THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 38, pp. 25697–25703, September 18, 2009

Vinyl Sulfones as Antiparasitic Agents and a Structural Basis for Drug Design*[S]

Received for publication, April 28, 2009, and in revised form, July 1, 2009 Published, JBC Papers in Press, July 20, 2009, DOI 10.1074/jbc.M109.014340

Iain D. Kerr[‡], Ji H. Lee[‡1], Christopher J. Farady[§], Rachael Marion[‡], Mathias Rickert[‡2], Mohammed Sajid[¶3], Kailash C. Pandey[‖], Conor R. Caffrey[¶], Jennifer Legac[‖], Elizabeth Hansell[¶], James H. McKerrow[¶], Charles S. Craik[§], Philip J. Rosenthal[‖4], and Linda S. Brinen[‡5]

From the [‡]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158, the [§]Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-2280, the [¶]Department of Pathology and Sandler Center for Basic Research in Parasitic Diseases, University of California, San Francisco, California 94143, and the [‖]Department of Medicine, San Francisco General Hospital, University of California, San Francisco, California 94143

Cysteine proteases of the papain superfamily are implicated in a number of cellular processes and are important virulence factors in the pathogenesis of parasitic disease. These enzymes have therefore emerged as promising targets for antiparasitic drugs. We report the crystal structures of three major parasite cysteine proteases, cruzain, falcipain-3, and the first reported structure of rhodesain, in complex with a class of potent, small molecule, cysteine protease inhibitors, the vinyl sulfones. These data, in conjunction with comparative inhibition kinetics, provide insight into the molecular mechanisms that drive cysteine protease inhibition by vinyl sulfones, the binding specificity of these important proteases and the potential of vinyl sulfones as antiparasitic drugs.

Sleeping sickness (African trypanosomiasis), caused by *Trypanosoma brucei*, and malaria, caused by *Plasmodium falciparum*, are significant, parasitic diseases of sub-Saharan Africa (1). Chagas' disease (South American trypanosomiasis), caused by *Trypanosoma cruzi*, affects approximately, 16–18 million people in South and Central America. For all three of these protozoan diseases, resistance and toxicity to current therapies makes treatment increasingly problematic, and thus the development of new drugs is an important priority (2–4).

T. cruzi, *T. brucei*, and *P. falciparum* produce an array of potential target enzymes implicated in pathogenesis and host cell invasion, including a number of essential and closely related papain-family cysteine proteases (5, 6). Inhibitors of cruzain and rhodesain, major cathepsin L-like papain-family cysteine proteases of *T. cruzi* and *T. brucei rhodesiense* (7–10) display considerable antitrypanosomal activity (11, 12), and some classes have been shown to cure *T. cruzi* infection in mouse models (11, 13, 14).

In *P. falciparum*, the papain-family cysteine proteases falcipain-2 (FP-2)[6] and falcipain-3 (FP-3) are known to catalyze the proteolysis of host hemoglobin, a process that is essential for the development of erythrocytic parasites (15–17). Specific inhibitors, targeted to both enzymes, display antiplasmodial activity (18). However, although the abnormal phenotype of FP-2 knock-outs is "rescued" during later stages of trophozoite development (17), FP-3 has proved recalcitrant to gene knockout (16) suggesting a critical function for this enzyme and underscoring its potential as a drug target.

Sequence analyses and substrate profiling identify cruzain, rhodesain, and FP-3 as cathepsin L-like, and several studies describe classes of small molecule inhibitors that target multiple cathepsin L-like cysteine proteases, some with overlapping antiparasitic activity (19–22). Among these small molecules, vinyl sulfones have been shown to be effective inhibitors of a number of papain family-like cysteine proteases (19, 23–27). Vinyl sulfones have many desirable attributes, including selectivity for cysteine proteases over serine proteases, stable inactivation of the target enzyme, and relative inertness in the absence of the protease target active site (25). This class has also been shown to have desirable pharmacokinetic and safety profiles in rodents, dogs, and primates (28, 29). We have determined the crystal structures of cruzain, rhodesain, and FP-3 bound to vinyl sulfone inhibitors and performed inhibition kinetics for each enzyme. Our results highlight key areas of

* This work was supported, in whole or in part, by National Institutes of Health Grants AI35707 and AI35800 (to C. R. C.). This work was also supported by the Sandler Foundation, the Medicines for Malaria Venture, and a Charles E. Culpeper Biomedical Pilot Initiative Award (to C. R. C.). Part of this research was performed at SSRL, a national user facility operated by Stanford University on behalf of the U. S. Dept. of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Dept. of Energy, Office of Biological and Environmental Research, and by the National Center for Research Resources, Biomedical Technology Program, NIH, and NIGMS, NIH.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) 2OZ2, 2P7U, and 3BWK.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S3 and Table S1.

[1] Present address: CrystalGenomics, Inc., 6F, 2nd Building, Asan Life Sciences Institute, Seoul 138-878, South Korea.

[2] Present address: Pfizer Rinat, South San Francisco, CA 94080.

[3] Present address: Leiden University Medical Centre, afd. Parasitologie, Albinusdreef 2, Kamer P4-35, Leiden 2333 ZA, Netherlands.

[4] A Doris Duke Charitable Foundation Distinguished Clinical Scientist.

[5] To whom correspondence should be addressed: Dept. of Cellular & Molecular Pharmacology and The Sandler Center for Basic Research in Parasitic Diseases, QB3/Byers Hall 508C, 1700 4th St., University of California, San Francisco, CA 94158-2550. Tel.: 415-514-3426; Fax: 415-502-8193; E-mail: brinen@cmp.ucsf.edu.

[6] The abbreviations used are: FP-2, falcipain-2; FP-3, falcipain-3; SAM, Stanford Automated Mounting; SSRL, Stanford Synchrotron Radiation Lightsource; r.m.s.d., root mean square distance; VSPh, phenyl vinyl sulfone; Hph, homophenylalanyl; Mu, morpholino urea; *N*-Mpip, *N*-methylpiperazine; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; Z, benzyloxycarbonyl; AMC, aminomethylcoumarin.

