S_2' substrate specificity and the role of His^{110} and His^{111} in the exopeptidase activity of human cathepsin B

Joanne C. KRUPA*, Sadiq HASNAIN†, Dorit K. NÄGLER‡[1], Robert MÉNARD‡ and John S. MORT*§[2]

*Joint Diseases Laboratory, Shriners Hospital for Children, 1529 Cedar Avenue, Montréal, Québec H3G 1A6, Canada, †National Research Council of Canada, 1500 Montréal Road, Ottawa, Ontario K1A 0R6, Canada, ‡Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montréal, Québec H4P 2R2, Canada, and §Department of Surgery, McGill University, Montréal, Québec H3A 2T6, Canada

The ability of the lysosomal cysteine protease cathepsin B to function as a peptidyldipeptidase (removing C-terminal dipeptides) has been attributed to the presence of two histidine residues (His^{110} and His^{111}) present in the occluding loop, an extra peptide segment located in the primed side of the active-site cleft. Whereas His^{111} is unpaired, His^{110} is present as an ion pair with Asp^{22} on the main body of the protease. This ion pair appears to act as a latch to hold the loop in a closed position. The exopeptidase activity of cathepsin B, examined using quenched fluorescence substrates, was shown to have a 20-fold preference for aromatic side chains in the P_2' position relative to glutamic acid as the least favourable residue. Site-directed mutagenesis demonstrated that His^{111} makes a positive 10-fold contribution to the exopeptidase activity, whereas His^{110} is critical for this action with the Asp^{22}–His^{110} ion pair stabilizing the electrostatic interaction by a maximum of 13.9 kJ/mol (3.3 kcal/mol). These studies showed that cathepsin B is optimized to act as an exopeptidase, cleaving dipeptides from protein substrates in a successive manner, because of its relaxed specificity in P_2' and its other subsites.

Key words: cysteine protease, dipeptidase, enzyme kinetics, occluding loop, quenched fluorescence substrates.

INTRODUCTION

Cathepsin B is a lysosomal cysteine protease of the papain superfamily that is generally believed to mediate intracellular protein turnover in the lysosome. However, it has been proposed that this peptidase also participates in a variety of physiological and pathological conditions [2,3]. Cathepsin B is believed to be involved in extracellular-matrix degradation in arthritis [4] and other pathologies, such as tumour invasion and metastasis [5]. In addition, recent studies using cathepsin $B^{-/-}$ mice demonstrated a function for this protease in tumour-necrosis-factor-α-mediated hepatocyte apoptosis [6]. These additional roles indicate that cathepsin B is an important target for drug development, where the distinctive features of the enzyme can be exploited in the design of selective inhibitors.

Unlike most other members of this family of proteases, cathepsin B displays both endo- (although relatively low compared with other cysteine proteases [7]) and exopeptidase activities. More specifically, as an exopeptidase, it removes dipeptides from the C-terminus of its substrates [8]. This peptidyldipeptidase activity is due to a structural feature particular to cathepsin B (when compared with other cysteine proteases), which is an extra peptide segment termed the occluding loop located in the primed side of the active site. On the basis of X-ray crystallographic studies, it was postulated that the exopeptidase activity is dependent on two histidine residues (His^{110} and His^{111}) located within the occluding loop and in the S_2' subsite [9]. Two salt-bridge interactions between the loop and the main body of the mature enzyme (Asp^{22}–His^{110} and Arg^{116}–Asp^{224}) maintain it in a closed position, so that these histidine residues form the outer boundary of the S_2' subsite (Figure 1). In contrast, the occluding loop adopts a drastically different conformation in procathepsin B where the propeptide passes directly through the active-site cleft, demonstrating the potential for mobility of this structural element in the active protease [10,11].

X-ray crystal structures of cathepsin B complexed with the epoxysuccinyl dipeptide inhibitors, CA030 and CA074 (the ethyl ester and propylamide respectively of L-3-carboxy-*trans*-2,3-epoxypropionyl-leucylamidoproline) [12,13], where the dipeptide portions of the inhibitors bind in the primed side of the active site, suggest an interaction between His^{110} and His^{111} should occur with the substrate's C-terminal carboxylate. Thus, as proposed by Musil et al. [9], these residues should provide an anchor for binding the substrate and in turn enhance the rate of substrate cleavage.

The role of the occluding loop has also been clarified by mutational analysis. It was demonstrated that deletion of the entire loop effectively eliminated exopeptidase activity while conserving endopeptidase activity [14]. Removal of the Asp^{22}–His^{110} and Arg^{116}–Asp^{224} ion pairs resulted in a dramatic increase in the endopeptidase activity of cathepsin B, as measured by extended quenched fluorescence substrates [7], suggesting that the loop is deflected when peptides bind for cleavage in the endopeptidase mode. In addition, the pH dependency of the binding of the cathepsin B propeptide was shown to be attributable to the Asp^{22}–His^{110} ion pair [15].

Building on these studies, our aims were 2-fold: first, to establish the nature of the S_2'-binding site through a systematic

Abbreviations used: Abz, 2-aminobenzoyl; Amc, 7-amino-4-methylcoumarin; Z-FR-Amc, benzyloxycarbonyl-L-phenylalanyl-L-arginine 4-methylcoumarin-7-ylamide; CA030 and CA074, the ethyl ester and propylamide respectively of L-3-carboxy-*trans*-2,3-epoxypropionyl-leucylamidoproline; $F(4NO_2)$, 4-nitro-L-phenylalanine; -EDDnp, -*N*-ethylenediamine-2,4-dinitrophenylamide; TFA, trifluoroacetic acid. The nomenclature used to describe the substrate residues and the subsites on the protease is that of Schechter and Berger [1]. Substrate C-terminal carboxylates are assumed to be free unless otherwise specified.

[1] Present address: Abteilung Klinische Chemie und Klinische Biochemie, Klinikum der Universität München, Chirurgische Klinik und Poliklinik-Innenstadt, Nussbaumstrasse 20, 80336 Munich, Germany.

[2] To whom correspondence should be addressed (e-mail jmort@shriners.mcgill.ca).

