

Identification of (4-(9H-fluoren-9-yl) piperazin-1-yl) methanone derivatives as falcipain 2 inhibitors active against *Plasmodium falciparum* cultures

Jorge E. Hernández-González^{a,b,1}, Emir Salas-Sarduy^{c,1}, Luisa F. Hernández Ramírez^d,
María J. Pascual^c, Diego E. Álvarez^c, Adriana Pabón^d, Vitor B.P. Leite^a, Pedro G. Pascutti^e,
Pedro A. Valiente^{b,*}

^a Departamento de Física, Universidade Estadual Paulista (UNESP), São José do Rio Preto, São Paulo CEP 15054-000, Brazil

^b Computational Biology and Biomolecular Dynamics Laboratory, Center for Protein Studies, Faculty of Biology, University of Havana, Havana, Cuba

^c Instituto de Investigaciones Biotecnológicas - Instituto Tecnológico de Chascomús (IIB-INTECH), Universidad Nacional de San Martín (UNSAM) - Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), San Martín, Buenos Aires, Argentina

^d Grupo de Malaria, Facultad de Medicina, Universidad de Antioquia, Medellín, Colombia

^e Laboratório de Dinâmica e Modelagem Molecular, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Ave. Carlos Chagas Filho, 373, CCS-Bloco D sala 30, Cidade Universitária Ilha de Fundão, Rio de Janeiro CEP 21941-902, Brazil

ARTICLE INFO

Keywords:

Falcipain 2
Plasmodium falciparum
Inhibition assay
Virtual screening
Molecular dynamics
Binding mode

ABSTRACT

Background: Falcipain 2 (FP-2) is the hemoglobin-degrading cysteine protease of *Plasmodium falciparum* most extensively targeted to develop novel antimalarials. However, no commercial antimalarial drugs based on FP-2 inhibition are available yet due to the low selectivity of most FP-2 inhibitors against the human cysteine proteases.

Methods: A structure-based virtual screening (SVBS) using Maybridge HitFinder™ compound database was conducted to identify potential FP-2 inhibitors. *In vitro* enzymatic and cell-growth inhibition assays were performed for the top-scoring compounds. Docking, molecular dynamics (MD) simulations and free energy calculations were employed to study the interaction of the best hits with FP-2 and other related enzymes.

Results and conclusions: Two hits based on 4-(9H-fluoren-9-yl) piperazin-1-yl) methanone scaffold, HTS07940 and HTS08262, were identified as inhibitors of FP-2 (half-maximal inhibitory concentration (IC₅₀) = 64 μM and 14.7 μM, respectively) without a detectable inhibition against the human off-target cathepsin K (hCatK). HTS07940 and HTS08262 inhibited the growth of the multidrug-resistant *P. falciparum* strain FCR3 in culture (half-maximal inhibitory concentrations (IC₅₀) = 2.91 μM and 34 μM, respectively) and exhibited only moderate cytotoxicity against HeLa cells (Half-maximal cytotoxic concentration (CC₅₀) = 133 μM and 350 μM, respectively). Free energy calculations reproduced the experimental affinities of the hits for FP-2 and explained the selectivity with respect to hCatK.

General significance: To the best of our knowledge, HTS07940 stands among the most selective FP-2 inhibitors identified by SVBS reported so far, displaying moderate antiplasmodial activity and low cytotoxicity against human cells. Hence, this compound constitutes a promising lead for the design of more potent and selective FP-2 inhibitors.

1. Introduction

Malaria is a mosquito-borne infectious disease prevalent in tropical regions, where > 3.2 billion people live under the risk of infection [1]. According to the World Health Organization (WHO), approximately 216 million cases of malaria and 445,000 deaths were reported in 2016. Alarmingly, after a period of success in global malaria control, the

progress has currently stalled [2]. Five species of parasites of the genus *Plasmodia*, i.e., *P. falciparum*, *P. vivax*, *P. ovale* and *P. knowlesi*, are the causative agents of malaria in humans, the former being responsible for the most common and lethal form of the disease [3,4].

So far, various antimalarials, e.g., quinine, chloroquine, artemisinin and atovaquone, have been discovered. However, the development of resistance to the available drugs is becoming a major health concern

* Corresponding author.

E-mail address: valiente@fbio.uh.cu (P.A. Valiente).

¹ These authors contributed equally to this work.

and, consequently, the identification of novel antimalarials is urgently needed [1,3]. In this sense, the development of new drugs has been directed to vital processes of the parasites, such as *i*) biosynthesis, *ii*) membrane transport and signalling and *iii*) hemoglobin catabolism [4]. The latter process is carried out by a battery of proteases active inside the food vacuole of parasites during the intraerythrocytic stage of their life cycle [4]. Of note, two C1 cysteine proteases, FP-2 (MEROPS ID: C01.046) and FP-3 (MEROPS ID: C01.063), have been identified as promising drug targets among the currently known hemoglobinases of *P. falciparum* [1,4].

Several falcipain (FP) inhibitors have been reported in literature, which can be classified into three main categories: *i*) peptide-based, *ii*) peptidomimetic and *iii*) non-peptidic compounds [1,4,5]. Although the most potent FP inhibitors identified so far are of peptide nature, their suitability for therapy is limited, since they are likely to undergo proteolytic cleavage *in vivo* and possess low permeability through cell membranes [5]. On the other hand, peptidomimetic and non-peptidic FP inhibitors, though typically less potent than the peptidic ones, can display improved pharmacokinetic and pharmacodynamic parameters, which makes them more adequate for therapy [5]. Remarkably, despite the important role of FPs, no commercial antimalarial drugs targeting these proteases are available yet; therefore, the search for new compounds capable of inhibiting FPs remains active [1].

Various approaches have been employed to identify FP inhibitors, some of them combining experimental and computational tools [1,3–5]. In particular, the SBVS of chemical databases has proven a widely-used and successful strategy [1,3]. Potential hits in the SBVS approach are typically identified by docking each compound of the chemical database into the binding site of the target molecule with known tridimensional (3D) structure. The compounds are then ranked according to their docking scores and subsequently subjected to experimental evaluation [1,3].

Previous works have shown that docking algorithms may fail to identify the correct binding mode of a ligand as the top-scoring pose [6,7]. Therefore, post-docking refinement steps are generally required in order to accurately predict the lowest-energy conformation. Computational methods such as Molecular Mechanics Poisson-Boltzmann (Generalized Born) Surface Area (MM-PB(GB)SA) free energy calculations, performed either on ensembles generated through molecular dynamics (MD) simulations or on single structures, can be employed for that purpose [6,8–11]. In addition, enhanced sampling techniques, *e.g.*, metadynamics and accelerated molecular dynamics (aMD) simulations, have also proven useful to identify the native binding modes of ligands [6,12].

The present work aims to discover new FP inhibitors through an integrated strategy that involved SBVS, experimental evaluation and *in silico* analyses of the interaction of the identified hits with FPs. We screened the Maybridge HitFinder™ (<http://www.maybridge.com/>) collection of compounds against the crystal structures of FP-2, FP-3 and also that of hCatK (EC 3.4.22.38), which was included to account for potential cross-inhibition effects in the human host. Nine compounds displaying high affinity and specificity for both parasitic proteases with respect to hCatK based on SBVS results, were purchased and subjected to experimental evaluation. From the experimental results, we identified two compounds, HTS07940 and HTS08262, displaying IC₅₀ values in the 10⁻⁶ to 10⁻⁵ M range against FP-2 and *P. falciparum* cultures, and low cytotoxicity and inhibition against HeLa cells and hCatK, respectively. Remarkably, HTS07940 and HTS08262 belong to a novel non-peptidic scaffold of FP-2 inhibitors. We also conducted a thorough computational study to determine the binding modes of HTS07940 and HTS08262 to FP-2, FP-3 and hCatK. Our predictions accurately reproduced the experimental affinities for FP-2 and the specificity with respect to hCatK. This provides a reliable structural knowledge that may help the future optimization of the identified inhibitors.

2. Material and methods

2.1. PSI-Blast and structural alignment

A PSI-Blast search for human homologs of FP-2 in the Protein Data Bank was conducted using the BLAST web server available at <https://www.ncbi.nlm.nih.gov/BLAST/> with default parameters [13]. The 3D structures retrieved from this step plus FP-3 (PDB: 3BWK) were structurally aligned to FP-2 (PDB: 2OUL) using the salign 3D routine implemented in Modeller9.9 [14]. Based on the structural alignment, the conservation degree of residues involved in ligand binding throughout the set of human proteases with respect to FP-2 and FP-3 was analyzed. The human homolog having the largest sequence identity (ID) [15] in the binding site region with both FPs was selected for potential cross-inhibition analysis *in silico*, as described below. The binding site was defined by the broadest set of residues contacting the ligands within a cutoff radius of 6 Å in various 3D structures of human cathepsins and FPs in complex with organic ligands (PDBs: 3BWK, 3BPF, 1AYV, 3BC3).

2.2. Structure-based virtual screening

SBVSs of the Maybridge HitFinder™ database, containing 14,400 drug-like compounds (<http://www.maybridge.com/>), were performed in parallel to target the active sites of FP-2, FP-3 and hCatK, the selected human off-target, using Autodock Vina 1.1.2 [16]. Of note, we incorporated FP-3 into the analysis seeking to identify potential dual inhibitors, but experimental assays were only carried out with FP-2 due to availability issues. The crystal structures of FP-2 (PDB: 2OUL), FP-3 (PDB: 3BWK) and hCatK (PDB: 1AYV) were transformed into pdbqt format using the *prepare_receptor4.py* tool of Autodock tools [17], after protonation at pH = 5.5 in the webserver PDB2PQR [18]. Then, a box centered on the catalytic sulfur atom of each protease, with X, Y and Z dimensions of 18.0, 23.62 and 24.38 Å, respectively, was created to dock the ligands. The compounds were protonated at pH = 5.5 using Open Babel [19] and transformed into pdbqt format with *prepare_ligand4.py* of Autodock tools [17]. The results of the SBVSs were processed to select compounds fulfilling the following criteria: *i*) high affinity for both FPs ($S_{vina} \leq -8.0$ kcal/mol) and *ii*) relative binding free energy ($\Delta\Delta G$) for at least one FP with respect to hCatK less than -1.0 kcal/mol ($\Delta S_{vina} = S_{vina}(\text{FP}) - S_{vina}(\text{hCatK}) < -1.0$ kcal/mol). The second condition sought to enhance the specificity of the putative inhibitors towards the parasitic targets, thereby potentially reducing cross-inhibition of the human off-target.

2.3. Expression and purification of recombinant enzymes

FP-2 (MEROPS ID: C01.046) was expressed as inclusion bodies in BL21(DE3) *Escherichia coli* strain, purified under denaturing conditions (final purity: 91%) and refolded to active enzyme [20]. A *Pichia pastoris* strain expressing a Ser149Ala (glycosylation -) hCatK (EC 3.4.22.38) mutant was kindly donated by Dr. Dieter Bromme (University of British Columbia, Vancouver, Canada). The enzyme was expressed, activated and purified (final purity: > 85%) as previously described [21].

