A model to explain the pH-dependent specificity of cathepsin B-catalysed hydrolyses

Henry E. KHOURI,* Céline PLOUFFE,* Sadiq HASNAIN,† Tomoko HIRAMA,† Andrew C. STORER* and Robert MÉNARD*‡

*Protein Engineering Section, Biotechnology Research Institute, National Research Council Canada, 6100 Royalmount Avenue, Montréal, Québec H4P 2R2, Canada and †Institute for Biological Sciences, National Research Council Canada, Ottawa, Ontario K1A 0R6, Canada

1. Three synthetic substrates of cathepsin B (EC 3.4.22.1) with various amino acid residues at the P₂ position (Cbz-Phe-Arg-NH-Mec, Cbz-Arg-Arg-NH-Mec and Cbz-Cit-Arg-NH-Mec, where Cbz represents benzyloxycarbonyl and NH-Mec represents 4-methylcoumarin-7-ylamide) were used to investigate the pH-dependency of cathepsin B-catalysed hydrolyses and to obtain information on the nature of enzyme–substrate interactions. 2. Non-linear-regression analysis of pH–activity profiles for these substrates indicates that at least four ionizable groups on cathepsin B with pK_a values of 3.3, 4.55, 5.46 and > 7.3 can affect the rate of substrate hydrolysis. 3. Ionization of the residue with a pK_a of 5.46 has a strong effect on activity towards the substrate with an arginine in P₂ (8.4-fold increase in activity) but has only a moderate effect on the rate of hydrolysis with Cbz-Cit-Arg-NH-Mec (2.3-fold increase in activity) and virtually no effect with Cbz-Phe-Arg-NH-Mec. The kinetic data are consistent with this group being an acid residue with a side chain able to interact with the side chains of an arginine or a citrulline in the P₂ position of a substrate. Amino acid sequence alignment and model building with the related enzyme papain (EC 3.4.22.2) suggest that Glu-245 of cathepsin B is a likely candidate. The relative importance of electrostatic and hydrophobic interactions in the S₂ subsite of cathepsin B is discussed. 4. For all three substrates, activity appears after ionization of a group with a pK_a of 3.3, believed to be the active-site Cys-29 of cathepsin B. The identity of the groups with pK_a values of 4.55 and > 7.3 remains unknown.

INTRODUCTION

Cathepsin B (EC 3.4.22.1) is a lysosomal cysteine proteinase that has been isolated from many mammalian tissues, including human liver (Barrett, 1977). Much of the information on the mechanism of cathepsin B has been obtained through comparisons with the well characterized cysteine proteinase papain (EC 3.4.22.2). Alignment and comparison of the sequences of related cysteine proteinases suggest similar three-dimensional structures for cathepsin B and papain (Kamphuis *et al.*, 1985; Dufour, 1988). Even though the basic features of the mechanisms are similar, major differences exist between the two enzymes. Cathepsin B is both an endopeptidase and a dipeptidyl carboxy dipeptidase, whereas papain works exclusively as an endopeptidase (Aronson & Barrett, 1978; Bond & Barrett, 1980; McKay *et al.*, 1983; Mason, 1989). The specificities of the two enzymes are also somewhat different. Papain shows a marked preference for peptides with a bulky non-polar side chain, especially L-phenylalanine, on the N-terminal side of the amino acid containing the peptide bond being hydrolysed, i.e. at the P₂ position (Berger & Schechter, 1970). This specificity is attributed to interaction of the phenylalanine side chain with hydrophobic residues at the S₂ subsite of papain, mainly Val-133 (173) and Val-157 (197) (Drenth *et al.*, 1976). (Cathepsin B numbering is used throughout the text except for residues of papain, where the position number for papain is followed in parentheses by the corresponding number for cathepsin B.) Cathepsin B activity towards substrates with a P₂ phenylalanine residue is still very

high, but the enzyme also has a strong affinity for substrates with an arginine residue at the P₂ position, such as Cbz-Arg-Arg-NH-Nap and Cbz-Arg-Arg-NH-Mec (Barrett & Kirschke, 1981). Ionization of a group with a pK_a of about 5.4, which is not observed in papain, seems to be essential for activity towards substrates such as Cbz-Arg-Arg-NH-Nap (Knight, 1980). On the basis of this result and the observed increase of activity towards reactivity probes by a deprotonation with a pK_a of 3.4, it has been proposed that formation of the catalytic thiolate–imidazolium ion-pair, which is believed to be the active form of papain and cathepsin B, is a necessary but not sufficient condition for catalysis by cathepsin B (Willenbrock & Brocklehurst, 1984, 1985). Cathepsin B, however, is active at low pH towards other substrates (Bajkowski & Frankfater, 1983; Katunuma *et al.*, 1983; Hirao *et al.*, 1984) and ionization of the group with a pK_a of about 5.4 is observed only when an arginine residue is present at the P₂ position of the substrate under study.