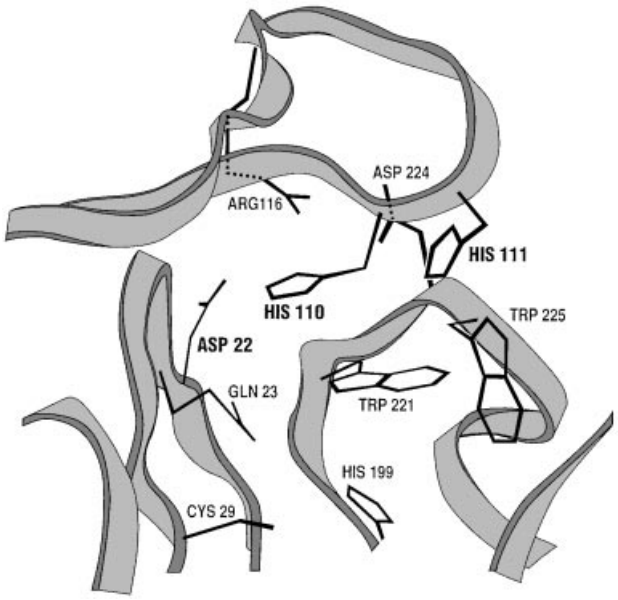

Figure 1 Representation of the occluding loop region in the human cathepsin B

Cys^{29} and His^{199} in the cathepsin B structure (Protein DataBank accession number 1HUC and [9]) act as the catalytic nucleophile and general base. Gln^{23} stabilizes the oxyanion tetrahedral intermediate, whereas Trp^{221} and Trp^{225} form a hydrophobic pocket around the active site. Residues mediating the peptidyldipeptidase activity are indicated in **bold**. Ion pairs are formed between Asp^{22} and His^{110}, and between Arg^{116} and Asp^{224}, whereas His^{111} is unpaired.

kinetic analysis using a series of internally quenched fluorescent substrates having the sequence 2-aminobenzoylphenylalanyl-argininylphenylalanyl-4-nitro-L-phenylalanine-X [Abz-FRF($4NO_2$)X], where X is the residue that interacts with S_2'; second, to investigate the contributions of Asp^{22}, His^{110} and His^{111} to the peptidyldipeptidase activity through site-specific mutations of cathepsin B, and to evaluate the activities of these mutants with the above-mentioned exopeptidase substrates and their amidated derivatives. The results indicate His^{111} makes a 10-fold contribution to the exopeptidase activity of cathepsin, whereas His^{110} is essential for this activity. In addition, the Asp^{22}–His^{110} ion pair stabilizes the electrostatic interaction between the imidazolium ion and the substrates C-terminal carboxylate by a maximum of 13.9 kJ/mol (3.3 kcal/mol).

EXPERIMENTAL

Materials

Benzyloxycarbonyl-L-phenylalanyl-L-arginine 4-methylcoumarin-7-ylamide (Z-FR-Amc) was purchased from Novabiochem (San Diego, CA, U.S.A.) and *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane ('E-64') (anhydrous) was obtained from Peptide Research Institute (Osaka, Japan). The pPIC9 vector and the yeast *Pichia pastoris* strain GS115 were purchased from Invitrogen (San Diego, CA, U.S.A.).

Synthesis of 2-t-butyloxycarbonylaminobenzoic acid

Anthranilic acid (40 mmol, 5.48 g) was dissolved in a mixture of tetrahydrofuran (30 ml) and water (20 ml). Di-t-butyl dicarbonate (44 mmol, 9.60 g) and K_2CO_3 (44 mmol, 6.08 g) were then added, and the solution was stirred for 16 h at 22 °C in the absence of light. After increasing the volume with water (100 ml), the solution was brought to pH 10 by the addition of 1 M NaOH. The aqueous layer was washed with dichloromethane (3×40 ml), brought to pH 4 with solid citric acid and then extracted with ethyl acetate (3×50 ml). The ethyl acetate layer was then washed with 0.2 M HCl (3×40 ml), water (1×25 ml) and saturated NaCl (1×25 ml). The organic layer was dried over Na_2SO_4 and the solvents were removed by rotary evaporation. The yellowish crude product was recrystallized from the ethanol/water mixture to yield 2-t-butyloxycarbonylaminobenzoic acid as white crystalline needles (6.70 g, 71 % yield); melting point 155–156 °C; published melting point 149–150 °C [16]; ^{1}H-NMR (200 MHz, $CDCl_3$) δ 1.53 (9H, s), 7.02 (1H, dt, $J = 1.1$, 7.0 Hz), 7.54 (1H, dt, $J = 1.6$, 7.1 Hz), 8.09 (1H, dd, $J = 1.6$, 8.0 Hz), 8.45 (1H, d, $J = 7.8$ Hz), 8.7–9.3 (NH proton difficult to detect, broad s), 10.00 (1H, s); ^{13}C-NMR (50 MHz, $CDCl_3$) δ 28.9, 80.8, 112.8, 118.5, 120.8, 131.2, 134.9, 142.1, 151.9, 171.7.

Synthesis of 3,4-dihydro-4-oxo-1,2,3-benzotriazol-3-yl 2-t-butyloxycarbonylaminobenzoate

The synthesis of the above compound was performed as described previously [16], with the exception of the solvents being anhydrous and the addition of a catalytic amount of 4-dimethylaminopyridine. Recrystallization produced the product in 60 % yield; decomposition point 153–154 °C; published melting point 155–156 °C [16]; ^{1}H-NMR (200 MHz, $CDCl_3$) δ 1.47 (9H, s), 7.12 (1H, dt, $J = 1.0$, 8.2 Hz), 7.66 (1H, dt, $J = 1.6$, 8.8 Hz), 7.86 (1H, dt, $J = 1.1$, 8.1 Hz), 8.03 (1H, dt, $J = 1.5$, 7.7 Hz), 8.25 (1H, d, $J = 8.0$ Hz), 8.34 (1H, dd, $J = 1.5$, 8.1 Hz), 8.40 (1H, dd, $J = 1.3$, 8.0 Hz), 8.53 (1H, d, $J = 8.7$ Hz), 9.54 (1H, s); ^{13}C-NMR (50 MHz, $CDCl_3$) δ 28.1, 81.1, 109.5, 119.1, 121.6, 122.2, 125.8, 129.1, 131.2, 132.9, 135.6, 136.9, 143.6, 144.3, 150.5, 152.3, 164.1.

