

Isolation and Characterization of Human Cathepsin V: A Major Proteinase in Corneal Epithelium

Wakako Adachi,^{1,2} Shoko Kawamoto,¹ Ikko Ohno,¹ Kohji Nishida,² Shigeru Kinoshita,² Kenichi Matsubara,¹ and Kousaku Okubo¹

PURPOSE. To isolate and characterize a novel cathepsin gene, as part of the systematic isolation of genes uniquely active in corneal epithelium.

METHODS. For the isolation of a full-length cDNA clone, a probe was selected from a set of expressed sequence tag clones classified as unique to corneal epithelium. Inserted cDNA was introduced into insect cells using a baculovirus expression system, and the secretion of recombinant protein was identified using antisera against a synthetic peptide. Proteolytic activity was determined using bovine serum albumin (BSA) as substrate. The expressions of the novel cathepsin in human cornea and other tissues were examined by reverse transcription-polymerase chain reaction (RT-PCR).

RESULTS. The full-length cDNA clone encoded a peptide of 334 amino acids with 82% identity with bovine cathepsin L and 77% identity with human cathepsin L when aligned. The recombinant protein produced in the baculovirus expression system cleaves BSA, and its activity was inhibited by the cysteine proteinase inhibitors E-64 and leupeptin, but not by pepstatin A, phenylmethylsulfonyl fluoride, and EDTA. By RT-PCR, a low level of expression was observed in some other epithelial tissues of ectodermal origin, but only in cornea was it higher than cathepsin L, which is known to be a general lysosomal cathepsin. Cathepsin V protein was detected in human corneal epithelium by western blot analysis, but not in tear fluid.

CONCLUSIONS. The amino acid homology and proteolytic activity of the recombinant protein indicate that the novel gene is a new member of the cathepsins that have features of cysteine proteinase. Its uniquely high expression in corneal epithelium strongly implies an important role in corneal physiology. (*Invest Ophthalmol Vis Sci.* 1998;39:1789-1796)

We have identified several candidates for genes uniquely active in corneal epithelium by a random 3'-directed expressed sequence tag (EST) collection and database comparison approach.¹⁻³ The most abundant EST clones among them, representing gene ID GS8025, encodes human keratin 12.³ From the remaining candidates, we selected GS8009 for further analysis. The nucleotide sequence of GS8009 appeared four times in 1062 isolates from a corneal epithelial library but was absent in the same type of data sets from 45 other tissues. The nucleotide sequence of GS8009 was highly similar to the 3' termini of human and porcine cathepsin L (71% in 149 bp and 86% in 177 bp, respectively), suggesting that the corresponding gene encodes protein of the same proteinase family.

Lysosomal cysteine proteinases, including cathepsin L, play a major role in protein degradation and turnover.^{4,5} In addition, they degrade extracellular matrix components, such as fibronectin,⁶ laminin,⁶ elastin,⁷ collagen,^{8,9} and proteoglycan,¹⁰ being secreted from cells under various conditions such as tumor invasion and tissue remodeling.¹¹⁻¹³ Its involvement in the incidence and progression of diseases such as pulmonary emphysema,⁷ rheumatoid arthritis,¹² and glomerulonephritis¹⁴ have also been reported. In the present study, we determined the complete nucleotide sequence of cDNA tagged by GS8009, and we named this novel cathepsin, cathepsin V. We showed the proteolytic activity of the recombinant protein and the tissue distribution of cathepsin V mRNA by reverse transcription-polymerase chain reaction (RT-PCR).

MATERIALS AND METHODS

Isolation and Sequence Analysis of Cathepsin V cDNA

Using the library of human corneal epithelium constructed previously, approximately 80,000 plaques were screened with ³²P-labeled random primed fragments of a 3'-directed EST, GS8009 (185 bp). Among 25 positive plaques, the longest cDNA insert was sequenced by the shotgun method (Fig. 1A).^{15,16}

Nucleotide and amino acid sequence alignments with human cathepsin L and other cysteine proteinases were performed with the CLUSTAL W program.¹⁷ The dendrogram was

From the ¹Institute for Molecular and Cellular Biology, Osaka University; and the ²Department of Ophthalmology, Kyoto Prefectural University of Medicine, Japan.

Supported in part by a grant-in-aid for the Human Genome Program; Research Grant 08457467 from the Ministry of Education, Science, and Culture of Japan; and a research grant from Kyoto Foundation for the promotion of Medical Science and the intramural research fund of Kyoto Prefectural University of Medicine, Japan.

Submitted for publication October 3, 1997; revised March 23, 1998; accepted May 22, 1998.

Proprietary interest category: N.

Reprint requests: Kousaku Okubo, Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamada-oka, Suita, Osaka, 565, Japan.