We have used a systematic approach to characterize the S₂ binding-site specificity of cathepsin B, based on three synthetic substrates of the enzyme. These substrates are of the type Cbz-Xaa-Arg-NH-Mec, where Xaa represents L-phenylalanine, L-arginine or L-citrulline. The use of the citrulline derivative, which is isosteric with arginine but carries a neutral ureido group instead of a positively charged guanidino group, allows evaluation of electrostatic contributions to enzyme–substrate interactions in the S₂ subsite of cathepsin B. We report here the detailed kinetic characterization of the influence of pH on the activity of cathepsin B towards these three substrates.

Abbreviations used: DTNB, 5,5'-dithiobis-(2-nitrobenzoic acid); Cbz-Phe-Arg-NH-Mec, benzyloxycarbonyl-L-phenylalanyl-L-arginine 4-methylcoumarin-7-ylamide; Cbz-Arg-Arg-NH-Mec, benzyloxycarbonyl-L-arginyl-L-arginine 4-methylcoumarin-7-ylamide; Cbz-Arg-Arg-NH-Nap, benzyloxycarbonyl-L-arginyl-L-arginine 2-naphthylamide; Cbz-Cit-Arg-NH-Mec, benzyloxycarbonyl-L-citrullyl-L-arginine 4-methylcoumarin-7-ylamide; E-64, 2-[(1-*trans*-3-carboxy-2,3-epoxypropionyl)-L-leucyl]amino-4-guanidinobutane; NH₂-Mec, 7-amino-4-methylcoumarin.

‡ To whom correspondence should be addressed.

MATERIALS AND METHODS

Materials

Cathepsin B was isolated and purified from human liver by following the procedure of Rich *et al.* (1986). The enzyme was activated by incubation with 100 mM-phosphate buffer, pH 6.0, containing 0.4 M-NaCl, 10 mM-EDTA and 20 mM-dithiothreitol for 1 h. Papain from Sigma Chemical Co. (St. Louis, MO, U.S.A.) was purified by using a mercurial-agarose column (Sluyterman & Wijdenes, 1970). Addition of 2-mercaptoethanol followed by gel filtration (Sephadex G-15) was used to activate the enzyme after purification. Both cathepsin B and papain were stored on ice following activation. DTNB was purchased from Sigma Chemical Co. NH₂-Mec was obtained from Aldrich Chemical Co. (Milwaukee, WI, U.S.A.). The substrates Cbz-Phe-Arg-NH-Mec and Cbz-Arg-Arg-NH-Mec, as well as the inhibitor E-64 were purchased from IAF Biochemical International (Laval, Quebec, Canada). Cbz-Cit-Arg-NH-Mec was generously given by Dr. Barbara J. Gour-Salin (Biotechnology Research Institute). H.p.l.c.-grade acetonitrile was used for kinetic experiments.

Kinetic measurements

The concentration of cathepsin B was determined by active-site titration with E-64 as described previously (Ménard *et al.*, 1990). For papain, titration with DTNB according to the method described by Ellman (1959) was used to obtain the concentration of active enzyme. Kinetic measurements were performed by

monitoring the fluorescence of NH₂-Mec released during hydrolysis of the substrate, as described previously (Ménard *et al.*, 1990). Kinetic parameters for substrate hydrolysis by cathepsin B were determined by measuring initial rates at various substrate concentrations. For the measurement of the pH-dependence of activity, $(k_{\text{cat.}}/K_m)_{\text{obs.}}$ (representing an experimentally observed value of the specificity constant) was obtained at substrate concentrations such that $[S] \ll K_m$. Under these conditions, the specificity constant can be obtained simply by dividing the initial rate by the enzyme and substrate concentrations. The standard assay conditions consisted of 50 mM-citrate (pH 3.0–5.7), phosphate (pH 5.7–7.9) or borate (pH 7.9–10.0) buffer consisting of 1–5 mM-EDTA and 0.2 M-NaCl. The reaction mixture also contained 5% (v/v) acetonitrile, unless specified otherwise, for substrate solubility. The pH-activity profiles for cathepsin B were determined in the acid range only, because of enzyme inactivation at pH values above about 7. All kinetic experiments were done at 25.0 °C.

Kinetic models for analysis of pH-activity profiles

Various models have been considered to interpret the kinetic data in the acid limb of pH-activity profiles for cathepsin B and papain.

Non-linear regression of experimental data to the appropriate equation was done using the program ENZFITTER written by R. J. Leatherbarrow (Elsevier-BIOSOFT). For each model, the calculated values were then compared with the experimental data.

