Activation of human prolegumain by cleavage at a C-terminal asparagine residue

Jinq-May CHEN¹, Mara FORTUNATO and Alan J. BARRETT

MRC Molecular Enzymology Laboratory, The Babraham Institute, Babraham, Cambs. CB2 4AT, U.K.

The processing and activation of prolegumain were studied using the recombinant protein synthesized by cells that had been stably transfected with a human legumain cDNA construct. A cell line termed C13 was selected for the high-level expression of prolegumain. C13 cells produced primarily 56 kDa prolegumain. The 56 kDa form was enzymically inactive but stable at neutral pH, unlike the 35 kDa mature pig legumain; it could be converted into a 46 kDa active form by incubation at pH 4.5. The 56 kDa pro-form and the 46 kDa active form were found to have the same N-terminal amino acid sequence, indicating that cleavage at the N-terminus was not necessary for prolegumain activation,

and that the decrease in molecular mass was due to a C-terminal cleavage. The C-terminal processing site was identified as Asn³²³. Replacement of Asn³²³ at the cleavage site with aspartate, serine, alanine or glutamate abolished the processing and activation of prolegumain. In contrast, mutation of other asparagine and aspartate residues near the cleavage site had no effect. These results demonstrate that Asn³²³ is essential for prolegumain activation.

Key words: asparaginyl endopeptidase, legumain, lysosomal endopeptidase, enzyme processing.

INTRODUCTION

Mammalian legumain (EC 3.4.22.34) is a recently discovered lysosomal protease that belongs to a distinct family of cysteine endopeptidases, C13. {Identifiers for peptidase families and clans are as defined in the MEROPS database, Release 5.0 (www.merops.co.uk/merops/merops.htm) [1].} Legumain requires acidic conditions for activity and has strict specificity for cleavage with an asparagine residue in the P1 site. Legumain activity can be inhibited by human cystatin C and ovocystatin, but not by compound E-64 [*trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane], a potent inhibitor of many cysteine peptidases of the papain family, including most lysosomal cathepsins [2–4]. In plants, there is abundant evidence that legumain functions as a processing enzyme that causes limited proteolysis of precursor proteins [5]. Mammalian legumain has also been shown to be involved in processing of the microbial tetanus toxin C fragment for MHC class II antigen presentation in the lysosomal/endosomal system [6,7]. Recently, legumain was implicated as an inhibitor of osteoclast formation and bone resorption [8].

It has been established that the genes for legumain in *Schistosoma mansoni* [9] and humans [10] encode inactive prolegumain that can be activated by incubation at acid pH. Prolegumain expressed in *Saccharomyces cerevisiae* from a plant legumain gene has also been demonstrated to be processed and activated in yeast vacuoles, known to be acidic compartments [11]. The activation appears to be autocatalytic, but for both the human [10] and plant [12] systems the authors have emphasized the importance of cleavage at aspartyl bonds, rather than the asparaginyl bonds that are far more readily cleaved by legumain.

On the basis of a common catalytic-site motif, legumain (C13) and four other families of endopeptidases, containing clostripain

(C11), gingipain (C25), caspases (C14) and separin (C50), form a distinct group of cysteine peptidases termed clan CD [13,14]. The substrate specificities of the peptidases in clan CD are dominated by the interactions of the S1 subsites. Many of the enzymes in clan CD are known to be synthesized as proenzymes, but for none of them is the mechanism of activation well understood. It is hoped that the present investigation of the route of activation of human prolegumain will contribute to the eventual understanding of the regulation of activity of this whole group of enzymes.

MATERIALS AND METHODS

Materials

Z-Ala-Ala-Asn-NHMec [where Z is benzyloxycarbonyl and NHMec is 7-(4-methyl) coumarylamide] was purchased from Dr C. Graham Knight (Department of Biochemistry, University of Cambridge, Cambridge, U.K.). Human embryonic kidney (HEK) 293 cells (cat. no. CRL 1573) were from American Type Culture Collection (Rockville, MD, U.S.A.). All reagents for tissue culture were purchased from either Sigma or Life Technologies (Paisley, Scotland, U.K.). The synthetic peptides acetyl (Ac)-DDPEDGG-amide (N-peptide), Ac-DLEESRQLTEEIQRHLD-amide (C-peptide), 9-fluorenylmethoxycarbonyl (Fmoc)-Ala-Glu-Gln-Lys-amide (AEQK) and Fmoc-Ala-Glu-Asn-Lys-amide (AENK) were purchased from MWG-Biotech (Milton Keynes, U.K.). All oligonucleotides were synthesized by Sigma-Genosys. Brefeldin A, NH₄Cl and protease inhibitors (E-64, leupeptin, PMSF, pepstatin, 1,10-phenanthroline and benzamidine) were purchased from Sigma. Endoglycosidase F (PNGase) was from New England Biolabs. Pig legumain and ovocystatin were as described previously [2].

Abbreviations used: E-64, *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane; NHMec, 7-(4-methyl) coumarylamide; Z, benzyloxycarbonyl; Ac, acetyl; Fmoc, 9-fluorenylmethoxycarbonyl; HEK, human embryonic kidney; FBS, fetal bovine serum; MEM, minimal Eagle's medium.

¹ To whom correspondence should be addressed (e-mail jinq-may.chen@bbsrc.ac.uk).

Table 1 Mutagenic primers

Codons for the mutated residues are underlined, and the emboldened characters represent the base changes introduced.

