

Development of α -keto-based inhibitors of cruzain, a cysteine protease implicated in Chagas disease

Youngchool Choe,^{a,†} Linda S. Brinen,^{b,†} Mark S. Price,^a Juan C. Engel,^c Meinolf Lange,^d Corinna Grisostomi,^d Scott G. Weston,^d Peter V. Pallai,^d Hong Cheng,^d Larry W. Hardy,^d David S. Hartsough,^d Marsha McMakin,^d Robert F. Tilton,^d Carmen M. Baldino^d and Charles S. Craik^{a,*}

^aDepartment of Pharmaceutical Chemistry, 600 16th Street, University of California at San Francisco, San Francisco, CA 94143-2280, USA

^bDepartment of Cellular and Molecular Pharmacology, University of California at San Francisco, 513 Parnassus Ave, San Francisco, CA 94143-0511, USA

^cDepartment of Pathology, School of Medicine, 513 Parnassus Avenue, University of California at San Francisco, San Francisco, CA 94143-0102, USA

^dDepartments of Biology, Chemistry, and Computational Design, ArQule, Inc., 19 Presidential Way, Woburn, MA 01801, USA

Received 17 November 2004; accepted 30 December 2004

Available online 2 February 2005

Abstract—*Trypanosoma cruzi*, a protozoan parasite, is the causative agent of Chagas disease, a major cause of cardiovascular disease in many Latin American countries. There is an urgent need to develop an improved therapy due to the toxicity of existing drugs and emerging drug resistance. Cruzain, the primary cysteine protease of *T. cruzi*, is essential for the survival of the parasite in host cells and therefore is an important target for the development of inhibitors as potential therapeutics. A novel series of α -ketoamide-, α -ketoacid-, α -ketoester-, and aldehyde-based inhibitors of cruzain has been developed. The inhibitors were identified by screening protease targeted small molecule libraries and systematically optimizing the P1, P2, P3, and P1' residues using specific structure-guided methods. A total of 20 compounds displayed picomolar potency in in vitro assays and three inhibitors representing different α -keto-based inhibitor scaffolds demonstrated anti-trypanosomal activity in cell culture. A 2.3 Å crystallographic structure of cruzain bound with one of the α -ketoester analogs is also reported. The structure and kinetic assay data illustrate the covalent binding, reversible inhibition mechanism of the inhibitor. Information on the compounds reported here will be useful in the development of new lead compounds as potential therapeutic agents for the treatment of Chagas disease and as biological probes to study the role that cruzain plays in the pathology. This study also demonstrates the validity of structure-guided approaches to focused library design and lead compound optimization.

© 2005 Elsevier Ltd. All rights reserved.

1. Introduction

Chagas disease or American trypanosomiasis, is one of the most threatening endemics in Central and South America. Approximately 16–18 million people are infected, resulting in adverse health events such as heart failure and more than 50,000 deaths each year.¹ It is

thought that another 100 million people are at risk of infection.¹ Due to the toxicity of existing drugs combined with the advent of drug resistance, treatment for this debilitating, often chronic illness is woefully inadequate.¹ *Trypanosoma cruzi*, a parasitic protozoan, is the causative agent of the disease. Its infectious trypomastigote form is transmitted to human hosts through the triatomine 'kissing bug' vectors. Trypomastigotes transform into the intracellular amastigotes after invading cardiac muscle cells through the bloodstream. Amastigotes complete the infectious life cycle by multiplying in the cell, transforming back to infectious trypomastigotes and rupturing the cell to release an

Keywords: Chagas disease; Cruzain; Cysteine protease inhibitors; Structure-based library design.

* Corresponding author. Tel.: +1 415 476 8146; fax: +1 415 502 8298; e-mail: craik@cgl.ucsf.edu

[†] These authors contributed equally to this work.

amplified quantity of the infectious form back into the bloodstream.²

Cruzain, the primary cysteine protease of *T. cruzi*, is expressed throughout the life cycle and is known to be essential for the survival of the parasite within host cells.³ In addition to its obvious role in parasite nutrition, cruzain has been proposed to be involved in the penetration of the parasite into the host cell and in digestion of immunoglobulins as a defense mechanism. Some inhibitors of cruzain have been shown to successfully treat animal models of Chagas disease by blocking the parasitic life cycle.⁴ Thus, cruzain presents itself as an interesting target for the development of potential therapeutics for the treatment of the disease. Cruzain was initially discovered from cell-free extracts,⁵ and subsequently heterologously expressed in *Escherichia coli*.⁶ The X-ray crystallographic structures of a recombinant truncated form of the enzyme in complex with several small molecule inhibitors have been determined.⁷ The overall folding pattern and the arrangement of the active site residues are similar to those in papain. As in most papain family proteases, the S2 subsite of the enzyme is known to be the predominating specificity determinant. Cruzain contains a unique C-terminal extension of 130 amino acids, which is retained in native forms. The enzyme is inhibited by organomercurial reagents, E-64, Tos-Lys-CH₂Cl, leupeptin, a number of peptidyl chloromethanes, and peptidyl fluoromethane derivatives, vinyl sulfones, thio semicarbazones, cystatins, stefins, and kininogens.^{2,4} Covalent inhibitors such as peptidyl epoxy ketones,⁸ aldehydes,⁹ and vinyl sulfones,⁹ have also been reported to inhibit cruzain. However, published information on cruzain inhibitory activity from α -keto based compounds has been very limited.^{10,11}

Due to its essential function in the parasite's life cycle and therefore in the propagation of Chagas disease, cruzain was elected as a target to develop inhibitors through in vitro screening of a series of protease specific small molecule libraries and structure-guided optimization. The libraries were designed to initially focus screening on known pharmacologically active and target class specific small molecules.¹⁰ Additional efficiencies could also be realized from this approach due to the comprehensive understanding of the automated process chemistry.^{12,13} The in vitro screening of these libraries for cruzain inhibitory activity led to the identification of novel α -ketoamide-based cruzain inhibitors. The screening also facilitated the exploration of closely related α -ketoamide-, α -ketoacid-, α -ketoester-, and aldehyde-based analogs. Identification of an inhibitor in the arrayed analog libraries provides rapid access to a database of structure–activity relationships (SAR). This information is then used to guide the systematic optimization of each inhibitor scaffold. Since the papain family cysteine proteases including cruzain (clan CA) share a common three-dimensional structure and very similar biochemical characteristics, this approach and the inhibitor scaffolds developed for cruzain can be readily applied to other physiologically important clan CA proteases.

2. Results

2.1. Identification of α -ketoamide-based inhibitors of cruzain

The in vitro screening of protease targeted libraries for inhibitory activity against cruzain identified novel peptidyl α -ketoamide-based compounds as covalent inhibitors of the enzyme (Fig. 1). Further exploration of this structural motif led to the synthesis of a small set of novel α -ketoester-, α -ketoacid-, and α -ketoamide-based compounds (Table 1). A modified Dakin–West reaction was used for the preparation of the compounds.¹⁴ It should be noted that this synthetic approach produces a mixture of diastereomers at the P1 α -carbon atom. The vinyl-based α -keto ester series was obtained by coupling of the peptide aldehyde (benzyloxycarbonyl (Cbz)-Phe-Ala-CHO) with triphenylphosphorane (Scheme 1). An examination of the SAR trends of this set revealed the need for the presence of an amide hydrogen atom. Kinetic analysis of these compounds demonstrated a competitive, reversible inhibition of cruzain by the members of this set (Fig. 2). In order to more effectively explore the P1' region of the enzyme, the α -ketoamides were selected as the chemical scaffold for further inhibitor optimization efforts. However, aldehyde-based compounds, which are easier to prepare than the analogous α -ketoamides were often used as the initial probes to establish inhibitory trends against the enzyme.

2.2. Optimization of the α -ketoamide scaffold at the P1 subunit

The optimization began with an analysis of the P1 subunit using peptidyl aldehydes, which were synthesized following the Reetz procedure,¹⁵ as surrogates for the α -ketoamide series. Various natural and unnatural amino acids including those lacking a chiral center were systematically incorporated at the P1 position. The SAR trends observed for this set clearly point to the need for a chiral center of the *S*-configuration (or L-amino acids) at the P1 and a preference for β -branched alkyl substituents at this position (Table 2). The structural reason underlying the observed SAR trends is not clear,

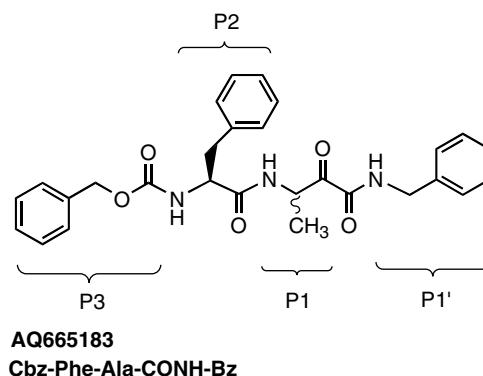
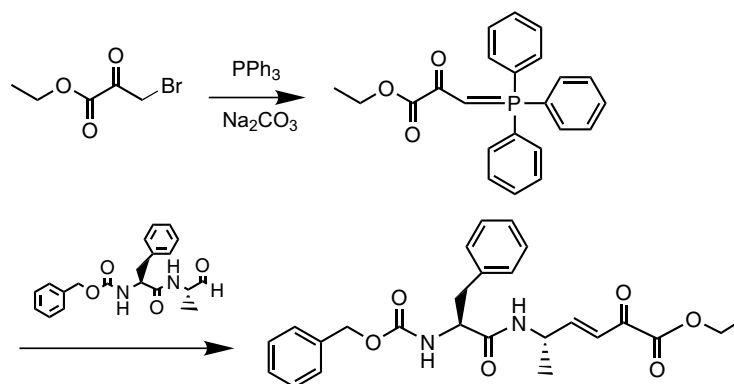


Figure 1. Chemical structure of AQ665183, an α -ketoamide identified from the initial screening. The P3, P2, P1, and P1' portions of the compound are denoted.

Table 1. Inhibition of cruzain by the initial set of α -ketoester, α -ketoamide, α -ketoacid compounds

	Compound	Cruzain inhibition	
		IC ₅₀ (nM)	K _i (nM)
	AQ616475	1000	ND
	AQ665184	39	13.8
	AQ616476	7.7	0.83
	AQ665183	5.5	3.4
	AQ713492	>5000	ND
	AQ581332	4.7	1102.9

ND = not determined.

