Structural Requirements of Procathepsin D Activation and Maturation*

(Received for publication, February 10, 1994, and in revised form, March 18, 1994)

Gary R. Richo and Gregory E. Conner†
From the Department of Cell Biology and Anatomy, University of Miami School of Medicine, Miami, Florida 33101

Cathepsin D biosynthesis involves several proteolytic events; however, the enzymology and sequence of these events are not known. Procathepsin D undergoes a pH-dependent, intramolecular proteolysis in vitro which removes 26 residues yielding an active form that is intermediate in size between procathepsin D and single-chain cathepsin D. This form, designated pseudocathepsin D, has not been shown to be an in vivo intermediate. The N-terminal sequence of the light chain of cathepsin D, isolated from human placenta, showed that 42 residues were removed as compared with 44 residues predicted by comparison with porcine cathepsin D. Site-directed mutations were generated at both processing sites within the propeptide of procathepsin D. Mutation at the autocatalytic site prevented in vitro autoactivation, but, after transfection of mouse Ltk cells, the mutant procathepsin D was transported to the lysosome and processed normally to the mature enzyme despite its inability to autoactivate in vitro. Mutation at the mature N terminus of cathepsin D prevented in vivo formation of the single-chain form of the enzyme; however, the protein was still processed to the two-chain form of human cathepsin D. This change at the mature N terminus did not prevent in vitro autoactivation. Procathepsin D with mutations at both cleavage sites was processed to the two-chain form despite the inability to undergo removal of the propeptide. These results indicated that stepwise autoactivation and propeptide removal were not necessary for later processing or delivery of human cathepsin D to the lysosome. The results also suggested that pseudocathepsin D was not a normal intermediate of procathepsin D processing in vivo.

Cathepsin D (EC 3.4.23.5) is the major intracellular aspartic proteinase of lysosomes and is related to the other aspartic proteinases such as renin, pepsin, and yeast proteinase A (for review, see Tang and Wong (1987) and Davies (1990)). Most aspartic proteinases are synthesized as preproenzymes containing an endoplasmic reticulum translocation signal and a propeptide for control of enzyme activity. Proteolytic removal of the propeptide results in the activation of the enzyme. This propeptide processing of aspartic proteinases has been well studied in vitro and varies among the different enzymes. Studies of pepsinogen show that, below pH 3.0, conversion to pepsin is unimolecular (Bustin and Conway-Jacobs, 1971; Al-Janabi et al., 1972; McPhie, 1972). The proposed first step in this mecha-

* This work was supported in part by grant GM35812 from the National Institutes of Health and in part by a grant-in-aid from the American Heart Association, Florida Affiliate. 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.
† To whom correspondence and reprint requests should be addressed. Tel.: 305-547-6926; Fax: 305-645-7166.

1 G. R. Richo and G. E. Conner, manuscript in preparation.
of procathepsin D proteolytic processing. These mutations allowed us to examine the role of autoactivation and propeptide removal in the maturation of human cathepsin D in vivo and to learn about sequence requirements for autocatalytic cleavage.

EXPERIMENTAL PROCEDURES

Materials—[35S]Methionine (700 Ci/ mol), [35S]Cysteine (1000 Ci/ mol), and [32P]dATP (1000 Ci/ mol) were purchased from DuPont NEN. All tissue culture media and components were from Life Technologies, Inc. The Sequenase 2.0 DNA sequencing kit and Pfu polymerase was obtained from United States Biochemical Corp. Sequencing and mutagenic primers were synthesized by either the University of Miami DNA Synthesis Facility or the DNA Synthesis Core Laboratory at the University of Florida, Gainesville. Restriction endonucleases, T4 DNA ligase, and Klenow DNA polymerase used for DNA manipulation were from New England Biolabs. T7 polymerase was from Perkin-Elmer Cetus. All other materials were from Sigma.

Growth and Transfection of Mouse Ltk" Cells—Mouse Ltk" cells were grown at 37 °C in 5% CO2 in air in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. Cells transfected with plasmids were maintained in Dulbecco's modified Eagle's medium at a concentration of 105 cells/ml. Cultures were transfected with 2 µg/ml unlabeled calf serum and antibiotics. Transfections were performed at room temperature in air based on the DEAE-dextran method of transfection (McCutchan and Pagano, 1968). Approximately 75% confluent cells were washed three times in serum-free medium and then incubated for 1 h in serum-free medium containing 200 µg/ml DEAE-dextran and 3-4 µg/ml DNA. Complete medium containing 10% dialyzed fetal calf serum was added, the plate was agitated, and the complete medium containing 80 µg chloroquine for 1 h. Cells were washed and allowed to recover for 24 h before the addition of sodium butyrate to the medium at a concentration of 10 mM. Cells were used for metabolic labeling studies 48-60 h after transfection.