TABLE 1
X-ray diffraction data and structure refinement statistics

	Cruzain·K11777	Rhodesain·K11777	FP-3·K11017
Data collection			
Space group	C2	P1	$P4_12_12$
Cell dimensions			
a, *b*, *c* (Å)	*a* = 134.3, *b* = 38.0, *c* = 95.2	*a* = 34.4, *b* = 39.7, *c* = 39.6	*a* = *b* = 114.0, *c* = 226.1
a, *b*, *g*	90°, 114°, 90°	120°, 94°, 101°	90°, 90°, 90°
Resolution (Å)	1.95 (2.02-1.95)	1.65 (1.70-1.65)	2.42 (2.55-2.42)
R_{merge}[a] (%)	7.1 (15.3)	4.7 (14.7)	9.1 (37.3)
$I/\sigma I$	26 (9.3)	14.1 (3.9)	21.8 (7.2)
Completeness (%)	97.7 (90.0)	89.7 (83.7)	100 (100)
Redundancy	3.6 (3.4)	1.8 (1.3)	14.2 (14.6)
Refinement			
R_{free}/R_{factor} (%)	20.7/15.7	17.5/13.5	20.9/17.5
Average *B*-factor (Å^2)	15.2	13.3	25.8
r.m.s.d.			
Bond length (Å)	0.019	0.020	0.015
Bond angle	1.6°	1.7°	1.5°
Ramachandran plot[b]			
Favored (%)	97.4	95.8	97.2
Allowed (%)	100	99.5	99.8
Outliers (%)	0	0.5	0.2
PDB ID	2OZ2	2P7U	3BWK

[a] $R_{merge} = \Sigma\Sigma I(h)j - (I(h))/\Sigma\Sigma I(h)j$, where $I(h)$ is the measured diffraction intensity, and the summation includes all observations.
[b] As defined by Molprobity (45).

interaction between proteases and inhibitors. These results help validate the vinyl sulfones as a class of antiparasitic drugs and provide structural insights to facilitate the design or modification of other small molecule inhibitor scaffolds.

EXPERIMENTAL PROCEDURES

Expression and Purification of the Cruzain·K11777 Complex—Recombinant cruzain was expressed in *Escherichia coli* and purified as described previously (8, 30, 31). Activated cruzain was incubated overnight with molar excess amounts of inhibitor dissolved in DMSO to prevent further proteolytic activity. Complete enzymatic inhibition was confirmed via fluorometric assay with the substrate Z-Phe-Arg-AMC. Excess inhibitor was removed by anion-exchange chromatography. Fractions containing pure, inhibited cruzain were pooled and concentrated to 8 mg/ml, with tandem buffer exchange to 2 mM Bis-Tris, pH 5.8, using a Viva-Spin (Viva Science) column (molecular mass of 15 kDa).

Crystallization and Structure Determination of the Cruzain·K11777 Complex—Crystals of maximum size were obtained after ~4 days via the hanging drop method, from a solution of 1.25 M ammonium sulfate and 100 mM HEPES, pH 7.5, at 22 °C. Crystals were cryoprotected in mother liquor containing 20% ethylene glycol, mounted in standard cryo loops, and loaded into a sample cassette used with the Stanford Automated Mounting (SAM) system (32).

All diffraction data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) Beamline 9-1, Menlo Park, CA, after selecting an optimal crystal from screening performed with the robotic SAM system (32). Data processing in the HKL2000 package (33) showed that crystals belonged to space group C2, and the structure was solved by molecular replacement using a model derived from cruzain bound to the vinyl sulfone, K11002 (PDB ID 1F29). Using MOLREP (34), two independent molecules were located with translation function scores of 14.49 and 14.03. Rigid body refinement of this solution yielded an R_{factor} of 46%. Clear and representative density for the entirety of both inhibitor molecules in the asymmetric unit was observed at better than 1.5σ above the noise level. The model was completed by interspersing iterative rounds of model building in COOT (35) and reciprocal space refinement in REFMAC5 (36). Waters were placed with COOT and manually assessed. Molecules of the cryoprotectant ethylene glycol and the crystallization precipitant ammonium sulfate were also discernable in final electron density maps and placed manually with COOT. This structure has been deposited in the Protein Data Bank (code 2OZ2). All statistics for data collection, structure solution, and refinement are given in Table 1.

Expression and Purification of the Rhodesain·K11777 Complex—Rhodesain (without the unusual C-terminal extension shared between trypanosomatid cathepsin Ls) was expressed in *P. pastoris* and purified as described previously (7) with a Ser > Ala mutation incorporated at position 172 of the protein sequence to remove a glycosylation site from the mature domain. Active rhodesain was incubated with molar excess of the inhibitor, dissolved in DMSO. Extinction of activity was confirmed by fluorometric assay with the Z-Phe-Arg-*N*-methylcoumarin substrate. Purified rhodesain was concentrated to ~7 mg/ml using vacuum dialysis in preparation for crystallization.

Crystallization and Structure Determination of the Rhodesain·K11777 Complex—Crystals of maximum size were obtained after ~6 days via the hanging drop method, from a solution of 100 mM imidazole, pH 8.0, and 1.0 M sodium citrate at 18 °C. Diffraction data were collected at room temperature on a Rigaku RU200 rotating anode source using CuK_α radiation at 1.54 Å and a Rigaku R-Axis IV detector. Data processing was performed in space group P1 with the HKL2000 software package (33). The structure was solved via molecular replacement in AMoRe (37), using cruzain (PDB ID 1F2A) as a search model. The top solution had a correlation coefficient of 64.3 and an R_{factor} of 37.7%. The inhibitor was manually placed and fit to the difference electron density using QUANTA (Accelrys). Clear

and representative density for the entirety of the inhibitor molecule was observed at better than 1.5σ above the noise level. Water molecules were placed with COOT (35) and then manually assessed. Final rounds of refinement were completed with REFMAC5 (36). This structure has been deposited in the Protein Data Bank (code 2P7U). All statistics for data collection, structure solution, and refinement are given in Table 1.

