Probing the specificity of cysteine proteinases at subsites remote from the active site: analysis of P₄, P₃, P₂' and P₃' variations in extended substrates

Fernada C. Vieira PORTARO*, Ana Beatriz F. SANTOS*, Maria Helena S. CEZARI†, Maria Aparecida JULIANO*, Luiz JULIANO†¹ and Euridice CARMONA*

*Department of Pharmacology, Instituto Butantan, 05503-900, São Paulo, Brazil, and †Departamento de Biofísica, Escola Paulista de Medicina, Rua Tres de Maio 100, São Paulo 04044-020, Brazil

We have determined the kinetic parameters for the hydrolysis by papain, cathepsin B and cathepsin L of internally quenched fluorescent peptides derived from the lead peptides Abz-AAFRSAQ-EDDnp [in which Abz and EDDnp stand for *o*-aminobenzoic acid and *N*-(2,4-dinitrophenyl)ethylenediamine respectively], to map the specificity of S₄ and S₃ subsites, and Abz-AFRSAAQ-EDDnp, to identify the specificity of S₂' and S₃'. Abz and EDDnp were the fluorescent quencher pair. These two series of peptides were cleaved at the Arg–Ser bond and systematic modifications at P₄, P₃, P₂' and P₃' were made. The S₄ to S₂' subsites had a significant influence on the hydrolytic efficiencies of the three enzymes. Only papain activity was observed to be dependent on S₃', indicating that its binding site is larger than those of cathepsins B and L. Hydrophobic amino acids were accepted at S₄, S₃, S₂' and S₃' of the three enzymes.

The best substrates for cathepsins L and B had Trp and Asn at P₂' respectively; variations at this position were less accepted by these enzymes. The best substrates for papain were peptides containing Trp, Tyr or Asn at P₃'. Basic residues at P₃ and P₄ were well accepted by cathepsin L and papain. We also explored the susceptibility of substrates Abz-AFRSXAQ-EDDnp, modified at P₂' (X), to human cathepsin B mutants from which one or two occluding loop contacts had been removed. The modifications at His¹¹¹ (H111A) and His¹¹⁰ (H110A) of cathepsin B led to an increase in *k*_{cat} values of one or two orders of magnitude. The hydrolytic efficiencies of these cathepsin B mutants became closer to those of papain or cathepsin L.

Key words: cathepsins B and L, fluorogenic peptide, limited proteolysis, papain, protease substrate.

INTRODUCTION

Cysteine proteinases of the papain superfamily are widely distributed in Nature; they can be found in both prokaryotes and eukaryotes, e.g. bacteria, parasites, plants, invertebrates and vertebrates [1]. In mammals, the major cysteine proteinases are the lysosomal cathepsins B and L. They are implicated in many physiological processes such as protein degradation [2], antigen presentation [3] and bone resorption [4] but they also have been implicated in a number of degradative and invasive processes such as arthritis [5], tumour invasion and metastasis [6] and muscular dystrophy [7]. The biological function of papain, which was first described in 1879 [8], is still not well defined but this enzyme has been extensively characterized and much information on cysteine proteases mechanism was obtained from papain studies (reviewed in [9–11]).

The kinetic properties of these different cysteine proteinases have been studied by using substrates and inhibitors, most extensively for papain, and cathepsins B and L. It has been well established that the primary determinant of specificity for papain and cathepsins B and L is the S₂ subsite [12,13]. Hydrophobic residues are preferred at the P₂ position of substrates for papain and cathepsin L, but cathepsin B also accepts basic residues there. This difference is due to the presence of a glutamic residue at S₂ of cathepsin B [14,15]. The other non-prime subsites were also studied with the use of small chromogenic or fluorogenic synthetic substrates and also by synthetic inhibitors. The S₁ subsites of papain and cathepsin L are not as selective as S₂, accepting a wide range of residues; however, again, cathepsin B

is different because it shows a preference for non-bulky residues [16–18]. S₃ and S₄ have been less studied and only a few results are available [19]. A similar lack of information exists for the prime subsites [20–22]. Several three-dimensional structures of cysteine proteinases of the papain superfamily are available and all share a common fold, consisting of two domains separated by the active-site cleft. A recent review of kinetic and structural data has reconsidered the definition of substrate-binding sites in papain-like cysteine proteases [23].

In the present study we revisited the specificity of papain and cathepsins B and L by using intramolecularly quenched fluorogenic substrates derived from the lead peptide Abz-AAFRSAQ-EDDnp [in which Abz and EDDnp stand for *o*-aminobenzoic acid and *N*-(2,4-dinitrophenyl)ethylenediamine respectively], to map the specificity of the S₄ and S₃ subsites, and Abz-AFRSAAQ-EDDnp, to identify the specificity of S₂' and S₃'. Abz was the fluorescent donor group and EDDnp the quencher group. Because these two peptides were cleaved at the Arg–Ser bond, systematic modifications in the substrates at P₄, P₃, P₂' and P₃' were performed, synthesizing four families of peptides. The activities of cathepsin B mutants in which occluding loop contacts had been modified [24] were examined with the substrate series Abz-AFRSXAQ-EDDnp. Two single-mutant enzymes, at His¹¹⁰ (H110A) and at His¹¹¹ (H111A), one double mutant at Asp²² and His¹¹⁰ (D22A/H110A) and one triple mutant at Asp²², His¹¹⁰ and R¹¹⁶ (D22A/H110A/R116A) were assayed. These mutations altered the stability of the occluding loop of cathepsin B because His¹¹⁰ and Arg¹¹⁶, which are part of the loop, form salt bridges with Asp²² and Asp²²⁴ respectively [25,26]. In addition,

