JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Computer-Aided Drug Design of Falcipain Inhibitors: Virtual Screening, Structure−Activity Relationships, Hydration Site Thermodynamics, and Reactivity Analysis

Falgun Shah,[†] Jiri Gut,[‖] Jennifer Legac,[‖] Devleena Shivakumar,[⊥] Woody Sherman,[⊥] Philip J. Rosenthal,[‖] and Mitchell A. Avery*[,†,§]

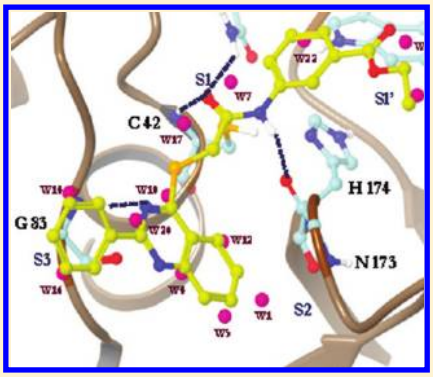
[†]Department of Medicinal Chemistry, School of Pharmacy, and [§]National Center for Natural Products Research, University of Mississippi, University, Mississippi 38677, United States

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

[⊥]Schrodinger, Inc., 120 West 45th Street, 17th Floor, New York, New York 10036, United States

Supporting Information

ABSTRACT: Falcipains (FPs) are hemoglobinases of *Plasmodium falciparum* that are validated targets for the development of antimalarial chemotherapy. A combined ligand- and structure-based virtual screening of commercial databases was performed to identify structural analogs of virtual screening hits previously discovered in our laboratory. A total of 28 low micromolar inhibitors of FP-2 and FP-3 were identified and the structure−activity relationship (SAR) in each series was elaborated. The SAR of the compounds was unusually steep in some cases and could not be explained by a traditional analysis of the ligand−protein interactions (van der Waals, electrostatics, and hydrogen bonds). To gain further insights, a statistical thermodynamic analysis of explicit solvent in the ligand binding domains of FP-2 and FP-3 was carried out to understand the roles played by water molecules in binding of these inhibitors. Indeed, the energetics associated with the displacement of water molecules upon ligand binding explained some of the complex trends in the SAR. Furthermore, low potency of a subset of FP-2 inhibitors that could not be understood by the water energetics was explained in the context of poor chemical reactivity of the reactive centers of these compounds. The present study highlights the importance of considering energetic contributors to binding beyond traditional ligand−protein interactions.



INTRODUCTION

Malaria remains a leading cause of morbidity and mortality around the globe.[1] This disease is caused by protozoa of the genus *Plasmodium*, with *P. falciparum* being the most lethal and widespread species. Recent estimates state that malaria accounts for hundreds of millions of illnesses and approximately 1 million deaths annually.[2] The control of this disease is complicated by emerging resistance to existing antimalarial compounds,[3] absence of a novel chemical class of antimalarials,[4] and difficulties in development of clinically approved vaccines.[5] In addition, early indications of decreasing clinical efficacy of front-line artemisinin combination therapies (ACTs) for the treatment of falciparum malaria are of significant concern.[6] These factors exemplify the need to discover new, effective, and affordable antimalarial chemotherapy acting through novel mechanisms.

We have devoted extensive efforts[7−13] to search for novel antimalarials by targeting falcipain (FP) cysteine proteases of *P. falciparum*. The best characterized of these papain-like cysteine proteases, FP-2 and FP-3, are responsible for hemoglobin degradation during the erythrocytic life-cycle of *P. falciparum*, providing essential nutrients for parasite growth.[14,15] Inhibitors of FP-2 and FP-3 are thus of interest as next generation antimalarials.[16]

Numerous studies have shown that FP inhibitors block the development of cultured *P. falciparum* parasites.[16−21] In particular, inhibition of FPs by covalent irreversible cysteine protease inhibitors blocked hemoglobin hydrolysis and caused the formation of characteristic swollen, dark-staining food vacuoles due to accumulation of unhydrolysed globin.[18,22] These morphological abnormalities are typical characteristics of cysteine protease inhibition not elicited by inhibitors of other food vacuole proteases such as aspartic proteases.[23] Although they are potent, the selectivity of irreversible covalent cysteine protease inhibitors between parasite and homologous mammalian proteases may limit their development as drug candidates. Hence, several efforts have been made to discover nonpeptidic, noncovalent inhibitors of FPs mainly by structure-guided drug design approaches.[7,8,13,24−27] Readers are referred to more comprehensive reviews to get a broader sense of the different classes of FP inhibitors.[28,29]

In our efforts toward the development of novel antimalarials,[30−37] we previously carried out structure-based virtual screening of a focused cysteine protease inhibitor library built

Received: November 19, 2011
Published: February 14, 2012

ACS Publications © 2012 American Chemical Society

on soft-electrophiles of interest.[11] We reported 21 diverse low-micromolar nonpeptidic hits against FP-2, originating from novel scaffolds. Compound **2**, with a tetrazole scaffold, was identified as a selective inhibitor of FP-2. In the first part of our current study, we describe efforts to explore the chemical space around the core group of **1**−**3** with low-micromolar affinity to FP-2 (Figure 1). A combined ligand- and structure-based

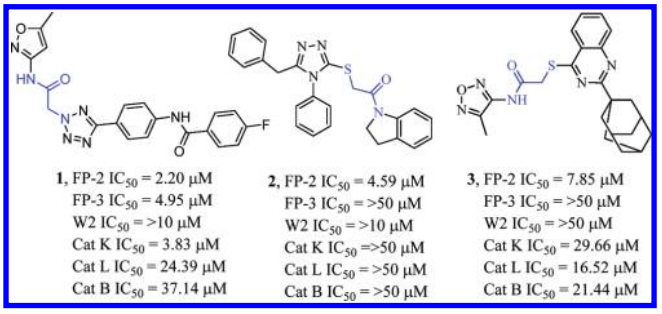


Figure 1. FP-2 inhibitors identified by structure-based virtual screening of the Focused Cysteine Protease Inhibitor library (IC_{50} < 10 μM). W2 stands for the chloroquine resistant strain of *P. falciparum* parasites. Cathepsin (Cat) K, L, and B are human cysteine proteases of the papain family. The α-heteroacetamide soft-electrophile is highlighted in blue.

virtual screening approach was adopted that involved mining of compounds bearing substructures of **1**−**3** from the Asinex and Chembridge databases, followed by docking of putative hits into the active site of FP-2. A number of active analogs were identified, some of which also showed activity against FP-3 and against development of cultured *P. falciparum*. Most of the active analogs showed the same order of potency as the parent compounds. However, the resulting SAR in these congeneric series of compounds was complex and could not be deciphered by only considering the steric, electrostatic, or van der Waals interactions of inhibitors in the protein active site. Therefore, the second part of this work considers additional factors involved in the binding process, such as the displacement of explicit water molecules and changes in the reactivity of electrophiles based on different chemical substitutions.

The solvent thermodynamics of the FP-2 and FP-3 ligand binding domain were computed using the recently published method referred to as WaterMap.[38,39] WaterMap predicts the location and thermodynamic properties of the active site solvent using a combination of explicit solvent molecular dynamics, spatial clustering, and inhomogeneous solvation theory.[40−42] The result of this procedure is a map containing water locations, occupancy, and thermodynamic characteristics (entropy, enthalpy, and free energy) that can be used to determine whether the water clusters (referred to as "hydration sites") are favorable or unfavorable in the binding site relative to bulk solvent. WaterMap has been successfully applied to a number of targets in a variety of contexts, such as assessing binding affinity of potent inhibitors,[43,44] understanding the SAR in congeneric series of compounds,[45,46] and investigating the selectivity of compounds within gene family targets such as kinases.[47] Recently, the method was used to rationalize accurate isothermal titration calorimetry data showing that the addition of lipophilic groups to a ligand resulted in increased potency that was gained via enthalpy from released water molecules, rather than the more traditional view that entropy gain from released water is the only driver of the hydrophobic effect.[48]

The water displacement patterns of FP-2 and FP-3 inhibitors aided in our understanding of key experimental observations of known FP-2 inhibitors and perplexing trends in the SAR of identified hits. However, the relatively modest affinity of some inhibitors (also displacing unstable hydration sites from FP-2) was not evident from the WaterMap analysis alone. We hypothesized that poor reactivity of soft-electrophiles could be responsible for the low affinity of some inhibitors toward FP-2. To further support our hypothesis, virtual screening of a small compound library enriched in diverse aryl and aliphatic nitriles with a poor chemical reactivity profile (based on the reactivity index of nitrile-containing compounds proposed by Oballa et al.[49]) was carried out against FP-2. The biological results showed good agreement between chemical reactivity and the inhibitory potency of these compounds against FP-2.

METHODS

Virtual Screening. The protocol used for virtual screening is shown in Figure 2. The substructure search on the

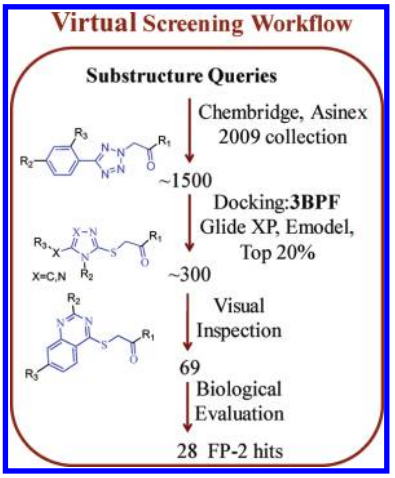


Figure 2. Virtual screening workflow to find hits related to compounds **1**−**3**. Substructure queries for **1**−**3** are shown in blue.

highlighted query atoms for compounds **1**−**3** was conducted in the Chembridge (http://www.hit2lead.com) and Asinex (http://www.asinex.com) 2009 collections using the built-in substructure search utilities of these databases. The resulting hits were prepared for docking using the LigPrep program (Schrodinger LLC). Approximately 1500 compounds matching subgraphs of **1**−**3** were subjected to docking using our previously validated Glide XP protocol.[11] The top 20% (~300) compounds were selected based on the top Emodel score. Docking poses were visually inspected for the following criteria prior to submitting them for biological evaluation: (a) placement of a soft-electrophile within 3.5 Å of the sulfur atom of the catalytic cysteine; (b) formation of hydrogen bonds by ligand atoms with the key residues of the FP-2 active site such as Gly83, Cys42, and Gln36; (c) energetically reasonable ligand geometry without obvious strain; and (d) commercial availability and diversity of putative hits. On the basis of these criteria, 69 analogs with diverse R_1−R_3 groups were purchased for analysis.

An additional substructure search was performed for compounds with aliphatic or aromatic nitriles, followed by a similar virtual screening protocol to that shown in Figure 2. Compounds with a match to the query "CC#N" (i.e., carbon connected to carbon connected to nitrogen via a triple bond)

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

Table 1. Biological Evaluation of Analogs of Virtual Screening Hit 1 with the 1,2,3,4-Tetrazole Core[a]

compound code	FP-2 IC_{50} (μM)	FP-3 IC_{50} (μM)	W2 IC_{50} (μM)	compound code	FP-2 IC_{50} (μM)	FP-3 IC_{50} (μM)	W2 IC_{50} (μM)
4	6.67	19.43	>10	**11**	38.13	>50	>10
5	11.35	43.89	>10	**12**	>50	>50	>10
6	19.15	>50	>10	**13**	>50	>50	>10
7	31.46	>50	>10	**14**	>50	>50	>10
8	33.04	>50	>10	**E-64**	0.03	0.16	-
9	35.92	>50	>10	artemisinin	-	-	0.092
10	37.81	>50	>10				

[a]W2 = chloroquine-resistant strain of *P. falciparum*.

were retrieved from the Chembridge database (2009 collection). The matching set of ~5000 compounds was then docked into the FP-2 binding site. The top 10% (~500) of compounds were selected based on Emodel score. Finally, a total of 14 compounds with aliphatic or aromatic nitriles were prioritized for biological evaluation based on two criteria (a) the compounds should make similar interactions in the binding site as predicted for pyrimidine nitrile compounds, a potent class of FP inhibitors,[50] and (b) they should displace the key unstable hydration sites from the FP-2 binding site.