Expression and Purification of the FP-3·K11017 Complex—FP-3 was expressed in *E. coli* strain M15(pREP4) transformed with the hexa-His-tagged FP-3-pQE-30 construct. Overexpression, refolding, and purification were carried out according to published protocols (38). The activity of FP-3 was tested with the substrate Z-Leu-Arg-AMC, as described (39), and completely abolished by the addition of vinyl sulfone inhibitor K11017 to a final concentration of 113 μM. Inhibited FP-3 was purified using a 10 ml of Q-Sepharose column and was eluted with a high salt buffer (20 mM Bis-Tris, pH 6.5, 0.5 M NaCl). Fractions that contained FP-3 were verified by SDS-PAGE, pooled, and buffer exchanged with 20 mM Bis-Tris, pH 6.5, and the enzyme was concentrated to ~10 mg/ml.

Crystallization and Structure Determination of the FP-3·K11017 Complex—Crystals were grown using the hanging drop, vapor-diffusion method (40) from a mixture of 1 μl of protein solution (10 mg/ml) and 1 μl of reservoir solution (1.26 M ammonium sulfate, 100 mM Tris-HCl, pH 8.5, 200 mM lithium sulfate) incubated at room temperature against 1 ml of reservoir solution. Crystals grew to a maximum size of 50 × 50 × 100 μm in 5 days.

Crystals of FP-3·K11017 grew as hexagonal rods. Cryoprotection was achieved by a brief soak in a solution containing mother-liquor solutions supplemented with 20% glycerol. All crystals were mounted in standard cryo loops and loaded into a sample cassette used with the SAM (32). Diffraction data were collected at SSRL Beamline 7-1. Reflection intensities were indexed and integrated using MOSFLM (41). Data were scaled and merged in space group $P4_12_12$ using SCALA (42).

The structure of FP-3·K11017 was determined by molecular replacement in PHASER (43) using the FP-3 component of the FP-3·leupeptin complex (PDB ID 3BPM). Four independent monomers were located in the asymmetric unit yielding a solution with an R_{factor} of 31% and a log-likelihood gain of 7214. After initial rounds of rigid body refinement and simulated annealing in CNS (44), K11017 was positioned in the active site of all four monomers according to $mF_o - DF_c$ SIGMAA-weighted electron density maps. Following several rounds of model building in COOT (35) interspersed with positional and *B*-factor refinement in CNS, waters were placed in difference map peaks greater than or equal to 3σ with reasonable hydrogen bonding. The final model shows excellent stereochemistry as assessed by MOLPROBITY (45). Statistics for this structure, which has been deposited in the PDB (3BWK), are summarized in Table 1.

Inhibition Kinetics of Cruzain, Rhodesain, and FP-3—Reactions were run in a 96-well Microfluor-1 U-bottom plate (Thermo Electron) and monitored in a SpectraMax Gemini fluorescence spectrometer (Molecular Devices) with excitation of 355 nm and emission at 460 nm, with a cutoff at 435 nm. Reactions were carried out in 100 mM sodium acetate, pH 5.5, 5 mM dithiothreitol, 0.001% Triton X-100, and 1% DMSO. For the inhibitors K11017 and K11777, 10 mM stock solutions in 100% DMSO were made by weighing out lyophilized compound. Inhibition constants were determined under pseudo-first order conditions using the progress curves method (46). Briefly, enzyme was added to a mixture of substrate and inhibitor, and the hydrolysis of an AMC substrate was monitored for 7 min (<10% total substrate hydrolysis). An observed rate constant, k_{obs}, was calculated at each inhibitor concentration by fitting the progress curve to the equation, $P = v_i/k_{obs}(1 - e^{(-k_{obs}t)})$, where P = product formation, v_i = initial velocity, and time = t. Second order rate constants (either k_a or K_{inact}/K_i) were determined depending on the kinetic behavior of the enzyme. If k_{obs} varied linearly with inhibitor concentration, the association constant k_a was determined by fitting to the linear equation, $k_{obs} = (k_a[I])/(1 + [S]/K_m)$. If k_{obs} varied hyperbolically with [I], the k_{inact}/K_I was determined by non-linear regression using the equation, $k_{obs} = k_{intact}[I]/([I] + K_I(1 + [S]/K_m))$. The K_m was experimentally determined using standard Michaelis-Menten kinetics, and all experiments were carried out in triplicate. For cruzain, the enzyme concentration was 2.5 nM, the substrate concentration was 8 μM Z-Phe-Arg-AMC (K_m = 0.65 μM), and inhibitor concentrations varied from 1 μM to 0.1 μM. For rhodesain, the enzyme concentration was 8 nM, the substrate concentration was 4 μM Z-Phe-Arg-AMC (K_m = 0.12 μM), and inhibitor concentrations varied from 10 μM to 0.1 μM. For FP-3, the enzyme concentration was 10 nM, the substrate concentration was 100 μM Z-Leu-Arg-AMC (K_m = 86 μM), and inhibitor concentrations varied from 2 μM to 30 μM. The FP-3 activity buffer also contained 3% glycerol.

RESULTS

Crystal Structure Determination—The cruzain·K11777 complex, which crystallized with two complete copies of the mature enzyme (residues 1–215) in the asymmetric unit, was determined to 1.95-Å resolution. The model was refined to an R_{free} of 20.7% and an R_{factor} of 15.9%. Both copies are essentially identical, and superimposition matched all 215 α-carbons of each chain with a root mean square distance (r.m.s.d.) of 0.19 Å. The rhodesain·K11777 complex crystallized with a single complete copy of the mature enzyme in the asymmetric unit (residues 1–215). The complex was refined to a resolution of 1.65 Å yielding an R_{free} of 17.5% and an R_{factor} of 13.5%. The structure of the FP-3·K11017 complex crystallized with four copies of the complex in the asymmetric unit and was determined to 2.42 Å. This represents residues 8–249 of the mature enzyme. The final model was refined to an R_{free} of 20.9% and R_{factor} of 17.5%. All four copies of the complex are very similar (supplemental Table S1) and superimposition matches, on average, 234 α-carbons with a mean r.m.s.d. of 0.26 Å.

