Self-activation of Recombinant Human Lysosomal Procathepsin D at a Newly Engineered Cleavage Junction, “Short” Pseudocathepsin D*

(Received for publication, March 1, 1996)

Brian M. Beyer and Ben M. Dunn‡
From the Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, Florida 32610-0245

To obtain a recombinant model of human cathepsin D with kinetic properties that are identical with native human liver enzyme, we have addressed the significant differences in structure and catalytic function between naturally occurring enzyme and bacterially derived pseudocathepsin D. Human procathepsin D was expressed in a baculovirus system to obtain correctly folded, glycosylated enzyme that upon acidification completely converts to the active intermediate, pseudocathepsin D. The oligosaccharide moieties of this recombinant enzyme contributed to about 5% of the apparent molecular mass of the enzyme, and the carbohydrate composition was quite similar to the native material. However, specificity constants ($k_{cat}/K_{m}$) of this glycosylated pseudoform for several synthetic chromogenic substrates were considerably less (33%-50%) than those for the native enzyme and were virtually identical with those observed with nonglycosylated pseudocathepsin D.

A cleavable junction suitable for self-processing at the normal maturation point of human cathepsin D was engineered into procathepsin D according to known specificity requirements of this enzyme, and the construct was expressed using baculovirus. Following experiments that demonstrated that the new proenzyme failed to process to the expected point, the new cleavage junction was moved 6 residues toward the amino terminus of procathepsin D and expressed in Escherichia coli. After refolding, the protein containing the newly engineered junction self-processed, generating a shortened mutant form of pseudocathepsin D that is 6 residues longer at the amino terminus than the native material. The kinetic properties of this newly engineered pseudoform proved to be identical with those of the native enzyme, thus establishing an improved recombinant model for this important aspartic protease.

The major intracellular aspartic protease cathepsin D (EC 3.4.23.5) is normally found in the endosomes and lysosomes of higher eukaryotes (1). This abundant endoprotease comprises approximately 10% of total protein in the lysosome and is expressed in all mammalian tissues with the exception of erythrocytes (2). In the lysosome, cathepsin D has been shown to be involved in protein turnover and has been postulated to play a role in antigen and prohormone processing (3, 4). Interest in cathepsin D as a target for drug design results from its correlation with several biological processes of pathological significance. Cathepsin D overexpression and activity have been implicated with breast cancer progression (5–7) and the release of β-amyloid peptides from the amyloid precursor in Alzheimer’s disease (8, 9).

Cathepsin D is synthesized on the rough endoplasmic reticulum as a preproenzyme that undergoes several proteolytic cleavages during biosynthesis to produce the mature form (10). Following the initial co-translational removal of the signal peptide to yield procathepsin D, sugars are attached at two N-linked glycosylation sites and the proenzyme is transported to the Golgi stacks. Procathepsin D is marked for the mannose 6-phosphate receptor and targeted to lysosomes (11–13), where 44 amino acid residues are ultimately removed from the amino terminus, yielding an active single-chain molecule with an apparent molecular mass of 44 kDa. This mature proteolytic activation event depends, most likely, on the action of thiol lysosomal proteinases (14). This single-chain species is eventually cleaved in the lysosome into a two-chain enzyme consisting of a light (15-kDa) amino-terminal domain and a heavy (30-kDa) carboxyl-terminal domain. Accompanying the conversion to the two-chain species, 7 amino acid residues between the light and heavy chains are removed (15). Furthermore, several more amino acids are also removed from the carboxyl terminus of the heavy chain (16).

Given the patho- and normal physiological implications, the availability of potent and specific inhibitors of cathepsin D becomes important for the further elucidation of the roles played by this enzyme in human disease. In order to facilitate this development, we are working to define the subtle differences in ligand/enzyme interactions that exist between cathepsin D and other aspartic proteinases. Previously, in order to examine active site requirements, human fibroblast procathepsin D has been overexpressed in Escherichia coli (17), refolded from solubilized inclusion bodies (18), and purified using pepsatin affinity chromatography (19). The resulting recombinant protein is active, nonglycosylated, not fully processed to the mature amino terminus, and retains 18 residues (27p-44p) of the pro-segment.

The catalytically active processing intermediate, pseudocathepsin D, has been used in order to study the structural, enzymatic, and biosynthetic intermediates of human cathepsin D (19, 20). This enzyme has been reported to be enzymatically and structurally similar to tissue-derived mature enzyme in many respects. However, one property of this recombinant protein that has plagued investigators is the approximately 3-fold lower specific activity (21) observed in the kinetic parameters for recombinant pseudocathepsin D and native enzyme. Several reasons may account for the differences in kinetic properties observed for this nonglycosylated, partially processed species and are the subject of this investigation.

