gastrointestinal (GI) absorption or human intestinal absorption (HIA), blood–brain barrier (BBB) penetration, volume of distribution (VD) at steady state, bioavailability ($F_{20\%}$ and $F_{30\%}$), plasma protein binding (PPB), the fraction unbound (Fu) in plasma, the human colon carcinoma (Caco-2) cell membrane

Table 3 ADMET-related properties of the hybrid molecules studied, as predicted by ADMETLab 2.0

Molecule	1A	1B	1C	1D	1E	2A	2B	2C
Absorption								
Caco-2 Permeability (Log unit)	−5.714	−5.664	−5.76	−5.914	−5.693	−5.297	−5.604	−5.856
MDCK Permeability (cm/s)	1.6×10^{-5}	2.0×10^{-5}	1.5×10^{-5}	1.6×10^{-5}	1.3×10^{-5}	1.4×10^{-05}	6.3×10^{-06}	7.7×10^{-6}
P-gp inhibitor	− − −	− − −	− − −	− − −	− − −	− − −	− − −	− − −
P-gp substrate	− − −	− − −	− − −	− − −	− − −	− − −	− − −	− − −
HIA	− − −	− − −	− − −	− − −	− − −	− − −	− − −	− − −
F20%	− − −	− − −	− − −	− − −	− − −	−	− − −	− − −
F30%	− −	+ +	− −	− − −	− − −	+ +	− −	−
Distribution								
PPB (%)	99.90	100	100	99.39	99.78	97.16	96.52	98.61
VD (L/kg)	0.963	1.41	0.621	0.345	0.924	1.478	1.339	0.669
BBB Penetration	−	− − −	− − −	+ +	− −	+ + +	+ +	+
Fu (%)	0.89	0.92	0.62	0.90	1.00	1.10	2.80	1.32
Metabolism								
CYP2 inhibitor	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +
CYP2 substrate	+ +	+ +	+ +	+	+ + +	+ + +	+ + +	+ + +
CYP2C19 inhibitor	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ +	+ + +
CYP2C19 substrate	− − −	− − −	− − −	− − −	− −	− − −	− −	− −
CYP2C9 inhibitor	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ +	+ + +
CYP2C9 substrate	+ +	+ +	+ +	+ +	+ +	+ + +	+ +	+ + +
CYP2C6 inhibitor	+	+ +	+	− − −	–	− −	− − −	− −
CYP2C6 substrate	+ +	+ +	+ +	–	+ +	+ +	+ +	+ + +
CYP3A4 inhibitor	+	+ +	+	+	+ +	+ + +	+	+ +
CYP3A4 substrate	− −	–	–	− −	+	− −	–	–
Excretion								
Clearance (ml/min/kg)	2.632	2.619	1.584	1.242	2.504	2.916	4.172	3.937
Half-life (h)	0.612	0.594	0.504	0.831	0.769	0.855	0.921	0.890
Toxicity								
hERG Blockers	− − −	− − −	− − −	− − −	− − −	− − −	− − −	− − −
H-HT	+ +	+	–	+	+	+	+	+ +
DILI	+ + +	+ + +	+ + +	+ + +	+ + +	+ +	+ +	+ + +
AMES Toxicity	–	− − −	− − −	− −	–	− −	− − −	+
ROA Toxicity	+ + +	+	+	+ +	+ + +	+ + +	+ + +	+ + +
Skin Sensitization	− − −	− − −	− − −	− − −	− − −	− −	− − −	− − −
Carcinogenicity	+ +	–	+	+	+	+ +	+ +	+ +
Eye Corrosion	− − −	− − −	− − −	− − −	− − −	− − −	− − −	− − −
Eye Irritation	− −	–	− −	− − −	–	− − −	− − −	–
Respiratory Toxicity	+ + +	+ +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +

The prediction probability values are replaced by the symbols: –(0–0.1), –(0.1–0.3), –(0.3–0.5), +(0.5–0.7), + +(0.7–0.9) and + + +(0.9–1.0), where + + +or + +indicates a more likely toxic or defective molecule, − − − or − −indicates a nontoxic or appropriate molecule, and + or −presents inconclusive information that warrants further assessment [22]

permeability, the madin-darby canine kidney (MDCK) cell membrane permeability, P-glycoprotein (P-gp) substrates and inhibitors, cytochrome P450 substrates and inhibitors, drug-induced liver injury (DILI), rat oral acute (ROA) toxicity, human hepatotoxicity (H-HT), the human ether-a-go-go-related gene (hERG) inhibition, and AMES toxicity [21, 38, 39, 41]. The calculated values of these properties, listed in Tables 3 and S2, were used to determine whether the hybrid molecules studied can be adequately absorbed into the body from their sites of administration, effectively distributed to their targets, metabolized in a way that does not quickly annihilate their therapeutic effects, properly eliminated from the body, and are of acceptable toxicity [22, 80].

It is evident from Tables 3 and S2 that the hybrids exhibit desirable intestinal absorption and permeability. Nearly all of the compounds show high Caco-2 cell permeability (greater than −5.15 Log unit) and high MDCK cell permeability (with nearly all values greater 20×10^{-6} cm/s). This suggests that their high intestinal absorption (revealed by high GI and HIA) can be mainly through passive transcellular transport across the epithelial cell lining. The compounds are potential inhibitors of P-gp (an efflux transporter of xenobiotics and drugs), which can lead to increased bioavailability as suggested by $F_{20\%}$ and somewhat by $F_{30\%}$. However, this may also lead to interaction with other drugs transported by the P-gp efflux pump. According to the drug distribution parameters, all of the hybrids are liable to high plasma protein binding (with PPB > 90% in all cases), which can block access to their therapeutic targets. However, their predicted VD values are within the optimal range of 0.04−20 l/kg, suggesting that a significant amount of the compounds can be contained in the tissues rather than in plasma. In all cases, the predicted fraction unbound to plasma proteins is low (Fu < 5%). Note that Fu affects the volume of distribution and potency because only unbound (free) drugs can diffuse across cell membranes from plasma to the tissues [77, 81].

In the case of drug metabolism, most of the molecules are potential inhibitors of five cytochrome P450 enzymes (CYP2, CYP2C9, CYP2C19, CYP2C6 and CYP3A4), which is likely to improve their metabolic stability. Inhibition of these enzymes can result in increased bioavailability, but may lead to drug-drug interactions when co-administered with prodrugs that require these P450 enzymes to be converted into the active metabolites. From the values in Table 3, the hybrids can be principally metabolized by CYP2C19 and to a lesser extent by CYP3A4, which is a means of mitigating their undesirable effects. After phase I metabolism of drugs by the cytochrome P450 or other enzymes, most of the drug metabolites undergo a conjugation reaction (phase II metabolism) that facilitates their irreversible elimination from the body, usually via excretion [82]. From the excretion assessment parameters, the compounds have a low clearance rate of ~ 5 ml/min/kg or lower. Nevertheless, they have a short elimination half-life of ~ 1 h or less, hence drug accumulation in the body is unlikely.

From the toxicological data, the hybrids may have a high risk of respiratory toxicity and liver injury. In addition, the hybrids apart from 1B and 1C are susceptible to acute oral toxicity. Moreover, 1A and 2C may lead to human hepatotoxicity, as well as 1A and 2A−2C are likely to be carcinogenic. Interestingly, none of the hybrids is predicted to have a high risk of hERG inhibition, skin sensitization, eye corrosion, and eye irritation. The compounds 1B, 1C, 1D, and 1E apparently have the lowest toxicity risk since they have the lowest probability of human hepatotoxicity and carcinogenicity, although further assessments in these regards are required because the prediction probabilities are + or − . Henceforth, attention herein is focused on these safer compounds (1B, 1C, 1D and 1E).

Table 4 Docking scores (in kcal/mol) of the most drug-like and safer benzimidazole-thiosemicarbazone ligands investigated, along with their interactions with FP2 as revealed by Biovia Discovery studio (in comparison with those of the native ligand, E64)

Ligand	Score	Noncovalent interactions	Interacting residues*
1B	- 7.900	Conventional Hydrogen-bond	Asn173, Cys42
		Pi-sulfur	Cys42
		Pi-Pi T-shaped stacking	Hsd174
		Alkyl and Pi-Alkyl	Trp206, Val152, Ala175, Trp43, Leu84, Ile85
1C	- 8.300	Conventional Hydrogen-bond	Asn81, Gly40, Cys42, Gly83
		Non-conventional Hydrogen-bond	Ser149
			Asp234
		Halogen bond	Trp43
		Pi-sulfur	Tyr78
		Pi-Pi stacking	Ala175, Leu84, Ile85
		Alkyl and Pi-Alkyl	
1D	- 8.200	Conventional Hydrogen-bond	Leu172, Gly83, Hsd174, Val150, Asp234, Asn173
		Non-conventional Hydrogen-bond	Gly40
			Cys42
		Pi-sulfur	Ala175
		Pi-sigma	Ile85, Leu84
		Alkyl and Pi-Alkyl	
1E	- 7.500	Conventional Hydrogen-bond	Gly83
		Pi-sigma	Leu84
		Alkyl and Pi-Alkyl	Ile85, Ala175, Cys42
E64	- 6.300	Conventional Hydrogen-bond	Cys42, Ser41, Asn173, Gly83, Asn81, Hsd174, Gln36
		Non-conventional Hydrogen-bond	Gly40, Trp43
			Tyr78
		Pi-cation	

* The residue color code: Green = S1 subsite; Purple = S1′ subsite; Blue = S2 subsite; red = S3 subsite of FP2

