JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

Identification of Novel Malarial Cysteine Protease Inhibitors Using Structure-Based Virtual Screening of a Focused Cysteine Protease Inhibitor Library

Falgun Shah,[†] Prasenjit Mukherjee,[∥,†] Jiri Gut,[‡] Jennifer Legac,[‡] Philip J. Rosenthal,[‡] Babu L. Tekwani,[§] and Mitchell A. Avery[*,†,§]

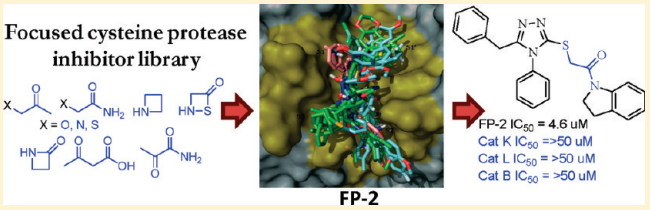
[†]Department of Medicinal Chemistry, School of Pharmacy, University of Mississippi, University, Mississippi 38677, United States

[‡]Department of Medicine, San Francisco General Hospital, University of California, San Francisco, California 94143, United States

[§]National Center for Natural Products Research, University of Mississippi, University, Mississippi 38677, United States

Supporting Information

ABSTRACT: Malaria, in particular that caused by *Plasmodium falciparum*, is prevalent across the tropics, and its medicinal control is limited by widespread drug resistance. Cysteine proteases of *P. falciparum*, falcipain-2 (FP-2) and falcipain-3 (FP-3), are major hemoglobinases, validated as potential antimalarial drug targets. Structure-based virtual screening of a focused cysteine protease inhibitor library built with soft rather than hard electrophiles was performed against an X-ray crystal structure of FP-2 using the Glide docking program. An enrichment study was performed to select a suitable scoring function and to retrieve potential candidates against FP-2 from a large chemical database. Biological evaluation of 50 selected compounds identified 21 diverse nonpeptidic inhibitors of FP-2 with a hit rate of 42%. Atomic Fukui indices were used to predict the most electrophilic center and its electrophilicity in the identified hits. Comparison of predicted electrophilicity of electrophiles in identified hits with those in known irreversible inhibitors suggested the soft-nature of electrophiles in the selected target compounds. The present study highlights the importance of focused libraries and enrichment studies in structure-based virtual screening. In addition, few compounds were screened against homologous human cysteine proteases for selectivity analysis. Further evaluation of structure—activity relationships around these nonpeptidic scaffolds could help in the development of selective leads for antimalarial chemotherapy.



INTRODUCTION

Malaria, a mosquito-borne parasitic disease, remains a serious health problem in the most of the developing world.[1] The World Health Organization has estimated that malaria accounts for close to a million deaths annually, nearly all due to *Plasmodium falciparum*, the most virulent human malaria parasite.[2] The regions where malaria is endemic include tropical and subtropical regions of Africa, Asia, and South America.[3] Falciparum malaria is a particular threat for Africans, especially children under the age of five. A number of drugs are currently available to treat malaria,[4] however, treatment is complicated by the toxicity, high cost, and diminishing efficacy of different agents. Recently, with increasing resistance to other agents, artemisinin-based combination therapy (ACT) has been recommended for the treatment of uncomplicated falciparum malaria in Africa and most other malaria-endemic countries.[5−7] However, ACTs remain in limited supply, are rather expensive, and may suffer from resistance to both the artemisinin component and the partner drugs. With concerns about the long-term efficacy of all available antimalarials, there is a great need for new therapies, ideally directed against new targets. Among potential new targets for antimalarial chemotherapy are cysteine proteases. In the past decade, several mammalian, viral, or protozoal cysteine proteases are targets of drug discovery efforts, including programs to treat osteoporosis and Chagas disease.[8] For example, odanacatib, Merck's investigational cathepsin K (mammalian cysteine protease) inhibitor, is currently in phase III clinical trials for osteoporosis and bone metastasis.[9] The best-studied cysteine proteases of *P. falciparum* are a set of papain-family proteases known as falcipains (FPs).

Among FPs, FP-2 and FP-3, encoded by nearby genes, are localized to the *P. falciparum* food vacuole, the site of hydrolysis of large quantities of hemoglobin by erythrocytic parasites. FP-2 and FP-3 have been shown to be hemoglobinases,[10] and they appear to be responsible for the hydrolysis of the host erythrocyte hemoglobin to provide amino acids required for protein synthesis in protozoa.[11−13] Knockout of FP-2 led to parasites with diminished hemoglobin hydrolysis.[11] Knockout of FP-3 was not possible, strongly suggesting that this protease is essential for erythrocytic parasites.[14] Functions of the other falcipains are less

Received: November 9, 2010
Published: March 23, 2011

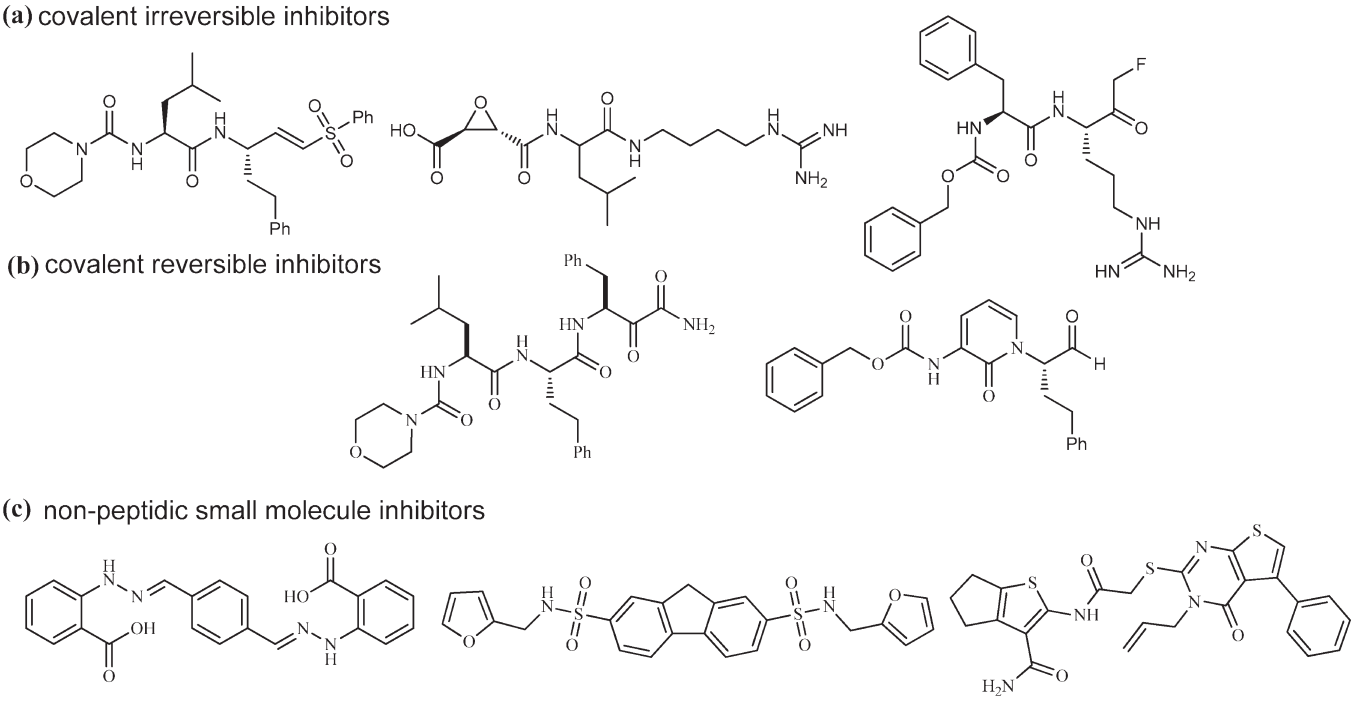


Figure 1. Representative FP-2 inhibitors: (a) covalent irreversible inhibitors;[20,23,34] (b) covalent reversible inhibitors;[20,25] and (c) nonpeptidic small molecule inhibitors.[36−38]

well understood. FP-1 activity was associated with erythrocyte invasion by *P. falciparum* merozoites,[15] but knockout of the protease did not alter erythrocytic parasites,[16] and its primary function is more likely in mosquito-stage parasites.[17] FP-2′ is a near-identical copy of FP-2, but knockout of FP-2′ had no apparent phenotype, indicating that the protease is not essential.[14] In any events, inhibition of FP-2 and FP-3 blocks parasite development, enabling these enzymes as viable targets for the development of antimalarial chemotherapy.

Extensive efforts were conducted for the discovery of inhibitors against FPs and are reviewed elsewhere.[18,19] In general, drug design efforts against FPs can be broadly classified in two main categories: peptide/peptidomimetic covalent inhibitors and nonpeptidic small molecule inhibitors. The covalent irreversible inhibitors contain reactive warheads or electrophiles, such as vinyl sulfones,[20,21] epoxy-succinate,[22,23] and fluoromethyl ketones (FMKs),[22,24] whereas compounds with electrophiles, such as α-keto amide[25] and aldehydes,[25] function as a reversible covalent inhibitor of the FPs (See Figure 1a and b). However, differentiation between an irreversible covalent and a reversible covalent category is very subtle, depending on the reactivity of electrophilic groups. For example, conjugate addition of thiol to enone can be reversible as reported in the case of cyclopropenone containing cysteine protease inhibitors[26] and for chalcone-based malarial cysteine protease[27] and proteasome inhibitors.[28] Similarly, nitriles containing inhibitors of cysteine proteases, although known to have covalent reversible binding mode, can exhibit irreversible binding to the proteins depending on their reactivity with cysteine sulfur.[29] The covalent modifiers with reactive warheads blocked the hydrolysis of hemoglobin by erythrocytic parasites, formed hemoglobin-filled food vacuoles in trophozoites and hampered parasites development.[22,24,30−33] Despite their excellent in vitro potency against cultured parasites, these compounds exhibited poor selectivity toward the parasitic cysteine protease over mammalian cysteine proteases.[33−35] In addition, potential toxicity associated with these compounds due to the formation of protein adducts restricted their further development as potential drug candidates. Moreover, peptidic nature of these inhibitors makes them susceptible to hydrolysis by host cell proteases. The other categories of FP-2 inhibitors include nonpeptidic small molecule inhibitors of the enzymes (See Figure 1c). Several structure-based virtual screening attempts were made to identify nonpeptidic reversible inhibitors against 3D models of FPs from the commercial databases.[36−38] However, to date, only a few nonpeptidic inhibitors of FP-2 are available with a lower micromolar (<10 μM) affinity as well as with the ability to inhibit cultured parasites. Thus, designing nonpeptidic, noncovalent inhibitors of FPs is an important goal, with the likelihood of being able to develop compounds with desirable druggable properties, good in vivo potency, and biochemical selectivity.