With observed differences in Michaelis constants, $K_{cat}$ values, and the resulting specificity constants ($k_{cat}/K_{m}$), our present studies were undertaken to generate a model for human cathepsin D that better represents the native enzyme in activ

Procathepsin D Self-activation to “Short” Pseudocathepsin D

ity and specificity. To this point, we have expressed cathepsin D in an eukaryotic expression system to assess the effect of glycosylation on these parameters. Furthermore, to determine the effect of the partially retained pro-segment, we have utilized known specificity requirements of human cathepsin D to engineer a cleavage junction that effectively removes 12 of the 18 amino acids in pseudocathepsin D. Our data indicate that the retained pro-segment, rather than the lack of glycosylation of recombinant human cathepsin D, is responsible for the differences in kinetic properties. The results presented in this paper detail the refinement of the recombinant model of human cathepsin D.

MATERIALS AND METHODS

Enzymes and Reagents—Enzymes for molecular biology were purchased from the Promega Corp. and used according to the manufacturer’s instructions. Sequencing radiochemicals were obtained from DuPont NEN. Sequencing kit was purchased from United States Biochemical Corp. Life Technologies Inc. provided tissue culture reagents, pepstatin agarose was from Sigma, and other biochemicals came from Fisher. The synthetic disaccharides were synthesized by the University of Florida, Interdisciplinary Center for Biotechnology Research (UF-ICBR) DNA Synthesis Core facility using an Applied Biosystems 421A DNA synthesizer. The disaccharides were synthesized by the solid phase method using an Applied Biosystems Model 430A in the UF-ICBR Protein Chemistry Core Facility. Native human liver cathepsin D was purchased from ART Biochemicals.

Expression, Refolding, and Purification of Nonglycosylated Pseudocathepsin D—Human fibroblast procathepsin D was expressed in E. coli using a derivative of the pET3a vector (pETPSD2) (17). The protein was isolated in the form of insoluble inclusion bodies as described previously (19). To optimize the recovery of active enzyme, previously reported procedures (19) were modified as follows. Inclusion bodies were solubilized and reduced at a concentration of 2 mg/ml in 50 mM CAPS, pH 10.7, 50 mM β-mercaptoethanol, and 8 M urea. Solubilized inclusions were incubated for 30 min at room temperature, and the sample was diluted 1:10 into a 50 mM sodium phosphate, pH 7.0, buffer. After addition of 1 mg/ml of dithiothreitol and 20 mM cysteine were then added at final concentrations of 1 mM, 0.1 mM, and 1 mM, respectively. After more hours of slow stirring at room temperature, the mixture was then dialyzed to pH 3.7 with 1 M sodium formate, and aliquots were taken and assayed every 8 h to monitor the emergence of hydrolytic activity. Immediately prior to pepstatin agarose chromatography, 0.1 volume of 1 M sodium formate, pH 3.5, 4 M NaCl, and 1% Brij 35 was added to the refolding solution, and the resulting mixture was applied to the affinity matrix column at 4°C. The column was washed with 10 column volumes (100 ml) of 10 mM sodium formate, pH 3.5, 1% NaCl, and 0.1% Brij 35, then eluted with 20 ml Tris-HCl, pH 8.1 at 4°C, 0.4 M NaCl, and 0.5% Brij 35. Column fractions exhibiting catalytic activity were pooled after analysis and stored in 10% glycerol at −70°C.

Expression and Purification of Glycosylated Procathepsin D—The baculovirus transfer vector pVL1392 encoding human procathepsin D (pVL1392CD), a generous gift from Dr. Gregery E. Conner (University of Miami), was co-transfected into Spodoptera frugiperda Sf9 cells. Following co-transfection of Sf9 cells with cesium chloride purified pVL1392CD and BaculGold® DNA, supernatants of the co-transfected cells were collected and spun at 600 × g for 10 min to remove unattached cells. Serial dilutions of the transfected supernatants (10−4, 10−2, and 10−1) were made, and plaque assays were then performed. A number of occlusion negative plaques were selected and expressed using a Pasteur pipette. Viral particles were eluted from the agarose plugs by incubating in Grace’s media supplemented with 10% fetal calf serum, and the resulting supernatants were amplified. Following three rounds of viral amplification, plaque assays were performed at dilutions of 10−2, 10−4, and 10−6 to determine the titer of the amplification solutions. Sf9 cells were then grown to a density of 2 × 107 cells/ml in a spinner flask and infected with recombinant virus at a multiplicity of infection of 10. Following infection, cells were incubated at 27°C for 72 h, pelleted at 600 × g, and the virus was removed by ultracentrifugation at 50,000 × g for 1 h. Glycosylated procathepsin D, secreted into the expression media, was purified using pepstatin aagarose chromatography and autoaffixed to pseudocathepsin D via acidification by the addition of 0.1 volume of 1 M sodium formate, pH 3.7, and incubated at 37°C for 30 min.

