Major Increase in Endopeptidase Activity of Human Cathepsin B upon Removal of Occluding Loop Contacts†

Dorit K. Nägler,‡ Andrew C. Storer,‡ Fernanda C. V. Portaro,§ Euridice Carmona,§ Luiz Juliano,|| and Robert Ménard*,‡

Biotechnology Research Institute, National Research Council of Canada, Montreal, Quebec, H4P2R2 Canada, Lab. Farmacologia, Instituto Butantan, São Paulo 05504, Brazil, and Department of Biophysics, Escola Paulista de Medicina, São Paulo 04044-020, Brazil

Received May 28, 1997; Revised Manuscript Received August 7, 1997⊗

ABSTRACT: The main feature distinguishing cathepsin B from other cysteine proteases of the papain family is the presence of a large insertion loop, termed the occluding loop, which occupies the S′ subsites of the enzyme. The loop is held in place mainly by two contacts with the rest of the enzyme, involving residues His110 and Arg116 on the loop that form salt bridges with Asp22 and Asp224, respectively. The influence of this loop on the endopeptidase activity of cathepsin B has been investigated using site-directed mutagenesis and internally quenched fluorogenic (IQF) substrates. Wild-type cathepsin B displays poor activity against the substrates Abz-AFRSAAQ-EDDnp and Abz-QVVAGA-EDDnp as compared to cathepsin L and papain. Appreciable increases in k_{cat}/K_M were observed for cathepsin B containing the single mutations D22A, H110A, R116A, and D224A. The highest activity however is observed for mutants where both loop to enzyme contacts are disrupted. For the triple-mutant D22A/H110A/R116A, an optimum k_{cat}/K_M value of 12×10^5 M^{-1} s^{-1} was obtained for hydrolysis of Abz-AFRSAAQ-EDDnp, which corresponds to a 600-fold increase relative to wild-type cathepsin B and approaches the level of activity observed with cathepsin L or papain. By comparison, the mutations have little effect on the hydrolysis of Cbz-FR-MCA. The influence of the mutations on the pH dependency of activity also indicates that the complexity of pH activity profiles normally observed for cathepsin B is related to the presence of the occluding loop. The major increase in endopeptidase activity is attributed to an increase in loop "flexibility" and suggests that the occluding loop might move when an endopeptidase substrate binds to the enzyme. The possible contribution of these interactions in regulating endopeptidase activity and the implications for cathepsin B activity in physiological or pathological conditions are discussed.

Cathepsin B is a lysosomal cysteine protease involved in intracellular protein turnover (Kirschke & Barrett, 1987). The enzyme has also been implicated in pathological conditions such as arthritis (Mort et al., 1984; Trabandt et al., 1991) and tumor metastasis [reviewed by Sloane (1990)]. Cathepsin B is the best characterized mammalian member of the papain family of cysteine proteases. A significant amount of functional (kinetic) data for the purified enzyme has been gathered since the late 1970s and the crystal structures of human and rat cathepsin B (Musil et al., 1991; Jia et al., 1995) as well as of complexes with inhibitors (Jia et al., 1995; Turk et al., 1995) are now available. The crystal structure of procathepsin B, the first structure of a cysteine protease zymogen, has been solved recently (Cygler et al., 1996; Turk et al., 1996). In addition, numerous functional properties of cathepsin B have been investigated by site-directed mutagenesis [e.g., Hasnain et al. (1992, 1993), Mach et al. (1993), Taralp et al. (1995), Fox et al. (1995), and Illy et al. (1997)]. The enzyme shows close structural and functional homology to other cysteine proteases of the papain family. One feature that distinguishes cathepsin B from other cysteine proteases is the position of the peptide bond being cleaved. While most cysteine proteases of the papain family are exclusively endopeptidases, cathepsin B has both exo- and endopeptidase activities (Barrett & Kirschke, 1981).

Despite the impressive amount of work done on cathepsin B, very few studies have dealt with the molecular determinants of the endo- and exopeptidase activities. The crystal structure of human cathepsin B (Musil et al., 1991) offered an explanation for the dipeptidyl carboxypeptidase activity of the enzyme. The presence of a large insertion loop found only in cathepsin B and positioned in the vicinity of the active site seems to favor dipeptidyl carboxypeptidase activity; two histidine residues (His110 and His111) in the loop provide positively charged anchors for the C-terminal carboxylate group of substrates. Accordingly, it was shown recently using site-directed mutagenesis that the presence of this loop is essential for the exopeptidase activity of the enzyme (Illy et al., 1997). The influence of the loop on the endopeptidase activity of cathepsin B however has not been investigated. It was determined many years ago that cysteine proteases of the papain family possess an extended binding site able to accommodate up to seven residues of a substrate [P_4 to P_3', Schechter & Berger (1967)]. For cathepsin B, the large insertion loop occupies the S′ region of the enzyme (S_2'−S_3') and should be detrimental to endopeptidase activity by

† NRCC Publication Number 39977. Supported in part by the Protein Engineering Network of Centres of Excellence.

* Address correspondence to this author at the Biotechnology Research Institute, 6100 Royalmount Avenue, Montreal, Canada H4P2R2. E-mail: robert.menard@nrc.ca.

‡ National Research Council of Canada.

§ Instituto Butantan.

|| Escola Paulista de Medicina.

⊗ Abstract published in *Advance ACS Abstracts,* October 1, 1997.

S0006-2960(97)01264-6 CCC: $14.00

interfering with binding of a large polypeptide substrate. Therefore, the insertion loop is often referred to as the "occluding loop". For the most commonly used cysteine protease substrate, Cbz-FR-MCA[1] , the rate of hydrolysis does not seem to be affected by the presence of the occluding loop since the level of activity of cathepsin B against this substrate is comparable to that observed with other cysteine proteases without an occluding loop (e.g., papain). In addition, hydrolysis of Cbz-FR-MCA is not affected by partial deletion of the occluding loop (Illy et al., 1997). This can be attributed to the fact that the aminomethylcoumarin portion of the substrate interacts only in the S_1' subsite of cathepsin B. Therefore, Cbz-FR-MCA is not an appropriate substrate to investigate the detailed contribution of the occluding loop to endopeptidase activity of cathepsin B.

