Structure
Article

The Structure of Chagasin in Complex with a Cysteine Protease Clarifies the Binding Mode and Evolution of an Inhibitor Family

Stephanie X. Wang,[1] Kailash C. Pandey,[2] Julio Scharfstein,[6] James Whisstock,[7] Rick K. Huang,[8] Jordan Jacobelli,[1] Robert J. Fletterick,[3] Philip J. Rosenthal,[2] Magnus Abrahamson,[9] Linda S. Brinen,[4] Andrea Rossi,[5] Andrej Sali,[5] and James H. McKerrow[1,*]
[1] Department of Pathology
[2] Department of Medicine
[3] Department of Biochemistry and Biophysics
[4] Department of Cell and Molecular Pharmacology
[5] Department of Biopharmaceutical Sciences
University of California, San Francisco, San Francisco, CA 94143, USA
[6] Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, CEP 21990-400 Rio de Janeiro, Brazil
[7] Department of Biochemistry and Molecular Biology, Victorian Bioinformatics Consortium, School of Biomedical Sciences, Monash University Clayton Campus, Wellington Road, Clayton, Victoria 3800, Australia
[8] Molecular and Cell Biology Undergraduate Program, University of California, Berkeley, Berkeley, CA 94720, USA
[9] Department of Clinical Chemistry, Institute of Laboratory Medicine, University of Lund, University Hospital, S-221 85 Lund, Sweden
*Correspondence: jmck@cgl.ucsf.edu
DOI 10.1016/j.str.2007.03.012

SUMMARY

Protein inhibitors of proteolytic enzymes regulate proteolysis and prevent the pathological effects of excess endogenous or exogenous proteases. Cysteine proteases are a large family of enzymes found throughout the plant and animal kingdoms. Disturbance of the equilibrium between cysteine proteases and natural inhibitors is a key event in the pathogenesis of cancer, rheumatoid arthritis, osteoporosis, and emphysema. A family (I42) of cysteine protease inhibitors (http://merops.sanger.ac.uk) was discovered in protozoan parasites and recently found widely distributed in prokaryotes and eukaryotes. We report the 2.2 Å crystal structure of the signature member of the I42 family, chagasin, in complex with a cysteine protease. Chagasin has a unique variant of the immunoglobulin fold with homology to human CD8α. Interactions of chagasin with a target protease are reminiscent of the cystatin family inhibitors. Protein inhibitors of cysteine proteases may have evolved more than once on nonhomologous scaffolds.

INTRODUCTION

Cysteine proteases are a large and diverse family of enzymes found throughout the plant and animal kingdoms, and represent the dominant protease family in invertebrates. Disturbance of the equilibrium between cysteine proteases and their natural inhibitors is a key event in the pathogenesis of cancer, rheumatoid arthritis, osteoporosis, and emphysema (Turk et al., 2003; Riese and Chapman, 2000). Chagasin is a protease inhibitor that was first identified in *Trypanosoma cruzi* as the physiological regulator of cruzain (also known as cruzipain), the major protease of this protozoan parasite (Monteiro et al., 2001; Rigden et al., 2002; Sanderson et al., 2003). Cruzain is a papain-like (Clan CA) cysteine protease that is expressed in all stages of the parasite life cycle. It is a key virulence factor of *T. cruzi*, the infectious agent responsible for the leading cause of heart disease in Latin America, Chagas disease (Scharfstein et al., 1986; Engel et al., 1998). Chagasin is associated with cruzain during its trafficking to specific compartments of the parasite cell, and accumulated evidence suggests that the primary role of chagasin is in posttranslational regulation of protease activity (Monteiro et al., 2001). Following the discovery of chagasin, homologous proteins were identified in numerous other eukaryotic and prokaryotic organisms (Rigden et al., 2002; Sanderson et al., 2003). In many cases, these related protease inhibitors likely regulate cysteine proteases produced by their cognate organism. However, in organisms such as the pathogenic bacterium *Pseudomonas aeruginosa*, a chagasin-like inhibitor is present but no gene encoding a cysteine protease target has been identified in the genome (Sanderson et al., 2003). An alternative function for this family of protease inhibitors was therefore proposed: inhibiting the activity of host cysteine proteases elaborated as part of the host defense against pathogens (Sanderson et al., 2003). The amino acid sequence of chagasin provides few clues to its function, as chagasin and other I42 family inhibitors share no sequence homology with any known protease inhibitors (Rigden et al., 2002).

Recently, NMR solution structures were solved for chagasin (Salmon et al., 2006) and the homolog of chagasin in

Table 1. Inhibition Constants for Chagasin versus Clan CA (Papain) Cysteine Proteases

Protease	Cystatin	Chagasin	Leupeptin
	K_i (nM)	K_i (nM)	K_i (nM)
Falcipain 2	6.5 ± 1.4	1.7 ± 0.53	0.20 ± 0.11
Falcipain 3	100 ± 8.6	0.62 ± 0.27	0.30 ± 0.09
Cathepsin B	101 ± 7.7	100 ± 9.5	0.37 ± 0.14
Cathepsin L	11.5 ± 3.6	0.35 ± 0.10	0.52 ± 0.22
Cathepsin K	25.4 ± 3.0	2.0 ± 0.28	0.64 ± 0.14
Cathepsin H	0.63 ± 0.24	15 ± 4.8	3.2 ± 1.4

Leishmania mexicana (Smith et al., 2006). Predictions were made about the possible mode of interaction between members of this new family of inhibitors and their protease targets. But without a protease-inhibitor complex, these remained speculative predictions. Solving of an X-ray structure of chagasin in complex with a target cysteine protease now allows clarification of the binding mode of this novel inhibitor.

RESULTS

Confirmation that Recombinant Chagasin Inhibits Target Proteases and Purification of the Chagasin-Falcipain 2 Complex

Before embarking on crystallization and structure analysis of the complex between chagasin and falcipain 2 (FP2), we confirmed that chagasin was indeed a tight-binding inhibitor of this parasite protease as well as homologous cathepsin L-like proteases (Table 1). Chagasin was, in fact, a more potent inhibitor than the cystatins against five cathepsin L-like proteases. As was the case with cystatin, the K_i versus cathepsin B was substantially higher due to the impediment of the occluding loop of cathepsin B to binding of protein inhibitors versus the small peptide inhibitor leupeptin. Having confirmed tight binding of chagasin to FP2, we then confirmed purification of that complex for crystallography and subsequent structural analysis (Figure 1).

