

# Molecular Cloning and Gene Expression of Chum Salmon Cystatin

Michiaki Yamashita<sup>1</sup> and Shiro Konagaya<sup>2</sup>

National Research Institute of Fisheries Science, 2-12-4 Fukuura, Kanazawa-ku, Yokohama 236

Received for publication, January 23, 1996

A salmon cystatin cDNA clone was isolated from a chum salmon cDNA library. The clone encoded a full-length extracellular-type cystatin and its signal peptide and included 5'- and 3'-untranslated regions. The deduced amino acid sequence showed a high degree of sequence similarity to mammalian cystatin C, chicken egg cystatin, and chum salmon pituitary cystatin. By Northern blot analysis, the salmon cystatin was found to show apparently non-tissue specific expression. Because platyfish EHS cells transfected with a cystatin expression vector produced a 13 kDa mature cystatin in the culture medium, the salmon cystatin was considered to act as an extracellular type of cystatin in the fish cells. These findings indicate that the salmon cystatin is a homolog of mammalian cystatin C.

**Key words:** chum salmon, cystatin, cysteine protease.

Cystatin is the name of the group of cysteine protease inhibitors, consisting of three types of molecular forms, families 1, 2, and 3 (1). Higher vertebrates are known to have some extracellular types of cystatin, i.e., family 2 cystatins, in body fluids, such as plasma, saliva, tears, and egg white. Cystatin C (2, 3), salivary type cystatins D, S, SA, and SN (4-7), and chicken egg cystatin (8-10) belong to the family 2 cystatins.

Gene expression of family 2 cystatins has been examined by several investigators previously (3, 4, 6, 7, 10, 11). The expression of the cystatin gene of chicken egg was shown to be regulated by estrogen in the oviduct and other tissues (10). The expression of salivary type cystatins is restricted to saliva and tears in mammals (4-7, 11, 12). On the other hand, mammalian cystatin C is present in many tissues and biological fluids (12). In the case of the human brain, cystatin C has a key role in the development of brain hemorrhage in a hereditary amyloidosis syndrome (1). These previous studies suggested that family 2 cystatins are related to regulation of intra- and extracellular proteolysis and antiviral or bacterial infections (1).

Studies on the structure and the physiological function of fish cystatin should be of special interest from the viewpoint of the evolution of the cystatin superfamily. The present study deals with the nucleotide sequence of cDNA and the deduced amino acid sequence of chum salmon cystatin, the gene expression in various tissues, and its functional expression in fish cells transfected with the cystatin cDNA.

## MATERIALS AND METHODS

**Materials**—Total RNA from the liver of chum salmon *Oncorhynchus keta* was extracted with guanidine thiocyanate (13). Poly(A) RNA was purified with oligo(dT)-resin (Oligotex, Takara). Double-stranded cDNA was

prepared using a cDNA synthesis kit (Pharmacia) with oligo(dT) primer and cloned into the *EcoRI* site of lambda gt11 (Pharmacia).

**PCR Cloning**—To isolate a partial cDNA sequence of chum salmon cystatin, PCR was conducted with degenerate oligonucleotide primers having the following sequences: the 20-mer 5' oligonucleotide set 5'-TTGCGCCTNSEN-GARTAYAA-3', corresponding to the amino acid sequence homologous to the family 2 cystatin, FA(M, V, I)(A, E, G)-EYN [residues 7-13 of chicken egg cystatin (8-10)]; the 26-mer 3' oligonucleotide set 5'-GAATTCTTAGTACCNG-CNACNASYTG-3', corresponding to the consensus amino acid sequence QVVAGTN [residues 46-52 of human cystatin B (14)], where Y = C + T, R = A + G, S = G + C, E = C + A + G, N = A + G + C + T. For PCR, 10 ng was taken as the unit amount, so 10 ng of salmon liver cDNA template was mixed with 5  $\mu$ l of 10 $\times$ PCR buffer (0.1 M Tris-HCl [pH 8.3] containing 0.5 M KCl, 15 mM MgCl<sub>2</sub> and 0.1% gelatin), 3  $\mu$ l of dNTP solution (5 mM each dATP, dCTP, dGTP, and dTTP), 1  $\mu$ l each of sense and antisense primer (from 100  $\mu$ M stock solutions), and 0.5  $\mu$ l of Tth DNA polymerase (Toyobo). H<sub>2</sub>O was added to give a total volume of 50  $\mu$ l per sample. The sample was then overlaid with 50  $\mu$ l of paraffin and subjected to 35 cycles of amplification (1 min at 93°C, 1 min at 60°C, and 1 min at 72°C). The amplified fragment was purified by excision from an agarose gel, cloned into the *SmaI* site of pUC118 (Takara), and sequenced.