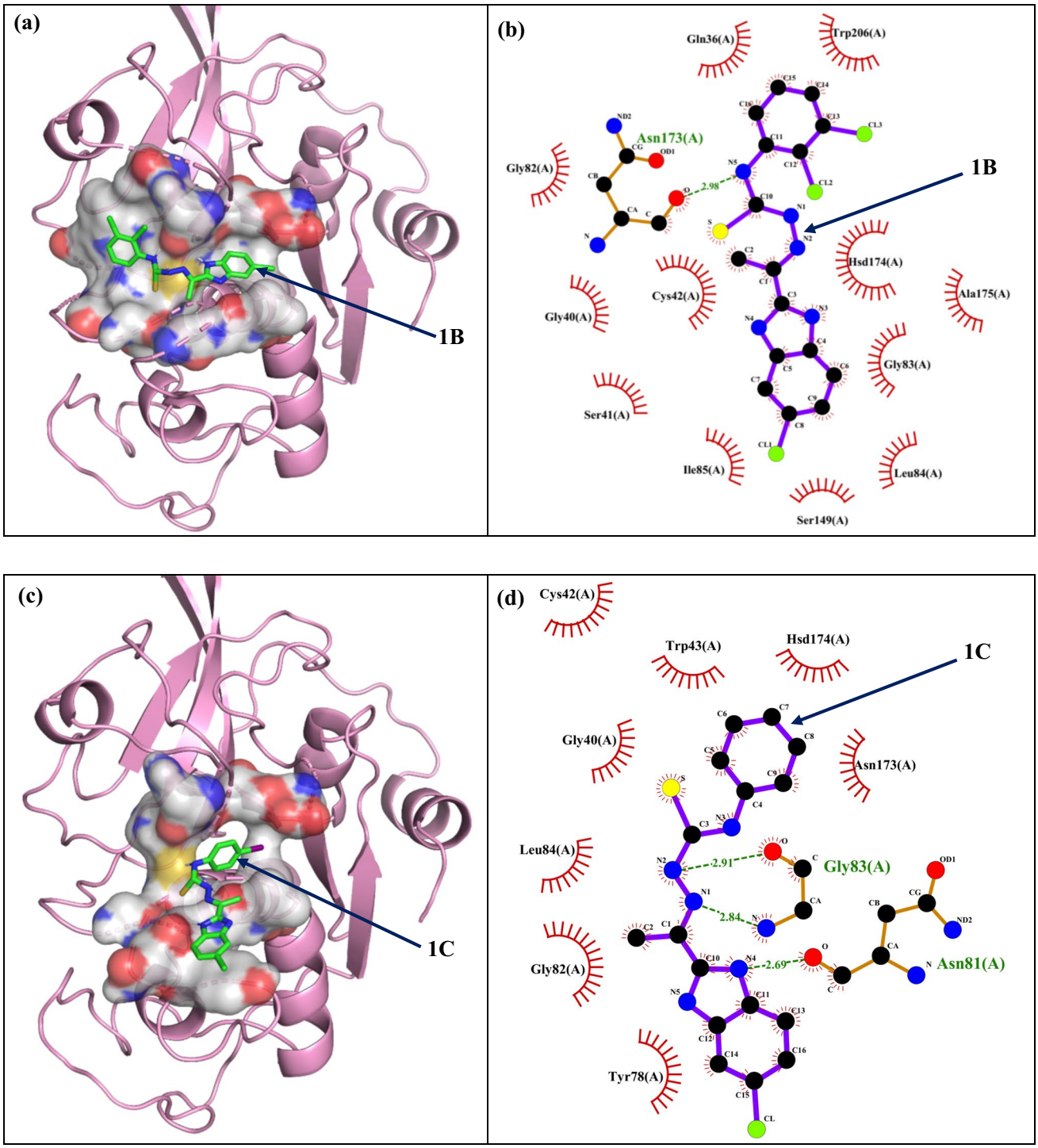


Fig. 2 The top-ranked docking pose of 1B (panel **a**) and 1C (panel **c**) in the binding site of FP2, along with the LigPlot^{+} ligand interaction diagram of 1B (panel **b**) and 1C (panel **d**). The yellow spot in the surface defining the active site cleft (panels **a** and **c**) corresponds to the position of the active site cysteine residue, Cys42

Molecular docking studies

Self-docking of FP2's native ligand (E64) was used as a means of selecting the best docking program between ADVina and ADFR. Note that ADFR is a new docking engine based on the Autodock4 scoring function, but has a new Genetic Algorithm (GA) [83]. The docking efficiency of the programs was validated by assessing the

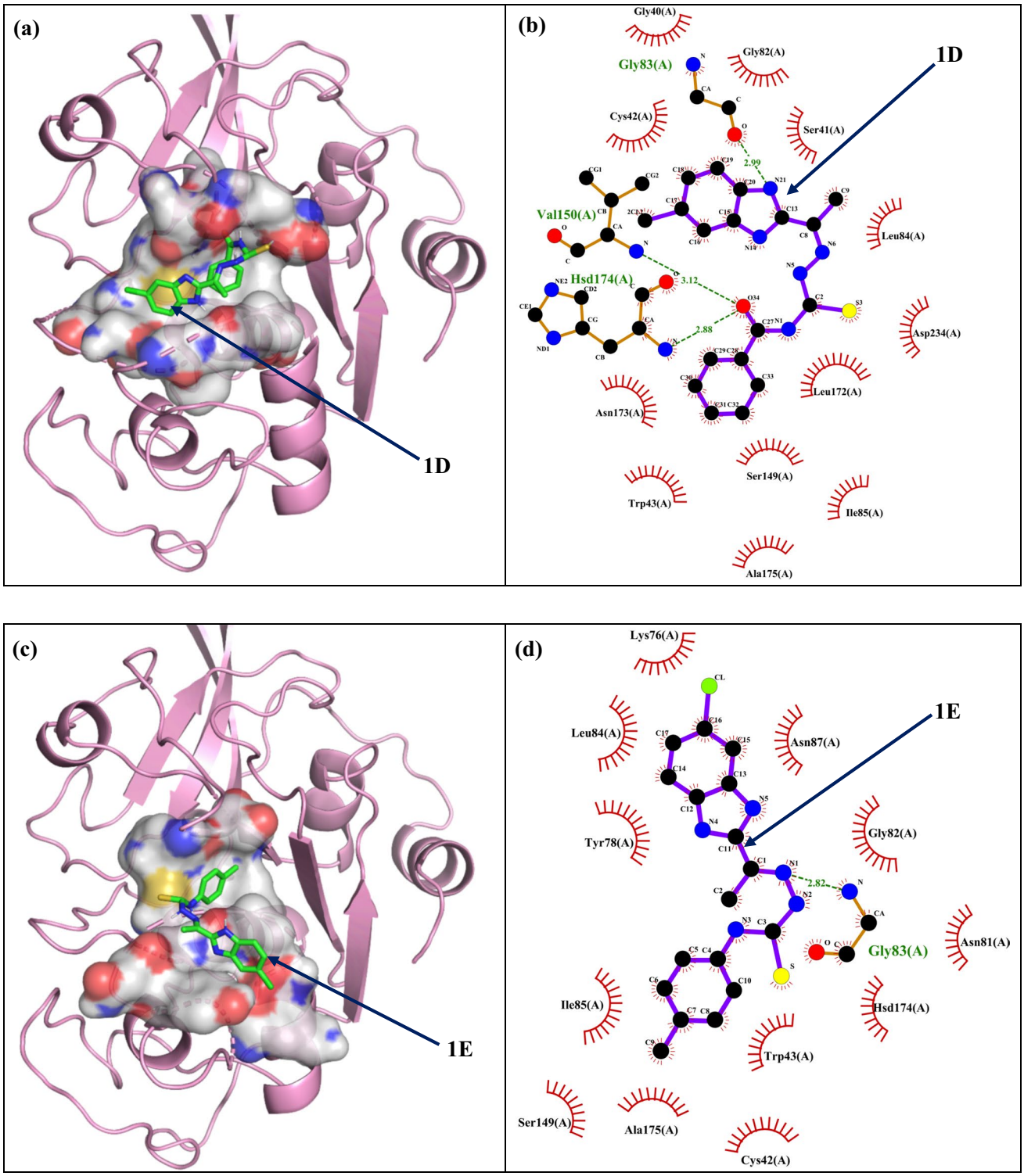


Fig. 3 The top-ranked docking pose of 1D (panel **a**) and 1E (panel **c**) in the binding site of FP2, along with the LigPlot^{+} ligand interaction diagram of 1D (panel **b**) and 1E (panel **d**). The yellow spot in the surface defining the active site cleft (panels **a** and **c**) corresponds to the position of the active site cysteine residue, Cys42

ability of each program to generate a top-ranked native-like pose. The top-ranked pose of E64 as generated by each program is compared with the native pose in Fig. S1. The heavy-atom RMSD value of the ADFR-generated top-ranked pose relative to the crystallographic structure of E64 is 1.7255 Å, which is in good agreement with a

Table 5 MM/GBSA-based binding free energies, free energy contributions, inhibitory constants, and ligand efficiency indices of the benzimidazole-thiosemicarbazone ligands studied

Parameter	Ligand−FP2 complex			
	1B−PF2	1C−PF2	1D−PF2	1E−PF2
ΔE_{bond}	1.13×10^{-06}	-4.67×10^{-07}	5.33×10^{-07}	6.67×10^{-07}
ΔE_{angle}	-1.20×10^{-06}	8.00×10^{-07}	-4.00×10^{-07}	-2.00×10^{-07}
$\Delta E_{torsion}$	9.33×10^{-07}	1.07×10^{-06}	4.67×10^{-07}	1.20×10^{-06}
ΔE_{vdW}	−36.86	−30.22	−26.34	−23.55
ΔE_{ele}	−15.19	−14.92	−12.60	−8.58
1–4 ΔE_{vdW}	-9.77×10^{-16}	-1.73×10^{-06}	-1.20×10^{-06}	7.33×10^{-07}
1–4 ΔE_{ele}	-4.00×10^{-07}	7.33×10^{-07}	-6.00×10^{-07}	6.67×10^{-07}
ΔG_{GB}	25.44	24.14	22.42	17.78
ΔG_{SA}	−3.71	−3.30	−3.04	−2.82
ΔG_{Gas}	−52.05	−45.14	−38.94	−32.13
ΔG_{solv}	21.73	20.85	19.38	14.96
ΔG_{bind}(kcal/mol)	**−30.32**	**−24.29**	**−19.57**	**−17.17**
K_i(nM)[#]	5.94×10^{-14}	1.56×10^{-09}	4.54×10^{-06}	2.59×10^{-04}
IC_{50} (μg/ml)[a]	0.005	0.012	0.023	0.025
LE (kcal/mol)[b]	1.213	1.012	0.783	0.715

[a]Experimental values obtained from the literature [14]

[b]Consult the supplementary material for details on the inhibitory constant (K_i) and ligand efficiency (LE)

Bold indicate Calculated binding free energies are highlighted in bold

previously reported value of 1.84 Å elsewhere [5] obtained using Autodock4. On the other hand, the heavy-atom RMSD value of the ADVina-generated top-ranked pose is 5.7298 Å. Generally, self-docking solutions are considered good when RMSD ≤ 2.0 Å, acceptable when RMSD is between 2.0 and 3.0 Å, and bad when RMSD ≥ 3.0 Å because such solutions are completely wrong [84, 85]. Clearly, ADFR outperforms ADVina as far as the FP2 protein is concerned, and was then chosen for cross-docking of the hybrid molecules earlier selected based on drug-likeness and ADMET predictions.