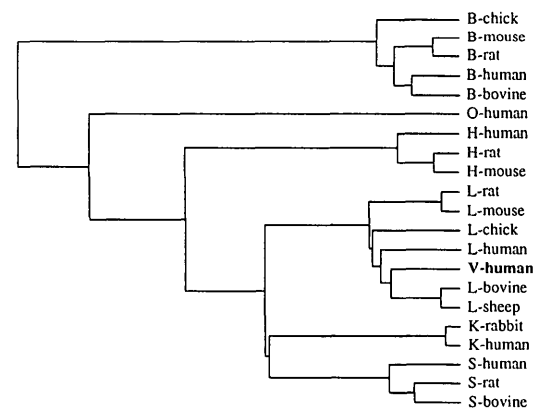
A

cathepsin V	1		C
cathepsin L	1	ACCTCCACGTGCCCTGTTTTTCGGAGGCACATCCTGGCCCTTCCACAGCTCTGGGT	
cathepsin V	2	GGACGCTGTAAATCTCAGAGGCTGTTCGAGGGTGCCTGCCAGCTGCCAGCGCT--	
cathepsin L	61	AAATGCTTGGGAGAATA-ATTTAAATATTTTATCTACCATTGGTGGCCCTAATTTTCA	
cathepsin V	60	GCTGCTTTTGAACATGAATCTTTGCTGCTCTGGCTGCCCTTTGCTTGGGAATAGCCT	
cathepsin L	120	GGGGCAGTAAGATATGAATCTTACATCATCTTGCTGCCCTTTGCTGGGAATGGCCT	
cathepsin V	120	CCGCTGTCTCAAAATTTGACCAAAATTTGGATACAAAGTGTCACGATGGAAGCAACAC	
cathepsin L	180	CAGCTACTCTAACATTTGATCACAGTTTAGAGGCACAGTGACCAGTGGAAGCGGATGC	
cathepsin V	180	ACAGAAGATTATATGGCGGAATGAAGAAGATGGAGGAGAGCAGTGCGGAAAGAATA	
cathepsin L	240	ACAACAGATTATACGGCATGAATGAAGAAGATGGAGGAGAGCAGTGCGGAGAAACA	
cathepsin V	240	TGAAAATGATGAACTGCACAATGGGGAATACAGCCAAGGAAACATGGCTTCAAAATGG	
cathepsin L	300	TGAAGATGATGAACTGCACAATCAGGAATACAGGGAAGGAAACACAGCTTCAAAATGG	
cathepsin V	300	CCATGAATGCTTTTGGTGACATGACCAATGAAGAATTCAGGCAGATGATGGTTGCTTTC	
cathepsin L	360	CCATGAACGCTTTGGAGACATGACCAATGAAGAATTCAGGCAGATGATGGTTGCTTTC	
cathepsin V	360	GAAACCAGAAATTCAGGAAGGGGAAAGTCTTCGTCAGCCTCTGTTTCTGATCTCCCA	
cathepsin L	420	AAAACCGTAAAGCCAGGAAGGGGAAAGTGTTCAGGAACCTCTGTTTATGAGGCCCAAC	
cathepsin V	420	AATCTCTGGATTGGAGAAAGAAAGGCTACGTGACCGCAGTGAAGAATCAGAAACAGTGT	
cathepsin L	480	GATCTCTGATTGGAGAGAGAAAGGCTACGTGACCTCTCTGAAGAATCAGGCTCAGTGTG	
cathepsin V	480	GTTCTGTGTGGGCTTTAGTGCAGCTGGTGCTCTGAAGGACAGATGTTCCGGAATACTG	
cathepsin L	540	GTTCTGTGTGGGCTTTAGTGCCTACTGGTGCTCTGAAGGACAGATGTTCCGGAATACTG	
cathepsin V	540	GGAAACTTGTCTCACTGAGCGAGCAGAACTGGTGGACTGTTCCGGCTCTCAAGGCAATC	
cathepsin L	600	GGAGGCTTATCTCACTGAGTGAAGCAATCTGGTAGACTGCTTGGGCTCTCAAGGCAATG	
cathepsin V	600	AGGGCTGCAATGGTGGCTTACATGGCTAGGCGCTTCAGTATGTCAGGAGAACGAGGCC	
cathepsin L	660	AAGGCTGCAATGGTGGCTTAAATGATTATGCTTTCAGTATGTCAGGAGAACGAGGCC	
cathepsin V	660	TGGACTCTGAGGAATCTTATCCATATGAGCAGTGGATGAAATCTGTAAATCAGACCTG	
cathepsin L	720	TGGACTCTGAGGAATCTTATCCATATGAGCAGCAGAGAAATCTGTAAATCAGATCCCA	
cathepsin V	720	AGAATTCTGTGCTAATGACACTGGCTTACAGATGGTCGCACCTGGAAGGAGAGAGGCC	
cathepsin L	780	AGTATTCTGTGCTAATGACACCGGCTTT--GTGGACATCCCTAAGCAGGAGAGAGGCC	
cathepsin V	780	TGATGAAACGACTGCGAAGCTGGGGCCATCTCCGTTGCTATGGAATGAGGCCATTCGT	
cathepsin L	837	TGATGAAAGCAGTGCACACTGGGGCCATTTCTGTGCTATGATGAGGTCATGAGT	
cathepsin V	840	CCTTCCAGTTCTACAAATCAGGCATTATTTTGAACAGACTGAGCAGCAAAAACCTGG	
cathepsin L	897	CCTTCTGCTTCTATAAGAGGCATTATTTTGAAGCAGACTGAGCAGTGAAGACATGG	
cathepsin V	900	ATCATCTGTCTCTGGTGTGGCTTGGCTTGAAGCAAGCAATTCGAAATCAGCAAGT	
cathepsin L	957	ATCATCTGTCTCTGGTGTGGCTTGGCTTGAAGCAAGCAATTCGAAATCAGCAATTAAT	
cathepsin V	960	ATTGGCTGCTCAAAAACAGCTGGGGTCCAGAAATGGGGTCCGAATGGCTATGTAATAATAG	
cathepsin L	1017	ATTGGCTGCTGAAGAACAGCTGGGGTGAAGAAATGGGGCATGGGTGGCTACGTAAGATGG	
cathepsin V	1020	CCAAAGACAGAACAACCACTCTGGAATCGCCACAGCAGCCAGCTACCCCAATCTGTAG	
cathepsin L	1077	CCAAAGACCGAGAAACCACTTGTGGAATGGCTCAGCAGCCAGCTACCCCACTGTGTAG	
cathepsin V	1080	CTGATGGATGGTGAGGAGGAAGCACTTAA--GGACAGCATCTGGGGAATTTTATC	
cathepsin L	1137	CTG-TGGACGGTGATGAGGAGGAAGCACTGACTGGGATGGCGCATGCACTGGGGAAT-TC	
cathepsin V	1136	TTGAAACTGACCAACCGCTTATTGTGTAAGATAAACCACTTGAATCATGAGGATCCAAG	
cathepsin L	1195	TTCACTTACC--AGCCCCGCTGTGTGGATACAC-ACTTGAATCATGAGGATCCGAG	
cathepsin V	1196	TTGAGATTTTAAATCTGTGACATTTTACAAGGTAATAATCTTACCACTACTTAAATATAT	
cathepsin L	1252	TTGATTTGCAATCTGTGATATTTTACACATGGTAAA--TGTTACCTCTATTTTAAATTAC	
cathepsin V	1256	TGTTATACACAGCTTTATGATATCAAGACTCATTTGCTTAATCTAAGACTTTTGAATTT	
cathepsin L	1310	TGCTATAAATAGGTTTATATATT--GATTCACTACTGACT--TTGCATTT	
cathepsin V	1316	TCATTTTTTAAAGATGTACAAAACAGTTTGAATAAATTTTAAATCTGATATAAA	
cathepsin L	1358	TGCTTTT--AAAAGGATGTATAAATTTTACCTGTTTAAATAAATATG	