Model 1:

$$(k_{\text{cat.}}/K_m)_{\text{obs.}} = \frac{k_{\text{cat.}}/K_{m,1}}{\frac{[\text{H}^+]^2}{K_1 K_2} + \frac{[\text{H}^+]}{K_2} + 1} \quad (1)$$

Model 2:

$$(k_{\text{cat.}}/K_m)_{\text{obs.}} = \frac{k_{\text{cat.}}/K_{m,1}}{\frac{[\text{H}^+]}{K_1} + 1 + \frac{K_2}{[\text{H}^+]}} + \frac{k_{\text{cat.}}/K_{m,2}}{\frac{[\text{H}^+]^2}{K_1 K_2} + \frac{[\text{H}^+]}{K_2} + 1} \quad (2)$$

Model 3:

$$(k_{\text{cat.}}/K_m)_{\text{obs.}} = \frac{k_{\text{cat.}}/K_{m,1}}{\frac{[\text{H}^+]}{K_1} + 1 + \frac{K_2}{[\text{H}^+]}} + \frac{k_{\text{cat.}}/K_{m,2}}{\frac{[\text{H}^+]^2}{K_1 K_2} + \frac{[\text{H}^+]}{K_2} + 1 + \frac{K_3}{[\text{H}^+]}} + \frac{k_{\text{cat.}}/K_{m,3}}{\frac{[\text{H}^+]^3}{K_1 K_2 K_3} + \frac{[\text{H}^+]^2}{K_2 K_3} + \frac{[\text{H}^+]}{K_3} + 1} \quad (3)$$

RESULTS

Determination of experimental pK_a values from pH-activity profiles

The pH-dependence of $(k_{cat.}/K_m)_{obs.}$ can be used to determine the pK_a values of groups on the free enzyme that, upon ionization, affect the activity towards specific substrates. It was found that the experimental conditions, in particular the organic solvent content of the medium, required for substrate solubility, have a very large influence on the overall shape of pH-activity profiles for cathepsin B. This is shown in Fig. 1, where the pH-activity profiles for hydrolysis of Cbz-Phe-Arg-NH-Mec by cathepsin B in the presence of 2.5% and 20.0% acetonitrile are shown. A certain number of 'transitions' are observed, reflecting the dependence of activity on ionization of various residues in cathepsin B. At greater than 15% acetonitrile these features become poorly defined owing to solvent effects on the pK_a values and reactivities of the various enzyme species. For that reason, experiments carried out at 5% acetonitrile were used for quantitative analysis, since better estimation of the kinetic parameters can be obtained under those conditions. In 5% acetonitrile the pH-activity profile for Cbz-Phe-Arg-NH-Mec is best described by Model 3, where the activity is modulated by three ionizable groups leading to three species of increasing reactivity. This is

Fig. 1. pH-dependence of $(k_{cat.}/K_m)_{obs.}$ for the reactions of Cbz-Phe-Arg-NH-Mec with cathepsin B in the presence of (a) 2.5% and (b) 20.0% acetonitrile

Experimental conditions were as described in the Materials and methods section.

Fig. 2. Comparison of pH-activity data for hydrolysis of Cbz-Phe-Arg-NH-Mec with calculated curves from various model systems

The experiment was done in the presence of 5% acetonitrile. The curves represent the best fit of the experimental data to Model 2 (---) and Model 3 (—) obtained by non-linear regression of the data to the appropriate equation. A theoretical curve is also illustrated for Model 1 (····) using pK_a values of 3 and 4 and a limiting $k_{cat.}/K_m$ value of $220000 \text{ M}^{-1} \cdot \text{s}^{-1}$ (see the text).

apparent from Fig. 2, where the best-fit curves for Models 2 and 3 are compared with the experimental results. Since no convergence could be obtained when carrying out non-linear regression of the pH-activity data to eqn. (1), a calculated curve with pK_a values of 3 and 4 and a limiting $k_{cat.}/K_m$ of $220000 \text{ M}^{-1} \cdot \text{s}^{-1}$ is also included in Fig. 2 to illustrate the pH behaviour predicted by Model 1. The apparent pK_a values of the three ionizable groups involved determined by non-linear regression of the pH-activity data to eqn. (3) are 3.2, 4.4 and > 7 (Table 1).

Table 1. Experimental pK_a values from pH-activity profiles of cathepsin B and papain with substrates of the type Cbz-Xaa-Arg-MH-Mec

Xaa	Minimum model*	Experimental pK_a †
Cathepsin B		
Phe	3	3.2 ± 0.3 ; 4.4 ± 0.1 ; > 7
Arg	2	5.4 ± 0.1 ; > 7
Cit	3	3.4 ± 0.5 ; 5.1 ± 0.1 ; > 7
Papain		
Phe	1	3.2 ± 0.6 ; 4.43 ± 0.05
Arg	1	3.7 ± 0.2 ; 4.30 ± 0.05
Cit	1	3.4 ± 0.1 ; 4.21 ± 0.02

* The minimum model is the simplest model that can be used to get a good agreement between experimental and calculated values.