Abbreviations used: Abz, *o*-aminobenzoic acid; E-64, *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane; EDDnp, *N*-(2,4-dinitrophenyl)-ethylenediamine; MALDI-TOF-MS, matrix-assisted laser desorption/ionization-time-of-flight MS; TFA, trifluoroacetic acid.

¹ To whom correspondence should be addressed (e-mail juliano.biof@epm.br).

the imidazole group of His¹¹⁰ and His¹¹¹ provide the positive charge for the C-terminal carboxylate interaction of inhibitors [27,28] and substrates [29].

MATERIALS AND METHODS

Enzymes

Human recombinant cathepsin B and rat recombinant cathepsin B (EC 3.4.22.1) were obtained as described previously [24,30]. Human cathepsin L (EC 3.4.22.15) and papain (EC 3.4.22.2) were obtained as described in [31] and [32] respectively. Human cathepsin B mutants were kindly provided by Dr Robert Ménard and Dr Dorit K. Nägler (Biotechnology Research Institute, NRC, Montreal, QC, Canada). The molar concentrations of the enzyme solutions were determined by active-site titration with *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane (E-64) as described in [33].

Peptide substrates

All the intramolecularly quenched fluorogenic peptides contained EDDnp attached to glutamine, a necessary result of the solid-phase peptide synthesis strategy employed, the details of which are provided elsewhere [34]. An automated benchtop simultaneous multiple solid-phase peptide synthesizer (a PSSM 8 system from Shimadzu) was used for the solid-phase synthesis of all the peptides by the Fmoc (fluoren-9-ylmethoxycarbonyl) procedure. The final deprotected peptides were purified by semipreparative HPLC with an Econosil C₁₈ column (10 μ m, 22.5 mm $\times$ 250 mm) and a two-solvent system: solvent A, trifluoroacetic acid (TFA)/water (1:1000, v/v); solvent B, TFA/acetonitrile/water (1:900:100, by vol.). The column was eluted at a flow rate of 5 ml/min with a 10% (or 30%) to 50% (or 60%) gradient of solvent B over 30 or 45 min. Analytical HPLC was performed with a binary HPLC system from Shimadzu with a SPD-10AV Shimadzu UV-visible detector and a Shimadzu RF-535 fluorescence detector, coupled to an Ultrasphere C₁₈ column (5 μ m, 4.6 mm $\times$ 150 mm) that was eluted with solvent A1 [H₃PO₄/water (1:1000, v/v)] and B1 [acetonitrile/water/H₃PO₄ (900:100:1, by vol.)] at a flow rate of 1.7 ml/min and a 10–80% gradient of B1 over 15 min. The HPLC column eluates were monitored by *A*₂₂₀ and by fluorescence emission at 420 nm after excitation at 320 nm. The molecular masses and purities of synthesized peptides were checked by matrix-assisted laser desorption ionization–time-of-flight (MALDI–TOF) MS (ToFSpec-E, Micromass).

HPLC analysis of the enzymic hydrolysis products of the synthetic fluorogenic substrates

The peptide solutions (50–100 μ M), in 50 mM phosphate buffer/200 mM NaCl/2.5 mM dithioerythritol/5 mM EDTA at pH 5.5 for cathepsin L, at pH 6.0 for cathepsin B and at pH 6.5 for papain, were incubated with the different proteinases at 37 °C. Samples (20 μ l) of the incubation reactions were removed periodically for HPLC analysis until 100% hydrolysis was reached. Substrate alone under the same reaction conditions was used as control. The hydrolysis products were separated by HPLC and subjected to amino acid analysis and MALDI–TOF-MS. The scissile bonds were deduced from the sequences of the substrate fragments. The HPLC conditions used for the analytical procedure were TFA in water (solvent A) and acetonitrile/solvent A (9:1, v/v) as solvent B. The separations were performed at a flow rate of 1 ml/min on a J. T. Baker C₁₈ column (4.5 mm $\times$ 300 mm).

Enzyme assays

Hydrolysis of the fluorogenic peptide substrates at 37 °C, in 50 mM phosphate buffer/200 mM NaCl/2.5 mM dithioerythritol/5 mM EDTA at pH 5.5 for cathepsin L, at pH 6.0 for cathepsin B and at pH 6.5 for papain, was monitored by measuring the fluorescence emission at 420 nm after excitation at 320 nm in a Hitachi F-2000 spectrofluorometer, as described previously [35].