General Characteristics of the Chagasin-Falcipain 2 Structure: Chagasin Is a New Variant of the Immunoglobulin Fold

The coordinates and structure of chagasin have been deposited in the Protein Data Bank (PDB) under ID code 2OUL. The X-ray structure of chagasin is found to differ from the structures of all known classes of protease inhibitors, including cystatin, staphostatin, and p41 (Bode et al., 1988; Dubin et al., 2003; Guncar et al., 1999). Chagasin adopts an immunoglobulin (Ig)-like β sandwich structure (Figure 2B). Most Ig-like β sandwich domains are categorized into four subgroups: *c*onstant-type, *v*ariable-type, switched-type, and *h*ybrid-type, as defined by the topological arrangement of the strands in the front and back β sheets (Bork et al., 1994). The hypervariable loops connecting the β strands are the complementarity determining regions (CDRs) that contribute to the versatility of Ig-like folds in numerous protein-protein interactions (Garcia et al., 1998, 1999). In chagasin, the CDR equivalent loops

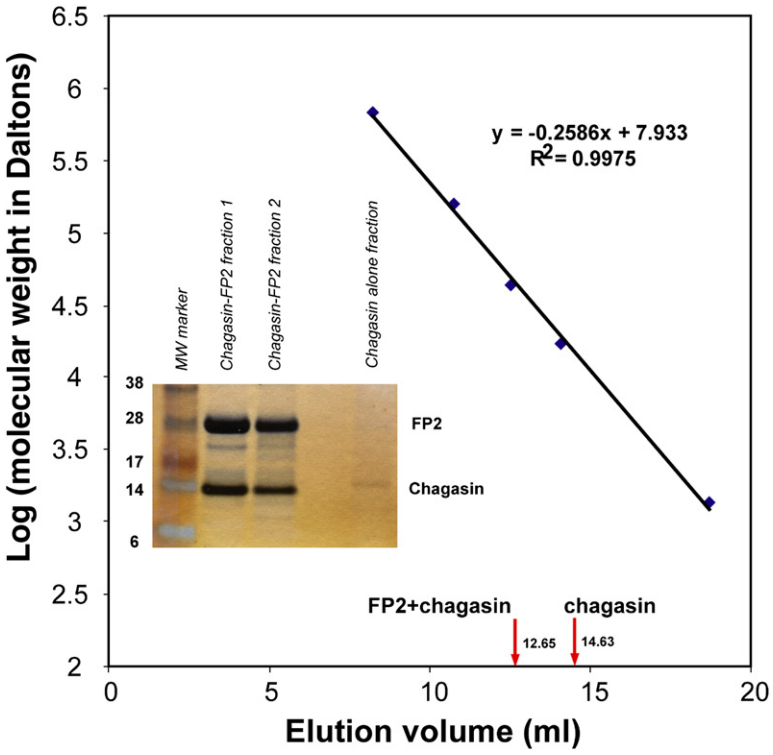

Figure 1. Stoichiometry Analysis of Chagasin-FP2 by Size-Exclusion Chromatography

The standard curve based on molecular weight standards and the elution volume for the chagasin-FP2 complex and chagasin alone are highlighted. The peak fraction containing the putative complex was analyzed by SDS-PAGE followed by silver staining. The apparent molecular weights for the chagasin-FP2 complex and chagasin alone were calculated as 46 kDa and 14 kDa, respectively. The molecular weight for FP2 is 28 kDa.

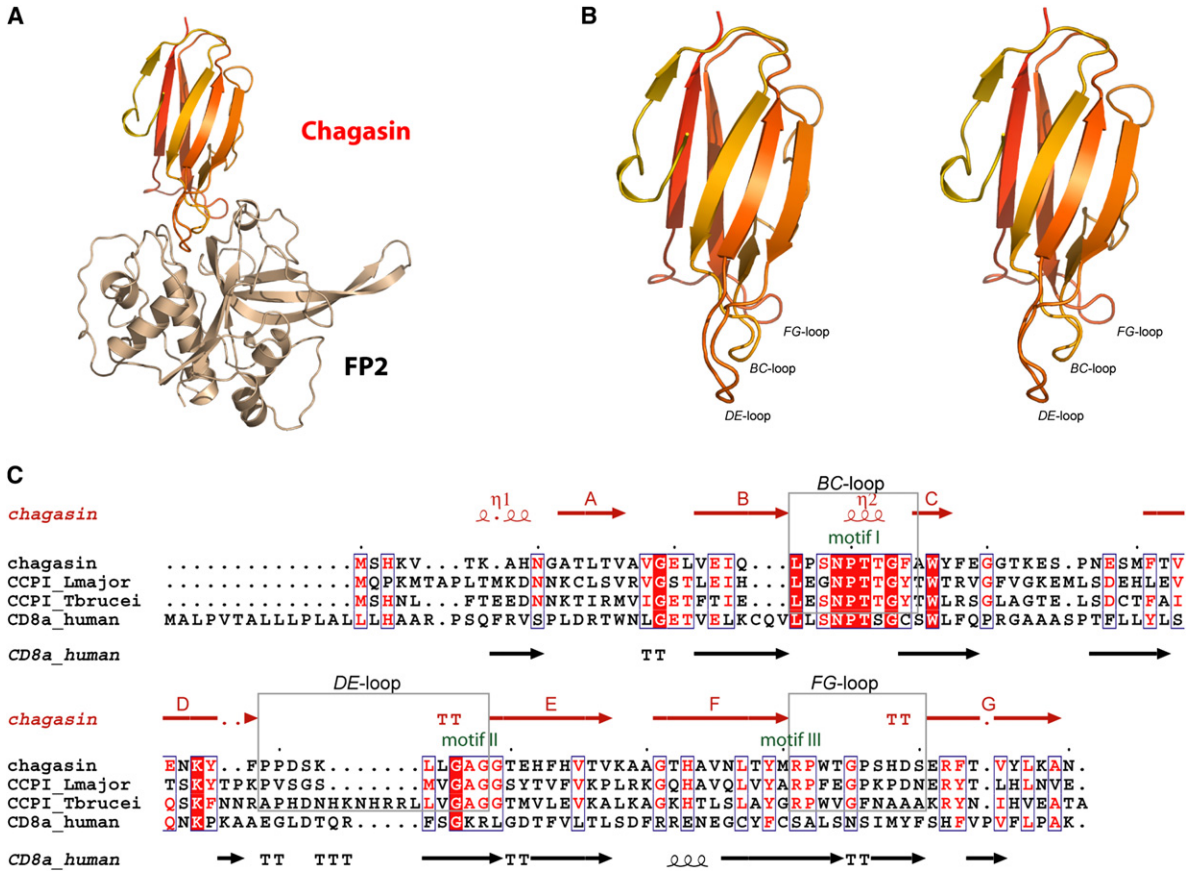

Figure 2. The Structure of Chagasin Bound to a Cysteine Protease Target, FP2

(A) Overall structure of chagasin-FP2 is shown with chagasin in red and FP2 in blue.

