Activity and Deletion Analysis of Recombinant Human Cathepsin L Expressed in *Escherichia coli* 

Spencer M. Smith$ and Michael M. Gottesman|&

From the $Howard Hughes Medical Institute and the Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

A cDNA clone encoding the human cysteine protease cathepsin L was expressed at high levels in *Escherichia coli* in a T7 expression system. The insoluble recombinant enzyme was solubilized in urea and refolded at alkaline pH. 38-kDa procathepsin L was purified by gel filtration at pH 8.0, and a 29-kDa form of the enzyme was purified by gel filtration after autocatalytic processing of the proenzyme at pH 6.5. The kinetic properties of the 29-kDa species of recombinant cathepsin L were similar to those published for the human liver enzyme (Mason, R. W., Green, G. D. J., and Barrett, A. J. (1985) *Biochem. J.* 226, 233–241), using benzyloxycarbonyl-Phe-Arg-7-(4-methyl)coumarylamide as substrate. However, the stability of the recombinant enzyme was decreased relative to the human enzyme, and its pH optimum for this substrate was shifted to a higher pH. Structure-function studies of cathepsin L were performed by constructing mutations in either the propeptide portion or the carboxy-terminal light chain portion of the protein. These constructions were expressed in the *E. coli* system, and enzymatic activities were assayed following solubilization, renaturation, and gel filtration chromatography of the mutated proteins. Deletions of increasing size in the propeptide resulted in large proportional losses of activity, indicating that the propeptide is essential for proper enzyme folding and/or processing in this renaturation system. Deletion of part of the light chain containing a disulfide-forming cysteine residue or a single amino acid substitution of alanine for this cysteine residue resulted in almost complete loss of activity. These data suggest that the disulfide bond joining the heavy and light chains of cathepsin L is essential for enzymatic activity.

Cathepsin L is the most active of the lysosomal cysteine proteases with regard to its ability to hydrolyze azocasein (Barrett and Kirschke, 1981), elastin (Mason *et al.*, 1986), and collagen (Kirschke *et al.*, 1982). Cathepsin L is synthesized as a preproenzyme, which is processed to a proenzyme and targeted to the lysosomes by the mannose 6-phosphate lysosomal recognition marker (von Figura and Hasilik, 1986). The enzyme is further processed to mature forms consisting of either a single polypeptide or an active site cysteine heavy chain and a light chain linked by a disulfide bond (Barrett and Kirschke, 1981).

Cathepsins appear to play a major role in intracellular protein degradation and turnover (Bohly *et al.*, 1979), bone remodeling (Delasalle *et al.*, 1986), and prohormone activation (Marx, 1987). It has been suggested that cathepsins are involved in a variety of disease processes such as glomerulonephritis (Thomas and Davies, 1989), arthritis (Van Noorden *et al.*, 1988), and cancer metastasis (Sloane and Honn, 1984). Frocathepsin L has been shown to interfere with antigen processing (McCoy *et al.*, 1986). In several transformed cells, cathepsin L is secreted in large amounts (Gottesman, 1978), presumably due to both high levels of synthesis and a low affinity of the proenzyme for the mannose 6-phosphate receptor (Dong *et al.*, 1989).

To obtain a ready source of easily purifiable cathepsin L for studies of its function in malignancy and other disease processes and to develop a system for analyzing mutant forms of the enzyme, we expressed a cDNA clone of human cathepsin L in an *Escherichia coli* expression system. Recombinant human procathepsin L was successfully expressed at high levels and purified as both procathepsin L and active processed cathepsin L forms. Moreover, information about the possible function of the propeptide in cathepsin L folding and or processing and about the necessity for the light chain of the enzyme for protease activity was obtained by expressing and purifying mutant enzymes carrying structural alterations in these regions.