**cDNA Cloning and Sequencing**—A cDNA library containing 10<sup>4</sup> independent recombinant phages was screened with a PCR-amplified cDNA probe. Hybridization was carried out at 65°C for 16 h in a solution containing 5 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl and 0.015 M sodium citrate, pH 7.6), 5 $\times$ Denhardt's solution, 100  $\mu$ g/ml heat-denatured salmon sperm DNA, 0.1% SDS, and the <sup>32</sup>P-labeled oligonucleotide probe that was prepared by the random primer method (BcaBEST DNA Labeling kit, Takara) (13). After hybridization, the filters were washed by immersion (4 times) in a sufficient volume of 2 $\times$ SSC containing 0.1% SDS at 65°C for 20 min. The filters were then dried and

<sup>1</sup> To whom correspondence should be addressed. Tel: +81-45-788-7643, Fax: +81-45-788-5001, E-mail: mic@nrifs.affrc.go.jp

<sup>2</sup> Present address: SEAFDEC, Changi Point, Singapore 1749, Singapore.

autoradiographed with an intensifying screen. A positive clone covering the 5'-sequence of cystatin was obtained, and its nucleotide sequence was analyzed by subcloning a suitable restriction endonuclease fragment into pUC118 DNA. The DNA sequence of the cDNA was determined with a Sequenase kit (U.S. Biochemical).

**Southern Blot Analysis**—DNA digested with restriction enzymes was fractionated by agarose gel electrophoresis. After transfer of the DNA to nitrocellulose paper, the filter was air-dried and baked at 80°C for 2 h. The DNA was hybridized to a <sup>32</sup>P-labeled cDNA probe prepared by the random primer method (BcaBEST DNA Labeling kit, Takara) at 68°C in 6×SSC, 0.5% SDS, 40 mM PIPES (pH 6.5), 0.02% polyvinylpyrrolidone, 0.02% Ficoll, and 100 µg/ml salmon DNA according to the usual method (13).

**Northern Blot Analysis**—Cells were lysed in 4 M guanidium thiocyanate, and RNA was harvested by pelleting through an ultracentrifuge with a cesium chloride gradient (13). For Northern analysis (13), 10 µg of RNA from each sample was denatured at 50°C in 50% formamide and 2.2 M formaldehyde and separated by 1% agarose-formamide gel electrophoresis. RNA was transferred onto a Hybond-N (Amersham) nylon membrane and fixed. A portion of the cDNA (nucleotides 432–620 of the sequence shown in Fig. 1) amplified by PCR with complementary oligonucleotides was used as a probe template. The RNA was hybridized to a <sup>32</sup>P-labeled cDNA probe prepared by the random primer method (BcaBEST DNA Labeling kit, Takara) at 42°C in 50% formamide, 6×SSC, 0.5% SDS, 40 mM PIPES (pH 6.5), 0.02% polyvinylpyrrolidone, 0.02% Ficoll, and 100 µg/ml salmon DNA according to the usual method (13).