Obtaining precise and reliable kinetic data for endopeptidase activity is a difficult task, and most kinetic studies with cysteine proteases have been carried out using small synthetic substrates instead of large protein substrates. This is a consequence of the existence of an extended binding site for the substrate, involving interactions in many subsites of the enzyme. The requirements for interaction in a given subsite can in the case of a polypeptide substrate be overruled by interactions of the polypeptide chain with other subsites of the enzyme. This gives rise to numerous possible cleavage patterns with large protein substrates. Intramolecularly quenched fluorogenic substrates (Yaron et al., 1979) offer a valid alternative to the use of protein substrates and are particularly well suited for the study of functional properties in the papain family of cysteine proteases. Abz and EDDnp are very good donor−acceptor groups for IQF substrates and can be separated by several amino acid residues (Hirata et al., 1994). They can therefore make use of the extended binding site of cysteine proteases, and with the proper choice of amino acid residues, these substrates can afford a reliable and well-defined assay for endopeptidase activity.

By using IQF substrates containing two or four amino acids on the C-terminal side of the scissile peptide bond, we have investigated the effect of the presence of the occluding loop on the endopeptidase activity of cathepsin B. The conformation and precise orientation of the occluding loop in cathepsin B is stabilized by contacts between residues within the loop itself as well as by interactions with the rest of the protein (Musil et al., 1991). The latter interactions involve two residues on the occluding loop, His110 and Arg116, which form salt bridges with Asp22 and Asp224, respectively. These two electrostatic interactions "lock down" the occluding loop in position, creating a compact structure. Recently, the crystal structure of procathepsin B showed that the occluding loop adopts a very different conformation in the proenzyme compared to the mature form (Cygler et al., 1996). The His110-Asp22 and Arg116-Asp224 contacts maintaining the loop in position in the mature enzyme are not present in the proenzyme. The relative importance of these interactions in defining the endopeptidase activity of cathepsin B has been investigated using site-directed mutagenesis. The results indicate that cathepsin B can possess a very high endopeptidase activity when the contacts between the loop and the enzyme are broken. The possible contribution of these interactions in regulating the endopeptidase activity of cathepsin B and the implications for cathepsin B activity in physiological or pathological conditions are discussed.

MATERIALS AND METHODS

Restriction endonucleases, T4 DNA ligase, and T4 DNA polymerase were purchased from New England Biolabs (Mississauga, Ontario). The vector (pPIC9) and *Pichia pastoris* strain GS115 were purchased from Invitrogen Corporation (San Diego, CA). The substrate Cbz-FR-MCA and the irreversible inhibitor E-64 were purchased from IAF Biochem International Inc. (Laval, Quebec). Abz-QVVAGA-EDDnp was from Bachem (King of Prussia, PA) Abz-AFRSAAQ-EDDnp was synthesized using solid phase chemistry as described previously (Hirata et al., 1994). Human cathepsin L was obtained as described previously (Carmona et al., 1996). Papain (crystallized suspension in sodium acetate) and immobilized pepsin were purchased from Sigma-Aldrich Canada Ltd. (Oakville, Ontario). Papain was further purified and activated, and the active site was titrated as described previously (Ménard et al., 1990). Recombinant human cystatin C was a generous gift from Dr. Irena Ekiel (Biotechnology Research Institute).

Site-Directed Mutagenesis and Recombinant Protein Expression. In vitro site-directed mutagenesis was performed using the cDNA insert for human procathepsin B [S115A] (Mach et al., 1994) in the vector pGEM-11Zf (Promega, Madison, WI) and the Transformer Site-Directed Mutagenesis Kit (CLONTECH Laboratories, Inc., Palo Alto, CA). The Ser115 to Ala mutation eliminates a glycosylation site, and all variants of cathepsin B will therefore be nonglycosylated. It has been shown that the carbohydrate moiety of cathepsin B has no influence on enzymatic activity (Mach et al., 1992). In the text, wild-type cathepsin B refers to an enzyme form containing no mutation other than S115A. The oligonucleotides used for mutagenesis (mutated bases are underlined) were as follows: 5′-CAA AGA GAT CAG AGC TCA GGG CTC CTG TG-3′ (D22A); 5′-CCC TGT GAG GCA CAT GTC AAC GGC G-3′ (H110A); 5′-CAT GTC AAC GGC GCT GCA CCC CCA TGC ACG G-3′ (R116A); 5′-CTC CTG GAA CAC TGC GTG GGG TGA CAA TGG-3′ (D224A). The changes introduced new *Sac*I (D22A), *Afl*III (H110A), and *Dra*III (D224A) sites and destroyed a *Nar*I site (R116A), features used for facile diagnosis of the mutations. Positive clones were identified by restriction site analysis and confirmed by DNA sequencing. The mutants were subcloned into the vector pPIC9 and expressed in the yeast *P. pastoris* as prepro-α-factor fusion constructs using the culture conditions recommended by Invitrogen.

Enzyme Activation and Purification. The proenzymes were expressed and secreted at levels of approximately 4−25 mg/L of initial culture medium. Yeast cells in suspension culture (1 L) were centrifuged at 3000*g* for 10 min, and the supernatant was concentrated 10-fold using an Amicon YM10 membrane in a stirred cell. The concentrated solution was diafiltered or dialyzed overnight at 4 °C against 50 mM

[1] Abbreviations: Cbz-FR-MCA, carbobenzoxy-L-phenylalanyl-L-arginine 4-methylcoumarinyl-7-amide; Abz-AFRSAAQ-EDDnp, *ortho*-aminobenzoyl-L-alanyl-L-phenylalanyl-L-arginyl-L-seryl-L-alanyl-L-alanyl-L-glutamine *N*-(ethylenediamine)-2,4-dinitrophenylamide; Abz-QVVAGA-EDDnp, *ortho*-aminobenzoyl-L-glutamyl-L-valyl-L-valyl-L-alanylglycyl-L-alanine *N*-(ethylenediamine)-2,4-dinitrophenylamide; IQF substrate, fluorogenic substrate with intramolecularly quenched fluorescence; E-64, 1-[[(L-*trans*-epoxysuccinyl)-L-leucyl]amino]-4-guanidino)butane.

sodium acetate, pH 5.0, to allow autoprocessing of proenzymes. Mutants carrying the mutation D22A or H110A autoprocessed slowly and were converted to the mature form at pH 4.7 with 50 U/mL of immobilized pepsin. Pepsin was removed by filtration. The processed enzymes were purified using an SP-Sepharose Phast Flow (Pharmacia Biotech Inc.) column (2.5 × 4 cm) equilibrated in 50 mM sodium acetate buffer, pH 5.0. The enzymes eluted between 150 and 250 mM NaCl. Purified proteins were stored at 4 °C in the elution buffer containing 100 μM−1 mM $HgCl_2$. N-terminal sequence analysis was performed on a Model 470A gas-phase sequencer equipped with an on-line Model 120A phenylthiohydantoin analyzer (Applied Biosystems Inc.). The pepsin cleavage site was found to be the same as previously reported for autoprocessed cathepsin B (Mach et al., 1994) and corresponds to a six-residue extension relative to the fully processed lysosomal form.