(B) Chagasin is shown in stereo view in the same orientation as in (A).

(C) Sequence and secondary structure alignment of chagasin and chagasin-like protease inhibitors (CCPI) from *T. brucei* and *L. major* versus human CD8α. The secondary structure distributions for chagasin and CD8α are depicted with the sequence alignment.

are the *BC*, *DE*, and *FG* loops (Figures 3B and 3C). These loops approximately correspond to the three most conserved motifs defined in the I42 family of inhibitors (Rigden et al., 2002), which indicates that the inhibition mechanism detailed in the current structure is likely to apply to the entire family of inhibitors.

The immunoglobulin fold (i.e., the Ig-like fold) derives its name from the core structure of the immunoglobulins themselves, but it is also found in other immune system proteins including interleukin 1, CD4, CD1, and CD8. The elucidation of the structure of chagasin confirms that this backbone has utility for protein-protein interactions that is much broader. A query on DBAli (Marti-Renom et al., 2001) (http://salilab.org/DBAli/), with no restriction on sequence identity, shows that there are 284 chains in the PDB with more than 50% equivalent positions to the structure of chagasin. Among the proteins identified with high similarity to chagasin, 11 of the 31 are immune system proteins. A structural core of four conserved β strands, *B*, *C*, *E*, and *F*, constitute one of the signature features of an Ig-like fold. This conserved core is found in all subtypes of Ig domains, but on first analysis it appeared to be absent in chagasin because the N-terminal strand *A* in chagasin is positioned differently from known Ig-like domains. If strand *A* were switched from its traditional front-sheet position to the back sheet, the topology arrangement of chagasin would be identical to that of a typical *c*-type Ig domain (Figure 3A). We therefore classify the novel topology arrangement in chagasin as an *N-sc*-type Ig domain (ami*N*o-terminal *s*witched *c*onstant-type).

Comparison of Chagasin X-Ray and Solution Structures

The Ig-like β sandwich structure of chagasin largely validates previous modeling predictions (Rigden et al., 2001) and two solution structures (Salmon et al., 2006; Smith et al., 2006). However, differences exist at the N terminus and among the CDR-like loops. Sequence alignment suggests that other chagasin-like protease inhibitors also adopt similar Ig-like folds, and this is confirmed at least for the NMR solution structure for *L. mexicana* ICP (Salmon et al., 2006). Structural variations are expected due to sequence inserts between strands *D* and *E* (Rigden

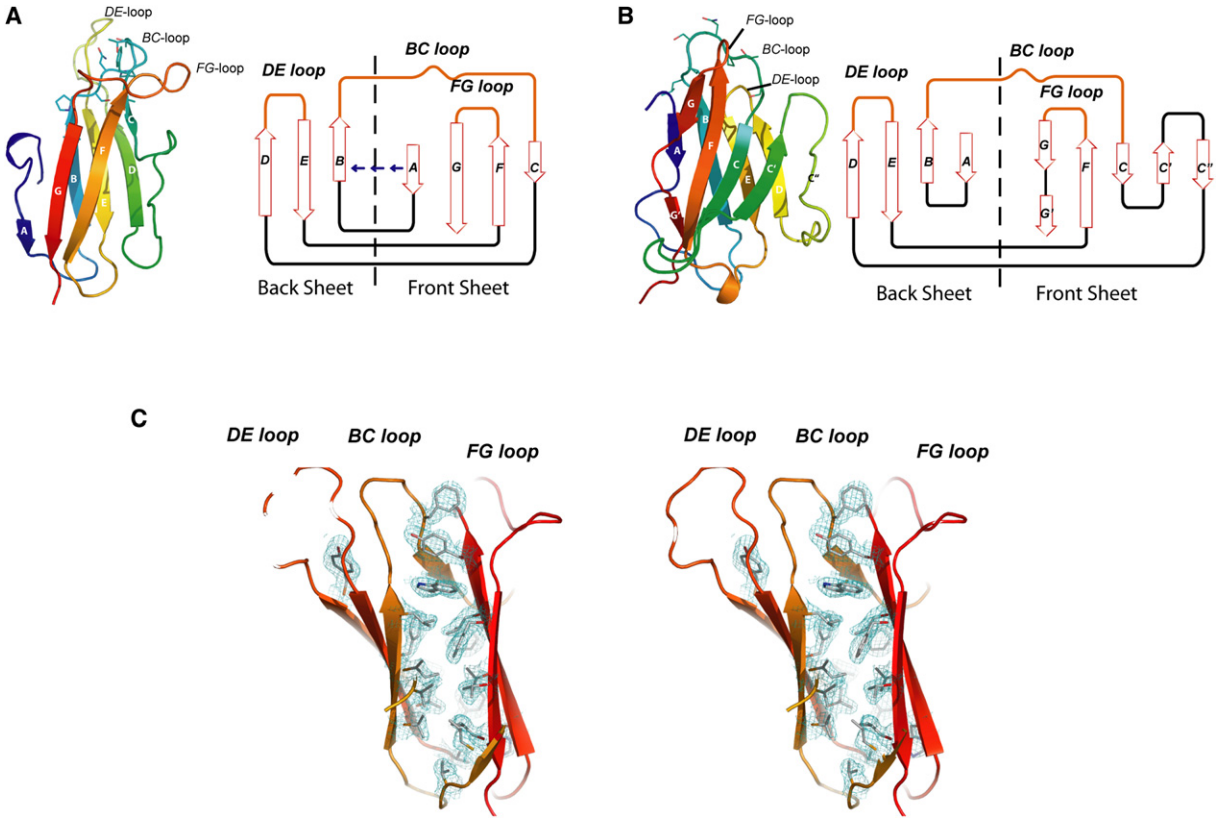

Figure 3. Comparison of Chagasin and CD8α

(A and B) Depiction of the topology arrangement of chagasin and CD8α, respectively. Both molecules are in a rainbow color scheme to reveal their β strands, with N termini in dark blue and C termini in bright red. Both molecules are oriented to reveal the conserved F, C, B, and E core strands. All structure figures were generated with PyMOL (http://www.pymol.org/; Delano Scientific LLC, South San Francisco, CA).

