S3 to S3′ subsite specificity of recombinant human cathepsin K and development of selective internally quenched fluorescent substrates

Marcio F. M. ALVES*, Luciano PUZER*, Simone S. COTRIN*, Maria Aparecida JULIANO*, Luiz JULIANO*, Dieter BRÖMME† and Adriana K. CARMONA*[1]

*Department of Biophysics, Escola Paulista de Medicina, UNIFESP, Rua Três de Maio 100, São Paulo 04044-020, Brazil, and †Department of Human Genetics, Mount Sinai School of Medicine, Box 1498, New York, NY 10029, U.S.A.

We have systematically examined the S3 to S3′ subsite substrate specificity requirements of cathepsin K using internally quenched fluorescent peptides derived from the lead sequence Abz-KLRFSKQ-EDDnp [where Abz is *o*-aminobenzoic acid and EDDnp is *N*-(2,4-dinitrophenyl)ethylenediamine]. We assayed six series of peptides, in which each position except Gln was substituted with various natural amino acids. The results indicated that the S3–S1 subsite requirements are more restricted than those of S1′–S3′. Cathepsin K preferentially accommodates hydrophobic amino acids with aliphatic side chains (Leu, Ile and Val) in the S2 site. Modifications at P1 residues also have a large influence on cathepsin K activity. Positively charged residues (Arg and Lys) represent the best accepted amino acids in this position, although a particular preference for Gly was found as well. Subsite S3 accepted preferentially basic amino acids such as Lys and Arg. A broad range of amino acids was accommodated in the remaining subsites. We further explored the acceptance of a Pro residue in the P2 position by cathepsin K in order to develop specific substrates for the enzyme. Two series of peptides with the general sequences Abz-KXPGSKQ-EDDnp and Abz-KPXGSKQ-EDDnp (where X denotes the position of the amino acid that is altered) were synthesized. The substrates Abz-KPRGSKQ-EDDnp and Abz-KKPGSKQ-EDDnp were cleaved by cathepsin K at the Arg–Gly and Gly–Ser bonds respectively, and have been shown to be specific for cathepsin K when compared with other lysosomal cysteine proteases such as cathepsins L and B and with the aspartyl protease cathepsin D.

Key words: lysosomal proteinase, cysteine protease, fluorogenic substrate.

INTRODUCTION

Mammalian lysosomal cysteine proteases are generally regarded as enzymes that randomly degrade proteins in lysosomes. However, this concept has been reviewed, and it was demonstrated that some of the human cathepsins are also involved in selective and controlled processes and have specific functions associated with their restricted tissue localization [1–3]. This is the case for human cathepsin K, which is highly expressed in osteoclasts and multinucleated giant cells and has been implicated in bone resorption and the degradation of foreign body material [4–6]. However, the expression of cathepsin K is not limited to osteoclasts and osteoclast-like cells. High expression levels have also been identified in synovial fibroblasts, which are involved in the pathological erosion of articular cartilage [5,7], and in various epithelial cells [8,9]. The biological function of cathepsin K in these cells is still unknown.

Remodelling of the human skeleton is a cyclical process that involves phases of bone resorption and replacement. This process is balanced through the activities of bone-generating osteoblasts and bone-resorbing osteoclasts. The specific role of cathepsin K in bone resorption was demonstrated by the observation that a deficiency of cathepsin K causes an inherited autosomal recessive bone dysplasia, pycnodysostosis [10]. Among the mammalian collagenases, cathepsin K is the only cysteine protease capable of cleaving type I collagen at multiple sites within the triple helical structure. The degradation of type I collagen is the major event during bone remodelling [11] and depends on the formation of collagenolytically active complexes between cathepsin K and chondroitin sulphate [12].

The critical role of cathepsin K in bone resorption has led to the belief that specific inhibitors of the enzyme will be useful in the treatment of osteoporosis. Recently, the substrate specificity profile of the S4 to S1 subsites of the enzyme (according to the Schechter and Berger nomenclature [13]) was explored using a positional scanning fluorogenic substrate library with the general sequence Z-peptidyl-ACC (where Z is benzyloxycarbonyl and ACC is 7-amino-4-carbamoylmethylcoumarin) [14]. The cleavage of these peptides can be monitored by measuring the hydrolytic release of the fluorogenic C-terminal leaving group, but this restricts the substrate specificity analysis to the mapping of S′ subsites. For this reason, the use of internally quenched fluorogenic substrates, in which fluorescence appears after the cleavage of any bond in the peptide chain, is welcomed for systematic studies of substrate specificities addressing S and S′ subsites as well as their interdependencies.

In the present work, we explored the S3 to S3′ subsite specificity requirements of cathepsin K using internally quenched fluorescent peptides derived from Abz-KLRFSKQ-EDDnp [where Abz is *o*-aminobenzoic acid and EDDnp is *N*-(2,4-dinitrophenyl)-ethylenediamine]. This general peptide sequence was designed based on results with a solid-phase substrate library for other cysteine proteases [15,16]. Six series of peptides were generated: Abz-XLRFSKQ-EDDnp, Abz-KXRFSKQ-EDDnp, Abz-KLXFSKQ-EDDnp, Abz-KLRXSKQ-EDDnp, Abz-KLRFXKQ-EDDnp and Abz-KLRFSXQ-EDDnp (where X

Abbreviations used: Abz, *o*-aminobenzoic acid; ACC, 7-amino-4-carbamoylmethylcoumarin; EDDnp, *N*-(2,4-dinitrophenyl)ethylenediamine; MCA, 7-amino-4-methylcoumarin; MALDI-TOF, matrix-assisted laser-desorption ionization–time-of-flight; Z-, benzyloxycarbonyl-.