**EXPERIMENTAL PROCEDURES**

**Bacterial Strains and Plasmids—** *E. coli* strain HB101 was used for subcloning and for purification of plasmid DNA for sequencing and enzyme digests (Boyer and Roulland-Dussoix, 1969). *E. coli* strain BL21(DE3) is a λ-lysogen containing the gene for T7 RNA polymerase in the chromosome under control of the IPTG-inducible lacUV5 promoter (Studier and Moffatt, 1986). Both bacterial strains were grown in Luria broth (Quality Biological) at 37 °C.

The triplet set of plasmids pAR2156, pAR2106, and pAR2113 was a gift from William Studier. They are modified pBR322 plasmids containing the phage T7 φ10 promoter and the beginning of the coding sequence of gene 10 inserted at the pBR322 BamHI site. Each plasmid allows insertion of genes at the BamH I site in a different reading frame. It is therefore possible to express any gene in the correct frame as a fusion protein with gene 10. Plasmid pH16 contains the full length human cathepsin L cDNA in an Okayama-Berg vector (Gal and Gottesman, 1988).

**Cathepsin L DNA Constructs—** DNA isolation and manipulations were carried out using standard procedures (Maniatis *et al.*, 1982). Restriction enzymes were obtained from Bethesda Research Laboratories or New England BioLabs. Oligonucleotides were synthesized either by a single polypeptide or an active site cysteine heavy chain and a light chain linked by a disulfide bond (Barrett and Kirschke, 1981).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

**Howard Hughes Medical Institute and National Institutes of Health Research Scholar.**

$ To whom correspondence should be addressed: NIH, Bldg. 37, Rm. 2F18, Bethesda, MD 20892. Tel.: 301-496-1530. Fax: 301-496-0260.

1 The abbreviations used are: IPTG, isopropyl-1-thio-β-D-galactopyranoside; SDS, sodium dodecyl sulfate; Z, benzyloxycarbonyl; NMeC, 7-(4-methyl)coumarylamide.
with an Applied Biosystems 380B DNA synthesizer.

The final constructions used in this work are shown in Fig. 5. To construct pCtSNL29, the pHu16 1138-base pair AvaiI-Drai human cathepsin L cDNA fragment was isolated and blunt ended with the Klenow fragment of DNA polymerase I (Bethesda Research Laboratories). Plasmid pAR213 was digested with 4.5 units of either AvaiI or DraiI and blunt-ended with Klenow fragment, dephosphorylated with calf intestinal phosphatase (Boehringer Mannheim), then ligated to the AvaiI-Drai cathepsin L cDNA fragment using T4 DNA ligase (Bethesda Research Laboratories). The resulting construct encodes CtsLNA22 cathepsin L, which includes the entire protein lacking the 17-amino acid signal peptide and the first 12 amino acids of the proenzyme. Plasmids pCtSNL87 and pCtSNL115 were constructed by ligating the pHu16 XhoII cDNA fragment and BamHI-digested, phosphatase-treated pAR2113.

Mutant CtsLNA115 cathepsin L lacks the entire 96-amino acid propeptide and the first two NH2-terminal amino acids of the mature cathepsin.

Plasmids pCtsLNA22 and pCtSLCC-A are both COOH-terminal modifications of pCtsLNA29 and therefore code for cathepsin L lacking the first 12 amino acids of the propeptide. pCtsLCA22 was constructed by digesting pCtSNL29 in the 3′-coding region with SnaBI, treating with phosphatase, then ligating with a T4 polynucleotide kinase-treated 34-base pair synthetic linker containing the stop codon in all reading frames. The linker premature stop codon causes deletion of the final 22 amino acids of the enzyme light chain and addition of lysine as the final amino acid of mutant CtsLCA22 cathepsin L.

