Investigation of the substrate specificity of cruzipain, the major cysteine proteinase of *Trypanosoma cruzi*, through the use of cystatin-derived substrates and inhibitors

Carole SERVEAU*, Gilles LALMANACH*, Maria A. JULIANO†, Julio SCHARFSTEIN‡, Luiz JULIANO† and Francis GAUTHIER*§

*Laboratory of Enzymology and Protein Chemistry, CNRS URA 1334, University François Rabelais, 2bis Bd Tonnellé, F37032 Tours Cedex, France, †Department of Biophysics, Escola Paulista de Medicina, Caixa Postal 20388, 04304 Sao Paulo, Brazil, and ‡Instituto de Biofisica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, 21944 Rio de Janeiro, Brazil

A panel of intramolecularly quenched fluorogenic substrates containing the conserved QVVA and LVG inhibitory sequences of cystatin inhibitors was used to describe the specificity of the major cysteine proteinase of *Trypanosoma cruzi* (cruzipain or cruzain). This approach was based on the observations that: (1) cruzipain is strongly inhibited by chicken cystatin and rat T-kininogen, two representative members of cystatin families 2 and 3; (2) the QVVA- and LVG-containing substrates are specifically hydrolysed by papain-like proteinases; and (3) the cystatin-like motifs are similar to the proteolytically sensitive sequences in cruzipain that separate the pro-region and/or the C-terminal extension from the catalytic domain. Specificity constants ($k_{cat.}/K_m$) were determined and compared with those of mammalian cathepsins B and L from rat liver lysosomes. Cruzipain and the mammalian proteinases cleaved cystatin-derived substrates at the same site, but their specificities differed significantly. Increased specificity for cruzipain was obtained by replacing amino acids at critical positions on both sides of the cleavage sites, especially at position P2′. The specificity constants ($k_{cat.}/K_m$) obtained for the two substrates with a prolyl residue at P2′ (*O*-aminobenzoyl-QVVAGP-ethylenediamine 2-4-dinitrophenyl and *O*-aminobenzoyl-VVGGP-ethylenediamine 2-4-dinitrophenyl) were about 50 times higher for cruzipain than for rat cathepsin L and about 100 times higher than for cathepsin B. Diazomethylketone derivatives, based on the non-prime sequence of cystatin-derived substrates, inhibited cruzipain irreversibly, but their inactivation rate constants were considerably lower than those for mammalian cathepsins B and L, confirming the importance of P′ residues for cruzipain specificity.

INTRODUCTION

Cysteine proteinases are involved in several aspects of host–parasite interactions and are the most abundant proteinases in most parasitic protozoa [1,2]. There has recently been great interest in this class of proteinases, since they may be new targets for chemotherapeutic drugs (see [3] for review). Cruzipain is the most abundant cysteine proteinase in *Trypanosoma cruzi* [4] and is the immunodominant antigen recognized during human infection [5]. It is encoded by more than 100 copies of polymorphic genes located on several chromosomes [6,7]. However, in spite of its genetic polymorphism, cruzipain is one of the best characterized parasite proteinases. All the molecular varieties reported so far have the catalytic triad characteristic of cysteine proteinases [8] and a structural organization similar to that of papain-like mammalian lysosomal proteinases [6,9,10]. Unlike lysosomal mammalian proteinases, however, the mature enzyme has a C-terminal extension of 131 residues that can be released by self-proteolysis [11,12]. This peculiar feature has also been reported in a cysteine proteinase from *Trypanosoma brucei* [13] and in a plant proteinase isolated from cold-treated tomato [14]. Neither the function of the catalytic N-terminal domain nor that of the C-terminal extension have been elucidated so far, even though cruzipain is thought to participate in host-cell invasion, binding to target cells, intracellular replication and differentiation of the parasite [5,15–18]. These putative functions, all at important stages of the life cycle of the parasite, make cruzipain an excellent target for specific, irreversible inhibitors that may be used to develop a new chemotherapy. Several successful attempts have been made in this direction, resulting in impairment of host-cell invasion by *T. cruzi* and intracellular development of the parasite [17–19]. Diazomethylketone and fluoromethylketone peptide inhibitors, which have been used for this purpose, were initially developed for mammalian lysosomal cysteine proteinases and are therefore not specific for cruzipain or related parasite proteinases. As part of our project to develop specific inhibitors that do not interfere with host cysteine proteinases, we have studied the substrate specificity of cruzipain isolated from the Dm28c clone of *T. cruzi* using peptide substrates with intramolecularly quenched fluorescence derived from the structure of natural cysteine proteinase inhibitors of the cystatin family [20]. These fluorogenic substrates are sensitive and specific for cysteine proteinases [21]. Their peptide moiety, which includes either the substrate-like N-terminal segment of cystatins [22] or the papain-sensitive QVVAG sequence [21], is similar to the hinge between the pro-region and the catalytic domain of cruzipain (VVGAP) and between the catalytic domain and the C-terminal extension (VVGGP). They therefore appear to be appropriate candidates for monitoring cruzipain activity.

MATERIALS AND METHODS

1-*trans*-Epoxysuccinyl-leucylamido(4-guanidino)butane (E64) and D,L-dithiothreitol (DTT) were purchased from Sigma (Saint-Quentin Fallavier, France). Benzyloxycarbonyl-Phe-Arg-7-amino-4-methylcoumarin (Z-Phe-Arg-AMC) and Z-Arg-Arg-AMC were from Bâle Biochimie (Voisins-Le-Bretonneux, France).