Synthesis and purification of cathepsin B substrates

The internally quenched fluorogenic substrates were synthesized on a 431A solid-phase synthesizer (Applied Biosystems), using standard Fastmoc chemistry. Either 2-t-butyloxycarbonylaminobenzoic acid (for which activation of the carbonyl for coupling was done via the synthesizer) or 3,4-dihydro-4-oxo-1,2,3-benzotriazol-3-yl 2-t-butyloxycarbonylaminobenzoate was used in the preparation of the peptides. The peptides were purified on a PerkinElmer PREP-10 octyl-reverse-phase column (1.0 cm × 25 cm) using a linear gradient of 0–80 % acetonitrile, containing 0.1 % trifluoroacetic acid (TFA), at a flow rate of 4.0 ml/min. The expected molecular mass of each peptide was confirmed by mass-spectral analysis using a matrix-assisted laser-desorption ionization–time-of-flight ('MALDI–TOF') mass spectrometer (LaserMat 2000, Finnigan).

Site-directed mutagenesis and expression of human recombinant cathepsin B

The human cathepsin B constructs for the mutations D22A ($Asp^{22} \rightarrow Ala$) and D22A/H110A ($Asp^{22} \rightarrow Ala/His^{110} \rightarrow Ala$) were prepared, and the mutant proteins purified as described previously [7]. The human cathepsin B H111A construct was prepared as described previously for similar mutations [17] with the following oligonucleotide (mutated bases are underlined): TGT-GAG-CAC-GCT-GTG-AAC-GCC-C. The mutation created a *Dra*III restriction site. The constructs for human cathepsin B H110Q/A were prepared as described below. Site-directed mutagenesis *in vitro* was performed on a previously prepared cDNA construct in which the glycosylation site at Asn^{113} in the mature region had been removed due to the introduction of a S115A mutation [18]. An *Xho*I and *Not*I fragment containing

the procathepsin B cDNA was subcloned into PGEM-11Zf+. Using the method of Kunkel [19] and single-stranded DNA of the above fragment (designated as pCBH30), His^{110} was transformed to Gln or Ala respectively with the following oligonucleotides (mutated bases are underlined): CC-TGT-GAG-CAA-CAC-GTC-AAC-G and T-CCC-TGT-GAG-GCA-CAC-GTC-AAC-G. These mutations introduced an *Aff*III restriction site. DNA sequencing was used to confirm the mutations. The mutant cDNAs were subcloned into the vector pPIC9 (an *Escherichia coli*/yeast shuttle vector), and introduced into the yeast *P. pastoris*, strain GS115, by electroporation. Recombinant human procathepsin B was expressed by the methylotropic yeast under conditions described by the manufacturer (Invitrogen).

Enzyme activation and purification

Culture media were centrifuged, and the resulting supernatants were concentrated by ultrafiltration using an Amicon YM10 membrane. The concentrated media were then dialysed against 20 mM sodium acetate/1 mM EDTA (pH 5.0) to allow for autoprocessing. The processed enzyme was reversibly inactivated with methylmethanethiolsulphonate and purified on a CM-Sepharose fast-flow column (Amersham Pharmacia Biotech, Montreal, Canada) (2.5 cm × 12 cm) using a linear gradient of 0–1.0 M NaCl in 20 mM sodium acetate, pH 5.0. The enzyme was eluted with approx. 300 mM NaCl. N-terminal sequencing performed on an Applied Biosystems 473A Protein Sequencer revealed either a three or six amino acid N-terminal extension for the mutant enzymes, similar to that observed for the wild-type protein.

Kinetic measurements

Kinetic measurements were carried out on a PerkinElmer LS-3 fluorescence spectrophotometer equipped with Flu-Sys software (version 1.02 [20]) and a temperature-controlled water bath. The fluorimeter was calibrated prior to use with a known concentration of either 7-amino-4-methylcoumarin (Amc) or Abz-FR-OH. Fluorescence increased upon cleavage of the intramolecularly quenched Abz substrates ($\lambda_{excitation}$, 335 nm; $\lambda_{emission}$, 415 nm) or upon release of the Amc moiety ($\lambda_{excitation}$, 380 nm; $\lambda_{emission}$, 440 nm). With the Abz substrates, a significant inner-filter effect limited the substrate concentration to $< 10\ \mu$M. The concentration of a substrate stock solution was determined by the fluorescent output upon total hydrolysis by excess protease. The fluorescence outputs of these substrates and their amidated derivatives was 16.5-fold and 10.5-fold respectively compared with background, making them sensitive kinetic probes.

Prior to an enzymic assay, protein preparations were subjected to activation on ice for 10 min in buffer containing 1 mM dithiothreitol. The buffer was composed of 25 mM sodium phosphate, pH 6.0, 0.5 mM EDTA and 0.0125 % Brij-35. The concentration of active wild-type and mutant enzymes were determined by titration with the irreversible cysteine protease inhibitor *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane [8]. The active-site titration required the enzyme to be incubated with the inhibitor for 1 h at 25 °C, prior to the measurement of any residual activity with the substrate Z-FR-Amc. All kinetic measurements were carried out at 25 °C, in the presence of the above buffer and 3 % DMSO in order to ensure substrate solubility.

For most of the kinetic trials with the internally quenched substrates, K_m was higher than the limiting substrate concentration (10 μM). As a result, only k_{cat}/K_m values were determined. The value of k_{cat}/K_m was obtained from the linear portion of plots of initial rate (< 5 % conversion) versus substrate concentration. When low enzymic activity was observed, the value of k_{cat}/K_m was obtained by dividing the initial rate by the substrate and enzyme concentrations. For those few cases in which it was possible to work in a range of substrate concentration above K_m, k_{cat} and K_m were determined using non-linear regression analysis [21]. However, in all cases, the results are reported as the ratio of k_{cat}/K_m. Each kinetic experiment was performed in duplicate. In addition, the values of the reported specificity constants were confirmed by following a complete progress curve {fitting to first-order kinetics ($k_{obs}/[E_0] = k_{cat}/K_m$)} generated under conditions where $[S] \ll K_m$. All parameters reported showed errors of less than 10 %. The thermostability of the different cathepsin B mutants was determined by incubation of the activated enzymes at 0 or 25 °C for periods of up to 3 h, followed by activity assay with Z-FR-Amc under conditions as described above.

Detection of potential amidase activity

The potential amidase activity of cathepsin B was evaluated using a sensitive fluorimetric assay for ammonia based on the reaction between *o*-phthaldialdehyde and β-mercaptoethanol [22]. An ammonia standard curve was generated using ammonium sulphate (0–0.27 μM) in 0.2 M sodium phosphate, pH 7.4, containing 3.72 mM *o*-phthaldialdehyde, 3.56 mM β-mercaptoethanol and 0.24 μM Abz-FRF($4NO_2$)F-NH_2 to reproduce as closely as possible the conditions in the mixture to be assayed. The assay was performed by adding 100 μl of the enzymic incubation mixture (initially containing 7 μM substrate) with 2.9 ml of the β-mercaptoethanol/*o*-phthaldialdehyde/phosphate buffer solution at the concentrations described above. The mixture was incubated at 22 °C for 45 min prior to any fluorescence measurement.