To construct pCtsLCC-A, plasmid pCtsLNA29 was digested at the SnaBI site in the cDNA 3′-coding region and at the downstream HindIII site within the vector. The plasmid was separated from the SnaBI-HindIII cleavage fragment and dephosphorylated. A 73-base pair synthetic linker with ends complementary to the plasmid’s SnaBI blunt end and HindIII sticky end was synthesized. This linker was identified in the cDNA 3′-coding region allowing the replacement of the normal stop codon with a UGA stop codon for the expression of the protein in the cytoplasm (pCtsLNA29, pCtsLNA87, pCtsLCA2.2, pCtsLCC-A) from BamHI linkers with the exception of a substi-
acids of the 96-amino acid propeptide (see Fig. 5). 3 h after induction of T7 polymerase with IPTG, recombinant CtsLNA29 cathepsin L represented approximately 20% of the total bacterial protein (Fig. 1, lane 9 h postinduction (data not shown). CtsLNA29 cathepsin L represented approximately 20% of the total protein continued to increase slightly 6–9 h postinduction (data not shown).

**Purification of Procathepsin L and Processed Cathepsin L**

Table I summarizes the steps involved in the purification of cathepsin L from bacterial lysates. Proteins overexpressed in *E. coli* frequently form insoluble aggregates or inclusion bodies in the bacterial cytoplasm (Marston, 1986). When plasmid pCtsLNA29 was expressed in this system, the recombinant protein proved to be very insoluble and could only be solubilized in strong detergents or denaturants above 37 °C. 8 M urea was the only solubilizing agent from which active cathepsin L was recovered following renaturation.

Several renaturation conditions were empirically found to be important for refolding of functional CtsLNA29 protein. 1) Glutathione was required in the renaturation buffer at an optimal reduced to oxidized ratio of 10:1. 2) Because mouse procathepsin L is known to be autocatalyzed at acid pH (Gal and Gottesman, 1986), renaturation was carried out at neutral or alkaline pH to ensure recovery of the procathepsin enzyme form. pH 10.7 was found to be the optimal pH for renaturation. Marston et al. (1984) reported a similar finding for the refolding of urea-solubilized recombinant calf prochymosin, another protease expressed in *E. coli*. 3) High dilution of the urea-solubilized extract in renaturation buffer by slow dropwise addition was beneficial. 200-fold dilution yielded 5-fold more activity than 50-fold dilution, and overnight dropwise dilution yielded 3.5-fold more activity than rapid dilution.

Fig. 1 represents the proteins obtained at various stages in the purification process. The majority of protein in the renaturation extract refolded into inactive insoluble aggregates that stuck to the ultrafiltration membrane or precipitated upon centrifugation of the ultrafiltrate. The remainder of the protein was primarily dimeric procathepsin L, as determined by gel filtration chromatography (data not shown) and 40- and 38-kDa forms of monomeric procathepsin L (Fig. 1, lane 3). Sephadex G-75 chromatography at pH 8.0 effectively separated monomeric from dimeric cathepsin L (lane 4). The procathepsin was somewhat unstable during chromatography at pH 8.0. The 40-kDa form, which is of the expected molecular mass for unglycosylated human cathepsin L (Smith et al., 1989), was converted to a 38-kDa protein. This species in turn was partially processed to a 30-kDa protein (lanes 3 and 4). After Sephadex G-75 chromatography at pH 6.5 of pooled procathepsin fractions, the enzyme was isolated almost entirely as a 29-kDa polypeptide (lane 5). Thus, the reducing gel in Fig. 1 appears to demonstrate sequential conversion of the enzyme from a 40-kDa to a 38-kDa single-chain pro-form and from a 30-kDa to a 29-kDa single-chain processed form as the pH was lowered from 8.0 to 6.5. Total enzyme activity against the synthetic substrate Z-Phe-Arg-NMec increased in parallel with this processing (Table I). Although less than 0.1% of the total cathepsin L produced in *E. coli* was recovered in purified active form (Table I), this procedure was rapid, and total yields were high (0.5 mg/4-liter bacterial culture) in comparison with the multistep purification protocols for cathepsin L from mammalian tissues (Mason et al., 1984, 1985).