B

V-human	60	NHSLVLAAP CLGIASATPK FDMGLDTRY QNKATRRRLY GNNEEWERRA VMEKNKHIE
L-human		NHPLILAAF CLGIASATLT FDSLEAQT KWKAMRNRLY GNNEEWERRA VMEKNKHIE
L-mouse		NHLLLLAVL CLGTALATPK FDOTFSAEMH QNKSTHRLY GNNEEWERRA VMEKNKHIE
L-rat		HTPLLLAVL CLGTALATPK FDOTFNAQNH QNKSTHRLY GNNEEWERRA VMEKNKHIE
L-bovine		NHPSFLLTVL CLGVASAAFK LDMPLDAHH QNKATRRRLY GNNEEWERRA VMEKNKHIE
L-sheep		
L-chick		
V-human	120	LHNGEYSGK HPTFMADAP GDMTNEEFQ VNGCFQNKFP RKQKVFPEPL FLVPSKVDW
L-human		LHNGEYREGK HPTFMADAP GDMTNEEFQ VNGCFQNKFP RKQKVFPEPL FYEAPRSVDW
L-mouse		LHNGEYSNQK HPTFMADAP GDMTNEEFQ VNGCFQNKFP RKQKVFPEPL FLVPSKVDW
L-rat		LHNGEYSNQK HPTFMADAP GDMTNEEFQ VNGCFQNKFP RKQKVFPEPL FLVPSKVDW
L-bovine		LHNGEYSEK HPTFMADAP GDMTNEEFQ VNGCFQNKFP RKQKVFPEPL FLVPSKVDW
L-sheep		
L-chick		
V-human	180	REKGYTPVK HQGQCGSCA FSATGALBQ MFRKTOKLV LSEQLVDCS RPDGNGCNG
L-human		REKGYTPVK HQGQCGSCA FSATGALBQ MFRKTOKLV LSEQLVDCS RPDGNGCNG
L-mouse		REKGYTPVK HQGQCGSCA FSATGALBQ MFRKTOKLV LSEQLVDCS RPDGNGCNG
L-rat		REKGYTPVK HQGQCGSCA FSATGALBQ MFRKTOKLV LSEQLVDCS RPDGNGCNG
L-bovine		REKGYTPVK HQGQCGSCA FSATGALBQ MFRKTOKLV LSEQLVDCS RPDGNGCNG
L-sheep		
L-chick		
V-human	240	GMADFAQYV QNGGLDSEE SVYPVAV-DEI CKYTPENVA NTOTFVDIPK -REKALKAV
L-human		GMADFAQYV QNGGLDSEE SVYPVAV-DEI CKYTPENVA NTOTFVDIPK -REKALKAV
L-mouse		GMADFAQYV QNGGLDSEE SVYPVAV-DEI CKYTPENVA NTOTFVDIPK -REKALKAV
L-rat		GMADFAQYV QNGGLDSEE SVYPVAV-DEI CKYTPENVA NTOTFVDIPK -REKALKAV
L-bovine		GMADFAQYV QNGGLDSEE SVYPVAV-DEI CKYTPENVA NTOTFVDIPK -REKALKAV
L-sheep		
L-chick		
V-human	300	ATVOPISVAH DAGHSFPQY KSGIYDFPC SKDLNHLVL VVGQFEGDN SINNKFWLVK
L-human		ATVOPISVAH DAGHSFPQY KSGIYDFPC SKDLNHLVL VVGQFEGDN SINNKFWLVK
L-mouse		ATVOPISVAH DAGHSFPQY KSGIYDFPC SKDLNHLVL VVGQFEGDN SINNKFWLVK
L-rat		ATVOPISVAH DAGHSFPQY KSGIYDFPC SKDLNHLVL VVGQFEGDN SINNKFWLVK
L-bovine		ATVOPISVAH DAGHSFPQY KSGIYDFPC SKDLNHLVL VVGQFEGDN SINNKFWLVK
L-sheep		
L-chick		
V-human	334	NHSGPFWGN QYVKAIDEN NHCOIATAAS YPTV
L-human		NHSGPFWGN QYVKAIDEN NHCOIATAAS YPTV
L-mouse		NHSGPFWGN QYVKAIDEN NHCOIATAAS YPTV
L-rat		NHSGPFWGN QYVKAIDEN NHCOIATAAS YPTV
L-bovine		NHSGPFWGN QYVKAIDEN NHCOIATAAS YPTV
L-sheep		
L-chick		

C



based on the amino acid sequences of the mature form of cysteine proteinases¹⁸⁻³⁶ and constructed with the unweighted pair group methods with averages clustering method.³⁷

Production of Preprocathepsin V in Baculovirus

The recombinant cathepsin V protein was produced using a commercial system (BAC-TO-BAC Baculovirus Expression Systems; GibcoBRL, Gaithersburg, MD). The cathepsin V cDNA containing the complete coding sequence was generated by PCR amplification with primers designed to include the *RsrII* or *SalI* site (5'-TGAACATGACGGTCCGTGATGAATCTTTCGCTCGTCCT-3' and 5'-GTGCAAGTCTAGGTCGACCATCCATCAGCTCACACATT-3') and the PCR product was cloned between the *RsrII* and *SalI* sites of the pFASTBAC donor plasmid. Starting from this plasmid, recombinant baculovirus with a titer of 10^3 pfu/ml was obtained according to the manufacturer's protocol. After a single round of amplification with Sf9 cells (Invitrogen, Carlsbad, CA), the titer of recombinant virus was raised to 10^7 pfu/ml. Hi5 cells (Invitrogen) at a density of 0.5×10^6 cells/ml were infected with recombinant cathepsin V baculovirus or wild-type AcNPV baculovirus (Invitrogen) of the same titer at a multiplicity of infection of five and were cultured at 27°C (Grace's Insect Medium; BRL) supplemented with 7% fetal calf serum (FCS). Twelve hours after infection, culture medium was replaced with serum-free medium (SF900-II; BRL). Suspension cultures were harvested at 24, 48, 72, and 96 hours after infection, and culture medium and cells were separated by centrifugation at 1500g for 10 minutes. Proteinase inhibitors at final concentrations of 20 μ M E-64, 20 μ M leupeptin, 20 μ M pepstatin A (Peptide Institute, Osaka, Japan), 1 mM phenylmethylsulfonyl fluoride (PMSF), and 1 mM EDTA were immediately added to the culture medium for time-course expression to prevent further processing.

Antiserum and Western Blot Analysis of Recombinant Protein

Antiserum for cathepsin V was made by immunizing a rabbit with synthetic peptide that corresponds to amino acid residues of 216 to 234 (ENSVANDTGFTVVAPGKEK; Fig. 1B) in preprocathepsin V and conjugated to keyhole limpet hemocyanin.