**Scheme 1.** Synthesis of the vinyl-based α -keto ester.

as there is no clearly defined S1 pocket in published cruzain crystallographic structures.^{7,16} Larger R groups such as the phenylethane side chain of homophenyl-

alanine have also been used at this position in other published series,⁹ but structural studies revealed no constructive interaction of these groups with the cruzain S1

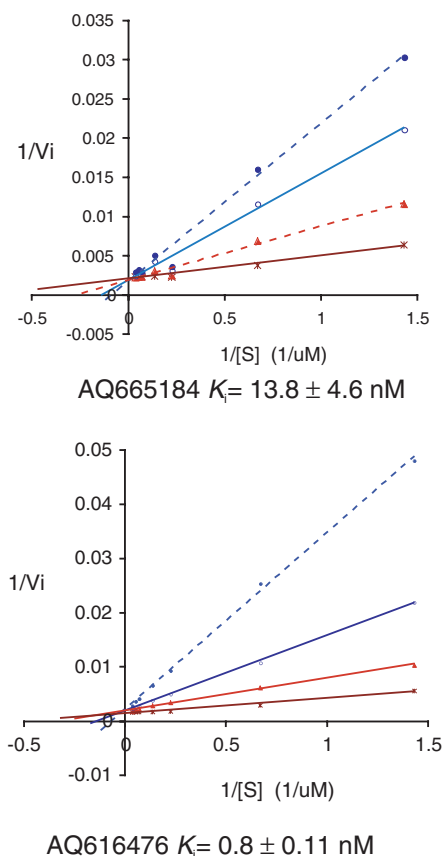


Figure 2. Lineweaver–Burk plots for the selected compounds.

subsite.⁷ Based upon the trends observed in the peptidyl aldehyde series, L-valine was selected as the preferred P1 residue. The corresponding α -ketoamide compound (AQ797526) was synthesized using Wassermann's procedure¹⁷ and found to be a subnanomolar inhibitor of cruzain while compound AQ752489 demonstrated the α -keto moiety is required for the activity (Table 3).

2.3. Optimization at the P2 subunit

The optimization of the P2 subunit was performed using the α -ketoamide series. A new synthetic scheme based on Wassermann's chemistry was employed in order to prepare the Cbz-Phe-Val- α -ketoesters with a desired chirality (S/L) at the P1 carbon (Scheme 2). All the compounds from this series have L-valine at the P1 position, a Cbz group at the P3 position, and an *N*-benzyl group at the P1' position as AQ797526 (Table 4). The SAR trends observed for this set correlate with the reported large and predominantly hydrophobic nature of the S2 pocket in cruzain.^{7,16} These results are also consistent with published data obtained using other series of cruzain inhibitors such as conformationally constrained γ -lactams containing an electrophilic aldehyde, vinyl sulfones, or fluoromethyl ketones.^{9,16} L-Phenylalanine was identified as the favored residue for the P2 subunit, followed by smaller hydrophobic amino acids such as L-leucine, L-valine, and L-methionine. Ultimately, L-phenylalanine was selected as the preferred P2 residue.

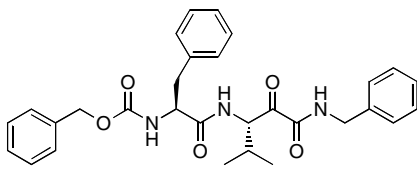
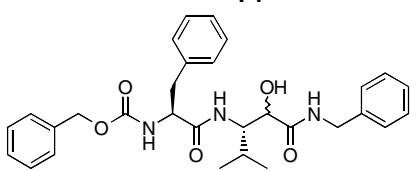
Table 2. Optimization of the P1 component in aldehyde series

Compound	IC ₅₀ (nM)
<p>P1</p>	7.2
	175
	5000
	854
	0.466
	50
	0.49

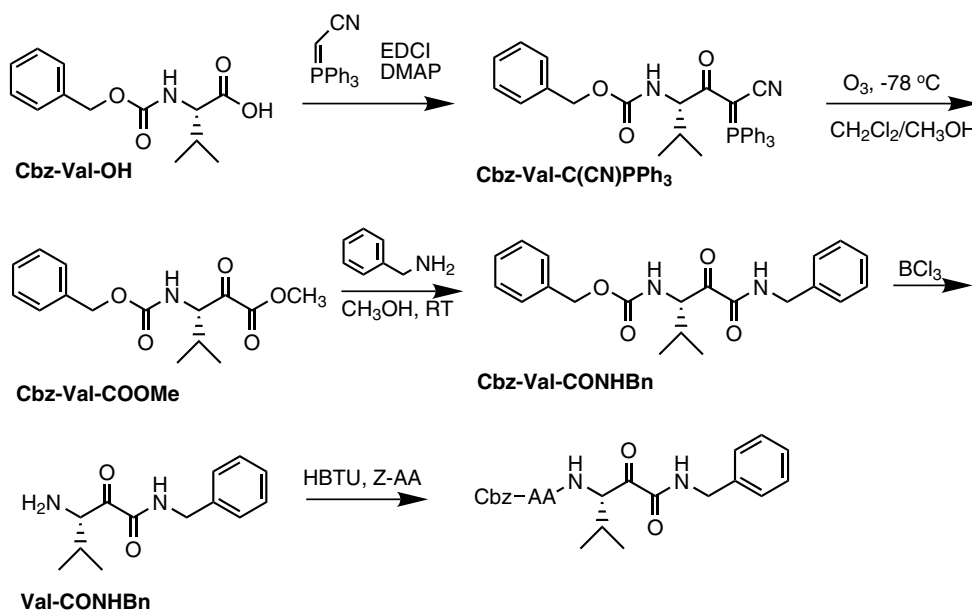
2.4. Optimization at the P3 and the P1' subunits

Evaluation of the P3 subunit quickly indicated that the removal of the Cbz-group caused dramatic loss of activity while the conversion to an acetyl group provided a similar but less pronounced impact. No further optimization at the P3 position was performed and Cbz was selected as the preferred P3 group for the remainder of the study. To optimize the P1' side chain of the α -keto-benzamide series, the P1, P2, and P3 positions were held constant with L-valine, L-phenylalanine, and Cbz, respectively. Given the hundreds of commercially available benzylamines, a structure-guided approach was

Table 3. Inhibition data for the amide-based compounds with L-valine at the P1 position

Compound	Cruzain inhibition		
	IC ₅₀ (nM)	K _i (nM)	
<div> P1</div>	AQ797526	0.797	0.878
<div> P1</div>	AQ752489	5000	ND

ND = not determined.

**Scheme 2.** Synthesis of the α -ketobenzamide library with variation at the P2 position.

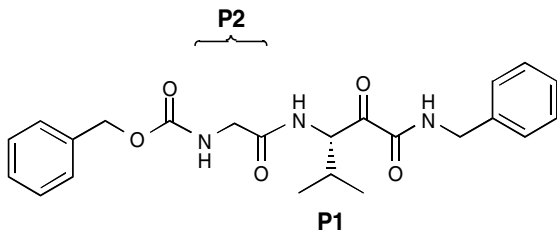
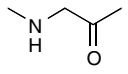
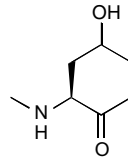
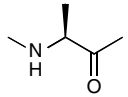
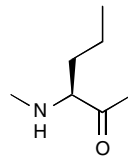
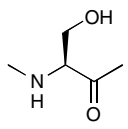
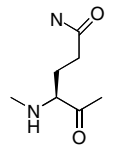
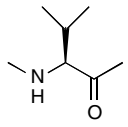
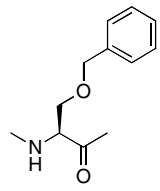
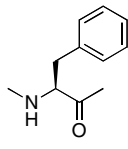
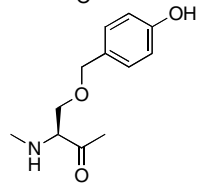
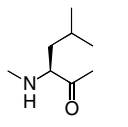
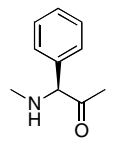
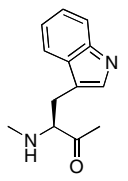
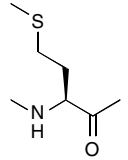
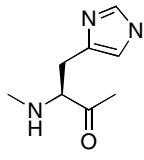
adopted to select a smaller, focused set of these reagents. A structural model of a covalently bound inhibitor (AQ665183) to cruzain was built and used to bias the reagent selection process using molecular docking methods (Fig. 3).⁷ Reagent rankings obtained by this molecular docking approach suggested that *p*-substituted phenyl rings should be favored over *m*- and *o*-substituted analogs (Table 5). Docking studies also suggested that α -substituted benzylamines would not fit into the active site as well as the unsubstituted analogs. A Cbz-Phe-Val- α -ketobenzamide library with variation at the P1' position was prepared (Scheme 3). The prediction based on structural modeling was consistent with the experimentally determined activity trends (Table 6). Both the observed trends and the modeling studies for the α -ketoamide series suggest that the primary structural justification for the increased affinity of

compounds containing a benzyl amine group is the aromatic/hydrophobic interaction with the indole side-chain of the nearby Trp177 residue (Fig. 3). This is consistent with observed SAR trends and structural studies of other cysteine protease inhibitor series such as vinyl sulfones.^{18,19}

2.5. Cell culture test for anti-trypanosomal activity

The α -ketoacid (AQ616476), α -ketoamide (AQ665183), and vinyl α -ketoester (AQ581332), all shown to be covalent binding, reversible inhibitors from the previously described series, were tested in cell culture test using *T. cruzi*-infected cells. At 30 μM , all three inhibitors demonstrated trypanostatic activity (Fig. 4). The difference in concentrations required to show activity in vitro (picomolar to nanomolar) and those needed for activity in cell culture (micromolar) can be explained, at least in

Table 4. Optimization of the P2 position in α -ketoamide series

<div style="text-align: center;">  </div>					
P2	Compound	Cruzain IC ₅₀ (nM)	P2	Compound	Cruzain IC ₅₀ (nM)
	ML-cruzain-1	10,000		ML-cruzain-10	>1000
	ML-cruzain-2	>1000		ML-cruzain-13	10
	ML-cruzain-3	>1000		ML-cruzain-16	525
	ML-cruzain-5	13		ML-cruzain-18	52
	ML-cruzain-6	3		ML-cruzain-19	2
	ML-cruzain-7	7		ML-cruzain-20	37
	ML-cruzain-8	11		ML-cruzain-21	13
	ML-cruzain-9	237			

part, by permeability issues arising from the peptidic nature of these compounds. AQ665183 (α -ketoamide) with a benzyl amide group at the P1' position displayed

better activity (66.7% longer survival than control) than AQ616476 (α -ketoacid) and AQ581332 (vinyl α -ketoester).