Over the past decade, structure-based virtual screening (SBVS) has gained attention and successfully identified novel scaffolds against different targets.[39] In brief, SBVS involves the computational docking of a commercial database into the active site of the 3D structure of the target, followed by the selection of putative binders using scoring functions.[19,40−42] With this approach, only a small set of compounds is selected and evaluated in a biological assay, as opposed to the high-throughput screening of an entire compound library.[43,44] As a part of our efforts to discover nonpeptidic reversible inhibitors of FP-2, we conducted SBVS against the crystal structure of FP-2 (PDB code: 2GHU)[45] using the Glide docking program. We first built a focused cysteine protease inhibitor (FCPI) library by substructure searching for compounds bearing soft electrophiles of interest from the pool of several commercial databases. The developed FCPI library was then screened against the crystal structure of FP-2 (PDB code: 2GHU, apo form of the enzyme) to discover novel noncovalent inhibitors of the enzyme. Evaluation of the selected

dx.doi.org/10.1021/ci200029y |*J. Chem. Inf. Model.* 2011, 51, 852–864

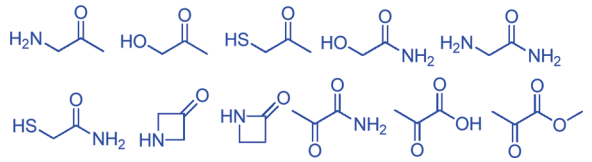


Figure 2. Soft electrophiles used in the substructure search to build FCPI library.

set of compounds identified 21 diverse nonpeptidic inhibitors of FP-2 with a lower micromolar potency and a potential for further structural optimizations in to the lead candidates. Four compounds also inhibited FP-3. Two compounds showed corresponding inhibition of the cultured parasites. In addition, selected compounds were evaluated against mammalian cysteine proteases (cathepsin B, K, and L) of the papain family. One compound was found to be a selective inhibitor of FP-2. Finally, an assessment of electrophilicity of electrophiles in selected target compounds was carried out by calculation of lowest unoccupied molecular orbital (LUMO) density of each atom. The results suggested the soft-nature of electrophiles in identified hits.

METHODS

Computational Tools. All calculations were performed on a Linux workstation equipped with four parallel Intel Xeon X5460 processors (3.16 GHZ) with 8 GB total RAM. LigPrep[46] was used to produce low-energy 3D structures of compounds. Glide docking program[47] was used for pose validation, enrichment study, and SBVS. QikProp[48] was used to calculate chemical properties of the FCPI library. Phase[49] was used to build a docking pose-based pharmacophore query. Diversity analyses of the final hits were performed using the Dissimilarity module available in Sybyl.[50] Geometry optimization and measurement of electrophilicity of identified compounds were performed using atomic Fukui indices implemented in Jaguar.[51] All images were created using Pymol.[52]

Building the Focused Cysteine Protease Inhibitor (FCPI) Library. The selected soft electrophiles of interest used for building of FCPI library include α-heteroatom-substituted ketones and amides, azetidinones, α-keto amide, α-keto acid, and α-keto ester (see Figure 2). Substructure searches were conducted with single line notations (SLN) of the soft electrophiles (shown in Figure 2) as a query in Asinex platinum (http://www.asinex.com), Chembridge (http://www.hit2lead.com), Specs (http://www.specs.net), Enamine (http://www.enamine.net), IBScreen (http://www.ibscreen.com), and Aurora Fine Chemicals Ltd. (http://www.aurorafinechemicals.com), using the dbsearch command from the Unity search module of Sybyl 8.1 (Tripos Inc., St. Louis, MO). The identical compounds obtained from a pool of different databases were removed. Thereafter, the database was filtered using chemical properties calculated by QikProp, such as molecular weight (MW): 250−550, H-bond acceptor (HBA): 1−10, H-bond donor (HBD): 0−5, and number of rotatable bonds: 0−10. In addition, compounds containing problematic groups, such as metals, N-oxides, chloramines, aldehydes, and peroxides, were removed from the FCPI library. The dbslnfilter utility of the Unity module of Sybyl 8.1 was used to carry out database filtration. Around 65 000 compounds with desired soft electrophiles were thus stored in the FCPI library.

Protein Preparation. In the present SBVS study, we utilized a recently published crystal structure of FP-2 available in the apo form (PDB code: 2GHU).[45] The residues within 14 Å of the Cys42 were included in the binding site definition. The PPREP utility of Maestro (Schrodinger, LLC, Portland, OR) was used for protein preparation. The protein preparation was carried out at pH = 5 to mimic the acidic environment of the *P. falciparum* food vacuole (pH = 4−6), where the enzyme is located. Acidic residues Asp and Glu and basic residues Arg and Lys were treated as charged unless they were surrounded by hydrophobic residues. Figure 3a shows the magnified view of the FP-2 active site with subsites S1, S1′, S2, and S3 labeled. Protein with all hydrogens added was then submitted to restrained molecular mechanics refinement using the OPLS2001 force field incorporated in the IMPREF (Schrodinger, LLC, Portland, OR) protein structure refinement utility, and minimization was continued until the root-mean-square deviation (rmsd) reached 0.3 Å. The final refined structure was used for the docking calculations.

Pose Validation, Enrichment Study, and Docking Pose-Based Pharmacophore Query. In the absence of the cocrystallized structure of FP-2, a preliminary docking study was carried out in the homologous cysteine protease cruzain, the major cysteine protease of *Trypanosoma cruzi* (*T. cruzi*), cocrystallized with a vinyl sulfone (VS) inhibitor (PDB ID: 1F2A).[53] The criteria for selecting cruzain as a surrogate platform to establish a docking protocol are as follows: (a) both cruzain and falcipains are papain family cysteine proteases; (b) the superimposition of FP-2 (2GHU), and cruzain (1F2A), using DaliLite server,[54] matches 209 α-carbons with an rmsd of 1.6 Å, a Z score of 30.4, and an overall sequence identity of 38%. Both proteins were found to have 62% of sequence identity and a rmsd of 2.6 Å in the ligand binding domain, as defined by 10 Å radius from catalytic cysteine (Cys42 in FP-2 and Cys25 in cruzain, see Figure S0, Supporting Information); (c) vinyl sulfone inhibitors, for example, shown in Figure 1a is a common inhibitor of both enzymes.[21,53]

Structural alignment of cruzain complexes, for example, with the irreversible covalent vinyl sulfone inhibitor (PDB code: 1F2A, 2OZ2) as well as the reversible covalent inhibitors, such as hydroxyl methyl ketones (PDB code: 1ME3), and a purine nitrile (PDB code: 3I06) using superimposition module of maestro showed an average rmsd of 0.45 Å. This suggests that receptor structure does not change substantially upon the binding of inhibitors from different chemotypes. Therefore, rigid receptor docking protocol implemented in Glide was used throughout the study. Once initial docking protocol is established in cruzain, subsequent enrichment and docking studies were performed in the crystal structure of FP-2.

A pose validation study was performed using the Glide XP mode without any docking constraints. The top ranking pose obtained for the VS inhibitor in cruzain using the Gscore scoring function was closest to its experimental binding mode (see Figure 3b). The rmsd between predicted and experimental poses of VS inhibitor was found to be 0.97 Å, which was quite satisfactory considering the number of rotatable bonds of VS (structure shown in Figures 3b and d). Subsequent docking of VS inhibitor in FP-2 active site reproduced the arrangements of P1−P3 groups, as observed in X-ray structure of cruzain. Moreover, H-bonds of VS inhibitor with Trp 177, Gln19, Gly66, and Asp158 of cruzain binding site were reproduced in the FP-2 active site.

The actives from the previous SBVS study[36,37] (having FP-2 $IC_{50} \leq 30\ \mu M$, see Figure S1 and Table S1, Supporting

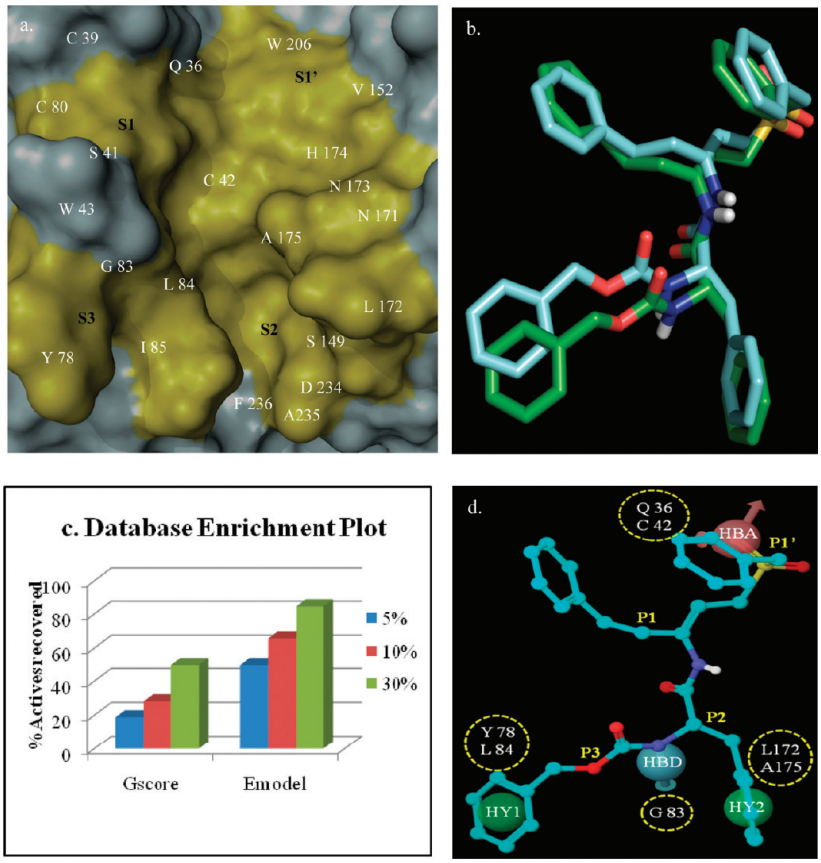


Figure 3. (a) Magnified view of the substrate-binding pocket of FP-2 (PDB code: 2GHU). Subsites S1, S1′, S2, and S3 of the enzyme are shown. (b) Pose validation study: superimposition of cocrystallized pose (green) and the top-ranked pose (cyan) obtained through docking. (c) Enrichment study: a bar graph showing percentage of actives recovered at 5% (blue), 10% (red), and 30% (green) of the ranked database with Gscore and Emodel scoring functions. (d) Docking pose-based pharmacophore query: HBD = H-bond donor; HBA = H-bond acceptor; HY = hydrophobic. P1, P1′, P2, and P3 side chains of VS and interacting residues Gln36, Cys42, Gly83, Tyr78, Leu84, Leu172, and Ala175 of FP-2 binding site are shown.