Mutant	Primer
N321Q	AGGAACTGATG C AGACCAATGATCTGGAGG
N323Q	ACTGATGAACACCC A AGATCTGGAGGAGTCC
N323S	ACTGATGAACACCA G TGATCTGGAGGAGTCC
N323D	ACTGATGAACACCC G ATGATCTGGAGGAGTCC
N323A/D324E	ACTGATGAACACCC GCTG AACTGGAGGAGTCCA
D324E	GATGAACACCAAT G AACTGGAGGAGTCCAGG
D340E	AGCGGCATCTGG A AGCCAGGCACCTCAT

Plasmids for the expression of wild-type and mutant legumain in mammalian cells

The full-length cDNA encoding human prolegumain was amplified using the forward primer 5'-GTGTCTGCAATTGATAT-CCAAGGTGCAG-3' and the reverse primer 5'-GTCTCTGATCAGCACACAGTCGG-3', and a human placenta Marathon-Ready™ cDNA (Clontech) as the template. The PCR product was subcloned into the mammalian expression vector pcDNA3.1(-) (Invitrogen) through 5' *EcoRV* and 3' *HindIII* sites. The construct (hsleg/pcDNA) was propagated in *Escherichia coli* DH5 α (Life Technologies) and the plasmid DNA was purified with a Qiagen Maxiprep kit. The insert was sequenced and found to be identical with the published human legumain sequence (accession number Y09862 [2]). Mutagenesis was performed by the PCR site-directed mutagenesis method originally described by Picard et al. [15], with modifications. Mutagenic primers based on the nucleotide sequence of human legumain were designed to generate seven constructs with putative processing sites mutated (Table 1). Polymerase *pfu* (Stratagene) was used for all PCR reactions, with a typical run of 94 °C for 3 min, followed by 30 cycles of 94 °C for 1 min, 48 °C for 1 min and 72 °C for 2 min. Mutated codons were introduced by performing PCR with a mutagenic primer and the reverse primer of pcDNA3.1 (5'-TAGAAGGCACAGTCGAGGC-3'), and hsleg/pcDNA as the template to generate a megaprimer. The megaprimer and a legumain-specific forward primer (5'-CTCCAGGAATTGTGATCAACAGGC-3') were used for the second run of PCR with the same template to generate a 1 kb product containing the mutated nucleotides. The 1 kb legumain fragment with mutation was double-digested with *Bam*HI and *Hind*III, and then ligated into hsleg/pcDNA plasmid that had also been digested with the same restriction enzymes to remove the wild-type legumain fragment. Plasmid DNA of mutants or wild type was propagated in DH5 α or XL1-Blue (Stratagene) cells and purified by standard procedures [16]. The sequences of all mutated constructs were confirmed by sequencing.

Production of recombinant legumain in mammalian cells

HEK293 cells maintained in minimal Eagle's medium (MEM) supplemented with 10% (v/v) fetal bovine serum (FBS) were transfected with the legumain construct by use of either a Calcium Phosphate Transfection Kit (Invitrogen) or DNA-liposome fusion with LIPOFECTAMINE™ reagent (Life Technologies).

To prepare stable transfectants, HEK293 cells transfected by incubation with the hsleg/pcDNA plasmid (20 μ g) in calcium phosphate precipitate for 24 h were subjected to a 30 s shock with glycerol/PBS (pH 7.4) solution. Cells were washed with

PBS and then fed with fresh complete growth medium and, 24 h later, were passaged with 1:6 dilution and grown under selection with 1 mg/ml antibiotic G418 (Geneticin; Life Technologies) for 2 weeks. Cells that were resistant to G418 were passaged into two 24-well plates and cultured under selection until confluent. Further passages were done to plate cells from the two 24-well plates to four 12-well and finally to eight 6-well plates. When these 48 cultures reached confluency, one-third of the cells from each well were passaged to fresh plates, and two-thirds were collected to prepare cell lysates. Positive clones were screened and selected on the basis of the protein expression and enzymic activity of recombinant legumain.

Transient expression of the wild-type and mutant legumain constructs was performed by transfection of plasmid DNA-liposome complexes into HEK293 cells. Briefly, cells were plated in 60 mm culture dishes and reached about 80% confluency the next day. Plasmid DNA (6 μ g) in 300 μ l of OPTI-MEM medium was first mixed with 25 μ l of LIPOFECTAMINE™, also diluted in 300 μ l of the same medium, and the mixture was incubated for 30 min to allow liposome-DNA complexes to form. The solution containing liposome-DNA complexes was brought up to 3 ml with OPTI-MEM medium and applied to cells that were rinsed with the same medium immediately before transfection. After incubation with the complexes for 5 h, the cells were changed into normal growth medium (MEM; 3 ml) containing 20% (v/v) FBS. This medium was replaced with fresh complete medium 24 h after transfection, and cells were harvested on the second day.

Cell culture and preparation of conditioned medium and cell lysate

C13 cells were grown in MEM supplemented with 10% (v/v) FBS and 0.5 mg/ml G418. In some instances, cultures were treated with 20 mM NH₄Cl or 20 μ M brefeldin A for 16 h before harvesting. To collect conditioned medium, confluent cultures were washed with PBS and incubated in serum-free MEM for 16 h. The serum-free conditioned medium was collected and concentrated 20-fold with a Centricon 30 (Amicon). The concentrated medium was dialysed into PBS containing 1 mM EDTA and 0.01% CHAPS, pH 7.2. Further separation of molecules greater than 30 kDa in the concentrated conditioned medium was achieved by gel filtration on a Superdex 75 column in the same buffer. Fractions containing prolegumain detectable by immunoblot were collected.