Samples for time-course expression were prepared as follows. Protein in 200 μ l culture medium was precipitated by centrifugation after adding 8 volumes acetone and dissolving in 20 μ l water. Cell pellets from 1 ml suspension culture were lysed in 100 μ l lysis buffer containing 50 mM Tris-HCl [pH 8.5], 10 mM 2-mercaptoethanol, 1% Nonidet P-40 (Sigma, St. Louis, MO), 1 mM PMSF, 10 μ M leupeptin, and 10 μ M pepstatin A. After brief centrifugation, the pellet was resuspended in 100 μ l lysis buffer and disrupted by sonication. Twenty microliters of

this lysate and concentrated culture medium were boiled with 6 μ l 4 $\times$ loading buffer (0.25 M Tris-HCl [pH 6.8], 8% sodium dodecyl sulfate [SDS], 40% glycerol, 20% 2-mercaptoethanol, and 0.004% bromophenol blue) and electrophoresed on 10% polyacrylamide gel in the presence of 0.1% SDS (10% SDS-PAGE) and transferred to a polyvinylidene difluoride membrane (Millipore, Bedford, MA). After blocking in buffer containing 20 mM Tris-HCl (pH 7.6), 137 mM NaCl, and 5% skim milk, the filter was incubated in 20 mM Tris-HCl (pH 7.6) 137 mM NaCl, 0.1% Tween 20, and 0.1% bovine serum albumin (BSA) with diluted 1:1000 anti-cathepsin V antisera. Immuno-reactive bands were visualized with peroxidase-conjugated goat anti-rabbit IgG (Bio-Rad, Richmond, CA) by immunostaining (HRP-1000; Konica, Tokyo, Japan). Culture medium and cells infected with wild-type AcNPV baculovirus were collected 72 hours after infection and examined by western blot analysis as described earlier.

Processing of Recombinant Protein

Two hundred microliters culture medium collected 72 hours after infection was incubated with an equal volume of assay buffer (200 mM sodium acetate [pH 6.0], 5 mM dithiothreitol, 5 mM EDTA, and 0.1% SDS) at 37°C for 0, 5, 15, 30, 45, or 60 minutes, and proteinase inhibitors were immediately added (final concentration): 20 μ M E-64, 20 μ M leupeptin, 20 μ M pepstatin A, 1 mM PMSF, and 1 mM EDTA. They were precipitated by centrifugation after adding 8 volumes acetone and dissolving in 20 μ l water and then were analyzed by western blot analysis, as described earlier.

Analysis of Proteolytic Activity of Recombinant Protein

Proteolytic activity of culture medium was examined using BSA (Sigma, St. Louis, MO) as substrate. For activation, 10 μ l culture medium collected 72 hours after infection was preincubated with an equal volume of assay buffer at 37°C for 30 minutes before 2 μ l 0.1% BSA was added for incubation at 37°C for 30 minutes. After adding 6 μ l of 4 $\times$ loading buffer, samples were fractionated by 10% SDS-PAGE and stained with Coomassie brilliant blue.³⁸ For inhibition studies, aliquots of activated culture medium were incubated with various proteinase inhibitors (10 μ M E-64, 10 μ M leupeptin, 10 μ M pepstatin A, 1 mM PMSF, and 1 mM EDTA) at 37°C for 5 minutes before adding BSA.

RT-PCR

All investigations followed the tenets of the Declaration of Helsinki; informed consent, and approval by the Institutional Review Board on Human Experiments of the Research Com-

FIGURE 1. (A) Alignment of human cathepsin V and human cathepsin L nucleotide sequences. Initiation codon, stop codon, and putative polyadenylation signal are underlined. Human cathepsin V-specific primers (cat V-S and cat V-AS), human cathepsin L-specific primers (cat L-S and cat L-AS), and common primers for unique competitive reverse transcription-polymerase chain reaction (common-S and common-AS) are indicated in **boldface**; arrows indicate direction of amplification. *Afl*III site is indicated in double underlines. (B) Alignment of amino acid sequences of human cathepsin V with cathepsin L (L). Two amino acids, cysteine and histidine, known as the active sites of cysteine proteinases, were conserved (arrows). Putative cleavage sites between signal peptides and proenzymes and between proenzymes and mature enzymes are indicated by **solid triangles**. Predicted N-glycosylation sites of cathepsin V (V) are indicated by **open triangles**. The region of synthetic peptide is underlined. (C) The dendrogram based on the amino acid sequence of the mature form of cysteine proteinases. B, cathepsin B; H, cathepsin H; L, cathepsin L; K, cathepsin K; O, cathepsin O; S, cathepsin S; V, cathepsin V.

mittee of the Kyoto Prefectural University of Medicine were obtained. Normal corneal epithelial cells were taken from eyes obtained in autopsy (1 man, 76 years old). Conjunctival epithelial cells were collected from 10 normal eyes (4 men, 6 women; age range, 23–65 years) by brush cytology technique, as described previously.³⁹ Approximately 1-cm × 5-cm segments of normal skin (1 man, 70 years old) were excised at orthopedic surgery. Total RNA from those tissues were prepared with a kit (RNeasy; Qiagen, Hilden, Germany) according to the manufacturer's instructions. Total RNA from other tissues (brain, heart, kidney, liver, lung, and placenta) were purchased from Clontech (Palo Alto, CA).

cDNA was generated from 1 µg total RNA according to a standard procedure.⁴⁰ PCR was performed for 25 cycles at 94°C for 30 seconds, 50°C for 1 minute, and 72°C for 1 minute, using a commercial system (GeneAmp; Perkin-Elmer Cetus, Norwalk, CT) using one eighth of the cDNA as a template in 50 µl reaction mixture containing 0.2 µM each of the primers, 200 µM deoxyribonucleoside triphosphate, and 1 unit DNA polymerase (AmpliTag; Perkin-Elmer Cetus, Norwalk, CT). Five microliters PCR products were electrophoresed on 1.2% agarose gel and stained with fluorochrome SYBR Green I (FMC Bioproducts, Rockland, ME).

For independent amplification, primers with the least conserved regions were selected from neighboring exons in cathepsin L (cat L-S and cat L-AS) and the corresponding regions of cathepsin V (cat V-S and cat V-AS; Fig. 1A). Glyceraldehyde 3-phosphate dehydrogenase (G3PDH) primers (5' primer: 5'-ACTGCCACCCAGAAGACTGT-3', 3' primer: 5'-TCCACCACCTGTTGCTGTA-3') were used as controls.

Competitive RT-PCR

Using the same sets of primers (common-S and common-AS; Fig. 1A), competitive RT-PCR was performed for 40 cycles at 94°C for 30 seconds, 50°C for 1 minute, and 72°C for 1 minute, using the same commercial system (GeneAmp; Perkin-Elmer Cetus). One eighth of the cDNA was used as a template in 50 µl total reaction mixture. After purification by phenolchloroform extraction and ethanol precipitation, each product was dissolved in 50 µl distilled water. Ten microliters of product was digested with 10 units *Afl*III at 37°C for 1.5 hours and electrophoresed on 6% polyacrylamide gel, along with the same volume of nondigested products, and stained with SYBR Green I. The intensity of each band was quantified by fluorescent scanner (FMBIO II; Hitachi Software Engineering, Yokohama, Japan), and ratios of cathepsins L and V were determined after normalization by band size.