Introduction of Cleavage Junctions by PCR—in order to remove the extra two acids retained in pseudocathepsin D, site-directed mutagenesis by PCR (23) using Taq Polymerase was employed with slight modifications (20). This procedure was utilized for the introduction of a new cleavage junction based on the amino acid sequence of the superior cathepsin D substrate Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu (* = site of cleavage, Nph = p-nitrophenylalanine) (21). Table I presents the primers used for introduction of codon changes and the resulting changes in the primary amino acid sequence. The PCR product for the mutagenized mature chain was cloned into the baculovirus transfer vector pVL1392CD, while the product for the “short” pseudoform mutant (cleavage junction moved 6 residues from the native maturation point toward the amino terminus of procathepsin D) was cloned into the baculovirus transfer vector pVL231CD. The vectors were cotransfected and expressed in Sf9 cells. The coding region was dideoxy-sequenced to confirm the presence of the mutations and to ensure no misincorporations from Taq polymerase. Following sequence confirmation, recombinant mutant proteins were expressed using the respective protocols for each vector.

Characterization of Expression Products—After pooling of active fractions from affinity chromatography, wild type and mutant samples were subjected to SDS-PAGE on 10% Tris-Tricine-SDS gels and stained with Coomassie or silver stains to assess purity. To determine molecular masses, samples were first mixed 1:1 with a saturated solution of 3,5-dimethoxy-4-hydroxycinnamic acid in 70% acetonitrile and then analyzed using a Perseptive Biosystems Voyager™ RP mass spectrometer ( Voyager™ MALDI-TOF). Amino-terminal sequence analysis was performed on procathepsin D and activated samples to determine the extent of processing during self-activation. Products were subjected to SDS-PAGE, transferred to Immobilon™-P, and amino-terminal residues were identified using an Applied Biosystems 470A Protein Sequencer using gas-phase chemistry by the Protein Chemistry Core Laboratory at the University of Florida.

Peptide Substrates and Native Enzyme—Stock peptide solutions were made in filtered distilled water and quantified by amino acid analysis. The purity of the peptides (≥ 90%) was verified by high pressure liquid chromatography and MALDI-TOF analysis. Native human liver cathepsin D was dissolved in ice-cold distilled water, and aliquots were stored in 10% glycerol at −70°C for future use. The kinetic parameters for cathepsin D, Kcat and kcat/Km, were determined by monitoring the decrease in the average absorbance from 284 to 324 nm upon substrate hydrolysis between PhεNph. This assay has been described extensively (21, 24), and values were determined in 0.1 M sodium formate, pH 3.7 at 37°C. The amount of active enzyme was determined by competitive titration with the bacterial fluorogenic substrate. N-Ac-NAHPPA, Nva-NAHPPA, and Nva-Naphthylene maleimide (gift from Kwan Hui, Lilly Research Laboratories) under assay conditions of 2% dimethyl sulfoxide. The resulting curve was fitted with the Henderson equation on the Enzfitter™ program (26).

Modeling of Recombinant Pseudocathepsin D Species—To aid in the interpretation of the data previously presented in the kinetic data between pseudocathepsin D and short pseudocathepsin D, a model was constructed using the high resolution three-dimensional structure of human liver cathepsin D complexed with pepstatin (27). The BUILDER command encoded in the BIOPOLYMER suite of SYBYL (supplied by

1 The abbreviations used are: CAPS, 3-(cyclohexylamino)propanesulfonic acid; Tricine, N-tris(hydroxymethyl)methylglycine; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; Ac, acetyl; NA, naphthylalanine; Nva, norvaline; AHPPA, aminohydroxyphe-nylpropanionic acid; WT, wild type.
Formation of Proteolytic Activity from Refolding Procathepsin D—When human procathepsin D was expressed in E. coli, inclusion bodies were purified, solubilized, and procathepsin D was refolded using the protocols described. Protodetlytic activity appeared 8 h after acidification of the refolding mixture, and pseudocathepsin D was purified by affinity chromatography when maximal activity was reached (7 days after acidification). A protein concentration of the pooled active fractions was conducted to determine the final amount of correctly refolded, autoactivated pseudocathepsin D recovered following refolding.