[1] To whom correspondence should be addressed (e-mail adriana.biof@infar.epm.br).

denotes the position of the amino acid that is altered). Recently, Lecaille et al. [14] demonstrated that the binding of a Pro residue in the S2 pocket of cathepsin K was required for its unique collagenase activity. In addition, the identification of the sites of cleavage by cathepsin K in collagens revealed the frequent presence of Pro residues in the P2 position [11,14,17,18]. In order to explore this unusual specificity, we synthesized two series of peptides with the general sequences Abz-KXPGSKQ-EDDnp and Abz-KPXGSKQ-EDDnp and studied their cleavage by cathepsin K. For comparison, these peptides were also assayed with cathepsins B and L.

MATERIALS AND METHODS

Enzymes

Recombinant human cathepsin K and cathepsin L were expressed in *Pichia pastoris* as described previously [19]. Cathepsin B was obtained as described in [20]. The molar concentrations of these enzymes were determined by active-site titration with E-64 [*trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane] according to [21]. Human liver cathepsin D was purchased from Calbiochem. Cathepsin K solutions were prepared in the presence of 0.01 % Tween 20 to increase the stability of the enzyme during the assays [22].

Peptide synthesis

The internally quenched fluorescent peptides containing the group EDDnp attached to a Gln residue were synthesized by the solid-phase synthesis method, as described elsewhere [23], using the Fmoc (fluoren-9-ylmethoxycarbonyl) procedure. An automated bench-top simultaneous multiple solid-phase peptide synthesizer (PSSM 8 system; Shimadzu) was used for the synthesis. All peptides obtained were purified by semi-preparative HPLC using an Econosil C-18 column. The molecular mass and purity of synthesized peptides were checked by amino acid analysis and by molecular mass determination with MALDI-TOF (matrix-assisted laser-desorption ionization–time-of-flight) MS, using a TOFSpec E instrument (Micromass, Manchester, U.K.). The stock solutions of EDDnp peptides were prepared in DMSO, and the concentrations were measured spectrophotometrically using a molar absorption coefficient of 17 300 $M^{-1} \cdot cm^{-1}$ at 365 nm.

Enzyme assays

The internally quenched fluorescent peptides were assayed in a Hitachi F-2000 spectrofluorimeter at 37 °C. Assays with the cysteine proteases cathepsins K, L and B and with the aspartyl protease cathepsin D were performed in 50 mM sodium acetate buffer, pH 5.5, containing 1 mM EDTA. Cathepsins K, L and B were pre-activated with 5 mM dithioerythritol for 5 min at 37 °C before the addition of the substrates. Fluorescence was measured continuously at $\lambda_{ex} = 320$ nm and $\lambda_{em} = 420$ nm (0.35–1.0 ml final volume). The enzyme concentration for initial-rate determinations was chosen at a level intended to hydrolyse less than 5 % of the amount of added substrate. The slope was converted into units of μmol of substrate hydrolysed per min, based on a calibration curve obtained from the complete hydrolysis of each peptide. The inner-filter effect was corrected for using an empirical equation as described previously [24]. The kinetic parameters K_m and k_{cat} were calculated by non-linear regression data analysis using the *Grafit* program [25], and the k_{cat}/K_m values were the ratios of these two parameters. The apparent second-order rate constant k_{cat}/K_m was determined under pseudo-first-order conditions, where $[S] \ll K_m$, and performed at two different substrate concentrations. For peptides hydrolysed at more than one site, the apparent k_{cat}/K_m values correspond to the sum of the individual values of k_{cat}/K_m for each cleavage site [26]. The errors were < 5 % for any kinetic parameter value obtained.

Determination of the peptide bonds cleaved

Fragments resulting from hydrolysis of the internally quenched fluorescent peptides by cathepsins K, L and B were isolated by HPLC in an Ultrasphere C-18 column (5 μm; 4.6 mm × 150 mm), and all of the cleavage sites were confirmed by MALDI-TOF MS (Tof-Spec-E; Micromass). This procedure was used in the present study for all peptides that were hydrolysed by any of the proteases.

RESULTS AND DISCUSSION

Noteworthy differences with regard to the specificity of cathepsin K were found when comparing our results obtained using internally quenched fluorogenic peptides with previously published data obtained using positional scanning fluorogenic substrate libraries [14]. We observed that the subsite specificities, mainly at S2, S1 and S1′, are not independent when the 'primed' substrate binding subsites are occupied by amino acids, and that significant changes in specificity can be observed. Synergic or hindrance effects on binding at different cathepsin K subsites that influence the specificity of the enzyme cannot be detected using positional scanning fluorogenic substrate libraries that have a fixed non-amino-acid residue in P1′. Thus the use of substrates spanning both substrate binding areas, N- and C-terminally from the scissile bond, permit a more informative analysis of the substrate specificities of proteases. In the following, our results using these substrates for the analysis of cathepsin K specificity are presented and discussed.