Lysates were prepared from confluent cultures of C13 cells, HEK293 cells transiently transfected with legumain constructs (wild type and mutated) or mock-transfected. Cells were disrupted by three cycles of freeze/thaw in a lysis buffer. Depending upon the experiment, one of two lysis buffers was used: 40 mM citric acid and 121 mM Na₂HPO₄ (pH 5.8), or 8 mM Na₂HPO₄, 1.47 mM KH₂PO₄, 136.8 mM NaCl and 2.68 mM KCl (pH 7.2); both included 1 mM EDTA and 1% n-octyl β -D-glucopyranoside. Cell lysates were collected after removal of material pelleted during centrifugation at 18000 g for 10 min.

Immunological and enzymic detection of recombinant legumain

C13 cells were fixed in 3% (w/v) paraformaldehyde in PBS and permeabilized in 0.4% Triton X-100 in PBS. They were then incubated with an affinity-purified anti-(pig legumain) antibody raised in a sheep [3], followed by FITC-conjugated anti-(sheep immunoglobulin) secondary antibody (Sigma). Cells were examined by the use of a Zeiss Axiophot microscope equipped with an epifluorescence light source and a filter set for fluorescein, and photographed with a Zeiss camera system.

Legumain in conditioned medium or cell lysates was detected by SDS/PAGE immunoblot. Proteins in the SDS/polyacrylamide gel were transferred to nitrocellulose or PVDF membrane according to the method of Towbin et al. [17]. Immunoreaction and colour development were as previously described [3].

To measure the enzymic activity of recombinant legumain, cell lysate (20 μ l) was placed in the wells of a 96-well microtitre plate, and the reaction was started by addition of 150 μ l of Z-Ala-Ala-Asn-NHMe solution in assay buffer (40 mM citric acid, 121 mM Na_2HPO_4 , pH 5.8, containing 1 mM dithiothreitol, 1 mM EDTA and 0.1% CHAPS) to give a final concentration of 10 μ M substrate. The plates were incubated at 25 °C and readings of fluorescence (excitation 360 nm; emission 460 nm) were taken for five cycles, with an interval of 7 min between cycles, in a PerSeptive Biosystems CytoFluor II fluorescence plate reader. The maximum rate, determined as the average of changes between cycles, was the measure of activity. Results were corrected by subtracting the value obtained with the extraction buffer control.

N-terminal microsequencing

Samples containing various forms of legumain and fragments were separated in 10% (w/v) NuPAGE Bis-Tris gels (Novex) in Mops/SDS running buffer (50 mM Mops, 50 mM Tris base, 3.5 mM SDS and 1 mM EDTA, pH 7.7). The separated proteins were transblotted on to PVDF membrane in 10 mM Caps buffer, pH 11, containing 10% (v/v) methanol, and stained with Coomassie Blue R solution. Bands of interest were excised and subjected to N-terminal sequence analysis. N-terminal sequencing was carried out using an Applied Biosystems Procise Model 492 Sequencer at the Microchemical Facility of the Babraham Institute.

RESULTS

Production of recombinant human prolegumain

Full-length cDNA for human legumain was stably transfected into cells of the HEK293 cell line that we had used previously for transient expression of mouse legumain [3]. Several clones resistant to G418 were screened and their expression of recombinant legumain was examined by immunoblot analysis and activity assay. One of the clones, designated C13, was selected on the basis of the high level of expression of prolegumain. An immunofluorescence image of the C13 cells labelled with a specific polyclonal antibody against pig legumain (Figure 1A) shows diffuse cytoplasmic staining, but also punctate staining consistent with the presence of legumain in lysosomes, as found previously for rat kidney cells [3]. Figure 1(B) demonstrates the species of legumain from C13 cells by immunoblot. Three bands at 56, 47 and 36 kDa were recognized by the anti-legumain antibody in a C13 lysate prepared by extracting cells with pH 5.8 lysis buffer (lane 1). Brefeldin A is known to block transit of proteins by disrupting the Golgi apparatus [18]. Following treatment of C13 cells with brefeldin A (20 μ M), the 56 kDa form was the predominant species in the lysate (lane 2). When C13 cultures were grown in the presence of 20 mM NH_4Cl for 16 h, a reagent known to increase the pH of the lysosomal system [19,20], the 56 kDa form could be found in the cell lysate (lane 3) and the conditioned medium (lane 4). A very small amount of the 47 kDa form was also detected in the lysate from either brefeldin A- or NH_4Cl -treated cultures (lanes 2 and 3). When pH 7.2 lysis buffer was used to extract C13 cells, only the 56 kDa form was seen in the lysate from cultures with or without NH_4Cl treatment (not shown here; see lanes 1 of Figures 3, 7 and 8). No enzymic activity was detected in lysates containing the 56 kDa form (see

Figure 1 Legumain expression in C13 cells

Recombinant legumain expressed in C13 cells, a cell line stably transfected with human legumain construct in pcDNA3.1 mammalian expression vector, was detected by (A) immunostaining of cells with an anti-legumain antibody followed by FITC-conjugated secondary antibodies, and (B) immunoblotting of the cell extract. In (B), the cell lysate (cl) from cultures without NH_4Cl treatment contained a 56 kDa form, a 47 kDa intermediate form and the 36 kDa mature enzyme (lane 1). The predominant 56 kDa legumain and a minor 47 kDa form were present in the cell extract of brefeldin A (BFA)-treated cultures (lane 2). The cell lysate (cl; lane 3) and conditioned medium (cm; lane 4) of C13 cells treated with NH_4Cl contained primarily a 56 kDa form. The 35 kDa legumain purified from pig kidney (pleg) as a reference is shown in lane 5.

