

Novel anti-*Plasmodial* hits identified by virtual screening of the ZINC database

Grace Mugumbate · Ana S. Newton ·
Philip J. Rosenthal · Jiri Gut · Rui Moreira ·
Kelly Chibale · Rita C. Guedes

Received: 5 June 2013 / Accepted: 12 October 2013 / Published online: 25 October 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Increased resistance of *Plasmodium falciparum* to most available drugs challenges the control of malaria. Studies with protease inhibitors have suggested important roles for the falcipain family of cysteine proteases. These enzymes act in concert with other proteases to hydrolyze host erythrocyte hemoglobin in the parasite food vacuole. In order to find potential new antimalarial drugs, we screened in silico the ZINC database using two different protocols involving structure- and ligand-based methodologies. Our search identified 19 novel low micromolar inhibitors of cultured chloroquine resistant *P. falciparum*. The most active compound presented an IC_{50} value of

0.5 μ M against cultured parasites and it also inhibited the cysteine protease falcipain-2 ($IC_{50} = 25.5 \mu$ M). These results identify novel classes of antimalarials that are structurally different from those currently in use and which can be further derivatized to deliver leads suitable for optimisation.

Keywords Malaria · Antimalarials · Falcipain inhibitors · Virtual screening

Introduction

Malaria remains one of the most severe parasitic diseases in humans, particularly in tropical regions of the world. The most recent World Health Organization (WHO) report states that 3.3 billion people were at risk of malaria in 2010, leading to nearly one million deaths per year, with over 90 % of these deaths occurring in pregnant women and children below the age of 5 years [1]. Malaria is caused by five different species of protozoan parasites, with *Plasmodium falciparum* the most dangerous and widespread species. The spread of multi-drug resistant *P. falciparum* strains combined with the lack of effective vaccines, and the recent absence of novel chemical classes of antimalarials with new mechanisms of action, create the urgent need to identify novel compounds with antimalarial potency [2–9].

Amongst the main potential *P. falciparum* drug targets are those involved in hemoglobin degradation, a process essential for the survival of the parasite [10]. Falcipain-2 (FP-2) and falcipain-3 (FP-3) are two well characterized cysteine proteases from *P. falciparum* that share approximately 68 % identity and act, along with other proteases, to hydrolyze host erythrocyte hemoglobin in the parasite digestive vacuole [11–13]. Inhibition of FP-2 and its

Grace Mugumbate and Ana S. Newton have equally contributed to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10822-013-9685-z) contains supplementary material, which is available to authorized users.

G. Mugumbate · K. Chibale
Department of Chemistry, University of Cape Town,
Rondebosch 7701, South Africa
e-mail: kelly.chibale@uct.ac.za

A. S. Newton · R. Moreira · R. C. Guedes (✉)
Research Institute for Medicines and Pharmaceutical Sciences
(iMed.UL), Faculty of Pharmacy, University of Lisbon,
Av. Prof. Gama Pinto, 1649-019 Lisbon, Portugal
e-mail: rguedes@ff.ul.pt

P. J. Rosenthal · J. Gut
Department of Medicine, San Francisco General Hospital,
University of California, San Francisco, Box 0811,
San Francisco, CA 94143, USA

K. Chibale
Institute of Infectious Disease and Molecular Medicine,
University of Cape Town, Rondebosch 7701, South Africa

analogues prevents parasite maturation, suggesting that these proteins are promising targets for the design of new antimalarial drugs. During the last decade, many efforts to discover and optimize falcipain inhibitors, from traditional chemical synthesis to in silico methods, have been implemented. In general, falcipain inhibitors can be classified in two main categories: peptide/peptidomimetic inhibitors and nonpeptidic small molecules inhibitors. [14] Many falcipain inhibitors have been identified, exemplified by peptide-based inhibitors such as peptidyl fluoromethyl ketones, [15, 16] peptidyl aldehydes, [17] vinyl sulfones [18], α -ketoamides [19], epoxysuccinyl derivatives [20] and peptidyl aziridines [21]. To date, only a few peptidomimetic inhibitors of FP-2 with activity in the micromolar range have been identified. These include compounds based on the 1,4 benzodiazepine and pyridine scaffolds [18, 22–24]. However, due to pharmacological constraints, in vivo efficacy of peptidic compounds has been limited, since other parameters apart from potency, e.g. solubility, stability and selectivity are important when designing an antimalarial. The development of nonpeptide molecules is expected to overcome these limitations. Some nonpeptide falcipain inhibitors [23, 25, 26], including chalcones, isoquinolines and others have shown antiparasitic activity in the low micromolar range. Most notably, pyrimidine carbonitrile derivatives [27] have recently been reported to possess outstanding activities, in the nanomolar and sub-nanomolar range, against FP-2 and FP-3, and against cultured *P. falciparum*.

Structure-based and ligand-based virtual screening (SBVS and LBVS) of commercially available small molecule databases have in recent years become useful tools in identifying promising hits in the initial stages of drug discovery [28, 29]. In brief, SBVS involves docking methods to screen a ligand database into a protein target, with the generation of different conformations (searching algorithms) of receptor-ligand complexes, followed by the application of a scoring function to estimate binding affinity. Desai et al. [30] have previously used this technique to screen the Chembridge and Available Chemical Directory (ACD) databases against homology models of FP-2 and FP-3. Another virtual screening study performed by Li et al. [31] used a crystal structure of FP-2 (PDB code: 2GHU) and the SPECS database. Based on the overall concept that similar compounds afford similar activity, LBVS is applied when one or more biologically active compounds are used to build a structural query or pharmacophore model. The query or model is a summary of the important properties of the ligands that is used to identify molecules from a database, which share properties with the known compounds.

As part of our on-going program to discover potential antiparasitodal compounds, we set out two different

protocols (Fig. 1) to virtually screen the ZINC database using the crystal structure of FP-2 (PDB code: 3BPF) to represent the protease family [32] and some of its inhibitors reported in the literature to validate this procedure. First, SBVS was performed, at Medicinal Chemistry Division, Research Institute for Medicine and Pharmaceutical Sciences, University of Lisbon, Portugal, using the GOLD docking program [33] to screen the freely available database [34]. The database contains over 13 million commercially available compounds and is maintained by the Shoichet group at the University of California, San Francisco. Second, an independent protocol involving a combination of LBVS using the software, Pentacle 1.0.6 (available on <http://www.moldiscovery.com>), [35] and SBVS using Autodock4 (available in the lab at the time of study) was used to screen the same database, at the Medicinal Chemistry Research Group, University of Cape Town, South Africa. From our methodology 412 compounds were identified as putative hits based on binding in the FP-2 active site. After a careful visual inspection, 60 compounds were selected for purchase and were tested against FP-2 and against cultured *P. falciparum*.