Abz-KLXFSKQ-EDDnp series for characterization of S1 specificity

Table 1 shows the kinetic parameters for the hydrolysis of the peptide series Abz-KLXFSKQ-EDDnp, with different amino acids in the X position. All of the peptides of this series were hydrolysed at the Xaa–Phe bond, with Xaa occupying the P1 position. The peptide Abz-KLRFSKQ-EDDnp was the most susceptible substrate for cathepsin K, with high k_{cat} and low K_m values. Analogues of Abz-KLRFSKQ-EDDnp with systematic substitution of Arg by each of 18 natural amino acids (Cys was excluded) resulted in substrates that were hydrolysed with lower catalytic efficiency (k_{cat}/K_m). Peptides of this series containing other basic amino acids (Lys and His) in the P1 position were also hydrolysed well by cathepsin K. Notably, the presence of Gly in the P1 position resulted in a substrate with the lowest K_m value among all peptides studied. The high affinity of Abz-KLGFSKQ-EDDnp was confirmed by using this peptide as a competitive inhibitor for the hydrolysis of Z-FR-MCA (where MCA is 7-amino-4-methylcoumarin), revealing a K_i value of 35 nM. This observation is in accordance with reported cathepsin K cleavage sites in the N-telopeptide-to-helix cross-linking domains of human bone type I collagen (e.g. cleavage of a Gly–Leu peptide bond [18]), suggesting that the presence of Gly in the P1 position may also play an important role in the hydrolysis of natural substrates. Peptides with Ala or aromatic residues (Trp, Tyr and Phe) in the P1 position also were susceptible to hydrolysis by

Table 1 Kinetic parameters for hydrolysis of the peptide series Abz-KLXFSKQ-EDDnp by recombinant human cathepsin K for the characterization of S1 subsite specificity

The assays were performed at 37 °C, in 50 mM sodium acetate buffer containing 1 mM EDTA, pH 5.5. The enzyme was preactivated with 5 mM dithioerythritol. * indicates that the k_{cat}/K_m value was determined under pseudo-first-order conditions. The cleavage site was at the Xaa–Phe bond, as determined by HPLC/MS.

X	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
R	0.14	4.2	30 000
G	0.04	0.8	20 000
K	0.27	4.1	15 185
H	0.34	4.9	14 412
A	0.14	1.5	10 714
W	0.32	3.2	10 000
Y	0.31	1.9	6129
F	0.82	4.0	4878
L	0.25	1.1	4400
E	2.33	3.1	1330
I	–	–	486*
Q	0.26	0.1	385
S	0.56	0.2	357
T	0.75	0.1	133
V	–	–	110*
N	1.04	0.1	96
D	3.27	0.3	92
M	4.40	0.1	23
P	–	–	18*

cathepsin K. The acidic amino acid Glu, but not Asp, fitted into the S1 subsite, permitting efficient hydrolysis. In contrast, peptides containing polar non-charged residues (Ser, Thr, Asn and Gln) in the P1 position were hydrolysed with low k_{cat} values. Peptides containing Met and Pro were poorly hydrolysed by cathepsin K, revealing the lowest k_{cat}/K_m values in the series.

These results agree with those obtained previously with positional scanning fluorogenic substrate libraries with regard to the preference of cathepsin K for basic amino acids in the P1 position [14]. However, the scanning libraries did not detect the hydrolysis of substrates containing His, Glu and particularly Gly in this position. It is likely that these previously undetected specificities are attributed to the fixation of Leu in the P2 position and/or the presence of additional interactions of the substrates with S′ subsites.

Abz-KXRFSKQ-EDDnp series for S2 specificity

Kinetic parameters for the hydrolysis of the peptide series Abz-KXRFSKQ-EDDnp with different amino acids in the X position are given in Table 2. Several substrates of this series were hydrolysed at more than one peptide bond, and the relative amount of each cleavage was determined using the areas from integrated HPLC chromatogram analysis. The apparent k_{cat}/K_m values for these substrates that were hydrolysed at more than one peptide bond are presented in Table 2, and correspond to the sums of the individual k_{cat}/K_m values for each bond cleaved.

Cathepsin K preferentially accommodates in the S2 subsite hydrophobic amino acids with aliphatic chains. The preference for these amino acids in the S2 pocket was very strict, as substrates containing Leu, Ile and Val in the P2 position were the only peptides that were hydrolysed exclusively at the Xaa–Phe bond. All other peptides of this series were hydrolysed by cathepsin K at two or three peptide bonds (e.g. Xaa–Arg, Arg–Phe or Phe–Ser). Although the S2 subsite of cysteine proteases is described as the substrate specificity determining binding site (reviewed in [27])

Table 2 Kinetic parameters for hydrolysis of the peptide series Abz-KXRFSKQ-EDDnp by recombinant human cathepsin K

Conditions for hydrolysis were as described in Table 1. N.H. no hydrolysis occurred at enzyme concentrations < 40 nM. The apparent k_{cat}/K_m values for substrates hydrolysed at more than one peptide bond ($^a k_{cat}/K_m$) represent the sum of the individual k_{cat}/K_m values for each cleavage. The cleavage sites were determined and quantified by HPLC/MS.