Western Blot Analysis of Cathepsin V Protein in Corneal Epithelium and Tears

Human corneal epithelial cells, scraped from an area approximately 6 mm in diameter in two eyes during laser photorefractive keratectomy, were solubilized in 50 µl lysis buffer (2% SDS, 0.2 mM PMSF, 10 µM E-64, 10 µM pepstatin A, and 5 mM EDTA). They were incubated at 55°C for 5 minutes, homogenized, and heated at 95°C for 5 minutes. Tears were collected from 20 eyes of 10 healthy, normal volunteers (6 men, 4 women; age range, 24–67 years) using 5 µl disposable microcapillaries (Accu-Fill90 Micropet; Clay Adams, Parsippany, NJ). Approximately 50 µg protein was fractionated by 10% SDS-PAGE, blotted to polyvinylidene difluoride membrane, and ana-

lyzed as described earlier. For a competitive experiment, anti-cathepsin V antisera were incubated with 5 µg/ml synthetic peptides for 1 hour at room temperature before use.

RESULTS

Sequence Analysis

The longest cDNA clone corresponding to *GS8009* contained an open reading frame encoding 334 amino acids starting with ATG proceeded by the Kozak consensus (Fig. 1A).^{41,42} A signal peptide sequence (amino acid position 1–17) and possible N-glycosylation site were predicted (Fig. 1B).⁴³ It had significant sequence similarities with cysteine proteinases, especially with cathepsin L¹⁸ and was designated cathepsin V (Fig. 1C). In the coding region, the highest similarity was found in bovine cathepsin L with which 82% of the entire peptide was identical. It was 77% identical with human cathepsin L. Among the members of the cathepsins, orthologues in various vertebrates are clustered by 72% to 95% identity, and different types of cathepsins are clearly segregated by a more than 40% difference (Fig. 1C). However, this novel cathepsin migrated in the middle of the cathepsin L cluster according to the mutual similarity among the mature peptides, suggesting a close evolutionary association with cathepsin L.

Production and Functional Assay of Preprocathepsin V Protein

To prove the enzymatic activity of cathepsin V protein, the complete coding sequence, including the signal peptide, was subcloned into a baculovirus expression vector for secretion into the culture medium in the form of procathepsin V, as described in the case of papain.⁴⁴ Using rabbit anti-sera raised against the synthetic peptide (amino acid residues 216–234), an approximately 39-kDa immunoreactive protein was detected in culture medium 24 hours after infection and accumulated for the next 24 hours (Fig. 2A). In the cells, in addition to 39-kDa protein, two immunoreactive bands, approximately 41 kDa and 37 kDa, were detected 48 hours to 96 hours after infection (Fig. 2A). The calculated molecular mass of cathepsin V in prepro, pro, and mature forms are 37 kDa, 35.5 kDa, and 24 kDa, respectively. Judging from the time course of appearance, it is reasonable to regard the 37 kDa protein detected only in cells as preprocathepsin V before modification. The remaining two bands, 41 kDa and 39 kDa, may correspond to differently glycosylated forms of procathepsin V. None of the immunoreactive bands was detectable in culture medium or in cells infected with wild-type AcNPV baculovirus (Fig. 2B).

Proteolytic activity in the culture medium was tested under conditions determined by modifying those for other cathepsins.^{18–36} In the presence of 0.05% SDS and 2.5 mM dithiothreitol at 37°C (pH 6.0) the secreted 39-kDa procathepsin V was autocatalytically converted to an approximately 31-kDa protein, which probably corresponded to the mature form, by intermediate proteins of heterogeneous sizes (Fig. 3). The same aliquot was used for the proteolytic assay by adding BSA (66 kDa) to a final concentration of 0.01% as substrate. By incubating at 37°C, BSA was completely digested into short peptides of heterogeneous sizes (Fig. 4; lane 3). Such proteolytic activity in the medium was inhibited specifically by the cysteine proteinase inhibitors E-64 and leupeptin and was absent in the culture medium infected with wild-type AcNPV baculovirus (Fig. 4).

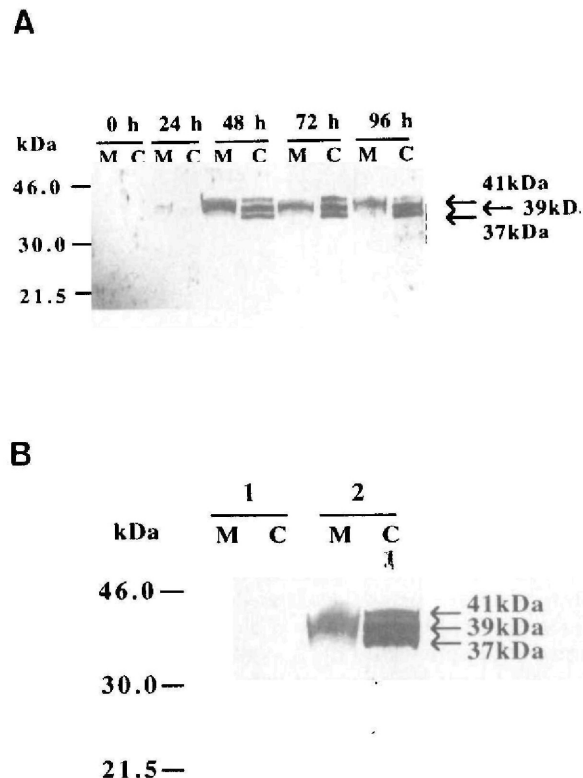


FIGURE 2. Western blot analysis of recombinant cathepsin V expression. (A) Time-course expression of recombinant cathepsin V. Culture medium (M) and cells (C) infected with recombinant cathepsin V baculovirus were collected 24, 48, 72, and 96 hours after infection. These samples were electrophoresed and transferred to polyvinylidene difluoride membrane. Western blot analysis was performed using anti-peptide antisera. (B) Culture medium (M) and cells (C) infected with wild-type AcNPV baculovirus (1) were examined by western blot analysis as described above. 2, recombinant cathepsin V baculovirus.