2.4. In vitro enzymatic assays

The nine top-scoring compounds from the SBVS results were purchased for experimental assessment. Inhibition assays against FP-2 and hCatK were conducted as described by Bertoldo et al. [22]. Briefly, stock solutions of the compounds were prepared at 10–20 mM in dimethyl-sulfoxide (DMSO). The enzymatic activity of both enzymes was measured by fluorimetric assays in a buffer (100 mM NaOAc, 10 mM dithiothreitol (DTT), pH = 5.5) employing Z-Phe-Arg-AMC (15 μM) as substrate, and fixed concentrations of FP-2 (0.7 nM) and hCatK (1 nM). Enzymes and compounds were incubated 15 min prior to the addition

of substrate and the release of AMC was monitored for 600 s at 30 °C with a Thermo Spectronic Aminco Bowman Series 2 spectrofluorometer (excitation at $\lambda = 355$ nm; emission at $\lambda = 460$ nm). All scans were corrected from the corresponding blanks, and controls (FP-2/hCatK + DMSO 2%, Z-FR-AMC + DMSO 2% and FP-2/hCatK + Z-FR-AMC + DMSO 2%). IC_{50} values were determined from dose-response curves varying the inhibitor concentration, while keeping fixed the concentrations of the enzyme and the substrate. The curves were fitted using GraphPad Prism version 5.03 [23]. All measurements were conducted in triplicate.

2.5. *In vitro* antiplasmodial activity assay

The antimalarial activities of the selected compounds were tested *in vitro* employing cultures of erythrocytes infected with a multi-resistant strain of *P. falciparum* (FCR3). The parasite cultures were incubated at 37 °C in a pure gas mixture of 5% O₂, 5% CO₂, and balanced N₂ environment, in RPMI 1640 medium supplemented with 25 mM HEPES, 5% (w/v) NaHCO₃, 0.1 mg/mL gentamicin, and 10% A+ heat-inactivated human serum, as previously described [24].

Several concentrations of each compound (from 200 μ M to 1.5 μ M) were evaluated. Diphosphate salt of chloroquine (CQ), $\geq 98\%$, SIGMA C6628, evaluated in a range of 0.0023 to 2.0 μ M, and *trans*-epoxysuccinyl-L-leucylamido(4-guanidino)butane (E64), 0.3125 to 20.0 μ M, were used as a treatment control in each assay. Culture medium in the absence of the compounds was used as a growth control. A suspension of parasitized erythrocytes with a hematocrit of 2%, a total parasitemia of 1%, and with a predominance of young forms, was prepared. The cultures with the treatments were incubated at 37 °C for 48 h in the presence of 5% CO₂, 5% O₂ and 90% N₂ [25]. After the incubation period, 50 μ L of 0.4% SYBR Green was added to each well and incubated for 10 min, so that the fluorochrome was intercalated into the DNA of the parasite. After a 10 min incubation, fluorescence emission detection was performed using the BD Accuri™ C6 flow cytometer (excitation at $\lambda = 485$ nm excitation, emission at $\lambda = 530$ nm) [26]. Each concentration was evaluated in duplicate, and two or more independent assays were performed. The percentages of parasitemia were analyzed using a non-linear slope-dependent regression with GraphPad Prism™ version 5.03 to estimate the IC_{50} values [23].

2.6. Cytotoxicity assays

Cytotoxicity assays were performed as described previously [27]. Briefly, HeLa cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) containing 10% FBS for 24 h at 37 °C under a 5% CO₂ atmosphere. Serial dilutions of each compound were prepared in DMEM containing 2% FPBS and 1% DMSO. Cells were incubated with compound dilutions for 72 h at 37 °C under a 5% CO₂ atmosphere. After the incubation period, viable cells were stained with crystal violet and absorption was measured at $\lambda = 590$ nm. The percentage of viability was calculated as follows:

$$\% \text{Viability} = [(A_{590\text{nm}})_{\text{sample}} / (A_{590\text{nm}})_{\text{control}}] \cdot 100$$

where $A_{590\text{nm}}$ stands for the absorbance at the indicated wavelength. Half-maximal cytotoxic concentration (CC_{50}) for each compound were estimated by nonlinear regression from dose-response curves using GraphPad Prism version 5.03 [23].

2.7. Compound parametrization for MD simulations

After experimental evaluation, the most promising compounds (= best hits) were prepared for MD simulations. Gaussian 09 [28] was employed to optimize the 3D structures of the compounds and to calculate the electrostatic surface potential (ESP), as detailed elsewhere [6]. Then, the partial atomic charges were calculated through the Restricted Electrostatic Potential (RESP) approach implemented in the

antechamber program of Amber14 package [29]. Finally, suitable non-bonded and Lennard-Jones parameters from the Generalized Amber Force Field (GAFF) were assigned to the atoms of each parametrized compound [29,30].

2.8. Molecular docking of the best hits

The binding modes of the best hits to FP-2, FP-3 and hCatK were predicted by conducting five independent random docking simulations with Autodock-vina, each generating 20 poses into the FP-2 active site [6]. Several docking poses, chosen by visual inspection, were then subjected to post-docking refinement steps as detailed below. The rationale behind the pose selection was to cover most of the conformation diversity generated by the docking simulations. Besides, we ensured that all the selected poses occluded, to some extent, the major selectivity pocket (S2 subsite) and the catalytic residues (S1 subsite) of FP-2.

Finally, the prediction of the binding modes to FP-3 and hCatK was focused only on the most favorable pose of each hit compound obtained for FP-2. The initial complex structure was generated by structural superposition using ModellerV9.9 [14]. Subsequent steps consisting of a combination of independent accelerated and conventional MD simulations were performed for each complex after appropriate equilibration steps; see below for more details.

2.9. Conventional MD simulations

The time stability of the selected FP-2 complexes was assessed through MD simulations in order to identify stable binding modes. Each complex was centered in an octahedral solvation box with a minimal distance of 10.5 Å between the solute surface and the box edges, and TIP3P water molecules were added. Electroneutrality was ensured by adding sufficient counterions to each solvation box (Na⁺ for FPs and Cl⁻ for hCatK). Protein parameters were derived from Amber 14SB force-field [31].

The solvated complex structures were relaxed through two consecutive energy minimization (EM), protocols. Then, each system was heated in the NVT ensemble and then equilibrated in the NPT ensemble at $p = 1$ bar and $T = 298$ K, keeping all complex heavy atoms restrained in both steps. Finally, production NPT runs were conducted for up to 100 ns. The setup of EM and MD simulations was similar to that reported in a previous work [6]. The overall protocol was also applied to carry out independent 100 ns MD simulations of free FP-2 and hCatK.

2.10. MM-GBSA free energy calculations

Average effective free energies (ΔG_{eff}) were calculated using the GB-neck and GB-neck2 models available in MMPBSA.py program of Amber 14 package [29,32]. Both models were selected as they yielded good correlations with experiment results in a previous work studying the interaction of FP-2 with peptide-based inhibitors [6]. The system setup for MM-GBSA free energy calculations can be found in more details in ref. [6].

In addition, the configurational entropy (ΔS_{conf}) was calculated by normal mode analysis using MMPBSA.py of Amber14 [29,32]. In total, 350 frames evenly distributed in the respective 100 ns MD simulations of the studied complexes were subjected to 50,000 cycles of conjugate-gradient EM in vacuum. All other parameters were set to the default values [29]. The number of frames selected for normal mode analysis was sufficient to attain a standard error of the mean ≤ 0.4 kcal/mol in the predicted entropic contribution values ($T\Delta S_{\text{conf}}$). The MM-GBSA free energies including the entropic contribution term will be herein-after denoted as ΔG_{bind} .

2.11. Accelerated MD simulations

aMD simulations were employed to enhance the phase space sampling of the protein-ligand complexes studied here. The lowest MM-GBSA energy poses of the ligands in complex with FP-2, and those inferred for the interaction with FP-3 and hCatK, were subjected to aMD simulations. The dihedral boosting parameter values (E_D and α_D) were calculated through empirical formulae published elsewhere [29,33]. No total potential energy boost was applied in this work.

Eleven independent 100 ns aMD simulations with identical boosting parameters were performed for each studied complex. Trajectories were then independently clustered as detailed below, and the central structures of each cluster were subjected to single-structure MM-GBSA free energy calculations [10,11] to identify the lowest-energy binding mode generated through aMD simulations. Finally, after the equilibration steps, 100 ns MD simulations were carried out for the identified lowest-energy complexes and mean MM-GBSA free energies were calculated based on the generated ensembles.

2.12. Trajectory analysis

Root-Mean-Square Deviation (RMSD) values and representative structures were determined with cpptraj module of Amber14 package [29,34]. The average-linkage algorithm [35] was employed to calculate representative structures, taking into account the RMSD values for the heavy atoms of the ligand and the protease interface residues, *i.e.*, those contacting the ligand at a distance ≤ 4 Å in the starting structure. For aMD simulations, each 100 ns aMD simulation was split into five clusters. Root mean square fluctuations (RMSF) for the FP-2 residues were also calculated during the MD simulations using cpptraj [29].

2.13. Principal component analysis

Principal Component Analysis (PCA) [36] were performed to compare the fluctuations of active site loops of FP-2 free and bound states in two-dimensional subspaces defined by the main eigenvectors, also known as principal components (PCs). The loops, termed L1, L2, L3, L4 and L5, comprised the following residues 35–41, 77–84, 150–153, 171–174 and 232–234, respectively. The PCA was conducted for the main-chain plus C β atoms of such loops, to take into account both, main-chain and side chain fluctuations. The gmx covar and gmx anaenig programs of Gromacs 5.1.4 were employed to conduct PCA [37].

2.14. Umbrella sampling free energy calculations

Standard binding free energies (ΔG°) for the formation of the complexes of interest were calculated through umbrella sampling (US), following the protocol described by Doudou et al. [38]. The previous protocol was adapted in order to run the MD simulations in Gromacs 5.1.4 (see Text S1 for more details).

3. Results

3.1. In silico identification of selective FP-2 inhibitors from Maybridge HitFinder™ database

The screening of large databases of chemical compounds against a parasite target and its human off-targets is a widely used strategy to identify selective inhibitors of relevant chemotherapeutic targets in neglected diseases [1,4,39,40]. In this work, the Maybridge Hitfinder™ database was screened to discover putative selective inhibitors of FPs. To identify a relevant off-target of FPs for SBVS among the eleven cathepsins expressed in humans [41], those with available 3D structures, *i.e.*, K, L, C, B, V, X, F and S, were retrieved from the PDB after a PSI-BLAST analysis. The structural alignment of the human cathepsins and FPs (Fig. S1) allowed the identification of hCatK as the human off-target

Table 1
Autodock-Vina scores for the selected compounds.^a

Compound	S_{vina} (FP3) (kcal/mol)	ΔS_{vina} (FP3- hCatK) (kcal/ mol)	ΔS_{vina} (FP2) (kcal/mol)	ΔS_{vina} (FP2- hCatK) (kcal/ mol)
HTS01959	-8.9	-1.2	-8.0	-0.3
SEW01979	-8.4	-1.4	-9.1	-2.1
HTS10627	-8.7	-1.5	-8.8	-1.6
BTB08650	-9.0	-1.1	-8.0	-0.1
HTS07377	-8.6	-1.3	-8.2	-0.9
HTS08262	-9.4	-1.1	-9.6	-1.3
HTS12139	-8.5	-1.4	-8.4	-1.3
HTS12945	-8.9	-1.3	-8.4	-2.1
HTS07940	-8.8	-0.9	-9.2	-1.3

^a The compounds were selected based on criteria presented in Materials and Methods.

having the largest ID with respect to both FPs in the active site region (see residues enclosed by dashed rectangles in Fig. S1). In fact, the IDs of hCatK with respect to FP-2 and FP-3 are 53.85% and 61.54%, respectively. Other proteases, such as hCatL and hCatB, were ranked in second and third places, respectively, with IDs ranging from 46.15% to 48.72% with respect to both FPs (Table S1). In addition, we took into account that FP-2, FP-3 and hCatK share the same specificity for Leu at P2; whereas hCatB and hCatL prefer Phe in that position [5,41–46]. Based on these results, we chose hCatK to assess the potential off-target inhibition of the compounds selected from SBVSs.