Abbreviations used: Abz, *O*-aminobenzoyl; AMC, 7-amino-4-methylcoumarin; E64, 1-*trans*-epoxysuccinyl-leucylamido-(4-guanidino)butane; EDDnp, ethylenediamine 2,4-dinitrophenyl; GlcA, gluconoyl; pNA, 4-nitroanilide; TFA, trifluoroacetic acid; Z, benzyloxycarbonyl; DTT, dithiothreitol.

§ To whom correspondence should be addressed.

Enzymes and inhibitors

Cathepsins B and L were purified from rat liver [23]. Cruzipain was purified from Dm28c clone epimastigotes of *T. cruzi* [24]. The active sites of all cysteine proteinases were titrated with E64 [25]. Rat T-kininogen (thiostatin) was purified as described in [23] and chicken egg white cystatin as described in [26]. Both cystatins were titrated with titrated commercial papain (Boehringer-Mannheim, Germany).

The activation buffers for enzyme assays were: 0.1 M phosphate buffer, pH 6.0, containing 1 mM EDTA/2 mM DTT/0.1 % Brij 35 for cathepsins B and L; and 0.1 M phosphate buffer, pH 6.0, containing 10 mM DTT for cruzipain. All proteinases were preincubated for 5 min at 30 °C before making kinetic measurements.

Synthesis of substrates with intramolecularly quenched fluorescence

The intramolecularly quenched fluorogenic peptides were synthesized by classical solution methods, essentially using a stepwise strategy from the C-terminal aminoacyl ethylenediamine 2-4-dinitrophenyl (aminoacyl EDDnp), with mixed anhydride coupling [27,28]. Substrates containing glutamine as the C-terminal residue were synthesized by solid-phase synthesis as follows: N^{α}-fluoren-α-ylmethoxycarbonyl-Glu(αCO)-*N*-EDDnp was obtained and attached to benzhydrylamino resin through *p*-{(*R*,*S*)-α-[1-(9*H*-fluoren-9-yl)methoxyformamido]-2,4-dimethoxybenzylphenoxyacetic acid} as a linker and the peptides were synthesized by fluoren-α-ylmethoxycarbonyl methodology using the multiple automated peptide synthesizer PSSM-8 (Shimadzu). Intramolecularly quenched fluorogenic substrates were prepared as 2 mM stock solution in dimethylformamide and diluted with activation buffer.

Synthesis of diazomethylketone inhibitors

Z-RLVG-CHN_2 was a gift from Dr. M. Abrahamson (University Hospital, Lund, Sweden) and gluconoyl-QVVA-CHN_2 (GlcA-QVVA-CHN_2) was synthesized from GlcA-QVVA-*O*-benzyl [29]. Inhibitors were dissolved in water and diluted in 0.1 M phosphate buffer, pH 6.8, containing 1 mM EDTA, 2 mM DTT and 0.1 % Brij 35. They were titrated using E64-titrated commercial papain.

Inhibition of cruzipain by cystatins

Measurements were made in the cruzipain activation buffer (final volume: 300 μl). Cruzipain (1.0 nM) was incubated at 30 °C for 30 min with various concentrations of rat T-kininogen (0.5–9.5 nM), and 5 μl of 0.2 mM Z-Phe-Arg-AMC was added to start the reaction. Enzyme activities were measured on a Hitachi F2000 spectrofluorimeter (excitation: 350 nm; emission: 460 nm). $K_{i,app}$ (K_i apparent) values were determined from the slope of a plot of $[I_0]/1-a$ against $1/a$ according to eqn. (1) for tight-binding inhibitors [30]:

$$\frac{[I_0]}{1-a} = \frac{1}{a}(K_{i,app}) + [E_0] \quad (1)$$

where: $a = v_i / v_0$, (v_i is the enzyme activity in the presence of inhibitor and v_0 is the activity without inhibitor) and

$$K_{i,app} = K_i\left(1 + \frac{[S_0]}{K_m}\right)$$

K_i constants were calculated from $K_{i,app}$ values using a K_m of 1.8 μM. Similar experiments were carried out using cruzipain (2.6 nM) and chicken cystatin (0.5–9 nM).

Kinetic measurements

The initial velocities of enzyme reactions with Z-Phe-Arg-AMC and Z-Arg-Arg-AMC as substrates were recorded spectrofluorimetrically with excitation at 350 nm and emission at 460 nm. Experiments were carried out using 0.5 μM AMC for calibration. Reactions were started by adding various concentrations of substrate to cruzipain (0.5 nM final concn.) and the increase in fluorescence was recorded continuously. K_m and $k_{cat.}$ values were determined from Hanes linear plots.

Second-order rate constants ($k_{cat.}/K_m$) were determined for intramolecularly quenched fluorogenic substrates under pseudo-first-order conditions, i.e. using a substrate concentration far below the K_m [31]. The time-course of the enzyme reaction was recorded spectrofluorimetrically (excitation: 320 nm; emission: 420 nm) at 30 °C in the corresponding activation buffer. The system was standardized using *O*-aminobenzoyl-Phe-Arg-OH (Abz-Phe-Arg-OH) prepared from total tryptic hydrolysis of Abz-Phe-Arg-4-nitroanilide (Abz-Phe-Arg-pNA) and its concentration was determined from the absorbance at 410 nm, assuming $\epsilon_{410} = 8800\ M^{-1}\cdot cm^{-1}$ for pNA. Final enzyme concentrations were in the 10^{-9} M range for cruzipain and cathepsin L and about 10^{-8} M for cathepsin B. The substrate-cleavage site was located by amino acid analysis (Applied Biosystems 420A) after fractionation of a proteinase–substrate lysate (molar ratio 1/10000, incubation 60 min at 30 °C) on a C_{18} reverse-phase column. Peptides were eluted at a flow rate of 0.2 ml/min using a 30 min linear (0–60 %) gradient of acetonitrile in 0.075 % trifluoroacetic acid (TFA).

pH–activity profiles

Cruzipain (0.50 nM final concn.) was incubated for 5 min at 30 °C in buffer containing 10 mM DTT, and the enzyme activity was measured by following the hydrolysis of the substrates Z-Phe-Arg-AMC (3.3 μM final concn.) and Z-Arg-Arg-AMC (10 μM final concn.). Buffers at pH 3.0–7.5 were prepared by mixing 0.2 M phosphate and 0.1 M citric acid; those at pH 8.0–9.0 were 0.1 M Tris/HCl.