† Obtained by non-linear regression to the equation corresponding to the minimum model. Only the acid limb of pH-activity profiles for papain has been considered.

Fig. 3. pH-dependence of $(k_{\text{cat}}/K_{\text{m}})_{\text{obs}}$ for the reactions of cathepsin B at 5% acetonitrile with (a) Cbz-Phe-Arg-NH-Mec, (b) Cbz-Arg-Arg-NH-Mec and (c) Cbz-Cit-Arg-NH-Mec

The points represent experimental data and the continuous lines correspond to the best fit of the data (up to pH 6.5) to eqn. (3). Kinetic parameters were obtained by simultaneous non-linear regression of the pH-activity data at pH ≤ 6.5 for all three substrates as described in the text and are given in Table 2.

The same type of analysis was performed with the substrates Cbz-Arg-Arg-NH-Mec and Cbz-Cit-Arg-NH-Mec and the results are given in Table 1. The pH-activity data are presented in Fig. 3. The pH-dependence of activity with Cbz-Arg-Arg-NH-Mec is somewhat less complex than what is observed with the substrate containing a phenylalanine residue in P_2 , and the minimum model describing the pH-activity data is Model 2, which considers two ionizable groups and two reactive forms of the enzyme. The data can also be described by Model 3, with $k_{\text{cat}}/K_{\text{m},1} \ll k_{\text{cat}}/K_{\text{m},2}$. A major increase in activity is observed upon ionization of a residue with an apparent pK_a of 5.4. This group does not seem to be involved in hydrolysis of Cbz-Phe-Arg-NH-Mec since the corresponding pH-activity profile does not show a similar transition with a pK_a of about 5.4. The citrulline residue in Cbz-Cit-Arg-NH-Mec is isosteric with arginine but lacks the positive charge on the side chain. Non-linear regression of the experimental data to eqns. (1)–(3) indicates that the pH behaviour of this substrate is best described by the same model as that for Cbz-Phe-Arg-NH-Mec, i.e. Model 3. The activity is modulated by the ionization of three groups in

Fig. 4. pH-dependence of $(k_{\text{cat}}/K_{\text{m}})_{\text{obs}}$ for the reactions of papain at 5% acetonitrile with (a) Cbz-Phe-Arg-NH-Mec, (b) Cbz-Arg-Arg-NH-Mec and (c) Cbz-Cit-Arg-NH-Mec

The points represent experimental data and the continuous lines correspond to the best fit of the data to an equation describing the full pH range of the profile (Ménard *et al.*, 1990). Kinetic parameters were obtained by non-linear regression. The $k_{\text{cat}}/K_{\text{m}}$ values are 932000, 1270 and 13800 $\text{M}^{-1}\cdot\text{s}^{-1}$ for Cbz-Phe-Arg-NH-Mec, Cbz-Arg-Arg-NH-Mec and Cbz-Cit-Arg-NH-Mec respectively. The pK_a values for the low-pH limb are given in Table 1. The pK_a values for the basic limb of the profiles are 8.32 ± 0.03 , 8.24 ± 0.02 and 8.27 ± 0.01 for the phenylalanine, arginine and citrulline derivatives respectively.

the enzyme with pK_a values of 3.4, 5.1 and > 7 . It must be noted that, as was the case with Cbz-Phe-Arg-NH-Mec, the hydrolysis of Cbz-Cit-Arg-NH-Mec cannot be described by Model 2 and three reactive species must be considered to reproduce the pH-activity profile adequately. For all three substrates an increase in $(k_{\text{cat}}/K_{\text{m}})_{\text{obs}}$ at high pH indicates that a group of pK_a higher than 7 participates in cathepsin B-catalysed hydrolysis of these synthetic substrates. Owing to rapid inactivation of cathepsin B at high pH, however, the variation in activity upon ionization of this group is poorly defined and its pK_a cannot be obtained with precision. Its presence cannot be doubted, however, and ionization of this group leads to a protonic form of higher activity.

By comparison, the pH-dependence of $(k_{\text{cat}}/K_{\text{m}})_{\text{obs}}$ for the substrates Cbz-Phe-Arg-NH-Mec, Cbz-Arg-Arg-NH-Mec and

Cbz-Cit-Arg-NH-Mec with papain are much simpler (Fig. 4). For this enzyme the acid limb of the profiles can be explained by Model 1, in which two pK_a values and one active form of the enzyme are present. The experimental values of pK_1 and pK_2 in Model 1 were obtained by non-linear regression of the data to eqn. (1) and the results are given in Table 1. It is important to note that, within the experimental error, the kinetically determined pK_a values are the same for all three substrates studied. Fig. 4 also shows that for papain, the activity decreases at $pH > 7.5$ due to deprotonation of the catalytic His-159(199) residue (Ménard *et al.*, 1990). A similar effect is probable in the case of cathepsin B but cannot be measured owing to rapid enzyme inactivation at high pH. Analysis of the full pH-activity profiles for papain as described previously (Ménard *et al.*, 1990) allows determination of the pK_a for His-159(199) in papain. Values of 8.32, 8.24 and 8.27 were obtained with Cbz-Phe-Arg-NH-Mec, Cbz-Arg-Arg-NH-Mec and Cbz-Cit-Arg-NH-Mec respectively, indicating once again that the pH-activity profiles are independent of the substrate with papain.