					Cleavage site (% of hydrolysis)		
X	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	$^a k_{cat}/K_m$ ($mM^{-1} \cdot s^{-1}$)	X↓R	R↓F	F↓S
L	0.14	4.2	30 000		0	100	0
F	–	–	–	23 810	14.5	77.4	8.1
I	0.42	9.0	21 428		0	100	0
R	–	–	–	21 176	76.0	0	24.0
V	0.18	2.6	14 444		0	100	0
Q	–	–	–	8019	49.4	27.5	23.1
P	–	–	–	7620	22.8	44.5	32.7
M	–	–	–	6666	7.0	93.0	0
N	–	–	–	6434	20.0	70.0	10.0
G	–	–	–	5968	65.1	0	34.9
A	–	–	–	5000	32.4	36.6	31.0
H	–	–	–	400	80.0	0	20.0
D	N.H.						

and thus as a widely autonomous binding pocket, the acceptance of Pro in the S2 subsite of cathepsin K seems to depend on the nature of the amino acids located at prime and non-prime positions of the substrate (see below). The ability of cathepsin K to cleave substrates with Pro in P2 [14,17,18,28] is unusual among the cysteine proteases, although it has been reported for cruzipain and cathepsin B [15,29]. The fact that peptide Abz-KPRFSKQ-EDDnp can be hydrolysed by cathepsin K at the Pro–Arg and Arg–Phe bonds demonstrates that Pro can also be accepted by the S1 subsite, depending on the substrate sequence. Peptides with a Met, Phe or Asn residue in the P2 position were hydrolysed preferentially at the Arg–Phe bond, indicating that these amino acids were also well accommodated in the S2 subsite of cathepsin K. In this series of substrates, the amino acids Arg, Gln, Gly and His preferentially occupied the S1 subsite, as these substrates were hydrolysed mainly at the Xaa–Arg bond, thus placing Lys into the S2 subsite. Similarly, the peptide Abz-KRRFSKQ-EDDnp was hydrolysed efficiently by cathepsin K at the Arg–Arg or Phe–Ser bond, again placing the basic residue Lys or Arg at S2 subsite. In addition, the peptide with Asp at P2 was resistant to hydrolysis. These are surprising results, because there is no negative charge localized in the S2 pocket of cathepsin K that would explain the accommodation of Arg and Lys or the repulsion of a negative charge (e.g. Asp side chain). Finally, the high k_{cat} value for the cleavage of the Gly–Arg bond in the peptide Abz-KGRFSKQ-EDDnp further confirmed the result shown above that Gly is well accepted in the P1 position.

Abz-XLRFSKQ-EDDnp series for S3 specificity

Table 3 shows the kinetic parameters for the hydrolysis of the Abz-XLRFSKQ-EDDnp series. All of the peptides of this series were hydrolysed at the Arg–Phe bond; thus Xaa occupied the P3 position in all of them. The S3 subsite showed a high preference for basic amino acid residues, such as Lys and Arg, as shown by the k_{cat}/K_m values. The substrates with Leu and Ala were also well hydrolysed by cathepsin K due to their low K_m values. The other substrates from this series were hydrolysed significantly, as demonstrated by the rather high k_{cat} values; however, the lower k_{cat}/K_m values were due to the higher K_m values. The best amino

Table 3 Kinetic parameters for hydrolysis of the peptide series Abz-XLRFSKQ-EDDnp by recombinant human cathepsin K for the characterization of S3 subsite specificity

Conditions for hydrolysis were as described in Table 1. *indicates that the k_{cat}/K_m value was determined under pseudo-first-order conditions. All cleavages were at the Arg–Phe bond, as determined by HPLC/MS.

X	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
K	0.14	4.2	30 000
R	0.10	2.4	24 000
L	0.10	1.3	13 000
A	0.21	2.7	12 857
H	0.34	2.6	7647
G	0.60	2.3	3833
N	0.62	1.8	2903
F	0.73	2.1	2877
S	–	–	1037*
D	2.00	1.1	550

Table 4 Kinetic parameters for hydrolysis of the peptide series Abz-KLRXSKQ-EDDnp by recombinant human cathepsin K for the characterization of S1′ subsite specificity

Conditions for hydrolysis were as described in Table 1. N.H. no hydrolysis occurred at enzyme concentrations < 40 nM. All cleavages (with one exception) were at the Arg–Xaa bond, as determined by HPLC/MS. The exception was the peptide containing Gly, for which there were two cleavage sites: R↓G (75 %) and G↓S (25 %); †indicates that the apparent k_{cat}/K_m value for this substrate represents the sum of the individual k_{cat}/K_m values for each cleavage.