Kinetic Measurements. Kinetic experiments were performed as previously described (Ménard et al., 1990). Fluorescence was monitored on a SPEX Fluorolog-2 spectrofluorometer. For Cbz-FR-MCA, the excitation and emission wavelengths were set at 380 and 440 nm, respectively. Hydrolysis of the IQF substrates resulted in an increase in Abz fluorescence, which was monitored at 420 nm using an excitation wavelength of 320 nm (Chagas et al., 1991). The assay with the IQF substrates was calibrated by monitoring the fluorescence of the Abz-containing hydrolysis product in presence of various concentrations of non-hydrolyzed substrate, therefore reproducing exactly the experimental conditions. In this way, the calibration corrects for the inner filter effect, which becomes significant only when the substrate concentration is higher than ca 2 μM. In addition, it was found that the Abz fluorescence intensity is pH-dependent, decreasing at low pH according to a single pK_a transition (pK_a = 3.6). This was taken into account in the assay calibration.

Prior to the assay, the mercury-inhibited enzymes were activated by incubation in 50 mM sodium phosphate, 0.2 M NaCl, 5 mM EDTA, and 2 mM DTT, pH 6.0. The concentration of active enzymes was determined by titration using E-64 (Barrett et al., 1982). Enzymes carrying the mutation H110A proved to be more difficult to titrate with E-64, and recombinant cystatin C was used for the titration. All kinetic measurements were done at 25 °C in presence of 2 mM DTT and 3% DMSO. The reaction mixtures also contained 50 mM sodium citrate (pH 3.0−5.9), 0.2 M NaCl, and 1 mM EDTA or 50 mM sodium phosphate (pH 5.8−7.9), 0.2 M NaCl, and 5 mM EDTA. The kinetic parameters were determined by measuring the initial rate of hydrolysis at various substrate concentrations. For most experiments with the IQF substrates, the variation in initial rate was linear up to the highest substrate concentrations used. Therefore, only k_{cat}/K_M values could be determined. For pH−activity profiles, initial rates were determined at substrate concentrations much lower than K_M, and k_{cat}/K_M values were obtained by dividing the initial rates by enzyme and substrate concentrations. The pH−activity profiles for cathepsin B and cathepsin B mutants are extremely complex, which precludes the use of well-defined models to curvefit the data. The continuous curves drawn through the data points were obtained by non-linear regression of the data to empirical equations and are included solely for clarity of the graphical display.

Table 1: Second-Order Rate Constant k_{cat}/K_M for Hydrolysis of Abz-AFRSAAQ-EDDnp, Abz-QVVAGA-EDDnp, and Cbz-FR-MCA by Human Cathepsin B, Human Cathepsin L, and Papain at pH 6.0

	k_{cat}/K_M (10^5 M^{-1} s^{-1})		
substrate	cathepsin B	cathepsin L	papain
Abz-AFRSAAQ-EDDnp	0.0195 ± 0.0005	495 ± 10	33.0 ± 3.0
Abz-QVVAGA-EDDnp	0.030 ± 0.002	5.73 ± 0.58	148.3 ± 8.7
Cbz-FR-MCA	4.25 ± 0.30	502 ± 24	8.63 ± 0.56

Identification of Cleavage Sites for IQF Substrates. Substrates and enzymes were incubated under conditions similar to that for the kinetic assays. The reactions were stopped by the addition of a TFA solution. The hydrolysis fragments were purified on a reverse phase C-18 column (150 × 1.8 mm, J. T. Baker) using a Shimadzu Model C-R7A HPLC system equipped with a UV detector (220 nm) and a fluorescent detector with excitation and emission wavelengths set at 320 and 420 nm, respectively. Analytical HPLC was carried out with a linear gradient from 10% CH_3-CN to 50% CH_3CN, all in 0.1% TFA in 25 min at a flow rate of 1.7 mL/min. The peaks were collected manually, freeze-dried, and used for amino acid analysis and mass spectrometry. The mass spectrometric analysis was performed on a triple quadrupole mass spectrometer (API III LC/MS/MS system; Sciex, Thornhill, Canada).

RESULTS AND DISCUSSION

Endopeptidase Activity of Cathepsin B. The kinetic parameters for hydrolysis of two different IQF substrates are reported in Table 1. Substrate Abz-AFRSAAQ-EDDnp was chosen in order to position Phe and Arg residues in P_2 and P_1, corresponding to preferred residues for interaction in the S subsites of cathepsin B (Barrett, 1980). In this manner, the substrate covers both the S and S′ subsites, and cleavage of the Arg-Ser bond is a measure of endopeptidase activity. The nature of the cleavage site has been confirmed by HPLC analysis of the hydrolysis products, i.e., initial cleavage occurs at the Arg-Ser bond. A minor cleavage site was also observed at the Ser-Ala bond due to the fact that Arg is also well accepted in the S_2 subsite of cathepsin B (Barrett & Kirschke, 1981). This does not constitute a problem for the present study since cleavage at the Ser-Ala bond is also the result of endopeptidase activity. Prolonged incubation leads to the appearance of additional products, mainly resulting from the dipeptidylpeptidase activity of cathepsin B, which releases the dipeptide FR from Abz-AFR. As shown in Table 1, Abz-AFRSAAQ-EDDnp is a very good substrate for both cathepsin L and papain with k_{cat}/K_M values of 495 × 10^5 and 33.0 × 10^5 M^{-1} s^{-1}, respectively. With cathepsin B, however, the k_{cat}/K_M is only 0.0195 × 10^5 M^{-1} s^{-1}, a value 25000-fold and 1700-fold lower than those observed with cathepsin L and papain.