A set of nine compounds were chosen after processing the results of the SBVSs conducted in parallel against FP-2, FP-3 and hCatK (Table 1 and Fig. 1). These hits were also filtered using the PAINS-Remover web server (<http://cbligand.org/PAINS>) to exclude PAN Assay Interference Compounds, *i.e.*, typical false-positive compounds that react with numerous biological molecules in a non-specific fashion [47]. Note that all the compounds shown in the table fulfill the conditions of affinity and selectivity stated in Section 2.2.

3.2. Experimental evaluation of the compounds selected from SBVS

To validate SBVS predictions, the compounds shown in Table 1 were purchased and their putative inhibition against FP-2 was measured *in vitro*. In order to confirm the proposed selectivity, we also performed hCatK cross-inhibition assays. Seven compounds inhibited FP-2 in a dose-dependent fashion (Fig. 2), confirming our previous computational predictions. Most of the compounds tested in our work displayed IC₅₀ values against FP-2 in the order of 10⁻⁵ M (Table 2); with the weakest inhibitor showing IC₅₀ in the order of 10⁻⁴ M and the most potent one, in the low micromolar range (10⁻⁶ M). On the contrary, the compounds showed negligible inhibitory activity against human off-target in the same range of concentrations assayed for FP-2. At the highest concentration tested for each compound, the maximum inhibition against the hCatK control ranged from 8 to 25% (data not shown), clearly indicating higher inhibitory potency of the selected compounds against the parasitic enzyme, as intended in this work.

In parallel, we assayed the antiparasitic activity of the selected compounds using E64 and CQ as positive controls. Five compounds showed no activity against *P. falciparum* at the tested concentrations. In contrast, four compounds (HTS07377, BTB08650, HTS08262 and HTS07940) possessed measurable inhibitory activities on *P. falciparum* growth, displaying IC₅₀ values in the 10⁻⁴ - 10⁻⁶ M range (Table 2). To further evaluate the selectivity of the observed cellular activity, the four active compounds were subjected to evaluation of general cytotoxicity activity against HeLa cells (Table 2).

After the experimental evaluations, two compounds displaying the greatest potential, *i.e.*, HTS07940 and HTS08262, were identified (Table 2). In addition to moderate and selective FP-2 inhibition with respect to hCatK, they showed suitable cellular selectivity indices (46

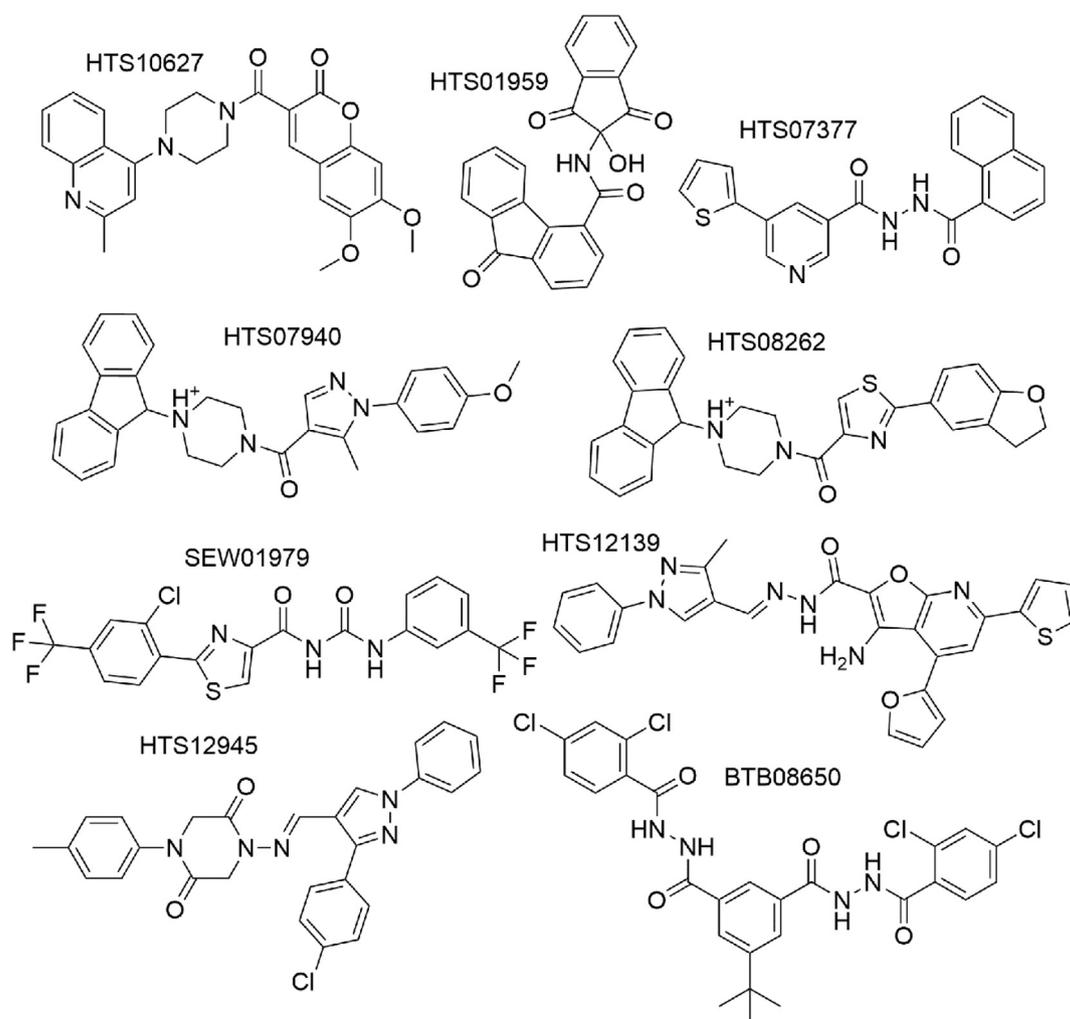


Fig. 1. Chemical structures, identifiers and IUPAC names of the compounds selected from the SBVS. The identifier of each compound is shown on top of its structure. The IUPAC names are the following: HTS10627, 6,7-dimethoxy-3-[4-(2-methylquinolin-4-yl)piperazine-1-carbonyl]chromen-2-one; HTS01959, N-(2-hydroxy-1,3-dioxinden-2-yl)-9-oxofluorene-4-carboxamide; HTS07377, N'-(1-naphthoyl)-5-(thiophen-2-yl)nicotinohydrazide; HTS07940, [4-(9H-fluoren-9-yl)piperazin-1-yl]-[1-(4-methoxyphenyl)-5-methylpyrazol-4-yl]methanone; HTS08262, [2-(2,3-dihydro-1-benzofuran-5-yl)-1,3-thiazol-4-yl]-[4-(9H-fluoren-9-yl)piperazin-1-yl]methanone; SEW01979, 2-[2-chloro-4-(trifluoromethyl)phenyl]-N-[[3-(trifluoromethyl)phenyl]carbamoyl]-1,3-thiazole-4-carboxamide; HTS12139, 3-amino-4-(furan-2-yl)-N-[(E)-(3-methyl-1-phenylpyrazol-4-yl)methylideneamino]-6-thiophen-2-ylfuro[2,3-b]pyridine-2-carboxamide; HTS12945, 1-[(E)-[3-(4-chlorophenyl)-1-phenylpyrazol-4-yl]methylideneamino]-4-(4-methylphenyl)piperazine-2,5-dione; BTB08650, 5-tert-butyl-1-N',3-N'-bis(2,4-dichlorobenzoyl)benzene-1,3-dicarboxamide.

and 10, respectively). Remarkably, the IC_{50} value of HTS07940 against *P. falciparum* culture was similar to that of E64, the prototypic inhibitor of C1 cysteine proteases [5], and one order-of-magnitude lower than that of CQ (Table 2). Overall, the inhibitory activities of HTS07940 and HTS08262 against *P. falciparum* cultures are in the typical order of magnitude shown by many non-peptidic FP-2 inhibitors identified so far [1]. However, the most promising feature of the identified hits is their suitable selectivity indices (Table 2).

It is noteworthy that HTS07940 and HTS08262 share a common scaffold based on (4-(9H-fluoren-9-yl) piperazin-1-yl) *R*-yl methanone. Here, *R* stands for different derivatives of a core structure comprising a five-membered and a six-membered aromatic ring linked together by a single covalent bond (see Fig. 1). Remarkably, we realized that a similar compound, Genz10850, in which *R* is the indol group, has been previously characterized *in vitro* as an inhibitor of reduced nicotinamide adenine dinucleotide (NADH)-dependent enoyl (acyl carrier protein) reductases (ENR) of *Mycobacterium tuberculosis* (InhA, IC_{50} = 0.16 μ M) and *P. falciparum* (PfENR, IC_{50} = 18 μ M) [48]. Genz10850 also inhibits *in vitro* the growth of different *P. falciparum* strains with IC_{50} values ranging from 14 to 31 μ M. The authors also reported the crystal

structure of InhA in complex with Genz10850 (PDB: 1P44). Using the bound conformation of this compound, we built a model of a putative PfENR:HTS07940:NAD⁺ ternary complex through structural superposition with a crystal structure of PfENR (PDB: 2OL4), see Text S2 and Fig. S2, for full details. The results obtained from the MD simulations of the previous complex suggest that HTS07940 (and, by inference, HTS08262) might interact with PfENR. Thus, the inhibition of *P. falciparum* growth exerted by HTS07940 and HTS08262 (Table 2) might arise from a combined inhibitory effect on both enzymes. This prediction constitutes an open question requiring future experimental validation and is beyond the scope of the current work. We also want to stress that, despite the similarities between Genz10850 and HTS07940/HTS08262, the former lacks the five-membered ring that connects the final aromatic group to the carbonyl carbon.