X	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
N	0.46	15.1	32 826
F	0.14	4.2	30 000
L	0.32	5.4	16 875
G	–	–	13 600†
S	0.26	2.6	10 000
A	–	–	9227
R	0.62	5.7	9193
D	1.12	1.9	1696
P	N.H.		

acid at P3 and the relative k_{cat}/K_m values of the substituted peptides examined in the present study are in accordance with the results found using the positional scanning fluorogenic Z-peptidyl-ACC substrate libraries [14].

Abz-KLRXSKQ-EDDnp series for S1′ specificity

The kinetic parameters for the hydrolysis by cathepsin K of peptides from the series Abz-KLRXSKQ-EDDnp are given in Table 4. Except for the substrate Abz-KLRGSKQ-EDDnp, which was hydrolysed at both the Arg–Gly bond (75 %) and the Gly–Ser bond (25 %), all the other peptides from this series were cleaved only at the Arg–Xaa bond. The significant susceptibility of substrates with Gly in the P1′ position is a feature that allows cathepsin K to degrade collagen helical domains, as reported previously [18]. The substrate Abz-KLRNSKQ-EDDnp presented the highest k_{cat} value in this series. However, the peptides with Phe and Leu were also very susceptible to hydrolysis. The substrate containing the positively charged amino acid Arg was less susceptible, whereas the peptide with a negatively charged Asp residue presented the highest K_m value, resulting in the lowest k_{cat}/K_m value of the series. The substrate containing Pro at P1′ was the only peptide that was resistant to hydrolysis. The specificity of the S1′ subsite of cathepsin K seems to be a hybrid of those of cathepsins B and L, since this site accepted hydrophobic (Phe and Leu) as well as hydrophilic (Asn and Ser) amino acids [30–33].

Table 5 Kinetic parameters for hydrolysis of the peptide series Abz-KLRFXKQ-EDDnp by recombinant human cathepsin K for the characterization of S2′ subsite specificity

Conditions for hydrolysis were as described in Table 1. *indicates that k_{cat}/K_m values were determined under pseudo-first-order conditions. All cleavages were at the Arg–Phe bond, as determined by HPLC/MS.

X	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
R	0.15	5.1	34 000
S	0.14	4.2	30 000
F	0.10	2.7	27 000
N	0.12	2.7	22 500
G	–	–	21 800*
A	–	–	19 100*
P	–	–	15 400*
D	0.45	3.6	8000

Table 6 Kinetic parameters for hydrolysis of the peptide series Abz-KLRFSXQ-EDDnp by recombinant human cathepsin K for the characterization of S3′ subsite specificity

Conditions for hydrolysis were as described in Table 1. Cleavage was at the Arg–Phe bond, as determined by HPLC/MS.

X	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
R	0.14	4.5	32 142
F	0.10	3.1	31 000
S	0.07	2.1	30 158
K	0.14	4.2	30 000
A	0.10	3.0	30 000
P	0.15	4.3	28 666
N	0.15	1.9	12 666
G	0.18	1.8	10 000
D	0.25	1.7	6814

Abz-KLRFXKQ-EDDnp and Abz-KLRFSXQ-EDDnp series for characterization of S2′ and S3′ specificities

Tables 5 and 6 present the kinetic parameters for the hydrolysis of the peptides from the series Abz-KLRFXKQ-EDDnp and Abz-KLRFSXQ-EDDnp respectively. All peptides of these two series were hydrolysed at the Arg–Phe bond. Cathepsin K did not show a clear preference for the different amino acids present at P2′ and P3′, revealing a broad specificity for these positions. However, the substrates containing a negatively charged Asp residue at the P2′ or P3′ position were hydrolysed with lower k_{cat}/K_m values.

Cathepsin K activity towards Pro-containing peptides with the general sequences Abz-KPXGSKQ-EDDnp and Abz-KXPGKQ-EDDnp, and towards fluorescent bradykinin derivatives

Lecaille et al. [14] demonstrated that the preference of cathepsin K for a Pro residue in the P2 substrate position distinguishes this enzyme from cathepsin L. Our results indicated that the enzyme could also hydrolyse peptides containing Pro in the P1 position (Table 2). These observations stimulated us to design specific substrates for cathepsin K. We generated two series of peptides, Abz-KPXGKQ-EDDnp and Abz-KXPGKQ-EDDnp, and tested them with cathepsin K. The kinetic parameters for hydrolysis are presented in Tables 7 and 8 respectively. Cathepsin K hydrolysed

Table 7 Kinetic parameters for hydrolysis of peptides with the general sequence Abz-KPXGSKQ-EDDnp by recombinant human cathepsins K, L and B

Experimental conditions were as described in Table 1. All the peptides were hydrolysed by cathepsin K at the Xaa–Gly bond, as determined by HPLC/MS.

	Cathepsin K			Cathepsin L	Cathepsin B
X	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
R	0.59	2.22	3763	60	7.5
Nle	0.86	1.68	1958	419	5
Q	1.61	1.93	1199	129	6
T	1.67	1.35	808	41	6
G	0.76	0.40	524	10	3

Table 8 Kinetic parameters for hydrolysis of peptides with the general sequence Abz-KXPGSKQ-EDDnp by recombinant human cathepsins K, L and B

Conditions for hydrolysis were as described in Table 1. All peptides were hydrolysed by cathepsin K at the Gly–Ser bond, as determined by HPLC/MS.