Biological Evaluation. The prioritized hits were evaluated against FP-2, FP-3, and cultured W2 strain (chloroquine-resistant) *P. falciparum*. A selected set of compounds was also evaluated against mammalian cysteine proteases cathepsin K (Cat K), cathepsin B (Cat B), and cathepsin L1(Cat L1). These biological assays were conducted using the same protocols as described previously.[11] The epoxy succinate inhibitor (E-64) was used as a positive control for FP and Cathepsin B, K, and L1 assay whereas the fast acting antimalarial drug artemisinin was used as a standard for cultured parasites assays.

Thermodynamic Characterization of Water Molecules in the FP Binding Site. The program WaterMap (Schrodinger, LLC) was used for all water analysis work. WaterMap predicts the location of water molecules around a protein and provides a quantitative estimation of enthalpy (ΔH), entropy ($-T\Delta S$), and free energy (ΔG) associated with the water molecules.[38,39] The approach combines explicit solvent molecular dynamics, solvent clustering, and statistical thermodynamics to compute properties for clusters of waters (hydration sites) around a protein. In short, a grand canonical Monte Carlo simulation is run to obtain the initial placement of the waters, followed by a series of equilibration stages and a 2.0 ns production run of explicit solvent molecular dynamics. The protein is constrained in the simulations, allowing for convergence in the sampling of water configurations around the protein conformation of interest. The waters from approximately 2000 equally spaced frames are clustered to determine regions of high solvent density, (i.e., "hydration sites").[38] The entropy for each hydration site is computed using inhomogeneous solvation theory.[41,42] The enthalpy of each hydration site is calculated using the nonbonded molecular mechanics interaction energy with the OPLS_2005 force field.[51] These energy terms are calculated relative to a hydration site in bulk water.

WaterMap profiles were generated for FP-2 in complex with an epoxy succinate (E-64) inhibitor (3BPF, PDB code)[52] and for FP-3 in complex with a vinyl sulfone (VS) inhibitor, k11017 (3BWK, PDB code).[53] Proteins were prepared using the Protein Preparation Wizard in Maestro (Schrodinger, LLC). The protonation states of the catalytic His174 in FP-2 and His183 in FP-3 were adjusted to HIE (histidine with hydrogen on the epsilon nitrogen). Finally, the proteins were minimized with a harmonic potential restraint as implemented in the Impref module from Impact to prepare the final systems for the WaterMap calculations. WaterMap calculations were performed with default parameters and all waters within 5 Å of the cocrystallized inhibitor were considered for the analysis.

RESULT AND DISCUSSION

The work presented here was initiated to establish structure–activity relationships (SAR) around the active core of compounds **1**–**3** (Figure 1). The combined substructure search and rigid-receptor docking identified 28 active analogs of **1**–**3**. The activity of these compounds against FP-2, FP-3, and cultured parasites is shown in Tables 1–3. The overlay of the cocrystallize epoxy succinate inhibitor (E-64) with representative hits from all three series to validate our structure-based drug design approach are shown in the Supporting Information (Figure S3). The SAR in each series is discussed below.

SAR in the 1,2,3,4-Tetrazole Series. In the tetrazole series, linear compounds showed greater inhibitory activity against FP-2 as compared to the branched compounds. For example, compounds with R_3 substitutions showed either a decrease (**8**, **9**, and **11**) in activity or no activity (**12**), whereas compounds with R_2 substitutions (**4**–**6**) showed moderate inhibition of FP-2 (Figure 3 and Table 1). In addition, compounds lacking R_2 substitutions (**13**) or containing smaller R_2 groups (**14**) were inactive. This trend was consistent with the SAR of compounds in this series recently reported by us using the combinatorial library synthesis approach.[12] Compounds **1**, **4**, and **9**, with heteroaromatic moieties at R_1, seem to be well tolerated in the S1–S1′ pocket, possibly because of additional interactions with residues of an oxyanion hole, such as Gln36 and Cys 42 in FP-2 (Figure 9a–d). Although linear compounds such as **1** and **4** showed decent potency against both falcipains, they also appeared to inhibit the mammalian cysteine proteases cathepsin B, K, and L1 and were nonselective (Table 4). In addition, none of the compounds in this series showed activity against cultured parasites at concentrations up to 10 μM (Table 1).

SAR of the 1,2,4-Triazole Series. Substituents at R_1 and R_2 had greater impact on the activity in the 1,2,4-triazole series. For example, compounds lacking R_2 substituents (**27**–**34**) or with shorter R_2 substituents (**35**–**38**) were inactive (Supporting Information Figure S1). The majority of the active analogs had an unsubstituted phenyl at the R_2 position. Compounds **43**–**48**, with substitutions on the R_2-phenyl moiety, were not tolerated either (Figure S1). At the R_1 position, the amino-substituted compounds (forming thioacetamide) were well tolerated (Figure 4 and Table 2). Analogs with the fused nitrogen heterocycles (**40**–**45**) or lacking polar –NH (**49**, **50**) at that position were inactive. This could be due to lack of

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696–710

Table 2. Biological Evaluation of Analogs of Virtual Screening Hit 2 with the 1,2,4-Triazole Core[a]

compound code	FP-2 IC_{50} (μM)	FP-3 IC_{50} (μM)	W2 IC_{50} (μM)	compound code	FP-2 IC_{50} (μM)	FP-3 IC_{50} (μM)	W2 IC_{50} (μM)
15	15.32	>50	1.61	**34**	>50	NT	NT
16	18.61	>50	5.91	**35**	>50	NT	NT
17	20.66	>50	NT	**36**	>50	NT	NT
18	23.21	>50	6.42	**37**	>50	NT	NT
19	26.45	>50	1.91	**38**	>50	NT	>10
20	27.43	>50	4.65	**39**	>50	NT	>10
21	28.61	31.33	>10	**40**	>50	NT	NT
22	29.40	>50	8.91	**41**	>50	NT	NT
23	36.95	>50	>10	**42**	>50	NT	NT
24	47.22	>50	6.29	**43**	>50	NT	NT
25	48.01	>50	4.34	**44**	>50	NT	NT
26	48.78	>50	>10	**45**	>50	NT	NT
27	>50	NT	NT	**46**	>50	NT	NT
28	>50	NT	NT	**47**	>50	NT	NT
29	>50	NT	NT	**48**	>50	NT	>10
30	>50	NT	NT	**49**	>50	NT	NT
31	>50	NT	NT	**50**	>50	NT	NT
32	>50	NT	NT	**E-64**	0.03	0.16	-
33	>50	NT	NT	artemisinin	-	-	0.018

[a]NT = not tested, W2 = chloroquine-resistant strain of *P. falciparum*.

Table 3. Biological Evaluation of Analogs of Virtual Screening Hit 3 with the Quinazoline Core[a]

compound code	FP-2 IC_{50} (μM)	FP-3 IC_{50} (μM)	W2 IC_{50} (μM)	compound code	FP-2 IC_{50} (μM)	FP-3 IC_{50} (μM)	W2 IC_{50} (μM)
52	4.64	>50	>10	**62**	>50	>50	4.14
53	5.93	>50	>10	**63**	>50	>50	8.00
54	8.57	>50	6.04	**64**	>50	>50	>10
55	9.19	>50	>10	**65**	>50	>50	>10
56	14.00	49.50	6.15	**66**	>50	>50	>10
57	15.81	>50	4.49	**67**	>50	>50	>10
58	18.82	>50	7.59	**68**	>50	>50	>10
59	>50	>50	>10	**69**	>50	>50	>10
60	>50	>50	>10	**E-64**	0.03	0.16	-
61	>50	>50	4.24	artemisinin	-	-	0.092

[a]W2 = chloroquine-resistant strain of *P. falciparum*.

ability of these compounds to form an H-bond with Asn173 (Figure 10c).

Substitution at R_3 seemed to be crucial for modulating selectivity against FP-2 for this series of compounds. From the multiple sequence alignment of falcipains and other homologous mammalian cysteine proteases of the papain family (cathepsin B, K, and L1), it was evident that the composition of the S2 and S3 pocket differs critically in these cysteine proteases.[11,12] More importantly, the gate-keeper Leu at the entrance of the S3 pocket in FP-2 and in cathepsin L1 is substituted by the bulkier Tyr in FP-3, Cat K, and Cat B. To further understand the selectivity of **2** against FP-2, preliminary docking studies were carried out in FP-2 and the cathepsins (Figure 5). As anticipated, the R_3 methyl phenyl group of **2** was aptly placed in the S3 pocket of FP-2 and Cat L, whereas this moiety was directed into the S2 pocket in FP-3, Cat K, and Cat B, leading to unfavorable placement of **2** (Figure 5). This might explain the exclusive selectivity of **2** for FP-2.

To further analyze the effects of R_3 substituents on FP-3 selectivity, we evaluated selected active compounds from this series against FP-3 and cathepsins B, K, and L1 (Figure 6). As expected, compounds with smaller R_3 substituents (e.g., pyridine in compound **21**) tended to be nonselective, showing activity against all the cysteine proteases (except Cat B). As the length of R_3 substituents increased, activity against FP-3, Cat K, and Cat B decreased, probably due to steric clash with Tyrosine at the entrance of the S3 pocket (for example, compounds **17**, **19**, and **2**). Interestingly, compounds with more flexible R_3 substituents (**16** and **18**) showed activity against FP-2 as well as cathepsins, probably due to the abilities of flexible R_3 substituents to enter into the S3 pocket. Most of the active compounds from this series were also active against cultured parasites, with an IC_{50} of <10 μM (Table 2).

SAR in Quinazoline Series. In the quinazoline series, activity was retained when aliphatic (*t*-butyl, isopropyl) or aryl (phenyl) substituents at the R_3 position were directly attached to the quinazoline core (e.g., compounds **52**−**54**). Alicyclic substitutions such as alkyl cyclopentane resulted in diminished activity (for example, compound **59**) or no activity (compounds **60** and **61**, Figure 7, Table 4) against FP-2. Substitutions on the quinazoline core at the R_2 position resulted in decreased activity (compound **58**) or inactive analogs (compound **62**), probably due to the steric clash in the narrow S2 pocket of FP-2. At the R_1 position, there was preference for ampiphilic aryl groups such as 1,2,5-oxadiazole (**3**), thiazole (**52**, **55**), 1,3,4-thiadiazole (**53**), isoxzaole (**56**, **59**), and furan (**57**). Also, the heterocyclic moiety with three heteroatoms (**3**, **53**) at R_1 was more potent than a two (**52**, **55**, **56**) or single (**57**) heteroatom containing ring.