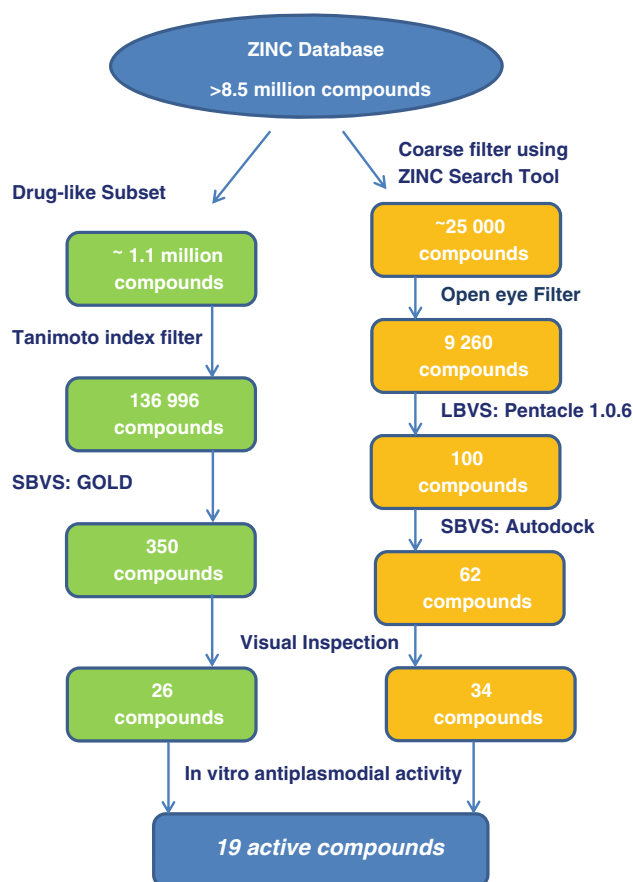


Fig. 1 Virtual screening workflow to find new lead compounds. Protocol 1 (left) and protocol 2 (right)

Methods

Virtual screening

The approach for virtual screening of the ZINC database is shown in Fig. 1. The two protocols implemented are described below.

Protein structure preparation

Only four crystallographic structures are available for FP-2: 3BPF (FP-2 covalently linked to E64 at 2.9 Å resolution), 1YVB (FP-2 covalently linked to cystatin at 2.7 Å resolution), 2GHU (not complexed at 3.1 Å resolution) and 2OUL (FP-2 covalently linked to chagasin at 2.2 Å resolution). Although both 1 YVB and 2OUL had better resolution than 3BPF, they are complexed to large inhibitors (over 100 amino acids). In contrast, 3BPF contains E64, a low molecular-weight, non-selective inhibitor that closely matches the libraries we intended to screen. To validate the 3D FP-2 structure, covalent- and non-covalent docking experiments were performed using E64 with the epoxide ring opened. GOLD and AutoDock 4 were used, from which resulted very similar poses and several numbers of genetic algorithm runs were used during docking calculations to optimize the screening time. The best pose obtained with a calculation with only 50 docking runs is shown in Fig. 2. In addition, we also allowed for protein flexibility (Cys42 and His174) with the set of test compounds. However, this methodology did not improve the docking scores, and thus we decided to perform rigid receptor docking throughout our virtual screening study. The receptor structure was obtained from the available FP-2 crystal structure in complex with the epoxysuccinate E64 (PDB code: 3BPF) [32] was retrieved from PDB databank and prepared using the protein preparation tools implemented in MOE 2010.10 software (Chemical Computing Group) [36]. The complex has four chains (A, B, C and D). For this study we decided to only work with the A chain. The E64 inhibitor and all crystallographic water molecules were removed from the coordinate set, hydrogen atoms were added to this reduced crystal structure and the protein was protonated to pH 5 to mimic the acidic environment of the *P. falciparum* food vacuole (pH 4–6), where the enzyme is located. The enzyme was then submitted to restrained molecular mechanics refinement using the AMBER99 force field implemented in MOE software [36]. To assess the suitability of this crystal structure-based model, a preliminary validation of the enzyme structure was carried out involving re-docking E64 to the prepared FP-2 structure with GOLD software version 4.0.1, with no restrictions on the flexibility of the linking options. The experimental pose of E64 was reproduced. The RMSD

between predicted and experimental crystallographic poses was 2 Å, which was quite satisfactory considering the 12 rotatable bonds of the E64 molecule and the fact that this pose was obtained with only 50 runs, the number of runs used in the virtual screening protocol (Fig. 2).

In addition, 7 reported inhibitors of FP-2 and 2 compounds with no activity were docked using both GOLD 4.0.1 and AutoDock4 software. This analysis discriminated the active (1–7) and inactive compounds (8, 9) (Table 1). The probe compounds were chemically diverse, from peptidyl vinyl sulfones (1, 3), [37] α -ketoamides (2, 4) [19], and nonpeptide inhibitors (5–7) [30, 38, 39].

As both softwares were able to reproduce the crystallographic geometry of E64 and the inhibitory activities of the set of test compounds (Table 1), we considered them suitable for the subsequent calculations in this study.

Protocol 1: ligand database preparation

The ZINC8 database [34] (<http://zinc.docking.org/>), which contains over 8.5 M compounds, was filtered to only include compounds with drug-like properties, based essentially on the Lipinski “rule of five” [40, 41]. ZINC webpage stated that the database had 13 million compounds. However, at the time we downloaded the compounds only ~8.5 million were available, which can be explained by the fact that ZINC databases are very dynamic and weekly updated. During filtering of the ZINC database, ~1.1 million drug-like compounds with octanol–water partition coefficients ($\log P$) ≤ 5 , molecular weights 150–500, rotatable bonds < 8 , polar surface area < 150 Å² and, hydrogen bond acceptors < 10 were collected. To avoid very similar compounds, database clustering was executed by analyzing similarities between the compounds. This was carried out with the algorithm of Bienfait to incrementally select compounds that differ from all previous considerations by the Tanimoto cutoff. A 90 % diversity set (136,966 compounds) was submitted for structure-based virtual screening (SBVS). This subset was