**In Vitro Translation**—An *in vitro* transcription-translation system (TNT T3 promoter system, Promega) was employed as recommended by the manufacturer. The entire salmon cystatin cDNA was inserted into the *Eco*RI site of the pBluescript II SK vector (Stratagene) adjacent to the promoter for T3 RNA polymerase. The reaction was run in a volume of 20 µl with 1 µg of the pBluescript construct, using L-[<sup>35</sup>S]methionine (Amersham, >600 Ci/mmol) as a label. Translation products were analyzed by SDS-PAGE with a 15% gel.

**Expression of the Cystatin cDNA in Fish Cells**—A portion of the cDNA (nucleotides 16–620 of the sequence shown in Fig. 1) was amplified by PCR with complementary oligonucleotides. The amplified product was purified and subcloned into a eukaryotic expression vector downstream from the human cytomegalovirus (CMV) promoter (pCEP4, Invitrogen). Platyfish fibroblast EHS cells (1×10<sup>6</sup> cells) established by Wakamatsu (15) were maintained in Leibovitz L-15 medium (Gibco BRL) containing 10% fetal bovine serum at 28°C and transfected with 10 µg of DNA with a Transfectum according to the manufacturer (Promega). The release of the expressed salmon cystatin into the medium was examined 24 h after transfer to serum-free L-15 medium.

For immunoblotting, rabbit polyclonal antibody was raised to a synthetic peptide, GLIGPMDANMNDQGT-RD, corresponding to amino acid residues 1–18 of the N-terminal sequence of the putative mature form of salmon cystatin (see Fig. 1).

## RESULTS AND DISCUSSION

A cDNA library containing 10<sup>4</sup> independent recombinant phages was screened and a positive clone covering the 5'-sequence of cystatin was obtained. The total nucleotide sequence was determined. The isolated cDNA of chum salmon cystatin consists of 706 nucleotides and a poly(A) tail (Fig. 1). The coding region for cystatin includes 396 nucleotides and corresponds to 132 amino acids. The 3'-noncoding region contains the putative polyadenylation signals, AATAAA, which are found 20, 24, and 128 nucleotides upstream of the poly(A) homopolymer stretch. The 5'- and 3'-non-coding sequences of the salmon cystatin clone are quite different from those of cystatins of other higher vertebrates.

The deduced amino acid sequence from the salmon liver cDNA is also similar to those of the members of cystatin family 2, such as chicken egg cystatin (8), mouse cystatin C (17), rat cystatin C (18), human cystatin C (2), cystatin D

TGACGATTAAAAAAGATATCAAACGGGAAA									
40	50	60	70	80	90				
ATGATCATGGAATGGAAGATCGTCGTTCCCTGCTGCTGCCGTG6CCTTTACGGTGGCAGAAC									
<u>MetIleMetGluTrpLysIleValValProLeuLeuAlaValAlaPheThrValAlaAsn</u>									
100	110	120	130	140	150				
GCCGGTTTGATCGGAGGGCCCCATGGACGCAAAATATGAACGACCAAGGAACGAGAGACGCC									
<u>AlaGlyLeuIleGlyGlyProMetAspAlaAsnMetAsnAspGlnGlyThrArgAspAla</u>									
160	170	180	190	200	210				
CTGCACTTCGCGGTGTGTCGAACACACAAGAAAAACAACGACATGTTTGTCAAGCAGGTG									
LeuGlnPheAlaValValGluHisAsnLysLysThrAsnAspMetPheValArgGlnVal									
220	230	240	250	260	270				
GCCAAGGTTGTCAATGCACAGAAAAAGTGGTATCTGGGATGAAGTACATCTTCAAGTG									
AlaLysValValAsnAlaGlnLysGlnValValSerGlyMetLysTyrIlePheThrVal									
280	290	300	310	320	330				
CAGATGGCGAGGACCCCATGCAAGGAAGGAGGTGTTGAGAAGATCTGCTCCGTGCACAAA									
GlnMetGlyArgThrProCysArgLysGlyGlyValGluLysIleCysSerValHisLys									
340	350	360	370	380	390				
GACCCGCAGATGGCTGCGCCCTACAAGTGCACCTTCGAGGTGTGGAGCCGCCCTGGATG									
AspProGlnMetAlaValProTyrLysCysThrPheGluValTrpSerArgProTrpMet									
400	410	420	430	440	450				
AGCGATATCCAGATGCTCAAGAACCAGTGTAAAGTTAAAGCCCTTAGTTAAGACCCAGTG									
SerAspIleGlnMetValLysAsnGlnCysGluSer***									
460	470	480	490	500	510				
GAGAGAACTTCCAATCAATGTCTAGTCTACCCAATAACTACTATTATCTAGTACTAGTGT									
520	530	540	550	560	570				
TATTTGTTAGTCTTACCAATGCAGTTCGACCTCCCTGTCTAGGGATGTGATTCAAGAA									
580	590	600	610	620	630				
TCGCACT <u>AAATAAT</u> ATGTTGTCAAATGTATTGCATGCCACATTAATATAAGCACTTAATGC									
640	650	660	670	680	690				
AAACATTGCTGCTTGAGAATGTAGTATTAATAAGGCAACAGTTAACTAA <u>ATAATAA</u>									
700	710	720							
ATGTTTGGAATGTTAAAAAAAAAAAAAAAAAAAA									