The standard hydrolysis conditions were strictly maintained for different substrates. The enzyme concentration varied from 0.006 nM for the best substrates to 12 nM for the least susceptible substrates. In most cases the concentrations of the substrates ranged from 0.1 *K*_m to 10 *K*_m. The kinetic parameters were calculated as described in [36]. The S.E.M. of *K*_m and *k*_{cat} determinations was in each case never greater than 5% of the value obtained.

Amino acid analysis

The amino acid compositions of the peptide substrates and their fragments were determined as follows: samples were freeze-dried and hydrolysed for 8 h at 110 °C in 6 M HCl containing 1% (w/v) phenol in vacuum-sealed tubes, then subjected to amino acid analysis with a pico Tag station [37].

RESULTS AND DISCUSSION

All the substrates assayed in the present study derived from the lead peptides Abz-AAFRSAQ-EDDnp and Abz-AFRSAAQ-EDDnp and were cleaved at the Arg-Ser bond by the three cysteine proteinases. All the kinetics of hydrolysis were fitted to the hyperbolic Michaelis–Menten rate equation; the kinetic parameters are presented in Figures 1–4 as bar graphs (*k*_{cat}/*K*_m) for better visualization of the differences in specificity for each amino acid substitution. The *K*_m values are shown on the tops of individual bars. In comparison with cathepsin L and papain, all the assayed substrates were poorly hydrolysed by cathepsin B, yielding *k*_{cat}/*K*_m values three orders of magnitude lower than those of the other two proteinases, mainly owing to the very low *k*_{cat}. Throughout the discussion below, papain numbering has been used on the residues of three-dimensional structures of the enzymes for a better comparison of the subsite characteristics.

S₂ subsite specificity

To verify whether the specificity previously described for S₂ remained for this class of substrates, we included in this study a limited series of quenched fluorogenic substrates with variations at P₂ (Abz-AXRSAAQ-EDDnp, where X represents Phe, Leu, Ala, Pro or Arg). Table 1 shows the kinetic parameters for hydrolysis of these substrates by cathepsin B, cathepsin L and papain. The *K*_m values were very similar for the hydrolysis of all studied substrates by the three proteinases, except the hydrolysis by cathepsin L of the substrate with X = Leu and the hydrolysis by cathepsin B of the substrate with X = Pro, which were an order of magnitude higher and lower respectively than the other substrates of the series. In contrast, *k*_{cat}/*K*_m ratios were very different owing to the differences in *k*_{cat} values. Cathepsin B cleaved the substrate with Arg and Phe at P₂ with similar efficiency, whereas cathepsin L and papain cleaved the former substrate with specificities approx. 1/160 those of the latter. The papain S₂ subsite seems to be more restrictive to Phe than cathepsin L, because papain hydrolysed the peptide containing Phe with a *k*_{cat}/*K*_m value 15-fold that with Leu, whereas with cathepsin L this difference was only 2-fold. In addition, papain hydrolysed peptides containing Leu or Ala with similar *k*_{cat}/*K*_m

Table 1 Kinetic parameters for the hydrolysis of substrates Abz-AXRSAQ-EDDnp (in which X represents Phe, Leu, Ala, Pro or Arg) by human cathepsin B, human cathepsin L and papain, at 37 °C in 50 mM phosphate (pH 6.0)/200 mM NaCl/5 mM EDTA/2.5 mM dithioerythritol

Enzyme	P ₂ residue	<i>K_m</i> (μM)	<i>k_{cat}</i> (s ⁻¹)	<i>k_{cat}/K_m</i> (mM ⁻¹ ·s ⁻¹)
Cathepsin B	Phe	4.1	0.068	17
	Leu	4.2	0.067	16
	Ala	6.2	0.032	5
	Pro	15	0.033	2
	Arg	2.0	0.029	14
Cathepsin L	Phe	1.9	6.4	3368
	Leu	0.56	1.0	1786
	Ala	1.8	0.7	389
	Pro	3.4	0.005	1.5
	Arg	3.1	0.07	22
Papain	Phe	1.1	3.4	3091
	Leu	2.1	0.5	238
	Ala	2.5	0.6	240
	Pro	4.5	0.09	20
	Arg	5.7	0.1	18

Figure 1 Specificity constant (k_{cat}/K_m) for the hydrolysis of Abz-AXFRSAQ-EDDnp by papain, rat cathepsin B and human cathepsin L

The amino acids at position X (P₃) are indicated on the x-axis. The numbers above the bars are the K_m values in μM.

values, indicating that for this enzyme the aromaticity of the substrate side chain is a determinant element for the P₂-S₂ interaction.