The low endopeptidase activity of cathepsin B against a IQF substrate has been reported previously using Abz-QVVAGA-EDDnp as a substrate (Lalmanach et al., 1995). Cleavage of this substrate by cathepsin B occurs at the Ala-Gly bond as determined by HPLC analysis, which corresponds to previously reported results (Lalmanach et al., 1995). The kinetic parameters obtained with Abz-QVVAGA-EDDnp in the present study are reported in Table 1. Both IQF substrates used in this study possess very different

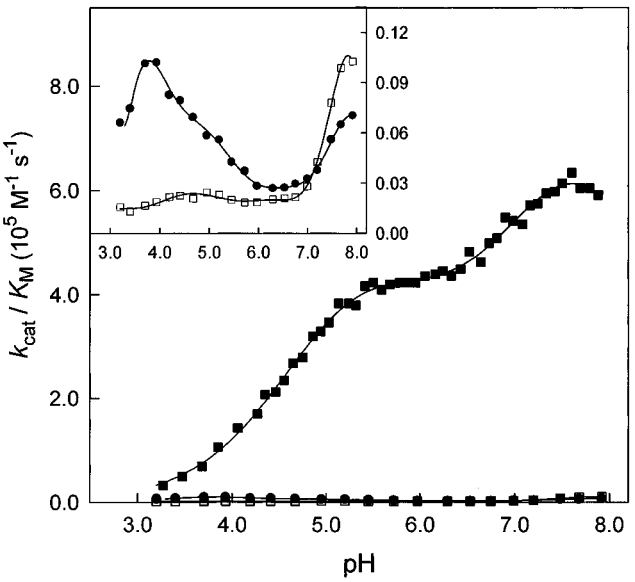

FIGURE 1: pH dependence of k_{cat}/K_M for hydrolysis of Abz-AFRSAAQ-EDDnp (□), Abz-QVVAGA-EDDnp (●), and Cbz-FR-MCA (■) by human cathepsin B. The inset shows the pH−activity profiles for the IQF substrates on a smaller scale.

peptide moieties, and the low endopeptidase activity observed with cathepsin B is unlikely to be due to specific residue−subsite interactions. Taken together, the results presented in Table 1 clearly show that, compared to cathepsin L and papain, cathepsin B is a poor endopeptidase.

pH Dependency of Endopeptidase Activity. In addition to their low rate of hydrolysis, the IQF substrates also display pH−activity profiles very different from those observed with small synthetic substrates like Cbz-FR-MCA (Figure 1). The pH dependency for hydrolysis of Cbz-FR-MCA by cathepsin B is characterized by a gradual increase in activity when the pH is raised from 3 to 8. The upper pH limit at which activity can be measured is dictated by enzyme instability at high pH. This upper limit can vary according to experimental conditions and the time required to measure the initial rate. The activity is modulated by the ionization of numerous residues on the enzyme, and the profile is complex as compared to that for example of papain (Khouri et al., 1991a). As shown in Figure 1, the level of activity against the IQF substrates remains much lower than that for Cbz-FR-MCA over the entire range of pH where cathepsin B is stable. The pH−activity profiles for Abz-AFRSAAQ-EDDnp and Abz-QVVAGA-EDDnp also contain features not observed with the smaller substrate (insert in Figure 1). With Abz-QVVAGA-EDDnp, increases in k_{cat}/K_M are observed both at low and high pH. The increase in activity at low pH is quite unusual for cysteine proteases of the papain family, which normally display maximum activity at near-neutral pH. The main feature observed with Abz-AFRSAAQ-EDDnp is an increase in activity at high pH. By comparison, the pH−activity profile for hydrolysis of Abz-QVVAGA-EDDnp by papain is bell-shaped and similar to that observed with Cbz-FR-MCA (data not shown).

The variations in activity observed in the pH profiles with the IQF substrates most likely reflect pH-dependent events related to the presence of the occluding loop and not found in other papain-like cysteine proteases. The detailed structural arrangement of the occluding loop can be obtained from the crystal structure of human cathepsin B (Musil et al., 1991). The loop is formed by residues Cys108-Cys119. It has been proposed that the Pro-Pro-Cys sequences flanking both ends of the loop and the existence of three inter-main chain hydrogen bonds within the loop can contribute to the stability of this structural element (Musil et al., 1991). In addition, electrostatic interactions (two salt bridges) exist between residues on the loop and within the rest of the enzyme (Figure 2a). The Asp224 side chain, which interacts with Arg116, is also in hydrogen bonding distance from Glu109. These interactions, which help in positioning the loop, involve ionizable groups and could therefore contribute to the pH dependency observed for the endopeptidase activity.

Contribution of Loop−Enzyme Interactions to Endopeptidase Activity. Site-directed mutagenesis has been used to investigate the importance of the His110-Asp22 and Arg116-Asp224 interactions for the endopeptidase activity of cathepsin B. A number of single and multiple mutants have been made and characterized against the IQF substrates as well as Cbz-FR-MCA (Table 2). It can be seen that removal of the His110-Asp22 salt bridge by replacing either His110 or Asp22 by Ala leads to 30- and 140-fold increases in k_{cat}/K_M, respectively, for hydrolysis of Abz-AFRSAAQ-EDDnp at pH 6.0. The fact that the increase in activity is more important for the D22A mutant as compared to H110A suggests that the presence of the negatively charged Asp22 with or without a hydrogen bonding partner in the mutant H110A has a detrimental effect on activity. This is supported by the observation that both mutants have the same activity at pH 4.0, i.e., 1.20×10^5 M^{-1} s^{-1} for H110A, 1.27×10^5 M^{-1} s^{-1} for D22A corresponding to a 60-fold increase in activity relative to wild-type cathepsin B. Similarly, mutation of either Arg116 or Asp224 to Ala causes increases in k_{cat}/K_M ranging between 12- and 24-fold at pH 6.0 and pH 4.0 for hydrolysis of Abz-AFRSAAQ-EDDnp. The same observations can be made for the results obtained with the substrate Abz-QVVAGA-EDDnp, i.e., the single mutants all display appreciable increases in k_{cat}/K_M at both pH 6.0 and pH 4.0 (Table 2). By comparison, the mutations have little effect on the activity against Cbz-FR-MCA.