Expression of Cathepsin V Gene in Human Tissues

The expressions of the human cathepsin V gene were examined in various tissues by RT-PCR, using cathepsin L, the gene's closest relative, as a reference. Starting with 125

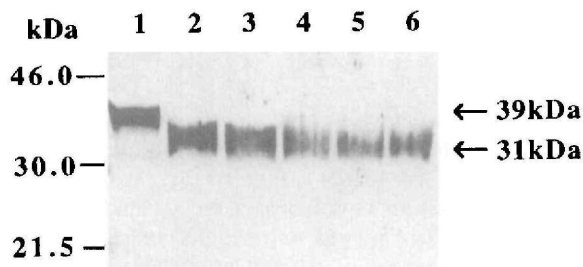


FIGURE 3. Processing of secreted procathepsin V. Culture medium containing procathepsin V was incubated as described in the Methods section and analyzed by western blot analysis. Lanes 1 to 6, samples after incubation for 0, 5, 15, 30, 45, and 60 minutes, respectively.

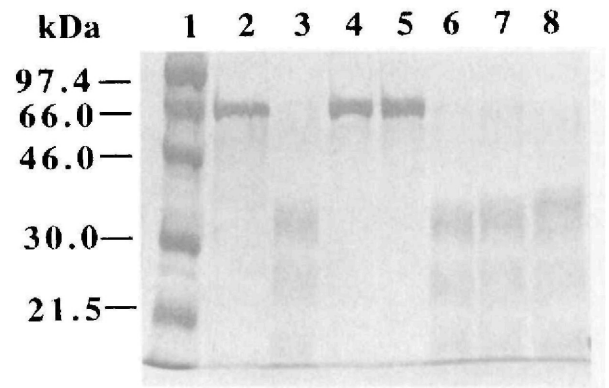


FIGURE 4. Proteolytic activity against bovine serum albumin (BSA; 66 kDa) and inhibition studies. After preincubation at 37°C for 30 minutes, the culture medium was incubated with 0.01% BSA at 37°C for 30 minutes. For inhibition studies, various proteinase inhibitors were added and incubated at 37°C for 5 minutes before addition of BSA. Samples were electrophoresed and stained with Coomassie brilliant blue. Lane 1, molecular size marker; lane 2, culture medium of wild-type AcNPV baculovirus; lanes 3 to 8, culture medium of recombinant cathepsin V baculovirus in the absence of inhibitor (lane 3), in the presence of 10 μ M E-64 (lane 4), 10 μ M leupeptin (lane 5), 10 μ M pepstatin A (lane 6), 1 mM phenylmethylsulfonyl fluoride (lane 7), and 1 mM EDTA (lane 8).

ng total RNA, various amounts of cathepsin L amplicon were detected in all tissues after 25 thermal cycles (Fig. 5A). In contrast, cathepsin V transcript was detectable only in stratified epithelia, clearly in cornea and faintly in conjunctiva and skin. Because preliminary studies showed that the amplification was still within the linear range for all sets of primers and templates under this condition, it was concluded that expression of cathepsin V was highest in cornea and almost limited to stratified epithelium, in contrast to the ubiquitous expression of cathepsin L.

To determine the relative expression levels of cathepsins L and V in tissues where coexpression was detected, another series of PCRs were performed in which these cathepsins competed for the same set of primers. For these reactions, primers (common-S and common-AS) were selected in the regions most conserved between cathepsins L and V. In positions of sequence discrepancy between the two, the nucleotides were replaced with those differing from either of them to equalize the amplifying efficiency (Fig. 1A). The expected PCR products from cathepsin V and cathepsin L are of almost the same size (291 bp and 290 bp), but the restriction site for *Afl*III exists only in the cathepsin V amplicon (Fig. 1A). As a reference, kidney was also included in the test, in which high-level expression of cathepsin L has been reported.^{45,46} After 40 cycles of amplification, products from all tissues were saturated to almost equal amounts. By selectively cleaving the cathepsin V fragment with *Afl*III and by quantification of the intensities of the bands, the ratios of cathepsins V to L were determined as 45, 1.2, 1.1, and 0.03, respectively, in cornea, conjunctiva, skin, and kidney (Fig. 5B).

By western blot analysis of human corneal epithelium, two immunoreactive bands that comigrate with processed

forms of recombinant proteins were observed (Fig. 6; lanes 2 and 4). They were completely competed by 5 μ g/ml peptide antigens (Fig. 6; lane 5). Such immunoreactive protein was not observed in tear fluid containing an equal amount of total protein (Fig. 6; lane 3).

DISCUSSION

By random 3'-directed EST collection and database comparison, we selected several cornea-specific genes.³ In this report, the nucleotide sequence of cDNA tagged by *GS8009* was shown to encode a novel cathepsin. In vitro, this novel cathepsin, designated cathepsin V, was autoactivated and digested BSA under empirically determined conditions (in

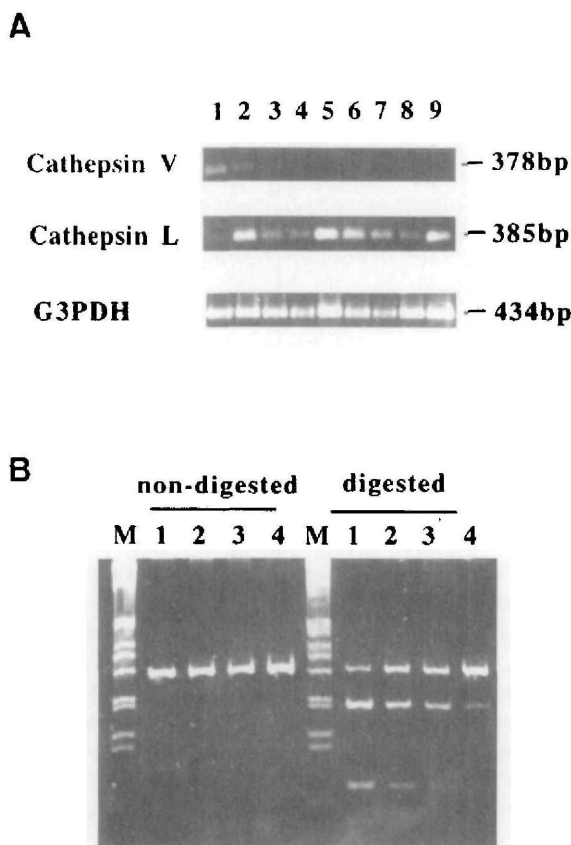


FIGURE 5. (A) Reverse transcription-polymerase chain reaction (RT-PCR) of human cathepsin V and human cathepsin L in corneal epithelium and other tissues. Amplified products were electrophoresed on 1.2% agarose gel and stained with SYBR Green I. Lane 1, corneal epithelium; lane 2, conjunctival epithelium; lane 3, skin; lane 4, brain; lane 5, heart; lane 6, kidney; lane 7, liver; lane 8, lung; and lane 9, placenta. (B) Comparison of expression of human cathepsin V and human cathepsin L by competitive RT-PCR. Coamplified products were electrophoresed on 6% polyacrylamide gel and stained by SYBR Green I, either before (nondigested) or after digestion with *Afl*III. Expected sizes of amplification products for cathepsin V and cathepsin L were 291 bp and 290 bp, respectively. Predicted sizes of *Afl*III-digested fragments for cathepsin V were 200 bp and 91 bp. Lane 1, corneal epithelium; lane 2, conjunctival epithelium; lane 3, skin; and lane 4, kidney.