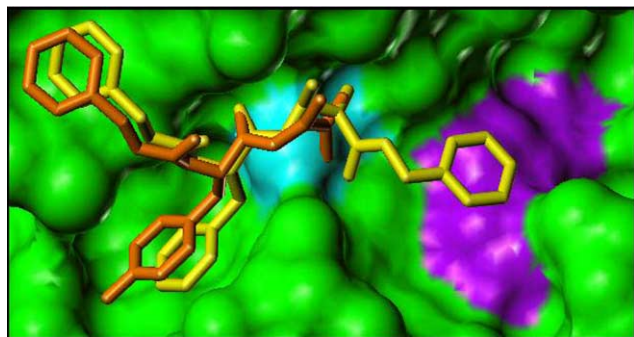
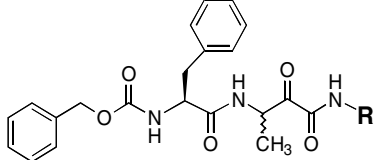
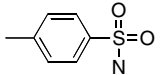
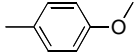
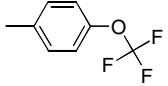

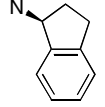
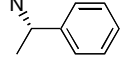
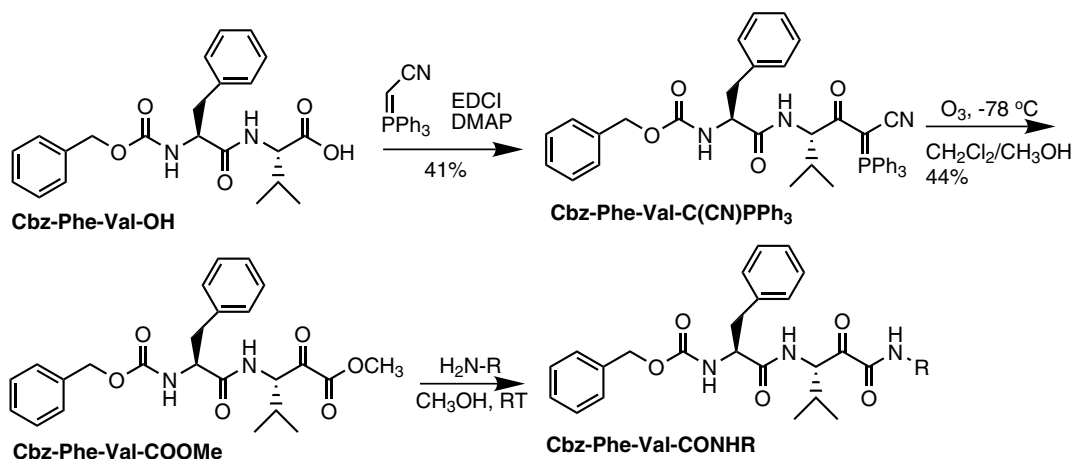


Figure 3. Comparison of the model structure of the AQ665183–cruzain complex with the crystallographic Cbz-Tyr-Ala-fluoromethylketone (FMK)–cruzain complex from the 1AIM structure. The solvent-accessible surface of cruzain (1AIM) is shown in green, the crystallographic structure of the FMK-based cruzain inhibitor from 1AIM is shown in orange, the molecular surface associated with the catalytic Cys25 residue is shown in cyan, the molecular surface associated with the nearby Trp177 residue is shown in violet, and the modeled α -ketoamide (Cbz-Phe-Ala-N-benzyl- α -ketoamide) is shown in yellow.

Table 5. Structure-guided p1' substituent selection based on DOCK prediction

		
R groups predicted to have better affinity	Parent Benzylamine (AQ665183)	R groups predicted to have worse affinity
  		 



Scheme 3. Synthetic approach to the L-valine-derived α -ketobenzamide library with variation at the P1' position.

2.6. Slow-binding kinetic analysis

The kinetic studies with the α -ketoester, AQ581332 indicated that the inhibitor is a slow-binding inhibitor of cruzain (Fig. 5). Slow-binding behavior of the inhibitor was evident at 6.25 μ M concentration of the inhibitor ($k_{\text{obs}} = 10.6 \times 10^{-3} \text{ s}^{-1}$), 5 nM concentration of cruzain, and 50 μ M concentration of substrate, Cbz-Phe-Arg-AMC. These conditions are more than 1000-fold excess of inhibitor over enzyme with saturating concentration of substrate, both requirements for monitoring the slow-binding behavior of the inhibitor.

2.7. Crystallographic data

A 2.3 Å crystallographic structure for one member of this series, AQ581332, a vinyl α -ketoester complexed to cruzain was determined. The inhibitor co-crystallized in the $P2_1$ space group with one molecule in the asymmetric unit. Diffraction was observed to a maximum reso-

lution of 2.2 Å and a complete data set was collected to 2.3 Å. Structure solution was accomplished via molecular replacement with subsequent refinement to 2.3 Å. Crystallographic and data collection parameters are summarized in Table 7. Cruzain is inhibited by AQ581332 when a covalent attachment is formed between the two molecules through a Michael addition, involving nucleophilic attack by the S_{γ} of the active site cysteine residue, Cys25, at the vinyl carbon indicated in Figure 6. Cbz is the P3 moiety of AQ581332, occupying the S3 subsite of the enzyme. The phenyl ring of the Cbz group makes weak hydrophobic contact with Leu67, 3.7 Å away while phenylalanine in the P2 subunit of the inhibitor takes advantage of cruzain's hydrophobic S2 pocket (Fig. 7). Constructive hydrophobic interactions are also made with Leu67, Ala133 and Leu157. Additionally, Glu205 at the base of the S2 pocket is bent out of the cleft and does not interfere with the aromatic ring. The ethoxy tail of the inhibitor that was expected to interact with the S1' region of cruzain¹⁸

Table 6. Inhibition of cruzain by members of P1' optimization library

P1' (R)	Compound	Cruzain IC50 (nM)	P1' (R)	Compound	Cruzain IC50 (nM)
	AQ797526	0.878		AQ903086	0.949
	AQ903084	0.644		AQ840162	0.955
	AQ840158	0.667		AQ840163	0.975
	AQ840159	0.673		AQ903087	1.115
	AQ840160	0.688		AQ840164	1.273
	AQ840161	0.720		AQ903085	1.507
	AQ903082	0.776		A840165	1.734
	AQ903091	0.794		AQ903083	2.405
	AQ903079	0.813		AQ903077	3.987
	AQ903090	0.817		AQ903088	4.529
	AQ903081	0.846		AQ840166	7.474
	AQ903080	0.860		AQ903089	10.509
	AQ903078	0.911			

was not visible in the crystal structure of the complex, which may suggest that the ethoxy group is exposed

freely to solution rather than bound to the enzyme at this site.

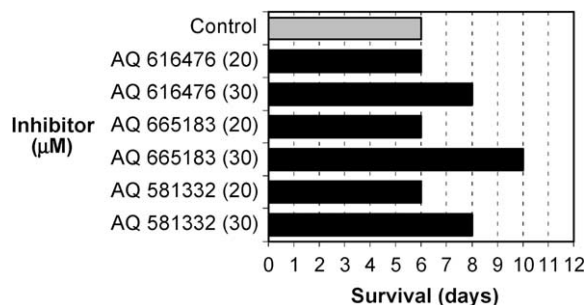


Figure 4. Cell culture test result for selected compounds.

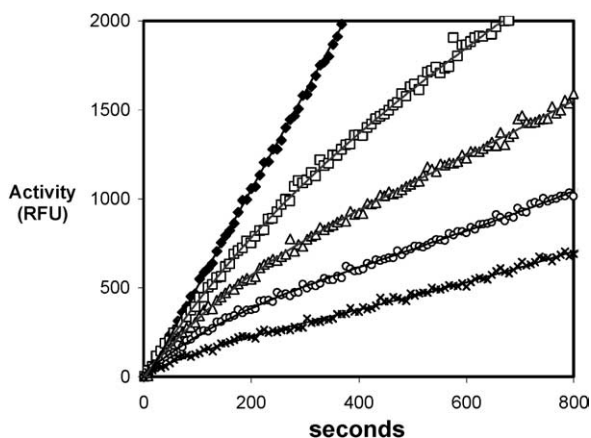


Figure 5. Progress curve of the hydrolysis of Cbz-Phe-Arg-AMC (50 μM) by cruzain displays slow-binding inhibition of cruzain by a α -ketoamide AQ581332. \blacklozenge : No inhibitor, \square : 3.125 μM, \triangle : 6.25 μM, \circ : 12.5 μM, \times : 25 μM.

Several interactions exist between the inhibitor and the main chain of cruzain. At distances of 3.4, 3.4, and 3.7 Å, the carbonyl oxygen of Gly66 contacts O1, O3, and N1 of the inhibitor, respectively. Contacts of 3.3 and 3.1 Å are also made from O1 and O3 to the peptide nitrogen of Gly66. N2 of the inhibitor generates a hydrogen bond to the carbonyl oxygen of Asp158 at a distance of 3.1 Å while an ordered water molecule situated near the active site cleft forms a 3.3 Å hydrogen bond to O2 of the inhibitor. The remaining constellation of electrostatic interactions is made between AQ581332 and side chains of amino acid residues of cruzain, as illustrated in Figure 8. Additional interactions include a weak 3.8 Å hydrogen bond between O4 and N δ 1 of His159, a second hydrogen bond between O4 and N ϵ 2 of Gln19 at a distance of 2.9 Å. Final coordinates have been deposited with the Protein Data Bank (PDB) and are accessible under the ID number 1U9Q.

3. Discussion

Herein, we have described the synthesis, identification, optimization, and crystallographically determined binding mode of a new series of α -keto-based cruzain inhibitors. As various essential roles are revealed that proteases play in both normal and pathological states, specific protease inhibitors are being developed and

Table 7. Crystallographic parameters, data collection, and refinement statistics

Crystallographic parameters	
Unit cell dimensions	45.720
a, b, c (Å) and (β)°	50.675
	45.576
	116.74
Space group	$P2_1$
Molecules per A/U	1
Data collection statistics	
Resolution limit (Å)	2.3
Number of unique reflections ($F > 0\sigma(F)$)	25,466
R_{merge} (%) ^a	9.2
Redundancy of reflections	
Overall	5.5
Highest resolution shell	2.3
$\langle I/\sigma(I) \rangle$	
Overall	8.8
Highest resolution shell	3.6
Completeness	
Overall (%)	93.3
Highest resolution shell (%)	88.2
Resolution range; highest shell (Å)	2.37–2.30
Refinement statistics	
Resolution range (Å)	40.8–2.3
R_{cryst} (%) ^b	13.3
R_{free} (%) ^c	21.8
Rmsd	
Bond lengths (Å)	0.02
Bond angles (°)	1.88
Water molecules	105
Average B factor (Å) ²	
Complex	14.12
Inhibitor	22.12

^a $R_{\text{merge}} = (\sum_h \sum_i |I_{hi} - \langle I_h \rangle|) / \sum_h I_h$, where I_h is the mean structure factor intensity of i observations of symmetry-related reflections with Bragg intensity index h .