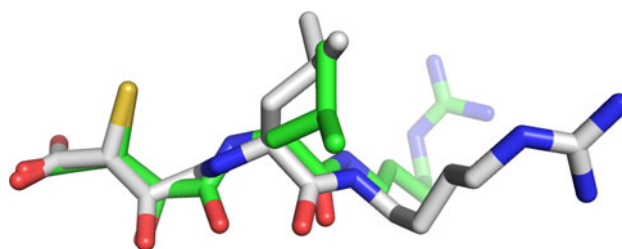
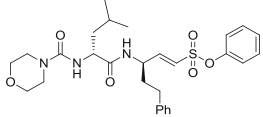
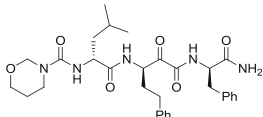
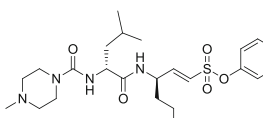
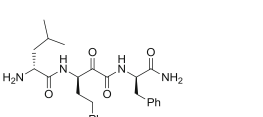
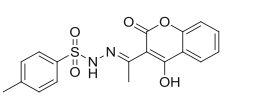
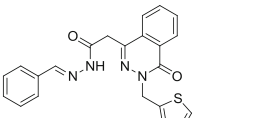
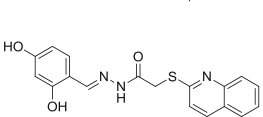
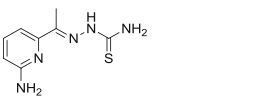



Fig. 2 Comparison of the covalently docked (green) and the experimental crystal structure [32] (grey) of E64. The RMSD for all heavy atoms 2 Å was calculated for the highest scored pose obtained with 50 docking runs. The sulphur in the figure belongs to the covalently bonded cysteine

Table 1 Structures, molecular weight, inhibition activity, predicted scores and binding energies from different docking software for seven known inhibitors and two decoys of the FP-2 enzyme

Compound	Structure	Molecular Weight	FP-2 IC ₅₀ (nM)	GOLD	AutoDock4
				GoldScore	Binding energy
1		543.67	0.7	71.95	-8.17
2		579.69	1.0	71.64	-7.00
3		556.72	350	69.04	-9.09
4		466.57	1500	68.45	-9.44
5		372.40	3900	52.93	-7.32
6		417.48	54300	57.54	-8.11
7		353.39	63400	54.82	-7.36
8		209.27	>20000	41.92	-5.93
9		147.13	n.e.	30.92	-5.63

n.e. not evaluated

directly downloaded from the ZINC webpage in a mol2 format.

SBVS

Structure-based virtual screening was performed by docking the ZINC “drug-like” library (136,996 compounds)

into the validated FP-2 enzyme structure with GOLD docking software (version 4.0.1). [33] Residues of the protein within 15 Å of the sulfur atom of the Cys42 were included in the binding site definition. GOLD software uses a genetic algorithm that adapts the principles of biological competition and population dynamics as a search technique. Calculations using the goldscore fitness function

[42] with 50 runs (population size of 100; 100,000 genetic algorithm operations; 5 islands) were conducted for each compound. This scoring function is the negative of the sum of five energy terms, i.e., protein–ligand hydrogen bond energy, protein–ligand van der Waals energy, ligand internal van der Waals energy, ligand torsional strain energy, and a constraint scoring contribution. Only the top 3 poses of each compound were stored. The 350 best ligands (i.e.; those with the highest score) were visually inspected with PyMol. [43] based on the following criteria: (a) adopted position and orientation of the ligand in the FP-2 binding site (essential); (b) distance between the Cys42 sulfur atom and the ligand (essential); (c) occupancy of the S2 pocket by the hydrophobic group of the ligand, since it has been demonstrated that the P2 position is a major determinant of FP-2 specificity (preferential); (d) formation of hydrogen bonds between the ligand and key residues of the FP-2 active site, like Gln36, Cys42, His42, and Asn204 (preferential); (e) and chemical diversity of compounds (preferential) [44]. Based on these criteria, 150 molecules were selected, but from these only 26 with different chemical scaffolds were available for purchase from respective vendors (Table 3 and Supporting Information Table S1). These were acquired and tested for activity against FP-2 and cultured *P. falciparum*.

Protocol 2: ligand database preparation

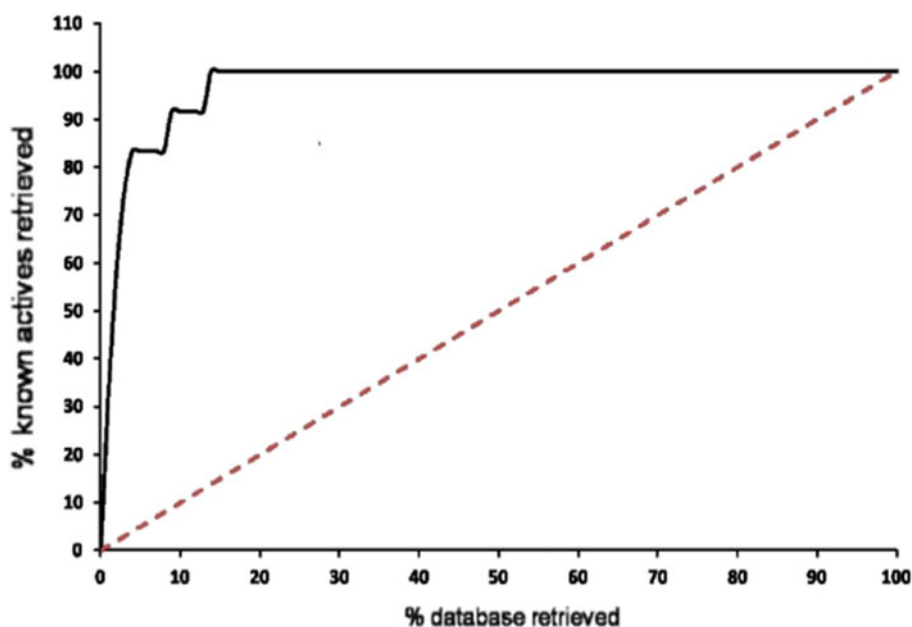
Using the Search tool in the ZINC database browser interface, more than 25,000 compounds were extracted from the drug-like subset, using a filter that included: molecular weights from 150 to 500, $\log P < 6$, rotatable bonds < 12 , hydrogen-donors < 6 , and hydrogen-acceptors < 10 .

The dataset was further filtered using the OpenEye drug-like FILTER v2.1.1 programme available on <http://www.eyesopen.com>. The filtering criteria were based on the properties of 9 nonpeptide inhibitors of FP-2 used in this study, including maximum H-acceptors of 8, maximum H-donors of 5, maximum rotatable bonds of 10 and polar surface area of 150 \AA^2 to collect 9,260 compounds.