Elaboration of a model to explain the pH-dependence of cathepsin B-catalysed reactions

The pK_a values determined from the influence of pH on the observed k_{cat}/K_m reflect the ionization of residues on the free enzyme and should be independent of the substrate under study (Fersht, 1985). This was observed in the case of papain (Fig. 4). However, as shown in Table 1 and Fig. 3, the pH-activity profiles for cathepsin B-catalysed hydrolyses vary greatly from substrate to substrate. This apparent substrate-dependence of the pH-activity profiles for hydrolysis of substrates with relatively similar structures can be explained by considering Model 4.

This model takes into account five protonic states of cathepsin B, four of which are active towards the substrates under consideration in this study. The four pK_a values are unique to the enzyme, but the k_{cat}/K_m values for each reactive species ($k_{cat}/K_{m,1}$ – $k_{cat}/K_{m,4}$) are dependent on the nature of the substrate. From the experimental pK_a values listed in Table 1, it can be argued that five ionizable groups, rather than four, should be used to explain the pH-dependence of $(k_{cat}/K_m)_{obs}$. However, we believe that the pK_a of 5.1 observed for Cbz-Cit-Arg-NH-Mec is a composite of the pK_a values of two ionizable groups (i.e. those with pK_a values of 4.4 and 5.4) (see below). It has been shown previously that such a composite pK_a can be indistinguishable from the two individual ionization steps if the pK_a values of these steps are relatively close (Brocklehurst *et al.*, 1983). Most of the parameters in Model 4 were obtained by non-linear regression of the data at $pH \leq 6.5$ to a simplified form of the model, where the effect of the last ionization is negligible and therefore pK_4 and $k_{cat}/K_{m,4}$ were not considered. This simplification of the system allows better evaluation of the individual parameters, since at the pH values where enzyme form EH in Model 4 contributes significantly to $(k_{cat}/K_m)_{obs}$ (i.e. $pH > 6.5$) the data become unreliable owing to enzyme inactivation. Ignoring the data at $pH > 6.5$ does not influence the

determination of other parameters, however, because, as is seen below, pK_3 and pK_4 differ by approx. 2 pH units. If enzyme form EH is not considered, then the influence of pH on the observed activity is represented by eqn. (3).

Rather than fitting the pH-activity profiles for each substrate individually, the parameters describing Model 4 were determined by curve-fitting the data for cathepsin B pH-activity profiles of the three substrates simultaneously to eqn. (3). This makes maximum use of the data and allows us to get more precise estimates of the parameters and to verify if pH-activity profiles for all three substrates can be accounted for by a single model. This was performed by using an improved version of the non-linear-regression program of Duggleby (1989). The data contained two independent variables: the pH and a second 'indicator' variable, which was equal to 1 for Cbz-Phe-Arg-NH-Mec, 2 for Cbz-Arg-Arg-NH-Mec and 3 for Cbz-Cit-Arg-NH-Mec. The data were also normalized to give equal importance to all substrates in the regression procedure. However, there is still a large number of parameters to adjust: three pK_a values (pK_1 , pK_2 and pK_3) and nine values of k_{cat}/K_m (three values for each substrate). Therefore, in a first round of curve-fitting, the pK_a values were held constant at the apparent pK_a values reported in Table 1, and estimates of the nine k_{cat}/K_m parameters were obtained by non-linear regression. In subsequent rounds, the three pK_a values and the k_{cat}/K_m for one reactive form of the enzyme were considered as parameters to adjust while the k_{cat}/K_m values of the remaining two species were held constant at values obtained in preceding curve-fitting rounds. This was done for each reactive form of the enzyme consecutively, for a total of four rounds of curve-fitting. In each case, the results of the

Table 2. Specificity constants for the reactive forms of cathepsin B with substrates of the type Cbz-Xaa-Arg-NH-Mec determined using Model 4

Xaa	Enzyme form*	$10^{-3} k_{cat}/K_m$ (mean $\pm$ S.D.)† ($M^{-1} \cdot s^{-1}$)	Relative k_{cat}/K_m ‡
Phe	EH ₄	73 $\pm$ 4	0.36
	EH ₃	203 $\pm$ 5	1.00
	EH ₂	220 $\pm$ 3	1.08
Arg	EH ₄	1.3 $\pm$ 1.1	0.19
	EH ₃	6.7 $\pm$ 1.2	1.00
	EH ₂	56.5 $\pm$ 0.6	8.44
Cit	EH ₄	9.1 $\pm$ 2.1	0.18
	EH ₃	51.4 $\pm$ 2.8	1.00
	EH ₂	116 $\pm$ 2	2.26

* Enzyme forms correspond to those represented in Model 4.