Overall Structures—All three enzymes share the common two-domain fold of papain superfamily cysteine proteases (Fig. 1). However, cruzain and rhodesain share a higher degree of structural similarity (214 α-carbons matching with an r.m.s.d. of 0.49 Å) than either does with FP-3 (191 and 190 α-carbons matching with an r.m.s.d. of 1.1 Å, respectively). The structure of FP-3 deviates slightly from the classic papain fold in having two insertions, one at either terminus, that are unique to plas-

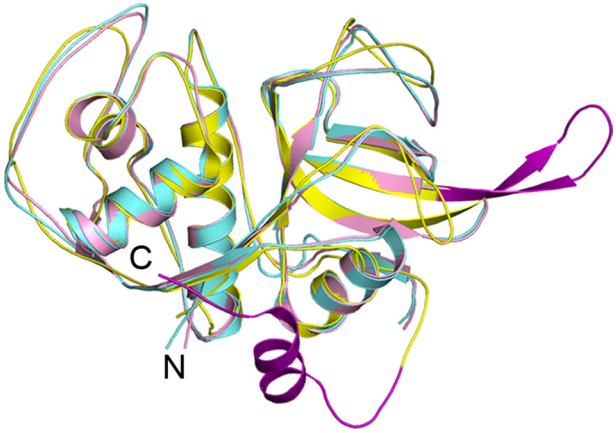

FIGURE 1. **The structures of cruzain, rhodesain, and falcipain-3.** Ribbon representation and superimposition of cruzain (*pink*), rhodesain (*aquamarine*), and falcipain-3 (*yellow*) are shown. Insertions in falcipain-3 are colored *purple*. All structure figures were prepared in PyMOL (DeLano, 2002).

modial cysteine proteases (supplemental Fig. S1) (47, 48). The N-terminal insertion (residues 1–25) is well ordered in our complex and has been implicated in the correct folding of the enzyme (49, 50). The C-terminal insertion is implicated in binding the *in vivo* substrate of FP-3, hemoglobin (39, 48), and is composed of residues 194–207. In our structure, this insertion is ordered in monomers A and B, but residues 195–203 in chain C and 194–204 in chain D were too flexible to be included in the final model. For the sake of simplicity, unless otherwise indicated, our analyses were performed using chain A of each model.

The chemical structures of K11017 (Mu-Leu-Hph-VSPh) and K11777 (*N*-Mpip-Phe-Hph-VSPh) are similar with phenyl vinyl sulfone (VSPh) at the P1′ position and homophenylalanyl (Hph) at the P1 position. Variation occurs at the P2 position, Leu and Phe, respectively, and the P3 position, morpholino urea (Mu) and *N*-methyl piperazine (*N*-Mpip) respectively (Fig. 2). The co-crystallized inhibitors span the respective S1′–S3 subsites and form an irreversible, covalent adduct with the sulfur of the active site cysteine thiol in each enzyme (Fig. 2). In each complex, there is a small conserved network of polar interactions between protein and inhibitor involving Gln-19, Gly-66, Asp-161, His-162, and Trp-184 (cruzain numbering, Fig. 3). These interactions serve to anchor the peptidyl backbone of the inhibitor in the protease active site and do not confer a preference for a particular substituent at any position (P1′–P3) of the bound inhibitor. Water-mediated and hydrophobic interactions also contribute to binding and are discussed in more detail below.

Inhibition of Cruzain, Rhodesain, and FP-3 by K11017 and K11777—To further investigate the utility of vinyl sulfones as inhibitors of papain family cysteine proteases we determined the inhibition kinetics of cruzain, rhodesain, and FP-3 in the presence of both K11017 and K11777 (Table 2). Inhibition was monitored using the progress curves method. For cruzain and FP-3 the observed inhibitory rate constants varied linearly with inhibitor concentration, and we therefore calculated k_a, the association constant. In the case of rhodesain, the rate of inhibition varied hyperbolically with inhibitor concentration and the second order inhibition constant k_{inact}/K_i was used (46).

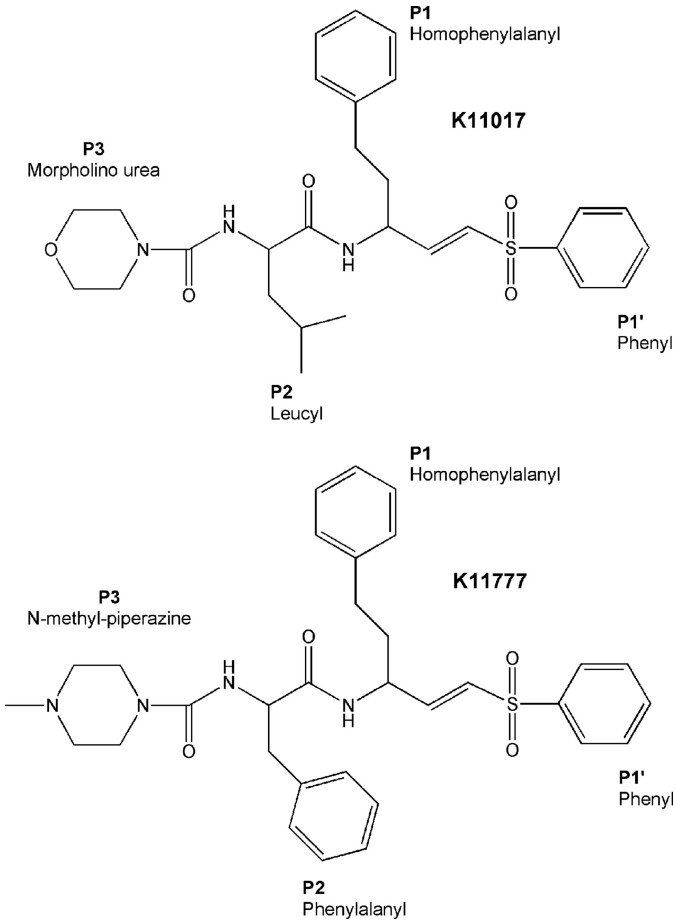

FIGURE 2. **The chemical structures of K11017 and K11777.** The P1′-P3 positions of K11017 (*top*) and K11777 (*bottom*) are labeled.