Determining cleavage sites in the exopeptidase and non-extended minimal endopeptidase substrates

Under similar conditions as those used to determine the kinetic parameters, the internally quenched fluorogenic substrates were subjected to almost complete hydrolysis by the various enzymes. The samples were kept at −20 °C until required. The reaction mixtures were diluted 1- or 2-fold in 1 % TFA prior to loading on to a reverse-phase HPLC column (C-8, PE XPRESS 3X3CR, 3.3 cm × 0.6 cm; PerkinElmer). The substrate and hydrolysis fragments were eluted from the column with a linear gradient of 25–40 % acetonitrile [in aqueous 3 mM SDS and 0.05 % (v/v) H_3PO_4] at a flow rate of 4.0 ml/min. The N-terminal fragment was detected by absorbance at 312 nm. Using synthetically prepared Abz-FR-OH as a standard, the point of cleavage was confirmed.

RESULTS

Exopeptidase substrates

To investigate the exopeptidase S_2'-subsite specificity of cathepsin B, quenched fluorogenic substrates were used that contained a fluorescent N-terminal Abz group internally quenched by interaction with a nitrophenylalanyl residue in the P_1' position [23]. The use of Phe and Arg in the P_2 and P_1 positions respectively was based on the previously demonstrated preferences of cathepsin B for these residues [7,24,25]. As a result, the Abz-FRF($4NO_2$)X derivatives (which are almost completely deprotonated at pH 6.0, the conditions used in the present study) and their amidated analogues can be used to investigate the P_2' preference of cathepsin B and the importance of the negatively charged C-terminal carboxylate. HPLC analysis confirmed that cleavage

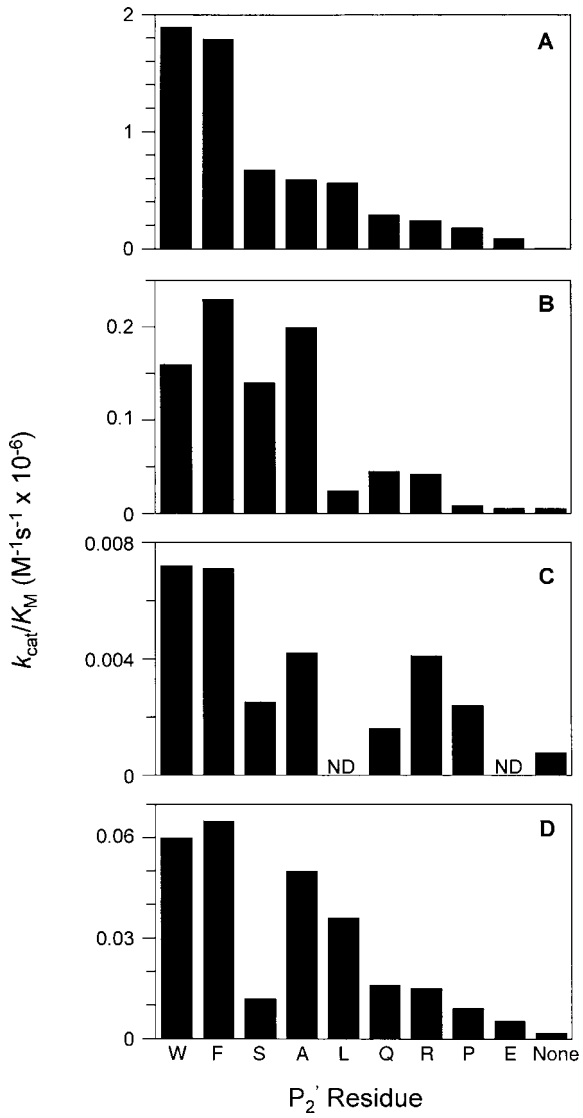

Figure 2 S_2' selectivity and effect of occluding loop mutations on cathepsin B activity

The specificity constants (k_{cat}/K_m) for wild-type and occluding loop cathepsin B mutants were determined at pH 6.0 using the exopeptidase substrate series, Abz-FRF(4NO$_2$)X where the P_2' residue (X) was as indicated. (**A**) Wild-type, (**B**) H111A, (**C**) D22A/H110A and (**D**) D22A. ND, no detectable activity; None, no P_2' residue.

occurred between the P_1 Arg and P_1' F(4NO$_2$), as expected and the same cleavage site was observed for the amidated derivatives. No appreciable amidase activity was detected.

Cathepsin B S_2' specificity

Examination of the second-order rate constants (k_{cat}/K_m) for cleavage of the Abz-FRF(4NO$_2$)X substrates with wild-type cathepsin B (Figure 2A) showed that the cathepsin B S_2' subsite prefers large aromatic residues. In the present study, the specificity constant for Abz-FRF(4NO$_2$) lacking the P_2' residue (represented as 'None' in Figure 2A) is two orders of magnitude lower than that for Abz-FRF(4NO$_2$)W, the most effective substrate {similar to the findings of Therrien et al. [26] with Abz-FRF(4NO$_2$)F}. However, activity with Abz-FRF(4NO$_2$) was only 8-fold lower than for Abz-FRF(4NO$_2$)E, the poorest member of the series, illustrating the broad specificity of the S_2' subsite. Although there was no major difference between the second-order rate constants for the aromatic amino acids in P_2', there was a 20-fold difference between substrates containing Trp and Glu.

Roles of occluding loop residues

The three-dimensional structure of cathepsin B and, in particular, of its complexes with CA030 and CA074 suggest that Asp^{22} and His^{110} play key roles in positioning the occluding loop, and His^{111} provides an anchor for the substrate's C-terminal carboxylate. The relative importance of these residues was investigated by site-directed mutagenesis. Cathepsin B containing mutations at Asp^{22}, His^{110} and His^{111} was expressed well in the *P. pastoris* system, and the proenzyme forms could be converted into the mature active enzyme by dialysis at low pH. All enzyme forms were stable at pH 6.0 for 3 h at either 0 °C or 25 °C, except for the H110A mutant, which showed a slight loss of activity at 0 °C, whereas at 25 °C it had a $t_{1/2}$ of approx. 9 h. The D22A mutant was sensitive to autocatalytic degradation at both 0 °C and 25 °C, where it had a $t_{1/2}$ of 12 h and 2 h respectively. All mutants had activity in the same order of magnitude as the wild-type enzyme with the standard cathepsin B substrate Z-FR-Amc.