Figure 2 pH dependence of prolegumain activation

Prolegumain in the conditioned medium (◆) or in the cell lysate (▲) of C13 cells was incubated at 30 °C in phosphate/citrate buffers in the pH range 2.5–7.5, and the activity was plotted between pH 2.5 and pH 6. Enzymic activity was determined as described in the Materials and methods section.

Figure 3 Processing of the 56 kDa legumain at pH 5.8 and pH 4.5

C13 cell lysate containing the 56 kDa legumain in pH 7.2 buffer (lane 1) was adjusted to pH 5.8 (lane 2) and 4.5 (lane 3) by the addition of 0.1 M citric acid containing 1 mM EDTA and 0.01 % CHAPS, and incubated at 30 °C for 3 h. After incubation, one-half of each sample was subjected to SDS/PAGE and immunoblotting to detect different forms of legumain (**A**), and the other half was used in the fluorimetric assay to detect legumain activity (**B**).

Figures 3B, 7B and 8B). The data suggest that the 56 kDa immunoreactive protein is the pro-form of legumain, which remains stable at pH 7.2.

Activation of prolegumain

The conditioned medium and cell extracts from NH_4Cl -treated C13 cultures containing 56 kDa prolegumain (Figure 1B, lanes 3 and 4), and showing no legumain activity, were incubated at 30 °C for 4 h in phosphate/citrate buffers in the range of pH 2.5–7.5 before assay of legumain activity at pH 5.8. As shown in Figure 2, active legumain was formed, the optimal pH for activation being approx. 4.5.

In the immunoblot (Figure 3A), it could be seen that, during incubation of the lysate prepared with the pH 7.2 buffer (lane 1) for 3 h at pH 5.8, the 56 kDa form was converted primarily into the 47 kDa form (lane 2), whereas pH 4.5 gave the 46 kDa form (lane 3). Appearance of the 46 kDa form in the cell lysate coincided with an increased amount of legumain activity, whereas only a low level of legumain activity was detected in the sample containing both the 56 and 47 kDa forms (Figure 3B). More 47 kDa form was generated in the sample containing both the 56 and 47 kDa forms after a 4 h incubation in the same pH 5.8 buffer, but again not much activity was detected (Figure 4A, lane 1; Figure 4B). Incubation at pH 4.5 for 4 h resulted in the conversion of both the 56 and 47 kDa forms into the 46 kDa

Figure 4 Generation of the 46 kDa active legumain from the 56 kDa and 47 kDa forms

C13 cell extract containing both the 56 kDa and 47 kDa forms of legumain at pH 5.8 was incubated at pH 5.8 and at pH 4.5 at 30 °C. The immunoblot in (**A**) shows the presence of 56 kDa and 47 kDa forms at pH 5.8 after 4 h of incubation (lane 1), and at pH 4.5 immediately before the incubation (lane 2). Lane 3 shows that both the 56 kDa and 47 kDa forms were converted into the 46 kDa form after 4 h of incubation at pH 4.5. The legumain activity of the three samples is shown in (**B**).

form, and in a 15-fold increase in enzymic activity (Figure 4A, lane 3; Figure 4B). These data demonstrate that the 56 kDa prolegumain could be converted into the 47 kDa intermediate form at pH 5.8 and into the 46 kDa active form at pH 4.5, and that the 46 kDa active form could also be generated from the 47 kDa intermediate form at pH 4.5, presumably by limited proteolysis.

Legumain purified from pig kidney was a 35 kDa glycoprotein [2], and a 36 kDa form was also detected in the C13 cell lysate prepared in the pH 5.8 buffer (Figure 1A, lane 1). In contrast, the active legumain derived from the recombinant 56 kDa prolegumain was a protein of 46 kDa. We asked whether the difference in apparent molecular mass between the two active forms of legumain was due to the degree of glycosylation. Treatment of the proteins with endoglycosidase F [2] caused a decrease in molecular mass of approx. 7 kDa for both (results not shown). This suggested that both proteins have the same carbohydrate component, and that the change in molecular mass is due to loss of polypeptide as a result of proteolysis. This possibility was examined further by prolonged (24 h) incubation of C13 cell lysate containing the 46 kDa active legumain in the

Figure 5 Sites of post-translational processing in prolegumain

Schematic representation of the 56 kDa human prolegumain showing the cleavage site of the signal peptide (position A) and the C-terminal processing site (position C) that produces active legumain. The inactive 56 kDa and 47 kDa forms had the same N-terminal residue, Val¹⁸, as the 46 kDa active legumain. The N-terminus of the 35 kDa legumain purified from pig kidney was identified as Gly²⁶ [2], showing cleavage at position B. His¹⁴⁸ and Cys¹⁸⁹ are the catalytic residues [13]. Two synthetic peptides, denoted N-peptide and C-peptide and used for the inhibitory study, were derived from the amino acid sequences as shown.

pH 4.5 buffer. There was a progressive conversion of the 46 kDa form into the 36 kDa form, with no large change in activity. The conversion could be inhibited by E-64 (10 μ M) and leupeptin (20 μ M). The 46 kDa form was not isolated in a pure state that would allow determination of specific activity, but it was determined that the K_i values for inhibition of both the activated 46 kDa recombinant legumain and the 35 kDa pig legumain by ovocystatin were similar (6 nM).