Information) were selected for the enrichment study. Active compounds were seeded with a total of 1000 decoys downloaded from the Directory of Useful Decoys (DUD) database.[55] Decoys were selected by considering the chemical properties such as MW: 250−550, number of HBD: 2−4, and number of HBA: 2−8 of known actives. Selection of the diverse sets of decoys resembling properties of known actives made the enrichment study quite challenging.

In addition, a receptor-based pharmacophore prefilter was generated based on the docking pose of the peptidic vinyl sulfone in the crystal structure of FP-2 (PDB code: 2GHU) encoding key receptor−ligand interactions. Phase 3.0 was used to build the pharmacophore query. The pharmacophore shown in Figure 3d consists of two hydrophobic features (green) mapped onto the P2 and P3 side chains of VS interacting with hydrophobic residues of S2 and S3 pockets: a HBD feature (cyan) mapped onto the −NH of the P3 side chain forming H-bond with Gly83 and a H-bond acceptor (pink) forming the H-bond with the residues of oxyanion hole (Cys42 and Gln36). As in the S2/S3 pockets, there is no strict requirement for the ring aromatic feature, the built-in definition for the directional ring aromatic feature was excluded for performing the search against the FCPI library, and the SMART patterns for aromatic rings were included in the hydrophobic feature definition.

SBVS Protocol. The FCPI library consisting of ∼65 000 compounds was prepared for virtual screening using the ligprep module of Schrodinger (see Figure 4). Compounds were subjected to hydrogen additions, removal of salt, ionization, and generation of low-energy ring conformations. The chiralities of the original compounds were preserved. Finally, the low-energy 3D structures of all compounds were produced. To curtail the database prior to docking, a prepared database was prefiltered using a four-point structure-based pharmacophore query based on the docking pose of vinyl sulfone. Compounds matching a minimum three out of four features of the pharmacophore query were retained. At this stage, filtered compounds were expected to possess soft electrophiles of interest and relevant pharmacophoric features of FP-2 inhibitors, which increase the probability for docking of more relevant compounds in the active site of FP-2. Approximately 30 000 compounds thus retained were subjected to docking in the crystal structure of FP-2 using the previously validated Glide XP protocol. The top 3000 molecules (top 10%) ranked on the basis of Emodel score were selected for eMBrAce minimization calculation (Macromodel). The eMBrAcE energy minimization in the energy difference mode was carried out for the protein−ligand complexes to study the association of ligands with the receptor. The energy difference was calculated as follows:

$$\Delta E = E_{complex} - E_{ligand} - E_{protein}$$

where ΔE = energy changes upon association; $E_{complex}$ = energy of the receptor−ligand complex; E_{ligand} = energy of ligand alone, and

dx.doi.org/10.1021/ci200029y |J. Chem. Inf. Model. 2011, 51, 852–864

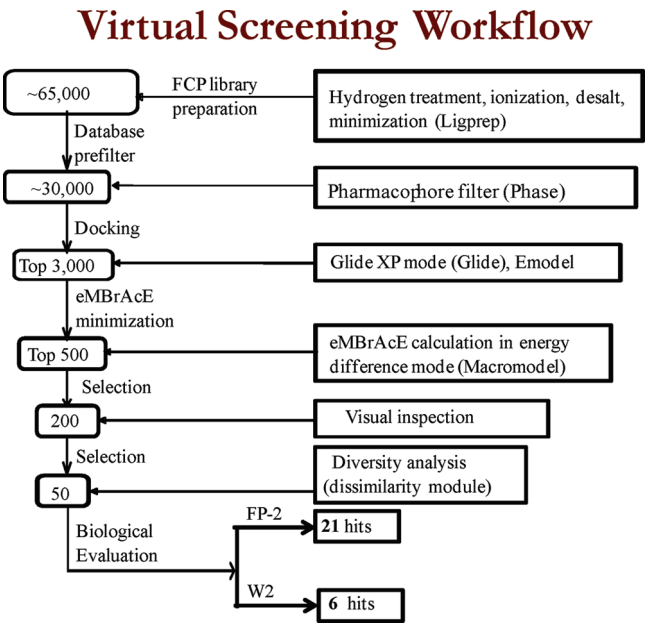


Figure 4. A flowchart depicting the virtual screening protocol utilized in the study. W2 stands for chloroquine resistance strain of *P. falciparum*.

$E_{protein}$ = energy of protein alone. Based on EMBrACE score, the top 500 molecules were selected.

Next, compounds were visually inspected for the following criteria: (a) presence of the soft electrophile in the vicinity of catalytic Cys42; (b) key hydrophobic (Leu172, Leu84) or H-bonding interactions (Gln36, Gly83) between ligand atoms and protein residues; (c) the internal geometry of the ligand in the binding site; (d) occupancy of the S2 subpocket by the hydrophobic group of the ligand (hemoglobin cleavage studies demonstrated strong preference for cleavage sites with Leu at P2 position as a major determinant of the specificity against this enzyme,[56] therefore, hydrophobic interactions of ligand in S2 subsite were considered essential); and (e) commercial availability of putative hits. Based on these criteria, 200 molecules were selected and subjected to diversity analysis using the Dissimilarity module of Sybyl 8.1 (Tripos Inc., St. Louis, MO). A total of 50 diverse compounds were purchased from respective vendors. The Tanimoto similarity matrix was generated to calculate an average pairwise Tanimoto similarity index of identified hits using similarity/distance matrix utility implemented in Canvas, version 1.2 (Schrödinger, LLC, New York, NY), using atom-pair fingerprints. The comparisons of the obtained hits in the present study with those from previous SBVS studies were carried out using the dbcmpr utility of Sybyl 8.1 (http://www.tripos.com). The comparison was done by constructing a UNITY database of each set using Sybyl 8.1.

Biological Evaluation. The selected 50 compounds were evaluated for inhibition of FP-2, FP-3 and against chloroquine resistant (W2 strain) *P. falciparum* parasites. To determine IC_{50} values, recombinant FP-2 and FP-3 were 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 Z-Leu-Arg-AMC (benzoxycarbonyl-Leu- Arg-7-amino-4-methyl-coumarin) in the same buffer was added to a final concentration of 25 μM. 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 activity over compound concentration using GraphPad Prism software.

The assays for selected human cysteine proteases, cathepsin B, K, and L were conducted in a similar way, as described above for FP-2. The assay concentration used for cathepsin B, K, and L were 7.94, 3.40, and 5.38 nM, respectively. The appropriate substrates (Z-Phe-Arg-AMC for cathepsin K and L and Z-Arg-Arg-MCA for cathepsin B) were added to a final concentration of 25 μM.

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 μM hypoxanthine in 96-well culture plates. Parasites were synchronized with 5% sorbitol.[57] All tested compounds were prepared as 10 mM stock solutions in DMSO and diluted in medium at least 1:1000, resulting in $\leq$0.1% final concentrations of DMSO. 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 μL aliquots were incubated for 15 min in the dark in 150 μL of 100 mM NH_4Cl, 0.1% Triton X-100 in PBS, and 10 nM YOYO-1 (Molecular Probes). Parasitemia was determined based on counts from a FACSort flow cytometer (Beckton Dickinson) using CellQuest software (Beckton Dickinson). IC_{50} values for growth inhibition were determined from plots of percent control parasitemia over inhibitor concentration using Prism software (GraphPad). Compounds were also evaluated for cytotoxicity on VERO (monkey kidney fibroblast) cells by the neutral red assay.[58] IC_{50} values for each compound were computed from the growth inhibition curve.

Calculation of Electrophilicity of Soft-Electrophiles in Identified Hits. Atomic Fukui indices, derived from Mulliken population of the highest occupied molecular orbital (HOMO) and the LUMO, were used to quantify electrophilicity of a molecule at a particular atomic site. Two different sets of Fukui indices calculated include the HOMO and the LUMO. Each index has two subscripts: N stands for the electron density and S stands for the spin density. Thus, a total of four indices were calculated for both, the HOMO and the LUMO: F_NN, F_NS, F_SN, and F_SS. The first index shows which property responds to a change in the property indicated by the second index, hence, F_NN indicates the change in the electron density when only the electron density is changing and the spin density is constant. Similarly, F_NS represents the change in the electron density when only the spin density is changing, etc. F_NN_LUMO was considered for measuring electrophilicity of electrophilic center in the target compounds. Structures were subjected to the geometry optimization using the hybrid density functional B3LYP[59,60] and a basis set 6-31G* in a gas phase using Jaguar. The default convergence criterion implemented in Jaguar was used for self-consistent field (SCF) calculations (accuracy level= quick, convergence criteria: maximum iteration = 48, and energy change = 5×10^{-5} hartree) and optimization (maximum steps = 100, convergence criteria = default, initial Hessian = Schlegel guess).[61] The fully optimized geometry of compounds were subjected to the calculation of atomic Fukui indices implemented in the property section of the Jaguar optimization panel, according to formalism presented elsewhere.[62,63] Fukui function for the electron density integrates to one, so the predicted reactivity here

Table 1. Biological Evaluation of Compounds Against Falcipains and Cultured Malarial Parasites

compound code[d]	FP-2[a] IC_{50} (μM)	FP-3[b] IC_{50} (μM)	W2[c] IC_{50} (μM)	compound code[d]	FP-2[a] IC_{50} (μM)	FP-3[b] IC_{50} (μM)	W2[c] IC_{50} (μM)
1	1.39	>50	NA[e]	**12**	23.81	>50	NA
2	2.18	4.95	NA	**13**	25.25	>50	NA
3	3.26	>50	NA	**14**	25.96	>50	NA
4	4.59	>50	NA	**15**	25.99	>50	NA
5	7.51	30.27	1.90	**16**	28.66	>50	NA
6	7.85	>50	4.34	**17**	32.85	>50	NA
7	10.32	>50	NA	**18**	37.34	>50	1.92
8	11.58	45.64	NA	**19**	44.94	>50	NA
9	11.86	>50	NA	**20**	46.97	>50	NA
10	15.11	>50	1.57	**21**	49.18	>50	5.83
11	20.87	36.05	NA	**22**	>50	>50	7.21
E-64	0.08	0.15	–	**E-64**	0.08	0.15	–
artemisinin	–		0.01	artemisinin	–		0.01
chloroquine	–		0.05	chloroquine	–		0.05

[a] Falcipain-2. [b] Falcipain-3. [c] Antimalarial activity against chloroquine resistant (W2) strain of *P. falciparum*. [d] All compounds were noncytotoxic at least up to 50 μM in VERO (monkey kidney fibroblast) cells. E-64, artemisinin, and chloroquine were used as positive controls. [e] NA = not active up to 50 μM.

is scaled from 0 to 1, 1 being the most reactive. It is worthwhile mentioning that in all cases, the SH group of the active site cysteine has been assumed as a nucleophile or HOMO component of a nucleophilic−electrophilic addition reaction.