Inactivation by peptidyl diazomethylketone inhibitors

Irreversible inactivation constants (k_{+0}) of cruzipain were measured by competition with the substrate for the catalytic site under first-order conditions by the method of Tian and Tsou [32]. Under these experimental conditions, product formation in the presence of an irreversible inhibitor is described by eqn. (2):

$$\log([P_{inf.}] - [P]) = \log[P_{inf.}] - 0.43\,A[Y]t \quad (2)$$

where $[P_{inf.}]$ is the product concentration formed at close to infinite time, [P] is the product concentration at time t, [Y] is the inhibitor concentration and A is the apparent inhibition rate constant in the presence of the substrate given by eqn. (3):

$$A = k_{+0}/(1 + K^{-1}[S]) \quad (3)$$

where k_{+0} is the rate constant between the inhibitor and the enzyme, K^{-1} is the inverted Michaelis constant and [S] the substrate concentration. As [Y] is known, plotting $\log([P_{inf.}] - [P])$ against t gives the apparent rate constant A from the slope and the second-order inhibition rate constant (k_{+0}) was deduced from eqn. (3).

Reactions were carried out under conditions in which less than 10 % of the substrate was hydrolysed before achievement of proteinase inhibition. Z-Phe-Arg-AMC (25 μM) was preincubated with various amounts of peptidyl diazomethylketone for 2 min at 30 °C. Enzyme reactions were started by adding cruzipain (0.2 nM). The rate of substrate hydrolysis was recorded

Table 1 Kinetic parameters of cruzipain from the Dm28c clone with Z-Phe-Arg-AMC and Z-Arg-Arg-AMC

Kinetic parameters for cruzipain were determined from Hanes linear plot. Experiments were carried out at 30 °C in 0.1 M phosphate buffer, pH 6.0, containing 10 mM DTT.

	K_m (μM)	$k_{cat.}$ (s^{-1})	$k_{cat.}/K_m$ ($mM^{-1} \cdot s^{-1}$)
Z-Phe-Arg-AMC	1.8	15.2	8590
Z-Arg-Arg-AMC	5.3	3.1	585

Table 2 Inhibition constants of cruzipain by members of cystatin families 2 and 3

Inhibition constants for cruzipain were determined using the method of Easson–Stedman [36]. Results are the means of at least two experiments which were carried out at 30 °C in 0.1 M phosphate buffer, pH 6.0, containing 10 mM DTT and using Z-Phe-Arg-AMC as substrate. Results taken from * [20] and † [23].

	K_i (nM)		
	Cruzipain	Cathepsin B	Cathepsin L
Chicken cystatin	0.22	1.7*	0.019*
Rat T-kininogen	0.06	80†	0.180†

continuously until complete inhibition. The Z-RLVG-CHN_2 concentration was 0–10 μM (final) and that of GlcA-QVVA-CHN_2 was 0–100 μM (final). Enzyme activity was assayed without inhibitor to check that the activated enzyme retained all its proteolytic activity.

RESULTS AND DISCUSSION

The reports demonstrating that inhibitors of cysteine proteinases can stop the development of protozoan parasites such as *T. cruzi*, *Plasmodium vinckei* and *Plasmodium falciparum* [17,18,33,34] have stimulated interest in these proteinases as they are putative targets for new chemotherapy. But such agents must be specific, so that they do not interfere with related proteinases from the host. The close structural and functional relationship between parasite cysteine proteinases and papain-related proteinases has meant that the substrates and inhibitors used, so far, to assay the activity of cruzipain are the same as those currently used for cathepsins B and L of mammalian lysosomes [1]. Better knowledge of the substrate specificity of parasitic cysteine proteinases and their sensitivity to inhibitors is therefore a prerequisite for the development of new drugs that discriminate between parasite and host cell proteinases.

Cruzipain activity and its inhibition by cystatin inhibitors

It has been shown recently that strain differences may occur in the overall structure of cruzipain isoforms [8]. Since non-conservative amino acid substitutions may occur in critical regions that may be involved in substrate recognition, the kinetic parameters for the hydrolysis of Z-Phe-Arg-AMC and Z-Arg-Arg-AMC by the proteinase isolated from the Dm28c clone were determined. The $k_{cat.}/K_m$ values for this proteinase (Table 1) are significantly higher than those found for cruzipain isolated from the Y strain, due to a higher $k_{cat.}$ value [24], and are in the same range as those reported for mammalian cathepsin L with the former substrate and for cathepsin B with the latter [35]. The temperature dependence of the hydrolysis (results not shown) confirmed previous studies showing that substrate inhibition occurred at 25 but not at 37 °C [24].