3.3. Prediction of the binding modes of HTS08262 and HTS07940 to FP-2

The binding modes to FP-2 of the two most promising compounds, HTS08262 and HTS07940, were predicted after subjecting docking poses to MD simulations. Since both compounds possess a common

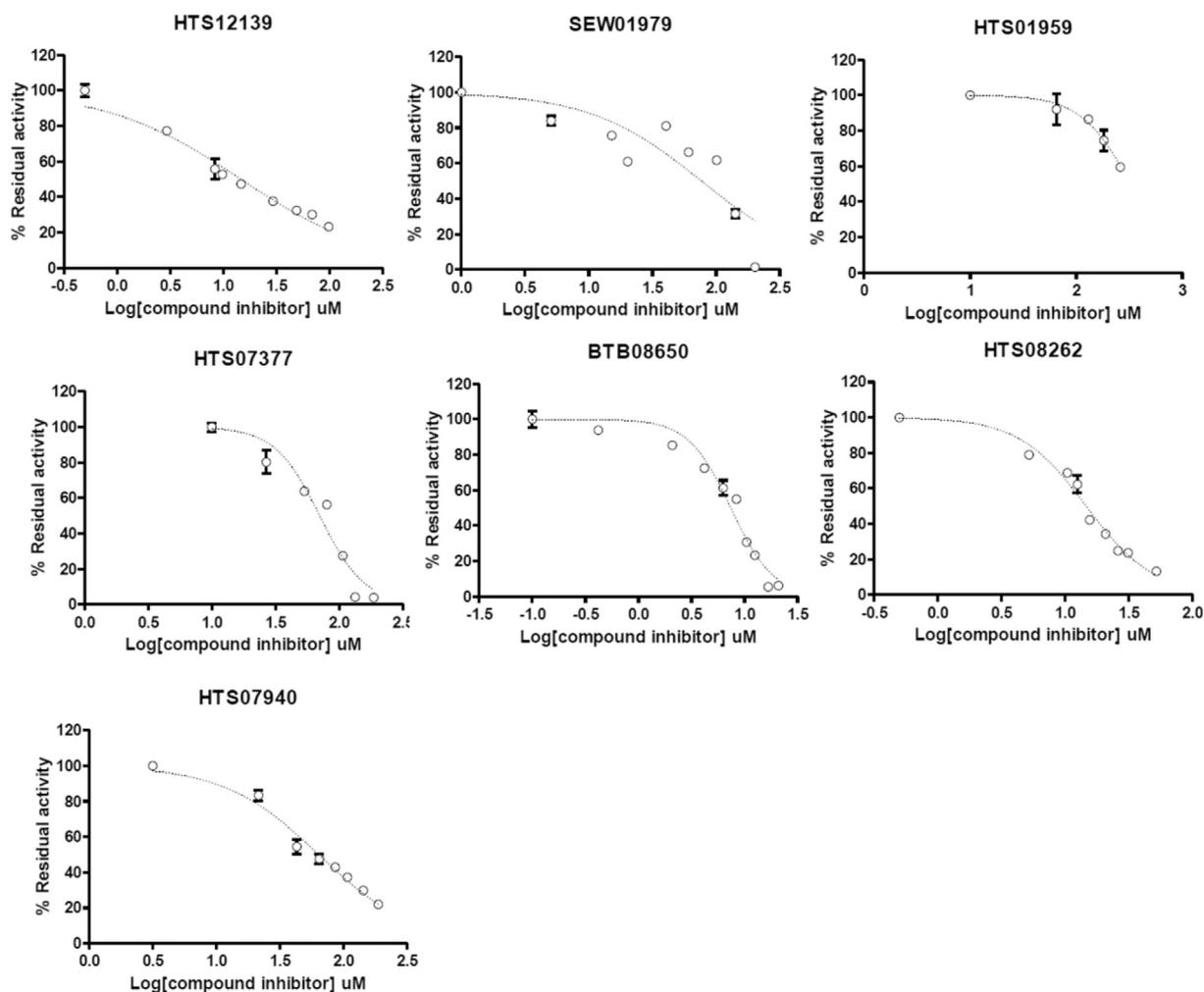


Fig. 2. Dose-response curves for seven FP-2 inhibitors identified by SBVS. The percents of FP-2 residual activity were measured using at least eight increasing concentrations of the inhibitors. IC_{50} values were determined by fitting experimental data to the four-parameter model of GraphPad Prism.

scaffold, we assumed that they must share a similar binding mode to the enzyme. Therefore, a refinement of HTS07940 docking poses into the FP-2 active site was conducted first to find the most stable binding mode. Then, an inferred lowest-energy pose of HTS08262 was analyzed taking into account the results obtained for HTS07940.

A set of 14 docking poses, clustered into six groups (labeled from A

to F) based on the ligand RMSD values (Fig. S3), were subjected to 100 ns MD simulations. Individual poses within each group were identified with a number following the cluster label. The final structures generated through MD simulations (Fig. S4), allowed the identification of two stable binding modes, *i.e.*, pose A3 and pose F1. Furthermore, pose F2 transitioned to a conformation similar to pose A3

Table 2

IC_{50} values and selectivity indices for the selected compounds.

Compound	IC_{50} (FP-2) (μ M)	IC_{50} (hCatK) (μ M)	Selectivity index ^a	IC_{50} (<i>Pf</i>) ^b (μ M)	CC ₅₀ (HeLa) (μ M)	Selectivity index ^c
HTS01959	326(56) ^d	> 261	> 1	Inactive	–	–
SEW01979	77(25)	> 142	> 1	Inactive	–	–
HTS10627	n.d. ^e	n.d.	n.d.	Inactive	–	–
BTB08650	7.4(0.5)	> 314	> 42	29(5)	34(5)	1.17(0.03)
HTS07377	71(8)	> 469	> 6	163(6)	< 33.5	< 0.2
HTS08262	14.7(0.8)	> 521	> 35	34(7)	350(14)	10(2)
HTS12139	15(2)	> 294	> 20	Inactive	–	–
HTS12945	n.d.	n.d.	–	Inactive	–	–
HTS07940	64(5)	> 565	> 8	2.91(0.04)	133(10)	46(3)
E64	– ^f	–	–	1.3(0.1)	–	–
CQ	–	–	–	0.14(0.01)	–	–

^a Selectivity index for FP-2 calculated as follows: Selectivity index = $IC_{50}(\text{hCatK})/IC_{50}(\text{FP-2})$.

^b *Pf* stands for a culture of *P. falciparum*-infected erythrocytes.

^c Selectivity index for parasite cells calculated as follows: Selectivity index = $CC_{50}(\text{HeLa})/IC_{50}(\text{Pf})$.

^d Standard deviations calculated from three independent measurements are shown between parentheses.

^e IC_{50} values were not measured due to low water solubility of the compounds.

^f Measurement not performed.

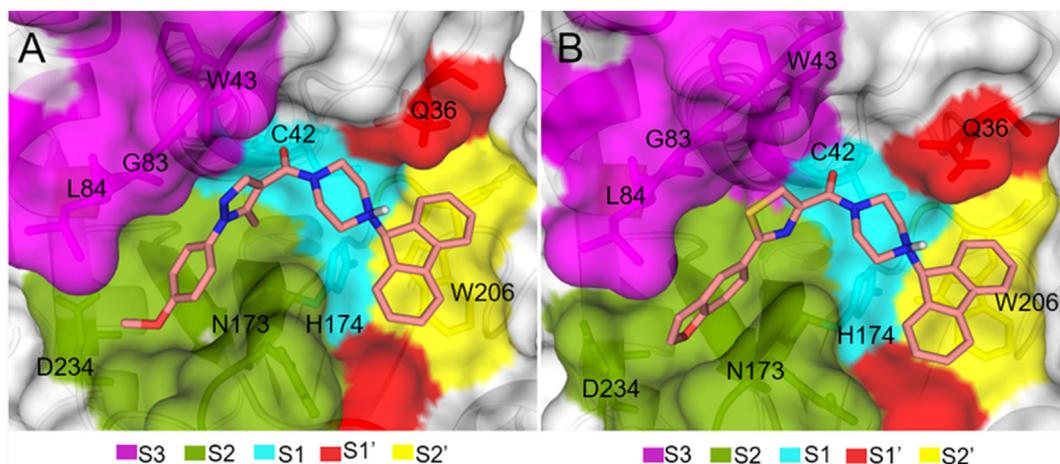


Fig. 3. Structures of FP-2 in complex with compounds HTS07940 and HTS8262. A) Three-dimensional representation of the FP2:HTS07940 complex interface. B) Three-dimensional representation of the FP2:HTS8262 complex interface. FP-2 subsites are colored according to the legend. The depicted interfaces correspond to the representative structures generated by clustering each 100 ns MD trajectory into a single cluster, and considering only the RMSD values for the heavy atoms of the protein interface residues and the ligand.

during the MD simulation (Fig. S4). Subsequent MM-GBSA free energy calculations based on the last 100 frames of each trajectory, identified pose A3 as the most stable one (Table S2).

To further enhance the sampling of the conformational space accessible to pose A3 of HTS07940, we subjected to aMD simulations eleven frames collected from the 100 ns MD simulation of this pose. Based on MM-GBSA ranking of central structures of the aMD simulations, we predicted pose A3b as the most likely binding mode of compound HTS07940 to FP-2 (Fig. S5). The structural features of the complex interface are depicted in Fig. 3A. Furthermore, the most-likely binding mode of compound HTS08262 to FP-2 inferred from pose A3b of HTS07940 is shown in Fig. 3B. The binding modes fulfill previous structural determinants for the interaction with FP-2. The condensed aromatic rings of both compounds occupy the S2' of the enzyme, where they are likely to establish π - π stacking interactions with the conserved residue W206, as has been described for other complexes involving C1 cysteine proteases [49]. In addition, the 2,3-dihydrobenzofuran-5-yl and *p*-methoxyphenyl moieties of HTS08262 and HTS07940, respectively, are placed within the S2 subsite, the major specificity determinant of FP-2 and most C1 cysteine proteases [5,49,50]. This is consistent with the well-known preference of FP-2 for hydrophobic (aliphatic and aromatic) amino acid side chains at P2 position [5,42,51]. Moreover, the carbonyl groups of both compounds lie within the S1 subsite, in agreement with the already-established proneness of electrophilic groups to interact with the catalytic residue (C42) [52].

Remarkably, the proposed binding mode of HTS07940 (Fig. 3A) was not found among the docking poses (Fig. S3); nor was it obtained after the subsequent 100 ns MD simulation (Fig. S4). In fact, we resorted to aMD simulations to generate more stable poses. By comparing the predicted representative structures of the FP-2:HTS08262 and FP-2:HTS07940 complexes with the crystal structure of FP-2 used as starting point for SBVS (PDB: 2OUL), we noticed some conformational changes in the active site region (Fig. 4A). These changes, which mainly involved five loops (L1, L2, L3, L4 and L5), led to a more open conformation of the active site in the complex structures (see arrows in Fig. 4A). The previous phenomenon, in turn, allowed a tighter fit of the compounds into the FP-2 binding cleft, as corroborated by MM-GBSA free energy calculations (Fig. S5).

We further investigated the motions of loops L1-L5 in free FP-2 and in the FP-2:HTS07940 complex through PCA. As shown in Fig. 4B, these loops correspond to the most flexible regions in FP-2 active site; therefore, they account for the most significant motions in this region of the protein. Moreover, the three main eigenvectors calculated through PCA account in total for nearly 50% of the L1-L5 fluctuation in free FP-2

(Fig. 4C). The 2D projections of L1-L5 fluctuations in free FP-2 (blue) and when bound to HTS07940 (red) onto the PC subsets 1,2 and 1,3 calculated from the MD simulations of free enzyme, indicate the overlap in the phase space accessible to both states (Fig. 4D and E). Furthermore, the fluctuation of loops L1-L5 in the bound state samples just a fraction of the phase space of free FP-2 according to the projections. This result, though consistent with a decreased loop flexibility in the FP-2 bound state, has to be taken with care, as the simulation time for the free and bound states are quite different. Our major purpose with the current analysis was to demonstrate that conformations similar to the bound state are sampled by free FP-2, as can be straightforwardly deduced from the phase space overlap in the chosen PCs.

We also projected the crystal structure of FP-2 (PDB: 2OUL) onto the PCs 1,2 and 1,3 of free FP-2 (see green dot, Fig. 4D and E), and noticed that the conformation of loops L1-L5 in the crystal structure lies outside the ensemble of conformations generated from the MD simulation of FP-2 bound to HTS07940. To reinforce this result, we compared the crystal structure and a frame extracted from the MD simulation of free FP-2, with PCs lying within the region that overlaps the bound state distribution in Fig. 4D and E (Fig. S6). Note that the frame extracted from the MD simulation largely matches the conformation of loops L1-L5 of FP-2 bound to HTS07940 (RMSD 1.57 Å), while the crystal structure differs more appreciably from the latter (RMSD 2.21 Å). The surface representations of the active sites of the MD frame and the crystal structure show that compound HTS07940 only fits well into the former (Figs. S6B and S6C). This corroborates our previous observation that the lowest-energy pose could not be obtained from docking simulations into the crystal structure binding site (Fig. S3). Overall, our results suggest that HTS07940 can bind to pre-existing free FP-2 conformations, which constitutes the basis of the so-called conformational selection mechanism [53]. The previous conclusion is also extensible to compound HTS08262 (data not shown).