The docking scores and the interaction types of the ligands 1B, 1C, 1D, and 1E with FP2 are given in Table 4, while the binding poses are depicted in Figs. 2 and 3. The listed interactions were identified using Biovia Discovery studio. The associated 2-dimensional diagrams are shown in supplementary Fig. S2.

The top-ranked docking scores of 1B, 1C, 1D, and 1E with FP2 are favorable and relatively lower than that of E64, suggesting that these compounds may be better inhibitors of FP2 than the native ligand. It is worth noting that E64 is a well-known potent and specific irreversible inhibitor of cysteine proteases [26, 86]. The principal noncovalent interactions between the ligands and FP2 are hydrogen bonds, Pi and alkyl interactions with the binding site amino acid residues such as Cys42, His174, Asn173, Trp43, Gly83, Asn81, Gly40, Tyr78, and Trp43. It can be seen from the LigPlot$^+$ maps in Figs. 2 and 3 that the binding interactions in the ligand−FP2 complexes are also more or less dominated by hydrophobic interactions between the ligands and the hydrophobic side chains of FP2, which should contribute significantly to the binding affinities of the ligands. Indeed, hydrophobic interactions usually contribute considerably to protein–ligand binding free energies, as they constitute the structural parameter that correlates best with binding free energy [87]. The LigPlot$^+$ maps further highlight the importance of the Gly83 residue as the main anchor of most of the ligands in the binding site of FP2. The yellow spot in the active site cleft in Figs. 2 and 3 indicates the position of the sulfur atom in the active site cysteine residue (Cys42) of the catalytic dyad. The distance between the sulfur atom and the thiocarbonyl carbon (the electrophilic center) is 3.90, 4.46, 8.10, and 4.52 Å in 1B, 1C, 1D, and 1E, respectively. The close proximity of the pre-reactive species in the 1B−PF2, 1C−PF2, and 1E−PF2 complexes can facilitate any covalent interaction between them. Indeed, the proximity of the ligand binding site to the catalytic residues is an important factor that influences the ease and rate of covalent bond formation between the ligand and the catalytic residues [30, 65]. Moreover, the ligands bind mainly in the S1 and S2 (as well as S1′ in some cases) subsites of FP2 (see Table 4) to form a noncovalent encounter complex necessary for covalent reaction with Cys42.

Molecular dynamics simulation

MD simulations were performed herein firstly to generate conformational snapshots for use in binding free energy

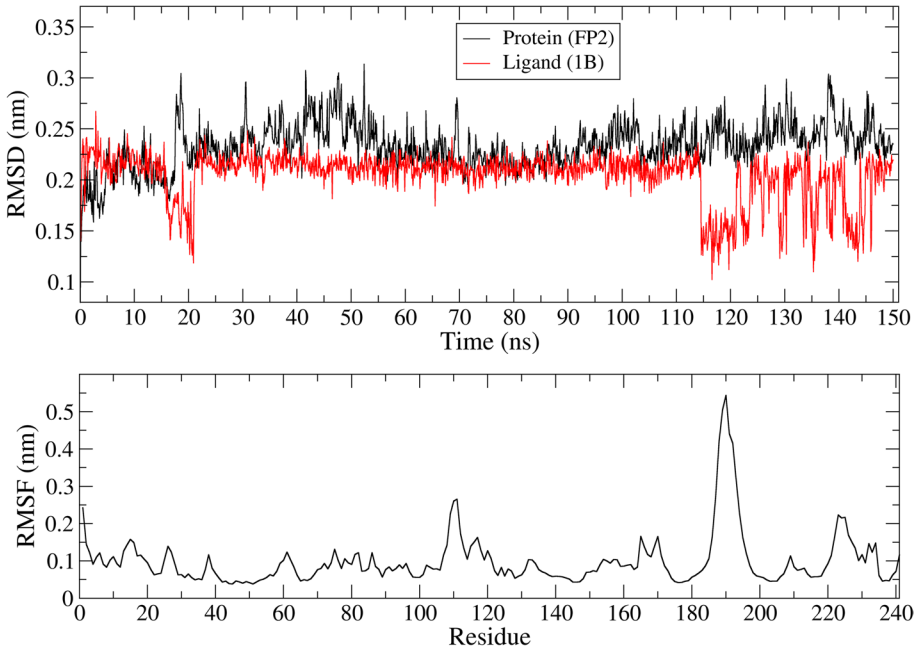


Fig. 4 Backbone RMSD plots of the protein and ligand (top), and the RMSF plot of the protein (bottom) for 150 ns MD simulation

calculations, and secondly to investigate the stability of the best ligand−FP2 complex.

Binding free energy analysis

The calculated binding free energies (ΔG_{bind}) in this work using the MM/GBSA method are presented in Table 5 (highlighted in bold). Note that the reported values were averaged over the 500 snapshots selected from each of the 5, 10, and 15 ns independent MD simulations per ligand-FP2 complex. The detailed results are reported in supplementary Table S3. Relatively short MD simulations were performed here because literature survey reveals that simulations longer than 5 ns are not necessarily beneficial [54–57, 88] to binding free energies calculated using the MM/GBSA and MM/PBSA methods. For some molecules, the binding free energies predicted based on short MD simulations have been found to be slightly better than those obtained from much longer simulations [88]. It is more effective to employ the MM/GBSA and MM/PBSA methods together with many short independent simulations than with a single long simulation [54,

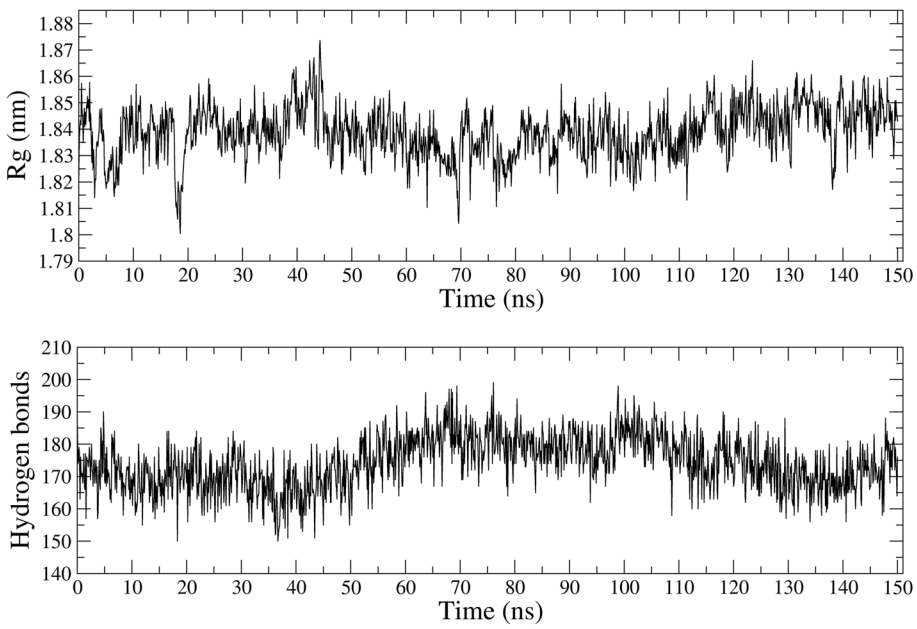


Fig. 5 The Radius of gyration (Rg) plot per residue (top) and the H-Bond plot (bottom) of the protein for 150 ns MD simulation

57]. As can be seen from Table 5, 1B exhibits the lowest average binding free energy of −30.32 kcal/mol and thus possesses the highest binding affinity for the FP2 enzyme. Clearly, the four ligands studied here are capable of non-covalent binding to FP2, but 1D and 1E have relatively high ΔG_{bind} values indicative of their weaker interaction with the protein's binding site residues. The inhibitory constants, K_i, of the ligands are listed in Table 5 together with the experimental values of IC_{50} gotten from the literature [14]. Also presented in Table 5 are ligand efficiency (LE) indices.