Table 3 Kinetic constants ($k_{cat.}/K_m$) for the hydrolysis of Abz-QVVAGA-EDDnp and Abz-LVGGA-EDDnp by cruzipain

$k_{cat.}/K_m$ constants were determined under pseudo-first-order conditions. Results are the means of at least three experiments which were carried out at 30 °C in 0.1 M phosphate buffer, pH 6.0, containing 10 mM DTT. Results taken from * [21] and † [29]. Substrate = Abz-peptidyl-EDDnp.

						$k_{cat.}/K_m$ ($mM^{-1} \cdot s^{-1}$)		
P4	P3	P2	P1′	P1′	P2′	Cruzipain	Cathepsin B	Cathepsin L
Q	V	V	A	G	A	3535	7*	515*
	L	V	G	G	A	11500	370†	735†

Cruzipain, like related cysteine proteinases, is inhibited by peptidyl diazomethylketones [17] and by peptidyl epoxides [10], but little is known about its inhibition by natural cystatin inhibitors [10]. The inhibitory properties of two representative members of the cystatin family (chicken cystatin and rat T-kininogen) [20] with respect to purified cruzipain were assayed using experimental conditions such that neither temperature nor substrate concentration were limiting [24]. K_i values were determined under tight-binding conditions using the Easson–Stedman plot [36]. Rat T-kininogen and chicken cystatin both inhibited cruzipain with K_i values lower than 1 nM (Table 2). Unlike mammalian cathepsins B, H and L and papain, which are all preferentially inhibited by chicken cystatin [20,23], cruzipain was more strongly inhibited by T-kininogen, indicating some differences in the mechanism of inhibition.

Hydrolysis of cystatin-derived, intramolecularly quenched, fluorogenic substrates by cruzipain

The crystal structures of chicken cystatin and the stefin B–papain complex, indicate that both the N-terminal glycine-containing sequence and the QXVXG consensus central region in cystatins, make intimate contact with the target proteinase [37,38]. These interactions contribute to the inhibitory activity of cystatins [22,39,40] but may vary from one proteinase to another [41,42]. Only the N-terminal sequence is substrate-like in cystatins [22], but the QVVAG segment was found to be sensitive to cleavage by papain at the A–G bond in synthetic peptides that mimicked cystatin surfaces [43]. We have exploited these features to develop new, fluorogenic peptide substrates of cysteine proteinases based on the sequences of these two segments. The resulting compounds are more specific for cysteine proteinases than for trypsin-like proteinases because they lack an arginyl residue at the P1 position, unlike the commonly used synthetic substrates. However, these substrates are efficiently cleaved by leucocyte elastase but not by other serine proteinases [21].They also have different specificities for individual cysteine proteinases from plants and mammals, suggesting that further modification or substitution in their peptide moiety may improve their specificity for particular cysteine proteinases. One approach is to use substrates with intramolecularly quenched fluorescence that are cleaved within their peptide moiety, so that the specificity may be investigated on both sides of the cleavage site. The activity of cruzipain was measured using the two leading substrates of the series (Abz-LVGGA-EDDnp and Abz-QVVAGA-EDDnp) designed for papain and mammalian lysosomal cysteine proteinases (Table 3). Both substrates were rapidly hydrolysed by cruzipain and were more sensitive to cruzipain hydrolysis than to hydrolysis by cathepsin L and cathepsin B (Table 3). Abz-QVVAGA-EDDnp

Table 4 Kinetic constants ($k_{cat.}/K_m$) for the hydrolysis of intramolecularly quenched fluorogenic substrates by cruzipain, and cathepsins B and L

$k_{cat.}/K_m$ constants were determined under pseudo-first-order conditions. Results are the means of at least three experiments which were carried out at 30 °C in 0.1 M phosphate buffer, pH 6.0, containing 10 mM DTT for cruzipain and in 0.1 M phosphate buffer, pH 6.0, containing 1 mM EDTA and 2 mM DTT for cathepsins B and L. * Results taken from [29]. Substrate = Abz-peptidyl-EDDnp.

						$k_{cat.}/K_m$ ($mM^{-1} \cdot s^{-1}$)		
P4	P3	P2	P1	P1′	P2′	Cruzipain	Cathepsin B	Cathepsin L
Q	V	V	A			380	24*	115*
Q	V	V	A	G		560	25*	110*
Q	V	V	A	G	G	340	6	55
Q	V	V	A	G	P	1270	5	25
	L	V	G	G	P	9090	17	1200
	V	V	G	G	P	4720	54	111

Table 5 Influence of the P1′ and P2′ residues

$k_{cat.}/K_m$ constants were determined under pseudo-first-order conditions. Results are the means of at least three experiments which were carried out at 30 °C in 0.1 M phosphate buffer, pH 6.0, containing 10 mM DTT for cruzipain and in 0.1 M phosphate buffer, pH 6.0, containing 1 mM EDTA and 2 mM DTT for cathepsins B and L Abbreviation: n.s.h., no significant hydrolysis. Substrate = Abz-peptidyl-EDDnp.

						$k_{cat.}/K_m$ ($mM^{-1} \cdot s^{-1}$)		
P3	P2	P1	P1′	P2′	P3′	Cruzipain	Cathepsin B	Cathepsin L
L	V	G	G	A	Q	1040	295	690
L	V	G	A	A	Q	5600	350	1240
L	V	G	F	A	Q	5210	295	450
L	V	G	D	A	Q	580	170	280
L	V	G	R	A	Q	27220	110	1380
L	V	G	G	R	Q	1210	110	330
L	V	G	G	D	Q	n.s.h	n.s.h	70
L	V	G	G	I	Q	730	310	1960
L	V	G	G	F	Q	1325	240	1555

was remarkably selective for cruzipain compared with mammalian cathepsin B which cleaved this substrate poorly. The catalytic site of cruzipain therefore differs from that of cathepsin B in spite of its reported cathepsin B-like properties [9,24]. The cleavage sites were identified by reverse-phase HPLC and amino acid analysis at the G–G and A–G bonds of Abz-LVGGA-EDDnp and Abz-QVVAGA-EDDnp respectively. The cleavage site within the peptide moiety of Abz-QVVAGA-EDDnp is identical with that observed for papain and lysosomal cathepsins [21].