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

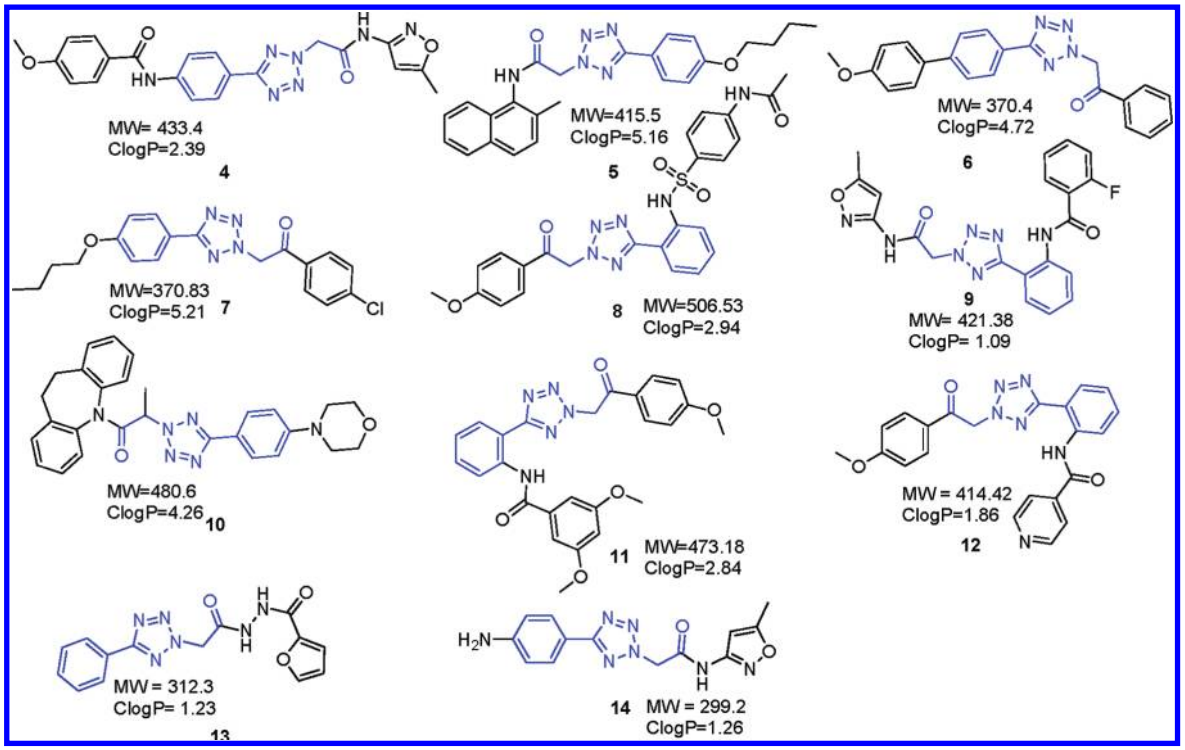


Figure 3. Chemical structures of compounds with 1,2,3,4-tetrazole core. The original subgraph is highlighted in blue.

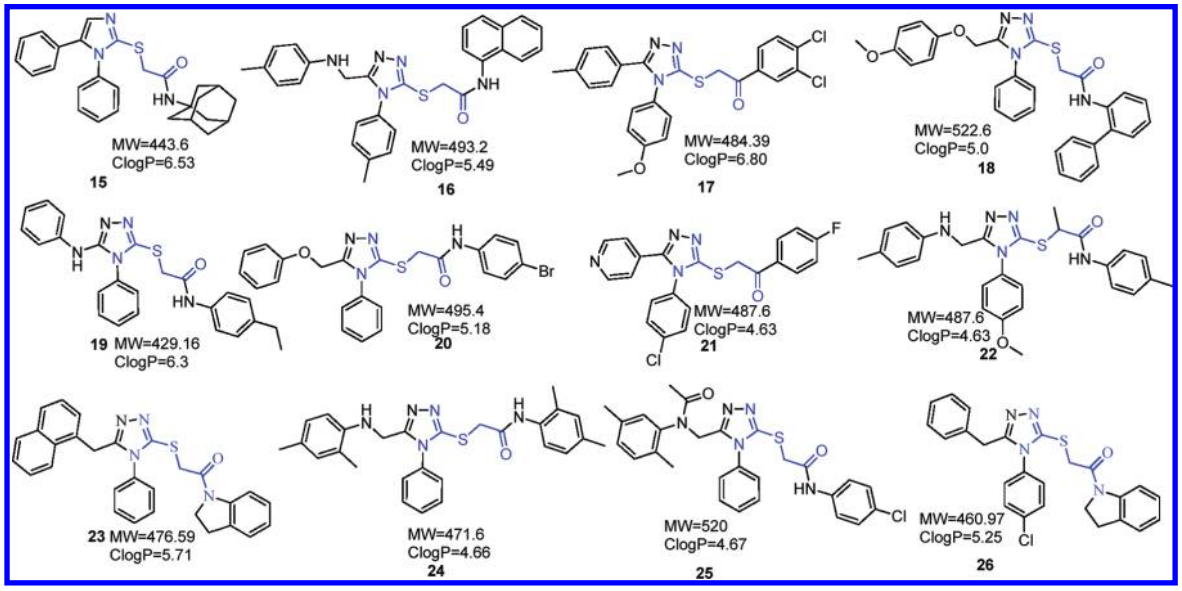


Figure 4. Chemical structures of compounds with the 1,2,4-triazole core. The original subgraph is highlighted in blue.

Substitutions at R_1 either with only hydrophobic groups, such as compounds **61** and **62** or polar groups at R_1 such as compound **64** and **65**, were found to be inactive. From the docking studies, it was suggested that these heterocycles exhibit hydrogen bonding interactions with the side chains of Gln 36 and Trp 206 in addition to hydrophobic interactions with Tyr206 (Figure 11a). The inactivity of **66** versus **54** might be explained by unfavorable placement of carboxylic acid in the S1′ pocket, which is composed predominantly of hydrophobic residues, whereas the inactivity of **67**−**69** is not clearly understood based on visual inspection of the interactions.

Among the active analogs, compound **54** showed an IC_{50} value of 8.5 μM against FP-2 and was also active against cultured parasites (W2 IC_{50} = 6 μM). Compound **54** moderately inhibited Cat K (IC_{50} = 39.2 μM) and Cat L (IC_{50} = 14.6 μM) and was inactive against Cat B at concentration up to 50 μM. In addition, as reported previously by our laboratory,[54] compound **54** was a low-micromolar inhibitor of severe acute respiratory syndrome (SARS) $3CL^{pro}$ (IC_{50} of 5.8 μM), which is the main protease of SARS, playing an important role in viral replication.[55] To the best of our knowledge, **54** is the first common inhibitor of malarial (protozoal) and SARS (viral) cysteine proteases reported to date. The discovery of **54** holds merit as these two cysteine proteases are quite different from each other, in terms of their protease family (FP-2 is a member of the papain-like family

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

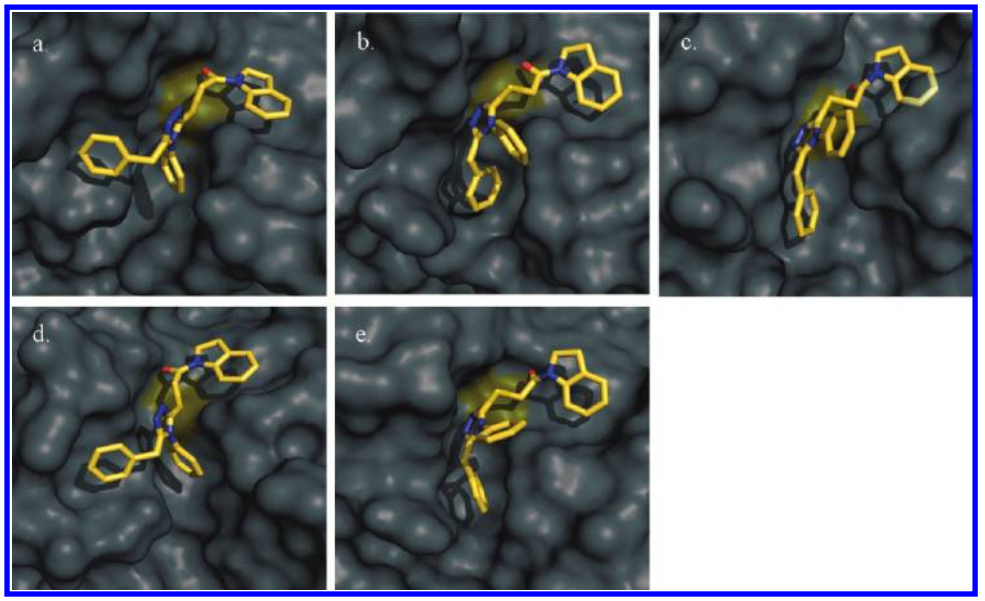


Figure 5. Proposed binding mode of **2** in (a) FP-2; (b) FP-3; (c) Cat K; (d) Cat L1; and (e) Cat B.

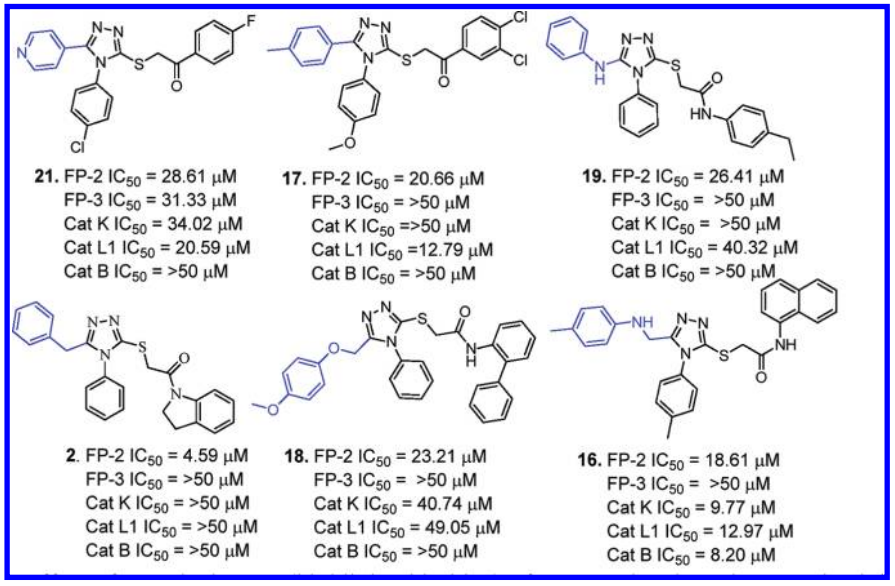


Figure 6. Effect of R_3 substituents (highlighted in blue) of 1,2,4-triazole series on selectivity to FPs and homologous mammalian cysteine proteases of the papain family: Cat K, Cat L1, and Cat B.

whereas SARS-3CL^{pro} is a chymotrypsin-like cysteine protease), and in their active site geometry.

In addition, we also correlated ClogP and FP-2 activity of the identified hits to consider the effects of lipophilicity of compounds in SAR. We found that active compounds (with selected threshold of IC_{50} < 30 μM) in all three series had an average ClogP of 5.0 whereas inactives (compounds with IC_{50} >50 μM) had an average ClogP value of 3.5 (See Supporting Information Table S3 for ClogP values of compounds **1**−**69**). These findings emphasize the preference for lipophilic compounds for apolar binding site of FP-2, although the computed ClogP values were insufficient to explain the details of SAR.