All of the above mutations will break one of the two salt bridges between the loop and the rest of the enzyme. Further increases in activity are observed when residues involved in both salt bridges are replaced by alanine, and the pH dependency of activity approaches bell-shaped profiles (Figure 3a,b). For the substrate Abz-AFRSAAQ-EDDnp, the double mutant H110A/R116A shows an important increase in k_{cat}/K_M as compared to H110A over the entire range of pH (Figure 3a). It must be noted however that a similar level of activity can be obtained by replacement of both His110 and Asp22 by Ala (D22A/H110A mutant in Figure 3a). This can be attributed to the unfavorable effect of the negatively charged Asp22 in the vicinity of the active site. Accordingly, the highest activity is observed for the triple-mutant D22A/H110A/R116A, where k_{cat}/K_M values are as high as 12×10^5 M^{-1} s^{-1}, corresponding to a 600-fold increase relative to the wild-type enzyme and approaching the level of activity observed with cathepsin L or papain (Table 1). The same observations can be made with the substrate Abz-QVVAGA-EDDnp (Figure 3b) with the exception that the double-mutant D22A/H110A and triple-mutant D22A/H110A/R116A now display similar levels of activity. Such a difference in the results with these two

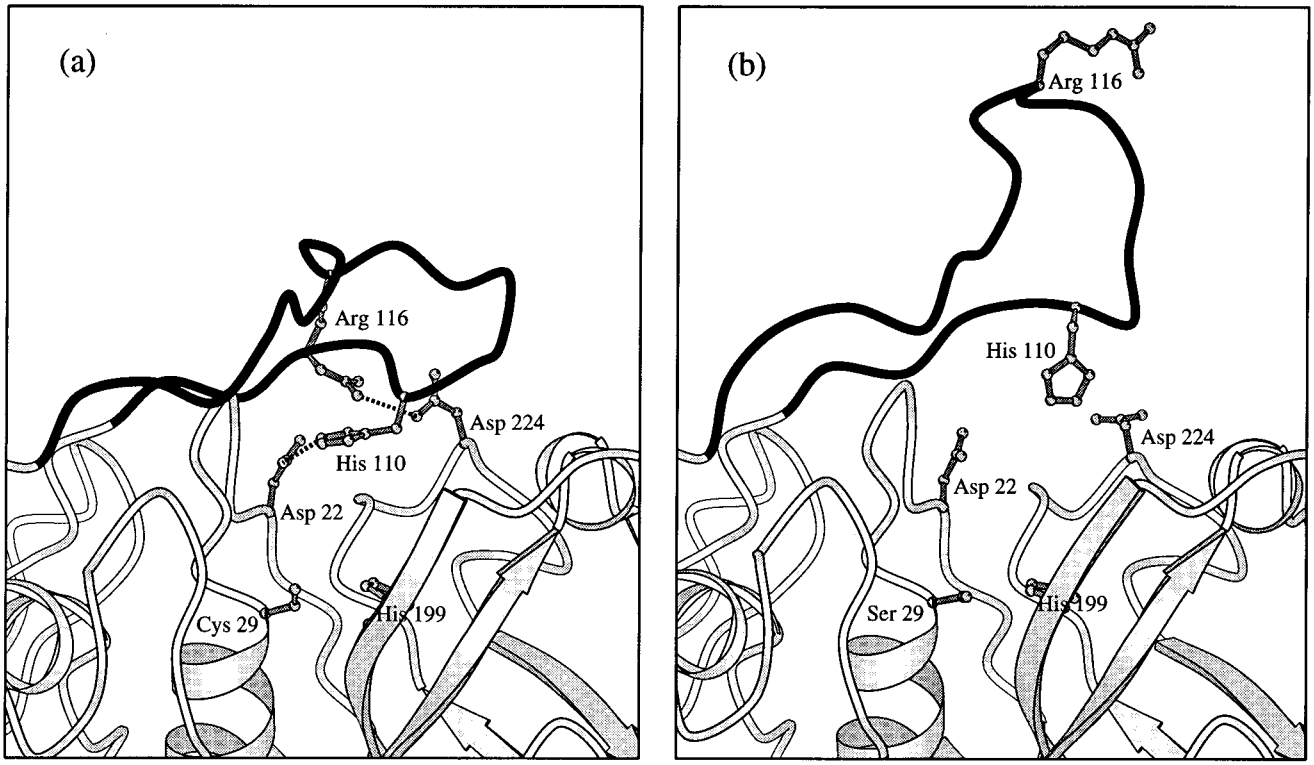

FIGURE 2: Conformation of the occluding loop in cathepsin B (a) and procathepsin B (b). The coordinates were obtained from the Brookhaven Protein Data Bank (1huc.pdb, 1mir.pdb). The occluding loop is represented in black. The side chains of the residues involved in the loop to enzyme contact in cathepsin B are shown (His110-Asp22, Arg116-Asp224). In addition, the active site residues Cys29 (Ser29 in the procathepsin B structure) and His199 are also displayed. For the procathepsin B structure, the proregion is not displayed to better visualize the conformation of the loop.

Table 2: Second-Order Rate Constant k_{cat}/K_M for Hydrolysis of Abz-AFRSAAQ-EDDnp, Abz-QVVAGA-EDDnp, and Cbz-FR-MCA at pH 6.0 and pH 4.0 by Wild-Type and Mutants of Human Cathepsin B

	k_{cat}/K_M (10^5 M^{-1} s^{-1})					
	pH 6.0			pH 4.0		
enzyme	Abz-AFRSAAQ-EDDnp	Abz-QVVAGA-EDDnp	Cbz-FR-MCA	Abz-AFRSAAQ-EDDnp	Abz-QVVAGA-EDDnp	Cbz-FR-MCA
wild type	0.0195 ± 0.0005	0.030 ± 0.002	4.25 ± 0.30	0.0215 ± 0.0005	0.075 ± 0.010	1.45 ± 0.10
H110A	0.62 ± 0.04	0.60 ± 0.10	3.63 ± 0.17	1.20 ± 0.05	3.75 ± 0.05	1.69 ± 0.01
D22A	2.78 ± 0.03	1.53 ± 0.47	13.40 ± 0.90	1.27 ± 0.07	2.60 ± 0.50	5.60 ± 0.50
R116A	0.46 ± 0.06	0.48 ± 0.10	23.0 ± 1.4	0.27 ± 0.03	0.90 ± 0.10	9.48 ± 0.03
D224A	0.277 ± 0.032	0.275 ± 0.045	8.82 ± 0.22	0.37 ± 0.03	1.10 ± 0.15	12.4 ± 1.2
H110A/R116A	2.74 ± 0.06	2.05 ± 0.05	4.24 ± 0.11	4.52 ± 0.32	10.75 ± 0.25	1.77 ± 0.03
H110A/D224A	0.285 ± 0.035	0.225 ± 0.055	2.45 ± 0.20	0.68 ± 0.04	2.10 ± 0.40	2.15 ± 0.05
D22A/H110A	3.90 ± 0.10	12.8 ± 1.4	8.02 ± 0.33	5.07 ± 0.03	39.2 ± 2.4	3.85 ± 0.75
D22A/H110A/R116A	9.28 ± 0.18	16.35 ± 0.85	10.55 ± 0.15	10.65 ± 0.85	28.8 ± 3.8	5.25 ± 0.25