It has been well established that the primary determinant of specificity for papain and cathepsins B and L is the S₂ subsite [12,13]. For cathepsin B, the residues that contribute to the enzyme-substrate interaction at the S₂ subsite are Pro⁶⁸, Ala⁶⁹, Ala¹³³, Gly¹⁵⁷, Ala¹⁶⁰ and Glu²⁰⁵ (papain numbering) [12,15]; this enzyme can therefore accommodate Arg at S₂, owing to the presence of Glu²⁰⁵. The papain S₂ subsite consists mainly of hydrophobic residues, Pro⁶⁸, Val¹³³, Val¹⁵⁷ and Phe²⁰⁷ [38]. The specificity of the papain S₂ subsite has recently been re-examined [39]; although the aromaticity of amino acids at the P₂ position of substrates was reported not to be essential, lower K_m values were obtained with substrates containing Phe, particularly those with a substitution on the phenyl group by Cl or NO₂. According to these observations, Tyr(NO₂) was selected at the P₂ position with high frequency by screening a substrate peptide library for papain [40]. Cathepsin L showed the same pattern as papain for the examined substitutions at the P₂ substrate sites [1]; however, Ala occupies position 207, in contrast with papain, in which the residue is Phe. The absence of an aromatic residue at S₂ of cathepsin L seems to decrease its selectivity for aromatic amino acids, as observed with papain. The present results are in accordance with those available on the characterization of S₂ for the enzymes studied. The effect of variations in P₂ was less pronounced than that observed for small substrates [14] but the overall pattern was maintained.

S₃ subsite specificity

Figure 1 shows the kinetic parameters for the hydrolysis of the quenched fluorogenic substrates with substitutions at P₃ in the series Abz-AXFRSAQ-EDDnp by papain and cathepsins B and L. In general, the k_{cat}/K_m values indicated that all three enzymes prefer hydrophobic residues at this position. The orders of preference, considering only the hydrophobic residues, were: Leu Tyr Phe, Trp for cathepsin B, Phe, Trp Leu Tyr for cathepsin L, and Phe, Leu Trp Tyr for papain. The substrates with His and Lys at the P₃ position were quite susceptible to hydrolysis by cathepsin L. In contrast, the peptide with His was the worst substrate for papain examined in this work owing to the very low k_{cat} . Acidic residues at P₃ decreased the hydrolysis of the substrate, particularly by cathepsin L. A proline residue interferes mainly with the substrate affinity, particularly towards papain, whose K_m was higher for the other substitutions.

The X-ray crystallographic structure of cathepsin B suggests that the S₃ pocket involves at least two residues, Asp⁶¹ and Tyr⁶⁷ [15,41,42]. We could therefore expect that at the S₃ subsite both electrostatic and hydrophobic interactions would occur. Indeed, our results with cathepsin B are in accordance with this view because the peptide with Lys at P₃ was hydrolysed by cathepsin B with the lowest K_m , and the Leu-containing or Phe-containing peptides were hydrolysed with the highest k_{cat}/K_m values in the series (Figure 1). These findings are corroborated by a previous characterization of the S₃ specificity of cathepsin B with tripeptidyl-7-amido-4-methylcoumarin and molecular modelling analysis [19], in which the higher k_{cat}/K_m values were obtained with substrates containing Tyr and Arg at the P₃ position; an aromatic-aromatic interaction was proposed between a substrate Tyr residue and enzyme Tyr⁶⁷. In addition, the crystallographic data support electrostatic interactions involving Asp⁶¹ [15] or Glu²⁰, as indicated in the proregion interaction in cathepsin B [41].

Figure 2 Specificity constant (k_{cat}/K_m) for the hydrolysis of Abz-XAFRSAQ-EDDnp by papain, rat cathepsin B and human cathepsin L

The amino acids at position X (P_4) are indicated on the x-axis. The numbers above the bars are the K_m values in μ M.

In papain, Tyr⁶¹ and Tyr⁶⁷ are part of its hydrophobic S_3 subsite [43–46], justifying the preference of papain for Leu and Phe at P_3 (Figure 1). Although papain hydrolysed the Lys-containing substrates quite efficiently, X-ray studies on the papain–E-64 complex [43] did not show electrostatic interaction at the papain S_3 subsite.

Our results with cathepsin L indicate a preference for substrates with bulky hydrophobic as well as positively charged residues, but not Asp (Figure 1). This is in accordance with the crystal structure of human cathepsin L complexed with E-64 [47], which shows a large S_3 pocket formed by the amino acids Asn⁶⁶, Glu⁶³ and Leu⁶⁹.

S_4 subsite specificity

Figure 2 shows the kinetic constants for the hydrolysis of the quenched fluorogenic substrates with substitutions at the P_4 position in the series Abz-XAFRSAQ-EDDnp by papain and cathepsins B and L. The substrates with hydrophobic residues Phe and Leu were hydrolysed with the highest k_{cat}/K_m values for the three proteinases, owing to lower K_m values for cathepsin B and higher k_{cat} values for papain and cathepsin L. Papain and

cathepsin L accepted Lys and His at the P_4 position quite well, with a preference for the latter. With regard to the susceptibility of papain, it is interesting to note that this very favourable effect of His at P_4 contrasts with the resistance of the peptide with His at the P_3 position. In contrast, the peptide with Asp at P_4 was poorly hydrolysed by cathepsin L and papain. However, no significant differences were observed in the hydrolysis by cathepsin B of the substrates containing either basic or acidic residues at P_4 . Little structural information is available from X-ray crystallographic studies on the S_4 subsites of papain and cathepsins B and L; it was uncertain whether this subsite would participate in the binding [15,40,41]. However, our results indicate that P_4 – S_4 interaction exists within the three proteinases because significant differences were observed in hydrolytic efficiency towards substrates with variations at the P_4 position. The S_4 subsite of cathepsin B seems to accept a hydrophobic residue in preference to any hydrophilic residue. X-ray studies of the complex of papain with Z-GPG-CHN₂ (in which Z stands for benzyloxycarbonyl) showed evidence of the interaction between the Z group and the papain R domain, in which the Z group at P_4 was interacting with Val¹⁵⁷ and Ser²⁰⁹ [48].