It is worthwhile to mention that compound series presented here can act on multiple targets in the malaria parasites. For example, compounds **24** and **25** from the 1,2,4-triazole series showed negligible potency against FP-2; however, these compounds exhibited single-digit micromolar potency against cultured parasites. Similarly, compounds **61**−**63** with a quinazoline core were inactive against FP-2 up to 50 μM test concentration; however, they showed moderate affinity (<10 μM) against malaria parasites. These findings emphasize the involvement of other targets in addition to FP-2 for the mode of action of these series of compounds.

Several trends in the SAR cannot be understood from the molecular docking studies with FP-2. For example, the experimental SAR of identified hits was very sensitive for compounds with R_2 substituents pointing toward the S2 pocket of FP-2, although no specific interactions in this pocket suggested such a significant change in potency. Also, there was a general preference for ampiphilic substituents in the S1−S1′ pocket that could not be explained based on molecular interactions. Moreover, most of the identified analogs showed the same log order of activity, in spite of enhanced interaction

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

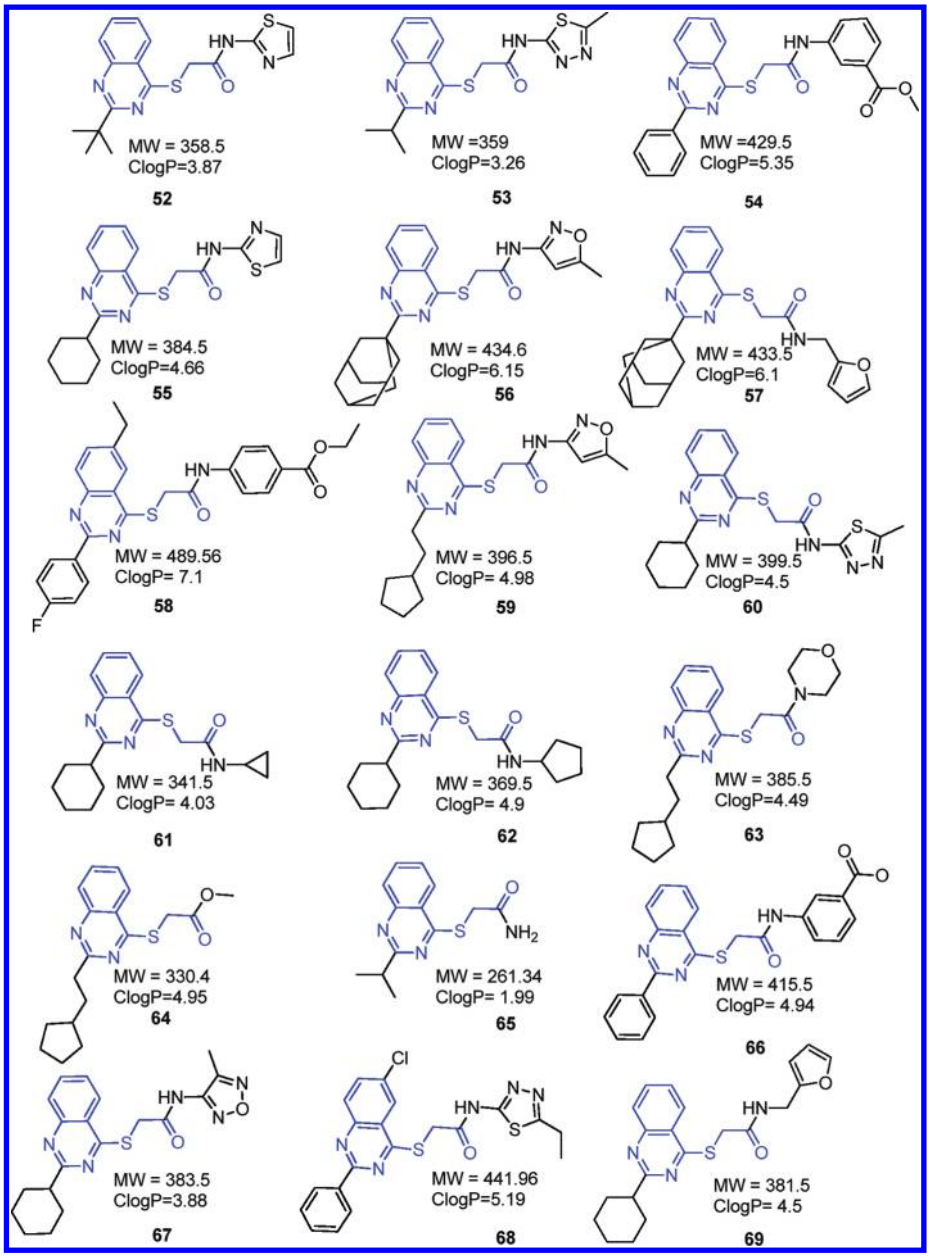


Figure 7. Chemical structures of compounds from the quinazoline core. The common substructure is highlighted in blue.

Table 4. Evaluation of Selected FP-2 Hits against the Mammalian Cysteine Proteases, Cathepsins K, L1, and B

compound no.	Cat K IC_{50} (μM)	Cat L1 IC_{50} (μM)	Cat B IC_{50} (μM)
4	4.22	3.12	10.07
17	>50	12.79	>50
18	40.74	49.05	>50
19	>50	40.32	>50
21	34.02	20.59	>50
54	39.21	14.55	>50
E-64	3.99	16.15	5.01

profiles, compared to the parent compounds **1**−**3** in the ligand binding domain of FPs. The steric and electrostatic interactions alone in addition to other terms of the Emodel scoring function (e.g., van der Waals, H-bond, lipophilic, ligand strain) were not sufficient to understand the steep SAR or moderate affinities of these compounds. Therefore, we anticipated a potentially important role of explicit solvent in the binding site and/or reactivity of the electrophiles in these compounds, both of which are generally not estimated by typical scoring functions.

Water molecules in protein binding sites have unfavorable entropy compared with those in bulk solvent due to their restrained movement in the active site, with the amount of entropy loss related to the degree of translational and rotational restriction.[38,39] They may have favorable or unfavorable enthalpy relative to bulk water, depending on the exact nature of the binding pocket and interaction with neighboring water molecules. Therefore, the thermodynamic properties of water molecules cannot be estimated by simple visual inspection. Accurately assessing the thermodynamics of water molecules is important for drug design, since displacement of high-energy water molecules from the protein binding site by complementary groups of ligands is considered to be a principle source of binding affinity.[43] A number of methods can account for solvation effects in the protein binding site, although many

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

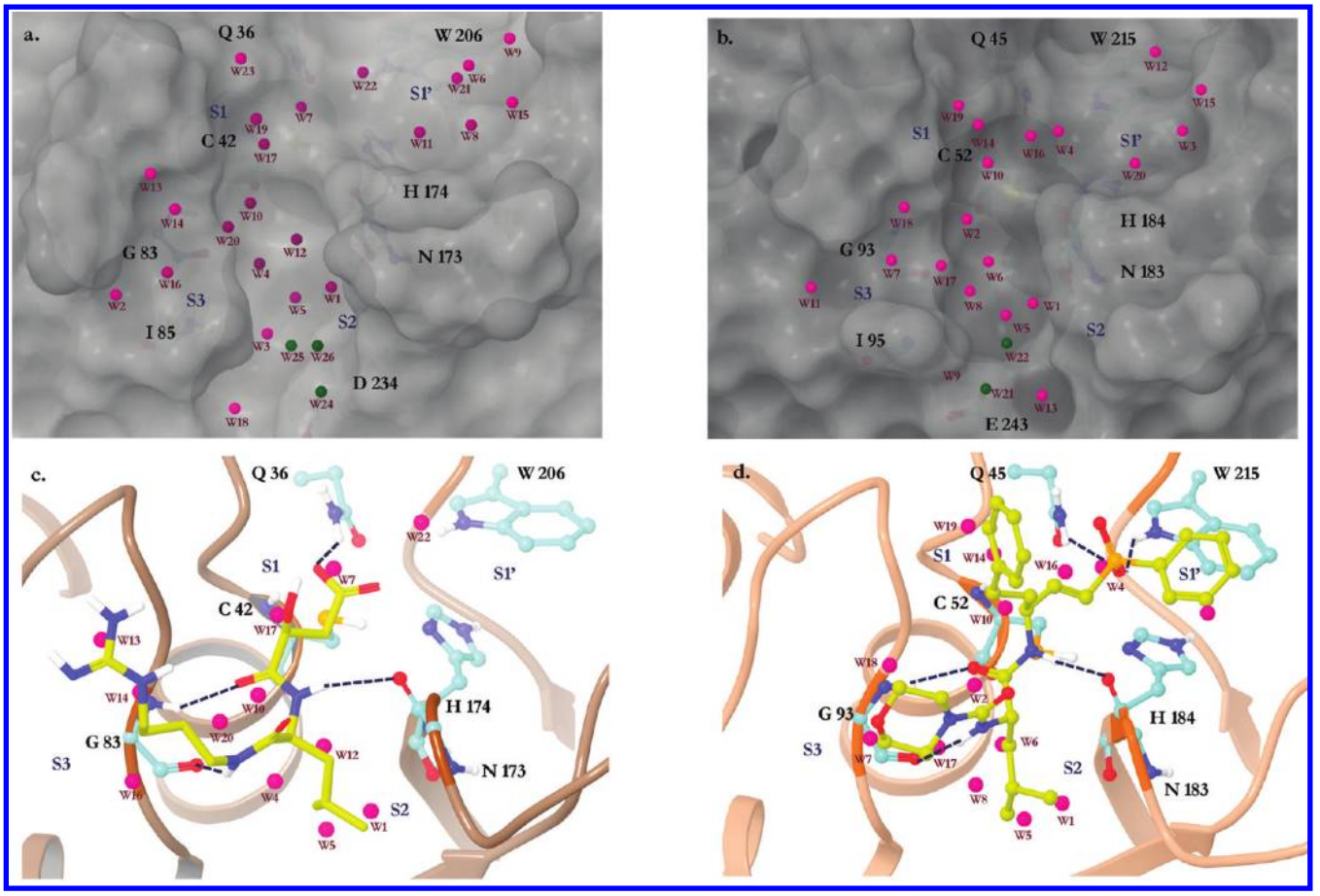


Figure 8. WaterMap profile of (a) FP-2, (b) FP-3, (c) FP-2 in complex with E-64, and (d) FP-3 in complex with vinyl sulfone. The thermodynamically interesting hydration sites important for the SAR are shown in spheres. Stable hydration sites ($\Delta G < 0$ kcal mol^{-1}) are colored in green, whereas significantly unstable hydration sites ($\Delta G > 1$ kcal mol^{-1}) are shown in purple. The water sites are labeled based on decreasing values of predicted ΔG. Key hydrogen bonding interactions of inhibitors (shown in yellow) with the residues of FP binding sites (shown in cyan) are displayed as dotted lines.

methods, such as MM-GBSA[56] and MM-PBSA,[57] are based on continuum electrostatics and do not account for the explicit nature of water. One method that accounts for waters explicitly, WaterMap, uses a combination of explicit solvent molecular dynamics (MD) simulation and statistical mechanics calculations to provide a quantitative description of enthalpy, entropy, and free energies of water molecules at the surface of proteins (see Methods).