Metabolic Labeling and Immunoprecipitation—For pulse-chase analysis, transfected cells were washed three times with calcium- and magnesium-free phosphate-buffered saline, starved for 20 min in methionine-free medium, and then labeled with either [35S]methionine (100 µCi/ml) alone in methionine-free Dulbecco's modified Eagle's medium supplemented with 5% dialyzed fetal calf serum or both [35S]methionine and [32P]cysteine in methionine- and cysteine-free medium. After pulse-labeling, the cells were chased for the indicated times by incubation with complete medium containing γ-glutamylcysteine-free calf serum, unlabeled methionine, and without glutamine. For continuous labeling, cells were incubated in Dulbecco's modified Eagle's medium supplemented with Life Technologies, Inc. GMS-G media supplement instead of γ-glutamylcysteine-free calf serum. For immunoprecipitation, the media were collected and the cells were harvested with SDS lysis buffer, and immunoprecipitation was carried out with antibodies against human cathepsin D (Conner and Uday, 1990) as previously described (Conner et al., 1989) or with a human specific anti-cathepsin D polyclonal antibody obtained from Trion Bioeciences (kind gift of Dr. Steven Wall). Modified Laemmli SDS-gel electrophoresis using 12% or 15% polyacrylamide gels was performed as described by Fisher et al. (1982). Dried gels were fluorographed and exposed to preflashed Kodak XAR-5 film to ensure a linear response of the film (Laskey and Mills, 1975).

Activation of Procathepsin D in Media—The media from cells continuously labeled in 17 µ chloroquine were collected and centrifuged at 10,000 x g. Sodium formate, pH 5.5, was added to a final concentration of 0.1 M. Following incubation at 37 °C for 0.5 h, the samples were dried down in a Speed-Vac, washed with 10% trichloroacetic acid, and resuspended in sample buffer for analysis on SDS-polyacrylamide gels.

Isolation and Sequencing of Human Cathepsin D—Human placental cathepsin D was isolated according to the protocol of Huang et al. (1979) with a few modifications. All steps were carried out at 4 °C. Placental tissue was homogenized in 10 ml Tris-HCl, pH 7.4, 0.5% Brij 35 with a Waring blender. Following centrifugation at 10,000 rpm for 0.5 h, the supernatant was adjusted to pH 4.7 with 5 M HCl and made to 0.1 M sodium formate. After 30 min, the supernatant was applied to a phosphatase-agarose column which had been prepared according to Huang et al. (1979) and equilibrated with 0.1 M sodium acetate, pH 3.5, 0.1% Brij 35, 1 M NaCl. The column was washed with the equilibration buffer and eluted with 50 mM Tris-HCl, pH 8.6, 0.1% Brij 35, 1 M NaCl. The eluted fraction was divided into aliquots and adjusted to 100 µM phosphatase, pH 7.0, 0.1% Brij 35, 0.1 M NaCl and purified human cathepsin D was then applied to a DEAE-Sephael column. The purified cathepsin D was found in the flow through of this column.

The Ilc27p.Asp mutant form of procathepsin D was purified by pepstatin affinity chromatography of conditioned media collected from eighteen 100-mm dishes of transfected Ltk" cells. The media were adjusted to 0.1 M sodium formate, pH 3.5, 0.4 M NaCl, 0.1% Brij 35 and applied to a phosphatase agarose column (0.25 ml) equilibrated with the same buffer. The column was washed with 100 ml of Tris-HCl, pH 3.5, 0.4 M NaCl, 0.1% Brij 35 and then eluted with 20 ml Tris-HCl, pH 8.3, 0.4 M NaCl, 0.02% Brij 35. The purified mutant form of procathepsin D was then treated as described above to allow for activation.

The purified human cathepsin D light chains and the Ilc27p.Asp activation intermediate were run on 12% SDS-polyacrylamide gels and transblotted to Immobilon P membranes according to the procedure of Matsudaira (1987). The protein bands were cut from the membranes and sequenced using a ABI Model 470A Gas Phase Protein Sequencer (Protein Chemistry Core Laboratory, University of Florida, Gainesville, FL).