Since pepstatin binds at pH 3.5 to the active site of both correctly folded procathepsin D and processed forms of cathepsin D (19), the absence of unactivated yet correctly folded procathepsin D in the column-eluted fractions indicates that the slow emergence of proteolytic activity is due to slow or defective refolding. This observation is in agreement with previous reports (19, 20), and it appeared that only approximately 1.1% of the procathepsin D was refolded during the 7-day course of the experiment.

Characterization and Glycosylation of Procathepsin D from Baculovirus—Approximately 1 mg of correctly folded, glycosylated inactive zymogen was recovered from a 1-liter expression culture utilizing affinity chromatography. Samples of glycosylated procathepsin D were then subjected to SDS-PAGE and mass spectrometry to assess the contributions of glycans to molecular mass. The apparent mass of procathepsin D from E. coli expression was 42.9 kDa and from baculovirus expression was 44.8 kDa; therefore, glycans accounted for 1.9 kDa or approximately 4.2% of the overall molecular mass. Glycan hydrazinolysis of this enzyme and native human liver cathepsin D determined that the glycans for both of these species were homogeneous. Monosaccharide composition analysis revealed that the glycans for both the recombinant species and native enzyme were similar in content. Approximately 5 mol of GlcNAc and 10 mol of Man were observed per mol of protein for both species. Furthermore, additional attachments were contributed by the S9f cell, and approximately 3 mol of GalNAc and 1 mol of Fuc were present in the N-linked glycans from the baculovirus expressed protein.

Upon autoactivation, the first 26 residues were removed yielding an amino terminus equivalent to pseudocathepsin D. The activation proceeded quite rapidly, and total conversion to the active intermediate was evident within 15 min from SDS-PAGE analysis.

Role of Glycosylation in Activity and Specificity—Since the baculovirus intermediate has an amino terminus identical with pseudocathepsin D refolded from bacterial inclusion bodies and glycan composition appears to be quite similar to the native enzyme, the role of glycosylation in activity and specificity was investigated. Table I compares the kinetic parameters of several synthetic chromogenic octapeptides with native human liver cathepsin D and the two intermediate species. The kinetic parameters indicate that several of these peptides are excellent substrates for cathepsin D; however, the $k_{cat}$ values were consistently lower (7%-56%) for the two intermediate forms. Furthermore, the Michaelis constants are slightly different resulting in consistently low specificity constants ($k_{cat}/K_m$) that are 33-64% less than those for the native enzyme. Cathepsin D does not cleave peptides with a positive charge in the S2 subsite (20), and the last peptide was therefore used as a negative control.

Activation and Characterization of Cleavage Junction Mutants—Table I details the amino acid substitutions for the introduction of a new cleavage junction at the native maturation point of human cathepsin D. The maturation mutant was expressed using the baculovirus system, purified in the inactive zymogen form, and autoactivated by acidification. Mass spectrometry and SDS-PAGE indicated that the molecular mass of the activated form was consistent with bacterially derived pseudocathepsin D (taking glycan contributions into consideration). Amino-terminal sequencing was performed and confirmed that only the first 26 amino acids were removed during self-activation yielding an amino terminus equivalent to pseudocathepsin D. Ionic strength, pH, incubation times, and temperatures were varied to promote further processing to no avail.

The short pseudoform mutant was refolded from bacterial inclusion bodies, and the emergence of proteolytic activity was monitored in comparison with wild type (WT) pseudocathepsin D at identical protein concentrations. Fig. 1 indicates that the initial emergence of activity from the short pseudocathepsin D mutant was approximately 10-fold that of the WT pseudocathersin. Activity peaked at 72 h following acidification and was roughly equal at that point for the two enzymes. Activity began to decrease following day 8; therefore, enzymes were purified in the activated state on day 7 by affinity chromatography. Samples of this mutant pseudoform were compared with WT pseudocathepsin D by mass spectrometry. Fig. 2 indicates that the apparent molecular mass of WT pseudocathepsin D was 39.9 kDa, and the mass for the short mutant pseudoform was 38.6 kDa. The approximate 1.3-kDa difference between the two pseudoforms corresponds to the anticipated mass difference resulting from further processing and the subsequent removal of the additional 12 residues. Amino-terminal sequence analysis gave the sequence listed in Fig. 3 and is indicative of peptide bond hydrolysis between the two phenylalanines of the newly engineered cleavage junction. Consequently, 12 of the 18 amino acids retained in pseudocathepsin D have been effectively removed, yielding the species that we have designated short pseudocathepsin D.