	Cathepsin K			Cathepsin L	Cathepsin B
X	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
K	0.85	1.29	1518	26	9
P	1.60	0.56	350	12	4
G	1.28	0.37	285	7.5	5
L	2.37	0.37	156	32	3
Q	3.55	0.21	59	12	4
I	2.15	0.11	51	15	3

all the peptides from Abz-KPXGKQ-EDDnp series at the Xaa–Gly bond, and all those from the Abz-KXPGSKQ-EDDnp series at the Gly–Ser bond. Therefore Pro was fixed in the P2 position in both series. The peptide Abz-KPRGSKQ-EDDnp was hydrolysed with the highest k_{cat}/K_m value, but almost 10 times less efficiently than the lead sequence Abz-KLRFSKQ-EDDnp. Pro-containing peptides with substitutions in the P3 position generated the series of Abz-KXPGSKQ-EDDnp substrates. The kinetic parameters presented in Table 8 demonstrate that the peptides from this series were hydrolysed with lower k_{cat}/K_m values than substrates containing Lys in P3, confirming the preference for a Lys residue in this position.

The capability of cathepsin K to hydrolyse Pro-containing peptides was explored further by testing the fluorescent bradykinin derivatives Abz-RPPGFSPFR-OH and Abz-RPPGFSPFR-EDDnp as substrates. HPLC analysis of the fragments and MS indicated that both peptides were cleaved at the Gly–Phe bond, placing Pro in P2 and Gly in P1. The kinetic parameters determined for the hydrolysis of Abz-RPPGFSPFR-EDDnp by cathepsin K were $K_m = 1.7\ \mu M$, $k_{cat} = 0.1\ s^{-1}$ and $k_{cat}/K_m = 62.8\ mM^{-1} \cdot s^{-1}$. These results indicated that the fluorogenic analogues of bradykinin are effective substrates for the enzyme, but are much less susceptible to hydrolysis than the other Pro-containing peptides tested (Tables 7 and 8). Cathepsin K has been described to present kininogenase activity, but the enzyme also hydrolysed bradykinin at the Gly–Phe bond [34], as we observed with the internally quenched fluorescent peptide derived from bradykinin.

Relative activities of cathepsins K, L and B towards Pro-containing peptides derived from the series Abz-KPXGSKQ-EDDnp and Abz-KXPGKQ-EDDnp

At pH 5.5, Abz-KPRGSKQ-EDDnp and Abz-KKPGSKQ-EDDnp are highly selective substrates for cathepsin K, displaying almost complete resistance towards cathepsin L- and B-mediated hydrolysis (Tables 7 and 8). In addition, no hydrolysis was observed by the aspartyl protease cathepsin D. In contrast, the best selective cathepsin K substrate presently available, Z-GPR-MCA, is also hydrolysed significantly by cathepsin B; the k_{cat}/K_m values for the hydrolysis of this substrate are 118 $mM^{-1} \cdot s^{-1}$ for cathepsin K and 10 $mM^{-1} \cdot s^{-1}$ for cathepsin B [35]. Therefore, due to their high selectivity, the peptides Abz-KPRGSKQ-EDDnp and Abz-KKPGSKQ-EDDnp may be particularly useful for studies of cathepsin K activity in tissues or cell extracts in which other proteases are present.

Positional scanning fluorogenic Z-peptidyl-ACC libraries and panels of internally quenched fluorescent substrates are complementary

All of the best accepted amino acids in the P3 to P1 positions identified using the fluorogenic Z-peptidyl-ACC series [14] were confirmed in the present study using internally quenched fluorescent substrates. However, these latter substrates also allowed us to explore the non-prime subsites of cathepsin K and the synergistic effects of prime-site interactions, demonstrating that the subsites are not independent. A noteworthy characteristic of cathepsin K that was not detected with positional scanning fluorogenic libraries was the acceptance of a Gly residue at the P1 position in the internally quenched fluorescent peptides. Since it lacks a side chain, Gly does not allow any side-chain-related interaction with the S1 subsite. The efficient hydrolysis of the Gly–Xaa bond thus requires interactions at prime sites of the substrate. In addition, the internally quenched fluorescent substrate containing Ala at P1 was also well hydrolysed by cathepsin K; however, similarly to Gly, Ala was not accepted in this position in the positional scanning fluorogenic library. This particular observation suggests that further investigation is warranted with regard to the substrate requirements for cleavage at Gly, because this is the site at which N-telopeptide-to-helix cross-linking domains of human bone type I collagen are cleaved by cathepsin K [18]. Other noteworthy internally quenched fluorescent substrates not detected with positional scanning fluorogenic libraries were those with His at P1 and Lys or Arg at P2. The reason for these differences is not clear, but, similar to the recently described plasticity of cathepsin S substrate binding sites [36], the occupancy of prime subsites by substrates in cathepsin K as described here may modulate the interactions of non-prime sites such as S1 and S2.