From the design perspective, water molecules with positive values of ΔG resulting from positive enthalpy (ΔG) should be displaced by complementary groups of the ligand, with expected gains in binding affinity resulting from an increase in entropy when water is released into bulk solvent. Conversely, water molecules with negative enthalpy (ΔH) should be considered carefully, as displacement will result in a loss of enthalpy. Therefore, such waters should be avoided, bridged, or judiciously displaced with polar groups to replace polar interactions of water molecules.[43]

The WaterMap profiles of FP-2 and FP-3 with and without cocrystallized inhibitors (E-64 in FP-2 and a vinyl sulfone inhibitor k11017 in FP-3) are shown in Figure 8a–d. The thermodynamic breakdown of predicted hydration sites are shown in Supporting Information Tables S2 and S3. The location and energetics values of water molecules in these homologous proteases are consistent. The solvent analysis identifies the triangle of enthalpically unstable hydration sites (W1, W5, W12 in FP-2 and W1, W5, W6 in FP-3), solvating the hydrophobic S2 pocket in both FPs. The hydration sites in the narrow hydrophobically enclosed region of the S2 pocket are enthalpically unfavorable (positive ΔH), as they are unable to form a full complement of hydrogen bonds with protein residues (ΔG = >2.3 kcal mol^{-1} for all three sites in FP-2 and ΔG = >3.5 kcal mol^{-1} for all three sites in FP-3).

WaterMap predicts two unfavorable hydration sites (W10 = 2.6 kcal mol^{-1} and W20 = 1.3 kcal mol^{-1} in FP-2; W2 = 4.2 kcal mol^{-1} and W17 = 1.4 kcal mol^{-1} in FP-3) near Gly83 in FP-2 and Gly92 in FP-3. The interaction of a ligand with the conserved glycine near the S3 pocket is considered essential for inhibitors of papain-family cysteine proteases binding to nonprime sites (S1–S3).[52,58,59] Compounds lacking these predicted interactions were void of activity. When cocrystallized inhibitors were mapped onto the generated WaterMap profile, they can be seen to selectively displace these unfavorable hydration sites. For example, the shape of the P2-Leu of E-64 in FP-2 and of k11017 (VS inhibitor) in FP-3 complements the arrangement of the unstable water molecules of the S2 pocket, expelling them to the solvent (Figure 8a and b). This result is consistent with the observation that FP-2 and FP-3 prefer substrates with leucine in the P2 position[60,61] and is consistent with previous findings with other targets that show strong preference for specific amino acids being driven by water displacement.[62] The amide carbonyl near the P2 side chain

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696–710

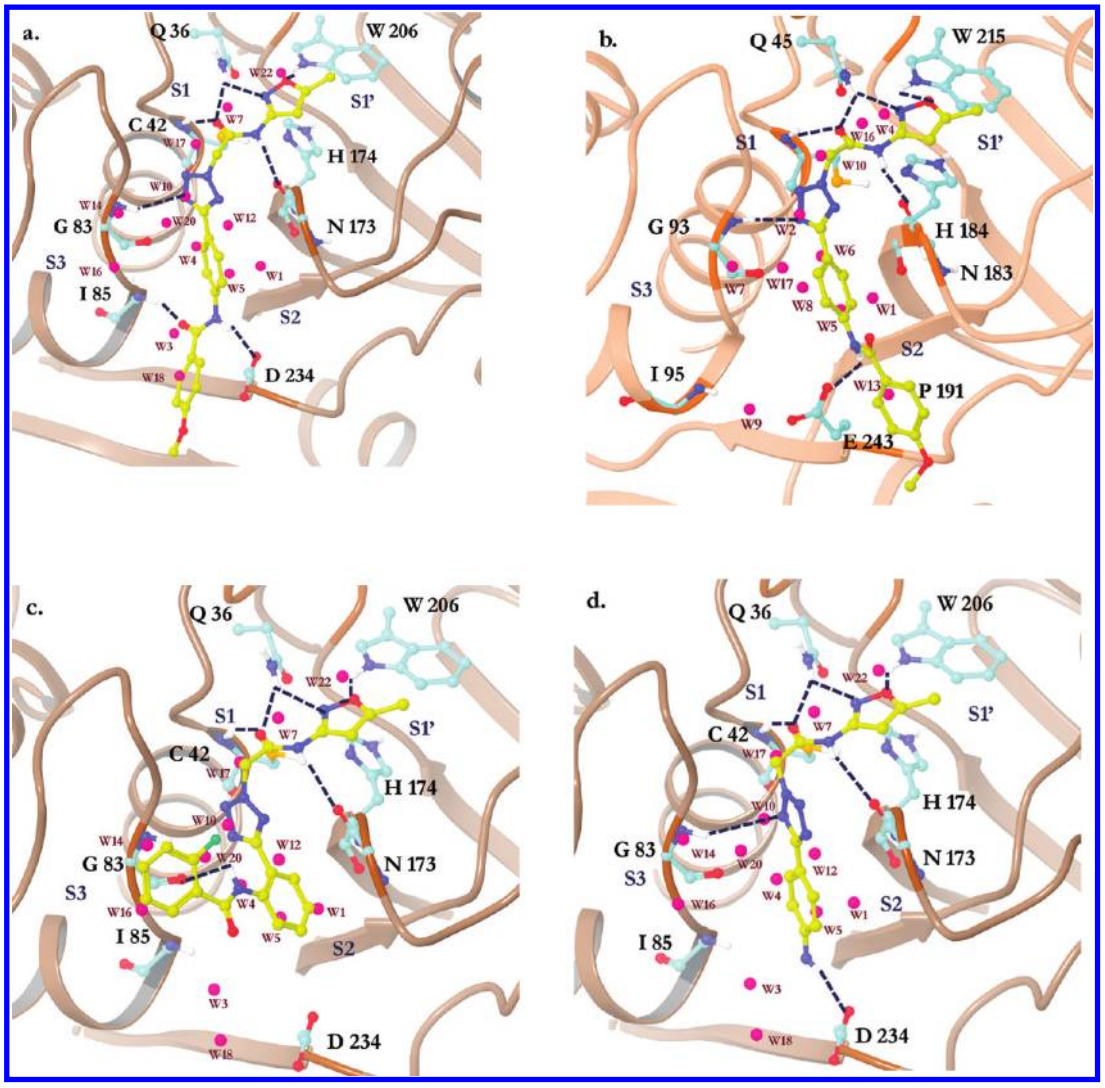


Figure 9. Predicted binding poses and thermodynamic hydration site profile of **4** in complex with (a) FP-2 and (b) FP-3; the predicted binding poses and thermodynamic profile of (c) **9** and (d) **14** in complex with FP-2. Water site labeling and color codes are the same as those in Figure 3.

displaces the unfavorable water (W10 in FP-2, ΔG = 2.6 kcal mol^{-1}; W2 in FP-3, ΔG = 4.2 kcal mol^{-1}) and replaces the H-bond with Gly83 in FP-2 and Gly92 in FP-3, thereby presumably imparting the observed gain in binding affinity.

WaterMap identifies unfavorable hydration sites (W7, ΔG = 3.0 kcal mol^{-1} and W22, ΔG = 1.1 kcal mol^{-1} in FP-2; W4, ΔG = 3.8 kcal mol^{-1} and W16, ΔG = 1.4 kcal mol^{-1} in FP-3) in the S1 pocket near the residues of an oxyanion hole (Cys42/51 and Gln36/45). W7 in FP-2 and both unstable waters (W4 and W16) in FP-3 are displaced by corresponding atoms of the cocrystallized inhibitors. In addition, the guanidine moiety of E-64 and morpholine side chain of the vinyl sulfone inhibitor displaces several unfavorable hydration sites from the S3 pocket of FP-2 and FP-3, respectively. In addition, WaterMap predicts other unfavorable hydration sites near the S3 pocket (W2, W3 in FP-2 and W9, W11 in FP-3) and in the S1′ pocket (W6, W8, W9, W11, W15, and W21 in FP-2; W3, W12, W15, and W20 in FP-3) that are out of the plane of these peptidic inhibitors. These sites could potentially be targeted, with anticipated gain in binding affinity, with a different ligand scaffold that is able to place functional groups to displace these hydration sites. Finally, WaterMap predicts stable clusters of water molecules (W24−W26 in FP-2; and W21, W22 in FP-3, ΔG and ΔH < 0 kcal mol^{-1}) in the vicinity of Asp234 and Ser149 of FP-2 and Glu243 and Ser158 of FP-3 of the S2 pocket that can potentially be bridged for water-mediated hydrogen bonding to gain binding affinity.

Mapping of the newly identified hits in this work on the WaterMap-generated thermodynamic hydration site profiles of FP-2 and FP-3 elucidated salient features of the SAR. For example, compound **4**, a linear analog from the tetrazole series, has a complementary shape to the binding site of both FP, and it displaces six unfavorable water molecules with ΔG > 2.0 kcal mol^{-1} of unfavorable hydration sites (Figure 9a and b). In particular, it displaces the unfavorable water W3 bound to Ile 85 in FP-2 (ΔG = 3.7 kcal mol^{-1}) and W18 (ΔG = 1.9 kcal mol^{-1}). Interestingly, compound **4** binds in a similar fashion to FP-3 and FP-2, but it could not displace the unfavorable water W9 (ΔG = 2.5 kcal mol^{-1}) near Ile 95 of FP-3, potentially explaining the log difference in potency of **4** to FP-3 over FP-2. This difference in hydration site displacement is possibly due to the narrow S2 pocket of FP-3 (Glu243 in FP-3 versus Asp234 in FP-2), directing the methoxy phenyl moiety of **4** close to Pro181.

Likewise, compounds **9** and **14** are unable to displace W3 and W18 from FP-2, and as a result, these compounds display