Construction of Mutants—pCPSD1, a human fibroblast cathepsin D cDNA (Conner et al., 1988), was used as the construction of the mutants. A 2130-base pair HindIII-Hind111 fragment containing 5'-untranslated sequence, the entire coding sequence, and some 3'-untranslated sequence was cloned into the M13mp18. Two EcoRI sites within the coding region and flanking the processing sites in the propeptide were introduced using the two-primer method for site-directed mutagenesis (Zoller and Smutka, 1984). The first EcoRI site was introduced by changing an BamHI site at position 655 (numbering beginning with the HindIII site of the original pCDX vector, Okayama and Berg (1983)) with the oligonucleotide primer (5' GCCTGGCTGTAAGTGGGAGGCTATAAGAAGGGG) and a primer to the M13mp18 vector. The second EcoRI site was introduced at position 1180 with the primer (5' CATGTTGTGGAGTTGAAAGGGG) and a primer to the M13mp18 vector. These mutants were transfected into the cells, and the primers which contained the Leu26p.Pro and the Leu26p.Val mutation alone and in tandem with the Val42p.Pro mutation were made using the polymerase chain reaction "megaprimer" method (Sarkar and Sommer, 1990). Using the pcPSD2 expression vector as a template, the primer for the Leu26p.Val mutation was (5' TGTTGAGCTGCTGATGAGGAGGCCCC) or (5' GCCTGGCTGAGCGGCGCCTAGGAGCCG). Following confirmatory sequencing of the mutations, the EcoRI cassette was cloned back into the pCPSD2 expression vector. The mutants which contained the Leu26p.Pro and the Leu26p.Val mutation alone and in tandem with the Val42p.Pro mutation were made using the polymerase chain reaction "megaprimer" method (Sarkar and Sommer, 1990). Using the pcPSD2 expression vector as a template, the primer for the Leu26p.Val mutation was (5' TCTTGGCCAGCCCGCTGTTGGAAGGGG) and for the Leu26p.Pro mutation the primer was (5' TCTTGGCCAGCCCGCTGTTGGAAGGGG). A primer complementary to the coding strand of cathepsin D and downstream of the second EcoRI site was used in the first polymerase chain reaction (5' AACAGCCGGCCTAGGACCTAGTT). The second polymerase chain reaction used the same primer and the first primer synthesized in the first reaction along with a primer complementary to cathepsin D sequence upstream from the first EcoRI site (5' ACTGCCCAGCCACCCACTCCCGGCTG). Polymerase chain reaction products were digested with EcoRI and cloned into the expression vector and sequenced. The double mutants with changes introduced at both processing sites were done also with the megaprimer method using the expression vector which already had the desired Val42p.Pro mutation as a template.

RESULTS

N-terminal Sequence of Human Cathepsin D Light Chain—Cathepsin D was isolated from human placenta by pepstatin affinity chromatography and ion exchange chromatography, and the purified protein was subjected to SDS-polyacrylamide gel electrophoresis and transblotted to Immobilon P membrane. The light chain of the two-chain enzyme migrated as two discreet bands on the gel. The bands were cut from the Immobilon P membrane and subjected to Edman degradation. Both bands showed N-terminal sequence heterogeneity (Fig. 1A). Three N-terminal sequences were found within each band. These results suggested that cleavage of the propeptide may be occurring at one of three different bonds or that cleavage of the propeptide is followed by aminopeptidase trimming, and that

Ilc27p.Asp denotes a site-directed change of Ile-27 in the propeptide to Asp.
Processing of Human Procathepsin D

A.

**Human procathepsin D**

**Human pseudocathepsin D**

**Human cathepsin D light chains**

**Porcine cathepsin D light chain**

---

The three distinct N termini which were detected (molar ratio in a, et al., 1980 Processing SDS gels and transferred to Immobilon. The light chains were sequenced at the ICBR Protein Core at the University of Florida (Gainesville, FL).

**Scissile Bond Does Not Block**

would take place between Va142p and Thrl or N-terminal to Leu26p-Ile27p cleavage site within the propeptide.

**Active site cleft with the amino acid side chains of residues at Ala28p bond.**

**Of the scissile bond by replacing Ile27p in the P', position with an alternative choice, that of changing Leu26p to a charged residue, such as glutamate (see Fig. 1A).** For the autocalytic activation of an aspartic proteinase to occur, the substrate must be able to reside in the extended active site cleft with the amino acid side chains of residues at the scissile bond fit into the S, and S', subsites so that the catalytic mechanism of the enzyme may work. Based on the sequence at the cleavage site and the substrate preference of cathepsin D for hydrophobic residues at the scissile bond (Powers et al., 1977), we decided to change the hydrophobic nature of the scissile bond by replacing Ile27p in the P' position with an aspartic acid (see Fig. 1B for all mutations constructed). The alternative choice, that of changing Leu26p to a charged residue, might have allowed cleavage at the adjacent Ile27p-Ala28p bond.