(C) The hydrophobic core of the chagasin structure is depicted. Selected residues involved in hydrophobic packing are shown with side chains which are highlighted by electron density at 1 (σ) from the 3fo2fc omit map.

et al., 2002). For example, the chagasin ortholog from *Plasmodium falciparum* contains an insert of 50 amino acids between its motifs I and II that may introduce significant structural deviation, or even include an additional domain.

Homology of Chagasin and CD8α

The chagasin X-ray structure confirms homology to the X-ray structure of the T cell (thymus-derived lymphocyte) surface protein CD8α (Leahy et al., 1992), a key element in antigen activation of T cells by antigen presenting cells. The *v*-type Ig-like domain in CD8α has a topology arrangement very similar to that of chagasin (Figures 3A and 3B). Like CD8α, the β strands in chagasin are held together by strong hydrophobic interactions. Residues such as Val, Phe, Trp, Leu, and Ile constitute a solid hydrophobic core that likely ensures the stability of chagasin in the absence of any disulfide bond (Figure 3C). More importantly, the *BC* loop in chagasin (LPSNPTTGFAW) closely resembles the sequence of one of the CDR loops in human CD8α (LLSNPTSGCSW) (Figure 2C). What is most remarkable is that, to date, the sequence motif found in the binding loop of CD8α and chagasin has not been found in any of the other immunoglobulin superfamily members, including the 284 chains in the PDB with at least 50% of their residues structurally equivalent to those in chagasin.

An analysis of the binding of the chagasin *BC* loop to a target protease, compared to the binding of CD8α to a major histocompatibility complex (MHC), shows that four of the seven conserved residues play very similar roles (protein-protein interactions, flexibility or stabilization of the loop, and stabilization of the core). It is noteworthy that this short yet important sequence homology could not be detected using sensitive search engines such as PSI-BLAST but only became apparent after the structural homology was revealed. The *BC* loop is one of the three signature motifs of the I42 family of inhibitors, and it at least partially accounts for the activity of the inhibitor against target proteases (Figure 3C). A synthetic version of the *BC* loop peptide in the chagasin-like protein from *Entamoeba histolytica* specifically blocked the activity of cysteine proteases that prefer Phe at the P1 site (Riekenberg et al., 2005). It is also noteworthy that Thr31 in the *BC* loop of chagasin binds to the catalytic Cys at the FP2 active site through water-mediated hydrogen bonds (Figure 4A). This highly conserved Thr likely serves as

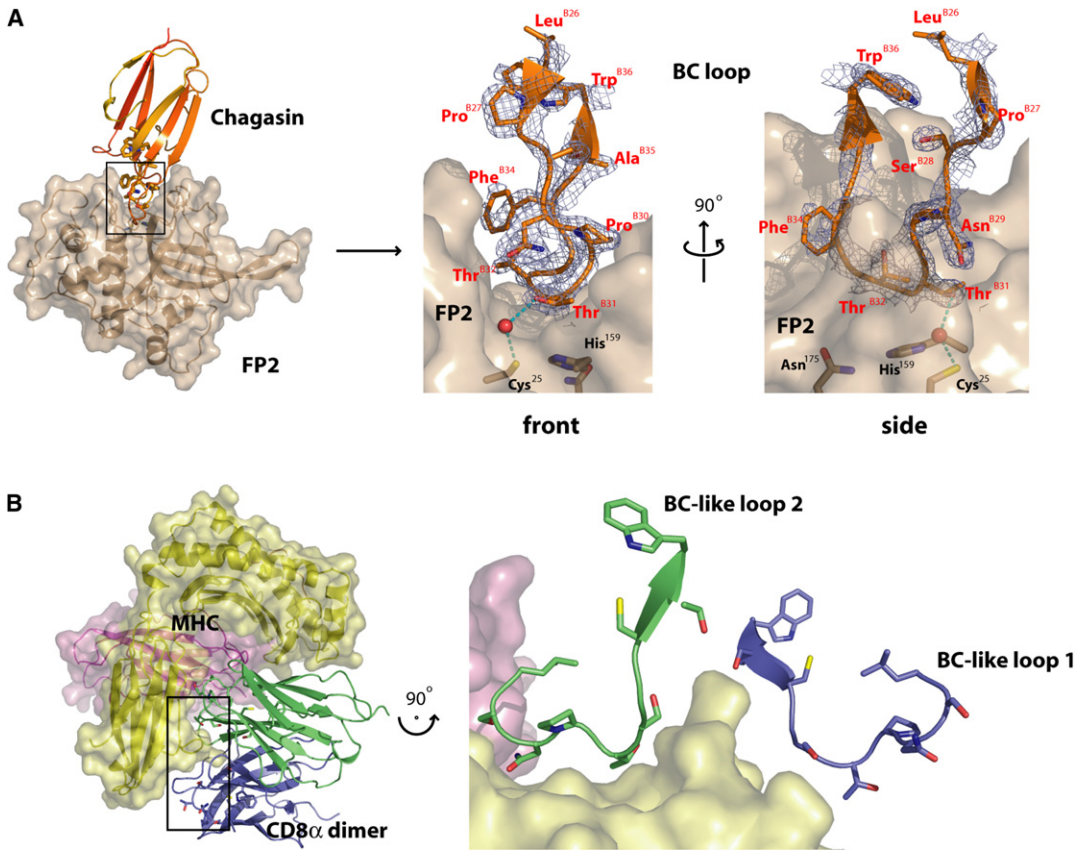

Figure 4. Key Binding Interactions between Chagasin and a Target Protease Compared to CD8α and HLA

Binding interactions between (A) the *BC* loop and FP2 are compared with those between (B) human CD8α and HLA (Protein Data Bank ID code 1AKJ). The protease and the MHC are rendered by surface presentation, while the *BC* loop in chagasin and the equivalent in CD8α are highlighted to reveal their conformational differences. The *BC* loop in the side view (A) and the *BC*-like loop 1 in (B) are oriented similarly.

a key functional residue among all chagasin-like inhibitors (identified to date) (Figure 2C).