Ligand-based virtual screening

Using Pentacle 1.0.6 available on <http://www.moldiscovery.com>, alignment-independent Grid Independent Descriptors (GRINDs) were used to generate a structural query from a dataset of 9 nonpeptide inhibitors of FP-2. The GRINDs were calculated using the hydrophobic (DRY), sp² carbonyl oxygen, (O), neutral flat NH (N1) and shape (TIP) probes, 15 principal components (PC) and an explained variance of 73 %. To test its performance, the structural query was used to retrieve compounds from a diverse set of decoys. The dataset consisted of 506 molecules extracted from the National Cancer Institute open database, presumed inactive to FP-2, collected using the Enhanced NCI database browser to obtain molecules with molecular weight range 200–500 (Supplementary information, Table S2). The dataset was contaminated with 13 assorted nonpeptide inhibitors of FP-2 listed by Ettari et al. [45]. About 84 % of actives were retrieved within ~ 4 % of database retrieval, as shown in Fig. 3. In addition, a BED-ROC value of 0.81 and an Enrichment factor (E^{10}) of 0.92 indicate high quality virtual screening and early recognition of actives by the structural query. Consequently, GRINDs of all compounds in the filtered database were generated using the same parameters, and the query was used to search

Fig. 3 The Relative Operating Characteristic (ROC) curve showing the percentage of known actives retrieved (black line) during retrieval of compounds from an in-house decoy database contaminated with compounds known to have FP-2 activity. The broken red line indicates the random ROC curve



the ZINC dataset, containing 9 260 molecules, for compounds with similar chemical features. A total of 346 compounds were found to have Principal Component Analysis t-scores (similarity scores) less than 2.0. From this list the top 100 compounds were extracted and used in SBVS by performing docking calculations to FP-2.

Structure-based virtual screening

During structure-based virtual screening, docking calculations of 9 inhibitors (Table 2) and 100 high ranked compounds from the LBSV were performed using Autodock4. Ligand and protein input files were prepared using AutoDockTools4 (ADT). All water molecules were removed from the coordinate file of the protein. Hydrogen atoms were added and all nonpolar hydrogens merged. Calculations were performed with a grid of $54 \times 50 \times 54$ xyz points, at grid centre of (xyz) $-56.01, -3.23, -17.46$ and with spacing of 0.39 \AA . For a population of 150, 50 genetic algorithm runs, and 2.5×10^6 evaluations and generations were used in addition to default Autodock4 parameters. AutoDock4 calculated for each ligand different alternative binding poses with an associated free energy of binding (ΔG_{bind}) which were then clustered according to their RMSD. Autodock converts this binding energy (kcal/mol) into inhibition constants K_i (μM). The calculated empirical free energy force field is based on molecular mechanics, which includes typical terms for dispersion and/or repulsion, hydrogen bonding, electrostatics, desolvation, and torsional entropy. The force field was calibrated against a large database of complexes with known structure and binding constant, allowing the force field to predict binding free energies. During the docking simulation, a grid-based method is used for energy evaluation. Autodock4 is Open Source software that makes use of the Lamarckian genetic search algorithm [46]. The molecular weights of the non-peptide inhibitors ranged from 147.13 to 470.98. As shown in Table 2, isatin (**9**) (MW = 147.13), with no reported inhibitory activity against FP-2, had the lowest FP-2 binding affinity ($\Delta G_b = -5.30$ kcal/mol, predicted $K_i = 129.72 \mu\text{M}$). The highest binding affinity ($\Delta G_b = -8.76$ kcal/mol, predicted $K_i = 0.38 \mu\text{M}$) was displayed by a thiosemicarbazone (**41**) (MW = 262.25). Except for three inhibitors, the predicted inhibition constants were $<10 \mu\text{M}$. This was the basis for choosing top compounds from our dataset to be experimentally tested.

From the docking calculations, binding modes and affinities of all structures were analyzed. Sixty-two compounds with predicted $K_i < 10 \mu\text{M}$ were identified and from this list 34 (**44–77**) were readily available for purchase (Table 4 and Supporting Information Table S1). These were acquired and tested for activity against FP-2 and cultured *P. falciparum*.

Biological activities

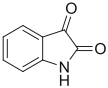
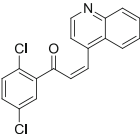
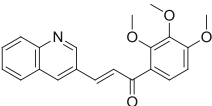
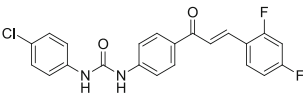
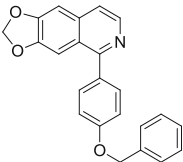
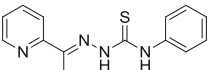
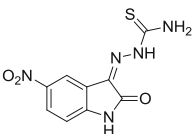
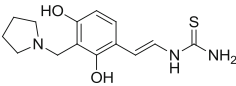
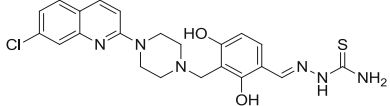
General

The selected compounds were evaluated for inhibition of FP-2 and activity against cultured *P. falciparum*. To determine IC_{50} values, recombinant FP-2 was incubated for 30 min at room temperature in 100 mM sodium acetate, pH = 5.5, and 10 mM dithiothreitol with different concentrations of inhibitors prepared from stocks in DMSO (maximum concentration of DMSO in the assay was 1 %). After 30 min, the substrate benzoxycarbonyl-Leu-Arg-7-amino-4-methyl-coumarin (Z-Leu-Arg-AMC) was added in the same buffer to a final concentration of $25 \mu\text{M}$. Fluorescence was monitored for 15 min at room temperature in a Lab Systems Fluoroskan Ascent spectrofluorometer. IC_{50} values were determined from plots of percent control activity over compound concentration using GraphPad Prism software. Activity against malaria parasites was tested against the chloroquine-resistant (W2) strain of *P. falciparum*, which was cultured at 2 % hematocrit of human erythrocytes in RPMI 1640 medium supplemented with 0.5 % Albumax (Gibco), 2 % heat inactivated human serum, and $100 \mu\text{M}$ hypoxanthine in 96-well culture plates. Parasites were synchronized with 5 % sorbitol. All tested compounds were prepared as 10 mM stock solutions in DMSO and diluted in medium at least 1:1,000, resulting in ≤ 0.1 % final concentrations of DMSO. Synchronized ring-stage parasites at 1 % parasitemia were incubated with compounds at 37°C under 3 % O_2 , 5 % CO_2 , and 92 % N_2 . After 48 h, the medium was removed, and cells were fixed with 2 % formaldehyde in PBS for 24 h. After fixation, $5 \mu\text{L}$ aliquots were incubated for 15 min in the dark in $150 \mu\text{L}$ of 100 mM NH_4Cl , 0.1 % Triton X-100 in PBS, and 10 nM YOYO-1 (Molecular Probes). Parasitemias were determined from dot plots (forward scatter vs. fluorescence) acquired on a FAC Sort flow cytometer using CELLQUEST software (Becton–Dickinson). IC_{50} values were determined from plots of percent control activity over compound concentration using GraphPad Prism software.