RESULTS AND DISCUSSION

Out of 50 compounds submitted for biological testing, a total of 21 diverse nonpeptidic hits were identified. The IC_{50} values of hits for FP-2 inhibition ranged from 1.4 to 49 μM (shown in Table 1). Four compounds **2**, **5**, **8**, and **11** were found to be dual inhibitors of FP-2 and FP-3. The structures of identified inhibitors are depicted in Figure 5.

A major bottleneck in SBVS is finding a suitable scoring function for the target of interest to enable the identification of potential active compounds prior to the screening of large chemical databases. Novel approaches to address this problem include the use of a target-biased scoring function, which takes into consideration the nature of the target (polar vs apolar binding site) as well as the nature (pharmacophoric patterns and similarity considerations) and key interaction of known ligands with the target.[64] In the present study, two scoring functions implemented in the Glide module of Schrodinger, Gscore and Emodel, were evaluated for their abilities to retrieve known FP-2 inhibitors from decoys downloaded from the DUD database. The recovery rates (Figure 3c) showed that among the scoring functions evaluated, Emodel performed better than Gscore in 5%, 10%, and 30% database screens. At the 5% database screened mark, the percentage of actives recovered was 50% for Emodel, while at the 10% and 30% database screened marks, the percentage of actives recovered was 66% and 84% by the Emodel scoring function as compared to 28% and 50%, respectively, for the Gscore. These results were quite encouraging considering the diversity of the actives used in this study and enhanced our confidence in using Emodel as a scoring function for the SBVS study. The higher enrichment rate associated with Emodel might be because of significant weighting of force field components such as the Coulomb and van der Waals energy in Emodel as compared to Gscore.[65] These interactions may be a major driving factor for the ligand binding for the apolar FP-2 binding site, and, as a result, Emodel performed better with the FP-2 active site than Gscore. A similar observation was made by Li et al. in their virtual screening study, suggesting the importance of Dock energy score (GAsDock) over a Chemscore-based (GScore) scoring function more suitable for the apolar FP-2 binding site.[38] The success ratio for the present study in finding novel chemotypes was 42%, which suggest the score-based active/inactive separation ability of current docking protocol against this target. The higher hit ratio also implied the importance of scoring function-based enrichment studies prior to SBVS.

Prior to virtual screening of FCPI library, the database was curtailed with a docking pose-based pharmacophore query. Key interactions of vinyl sulfone inhibitor with residues of FP-2 active site, such as Gly83, Cys42, Gln36, along with the hydrophobic residues of S2 and S3 pocket (shown in Figure 3d) were used to build the pharmacophore query, consisting of four descriptors: HBD, HBA, HY1, and, HY2. The selection of pharmacophoric descriptors were based on a thoughtful consideration of crucial interactions of cysteine protease inhibitors with the conserved residues of the papain family cysteine proteases. For example, Gly83 is highly conserved in the S3 subsite of clan CA cysteine proteases[13,66] (Figure S4, Supporting Information), and the lack of potency was reported for the cysteine protease inhibitors particularly binding to the nonprime (S1−S4) site lacking H-bond with the carbonyl oxygen of glycine.[67,68] Similarly, inhibitors of the papain family of enzymes are commonly shown to form H-bonding interactions with residues forming an oxyanion hole, such as Gln36 and the backbone amine of catalytic cysteine Cys42 in FP-2[67,69,70] (Figure S4, Supporting Information). Moreover, as with the most other papain family of proteases, analysis of the specificity of FP-2 indicated that the amino acid at the P2 position plays a key role in mediating substrate specificity, at least with peptide substrates.[13] In addition, substrate mapping and inhibitor profiling study against FP-2 revealed strong preference for hydrophobic residue at P2 (Leu > Phe > Val) in the S2 Pocket.[13,35] Finally, the recently published X-ray structure of FP-2 bound to a covalent irreversible epoxysuccinate inhibitor (E-64)[23] gives an additional validity to our selection of docking pose-based pharmacophoric descriptors (See Figure S2, Supporting Information).

dx.doi.org/10.1021/ci200029y |*J. Chem. Inf. Model.* 2011, 51, 852–864

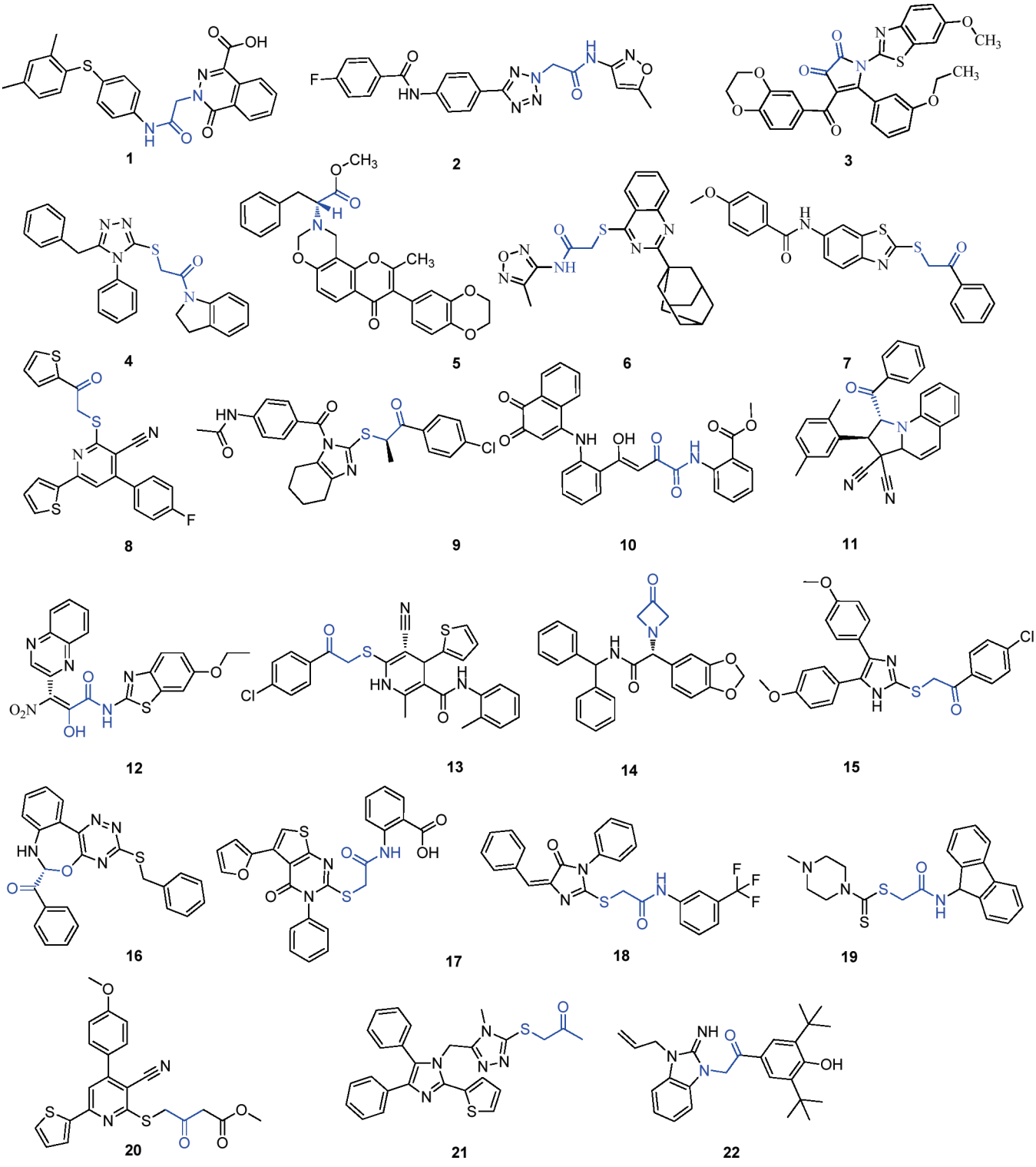


Figure 5. Structures of the FP-2 inhibitors identified through SBVS of FCPI library. Substructure containing electrophiles of interest are highlighted in blue.

The carboxyl group of E-64 occupies the oxyanion hole formed by Gln36 and Cys42, resembling HBA feature of our pharmacophore query. Similarly, peptidyl backbone amine of E-64 forms H-bond with a backbone carbonyl of Gly83, similar to assigned HBD descriptor of pharmacophore query. The P2-leucine of E-64 exhibits hydrophobic interactions in the S2 pocket with Ile85, L172, and Ala175, whereas P3 side chain of E-64 exhibited hydrophobic interaction with Gly83, Leu84, and, cation$-\pi$ interactions with the Tyr78 in S3 pocket, relating to the HY1 and HY2 descriptors of docking pose-based pharmacophore query.