Mouse Ltk<sup>−</sup> cells were transfected with the mutant cDNA, and, 48 h later, the cells were continuously labeled for 8 h by the addition of [35S]methionine to serum-free medium containing chloroquine to increase the secretion of procathepsin D into the medium (Gonzalez-Noriega et al., 1980). The medium was collected and incubated for 0.5 h at 37 °C at pH 3.5. The samples were dried down and run on a SDS-polyacrylamide gel, and the gel was fluorographed (Fig. 2). These data showed no difference between the autocalytic activations of wild type human procathepsin D and mutant human procathepsin D which contains the change of Ile27p.Asp (compare lanes 2 of the wild type transfection and the mutant Ile27p.Asp). The addition of soluble pepstatin to one sample prior to acidification and incubation prevented cleavage and ensured that an aspartic proteinase was responsible for the activation (lane 3 of each transfection).

The results suggested that either the mutation was not sufficient to prevent autoactivation and cleavage still occurred between Leu26p and Asp27p or that procathepsin D was capable of autoactivation at an adjacent site. To distinguish between these two possibilities, the mutant protein was purified by pepstatin affinity chromatography and allowed to autoactivate at pH 3.5. The protein was then run on a SDS-polyacrylamide gel and transblotted to Immobilon P membrane for amino acid sequencing. The sequencing data confirmed that cleavage had occurred between Leu26p and Asp27p. Although cathepsin D is thought to participate in general lysosomal proteolysis and there are few known specific protein substrates for the enzyme, in vitro digesta have been performed on several polypeptides to ascertain cleavage site preferences and to our knowledge this is the first report of cathepsin D cleavage with an aspartic acid residue in the P' site.

**Substitution of Proline at Leu26p Prevents Autoactivation of Procathepsin D**—Studies with a systematic series of synthetic substrates show that human cathepsin D poorly cleaves peptide substrates with a β-branched amino acid in the P<sub>1</sub> position (Scarborough et al., 1991). Two new mutations were created at the P<sub>1</sub> position of the procathepsin D autocalytic cleavage site. One mutation changed Leu26p to valine to introduce a β-branched amino acid at P<sub>1</sub>. The other mutation replaced Leu26p with a proline to alter scissile peptide bond chemistry

---

**Fig. 1. N-terminal sequence of isolated human cathepsin D light chain.** A, human cathepsin D, purified from placenta, was separated on SDS gels and transferred to Immobilon. The light chains were sequenced at the ICBR Protein Core at the University of Florida (Gainesville, FL). The three distinct N termini which were detected (molar ratio in parentheses) in each light chain band are shown in comparison to human cathepsin D (Faust et al., 1985), pseudocathepsin D made by in vitro autoactivation (Conner and Richo, 1992), and porcine cathepsin D (Huang et al., 1979). B, cleavage site mutations. Site-directed mutations at the cleavage sites are shown in boldface type and were confirmed by DNA sequencing. In the case of Ile27p.Asp, the amino acid sequence of the activation intermediate was additionally confirmed by Edman degradation.

---

**Fig. 2. Autoactivation of the Ile27p.Asp mutant.** The cDNA encoding for human cathepsin D was mutagenized to change Ile27p to Asp. Mouse Ltk<sup>−</sup> cells were either mock-transfected (No DNA) or transfected with normal cathepsin D cDNA (wild type) or the mutant cDNA (Ile27p.Asp). The transfected cells were preincubated in the presence of 17 μM chloroquine for 3 h and then labeled continuously with [35S]methionine for 8 h in the presence of chloroquine. Media were collected, acidified to pH 3.5, and incubated at 37 °C to allow autoactivation. Lane 1 of each transfection shows medium prior to acidification. Lane 2 shows medium following acidification and incubation at 37 °C. Lane 3 shows an incubation done in the presence of soluble pepstatin to demonstrate that cleavage was due to an aspartic proteinase. Lane 4 shows the result of incubation at 37 °C without acidification to pH 3.5. The molecular mass standards correspond to: bovine serum albumin, 66 kDa, and ovalbumin, 45 kDa. Despite the mutation at the autocalytic cleavage site (Ile27p > Asp), autoactivation occurred as seen by the shift of procathepsin D (arrow) to pseudocathepsin D (arrowhead). The inability of the mutation to prevent autoactivation showed that 2 hydrophobic residues are not required at the scissile bond of aspartic proteinase substrates.
and to potentially alter the structure within the propeptide with the idea of possibly preventing proper binding within the active site cleft for cleavage (McCaman and Cummings, 1986). The cDNAs of both mutants were sequenced, and Ltk" cells were transfected with the mutant cDNAs and radiolabeled as above. The secreted mutant proteins were examined for autoactivation. Fig. 3 showed that Leu26p.Val autoactivated to pseudocathepsin D but Leu26p.Pro did not. The Leu26p.Pro mutant specifically bound pepstatin agarose eliminating the possibility that the entire protein was folded incorrectly (data not shown).