substrates is not surprising if one considers that Abz-AFRSAAQ-EDDnp contains four amino acids on the C-terminal side of the scissile peptide bond, while Abz-QVVAGA-EDDnp contains only two. These two substrates might interact to a different extent in the S′ subsites of cathepsin B and therefore respond differently to mutations in the region of the occluding loop. The overall results however are still comparable for the two substrates, indicating that the large increases in activity observed for mutations in the loop region are likely to reflect events at the enzyme level independent of the peptide sequence of the substrate. The large increase in activity observed for these mutants can be attributed to an increased "flexibility" of the occluding loop, which can move out when the substrate binds to the enzyme. Alternatively, for certain mutants, the loop might no longer be blocking access to the substrate binding site. It must be noted that no increase in activity was observed for hydrolysis of azocasein when the occluding loop was completely removed by mutagenesis (Illy et al., 1997). However, as pointed out by the authors, there was considerable loss in activity of the mutant enzyme during the incubation time used in the assay, and as such, the activity of the loop-deleted mutant against azocasein is underestimated.

Compared to the results with the IQF substrates, the single- and multiple-mutants have less pronounced effects on the pH−activity profiles for hydrolysis of Cbz-FR-MCA (Figure 3c). The small increase in activity observed at around pH 7 in cathepsin B is absent in mutants where His110 is replaced by Ala. The disappearance of this transition can also be noted in the profiles with the IQF substrates, suggesting that deprotonation of His110 might be responsible for the increase in activity at high pH in the wild-type enzyme. His110 forms a buried salt bridge with Asp22 and is expected to have a pK_a higher than normally observed for a histidine residue. It is interesting to note that the pH−activity profile for

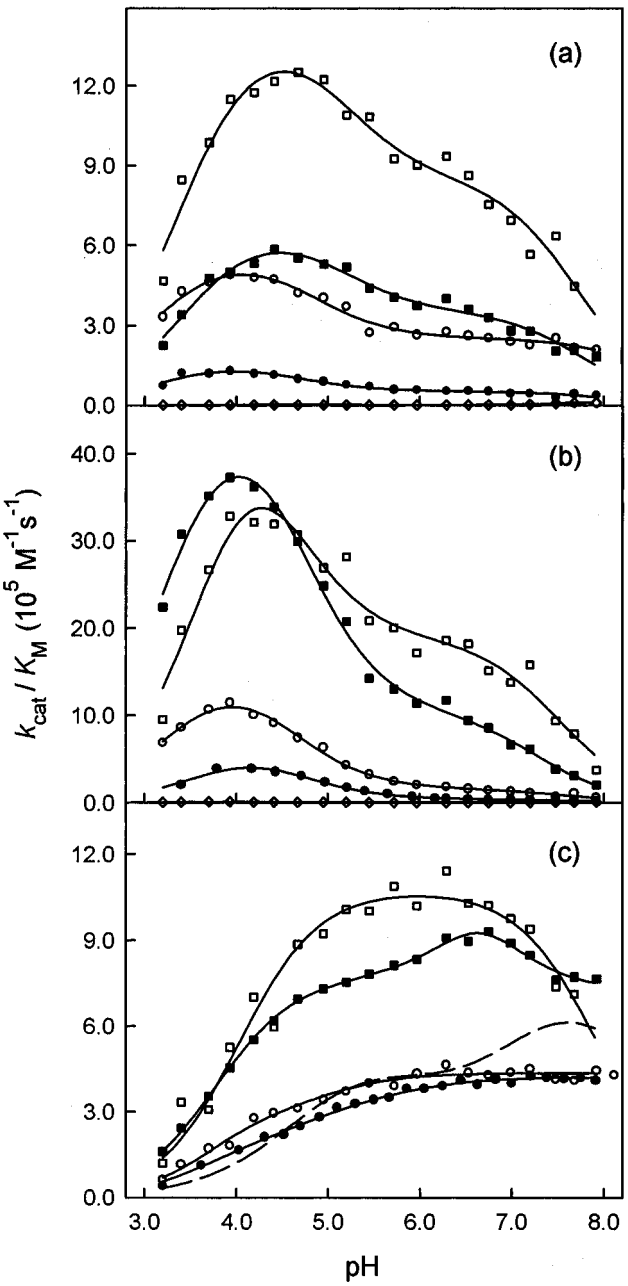

FIGURE 3: pH dependence of k_{cat}/K_M for hydrolysis of Abz-AFRSAAQ-EDDnp (a), Abz-QVVAGA-EDDnp (b), and Cbz-FR-MCA (c) by wild-type (◊), H110A (●), H110A/R116A (○), D22A/H110A (■), and D22A/H110A/R116A (□) cathepsin B. The data points have been omitted for hydrolysis of Cbz-FR-MCA by wild-type cathepsin B for clarity.

hydrolysis of Cbz-FR-MCA by the triple-mutant is almost identical to that for hydrolysis by papain. This indicates that the complexity of pH−activity profiles observed for hydrolysis of small synthetic substrates by cathepsin B can in large part be attributed to features related to the occluding loop. It is therefore becoming clear that many of the differences in functional properties between cathepsin B and other cysteine proteases of the papain family are related to the presence of this occluding loop.

An interesting result was obtained when Asp224 is mutated instead of Arg116 in order to break the Arg116-Asp224 interaction. As shown in Figure 4 for hydrolysis of Abz-QVVAGA-EDDnp, the single mutants H110A and D224A both display significant increases in activity. However, the double-mutant H110A/D224A does not show an additional

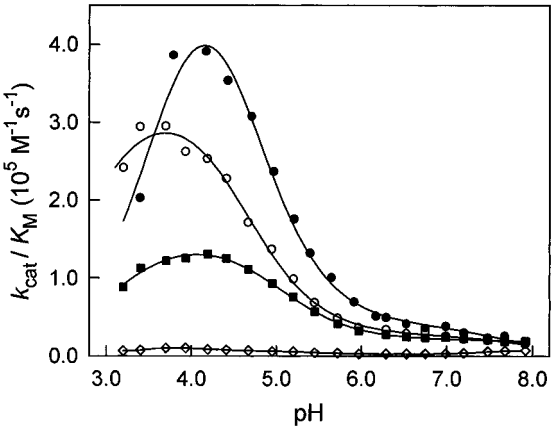

FIGURE 4: pH dependence of k_{cat}/K_M for hydrolysis of Abz-QVVAGA-EDDnp by wild type (◊), H110A (●), D224A (■), and H110A/D224A (○) cathepsin B.

important increase in activity as observed with the double- and triple-mutants presented in Figure 3b. Instead, the level of activity with the H110A/D224A mutant is lower than for the single-mutant H110A. Careful examination of the cathepsin B structure offers an explanation for this result. It can be seen that, when both His110 and Asp224 are replaced by alanine residues, the side chain of Arg116 can reorient slightly and form a new salt bridge with the side chain of Asp22. Therefore, Arg116 is a better candidate for mutation in order to break the loop to enzyme interactions. This result strongly supports however the hypothesis that the large increase in endopeptidase activity observed in this study are due to an increase in loop flexibility.