S_2' subsite specificity

Figure 3 shows the kinetic constants for the hydrolysis of the quenched fluorogenic substrates with substitutions at P_2' in the series Abz-AFRSXAQ-EDDnp. The most significant observation with cathepsin B is the variation in K_m values with the different amino acids at the X position. The lowest K_m with cathepsin B (0.81 μ M) was observed with the hydrolysis of substrate containing Asn; the highest (7.6 μ M) was observed with the substrate containing Pro, although His, Lys and Phe at P_2' also resulted in substrates hydrolysed with K_m values higher than 4 μ M. Cathepsin L hydrolysed the peptide Abz-AFRSWAQ-EDDnp with the highest k_{cat}/K_m value of all substrates assayed in this study. In comparison with other substrates of this series, the k_{cat}/K_m for this peptide containing Trp at P_2' was due mainly to a low K_m . The nature of amino acid at position P_2' therefore seems to be critical for the affinity of substrate for cathepsins B and L. The latter enzyme also accepted the other hydrophobic amino acids, Tyr, Phe and Leu, quite well. Papain also showed a preference for the substrates with hydrophobic residues (Trp, Phe, Leu and Tyr), followed by the basic amino acids Lys and His. The S_2' subsites of cathepsin L and papain, as with S_3 and S_4 , did not accept an acidic residue. It is well known, on the basis of all the X-ray structures available for cathepsin B, that the occluding loop is located towards the prime region of the active site of this enzyme. The contacts between compound CA-030, an E-64 epoxy-analogue, and the occluding loop have been reported [27]; the interaction of His¹¹⁰ and His¹¹¹ with the P_2' residue of the inhibitor, particularly the C-terminal carboxy group, was clearly demonstrated. In addition, it is also accepted that the occluding loop is responsible for the low k_{cat} values for cathepsin B endopeptidase activity observed in this present paper as well as in previous reports [24,26,49,50]. Hydrophobic interactions might also occur on S_2' because substrates with Trp and Tyr at P_2' were well accepted by cathepsin B (Figure 3). In accordance with this view, the procathepsin B X-ray structure revealed an interaction of Leu⁴¹ of the propeptide with Trp²²¹ and His¹¹⁰ of the occluding loop [42].

S_2' of cathepsin L is preferentially hydrophobic, accepting mainly Trp, followed by Leu and Tyr. This pattern was also observed on the X-ray structure of procathepsin L, in which Met⁷⁵ of the propeptide interacted with Trp¹⁷⁷ on the core of the enzyme molecule [50], and also by kinetic studies [51]. Papain S_2'

Figure 3 Specificity constant (k_{cat}/K_m) for the hydrolysis of Abz-AFRSXAQ-EDDnp by papain, rat cathepsin B and human cathepsin L

The amino acids at position X (P_2') are indicated on the x-axis. The numbers above the bars are the K_m values in μM .

seems to be similar to that of cathepsin L; it preferred hydrophobic residues but also accepted basic amino acids such as Lys and His. The available X-ray structures of papain did not show the S' sites but by analogy with the procathepsin L structure it can be suggested that Met¹⁷⁷ would participate on this pocket, giving it its hydrophobic character.

S_3' subsite specificity

Figure 4 shows the kinetic constants for the hydrolysis of the quenched fluorogenic substrates with substitutions at P_3' in the series Abz-AFRSXAQ-EDDnp by papain and cathepsins B and L. At this position, cathepsins B and L did not show any clear preference between the different amino acids used, suggesting that substrate P_3' does not interact with these enzymes. In contrast, the best substrates for papain were observed in this series. The highest k_{cat}/K_m values were obtained with the hydrophobic residues Trp, Phe, Leu and Tyr at P_3' . The peptides with Asn and Asp were also well hydrolysed by papain. In accordance with this observation is the high-frequency selection of acidic amino acids at the P_3' position on screening of a substrate library for papain [40]. Although there are no available

Figure 4 Specificity constant (k_{cat}/K_m) for the hydrolysis of Abz-AFRSXAQ-EDDnp by papain, rat cathepsin B and human cathepsin L

The amino acids at position X (P_3') are indicated on the x-axis. The numbers above the bars are the K_m values in μM .

structural data related to the S_3' subsite, our results indicate that papain's binding site is larger than that of cathepsin L or B, which is in accordance with the classical proposal of an extended binding site (P_4 to P_3') for papain [51].