Prolegumain processing site

Sequencing of the N-termini of the 56, 47 and 46 kDa legumain forms showed the same sequence from all three, i.e. Val-Pro-Ile-Asp-Asp-Pro-Glu, with the first residue, Val¹⁸, immediately C-terminal to the predicted signal peptide (Figure 5; [2]). Since the 46 kDa form was catalytically active, it is evident that N-terminal processing was not required for prolegumain activation, and that the conversion of the 56 and 47 kDa forms into the active 46 kDa legumain resulted from C-terminal processing.

To identify the site of processing at the C-terminus, we incubated the 56 kDa prolegumain at pH 4.5 for 4 h and analysed the products by SDS/PAGE and activity assays. The 46 kDa active legumain and an approx. 11 kDa fragment were generated. N-terminal sequencing of the 11 kDa band showed the first seven residues of this fragment to be Asp-Leu-Glu-Glu-Ser-Arg-Gln. This sequence exactly matches residues Asp³²⁴–Gln³³⁰ in human prolegumain. The data indicate that, at pH 4.5, the 46 kDa active legumain was generated directly from the 56 kDa pro-form by a single cleavage at the Asn³²³–Asp³²⁴ peptide bond (Figure 5).

Legumain has a very strict specificity for cleavage of asparaginyl bonds at pH 5.8 [2,4], although it has also been shown to cleave aspartyl bonds slowly, especially under acidic conditions [10,12]. There are Asn and Asp residues close to the identified processing site. In order to test whether Asn³²³ is crucial for activation of prolegumain, Asn and Asp residues at and near the site were mutated. Seven mutant constructs, N321Q, N323Q, N323D, N323S, N323A/D324E, D324E and D340E, were transiently expressed in HEK293 cells. Cells were treated with 20 mM NH₄Cl for 16 h (a condition previously described for obtaining mostly prolegumain) before harvesting, and extracted with pH 7.2 buffer. Cell lysates from mutants and the wild type containing 56 kDa prolegumain were subjected to activation by incubation at pH 4.5. Prolegumain from mutants N323Q, N323D, N323S and N323A/D324E was found not to be processed (Figure 6A) and only a low level of activity was detected from these mutants on the basis of fluorimetric assays (Figure 6B). Under the same conditions, the recombinant prolegumain from the wild type and mutants of other putative

Figure 6 Asn³²³ is essential for the processing and activation of prolegumain

The wild-type and mutant forms of legumain were transiently expressed in HEK293 cells, as described in the Materials and methods section. Cells were extracted with pH 7.2 lysis buffer and incubated at pH 7.2 or pH 4.5 for 3 h. After incubation, samples from each preparation were used for immunoblotting (A) to analyse prolegumain expression and processing. Samples from each preparation were also assayed in triplicate for legumain activity (B). It can be seen that 56 kDa prolegumain was expressed from the wild type and all mutants, and it remained stable at pH 7.2 (odd-numbered lanes). Processing of the 56 kDa form to the 46 kDa form at pH 4.5 was detected only with mutants N321Q (lane 2), D324E (lane 12) and D340E (lane 14), and the wild type (WT; lane 16). Only these four samples acquired enzymic activity after incubation at pH 4.5, as shown in (B). In the immunoblot (A), the band between 56 kDa and 46 kDa (block arrow; lanes 1–18) was an artifact, seen also in the vector control (VEC; lanes 17 and 18). In lanes 7 and 8, prolegumain expressed from mutant N323S was larger than 56 kDa. It is likely that additional glycosylation occurred in this mutant, as a result of the creation of a new potential glycosylation site (Asn-Thr-Ser) at Asn³²¹ when Asn³²³ was mutated to Ser³²³. The data shown are representative of the results from four transfection experiments.

processing sites (N321Q, D324E and D340E) were processed and activated (Figure 6). The results clearly demonstrate that Asn³²³ is the obligatory cleavage site for processing of prolegumain to the active form.

Autocatalytic activation

The finding that the C-terminal residue of mature legumain is asparagine suggested that the activation of prolegumain is an autoprocessing event. We tested this hypothesis by the use of proteinase inhibitors. A mixture of inhibitors of aspartic proteinases (pepstatin), metalloproteinases (1,10-phenanthroline), serine proteinases (PMSF and benzamidin) and cysteine proteinases (E-64 and leupeptin) did not prevent the conversion of prolegumain when the 56 kDa form (Figure 7, lane 4) was incubated at either pH 5.8 (lane 5) or pH 4.5 (lane 6). Only when ovocystatin was added to the inhibitor mixture was the conversion inhibited (lanes 8 and 9). Legumain is the only peptidase known so far to be inhibited by cystatin but not by E-64 [21]. We conclude that ovocystatin blocked the processing of prolegumain into the 47 kDa intermediate form as well as into the 46 kDa active enzyme.