How Chagasin Binds to a Target Protease

Models of the binding loops and predictions of interaction of these loops with target proteases were made based on two solution structures, one of chagasin itself (Salmon et al., 2006) and the other of the homolog from *L. mexicana* (Smith et al., 2006). As expected from the observed flexibility of these loops in solution, the modeled structures differ from those determined here by X-ray analysis. Nevertheless, the *BC* loop, the highly mobile *DE* loop, and the RPW/F motif in the *FG* loop are all confirmed as key elements for binding. The previous report of the lack of inhibitor activity by a mutant form of the *L. mexicana* ICP highlights the importance of the GXG motif in the *DE* loop (Smith et al., 2006). However, as shown in the X-ray analysis, a wedge-like interaction with the target by all three loops is key to binding and, by inference, full inhibition activity of chagasin.

Despite its clear homology to the LLSNPTSGCSW motif in human CD8α, the *BC* loop in chagasin appears to function via a different mechanism from what has been established for the CDR loops. Chagasin demonstrates a strict 1:1 binding to a target protease, FP2, in structure (Figure 4A) and in gel-filtration analysis (Figure 1). Both human and mouse CD8α form dimers that bind MHCs (Garcia et al., 1996; Gao et al., 1997; Liu et al., 2003) (Figure 4B). In CD8α, these interactions rely upon the CDR loops and include the *BC* loop motif from each CD8α monomer (Gao et al., 1997; Liu et al., 2003). By definition, the CDR loops adopt multiple conformations to complement antigen binding; thus, it is no surprise that the *BC* loop in chagasin and the equivalent loops in human and mouse CD8α have different structural features (Figures 4A and 4B). Mouse CD8α, at concentrations up to 1 mM, did not inhibit the proteases cruzain or FP2 (data not shown). Conversely, chagasin does not compete with binding of CD8α to an MHC by biacore analysis (H. Cheroutre, personal communication) or in a T cell activation assay.

DISCUSSION

The structure of chagasin was determined bound to the cathepsin L-like cysteine protease FP2 (Figure 2A). The original FP2 structure was determined in a cystatin-FP2

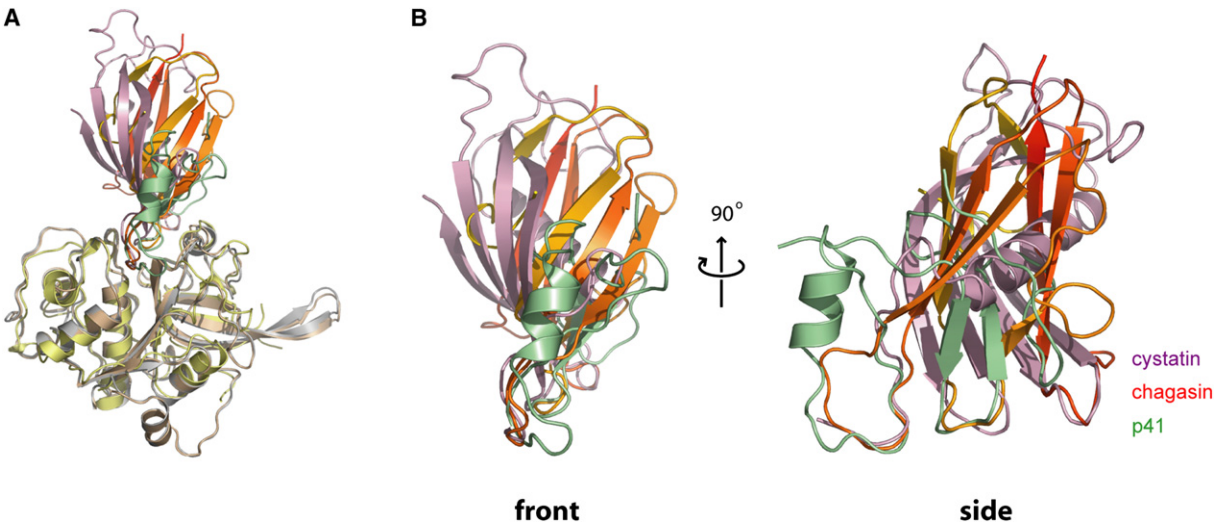

Figure 5. Comparison of Chagasin Target Interactions to that of a Cystatin Inhibitor and the p41 Inhibitory Fragment
(A) The complexes of p41-catL (Protein Data Bank ID code 1ICF; p41 in green and catL in light yellow), cystatin-FP2 (Protein Data Bank ID code 1YVB; cystatin in violet and FP2 in gray), and chagasin-FP2 (chagasin in a color gradient from yellow at the N terminus to dark red at the C terminus and FP2 in wheat) are superimposed by the 110 most conserved residues in the protease domains.
(B) Front and side views of the superimposed inhibitors.

complex (Wang et al., 2006). Overall, the structural difference between chagasin-bound FP2 and cystatin-bound FP2 is minimal, with a root-mean-square deviation (rmsd) around 0.5 Å when compared by all 241 Cα atoms, suggesting that both inhibitors interact with a consistently folded protease.

The protease binding loops (*BC*, *DE*, and *FG*) in chagasin form a well-aligned wedge that fills the active site groove of target cysteine proteases to obstruct substrate binding. The tripartite binding in chagasin is reminiscent of that found in the cystatins (Bode et al., 1988; Wang et al., 2006). Similar interactions are found between human cathepsin L and the MHC class II-associated p41 Ii fragment (Guncar et al., 1999). Kinetic analysis confirms the advantage of tripartite binding. Chagasin inhibits both papain and cruzain with picomolar affinity (Monteiro et al., 2001). It also inhibits the malarial cysteine proteases FP2 and falcipain 3 (FP3), as well as the human cathepsins B, H, K, and L (Table 1).