The H111A substitution decreased the exopeptidase activity of cathepsin B by an order of magnitude (Figure 2B), indicating that this residue contributes to substrate binding, but suggesting that it is not critical for exopeptidase activity. In general the pattern of activity with different P_2' residues was similar to that shown by the wild-type enzyme, although there was relatively greater activity with Ser and Ala at this position. In contrast to His^{111}, which is unpaired in the closed form of cathepsin, His^{110} is normally salt bridged to Asp^{22} (Figure 1). The double mutation D22A/H110A showed a dramatic loss of exopeptidase activity (Figure 2C). For example a 250-fold reduction was observed for the Phe-containing substrate when compared with the wild-type enzyme. This translates to a maximum contribution of 13.9 kJ/mol to transition-state stabilization by this ion pair [27]. It would be expected that, in the case of the H110A mutant, loss of this charge pairing and the resulting presence of the free negative charge on Asp^{22} would produce an environment that is averse to substrate binding. The double mutant D22A/H110A eliminates this negative charge environment, and therefore demonstrates on average an increase in the specificity constant of 5-fold over H110A (Figure 3A). The specificity constants for the single H110A mutant was three orders of magnitude lower compared with the wild-type enzyme, indicating that the imidazolium ion of His^{110} is critical for this exopeptidase activity (Figure 3A). The single D22A mutant showed a 10-fold increase in activity relative to that of the double mutant D22A/H110A (Figure 2D), indicating flexibility in the occluding loop and the ability of His^{111} to periodically interact with the substrate's carboxylate when the loop is mobile. In all cases, a similar trend in selectivity for the P_2' residue was seen.

Activity against amidated substrates

The effect of the removal of the charge-acceptor mechanism in the occluding loop on the action of cathepsin B was also investigated against substrates lacking a C-terminal carboxylate (Figure 3). In the case of the wild-type enzyme, the Abz-FRF(4NO$_2$)A-NH$_2$ substrate showed somewhat reduced activity compared with to its non-amidated counterpart (k_{cat}/K_m = 160000 $M^{-1} \cdot s^{-1}$, compared with 600000 $M^{-1} \cdot s^{-1}$). In the case of the H111A and D22A mutants, the activity was similar for the amidated and non-amidated substrates (k_{cat}/K_m, 50000 $M^{-1} \cdot s^{-1}$ and 69000 $M^{-1} \cdot s^{-1}$ respectively). In contrast, for the His^{110}

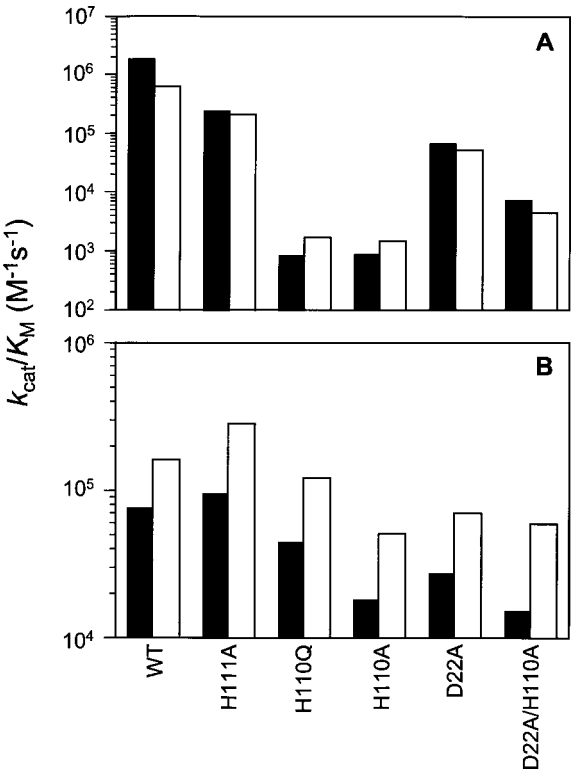

Figure 3 Comparison of activities of cathepsin B occluding loop mutants against amidated substrates

The specificity constant (k_{cat}/K_m) was determined for wild-type and cathepsin B occluding loop mutants, as indicated using the substrates Abz-FRF(4NO$_2$)F (closed bars) and Abz-FRF(4NO$_2$)A (open bars) (**A**) and their amidated forms (**B**).

mutants, which showed almost no activity with the non-amidated substrates, activity with the amidated substrates was comparable with that of the wild-type.

Activities with the amidated substrates also demonstrated a marked change in the response of cathepsin B with respect to the nature of the P'_2 residue. Whereas the wild-type enzyme showed a 3-fold preference for Phe compared with Ala at this position in the case of the non-amidated substrate, the occluding loop mutants showed much less of a preference. In the case of the amidated substrates all forms of the enzyme showed a similar 2-fold or greater preference for Ala relative to Phe at this position (Figure 3). In general, mutation of occluding loop residues caused a slight reduction in activity compared with the wild-type. However, the H111A mutant was clearly an exception for substrates lacking the C-terminal carboxylate. Activity with the amidated substrate Abz-FRF(4NO$_2$)A-NH$_2$ was 1.8 compared with that of wild-type, but even more impressive was the 4.8-fold increase in activity of the H111A mutant compared with the wild-type against Z-FR-Amc, the substrate most commonly used to assay members of the papain superfamily. In contrast, activity against this substrate was slightly lower than wild-type for the other occluding loop mutants.

DISCUSSION

The S'_2 subsite of cathepsin B, as described previously by Musil et al. [9] is a shallow hydrophobic depression, and crystallographic data for the epoxysuccinyl-dipeptides CA030 and CA074 bound into the active site of this enzyme, confirmed these findings [12,13]. The suggested minimal topography explains the relatively small difference in the second-order rate constants between the Trp- and Glu-containing substrates. A trend towards a preference for aromatic residues was observed, similar to previous findings using a variety of protein and small peptide substrates [28,29]. In addition, simple carboxypeptidase activity was observed, as has been documented previously in the autoprocessing of the C-terminal extension of cathepsin B [30]. The efficiency of removing single residues ranged from 8- to 150-fold less than the peptidyldipeptidase activity, depending on the P'_2 residue used for comparison. Turk et al. [12] suggested that the S'_2 subsite is filled by a proline residue; however, the groove could accommodate other side chains as well. In contrast, our results suggest that Pro is not the preferred residue in the S'_2 subsite. This is not a major discrepancy since the location of its side-chain in CA030, relative to a normal substrate, would be displaced by the distance of a carbonyl moiety. As shown previously by Berti et al. [31] for papain, modifications of the main chain can influence the interaction energies between the substrate and enzyme. Thus, CA030 can only approximate the P'_2 substrate.