3.4. Prediction of the binding modes of HTS07940 to FP-3 and to hCatK

The binding mode of HTS07940 to FP-3 was predicted by subjecting pose A3' (initially obtained by superposition of pose A3 bound to FP-2 onto FP-3) to multiple independent aMD simulations. Five central structures obtained from these simulations are shown in Fig. S7. Note that pose A3'b is similar to the predicted binding mode of HTS07940 to FP-2, and has the lowest ΔG_{eff} value (compare Figs. S5 and S7). The subsequent refinement of pose A3'b through a 100 ns MD simulation yielded the representative structure depicted in Fig. 5A. The predicted structure of HTS07940 in complex with hCatK was obtained by means

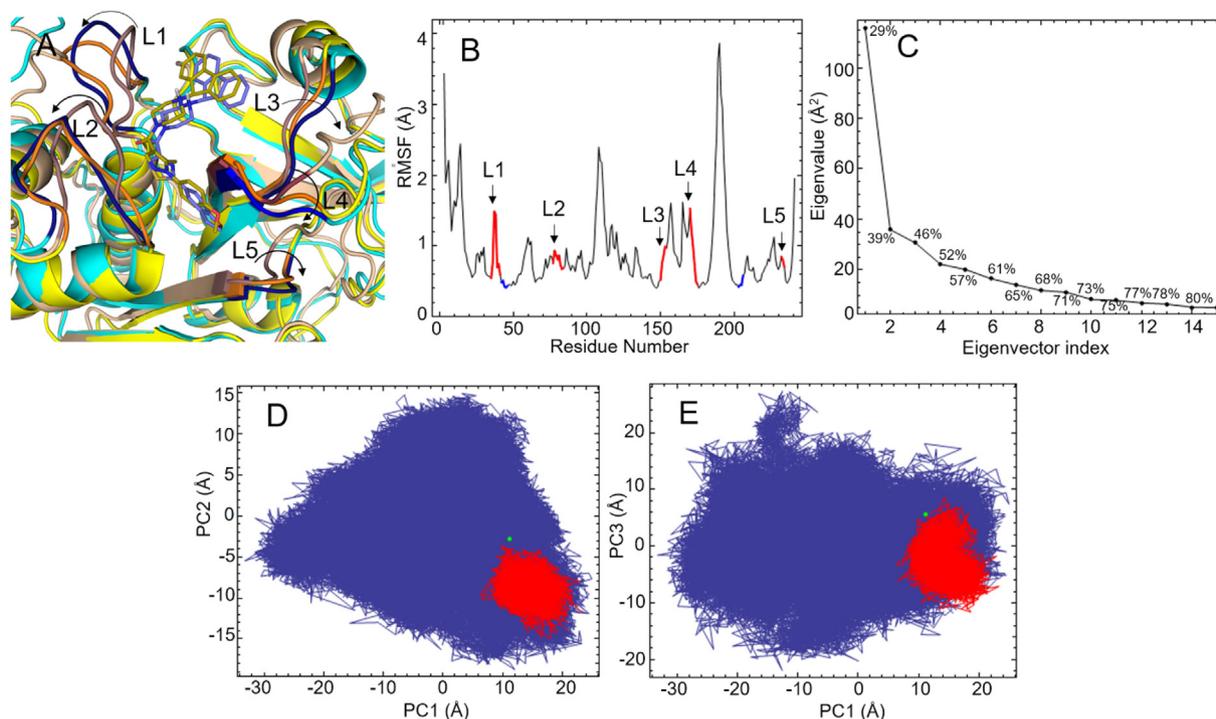


Fig. 4. Analysis of the dynamics of loops L1-L5 in free and bound FP-2. **A)** Superposition of the crystal structure of FP-2 (salmon, PDB: 2OUL) onto the representative structures of FP-2 bound to HTS07940 (yellow and lemon) and HTS08262 (cyan and blue). Arrows indicate the main backbone conformational changes in the active site region of superimposed structures involving loops L1, L2, L3, L4 and L5, depicted in dark colors. **B)** Per residue main-chain + C β RMSF values calculated during the MD simulation of free FP-2. Curve sections corresponding to RMSF values for the active site loops L1-L5 are colored in red, whereas those corresponding to less flexible interface residues are depicted in blue. **C)** Eigenvalues for the 15 top-ranked eigenvectors obtained from PCA of loops L1-L5 during a concatenated 1.2 μ s MD simulation of free FP-2. The percent of cumulative fluctuation (CF) is shown beside each point in the graph for eigenvectors 1 to 14. **D)** and **E)** Two-dimensional projections of L1-L5 fluctuations in free FP-2 (blue) onto the PCs 1, 2, and 1,3, respectively. In both graphs, the fluctuations of loops L1-L5 sampled during a 100 ns MD simulation of the FP-2:HTS07940 complex are projected onto the corresponding PCs of free FP-2 (red). The green dot in each graph corresponds to the equivalent projection of FP-2 crystal structure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of a similar strategy, fully described in Text S3. The interface of this complex is depicted in Fig. 5B.

Residues Y67 and L209 have long been recognized as crucial to determine the specificity of hCatK S2 subsite [54]. The steric wall created by these residues at the bottom of the S2 subsite (Fig. S10A) seems to limit the size of the P2 moieties that can fit into it. Remarkably, our current predictions suggest that, instead of a rock-solid wall, the side chains of Y67 and L209 can act like a gate controlling the

access to a distal region of the S2 subsite in hCatK (compare Figs. 5B and S10A). A similar ‘gating’ phenomenon, involving the equivalent pairs of residues W69-V209 and L67-E205, was predicted for the papain-like proteases *Fasciola hepatica* cathepsin L3 and cruzain, respectively [55,56]. Nonetheless, one must take into account that the eventual opening the Y67-L209 gate in hCatK has an energetic cost of ~ 2.6 kcal/mol that reduces the binding affinity for the ligand (Table S3). It is noteworthy that the equivalent residues Y86 and E236 of FP-3

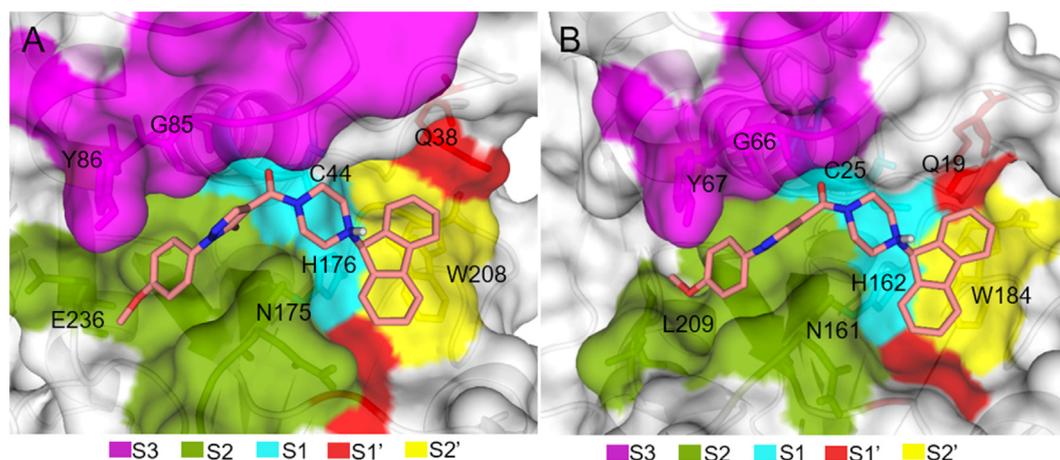


Fig. 5. Structural representation of the interfaces of compound HTS07940 bound to FP-3 and to hCatK. **A)** FP-3 in complex with HTS07940. **B)** hCatK in complex with HTS07940. Enzyme subsites are colored according to the legend. Representative structures were generated by clustering each 100 ns MD trajectory into a single cluster, and considering only the RMSD values for the heavy atoms of the protein interface residues and the ligand.

Table 3
MM-GBSA free energy values for the studied complexes.

Energy components ^a (kcal/mol)	Complexes			
	FP-2:HTS08262	FP-2:HTS07940	FP-3:HTS07940	hCatK:HTS07940
ΔG_{npol}^b	-55.21	-54.94	-54.98	-54.12
ΔG_{pol}^c (GBn) ^c	17.76	18.16	20.10	20.06
ΔG_{pol}^c (GBn2) ^c	18.45	19.20	21.09	20.34
$-T\Delta S_{conf}^d$	21.3(0.4) ⁱ	23.6(0.4)	24.2(0.4)	22.6(0.3)
ΔG_{eff}^e (GBn) ^e	-37.45(0.05)	-36.78(0.08)	-34.88(0.05)	-34.06(0.04)
ΔG_{eff}^e (GBn2) ^e	-36.76(0.04)	-35.74(0.08)	-33.89(0.05)	-33.78(0.04)
ΔG_{bind}^f (GBn) ^f	-16.2(0.4)	-13.1(0.4)	-10.7(0.4)	-11.5(0.3)
ΔG_{bind}^f (GBn2) ^f	-15.5(0.4)	-12.1(0.4)	-9.7(0.4)	-11.2(0.3)
$\Delta\Delta G_{eff}^g$ (GBn) ^g	-0.67(0.09)	-	1.90(0.09)	2.72(0.09)
$\Delta\Delta G_{eff}^g$ (GBn2) ^g	-1.02(0.09)	-	1.85(0.09)	1.96(0.09)
$\Delta\Delta G_{bind}^h$ (GBn) ^h	-3.1(0.6)	-	2.4(0.6)	1.6(0.5)
$\Delta\Delta G_{bind}^h$ (GBn2) ^h	-3.4(0.6)	-	2.4(0.6)	0.9(0.5)

^a All energy components were calculated based on 100 ns MD simulations of the complexes.

^b Non-polar free energy calculated as follows: $\Delta G_{npol} = \Delta E_{vw} + \Delta G_{SA}$, where ΔE_{vw} and ΔG_{SA} are the variations of the van der Waals energy and of the surface tension, respectively, associated to the ligand binding.

^c Polar free energy calculated as follows: $\Delta G_{pol} = \Delta E_{el} + \Delta G_{GB}$, where ΔE_{el} and ΔG_{GB} are the variations of the electrostatic energy in vacuum and of the polar solvation energy calculated with the indicated GB models, respectively.

^d Contribution of configurational entropy at $T = 298$ K to the binding free energy calculated from Normal Mode Analysis.

^e Effective free energy calculated by the following equation: $\Delta G_{eff} = \Delta G_{pol} + \Delta G_{npol}$. The ΔG_{pol} term was determined using the GBn and GBn2 models.

^f Binding free energy calculated as follows: $\Delta G_{bind} = \Delta G_{eff} + (-T\Delta S_{conf})$.

^g Relative effective free energy calculated as follows: $\Delta\Delta G_{eff} = \Delta G_{eff}(X) - \Delta G_{eff}(FP-2:HTS07940)$, where X stands for the FP-2:HTS08262, FP-3:HTS07940 or hCatK:HTS07940 complexes.