Chagasin, cystatins, and p41 inhibitory fragment come from protist, vertebrate, and human sources, respectively. They share no overall sequence homology and contain no conserved binding motifs. Nevertheless, they have evolved to bind to the same class of enzymes via remarkably similar binding interactions (Figures 5A and 5B). The tripartite mode of inhibition effectively blocks enzymatic activities in papain-like cysteine proteases and provides unique advantages that may have driven such evolutionary convergence. First, all tripartite binding results in very large protein-protease interaction surfaces that often correspond to strong binding energy. All known tripartite inhibitors bind to target cysteine proteases over buried surface areas larger than 2000 $Å^2$ (Figure 5A), whereas most other enzyme-inhibitor complexes only interact over surfaces averaging approximately 950 $Å^2$ in size (Jones and Thornton, 1996). Second, the three protease contact sites are discontinuous in amino acid sequence and relatively independent from each other (Figure 5B). This organization, which is reminiscent of the versatility of antibodies in antigen binding, enables chagasin to adapt to the active sites of different cysteine proteases and likely accounts for its broad inhibitory activities. Third, it is intriguing that such multiloop binding interactions at the protease active site are only found between papain-like enzymes and their inhibitors (Stubbs et al., 1990; Jenko et al., 2003; Guncar et al., 1999; Wang et al., 2006). By comparison to other protease classes, papain-like cysteine proteases harbor many characteristics that facilitate protein-protein interactions, including an accessible active site, a relatively flat substrate binding cleft, and the predominantly hydrophobic nature of the binding surface (Jones and Thornton, 1996; DeLano et al., 2000).

The convergent evolution of chagasin, p41 inhibitory fragment, and cystatin reflects the importance of post-translational regulation of cysteine proteases in organisms as ancient as protists and as complex as primates. In both *T. cruzi* and mammalian cells, these inhibitors have a commonality of function and location in that they interact with their target proteases in protein-trafficking pathways between the Golgi (chagasin) and the lysosome (cystatins). While the chagasin family proteins commonly function to inhibit endogenous papain-like cysteine proteases (Santos et al., 2005), there are clearly examples of organisms without such enzymes in which a chagasin family inhibitor is produced (*Pseudomonas aeruginosa* and *Thermobifida fusca*, for example). In situations where a pathogen is involved, Sanderson et al. (2003) suggested that the chagasin homolog may function to protect the organism from

host proteases elaborated as part of an innate immune response or other defense mechanism. Nevertheless, there are also examples of chagasin family inhibitors being expressed by nonpathogenic organisms such as the archae *Methanosarcina acetivorans*.

Functional convergence may account for the binding of an Ig-like inhibitor such as chagasin to its target proteases. It is more challenging to explain why distant proteins such as chagasin and CD8α share structural and sequence homology. It is possible that they share a common ancestor but have diverged in function. The highly homologous loops evolved to meet different functional requirements in protease binding and antigen presentation. An alternative hypothesis has also been proposed, invoking a horizontal gene transfer mechanism from human to the parasite, but it remains speculative without significant supporting nucleotide sequence data (Rigden et al., 2001, 2002). Alternatively, the discovery of homology between chagasin and CD8α suggests that a very effective and commonly used scaffold for protein binding, the Ig-like fold, arose in ancient eukaryotes. Because of the physical principles that hold β sheets within the Ig-like fold, the overall scaffold can remain stable yet diversify in the loop structures to provide novel biologic functions with specific binding partners. In the case of CD8α and chagasin, this led to a very similar motif presented by two different loops but interacting with the target ligand using four of seven similar amino acid side-chain modes. Protein-protein interactions using this scaffold likely diverged into distinct functional entities (enzyme inhibitors versus receptor binding cofactors) prior to or concurrent with the evolution of the adaptive immune system in vertebrates.

Conclusion

In conclusion, the specific mode of binding that confers inhibitory activity to chagasin and by analogy other members of this new inhibitor family is now clarified by X-ray structure. Predictions made from solution structures and mutagenesis studies are confirmed or corrected. Two important aspects of the molecular evolution of cysteine protease inhibitors are also mirrored in the structure. First, the homology between chagasin and CD8α suggests that a very effective and commonly used scaffold for protein binding arose in agent eukaryotes, but diverged into distinct entities (inhibitors versus binding cofactors) with the evolution of the adaptive immune system in vertebrates. Second, the tripartite loop interaction demonstrated between chagasin and a target protease appears to be a particularly effective and commonly used binding mode reflected in the convergent evolution of similar target binding by chagasin, p41 inhibitory fragment, and cystatin.

EXPERIMENTAL PROCEDURES

Cloning, Expression, Purification, Gel Filtration, and Kinetic Analysis of Chagasin

Two primers (5′-CTTAAAATCGGATCCCACAAGGTGACGAAAGCCCATAAC-3′ and 5′-CCCAAGCTTGGGTCAGTTTGCCTTGAGATATACAGTGAA-3′) were used to amplify the chagasin gene from *T. cruzi* chromosomal DNA. It was subsequently inserted between the BamHI and HindIII restriction sites in the pQE30 vector (QIAGEN). Upon transformation into m15 (pREP4) cells, the 12 kDa chagasin was expressed as an insoluble protein at 37°C, and was partially solubilized by slower cell growth in LB at 15°C. Soluble chagasin was then purified by an N-terminal His tag on an Ni-NTA column (QIAGEN) preequilibrated with 25 mM imidazole, 50 mM phosphate (pH 7.0), and 200 mM NaCl. Pure chagasin was eluted with a 100 mM–1 M imidazole gradient in the presence of equilibrium buffer.

The apparent K_i of chagasin against FP2, FP3, and cathepsin B, H, K, and L was determined by standard protocols (Wang et al., 2006).

Purification and Crystallization of the Chagasin-FP2 Complex

Purified and activated FP2 was incubated with recombinant chagasin at 4°C and further purified as previously described (Wang et al., 2006). The purified chagasin-FP2 complex was dialyzed into 50 mM phosphate buffer (pH 7.0) and 200 mM NaCl and concentrated to 6 mg/ml. The complex was subsequently crystallized by the sitting drop vapor diffusion method by mixing 0.8 ml of chagasin-FP2 with 0.8 ml of well solution at 4°C in the presence of 20% (v/v) PEG300, 0.1 M Tris (pH 8.5), 5% (w/v) PEG8000, and 10%–15% (v/v) glycerol. Molecular weight standards (Bio-Rad) and a preincubated mixture of chagasin-FP2 were analyzed by size-exclusion chromatography on a Superose 12 column (Pharmacia) in 50 mM Bis-Tris (pH 5.6) with 200 mM NaCl at a flow rate of 0.5 ml/min.