**Catalytic Activity of Recombinant Cathepsin L**

Action on Z-Phe-Arg-NMec—The 29-kDa processed form of CtsLNA29 cathepsin L demonstrated high activity against Z-Phe-Arg-NMec, the optimal synthetic substrate for cathepsin L (Mason et al., 1985). The specific activity of pooled pH 6.5 Sephadex G-75 column enzyme fractions for this substrate was 22,500 nmol/min/mg (Table I). This activity is similar to the 34,400 nmol/min/mg value obtained for rabbit cathepsin L (Mason et al., 1984). Likewise, the *K*<sub>a</sub> (1.1 μM) and *K*<sub>cat</sub> (39/s) values for the recombinant enzyme are comparable to those reported for cathepsin L from five different species (Table II). The finding that unglycosylated recombinant human cathepsin L is enzymatically active corroborates a previous study of deglycosylated mouse cathepsin L (Smith et al., 1989) which suggested that glycosylation of cathepsin L is not essential for enzymatic function.

<table>
<thead>
<tr>
<th>Purification step</th>
<th>Protein</th>
<th>Total activity</th>
<th>Specific activity</th>
<th>Purification factor</th>
<th>Yield % activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total bacterial lysate</td>
<td>6,730</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urea-solubilized renatured extract</td>
<td>168</td>
<td>18</td>
<td>111</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Ultrafiltration retentate</td>
<td>19</td>
<td>7.84</td>
<td>413</td>
<td>3.7</td>
<td>42.1</td>
</tr>
<tr>
<td>Sephadex G-75 fractions, pH 8.0</td>
<td>1.34</td>
<td>7.230</td>
<td>7.230</td>
<td>65.1</td>
<td>51.8</td>
</tr>
<tr>
<td>Sephadex G-75 fractions, pH 6.5</td>
<td>0.521</td>
<td>11.70</td>
<td>22,500</td>
<td>203</td>
<td>62.8</td>
</tr>
</tbody>
</table>
Recombinant Human Cathepsin L

pH Stability—Fig. 2 demonstrates the stability of recombinant CtsLNA29 enzyme activity against Z-Phe-Arg-NMec as a function of pH for both the processed form and the pro-form of the protein. Activity of 29-kDa processed enzyme was unstable at 37 °C. Following incubation at 37 °C for 1 h, processed cathepsin L lost over 95% of its hydrolyzing ability at all pH values. Mature cathepsin L purified from human, rabbit, sheep, or ox liver, on the other hand, retains over 90% activity at pH 6.0 under these same conditions (Mason, 1986). On the other hand, 38-kDa procathepsin L samples retained 90% activity after 1 h at 37 °C at pH 3.5-4.0 but rapidly lost activity as the pH was raised above 4.0. Less than 7% activity was retained above pH 5.5. Interpretation of this result is difficult since procathepsin L is rapidly autoprocessed to mature cathepsin L at low pH (see below), which might increase its activity. However, in contrast, native mouse procathepsin L is most stable at pH 6.0-6.5 and retains over 90% activity at pH 8.0 (Mason et al., 1987).

The time course of recombinant procathepsin L pH-dependent autoprocessing at 37 °C is seen in Fig. 3. This Western blot shows nearly complete processing to both 30- and 29-kDa mature forms in 1 min at pH 3.0 and in 5 min at pH 5.0. Processing was not observed at pH 8.0; processing at pH 6.0 and 7.0 was incomplete at 60 min and appeared limited to formation of 30-kDa cathepsin (Fig. 3). This supports the observation made during the purification process that the procathepsin is sequentially processed from 38 to 30 to 29 kDa (Fig. 1). In contrast to the processing of recombinant human procathepsin L seen below pH 8.0, no autoprocessing of mouse procathepsin L is observed at 37 °C for 30 min at pH 5.0 (Mason et al., 1987). Because recombinant procathepsin L was autoprocessed within 5 min below pH 5.0 (Fig. 3) in this assay, newly processed mature cathepsin L must be stable for at least 55 min at low pH and 37 °C in contrast to the autoprocessed recombinant enzyme purified and stored at pH 6.5 (Fig. 2).