The kinetic data show potent inhibition of cruzain and rhodesain by each inhibitor, with K11017 showing slightly better inhibition of both enzymes. Cruzain and rhodesain tolerate a range of hydrophobic residues in their S2 subsites (51, 52), and, whereas a minor effect, we were unable to reconcile a slight preference for K11017 by consideration of the P2 residue alone. These results support the suggestion that interactions at other subsites are also important. FP-3 has a clear preference for Leu at P2 (38), and therefore, as expected, FP-3 is preferentially inhibited by K11017, which has an 8-fold higher k_a compared with K11777. Both vinyl sulfones inhibited FP-3 less efficiently than cruzain and rhodesain, with second order inhibition constants for K11017 and K11777 being two orders of magnitude lower. This reduced activity is at least partly attributable to the lower catalytic efficiency of FP-3 in the presence of peptides when compared with the *Trypanosoma* enzymes (38).

DISCUSSION

We present the crystal structures of cruzain·K11777, rhodesain·K11777, and FP-3·K110117. This is the first structure reported for rhodesain and the first structure of an FP-3·vinyl sulfone inhibitor complex. Cruzain, rhodesain, and FP-3 all share the active site catalytic triad (His/Cys/Asn) of papain-family cysteine proteases. Given the hydrophobic nature of the P1 and P2 substituents, it is not surprising that the active site in each complex is lined with a number of residues that are able to

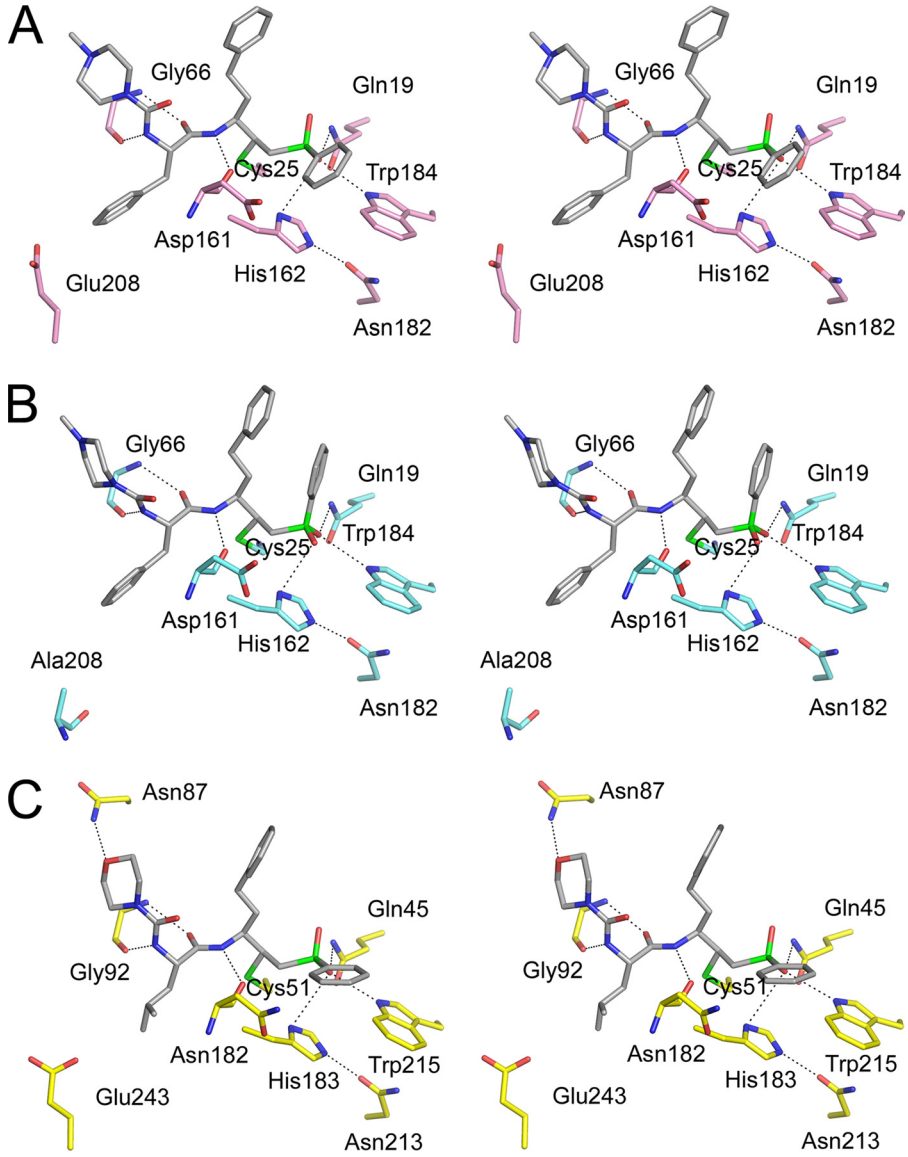

FIGURE 3. **The active sites of cruzain (*A*), rhodesain (*B*), and falcipain-3 (*C*).** *Ball and stick* representation shows the conserved catalytic triad and other important residues. The figure is colored as in Fig. 1 with each inhibitor in *gray*.

TABLE 2
Vinyl sulfone second order inhibition constants

	Cruzain, k_a	Rhodesain, k_{inact}/K_i	FP-3, k_a
		$M^{-1}\,s^{-1}$	
K11017	676,000	264,000	8,800
K11777	517,000	150,000	1,050

make non-polar contacts with their respective inhibitor (Fig. 4). At the primary sequence level, similarity at these positions allows the P1, P2, and P3 substituents in each complex to adopt similar conformations. As is seen in other crystal structures of cruzain (53, 54), the residue at the bottom of the S2 subsite (Glu-208) points out of the pocket to avoid a potentially unfavorable interaction with the bulky Phe residue at the P2 position of K11777. A similar situation is seen in the FP-3·K11017 structure (Glu-243), whereas in rhodesain, the residue at this position is Ala and steric clash is mitigated.

The conserved structure of the papain fold allows facile superimposition of the three protozoan proteases and reveals a striking difference at the S1′ subsite. In the rhodesain complex, the conserved phenyl-sulfone group at the P1′ position is flipped ~90° out of the active site in relation to the cruzain·K11777 and FP-3·K11017 complexes. In comparison with cruzain, the substitution of Trp for the slightly less bulky Phe at position 144 in rhodesain allows this residue to more readily access a deep, buried pocket in the bottom of the S1′ subsite. As a consequence, the neighboring Met-145 is able to penetrate deeper into the S1′ subsite and prevent the phenyl sulfone substituent from lying flat (Fig. 5).