^b $R_{\text{cryst}} = (\sum_h \sum_i ||F_{\text{obs}} - |F_{\text{calc}}||) / \sum_h F_{\text{obs}}$, where F_{obs} and F_{calc} are the observed and the calculated structure factor magnitudes.

^c $R_{\text{free}} = \sum_{(hkl)\tau} ||F_{\text{obs}(hkl)} - |F_{\text{calc}}|| / \sum_{(hkl)\tau} F_{\text{obs}(hkl)}$, where the τ set of reflections is omitted from the refinement process; 10% of the data were included in the τ set.

are proving to be promising lead compounds for drug development. The approach has been greatly augmented by the advance of structural biology, high-speed parallel synthesis, and computational chemistry. In the case of HIV-1 protease inhibitor development, which is considered a successful example, structure-based inhibitor design has focused on either transition state analogs, substrate analogs, or de novo designed compounds. Transition-state analog inhibitors are generally peptide derivatives containing an electrophilic carbonyl group in place of the scissile amide bond of the substrate.²⁰ The design of transition state analogue inhibitors of HIV-1 protease has also centered on attempts to replace the scissile amide bond with nonhydrolyzable peptide isosteres such as reduced amide, hydroxyethylamines, and statin derivatives.²¹

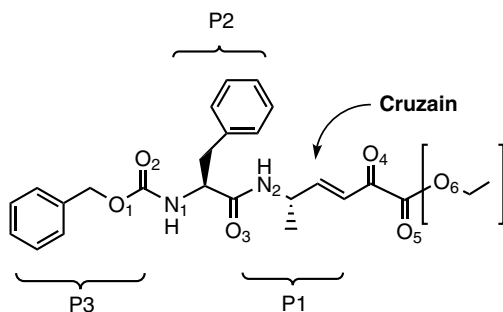


Figure 6. AQ581332. Point of attachment of cruzain is indicated. Atoms within brackets were not visible in X-ray structure.

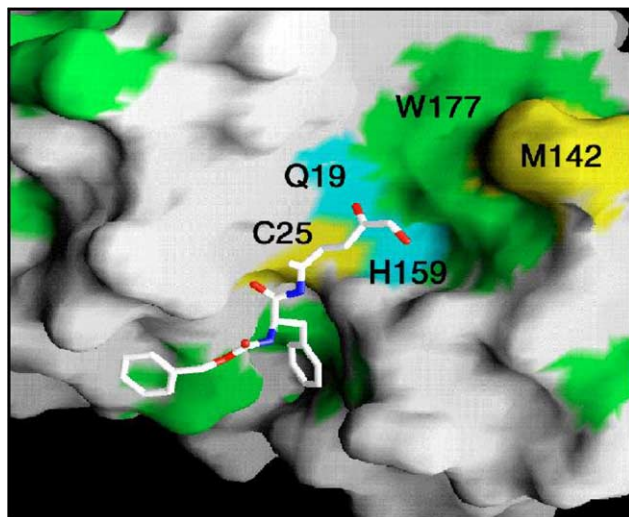


Figure 7. Surface representation of the cruzain active site region. Hydrophobic residues are indicated in green. (Note that the sides of the S2 pocket are lined with hydrophobic patches.) This figure was prepared with GRASP.

Initially, the peptidyl α -ketoester functional scaffold was reported as a potent transition state analog inhibitor of serine proteases. The design was based on the crystal structure of trypsin complexed with an α -ketoacid (4-amidinophenyl pyruvate, APPA).²² The active site serine adds to the ketone carbonyl group of APPA to give a tetrahedral adduct with the oxyanion stabilized by interacting with trypsin's oxyanion hole. In addition, cyclotheonamide A and B, the natural thrombin inhibitors isolated from a marine sponge, both contain the α -keto amide functional group.²⁰ Since peptide aldehydes, with peptide sequences based on the substrate specificity of calpain, were found to be potent reversible inhibitors of calpain I and II with a K_i as low as 36 nM,²³ various transition state inhibitors for cysteine proteases including peptide aldehydes, peptide fluoroalkyl ketones, dicarbonyl derivatives, peptide α -ketoesters, α -ketoamides, and α -ketoacids have been developed. In each case, the inhibitors are known to inhibit cysteine proteases and serine proteases by the reversible formation of a covalent tetrahedral hemithioacetal or hemiketal. They result from the attack of the sulfhydryl or hydroxy group of the active site cysteine residue or serine residue at an electrophile center of the inhibitor.²⁴ It became

apparent that reactive ketone-containing compounds could serve as inhibitors for a broad range of enzymes. Peptidyl α -ketoester inhibitors of elastase and peptidyl α -ketoamide inhibitors of calpains and other cysteine proteases such as cathepsin B and papain have been developed.²⁵ More recently, α -ketoesters and α -diketones have been shown to be very effective inhibitors of α -chymotrypsin and calpain.²⁴ A substrate-derived ketone inhibitor was also reported to strongly inhibit angiotensin converting enzyme (ACE), a metalloprotease, while efficient inhibitors of an aspartic protease, renin, were prepared by replacing the scissile amide bond of its substrate with a keto-methylene isostere.^{26,27}

The screening of protease targeted libraries coupled with the exploration of SAR and the structure-guided optimization of the lead compound series led to a novel series of potent small-molecular peptidyl inhibitors of cruzain. Initially α -ketoamide compounds were identified and subsequently various α -keto analogs such as α -ketoesters, α -ketoacids, and aldehydes with IC_{50} values in the low nanomolar range were synthesized. Through the optimization effort at each of the P1, P2, P3, and P1' subunits, L-valine, L-phenylalanine, Cbz-group, and *p*-substituted benzylamines were determined to be the optimal substituents respectively. A total of 20 compounds displayed picomolar potency in *in vitro* assays. Three inhibitors that were tested in cell culture demonstrated anti-trypansomal activity in cell culture test. It is worth noting that a major determinant of potency is the phenylalanine at P2. This is in agreement with the well known preference for phenylalanine in clan CA cysteine proteases. The S2 pocket of cruzain is lined with hydrophobic residues such as Leu67, Ala133, and Leu157 that readily accommodate the aromatic side-chain of the substrates and inhibitors.^{7,16}

Kinetics studies and the crystallographic structure of the α -ketoester, AQ581332 bound to cruzain clearly demonstrate the inhibition mechanism of this series of inhibitors. The inhibitor forms a tetrahedral enzyme-inhibitor complex that results from the nucleophilic addition of the active site cysteine sulfhydryl to the electrophilic vinyl carbon of the inhibitor. This mimics the catalytic mechanism of cysteine proteases in which the cysteine sulfhydryl attacks the amide carbonyl of the scissile bond. Also, the kinetic studies indicate AQ581332 is a slow-binding inhibitor. Slow-binding inhibitors do not instantly show a steady-state initial velocity. Instead, these inhibitors exhibit a slow onset of inhibition with time until a steady-state initial velocity is attained.²⁸ The slow-binding phenomenon could be due to a conformational change of the enzyme that is required for the optimum binding of the inhibitor and has previously been observed with other transition-state inhibitors of proteases. Examples include the binding of peptide ketoacids and ketoesters to papain and cathepsin B.²⁵

Three inhibitors representing the different inhibitor chemotypes were tested in cell culture using macrophage infected with *T. cruzi*. Each of the inhibitors that were tested exhibited a dose dependent inhibition of trypano-

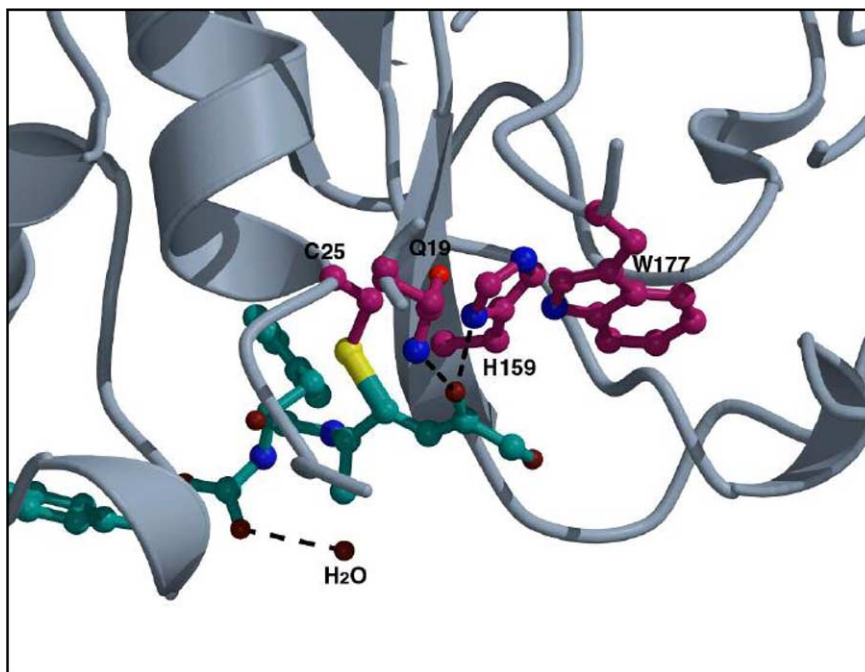


Figure 8. Cruzain active site with an inhibitor bound. Cruzain is shown in ribbon representation and the inhibitor is shown in ball-and-stick representation. Amino acid side-chains involved in binding are highlighted in magenta. Bound water molecule is displayed in red. Hydrogen bonds are indicated by dashed lines. This figure was prepared with the programs MOLSCRIPT and RASTER 3D.

somal activity. It appears the intracellular cycle of the parasite was slowed by the inhibitors and, although the exact mechanism of the action is yet to be clarified, cells survived longer. In order to effectively inhibit cruzain *in vivo*, an inhibitor must penetrate cell membranes. Among the inhibitors tested, α -ketoamide (AQ665183), which is more hydrophobic and possibly more membrane permeable displayed stronger trypanostatic activity than α -ketoester (AQ581332) or α -keto acid (AQ616476), which has poor membrane permeability as illustrated in a platelet membrane permeability assay.^{14,28} It is not uncommon that potent inhibitors from *in vitro* screening prove to be ineffective against intracellular proteases due to poor membrane permeability. To enhance membrane permeability, further optimization of the inhibitors will be necessary. This effort will focus on the preparation of more neutral, hydrophobic compounds. In addition, it has been observed that the ester bond of α -ketoester tends to be rapidly metabolized *in vivo*.²⁰ Therefore, further optimization to improve the specificity and potency of inhibitors may focus on mainly peptidyl α -ketoamides for better membrane permeability. However, the hydrophobicity of the inhibitors will also have to be carefully balanced with aqueous solubility to achieve a potential therapeutic agent.