We also examined the effect of a competing tetrapeptide substrate on prolegumain activation (Figure 8). When Fmoc-

Figure 7 Ovocystatin blocks the conversion of the 56 kDa prolegumain into the intermediate and active forms

The lysate of C13 cells was incubated at pH 7.2, 5.8 and 4.5 at 30 °C for 4 h, without inhibitors (lanes 1–3), with a mixture of inhibitors (Mix; lanes 4–6), or with the same mixture of inhibitors plus 1 μ M ovocystatin (Mix + Cy; lanes 7–9). The inhibitor mixture contained E-64 (10 μ M), leupeptin (50 μ M), PMSF (1 mM), pepstatin (1 μ M), 1,10-phenanthroline (1 mM) and benzamide (5 mM). Samples from each incubation mixture were subjected to (A) immunoblot analysis, and (B) assay of legumain activity (in triplicate).

Ala-Glu-Asn-Lys-amide (AENK), a specific peptide substrate of legumain shown previously to block the processing of the tetanus toxin C fragment by legumain [6], was included in the pH 4.5 incubation buffer, the conversion of prolegumain into the active form was almost completely prevented (Figure 8, lane 3). The same concentration of the glutamine analogue Fmoc-Ala-Glu-Gln-Lys-amide (AEQK; lane 4) had no effect on the processing and activation of prolegumain under the same conditions.

Slight inhibition of legumain activity by C-peptide

There is ample evidence that the propeptides of proteinases often act as inhibitors of the mature active enzymes. Examples are the N-terminal propeptides of cathepsins B [22,23], L [24,25] and S [26], *Pseudomonas aeruginosa* elastase [27], and a metallo-proteinase from *Brevibacillus brevis* [28]. The data described above show that activation of prolegumain depends upon cleavage of a C-terminal peptide, and especially that the inactive 47 kDa intermediate form is converted into the active 46 kDa form by removal of a C-terminal peptide, since both forms share the same N-terminus (Figures 3–5). To evaluate the possibility that the residues immediately C-terminal to Asn³²³ of prolegumain may block the catalytic site, C-peptide consisting of 17 amino acids (Ac-DLEESRQLTEEIQRHLD-amide; Figure 5) was synthesized and tested for its inhibitory activity against the active mature pig legumain at pH 5.8. An N-peptide composed of seven amino acids from Asp²¹ to Gly²⁷ (Ac-DDPEDGG-amide; Figure 5) was also used. N-peptide was tested because the first residue of mature pig legumain from kidney was Gly²⁶ [2], despite the fact that no N-terminal processing was observed during the activation of recombinant prolegumain. C-peptide

Figure 8 Competing peptide substrate of legumain inhibits autoactivation of prolegumain

C13 cell extract was incubated for 4 h in the pH 7.2 buffer (lane 1), and also in the pH 4.5 buffer alone (lane 2), with a specific peptide substrate of legumain (AENK; lane 3), or with a control tetrapeptide (AEQK; lane 4). The incubation products were analysed in (A) immunoblots and (B) activity assays (in triplicate). The processing and activation of the 56 kDa prolegumain to the 46 kDa active enzyme at pH 4.5 were inhibited by 1 mM AENK, but not by the analogue AEQK.

inhibited approx. 75 % of pig legumain activity at 1 mM concentration ($K_i \cong 630 \mu$ M), whereas N-peptide inhibited only about 10 % of the activity at the same concentration.

DISCUSSION

Most lysosomal cysteine peptidases are synthesized as zymogens and can be autoactivated under acidic conditions, such as those found within the lysosomal compartment. Recombinant legumain overexpressed in C13 cells consisted of the 56 kDa pro-form, a 47 kDa intermediate form and a 36 kDa active form when cells were grown under normal conditions. The 36 kDa form was similar in size to the 35 kDa legumain purified from pig kidney [2], and the 36 kDa legumain from mouse tissues and recombinant mouse legumain expressed in HEK293 cells [3]. If the culture was exposed to NH_4Cl , which would be expected to induce a prolonged rise in pH in the lysosomes, only the 56 kDa pro-form was detected. Furthermore, treatment of the C13 cell culture with brefeldin A, which perturbs membrane trafficking, resulted in the presence of only 56 kDa prolegumain (Figure 1). These findings suggest that *in vivo*, once prolegumain is synthesized, it is routed via the Golgi to the lysosomal system for processing.

Reduced pH is crucial for the activation of the 56 kDa prolegumain. As shown in Figure 3, an inactive 47 kDa form of legumain was generated when prolegumain was incubated at pH 5.8, whereas incubation at pH 4.5 produced the 46 kDa

active legumain. The optimal pH for prolegumain activation was found to be 4.5 (Figure 2); this is in the pH range of lysosomal compartments measured from macrophages, kidney cells and some fibroblasts [20,29–31].

Since the 47 kDa and 46 kDa forms share the same N-terminus, the difference in molecular mass is attributable to a C-terminal extension of 10–20 amino acids beyond Asn³²³ in the 47 kDa form. The extension may cause the inactivity of the 47 kDa form by blocking the active-site cleft, as N-terminal propeptides have been shown to do in other endopeptidases ([25]; see [32] for a review). The synthetic 17-mer peptide mimicking region Asp³²⁴–Asp³⁴⁰ was weakly inhibitory, and this inhibition was not due to substrate competition, since C-peptide contains neither asparagine residues nor cleavable aspartyl bonds. Nevertheless, no clear conclusion can be reached from this. It is conceivable that the short synthetic peptide could not adopt the necessary conformation to exert inhibitory activity.