pH Activity Profiles—Fig. 4 shows the enzyme activities of recombinant procathepsin L and processed cathepsin L against Z-Phe-Arg-NMec as a function of pH. Neither purified species demonstrated significant substrate hydrolysis at pH 8.0. However, the pH optimum of recombinant mature cathepsin L (pH 6.5) was higher than that published for human liver cathepsin L under similar conditions (pH 5.5) (Mason et al., 1987). Moreover, the pH profile of recombinant processed cathepsin L more closely resembled that of mature rabbit liver cathepsin L than that of mature human liver cathepsin L (Mason et al., 1984) but had a higher percent maximum activity below pH 5.0 (70%) than did either of the other two enzymes (30 and 40% for rabbit and human, respectively).

Fig. 4 demonstrates that recombinant procathepsin L had peak activity at pH 4.0. Activity fell to 38% maximum at pH 5.0 and remained above 30% maximum to pH 7.0. Since procathepsin L is rapidly converted to mature cathepsin L at pH 7.0 and below (Fig. 3), this activity profile may be entirely attributable to the presence of enzymatically active mature cathepsin L.

Table II

<table>
<thead>
<tr>
<th>Cathepsin L source</th>
<th>K_a (μM)</th>
<th>K_m (s^-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recombinant human</td>
<td>1.1</td>
<td>39</td>
</tr>
<tr>
<td>Human liver</td>
<td>2.4</td>
<td>17</td>
</tr>
<tr>
<td>Rabbit liver</td>
<td>1.8</td>
<td>20</td>
</tr>
<tr>
<td>Rat liver</td>
<td>4.5</td>
<td>26</td>
</tr>
<tr>
<td>Ox liver</td>
<td>2.2</td>
<td>10</td>
</tr>
<tr>
<td>Sheep liver</td>
<td>1.8</td>
<td>30</td>
</tr>
</tbody>
</table>

* Mason (1986).

Structure-Function Analysis of Recombinant Cathepsin L

Expression of Mutant Cathepsin L Constructs—To assess the possible functional roles of the propeptide and of the light chain, new cDNA constructs were engineered bearing alterations in the protein-coding sequences in the NH_2 and COOH

---

**Fig. 2.** Recombinant human procathepsin L and processed cathepsin L pH stability. Enzyme samples were preincubated for 1 h at 37 °C in a range of buffers (pH 3.0-8.0) as described under "Experimental Procedures." Activity remaining was assayed at pH 5.5 with the substrate Z-Phe-Arg-NMec and expressed as a percentage of activity measured without preincubation. □, recombinant procathepsin L; □, recombinant processed cathepsin L.
FIG. 3. Time course and pH dependence of human procathepsin L autocatalytic processing. Purified recombinant procathepsin L samples (10 μg) were preincubated at 37 °C for different lengths of time (0–60 min) in a range of buffers (pH 3.0–8.0), then placed on ice. Samples were then subjected to SDS-polyacrylamide gel electrophoresis under reducing conditions and Western transfer as described under “Experimental Procedures.” Arrows indicate procathepsin L and processed cathepsin L.

![Diagram of time course and pH dependence of human procathepsin L autocatalytic processing]

The altered forms of the protein are illustrated in Fig. 5. Plasmids pCtsLNA87 and pCtsLNA115 have propeptide deletions, pCtsLCΔ22 has a deletion in the light chain, and pCtsLCC→A encodes a cysteine to alanine single amino acid substitution in the light chain. These constructions were highly expressed in E. coli as proteins of the expected molecular masses (Fig. 6). The new proteins appearing in the bacterial lysates in Fig. 6 all were reactive on Western blots (Fig. 7) with a polyclonal rabbit antiserum that specifically recognizes human procathepsin L.