To confirm that k_{cat}/K_M values reported for the mutants still reflect endopeptidase activity, the nature of the cleavage sites were identified by HPLC analysis as described under Materials and Methods. Cleavage occurred either at the Arg-Ser or Ser-Ala bond for Abz-AFRSAAQ-EDDnp and at the Ala-Gly or Val-Ala bond for Abz-QVVAGA-EDDnp for the mutant enzymes. Free EDDnp was not detected. It is not known exactly why the cleavage pattern for some mutant enzymes is different from that of wild-type cathepsin B. However, it is important to note that for the cleavage sites indentified there is always two to four residues on the substrate able to interact with the S′ subsites of cathepsin B and that no free EDDnp was detected. Therefore, the mutations reported in the present study do reflect increases in the endopeptidase activity of cathepsin B.

The shapes of the pH−activity profiles with the IQF substrates in Figures 3 and 4 are still different from those observed with enzymes that do not contain an occluding loop. Even with the triple-mutant D22A/H110A/R116A, a small decrease in activity is observed with IQF substrates around pH 5−6 that is not present in papain (data not shown). This decrease in activity is possibly attributable to an event occurring in the S′ subsites, since this transition is not observed for hydrolysis of Cbz-FR-MCA by the triple-mutant (Figure 3c). In general, it is difficult to ascribe the modulations of activity observed in the pH profiles to particular pH-dependent events at the molecular level due to the complexity of the system. Cathepsin B possesses several acid residues in the vicinity of the active site able to modulate the electrostatic environment (Glu171, Glu36, Glu122, Glu245), which can explain why many mutants display a higher level of activity at low pH. The fact that

for the mutant enzymes cleavage can occur at both the Arg-Ser and Ser-Ala bonds in Abz-AFRSAAQ-EDDnp can also add to the complexity of pH−activity profiles. Cleavage at the Ser-Ala bond places an arginine in the P_2 position, which can interact in a pH-dependent manner with Glu245 in the S_2 subsite of cathepsin B (Khouri et al., 1991b; Hasnain et al., 1993). It must be noted however that the pH−activity profiles obtained with Abz-AFRSAAQ-EDDnp are relatively similar to those obtained with Abz-QVVAGA-EDDnp, indicating that this interaction does not have a major contribution to the profiles. It might however explain the smaller difference in activity between pH 6 and pH 4 observed with Abz-AFRSAAQ-EDDnp as compared to Abz-QVVAGA-EDDnp for some of the mutant enzymes.

Does Endopeptidase Activity of Cathepsin B Involve Movement of the Occluding Loop Under In Vitro or In Vivo Conditions? Cathepsin B is generally considered to possess both endo- and exopeptidase activities (Barrett & Kirschke, 1981). However as shown in this work, the endopeptidase activity of cathepsin B is relatively poor as compared to other cysteine proteases of the papain family. The present study has established a relationship between the occluding loop and the endopeptidase activity of cathepsin B. Previous work has shown that the occluding loop is also essential for the dipeptidyl carboxypeptidase activity of the enzyme (Illy et al., 1997). This is due to the fact that the loop contains residues His110 and His111 that can interact with the C-terminal carboxylate of a substrate (Musil et al., 1991). A more detailed investigation of the contribution of individual residues to the exopeptidase activity of cathepsin B is in progress (John Mort, personal communication). We have shown by using site-directed mutagenesis that when the contacts between the occluding loop and the enzyme are broken or removed, the endopeptidase activity of cathepsin B increases dramatically to a level similar to that observed with other cysteine proteases that do not contain an occluding loop. For the mutant enzymes, the large increase in activity could be due to the increased "flexibility" of the occluding loop, which might even no longer block access to the substrate binding site. Electrostatic factors also contribute to the observed variations but do not seem to be the dominant factor. This raises the question of whether or not the loop has to "move out" when substrate binds for the endopeptidase activity of wild-type cathepsin B, where several interactions create a compact structure and hold the loop in position. Is the low-level endopeptidase activity due to non-optimal binding of the substrate "around" the loop or to the fact that the substrate has to displace the loop before being hydrolyzed? It is impossible to give a definitive answer to this question with the experiments performed in this study. However, various observations suggest that the occluding loop could be displaced by a substrate binding in the endopeptidase mode. It has been shown that in procathepsin B the loop can exist in an "open" conformation (Cygler et al., 1996). In this case, the presence of the proregion binding at the active site and in the S′ subsites of cathepsin B is responsible for the different conformation adopted by the occluding loop. Contrary to the mature enzyme, the structure of the loop in the open conformation is less compact, and there is no evidence for a preferred conformation. In fact, the shape of the occluding loop in procathepsin B is dictated by the shape of the prosegment, suggesting that the loop in the open conformation is more flexible and could adapt to accommodate a protein substrate (Cygler et al., 1996). It has been shown also that a peptide corresponding to the proregion of procathepsin B is a very potent inhibitor of the mature enzyme (Fox et al., 1992). Given the binding mode of the proregion in procathepsin B, inhibition of cathepsin B by this propeptide is very likely to involve displacement of the occluding loop by the peptide. Accordingly, removal of the occluding loop by site-directed mutagenesis was shown to enhance binding of the propeptide by more than 50-fold (Illy et al., 1997). Further work is necessary to confirm that the endoproteolytic activity of cathepsin B involves movement of the occluding loop. For example, a mutation that would lock down the loop to the enzyme could completely inhibit the endopeptidase activity of cathepsin B. If this is the case, exploitation of this aspect could lead to novel strategies for designing cathepsin B inhibitors.