The values of K_i and IC_{50} are commonly used to compare the relative potencies of inhibitor molecules. Typically, smaller K_i and IC_{50} values indicate tighter ligand binding to the receptor and consequently better inhibitory activities [89]. It can be seen that the K_i values are well below single-digit *n*M, showing that the ligands can bind very tightly to the FP2 enzyme. Interestingly, the trend in the K_i values (1B < 1C < 1D < 1E) is the same as that in the IC_{50} values. Accordingly, 1B is the strongest binder of FP2 and also the most biologically active compound among the hybrid molecules studied. All calculated LE indices are favorable since they are greater than 0.3 kcal/mol in each case, thus confirming the suitability of the benzimidazole-thiosemicarbazone hybrids as drug candidates. It is also clear from Table 5 that the van der Waals and electrostatic interactions contribute significantly to the binding free energies. It can be seen from Table S3 that the IE standard deviations, σ(Int. Energy), are all greater than the highest allowable value of 3.6 kcal/mol (~ 15 kJ/mol). Therefore, the calculated interaction

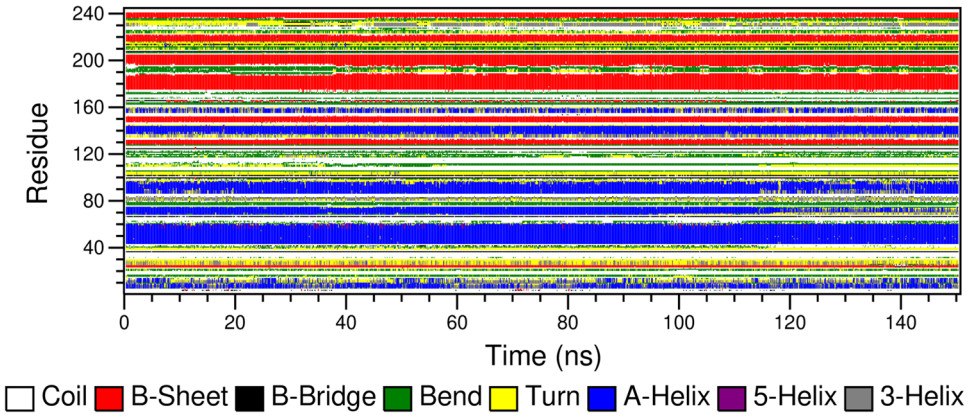


Fig. 6 The DSSP plot of PF2 in complex with 1B for 150 ns MD simulation

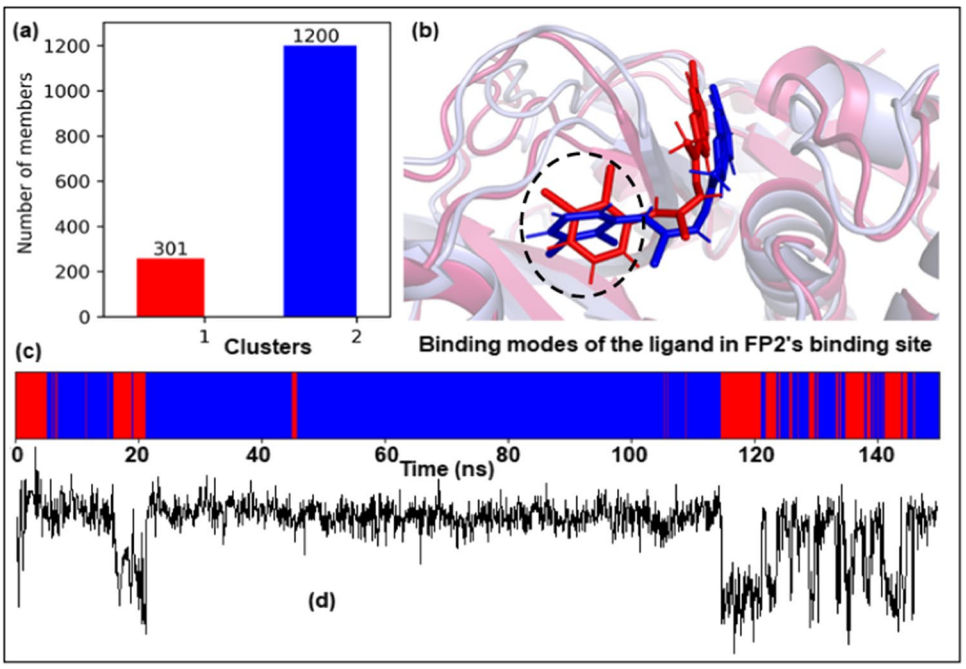


Fig. 7 **a** the clustering histogram, **b** the principal binding modes of 1B, **c** the timeline bar plot, and **d** the RMSD plot of ligand 1B for 150 ns MD simulation

entropy values were considered unreliable [90] and were excluded from the reported ΔG_{bind} values.

Stability analysis of FP2 in complex with 1B

MD simulation of the best noncovalently docked ligand-FP2 complex was performed to verify its stability and correctness. The 150 ns MD trajectory of the complex was used to calculate the RMSD and RMSF of the protein, as well as the RMSD of the ligand, all of which are depicted graphically in Fig. 4. It is important to note that the reported RMSD values of the protein were calculated with respect to its crystallographic structure after least square fitting to the backbone atoms, while those of the ligand were calculated relative to its docked conformation within FP2's binding site. The RMSD plot is often used to validate the stability of a protein–ligand complex and to predict any ligand-induced perturbations, while the RMSF plot is used to assess the local structural flexibility of the amino acid residues of a protein. FP2 in complex with 1B (black) showed the most stable backbone RMSD between 55 and 150 ns of the simulation, and no sharp fluctuations were observed during this time frame. On the other hand, the ligand (red) showed a fairly stable RMSD between 5 and 114 ns of the simulation. From the 115th ns to the end of the simulation, the structure of the ligand alternated between two binding states with average RMSD values of about 0.21 and 0.15 nm. The RMSD of the protein backbone increased to a maximum value at the 52nd ns and eventually stabilized at an average value of 0.23 ± 0.02 nm (or 2.3 ± 0.2 Å) during the rest of the simulation time. Interestingly, the average RMSD of the protein in complex with 1B remained below 3 Å, indicating the stability of the 1B−PF2 complex in which the ligand is tightly bonded to the protein without significantly perturbating its secondary structure.

To validate the predictions of the RMSD analysis, RMSF values were computed and analyzed. It is worth noting that the RMSF plot of a protein provides important information regarding its stability. Generally, high fluctuations in the RMSF plot indicate more flexible and relatively unstable bonds, while low fluctuations indicate well-structured regions with less distortion. As can be seen from Fig. 4, the most fluctuating residues of FP2 bound to 1B are within the range 185−195, which corresponds to a loop and randomly coiled region of the protein. Hence, such flexibility is considered ordinary. The low average RMSF value of the protein (0.10 ± 0.07 nm) confirms the stability of the 1B−PF2 complex as predicted by the RMSD plots.

To determine the compactness of the 1B−PF2 complex, its radius of gyration (Rg) was calculated and plotted against simulation time as shown in Fig. 5. In general, high Rg values signify less compactness, while low values indicate more folded and hence highly stable protein structures. As can be seen from Fig. 5, the calculated Rg values of FP2 bound

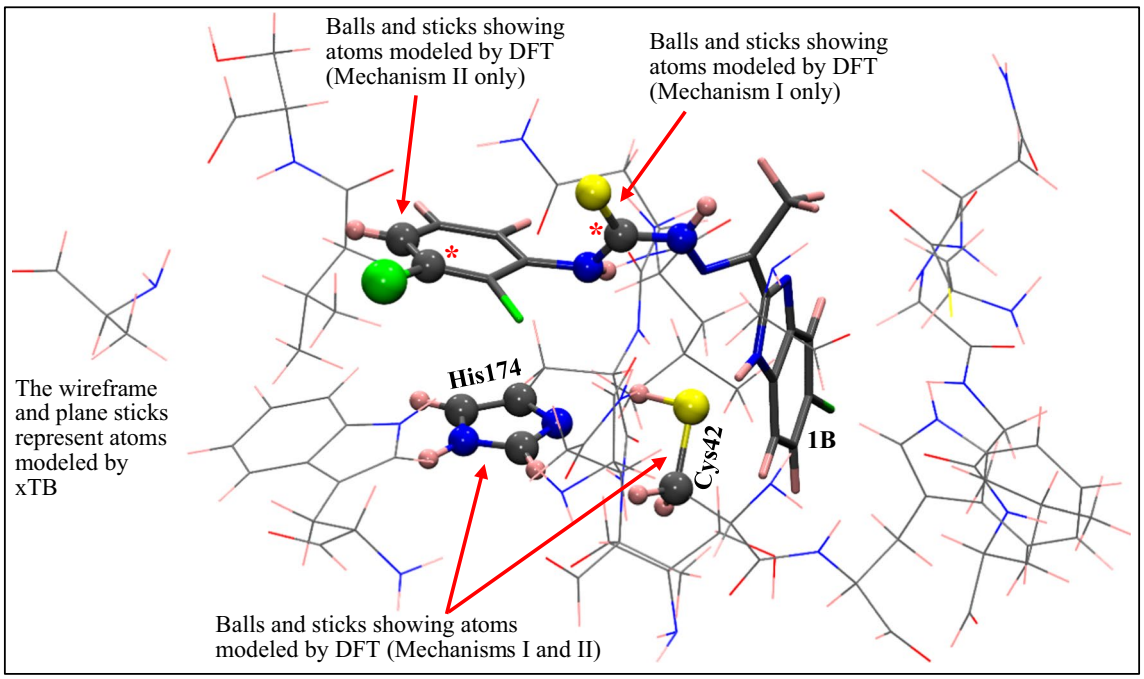
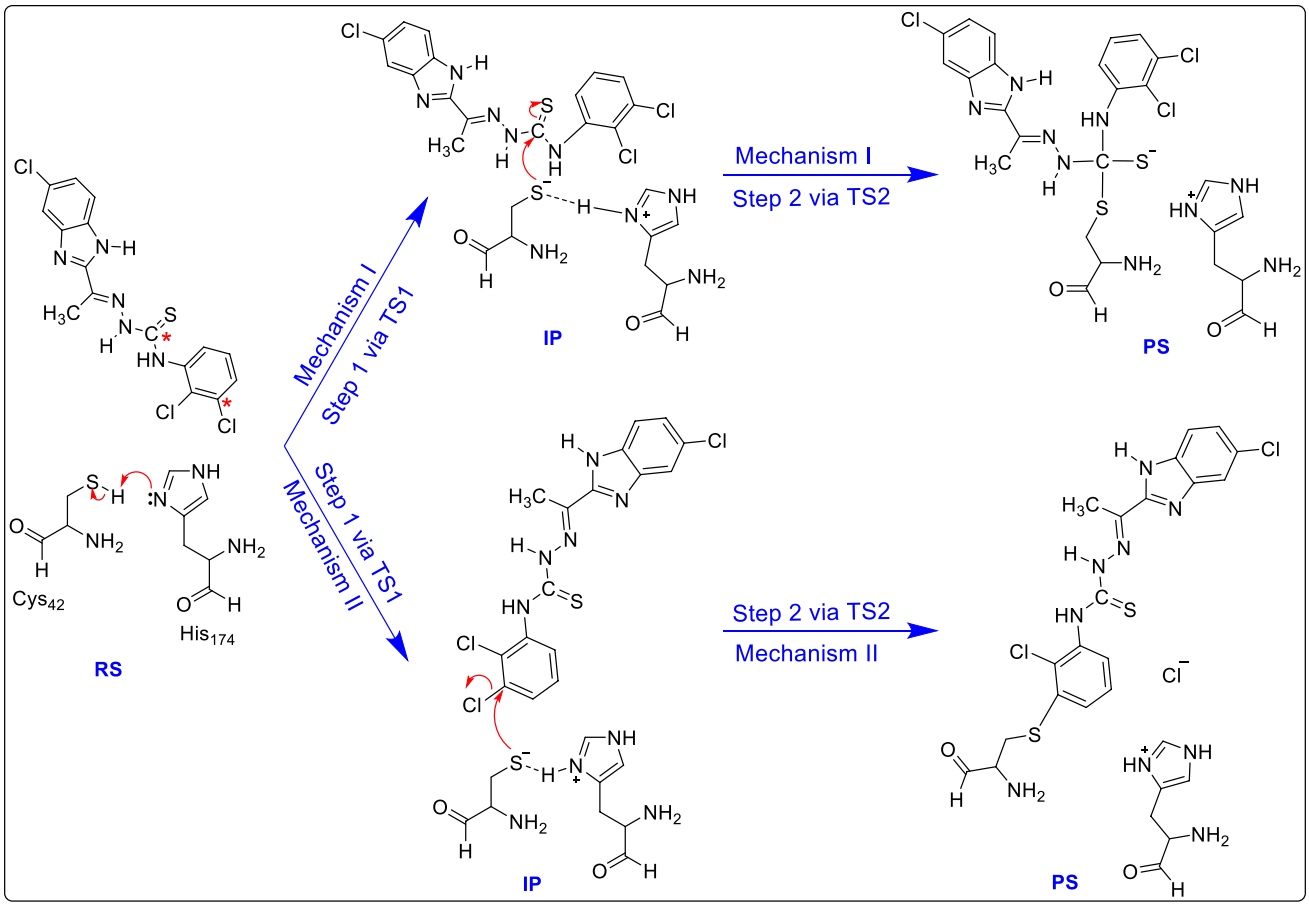