The same solubilization/renaturation/gel filtration protocol outlined above was applied to each bacterial lysate containing mutant enzyme after determination that each mutant was expressed in insoluble form (data not shown). Sephadex G-75 fractions were assayed for activity against Z-Phe-Arg-NMec at pH 5.5 following a 30-s incubation at pH 3.0. Fig. 8 demonstrates that Sephadex G-75, pH 8.0, protein patterns were similar for CtsLNA29 cathepsin L and for the mutant enzymes, although specific activities were markedly different. A large peak of high molecular mass protein was followed by a smaller peak at approximately 30 kDa, the early part of which coeluted with enzyme activity. Fig. 5 summarizes the peak specific activities for each enzyme variant. All protease activities were >98% inhibitable with 2.5 μM E-64, a specific irreversible cysteine protease inhibitor.

Propeptide Mutations—As demonstrated above, although CtsLNA29 cathepsin L is a fusion protein, with 14 amino acids from the bacterial expression plasmid in place of the first 12 amino acids of the propeptide, it is highly active enzymatically. Deletion of the first 70 amino acids (73%) of the propeptide (CtsLNA87 cathepsin L) resulted in a 278-fold loss of specific activity (Figs. 5 and 8). Furthermore, deletion of the entire propeptide plus the first 2 amino acids of the heavy chain (CtsLNA115 cathepsin L) resulted in an addi-

---

2 S. M. Smith and M. M. Gottesman, unpublished data.
**Recombinant Human Cathepsin L**

**FIG. 5.** Structure and enzymatic activity of recombinant human cathepsin L mutants. DNA constructions for each mutant are described under "Experimental Procedures." Peak specific activities of each solubilized renatured enzyme against Z-Phe-Arg-NMec are reported in nmol/min/mg. The amino acid numbers at the junctions of the pre-, pro-, heavy chain, and light chain peptides are given at the top of the figure. Fusion, pro-, heavy chain, and light chain peptide portions of each protein are indicated by striped, lightly dotted, black, and heavily dotted rectangles, respectively. Cathepsin L cDNA restriction enzyme sites utilized in engineering mutant constructs are A, AvaII; E, EcoRI; X, XhoII; and S, SnaBI. C denotes a putative disulfide bridge. Heavy chain active site cysteinyl (white C) and histidyl (white H) residues are indicated.

### Specific Activity (nmol/min/mg)

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Pro</th>
<th>Heavy</th>
<th>Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>CtsLNA29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CtsLNA87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CtsLNA115</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CtsLCΔ22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CtsLCC→A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIG. 6.** Expression of recombinant human cathepsin L mutants in E. coli. Induced whole cell bacterial lysates were subjected to SDS-polyacrylamide gel electrophoresis as outlined under "Experimental Procedures." Lane 1, protein standards. Lane 2, BL21(DE3). Lane 3, BL21(DE3) + pCtsLNΔ29. Lane 4, BL21(DE3) + pCtsLCC→A. Lane 5, BL21(DE3) + pCtsLNΔ87. Lane 6, BL21(DE3) + pCtsLNΔ115. Lane 7, BL21(DE3) + pCtsLCΔ22. The arrow indicates the position of CtsLNΔ29 procathepsin L.

**FIG. 7.** Immunoblotting of recombinant human cathepsin L mutants expressed in E. coli. Induced whole cell BL21(DE3) bacterial lysates were electrophoresed, transblotted onto nitrocellulose, immunolabeled, and autoradiographed as described under "Experimental Procedures." Lane 1, BL21(DE3). Lane 2, BL21(DE3) + pCtsLNΔ29. Lane 3, BL21(DE3) + pCtsLCC→A. Lane 4, BL21(DE3) + pCtsLNΔ87. Lane 5, BL21(DE3) + pCtsLNA115. Lane 6, BL21(DE3) + pCtsLCΔ22. The arrow indicates CtsLNΔ29 procathepsin L. The 70-kDa band in each lane and other minor bands probably represent bacterial proteins that cross-react with the polyclonal rabbit antisera.