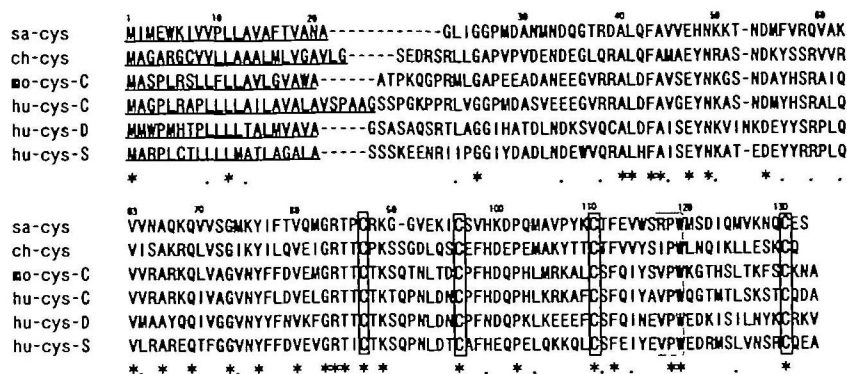


Fig. 2. Comparison of amino acid sequences of chum salmon cystatin (sa-cys) with other members of the cystatin family 2, i.e., chicken cystatin (ch-cys), mouse cystatin C (mo-cys-C), human cystatin C (hu-cys-C), human cystatin D (hu-cys-D), and human cystatin S (hu-cys-S). Asterisks indicate identical amino acid residues among all sequences and the dots indicate conserved residues. Residues corresponding to the reactive center are shadowed and cysteine residues in disulfide bonds are boxed. Signal sequences are underlined.

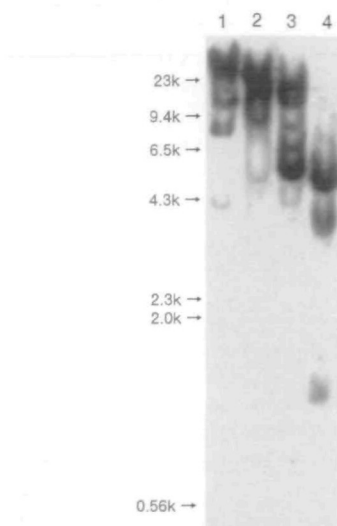


Fig. 3. Southern blot analysis of rainbow trout genomic DNA. Approximately 6  $\mu$ g of the genomic DNA was digested with *Bam*HI (lane 1), *Eco*RI (lane 2), *Hind*III (lane 3), and *Pst*I (lane 4) and separated by electrophoresis on a 1% agarose gel. The DNA was then transferred to nitrocellulose filters, hybridized with  $^{32}$ P-labeled chum salmon cystatin cDNA at 68°C and washed in 2 $\times$  SSC containing 0.1% SDS at 60°C for 20 min. The size markers are from a *Hind*III digest of phage lambda DNA and their sizes are indicated in nts on the left of the autoradiograph.