The substrate specificity of cruzipain was further investigated by changing the length of the peptide moiety in the model substrates or by changing amino acid residues at critical positions on either side of the scissile bond. Parallel assays were done with lysosomal cathepsins B and L for comparison. The influence of P′ residues on the rate of cleavage was studied by shortening the peptide moiety of Abz-QVVAGA-EDDnp C-terminally, then replacing the P2′ Ala with Gly or Pro. A prolyl residue was inserted at the P2′ position because this residue is present at the cleavage site of the pro-region of cruzipain [6] and maturation of the proenzyme occurs through an autocatalytic process, as deduced from studies with recombinant cruzipain [6]. A proline at the P2′ position is also found at the presumed autocatalysed cleavage site in the hinge region that joins the catalytic domain to the C-terminal extension [11,12].

Cruzipain and cathepsin L behaved quite differently from cathepsin B when the QVVA-derived substrate was shortened C-terminally (Table 4). The $k_{cat.}/K_m$ of the parasite proteinase decreased significantly when the P2′ Ala or the P1′-P2′ subsites (i.e. Gly-Ala) were removed from Abz-QVVAGA-EDDnp, while $k_{cat.}/K_m$ for cathepsin B increased slightly. Replacing the P2′ Ala with Gly in Abz-QVVAGA-EDDnp resulted in a 10-fold decrease in the specificity constant for both cruzipain and cathepsin L (Table 4), indicating the importance of this subsite in the interaction. Remarkably, replacing the P2′ residue Ala with a Pro produced a more specific substrate for cruzipain than for cathepsins B and L. This is consistent with the presence of a prolyl residue at the P2′ position with respect to the two putative autocatalytic sites in the complete sequence of cruzipain [6]. However, assays with other substrates revealed that this same substitution did not induce any dramatic change in the specificity constant for cruzipain, as shown when replacing the P2′ Ala with Pro in Abz-LVGGA-EDDnp (Table 4).

The sequence at the hinge between the catalytic domain and the C-terminal extension of cruzipain [6] is very similar to the LVGG substrate-like region [22] in the cystatin-derived fluorogenic substrates. The only difference is a Val at position P3, where there is a leucyl residue in the cruzipain sequence. The corresponding substrate, Abz-VVGGP-EDDnp, was prepared and assayed as before. Substituting Leu for Val only slightly decreased the specificity constant for cruzipain and cathepsin B (Table 4), in agreement with the fact that Val and Leu have similar physicochemical properties. However, this change significantly decreased the specificity constant for cathepsin L by more than one order of magnitude (Table 4). This marked effect makes this substrate one of the two most selective for cruzipain, together with the corresponding QVVA derivative (Abz-QVVAGP-EDDnp). The specificity constant for Abz-QVVAGP-EDDnp is about 50-fold higher for cruzipain than for cathepsin L and 250-fold higher than for cathepsin B. Using Abz-VVGGP-EDDnp the increase is about 40-fold as compared with cathepsin L and 90-fold as compared with cathepsin B.

The P1′ and P2′ of peptide substrates deduced from the conserved sequences in cystatins and from presumed autocleavable sequences in cruzipain were then replaced by residues representative of the main amino acid groups. These two series of substrates were prepared by solid-phase synthesis in which an EDDnp-substituted glutamic acid was first coupled to the activated resin as described in the Materials and methods section. As a result, there was a Gln residue at position P3′ in all the substrates. The results show that placing an Arg residue at P1′ significantly increases the specificity constant for cruzipain, making the corresponding substrate the most sensitive of all substrates assayed here for this proteinase (Table 5). However, it does not allow further discrimination between the parasite proteinase and its mammalian homologues, as cathepsin L also accomodates this residue in P1′. Therefore the $k_{cat.}/K_m$ obtained with cruzipain is less than 20 times higher than that for cathepsin L. This is in agreement with previous findings showing that cathepsin L accomodates positively charged residues in P1′ [44]. Furthermore, this substitution introduces an additional cleavage site for both serine and cysteine proteinases. While cruzipain specifically cleaves Abz-LVGRAQ-EDDnp at the G–R site, as shown after C_{18} HPLC fractionation of the hydrolysis product and sequencing of the resulting peptides, trypsin hydrolyses this substrate at the R–A site. All three cysteine proteinases can accept a broad range of amino acids at P2′ since positively