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

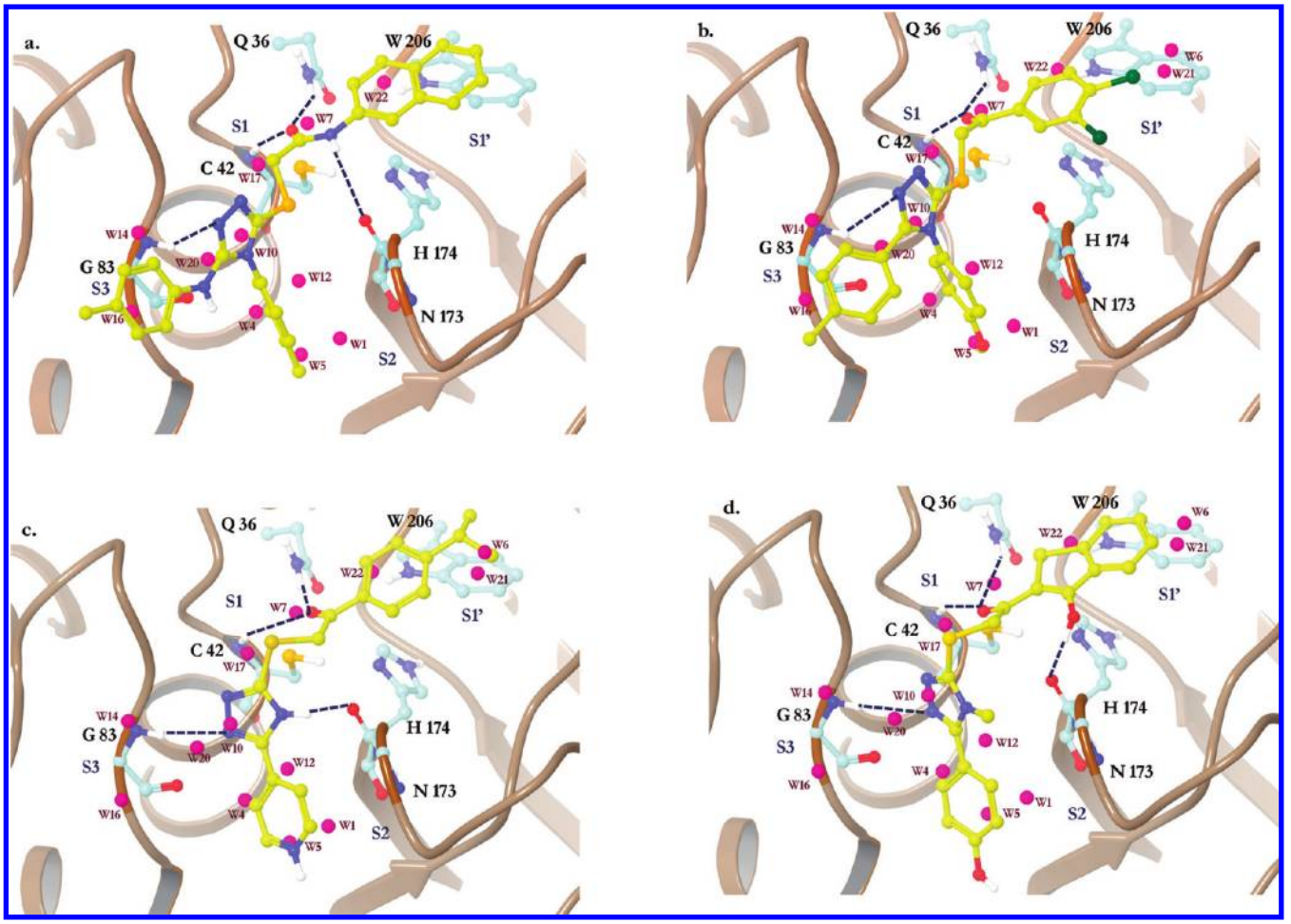


Figure 10. Predicted binding poses and thermodynamic hydration site profile of (a) **16**, (b) **17**, (c) **29**, and (d) **36** in complex with FP-2. Water site labeling and color codes are the same as those in Figure 3.

either diminished activity or inactivity against FP-2 (Figure 9c and d). In FP-3 these compounds also miss key unfavorable hydration sites and hence are devoid of activity (not shown). This may explain why linear compounds in the tetrazole series were more active than the branched compounds. The oxazole moiety of **4** and **9** displaces the unfavorable water (W22 in FP-2, ΔG = 1.1 kcal mol^{-1}, W4 in FP-3, ΔG= 3.8 kcal mol^{-1}) from the S1 pocket and forms an additional H-bond with Tyr 206, potentially explaining why heteroaromatic substituents were preferred in the S1′ pocket. These interactions and water displacements profiles together likely contribute to the higher affinity of **1** and **4** for both falcipains compared to other analogs in the series.

In the triazole series, compounds lacking R_2 substituents (**29**) or with shorter R_2 substituents (**36**) were unable to displace the hydrophobically enclosed unstable waters from the S2 pocket, and hence, they are inactive (Figure 10c and d). Moreover, R_2 substituents of this series projected differently than the P2-Leu of E-64. For example, compounds with a phenyl (**16**) or substituted phenyl (**17**) at R_2 could not displace the least favorable hydration sites from the S2 pocket, and hence, they showed reduced affinity for FP-2 (Figure 10a and b).

Compounds **3** and **57** from the quinazoline series with a heteroaromatic moiety are predicted to displace the unfavorable hydration sites (W7 and W22) from the S1−S1′ pocket as well as form an H-bond with the residues of the oxyanion hole such as Gln36, Cys42, and Tyr206 (Figure 11b). Compounds such as **32** and **38,** lacking corresponding heteroaromatic substitutions or unable to displace the unfavorable water molecules of the S1′ pocket, were inactive (Figure 11c). Compound **54** additionally displaces the unfavorable water molecules (W8, ΔG = 2.9 kcal mol^{-1}, W21, ΔG = 1.2 kcal mol^{-1}) from the S1′ pocket. The quinazoline moiety of **54** displaces the unstable hydration sites from the S2 pocket (except the least favorable W1) as well as from the S3 pocket (W14, ΔG = 2.2 kcal mol^{-1} and W16, ΔG = 2.1 kcal mol^{-1}), which likely accounts for moderate affinity of this compound for FP-2.

A few analogs with the ability to displace the high-energy water molecules of the S2 pocket, such as compound **68** from the quinazoline series (Figure 11d) or compound **30** (Figure 12d) from the triazole series, did not show additional gains in binding affinity. Also, none of the analogs of this series were more potent than compounds **1**−**3**, suggesting that there was an additional factor reducing the binding affinity of these compounds. On the basis of prior work in our group, we reasoned that the soft-nature of electrophiles present in these hits could account for their moderate affinity against FPs. In our previous study,[11] we utilized the LUMO density variant of atomic Fukui indices to locate the most electrophilic center and to estimate the reactivity of electrophiles present in the virtual screening hits. The predicted LUMO density of α-hetero amide electrophiles were in the range of 0.02−0.32 (on the scale of zero to one, one being the most electrophilic), lower than the LUMO density of compounds containing hard

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

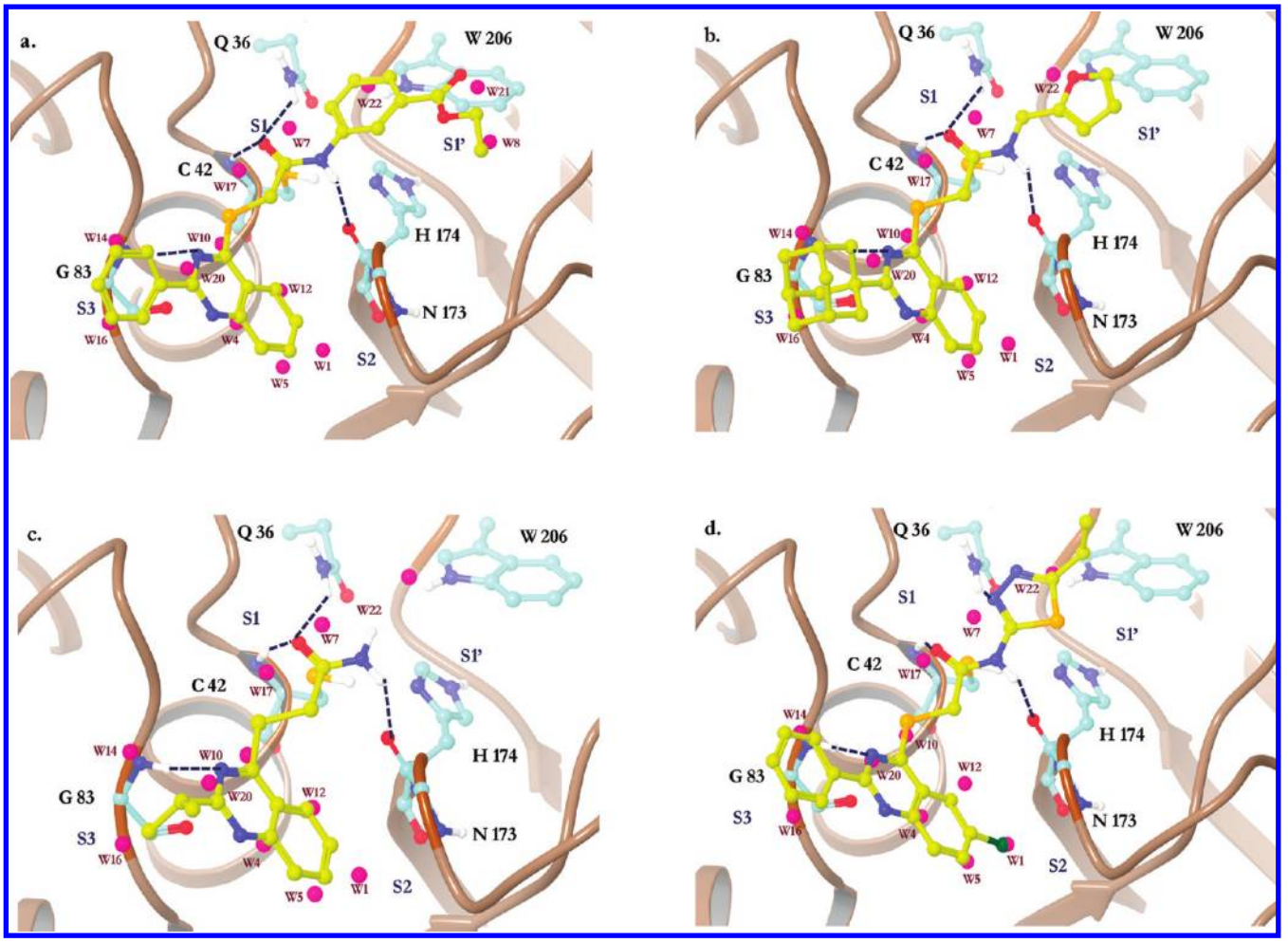


Figure 11. Predicted binding poses and thermodynamic profiles of (a) **54**, (b) **57**, (c) **67**, and (d) **68** in complex with FP-2. Water site labeling and color codes are the same as shown in Figure 3.

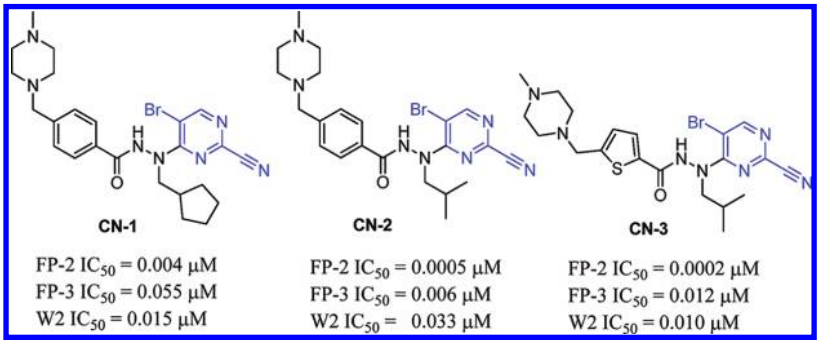


Figure 12. Potent inhibitors of FP-2, FP-3, and W2 from the 2-pyrimidinecarbonitrile series reported by Coteron et al.[50]

electrophiles, such as α-fluoro ketone (0.52) or vinyl sulfone (0.43). In a similar reactivity analysis, Obella and co-workers assessed the reactivity of diverse nitrile containing compounds using density functional theory calculations.[49] Their method involved approximating the free energy of formation of thioimidate by calculating the difference between the energy of thioimidate adduct [E(Adduct)], the nitrile molecule [E(Nitrile compound)], and the precursor methanethiol [E(Methanethiol)]. The generated reactivity index was scaled from 0 to −10 kcal/mol. According to their results, the heteroaromatic nitriles, in particular triazine and pyrimidine nitriles, were excellent electrophiles (predicted reactivity of −10 kcal mol^{-1}), whereas the aliphatic and the aryl nitriles were predicted as poor electrophiles (0 to −2 kcal mol^{-1}). Furthermore, the authors validated their findings by measuring the ability of representative aliphatic, aromatic, and heteroaromatic nitriles to form irreversible thiazoline adducts upon incubation with cysteine. The calculations showed that poorly electrophilic aryl nitriles (such as naphthalene or pyridine nitrile) have a reduced ability to form cysteine adducts, whereas more electrophilic pyrimidine nitriles or cyanamide have enhanced cysteine adduct formation. Their results showed excellent correlation between predicted reactivity and the ability to form irreversible cysteine adducts measured experimentally.[49]

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

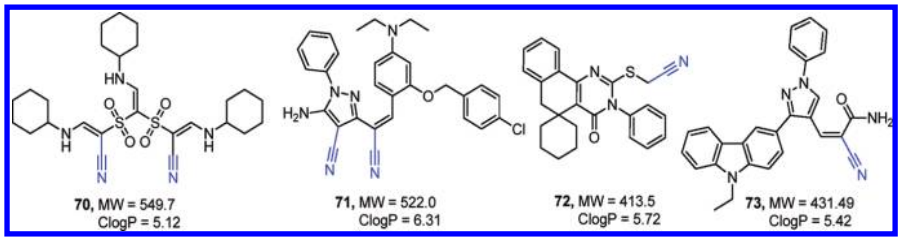


Figure 13. Aryl- or aliphatic nitrile-containing compounds discovered by substructure search in the Chembridge database. Alkyl nitrile electrophiles are highlighted in blue.