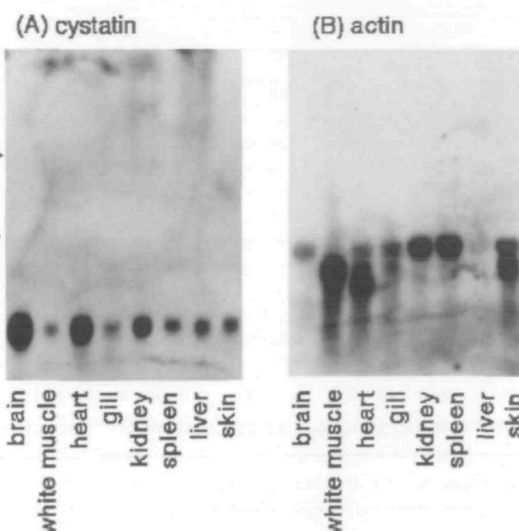


Fig. 4. Analysis of cystatin mRNA in various tissues. The RNA samples were prepared from the brain, white muscle, heart, kidney, liver, skin, spleen, and gill. Each RNA sample (10  $\mu$ g) was separated by 1% agarose-formamide gel electrophoresis, transferred onto a nylon membrane and fixed. A fragment of the 3'-noncoding region (nucleotides 432-620 of the sequence shown in Fig. 1) of chum salmon cystatin cDNA was used as the probe. The blot was reprobed with a human  $\beta$ -actin cDNA probe (Wako) at 37°C and washed in 6 $\times$  SSC containing 0.1% SDS at 55°C for 20 min. The positions of 28 and 18 S rRNA are indicated.

(4), and cystatin S (5); the homologies in total amino acid sequence are 38.6, 34.8, 36.0, 34.1, 28.0, and 31.1%, respectively (Fig. 2). Furthermore, the salmon cystatin has four cysteine residues which are assumed to form disulfide bonds in the C-terminal region. These structures of the salmon cystatins are characteristic of family 2 cystatins.

The deduced amino acid sequence based on the cDNA is very similar to the sequence of cystatin purified from the pituitary of chum salmon (16). When our sequence of the chum salmon cystatin is compared with the published sequence of the mature form of pituitary cystatin (16), 97.3% of the residues are identical; among the residues in the deduced sequence, Arg-117, Asp-122, and Gln-124 were replaced with Ile, Gly, and Lys in the pituitary cystatin, respectively.

Southern blot analysis was performed to determine whether the gene encoding the chum salmon cystatin was present at high or low copy number. The hybridization pattern obtained with the isolated cystatin cDNA insert

shows that chum salmon has several related genes to that encoding the cystatin (Fig. 3).

Among human cystatins, family 2 has several members, such as cystatin C and salivary-type cystatins (2-7). The two forms of the salmon cystatin, i.e., the cystatin encoded by the cDNA isolated in this study and the pituitary cystatin found by Koide and Noso (16), are considered to be genetically distinct members of cystatin family 2 of chum salmon. A detailed study on the genomic structure of chum salmon cystatin genes remains necessary.

Figure 4 shows the results of Northern blots with the various salmon tissues probed with the salmon cystatin cDNA clone. The 189 bp fragment of the 3'-non-coding region was used as a probe, since it shows relatively low homology between cystatin family 1 and family 2. The expression of actin mRNA was tissue-specific, though a cystatin mRNA of 0.7 kb was evident in the RNA isolated from every tissue examined, being most abundant in the