^h Same as ^g, but using the ΔG_{bind} values instead of the ΔG_{eff} ones.

ⁱ Standard errors of the mean written between parentheses. The standard errors of ΔG_{npol} and ΔG_{pol} are not shown, since these energy components are not the standard output of the MMPBSA.py calculations and, thus, we did not have access to the error values.

tend to adopt a more open conformation in the crystal structure of the enzyme complexed with an organic inhibitor bearing Leu at P2 (PDB: 3BWK, see Fig. S10B). In fact, residue E236 of FP-3 is not encroaching into the S2 subsite, as L209 of hCatK does. Instead, it bends away to form a hydrogen bond with Y238, both in the crystal structure 3BWK and during nearly 72% of the MD simulation of FP-3 in complex with HTS07940 (Figs. S10C and S10D). Additionally, a close inspection of the predicted hCatK:HTS0940 structure (Fig. 5B) reveals that, even in the open state of the 'gate', L209 is poised to sterically hinder a deeper accommodation of the *p*-methoxyphenyl moiety of the compound into the S2 bottom. In agreement with this observation, some authors have attributed a key role to this sole residue in preventing the accommodation of bulky groups within the S2 subsite of hCatK [57]. The latter may be the most important factor in determining the lower affinity of HTS07940 and HTS08262 for hCatK with respect to that for FP-2 (Table 2).

3.5. Calculation of the binding free energies for the studied complexes

After the binding mode prediction, we carried out MM-GBSA free energy calculations for the studied complexes using two GB models (Table 3). Remarkably, the relative effective free energies between the FP-2:HTS08262 and the FP-2:HTS07940 complexes ($\Delta\Delta G_{eff} = -0.67$ kcal/mol), correctly predicted the higher affinity of the former complex (compare Table 2 and Table 3) and were also close to the experimental value (-0.87 kcal/mol). When including the contribution of configurational entropy, the prediction does not change qualitatively, but the relative affinity between both complexes was overestimated by > 2 kcal/mol (see $\Delta\Delta G_{bind}$ values in Table 3). Based on the literature [58,59], it is not surprising that the inclusion of entropy in MM-GBSA calculations did not improve the results, especially when dealing with very similar complexes.

We also extended the MM-GBSA free energy calculations to the complexes involving compound HTS07940 bound to FP-3 and hCatK (Table 3). This compound was predicted to possess lower affinities for FP-3 than for FP-2 in a consistent fashion by both GB models. The

inclusion of the configurational entropy term did not alter the previous result. Hence, FP-2 seems to be a more important target of compound HTS07940 in *P. falciparum* than FP-3. Moreover, HTS07940 shows higher selectivity for FP-2 than for the human off-target hCatK, with relative free energies ranging from 0.9 to 2.7 kcal/mol, depending on the GB model and the inclusion of the entropy term (Table 3). Of note, the previous predictions are in agreement with the experimental relative affinity of HTS07940 for hCatK and FP-2, which is > 1.2 kcal/mol favorable to the latter enzyme (value calculated from selectivity index shown in Table 2).

In addition, we employed US calculations to complement the previous predictions using a method based on more solid theoretical principles. The PMFs for the four complexes studied here are shown in Fig. 6. The reaction coordinate, *i.e.*, the Cartesian *z* component of the distance vector between C α of the catalytic cysteine and the amide nitrogen of HTS07940 and HTS08262 (Fig. 1), was increased up to ~ 42 Å in all cases (Fig. 6). All PMFs reached a plateau beyond $z \approx 20$ Å, which was used as a cutoff to define the bound ($z \leq 20$ Å) and unbound ($20 \text{ Å} < z \leq 42 \text{ Å}$) regions, following the same criterion used by the authors of the method [38]. Remarkably, the predicted ΔG° values for the binding of HTS07940 and HTS08262 to FP-2 are in good agreement with the experimental ones (Table 4). In fact, they overestimated the experimentally-determined affinities only by 0.93 and 0.53 kcal/mol, respectively, which lie within the range of accuracy of the method reported in literature [38]. Furthermore, the relative affinity of both compounds for FP-2 ($\Delta\Delta G^\circ = 0.46$ kcal/mol) matches the experimental value (0.87 kcal/mol). Overall, both the MM-GBSA and US free energy calculations consistently predicted the higher affinity of compound HTS08262 for FP-2 and relative free energies that correlated well with the experimental values. This, in turn, suggests the validity of the binding modes of HTS07940 and HTS08262 to FP-2 proposed in our work.

The ΔG° values for the binding of HTS07940 to FP-3 and hCatK were also calculated through US (Fig. 6 and Table 4). The compound was predicted to have a lower affinity for FP-3 and hCatK than for FP-2. This is agreement with the previous MM-GBSA calculations (Table 3) and, in

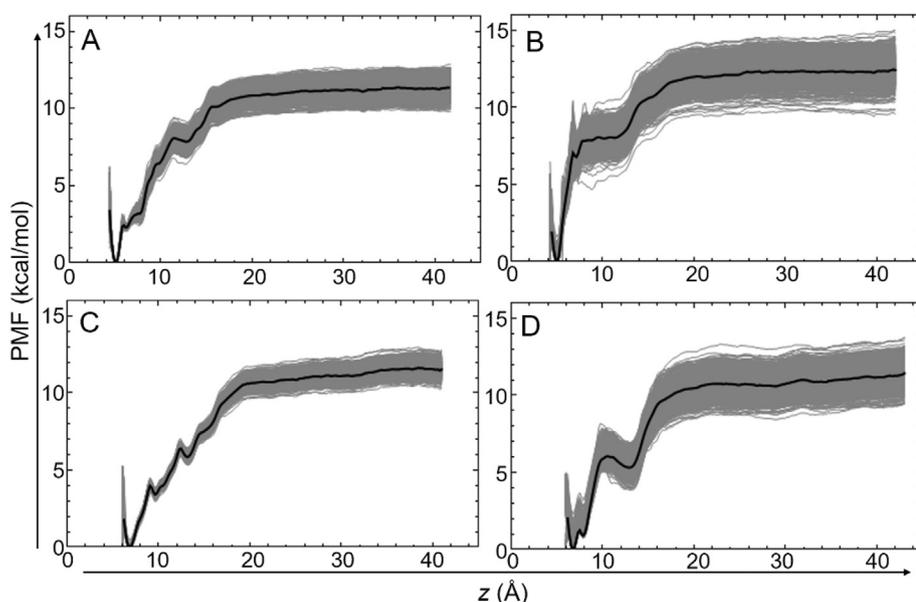


Fig. 6. PMFs along the z coordinate for the following complexes: A) FP-2:HTS07940, B) FP-2:HTS08262, C) FP-3:HTS07940 and D) hCatK:HTS07940. The thick black line in each graph represents the average PMF calculated from 1000 individual PMFs obtained by bootstrapping new trajectories from umbrella histograms (gray lines).

the case of hCatK, with the experimental measurements (Table 2). Note, however, that the $\Delta\Delta G^\circ$ values with respect to FP-2 (0.26 kcal/mol in both cases) are rather small if compared to the relative affinities predicted through MM-GBSA free energy calculations, and with the experimental value for hCatK (> 1.2 kcal/mol). It is important to highlight that the US method employed here led to variations of 3.5 kcal/mol in the ΔG° values for the same system when varying the orthogonal restraint constants (see Table 2 in ref. [38]). These variations can even become larger when changing the path orientation (see Fig. 6 in ref. [38]). In our work, we tried to minimize the divergence caused by the previous variables on the individual ΔG° values by structurally aligning the complexes prior to pulling on the z direction and by applying the same orthogonal restraints to all complexes. Nonetheless, the accuracy of this method in ranking affinities has not been assessed yet. Therefore, we considered a good result the fact of having predicted ΔG° values for the interaction of HTS07940 and HTS08262 with FP-2 close to the experimental ones, and the correct experimental trend of the relative affinity of the former compound for FP-2 and hCatK.

4. Discussion

Two promising compounds (HTS07940 and HTS08262) in terms of

their inhibitory activity against *P. falciparum* and lower toxicity for HeLa cells, were identified in the current work. The compounds were selected from the results of SBVSs using FP-2 and FP-3 as target molecules. We kept into account the specificity issues since the beginning of our *in silico* strategy by excluding compounds with potential strong affinity for hCatK from our selection. Compounds HTS07940 and HTS08262 share a common scaffold consisting of a (4-(9H-fluoren-9-yl) piperazin-1-yl) methanone moiety that binds to a variable substituent, whose core structure comprises a five-membered and a six-membered aromatic ring attached to each other through a single covalent bond. The conserved (4-(9H-fluoren-9-yl) piperazin-1-yl) methanone moiety have been found in a compound (Genz10850) previously characterized as an inhibitor of PfENR and of *P. falciparum* growth in *in vitro* conditions [48]. However, the substituent of Genz10850 attached to the (4-(9H-fluoren-9-yl) piperazin-1-yl) methanone moiety is an indol group, which differs from the core structure found in the compounds identified here.

The structural similarities between Genz10850 and the inhibitors reported here encouraged us to assess *in silico* the capacity of HTS07940 to inhibit PfENR in addition to FP-2. Our results based on MD simulations suggested that HTS07940 might bind the former enzyme as well, a conclusion extensible, in principle, to HTS08262. Therefore, HTS07940

Table 4
Standard binding free energy and its components obtained from US calculations on the studied complexes.

	FP-2:HTS08262	FP-2:HTS07940	FP-3:HTS07940	hCatK:HTS07940
l_b^a (Å)	0.52	0.59	0.79	0.67
ΔW_R^b (kcal/mol)	-12.31(0.02)	-11.22(0.02)	-11.21(0.01)	-11.01(0.02)
ΔG_{PMF}^c (kcal/mol)	-10.08(0.02)	-9.07(0.02)	-9.25(0.01)	-8.93(0.02)
ΔG_R^d (kcal/mol)	-0.52(0.01)	-1.06(0.02)	-0.63(0.01)	-0.98(0.02)
ΔG^{re} (kcal/mol)	-7.48(0.02)	-7.02(0.03)	-6.76(0.01)	-6.76(0.03)
ΔG^{exp} (kcal/mol)	-6.95(0.02)	-6.09(0.03)	-	> -4
$\Delta\Delta G^{re}$ (kcal/mol)	0.46(0.04)	-	0.26(0.03)	0.26(0.04)

^a Configurational integral of the PMF in the bound region, calculated through eq. 4 in ref. [38].

^b PMF depth defined by Eq. (3) in ref. [38].

^c Binding free energy change between the bound and unbound sections of the PMF, see Eq. (1) in ref. [38].

^d Free energy associated to orthogonal xy restraint removal when the ligand is bound, calculated through Eq. (10) in ref. [38].

^e Standard binding free energy.

^f Experimental binding free energy calculated through the following equation: $\Delta G^\circ = RT \ln K_i$. The K_i values were calculated, in turn, based on the IC_{50} values shown in Table 2 through the expression relating the K_i to the IC_{50} value for a competitive inhibitor [60]. The K_M of FP-2 used in the calculations was 17.63 μ M and the initial substrate concentration in the assays was 15 μ M.