Cathepsin B is a member of the mammalian family of papain-like cysteine proteases, of which there are presently eleven human members. Most of these enzymes consist of a simple catalytic unit and act as endopeptidases [32]. However, structural modifications of the basic papain framework are present in four members of this family, and these render cathepsins H, C, B and X into an aminopeptidase, dipeptidylpeptidase, peptidyldipeptidase and carboxypeptidase, which remove single residues or dipeptides from the N- or C-terminus respectively. In the case of cathepsins H and C, regions of the propeptide remaining following processing of the proenzyme are responsible for the ability of these enzymes to remove one or two residues respectively from the N-terminus [33,34]. In the case of cathepsins X and B, an additional peptide loop has been grafted on to the basic structure. The carboxydipeptidase activity of cathepsin B has been attributed to the occluding loop which provides an appropriately spaced acceptor for the negatively charged C-terminal carboxylate of the substrate P'_2 residue [9]. In cathepsin X, a mini-loop is present in which His^{23} has been speculated to have a similar role in defining the carboxypeptidase of this enzyme. In this case, the arrangement of the loop allows only a single residue in the primed subsites [35,36].

Two histidine residues are strategically positioned in the cathepsin B occluding loop and have been implicated in the exopeptidase activity. The present study demonstrates that mutation of His^{111} results in a 10-fold reduction of exopeptidase activity validating this hypothesis. However, the presence of the Asp^{22}–His^{110} ion pair plays a much more dramatic role. The assumption that this interaction acts to constrain the occluding loop into a functional conformation is supported by several findings. Binding of the cysteine protease inhibitor cystatin C to cathepsin B has been shown kinetically to require a two-step mechanism [37], whereas only a single-step mechanism was required for the H110A mutant [38]. Similarly, the binding of the cathepsin B propeptide, which requires relocation of the occluding loop [10], was dramatically enhanced in the case of this mutation [15]. Thus when either member of the ion pair (Asp^{22} or His^{110}) is mutated, the loop is expected to be more flexible. Therefore, the orientation of His^{111} should be displaced from its position in the wild-type enzyme and only able to interact favourably with the substrate carboxylate for a fraction of the time. With the His^{110} mutant, Asp^{22} would be expected to interact unfavourably with the substrate carboxylate through charge repulsion. Due to a combination of loop displacement

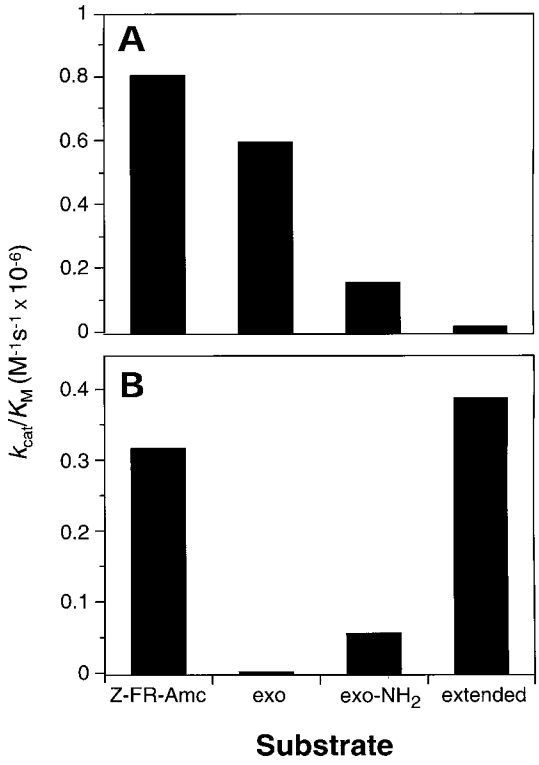

Figure 4 Summary of the effects of deletion of the Asp^{22}–His^{110} ion pair on cathepsin B activities

Specificity constants for wild-type cathepsin B (**A**) and the D22A/H110A mutant (**B**) using the general cysteine protease substrate Z-FR-Amc, Abz-FRF(4NO_2)A (exo), Abz-FRF(4NO_2)A-NH_2 (exo-NH_2), and the endopeptidase substrate Abz-AFRSAAQ-EDDnp (extended; results from [7]).

and charge repulsion, ion-pair mutants do not directly quantify the contribution of these residues to the exopeptidase activity. However, it is clear that His^{110} is essential for the exopeptidase activity and that a maximum contribution of 13.9 kJ/mol can be applied to the transition-state stabilization by this Asp^{22}–His^{110} ion pair.

The role of the Asp^{22}–His^{110} ion pair in the ability of cathepsin B to cleave different types of substrates is summarized in Figure 4, where the activity of the D22A/H110A cathepsin B mutant is compared with that of the wild-type enzyme. Essentially no effect is seen for the generic cysteine protease substrate Z-FR-Amc or for the amidated form of the quenched fluorescence substrate Abz-FRF(4NO_2)A. These compounds are neither typical exo- nor endo-peptidase substrates since, although they are short (thus mimicking exopeptidase substrates), no terminal charge is present, and in the case of Z-FR-Amc, the leaving group does not reach into the S_2' subsite [39]. Correct orientation of the loop is critical for exopeptidase action, as demonstrated by the almost complete lack of activity of the D22A/H110A mutant with the Abz-FRF(4NO_2)A exopeptidase substrate, which is an excellent wild-type cathepsin B substrate. In contrast, the extended quenched fluorescence substrate Abz-AFRSAAQ-EDDnp (where -EDDnp denotes -*N*-ethylenediamine-2,4-dinitrophenylamide) [7], an endopeptidase substrate, is poorly cleaved by wild-type cathepsin B, but is an excellent substrate in the case of the D22A/H110A mutant, where the loop can be deflected easily to allow access to the extended substrate [7]. These combined results clearly demonstrate the role of the occluding loop in tailoring cathepsin B to act as an exopeptidase at acidic pH.