Results and discussion

In our effort to identify new chemical entities from the ZINC database with a high probability to bind to FP-2 and inhibit *P. falciparum*, two protocols (Fig. 1) with different filtering methods were followed. We envisaged that combining ligand- and structure-based approaches under the same procedure scheme could overcome some of the inherent weaknesses of these computational modelling tools.

Table 2 Molecular weight, inhibition activity, predicted affinities and inhibition constant for seven known non-peptide inhibitors and two decoys of FP-2 and W2 used in protocol 2

Compound	Structure	Molecular weight	IC ₅₀ (μM) W2/FP-2	AutoDock4 results	
				Affinity (Kcal/mol)	Inhibition constant μM
9		147.13	n.e	-5.30	129.72
36		328.19	0.20(W2)	-8.10	1.16
37		349.38	2.00(W2)	-7.55	2.92
38		412.82	1.80(FP-2)	-8.19	1.00
39		355.39	3.00(FP-2)	-7.84	1.78
40		270.35	n.e	-6.45	18.85
41		265.25	4.40(FP-2)	-8.76	0.38
42		293.38	5.80(FP-2)	-6.14	31.59
43		470.98	2.25(FP-2)	-6.87	9.19

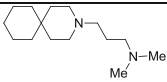
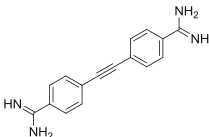
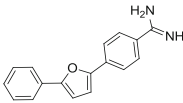
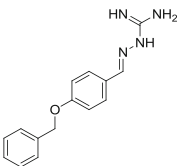
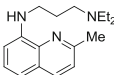
n.e. not evaluated

Out of the sixty compounds tested against cultured *P. falciparum*, nineteen hits were identified with IC₅₀ values of 0.46–9 μM (Table 3, 4). However, only one of these hits was active against FP-2 (IC₅₀ = 25.5 μM) at least at concentrations up to 50 μM.

Inhibition of FP-2

An analysis of the ligand docking poses resulting from the protocol 1 approach revealed that all 26 selected molecules fit well within the FP-2 binding site pocket, occupying S1'

Table 3 Structure, molecular weight (MW), predicted scores, and biological evaluation (Pf-W2 and FP-2) of virtual screening hits from protocol 1

Compound	Structure	MW	score	IC ₅₀ Pf-W2 (μM)	IC ₅₀ FP-2 (μM)
15		238.41	55.26	1.32 ± 0.40	> 50
17		262.31	54.74	0.87 ± 0.28	> 50
24		262.31	53.87	2.73 ± 1.02	> 50
26		268.30	53.72	1.05 ± 0.69	> 50
27		271.42	53.60	3.93 ± 1.25	> 50

and S2 subsites, which was encouraging for a good enzyme ligand interaction. However, none of the 26 tested compounds were active against FP-2 at concentrations up to 50 μM. This lack of activity was not completely unexpected at this range of concentrations. In fact, the docking scores obtained for these compounds, ranging from 59.2 to 53.3, suggest moderate inhibition (from 3.9 to 63.4 μM) when compared to the scores obtained on the validation test (Table 1). The fact that these low molecular weight (between 238 and 271) compounds seem to fit into the FP-2 active site suggest that they could potentially have favourable poses inside the binding pocket, occupying the S1' and S2 key subsites, but that they lack sufficient size to establish the necessary interactions to have potent inhibition. The apparent failure to predict FP-2 inhibitory activity could also be ascribed to the important role played by water molecules inside the binding pocket or to an increased flexibility of FP-2 binding pockets that is not considered in our calculations. On the other hand, five compounds (**15**, **17**, **24**, **26** and **27**) were active (IC₅₀ < 4 μM) against *P. falciparum* (IC₅₀ 0.87 and 3.93 μM, Table 3). Compound **17** was also tested against FP-3 and showed an inhibitory activity of 0.99 μM, an activity very similar to the one obtained against *P. falciparum*, suggesting that this could be a possible target. This may also suggest that the compounds acted against other targets, for example on other cysteine

proteases, not contained in our screen. This was also previously suggested by a recent work of Shah et al. [47].

The highest activity against *P. falciparum* (IC₅₀ = 0.87 μM) was displayed by compound **17**. Of the five active compounds, only the 8-aminoquinoline compound **27** is similar to a well-known antimalarial, primaquine, whilst the rest are based on scaffolds not represented in clinically used antimalarials. These include spiroamine (**15**), azaguanidine (**26**), and amidines (**15** and **24**).

For the hit compounds obtained from protocol 2 (Table 4), only compound **47** showed weak inhibitory activity against FP-2 (IC₅₀ ~ 25 μM) when tested at concentrations up to 50 μM. This was nevertheless encouraging since, as pointed out before, the lack of inhibitory activity against FP-2 was not a surprise at this range of concentrations. Therefore compound **47** could provide a new template for the development of novel inhibitors of FP-2.

Binding mode of compound 47 to FP-2

The weak inhibitory activity exhibited by compound **47** may be rationalised on the basis of its binding to FP-2. This compound, which showed a similarity score of 1.42 on Pentacle, is based on a thiozolidine-2,4-dione (TDZ) scaffold linked on either side to 2,4,6-trimethoxybenzylidene and the acylated phenylpiperazine moieties.