The Leu26p.Pro Mutant Is Processed in Vivo to Mature Forms of Cathepsin D—The biosynthetic pathway of human cathepsin D is well characterized (Hasilik and Neufeld, 1980; Gieselmann et al., 1983). The co-translational import of human cathepsin D into the lumen of the rough endoplasmic reticulum is accompanied by the proteolytic cleavage of the signal peptide and subsequent oligosaccharide and proteolytic modifications of procathepsin D. During delivery to or upon its arrival in a lysosomal compartment, the proprotein is processed to the active single-chain form of the enzyme. Subsequently, human cathepsin D is processed to the two-chain form of the enzyme by the action of a cysteine proteinase in the lysosome (Gieselmann et al., 1985). In mouse Ltk" cells, the endogenous cathepsin D is processed only to the single-chain form of the enzyme. The cells, however, are able to process human cathepsin D to the two-chain form (Conner et al., 1989). We used the species-specific processing to analyze the processing of human procathepsin D in mouse Ltk" cells.

To determine if pseudocathepsin D is an obligate intermediate in the in vivo processing of procathepsin D, Ltk" cells were transfected with the cDNAs encoding the wild type procathepsin D and the Leu26p.Pro mutant which was unable to autoactivate. The cells were pulse-labeled with \(^{35}\)S)methionine for 0.5 h and chased for 1, 3, and 5 h in the presence of unlabeled methionine. After harvesting, the cells were solubilized, immunoprecipitated with a monoclonal antibody specific for human cathepsin D, and analyzed by SDS-polyacrylamide gel electrophoresis. The results, shown in Fig. 4, demonstrated that the Leu26p.Pro mutant was processed to the single-chain and two-chain form of the enzyme despite the previously observed inability of the protein to autoactivate. These data indicated that activation of procathepsin D to pseudocathepsin D was not a required step for removal of the propeptide or for the maturation of cathepsin D to the single-chain form.

Propeptide Removal Is Not Necessary for Maturation of Human Procathepsin D—It is generally believed that removal of the propeptide is primarily for the activation of the proteinase. It is not known whether propeptide removal is necessary for maturation or delivery of the enzyme to the lysosome. It has also been proposed that propeptide processing may be important in mannose 6-phosphate-independent membrane dissociation (McIntyre and Erickson, 1991). Val42p was mutated to a proline to prevent removal of the propeptide during processing of procathepsin D in vivo. The cDNA encoding for the mutation was sequenced and transfected into Ltk" cells, and the processing of the human enzyme was analyzed using pulse-chase methods and immunoprecipitation as described above. The results are shown in Fig. 5. Comparison between the 1-h and 5-h chase points of the cells transfected with the wild type procathepsin D and cells transfected with the Val42p.Pro mutant clearly demonstrated that the single-chain form of the enzyme lacking the propeptide was not present in cells transfected with the mutant, despite the formation of the two-chain form of the enzyme.

If the propeptide was not removed as is suggested by the absence of the single-chain, then there should have been a concurrent increase in the size of the light chain of cathepsin D. Normally, the light chain of cathepsin D is not visible on the autoradiogram due to the presence of only one methionine and its faster migration rate in 15% SDS-acrylamide gels. The light chain of cathepsin D does include 4 cysteine residues. Transfected Ltk" cells were starved in the absence of methionine and cysteine and labeled with \(^{35}\)S)methionine and \(^{35}\)S)cysteine for 0.5 h for visualization of the light chain. The cells were then chased in the presence of unlabeled methionine and cysteine for 1, 3, and 5 h and harvested for immunoprecipitation with cathepsin D antibodies. The immunoprecipitates were run on a 15% SDS-acrylamide gel and fluorographed. The data in Fig. 6 clearly showed that the light chains of the Val42p.Pro mutant cathepsin D migrated slower than the light chains of wild type cathepsin D indicating an increase in size. These data showed that the mutant protein was being sorted correctly to the lysosome and appeared to be as stable as the wild type protein.