Kinetic parameters for this new mutant pseudoform of human cathepsin D were determined using the same peptides as described before. Table III reports these values for native human liver cathepsin D and compares them with those for the short pseudoform. The analyses with native enzyme were real.
determined simultaneously with those for the short pseudo-form. Therefore, some differences can be noted between the values for the native enzyme in Tables II and III. In all cases, values of $k_{\text{cat}}/K_{\text{m}}$ are identical within error limits when comparing native enzyme in Tables II and III. The kinetic parameters indicate that some unique feature of the 12 removed residues was responsible for the apparent differences in specific activity and apparent binding affinity between native human liver cathepsin D and WT pseudocathepsin D.

Models of Pseudocathepsin D and Short Pseudocathepsin D—Fig. 4 illustrates a superposition of the models generated for the two pseudoforms investigated in this study. The conformations displayed here are energetically favorable, as determined from energy minimization calculations, and clearly indicates that the remaining 18 residues of the pro-segment in pseudocathepsin D have the potential to interact with the active site. This retained region is displayed in white and shown entering the active site from the non-prime side. A potential electrostatic interaction between the side chain of Lys31p (cyan) and the side chains of the catalytic aspartates (red/white) appears quite favorable. On the other hand, the modeling exercise revealed that the 6 residues retained in short pseudocathepsin D (yellow) do not form an extension long enough to interact with the active site.

**DISCUSSION**

The primary function of attached oligosaccharides in human cathepsin D has been attributed to their role in phosphotransferase recognition and targeting of cathepsin D to lysosomes. Alternatively, the presence of N-linked oligosaccharide chains at Asn-70 and Asn-199 in native enzyme may reflect the contribution of glycosylation to proteolytic activity or enzyme stability. Evidence has been presented (28) that glycosylation may enhance the proteolytic activity of a related aspartic proteinase, Mucor rennin. Furthermore, the absence of glycans in bacterially derived pseudocathepsin D could possibly contribute to the sensitivity of the recombinant enzyme to autoproteolysis. A total disappearance of recombinant cathepsin D has been observed during attempts to process to fully mature enzyme.

With these observations in mind, we successfully expressed preprocathepsin D using a baculovirus expression system and obtained recombinant enzyme with glycans similar to those of native human liver cathepsin D. Oligosaccharides in N-glycosidic attachments usually have a distinctive core consisting of 2 mol of GlcNAc and 3 mol of Man with the peripheral Man residues linked to either Man or GlcNAc. The latter residues may, in turn, be linked to yet other sugar residues creating diversity. Our data suggested that with approximately 5 mol of GlcNAc and 10 mol of Man observed per mol of protein for the recombinant and native species, the core structure of these glycans should be virtually identical. The small amounts of Fuc and GalNAc, in turn, should lie in the periphery of the glycans of this recombinant species.

Table II was constructed to assess the role of glycosylation in activity and specificity by comparing the kinetic parameters of seven octapeptides with native enzyme and the two recombinant pseudoforms. The results indicate that, although the glycan composition of the baculovirus recombinant species is quite similar to native enzyme, $k_{\text{cat}}$ values are consistently lower (15–47%) and the Michaelis constants are somewhat different in comparison with native material. Furthermore, the resulting parameters for this glycosylated pseudoform are virtually identical with the nonglycosylated bacterially expressed enzyme. As glycosylation is the only difference between these two recombinant species, the data appear to rule out the role of
glycosylation in specific activity and apparent binding affinity for this enzyme.

A more likely reason for the kinetic differences seen between the recombinant proteins and the native enzyme is the presence of the retained portion of the prosequence. Within these 18 residues, a highly conserved dipeptide exists (Lys36p-Tyr37p), that has been observed in x-ray crystal structures of porcine pepsinogen (29) and human progastricsin (30) to interact electrostatically with the catalytic apparatus. This interaction is thought to contribute to the anchoring of the pro-region into the