Superimposition of the FP-3 and rhodesain complexes shows a similar Trp to Phe substitution (Phe-165 in FP-3), however, the amino acid equivalent to Met-145 in rhodesain is the less bulky Ala-166, which allows the phenyl sulfone of the inhibitor to rest on the floor of the S1′ subsite, as is normally seen in structures of cruzain with vinyl sulfones. Interestingly, flipping of the phenyl sulfone at the P1′ position seems to be a consistent structural feature in rhodesain. The high resolution crystal structure of rhodesain in complex with the vinyl sulfone K11002 (PDB ID 2P86, Mu-Phe-Hph-VSPh)[7] reveals that this flipping may be transient and in this complex the P1′ moiety is modeled at half occupancy in both the "in" and "out" conformations.

Cruzain and rhodesain have a strong preference for large hydrophobic residues and Leu at the P2 position of peptide substrates (31, 51, 55). In papain-family cysteine proteases the P2 position can be a key determinant of specificity. Our kinetic data show that K11017 and K11777 display very strong inhibition against these two parasite proteases, and this result can be correlated with the fact that K11777 and K11017 have hydrophobic P2 groups (Phe and Leu, respectively). A more striking difference is seen at the P3 position where *N*-Mpip in K11777 is substituted for Mu in K11017. The P3 substituent of the vinyl sulfones has recently been a point of particular interest, and modification of this position has been shown to influence a number of properties, including lysosomotropism, hepatotox-

[7] I. D. Kerr, P. Wu, R. Marion-Tsukemaki, Z. B. Mackey, and L. S. Brinen, manuscript in preparation.

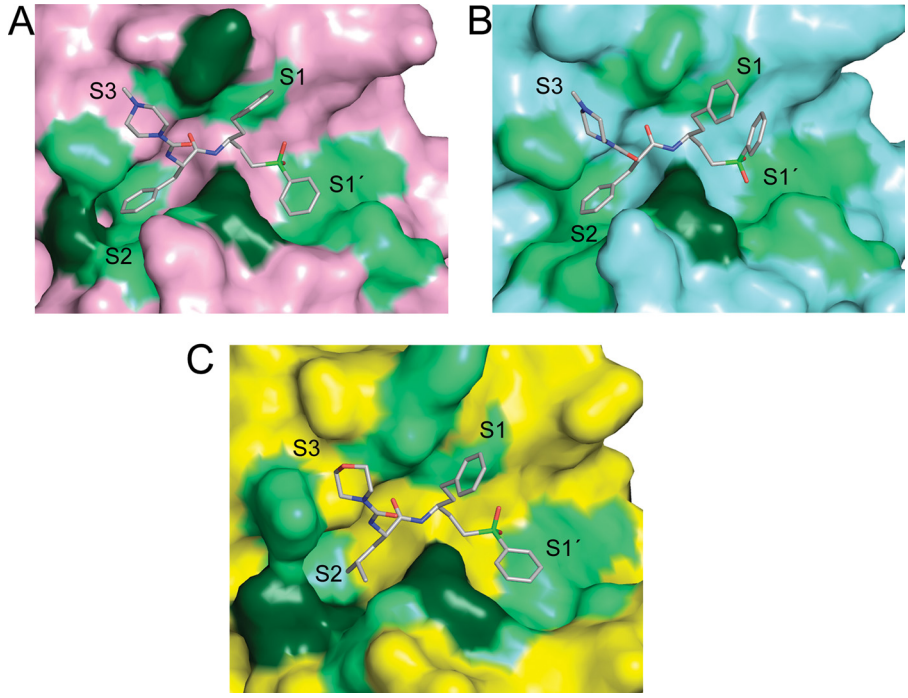

FIGURE 4. **The hydrophobic environment of the active sites.** Surface representation of the substrate binding sites of cruzain (*A*), rhodesain (*B*), and falcipain-3 (*C*). Hydrophobic residues are colored *light green*, and polar residues that interact with the ligand through their non-polar C–C bonds are colored *dark green*.

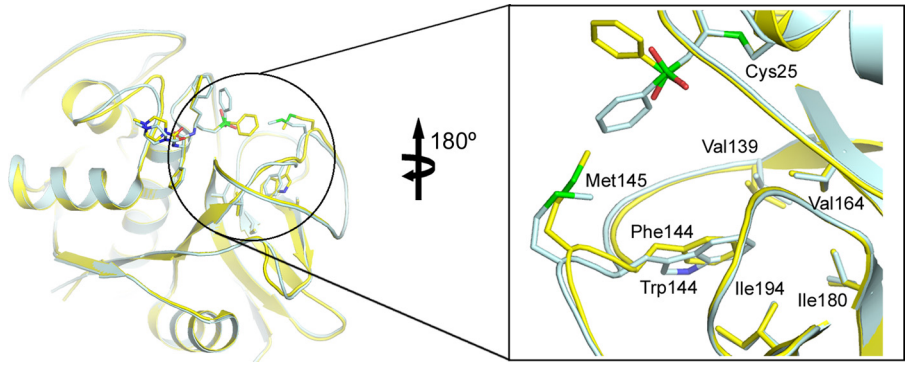

FIGURE 5. **Phenyl sulfone flipping in the rhodesain·K11777 complex.** Ribbon representation and detailed inlay of the S1′ subsites of cruzain (*pale blue*) and rhodesain (*yellow*) are shown. Important residues are depicted in *ball and stick* representation. Bound ligands are colored as their respective protease partners.

icity, and pharmacokinetics (56, 57). Our earlier determination of the cruzain·K11002 (Mu-Phe-Hph-VSPh) complex showed that this enzyme is well suited to accommodate the Mu substituent, with a network of bridging water molecules anchoring the morpholino oxygen of the inhibitor to the solvent-exposed Asp-60 and Ser-61 in the S3 subsite (53) (supplemental Fig. S2). The *N*-Mpip substituent of K11777 excludes some of this water upon binding to cruzain, and the P3 position of the inhibitor is therefore unable to form the same polar interactions with the enzyme that Mu is. The absence of these polar interactions in the cruzain·K11777 complex may account for the slight preference seen for inhibition by K11017.