charged, aromatic or hydrophobic residues are accommodated at that position (Table 5). Neither mammalian cathepsins nor cruzipain, however, accommodate a negatively charged residue in P2′ (Table 5). Incidentally the introduction of a Gln residue at P3′ in the lead substrate (Abz-LVGGAQ-EDDnp) induces a significant decrease in the $k_{cat.}/K_m$ constant for cruzipain but not for mammalian cathepsins (Table 5). This corroborates the observation of Ménard et al. [44] that the nature of residues in neighbouring positions may influence the binding at other subsites of the enzyme. These initial studies have therefore shown that cystatin-derived substrates more clearly discriminate between cruzipain and cathepsin B than between the parasite enzyme and cathepsin L. This limitation was partially overcome by changing amino acids at positions P3 and P2′ of the original substrate, to produce compounds with greater specificity for cruzipain. The similarity of cruzipain and cathepsin L is in agreement with previous studies that identified dipeptidyl diazomethylketone derivatives containing aromatic residues in the P1 and P2 positions as the most active inhibitors of purified cruzipain from Y strain [17]. Their effects were correlated with trypanocidal activity *in vitro* [17]. From our ability to discriminate cruzipain from mammalian cathepsin B with the cystatin-derived substrates, and from pH profiles of the proteinase isolated from the Y strain [24], it can be inferred that the presence of a Glu residue at the S2 subsite (position 205, papain numbering) in several genes encoding cruzipain [6–8] is not of major importance in the interaction with our substrates, whereas it is proposed to give to cathepsin B its specificity for positively charged residues at position P2 [45]. The kinetic parameters obtained with cruzipain isolated from the Dm28c clone of *T. cruzi* were significantly different from those for the Y-strain enzyme [24], as defined by assays with Z-Phe-Arg-AMC and Z-Arg-Arg-AMC. The enzyme from the Dm28c clone hydrolyses the fluorogenic substrate Z-Arg-Arg-AMC at a significant rate, but the specificity constant is about 15 times lower than that with Z-Phe-Arg-AMC, whereas this ratio is only about 3 for cathepsin B [35]. The ratio for the Y-strain enzyme is about 35 [24]. However the pH–activity profile of the Dm28c enzyme is significantly shifted with Z-Arg-Arg-AMC (Figure 1), as was previously shown for the Y-strain enzyme [24]. These observations strongly suggest that a Glu residue is present at position 205 in both proteinases [24], as this glutamyl residue may interact with the positively charged side-chain of the arginyl residue in position P2 of Z-Arg-Arg-AMC [45,46]. However, Glu-205 probably contributes less to the S2 subsite specificity of the Dm28c proteinase than to the proteinase purified from the Y strain. These kinetic differences may reflect differences in the repertoire of cruzipains expressed by *T. cruzi* strains or clonal populations [8], since non-conservative substitutions have been more frequently found in the catalytic domain than in the pro-region or C-terminal extension sequences. As some of the isoforms expressed by a given strain do not present Glu at position 205 (papain numbering), they may not display the typical cathepsin B-like specificity predicted for other isoforms [24]. The recently published phylogeny of the papain superfamily of cysteine proteinases [47] showed that the cathepsin B class of cysteine proteinases diverged very early from other classes of cysteine proteinases and that proteinases of *T. cruzi* and *T. brucei* belong to a class distinct from that of plant proteinases and of cathepsin L. This feature may facilitate the design of specific substrates of parasite proteinases. Accordingly, sequence comparison of all the main cysteine proteinases described so far, identified two distinct gene families from the presence of an interspersed amino acid motif that occurs in all pro-sequences, but not in those of cathepsin B-like proteinases [48]. This ERFNIN motif is found in the propeptide region of cruzipain, except that the two last residues are Ala in both cruzipain and its *T. brucei* homologue [13].

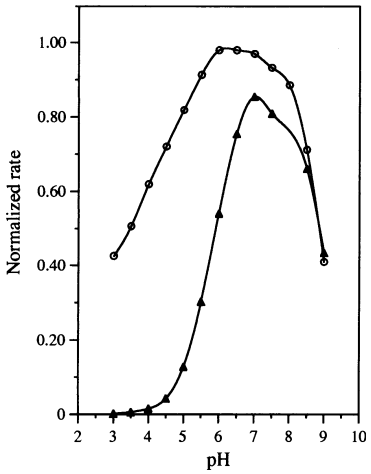

Figure 1 pH–activity profiles of cruzipain isolated from the Dm28c clone

Cruzipain (0.50 nM) was incubated for 5 min at 30 °C in pH 2–9 buffers and the enzyme activity was measured using Z-Phe-Arg-AMC (3.3 μM) (○) and Z-Arg-Arg-AMC (10 μM) (▲).

Cystatin-derived inhibitors of cruzipain

The presence of cruzipain in both the intracellular and extracellular developmental forms of *T. cruzi* has stimulated efforts to develop first-generation inhibitors and these have proved to be highly effective *in vitro* [17–19]. Some of these compounds have recently given encouraging results in mice infected with *T. cruzi* [49]. Diazomethylketone inhibitors have been used in some of these studies [17] because they are selective for cysteine proteinases [50]. An irreversible diazomethylketone inhibitor has also

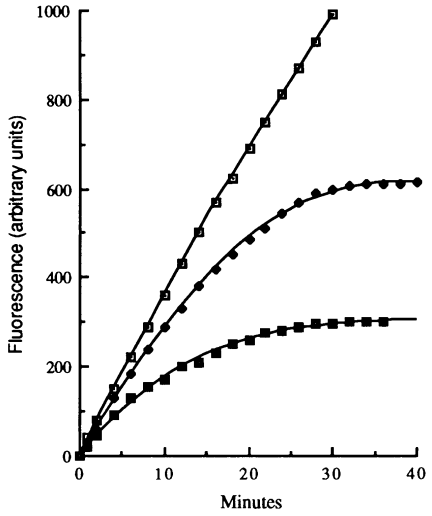

Figure 2 Progress curves for the generation of AMC from the cruzipain-catalysed hydrolysis of Z-Phe-Arg-AMC in the presence of the peptidyl diazomethylketone GlcA-QVVA-CHN$_2$

The inactivation of cruzipain was measured by monitoring continuously the increase of fluorescence, in the presence of a fixed concentration of substrate and various amounts of inhibitors (□, control; ◆, 50 μM; ■, 100 μM), as described in the Materials and methods section.