4. Conclusion

A novel series of peptide based α -ketoamide, α -ketoester, α -ketoacid, and aldehyde analogs has been developed and shown to be potent inhibitors of cruzain, a major cysteine protease involved in Chagas disease. To aid the development and the optimization of the inhibitors at the P1, P2, P3, and P1' positions, a specific

SAR has been established using high speed chemistry, three-dimensional enzyme structure guided modeling, and docking approaches. A total of 20 compounds displayed picomolar potency in *in vitro* assays and three inhibitors representing different α -keto-based inhibitor chemotypes demonstrated anti-trypanosomal activity in cell culture test, which supports the validity and usefulness of the approach. A 2.3 Å crystal structure of cruzain covalently bound to one of the compounds is also reported. The compounds are potent transition state analog inhibitors, which reversibly inhibit cruzain by covalently binding to the sulfhydryl moiety of active site cysteine. These potent and novel inhibitors of cruzain are expected to provide a promising entry point for the development of new anti-trypanosomal chemotherapy and useful probes for the assessment of the enzyme's role in Chagas disease. So far, various chemical backbones such as fluoromethyl ketones, vinyl sulfones, thio semicarbazones have been developed as cruzain inhibitors. The potent α -keto based inhibitors which have been developed from this study expand the diversity of chemical scaffolds available for developing cruzain inhibitors and may be useful to overcome ever increasing drug resistance of *T. cruzi*.

5. Experimental

5.1. Materials

Cruzain was expressed, purified as previously described and stored in 20 mM bis-tris buffer pH 5.8 at 4 °C.^{6,16,29,30} All chemicals for buffer preparation were purchased from Sigma unless described otherwise. The substrate, Cbz-Phe-Arg-AMC, was purchased from

molecular probes and 10 mM stock solution in neat DMSO was stored at -20°C .

5.2. In vitro cruzain activity assay

A 96-well formatted fluorescent assay was used to monitor the time course of the activity assay. For K_m determination, initial velocities at a total of 11 substrate concentrations in the range from 0.15 to 30 μM were measured. The stock solutions were prepared in 100% DMSO, and 5 μL of each were added to the reaction mixture to produce a final mixture with 5% DMSO total. In the case of initial screening experiments, a substrate concentration of 3 μM ($2 \times K_m$) was used. Briefly, inhibitors in DMSO 5 μL were added to 175 μL of substrate solution (substrate in 5 μL 100% DMSO) and assay buffer (100 mM NaOAc, pH 5.5, 10 mM EDTA, and 10 mM DTT) and the well was thoroughly mixed. The reaction was initiated by the addition of cruzain (30 μL , 1.4 nM in assay buffer, preincubated in a siliconized microcentrifuge tube on ice for at least 10 min). The final assay volume was 210 μL and the final DMSO concentration was less than 5%. The initial velocity was monitored for 5 min at 22°C on a BMG FluoStar spectrofluorimeter (BMG Labtechnologies) by following the increase in the fluorescent signal (excitation at 390 nm and emission at 460 nm). Since cruzain is highly hydrophobic like other members of the papain family, Costar NBS (nonbinding surface) plates were used or 0.005% Brij35 was included in the assay buffer in order to prevent any potential aggregation.

5.3. Inhibition assays

A solution of substrate (3 μM , 5 μL) was premixed with inhibitor (1 nM to 10 μM , 5 μL /concentration/well) and assay buffer (180 μL). The reaction was initiated by the addition of Cruzain (30 μL , 1.4 nM). Initially, cruzain IC_{50} values for all compounds were determined and then K_i was determined for the most active analogs. The K_i values were determined by measuring the initial velocity at varying substrate concentrations in the presence of different inhibitor concentrations and analyzing the result using the program developed by Cleland.³¹ A double reciprocal plot was generated in order to demonstrate the mode of inhibition. The kinetic fluorescence measurement for the slow-binding inhibitor AQ581332 was carried out twice in duplicate using a SpectraMax Gemini fluorescence spectrometer (Molecular Devices) with an excitation wavelength of 380 nm, an emission wavelength of 460 nm, and a 435 nm cutoff filter. Progress curves for the inhibition of cruzain by AQ581332 followed typical slow-binding kinetics. Resulting steady-state rates were fit by nonlinear regression to the following equation for several substrate concentrations (5, 10, 25, 50, 75, and 100 μM) using the program KALEIDAGRAPH (Synergy software) according to Eggers et al.³²

$$v_i/v_0 = 1$$

$$= \frac{[E_0] + [I_0] + K_i^* - \sqrt{([E_0] + [I_0] + K_i^*)^2 - 4[E_0][I_0]}}{2[E_0]}$$

where K_i^* is the apparent dissociation constant for the enzyme–inhibitor complex, v_0 is the control (uninhibited) velocity, v_i is the inhibited steady-state velocity, $[I_0]$ is the total inhibitor concentration, and $[E_0]$ is the total cruzain concentration. The K_i^* values for various substrate concentrations were plotted against the substrate concentration and a linear regression line displayed a y -intercept, which is an estimate of K_i of AQ581332 for cruzain.

5.4. Cell culture test

Growth inhibition of *T. cruzi* amastigotes and toxicity of the inhibitors were examined by checking the survival of a mammalian cell culture model infected by *T. cruzi* trypomastigotes. J774 mouse macrophage cell line was cultured in RPMI 1640 medium with 5% heat-inactivated FBS (RPMI medium). For growth inhibition studies, J774 macrophages were irradiated (3000 rad) to arrest the cell cycle and cultured in six-well plates for 24 h at 37°C . The cultures were infected with 10^6 *T. cruzi* trypomastigotes/well of the Y strain for three hours. Monolayers were washed with RPMI medium and treated with the inhibitors at 20–30 μM in RPMI medium (triplicate wells per inhibitor). Inhibitor stocks were made at 100 mM in DMSO and control without inhibitors included DMSO (0.01–0.02% v/v). Trypomastigotes output, an indicative of the completion of the parasite intracellular cycle, was counted in inhibitor-treated and untreated cultures to measure the growth inhibition of intracellular amastigote stage of the parasite by the inhibitors. The culture medium containing inhibitor or DMSO as the control was replaced daily and inhibitor efficacy was monitored daily. Survival time is defined as the time before the cell monolayer is destroyed by the infection. Untreated cells were completely destroyed after 6 days in this assay indicating the parasite completed the intracellular cycle and released from the host cells at the end of the 6 day cycle.³³

5.5. Modeling

The modeling initially began with the published structure of cruzain complexed to Cbz-Tyr-Ala-FMK (PDB code: 1AIM).¹⁶ Other cysteinyl protease-covalent inhibitor structures, such as a cathepsin K-hydrazide complex (PDB code: 1AYW), were used to guide the development of a ‘prime–nonprime’ spanning model for α -ketoamide binding to cruzain. Additional parameters for inter- and intra-atomic torsional, angle, and distance constraints for the bound α -ketoamide group were taken from serine protease- α -ketoamide complexes, such as the trypsin–cyclotheonamide A complex (PDB code: 1TYN). The structural model for the covalent bound AQ665183 to cruzain (Fig. 3) was constructed using the backbone of the crystallographic ligand from the 1AIM structure and making modifications as needed to construct the AQ665183 compound. Torsional, angle, and distance constraints derived from other complexes were applied and the AQ665183–cruzain complex constructed was subjected to several cycles of molecular mechanics-based energy minimization using SYBYL (Tripos Inc.) in order to arrive at the final structural model.

5.6. Preparation, purification, and crystallization of cruzain bound to inhibitors

Cruzain was expressed and purified as previously described.^{6,16,34} Active cruzain was reversibly inhibited with excess methyl methanethiosulfonate (MTS). This prevented further proteolytic activity of the enzyme while purification was completed. After passage over a MonoQ column on an ÄKTA Explorer system (Amersham Pharmacia Biotech Inc.), fractions containing pure cruzain bound to MTS were concentrated via vacuum dialysis against 2 mM bis-tris pH 5.8 to a final concentration of 10–12 mg/mL for crystallization. During concentration, DTT was added to a concentration of 5 mM to remove the MTS. The inhibitor AQ581332, dissolved in DMSO, was added in molar excess quantities to achieve complexation. Crystals were obtained by the hanging drop method against a grid of 0.6–1.0 M sodium citrate, pH 5.5–7.0. Drops of cruzain bound inhibitor that were equilibrated at 18 °C for 24 h were microseeded with existing crystals of cruzain bound to another covalently bound inhibitor.¹⁶ After 7–12 days, small crystals of cruzain bound to a vinyl sulfone-based inhibitor appeared. These crystals were crushed and diluted 10⁴ or 10⁵-fold to produce seeding solutions. New hanging drops of cruzain bound to inhibitor that had been equilibrated at 18 °C for 24 h were microseeded with one microliter of diluted seeding stock solution. Data collection quality crystals (100–200 μ m per edge) grew to maximum size within 3 weeks.

5.7. Data collection, structure solution, and crystallographic refinement

Data were collected at room temperature on an *R*-axis IIC detector fitted with Yale MSC mirrors. A Rigaku RU-H3R rotating anode operating at 100 mA, 50 kV was used for this complex. Data were integrated and scaled using the HKL package of software.³⁵ Table 5 summarizes data processing statistics. The structure was solved by molecular replacement using AMoRe.³⁶ The 1.6 Å structure of cruzain bound to a covalently bound peptidyl inhibitor was used as a search model, excluding the inhibitor molecule and all water molecules.¹⁶ One unique solution was easily found for the complex, with a resulting *R*-factor prior to refinement of 0.296% and correlation coefficient of 0.758. The starting model and parameter set for the inhibitor molecule was generated with Moloc.³⁷ Initial placement of the inhibitor model into the structure was made via inspection of $F_o - F_c$ difference electron density maps calculated from the molecular replacement solution and contoured at 3 σ . Subsequent model adjustment and refinement (positional and individual *B*-factor) were accomplished with CHAIN and X-PLOR.³⁸ A simulated annealing omit map calculated with a 5 Å radius around the omitted inhibitor molecule were used to confirm placement and orientation of the inhibitor molecule. Final rounds of refinement were completed with the Refmac module of CCP4i³⁹ and electron density maps were visualized with Quanta.⁴⁰ Water molecules were added automatically with the X-solvate module of Quanta as well as manually with the requirement that they be located in

strong difference electron density peaks ($\geq 3\sigma$) and within acceptable hydrogen-bonding distance and geometry.