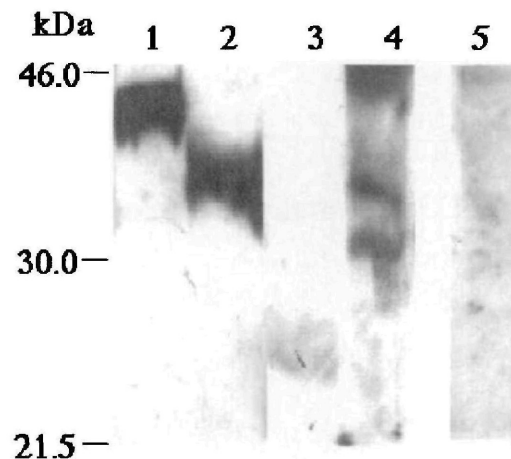


FIGURE 6. Western blot analysis of cathepsin V protein in corneal epithelium and tears. Human corneal epithelial lysates and tears, both containing approximately 50 μ g protein, were electrophoresed on 10% sodium dodecyl sulfate-polyacrylamide gels and transferred to polyvinylidene difluoride membranes. Western blot analysis was performed using antipeptide antisera. A competitive experiment was performed with 5 μ g/ml synthetic peptides. Lanes 1 and 2, culture medium infected with recombinant cathepsin V baculovirus before and after incubation for 45 minutes in processing conditions; lane 3, tear fluid; lane 4, corneal epithelial lysates; lane 5, competitive experiment with equal amounts of human corneal lysates.

the presence of 0.05% SDS and 2.5 mM dithiothreitol at pH 6.0) and was selectively inhibited by cysteine proteinase inhibitors.

By RT-PCR, it was shown that the expression of cathepsin V was almost limited to stratified epithelium and was strongest in cornea (Fig. 5). Moreover, it is probably the most abundant proteinase in corneal epithelium, because other known proteinase cDNA was found in the EST clones collected from this tissue.³

The most important question is what the physiologic roles or principal substrates of cathepsin V are. Cathepsins had been regarded as general lysosomal proteinases that cooperate in terminal degradation of endocytized proteins.⁴⁷ However, recent studies have shown that some cathepsins, especially those with uneven tissue distributions, have specific functions.⁴⁷ For example, cathepsin S, abundant in spleen and thyroid, is believed to be involved in antigen processing⁴⁸ and in thyroglobulin processing.³³ Another example is cathepsin K, which is abundant in osteoclasts.⁴⁹ It has been determined by biochemistry to be a major proteinase for bone resorption^{50,51} and by its elucidation as a candidate gene for hereditary osteochondrodysplasia.⁵² In either case, cathepsins with unique tissue distributions process substrates unique to each tissue.

Because western blot analysis showed that this enzyme stayed within cells and was not secreted in tear fluid, some structural components of corneal epithelium are the likely substrates. Biochemical and histochemical studies are awaited for further speculations.

From the structural aspect, cathepsin V is exceptional among cathepsins, in that it has high sequence similarity to cathepsin L, disturbing the order in similarity clustering of

family members. Chromosomal localization and the analysis of orthologues in different vertebrates will be the next step in clarifying the evolution of the cathepsin V gene.