As mentioned before, the main drawback of covalent modifiers (for example, compounds bearing electrophiles such as α,

dx.doi.org/10.1021/ci200029y |*J. Chem. Inf. Model.* 2011, 51, 852–864

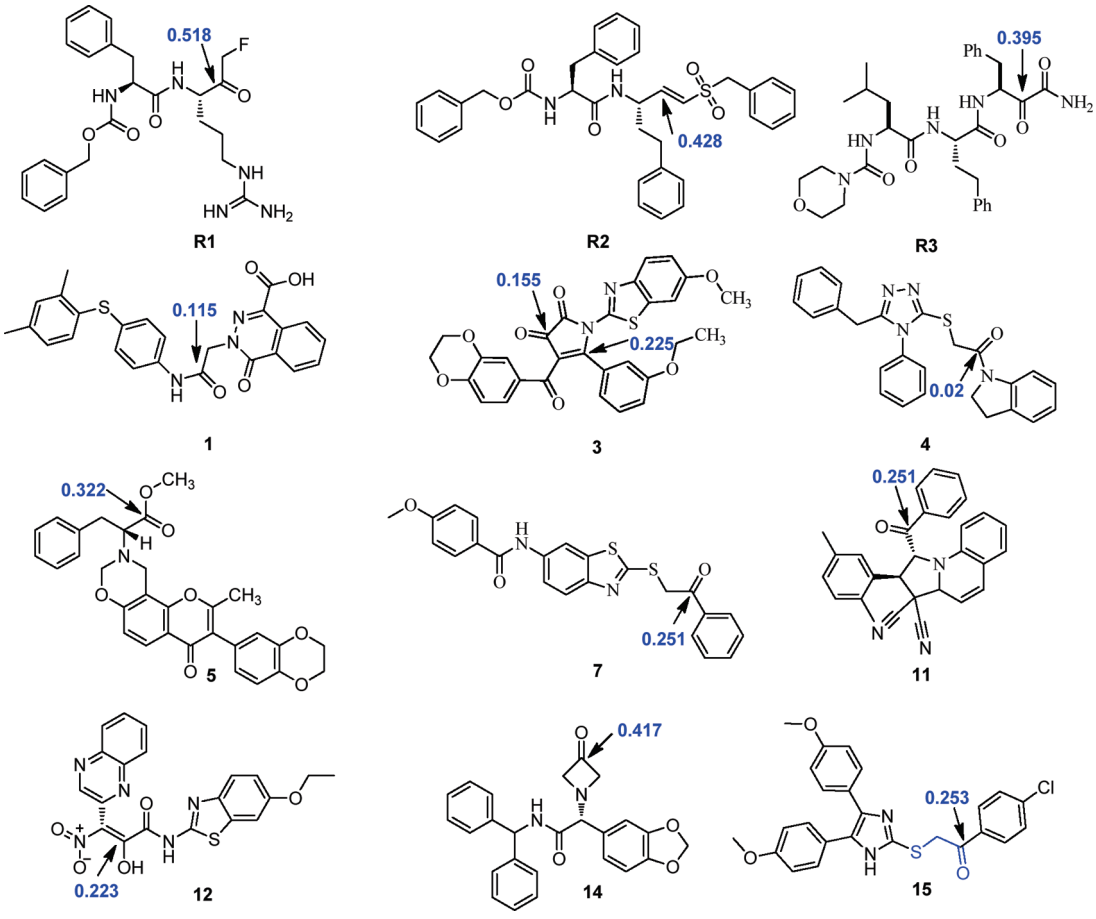


Figure 6. An evaluation of regional Fukui functions to predict the electrophilicity of atomic site of reference compounds (**R1**−**R3**) and the selected FP-2 inhibitors is shown here. The most electrophilic center in the compounds is shown by the arrows. The corresponding value of the Fukui index, F_NN_LUMO, for the reactive atoms (scaled from 0 to 1) is shown in blue.

β-unsaturated ketones, α-halo ketones, cynamide) is lack of target specificity. In addition, nonspecific binding of covalent modifiers to protein, DNA, or glutathione (GSH) may lead to unfavorable toxicological events, such as immunogenic response or life-threatening idiosyncratic adverse drug reactions (ADRs).[71,72] Soft-electrophiles attenuate the reactive nature of hard electrophiles and might reduce the risk of toxicity associated with covalent modifiers. Thus, use of the soft-electrophiles based FCPI library for SBVS provides a strategic approach to discover novel, selective inhibitors, preferably modulating the enzyme through noncovalent interactions. The soft-electrophiles selected here had previously shown to form tetrahedral intermediates with cysteine proteases, for example, α-heteroatom substituted ketones,[70,73−75] α-keto amide,[76,77] α-keto acid,[77] α-keto ester,[77] and azetidinones.[68] To assess the soft-nature of electrophiles present in target compounds, we performed calculations of the atomic Fukui indices on a few selected compounds (**1**, **3**−**5**, **7**, **11**, **12**, **14**, and **15**) containing diverse soft-electrophiles. Fukui indices are partial second derivatives of the electron or spin density with respect to a change in the electron count or the unpaired spin count. They are regional descriptors of site reactivity predicting which atoms in a molecule are most reactive toward electrophilic or nucleophilic attack based on the analysis of frontier molecular orbitals (FMOs) in a molecule.[62] The greater the magnitude of an index, the greater change in electron or spin density near the atoms of interest and thus the higher reactivity of a molecule at that atomic site for an electrophilic or a nucleophilic reaction.[78] Generally, a nucleophilic reaction takes place preferably in the substrate sites with the largest values for the LUMO density. Likewise, an electrophilic reaction is most likely to occur at the largest HOMO electron density in the molecular site.[63,79] Since our goal was to probe the most electrophilic center in identified FP-2 hits and to predict its electrophilicity, we considered calculation of Fukui indices, F_NN_LUMO at a constant spin density in selected compounds. The electrophilicity of electrophilic centers present in identified FP-2 hits were compared with the known cysteine protease inhibitors with an irreversible covalent (α-halo ketone, **R1**; vinyl sulfone, **R2**) and a reversible covalent (α-ketoamide, **R3**) warhead group. The values of Fukui index F_NN_LUMO, of the most electrophilic atoms of the FP-2 inhibitors, and of reference compounds (**R1**−**R3**) are shown in Figure 6. The detailed calculations of all four Fukui indices are available in Tables S3−S14, Supporting Information.

The result from the calculation (see Figure 6) demonstrates that the site-specific reactivity of the identified FP-2 inhibitors is less than the irreversible covalent inhibitors: **R1** with an α-halo ketone (LUMO = 0.518) and **R2** with a vinyl sulfone warhead group (LUMO = 0.428) and are comparable to the known reversible covalent inhibitor, **R3** with an α-keto amide (LUMO = 0.395) moiety. These calculations suggest the soft-nature of the electrophilic atoms in the identified FP-2 hits. However, one

dx.doi.org/10.1021/ci200029y |*J. Chem. Inf. Model.* 2011, 51, 852–864

exception here is compound **14** with a higher than expected electrophilicity value, presumably because of the ring strain of an azetidin-3 ones moiety, which is reflected in the calculated LUMO value (0.417). On the basis of the predicted site reactivity of soft-electrophiles, compounds with an α-hetero ester (**5**, LUMO = 0.322) were found to be more electrophilic than those with an α-hetero ketone (**7**, **11**, LUMO = 0.251; **15**, LUMO = 0.253), which in turn had a higher electrophilicity than hits with an α-hetero amide functionality (**1**, LUMO = 0.115; **4**, LUMO = 0.020). One point worth mentioning here is that the calculated electrophilicity is influenced by the complex interplay of factors, such as a neighboring group effects, the internal geometry of the compound, intermolecular H-bonding and the solvation effects. For example, the predicted electrophilicity of the α-keto amide **12** (LUMO = 0.223) was found to be lower than the **R3** LUMO (0.395). This difference may be due to the presence of an adjacent heteroaromatic system leading to extensive conjugation as well as intramolecular hydrogen bonding between the nitro group and the enolic OH forming a six-membered chelate. Similarly, in case of hits with α-heteroatom substituted ketones, a slightly higher predicted electrophilicity of **15** than **7** or **11** might be due to the presence of the electron-withdrawing *p*-chloro group, increasing the electrophilicity of the carbonyl carbon. However, a comparison of predicted site-specific reactivity of identified hits with FP-2 inhibition data suggests a lack of correlation of electrophilicity with the enzymatic inhibition. For instance, compound **14** with an azetidine-3 one had the highest predicted LUMO-based electrophilicity among the identified hits, however, it was only moderately active in the FP-2 inhibition assay (IC_{50} = 25.9 μM). By comparison, compounds with an α-hetero amide (compounds **1**, **4**) were predicted to have a negligible electrophilicity but showed a relatively higher experimental inhibition of FP-2 (IC_{50} = 1−5 μM). This observation emphasizes the obvious involvement of other factors in enzyme inhibition, such as the overall geometry and interaction patterns of the compounds in the FP-2 binding site.

Figure 7 shows predicted interaction profiles of representative compounds **1**, **2**, and **6** in FP-2 active site. The 4-oxo-phthalazine core in molecule **1** appears to protrude into the S1′ pocket with the carboxylic acid group interacting with the residues of an active site-forming oxyanion hole, such as the backbone −NH of Cys 42 and the amide side chain of Gln 36 (see Figure 7a). Also, Trp206 and Val152 of the S1′ pocket are involved in the van der Waals interactions with the pthalazine moiety wherein the indole −NH of Trp206 participates in hydrogen bonding with the carboxylate moiety of **1**. The α-ketoamide soft-electrophile exhibits H-bonding interaction with the backbone −NH of Gly82 and the backbone carbonyl of Asn173. The phenyl ring at the terminus of the amide linker can undergo hydrophobic interaction with the side chains of Ala175. The terminal 2,4-dimethyl phenyl moiety is involved in hydrophobic interactions with the side chains of Leu84 and Ile85 at the entrance of the S3 pocket as well as with the Phe236 buried deeper in the S2 pocket.

In case of compound **2**, the 5-methyl-isooxazole moiety was well placed in the S1′ pocket of FP-2, exhibiting hydrophobic interactions with the side chains of Tyr206 and Val152 (see Figure 7b). Also, the isoxazoyl moiety of **2** was involved in multiple H-bonding interactions with the side chain amide proton of Gln36 as well as the indole −NH of Trp206. The carbonyl group of the α-keto amide electrophile located over the catalytic Cys42 formed a critical H-bonding interaction with the backbone amine of Cys42, whereas the −NH of the α-keto amide electrophile acts as a HBD for the backbone carbonyl of Asn 173. In addition, the α-keto amide linker was involved in hydrophobic interactions with Gly82. The nitrogen of the tetrazole moiety acted as a HBA with the backbone amine of Gly83. The phenyl ring attached to the tetrazole moiety of **2** vectors into the critical S2 pocket, forming hydrophobic interactions with Leu172 at the entrance of the S2 pocket as well as with the Leu84 side chain lining the S3 pocket. The anilinic −NH