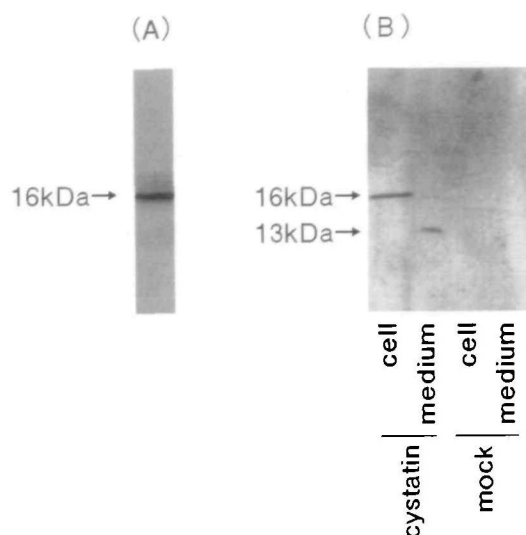


Fig. 5. Expression of chum salmon cystatin. (A) *In vitro* translation product radiolabeled with L-[<sup>35</sup>S]methionine. (B) Platyfish EHS cells were transiently transfected with the cloned chum salmon cystatin cDNA downstream from the human CMV promoter. The expression of the salmon cystatin into the medium was examined 24 h after changing in serum-free L-15 medium. The cystatin produced in the medium or cells was detected by immunoblotting with antiserum against chum salmon cystatin, while no expression of cystatin occur with a mock vector. A 15% gel was used for SDS-PAGE.

brain. Signal intensities of cystatin mRNA of the heart and kidney followed that of the brain in decreasing order. The least amount of cystatin mRNA was found in the gill. Therefore, the salmon cystatin showed apparently non-tissue specific expression.

Recent studies on the gene expression of cystatins in higher vertebrates revealed that human cystatin C (12) and chicken egg cystatin (10) were expressed in most of the tissues, while expression of salivary cystatins (4-7, 10) was restricted to the secretory glands. The chum salmon cystatin was expressed in all tissues tested, especially in the brain, heart, and kidney; therefore the salmon cystatin is considered to be a homolog of mammalian cystatin C. It may have a function in the regulation of proteolytic activity in the body similar to that of mammalian cystatin C.

To confirm that the salmon cystatin cDNA contains a functional open reading frame for cystatin, the total length of the salmon cystatin cDNA was inserted into pBluescript II SK vector adjacent to the promoter for T3 RNA polymerase. When the cDNA construct was subjected to translation *in vitro* in the presence of [<sup>35</sup>S]methionine, a radiolabeled protein of the expected size (16 kDa) was synthesized (Fig. 5A). Furthermore, an expression system of salmon cystatin was constructed in platyfish EHS cells in order to characterize biosynthesis, processing and secretion of cystatin in the fish cells. After changing to a serum-free medium, a 16 kDa preform of cystatin was found to be accumulated in the cells, and a 13 kDa mature cystatin secreted into the medium was detected by immunoblotting with anti-cystatin antiserum (Fig. 5B). The identity between the deduced amino acid sequence from the isolated cDNA in this study and the N-terminal amino acid sequence of pituitary cystatin determined by Koide and Noso (16) suggests that a signal peptidase probably cleaves at Ala-Gly (residues

21-22) to give rise to the mature protein.

These findings indicate that the salmon cystatin cDNA clone is normally expressed in the fish cells. In addition to the fact that the signal sequence was observed in the deduced amino acid sequence of the salmon cDNA, the detection of mature cystatin in the culture medium of the cDNA-transfected cells shows that the salmon cystatin acts as an extracellular type of cystatin in the fish cells. Since the serum-free medium in which the cDNA-transfected cells was cultured showed a weak papain-inhibitory activity (data not shown), cystatin was found to be functionally produced from the cDNA clone.

The properties of salmon cystatin, such as amino acid sequence and gene expression, were similar to those of cystatin C, suggesting that cystatin C is widely distributed in vertebrates. As postulated by Barrett *et al.* (1), the possible defensive role of extracellular-type cystatins in body fluids against the invasion of pathogens and parasites is of special interest. The availability of the salmon cystatin cDNA will facilitate studies to elucidate the mechanisms of physiological functions of cystatin in fish.