In the present study, interactions that hold the occluding loop in position have been altered by site-directed mutagenesis. Other more physiologically relevant methods to weaken or break these contacts could exist. For example, elements within the pH-activity profiles for wild-type cathepsin B with IQF substrates can be interpreted to reflect a pH control of endopeptidase activity and possibly of loop−enzyme contacts. Therefore, the pH of the environment could affect the "stability" of the loop. In addition, the interactions with components present in a complex biological environment could also affect the positioning of the occluding loop in cathepsin B. In certain pathological conditions such as tumor cell invasion and metastasis, cathepsin B can exist in association with the plasma membrane (Sloane et al., 1986; Sloane, 1990). It has been shown that membrane-bound forms of cathepsin B display modified properties, e.g., resistance to inactivation at neutral pH (Sloane et al., 1988). If interactions with membranes affect the positioning of the occluding loop, one would expect based on the results of this study that the enzyme could have very good endopeptidase activity and might also be poorly inhibited by compounds like CA-030, which require interactions with residues present on the loop (Turk et al., 1995). However, such a change might also render the enzyme more susceptible to inhibition by endogenous protein protease inhibitors (e.g., cystatin C). If the functional properties of cathepsin B under physiological or pathological conditions are significantly different from those of cathepsin B in vitro, the common strategies now being used to develop cathepsin B inhibitors might have to be reviewed. Studies addressing these points remain to be done.

ACKNOWLEDGMENT

We would like to thank Dr John S. Mort (Shriners Hospital for Crippled Children, Montreal) for helpful discussions and for providing the cDNA for human procathepsin B and René Coulombe for helping in the preparation of figures.

REFERENCES

Barrett, A. J. (1980) *Biochem. J. 187*, 909−912.

Barrett, A. J., & Kirschke, H. (1981) *Methods Enzymol. 80*, 535−561.

Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M., & Hanada, K. (1982) *Biochem. J. 201*, 189−198.

Carmona, E., Dufour, E., Plouffe, C., Takebe, S., Mason, P., Mort, J. S., & Ménard, R. (1996) *Biochemistry 35*, 8149−8157.

Chagas, J. R., Juliano, L., & Prado, E. S. (1991) *Anal. Biochem. 192*, 419−425.
Cygler, M., Sivaraman, J., Grochulski, P., Coulombe, R., Storer, A. C., & Mort, J. S. (1996) *Structure 4*, 405−416.
Fox, T., deMiguel, E., Mort, J. S., & Storer, A. C. (1992) *Biochemistry 31*, 12571−12576.
Fox, T., Mason, P., Storer, A. C., & Mort, J. S. (1995) *Protein Eng. 8*, 53−57.
Hasnain, S., Huber, C. P., Muir, A., Rowan, A. D., & Mort, J. S. (1992) *Biol. Chem. Hoppe-Seyler 373*, 413−418.
Hasnain, S., Hirama, T., Huber, C. P., Mason, P., & Mort, J. S. (1993) *J. Biol. Chem. 268*, 235−240.
Hirata, I. Y., Cezari, M. H. S., Nakaie, C. R., Boschcov, P., Ito, A. S., Juliano, M. A., & Juliano, L. (1994) *Lett. Pept. Sci. 1*, 299−308.
Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T., & Mort, J. (1997) *J. Biol. Chem. 272*, 1197−1202.
Jia, Z., Hasnain, S., Hirama, T., Lee, X., Mort, J. S., To, R., & Huber, C. P. (1995) *J. Biol. Chem. 270*, 5527−5533.
Khouri, H. E., Plouffe, C., Hasnain, S., Hirama, T., Storer, A. C., & Ménard, R. (1991a) *Biochem. J. 275*, 751−757.
Khouri, H. E., Vernet, T., Ménard, R., Parlati, F., Laflamme, P., Tessier, D. C., Gour-Salin, B., Thomas, D. Y., & Storer, A. C. (1991b) *Biochemistry 30*, 8929−8936.
Kirschke, H., & Barrett, A. J. (1987) in *Lysosomes: their role in protein breakdown* (Glaumann, H., & Ballard, F. J., Eds.) pp 193−238, Academic Press, London.
Lalmanach, G., Serveau, C., Brillard-Bourdet, M., Chagas, J. R., Mayer, R., Juliano, L., & Gauthier, F. (1995) *J. Protein Chem. 14*, 645−653.
Mach, L., Stüwe, K., Hagen, A., Ballaun, C., & Glössl, J. (1992) *Biochem. J. 282*, 577−582.
Mach, L., Schwihla, H., Stuwe, K., Rowan, A. D., Mort, J. S., & Glössl, J. (1993) *Biochem, J. 293*, 437−442.
Mach, L., Mort, J. S., & Glössl, J. (1994) *J. Biol. Chem. 269*, 13036−13040.
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.
Mort, J. S., Recklies, A. D., & Poole, A. R. (1984) *Arthritis Rheum. 27*, 509−515.
Musil, D., Zucic, D., Turk, D., Engh, R. A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., Katunuma, N., & Bode, W. (1991) *EMBO J. 10*, 2321−2330.
Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
Sloane, B. F., Rozhin, J., Johnson, K., Taylor, H., Crissman, J. D., & Honn, K. V. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2483−2487.
Sloane, B. F., Rozhin, J., Lah, T. T., Day. N. A., Buck, M., Ryan, R. E., Crissman, J. D., & Honn, K. V. (1988) *Adv. Exp. Med. Biol. 233*, 259−268.
Sloane, B. F. (1990) *Semin. Cancer Biol. 1*, 137−152.
Taralp, A., Kaplan, H., Sytwu, I.-I., Vlattas, I., Bohacek, R., Knap, A. K., Hirama, T., Huber, C. P., & Hasnain, S. (1995) *J. Biol. Chem. 270*, 18036−18043.
Trabandt, A., Gay, R. E., Fassbender, H. G., & Gay, S. (1991) *Arthritis Rheum. 34*, 1444−1451.
Turk, D., Podobnik, Popovič, T., Katunuma, N., Bode, W., Huber, R., & Turk, V. (1995) *Biochemistry 34*, 4791−4797.
Turk, D., Podobnik, M., Kuhelj, R., Dolinar, M., & Turk, V. (1996) *FEBS Lett. 384*, 211−214.
Yaron, A., Carmel, A., & Katchalski-Katzir, E. (1979) *Anal. Biochem. 95*, 228−235.

BI971264+