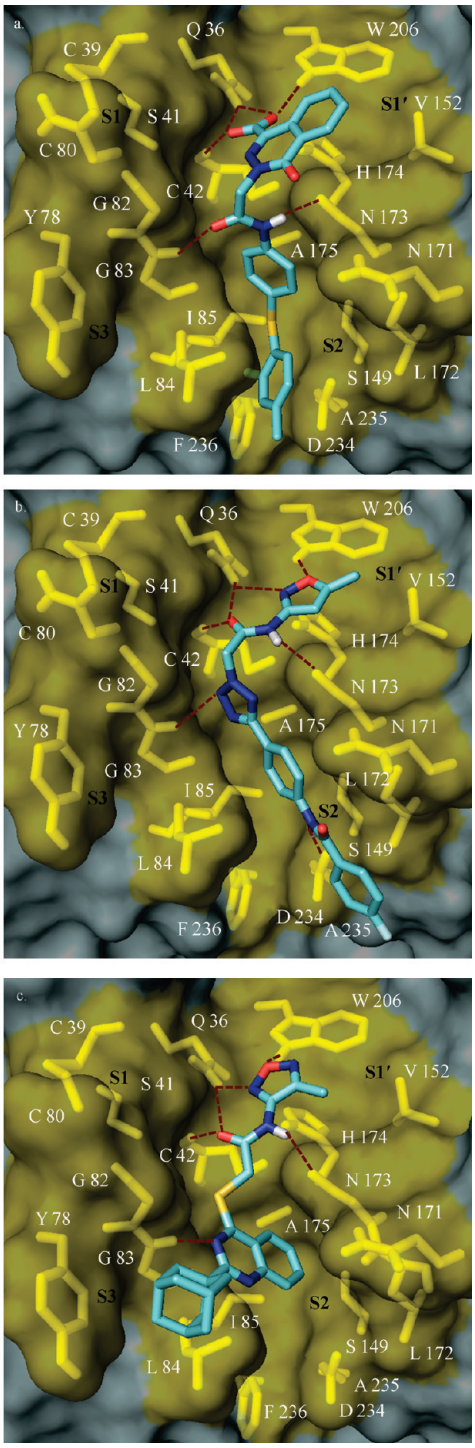


Figure 7. Predicted binding mode of virtual screening hits: (a) **1**, (b) **2**, and (c) **6** in the FP-2 binding site.

dx.doi.org/10.1021/ci200029y |*J. Chem. Inf. Model.* 2011, 51, 852–864

of **2** donated a H-bond to the side chain carboxylic group of Asp234 buried in the hydrophobic environment of the S2 pocket. The terminal phenyl ring was found to exhibit van der Waals interactions with the side chain of Leu172, exposing the paraf-luoro group to the solvent.

A similar observation was made for compound **6** near the S1′ pocket, especially for the 5-methyl-isooxazole moiety (see Figure 7c). The thioacetamide electrophile is placed over the catalytic cysteine sulfur, and its carboxamide group participates in the H-bonding interactions with Cys42 and Asn173 in a similar manner as observed for **2**. The quinazoline core of **6** is placed aptly in the S2 pocket, forming van der Waals interactions with Ala175 and Leu172. In addition, nitrogen on the quinazoline ring of **6** accepts a H-bond from the backbone of Gly83. Moreover, the admantyl moiety of **6** protrudes into the S3 pocket and participates in the van der Waals interactions with the Tyr78 of the S3 pocket.

The average pairwise Tanimoto similarity index of the 21 FP-2 inhibitors was 0.24, which points toward the diverse nature of the identified hits (see similarity matrix, Table S1 in Supporting Information). The diversity of hits obtained in the present study with previously identified FP-2 inhibitors was analyzed using the dbcmpr module of Sybyl; dbcmpr utility calculates the Tanimoto similarities using UNITY 2D fingerprints between all compounds identified in current study (test set) with their nearest neighbors in the previous study (reference set). The result of this comparison showed that 95% of compounds from the test set has the Tanimoto similarity of less than or equal to 0.5 to its nearest neighbor in the reference set (see database comparison, Figure S3 in the Supporting Information). The results clearly suggest novelty of the hits identified in the present study.

In addition to inhibition of FP-2, compounds **5**, **6**, **10**, **18**, **21**, and **22** exhibited in vitro antimalarial activity. Compounds **5** and **6** showed the same log order IC_{50} values in whole-cell assays as in the enzyme assay. Other compounds (**10**, **18**, **21**, and **22**) displayed greater activity against cultured parasites than FP-2. These observations might be explained in a number of ways: First, the compounds might be acting both against FP-2 and non-protease targets, as suggested by results with phenothiazines.[80] For example, compound **10** contains napthoquinone scaffold conferring its redox potential[81] and may exert its activity via protein cross-link in the cultured parasites in addition to inhibition of FP-2. Second, the compounds may be acting against targets other than FP-2. This possibility is suggested by the fact that compounds **18**, **21**, and **22** did not form a swollen, hemoglobin-filled food vacuole, as seen with highly potent covalent falcipain inhibitors.[30,82] Third, the compounds might concentrate in the food vacuole, allowing activity at culture concentrations below those that inhibit the protease. In any event, our screens identified a number of compounds with low micromolar antimalarial activity. Of note, all the compounds shown in Figure 5 were nontoxic in cytotoxicity assay, suggesting potential for development of related compounds as drugs.

Next, we screened identified FP-2 hits (compounds **1**−**7**, FP-2 $IC_{50} \leq 10\ \mu M$) against selected human papain family cysteine proteases (cathepsins B, K, and L), for possible off-target activity (Table 2). As expected, the evaluated hits (**1**−**3**, **5**−**7**) were found to inhibit selected cathepsins that are closely related to the parasitic cysteine proteases. Although, homology exists between the parasitic cysteine protease and its human orthologues, there are significant differences in the binding site residues, especially near the S2 pocket of these enzymes. For instance, sequence comparison (Figure S4, Supporting Information) suggested Ser149 in the S2 pocket of FPs is replaced by Ala in human orthologues. Ser149A mutation in FP-2 was reported to decrease the proteolytic activity by 91%[83] and can be targeted for design of selective protozoal cysteine protease inhibitors. In addition, Ile85 of the S2 pocket of FP-2 and FP-3 is replaced by methionine in cathepsin K and L and proline in cathepsin B. Moreover, the polar residues, such as Asp234 and Glu243 in the S2 pocket of the FP-2 and FP-3 binding sites, respectively, are replaced by hydrophobic residues, such as Leu and Ala in cathepsin K and L.[84] These differences between the proteases can be utilized for the design of selective inhibitors of FPs. To our surprise, compound **4** did not show any inhibition up to 50 μM against the evaluated mammalian cysteine proteases. Further docking studies of **4** in the selected mammalian cysteine protease panel might provide valuable insights for the binding mode and selectivity of this compound in FP-2.

Table 2. IC_{50} Values of FP-2 Hits against Selected Human Cathepsin Peptidases

comp no.	cathepsin K IC_{50} (μM)	cathepsin L IC_{50} (μM)	cathepsin B IC_{50} (μM)
1	20.00	9.30	10.72
2	3.83	24.39	37.14
3	25.60	41.57	14.74
4	>50	>50	>50
5	26.70	5.86	49.08
6	29.66	16.52	21.44
7	6.12	31.63	17.66
E-64	0.004	0.016	0.007

CONCLUSION

The focused cysteine protease inhibitor library consisting of ∼65 000 compounds was screened against the X-ray crystal structure of FP-2. The attractive components of the virtual screening approach were the FCPI library based on soft-electrophiles of interest, the enrichment study to select the best scoring function, the docking pose-based pharmacophore query as a filter in the SBVS protocol, and consideration of binding energy of the receptor−ligand complex for the selection of putative hits. The study reported a high success ratio with novel antimalarial scaffolds. Twenty one compounds were active against FP-2, and four were also active against FP-3. Two FP-2 hits, compound **5** and **6**, showed corresponding inhibition of the cultured parasites, showing promise for further structural optimization. Moreover, the predicted site-specific electrophilicity, as measured by calculating the LUMO electron density of each atom in the form of the atomic Fukui indices, suggested a soft-nature of the electrophiles present in identified hits. Thus, any potential risk of toxicity associated with the irreversible inhibition of the cysteine protease is attenuated. Assessment of the selectivity of identified hits against related mammalian cysteine proteases (cathepsins B, K, and L) revealed compound **4** as a selective inhibitor of FP-2 ($IC_{50} > 50\ \mu M$ against tested mammalian proteases). The hits obtained through the present screening efforts can be optimized via substructure/similarity search approaches to identify structural analogs of the active hits, taking into consideration the key residues of the S2 pocket for the biochemical selectivity as discussed above. In addition, combinatorial library synthesis for

selected hits (for example, compounds **2**, **4**, **6**, and **7**) using commercially available building blocks can be pursued taking insights from the recently published cocrystallized structure of FP-2,[23] for enhanced binding site occupancy and interaction profiles, and to establish an SAR.

■ ASSOCIATED CONTENT

Supporting Information. The chemical structures, biological activities of previously reported FP-2 hits used in the enrichment study, similarity matrix of obtained virtual screening hits, comparison of similarity of hits obtained in the present study with the previously identified hits, multiple sequence alignment of papain family cysteine proteases, and the calculations of the atomic Fukui indices. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: mavery@olemiss.edu; telephone: 662-915-5879.

Present Addresses

[||]Novartis Institute for Biomedical Research, Emeryville, CA 94608, United States.

■ ACKNOWLEDGMENT

This work was supported by CDC Cooperative Agreement UR3/CCU4186520-03. Authors are thankful to Claire Mischker in the Graduate Writing Center, University of Mississippi, MS, for her timely help in refining the language of the manuscript. Falgun Shah is thankful to Dr. Yakambram Pedduri for useful discussion. We thank Schrodinger technical support team especially Madhavi Sastry, Dale Braden, and Stuart Murdock for their technical inputs.

■ REFERENCES

(1) Mital, A. Recent advances in antimalarial compounds and their patents. *Curr. Med. Chem.* **2007**, *14*, 759–773.

(2) *World Malarial Report*; World Health Organization: Geneva, Switzerland; http://malaria.who.int/wmr2008/. Accessed May 1, 2010.

(3) Sachs, J.; Malaney, P. The economic and social burden of malaria. *Nature* **2002**, *415*, 680–685.

(4) Fidock, D. A. Drug discovery: Priming the antimalarial pipeline. *Nature* **2010**, *465*, 297–298.

(5) Nosten, F.; White, N. J. Artemisinin-based combination treatment of falciparum malaria. *Am. J. Trop. Med. Hyg.* **2007**, *77*, 181–192.

(6) Yeung, S.; Pongtavornpinyo, W.; Hastings, I. M.; Mills, A. J.; White, N. J. Antimalarial drug resistance, artemisinin-based combination therapy, and the contribution of modeling to elucidating policy choices. *Am. J. Trop. Med. Hyg.* **2004**, *71*, 179–186.

(7) Sinclair, D.; Zani, B.; Donegan, S.; Olliaro, P.; Garner, P. Artemisinin-based combination therapy for treating uncomplicated malaria. *Cochrane Database Syst. Rev.* **2009**, CD007483.

(8) Leung-Toung, R.; Zhao, Y.; Li, W.; Tam, T. F.; Karimian, K.; Spino, M. Thiol proteases: inhibitors and potential therapeutic targets. *Curr. Med. Chem.* **2006**, *13*, 547–581.