Susceptibility of the Abz-AFRSXAQ-EDDnp substrate series to human cathepsin B mutants

The mutants of cathepsin B examined in this work have in common modifications on the occluding loop (H111A and H110A) as well in residues outside the loop (D22A/H110A and D22A/H110/R116A) that anchor it in position in the native enzyme [29]. The kinetic constants for the hydrolysis of substrates from the series Abz-AFRSXAQ-EDDnp by human wild-type cathepsin B (Table 2) were very similar to those obtained with rat cathepsin B (Figure 3). The modification at His¹¹¹ (H111A) and His¹¹⁰ (H110A) of cathepsin B led to increases in the hydrolytic activity towards all peptides of the series, which was due mainly to an increase in k_{cat} values by one or two orders of magnitude.

H110A and the double (D22A/H110A) and triple (D22A/H110A/R116A) mutants hydrolysed the substrates containing

Table 2 Kinetic parameters for hydrolysis of substrates Abz-AFRSXAQ-EDDnp by human wild-type and mutant cathepsins B at 37 °C in 50 mM phosphate (pH 6.0)/200 mM NaCl/5 mM EDTA/2.5 mM dithioerythritol

Variable residue	Cathepsin B					
	Kinetic parameter	Wild-type	H111A	H110A	D22A/H110A	D22A/H110A/R116A
Trp	K_m (μ M)	2.2	2.6			
	k_{cat} (s^{-1})	0.039	1.1			
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	18	423	722*	7583*	24 083*
Tyr	K_m (μ M)	3.8	2.3			
	k_{cat} (s^{-1})	0.051	2.7			
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	13	1174	976*	7805*	23 662*
Phe	K_m (μ M)	4.2	3.9			
	k_{cat} (s^{-1})	0.045	2.7			
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	11	1421	1036*	12 187*	25 541*
Leu	K_m (μ M)	2.8	3.2			
	k_{cat} (s^{-1})	0.032	3.6			
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	11	1125	2466*	18 835*	20 583*
Ala	K_m (μ M)	4.1	6.3			
	k_{cat} (s^{-1})	0.068	1.7			
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	17	270	1230*	9000*	10 900*
Ser	K_m (μ M)	1.8	1.4	0.62	0.53	0.42
	k_{cat} (s^{-1})	0.031	1.5	1.4	4.3	6.5
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	17	1071	2258	8113	15 476
Asn	K_m (μ M)	0.7	1.3	0.54	0.41	0.38
	k_{cat} (s^{-1})	0.027	1.0	1.0	3.1	6.2
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	38	769	1852	7561	16 316
Asp	K_m (μ M)	1.6	2.1	0.71	0.53	0.64
	k_{cat} (s^{-1})	0.011	0.13	0.11	2.1	3.1
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	7	62	155	3962	4844
His	K_m (μ M)	7.1	5.1	1.2	0.92	1.1
	k_{cat} (s^{-1})	0.046	0.62	0.35	2.9	5.8
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	6	122	292	3152	5273
Pro	K_m (μ M)	9	9	3.5	2.8	2.3
	k_{cat} (s^{-1})	0.070	0.12	0.11	1.1	2.3
	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	8	15	31	393	1000

* Double cleavage, as indicated in Figure 5.

Abz- A-F-R- $\uparrow$ -S- $\downarrow$ -W-A-Q-EDDnp ($\uparrow$)46 % ($\downarrow$)60 %Abz- A-F-R- $\uparrow$ -S- $\downarrow$ -Y-A-Q-EDDnp ($\uparrow$)45 % ($\downarrow$)55 %Abz- A-F-R- $\uparrow$ -S- $\downarrow$ -F-A-Q-EDDnp ($\uparrow$)64 % ($\downarrow$)36 %Abz- A-F-R- $\uparrow$ -S- $\downarrow$ -L-A-Q-EDDnp ($\uparrow$)78 % ($\downarrow$)22 %Abz- A-F-R- $\uparrow$ -S- $\downarrow$ -A-A-Q-EDDnp ($\uparrow$)45 % ($\downarrow$) 55 %**Figure 5** Peptide bonds cleaved by human H110A cathepsin B mutant on the substrates Abz-AFRSXAQ-EDDnp (in which X represents Trp, Tyr, Phe, Leu and Ala)

An upward arrow ($\uparrow$) indicates cleavage at the Arg–Ser bond and a downward arrow ($\downarrow$) indicates cleavage at the Ser–Xaa bond. The numbers represent the relative percentages of cleavage. The same cleavage pattern was observed for D22A/H110A and D22A/H110A/R116A cathepsin B mutants.

the hydrophobic amino acids Trp, Tyr, Phe, Leu and Ala at two peptide bonds, as shown in Figure 5. The hydrolysis at the carboxy group of Ser puts Arg at the S_2 enzyme subsite and the hydrophobic amino acids at S_1' . These results indicated that on removal of the occluding loop contacts, this segment becomes more flexible, giving more room for substrate interaction with

the extended binding site of cathepsin B. In this condition, the S_1' – P_1' interaction takes a determinant role; the preference for Trp, Tyr, Phe, Leu and Ala to occupy the S_1' subsite direct the cleavage to the carboxy group of Ser. In accordance with this interpretation, Trp, Tyr, Phe, Leu have been described to be present at P_1' of the best substrates for cathepsin B in the series Dns-Phe-Arg-Xaa-Trp-Ala [22]. The apparent k_{cat}/K_m values for these peptides with two cleavage sites are presented in Table 2 to give the magnitude of hydrolytic activity of the mutants. All of the substrates assayed were hydrolysed by all mutants and a progressive increase in k_{cat}/K_m values was observed from the single-mutant to the triple-mutant enzymes. This occurred owing to a progressive increase in k_{cat} values, at least in the peptides containing Ser, Asn, Asp, His and Pro at the P_2' site, which were hydrolysed only at the Arg–Ser bond.