† Specificity constants were obtained by simultaneous non-linear regression of pH-activity data for the three substrates to eqn. (3) as described in the text. k_{cat}/K_m of enzyme forms EH₄, EH₃ and EH₂ correspond to $k_{cat}/K_{m,1}$, $k_{cat}/K_{m,2}$ and $k_{cat}/K_{m,3}$ respectively in Model 4.

‡ Relative to enzyme form EH₃ (= 1.00).

Model 4:

previous round were used as initial estimates for the curve-fitting routine. The entire process was repeated a second time, but very little improvement was obtained between the eighth and ninth rounds of curve-fitting. The agreements between calculated and experimental data are represented in Fig. 3 and the corresponding parameters for each substrate are given in Table 2. The pK_a values of the groups modulating activity of cathepsin B are: $pK_1 = 3.3 \pm 0.2$; $pK_2 = 4.55 \pm 0.07$; $pK_3 = 5.46 \pm 0.02$. In addition, the value of pK_4 can be estimated to be higher than 7.3.

DISCUSSION

The influence of pH on the specificity constant (k_{cat}/K_m)_{obs} for hydrolysis of Cbz-Phe-Arg-NH-Mec, Cbz-Arg-Arg-NH-Mec and Cbz-Cit-Arg-NH-Mec by cathepsin B reveals the existence of four different active forms of the enzyme. This situation contrasts sharply with reactions of papain, where pH-activity profiles are much simpler since only two ionizable groups modulate activity of a single reactive form of the enzyme in the acid limb. Moreover, the complexity of pH-activity profiles for cathepsin B is increased because of variations in relative reactivities of the enzyme forms towards different substrates (Table 2). Specifically, the pH-dependence of Cbz-Arg-Arg-NH-Mec hydrolysis is dominated by two groups with pK_a values of 5.46 and > 7.3 , because enzyme forms EH_4 and EH_3 in Model 4 have very low reactivity towards this substrate. The reaction of cathepsin B with Cbz-Phe-Arg-NH-Mec is modulated by only three groups, since ionization of the group with a pK_a of 5.46 does not affect the activity of the enzyme, i.e. k_{cat}/K_m is approximately the same for enzyme forms EH_3 and EH_2 with Cbz-Phe-Arg-NH-Mec (Table 2). It was shown also that experimental conditions, in this case the acetonitrile concentration, can complicate the interpretation of pH-activity profiles because of differential solvent effects on the pK_a values and reactivities of the various enzyme species involved (see Fig. 1). It is therefore not surprising that widely varying pH-activity profiles have been reported in the literature for different substrates (Knight, 1980; Bajkowski & Frankfater, 1983; Katunuma *et al.*, 1983; Hirao *et al.*, 1984; Willenbrock & Brocklehurst, 1984).

The most striking differences between pH-activity profiles of cathepsin B and papain is the presence of a group with a pK_a of about 5.4 modulating the activity of the former enzyme towards substrates bearing an arginine residue in P_2 . From our study, we can add that ionization of this residue also has a moderate effect on the rate of hydrolysis of Cbz-Cit-Arg-NH-Mec and virtually no effect on Cbz-Phe-Arg-NH-Mec. This must reflect the particular amino acid composition of the S_2 subsite of cathepsin B. Comparison of the amino acid sequences of papain and cathepsin B indicates that Ser-205(245) located in the hydrophobic cavity of papain is replaced by a glutamic acid residue in human cathepsin B. By use of the crystal structure of papain and molecular graphics, it can be seen that if residue 205(245), which is located in the back of the cavity, is a glutamic acid residue it will be in a position to interact with the side chain of a bound substrate containing an arginine (or a citrulline) residue in the P_2 position. It is thus possible that the affinity of cathepsin B for substrates bearing an arginine in P_2 could be due to an interaction between this glutamic acid residue and the positively charged side chain of arginine. The pK_a of 5.46 determined for this residue in our study is a reasonable value for a glutamic acid residue in a hydrophobic environment. As shown in Table 2, ionization of this residue causes a 8.44-fold increase in reactivity towards Cbz-Arg-Arg-NH-Mec. Since the values of k_{cat}/K_m for enzyme forms EH_3 and EH_2 were obtained with very good

precision, a more quantitative analysis of their significance can be performed. By using the relationship

$$\Delta\Delta G = -RT \cdot \ln[k_{cat}/K_m (EH_2 \text{ form})/k_{cat}/K_m (EH_3 \text{ form})]$$