5.8. Synthesis and analysis

All commercial reagents were used as received from their respective suppliers unless otherwise specified. Flash column chromatography was performed using silica gel 60 (Merck Art 9385). ¹H NMR spectra were recorded on a Varian Gemini 300 MHz NMR system. Deuterated solvents were obtained from Cambridge Isotope Laboratories, Inc. and tetramethylsilane was used as an internal standard. Chemical shifts are reported in ppm (δ) downfield relative to tetramethylsilane and the coupling constants are given in Hertz. HPLC purity analysis was run using UV detection at 254 nm on a Shimadzu 10 A analytical chromatography system fitted with a YMC reverse phase (C18, 3 μ m, 3 \times 100 mm) column and/or ELSD on a Sedex. The retention times reported in this experimental were obtained using the following conditions. Samples of the reaction product (5 μ L, 1 mg/mL) in DMSO were injected onto the column. The flow rate was 0.75 mL per minute. The linear gradient elution for the acetonitrile/water eluent was from 20% to 100% acetonitrile over 2.5 min. The elution solvents contained 0.1% trifluoroacetic acid. Mass spectra were obtained on a Micromass Platform 2 using a Gilson 215 autosampler and an HP 1050 LC. Ionization of the compounds was achieved either by electrospray (ES+) or chemical ionization (APCI, N₂) techniques. The following abbreviations are used: TFA (trifluoroacetic acid), DIEA (*N,N*-diisopropylethylamine), HBTU (*O*-benzotriazol-1-yl-*N,N,N',N'*-tetramethyluronium hexafluorophosphate).

5.8.1. Preparation of 2-(ethoxycarbonyl)-2-oxoethylidene-triphenylphosphorane. A solution of ethyl bromopyruvate (100 g, 0.5127 mol) in dry carbon tetrachloride (200 mL) was added dropwise over 3 h to a stirred and cooled (0 °C) solution of PPh₃ (135 g, 0.514 mol) in dry carbon tetrachloride (1.5 L). The reaction warmed to rt and then stirred for an additional 36 h. The supernatant was decanted from the yellow hygroscopic crystals, which were then washed with anhydrous ether (3 \times 400 mL) by trituration and decantation. The resulting dried sticky solid was dissolved in methanol (600 mL) and the solution cooled to 0 °C. The pH was adjusted to 10 by gradual addition of iced aqueous sodium carbonate (1 N). The solution was diluted to 4 L with ice water and stirred for 1 h at 0 °C, after which the precipitate was collected and washed with cold water. It was then recrystallized from hot ethanol and water. The resulting crystals separated out as plates (120 g). A second crop (11 g) was obtained as described. The total yield was 131 g (75%). Mp 182–183 °C.

LC/MS (ES+) 378.5. ¹H NMR (300 MHz, CDCl₃): δ 1.35 (t, $J_1 = 7.9$ Hz, 3H, –CH₃), 4.26 (dd, $J_1 = 7.9$ Hz, 2H, –CH₂–), 4.84 (d, $J_1 = 25.7$ Hz, 1H, CH), 7.42–7.73 (m, 15H, –C₆H₅). C₂₃H₂₁O₃P (376.38).

5.8.2. Preparation of peptidyl vinyl- α -keto esters. A solution of the commercially available aldehyde Z-Phe CHO (0.200 g, 0.706 mmol, 1.0 equiv) and the

(ethoxyoxalyl)methylene-triphenylphosphorane (0.533 g, 1.412 mmol, 2.0 equiv) in anhydrous dioxane (10 mL) was heated under reflux for 10 h. The pale yellow solution was reduced in vacuo, the residue triturated with diethyl ether (50 mL) and the mixture cooled to 0 °C. The precipitated triphenylphosphine oxide was filtered off and washed with cold diethyl ether (2 × 20 mL) and the combined filtrate and washings were concentrated in vacuo to a thick oil, which after purification over silica (EtOAc–hexane, 1:3) yielded the vinyl α -keto ester as a white solid (0.1885 g, 70%).

LC/MS (AP+) 382.3. ^1H NMR (300 MHz, CDCl_3): δ 1.36 (t, $J = 7.9$ Hz, 3H, $-\text{CH}_3$), 2.95 (d, $J = 3$ Hz, 2H, $-\text{CH}_2-\text{C}_6\text{H}_5$), 4.33 (q; $J = 7.9$ Hz, $-\text{CH}_2$), 4.71–4.83 (m, 1H, $-\text{CH}-$), 4.88 (s, 1H, NH), 5.08 (s, 2H, $-\text{CH}_2-\text{C}_6\text{H}_5$), 6.72 (d, $J = 17.8$ Hz, 1H, $-\text{CH}$), 7.08–7.41 (m, 11H, $-\text{CH}$ and $-\text{C}_6\text{H}_5$).

A solution of the aldehyde Z-Phe-Ala-CHO (0.3 g, 0.847 mmol, 1.0 equiv) and (ethoxyoxalyl)methylene-triphenylphosphorane (0.639 g, 1.693 mmol, 2.0 equiv) in anhydrous dioxane (10 mL) were treated as described earlier. Purification over silica (EtOAc–hexane, 1:3) yielded the corresponding vinyl α -keto ester as a white solid (0.2491 g, 65%) of the vinyl α -keto ester.

LC/MS (ES+) 452.9. ^1H NMR (300 MHz, CDCl_3): δ 1.21 (d, $J = \text{Hz}$, 3H, $-\text{CH}_3$), 1.42 (dd, $J_1 = J_2 = 7.9$ Hz, 3H, $-\text{CH}_3$), 3.10 (dd, $J_1 = \text{Hz}$, $J_2 = \text{Hz}$, 2H, $-\text{CH}_2-\text{C}_6\text{H}_5$), 4.38 (dd, $J_1 = 15.9$ Hz, $J_2 = 7.9$ Hz, $-\text{CH}_2$), 4.45 (q, $J = 7.5$ Hz, 1H, $-\text{CH}$), 4.72 (q, $J = 7.5$ Hz, 1H, $-\text{CH}$), 5.11 (s, 2H, $-\text{CH}_2-\text{C}_6\text{H}_5$), 5.60 (d, $J = 7.5$ Hz, 1H, NH), 6.24 (d, $J = 7.5$ Hz, 1H, NH), 6.59 (d, $J = 17.7$ Hz, 1H, $-\text{CH}$), 6.89 (dd, $J_1 = 17.7$ Hz, $J_2 = 5.9$ Hz, 1H, $-\text{CH}$), 7.16–7.45 (m, 10H, $-\text{C}_6\text{H}_5$).

5.8.3. Preparation of ML-library. To a cooled solution (0 °C) of the Cbz-protected α -ketoamide (0.6 g, 2.56 mmol, 1.0 equiv) in dry dichloromethane (10 mL) was added BCl_3 (1.20 g, 10.24 mmol, 4.0 equiv). The reaction mixture was stirred at room temperature for 14 h. The solvent was removed under reduced pressure and the resulting residue was dissolved in DMF (10.25 mL) to obtain a 0.25 M reaction solution.

A solution of Cbz-protected amino acid in DMF was stirred with 1.1 equiv of DIEA (0.25 M DMF solution) and 1.05 equiv HBTU (0.25 M DMF solution) at room temperature. After 15 min a 1.1 equiv of the Cbz-deprotected α -keto amide (0.25 M DMF solution) was added and stirred for 12 h at room temperature. The resulting mixtures were diluted with ethyl acetate (10 mL) and extracted with 1 N NaHCO_3 (10 mL) solution and water (10 mL). After removal of the solvent in vacuo the residues were analyzed by LC–MS and the crude products were isolated by preparative HPLC.

5.8.4. Characterization of the Z-Phe-Ala- α -keto inhibitors

5.8.4.1. Z-Phe-Ala- α -keto-ethylester. Purified by column chromatography (EtOAc/hexane = 1/1). Yield 75%. ^1H NMR (300 MHz, CDCl_3): δ 1.25 (t,

$J_1 = 7.14$ Hz, 3H, $-\text{CH}-\text{CH}_3$), 1.32–1.4 (m, 3H, $\text{O}-\text{CH}_2-\text{CH}_3$), 2.9–3.15 (m, 2H, $-\text{CH}_2-\text{Ph}$), 4.32–4.38 (m, 2H, $\text{O}-\text{CH}_2-\text{CH}_3$), 4.45 (m, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.09 (s, 2H, $\text{O}-\text{CH}_2-\text{Ph}$), 5.32 (br s, 1H, N–H), 6.2 (br s, 1H, N–H), 7.20–7.35 (m, 10H, aromatic). MS (EI), m/z (relative intensity): 428 ($[\text{M}^+]$, 2).

5.8.4.2. Z-Phe-Ala- α -keto-acid. Purified by reverse phase chromatography. Yield 69%. ^1H NMR (300 MHz, CDCl_3): δ 1.21 (t, $J_1 = 7.71$ Hz, 3H, $\text{CH}-\text{CH}_3$), 3.04 (m, 2H, $-\text{CH}_2-\text{Ph}$), 4.48 (m, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.05 (m, 3H, $\text{O}-\text{CH}_2-\text{Ph}$ and $\text{CH}-\text{CH}_2-\text{Ph}$), 7.16–7.32 (m, 10H, aromatic). MS (EI), m/z (relative intensity): 399 ($[\text{M}^+]$, 2).

5.8.4.3. Z-Phe-Ala- α -keto-methylamide. Purified by column chromatography (3% MeOH in CH_2Cl_2). Yield 50%. ^1H NMR (300 MHz, CDCl_3): δ 1.30 (d, $J_1 = 7.23$ Hz, 1.5H, $-\text{CH}-\text{CH}_3$), 1.45 (d, $J_1 = 7.15$ Hz, 1.5H, $-\text{CH}-\text{CH}_3$), 3.05–3.19 (m, 2H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 4.32–4.51 (m, 3H, $\text{O}-\text{CH}_2-\text{CH}_3$ and $\text{NH}-\text{CH}_2-\text{Ph}$), 5.12 (m, 2H, $\text{O}-\text{CH}_2-\text{Ph}$), 5.24 (m, 1H, $-\text{CH}-$), 5.42 (br s, 1H, N–H), 6.3 and 6.5 (br s, 1H, N–H), 7.23–7.38 (m, 15H, aromatic). MS (EI), m/z (relative intensity): 488 ($[\text{M}^+]$, 2).