Conclusions

In conclusion, we have described in detail the specificity of cathepsin K and have obtained evidence that the subsites of the active centre in this important enzyme are interdependent. Aliphatic side chains in the P2 position of substrates direct fitting of the substrate to the catalytic groove of the enzyme. Proline preferentially occupies the S2 subsite if Gly is present at the P1 or P1′ position. Furthermore, the S2 subsite of cathepsin K accepted basic amino acids in some of the examined peptides and excluded Asp. In addition, we have developed efficient and highly selective substrates for cathepsin K.

This work was supported by Fundação de Amparo Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Human Frontiers for Science Progress (RG 00043/2000-M). This research was also supported in part by National Institutes of Health grants AR 46182 and AR 48669 (D. B.).

REFERENCES

1 Turk, B., Turk, D. and Turk, V. (2000) Lysosomal cysteine proteases: more than scavengers. Biochim. Biophys. Acta **1477**, 98–111
2 Buhling, F., Fengler, A., Brandt, W., Welte, T., Ansorge, S. and Nagler, D. K. (2000) Review: novel cysteine proteases of the papain family. Adv. Exp. Med. Biol. **477**, 241–254
3 Lecaille, F., Kaleta, J. and Brömme, D. (2002) Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitors design. Chem. Rev. **102**, 4459–4488
4 Tezuka, K., Tezuka, Y., Maejima, A., Sato, T., Nemoto, K., Kamioka, H., Hakeda, Y. and Kumugawa, M. (1994) Molecular cloning of a possible cysteine proteinase predominantly expressed in osteoclasts. J. Biol. Chem. **269**, 1106–1109
5 Brömme, D. and Okamoto, K. (1995) Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary: molecular cloning, sequencing and tissue distribution. Biol. Chem. Hoppe Seyler **376**, 379–384
6 Drake, F. H., Dodds, R. A., James, I. E., Connor, J. R., Debouck, C., Richardson, S., Lee, E., Rieman, D., Barthlow, R., Hastings, G. and Gowen, M. (1996) Cathepsin K but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts. J. Biol. Chem. **271**, 12511–12516
7 Bossard, M. J., Tomaszek, T. A., Thompson, S. K., Amegadzie, B. Y., Hanning, C. R., Jones, C., Kurdyla, J. T., McNulty, D. E., Drake, F. H., Gowen, M. and Levy, M. A. (1996) Proteolytic activity of human osteoclast cathepsin K. Expression, purification, activation, and substrate identification. J. Biol. Chem. **271**, 12517–12524
8 Buhling, F., Gerber, A., Hackel, C., Kruger, S., Kohnlein, T., Brömme, D., Reinhold, D., Ansorge, S. and Welte, T. (1999) Expression of cathepsin K in lung epithelial cells. Am. J. Respir. Cell Mol. Biol. **20**, 612–619
9 Haeckel, C., Krueger, S., Buehling, F., Brömme, D., Franke, K., Schuetze, A., Roese, I. and Roessner, A. (1999) Expression of cathepsin K in the human embryo and fetus. Dev. Dyn. **216**, 89–95
10 Gelb, B. D., Shi, G. P., Chapman, H. A. and Desnick, R. J. (1996) Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. Science **273**, 1236–1238
11 Garnero, P., Borel, O., Byrjalsen, I., Ferreras, M., Drake, F. H., McQueney, M. S., Foged, N. T., Delmas, P. D. and Delaisse, J. M. (1998) The collagenolytic activity of cathepsin K is unique among mammalian proteinases. J. Biol. Chem. **273**, 32347–32352
12 Li, Z., Hou, W. S., Escalante-Torres, C. R., Gelb, B. D. and Brömme, D. (2002) Collagenase activity of cathepsin K depends on complex formation with chondroitin sulfate. J. Biol. Chem. **277**, 28669–28676
13 Schechter, I. and Berger, A. (1968) On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. **27**, 157–162
14 Lecaille, F., Choe, Y., Brandt, W., Li, Z., Craik, C. S. and Brömme, D. (2002) Selective inhibition of the collagenolytic activity of human cathepsin K by altering its S2 subsite specificity. Biochemistry **41**, 8447–8454
15 Del Nery, E., Juliano, M. A., Meldal, M., Svedensen, I., Scharfstein, J., Walmsley, A. and Juliano, L. (1997) Characterization of the substrate specificity of the major cysteine protease (cruzipain) from *Trypanosoma cruzi* using a portion-mixing combinatorial library and fluorogenic peptides. Biochem. J. **323**, 427–433
16 St Hilaire, P. M., Alves, L. C., Sanderson, S. J., Mottram, J. C., Juliano, M. A., Juliano, L., Coombs, G. H. and Meldal, M. (2000) The substrate specificity of a recombinant cysteine protease from *Leishmania mexicana*: application of a combinatorial peptide library approach. Chembiochem **1**, 115–122
17 Kafienah, W., Brömme, D., Buttle, D. J., Croucher, L. J. and Hollander, A. P. (1998) Human cathepsin K cleaves native type I and II collagens at the N-terminal end of the triple helix. Biochem. J. **331**, 727–732