we can estimate the variation in enzyme transition-state binding energy upon ionization of the residue with a pK_a of 5.46 (see Scheme 1). A value of -5.29 kJ/mol (-1.26 kcal/mol) is obtained with the substrate Cbz-Arg-Arg-NH-Mec, -2.02 kJ/mol (-0.48 kcal/mol) with Cbz-Cit-Arg-NH-Mec and approximately no stabilization [$\Delta\Delta G = -0.19$ kJ/mol (-0.05 kcal/mol)] with Cbz-Phe-Arg-NH-Mec. These values are in order of what one would expect for the type of interactions suggested to be present in the S_2 subsite with these substrates (Scheme 1). With an arginine in P_2 , the interaction changes from a charged hydrogen bond with enzyme form EH_3 to a salt bridge with enzyme form EH_2 and the value of -5.29 kJ/mol (-1.26 kcal/mol) reflects the relative strengths of these two types of interaction. The neutral ureido side chain of the citrulline derivative cannot form a salt bridge and the experimentally measured $\Delta\Delta G$ of -2.02 kJ/mol (-0.48 kcal/mol) is an indication of the smaller stabilization effect obtained upon ionization of Glu-245 with this substrate. The very low value of $\Delta\Delta G$ obtained with Cbz-Phe-Arg-NH-Mec indicates that the phenylalanine side chain does not interact with Glu-245, which is reasonable since this side chain cannot reach to the back of the S_2 cavity of cathepsin B. These considerations support the view that the pK_a of about 5.4, which affects hydrolysis of substrates with an arginine in P_2 , is due to an interaction with an acid

Scheme 1. Postulated interactions between a glutamic acid residue in the S_2 subsite of cathepsin B and the side chain of residues in the P_2 position of substrates (a) Cbz-Phe-Arg-NH-Mec, (b) Cbz-Arg-Arg-NH-Mec and (c) Cbz-Cit-Arg-NH-Mec

EH_3 and EH_2 correspond to enzyme forms described in Model 4. $\Delta\Delta G$ was calculated as described in the text.

residue in the back of the S_2 subsite of cathepsin B, and comparison of amino acid sequences along with molecular modelling suggests that Glu-245 is the most likely candidate to participate in such an interaction.

The use of the synthetic substrate Cbz-Cit-Arg-NH-Mec is instrumental in understanding the system. It allows us to state that four ionizable groups modulate the activity of cathepsin B for the reactions considered in this study and it is essential for obtaining reasonably precise estimates of the parameters describing the system. Since the citrulline residue is isosteric with arginine but lacks the positive charge on the side chain, it can also be used to evaluate the contribution of electrostatic interactions in the enzyme-substrate systems. As seen in Table 2, the ratios of the specificity constants for the citrulline and arginine derivatives [$k_{\text{cat.}}/K_m$ (Cbz-Cit-Arg-NH-Mec)/ $k_{\text{cat.}}/K_m$ (Cbz-Arg-Arg-NH-Mec)] are 7.0 and 7.7 for enzyme forms EH_4 and EH_3 respectively, where Glu-245 is present as the neutral acid form. This ratio is relatively high if we consider that the difference between the two substrates is that arginine forms a charged hydrogen bond with Glu-245 whereas citrulline forms a non-charged hydrogen bond. This suggests that the positive charge on the arginine side chain interacts unfavourably with other residues in the S_2 subsite. When the side chain of Glu-245 is deprotonated and negatively charged (i.e. enzyme form EH_2) this ratio becomes 2.0, which shows that the electrostatic contributions to the binding of Cbz-Arg-Arg-NH-Mec are important and can partly overcome the unfavourable effect of the positively charged guanidinium group of arginine.

It has been suggested that a single group with a pK_a of about 5.5 controls the pH-dependence of both exo- and endo-peptidase activities of cathepsin B (Polgár & Csoma, 1987). However, subsite specificity was not considered in choosing the substrates used for monitoring exo- and endo-peptidase activities (glycine in P_2 for the exopeptidase substrate and arginine in P_2 for the endopeptidase substrate), and the influence of pH could reflect different phenomena for these two substrates. Polgár & Csoma (1987) also proposed that at $\text{pH} > 5.5$ the carboxylate side chains of Glu-171 and Glu-245 interact with the positively charged guanidinium group of Arg-202, keeping it away from the thiolate-imidazolium ion-pair of cathepsin B. Below $\text{pH} 5.5$, however, the Arg-202 side chain would move closer to the thiolate ion of Cys-29, causing a perturbation of the active-site ion-pair that would affect the relative exo- and endo-peptidase activities. If such a perturbation of the ion pair occurs, it should also influence the kinetic parameters with the substrate Cbz-Phe-Arg-NH-Mec. Since hydrolysis of this substrate is not affected by ionization of the group with a pK_a of 5.46, this hypothesis does not seem to be valid.