^g Relative binding free energy calculated as follows: $\Delta\Delta G^\circ = \Delta G^\circ(X) - \Delta G^\circ(\text{FP-2:HTS07940})$, where X stands for the FP-2:HTS08262, FP-3:HTS07940 and hCatK:HTS07940 complexes.

and HTS08262 could act on two *P. falciparum* enzymes important for the parasite growth and involved in orthogonal metabolic pathways, i.e., FP-2 degrades hemoglobin, whereas PfENR participates in fatty acid biosynthesis [51,61]. Interestingly, PfENR is an attractive target for the design of specific antimalarials, as it is involved in a biosynthetic pathway absent in humans [48]. Therefore, a putative interaction of the identified hits with PfENR may not be, in principle, detrimental for their higher specificity for the parasitic cells. Future assays will be required to validate the *in vivo* target(s) of the identified hits. In this sense, the microscopic inspection of developmental and morphological changes in the parasite cells is usually conducted to assess the capacity of the compounds to impair the hemoglobin degradation [62]. On the other hand, the incorporation of radiolabeled acetate into the fatty acids chains has been used to detect the inhibition of enzymes involved in the fatty acid biosynthesis [48].

We also assessed the putative interaction of Genz10850 with FP-2 through 100 ns MD simulations (data not shown). In this case, the FP-2:Genz10850 complex, built using the predicted structure of FP2 bound to HTS08262 as a template, was unstable and the compound readily dissociated the active site during the MD simulation. Thus, the shorter substituent of Genz10850 (indol ring) attached to the carbonyl group does not seem to be suitable to interact with FP-2. This result reinforces the importance of the core structure of the variable substituent present in HTS07940 and HTS08262 for FP-2 inhibition.

Interestingly, we also found a previously reported FP-2 inhibitor ($IC_{50} = 44.94 \mu M$), i.e., 2-((9H-fluoren-9-yl)amino)-2-oxoethyl 4-methylpiperazin-1-carbodithiolate, labeled as compound 19 in ref. [63], which shares some structural similarities with the best hits identified here. In fact, compound 19 bears a 9H-fluoren-9-yl moiety that might bind FP-2 in a similar fashion to that of HTS07940 and HTS08262, i.e., by interacting with W206 at the S2' subsite. Compound 19 also possesses a piperazinyl moiety, a common structural substituent present in various FP-2 inhibitors [5], but in this case it is not directly bonded to the 9H-fluoren-9-yl ring, as it does in the compounds reported here. Excluding the aforementioned similarities, the structures of compounds HTS07940 and HTS08262 differ appreciably from that of 19 in other aspects; thus, they do not belong to the same scaffold. Such structural differences seem to be quite relevant, since, in contrast to HTS07940 and HTS08262, compound 19 did not show measurable inhibition of *P. falciparum* growth *in vitro* [63]. From the information discussed in this paragraph, we concluded that the 9H-fluoren-9-yl ring is an important structural feature for the interaction with FP-2, something not highlighted in literature so far.

It is worth noting that the IC_{50} values of HTS07940 and HTS08262 against *P. falciparum* cultures are either of the same order or even lower than their respective IC_{50} values measured against FP-2. Since the compounds have to diffuse through various membranes to finally reach the target enzyme inside the parasite food vacuole, the high efficiency of these compounds displayed in the *P. falciparum* growth inhibition assays may arise from various factors. For example, the interaction with various targets, such as other cysteine proteases, although in the case of FP-3, the predicted affinity is lower than that for FP-2, and PfENR. In addition, both compounds possess amine groups that become protonated in the acidic environment of the food vacuole. Interestingly, it has been reported before that positively-charged compounds can concentrate in acidic cell bodies such as lysosomes and reservosomes [64]. This phenomenon can contribute to enhance the potency of HTS07940 and HTS08262 against FP-2 in the live parasites.

Our experimental results have shown that HTS07940 and HTS08262 are very weak inhibitors of hCatK, a property that is thought to minimize unwanted cross-inhibition effects in the host. Of note, the evaluation of selectivity of SBVS-identified FP inhibitors against human off-targets is not a common practice, limiting the possibility to contextualize our results. Using a compound library potentially-enriched in cysteine protease inhibitors, Shah et al. identified 21 compounds carrying FP-2 inhibitory activity ($IC_{50} \leq 50 \mu M$); with six of

them also inhibiting *P. falciparum* growth ($IC_{50} \leq 50 \mu M$) [63]. The seven most potent inhibitors exhibited moderate selectivity for FP-2 over hCatK, hCatL and hCatB; with IC_{50} values for off-targets being 2.75 to 36-fold higher than those for the plasmodial enzyme. The most selective inhibitor (compound 4) proved to be > 36-fold more selective for FP-2 than for human cathepsins, but it was inactive against *P. falciparum*. This selectivity is very similar to that of HTS07940, the most selective inhibitor identified here. Using three of the seven compounds identified in the previous work, the same group searched for similar compounds in commercial databases through computational methods [65]. Although 28 FP-2/FP-3 low micromolar compounds were identified, their selectivity over mammalian proteases was not improved in comparison with parental compounds. Thus, to the best of our knowledge, HTS07940 stands to date among the most selective FP-2 inhibitors identified by SBVS.

The *in silico* structural and energetic analyses carried out here confirmed that the affinity of both compounds for hCatK is lower than that for FP-2. According to the information available in literature [54,57,66], the presence of residues Y67 and L209 at the S2 bottom of hCatK precludes the accommodation of bulky moieties into this subsite. In agreement with the previous statement, we predicted that the accommodation of the *p*-methoxyphenyl ring of HTS07940 at the S2 bottom would require the opening of a van der Waals wall formed by the side chains of Y67 and L209, which is associated to an energetic penalty. Furthermore, the similarities between hCatK and FP-3 at the S2 bottom, may also determine the smaller affinity of HTS07940 for the latter enzyme when compared to FP-2, as predicted in our work. However, other structural features of FP-3 active site might have a negative impact on its affinity for HTS07940. In fact, we observed that the S2 bottom of latter enzyme is naturally more open than that of hCatK, but this did not lead to a lower predicted binding free energy of FP-3 for HTS07940. Overall, the abovementioned findings suggest that other human cathepsins bearing bulky residues at positions 67 and/or 205 in the papain numbering scheme, e.g., hCatC, hCatF, hCatX and hCatS, may also have low affinities for HTS07940 and HTS08262.

5. Conclusions

We here reported the identification of two FP-2 inhibitors, i.e., HTS07940 and HTS08262, by employing an SBVS approach. The compounds showed micromolar inhibition of *in vitro* *P. falciparum* cultures and suitable selectivity indices when tested against HeLa cells and hCatK. Remarkably, both compounds belong to a novel scaffold of FP-2 inhibitors, thus increasing the chemical diversity of known compounds with the potential to become future antimalarials. Our work constitutes a successful example of an integrated workflow of *in silico* analyses and experimental procedures useful for drug discovery in a broader context.

Acknowledgments

We would like to thank Dr. Dieter Bromme (University of British Columbia, Vancouver, Canada) for providing a *Pichia pastoris* strain expressing human cathepsin K and Dr. Chris Oostenbrink (University of Natural Resources and Life Sciences, Vienna, Austria) for his help in Umbrella Sampling free energy calculations.

Funding

This work was partially supported by the International Foundation for Science (Sweden) [grants F/5198-1, F/4081-1 and F/4081-2]; fellowships from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, Argentina); São Paulo Research Foundation (FAPESP) [grant number 2016/24587-9] (Brazil) and Coordination for the Improvement of Higher Education Personnel (CAPES) [grant number 031/2013 PRÓ-DEFESA 3] (Brazil).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bbagen.2018.09.015>.