It is interesting to note that the H111A cathepsin B mutant hydrolysed all substrates assayed at the Arg–Ser bond but with k_{cat}/K_m values one to two orders of magnitude higher than those of the wild-type enzyme. Therefore, by removing the imidazole ring of His¹¹¹, modifications are introduced in the hydrolytic centre of the enzyme that are distinct from those of the H110A mutation. It is possible that in the H111A cathepsin B mutant there is more space for substrate fitting without displacement of the occluding loop, justifying the increase in k_{cat}/K_m values but maintaining the hydrolysis at the Arg–Ser bond.

We thank Dr Robert Ménard and Dr Dorit K. Nägler (Biotechnology Research Institute, NRC, Montreal, QC, Canada) for providing the cathepsin B mutants.

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), PADCT: Biotecnologia III and the INCO-DC programme (EU Contract Number ERBIC18CT970225).

REFERENCES

- Berti, P. J. and Storer, A. C. (1995) *J. Mol. Biol.* **246**, 273–283
- Kirschke, H., Barrett, A. J. and Rawlings, N. D. (1995) *Protein Profile* **2**, 1587–1644
- Authier, F., Posner, B. I. and Bergeron, J. J. (1996) *FEBS Lett.* **389**, 55–60
- Kakegawa, H., Nikawa, T., Tagami, K., Kamioka, H., Sumitani, K., Kawata, T., Drobnic-Kosorok, M., Lenarcic, B., Turk, V. and Katunuma, N. (1993) *FEBS Lett.* **321**, 247–250
- Esser, R. E., Angelo, R. A., Murphey, M. D., Watt, L. M., Thornburg, L. P., Palmer, J. T., Talhouk, J. W. and Smith, R. E. (1994) *Arthritis Rheum.* **37**, 236–247
- Duffy, M. J. (1992) *Biotherapy* **4**, 45–52
- Katunuma, N., Kominomi, E. and Li, K. (1987) *Am. J. Pathol.* **127**, 681–688
- Wurtz, A. and Bouchut, E. (1879) *Compt. Rend.* **89**, 425–430
- Brocklehurst, K., Willenbrock, F. and Salih, E. (1987) in *Hydrolytic Enzymes* (Neuberger, A. and Brocklehurst, K., eds.), pp. 39–158, Elsevier Biomedical Press, Amsterdam
- Storer, A. C. and Ménard, R. (1994) *Methods Enzymol.* **244**, 486–500
- Ménard, R. and Storer, A. C. (1998) in *Handbook of Proteolytic Enzymes* (Barrett, A. J., Rawling, N. D. and Woessner, J. F., eds.), pp. 555–557, Academic Press, New York
- Berger, A. and Schechter, I. (1970) *Phil. Trans. R. Soc. Lond. B* **257**, 249–264
- Kärgel, H. J., Dettmer, R., Etzold, G., Kirschke, H., Bohley, P. and Langner, J. (1980) *FEBS Lett.* **114**, 257–260
- Hasnain, S., Hiram, T., Huber, C. P., Mason, P. and Mort, J. S. (1993) *J. Biol. Chem.* **268**, 235–240
- Jia, Z., Hasnain, S., Hiram, T., Lee, X., Mort, J. S., To, R. and Huber, C. P. (1995) *J. Biol. Chem.* **270**, 5527–5533
- Shaw, E. (1990) *Adv. Enzymol. Relat. Areas Mol. Biol.* **63**, 271–347
- Krantz, A., Copp, L. J., Coles, J., Smith, R. A. and Heard, S. B. (1991) *Biochemistry* **30**, 4678–4687
- Pliura, D. H., Bonaventura, B. J., Smith, R. A., Coles, P. J. and Krantz, A. (1992) *Biochem. J.* **288**, 759–762
- Taralp, A., Kaplan, H., Sytwu, I.-I., Vlattas, I., Bohacek, R., Knap, A. K., Hiram, T., Huber, C. P. and Hasnain, S. (1995) *J. Biol. Chem.* **270**, 18036–18043
- García-Echeverría, C. and Rich, D. H. (1992) *FEBS Lett.* **297**, 100–102
- García-Echeverría, C. and Rich, D. H. (1992) *Biochem. Biophys. Res. Commun.* **187**, 615–619
- Ménard, R., Carmona, E., Plouffe, C., Brömme, D., Konishi, Y., Lefebvre, J. and Storer, A. C. (1993) *FEBS Lett.* **328**, 107–110
- Turk, D., Guncar, G., Podobnik, M. and Turk, B. (1998) *Biol. Chem.* **379**, 137–147