Whereas, chemically, the quenched fluorescence compounds used in this work can be considered to represent 'exopeptidase substrates', our previous study [23] showed that they are well cleaved by other members of the mammalian cysteine proteases of the papain family, which are usually considered to be only endopeptidases. It was shown that the ability of unspecialized cysteine proteases to cleave such substrates was dependent on the presence of optimal residues in the P_2 and P_1 subsites (Phe and Arg respectively). When suboptimal residues were present at these positions (for example, Ala) activity was dramatically reduced with these enzymes, whereas high activities were seen in the case of the specialized exopeptidases. In the present study, cathepsin B was shown to have relatively low (20-fold) specificity in the P_2' subsite. Cathepsin B has been shown to have a similar level of specificity in the P_3 subsite [39]. As a general principle, the major specificity determinant of the papain family is an extremely high selectivity for large hydrophobic residues in P_2, for example 2700- and 12000-fold higher activity for Leu compared with Arg in the case of cathepsins S and cathepsin K respectively [40,41], and similar patterns for cathepsin L and cathepsin L2 (cathepsin V) [42]. In contrast, the S_2 subsite in cathepsin B is modified so that, in addition to accepting 'standard' hydrophobic side chains, it will also accept basic residues [43], generalizing selectivity. Thus cathepsin B's relaxed specificity at all subsites enables it to be less selective in the substrates in chooses to digest, which in turn allows for a processive removal of C-terminal dipeptides. This behaviour has been demonstrated with various protein substrates [30,44–46], showing the effectiveness of the overall design strategy.

We thank Marie-Claude Magny and Patrizia Mason for the expression and purification of wild-type cathepsin B. We also thank Boehringer Ingelheim Canada, Research and Development, for the use of their HPLC and their MALDI–TOF mass spectrometer. N-terminal sequencing and peptide synthesis were carried out by Elisa de Miguel. Edwin Wan synthesized the oligonucleotides used in the mutagenesis experiments. This work was supported by the Shriners of North America and the Protein Engineering Network of Centers of Excellence.

REFERENCES

1 Schechter, I. and Berger, A. (1967) On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. **27**, 157–162
2 Mort, J. S. and Buttle, D. J. (1997) Molecules in focus. Cathepsin B. Int. J. Biochem. Cell Biol. **29**, 715–720
3 Halangk, W., Lerch, M. M., Brandt-Nedelev, B., Roth, W., Ruthenbuerger, M., Reinheckel, T., Domschke, W., Lippert, H., Peters, C. and Deussing, J. (2000) Role of cathepsin B in intracellular trypsinogen activation and the onset of acute pancreatitis. J. Clin. Invest. **106**, 773–781
4 Mort, J. S., Recklies, A. D. and Poole, A. R. (1984) Extracellular presence of the lysosomal proteinase cathepsin B in rheumatoid synovium and its activity at neutral pH. Arthritis Rheum. **27**, 509–515
5 Sloane, B. F. (1990) Cathepsin B and cystatins: evidence for a role in cancer progression. Semin. Cancer Biol. **1**, 137–152
6 Guicciardi, M. E., Deussing, J., Miyoshi, H., Bronk, S. F., Svingen, P. A., Peters, C., Kaufmann, S. H. and Gores, G. J. (2000) Cathepsin B contributes to TNF-α-mediated hepatocyte apoptosis by promoting mitochondrial release of cytochrome *c*. J. Clin. Invest. **106**, 1127–1137
7 Nägler, D. K., Storer, A. C., Portaro, F. C. V., Carmona, E., Juliano, L. and Ménard, R. (1997) Major increase in endopeptidase activity of human cathepsin B upon removal of occluding loop contacts. Biochemistry **36**, 12608–12615
8 Barrett, A. J. and Kirschke, H. (1981) Cathepsin B, cathepsin H and cathepsin L. Methods Enzymol. **80**, 535–561
9 Musil, D., Zucic, D., Engh, R. A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., Katunuma, N. and Bode, W. (1991) The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. EMBO J. **10**, 2321–2330

10 Cygler, M., Sivaraman, J., Grochulski, P., Coulombe, R., Storer, A. C. and Mort, J. S. (1996) Structure of rat procathepsin B. Model for inhibition of cysteine protease activity by the proregion. Structure **4**, 405–416

11 Podobnik, M., Kuhelj, R., Turk, V. and Turk, D. (1997) Crystal structure of the wild-type human procathepsin B at 2.5 Å resolution reveals the native active site of a papain-like cysteine protease zymogen. J. Mol. Biol. **271**, 774–788

12 Turk, D., Podobnik, M., Popovic, T., Katunuma, N., Bode, W., Huber, R. and Turk, V. (1995) Crystal structure of cathepsin B inhibited with CA030 at 2.0 Å resolution: a basis for the design of specific epoxysuccinyl inhibitors. Biochemistry **34**, 4791–4797

13 Yamamoto, A., Hara, T., Tomoo, K., Ishida, T., Fujii, T., Hata, Y., Murata, M. and Kitamura, K. (1997) Binding mode of CA074, a specific irreversible inhibitor, to bovine cathepsin B as determined by X-ray crystal analysis of the complex. J. Biochem. (Tokyo) **121**, 974–977

14 Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T. and Mort, J. S. (1997) Role of the occluding loop in cathepsin B activity. J. Biol. Chem. **272**, 1197–1202

15 Quraishi, O., Nägler, D. K., Fox, T., Sivaraman, J., Cygler, M., Mort, J. S. and Storer, A. C. (1999) The occluding loop in cathepsin B defines the pH dependence of inhibition by its propeptide. Biochemistry **38**, 5017–5023

16 Meldal, M. and Breddam, K. (1991) Anthranilamide and nitrotyrosine as a donor–acceptor pair in internally quenched fluorescence substrates for endopeptidases: multicolumn peptide synthesis of enzyme substrates for subtilisin Carlsberg and pepsin. Anal. Biochem. **195**, 141–147

17 Kunkel, T. A., Roberts, J. D. and Zakour, R. A. (1987) Rapid and efficient site-directed mutagenesis without phenotypic selection. Methods Enzymol. **154**, 367–382

18 Mach, L., Schwihla, H., Stüwe, K., Rowan, A. D., Mort, J. S. and Glössl, J. (1993) Activation of procathepsin B in human hepatoma cells: the conversion to the mature enzyme relies on the action of cathepsin B itself. Biochem. J. **293**, 437–442

19 Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. U.S.A. **82**, 488–492

20 Rawlings, N. D. and Barrett, A. J. (1990) FLUSYS: A software package for the collection and analysis of kinetic and scanning data from PerkinElmer fluorimeters. Comput. Appl. Biosci. **6**, 118–119