Although the first residue of active legumain purified from kidney tissues was identified as Gly²⁶ [2], the first amino acid of the 46 kDa active legumain generated by incubating the 56 kDa prolegumain at pH 4.5 was Val¹⁸ (Figure 5), immediately C-terminal to the putative signal peptide of human legumain. Conversion of the 56 kDa form into active 46 kDa legumain was due to C-terminal processing, removing 110 amino acids. The data indicate that the removal of the N-terminal peptide from prolegumain is not required for legumain activity. This activation profile may represent one of the features of members of clan CD, which also includes the caspase family. In caspases, activation of the pro-enzymes is achieved primarily by internal cleavages to produce two subunits with the catalytic dyad in the N-terminal subunit, and removal of the N-terminal peptide is not required for activation [33]. In contrast, conversion and activation of the proenzymes of the papain family (in clan CA), including cathepsins B, K and L, involve limited proteolysis leading to the removal of the N-terminal peptide [34–36].

Conversion of prolegumain into the 46 kDa active form by C-terminal cleavage was attributable to the action of legumain itself, as it could be inhibited by a legumain inhibitor, ovocystatin, and a competing legumain substrate, AENK, but not the analogue AEQK (Figures 7 and 8). This is consistent with the recent report of Caffrey et al. [9] that a form of recombinant prolegumain from *S. mansoni*, in which asparagine replaced the catalytic cysteine, could not self-process to yield the active form.

Our finding that the activation of prolegumain occurs by cleavage of an asparaginyl bond contrasts with two previous reports implicating cleavage of aspartyl bonds. A human prolegumain–IgG fusion protein was shown to be activated at pH 3, and the sequences of fragments generated during the activation were interpreted as showing cleavage at Asp³²¹ and Asp³⁴⁰ [10]. Hiraiwa et al. [12] also provided evidence to suggest that the C-terminal processing of plant legumain occurred at the second aspartyl bond of a di-aspartate site. Our data from site-directed mutagenesis (Figure 6) clearly show that, at the optimal pH (4.5) for human prolegumain activation, the crucial residue is Asn³²³. Processing and activation of prolegumain were not affected by replacement of Asn³²¹, Asp³²⁴ or Asp³⁴⁰ with other amino acids. Only when Asn³²³ was mutated was the processing prevented. Moreover, with prolegumain from the N324D mutant, in which a di-aspartate site was created by the substitution of Asn³²³ with Asp, neither conversion nor activation of the mutant prolegumain occurred.

The autolytic activation of prolegumain gives rise to a 46 kDa active form, but a 35/36 kDa form of legumain is commonly present in cells and tissues [2,3]. The change from 46 kDa to

36 kDa can be mediated at pH 4.5 by one or more peptidases that are inhibited by E-64 and leupeptin. This further processing is therefore likely to be due to the action of other lysosomal peptidases in the acidic conditions of the lysosomal compartment, and may be a step towards the ultimate degradation of the enzyme.