(9) Gauthier, J. Y.; Chauret, N.; Cromlish, W.; Desmarais, S.; Duong le, T.; Falgueyret, J. P.; Kimmel, D. B.; Lamontagne, S.; Leger, S.; LeRiche, T.; Li, C. S.; Masse, F.; McKay, D. J.; Nicoll-Griffith, D. A.; Oballa, R. M.; Palmer, J. T.; Percival, M. D.; Riendeau, D.; Robichaud, J.; Rodan, G. A.; Rodan, S. B.; Seto, C.; Therien, M.; Truong, V. L.; Venuti, M. C.; Wesolowski, G.; Young, R. N.; Zamboni, R.; Black, W. C. The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 923–928.

(10) Rosenthal, P. J.; Sijwali, P. S.; Singh, A.; Shenai, B. R. Cysteine proteases of malaria parasites: targets for chemotherapy. *Curr. Pharm. Des.* **2002**, *8*, 1659–1672.

(11) Sijwali, P. S.; Rosenthal, P. J. Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by Plasmodium falciparum. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 4384–4389.

(12) Sijwali, P. S.; Shenai, B. R.; Gut, J.; Singh, A.; Rosenthal, P. J. Expression and characterization of the Plasmodium falciparum haemoglobinase falcipain-3. *Biochem. J.* **2001**, *360*, 481–489.

(13) Shenai, B. R.; Sijwali, P. S.; Singh, A.; Rosenthal, P. J. Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of Plasmodium falciparum. *J. Biol. Chem.* **2000**, *275*, 29000–29010.

(14) Sijwali, P. S.; Koo, J.; Singh, N.; Rosenthal, P. J. Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of Plasmodium falciparum. *Mol. Biochem. Parasitol.* **2006**, *150*, 96–106.

(15) Greenbaum, D. C.; Baruch, A.; Grainger, M.; Bozdech, Z.; Medzihradszky, K. F.; Engel, J.; DeRisi, J.; Holder, A. A.; Bogyo, M. A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science* **2002**, *298*, 2002–2006.

(16) Sijwali, P. S.; Kato, K.; Seydel, K. B.; Gut, J.; Lehman, J.; Klemba, M.; Goldberg, D. E.; Miller, L. H.; Rosenthal, P. J. Plasmodium falciparum cysteine protease falcipain-1 is not essential in erythrocytic stage malaria parasites. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8721–8726.

(17) Eksi, S.; Czesny, B.; Greenbaum, D. C.; Bogyo, M.; Williamson, K. C. Targeted disruption of Plasmodium falciparum cysteine protease, falcipain 1, reduces oocyst production, not erythrocytic stage growth. *Mol. Microbiol.* **2004**, *53*, 243–250.

(18) Ettari, R.; Bova, F.; Zappala, M.; Grasso, S.; Micale, N. Falcipain-2 inhibitors. *Med. Res. Rev.* **2010**, *30*, 136–167.

(19) Shah, F.; Mukherjee, P.; Desai, P.; Avery, M. Computational approaches for the discovery of cysteine protease inhibitors against malaria and SARS. *Curr. Comput.-Aided Drug Des.* **2010**, *6*, 1–23.

(20) Verissimo, E.; Berry, N.; Gibbons, P.; Cristiano, M. L.; Rosenthal, P. J.; Gut, J.; Ward, S. A.; O'Neill, P. M. Design and synthesis of novel 2-pyridone peptidomimetic falcipain 2/3 inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4210–4214.

(21) Shenai, B. R.; Lee, B. J.; Alvarez-Hernandez, A.; Chong, P. Y.; Emal, C. D.; Neitz, R. J.; Roush, W. R.; Rosenthal, P. J. Structure-activity relationships for inhibition of cysteine protease activity and development of Plasmodium falciparum by peptidyl vinyl sulfones. *Antimicrob. Agents Chemother.* **2003**, *47*, 154–160.

(22) Rosenthal, P. J.; Wollish, W. S.; Palmer, J. T.; Rasnick, D. Antimalarial effects of peptide inhibitors of a Plasmodium falciparum cysteine proteinase. *J. Clin. Invest.* **1991**, *88*, 1467–1472.

(23) Kerr, I. D.; Lee, J. H.; Pandey, K. C.; Harrison, A.; Sajid, M.; Rosenthal, P. J.; Brinen, L. S. Structures of Falcipain-2 and Falcipain-3 Bound to Small Molecule Inhibitors: Implications for Substrate Specificity. *J. Med. Chem.* **2009**, *52*, 852–857.

(24) Rosenthal, P. J.; Lee, G. K.; Smith, R. E. Inhibition of a Plasmodium vinckei cysteine proteinase cures murine malaria. *J. Clin. Invest.* **1993**, *91*, 1052–1056.

(25) Lee, B. J.; Singh, A.; Chiang, P.; Kemp, S. J.; Goldman, E. A.; Weinhouse, M. I.; Vlasuk, G. P.; Rosenthal, P. J. Antimalarial activities of novel synthetic cysteine protease inhibitors. *Antimicrob. Agents Chemother.* **2003**, *47*, 3810–3814.

(26) Ando, R.; Sakaki, T.; Morinaka, Y.; Takahashi, C.; Tamao, Y.; Yoshii, N.; Katayama, S.; Saito, K.; Tokuyama, H.; Isaka, M.; Nakamura, E. Cyclopropenone-containing cysteine proteinase inhibitors. Synthesis and enzyme inhibitory activities. *Bioorg. Med. Chem.* **1999**, *7*, 571–579.

(27) Li, R.; Chen, X.; Gong, B.; Dominguez, J. N.; Davidson, E.; Kurzban, G.; Miller, R. E.; Nuzum, E. O.; Rosenthal, P. J.; et al. In Vitro

Antimalarial Activity of Chalcones and Their Derivatives. *J. Med. Chem.* **1995**, *38*, 5031–5037.

(28) Bazzaro, M.; Anchoori, R. K.; Mudiam, M. K. R.; Issaenko, O.; Kumar, S.; Karanam, B.; Lin, Z.; Isaksson Vogel, R.; Gavioli, R.; Destro, F. α,β-Unsaturated Carbonyl System of Chalcone-Based Derivatives Is Responsible for Broad Inhibition of Proteasomal Activity and Preferential Killing of Human Papilloma Virus (HPV) Positive Cervical Cancer Cells. *J. Med. Chem.* **2010**, 27–53.

(29) Oballa, R. M.; Truchon, J. F.; Bayly, C. I.; Chauret, N.; Day, S.; Crane, S.; Berthelette, C. A generally applicable method for assessing the electrophilicity and reactivity of diverse nitrile-containing compounds. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 998–1002.

(30) Rosenthal, P. J.; McKerrow, J. H.; Aikawa, M.; Nagasawa, H.; Leech, J. H. A malarial cysteine proteinase is necessary for hemoglobin degradation by Plasmodium falciparum. *J. Clin. Invest.* **1988**, *82*, 1560–1566.

(31) Palmer, J. T.; Rasnick, D.; Klaus, J. L.; Bromme, D. Vinyl Sulfones as Mechanism-Based Cysteine Protease Inhibitors. *J. Med. Chem.* **1995**, *38*, 3193–3196.

(32) Gamboa de Dominguez, N. D.; Rosenthal, P. J. Cysteine proteinase inhibitors block early steps in hemoglobin degradation by cultured malaria parasites. *Blood* **1996**, *87*, 4448–4454.

(33) Rosenthal, P. J.; Olson, J. E.; Lee, G. K.; Palmer, J. T.; Klaus, J. L.; Rasnick, D. Antimalarial effects of vinyl sulfone cysteine proteinase inhibitors. *Antimicrob. Agents Chemother.* **1996**, *40*, 1600–1603.

(34) Ettari, R.; Nizi, E.; Di Francesco, M. E.; Micale, N.; Grasso, S.; Zappala, M.; Vicik, R.; Schirmeister, T. Nonpeptidic vinyl and allyl phosphonates as falcipain-2 inhibitors. *ChemMedChem* **2008**, *3*, 1030–1033.

(35) Ramjee, M. K.; Flinn, N. S.; Pemberton, T. P.; Quibell, M.; Wang, Y.; Watts, J. P. Substrate mapping and inhibitor profiling of falcipain-2, falcipain-3 and berghepain-2: Implications for peptidase antimalarial drug discovery. *Biochem. J.* **2006**, *399*, 47–57.

(36) Desai, P. V.; Patny, A.; Sabnis, Y.; Tekwani, B.; Gut, J.; Rosenthal, P.; Srivastava, A.; Avery, M. Identification of Novel Parasitic Cysteine Protease Inhibitors Using Virtual Screening. 1. The ChemBridge Database. *J. Med. Chem.* **2004**, *47*, 6609–6615.

(37) Desai, P. V.; Patny, A.; Gut, J.; Rosenthal, P. J.; Tekwani, B.; Srivastava, A.; Avery, M. Identification of Novel Parasitic Cysteine Protease Inhibitors by Use of Virtual Screening. 2. The Available Chemical Directory. *J. Med. Chem.* **2006**, *49*, 1576–1584.

(38) Li, H.; Huang, J.; Chen, L.; Liu, X.; Chen, T.; Zhu, J.; Lu, W.; Shen, X.; Li, J.; Hilgenfeld, R.; Jiang, H. Identification of Novel Falcipain-2 Inhibitors as Potential Antimalarial Agents through Structure-Based Virtual Screening. *J. Med. Chem.* **2009**, *52*, 4936–4940.

(39) Oprea, T. I.; Matter, H. Integrating virtual screening in lead discovery. *Curr. Opin. Chem. Biol.* **2004**, *8*, 349–358.

(40) Bissantz, C.; Folkers, G.; Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, *43*, 4759–4767.

(41) Mukherjee, P.; Desai, P.; Zhou, Y. D.; Avery, M. Targeting the BH3 Domain Mediated Protein-Protein Interaction of Bcl-xL through Virtual Screening. *J. Chem. Inf. Model.* **2010**, *50*, 906–923.

(42) Vijayan, R. S. K.; Prabu, M.; Mascarenhas, N. M.; Ghoshal, N. Hybrid Structure-Based Virtual Screening Protocol for the Identification of Novel BACE1 Inhibitors. *J. Chem. Inf. Model.* **2009**, *49*, 647–657.

(43) Gamo, F. J.; Sanz, L. M.; Vidal, J.; de Cozar, C.; Alvarez, E.; Lavandera, J. L.; Vanderwall, D. E.; Green, D. V.; Kumar, V.; Hasan, S.; Brown, J. R.; Peishoff, C. E.; Cardon, L. R.; Garcia-Bustos, J. F. Thousands of chemical starting points for antimalarial lead identification. *Nature* **2010**, *465*, 305–310.