Fig. 8 The cluster model of the 1B−PF2 complex used herein. The C, H, O, N, S, and Cl atoms are shown in gray, pink, red, blue, yellow, and green colors, respectively. Electrophilic centers of the chemical warhead are indicated by the red-colored asterisks. Mechanisms I and II are described below



Scheme 1 Two pathways of the reaction between 1B and PF2 suggested in this work. The electrophilic centers of the chemical warhead are indicated by the red-colored asterisks

to 1B lie within the narrow range 1.80–1.87 nm throughout the MD simulation, implying that the binding of 1B to FP2 does not induce any noticeable structural changes. This further affirms the stability of the 1B−PF2 complex. Hydrogen bonds on their part play a crucial role in determining the protein–ligand binding strength. It is clear from Fig. 5 that the number of hydrogen bonds in the 1B−PF2 complex remained fairly constant during the simulation, suggesting that the secondary structure of FP2 is minimally impacted by 1B.

The definition secondary structure of protein (DSSP) plot of PF2 in complex with 1B is shown in Fig. 6. A DSSP plot can provide valuable information on the nature of the secondary structure elements of a protein (such as the alpha-helices, beta-sheets, random coils, bends, and turns) during MD simulation. It can be seen from Fig. 6 that nearly all of the alpha-helices and beta-sheets of PF2 were conserved during the simulation. Some interconversion among the random coils, bends, and turns is observed, but cannot be attributed to the presence of the ligand because these secondary structure elements are usually composed of highly flexible amino acid residues. Generally, the DSSP plot has not revealed any alterations in the secondary structure elements of PF2 that can lead to protein instability.

Clustering of the trajectory

The trajectory clustering results are presented in Fig. 7. Two clusters comprising different binding modes of ligand 1B in the binding site of FP2 have been generated by the TTClust program.

Recall that the trajectory was clustered based on RMSD values calculated using the ligand atoms. The histogram in panel (a) describes the population of the clusters and the timeline bar plot in panel (c) shows the distribution of cluster members along the entire MD trajectory. The two binding orientations of the ligand are shown in panel (b). It is clear from both the RMSD plot of 1B in panel (d) and the timeline bar plot that the ligand was predominantly stable in the interval 5–114 ns of the simulation in the binding mode depicted in blue. This binding mode is therefore the ligand's dominant configuration, since it pertains to the most

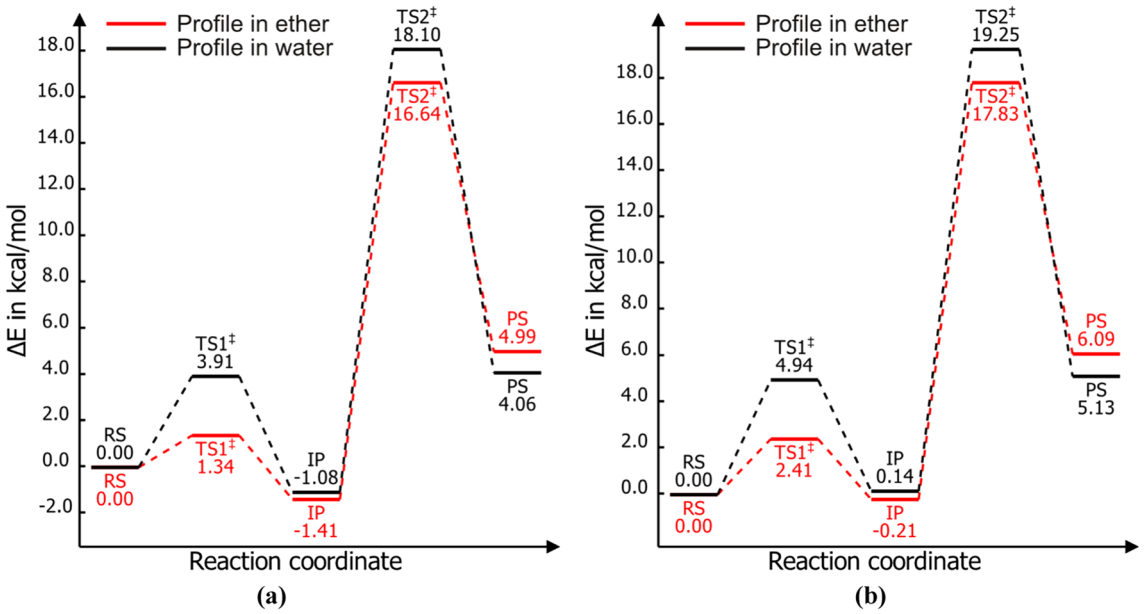


Fig. 9 The energy profiles for mechanism I computed at **a** (M06-2X/def2-TZVPPD)/xTB and **b** (DSD-PBEB95/def2-TZVPPD)/xTB levels of theory. All relative energy values are in kcal/mol

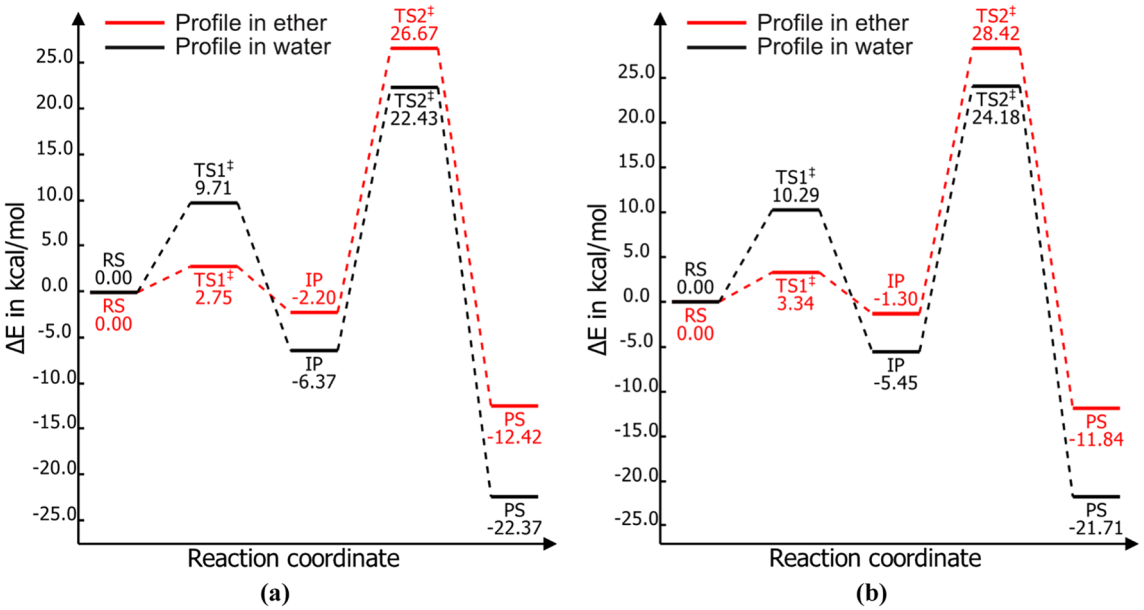


Fig. 10 The energy profiles for mechanism II computed at **a** (M06-2X/def2-TZVPPD)/xTB and **b** (DSD-PBEB95/def2-TZVPPD)/xTB levels of theory. All relative energy values are in kcal/mol

populated cluster. The timeline bar plot also confirms that the structure of the ligand alternated between two binding states from the 115th ns to the end of the simulation. Notice that the two binding modes differ mainly in the orientation of the *dichloro phenyl* group enclosed in the dashed-line circle. Hence, the orientation of the ligand's reactive warhead in both binding states are virtually the same, implying that the two states are likely to follow the same covalent binding mechanism to the reactive active site cysteine residue of FP2. The cluster representative of the dominant

configuration of the 1B−PF2 complex was chosen for mechanistic and kinetic studies described below.