Data Collection and Structure Determination of the Chagasin-FP2 Complex

Crystals of chagasin-FP2 were directly flash-cooled in liquid N_2 for cryoprotection against synchrotron radiation. Complete data sets of chagasin-FP2 were collected at 100K on beamline 8.3.1 at the Advanced Light Source (ALS), using a CCD camera. The data sets were indexed and integrated using DENZO/SCALEPACK (Otwinowski and Minor, 1997) to yield a data set 96.7% complete at 2.2 Å (R_{merge} = 10.8%). The chagasin-FP2 crystals belong to the space group $P4_32_12$, with cell dimensions a = b = 94.236 Å, c = 119.764 Å, and $\alpha = \beta = \gamma = 90°$. Each asymmetric unit contains only one chagasin-FP2 complex. Data up to 3.5 Å were included in the molecular replacement search with an FP2 model (PDB ID code 1YVB), using the rotational and translational functions from the Crystallography & NMR System software suites (CNS 1.1) (Brunger et al., 1998). Following a rigid body refinement, the FP2 model gave an initial $R_{working}$ of 42.9%. Chagasin was built into the extra electron densities by de novo building using QUANTA 2000 (Molecular Simulations). The chagasin-FP2 structure was refined by alternate cycles of energy minimization, simulated annealing, and group B factor refinements in the CNS suites. Model building and fitting were done using QUANTA 2000. The final chagasin-FP2 structure has been refined to 2.2 Å. Complete data and refinement statistics are listed in Table 2.

Sequence and Structure Homology Analyses

Various segments of the chagasin amino acid sequence were used to search against the general nonredundant database as well as the PDB database using the NCBI BLAST site for possible protein homologs. Results from the first search include similar protease inhibitors from other parasitic organisms, for example *T. brucei* and *P. falciparum*. Although a search against the PDB database identified scores of proteins from the immune system, regardless of which segment was used to search the PDB database, CD8α was always identified as one of the top hits. Alignment of the chagasin and CD8α sequences revealed that CD8α harbors a short sequence motif that is highly similar to motif I in chagasin.

The structure of chagasin was searched against the DABli database (http://www.salilab.org/DBAli/) to identify all known structures with significant structural similarity to chagasin (i.e., with p value > 8.0). DBAli contains 1,086,905,585 pairwise structural alignments and family-based multiple structure alignments for 26,950 nonredundant chains in the PDB. Three of the six matches are classified as immunoglobulins in SCOP (http://scop.mrc-lmb.cam.ac.uk/scop/; including

Table 2. Diffraction Data Processing and Refinement Statistics

	Chagasin-FP2
Structure	
Space group	$P4_32_12$
Cell parameters (Å)	
a = b	94.236
c	119.764
$\alpha = \beta = \gamma$ (°)	90
Data resolution (Å)	50–2.2 (2.28–2.2)[a]
Data Processing	
Total reflections	699,566
Unique reflections	27,127
Redundancy	9.9 (2.5)[a]
Completeness (%)	96.7 (80.2)[a]
I/σI	12.6
R_{merge} (%)[b]	10.8 (37.8)[a]
Refinement	
$R_{working}$ (%)[c]	22.3
R_{free} (%)[d]	24.0
Average B factor (Å^2)	24.1
Rmsd bond (Å)	0.009
Rmsd angle (°)	1.6

[a] Statistics for the highest-resolution shell.
[b] $R_{merge} = \sum|(I - <I>)|/\sum(I)$.
[c] $R = \sum_{h,k,l}(|F_{obs}(h,k,l)| - k|F_{calc}(h,k,l)|)/\sum_{h,k,l}|F_{obs}(h,k,l)|$.
[d] R_{free}: crossvalidation R calculated by omitting 5% of the reflections (Kleywegt and Brunger, 1996).

the highest significance hit): one belongs to the PaPD superfamily of the Ig-like fold, and the remaining two are classified as non-Ig-like but β sheet-containing folds. Therefore, the available evidence suggests that chagasin belongs to the immunoglobulin superfamily with a very ancient divergence.

T Cell Cytotoxicity Assay

Cytotoxic activity was measured in a standard ^{51}Cr release assay. Activated CD8$^+$ T cells from OT-I transgenic mice were used as effectors, generated as in Krummel et al. (1999), and LB27.4 B cells were used as targets for cytotoxicity. Target B cells were labeled with 50 milliCi of ^{51}Cr per 1 × 10^6 cells for 1 hr. One milligram/milliliter of SIINFEKL cognate ovalbumin-derived peptide (pOVA) was added to the target cells during ^{51}Cr labeling followed by extensive washing. Different concentrations of chagasin protein were diluted in the wells for the cytotoxicity assay. In addition, one group of target cells was preincubated with 50 mg/ml of chagasin protein for 1 hr. Various dilutions of OT-I effector cells were incubated with 1 × 10^4 ^{51}Cr-labeled target B cells in round-bottom 96-well plates for 4 hr at 37°C. Subsequently, 100 ml of supernatant was removed and counted on a scintillation counter to determine experimental ^{51}Cr release (ER). Spontaneous release (SR) was determined using the target cells alone, while maximum release (MR) was determined by lysing target cells with 1% SDS and measuring the amount of ^{51}Cr release. Specific lysis (in %) was calculated as follows: (ER − SR)/(MR − SR)*100. All samples were set up in triplicate.

ACKNOWLEDGMENTS

This work was supported by NIH grants AI35707 and AI35800, the Sandler Family Supporting Foundation, and the Medicines for Malaria Venture. S.X.W. was a recipient of a postdoctoral fellowship from the American Heart Association.

Received: November 21, 2006
Revised: March 1, 2007
Accepted: March 20, 2007
Published: May 15, 2007

REFERENCES

Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., and Turk, V. (1988). The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. EMBO J. *7*, 2593–2599.

Bork, P., Holm, L., and Sander, C. (1994). The immunoglobulin fold. Structural classification, sequence patterns and common core. J. Mol. Biol. *242*, 309–320.

Brunger, 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., et al. (1998). Crystallography & NMR System: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

DeLano, W.L., Ultsch, M.H., de Vos, A.M., and Wells, J.A. (2000). Convergent solutions to binding at a protein-protein interface. Science *287*, 1279–1283.

Dubin, G., Krajewski, M., Popowicz, G., Stec-Niemczyk, J., Bochtler, M., Potempa, J., Dubin, A., and Holak, T.A. (2003). A novel class of cysteine protease inhibitors: solution structure of staphostatin A from *Staphylococcus aureus*. Biochemistry *42*, 13449–13456.

Engel, J.C., Doyle, P.S., Hsieh, I., and McKerrow, J.H. (1998). Cysteine protease inhibitors cure an experimental *Trypanosoma cruzi* infection. J. Exp. Med. *188*, 725–734.