To determine if the Val42p.Pro mutant protein folded incor-
beled methionine for 1 and 5 h (times indicated above the lanes). The samples were pulsed and chased with unlabeled methionine and chased with unlabel methionine for 1 and 5 h (times indicated above the lanes). The samples were immunoprecipitated with anti-cathepsin D antibodies (h). The single-chain form of the enzyme is absent in cells transfected with the Val42p.Pro mutant. The transfected cells were either mock-transfected (No DNA) or transfected with a cDNA encoding normal human procathepsin D (wild type) or the mutant with a Pro substitution at Val42p.Pro (Val42p.Pro). The transfected cells were labeled for 0.5 h with [33S]methionine and [35S]cysteine and chased with unlabeled methionine and cysteine. The samples were immunoprecipitated with antibodies to cathepsin D, and the immunoprecipitates were subjected to autoradiography. The position of the two-light chain bands of different apparent molecular weights is indicated. The molecular mass standards are bovine serum albumin, 66 kDa; ovalbumin, 45 kDa; glyceraldehyde-3-phosphate dehydrogenase, 36 kDa; carbonic anhydrase, 29 kDa; trypsinogen, 24 kDa; and soybean trypsin inhibitor, 20 kDa. The single-chain form of the enzyme is absent in cells transfected with the Val42p.Pro mutant.

FIG. 5. Pulse-chase analysis of the Val42p.Pro mutation. Mouse Ltk-cells were either mock-transfected (No DNA) or transfected with a cDNA encoding normal human procathepsin D (wild type) or the mutant with a Pro substitution at Val42p.Pro (Val42p.Pro). The transfected cells were labeled for 0.5 h with [33S]methionine and [35S]cysteine and chased with unlabeled methionine and cysteine. The samples were immunoprecipitated with anti-cathepsin D antibodies and electrophoresed on 12% SDS-polyacrylamide gels. The positions of procathepsin D (p), single-chain cathepsin D (s), and the heavy chain (h) of the two-chain form of the enzyme are indicated. The molecular mass standards are bovine serum albumin, 66 kDa; ovalbumin, 45 kDa; glyceraldehyde-3-phosphate dehydrogenase, 36 kDa; carbonic anhydrase, 29 kDa; trypsinogen, 24 kDa; and soybean trypsin inhibitor, 20 kDa. The single-chain form of the enzyme is absent in cells transfected with the Val42p.Pro mutant.

FIG. 6. SDS-gel migration of Val42p.Pro light chains. Mouse Ltk-cells were either mock-transfected (No DNA) or transfected with a cDNA encoding normal human procathepsin D (wild type) or the mutant with Pro substituted at Val42p.Pro (Val42p.Pro). The transfected cells were labeled for 0.5 h with [33S]methionine and [35S]cysteine and chased for 1, 3, and 5 h (indicated above the lanes) in medium containing unlabeled methionine and cysteine. The samples were immunoprecipitated with anti-cathepsin D antibodies and electrophoresed on 12% SDS-polyacrylamide gels. The molecular mass standards are bovine serum albumin, 66 kDa; ovalbumin, 45 kDa; glyceraldehyde-3-phosphate dehydrogenase, 36 kDa; carbonic anhydrase, 29 kDa; trypsinogen, 24 kDa; and soybean trypsin inhibitor, 20 kDa. A clear shift in the migration of the light chains of the mutant was seen when compared to the wild type protein indicating that the propeptide had not been cleaved.

rectly because of the low conformational freedom of the prolyl residue, mouse Ltk-cells were transfected with the Val42p.Pro cDNA and continuously labeled with [33S]methionine and [35S]cysteine for 4 h so that all forms of the enzyme would be labeled. The cells were extracted with Triton lysis buffer and subjected to autoactivation at pH 3.5. Fig. 8 showed that the Val42p.Pro mutant was able to autoactivate to pseudocathepsin D. The fact that processing to the single-chain form of the enzyme did not occur and yet there was no accumulation of the pseudocathepsin D intermediate in the pulse-chase analysis strengthens the argument that pseudocathepsin D is not a normal in vivo intermediate of the enzyme.

It was possible that autoactivation of Val42p.Pro occurred and was followed by rapid processing to the two-chain form of the enzyme. In this case, pseudocathepsin D would not be detected with our gel system, and processing to pseudocathepsin D might still occur as a prerequisite to processing to the two-chain molecule. To rule out this possibility, a mutant which possessed both the Leu26p.Pro and the Val42p.Pro mutations was constructed and expressed in Ltk-cells. As expected, this double mutant procathepsin D was not able to autoactivate at pH 3.5 in vitro (Fig. 9). Pulse-chase analysis demonstrated that the double mutant was still processed to the two-chain form of cathepsin D despite the inability to remove any part of the propeptide (Fig. 10).

DISCUSSION

In vivo processing of aspartic proteinases could occur by several mechanisms. Sequence analysis of different forms of cathepsin D (Fig. 1A and Conner and Richo (1992)) indicated that removal of the propeptide could occur by autocatalytic cleavage to pseudocathepsin D, followed by a second autocatalytic cleavage or by another proteinase to remove the remainder of the propeptide. Alternatively, the formation of pseudocathepsin D may only occur in vitro, and conversion of the precursors in vivo may occur as a single step either autocatalytically or by another proteinase. The data presented here strongly suggest that formation of pseudocathepsin D occurs only in vitro and that in vivo the propeptide is removed in a single step. The site of this cleavage is N-terminal of the previously identified cleavage site of the homologous porcine enzyme (Huang et al. 1979; Takahashi and Tang, 1983).