Rhodesain also shows a slight preference for inhibition by K11017. Although we are lacking a rhodesain·K11017 structure, superimposition of the rhodesain·K11002 and rhodesain·K11777 structures shows that the residue equivalent to Ser-61 in rhodesain (Phe-61) is able to largely exclude the solvent that would otherwise be available to allow Mu to interact with S3 residues (supplemental Fig. S2). Meanwhile, in the rhodesain·K11777 complex, Phe-61 makes torsional re-adjustments about $\chi 1$ to swing ~28° out of the S3 subsite and provide room for the branched *N*-Mpip (supplemental Fig. S3). Therefore, although rhodesain is unable to form any specific polar interactions with Mu, this moiety may nonetheless be preferred to the slightly larger *N*-Mpip.

In contrast to cruzain, rhodesain and the closely related FP-2, FP-3 is typically far less catalytically active against peptide substrates and less responsive to inhibition by peptidyl-based small molecules (38, 47). These observations are underscored by our kinetic results, which show that FP-3 is markedly less sensitive to inhibition by both K11017 and K11777 than either cruzain or rhodesain. We have previously speculated that the S2 subsite in FP-3 site is particularly restricted for a cathepsin L-like protease through a combination of two "gatekeeper" residues (Tyr-93 and Pro-181) and the Glu at the bottom of the S2 subsite (58). Our FP-3·K11017 structure provides four independent views of the complex in the asymmetric unit and in at least one (monomer A) the entrance to the S2 subsite appears to be almost completely occluded. Indeed, the structural data correlate well with previous biochemical studies showing that the enzyme has a very narrow and a clear preference for substrates with Leu *versus* the more bulky Phe at the P2 position (38). Consistent with this substrate preference, FP-3 was therefore much less sensitive to K11777 than K11017.

Our kinetic and structural data show that cruzain and rhodesain can be targeted for inhibition by the vinyl sulfones. Indeed, K11777 has been shown in pre-clinical trials to be non-mutagenic, to be well tolerated with an acceptable pharmacokinetic profile, and to demonstrate efficacy in models of acute and chronic Chagas disease in both mice and dogs. On the basis of these results a pre-filing for an Investigational New Drug application is in preparation to allow the inhibitor to enter Phase I trials in human subjects. In comparison with the trypanosomal enzymes, consideration of the vinyl sulfones as effective FP-3 inhibitors may prove more challenging, especially in light of the structural restrictions on the S2 subsite. However, this may allow us to engineer a certain amount of selectivity that may be

lacking in the case of both cruzain and rhodesain. We believe that plasmodial cysteine proteases are still very promising drug targets, and we are hopeful that our structural insights will aid in the design of small molecules that better inhibit these enzymes.

Acknowledgments—We thank Irimpan Mathews for beamline support at the SSRL. We also thank Dr. Thomas Stout and Prof. Robert Fletterick for critical appraisal of the manuscript and Christopher Franklin for support with figure preparation.