Table 6 Apparent second-order inhibition constants (k_{+0}) of cruzipain by peptidyl diazomethylketones

Apparent second-order constants (k_{+0}) were determined under first-order conditions. Experiments were carried out at 30 °C in 0.1 M phosphate buffer, pH 6.0, containing 10 mM DTT. * Results taken from [29].

	k_{+0} ($M^{-1} \cdot s^{-1}$)		
	Cruzipain	Cathepsin B	Cathepsin L
GlcA-QVVA-CHN_2	1900	5000*	18000*
Z-RLVG-CHN_2	5900	31000*	241000*

been designed from the LVG sequence of human cystatin C [51]. We synthesized a diazomethylketone inhibitor from the QVVA sequence and compared its ability to inhibit cruzipain and related mammalian proteinases with that of the LVG-containing inhibitor [51] (Figure 2). The water-solubility of the QVVA derivative was improved by adding a GlcA group to its N-terminus [52]. The k_{+0} constants for the inhibition of cruzipain by Z-RLVG-CHN_2 and GlcA-QVVA-CHN_2 are shown in Table 6. Both compounds inhibited cruzipain and cathepsins B and L quite efficiently. In contrast to the high sensitivity of cystatin-derived substrates for cruzipain, as compared with mammalian cathepsins, the inactivation of the parasite proteinase by both diazomethylketone inhibitors was far less efficient than that of cathepsins B and L. These results reinforce our previous observations that the prime region in the structurally-related substrates was responsible for the increased specificity towards cruzipain. This information, together with refinement of contributions determined by the non-prime specificity, will be useful for designing more selective inhibitors for the parasite proteinases, especially those containing a pseudo-peptide bond at the P1–P1′ site to take advantage of the prime specificity of cruzipain.

We thank Dr. Roger Mayer (Centre de Biophysique Moléculaire, Orléans, France) for synthesizing GlcA-QVVA-CHN_2 and Dr Magnus Abrahamson (University of Lund, Sweden) for providing Z-RLVG-CHN_2. This work was supported by PADCT-CNPq, FAPERJ, FINEP, CEPG-UFRJ, INSERM-CNPq. C.S holds an MRES doctoral fellowship.