Table 4 Structure, molecular weight, Pentacle similarity scores (Pss), predicted binding energy and inhibition constants, and biological evaluation (Pf-W2 and FP-2) of virtual screening hits from protocol 2

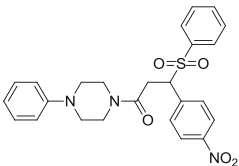
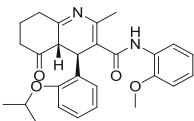
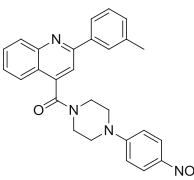
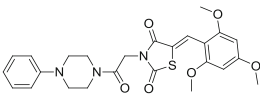
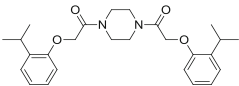
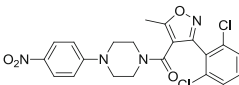
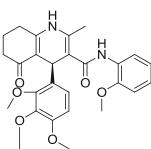
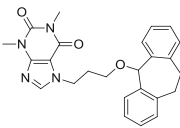
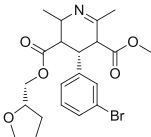
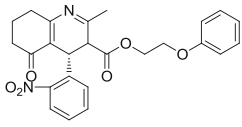
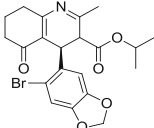
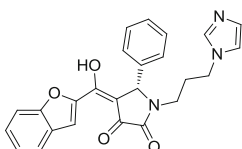
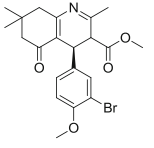
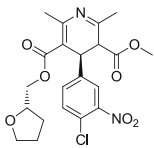
Compd	Structure	MW	Pss	ΔG_b (Kcal/mol)	K_i (μ M)	IC ₅₀ Pf-W2 (μ M)	IC ₅₀ FP-2 (μ M)
44		480.55	1.66	-8.39	0.71	7.87 ± 1.46	> 50
45		446.55	1.35	-8.16	1.05	9.04 ± 0.55	> 50
46		452.51	1.48	-8.12	1.11	8.50 ± 0.06	> 50
47		497.57	1.42	-7.94	1.50	0.46 ± 0.02	25.54 ± 0.18
48		438.57	1.68	-7.83	1.82	4.25 ± 1.08	> 50
49		461.31	1.69	-7.80	1.92	5.56 ± 0.53	> 50
50		478.55	1.65	-7.61	2.64	8.52 ± 0.12	> 50
51		430.51	1.49	-7.40	3.76	6.12 ± 1.08	> 50
52		450.33	1.04	-7.21	5.17	7.96 ± 0.05	> 50

Table 4 continued

53		448.48	1.47	-7.18	5.44	2.59 ± 0.93	> 50
54		448.31	1.72	-7.13	5.92	6.29 ± 1.29	> 50
55		426.45	1.52	-7.09	6.32	5.39 ± 3.27	> 50
56		434.33	1.67	-6.89	8.89	7.83 ± 0.43	> 50
57		450.88	0.96	-6.46	18.70	2.40 ± 0.39	> 50

Docking calculations of **47** indicate similar affinity for FP-2 (binding energy = -7.94 kcal/mol) and orientation. The TDZ linker is centrally positioned in the binding site of FP-2 (Fig. 4a) and directs the 2,4,6-trimethoxybenzylidene scaffold into the S2 subsite, which is important for specificity of cysteine proteases [48]. Consequently, the oxo-2,4-phenyl-1-piperazinyl scaffold of **47** is placed in a cleft just below the S3 subsite and interacts with Gly83. In this position it forms a hydrogen bond with Ile85 (Fig. 4b). Occupation of the S2 subsite and interaction with Ile85 was also found with inhibitor **37** (Fig. 4c) and is a position also occupied by the leucine side chain in E64 [49]. In contrast to **47**, inhibitor **37** (shown in table 2) also forms a hydrogen bond with His174 as shown in Fig. 4d. Cluster analysis of the 50 docked conformations of **47** at a cutoff 2.0 Å indicates that this was the most favoured orientation of the compound in this binding pocket. The lowest energy conformation of **47** shown in Fig. 4 belongs to this popular cluster containing 9 hits. The large distance between Cys42 and compound **47** implies that covalent binding to FP-2 is unlikely.

The different binding modes and interactions with the binding site residues between E64 and compound **47** implies that modification of this compound could lead to potential reversible inhibitors of FP-2. However, there is need to perform experiments to determine the specific binding of compound **47** to FP-2. The specific activity of **47** against FP-2 could be confirmed by performing further in vitro experiments in order to rule out possibilities of it being an aggregator.

Inhibition of *P. falciparum* W2 strain

Although only one compound showed inhibitory activity against FP-2, it was encouraging to identify nineteen active compounds (**15**, **17**, **24**, **26**, **27** and **44–57**), ($IC_{50} < 10$ μ M) against *P. falciparum* (Tables 3, 4). Amongst them was compound **47** which gave the highest activity ($IC_{50} \sim 0.5$ μ M). Recently, GlaxoSmithKline (GSK) reported the release of antimalarial compounds into the public domain [50] and Zhang et al. [51] published yet

