AN ABSTRACT OF THE THESIS OF

Xianbin F. Nickel for the degree of Master of Science in Food Science and Technology presented on April 25. Title: Characterization of Pacific Whiting Proteinase P-II and Partial Cloning of Cathepsins L and K cDNA from Rainbow Trout Liver.

Abstract approved by:

Haejung An

Proteinase P-II purified from parasitized Pacific whiting muscle was previously identified to be one form of cathepsin L. It appeared to be present in three isozymatic forms on non-denaturing PAGE gel stained for activity. Its autolytic degradation was observed on SDS-PAGE gel under its optimum condition, 55°C and pH 5.5, in the absence of substrate. Amino acid composition analysis revealed that this enzyme had a considerably greater proportion of hydrophobic amino acids than cathepsin L from other fish species, and monosaccharide analysis showed it was not glycosylated. The N-terminal amino acid sequence of the enzyme was 60-65% identical with cathepsin L from chicken and mammalian species, but only 39% identical with mammalian cathepsin B. The moderate identity of the N-terminal amino acid sequence of P-II with other cathepsin L revealed that this cysteine proteinase from Pacific whiting might be encoded by a cathepsin L-related gene.

Two degenerate primers were designed to clone cathepsins cDNA from rainbow trout. The 500-bp PCR product from rainbow trout liver cDNA contained at least three

different cysteine proteinase sequences, referred to as SFL2, SFL5, and SFL17. SFL5 was the partial cDNA of trout cathepsin L, which was over 80% identical with chicken cathepsin L amino acid sequence. SFL5 was labeled with Dig-11-dUTP and used to screen a trout liver cDNA library. One positive clone referred to as LC was identified and contained a 700-bp insertion overlapping with SFL5. By combining the two overlapping sequences, a 895-bp cDNA sequence was identified, which included 88% of the mature enzyme and a 307-bp 3' end untranslated part. Its deduced amino acid sequences had 83% identity, 91% similarity with chicken cathepsin L and 73% identity, 86% similarity with human cathepsin L. SFL2 might be the partial cDNA of a novel cathepsin L-related cysteine proteinase. SFL17 may be the partial cDNA of trout cathepsin K. It had 70% identity and 89% similarity with rabbit and human cathepsin K at the amino acid level.

Characterization of Pacific Whiting Proteinase P-II and Partial Cloning of Cathepsins L and K cDNA from Rainbow Trout Liver

by

Xianbin F. Nickel

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirement for the
degree of

Master of Science

Completed April 25, 1996
Commencement June 1996

Master of Science thesis of Xianbin F. Nickel presented on April 25, 1996.

APPROVED:

Major Professor, representing Food Science and Technology

Head of the Department of Food Science and Technology

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorized release of my thesis to any reader upon request.

Xianbin F. Nickel, Author

ACKNOWLEDGMENT

I am deeply grateful to Fugen Li for providing trout liver cDNA, the cDNA library, and all his patient technical support. I sincerely thank my major professor, Dr. Heajung An, for her invaluable encouragement and guidance. I also extend my special gratitude to Dr. Michael T. Morrissey for providing me with the opportunity to study at the Seafood Lab, to Lewis Richardson (Daddy Lew) for helping me setup the research facilities, and to Dr. David W. Barnes for his invaluable advice and understanding. I am also grateful to all of my colleagues and others whose names are not mentioned, for their support and advice.

I wish to thank my parents, parents-in-law and my husband for their love and endless encouragement and support.

TABLE OF CONTENTS

	Page
1. INTRODUCTION	1
2. LITERATURE REVIEW	4
Lysosomal Cysteine Protease	4
Biochemical Properties of Cathepsins B, H, L, and S	4
Tissue Distribution of Cathepsins B, L, and S	6
Sequence Homology and Molecular Structure of Cysteine Proteinases	7
Processing of Procathepsins	8
Polymerase Chain Reaction (PCR)	11
cDNA Cloning of Cathepsins	12
Expression and Regulation of Procathepsin L	13
3. CHARACTERIZATION OF PROTEINASE P-II FROM PACIFIC WHITING	15
Abstract	15
Introduction	15
Materials and Methods	17
Materials	17
Enzyme purification	17
Autolysis	18
Native-polyacrylamide gel electrophoresis (PAGE) and activity staining	18
Carbohydrate staining and analysis	18