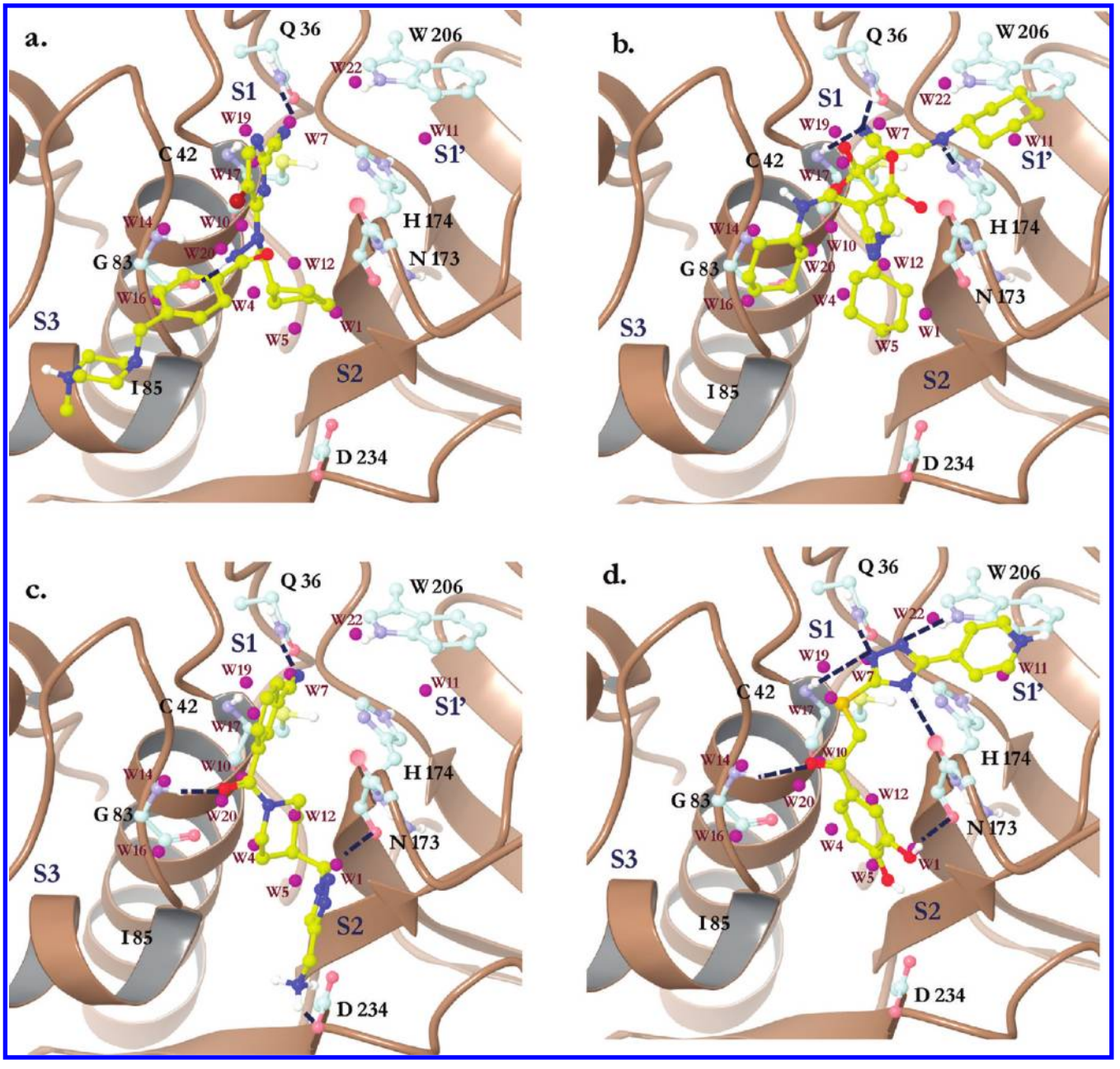


Figure 14. Predicted binding poses and thermodynamic hydration site profiles of (a) **CN-1**, (b) **70**, (c) **74**, and (d) **30** in complex with FP-2. Water site labeling and color codes are identical to those shown in Figure 3.

Recently, Coteron et al. published data for a 2-pyridinecarbonitrile series, showing subnanomolar activity against FPs as well as cultured parasites.[50] The representative compounds from this series are shown in Figure 12. As discussed above, the strong electrophile pyrimidine nitrile could be a major contributor in the high affinity of these compounds against FPs. This motivated us to further assess the importance of reactivity of electrophiles in the inhibition of cysteine proteases by evaluating a subset of aliphatic or aryl nitriles (purchased from Chembridge corporation) with poor electrophilicity

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

Table 5. Biological Evaluation of Selected Aryl or Aliphatic Nitriles against FP[a]

compound code	FP-2 IC_{50} (μM)	FP-3 IC_{50} (μM)	W2 IC_{50} (μM)	compound code	FP-2 IC_{50} (μM)	FP-3 IC_{50} (μM)	W2 IC_{50} (μM)
70	18.97	>50	7.89	**78**	>50	NT	>10
71	10.33	>50	5.92	**79**	>50	NT	>10
72	18.39	>50	>10	**80**	>50	NT	>10
73	11.21	>50	6.75	**81**	>50	NT	>10
74	>10	NT	>10	**82**	>50	NT	>10
75	>50	NT	>10	**83**	>50	NT	>10
76	>50	NT	>10	**E-64**	0.07	0.16	-
77	>50	NT	>10	artemisinin	-	-	0.092

[a]NT = not tested, W2 = chloroquine-resistant strain of *P. falciparum*.

(as proposed by Oballa et al.) against FP-2. If the reactivity of electrophiles plays an important role in the affinity of FP inhibitors discussed in the present study, then these compounds were expected to show moderate or less binding affinity as compared to **CN-1** to **3**.

As anticipated, a majority of compounds (**74**−**83**) did not show any activity against FP-2 up to 50 μM test concentration (Supporting Information Figure S2). The identified actives **70**−**73**, although having similar binding modes and water displacement profiles as **CN-1** (Figure 14), displayed moderate affinity toward FP-2 (IC_{50} = 10−20 μM, Figure 13). This study supports our hypothesis that the chemical reactivity of electrophiles has pronounced effects on the potency of cysteine protease inhibitors as well as justifies the low potency of identified FP-2 hits with soft-electrophiles. Our data clearly demonstrates that it can be difficult to achieve potent inhibition of FPs with compounds lacking reactive electrophilic centers (Table 5).

CONCLUSIONS

A combined ligand- and structure-based virtual screening was performed against the Asinex and Chembridge databases to discover nonpeptidic inhibitors of falcipains-2 (FP-2). The core structures of previously discovered virtual screenings hits were used as queries to mine commercial databases for structural analogs. A total of 69 putative hits with diverse R-groups, prioritized based on docking calculations, were experimentally evaluated against FP-2. Of the 69 compounds tested 28 inhibited FP-2, with IC_{50} values of 5−48 μM. Some of these compounds were also active against cultured malaria parasites, with IC_{50} < 10 μM. While most of these compounds were shown to bind to FP-2, it is possible that the active parasite inhibition is a result of promiscuous binding to other targets in the parasite, thus showing higher activity in culture than the FP assay.

The motivation for analyzing the water energetics with WaterMap was that all previous analyses using more traditional structure-based computational approaches (MM-GBSA, docking scores, and molecular mechanics interaction energies) as well as individual descriptors (hydrogen bonds, van der Waals, electrostatics, and ligand strain) could not explain the complex SAR of the compounds reported here. Furthermore, the majority of the modifications in this report do not change the hydrogen bond donor/acceptor capabilities of the ligands, yet we still observe significant changes in IC_{50} values. We found that the thermodynamic contributions of explicit solvent in the ligand binding domain of FP-2 and FP-3 could be used to explain most of the SAR. The hydration site thermodynamics revealed multiple unfavorable hydration sites, most notably in the S2 pocket of FP-2 and FP-3. Modeling predicted that the leucine moiety at the P2 position of the cocrystallized inhibitors E-64 and vinyl sulfone (k11017) in the S2 pockets of FP-2 and FP-3 complemented the arrangement of these hydrophobically enclosed waters and, as a result, displaced them from the S2 pocket of both FPs for increased binding affinity. This also reflected the preference for a P2-Leu as a recognition element for FP inhibition. The most potent hits generally displaced the most unfavorable hydration sites; however, a majority of the less potent hits were unable to displace the least favorable water of the S2 pocket (in a similar pattern to leucine), which accounts for the moderate affinity observed for most of these compounds.

In a few cases, displacement of the most unfavorable hydration sites did not correlate with the binding affinity. We were able to explain this based on the poor reactivity of soft electrophiles toward the catalytic cysteine in these weak hits. We evaluated moderately electrophilic alkyl and aryl nitrile compounds against FP-2 to further highlight the effect of the reactivity on cysteine protease inhibition. The aryl and alkyl nitriles only showed moderate affinity (10−20 μM) toward FP-2 as opposed to compounds with strong electrophiles such as pyrimidine nitriles, which have subnanomolar affinity.

The present study highlights three important factors that contribute to potent inhibitors of FP-2 and FP-3: (a) strong hydrogen bond networks with the key residues of the FP-2 active site such as Gly83, Cys42, Gln36, and Asn173; (b) displacement of unfavorable hydration sites (in particularly the triangle of unstable waters in the S2 pocket) by complementary groups of the ligands; and (c) chemical reactivity of electrophiles to the catalytic cysteine of the active site. In general, computing binding energies is a complex process with many factors involved. In the study here, the congeneric nature of the ligands simplified the problem because a large part of the structure is constant, thereby canceling in the calculation of binding energies for different ligands. However, computing differences in SAR series is still challenging and can be made increasingly difficult as receptor flexibility becomes significant within the series. For the ligands studied here, it appears that substantial receptor flexibility was not needed to accommodate the ligands within the series, which likely contributes to the good results obtained in this work.

In summary, we found that terms that are often neglected from scoring functions (i.e., explicit water energies and chemical reactivity) can play an important role in binding. While the chemical reactivity of a ligand is very important in cases of covalent bond formation, as was the case for the FPs studied here, the hydration site thermodynamics are expected to be important in most ligand binding processes, as this is the underlying mechanism for the hydrophobic effect. The combined analysis presented in this work using multiple energetic contributors to binding was able to explain the SAR for the

dx.doi.org/10.1021/ci2005516 | J. Chem. Inf. Model. 2012, 52, 696−710

FP inhibitors. However, more work is needed to develop a unified scoring function that can simultaneously account for all of these terms.