5.8.4.4. Z-Phe-Ala- α -keto-benzylamide. Purified by column chromatography (3% MeOH in CH_2Cl_2). Yield 72%. ^1H NMR (300 MHz, CDCl_3): δ 1.30 (d, $J_1 = 7.23$ Hz, 1.5H, $-\text{CH}-\text{CH}_3$), 1.41 (d, $J_1 = 7.15$ Hz, 1.5H, $-\text{CH}-\text{CH}_3$), 3.05–3.19 (m, 2H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 4.32–4.51 (m, 3H, $\text{O}-\text{CH}_2-\text{CH}_3$ and $\text{NH}-\text{CH}_2-\text{Ph}$), 5.12 (m, 2H, $\text{O}-\text{CH}_2-\text{Ph}$), 5.24 (m, 1H, $-\text{CH}-$), 5.42 (br s, 1H, N–H), 6.3 and 6.5 (br s, 1H, N–H), 7.23–7.38 (m, 15H, aromatic). MS (EI), m/z (relative intensity): 488 ($[\text{M}^+]$, 2).

5.8.4.5. Z-Phe-Val- α -keto-benzylamide. Purified by recrystallization from (CH_2Cl_2 –hexane). Yield 60%. ^1H NMR (300 MHz, CDCl_3): δ 1.30 (d, $J_1 = 7.23$ Hz, 1.5H, $-\text{CH}-\text{CH}_3$), 1.41 (d, $J_1 = 7.15$ Hz, 1.5H, $-\text{CH}-\text{CH}_3$), 3.05–3.19 (m, 2H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 4.32–4.51 (m, 3H, $\text{O}-\text{CH}_2-\text{CH}_3$ and $\text{NH}-\text{CH}_2-\text{Ph}$), 5.12 (m, 2H, $\text{O}-\text{CH}_2-\text{Ph}$), 5.24 (m, 1H, $-\text{CH}-$), 5.42 (br s, 1H, N–H), 6.3 and 6.5 (br s, 1H, N–H), 7.23–7.38 (m, 15H, aromatic). MS (EI), m/z (relative intensity): 488 ($[\text{M}^+]$, 2).

5.8.5. General method for the synthesis of peptide alcohols. A solution of an amino alcohol (1.26 mmol) in 1,4-dioxane (5 mL) was added dropwise to a solution of a commercially available activated *N*-protected amino acid succinimide ester (1.26 mmol) in 1,4-dioxane (5 mL). It was stirred at room temperature for 12 h. Diethylether was added (40 mL) and the organic layer washed with 1 N HCl (10 mL), saturated NaHCO_3 (10 mL), saturated NaCl (10 mL), and dried over MgSO_4 . After removal of the solvent, the crude products were characterized and used without further purification.

5.8.5.1. 3-[*N*-(Benzyloxycarbonyl-(*S*)-phenylalanyl)-aminopropan-1-ol]. Yield 76%. ^1H NMR (300 MHz, CDCl_3): δ 1.54 (dd, $J_1 = J_2 = 6.03$ Hz, 2H, $-\text{CH}_2-$), 2.56 (br s, 1H, O–H), 3.05 (m, 2H, $-\text{CH}_2-$), 3.30 (dd,

$J_1 = J_2 = 6.03$ Hz, 2H, $-\text{CH}_2-$), 3.48 (dd, $J_1 = J_2 = 6.03$ Hz, 2H, $-\text{CH}_2-\text{OH}$), 3.48 (dd, $J_1 = J_2 = 5.49$ Hz, 2H, $-\text{CH}_2-$), 4.47 (dd, $J_1 = J_2 = 6.06$ Hz, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.08 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.4 (br s, 1H, N-H), 6.17 (br s, 1H, N-H), 7.18–7.35 (m, 10H, aromatic). MS (EI), m/z (relative intensity): 357 ($[\text{M}^+]$, 2).

5.8.5.2. 2-[N-(Benzyloxycarbonyl-(S)-phenylalanyl)-amino]ethan-1-ol. Purified by recrystallization (EtOAc–pentane). Yield 83%. ^1H NMR (300 MHz, CDCl_3): δ 2.15 (br s, 1H, O-H), 3.01 (dd, $J_1 = J_2 = 7.83$ Hz, 1H, $-\text{CH}_2-\text{Ph}$), 3.14 (dd, $J_1 = J_2 = 6.5$ Hz, 1H, $-\text{CH}_2-\text{Ph}$), 3.26–3.35 (m, 2H, $-\text{CH}_2-\text{NH}-$), 3.48–3.56 (m, 2H, $-\text{CH}_2-\text{OH}$), 4.35 (dd, $J_1 = J_2 = 6.6$ Hz, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.08 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.35 (br s, 1H, N-H), 6.00 (br s, 1H, N-H), 7.19–7.37 (m, 10H, aromatic). MS (EI), m/z (relative intensity): 343 ($[\text{M}^+]$, 2).

5.8.5.3. (2S)-2-[N-(Benzyloxycarbonyl-(S)-phenylalanyl)amino]-3-methylpentan-1-ol. Yield 95%. ^1H NMR (300 MHz, CDCl_3): δ 0.7–0.84 (m, 6H, $2 \times \text{CH}_3$) 0.9–1.03 (m, 1H, $\text{CH}_3-\text{CH}-\text{CH}_2-$) 1.25–1.35 (m, 1H, $-\text{CH}-\text{CH}_2-\text{CH}_3$), 1.4–1.52 (m, 1H, $-\text{CH}-\text{CH}_2-\text{CH}_3$), 1.8 (br s, 1H, O-H), 3.00 (dd, $J_1 = J_2 = 7.83$ Hz, 1H, $-\text{CH}_2-\text{Ph}$), 3.14 (dd, $J_1 = J_2 = 6.5$ Hz, 1H, $-\text{CH}_2-\text{Ph}$), 3.43 (m, 2H, $-\text{CH}_2-\text{OH}$), 3.65 (m, 1H, $-\text{CH}-$), 4.35 (dd, $J_1 = J_2 = 6.6$ Hz, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.11 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.39 (br s, 1H, N-H), 5.62 (br s, 1H, N-H), 7.23–7.34 (m, 10H, aromatic). MS (EI), m/z (relative intensity): 389 ($[\text{M}^+]$, 2).

5.8.5.4. (2S)-2-[N-(Benzyloxycarbonyl-(S)-phenylalanyl)amino]-3-phenylpropan-1-ol. Yield 53%. ^1H NMR (300 MHz, CDCl_3): δ 1.78 (br s, 1H, O-H), 2.97 (dd, $J_1 = J_2 = 8.16$ Hz, 1H, $-\text{CH}_2-\text{Ph}$), 3.02 (m, 1H, $-\text{CH}_2-\text{Ph}$), 2.71 (m, 2H, $-\text{CH}_2-\text{Ph}$), 3.39 (m, 2H, $-\text{CH}_2-\text{OH}$), 4.05 (m, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 4.28 (dd, $J_1 = J_2 = 6.6$ Hz, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.08 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.25 (br s, 1H, N-H), 5.74 (br s, 1H, N-H), 7.08–7.36 (m, 15H, aromatic). MS (EI), m/z (relative intensity): 433 ($[\text{M}^+]$, 2).

5.8.5.5. (2S)-2-[N-(Benzyloxycarbonyl-(S)-phenylalanyl)amino]-3-cyclohexylpropan-1-ol. Yield 99%. ^1H NMR (300 MHz, CDCl_3): δ 0.8–0.95 (m, 2H, $-\text{CH}_2-$, cyclohexyl), 1.1–1.28 (m, 5H, $-\text{CH}-\text{CH}_2-\text{CH}_2-$, cyclohexyl), 1.53–1.72 (br m, 4H, $-\text{CH}_2-\text{CH}_2-$, cyclohexyl), 3.0 (m, 1H, $-\text{CH}_2-\text{Ph}$), 3.18–3.0 (m, 1H, $-\text{CH}_2-\text{Ph}$), 3.32 (m, 1H, $-\text{CH}_2-\text{OH}$), 3.45 (m, 1H, $-\text{CH}_2-\text{OH}$), 1.8 (br s, 1H, O-H), 3.95 (m, 1H, $-\text{CH}-$), 4.35 (dd, $J_1 = J_2 = 6.6$ Hz, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.12 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.35 (br m, 1H, N-H), 5.51 (br m, 1H, N-H), 7.26–7.36 (m, 10H, aromatic). MS (EI), m/z (relative intensity): 439 ($[\text{M}^+]$, 2).

5.8.6. General method for the Swern oxidation of peptide alcohols under formation of peptide aldehydes. Dimethylsulfide anhydrous (178 mL, 2.49 mmol, 1.5 equiv) in dry CH_2Cl_2 (3.6 mL) was added dropwise over 10 min to a cooled (-63°C) solution of oxalyl chloride (1.78 mol, 1.5 equiv) in dry CH_2Cl_2 (3.6 mL). After additional 20 min at -63°C , the reaction mixture was treated

dropwise over 20 min with the peptide alcohol (1.58 mmol) in dry CH_2Cl_2 (9 mL) and stirred for one additional hour at -63°C . Triethylamine (0.6672 mL, 4.632 mmol, 4 equiv) was finally added dropwise over 20 min and it was stirred additional 30 min at -63°C . The reaction mixture was let warm up to rt, poured into a citric acid solution (0.24 N, 10 mL), the organic layer was separated and the aqueous layer was extracted with EtOAc (3×10 mL). The combined organic extracts were washed with saturated NaHCO_3 (10 mL), saturated NaCl (10 mL) and dried over MgSO_4 . The solvent was removed under reduced pressure to yield the crude aldehyde.

5.8.6.1. 3-[N-(Benzyloxycarbonyl-(S)-phenylalanyl)-amino]propan-1-al. Yield 96%. ^1H NMR (300 MHz, CDCl_3): δ 2.0 (m, 2H, $-\text{CH}_2-$), 2.93 (m, 1H, $-\text{CH}_2-$), 3.05 (m, 1H, $-\text{CH}_2-$), 3.32–3.48 (br m, 2H, $-\text{CH}_2-\text{Ph}$), 4.26 (m, 2H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.11 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.25 (br s, 1H, N-H), 6.02 (br s, 1H, N-H), 7.17–7.41 (m, 10H, aromatic), 9.51 (s, 1H, $-\text{CHO}$). MS (EI), m/z (relative intensity): 355 ($[\text{M}^+]$, 2).