ONIOM-based mechanistic and kinetics studies of covalent modification of FP2 by 1B

The optimized structure of the truncated cluster model of FP2 noncovalently bound to 1B for the reactant state is shown in Fig. 8. Remember that the model was constructed by extracting ligand 1B, along with all amino acid residues in the 1B−PF2 complex with atoms within 4.0 Å of 1B (Gln36, Gly40, Cys42, Trp43, Cys80, Asn81, Gly82, Gly83, Leu84, Ser149, Val152, Ser153, Ala157, Leu172, Asn173, His174, Ala175, and Trp206).

All broken N-terminal and C-terminal bonds were manually capped with hydrogen atoms, and the original positions of the peripheral atoms of the model at the truncation points were constrained. The constraints were intended to prevent spurious displacements of atoms during geometry optimization and NEB-TS calculations in order to preserve the structural integrity of the enzyme's binding site. The presence of the entire protein environment of a catalytic site has a significant impact on the energetics of a reaction, but is unlikely to alter the kinetics and mechanistic features of the reaction [30, 91]. Implicit solvation was applied in all ONIOM calculations, and the solvent effects were included via the Analytical Linearized Poisson-Boltzmann (ALPB) solvation model along with diethyl ether and water as solvents, considered individually. Diethyl ether was chosen so as to include the effects of the low dielectric constant (usually 4.0) of the protein environment, and water was chosen because many biochemical reactions in the human body occur in the aqueous medium. The cluster model was made up of 303 atoms and a total charge of zero. The link hydrogen atoms between the ONIOM layers were automatically generated and charge alteration applied by the ORCA program to prevent overpolarization of the electron density.

Two pathways of the reaction between 1B and PF2, depicted in Scheme **1**, were studied, each comprising two sequential steps: (1) the deprotonation of the thiol group of Cys42 by His174 leading to the formation of $Cys42^{-}/His174^{+}$, called the thiolate-imidazolium ion pair (IP) and (2) covalent modification of the thiolate sulfur of $Cys42^{-}$ by the carbon atom serving as the electrophilic center of the chemical warhead.

In step 1 of each mechanism, the reactant state (RS) is transformed into the IP state through a transition state TS1, while in step 2, the thiolate sulfur of $Cys42^{-}$ and the corresponding electrophilic center combine covalently yielding the product state (PS) via a transition state TS2. The optimized geometries of TS1 and TS2 for each mechanism are shown in supplementary Fig. S3. The energy profiles for Mechanisms I and II computed at (DSD-PBEB95/def2-TZVPPD)/xTB and (M06-2X/def2-TZVPPD)/xTB theoretical levels are displayed in Figs. 9 and 10, respectively. For each mechanism, the reaction coordinate for step 1 (RS → TS1 → IP) is the proton transfer path between the gamma sulfur (SG) of Cys42 and the epsilon nitrogen (NE) of His174. The reaction coordinate for step 2 (IP → TS2 → PS) of Mechanism I is the path linking the thiocarbonyl carbon of the TSC moiety and the SG atom of Cys42. For Mechanism II, the reaction coordinate for step 2 is the path linking the *meta*-carbon atom of the ^{4}N-substituent and the SG atom of Cys42. Notice that in the latter case, a chloride ion is eliminated as a leaving group.

It is clear from Figs. 9 and 10 that similar energy profiles are obtained at the two levels of theory used, although the Gibbs free energy changes calculated using the double-hybrid functional DSD-PBEB95 are slightly higher than those calculated using the global hybrid *meta*-GGA functional M06-2X. Since the double-hybrid functionals often outperform the global hybrid counterparts in typical thermochemical applications, attention herein is henceforth shifted toward the energy profiles obtained using the DSD-PBEB95 functional. It is clear from the energy profiles that the reaction described by Mechanism I is endergonic in both water and ether ($\Delta G_{rxn} = 5.13$ and 6.09 kcal/mol), while that described by Mechanism II is exergonic ($\Delta G_{rxn} = -21.71$ and −11.84 kcal/mol) in both solvents. As a consequence, Mechanism II is thermodynamically favorable, whereas Mechanism I is thermodynamically unfeasible. The nonviability of Mechanism I can be partly attributed to the steric effects of the bulky ^{4}N-substituent, which probably render the thiocarbonyl group of the TSC moiety less attractive as a chemical warhead. From our results, the deprotonation of the thiol group of Cys42 by the NE atom of His174 occurs across a very low energy barrier in diethyl ether (2.41 kcal/mol for Mechanism I and 3.34 kcal/mol for Mechanism II), which either doubles or triples in water. Interestingly, the barrier heights associated with TS1 as calculated herein is roughly in good agreement with the value 5.3 kcal/mol obtained for a similar reaction elsewhere [24]. The foregoing results suggest that the transformation of the neutral active site residues into the IP state is not the rate-determining step in both mechanisms studied. The rate-determining step is undoubtedly that involving the formation of a covalent bond between Cys42 and the chemical warhead in 1B, which in the case of Mechanism II, occurs across a relatively high barrier height of about 28.42 kcal/mol in diethyl ether and 24.18 kcal/mol in water. In the case of the thermodynamically favorable mechanism, the aqueous medium stabilizes both the IP and PS states and also reduces the energy barrier between them. As a result of such stabilization, the IP state is appreciably lower in energy than the neutral form, which is in good qualitative and quantitative agreement with the results of some previous works [92]. Our energy barrier

for the rate-determining step of Mechanism II in aqueous medium (24.18 kcal/mol) is in excellent agreement with those reported in the literature [26] for similar covalent interactions between FP2 and E64 (24.3 and 23.6 kcal/mol) obtained using the Potential of Mean Force (PMF) computed at AM1d/MM level.

Arguably, the cluster model and ONIOM methodology used in this work correctly describe the reaction between PF2 and 1B. Interestingly, the ONIOM results have revealed that ligand 1B may act as a covalent inhibitor of FP2 via Mechanism II suggested in this work. This observation may be extended to ligand 1C (containing a *para*-carbon in its ^{4}N-substituent bonded to a fluorine atom) and ligand 1D (possessing a carbonyl carbon in its ^{4}N-substituent). Unexpectedly, the results show that the *meta*-carbon of the aromatic ring in the ^{4}N-substituent of ligand 1B is more susceptible to nucleophilic attack by the thiolate sulfur of Cys42 than the thiocarbonyl carbon. The said aromatic ring forms a Pi-sulfur interaction (of length 5.68 Å) with the Cys42 sulfur atom, as shown in Fig. S2, implying that it is within the reach of the Cys42 sulfur atom. The fact that the covalent interaction between the sulfur atom of Cys42 and the *meta*-carbon of the aromatic ring is thermodynamically and kinetically favorable as opposed to that between the sulfur and the thiocarbonyl carbon (which is more proximal to the Cys42) shows that the aromatic ring is a fast-reacting warhead. The electrophilic center of the ring is also less sterically hindered than that of the thiocarbonyl group, since the attached chlorine atom is eliminated as an anion prior to attack on the Cys42 sulfur atom. It can be inferred from the foregoing observations that ^{4}N-substituents which improve the antiplasmodial activity of the parent ligand 1A are those equipped with adequately reactive and less sterically hindered chemical warheads.

Conclusions

The effectiveness of Artemisinin Combination Therapies (ACTs), the first-line treatments for malaria, is declining in some malaria-endemic areas due to the emergence of artemisinin-resistant *Plasmodium falciparum* species. This warrants the urgent search for alternative antimalarial drugs with novel modes of action. Against this background, an in silico study to screen antimalarial drug candidates from a series of benzimidazole-thiosemicarbazone hybrid molecules with significant antiplasmodial properties and explore their falcipain-2 inhibitory potentials has been undertaken herein. Pharmacokinetic properties, ADMET profiles, MM/GBSA-based binding free energies, reaction mechanisms, and associated barrier heights have been investigated. DFT, molecular dynamics simulation, molecular docking, and ONIOM methods were employed. Computational chemistry methods were chosen because in silico predictions of receptor−ligand binding affinities and interactions have become an excellent approach in drug discovery and development. From the results obtained, four ^{4}N-substituted derivatives of the hybrid molecule (*E*)-2-(1-(5-chloro-1*H*-benzo[*d*]imidazol-2-yl)ethylidene)hydrazine-1-carbothioamide (1A) denoted 1B, 1C, 1D, and 1E are drug-like and promising covalent binders of falcipain-2, exhibiting remarkably small inhibitory constants ($5.94\times10^{-14}-2.59\times10^{-04}$ *n* M) and favorable binding free energies (−30.32 to −17.17 kcal/mol). The rate-determining step of the thermodynamically favorable mechanism occurs across a surmountable barrier height of 24.18 kcal/mol in water and 28.42 kcal/mol in diethyl ether. It is also apparent from our results that the ^{4}N-substituents which improve the antiplasmodial activity of 1A are those equipped with adequately reactive and less sterically hindered chemical warheads.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11030-022-10594-3.