■ ASSOCIATED CONTENT

Supporting Information

Inactive compounds from 1,2,4-triazole and aryl and aliphatic nitrile series, overlay of E-64 with representative compounds from each series, thermodynamic properties of selected water molecules in FP-2 and FP-3 binding site, and ClogP values of compounds **1**−**69**. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 662-915-5879. Fax: 662-915-5638. E-mail: mavery@olemiss.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by CDC Cooperative Agreement UR3/CCU4186520-03.

■ REFERENCES

(1) Dluzewski, A. R.; Rangachari, K.; Wilson, R. J.; Gratzer, W. B. Plasmodium falciparum: protease inhibitors and inhibition of erythrocyte invasion. *Exp Parasitol* **1986**, *62*, 416−422.

(2) World Health Organization. *World Malaria Report 2009* (accessed October 8th, 2010).

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

(4) Ekland, E. H.; Fidock, D. A. In vitro evaluations of antimalarial drugs and their relevance to clinical outcomes. *Int. J. Parasitol.* **2008**, *38*, 743−747.

(5) Wisniewski, M.; Zak, D. J. Achievements and perspectives of research into development of a vaccine against malaria. *Wiad. Parazytol.* **2010**, *56*, 133−140.

(6) Dondorp, A. M.; Nosten, F.; Yi, P.; Das, D.; Phyo, A. P.; Tarning, J.; Lwin, K. M.; Ariey, F.; Hanpithakpong, W.; Lee, S. J.; Ringwald, P.; Silamut, K.; Imwong, M.; Chotivanich, K.; Lim, P.; Herdman, T.; An, S. S.; Yeung, S.; Singhasivanon, P.; Day, N. P.; Lindegardh, N.; Socheat, D.; White, N. J. Artemisinin resistance in Plasmodium falciparum malaria. *N. Engl. J. Med.* **2009**, *361*, 455−467.

(7) 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.

(8) 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.

(9) Goud, P. M.; Sheri, A.; Desai, P. V.; Watkins, E. B.; Tekwani, B.; Sabnis, Y.; Gut, J.; Rosenthal, P. J.; Avery, M. A. Design, synthesis and evaluation of trisubstituted thiazoles targeting Plasmodium falciparum cysteine proteases. *Med. Chem. Res.* **2005**, *14*, 74−105.

(10) Sabnis, Y.; Rosenthal, P. J.; Desai, P.; Avery, M. A. Homology modeling of falcipain-2: validation, de novo ligand design and synthesis of novel inhibitors. *J. Biomol. Struct. Dyn.* **2002**, *19*, 765−774.

(11) Shah, F.; Mukherjee, P.; Gut, J.; Legac, J.; Rosenthal, P. J.; Tekwani, B. L.; Avery, M. A. Identification of Novel Malarial Cysteine Protease Inhibitors Using Structure-Based Virtual Screening of a Focused Cysteine Protease Inhibitor Library. *J. Chem. Inf. Mod.* **2011**, *51*, 852−864.

(12) Shah, F.; Wu, Y.; Gut, J.; Pedduri, Y.; Legac, J.; Rosenthal, P. J.; Avery, M. A. Design, synthesis and biological evaluation of novel benzothiazole and triazole analogs as falcipain inhibitors. *Med. Chem. Commun.* **2011**, DOI: 10.1039/C1031MD00129A.

(13) Batra, S.; Sabnis, Y. A.; Rosenthal, P. J.; Avery, M. A. Structure-based approach to falcipain-2 inhibitors: synthesis and biological evaluation of 1,6,7-Trisubstituted dihydroisoquinolines and isoquinolines. *Biorg. Med. Chem.* **2003**, *11*, 2293−2299.

(14) 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. USA* **2004**, *101*, 4384−4389.

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

(16) 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.

(17) Rockett, K. A.; Playfair, J. H. L.; Ashall, F.; Target, G. A. T.; Angliker, H.; Shaw, E. Inhibition of intraerythrocytic development of Plasmodium falciparum by proteinase inhibitors. *FEBS Lett.* **1990**, *259*, 257−259.

(18) 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.

(19) 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.

(20) Olson, J. E.; Lee, G. K.; Semenov, A.; Rosenthal, P. J. Antimalarial effects in mice of orally administered peptidyl cysteine protease inhibitors. *Biorg. Med. Chem.* **1999**, *7*, 633−638.

(21) 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.

(22) Rosenthal, P. J. Cysteine proteases of malaria parasites. *Int. J. Parasitol.* **2004**, *34*, 1489−1499.

(23) Rosenthal, P. J. Plasmodium falciparum: effects of proteinase inhibitors on globin hydrolysis by cultured malaria parasites. *Exp. Parasitol.* **1995**, *80*, 272−281.

(24) 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.

(25) 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.

(26) 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.

(27) Ring, C. S.; Sun, E.; McKerrow, J. H.; Lee, G. K.; Rosenthal, P. J.; Kuntz, I. D.; Cohen, F. E. Structure-based inhibitor design by using protein models for the development of antiparasitic agents. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3583−3587.

(28) 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.

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

(30) Avery, M. A.; Alvim-Gaston, M.; Rodrigues, C. R.; Barreiro, E. J.; Cohen, F. E.; Sabnis, Y. A.; Woolfrey, J. R. Structure-activity relationships of the antimalarial agent artemisinin. 6. The development of predictive in vitro potency models using CoMFA and HQSAR methodologies. *J. Med. Chem.* **2002**, *45*, 292−303.

(31) Avery, M. A.; Alvim-Gaston, M.; Vroman, J. A.; Wu, B.; Ager, A.; Peters, W.; Robinson, B. L.; Charman, W. Structure-activity relation-

ships of the antimalarial agent artemisinin. 7. Direct modification of (+)-artemisinin and in vivo antimalarial screening of new, potential preclinical antimalarial candidates. *J. Med. Chem.* **2002**, *45*, 4321−4335.
(32) Avery, M. A.; Bonk, J. D.; Chong, W. K.; Mehrotra, S.; Miller, R.; Milhous, W.; Goins, D. K.; Venkatesan, S.; Wyandt, C.; Khan, I.; et al. Structure-activity relationships of the antimalarial agent artemisinin. 2. Effect of heteroatom substitution at O-11: synthesis and bioassay of N-alkyl-11-aza-9-desmethylartemisinins. *J. Med. Chem.* **1995**, *38*, 5038−5044.
(33) Avery, M. A.; Fan, P.; Karle, J. M.; Bonk, J. D.; Miller, R.; Goins, D. K. Structure-activity relationships of the antimalarial agent artemisinin. 3. Total synthesis of (+)-13-carbaartemisinin and related tetra- and tricyclic structures. *J. Med. Chem.* **1996**, *39*, 1885−1897.
(34) Avery, M. A.; Gao, F.; Chong, W. K.; Mehrotra, S.; Milhous, W. K. Structure-activity relationships of the antimalarial agent artemisinin. 1. Synthesis and comparative molecular field analysis of C-9 analogs of artemisinin and 10-deoxoartemisinin. *J. Med. Chem.* **1993**, *36*, 4264−4275.
(35) Avery, M. A.; Mehrotra, S.; Bonk, J. D.; Vroman, J. A.; Goins, D. K.; Miller, R. Structure-activity relationships of the antimalarial agent artemisinin. 4. Effect of substitution at C-3. *J. Med. Chem.* **1996**, *39*, 2900−2906.
(36) Avery, M. A.; Mehrotra, S.; Johnson, T. L.; Bonk, J. D.; Vroman, J. A.; Miller, R. Structure-activity relationships of the antimalarial agent artemisinin. 5. Analogs of 10-deoxoartemisinin substituted at C-3 and C-9. *J. Med. Chem.* **1996**, *39*, 4149−4155.
(37) Sun, L.; Shah, F.; Helal, M. A.; Wu, Y.; Pedduri, Y.; Chittiboyina, A. G.; Gut, J.; Rosenthal, P. J.; Avery, M. A. Design, synthesis, and development of novel guaianolide-endoperoxides as potential antimalarial agents. *J. Med. Chem.* **2010**, *53*, 7864−7868.
(38) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the active-site solvent in the thermodynamics of factor Xa ligand binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817−2831.
(39) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein−ligand binding. *Proc. Natl. Acad. Sci.* **2007**, *104*, 808.
(40) Li, Z.; Lazaridis, T. Thermodynamic contributions of the ordered water molecule in HIV-1 protease. *J. Am. Chem. Soc.* **2003**, *125*, 6636−6637.
(41) Lazaridis, T. Inhomogeneous fluid approach to solvation thermodynamics. 2. Applications to simple fluids. *J. Phys. Chem. B* **1998**, *102*, 3542−3550.
(42) Lazaridis, T. Inhomogeneous fluid approach to solvation thermodynamics. 1. Theory. *J. Phys. Chem. B* **1998**, *102*, 3531−3541.
(43) Pellicciari, R.; Camaioni, E.; Gilbert, A. M.; Macchiarulo, A.; Bikker, J. A.; Shah, F.; Bard, J.; Costantino, G.; Gioiello, A.; Robertson, G. M. Discovery and characterization of novel potent PARP-1 inhibitors endowed with neuroprotective properties: From TIQ-A to HYDAMTIQ. *Med. Chem. Commun.* **2011**, *2*, 559−565.
(44) Chrencik, J. E.; Patny, A.; Leung, I. K.; Korniski, B.; Emmons, T. L.; Hall, T.; Weinberg, R. A.; Gormley, J. A.; Williams, J. M.; Day, J. E. Structural and Thermodynamic Characterization of the TYK2 and JAK3 Kinase Domains in Complex with CP-690550 and CMP-6. *J. Mol. Biol.* **2010**, *400*, 413−433.
(45) Abel, R.; Salam, N. K.; Shelley, J.; Farid, R.; Friesner, R. A.; Sherman, W. Contribution of Explicit Solvent Effects to the Binding Affinity of Small Molecule Inhibitors in Blood Coagulation Factor Serine Proteases. *ChemMedChem* **2011**, *6*, 1049−1066.
(46) Higgs, C.; Beuming, T.; Sherman, W. Hydration Site Thermodynamics Explain SARs for Triazolylpurines Analogues Binding to the A2A Receptor. *ACS Med. Chem. Lett.* **2010**, *1*, 160−164.
(47) Robinson, D. D.; Sherman, W.; Farid, R. Understanding kinase selectivity through energetic analysis of binding site waters. *ChemMedChem* **2010**, *5*, 618−627.
(48) Snyder, P. W.; Mecinovia, J.; Moustakas, D. T.; Thomas Iii, S. W.; Harder, M.; Mack, E. T.; Lockett, M. R.; Haroux, A.; Sherman, W.; Whitesides, G. M. Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. *Proc. Natl. Acad. Sci.* **2011**, *108*, 17889−17894.
(49) 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.
(50) 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.
(51) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(52) 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.
(53) 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.
(54) Mukherjee, P.; Desai, P.; Ross, L.; White, E. L.; Avery, M. A. Structure-based virtual screening against SARS-3CL(pro) to identify novel non-peptidic hits. *Bioorg. Med. Chem.* **2008**, *16*, 4138−4149.
(55) Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **2003**, *300*, 1763.
(56) Massova, I.; Kollman, P. A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect. Drug Discovery Des.* **2000**, *18*, 113−135.
(57) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(58) 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.
(59) 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.
(60) 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 anti-malarial drug discovery. *Biochem. J.* **2006**, *399*, 47.
(61) 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.
(62) Beuming, T.; Farid, R.; Sherman, W. High energy water sites determine peptide binding affinity and specificity of PDZ domains. *Protein Sci.* **2009**, *18*, 1609−1619.