The nature of the other ionizations observed in our study is not as well understood. For all three substrates, activity appears after ionization of a group with a pK_a of 3.3. This pK_a probably reflects ionization of the active-site Cys-29 leading to the thiolate-imidazolium form of cathepsin B. This is in agreement with the value of 3.4 proposed by Willenbrock & Brocklehurst (1984) for ionization of Cys-29, based on the pH-dependence of the reaction of cathepsin B towards 2,2'-dipyridyl disulphide. The activity of cathepsin B is modulated by an additional group with a pK_a of 4.55 that cannot be attributed to any particular amino acid residue at the moment. The residue in cathepsin B with a pK_a higher than 7.3 is also unknown. Ionization of this

group leads to an enzyme form of higher activity towards all three substrates studied. This increase in activity at high pH has been observed previously for cathepsin B (Willenbrock & Brocklehurst, 1985), but is not present with the enzyme papain and constitutes another major difference between cathepsin B- and papain-catalysed hydrolyses. The presence of many ionizing groups affecting the activity of cathepsin B is not surprising if we consider that cathepsin B contains 40 residues with side chains that could ionize in the pH range 3–10 (13 aspartic acid, 18 glutamic acid and nine histidine residues) whereas papain has only 18 such residues (six aspartic acid, ten glutamic acid and two histidine residues). In addition, cathepsin B contains two cysteine residues not involved in disulphide bridges (one of them being the active-site Cys-29) whereas in papain only the active-site cysteine residue is free. The pH-dependence of cathepsin B is therefore much more complex and difficult to understand. Site-directed mutagenesis has been used to produce variants of cathepsin B and papain that are at present being characterized in our group. Results of these studies should lead to a better understanding of this complex system.

We thank Dr. Ronald G. Duggleby for providing us with a copy of his program DNRP53 for non-linear-regression analysis and Dr. Barbara J. Gour-Salin for the Cbz-Cit-Arg-NH-Mec substrate. This paper is N.R.C.C. publication no. 32424.

REFERENCES

- Aronson, N. N. & Barrett, A. J. (1978) *Biochem. J.* **171**, 759–765
- Bajkowski, A. S. & Frankfater, A. (1983) *J. Biol. Chem.* **258**, 1650–1655
- Barrett, A. J. (1977) in *Proteinases in Mammalian Cells and Tissues* (Barrett, A. J., ed.), pp. 181–207, North-Holland, Amsterdam
- Barrett, A. J. & Kirschke, H. (1981) *Methods Enzymol.* **80**, 535–561
- Berger, A. & Schechter, I. (1970) *Philos. Trans. R. Soc. London B* **257**, 249–264
- Bond, J. S. & Barrett, A. J. (1980) *Biochem. J.* **189**, 17–25
- Brocklehurst, K., Willenbrock, S. J. F. & Salih, E. (1983) *Biochem. J.* **211**, 701–708
- Drenth, J., Kalk, K. H. & Swen, H. M. (1976) *Biochemistry* **15**, 3731–3738
- Dufour, E. (1988) *Biochimie* **70**, 1335–1342
- Duggleby, R. G. (1989) *Biochem. J.* **257**, 419–424
- Ellman, G. L. (1959) *Arch. Biochem. Biophys.* **82**, 70–77
- Fersht, A. R. (1985) *Enzyme Structure and Mechanism*, pp. 155–175, W. H. Freeman, New York
- Hirao, T., Hara, K. & Takahashi, K. (1984) *J. Biochem. (Tokyo)* **95**, 871–879
- Kamphuis, I. G., Drenth, J. & Baker, E. N. (1985) *J. Mol. Biol.* **182**, 317–329
- Katunuma, N., Towatari, T., Tamai, M. & Hanada, K. (1983) *J. Biochem. (Tokyo)* **93**, 1129–1135
- Knight, C. G. (1980) *Biochem. J.* **189**, 447–453
- Mason, R. W. (1989) *Arch. Biochem. Biophys.* **273**, 367–374
- McKay, M., Offermann, M. K., Barrett, A. J. & Bond, J. S. (1983) *Biochem. J.* **213**, 467–471
- Ménard, R., Khouri, H. E., Plouffe, C., Dupras, R., Ripoll, D., Vernet, T., Tessier, D. C., Laliberté, F., Thomas, D. Y. & Storer, A. C. (1990) *Biochemistry* **29**, 6706–6713
- Polgár, L. & Csoma, C. (1987) *J. Biol. Chem.* **262**, 14448–14453
- Rich, D. H., Brown, M. A. & Barrett, A. J. (1986) *Biochem. J.* **235**, 731–734
- Sluyterman, L. A. A. & Wijdenes, J. (1970) *Biochim. Biophys. Acta* **200**, 593–594
- Willenbrock, F. & Brocklehurst, K. (1984) *Biochem. J.* **222**, 805–814
- Willenbrock, F. & Brocklehurst, K. (1985) *Biochem. J.* **227**, 521–528