(44) Guiguemde, W. A.; Shelat, A. A.; Bouck, D.; Duffy, S.; Crowther, G. J.; Davis, P. H.; Smithson, D. C.; Connelly, M.; Clark, J.; Zhu, F.; Jimenez-Diaz, M. B.; Martinez, M. S.; Wilson, E. B.; Tripathi, A. K.; Gut, J.; Sharlow, E. R.; Bathurst, I.; El Mazouni, F.; Fowble, J. W.; Forquer, I.; McGinley, P. L.; Castro, S.; Angulo-Barturen, I.; Ferrer, S.; Rosenthal, P. J.; Derisi, J. L.; Sullivan, D. J.; Lazo, J. S.; Roos, D. S.; Riscoe, M. K.; Phillips, M. A.; Rathod, P. K.; Van Voorhis, W. C.; Avery, V. M.; Guy, R. K. Chemical genetics of Plasmodium falciparum. *Nature* **2010**, *465*, 311–315.

(45) Hogg, T.; Nagarajan, K.; Herzberg, S.; Chen, L.; Shen, X.; Jiang, H.; Wecke, M.; Blohmke, C.; Hilgenfeld, R.; Schmidt, C. L. Structural and functional characterization of Falcipain-2, a hemoglobinase from the malarial parasite Plasmodium falciparum. *J. Biol. Chem.* **2006**, *281*, 25425–25437.

(46) *LigPrep*, version 2.2; Schrodinger, LLC: New York, 2008.

(47) *Glide*, version 5.0; Schrodinger, LLC, New York, 2008.

(48) *QikProp*, version 3.1; Schrodinger, LLC, New York, 2008.

(49) *Phase*, version 3.0; Schrodinger, LLC, New York, 2008.

(50) *Sybyl*, version 8.1; Tripos, Inc., St. Louis, MO, 2008.

(51) *Jaguar*, version 7.6; Schrödinger, LLC, New York, 2009.

(52) *Pymol*, Schrodinger, LLC, New York, 2008.

(53) Brinen, L. S.; Hansell, E.; Cheng, J.; Roush, W. R.; McKerrow, J. H.; Fletterick, R. J. A target within the target: probing cruzain's P1′ site to define structural determinants for the Chagas' disease protease. *Structure* **2000**, *8*, 831–840.

(54) Holm, L.; Park, J. DaliLite workbench for protein structure comparison. *Bioinformatics* **2000**, *16*, 566.

(55) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789–6801.

(56) Subramanian, S.; Hardt, M.; Choe, Y.; Niles, R. K.; Johansen, E. B.; Legac, J.; Gut, J.; Kerr, I. D.; Craik, C. S.; Rosenthal, P. J. Hemoglobin cleavage site-specificity of the Plasmodium falciparum cysteine proteases falcipain-2 and falcipain-3. *PLoS One* **2009**, *4*, e5156.

(57) Lambros, C.; Vanderberg, J. P. Synchronization of Plasmodium falciparum erythrocytic stages in culture. *J. Parasitol.* **1979**, *65*, 418–420.

(58) Babich, H.; Borenfreund, E. Cytotoxicity of T-2 toxin and its metabolites determined with the neutral red cell viability assay. *Appl. Environ. Microbiol.* **1991**, *57*, 2101–2103.

(59) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

(60) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *Chem. Phys.* **1993**, *98*, 5648–5652.

(61) Bernhard Schlegel, H. Estimating the Hessian for gradient-type geometry optimizations. *Theor. Chim. Acta* **1984**, *66*, 333–340.

(62) Contreras, R. R.; Fuentealba, P.; Galván, M.; Pérez, P. A direct evaluation of regional Fukui functions in molecules. *Chem. Phys. Lett.* **1999**, *304*, 405–413.

(63) Bulat, F. A.; Chamorro, E.; Fuentealba, P.; Toro-Labbé, A. Condensation of frontier molecular orbital Fukui functions. *J. Phys. Chem. A* **2004**, *108*, 342–349.

(64) Jansen, J. M.; Martin, E. J. Target-biased scoring approaches and expert systems in structure-based virtual screening. *Curr. Opin. Chem. Biol.* **2004**, *8*, 359–364.

(65) *Glide FAQ;* http://www.schrodinger.com/supportfaq/18/5/#faq1_33. Accessed May 1, 2010).

(66) Nägler, D. K.; Ménard, R. Human cathepsin X: A novel cysteine protease of the papain family with a very short proregion and unique insertions1. *FEBS Lett.* **1998**, *434*, 135–139.

(67) Marquis, R. W.; Ru, Y.; Zeng, J.; Trout, R. E.; LoCastro, S. M.; Gribble, A. D.; Witherington, J.; Fenwick, A. E.; Garnier, B.; Tomaszek, T.; Tew, D.; Hemling, M. E.; Quinn, C. J.; Smith, W. W.; Zhao, B.; McQueney, M. S.; Janson, C. A.; D'Alessio, K.; Veber, D. F. Cyclic ketone inhibitors of the cysteine protease cathepsin K. *J. Med. Chem.* **2001**, *44*, 725–736.

(68) Setti, E. L.; Davis, D.; Chung, T.; McCarter, J. 3,4-disubstituted azetidinones as selective inhibitors of the cysteine protease cathepsin K. Exploring P2 elements for selectivity. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2051–2053.

(69) Kerr, I. D.; Lee, J. H.; Farady, C. J.; Marion, R.; Rickert, M.; Sajid, M.; Pandey, K. C.; Caffrey, C. R.; Legac, J.; Hansell, E. Vinyl sulfones as antiparasitic agents and a structural basis for drug design. *J. Biol. Chem.* **2009**, *284*, 25697.

(70) Brak, K.; Kerr, I. D.; Barrett, K. T.; Fuchi, N.; Debnath, M.; Ang, K.; Engel, J. C.; McKerrow, J. H.; Doyle, P. S.; Brinen, L. S. Nonpeptidic

tetrafluorophenoxymethyl ketone cruzain inhibitors as promising new leads for chagas disease chemotherapy. *J. Med. Chem.* 53, 1763−1773.

(71) Johnson, D. S.; Weerapana, E.; Cravatt, B. F. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. *Future Med. Chem.* **2010**, *2*, 949–964.

(72) Potashman, M. H.; Duggan, M. E. Covalent modifiers: an orthogonal approach to drug design. *J. Med. Chem.* **2009**, *52*, 1231–1246.

(73) Marquis, R. W.; Ru, Y.; Yamashita, D. S.; Oh, H. J.; Yen, J.; Thompson, S. K.; Carr, T. J.; Levy, M. A.; Tomaszek, T. A.; Ijames, C. F.; Smith, W. W.; Zhao, B.; Janson, C. A.; Abdel-Meguid, S. S.; D'Alessio, K. J.; McQueney, M. S.; Veber, D. F. Potent dipeptidylketone inhibitors of the cysteine protease cathepsin K. *Bioorg. Med. Chem.* **1999**, *7*, 581–588.

(74) LaLonde, J. M.; Zhao, B.; Smith, W. W.; Janson, C. A.; DesJarlais, R. L.; Tomaszek, T. A.; Carr, T. J.; Thompson, S. K.; Oh, H. J.; Yamashita, D. S.; Veber, D. F.; Abdel-Meguid, S. S. Use of papain as a model for the structure-based design of cathepsin K inhibitors: crystal structures of two papain-inhibitor complexes demonstrate binding to S'-subsites. *J. Med. Chem.* **1998**, *41*, 4567–4576.

(75) Steert, K.; Berg, M.; Mottram, J. C.; Westrop, G. D.; Coombs, G. H.; Cos, P.; Maes, L.; Joossens, J.; Van der Veken, P.; Haemers, A.; Augustyns, K. alpha-ketoheterocycles as inhibitors of Leishmania mexicana cysteine protease CPB. *ChemMedChem 5*, 1734−1748.

(76) Lee, B. J.; Singh, A.; Chiang, P.; Kemp, S. J.; Goldman, E. A.; Weinhouse, M. I.; Vlasuk, G. P.; Rosenthal, P. J. Antimalarial activities of novel synthetic cysteine protease inhibitors. *Antimicrob. Agents Chemother.* **2003**, *47*, 3810–3814.

(77) Li, Z.; Patil, G. S.; Golubski, Z. E.; Hori, H.; Tehrani, K.; Foreman, J. E.; Eveleth, D. D.; Bartus, R. T.; Powers, J. C. Peptide alpha-keto ester, alpha-keto amide, and alpha-keto acid inhibitors of calpains and other cysteine proteases. *J. Med. Chem.* **1993**, *36*, 3472–3480.

(78) *Jaguar user manual*, version 7.6; Schrödinger, LLC, New York, 2009.

(79) Rietjens, I.; Soffers, A.; Hooiveld, G.; Veeger, C.; Vervoort, J. Quantitative structure-activity relationships based on computer calculated parameters for the overall rate of glutathione S-transferase catalyzed conjugation of a series of fluoronitrobenzenes. *Chem. Res. Toxicol.* **1995**, *8*, 481–488.

(80) Dominguez, J. N.; Lopez, S.; Charris, J.; Iarruso, L.; Lobo, G.; Semenov, A.; Olson, J. E.; Rosenthal, P. J. Synthesis and antimalarial effects of phenothiazine inhibitors of a Plasmodium falciparum cysteine protease. *J. Med. Chem.* **1997**, *40*, 2726–2732.

(81) Pinto, A. V.; de Castro, S. L. The trypanocidal activity of naphthoquinones: a review. *Molecules* **2009**, *14*, 4570–4590.

(82) Coteron, J. M.; Catterick, D.; Castro, J.; Chaparro, M. J.; Diaz, B.; Fernandez, E.; Ferrer, S.; Gamo, F. J.; Gordo, M.; Gut, J.; de las Heras, L.; Legac, J.; Marco, M.; Miguel, J.; Munoz, V.; Porras, E.; de la Rosa, J. C.; Ruiz, J. R.; Sandoval, E.; Ventosa, P.; Rosenthal, P. J.; Fiandor, J. M. Falcipain inhibitors: optimization studies of the 2-pyrimidinecarbonitrile lead series. *J. Med. Chem.* **2010**, *53*, 6129–6152.

(83) Goh, L. L.; Sim, T. S. Homology modeling and mutagenesis analyses of Plasmodium falciparum falcipain 2A: implications for rational drug design. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 565–572.

(84) Sabnis, Y. A.; Desai, P. V.; Rosenthal, P. J.; Avery, M. A. Probing the structure of falcipain 3, a cysteine protease from Plasmodium falciparum: Comparative protein modeling and docking studies. *Protein Sci.* **2003**, *12*, 501–509.