- Nägler, D. K., Storer, A. C., Portaro, F. C. V., Carmona, E., Juliano, L. and Ménard, R. (1997) *Biochemistry* **36**, 12608–12615
- Musil, D., Zucic, D., Turk, D., Engh, R. A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., Katunuma, N. and Bode, W. (1991) *EMBO J.* **10**, 2321–2330
- Illy, C., Quraishi, O., Wang, J., Purísima, E., Vernet, T. and Mort, J. (1997) *J. Biol. Chem.* **272**, 1197–1202
- Turk, D., Podobnik, M., Popovic, T., Katsunuma, N., Bode, W., Huber, R. and Turk, V. (1995) *Biochemistry* **34**, 4791–4797
- Yamamoto, A., Hara, T., Tomoo, K., Ishida, T., Fujii, T., Hata, Y., Murata, M. and Kitamura, K. (1997) *J. Biochem. (Tokyo)* **121**, 974–977
- Nägler, D. K., Tam, W., Storer, A. C., Krupa, J. C., Mort, J. S. and Ménard, R. (1999) *Biochemistry* **38**, 4868–4874
- Rowan, A. D., Mason, P., Mach, L. and Mort, J. S. (1992) *J. Biol. Chem.* **267**, 15993–15999
- Carmona, E., Dufour, É., Plouffe, C., Takebe, S., Manson, P., Mort, S. M. and Ménard, R. (1996) *Biochemistry* **35**, 8149–8157
- Ménard, R., Khouri, H. E., Plouffe, C., Dupras, R., Rippoli, D., Vernet, T., Tessier, D. C., Laliberté, F., Thomas, D. Y. and Storer, A. C. (1990) *Biochemistry* **29**, 6706–6713
- Barrett, A. J. and Kirschke, H. (1981) *Methods Enzymol.* **80**, 771–778
- Hirata, Y., Cezari, M. H. S., Nakaie, C. R., Boschcov, P., Ito, A. S., Juliano, M. A. and Juliano, L. (1994) *Lett. Peptide Sci.* **1**, 299–308
- Chagas, J. R., Juliano, L. and Prado, E. S. (1990) *Anal. Biochem.* **192**, 419–425
- Wilkinson, G. N. (1961) *Biochem. J.* **80**, 324–332
- Heinrikson, R. L. and Meredith, S. C. (1984) *Anal. Biochem.* **436**, 65–74
- Drenth, J., Jansonius, J. N., Koekoek, R., Swen, H. M. and Wolthers, B. G. (1968) *Nature (London)* **218**, 929–932
- Lecaille, F., Serveau, C., Gauthier, F. and Lalmanach, G. (1999) *FEBS Lett.* **445**, 311–314
- Hilaire, P. M., Willert, M., Juliano, M. A., Juliano, L. and Morten, M. J. (1999) *J. Combin. Chem.* **1**, 509–523
- Turk, D., Podobnik, M., Kuhelj, R., Dolinar, M. and Turk, V. (1996) *FEBS Lett.* **384**, 211–214
- Cygler, M., Sivaraman, J., Grochulski, P., Coulombe, R., Storer, A. C. and Mort, J. S. (1996) *Structure* **4**, 405–416
- Drenth, J., Kalk, K. H. and Swen, H. M. (1976) *Biochemistry* **15**, 3731–3738
- Varughese, K. I., Ahmed, F. R., Carey, P. R., Hasnain, S., Huber, C. P. and Storer, A. C. (1989) *Biochemistry* **28**, 1330–1332
- Kim, M., Yamamoto, D., Matsumoto, K., Inoue, M., Ishida, T., Mizuno, H., Sumiya, S. and Kitamura, K. (1992) *Biochem. J.* **287**, 797–803
- Schröder, E., Phillips, C., Garman, E., Marlos, K. and Grawford, C. (1993) *FEBS Lett.* **315**, 38–42
- Fujishima, A., Imai, Y., Toshiyuki, N., Fujisawa, Y., Yamamoto, Y. and Sugawara, T. (1997) *FEBS Lett.* **407**, 47–50
- O'Hara, B. P., Hemmings, A. M., Buttle, D. J. and Pearl, L. H. (1995) *Biochemistry* **34**, 13190–13195
- Gauthier, F., Moreau, T., Lalmanach, G., Brillard-Bourdet, M., Ferrer, D., Martino, M. and Juliano, L. (1993) *Arch. Biochem. Biophys.* **306**, 304–308
- Coulombe, R., Grochulski, P., Sivaraman, J., Ménard, R., Mort, J. S. and Cygler, M. (1996) *EMBO J.* **15**, 5492–5503
- Servau, C., Lalmanach, G., Juliano, M. A., Scharfstein, J., Juliano, L. and Gauthier, F. (1996) *Biochem. J.* **313**, 951–956

Received 27 September 1999/3 December 1999; accepted 20 January 2000