residues and is believed to form three disulfide bonds (Ritonja et al., 1988). One disulfide bridge links the cathepsin L heavy and light chains (Kirschke and Barrett, 1985). Invariably conserved cysteine 322 is the only light chain cysteine and must therefore be involved in this linkage (Fig. 9). In order to assess enzymatic activity in the absence of this disulfide bridge, amino acid 322 was changed from cysteine to structurally similar alanine in CtsLCC→A cathepsin L. This substitution resulted in a 12,700-fold loss of activity (Figs. 5 and 7).
The dramatic effect of this single amino acid substitution on enzymatic activity strongly suggests that the bridging disulfide bond is necessary for proper orientation of the light chain relative to the heavy chain.

No substrate hydrolysis was detected in Sephadex G-75 fractions for CtsLCA22 cathepsin L, from which the COOH-terminal 22 amino acids (51%) were deleted (Figs. 5 and 8). This again implicates the light chain, or a portion of it, in enzymatic function in vitro.

DISCUSSION

We have shown that recombinant human cathepsin L can be expressed at high levels in E. coli and can be purified as either a 38-kDa proenzyme or a 29-kDa processed enzyme. The 29-kDa processed form has a specific activity and kinetic constants comparable to those described for human liver cathepsin L when assayed with the substrate Z-Phe-Arg-NMec. Mort et al. (1988) have reported expression of the lysosomal cysteine proteases, rat and mouse cathepsin B, in E. coli. However, expression was at very low levels, and no cysteine protease activity was recovered.

The products of pH-dependent processing of cathepsin L detected in vitro differ from the products of processing in vivo. Purified recombinant human procathepsin L and purified mouse procathepsin L both undergo similar sequential processing at low pH in vitro to proteins of lower molecular mass which are each believed to consist of a single polypeptide chain. The 40-kDa recombinant human proenzyme is converted to a 38-kDa species, then to 30- and 29-kDa species (Figs. 1 and 3). Likewise, the 39-kDa mouse proenzyme is converted to a 37-kDa species, then to 29- and 27-kDa species (Gal and Gottesman, 1986). In contrast, in vivo processing of procathepsin L results in the appearance of a single one-chain mature protein followed by a single two-chain mature protein that becomes the major intracellular species (Smith et al., 1989; Gal et al., 1985; Nishimura et al., 1988). Processing of procathepsin L to mature products in vivo is thought to be sequentially mediated by cellular carboxyl protease(s) (Nishimura et al., 1988a) and cysteine protease(s) (Hara et al., 1988). The differences between the processing products formed in vivo and in vitro may therefore be due to differences in the cleavage site specificities of the cellular proteases in vivo and in the autocatalytic cleavage specificity of cathepsin L at low pH. The close similarity between the processing patterns of recombinant human procathepsin L and purified mouse procathepsin L in vitro argues that the pH-dependent processing pattern observed for the recombinant enzyme is not an aberration due to alterations in the bacterially expressed protein.

Although liver cathepsin L has only been isolated from...
tissue extracts from several species in the two-chain mature form (Mason, 1986), recombinant human cathepsin L and chicken liver cathepsin L isolated from lysosomes (Dufour et al., 1987) have now both been purified in active single-chain forms. Moreover, a radiolabeled diazomethane active site inhibitor binds both single- and double-chain forms of mouse cathepsin L in vivo (Mason et al., 1989). These data suggest that both forms of processed cathepsin L are proteolytically active.

Secreted procathepsin L may have important functions in malignancy whether or not it is intrinsically active. In tissue culture, malignantly transformed mouse fibroblasts synthesizes and secretes large amounts of procathepsin L, also known as MIEP, for major excreted protein (Gottesman, 1978; Gottesman and Sobel, 1980; Doherty et al., 1985; Rabin et al., 1986). Many human tumors have increased levels of mRNA for procathepsin L, which also has been documented as increased gene expression. This has been presumed to reflect increased gene expression in malignant cells.