The two light chain bands of different apparent molecular mass, found on SDS gels, had identical N termini, and endoglycosidase H treatment of the bands did not result in a single species suggesting that the difference was not carbohydrate modifications (data not shown). Since 11 amino acids are removed between the light and heavy chains (Faust et al. 1985; Yonezawa et al. 1988), the two light chains probably represent molecules with different C termini. The heterogeneity of the...
Val42p.Pro mutant and continuously labeled for 12 h with [%]methionine and chloroquine as described in Fig. 2. The media were collected, acidified to pH 3.5, and incubated at 37 °C to allow autoactivation. The first lane of each transfection shows medium prior to acidification. Lane 2 is medium following acidification and incubation at 37 °C. Lane 3 of each transfection is an incubation done in the presence of soluble pepstatin. Lane 4 of each transfection shows the result of incubation at 37 °C without acidification to pH 3.5. The arrow indicates the position of procathepsin D, and the arrowhead the position of pseudocathepsin D. The molecular mass standards are as described in the legend of Fig. 2. The Val42p.Pro mutant was able to auto-catalytically activate to the pseudoform despite the absence of processing in vivo to the single-chain enzyme.

**Fig. 8.** Autoactivation of the Val42p.Pro mutant protein. Mouse Ltk" cells were transfected with the cDNA which encodes the Val42p.Pro mutant and continuously labeled for 12 h with [%]methionine and chloroquine as described in Fig. 2. The media were collected, acidified to pH 3.5, and incubated at 37 °C to allow autoactivation. The first lane of each transfection shows medium prior to acidification. Lane 2 is medium following acidification and incubation at 37 °C. Lane 3 of each transfection is an incubation done in the presence of soluble pepstatin. Lane 4 of each transfection shows the result of incubation at 37 °C without acidification to pH 3.5. The arrow indicates the position of procathepsin D, and the arrowhead the position of pseudocathepsin D. The molecular mass standards are as described in the legend of Fig. 2. The Val42p.Pro mutant was able to auto-catalytically activate to the pseudoform despite the absence of processing in vivo to the single-chain enzyme.

**Fig. 9.** Autoactivation of procathepsin D with mutations at both propeptide cleavage sites. Mouse Ltk" cells were transfected with cDNAs carrying the Val42p.Pro mutation of human procathepsin D and either the Pro (Leu26p.Pro) or Val (Leu26p.Val) mutation at Leu26p. The cells were continuously labeled for 12 h with [%]methionine in the presence of chloroquine as described in Fig. 2. The media were collected (lane 1), acidified to pH 3.5, and incubated at 37 °C for 0.5 h to the absence of pepstatin (lane 2) or presence of pepstatin (lane 3). Lane 4 shows a control incubation without acidification. The arrow indicates the position of procathepsin D, and the arrowhead the position of pseudocathepsin D. The molecular mass standards are as described in the legend of Fig. 2. Procathepsin D which contains both the Leu26p.Pro and Val26p.Pro mutations was not capable of autoactivation.

Light chain N termini could be a result of aminopeptidase action in vivo or during purification of the protein. Regardless, the scissile bond did not resemble those thought to be preferred by cathepsin D or other aspartic proteinases (Powers et al., 1977). A comparison of the sequences surrounding the propeptide cleavage sites of four major lysosomal proteinases (cathepsins D, B, L, and H) reveals that in each case the penultimate residue of the mature N terminus is a proline. This suggests that cleavage of these propeptides may be followed by lysosomal aminopeptidase trimming (e.g. lysosomal cathepsin C) which does not normally cleave on the carboxyl side of a proline.

Although it is believed that autoactivation is a normal cleavage event in the stepwise, proteolytic processing of aspartic proteinases to their mature forms because of the existence of in vitro autoactivation products, only the maturation of yeast proteinase A has been studied in vivo. Using pep4 strains and site-directed mutagenesis, two groups have identified a pseudoform of proteinase A, intermediate in size between the proprotein and mature proteinase A (van den Hazel et al., 1992; Woolford et al., 1993). The generation of this form is dependent upon the integrity of the active site and thus is due to in vivo auto-catalytic processing. Interestingly, processing to the mature form of proteinase A still requires the yeast serine proteinase, proteinase B (Woolford et al., 1993). Despite the existence of pseudoproteinase A in vivo, autoactivation does not appear to be obligatory for processing to mature enzyme. Mutations which prevent the formation of pseudoproteinase D (Leu26p.Pro) did not block the in vivo formation of the single-chain or two-chain form of proteinase D indicating that, similar to proteinase A, pseudoproteinase D is not an obligate intermediate in propeptide processing. Mutation of the residue preceding the N terminus of the light chain (Val42p) to a proline blocked formation of single-chain cathepsin D, yet no pseudoform of the enzyme was observed in the cell. This differs from yeast in that pseudoproteinase A was easily detectable in strains unable to remove the proteinase A propeptide (Woolford et al., 1993)). Finally, expression of a procathepsin D mutant which prevented both cleavage events demonstrated that no processing of procathepsin D was required before cleavage to the two-chain enzyme or for transport to the lysosome.