TABLE OF CONTENTS (Continued)

Page

Amino acid composition analysis 19

N-terminal amino acid sequence 19

Result and Discussion 20

Autolysis 20

Multi-isozymatic forms 20

Amino acid composition 23

Glycosylation 23

N-terminal amino acid sequence 26

4. PARTIAL CLONING OF CATHEPSIN L cDNA FROM RAINBOW TROUT 28

Abstract 28

Introduction 28

Materials and Methods 30

Materials 30

Designing primers for PCR amplification 30

PCR amplification of cathepsin cDNA fragments 31

Separation of PCR products 31

Recovery of 500-bp PCR fragments from the agarose gel 33

Isolation of pUC19 plasmid from *E.coli* 33

Preparation of EcoRI-digested fragment for ligation reaction 33

Subcloning the PCR product into pUC19 34

Nonradioactive probe labeling 34

TABLE OF CONTENTS (Continued)

	Page
Detection of labeling efficiency	35
cDNA library screening	35
Preparation of phage lysate	36
Extraction of phage DNA from lysate and PCR amplification of the insertion	37
Result and Discussion	37
Generation of cathepsin cDNA fragments by PCR	37
SFL5 is the partial sequence of trout liver cathepsin L	38
SFL2 is the partial sequence of a cysteine proteinase gene	49
SFL17 is the partial sequence of cathepsin K	50
5. SUMMARY	56
BIBLIOGRAPHY	58

LIST OF FIGURES

Figure		Page
1.	The primary structure of a cysteine proteinase.	9
2.	The autolysis of P-II on SDS-PAGE in the absence of substrates.	21
3.	The isozymatic forms of P-II on PAGE.	22
4.	Monosaccharide analysis of Pacific whiting P-II.	25
5.	The codon usage of conservative amino acid sequences QGQCGSCW.	32
6.	The codon usage of conservative amino acid sequences WLVKNSWG.	32
7.	PCR amplification of cysteine cathepsins cDNA with degenerate primers.	39
8.	EcoRI-digested plasmid DNAs of 5 white clones showing the 500-bp insertion.	40
9.	Nucleotide sequence and deduced amino acid sequence of clone SFL 2, 7, and 10.	41
10.	Partial cathepsin L cDNA sequence and its deduced amino acid sequence.	42
11.	Nucleotide sequence and deduced amino acid sequence of clone SFL 17.	43
12.	Positive clone LA from secondary screening of cDNA library.	45
13.	PCR products from the six positive clones with λgt11 forward and reverse sequence primers.	46
14.	The restriction maps of SFL 17 and partial trout cathepsin L cDNA sequences.	47

LIST OF TABLES

Table		Page
1.	Amino acid composition of Pacific whiting P-II.	24
2.	Comparison of N-terminal amino acid sequence of Pacific whiting P-II with other cysteine protease superfamily.	27
3.	Comparison of amino acid sequences of partial trout cathepsin L and chicken, human cathepsin L.	48
4.	Alignment of amino acid sequences of SFL 2 and human cathepsin L.	51
5.	Alignment of amino acid sequences of SFL 17, rabbit cathepsin K, and human cathepsin K.	52
6.	Comparison of amino acid sequences of SFL 17, chicken cathepsin L and bovine cathepsin S.	54

1CHARACTERIZATION OF PACIFIC WHITING PROTEINASE P-II AND PARTIAL CLONING OF CATHEPSINS L AND K cDNA FROM RAINBOW TROUT LIVER

Chapter 1. INTRODUCTION

The post-mortem muscle softening observed in vertebrates is generally attributed to the proteolytic degradation of muscle structural proteins. The changes associated with post-mortem tenderization of muscle tissue are weakening of Z-disk followed by myofibril fragmentation, disappearance of troponin-T, desmin, and degradation of myosin and actin (Koohmaraie et al., 1986; Yamashita and Konagaya, 1991a). These changes are normally associated with the activity of proteases, such as cathepsins and calcium-dependent proteases (Yamashita and Konagaya, 1991a).