| P₅ | P₄ | P₃ | P₂ | P₁ | P₁' | P₂' | P₃' | Human liver cathepsin D | Bacterial pseudocathepsin D | Baculovirus pseudocathepsin D |
|----|----|----|----|----|----|----|----|-----------------|-----------------|-----------------|
|    |    |    |    |    |    |    |     | kₗₜₐ | kₘ | kₗₜₐ/Kₘ | kₗₜₐ | kₘ | kₗₜₐ/Kₘ | kₗₜₐ | kₘ | kₗₜₐ/Kₘ |
|    |    |    |    |    |    |    |     | s⁻¹ | μM⁻¹ | s⁻¹ | μM⁻¹ | s⁻¹ | μM⁻¹ | s⁻¹ | μM⁻¹ | s⁻¹ | μM⁻¹ |
| P P I E F X R L | 39 ± 4 | 26 ± 2 | 1.5 ± 0.2 | 33 ± 5 | 46 ± 8 | 0.7 ± 0.1 | 25 ± 4 | 33 ± 5 | 0.8 ± 0.1 |
| P P F E F X R L | 33 ± 6 | 23 ± 2 | 1.4 ± 0.2 | 25 ± 4 | 49 ± 7 | 0.5 ± 0.1 | 21 ± 3 | 32 ± 5 | 0.7 ± 0.1 |
| P P V E F X R L | 62 ± 7 | 68 ± 8 | 0.9 ± 0.1 | 40 ± 4 | 80 ± 9 | 0.5 ± 0.1 | 38 ± 4 | 59 ± 8 | 0.6 ± 0.1 |
| P P I E F X S L | 20 ± 4 | 24 ± 2 | 0.8 ± 0.1 | 11 ± 1 | 27 ± 3 | 0.4 ± 0.1 | 10 ± 1 | 19 ± 2 | 0.5 ± 0.1 |
| P P I A F X R L | 27 ± 5 | 76 ± 9 | 0.4 ± 0.1 | 15 ± 1 | 82 ± 9 | 0.2 ± 0.1 | 12 ± 1 | 69 ± 9 | 0.2 ± 0.1 |
| P P I S F X R L | 15 ± 3 | 42 ± 7 | 0.4 ± 0.1 | 8 ± 1 | 51 ± 7 | 0.2 ± 0.1 | 14 ± 1 | 72 ± 8 | 0.2 ± 0.1 |
| P P I K F X R L | —    | —    | <0.01  | —    | —    | <0.01  | —    | —    | <0.01  | —    | —    | <0.01  |

| P₅ | P₄ | P₃ | P₂ | P₁ | P₁' | P₂' | P₃' | Human liver cathepsin D | "Short" pseudocathepsin D |
|----|----|----|----|----|----|----|----|-----------------|-----------------|
|    |    |    |    |    |    |    |     | kₗₜₐ | kₘ | kₗₜₐ/Kₘ | kₗₜₐ | kₘ | kₗₜₐ/Kₘ |
|    |    |    |    |    |    |    |     | s⁻¹ | μM⁻¹ | s⁻¹ | μM⁻¹ | s⁻¹ | μM⁻¹ | s⁻¹ | μM⁻¹ |
| P P I E F X R L | 45 ± 5 | 29 ± 3 | 1.6 ± 0.1 | 40 ± 4 | 27 ± 3 | 1.5 ± 0.1 |
| P P F E F X R L | 29 ± 4 | 24 ± 4 | 1.2 ± 0.1 | 30 ± 5 | 22 ± 3 | 1.4 ± 0.1 |
| P P V E F X R L | 70 ± 6 | 64 ± 5 | 1.1 ± 0.1 | 68 ± 5 | 69 ± 6 | 1.0 ± 0.1 |
| P P I E F X S L | 48 ± 5 | 50 ± 8 | 1.0 ± 0.1 | 52 ± 5 | 55 ± 7 | 1.0 ± 0.1 |
| P P I A F X R L | 18 ± 3 | 47 ± 9 | 0.4 ± 0.1 | 20 ± 4 | 44 ± 6 | 0.5 ± 0.1 |
| P P I S F X R L | 22 ± 3 | 56 ± 4 | 0.4 ± 0.1 | 19 ± 3 | 52 ± 4 | 0.4 ± 0.1 |
| P P I K F X R L | —    | —    | <0.01  | —    | —    | <0.01  | —    | —    | <0.01  |

FIG. 4. Superposition of models generated for pseudocathepsin D and short pseudocathepsin D. Ribbon trace of common structures is displayed in purple, and extra residues of the retained pro-segment of pseudocathepsin D (white tube) and those of short pseudocathepsin D (yellow tube) are shown. A potential favorable electrostatic interaction between the side chains of the catalytic aspartates (white/red) and Lys33p (cyan) is shown as well.
active site rendering the enzyme inactive. The presence of this sequence (Lys-Tyr) and another positive charge contributed by Lys31p, in pseudocathepsin D, may promote competition at the active site between exogenous substrate and the remaining pro-peptide.

