Vinyl Sulfones as Antiparasitic Agents and a Structural Basis for Drug Design

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, Christopher J. Farady, Rachael Marion, Mathias Rickert, Mohammed Sajid, Kailash C. Pandey, Conor R. Caffrey, Jennifer Legac, Elizabeth Hansell, James H. McKerrow, Charles S. Craik, Philip J. Rosenthal, and Linda S. Brinen

From the 1Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158, the 2Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-2280, the 3Department of Pathology and Sandler Center for Basic Research in Parasitic Diseases, University of California, San Francisco, California 94143, and the 4Department 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.

Vinyl sulfones have many desirable attributes, including selectivity for cysteine proteases over serine proteases, stability inactivity for cysteine proteases over serine proteases, stability inacti-

Vinyl sulfones have many desirable attributes, including selectivity for cysteine proteases over serine proteases, stability inactivity for cysteine proteases over serine proteases, stability inacti-

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 antityranosomal 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) 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 knock-out (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
Vinyl Sulfone Inhibition of Parasite Proteases

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 CruzainK11777 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 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 using vacuum dialysis in preparation for crystallization.

Crystallization and Structure Determination of the CruzainK11777 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 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).

Data collection

| Data collection | CruzainK11777 | RhodesainK11777 | FP-3K11017 |
|-----------------|--------------|----------------|------------|
| Space group     | C2           | P1             | P4,2,2     |
| Cell dimensions |              |                |            |
| \(a, b, c (\text{Å})\) | 134.3, 38.0, 95.2 | 34.4, 39.7, 39.6 | 114.0, 226.1 |
| Resolution (Å)  | 1.95 (2.02-1.95) | 1.65 (1.70-1.65) | 2.42 (2.55-2.42) |
| \(R_{	ext{merge}}\) (%) | 7.1 (15.3) | 4.7 (14.7) | 9.1 (37.3) |
| \(I/\sigma\)    | 26 (9.3) | 14.1 (3.9) | 21.8 (7.2) |
|Completeness (%)  | 97.7 (90.0) | 89.7 (82.7) | 100 (100) |
|Redundancy       | 3.6 (3.4) | 1.8 (1.3) | 14.2 (14.6) |

Refinement

| Refinement | CruzainK11777 | RhodesainK11777 | FP-3K11017 |
|------------|--------------|----------------|------------|
| \(R_{	ext{merge}}/R_{	ext{free}}\) (%) | 20.7/15.7 | 17.5/13.5 | 20.9/17.5 |
|Average B-factor (Å²) | 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\(^a\) | Favored (%) | 97.4 | 95.8 | 97.2 |
| Allowed (%) | 100 | 99.5 | 99.8 |
| outliers (%) | 0 | 0.5 | 0.2 |
|PDB ID | 2OZ2 | 2P7U | 3BWK |

\(^a\) \(R_{	ext{merge}} = \Sigma f(h) - \Sigma f(h)\), where \(f(h)\) is the measured diffraction intensity, and the summation includes all observations.

\(^b\) As defined by Molprobity (45).
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 published protocols (38). The activity of FP-3 was tested with a high salt buffer (20 mM Bis-Tris, pH 6.5). Inhibited FP-3 was determined by molecular replacement in PHASER (43) using the FP-3 component of the FP-3-K11017 complex crystallized with a single complete copy of the mature enzyme in the asymmetric unit (residues 1–215) in the asymmetric unit and was determined to 2.42 Å. This structure in the asymmetric unit and was determined to 2.42 Å. This structure was refined to 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 - DF, 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.

**Vinyl Sulfone Inhibition of Parasite Proteases**

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 Rfactor of 20.7% and an Rfree 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 Rfree of 17.5% and an Rfactor 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 Rfree of 20.9% and an Rfactor of 17.5%. All four copies of the complex are very similar (supplemental Table S1) and superimposition matches, on average, 234 α-carbons with an 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-
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 $k_{oi}$, 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).

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 $k_{oi}$ 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
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 superposition 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) 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-

TABLE 2
Vinyl sulfone second order inhibition constants

|       | Cruzain, $k_a$ | Rhodesain, $k_{max}/k_i$ | FP-3, $k_a$ |
|-------|---------------|--------------------------|-------------|
| K11017 | 676,000       | 264,000                  | 8,800       |
| K11777 | 517,000       | 150,000                  | 1,050       |

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.

7 I. D. Kerr, P. Wu, R. Marion-Tsukemaki, Z. B. Mackey, and L. S. Brinen, manuscript in preparation.
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 χ1 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 “gate-keeper” 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, L. 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., Djimde`e, A. A., and Plowe, C. V. (2007) Parasitol Res 96, 90–106
4. Engle, J. C., Doyle, P. S., and McKerrow, J. H. (1999) J. Cell Sci. 112, 2171–2175
5. Sugiyama, S., Kornfeld, R., and Gething, P. W. (1990) J. Biol. Chem. 265, 11931–11937
6. Rosenthal, P. J. (2000) J. Biol. Chem. 275, 29132–29135
7. Yagin, A., and Teplyakov, A. (1997) J. Appl. Crystallogr. 30, 1022–1025
8. Ackermann, R. J., and Rich, S. M. (1997) Acta Crystallogr. D Biol. Crystallogr. 53, 240–255
9. Navaza, J. (1994) Acta Crystallogr. A 50, 157–163
10. Siwai, F., Gomaa, A., Sirikul, V., and McKerrow, J. H. (2000) Bioorg. Med. Chem. Lett. 10, 597–606
11. J. Biol. Chem. 277, 12824–12832
12. J. Biol. Chem. 276, 28385–28394
13. J. Biol. Chem. 283, 18430–18437
14. J. Biol. Chem. 281, 12824–12832
15. J. Biol. Chem. 281, 12824–12832
16. J. Biol. Chem. 279, 29494–29502
17. J. Biol. Chem. 278, 28415–28422