18 Atley, L. M., Mort, J. S., Lalumiere, M. and Eyre, D. R. (2000) Proteolysis of human bone collagen by cathepsin K: characterization of the cleavage sites generating by cross-linked N-telopeptide neoepitope. Bone **26**, 241–247
19 Linnevers, C. J., McGrath, M. E., Armstrong, R., Mistry, F. R., Barnes, M. G., Klaus, J. L., Palmer, J. T., Katz, B. A. and Brömme, D. (1997) Expression of human cathepsin K in *Pichia pastoris* and preliminary crystallographic studies of an inhibitor complex. Protein Sci. **6**, 919–921
20 Nagler, D. K., Storer, A. C., Portaro, F. C., Carmona, E., Juliano, L. and Menard, R. (1997) Major increase in endopeptidase activity of human cathepsin B upon removal of occluding loop contacts. Biochemistry **36**, 12608–12615
21 Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M. and Hanada, K. (1982) L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. Biochem. J. **201**, 189–198
22 Krupa, J. C. and Mort, J. S. (2000) Optimization of detergents for the assay of cathepsins B, L, S and K. Anal. Biochem. **283**, 99–103
23 Hirata, I. Y., Cezari, M. H. C., Nakaie, C. R., Boshcov, P., Ito, A. S., Juliano, M. A. and Juliano, L. (1994) Internally quenched fluorogenic protease substrates: Solid-phase synthesis and fluorescent spectroscopy of peptides containing *ortho*-aminobenzoil-dinitrophenyl groups as donor-acceptor pairs. Lett. Pept. Sci. **1**, 299–308
24 Araujo, M. C., Melo, R. L., Cesari, M. H., Juliano, M. A., Juliano, L. and Carmona, A. K. (2000) Peptidase specificity characterization of C- and N-terminal catalytic sites of angiotensin I-converting enzyme. Biochemistry **39**, 8519–8525
25 Leatherbarrow, R. J. (1992) *Grafit* Version 3.0, Erithacus Software Ltd, Staines, U.K.
26 Fersht, A. (1985) Enzyme Structure and Mechanism, W. H. Freeman and Co., New York
27 Turk, D., Guncar, G., Podobnik, M. and Turk, B. (1998) Revised definition of substrate binding sites of papain-like cysteine proteases. Biol. Chem. **379**, 137–147
28 Aibe, K., Yazawa, H., Abe, K., Teramura, K., Kumegawa, M., Kawashima, H. and Honda, K. (1996) Substrate specificity of recombinant osteoclast-specific cathepsin K from rabbits. Biol. Pharm. Bull. **19**, 1026–1031
29 Del Nery, E., Alves, L. C., Melo R. L., Cezari, M. H. S., Juliano, L. and Juliano, M. A. (2000) Specificity of cathepsin B to fluorescent substrates containing benzyl side-chain-substituted amino acids at P1 subsite. J. Protein Chem. **19**, 33–38
30 Ménard, R., Carmona, E., Pouffle, C., Brömme, D., Konishi, Y., Lefebvre, J. and Storer, A. C. (1993) The specificity of the S1′ subsite of cysteine proteases. FEBS Lett. **328**, 107–110
31 Portaro, F. C. V., Santos, A. B. F., Cezari, M. H. S., Juliano, M. A., Juliano, L. and Carmona, E. (2000) Probing the specificity of cysteine proteinases at subsites remote from the active site: analysis of P4, P3, P2′ and P3′ variations in extended substrates. Biochem. J. **347**, 123–129
32 Melo, R. L., Pozzo, R. C. B., Alves, L. C., Perissutti, E., Caliendo, G., Santagada, V., Juliano, L. and Juliano, M. A. (2001) Synthesis and hydrolysis by cathepsin B of fluorogenic substrates with the general structure benzoyl-X-Arg-MCA containing non-natural basic amino acids at position X. Biochim. Biophys. Acta **1547**, 82–94
33 Cezari, M. H. S, Puzer, L., Juliano, M. A., Carmona, A. K. and Juliano, L. (2002) Cathepsin B carboxydipeptidase specificity analysis using internally quenched fluorescent peptides. Biochem. J. **368**, 365–369
34 Desmazes, C., Galineau, L., Gauthier, F., Brömme, D. and Lalmanach, G. (2003) Kininogen-derived peptides for investigating the putative vasoactive properties of human cathepsins K and L. Eur. J. Biochem. **270**, 171–178
35 Xia, L., Kilb, J., Wex, H., Lipyansky, A., Breuil, V., Stein, L., Palmer, J. T., Dempster, D. W. and Brömme, D. (1999) Localization of rat cathepsin K in osteoclasts and resorption pits: Inhibition of bone resorption cathepsin K-activity by peptidyl vinyl sulfones. Biol. Chem. **380**, 679–687
36 Pauly, T. A., Sulea, T., Ammirati, M., Sivaraman, J., Danley, D. E., Griffor, M. C., Kamath, A. V., Wang, I. K., Laird, E. R., Seddon, A. P. et al. (2003) Specificity determinants of human cathepsin S revealed by crystal structures of complexes. Biochemistry **42**, 3203–3213

Received 20 March 2003/2 May 2003; accepted 6 May 2003
Published as BJ Immediate Publication 6 May 2003, DOI 10.1042/BJ20030438