Acknowledgements The authors are thankful to the Ministry of Higher Education of Cameroon for supporting this work through the research modernization allowance to University Lecturers.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

References

1. Olivier M, Van Den Ham K, Shio MT, Kassa FA, Fougeray S (2014) Malarial pigment hemozoin and the innate inflammatory response. Front Immunol. https://doi.org/10.3389/fimmu.2014.00025
2. Mishra M, Singh V, Singh S (2019) Structural insights into key plasmodium proteases as therapeutic drug targets. Front Microbiol. https://doi.org/10.3389/fmicb.2019.00394
3. Rifaie-Graham O, Pollard J, Raccio S, Balog S, Rusch S, Hernández-Castañeda MA, Mantel P-Y, Beck H-P, Bruns N (2019) Hemozoin-catalyzed precipitation polymerization as an assay for malaria diagnosis. Nat Commun. https://doi.org/10.1038/s41467-019-09122-z
4. Coronado LM, Nadovich CT, Spadafora C (2014) Malarial hemozoin: from target to tool. Biochimica et Biophysica Acta (BBA) 1840(6):2032–2041. https://doi.org/10.1016/j.bbagen.2014.02.009
5. Alberca LN, Chuguransky SR, Álvarez CL, Talevi A, Salas-Sarduy E (2019) In silico guided drug repurposing: discovery of new competitive and non-competitive inhibitors of falcipain-2. Front Chem. https://doi.org/10.3389/fchem.2019.00534
6. WHO (2021) World malaria report 2021. World Health Organization. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021. Accessed 20 Dec 2021
7. Pandey KC, Dixit R (2012) Structure-function of falcipains: malarial cysteine proteases. J Trop Med 2012:1–11. https://doi.org/10.1155/2012/345195

8. Lopes F, Santos MMM, Moreira R (2016) Designing covalent inhibitors: a medicinal chemistry challenge. In Biomedical chemistry: current trends and developments. De Gruyter Open Poland, pp. 44–60. https://doi.org/10.1515/9783110468755-002
9. Sharma M, Prasher P (2020) An epigrammatic status of the 'azole'-based antimalarial drugs. RSC Med Chem 11(2):184–211. https://doi.org/10.1039/c9md00479c
10. Yuthavong Y (2013) Antifolate drugs. In Encyclopedia of malaria. pp 1–12. https://doi.org/10.1007/978-1-4614-8757-9_2-1
11. Nixon GL, Moss DM, Shone AE, Lalloo DG, Fisher N, O'Neill PM, Ward SA, Biagini GA (2013) Antimalarial pharmacology and therapeutics of atovaquone. J Antimicrob Chemother 68(5):977–985. https://doi.org/10.1093/jac/dks504
12. Matsa R, Makam P, Kaushik M, Hoti SL, Kannan T (2019) Thiosemicarbazone derivatives: design, synthesis and in vitro antimalarial activity studies. Eur J Pharm Sci. https://doi.org/10.1016/j.ejps.2019.104986
13. Darrell OT, Hulushe ST, Mtshare TE, Beteck RM, Isaacs M, Laming D, Hoppe HC, Krause RWM, Khanye SD (2018) Synthesis, antiplasmodial and antitrypanosomal evaluation of a series of novel 2-oxoquinoline-based thiosemicarbazone derivatives. S Afr J Chem 71:174–181. https://doi.org/10.17159/0379-4350/2018/v71a23
14. Divatia SM, Rajani DP, Rajani SD, Patel HD (2019) Novel thiosemicarbazone derivatives containing benzimidazole moiety: green synthesis and anti-malarial activity. Arab J Chem 12(7):1641–1651. https://doi.org/10.1016/j.arabjc.2014.09.007
15. Beteck RM, Seldon R, Jordaan A, Warner DF, Hoppe HC, Laming D, Khanye SD (2019) New quinolone-based thiosemicarbazones showing activity against plasmodium falciparum and mycobacterium tuberculosis. Molecules 24(9):1740. https://doi.org/10.3390/molecules24091740
16. Nigam A, Kawathekar N, Zaveri M, Jain G (2017) Thiosemicarbazone and benzimidazole hybrid molecules: the privileged scaffolds for antimalarial activity. IJISET—Int J Innov Sci Eng Technol 4(10):6
17. Shibeshi MA, Kifle ZD, Atnafie SA (2020) Antimalarial drug resistance and novel targets for antimalarial drug discovery. Infect Drug Resist 13:14
18. Chellan P, Land KM, Shokar A, Au A, An SH, Clavel CM, Dyson PJ, Kock Cd, Smith PJ, Chibale K, Smith GS (2012) Exploring the versatility of cycloplatinated thiosemicarbazones as antitumor and antiparasitic agents. Organometallics 31(16):5791–5799. https://doi.org/10.1021/om300334z
19. Nkungli NK, Ghogomu JN (2017) Theoretical analysis of the binding of iron(III) protoporphyrin IX to 4-methoxyacetophenone thiosemicarbazone via DFT-D3, MEP, QTAIM, NCI, ELF, and LOL studies. J Mol Model. https://doi.org/10.1007/s00894-017-3370-4
20. Patel HD, Divatia SM, de Clercq E (2013) Synthesis of some novel thiosemicarbazone derivatives having anti-cancer, anti-HIV as well as anti-bacterial activity. Indian J Chem (B) 52(4):11
21. Ismail MI, Ragab HM, Bekhit AA, Ibrahim TM (2021) Targeting multiple conformations of SARS-CoV2 Papain-like protease for drug repositioning: an in-silico study. Comput Biol Med. https://doi.org/10.1016/j.compbiomed.2021.104295
22. Xiong G, Wu Z, Yi J, Fu L, Yang Z, Hsieh C, Yin M, Zeng X, Wu C, Lu A, Chen X, Hou T, Cao D (2021) ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. Nucleic Acids Res 49(W1):W5–W14. https://doi.org/10.1093/nar/gkab255
23. Yañez O, Osorio MI, Uriarte E, Areche C, Tiznado W, Pérez-Donoso JM, García-Beltrán O, González-Nilo F (2021) In silico study of coumarins and quinolines derivatives as potent inhibitors of SARS-CoV-2 main protease. Front Chem. https://doi.org/10.3389/fchem.2020.595097
24. Grazioso G, Legnani L, Toma L, Ettari R, Micale N, De Micheli C (2012) Mechanism of falcipain-2 inhibition by α, β-unsaturated benzo[1,4]diazepin-2-one methyl ester. J Comput Aided Mol Des 26(9):1035–1043. https://doi.org/10.1007/s10822-012-9596-4
25. Rajguru T, Bora D, Modi MK (2022) Identification of promising inhibitors for plasmodium haemoglobinase Falcipain-2, using virtual screening, molecular docking, and MD simulation. J Mol Struct. https://doi.org/10.1016/j.molstruc.2021.131427
26. Arafet K, Ferrer S, Martí S, Moliner V (2014) Quantum mechanics/molecular mechanics studies of the mechanism of falcipain-2 inhibition by the epoxysuccinate E64. Biochemistry 53(20):3336–3346. https://doi.org/10.1021/bi500060h
27. Greenbaum DC, Mackey Z, Hansell E, Doyle P, Gut J, Caffrey CR, Lehrman J, Rosenthal PJ, McKerrow JH, Chibale K (2004) Synthesis and structure−activity relationships of parasiticidal thiosemicarbazone cysteine protease inhibitors against plasmodium falciparum, trypanosoma brucei, and trypanosoma cruzi. J Med Chem 47(12):3212–3219. https://doi.org/10.1021/jm030549j
28. Schröder J, Noack S, Marhöfer RJ, Mottram JC, Coombs GH, Selzer PM (2013) Identification of semicarbazones, thiosemicarbazones and triazine nitriles as inhibitors of leishmania mexicana Cysteine Protease CPB. PLoS ONE 8(10):e77460. https://doi.org/10.1371/journal.pone.0077460
29. Pitchumani VMC, Shankar R, Vijayakumar S (2019) Mechanistic insights into the inhibition mechanism of cysteine cathepsins by chalcone-based inhibitors—a QM cluster model approach. Struct Chem 30(5):1779–1793. https://doi.org/10.1007/s11224-018-1273-3
30. Madabeni A, Nogara PA, Omage FB, Rocha JBT, Orian L (2021) Mechanistic insight into SARS-CoV-2 Mpro inhibition by organoselenides: the Ebselen case study. Appl Sci 11(14):6291. https://doi.org/10.3390/app11146291
31. Jo S, Cheng X, Islam SM, Huang L, Rui H, Zhu A, Lee HS, Qi Y, Han W, Vanommeslaeghe K, MacKerell AD, Roux B, Im W (2014) CHARMM-GUI PDB manipulator for advanced modeling and simulations of proteins containing nonstandard residues. In Biomolecular modelling and simulations. Adv Protein Chem Struct Biol. 235–265. doi:https://doi.org/10.1016/bs.apcsb.2014.06.002
32. Anandakrishnan R, Aguilar B, Onufriev AV (2012) H + + 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic Acids Res 40(W1):W537–W541. https://doi.org/10.1093/nar/gks375
33. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J Cheminf 4(1):17. https://doi.org/10.1186/1758-2946-4-17
34. Grimme S, Hansen A, Ehlert S, Mewes J-M (2021) r2SCAN-3c: a "Swiss army knife" composite electronic-structure method. J Chem Phys 154(6):064103. https://doi.org/10.1063/5.0040021
35. Neese F (2022) Software update: The ORCA program system–version 5.0. WIREs Comput Mol Sci. https://doi.org/10.1002/wcms.1606
36. Guan L, Yang H, Cai Y, Sun L, Di P, Li W, Liu G, Tang Y (2019) ADMET-score—a comprehensive scoring function for evaluation of chemical drug-likeness. MedChemComm 10(1):148–157. https://doi.org/10.1039/c8md00472b
37. Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. https://doi.org/10.1038/srep42717
38. Athar M, Sona AN, Bekono BD, Ntie-Kang F (2019) Fundamental physical and chemical concepts behind "drug-likeness" and