REFERENCES

- Rawlings, N. D. and Barrett, A. J. (2000) MEROPS: the peptidase database. *Nucleic Acids Res.* **28**, 323–325
- Chen, J.-M., Dando, P. M., Rawlings, N. D., Brown, M. A., Young, N. E., Stevens, R. A., Hewitt, E., Watts, C. and Barrett, A. J. (1997) Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. *J. Biol. Chem.* **272**, 8090–8098
- Chen, J.-M., Dando, P. M., Stevens, R. A. E., Fortunato, M. and Barrett, A. J. (1998) Cloning and expression of mouse legumain, a lysosomal endopeptidase. *Biochem. J.* **335**, 111–117
- Dando, P. M., Fortunato, M., Smith, L., Knight, C. G., McKendrick, J. E. and Barrett, A. J. (1999) Pig kidney legumain: an asparaginyl endopeptidase with restricted specificity. *Biochem. J.* **339**, 743–749
- Hara-Nishimura, I., Kinoshita, T., Hiraiwa, N. and Nishimura, M. (1998) Vacuolar processing enzymes in protein-storage vacuoles and lytic vacuoles. *J. Plant Physiol.* **152**, 668–674
- Manoury, B., Hewitt, E. W., Morrice, N., Dando, P. M., Barrett, A. J. and Watts, C. (1998) An asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation. *Nature (London)* **396**, 695–699
- Antoniou, A. N., Blackwood, S. L., Mazzeo, D. and Watts, C. (2000) Control of antigen presentation by a single protease cleavage site. *Immunity* **12**, 391–398
- Choi, S. J., Reddy, S. V., Devlin, R. D., Menaa, C., Chung, H., Boyce, B. F. and Roodman, G. D. (1999) Identification of human asparaginyl endopeptidase (legumain) as an inhibitor of osteoclast formation and bone resorption. *J. Biol. Chem.* **274**, 27747–27753
- Caffrey, C. R., Mathieu, M. A., Gaffney, A. M., Salter, J. P., Sajid, M., Lucas, K. D., Franklin, C., Bogoy, M. and McKerrow, J. H. (2000) Identification of a cDNA encoding an active asparaginyl endopeptidase of *Schistosoma mansoni* and its expression in *Pichia pastoris*. *FEBS Lett.* **466**, 244–248
- Halfon, S., Patel, S., Vega, F., Zurawski, S. and Zurawski, G. (1998) Autocatalytic activation of human legumain at aspartic acid residues. *FEBS Lett.* **438**, 114–118
- Hiraiwa, N., Nishimura, M. and Hara-Nishimura, I. (1997) Expression and activation of the vacuolar processing enzyme in *Saccharomyces cerevisiae*. *Plant J.* **12**, 819–829
- Hiraiwa, N., Nishimura, M. and Hara-Nishimura, I. (1999) Vacuolar processing enzyme is self-catalytically activated by sequential removal of the C-terminal and N-terminal propeptides. *FEBS Lett.* **447**, 213–216
- Chen, J.-M., Rawlings, N. D., Stevens, R. A. E. and Barrett, A. J. (1998) Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases. *FEBS Lett.* **441**, 361–365
- Nasmyth, K., Peters, J. M. and Uhlmann, F. (2000) Splitting the chromosome: cutting the ties that bind sister chromatids. *Science* **288**, 1379–1385
- Picard, V., Ersdal-Badju, E., Lu, A. and Bock, S. C. (1994) A rapid and efficient one-tube PCR-based mutagenesis technique using *Pfu* DNA polymerase. *Nucleic Acids Res.* **22**, 2587–2591
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350–4354
- Klausner, R. D., Donaldson, J. G. and Lippincott-Schwartz, J. (1992) Brefeldin A: insights into the control of membrane traffic and organelle structure. *J. Cell Biol.* **116**, 1071–1080
- Isidoro, C., De Stefanis, D., Demoz, M., Ogier-Denis, E., Codogno, P. and Baccino, F. M. (1997) Expression and posttranslational fate of cathepsin D in HT-29 tumor cells depend on their enterocytic differentiation state. *Cell Growth Differ.* **8**, 1029–1037
- Ling, H., Ardjomand, P., Samvakas, S., Simm, A., Busch, G. L., Lang, F., Sebekova, K. and Heidland, A. (1998) Mesangial cell hypertrophy induced by NH₄Cl: role of depressed activities of cathepsins due to elevated lysosomal pH. *Kidney Int.* **53**, 1706–1712
- Fernandez, M. A., Barrett, A. J., Gerhartz, B., Dando, P. M., Ni, J. and Abrahamson, M. (1999) *J. Biol. Chem.* **274**, 19195–19203
- Cyglar, M. and Mort, J. S. (1997) Proregion structure of members of the papain superfamily. Mode of inhibition of enzymatic activity. *Biochimie (Paris)* **79**, 645–652

- 23 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
- 24 Carmona, E., Dufour, É., Plouffe, C., Takebe, S., Mason, P., Mort, J. S. and Ménard, R. (1996) Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. *Biochemistry* **35**, 8149–8157
- 25 Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J. S. and Cygler, M. (1996) Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J.* **15**, 5492–5503
- 26 Maubach, G., Schilling, K., Rommerskirch, W., Wenz, I., Schultz, J. E., Weber, E. and Wiederanders, B. (1997) The inhibition of cathepsin S by its propeptide—specificity and mechanism of action. *Eur. J. Biochem.* **250**, 745–750
- 27 Kessler, E. and Safrin, M. (1994) The propeptide of *Pseudomonas aeruginosa* elastase acts as an elastase inhibitor. *J. Biol. Chem.* **269**, 22726–22731
- 28 Serkina, A. V., Gorozhankina, T. F., Shevelev, A. B. and Chestukhina, G. G. (1999) Propeptide of the metalloprotease of *Brevibacillus brevis* 7882 is a strong inhibitor of the mature enzyme. *FEBS Lett.* **456**, 215–219
- 29 Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J. D. (1994) *Molecular Biology of the Cell*, Garland Publishing, Inc., New York
- 30 Ohkuma, S. and Poole, B. (1978) Fluorescence probe measurement of the intralysosomal pH in living cells and the perturbation of pH by various agents. *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3327–3331
- 31 Kundra, R. and Kornfeld, S. (1999) Asparagine-linked oligosaccharides protect Lamp-1 and Lamp-2 from intracellular proteolysis. *J. Biol. Chem.* **274**, 31039–31046
- 32 Khan, A. R. and James, M. N. G. (1998) Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Sci.* **7**, 815–836
- 33 Salvesen, G. S. and Dixit, V. M. (1999) Caspase activation: The induced-proximity model. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10964–10967
- 34 Mach, L., Mort, J. S. and Glössl, J. (1994) Maturation of human procathepsin B. Proenzyme activation and proteolytic processing of the precursor to the mature proteinase, *in vitro*, are primarily unimolecular processes. *J. Biol. Chem.* **269**, 13030–13035
- 35 McQueney, M. S., Amegadzie, B. Y., D'Alessio, K., Hanning, C. R., McLaughlin, M. M., McNulty, D., Carr, S. A., Ijames, C., Kurdyla, J. and Jones, C. S. (1997) Autocatalytic activation of human cathepsin K. *J. Biol. Chem.* **272**, 13955–13960
- 36 Ménard, R., Carmona, E., Takebe, S., Dufour, É., Plouffe, C., Mason, P. and Mort, J. S. (1998) Autocatalytic processing of recombinant human procathepsin L—contribution of both intermolecular and unimolecular events in the processing of procathepsin L *in vitro*. *J. Biol. Chem.* **273**, 4478–4484

Received 23 June 2000/23 August 2000; accepted 14 September 2000