REFERENCES

1. Balakrishnan, I., and Gillespie, S. H. (2002) in *Principles and Practice of Travel Medicine* (Zuckerman, J. N., ed) pp. 91–124, John Wiley & Sons, New York
2. Docampo, R., and Moreno, S. N. (2003) *Parasitol Res* **90,** Suppl. 1, S10–S13
3. Laufer, M. K., Djimdé, A. A., and Plowe, C. V. (2007) *Am. J. Trop. Med. Hyg.* **77,** 160–169
4. Ouellette, M. (2001) *Trop. Med. Int. Health* **6,** 874–882
5. Caffrey, C. R., Scory, S., and Steverding, D. (2000) *Curr. Drug Targets* **1,** 155–162
6. Rosenthal, P. J. (2004) *Int. J. Parasitol.* **34,** 1489–1499
7. Caffrey, C. R., Hansell, E., Lucas, K. D., Brinen, L. S., Alvarez Hernandez, A., Cheng, J., Gwaltney, S. L., 2nd, Roush, W. R., Stierhof, Y. D., Bogyo, M., Steverding, D., and McKerrow, J. H. (2001) *Mol. Biochem. Parasitol.* **118,** 61–73
8. Eakin, A. E., Mills, A. A., Harth, G., McKerrow, J. H., and Craik, C. S. (1992) *J. Biol. Chem.* **267,** 7411–7420
9. Nkemgu, N. J., Grande, R., Hansell, E., McKerrow, J. H., Caffrey, C. R., and Steverding, D. (2003) *Int. J. Antimicrob. Agents* **22,** 155–159
10. Steverding, D., Caffrey, C. R., and Sajid, M. (2006) *Mini. Rev. Med. Chem.* **6,** 1025–1032
11. Engel, J. C., Doyle, P. S., and McKerrow, J. H. (1999) *Medicina* **59,** Suppl. 2, 171–175
12. Vicik, R., Hoerr, V., Glaser, M., Schultheis, M., Hansell, E., McKerrow, J. H., Holzgrabe, U., Caffrey, C. R., Ponte-Sucre, A., Moll, H., Stich, A., and Schirmeister, T. (2006) *Bioorg. Med. Chem. Lett.* **16,** 2753–2757
13. Engel, J. C., Doyle, P. S., Hsieh, I., and McKerrow, J. H. (1998) *J. Exp. Med.* **188,** 725–734
14. Engel, J. C., Doyle, P. S., Palmer, J., Hsieh, I., Bainton, D. F., and McKerrow, J. H. (1998) *J. Cell Sci.* **111,** 597–606
15. Rosenthal, P. J., McKerrow, J. H., Aikawa, M., Nagasawa, H., and Leech, J. H. (1988) *J. Clin. Invest.* **82,** 1560–1566
16. Sijwali, P. S., Koo, J., Singh, N., and Rosenthal, P. J. (2006) *Mol. Biochem. Parasitol.* **150,** 96–106
17. Sijwali, P. S., and Rosenthal, P. J. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101,** 4384–4389
18. Schulz, F., Gelhaus, C., Degel, B., Vicik, R., Heppner, S., Breuning, A., Leippe, M., Gut, J., Rosenthal, P. J., and Schirmeister, T. (2007) *ChemMedChem* **2,** 1214–1224
19. Chen, Y. T., Lira, R., Hansell, E., McKerrow, J. H., and Roush, W. R. (2008) *Bioorg. Med. Chem. Lett.* **18,** 5860–5863
20. Fujii, N., Mallari, J. P., Hansell, E. J., Mackey, Z., Doyle, P., Zhou, Y. M., Gut, J., Rosenthal, P. J., McKerrow, J. H., and Guy, R. K. (2005) *Bioorg. Med. Chem. Lett.* **15,** 121–123
21. González, F. V., Izquierdo, J., Rodriguez, S., McKerrow, J. H., and Hansell, E. (2007) *Bioorg. Med. Chem. Lett.* **17,** 6697–6700
22. Greenbaum, D. C., Mackey, Z., Hansell, E., Doyle, P., Gut, J., Caffrey, C. R., Lehrman, J., Rosenthal, P. J., McKerrow, J. H., and Chibale, K. (2004) *J. Med. Chem.* **47,** 3212–3219
23. Ettari, R., Nizi, E., Di Francesco, M. E., Dude, M. A., Pradel, G., Vicík, R., Schirmeister, T., Micale, N., Grasso, S., and Zappalà, M. (2008) *J. Med. Chem.* **51,** 988–996
24. Jaishankar, P., Hansell, E., Zhao, D. M., Doyle, P. S., McKerrow, J. H., and Renslo, A. R. (2008) *Bioorg. Med. Chem. Lett.* **18,** 624–628
25. Palmer, J. T., Rasnick, D., Klaus, J. L., and Brömme, D. (1995) *J. Med. Chem.* **38,** 3193–3196
26. Rosenthal, P. J., Olson, J. E., Lee, G. K., Palmer, J. T., Klaus, J. L., and Rasnick, D. (1996) *Antimicrob. Agents Chemother.* **40,** 1600–1603
27. Shenai, B. R., Lee, B. J., Alvarez-Hernandez, A., Chong, P. Y., Emal, C. D., Neitz, R. J., Roush, W. R., and Rosenthal, P. J. (2003) *Antimicrob. Agents Chemother.* **47,** 154–160
28. McKerrow, J. H., Rosenthal, P. J., Swenerton, R., and Doyle, P. (2008) *Curr. Opin. Infect. Dis.* **21,** 668–672
29. Renslo, A. R., and McKerrow, J. H. (2006) *Nat. Chem. Biol.* **2,** 701–710
30. Eakin, A. E., McGrath, M. E., McKerrow, J. H., Fletterick, R. J., and Craik, C. S. (1993) *J. Biol. Chem.* **268,** 6115–6118
31. Gillmor, S. A., Craik, C. S., and Fletterick, R. J. (1997) *Protein Sci.* **6,** 1603–1611
32. Cohen, A. E., Ellis, P. J., Deacon, A. M., Miller, M. D., and Phizackerley, R. P. (2002) *J. Appl. Crystallogr.* **35,** 720–726
33. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
34. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30,** 1022–1025
35. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
36. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
37. Navaza, J. (1994) *Acta Crystallogr. A* **50,** 157–163
38. Sijwali, P. S., Shenai, B. R., Gut, J., Singh, A., and Rosenthal, P. J. (2001) *Biochem. J.* **360,** 481–489
39. Pandey, K. C., Wang, S. X., Sijwali, P. S., Lau, A. L., McKerrow, J. H., and Rosenthal, P. J. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102,** 9138–9143
40. McPherson, A. (1982) *Preparation and Analysis of Protein Crystals*, John Wiley & Sons, New York
41. Leslie, A. G. (1992) *Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography* **26**
42. Evans, P. R. (1997) *Joint CCP4 and ESF-EAMCB Newsletter on protein Crystallography* **33,** 22–24
43. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *J. Appl. Crystallogr.* **40,** 658–674
44. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921
45. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) *Nucleic Acids Res.* **35,** W375–383
46. Bieth, J. G. (1995) *Methods Enzymol.* **248,** 59–84
47. Wang, S. X., Pandey, K. C., Scharfstein, J., Whisstock, J., Huang, R. K., Jacobelli, J., Fletterick, R. J., Rosenthal, P. J., Abrahamson, M., Brinen, L. S., Rossi, A., Sali, A., and McKerrow, J. H. (2007) *Structure* **15,** 535–543
48. Wang, S. X., Pandey, K. C., Somoza, J. R., Sijwali, P. S., Kortemme, T., Brinen, L. S., Fletterick, R. J., Rosenthal, P. J., and McKerrow, J. H. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103,** 11503–11508
49. Pandey, K. C., Sijwali, P. S., Singh, A., Na, B. K., and Rosenthal, P. J. (2004) *J. Biol. Chem.* **279,** 3484–3491
50. Sijwali, P. S., Shenai, B. R., and Rosenthal, P. J. (2002) *J. Biol. Chem.* **277,** 14910–14915
51. Choe, Y., Leonetti, F., Greenbaum, D. C., Lecaille, F., Bogyo, M., Brömme, D., Ellman, J. A., and Craik, C. S. (2006) *J. Biol. Chem.* **281,** 12824–12832
52. O'Brien, T. C., Mackey, Z. B., Fetter, R. D., Choe, Y., O'Donoghue, A. J., Zhou, M., Craik, C. S., Caffrey, C. R., and McKerrow, J. H. (2008) *J. Biol. Chem.* **283,** 28934–28943
53. Brinen, L. S., Hansell, E., Cheng, J., Roush, W. R., McKerrow, J. H., and Fletterick, R. J. (2000) *Structure* **8,** 831–840
54. Huang, L., Brinen, L. S., and Ellman, J. A. (2003) *Bioorg. Med. Chem.* **11,** 21–29
55. Harris, J. L., Backes, B. J., Leonetti, F., Mahrus, S., Ellman, J. A., and Craik, C. S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97,** 7754–7759
56. Jacobsen, W., Christians, U., and Benet, L. Z. (2000) *Drug Metab. Dispos.* **28,** 1343–1351
57. Zhang, Y., Guo, X., Lin, E. T., and Benet, L. Z. (1998) *Drug Metab. Dispos.* **26,** 360–366
58. Kerr, I. D., Lee, J. H., Pandey, K. C., Harrison, A., Sajid, M., Rosenthal, P. J., and Brinen, L. S. (2009) *J. Med. Chem.* **52,** 852–857