"natural product-likeness." Phys Sci Rev. https://doi.org/10.1515/psr-2018-0101
39. Eberhardt J, Santos-Martins D, Tillack AF, Forli S (2021) AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. J Chem Inf Model 61(8):3891–3898. https://doi.org/10.1021/acs.jcim.1c00203
40. Trott O, Olson AJ (2009) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. https://doi.org/10.1002/jcc.21334
41. Forli S, Olson AJ (2012) A force field with discrete displaceable waters and desolvation entropy for hydrated ligand docking. J Med Chem 55(2):623–638. https://doi.org/10.1021/jm2005145
42. Dassault Systèmes Biovia (2021) Discovery studio molecular visualizer, version 4.5. Dassault Systèmes, San Diego, p 2021
43. Laskowski RA, Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model 51(10):2778–2786. https://doi.org/10.1021/ci200227u
44. The PyMOL Molecular Graphics System Version 2.6 (2021) Molecular graphics system version 2.6. Schrödinger, LLC
45. DeLano WL (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos
46. 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
47. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78(8):1950–1958. https://doi.org/10.1002/prot.22711
48. Jämbeck JPM, Lyubartsev AP (2014) Update to the general amber force field for small solutes with an emphasis on free energies of hydration. J Phys Chem B 118(14):3793–3804. https://doi.org/10.1021/jp4111234
49. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79(2):926–935. https://doi.org/10.1063/1.445869
50. He X, Man VH, Yang W, Lee T-S, Wang J (2020) A fast and high-quality charge model for the next generation general AMBER force field. J Chem Phys 153(11):114502. https://doi.org/10.1063/5.0019056
51. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 25(2):247–260. https://doi.org/10.1016/j.jmgm.2005.12.005
52. Case DA, Aktulga HM, Belfon K, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TE, Cruzeiro VWD, Darden TA, Duke RE, Giambasu G, Gilson MK, Gohlke H, Goetz AW, Harris R, Izadi S, Izmailov SA, Jin C, Kasavajhala K, Kaymak MC, King E, Kovalenko A, Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Machado M, Man V, Manathunga M, Merz KM, Miao Y, Mikhailovskii O, Monard G, Nguyen H, O'Hearn KA, Onufriev A, Pan F, Pantano S, Qi R, Rahnamoun A, Roe DR, Roitberg A, Sagui C, Schott-Verdugo S, Shen J, Simmerling CL, Skrynnikov NR, Smith J, Swails J, Walker RC, Wang J, Wei H, Wolf RM, Wu X, Xue Y, York DM, Zhao S, Kollman PA (2021) Amber 2021. University of California, San Francisco
53. Sousa da Silva AW, Vranken WF (2012) ACPYPE—AnteChamber PYthon parser interfacE. BMC Res Notes 5(1):367. https://doi.org/10.1186/1756-0500-5-367
54. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov 10(5):449–461. https://doi.org/10.1517/17460441.2015.1032936
55. Huang K, Luo S, Cong Y, Zhong S, Zhang JZH, Duan L (2020) An accurate free energy estimator: based on MM/PBSA combined with interaction entropy for protein–ligand binding affinity. Nanoscale 12(19):10737–10750. https://doi.org/10.1039/c9nr10638c
56. Poli G, Granchi C, Rizzolio F, Tuccinardi T (2020) Application of MM-PBSA methods in virtual screening. Molecules 25(8):1971. https://doi.org/10.3390/molecules25081971
57. Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZH, Hou T (2019) End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. Chem Rev 119(16):9478–9508. https://doi.org/10.1021/acs.chemrev.9b00055
58. Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA, Moreno E (2021) gmx_MMPBSA: a new tool to perform end-state free energy calculations with GROMACS. J Chem Theory Comput 17(10):6281–6291. https://doi.org/10.1021/acs.jctc.1c00645
59. Miller BR, McGee TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE (2012) MMPBSA.py: an efficient program for end-state free energy calculations. J Chem Theor Comput 8(9):3314–3321. https://doi.org/10.1021/ct300418h
60. Wang C, Greene DA, Xiao L, Qi R, Luo R (2018) Recent developments and applications of the MMPBSA method. Front Mol Biosci. https://doi.org/10.3389/fmolb.2017.00087
61. Duan L, Liu X, Zhang JZH (2016) Interaction entropy: a new paradigm for highly efficient and reliable computation of protein-ligand binding free energy. J Am Chem Soc 138(17):5722–5728. https://doi.org/10.1021/jacs.6b02682
62. Tuccinardi T (2021) What is the current value of MM/PBSA and MM/GBSA methods in drug discovery? Expert Opin Drug Discov 16(11):1233–1237. https://doi.org/10.1080/17460441.2021.1942836
63. Loo JSE, Yong AYY, Yong YN (2020) The effect of multiple simulation parameters on MM/PBSA performance for binding affinity prediction of CB1 cannabinoid receptor agonists and antagonists. Chem Biol Drug Des 96(5):1244–1254. https://doi.org/10.1111/cbdd.13733
64. Bernetti M, Cavalli A, Mollica L (2017) Protein–ligand (un)binding kinetics as a new paradigm for drug discovery at the crossroad between experiments and modelling. MedChemComm 8(3):534–550. https://doi.org/10.1039/c6md00581k
65. Luzik DA, Rogacheva ON, Izmailov SA, Indeykina MI, Kononikhin AS, Skrynnikov NR (2019) Molecular dynamics model of peptide-protein conjugation: case study of covalent complex between Sos1 peptide and N-terminal SH3 domain from Grb2. Sci Rep. https://doi.org/10.1038/s41598-019-56078-7
66. Tubiana T, Carvaillo J-C, Boulard Y, Bressanelli S (2018) TTClust: a versatile molecular simulation trajectory clustering program with graphical summaries. J Chem Inf Model 58(11):2178–2182. https://doi.org/10.1021/acs.jcim.8b00512
67. Bannwarth C, Ehlert S, Grimme S (2019) GFN2-xTB—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. J Chem Theory Comput 15(3):1652–1671. https://doi.org/10.1021/acs.jctc.8b01176
68. Ásgeirsson V, Birgisson BO, Bjornsson R, Becker U, Neese F, Riplinger C, Jónsson H (2021) Nudged elastic band method for molecular reactions using energy-weighted springs combined with eigenvector following. J Chem Theory Comput 17(8):4929–4945. https://doi.org/10.1021/acs.jctc.1c00462
69. Zhao Y, Truhlar DG (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function. Theor Chem Acc 120:215–241. https://doi.org/10.1007/s00214-007-0310-x

70. Weigend F, Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. Phys Chem Chem Phys 7(18):3297. https://doi.org/10.1039/b508541a
71. Kozuch S, Martin JML (2013) Spin-component-scaled double hybrids: an extensive search for the best fifth-rung functionals blending DFT and perturbation theory. J Comput Chem. https://doi.org/10.1002/jcc.23391
72. Grimme S (2011) Density functional theory with London dispersion corrections. WIREs Comput Mol Sci 1(2):211–228. https://doi.org/10.1002/wcms.30
73. Becke AD, Johnson ER (2005) A density-functional model of the dispersion interaction. J Chem Phys 123(15):154101. https://doi.org/10.1063/1.2065267
74. Johnson ER, Becke AD (2005) A post-Hartree–Fock model of intermolecular interactions. J Chem Phys 123(2):024101. https://doi.org/10.1063/1.1949201
75. Lu T, Chen Q (2021) Shermo: a general code for calculating molecular thermochemistry properties. Comput Theoret Chem. https://doi.org/10.1016/j.comptc.2021.113249
76. Jia C-Y, Li J-Y, Hao G-F, Yang G-F (2020) A drug-likeness toolbox facilitates ADMET study in drug discovery. Drug Discov Today 25(1):248–258. https://doi.org/10.1016/j.drudis.2019.10.014
77. Shahryari S, Mohammadnejad P, Noghabi KA (2021) Screening of anti-acinetobacter baumannii phytochemicals, based on the potential inhibitory effect on OmpA and OmpW functions. R Soc Open Sci 8(8):201652. https://doi.org/10.1098/rsos.201652
78. Härter MW, Keldenich J, Schmitt W (2002) Estimation of physicochemical and ADME parameters. In: Nicolaou KC, Hanko R, Hartwig W (eds) Handbook of combinatorial chemistry. Wiley, pp 743–760
79. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 23(1–3):3–25. https://doi.org/10.1016/s0169-409x(96)00423-1
80. Rudrapal M, Chetia D, Singh V (2017) Novel series of 1,2,4-trioxane derivatives as antimalarial agents. J Enzyme Inhib Med Chem 32(1):1159–1173. https://doi.org/10.1080/14756366.2017.1363742
81. Domínguez-Villa FX, Durán-Iturbide NA, Ávila-Zárraga JG (2021) Synthesis, molecular docking, and in silico ADME/Tox profiling studies of new 1-aryl-5-(3-azidopropyl)indol-4-ones: potential inhibitors of SARS CoV-2 main protease. Bioorg Chem. https://doi.org/10.1016/j.bioorg.2020.104497
82. van der Kamp MW, Mulholland AJ (2013) Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. Biochemistry 52(16):2708–2728. https://doi.org/10.1021/bi400215w
83. Ravindranath PA, Forli S, Goodsell DS, Olson AJ, Sanner MF (2015) AutoDockFR: advances in protein-ligand docking with explicitly specified binding site flexibility. PLOS Comput Biol 11(12):e1004586. https://doi.org/10.1371/journal.pcbi.1004586
84. Ramírez D, Caballero J (2018) Is it reliable to take the molecular docking top scoring position as the best solution without considering available structural data? Molecules 23(5):1038. https://doi.org/10.3390/molecules23051038
85. Sallum LO, Vaz WF, Borges NM, de Campos CEM, Bortoluzzi AJ, Franco CHJ, Ramos LM, Napolitano HB (2019) Synthesis, conformational analysis and molecular docking studies on three novel dihydropyrimidine derivatives. J Mol Struct 1192:274–287. https://doi.org/10.1016/j.molstruc.2019.04.100
86. Salawu EO (2018) In silico study reveals how e64 approaches, binds to, and inhibits falcipain-2 of plasmodium falciparum that causes malaria in humans. Sci Rep. https://doi.org/10.1038/s41598-018-34622-1
87. Bronowska AK (2011) Thermodynamics of ligand-protein interactions: implications for molecular design. In: Pirajn JCM (ed) Thermodynamics—interaction studies—solids, liquids and gases. InTech
88. Hou T, Wang J, Li Y, Wang W (2010) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1: the accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 51(1):69–82. https://doi.org/10.1021/ci100275a
89. Burlingham BT, Widlanski TS (2003) An intuitive look at the relationship of K_i and IC_{50}: a more general use for the Dixon plot. J Chem Educ 80(2):5
90. Ekberg V, Ryde U (2021) On the use of interaction entropy and related methods to estimate binding entropies. J Chem Theor Comput 17(8):5379–5391. https://doi.org/10.1021/acs.jctc.1c00374
91. Prabhakar R, Vreven T, Frisch MJ, Morokuma K, Musaev DG (2006) Is the protein surrounding the active site critical for hydrogen peroxide reduction by selenoprotein glutathione peroxidase? An ONIOM study. J Phys Chem B 110(27):13608–13613. https://doi.org/10.1021/jp0619181
92. Arafet K, Świderek K, Moliner V (2018) Computational study of the Michaelis complex formation and the effect on the reaction mechanism of Cruzain cysteine protease. ACS Omega 3(12):18613–18622. https://doi.org/10.1021/acsomega.8b03010