Gao, G.F., Tormo, J., Gerth, U.C., Wyer, J.R., McMichael, A.J., Stuart, D.I., Bell, J.I., Jones, E.Y., and Jakobsen, B.K. (1997). Crystal structure of the complex between human CD8α (α) and HLA-A2. Nature *387*, 630–634.

Garcia, K.C., Scott, C.A., Brunmark, A., Carbone, F.R., Peterson, P.A., Wilson, I.A., and Teyton, L. (1996). CD8 enhances formation of stable T-cell receptor/MHC class I molecule complexes. Nature *384*, 577–581.

Garcia, K.C., Degano, M., Pease, L.R., Huang, M., Peterson, P.A., Teyton, L., and Wilson, I.A. (1998). Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. Science *279*, 1166–1172.

Garcia, K.C., Teyton, L., and Wilson, I.A. (1999). Structural basis of T cell recognition. Annu. Rev. Immunol. *17*, 369–397.

Guncar, G., Pungercic, G., Klemencic, I., Turk, V., and Turk, D. (1999). Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. EMBO J. *18*, 793–803.

Jenko, S., Dolenc, I., Guncar, G., Dobersek, A., Podobnik, M., and Turk, D. (2003). Crystal structure of Stefin A in complex with cathepsin H: N-terminal residues of inhibitors can adapt to the active sites of endo- and exopeptidases. J. Mol. Biol. *326*, 875–885.

Jones, S., and Thornton, J.M. (1996). Principles of protein-protein interactions. Proc. Natl. Acad. Sci. USA *93*, 13–20.

Kleywegt, G.J., and Brunger, A.T. (1996). Checking your imagination: applications of the free R value. Structure *4*, 897–904.

Krummel, M.F., Heath, W.R., and Allison, J. (1999). Differential coupling of second signals for cytotoxicity and proliferation in CD8$^+$

T cell effectors: amplification of the lytic potential by B7. J. Immunol. *163*, 2999–3006.

Leahy, D.J., Axel, R., and Hendrickson, W.A. (1992). Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6 Å resolution. Cell *68*, 1145–1162.

Liu, Y., Xiong, Y., Naidenko, O.V., Liu, J.H., Zhang, R., Joachimiak, A., Kronenberg, M., Cheroutre, H., Reinherz, E.L., and Wang, J.H. (2003). The crystal structure of a TL/CD8$\alpha\alpha$ complex at 2.1 Å resolution: implications for modulation of T cell activation and memory. Immunity *18*, 205–215.

Marti-Renom, M.A., Ilyin, V.A., and Sali, A. (2001). DBAli: a database of protein structure alignments. Bioinformatics *17*, 746–747.

Monteiro, A.C., Abrahamson, M., Lima, A.P., Vannier-Santos, M.A., and Scharfstein, J. (2001). Identification, characterization and localization of chagasin, a tight-binding cysteine protease inhibitor in *Trypanosoma cruzi*. J. Cell Sci. *114*, 3933–3942.

Otwinowski, Z., and Minor, W. (1997). Processing of X-Ray Diffraction Data Collected in Oscillation Mode, *Volume 276* (New York: Academic Press).

Riekenberg, S., Witjes, B., Saric, M., Bruchhaus, I., and Scholze, H. (2005). Identification of EhICP1, a chagasin-like cysteine protease inhibitor of *Entamoeba histolytica*. FEBS Lett. *579*, 1573–1578.

Riese, R.J., and Chapman, H.A. (2000). Cathepsins and compartmentalization in antigen presentation. Curr. Opin. Immunol. *12*, 107–113.

Rigden, D.J., Monteiro, A.C., and Grossi de Sa, M.F. (2001). The protease inhibitor chagasin of *Trypanosoma cruzi* adopts an immunoglobulin-type fold and may have arisen by horizontal gene transfer. FEBS Lett. *504*, 41–44.

Rigden, D.J., Mosolov, V.V., and Galperin, M.Y. (2002). Sequence conservation in the chagasin family suggests a common trend in cysteine proteinase binding by unrelated protein inhibitors. Protein Sci. *11*, 1971–1977.

Salmon, D., do Aido-Machado, R., Diehl, A., Leidert, M., Schmetzer, O., de A. Lima, A.P.C., Scharfstein, J., Oschkinat, H., and Pires, J.R. (2006). Solution structure and backbone dynamics of the *Trypanosoma cruzi* cysteine protease inhibitor chagasin. J. Mol. Biol. *357*, 1511–1521.

Sanderson, S.J., Westrop, G.D., Scharfstein, J., Mottram, J.C., and Coombs, G.H. (2003). Functional conservation of a natural cysteine peptidase inhibitor in protozoan and bacterial pathogens. FEBS Lett. *542*, 12–16.

Santos, C.C., Sant'anna, C., Terres, A., Cunha-e-Silva, N.L., Scharfstein, J., and de A. Lima, A.P.C. (2005). Chagasin, the endogenous cysteine-protease inhibitor of *Trypanosoma cruzi*, modulates parasite differentiation and invasion of mammalian cells. J. Cell Sci. *118*, 901–915.

Scharfstein, J., Schechter, M., Senna, M., Peralta, J.M., Mendonca-Previato, L., and Miles, M.A. (1986). *Trypanosoma cruzi*: characterization and isolation of a 57/51,000 m.w. surface glycoprotein (GP57/51) expressed by epimastigotes and bloodstream trypomastigotes. J. Immunol. *137*, 1336–1341.

Smith, B.O., Picken, N.C., Westrop, G.D., Bromek, K., Mottram, J.C., and Coombs, G.H. (2006). The structure of *Leishmania mexicana* ICP provides evidence for convergent evolution of cysteine peptidase inhibitors. J. Biol. Chem. *281*, 5821–5828.

Stubbs, M.T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., and Turk, V. (1990). The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. EMBO J. *9*, 1939–1947.

Turk, D., Turk, B., and Turk, V. (2003). Papain-like lysosomal cysteine proteases and their inhibitors: drug discovery targets? Biochem. Soc. Symp. (70), 15–30.

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). Structural basis for unique mechanisms of folding and hemoglobin binding by a malarial protease. Proc. Natl. Acad. Sci. USA *103*, 11503–11508.

Accession Numbers

The coordinates and structure of chagasin have been deposited in the Protein Data Bank under ID code 2OUL.