References

- Okubo K, Hori N, Matoba R, Niiyama T, Matsubara K. A novel system for large-scale sequencing of cDNA by PCR amplification. *DNA Seq.* 1991;2:137-144.
- Okubo K, Hori N, Matoba R, et al. Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nat Genet.* 1992;2:173-179.
- Nishida K, Adachi W, Shimizu-Matsumoto A, et al. A gene expression profile of human corneal epithelium and the isolation of human keratin 12 cDNA. *Invest Ophthalmol Vis Sci.* 1996;37:1800-1809.
- Barrett AJ, Kirschke H. Cathepsin B, cathepsin H, and cathepsin L. *Methods Enzymol.* 1981;80:535-542.
- Katunuma N, Kominami E. Structures and functions of lysosomal thiol proteinase and their endogenous inhibitors. *Curr Top Cell Regul.* 1983;22:71-101.
- Guinec N, Dalet-Fumeron V, Pagano M. "In vitro" study of basement membrane degradation by the cysteine proteinases, cathepsins B, B-like and L. Digestion of collagen IV, laminin, fibronectin, and release of gelatinase activities from basement membrane fibronectin. *Biol Chem Hoppe Seyler.* 1993;374:1135-1146.
- Mason RW, Johnson DA, Barrett AJ, Chapman HA. Elastolytic activity of human cathepsin L. *Biochem J.* 1986;233:925-927.
- Kirschke H, Kambhavi AA, Bohley P, Barrett AJ. Action of rat liver cathepsin L on collagen and other substrates. *Biochem J.* 1982;201:367-372.
- Maciewicz RA, Wotton SF, Etherington DJ, Duance VC. Susceptibility of the cartilage collagens types II, IX and XI to degradation by the cysteine proteinases, cathepsins B and L. *FEBS Lett.* 1990;269:189-193.
- Nguyen Q, Mort JS, Roughley PJ. Cartilage proteoglycan aggregate is degraded more extensively by cathepsin L than by cathepsin B. *Biochem J.* 1990;266:569-573.
- Sloane BF, Honn KV. Cysteine proteinases and metastasis. *Cancer Metastasis Rev.* 1984;3:249-263.
- Etherington DJ, Taylor MA, Henderson B. Elevation of cathepsin L levels in the synovial lining of rabbits with antigen-induced arthritis. *Br J Exp Pathol.* 1988;69:281-289.
- Reilly J Jr, Mason RW, Chen P, et al. Synthesis and processing of cathepsin L, an elastase, by human alveolar macrophages. *Biochem J.* 1989;257:493-498.
- Baricos WH, Cortez SL, Le QC, et al. Evidence suggesting a role for cathepsin L in an experimental model of glomerulonephritis. *Arch Biochem Biophys.* 1991;288:468-472.
- Frischauf AM, Garoff H, Lerach H. A subcloning strategy for DNA sequence analysis. *Nucleic Acids Res.* 1980;8:5541-5549.
- Bankier AT, Weston KM, Barrell BG. Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. *Methods Enzymol.* 1987;155:51-93.
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994;22:4673-4680.
- Gal S, Gottesman MM. Isolation and sequence of a cDNA for human pro-(cathepsin L). *Biochem J.* 1988;253:303-306.
- Chan SJ, San Segundo B, McCormick MB, Steiner DF. Nucleotide and predicted amino acid sequences of cloned human and mouse preprocathepsin B cDNAs. *Proc Natl Acad Sci USA.* 1986;83:7721-7725.
- Bechet DM, Ferrara MJ, Mordier SB, et al. Expression of lysosomal cathepsin B during calf myoblast-myotube differentiation. Characterization of a cDNA encoding bovine cathepsin B. *J Biol Chem.* 1991;266:14104-14112.
- Dong SS, Stransky GI, Whitaker CH, et al. Avian cathepsin B cDNA: sequence and demonstration that mRNAs of two sizes are produced in cell types producing large quantities of the enzyme. *Biochim Biophys Acta.* 1995;1251:69-73.
- Takio K, Towatari T, Katunuma N, Teller DC, Titani K. Homology of amino acid sequences of rat liver cathepsins B and hours with that of papain. *Proc Natl Acad Sci USA.* 1983;80:3666-3670.
- Fuchs R, Gassen HG. Nucleotide sequence of human preprocathepsin H, a lysosomal cysteine proteinase. *Nucleic Acids Res.* 1989;17:9471.
- Lafuse WP, Brown D, Castle L, Zwilling BS. IFN-gamma increases cathepsin hours mRNA levels in mouse macrophages. *J Leukoc Biol.* 1995;57:663-669.
- Ishidoh K, Imajoh S, Emori Y, et al. Molecular cloning and sequencing of cDNA for rat cathepsin H. Homology in pro-peptide regions of cysteine proteinases. *FEBS Lett.* 1987;226:33-37.
- Shi GP, Chapman HA, Bhairi SM, et al. Molecular cloning of human cathepsin O, a novel endoproteinase and homologue of rabbit OC2. *FEBS Lett.* 1995;357:129-134.
- Tezuka K, Tezuka Y, Maejima A, et al. Molecular cloning of a possible cysteine proteinase predominantly expressed in osteoclasts. *J Biol Chem.* 1994;269:1106-1109.
- Ishidoh K, Towatari T, Imajoh S, et al. Molecular cloning and sequencing of cDNA for rat cathepsin L. *FEBS Lett.* 1987;223:69-73.
- Joseph LJ, Chang LC, Stamenkovich D, Sukhatme VP. Complete nucleotide and deduced amino acid sequences of human and murine preprocathepsin L. An abundant transcript induced by transformation of fibroblasts. *J Clin Invest.* 1988;81:1621-1629.
- Wada K, Takai T, Tanabe T. Amino acid sequence of chicken liver cathepsin L. *Eur J Biochem.* 1987;167:13-18.
- Velasco G, Ferrando AA, Puente XS, Sanchez LM, Lopez-Otin C. Human cathepsin O. Molecular cloning from a breast carcinoma, production of the active enzyme in *Escherichia coli*, and expression analysis in human tissues. *J Biol Chem.* 1994;269:27136-27142.
- Shi GP, Munger JS, Meara JP, Rich DH, Chapman HA. Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastolytic cysteine protease. *J Biol Chem.* 1992;267:7258-7262.
- Petanceska S, Devi L. Sequence analysis, tissue distribution, and expression of rat cathepsin S. *J Biol Chem.* 1992;267:26038-26043.
- Ritonja A, Colic A, Dolenc I, et al. The complete amino acid sequence of bovine cathepsin S and a partial sequence of bovine cathepsin L. *FEBS Lett.* 1991;283:329-331.
- Wiederanders B, Broemme D, Kirschke H, et al. Primary structure of bovine cathepsin S. Comparison to cathepsins L, H, B and papain. *FEBS Lett.* 1991;286:189-192.
- Ritonja A, Coetzer TH, Pike RN, Dennison C. The amino acid sequences, structure comparisons and inhibition kinetics of sheep cathepsin L and sheep stefin B. *Comp Biochem Physiol B Biochem Mol Biol.* 1996;114:193-198.
- Nei M. *Molecular Evolutionary Genetics*. New York: Columbia University Press; 1987.
- Sambrook J, Fritsch EF, Maniatis T. Detection and analysis of proteins expressed from cloned genes. In: *Molecular Cloning: A Laboratory Manual*. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1989;18:1-18.88.
- Tsubota K, Kajiwaru K, Ugajin S, Hasegawa T. Conjunctival brush cytology. *Acta Cytol.* 1990;34:233-235.
- Sambrook J, Fritsch EF, Maniatis T. Construction and analysis of cDNA library. In: *Molecular Cloning: A Laboratory Manual*. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1989;8:1-8.86.
- Kozak M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* 1984;12:857-872.
- Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 1987;15:8125-8148.
- Kamphuis IG, Drenth J, Baker EN. Thiol proteases. Comparative studies based on the high-resolution structures of papain and

- actinidin, and on amino acid sequence information for cathepsins B and H, and stem bromelain. *J Mol Biol.* 1985;182:317-329.
44. Vernet T, Tessier DC, Richardson C, et al. Secretion of functional papain precursor from insect cells. Requirement for N-glycosylation of the pro-region. *J Biol Chem.* 1990;265:16661-16666.
45. Bando Y, Kominami E, Katunuma N. Purification and tissue distribution of rat cathepsin L. *J Biochem.* 1986;100:35-42.
46. Mason RW, Bartholomew LT, Hardwick BS. The use of benzyl-oxycarbonyl[125I]iodotyrosylalanyldiazomethane as a probe for active cysteine proteinases in human tissues. *Biochem J.* 1989;263:945-949.
47. Chapman HA, Munger JS, Shi GP. The role of thiol proteases in tissue injury and remodeling. *Am J Respir Crit Care Med.* 1994;150:S155-S159.
48. Shi GP, Webb AC, Foster KE, et al. Human cathepsin S: chromosomal localization, gene structure, and tissue distribution. *J Biol Chem.* 1994;269:11530-11536.
49. Drake FH, Dodds RA, James IE, et al. Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts. *J Biol Chem.* 1996;271:12511-12516.
50. Bromme D, Okamoto K, Wang BB, Biroc S. Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. *J Biol Chem.* 1996;271:2126-2132.
51. Bossard MJ, Tomaszek TA, Thompson SK, et al. Proteolytic activity of human osteoclast cathepsin K: expression, purification, activation, and substrate identification. *J Biol Chem.* 1996;271:12517-12524.
52. Gelb BD, Shi GP, Chapman HA, Desnick RJ. Pyknodysostosis, a lysosomal disease caused by cathepsin K deficiency. *Science.* 1996;273:1236-1238.