Attempts have been made previously to fully process pseudocathepsin D autocatalytically to its mature amino terminus but have so far been unsuccessful. This is probably due to the lack of a suitable cleavage site closer to the maturation point of this enzyme. Furthermore, other approaches have attempted to express cathepsin D without the propeptide. Conner (31) reported that an imperfect deletion of the propeptide resulted in the abolishment of stable expression and isolation of the transfected protein in mammalian cells. This was attributed to the suspected role of the propeptide in folding of the enzyme. In contrast to this report, a precise deletion created by Frotenberry and Chirgwin (32) resulted in correctly folded mature cathepsin D with yields, however, that are inappropriate for structure/function studies. With these observations and the documented importance of pro-parts in refolding of aspartic proteinases (33) in mind, the introduction of a new cleavage junction at the maturation point of the native enzyme appeared necessary and promising.

The data indicated that the maturation mutant was expressed and correctly folded, as assessed by affinity chromatography, at levels identical with the WT form. The autocatalytic activation of this mutant to a WT pseudocathepsin D amino terminus was rapid and terminated the processing event. Molecular modeling studies utilizing the crystal structure of porcine pepsinogen indicated that the engineered cleavage junction should be accessible to solvent and could have been readily cleaved through a bimolecular interaction at the concentrations studied. Surprisingly, further processing was not observed and may reflect a unique condition found in the lysosomal compartment that contributes to accessibility of this region and was not reproduced in our in vitro conditions.

Because other aspartic proteinases self-process to short pseudofoms (34, 35), the potential cleavage junction was moved six amino acids from the maturation point toward the amino terminus of procathepsin D. For rapid screening of this mutant, and because experiments described above demonstrated the non-importance of carbohydrate attachment, expression was done using the E. coli system. In Fig. 1, the initial enhancement of catalytic activity following refolding was encouraging; however, this could have meant one of two things. Since the propeptides of aspartic proteinases have been described as intramolecular chaperones (36), it was thought that the mutations in this region could have altered refolding efficiency. The corresponding increase in catalytic activity could therefore be due to the presence of more correctly folded autocatalytically recombinant enzyme. Alternatively, if the new cleavage junction was a superior substrate for autocatalysis, the initial amount of catalytic activity observed would be greater for the mutant form than for WT. However, the amount of final activity observed in the refolding solution should be approximately equal if the overall refolding efficiency has not been affected. The data support the latter hypothesis as overall activity in refolding solutions of wild type sequence and the mutant sequence is approximately equal from day 6 on. Activity decreases after day 8 and reflects the similarities in stability between the two enzymes.

Mass spectrometry clearly resolved the expected 1.3-kDa difference, and successful processing to the cleavage junction was confirmed by amino-terminal sequencing. Activation appears to occur through an intramolecular mechanism since the refolding/activation concentration was 20 μg/mL (34). To assess the role of these otherwise retained 12 amino acids, kinetic analysis of this new recombinant species, short pseudocathepsin D, was done (Table I) using the same chromogenic octapeptide.

The data indicate that a feature of this normally retained propeptide region was responsible for the differences previously observed in specific activity and Michaelis constants.

Consequently, models of pseudocathepsin D and short pseudocathepsin D were constructed to assess the feasibility of competition at the active site between exogenous substrate and the remaining pro-segments. Ribbon traces of these two diagrams were superimposed, and the resulting structure is illustrated in Fig. 4. The extra 18 residues of pseudocathepsin D form an extension long enough to wrap around the front of this protease penetrating well into the active site. The side chain of Lys31p, in the longer pseudoform, is shown to be in close proximity with the two catalytic aspartates creating a favorable interaction. Furthermore, the anchoring dipeptide Lys36p-Tyr37p (not shown) was found to lie within the periphery of the active site and may promote other interactions between the active site and the retained pro-segment of the longer pseudoform. Short pseudocathepsin D, consisting of only 6 additional residues, is not long enough to come in close proximity with the active site rendering this extension apparently harmless.

In spite of the remaining structural difference of a single chain versus two-chain enzyme, the present results provide an improved recombinant model for human cathepsin D with kinetic properties virtually identical with the native enzyme. In order to facilitate inhibitor development, future work will utilize this model to define subtle differences in ligand-enzyme interactions that exist between cathepsin D and other aspartic proteinases. Furthermore, the approach detailed here could be used to activate other inactive zymogens whose activation conditions have not yet been determined.