21 Leatherbarrow, R. J. (1992) GraFit Version 3.0, Erithacus Software Ltd., Staines, U.K.

22 Taylor, S., Ninjoor, V., Dowd, D. M. and Tappel, A. L. (1974) Cathepsin B2 measurement by sensitive fluorometric ammonia analysis. Anal. Biochem. **60**, 153–162

23 Nägler, D. K., Tam, W., Storer, A. C., Krupa, J. C., Mort, J. S. and Ménard, R. (1999) Interdependency of sequence and positional specificities for cysteine proteases of the papain family. Biochemistry **38**, 4868–4874

24 Kirschke, H., Locnikar, P. and Turk, V. (1984) Species variations amongst lysosomal cysteine proteinases. FEBS Lett. **174**, 123–127

25 Ménard, R., Carmona, E., Plouffe, C., Brömme, D., Konishi, Y., Lefebvre, J. and Storer, A. C. (1993) The specificity of the S_1' subsite of cysteine proteases. FEBS Lett. **328**, 107–110

26 Therrien, C., Lachance, P., Sulea, T., Purisima, E. O., Qi, H., Ziomek, E., Alvarez-Hernandez, A., Roush, W. R. and Ménard, R. (2001) Cathepsins X and B can be differentiated through their respective mono- and dipeptidyl carboxypeptidase activities. Biochemistry **40**, 2702–2711

27 Wilkinson, A. J., Fersht, A. R., Blow, D. M. and Winter, G. (1983) Site-directed mutagenesis as a probe of enzyme structure and catalysis: tyrosyl-tRNA synthetase cysteine-35 to glycine-35 mutation. Biochemistry **22**, 3581–3586

28 Koga, H., Yamada, H., Nishimura, Y., Kato, K. and Imoto, T. (1991) Multiple proteolytic action of rat liver cathepsin B: Specificities and pH-dependences of the endo- and exopeptidase activities. J. Biochem. (Tokyo) **110**, 179–188

29 Takahashi, T., Dehdarani, A. H., Yonezawa, S. and Tang, J. (1986) Porcine spleen cathepsin B is an exopeptidase. J. Biol. Chem. **261**, 9375–9381

30 Rowan, A. D., Feng, R., Konishi, Y. and Mort, J. S. (1993) Demonstration by electrospray mass spectrometry that the peptidyldipeptidase activity of cathepsin B is capable of rat cathepsin B C-terminal processing. Biochem. J. **294**, 923–927

31 Berti, P. J., Faerman, C. H. and Storer, A. C. (1991) Cooperativity of papain–substrate interaction energies in the S2 to S2′ subsites. Biochemistry **30**, 1394–1402

32 Turk, B., Turk, D. and Turk, V. (2000) Lysosomal cysteine proteases: more than scavengers. Biochim. Biophys. Acta **1477**, 98–111

33 Guncar, G., Podobnik, M., Pungercar, J., Strukelj, B., Turk, V. and Turk, D. (1998) Crystal structure of porcine cathepsin H determined at 2.1 Å resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. Structure **6**, 51–61

34 Cigic, B., Krizaj, I., Kralj, B., Turk, V. and Pain, R. H. (1998) Stoichiometry and heterogeneity of the pro-region chain in tetrameric human cathepsin C. Biochim. Biophys. Acta **1382**, 143–150

35 Sivaraman, J., Nägler, D. K., Zhang, R., Ménard, R. and Cygler, M. (2000) Crystal structure of human procathepsin X: a cysteine protease with the proregion covalently linked to the active site cysteine. J. Mol. Biol. **295**, 939–951

36 Guncar, G., Klemencic, I., Turk, B., Turk, V., Karaoglanovic-Carmona, A., Juliano, L. and Turk, D. (2000) Crystal structure of cathepsin X: a flip-flop of the ring of His23 allows carboxy-monopeptidase and carboxy-dipeptidase activity of the protease. Structure Fold. Des. **8**, 305–313

37 Nycander, M., Estrada, S., Mort, J. S., Abrahamson, M. and Björk, I. (1998) Two-step mechanism of inhibition of cathepsin B by cystatin C due to displacement of the proteinase occluding loop. FEBS Lett. **422**, 61–64

38 Pavlova, A., Krupa, J. C., Mort, J. S., Abrahamson, M. and Björk, I. (2000) Cystatin inhibition of cathepsin B requires dislocation of the proteinase occluding loop. Demonstration of release of loop anchoring through mutation of His^{110}. FEBS Lett. **487**, 156–160

39 Taralp, A., Kaplan, H., Sytwu, I.-I., Vlattas, I., Bohacek, R., Knap, A. K., Hirama, T., Huber, C. P. and Hasnain, S. (1995) Characterization of the S_3 subsite specificity of cathepsin B. J. Biol. Chem. **270**, 18036–18043

40 Brömme, D., Bonneau, P. R., Lachance, P. and Storer, A. C. (1994) Engineering the S_2 subsite specificity of human cathepsin S to a cathepsin L- and cathepsin B-like specificity. J. Biol. Chem. **269**, 30238–30242

41 Brömme, D., Okamoto, K., Wang, B. B. and Biroc, S. (1996) Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. Functional expression of human cathepsin O2 in *Spodoptera frugiperda* and characterization of the enzyme. J. Biol. Chem. **271**, 2126–2132

42 Brömme, D., Li, Z., Barnes, M. and Mehler, E. (1999) Human cathepsin V functional expression, tissue distribution, electrostatic surface potential, enzymatic characterization, and chromosomal localization. Biochemistry **38**, 2377–2385

43 Hasnain, S., Hirama, T., Huber, C. P., Mason, P. and Mort, J. S. (1993) Characterization of cathepsin B specificity by site-directed mutagenesis. The importance of Glu^{245} in the S_2–P_2 specificity for arginine and its role in transition state stabilization. J. Biol. Chem. **268**, 235–240

44 Aronson, N. N. and Barrett, A. J. (1978) The specificity of cathepsin B. Hydrolysis of glucagon at the C-terminus by a peptidyldipeptidase mechanism. Biochem. J. **171**, 759–765

45 Mort, J. S., Magny, M.-C. and Lee, E. R. (1998) Cathepsin B: an alternative protease for the generation of an aggrecan 'metalloproteinase' cleavage neoepitope. Biochem. J. **335**, 491–494

46 Authier, F., Métioui, M., Bell, A. W. and Mort, J. S. (1999) Negative regulation of epidermal growth factor signaling by selective proteolytic mechanisms in the endosome mediated by cathepsin B. J. Biol. Chem. **274**, 33723–33731

Received 18 May 2001/17 October 2001; accepted 9 November 2001