REFERENCES

1 North, M. J., Mottram, J. C. and Coombs G. H. (1990) Parasitol. Today **6**, 270–275
2 Mac Kerrow, J. H., Sun, E., Rosenthal, P. J. and Bouvier, J. (1993) Annu. Rev. Microbiol. **47**, 821–853
3 Mac Kerrow, J. H. (1989) Exp. Parasitol. **68**, 111–115
4 Bontempi, E., Franke de Cazzulo, B. M., Ruiz, A. M. and Cazzulo, J. J. (1984) Comp. Biochem. Physiol. **77B**, 599–604
5 Scharfstein, J., Schechter, M., Senna, M., Peralta, J. M., Mendonça-Previato, L. and Miles, M. A. (1986) J. Immunol. **137**, 1336–1341
6 Eakin, A. E., Mills, A. A., Harth, G., Mac Kerrow, J. H. and Craik, C. S. (1992) J. Biol. Chem. **267**, 7411–7420
7 Campetella, O., Henriksson, J., Aslund, L., Frasch, A. C. C., Petterson, U. and Cazzulo, J. J. (1992) Mol. Biochem. Parasitol. **50**, 225–234
8 Lima, A. C. A., Tessier, D. C., Thomas, D. Y., Scharfstein, J., Storer, A. C. and Vernet, T. (1994) Mol. Biochem. Parasitol. **67**, 333–338
9 Cazzulo, J. J., Couso, R., Raimondi, A., Wernstedt, C. and Hellman, U. (1989) Mol. Biochem. Parasitol. **33**, 33–42
10 Murta, A. C. M., Persechini, P. M., Souto-Padron, T., De Souza, W., Guimaraes, J. A. and Scharfstein, J. (1990) Mol. Biochem. Parasitol. **43**, 27–38
11 Aslund, L., Henriksson, J., Campetella, O., Frasch, A. C. C., Pettersson, U. and Cazzulo, J. J. (1991) Mol. Biochem. Parasitol. **45**, 345–348
12 Hellman, U., Wernstedt, C. and Cazzulo, J. J. (1991) Mol. Biochem. Parasitol. **44**, 15–22
13 Mottram, J. C., North, M. J., Barry, J. D. and Coombs, G. H. (1989) FEBS Lett. **258**, 211–215
14 Schaffer, M. A. and Fischer, R. L. (1988) Plant Physiol. **87**, 431–436
15 Souto-Padron, T., Campetella, O. E., Cazzulo, J. J. and De Souza, W. (1990) J. Cell Sci. **96**, 485–490
16 Campetella, O., Martinez, J. and Cazzulo, J. J. (1990) FEMS Microbiol. Lett. **67**, 145–150
17 Meirelles, M. N. L., Juliano, L., Carmona, E., Silva, S. G., Costa, E. M., Murta, A. C. M. and Scharfstein, J. (1992) Mol. Biochem. Parasitol. **52**, 175–184
18 Harth, G., Andrews, N., Mills, A. A., Engel, J. C., Smith, R. and Mac Kerrow, J. H. (1993) Mol. Biochem. Parasitol. **58**, 17–24
19 Bonaldo, M. C., D'Escoffier, L. N., Salles, J. M. and Goldenberg, S. (1991) Exp. Parasitol. **73**, 44–51
20 Barrett, A. J., Rawlings, N. D., Davies, M. E., Machleidt, W., Salvesen, G. and Turk, V. (1986) in Proteinase Inhibitors: Cysteine Proteinase Inhibitors of The Cystatin Superfamily (Barrett, A. J. and Salvesen, G., eds.), pp. 515–569, Elsevier, Amsterdam
21 Gauthier, F., Moreau, T., Lalmanach, G., Brillard-Bourdet, M., Ferrer-Di Martino, M. and Juliano, L. (1993) Arch. Biochem. Biophys. **306**, 304–308
22 Machleidt, W., Thiele, U., Laber, B., Assfalg-Machleidt, I., Esterl, A., Wiegand, G., Kos, J., Turk, V. and Bode, W. (1989) FEBS Lett. **243**, 234–238
23 Moreau, T., Gutman, N., El Moujahed, A., Esnard, F. and Gauthier, F. (1988) Eur. J. Biochem. **173**, 185–190
24 Lima, A. C. A., Scharfstein, J., Storer, A. C. and Ménard, R. (1992) Mol. Biochem. Parasitol. **56**, 335–338
25 Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M. and Hanada, K. (1982) Biochem. J. **201**, 189–198
26 Anastasi, A., Brown, M. A., Kembhavi, A. A., Nicklin, M. J. H., Sayers, C. A., Sunter, D. C. and Barrett, A. J. (1983) Biochem. J. **211**, 129–138
27 Chagas, J. R., Juliano, L. and Prado, E. (1991) Anal. Biochem. **192**, 419–425
28 Oliveira, M. C. F., Hiraya, L. Y., Chagas, J. R., Boschcov, P., Gomes, R. A. S., Figueredo, A. F. S. and Juliano, L. (1992) Anal. Biochem. **203**, 39–46
29 Lalmanach, G., Serveau, C., Brillard-Bourdet, M., Chagas, J. R., Juliano, L., Mayer, R. and Gauthier F. J. (1995) Protein. Chem. **14**, 645–653
30 Henderson, P. J. F. (1972) Biochem. J. **127**, 321–333
31 El Moujahed, A., Gutman, N., Brillard, M. and Gauthier, F. (1990) FEBS Lett. **265**, 137–140
32 Tian, W. X. and Tsou, C. L. (1982) Biochemistry **21**, 1028–1032
33 Rosenthal, P. J., Lee, G. K. and Smith, R. E. (1993) J. Clin. Invest. **91**, 1052–1056
34 Mayer, R., Picard, I., Lawton, P., Grellier, P., Barrault, C., Monsigny, M. and Schrével, J. (1991) J. Med. Chem. **34**, 3029–3035
35 Barrett, A. J. and Kirschke, H. (1981) Methods Enzymol. **80**, 535–560
36 Easson, L. H. and Steadman, E. (1936) Proc. R. Soc. London Ser. B **121**, 141–164
37 Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J. and Turk, V. (1988) EMBO J. **7**, 2593–2599
38 Stubbs, M. Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B. and Turk, V. (1990) EMBO J. **9**, 1939–1947
39 Arai, S., Watanabe, H., Kondo, H., Emori, Y. and Abe, K. (1991) J. Biochem. (Tokyo) **109**, 294–298
40 Bjork, I., Pol, E., Raub-Segall, E., Abrahamson, M., Rowan, A. D. and Mort, J. (1994) Biochem. J. **299**, 219–225
41 Genenger, G., Lenzen, S., Mentele, R., Assfalg-Machleidt, I. and Auerswald, E. A. (1991) Biomed. Biochim. Acta **50**, 4–6
42 Auerswald, E. A., Genenger, G., Assfalg-Machleidt, I., Machleidt, W., Engh, R. and Fritz, H. (1992) Eur. J. Biochem. **209**, 837–845
43 Lalmanach, G., Hoebeke, J., Moreau, T., Brillard-Bourdet, M., Ferrer-Di Martino, M., Borras-Cuesta, F. and Gauthier, F. (1993) J. Protein Chem. **12**, 23–31
44 Ménard, R., Carmona, E., Plouffe, C., Brömme, D., Konishi, Y., Lefebvre, J. and Storer, A. C. (1993) FEBS Lett. **328**, 107–110
45 Khouri, H. E., Plouffe, C. Hasnain, S., Hirama, T., Storer A. C. and Ménard, R. (1991) Biochem. J. **275**, 751–757
46 Khouri, H. E., Vernet, T., Ménard, R., Parlati, F., Laflamme, P., Tessier, D. C., Gour-Salin, B., Thomas, D. Y. and Storer, A. C. (1991) Biochemistry **30**, 8929–8936
47 Berti, P. J. and Storer, A. C. (1995) J. Mol. Biol. **246**, 273–283
48 Karrer, K. M., Peiffer, S. L. and DiTomas, M. E. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 3063–3067
49 McGrath, M. E., Eakin, A. E., Engel, J. C., McKerrow, J.H, Craik, C. S. and Fletterick, R. J. (1995) J. Mol. Biol. **247**, 251–259
50 Green, G. D. J. and Shaw, E. (1981) J. Biol. Chem. **256**, 1923–1928
51 Hall, A., Abrahamson, M., Grubb, A., Trojnar, J., Kania, P., Kasprzykowska, R. and Kasprzykowski, F. (1992) J. Enzyme Inhib. **6**, 113–123
52 Harnois-Pontoni, M., Monsigny, M. and Mayer, R. (1991) Anal. Biochem. **193**, 248–255

Received 5 June 1995/25 September 1995; accepted 4 October 1995