5.8.6.2. 2-[N-(Benzyloxycarbonyl-(S)-phenylalanyl)amino]ethan-1-al. Yield 98%. ^1H NMR (300 MHz, CDCl_3): δ 3.01–3.12 (br m, 2H, $-\text{CH}_2-\text{Ph}$), 3.26 (m, 2H, $-\text{CH}_2-\text{NH}-$), 4.45 (m, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.08 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.35 (br s, 1H, N-H), 5.82 (br s, 1H, N-H), 7.17–7.41 (m, 10H, aromatic), 9.53 (s, 1H, $-\text{CHO}$). MS (EI), m/z (relative intensity): 341 ($[\text{M}^+]$, 2).

5.8.6.3. (2S)-2-[N-(Benzyloxycarbonyl-(S)-phenylalanyl)amino]-3-methylpentan-1-al. Yield 98%. ^1H NMR (300 MHz, CDCl_3): δ 0.82–0.91 (m, 6H, $2 \times \text{CH}_3$) 1.1–1.18 (m, 1H, $\text{CH}_3-\text{CH}-\text{CH}_2-$) 1.32–1.45 (m, 1H, $-\text{CH}-\text{CH}_2-\text{CH}_3$), 1.9–2.02 (m, 1H, $-\text{CH}-\text{CH}_2-\text{CH}_3$), 3.08 (m, 2H, $-\text{CH}_2-\text{Ph}$), 4.42–4.58 (2m, 2H, $-\text{CH}-$ and $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.08 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.45 (br s, 1H, N-H), 6.5 (br s, 1H, N-H), 7.16–7.33 (m, 10H, aromatic), 9.47 (s, 1H, $-\text{CHO}$). MS (EI), m/z (relative intensity): 397 ($[\text{M}^+]$, 2).

5.8.6.4. (2S)-2-[N-(Benzyloxycarbonyl-(S)-phenylalanyl)amino]-3-phenylpropan-1-al. Yield 93%. ^1H NMR (300 MHz, CDCl_3): δ 3.03–3.18 (m, 4H, $2 \times -\text{CH}_2-\text{Ph}$), 4.41 (m, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 4.6 (m, 1H, $-\text{CH}-\text{CH}_2-\text{Ph}$), 5.07 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.23 (br s, 1H, N-H), 6.25 (br s, 1H, N-H), 7.00–7.34 (m, 15H, aromatic), 9.46 (s, 1H, $-\text{CHO}$). MS (EI), m/z (relative intensity): 431 ($[\text{M}^+]$, 2).

5.8.6.5. (2S)-2-[N-(Benzyloxycarbonyl-(S)-phenylalanyl)amino]-3-cyclohexylpropan-1-al. Yield 98%. ^1H NMR (300 MHz, CDCl_3): δ 0.85–0.97 (m, 2H, $-\text{CH}_2-$, cyclohexyl), 1.13–1.26 (m, 4H, $-\text{CH}_2-\text{CH}_2-$, cyclohexyl), 1.59–1.66 (br m, 5H, $-\text{CH}-\text{CH}_2-\text{CH}_2-$, cyclohexyl), 3.02 (m, 2H, $-\text{CH}_2-\text{Ph}$), 4.45 (m, 2H, $-\text{CH}-\text{CH}_2-\text{Ph}$ and $-\text{CH}-\text{CH}_2-\text{cyclohexyl}$), 5.10 (s, $\text{O}-\text{CH}_2-\text{Ph}$), 5.31 (br m, 1H, N-H), 6.1 (br m, 1H, N-H), 7.18–7.36 (m, 10H, aromatic), 9.39 (s, 1H, $-\text{CHO}$). MS (EI), m/z (relative intensity): 439 ($[\text{M}^+]$, 2).

Acknowledgements

The authors would like to thank the Analytical and Production Groups at ArQule, Inc. for their support. This work was supported by a NIH grant P01-AI35707 to C.S.C. and J.C.E.

References and notes

1. World Health Organization: *Special Program for Research and Training in Tropical Diseases (TDR)*.
2. Du, X.; Guo, C.; Hansell, E.; Doyle, P. S.; Caffrey, C. R.; Holler, T. P.; McKerrow, J. H.; Cohen, F. E. *J. Med. Chem.* **2002**, *45*, 2695.
3. Engel, J. C.; Doyle, P. S.; Palmer, J.; Hsieh, I.; Bainton, D. F.; McKerrow, J. H. *J. Cell Sci.* **1998**, *111*(Pt5), 597.
4. Engel, J. C.; Doyle, P. S.; Hsieh, I.; McKerrow, J. H. *J. Exp. Med.* **1998**, *188*, 725.
5. Itow, S.; Camargo, E. P. *J. Protozool.* **1977**, *24*, 591.
6. Eakin, A. E.; Mills, A. A.; Harth, G.; McKerrow, J. H.; Craik, C. S. *J. Biol. Chem.* **1992**, *267*, 7411.
7. McGrath, M. E.; Eakin, A. E.; Engel, J. C.; McKerrow, J. H.; Craik, C. S.; Fletterick, R. J. *J. Mol. Biol.* **1995**, *247*, 251.
8. Roush, W. R.; Gonzalez, F. V.; McKerrow, J. H.; Hansell, E. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2809.
9. Scheidt, K. A.; Roush, W. R.; McKerrow, J. H.; Selzer, P. M.; Hansell, E.; Rosenthal, P. J. *Bioorg. Med. Chem.* **1998**, *6*, 2477.
10. Coffen, D. L.; Baldino, C. M.; Lange, M.; Tilton, R. F.; Tu, C. *Med. Chem. Res.* **1998**, *8*(4/5), 206.
11. Baldino, C. M.; Coffen, D. L.; Chipman, S. D.; Cheng, H. U.S. Patent, 2000.
12. Baldino, C. M. *J. Comb. Chem.* **2000**, *2*, 89.
13. Baldino, C. M.; Caserta, J.; Goetzinger, W.; Harris, M.; Hartsough, D.; Yohannes, D.; Yu, L.; Kyranos, J. N. *Curr. Drug Discov.* **2004**, *15*.
14. Li, Z.; Ortega-Vilain, A. C.; Patil, G. S.; Chu, D. L.; Foreman, J. E.; Eveleth, D. D.; Powers, J. C. *J. Med. Chem.* **1996**, *39*, 4089.
15. Reetz, M. T.; Griebenow, N. *Liebigs Ann.* **1996**, 335.
16. Gillmor, S. A.; Craik, C. S.; Fletterick, R. J. *Protein Sci.* **1997**, *6*, 1603.
17. Wassermann, H. H.; Petersen, A. K. *Tetrahedron Lett.* **1997**, *38*, 953.
18. Brinen, L. S.; Hansell, E.; Cheng, J.; Roush, W. R.; McKerrow, J. H.; Fletterick, R. J. *Struct. Fold. Des.* **2000**, *8*, 831.
19. Shoichet, B. K.; Kuntz, I. D. *Protein Eng.* **1993**, *6*, 723.
20. Li, Z.; Patil, G. S.; Golubski, Z. E.; Hori, H.; Tehrani, K.; Foreman, J. E.; Eveleth, D. D.; Bartus, R. T.; Powers, J. C. *J. Med. Chem.* **1993**, *36*, 3472.
21. Rutenber, E. E.; McPhee, F.; Kaplan, A. P.; Gallion, S. L., Jr.; Hogan, J. C., Jr.; Craik, C. S.; Stroud, R. M. *Bioorg. Med. Chem.* **1996**, *4*, 1545.
22. Walter, J.; Bode, W. *Hoppe Seylers Z. Physiol. Chem.* **1983**, *364*, 949.
23. Sasaki, T.; Kishi, M.; Saito, M.; Tanaka, T.; Higuchi, N.; Kominami, E.; Katunuma, N.; Murachi, T. *J. Enzyme Inhib.* **1990**, *3*, 195.
24. Angelastro, M. R.; Mehdi, S.; Burkhart, J. P.; Peet, N. P.; Bey, P. *J. Med. Chem.* **1990**, *33*, 11.
25. Hu, L. Y.; Abeles, R. H. *Arch. Biochem. Biophys.* **1990**, *281*, 271.
26. Almquist, R. G.; Chao, W. R.; Ellis, M. E.; Johnson, H. L. *J. Med. Chem.* **1980**, *23*, 1392.
27. Szelke, M.; Leckie, B.; Hallett, A.; Jones, D. M.; Sueiras, J.; Atrash, B.; Lever, A. F. *Nature* **1982**, *299*, 555.
28. Harbeson, S. L.; Abelleira, S. M.; Akiyama, A.; Barrett, R., III; Carroll, R. M.; Straub, J. A.; Tkacz, J. N.; Wu, C.; Musso, G. F. *J. Med. Chem.* **1994**, *37*, 2918.
29. LaLonde, J. M.; Zhao, B.; Smith, W. W.; Janson, C. A.; DesJarlais, R. L.; Tomaszek, T. A.; Carr, T. J.; Thompson, S. K.; Oh, H. J.; Yamashita, D. S.; Veber, D. F.; Abdel-Meguid, S. S. *J. Med. Chem.* **1998**, *41*, 4567.
30. Thompson, S. K.; Smith, W. W.; Zhao, B.; Halbert, S. M.; Tomaszek, T. A.; Tew, D. G.; Levy, M. A.; Janson, C. A.; D'Alessio, K. J.; McQueney, M. S.; Kurdyla, J.; Jones, C. S.; DesJarlais, R. L.; Abdel-Meguid, S. S.; Veber, D. F. *J. Med. Chem.* **1998**, *41*, 3923.
31. Cleland, W. W. In *Methods in Enzymology*; Purich, D. L., Ed.; Academic, 1979; Vol. 103.
32. Eggers, C. T.; Wang, S. X.; Fletterick, R. J.; Craik, C. S. *J. Mol. Biol.* **2001**, *308*, 975.
33. Du, X.; Hansell, E.; Engel, J. C.; Caffrey, C. R.; Cohen, F. E.; McKerrow, J. H. *Chem. Biol.* **2000**, *7*, 733.
34. Eakin, A. E.; McGrath, M. E.; McKerrow, J. H.; Fletterick, R. J.; Craik, C. S. *J. Biol. Chem.* **1993**, *268*, 6115.
35. Otwinowski, Z. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*; Sawyer, L. Isaacs N., Bailey, S., Eds.; Daresbury Laboratory: Warrington, UK, 1993.
36. Navaza, J. *Acta Crystallogr.* **1994**, *50*, 157.
37. Gerber, P. R.; Muller, K. *J. Comput. Aided Mol. Des.* **1995**, *9*, 251.
38. Sack, J. S. *J. Mol. Graphics* **1988**, *6*, 244.
39. Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E. *Acta Crystallogr.* **2003**, *D59*, 1131.
40. Accelrys; San Diego, CA, USA, 2000.