The finding of significant activity of mature cathepsin L at neutral pH (Mason et al., 1989) argues for a possible role of cathepsin L in tumor invasiveness or metastasis. Furthermore, secreted procathepsin L may have the ability to function in malignancy without extracellular activation. McCoy et al. (1988) reported that purified mouse procathepsin L interferes with antigen presentation of pigeon cytochrome c to T cells after being internalized into an acidic compartment within antigen-presenting cells.

Mature cathepsin L is not stable for very long at neutral pH (Mason et al., 1987). The stable latent activity of procathepsin L produced by mammalian cells at neutral pH (Nishimura et al., 1988a) argues for a role of the propeptide in stabilizing the enzyme, perhaps allowing the enzyme to pass intact through the less acidic Golgi apparatus before transport to the lysosomes or before secretion (Mason et al., 1987). The pro-sequence of cathepsin L might also be essential for proper protein folding, as demonstrated for the serine protease subtilisin in vitro (Zhu et al., 1989). The in vitro procedure described here for renaturing solubilized recombinant enzyme at alkaline pH does not duplicate the folding and transport conditions within mammalian cells. Nevertheless, the correlation between increased size of the propeptide deletion and decreased recovery of enzymatic function following renaturation suggests that the propeptide has a crucial role in protein folding and/or stability. It is possible that the yield of active enzyme following renaturation of CtsLΔ29 cathepsin L could be improved if full length proenzyme without fusion residues could successfully be expressed in E. coli.

The crystal structures of the plant proteases papain and actinidin indicate involvement of COOH-terminal region amino acid residues in the hydrophobic core and the active site pocket. Because these enzymes share 35-50% overall identity with the lysosomal cysteine proteases, it has been assumed that the thiol cathepsin COOH termini are likewise involved in enzyme specificity and function (Ritonja et al., 1988). Papain consists of an NH2-terminal domain of three helical regions folded around a hydrophobic core and a more stable COOH-terminal domain of antiparallel β-sheets and α-helices folded around a hydrophobic core (Hernandez-Arana and Soriano-Garcia, 1988). The 22 amino acids of the light chain deleted from CtsLCΔ22 cathepsin L (Fig. 5) were predicted to be important for enzymatic function for several reasons (Fig. 9). 1) These amino acids include hydrophobic residues at all five positions occupied by hydrophobic core residues in papain and actinidin; 2) alanine 327 of cathepsin L occupies a position near the S2 active site pocket opening postulated to determine substrate-binding specificity (Barrett and Kirschke, 1981; Joseph et al., 1989); and 3) these amino acids include 10 of 13 residues predicted by circular dichroism analysis of papain and cathepsin L to form an α-helix unique to cathepsin L (Dufour et al., 1988). It is therefore not surprising that deletion of these amino acids abolished enzyme function.

Replacement of the cysteine at position 322 by an alanine should conserve structure but eliminate the ability to form a disulfide bridge between Lys326 and Cys322 (Dufour, 1988). This mutation was designed to assess the role of this conserved amino acid residue in enzymatic function (Fig. 5). Since this is a single amino acid substitution, loss of enzymatic activity probably would not be secondary to aberrant folding in the artificial renaturation system. Rather, loss of activity should indicate a direct role for the substituted cysteine residue (or the disulfide linkage) in protein function. Hence,
we believe that the dramatic loss of activity observed with this mutation indicates that cysteine 322 is essential for enzymatic activity and that a disulfide bond between the two peptides is a critical structural feature.

The ability to produce enzymatically active cathepsin L in E. coli will make it possible to produce large amounts of enzyme for structural studies and will also make it possible to perform detailed analysis on additional mutant forms of this acid protease.

Acknowledgments—We thank Steven Neal for photographic assistance, Joyce Sharrar for secretarial assistance, and Dr. Susan Kane for critical comments on the manuscript.

REFERENCES