References

- [1] B.D. Bekono, F. Ntie-Kang, L.C. Owono Owono, E. Megnassan, Targeting cysteine proteases from *Plasmodium falciparum*: a general overview, rational drug design and computational approaches for drug discovery, *Curr. Drug Targets* 19 (2018) 501–526.
- [2] World Malaria Report, <http://www.who.int/malaria/publications/world-malaria-report-2017/report/en/>, (2017).
- [3] C. Teixeira, J.R. Gomes, P. Gomes, Falcipains, *Plasmodium falciparum* cysteine proteases as key drug targets against malaria, *Curr. Med. Chem.* 18 (2011) 1555–1572.
- [4] K.K. Roy, Targeting the active sites of malarial proteases for antimalarial drug discovery: approaches, progress and challenges, *Int. J. Antimicrob. Agents* 50 (2017) 287–302.
- [5] R. Ettari, F. Bova, M. Zappala, S. Grasso, N. Micale, Falcipain-2 inhibitors, *Med. Res. Rev.* 30 (2010) 136–167.
- [6] J.E. Hernandez Gonzalez, L. Hernandez Alvarez, P.G. Pascutti, P.A. Valiente, Predicting binding modes of reversible peptide-based inhibitors of falcipain-2 consistent with structure-activity relationships, *Proteins* 85 (2017) 1666–1683.
- [7] E. Stjernschantz, C. Oostenbrink, Improved ligand-protein binding affinity predictions using multiple binding modes, *Biophys. J.* 98 (2010) 2682–2691.
- [8] D.C. Thompson, C. Humblet, D. Joseph-McCarthy, Investigation of MM-PBSA re-scoring of docking poses, *J. Chem. Inf. Model.* 48 (2008) 1081–1091.
- [9] A. Lindstrom, L. Edvinsson, A. Johansson, C.D. Andersson, I.E. Andersson, F. Raubacher, A. Linusson, Postprocessing of docked protein-ligand complexes using implicit solvation models, *J. Chem. Inf. Model.* 51 (2011) 267–282.
- [10] T. Hou, J. Wang, Y. Li, W. Wang, Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized born surface area methods. II. The accuracy of ranking poses generated from docking, *J. Comput. Chem.* 32 (2011) 866–877.
- [11] G. Rastelli, A. Del Rio, G. Degliesposti, M. Sgobba, Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA, *J. Comput. Chem.* 31 (2010) 797–810.
- [12] A.J. Clark, P. Tiwary, K. Borrelli, S. Feng, E.B. Miller, R. Abel, R.A. Friesner, B.J. Berne, Prediction of protein-ligand binding poses via a combination of induced fit docking and metadynamics simulations, *J. Chem. Theory Comput.* 12 (2016) 2990–2998.
- [13] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* 25 (1997) 3389–3402.
- [14] M.S. Madhusudhan, B.M. Webb, M.A. Marti-Renom, N. Eswar, A. Sali, Alignment of multiple protein structures based on sequence and structure features, *Protein Eng. Des. Sel.* 22 (2009) 569–574.
- [15] R.F. Doolittle, Similar amino acid sequences: chance or common ancestry? *Science* 214 (1981) 149–159.
- [16] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.* 31 (2010) 455–461.
- [17] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, *J. Comput. Chem.* 30 (2009) 2785–2791.
- [18] T.J. Dolinsky, P. Czodrowski, H. Li, J.E. Nielsen, J.H. Jensen, G. Klebe, N.A. Baker, PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations, *Nucleic Acids Res.* 35 (2007) W522–W525.
- [19] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open Babel: an open chemical toolbox, *J. Cheminform.* 3 (2011) 33.
- [20] E.S. Sarduy, A.C. Munoz, S.A. Trejo, A.C.P.M. de Los, High-level expression of Falcipain-2 in *Escherichia coli* by codon optimization and auto-induction, *Protein Expr. Purif.* 83 (2012) 59–69.
- [21] C.J. Linnevers, M.E. McGrath, R. Armstrong, F.R. Mistry, M.G. Barnes, J.L. Klaus, J.T. Palmer, B.A. Katz, D. Bromme, Expression of human cathepsin K in *Pichia pastoris* and preliminary crystallographic studies of an inhibitor complex, *Protein Sci.* 6 (1997) 919–921.
- [22] J.B. Bertoldo, L.D. Chiaradia-Delatorre, A. Mascarello, P.C. Leal, M.N. Cordeiro, R.J. Nunes, E.S. Sarduy, P.J. Rosenthal, H. Terenzi, Synthetic compounds from an in house library as inhibitors of falcipain-2 from *Plasmodium falciparum*, *J. Enzyme Inhib. Med. Chem.* 30 (2015) 299–307.
- [23] One-way ANOVA followed by Dunnett's multiple comparisons test was performed using GraphPad Prism version 5.03 for Windows, in, GraphPad Software, La Jolla California USA, www.graphpad.com.
- [24] W. Trager, J.B. Jensen, Human malaria parasites in continuous culture, *Science* 193 (1976) 673–675.
- [25] A. Pabon, O. Ramirez, A. Rios, E. Lopez, B. de Las Salas, F. Cardona, S. Blair, Antiplasmodial and cytotoxic activity of raw plant extracts as reported by knowledgeable indigenous people of the Amazon Region (Vaupes Medio in Colombia), *Planta Med.* 82 (2016) 717–722.
- [26] A.E. Bianco, J.M. Favaloro, T.R. Burkot, J.G. Culvenor, P.E. Crewther, G.V. Brown, R.F. Anders, R.L. Coppel, D.J. Kemp, A repetitive antigen of *Plasmodium falciparum* that is homologous to heat shock protein 70 of *Drosophila melanogaster*, *Proc. Natl. Acad. Sci. U. S. A.* 83 (1986) 8713–8717.
- [27] M.J. Pascual, F. Merwaiss, E. Leal, M.E. Quintana, A.V. Capozzo, C.N. Cavasotto, M. Bollini, D.E. Alvarez, Structure-based drug design for envelope protein E2 uncovers a new class of bovine viral diarrhea inhibitors that block virus entry, *Antivir. Res.* 149 (2018) 179–190.
- [28] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M.J. Bearpark, J. Heyd, E.N. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A.P. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N.J. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09, Gaussian, Inc., Wallingford, CT, USA, 2009.
- [29] D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, I.T.E. Cheatham, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, P.A. Kollman, AMBER14, University of California, San Francisco, 2014.
- [30] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, *J. Comput. Chem.* 25 (2004) 1157–1174.
- [31] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, *J. Chem. Theory Comput.* 11 (2015) 3696–3713.
- [32] B.R. Miller 3rd, T.D. McGee Jr., J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA.py: an efficient program for end-state free energy calculations, *J. Chem. Theory Comput.* 8 (2012) 3314–3321.
- [33] J. Wereszczynski, J.A. McCammon, Accelerated molecular dynamics in computational drug design, *Methods Mol. Biol.* 819 (2012) 515–524.
- [34] D.R. Roe, T.E. Cheatham 3rd, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, *J. Chem. Theory Comput.* 9 (2013) 3084–3095.
- [35] J. Shao, S.W. Tanner, N. Thompson, T.E. Cheatham, Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms, *J. Chem. Theory Comput.* 3 (2007) 2312–2334.
- [36] A. Amadi, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, *Proteins* 17 (1993) 412–425.
- [37] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX* 1 (2015) 19–25.
- [38] S. Doudou, N.A. Burton, R.H. Henchman, Standard free energy of binding from a one-dimensional potential of mean Force, *J. Chem. Theory Comput.* 5 (2009) 909–918.
- [39] C. Herrera Acevedo, L. Scotti, M. Feitosa Alves, M.F. Formiga Melo, M.T. Scotti Diniz, Computer-aided drug design using sesquiterpene lactones as sources of new structures with potential activity against infectious neglected diseases, *Molecules* 22 (2017).
- [40] A.C. Mafud, L.G. Ferreira, Y.P. Mascarenhas, A.D. Andricopulo, J. de Moraes, Discovery of novel antischistosomal agents by molecular modeling approaches, *Trends Parasitol.* 32 (2016) 874–886.
- [41] V. Turk, V. Stoka, O. Vasiljeva, M. Renko, T. Sun, B. Turk, D. Turk, Cysteine cathepsins: from structure, function and regulation to new frontiers, *Biochim. Biophys. Acta* 1824 (2012) 68–88.
- [42] S.S. Cotrin, I.E. Gouvea, P.M. Melo, P. Bagnaresi, D.M. Assis, M.S. Araujo, M.A. Juliano, M.L. Gazarini, P.J. Rosenthal, L. Juliano, A.K. Carmona, Substrate specificity studies of the cysteine peptidases falcipain-2 and falcipain-3 from *Plasmodium falciparum* and demonstration of their kininogenase activity, *Mol. Biochem. Parasitol.* 187 (2013) 111–116.
- [43] M.F. Alves, L. Puzer, S.S. Cotrin, M.A. Juliano, L. Juliano, D. Bromme, A.K. Carmona, S3 to S3' subsite specificity of recombinant human cathepsin K and development of selective internally quenched fluorescent substrates, *Biochem. J.* 373 (2003) 981–986.
- [44] S.S. Cotrin, L. Puzer, W.A. de Souza Judice, L. Juliano, A.K. Carmona, M.A. Juliano, Positional-scanning combinatorial libraries of fluorescence resonance energy transfer peptides to define substrate specificity of carboxypeptidases: assays with human cathepsin B, *Anal. Biochem.* 335 (2004) 244–252.
- [45] L. Puzer, S.S. Cotrin, M.F. Alves, T. Egborge, M.S. Araujo, M.A. Juliano, L. Juliano, D. Bromme, A.K. Carmona, Comparative substrate specificity analysis of recombinant human cathepsin V and cathepsin L, *Arch. Biochem. Biophys.* 430 (2004) 274–283.
- [46] D. Bromme, P.R. Bonneau, P. Lachance, A.C. Storer, Engineering the S2 subsite specificity of human cathepsin S to a cathepsin L- and cathepsin B-like specificity, *J. Biol. Chem.* 269 (1994) 30238–30242.
- [47] J.B. Baell, G.A. Holloway, New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays, *J. Med. Chem.* 53 (2010) 2719–2740.
- [48] M.R. Kuo, H.R. Morbidoni, D. Alland, S.F. Sneddon, B.B. Gourlie, M.M. Staveski, M. Leonard, J.S. Gregory, A.D. Janjigian, C. Yee, J.J.M. Musser, B. Kreiswirth, H. Iwamoto, R. Perozzo, W.R. Jacobs Jr., J.C. Sacchetti, D.A. Fidock, Targeting tuberculosis and malaria through inhibition of Enoyl reductase: compound activity

- and structural data, *J. Biol. Chem.* 278 (2003) 20851–20859.
- [49] J.C. Powers, J.L. Asgian, O.D. Ekici, K.E. James, Irreversible inhibitors of serine, cysteine, and threonine proteases, *Chem. Rev.* 102 (2002) 4639–4750.
- [50] A.C. Storer, R. Ménard, Recent insights into cysteine protease specificity: lessons for drug design, *Percept. Drug Discov. Des.* 6 (1996) 33–46.
- [51] B.R. Shenai, P.S. Sijwali, A. Singh, P.J. Rosenthal, Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of *Plasmodium falciparum*, *J. Biol. Chem.* 275 (2000) 29000–29010.
- [52] F. Shah, P. Mukherjee, J. Gut, J. Legac, P.J. Rosenthal, B.L. Tekwani, M.A. Avery, Identification of novel malarial cysteine protease inhibitors using structure-based virtual screening of a focused cysteine protease inhibitor library, *J. Chem. Inf. Model.* 51 (2011) 852–864.
- [53] J.P. Changeux, S. Edelstein, Conformational selection or induced fit? 50 years of debate resolved, *F1000 Biol. Rep.* 3 (2011) 19.
- [54] F. Lecaillon, S. Chowdhury, E. Purisima, D. Bromme, G. Lalmanach, The S2 subsites of cathepsins K and L and their contribution to collagen degradation, *Protein Sci.* 16 (2007) 662–670.
- [55] L. Hernandez Alvarez, D. Naranjo Feliciano, J.E. Hernandez Gonzalez, R.O. Soares, D.E. Barreto Gomes, P.G. Pascutti, Insights into the interactions of *Fasciola hepatica* cathepsin L3 with a substrate and potential novel inhibitors through in silico approaches, *PLoS Negl. Trop. Dis.* 9 (2015) e0003759.
- [56] J.D. Durrant, H. Keranen, B.A. Wilson, J.A. McCammon, Computational identification of uncharacterized cruzain binding sites, *PLoS Negl. Trop. Dis.* 4 (2010) e676.
- [57] C.P. Brady, R.I. Brinkworth, J.P. Dalton, A.J. Dowd, C.K. Verity, P.J. Brindley, Molecular modeling and substrate specificity of discrete cruzipain-like and cathepsin L-like cysteine proteinases of the human blood fluke *Schistosoma mansoni*, *Arch. Biochem. Biophys.* 380 (2000) 46–55.
- [58] L. Xu, H. Sun, Y. Li, J. Wang, T. Hou, Assessing the performance of MM/PBSA and MM/GBSA methods. 3. The impact of force fields and ligand charge models, *J. Phys. Chem. B* 117 (2013) 8408–8421.
- [59] S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, *Expert Opin. Drug Discovery* 10 (2015) 449–461.
- [60] Y. Cheng, W.H. Prusoff, Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I₅₀) of an enzymatic reaction, *Biochem. Pharmacol.* 22 (1973) 3099–3108.
- [61] R.F. Waller, P.J. Keeling, R.G. Donald, B. Striepen, E. Handman, N. Lang-Unnasch, A.F. Cowman, G.S. Besra, D.S. Roos, G.I. McFadden, Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*, *Proc. Natl. Acad. Sci. U. S. A.* 95 (1998) 12352–12357.
- [62] M. Duran-Lengua, E. Salas-Sarduy, L.M. Cano-Duran, E. Carlos, D.M. Mendez-Cuadro, J. Lorenzo-Rivera, J. Piermattey-Ditta, J. Montalvo-Acosta, J.M. Bautista-Santa Cruz, R.G. Ibarra, Quinoid compounds cause inhibition of falcipain 2, and arrest *Plasmodium falciparum* growth in vitro, *Lat. Am. J. Pharm.* 33 (2014) 666–674.
- [63] F. Shah, P. Mukherjee, J. Gut, J. Legac, P.J. Rosenthal, B.L. Tekwani, M.A. Avery, Identification of novel malarial cysteine protease inhibitors using structure-based virtual screening of a focused cysteine protease inhibitor library, *J. Chem. Inf. Model.* 51 (2011) 852–864.
- [64] D.A. Nicoll-Griffith, Use of cysteine-reactive small molecules in drug discovery for trypanosomal disease, *Expert Opin. Drug Discovery* 7 (2012) 353–366.
- [65] F. Shah, J. Gut, J. Legac, D. Shivakumar, W. Sherman, P.J. Rosenthal, M.A. Avery, Computer-aided drug design of falcipain inhibitors: virtual screening, structure-activity relationships, hydration site thermodynamics, and reactivity analysis, *J. Chem. Inf. Model.* 52 (2012) 696–710.
- [66] F. Lecaillon, Y. Choe, W. Brandt, Z. Li, C.S. Craik, D. Bromme, Selective inhibition of the collagenolytic activity of human cathepsin K by altering its S2 subsite specificity, *Biochemistry* 41 (2002) 8447–8454.