The mutations constructed at the auto-catalytic site suggest that the subsite preferences and requirements for autoactivation of procathepsin D are different from those for cathepsin D cleavage of peptide substrates (Scarborough et al., 1993). These studies of peptide substrates indicate that the presence of a β-branched side chain in the P1 position hinders cleavage (Scarborough et al., 1991) and confirms the previously recognized preference for hydrophobic amino acids in the P1 position. A model based on the structures of homologous regions of other aspartic proteinases predicts that the amino acids which line the S1′ subsite are very hydrophobic (Scarborough et al., 1993). A recent crystal structure of the protein (Baldwin et al., 1993) is in agreement with the predictions of subsite preferences that were based on the modeled structure. The change of Ile27p to an aspartic acid in P1′ did not prevent autoactivation despite the replacement of the hydrophobic residue with a charged residue. Since the autoactivation was done at pH 3.5, we cannot rule out the possibility that the aspartic acid was uncharged and therefore was more easily accommodated in the S1′ subsite. Although the pK_a of the R′ group of the free amino acid is between 3.65 and 3.86, the local environment of the protein may cause a change in this value. Replacement of Leu26p at P1 with the β-branched valine also did not prevent autoactivation of procathepsin D. Studies done with the HIV proteinase showed that peptide substrates which contained either valine or isoleucine in the P1 position were hydrolyzed extremely slowly (Richards et al., 1990). Also, HIV proteinase cleavage of human procathepsin D was significantly inhibited by pepstatin and not by aminopeptidase A (Woolford et al., 1993)
reverse transcriptase-integrase junction of the HIV pol polyprotein was blocked by replacement of Leu at P$_1$ to an Ile (Jupp et al., 1991). The assay we used did not examine the kinetics of the reaction, and it is possible that cleavage occurred at a much slower rate.

Replacement of Leu26p to a proline blocked autocatalytic activity of cathepsin D. This may be due to the inability of cathepsin D to cleave any substrate with a proline in P$_1$ perhaps because of the effects of the proline on local folding or flexibility or perhaps because of the structure of the prolyl peptide bond. Prochymosin has a similar inability to autoactivate at a proline (McCaman and Cummings, 1986). It is important to note that our use of prolines to prevent processing of the propeptide did not affect the formation of the active site as indicated by the ability of the mutant protein to bind pepstatin in a pH-dependent manner. The results obtained with the three autocatalytic cleavage site mutants clearly emphasize the role of secondary specificity in the binding of cathepsin D substrates and the catalytic efficiency of the enzyme (Fruton, 1970; Scarborough et al., 1983).

If a favorable site for cleavage by procathepsin D requires hydrophobic residues in the P$_2$, P$_3$, and P$_1'$ positions as suggested by the studies with peptide substrates, then inspection of the propeptide sequence of procathepsin D reveals no "favorable" sites for cleavage. As further evidence that autoactivation of cathepsin D is different from peptide substrate cleavage, the Leu26p-Val mutant had unfavorable residues in both P$_1$, and P$_2$, and P$_3$ with valine in P$_2$ and aspartic acid in P$_1'$. It is possible that the very high local concentration of the propeptide is responsible for some of these differences. These differences in specificity for peptide substrates and autoactivation sites are also apparent in studies of pepstatin (Pohl and Dunn, 1988). The data presented in this report suggest that the sequence specificity of cathepsin D for autocatalytic cleavage was broader than previously assumed and that peptide bond structure or its position relative to the active site may be of greater importance in the autocatalytic process than the primary structure of the propeptide.

Acknowledgments—We thank Drs. Ben Dunn and Paula Scarborough for helpful discussions, Dr. Steven Weitz for human specific anti-cathepsin D monoclonal antibodies, and Dr. Keith Brew for critical reading of the manuscript.

REFERENCES

Guagliardi, L. E., Koppelmann, B., Blum, J. S., Marks, M. S., Cresswell, P., and Brodsky, F. M. (1980) Nature 343, 133–139