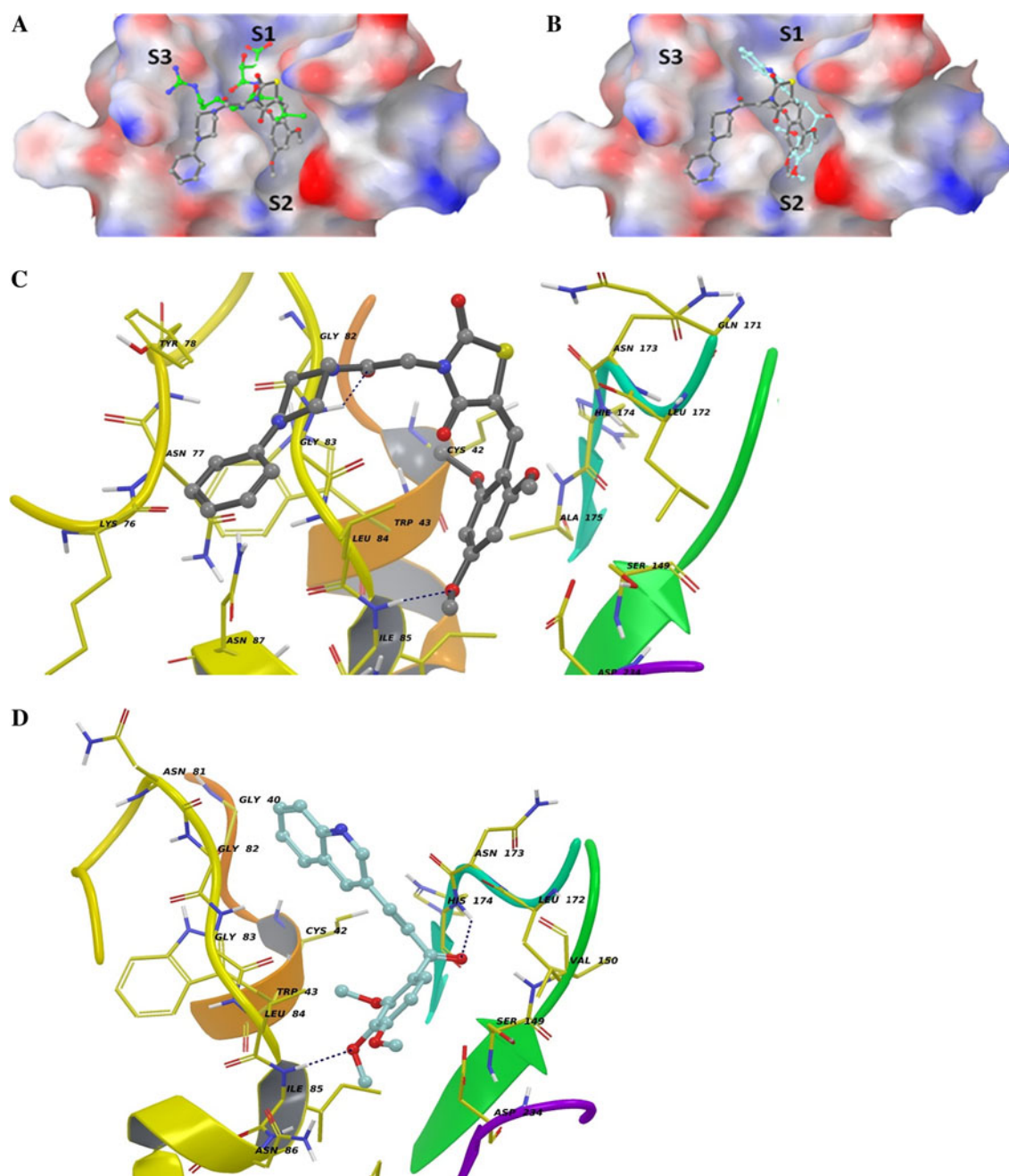


Fig. 4 **a** Electrostatic potential (red negative and blue positive) surface structure of FP-2 showing the subsites *S1*, *S2*, and *S3*, occupied by its inhibitor E64. A comparison between the binding orientations of E64 (green sticks) and compound **47** (grey sticks) is given. **b** A comparison of binding orientation of **47** and inhibitor **37**

(blue sticks). **c** Compound **47** forms a hydrogen bond (blue dots) with Ile85 in the *S2* subsite and with Gly83. **d** Similarly, inhibitor **37** forms an H-bond with Ile85 and His174 of FP-2. Yellow sticks and ribbons depict enzyme residues within a radius of 4 Å from the ligand

another set of active compounds from virtual screening of the ChemBridge database. Comparison of the two datasets with our active compounds, identified that only compound **26** contains the 1-benzyloxy-4-methylbenzene- scaffold also found in two compounds from Zhang et al. 's dataset (SJ000565115 and SJ000565094).

The activity of these compounds against W2 strain could suggest that possibly these compounds are interacting on

other cysteine proteases or on other targets in the malaria parasite. This was also previously suggested by a recent work of Desai et al. [30]. In this post genomic era and with developments in Bioinformatics, the mode of action of these compounds will have to be determine as part of ongoing work.

Our results clearly show that different protocols will generate new sets of hits for a given enzyme and database. For

example Table 3 contains compounds with molecular weight between 200 and 300 whereas Table 4 has larger and greasy compounds (Mwt 400–500). In addition, Table 4 contains five compounds with the oxo-2,4-phenyl-1-piperazinyl scaffold and five quinoline derivatives as well compounds containing the pyridine and imidazole moieties. In Table 3 only the 8-aminoquinoline compound **27** is similar to a well-known antimalarial, primaquine whilst the rest display new scaffolds which include a spiroamine, an azaguanidine, thiozolidine-2,4-dione group, and an amidine group. This implies that virtual screening of a database as big as the ZINC is never exhaustive. Application of different VS approaches will enhance the identification of new and diverse chemical entities that can be used in the development of new anti-malarial drugs.

Conclusion

In this study, two independent virtual screening strategies for *P. falciparum* inhibitors were proposed. The model and screening protocols were first validated theoretically on a literature database of known FP-2 inhibitors and used subsequently to screen the ZINC database, a large library of commercially available compounds. Finally, the screening results obtained were validated experimentally by testing the most promising VS hits in an in vitro biochemical assay against the FP-2 and cultured *P. falciparum*. A total of sixty compounds with chemically diverse structures were tested. From these, nineteen *P. falciparum* inhibitors were identified, with IC₅₀ values ranging between 0.46 and 9.04 μ M. The most active compound against the cultured parasite was also active against the FP-2 enzyme, although the fact that it was 50-times more active against cultured parasites than against FP-2 suggests that its primary target was not FP-2.

Acknowledgments ASN's work was financially supported by Fundação para a Ciência e Tecnologia, through the doctoral Grant SFRH/BD/41276/2007. GM's work was financially supported by the South African National Research Foundation (NRF), Medical Research Council, and University of Cape Town. The University of Cape Town, South African Medical Research Council and South African Research Chairs initiative of the Department of Science and Technology administered through the NRF are gratefully acknowledged for support (KC).