## REFERENCES

- Barrett, A.J., Rawlings, N.D., Davies, M.E., Macleidt, W., Salvesen, G., and Turk, V. (1986) Cysteine protease inhibitors of the cystatin superfamily in *Proteinase Inhibitors* (Barrett, A.J. and Salvesen, G., eds.) pp. 515-569, Elsevier, Amsterdam
- Abrahamson, M., Grubb, A., Olafsson, I., and Lundwall, A. (1987) Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C. *FEBS Lett.* **216**, 229-233
- Cole, T., Dickson, P.W., Esnard, F., Averill, S., Risbridger, G.P., Gauthier, F., and Schreiber, G. (1989) The cDNA structure and expression analysis of the genes for the cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin in rat brain. *Eur. J. Biochem.* **186**, 35-42
- Freije, J.P., Abrahamson, M., Olafsson, I., Velasco, G., Grubb, A., and Lopez-Otin, C. (1991) Structure and expression of the gene encoding cystatin D, a novel human cysteine proteinase inhibitor. *J. Biol. Chem.* **266**, 20538-20543
- Saitoh, E., Kim, H.S., Smithies, O., and Maeda, N. (1987) Human cysteine-proteinase inhibitors: Nucleotide sequence analysis of three members of the cystatin gene family. *Gene* **61**, 329-338
- Al-Hashim, I., Dickinson, D.P., and Levine, M.J. (1988) Purification, molecular cloning, and sequencing of salivary cystatin SA-1. *J. Biol. Chem.* **263**, 9381-9387
- Sabatini, L.M., Warner, T.F., Saitoh, E., and Azen, E.A. (1989) Tissue distribution of RNAs for cystatins, histatins, statherin, and proline-rich salivary proteins in humans and macaques. *J. Dent. Res.* **68**, 1138-1145
- Turk, V., Brzin, J., Longer, M., Ritonja, A., Eropkin, M., Borchart, U., and Machleidt, W. (1983) Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin from chicken egg white. *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 1487-1496
- Schwabe, C., Anastasi, A., Crow, H., McDonald, J.K., and Barrett, A.J. (1984) Cystatin. Amino acid sequence and possible secondary structure. *Biochem. J.* **217**, 813-817
- Colella, R., Sakaguchi, Y., Nagase, H., and Bird, J.W. (1989) Chicken egg white cystatin. Molecular cloning, nucleotide sequence, and tissue distribution. *J. Biol. Chem.* **264**, 17164-17169
- Bobek, L.A., Aguirre, A., and Levine, M.J. (1991) Human salivary cystatin S. Cloning, sequence analysis, hybridization *in situ* and immunocytochemistry. *Biochem. J.* **278**, 627-635
- Abrahamson, M., Olafsson, I., Palsdottir, A., Ulvback, M., Lundwall, A., Jensson, O., and Grubb, A. (1990) Structure and

- expression of the human cystatin C gene. *Biochem. J.* **268**, 287-294
13. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
  14. Ritonja, A., Machleidt, W., and Barrett, A.J. (1985) Amino acid sequence of the intracellular cysteine proteinase inhibitor cystatin B from human liver. *Biochem. Biophys. Res. Commun.* **131**, 1187-1192
  15. Wakamatsu, Y. (1981) Establishment of a cell line from the platyfish-swordtail hybrid melanoma. *Cancer Res.* **41**, 679-680
  16. Koide, Y. and Noso, T. (1994) The complete amino acid sequence of pituitary cystatin from chum salmon. *Biosci. Biotechnol. Biochem.* **58**, 164-169
  17. Solem, M., Rawson, C., Lindburg, K., and Barnes, D. (1990) Transforming growth factor beta regulates cystatin C in serum-free mouse embryo (SFME) cells. *Biochem. Biophys. Res. Commun.* **172**, 945-951
  18. Bedi, G.S. (1989) Amino acid sequence of an inducible cysteine proteinase inhibitor (cystatin) from submandibular glands of isoproterenol-treated rats. *Arch. Biochem. Biophys.* **273**, 245-253