Acknowledgments—We thank Dr. Gregory E. Conner and Yunxiang Zhu for generously supplying the baculovirus transfer vector encoding human proprocathepsin D (pVL1392CD) and for their helpful discussions. We also gratefully acknowledge the critical assistance of the Uf-IFCR Protein Chemistry and the Glycobiology Core Facilities.

REFERENCES

1. Barrett, A. J. (ed) (1977) in Proteins in Mammalian Cells and Tissues, pp. 209–248, North Holland Publishing Co., New York
2. Dean, R. T., and Barrett, A. J. (1976) Essays Biochem. 1, 39
3. Pillai, S., and Zulli, J. (1986) J. Biol. Chem. 261, 14919–14923
4. Diment, S., Leech, M., and Stahl, P. (1988) J. Biol. Chem. 263, 6901–6907
5. Rochefort, H. (1990) Science 250, 153–160
6. Sanchez, L. M., Ferrando, A. A., Diez-Itza, I., Vizoso, F., and Lopez-Otin, C. (1993) Br. J. Cancer 67, 1076–1081
7. Acs, S., Lippuner, P., Kosma, M., Aaltonaa, S., and Syrjanen, K. (1995) Anticancer Res. 15, 1033–1038
8. Siman, R., Mistretta, J. T., Durkin, M. J., Savage, T., Loh, S., Trusko, S., and Scott, R. W. (1993) J. Biol. Chem. 268, 16602–16609
9. Ladrir, U., Snyder, S., Wang, G., Hidamna, T., and Krafft, G. (1994) J. Biol. Chem. 269, 18422–18428
10. Conner, G. E., Bloomberg, G., and Erickson, A. H. (1987) in Uf-IFCR Protein Chemistry and the Glycobiology Core Facilities.
24. Dunn, B. M., Scarborough, P. E., Davenport, R., and Swietnicki, W. (1994) in Peptide Analysis Protocols (Pennington, M., and Dunn, B. M., eds) pp. 225–243, Humana Press Inc., Totowa, NJ

25. Hui, K. Y., Hermann, R. B., Manetta, J. V., Gygi, T., and Angleton, E. L. (1993) FEBS Lett. 327, 355–360

26. Henderson, P. (1972) J. Biol. Chem. 247, 321–333

27. Baldwin, E. T., Bhat, T. N., Gulnik, S., Hosur, M. V., Sowder, R. C., Cachau, R. E., Collins, J., Silva, A. M., and Erickson, J. W. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6796–6800

28. Aikawa, J., Yamashita, T., Nishiyama, M., Horinouchi, S., and Beppu, T. (1990) J. Biol. Chem. 265, 13955–13959

29. James, M. G., and Sielecki, A. (1986) Nature 319, 33–38

30. James, M., Moore, S., Sielecki, A., Chernaia, M., and Tarasova, N. (1995) in Aspartic Proteases: Structure, Function, Biology, and Biomedical Implications (Takahashi, K., ed) Vol. 362, pp. 11–18, Plenum Press, New York

31. Conner, G. E. (1992) J. Biol. Chem. 267, 21738–21745

32. Fotenberry, S., and Chirgwin, J. (1995) J. Biol. Chem. 270, 9778–9782

33. Koelsch, G., Mares, M., Metcalf, P., and Fused, M. (1994) FEBS Lett. 343, 6–10

34. Marciniszyn, J., Jr., Huang, J. S., Hartsuck, J. A., and Tang, J. (1976) J. Biol. Chem. 251, 7095–7102

35. Chen, Z., Koelsch, G., Han, H., Wang, X. J., Lin, X., Hartsuck, J. A., and Tang, J. (1991) J. Biol. Chem. 266, 11718–11725

36. Fukuda, R., Horiiuchi, H., Ohta, A., and Takag, M. (1994) J. Biol. Chem. 269, 9556–9561
Self-activation of Recombinant Human Lysosomal Procathepsin D at a Newly Engineered Cleavage Junction, "Short" Pseudocathepsin D
Brian M. Beyer and Ben M. Dunn

J. Biol. Chem. 1996, 271:15590-15596.
doi: 10.1074/jbc.271.26.15590

Access the most updated version of this article at http://www.jbc.org/content/271/26/15590

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 31 references, 15 of which can be accessed free at http://www.jbc.org/content/271/26/15590.full.html#ref-list-1