References

- WHO (2011) World malaria report 2011, in WHO library cataloguing-in-publication data 2011. World Health Organization WHO, Geneva
- Greenwood BM, Fidock DA, Kyle DE, Kappe SHI, Alonso PL, Collins FH, Duffy PE (2008) Malaria: progress, perils, and prospects for eradication. *J Clin Invest* 118:1266–1276
- Ballou WR (2009) The development of the RTS, S malaria vaccine candidate: challenges and lessons. *Parasite Immunol* 31(9):492–500
- Wu T, Nagle AS, Chatterjee AK (2011) Road towards new antimalarials—overview of the strategies and their chemical progress. *Curr Med Chem* 18(6):853–871
- Hyde JE (2002) Mechanisms of resistance of *Plasmodium falciparum* to antimalarial drugs. *Microbes Infect* 4(2):165–174
- White NJ (2004) Antimalarial drug resistance. *J Clin Invest* 113:1084–1092
- Dondorp AM, Nosten F, Yi P, Das D, Phyo AP, Tarning J, Lwin KM, Arie F, Hanpithakpong W, Lee SJ, Ringwald P, Silamut K, Imwong M, Chotivanich K, Lim P, Herdman T, An SS, Yeung S, Singhasivanon P, Day NPJ, Lindegardh N, Socheat D, White NJ (2009) Artemisinin resistance in *Plasmodium falciparum* malaria. *N Engl J Med* 361:455–467
- Ridley RG (2002) Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature* 415:686–693
- Pink R, Hudson A, Mouries MA, Bendig M (2005) Opportunities and challenges in antiparasitic drug discovery. *Nat Rev Drug Discov* 4:727–740
- Jackson KE, Habib S, Frugier M, Hoen R, Khan S, Pham JS, de Pouplana LR, Royo M, Santos MAS, Sharma A, Ralph SA (2011) *Trends Parasitol* 27:467–476
- Shenai BR, Sijwali PS, Singh A, Rosenthal PJ (2000) *J Biol Chem* 275:29000–29010
- Sijwali PS, Shenai BR, Gut J, Singh A, Rosenthal PJ (2001) *Biochem J* 360:481–489
- Rosenthal PJ (2011) Cysteine proteases of pathogenic organisms. Springer, US, pp 30–48
- Shah F, Mukherjee P, Gut J, Legac J, Rosenthal PJ, Tekwani BL, Avery MA (2011) *J Chem Inf Model* 51(4):852–864
- Rosenthal PJ, Wollish WS, Palmer JT, Rasnick D (1991) *J Clin Invest* 88:1467–1472
- Rosenthal PJ, Lee GK, Smith RE (1993) *J Clin Invest* 91:1052–1056
- Rosenthal PJ, Olson JE, Lee GK, Palmer JT, Klaus JL, Rasnick D (1996) *Antimicrob Agents Chemother* 40:1600–1603
- Verissimo E, Berry N, Gibbons P, Cristiano MLS, Rosenthal PJ, Gut J, Ward SA, O'Neill PM (2008) *Bioorg Med Chem Lett* 18:4210–4214
- Lee BJ, Singh A, Chiang P, Kemp SJ, Goldman EA, Weinhouse MI, Vlasuk GP, Rosenthal PJ (2003) *Antimicrob Agents Chemother* 47:3810–3814
- Schulz F, Gelhaus C, Degel B, Vicik R, Heppner S, Breuning A, Leippe M, Gut J, Rosenthal PJ, Scheirmeister T (2007) *ChemMedChem* 2:1214–1224
- Martichonok V, Plouffe C, Storer A, Ménard R, Jones B (1995) *J Med Chem* 38:3078–3085
- Micale N, Kozikowski AP, Ettari R, Grasso S, Zappalà M, Jeong J, Kumar A, Hanspal M, Chishti AH (2006) *J Med Chem* 49:3064–3067
- Ettari R, Micale N, Schirmeister T, Gelhaus C, Leippe M, Nizi E, Di Francesco ME, Grasso S, Zappalà M (2009) *J Med Chem* 52:2157–2160
- Ettari R, Zappalà M, Micale N, Grazioso G, Giofrè S, Schirmeister T, Grasso S (2011) *Eur J Med Chem* 46:2058–2065
- Domínguez JN, López S, Charris J, Iarruso L, Lobo G, Semenov A, Olson JE, Rosenthal PJ (1997) *J Med Chem* 40:2726–2732
- Ring CS, Sun E, McKerrow JH, Lee GK, Rosenthal PJ, Kuntz ID, Cohen FE (1993) *Proc Natl Acad Sci* 90:3583–3587
- Coterón JM, Catterick D, Castro J, Chaparro MJ, Díaz B, Fernández E, Ferrer S, Gamo FJ, Gordo M, Gut J, de las Heras L, Legac J, Marco M, Miguel J, Muñoz V, Porras E, de la Rosa JC, de la Ruiz JR, Sandoval E, Ventosa P, Rosenthal PJ, Fiandor JM (2010) *J Med Chem* 53:6129–6152

28. Ekins S, Mestres JJ, Testa B (2007) *Br J Pharmacol* 152:9–20
29. Miteva M (2008) *Biotechnol Biotechnol Equip* 22(1):634–638
30. Desai PV, Patny A, Gut J, Rosenthal PJ, Tekwani B, Srivastava A, Avery M (2006) *J Med Chem* 49:1576–1584
31. Li H, Huang J, Chen L, Liu X, Chen T, Zhu J, Lu W, Shen X, Li J, Hilgenfeld R, Jiang H (2009) *J Med Chem* 52:4936–4940
32. Kerr ID, Lee JH, Pandey KC, Harrison A, Sajid M, Rosenthal PJ, Brinen LS (2009) *J Med Chem* 52:852–857
33. Jones G, Willett P, Glen RC (1995) *J Mol Biol* 245:43–53
34. Irwin JJ, Schoichet BK (2005) *J Chem Inf Model* 45:177–182
35. Durán A, Zamora I, Pastor M (2009) *J Chem Inf Model* 49:2129–2138
36. MOE, Chemical Computing Group Inc Montreal, <http://www.chemcomp.com>
37. Shenai BR, Lee BJ, Alvarez-Hernandez A, Chong PY, Emal CD, Neitz RJ, Roush WR, Rosenthal PJ (2003) *Antimicrob Agents Chemother* 47:154–160
38. Desai PV, Patny A, Sabnis Y, Tekwani B, Gut J, Rosenthal PJ, Srivastava A, Avery M (2004) *J Med Chem* 47:6609–6615
39. Chiyanzu I, Hansell E, Gut J, Rosenthal PJ, McKerrow JH, Chibale K (2003) *Bioorg Med Chem Lett* 13:3527–3530
40. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) *Adv Drug Deliv Rev* 23:3–25
41. Lipinski CA (2000) *J Pharmacol Toxicol Methods* 44:235–249
42. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) *J Mol Biol* 267:727–748
43. DeLano WL (1997) The pymol molecular graphic system. DeLano Scientific, Palo Alto
44. Subramanian S, Hardt M, Choe Y, Niles RK, Johansen EB, Legac J, Gut J, Kerr ID, Craik CS, Rosenthal PJ (2009) *PLoS One* 4:e5156
45. Ettari R, Bova F, Zappalà M, Grasso S, Micale N (2010) *Med Res Rev* 30:136–167
46. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) *J Comput Chem* 30:2785–2791
47. Shah F, Gut J, Legac J, Shivakumar D, Sherman W, Rosenthal PJ, Avery MA (2012) *J Chem Inf Model* 52:696–710
48. Teixeira C, Gomes JRB, Gomes P (2011) *Curr Med Chem* 18:1555–1572
49. Kerr ID, Pandey KC, Harrison A, Sajid M, Rosenthal PJ, Brinen LS (2009) *J Med Chem* 52:852–857
50. Gamo FJ, Sanz LM, Vidal J, de Cozar C, Alvarez E, Lavandera JL, Vanderwall DE, Green DVS, Kumar V, Hasan S, Brown JR, Peishoff CE, Cardon LR, Garcia-Bustos JF (2010) *Nature* 465(7296):305–310
51. Zhang L, Fourches D, Sedykh A, Zhu H, Golbraikh A, Ekins S, Clark J, Connelly MC, Sigal M, Hodges D, Guiguemde A, Guy RK, Tropsha A (2013) *J Chem Inf Model* 53(2):475–492