The phenomenon of muscle autolysis has been observed in several fish species. Cathepsin L (EC 3.4.22.15) was identified for its major role in Pacific whiting (Masaki et al., 1993; An et al., 1994a; Seymour et al., 1994), matured chum salmon (Yamashita and Konagaya, 1990a, b), and mackerel (Lee et al., 1993). Cathepsin L purified from fish shared similarities with those from mammalian tissues in structure and function. Lee et al. (1993) reported that cathepsin L from mackerel resembled mammalian enzymes in molecular weight, carbohydrate content, and many functional properties. Yamashita and Konagaya (1990b) reported that cathepsin L from chum salmon hydrolyzed various protein substrates efficiently, and bond cleavage specificity to insulin B chain was similar to those of rat and mouse cathepsin L. In contrast to these similarites, fish cathepsin L are composed of a single polypeptide, unlike mammalian enzymes which are composed

of light and heavy chains (Kirschke et al., 1977; Mason et al., 1984, 1986; Mason, 1986).

It has been reported that the activity of muscle cathepsin L varies in different fish species (Makinodan et al., 1984) and is associated with different physiological and pathological state. The white muscle of chum salmon during spawning migration shows enhanced catheptic activity and the extensive muscle softening in mature salmon is suggested to be due to the activity of cathepsin L liberated from the phagocytes in the muscle (Yamashita and Konogaya, 1991a, b). The soft texture of Pacific whiting has been associated with an infection of its muscle tissue by the Myxosporean spores (Erickson et al., 1983; Toyohara et al., 1993), and the proteolytic activity varies with the degree and stage of the infection between individual Pacific whiting fish (Konagaya and Aoki, 1981). The commercial importance of muscle softening in fish has generated increased interest in the biochemical and molecular properties of cathepsin L. While the biochemical properties of mammalian cathepsin L has been extensively studied (Kirschke et al., 1977; Okitani et al., 1980; Mason et al., 1984, 1985; Mason, 1986), information on the fish muscle enzyme is still limited. In addition, previous studies related to the abnormal expression of cathepsin L have almost been confined to mammalian diseases (Gottesman, 1978; Doherty et al., 1985; Tryggvason et al., 1987), and there has not been a report on enhanced expression of fish cathepsin L in the molecular genetics field. Even though cathepsin L or cathepsin L-related cysteine proteinase cDNA gene has been cloned from mammals, anthropods, and some parasites (Ishidoh et al., 1987; Joseph et al., 1988; Laycock et al., 1991; Heussler et al., 1994; Mallinson et al., 1994), no cathepsin gene has been cloned from a fish source. To better understand the

post-mortem muscle softening and relationship between molecular structure and function of fish cathepsin L, the objectives of the study were to characterize the molecular properties of cathepsin L from Pacific whiting and clone cathepsin L cDNA from fish.

Chapter 2. LITERATURE REVIEW

Lysosomal Cysteine Proteinase

Cathepsins are lysosomal proteinases which include the cysteine proteinases (cathepsins B, H, L, and S), serine proteinase (cathepsin G), and aspartyl proteinase (cathepsin D) (Kirschke et al., 1980). Lysosomes are the principal sites of intracellular turnover. They maintain a pH of about 5 in their interior and contain about 40 types of acid hydrolases, including proteases, nucleases, and glycosidases (Alberts et al., 1994).

Cathepsins have been considered to be the most active proteinases responsible for intracellular protein turnover (Barrett and Kirschke, 1981). They have been participated in many different cellular events, including regulatory peptide release, receptor internalization, antigen processing, enzyme activity regulation, and secretory protein processing (Katunuma and Kominami, 1983; Bond et al., 1987; Devi, 1991; Lafuse et al., 1995). Due to their vital role in cellular protein metabolism regulation, cathepsins have been found to be related with various diseases, such as, cancer, arthritis, muscle dystrophy and Alzheimer's disease (Katunuma and Kaminami, 1983; Tryggvason et al., 1987; Golde et al., 1992; Petanceska and Devi, 1992).

Biochemical Properties of Cathepsins B, H, L, and S

Cathepsins B, H, L, and S are major enzymes in the group of lysosomal cysteine proteinases (Mason et al., 1984; Kirschke et al., 1989). They are identified by their ability to degrade small peptides and by their broad tissue distribution (Barrett, 1981). Cathepsins B and L have been purified from the liver and the muscle of numerous

mammalian and fish species (Kirschke et al., 1977; Kirschke et al., 1980; Mason et al., 1984, 1985; Okitani et al., 1988; Yamashita and Konagaya 1990a, c; Lee et al., 1993; Jiang et al., 1994; Seymour et al., 1994); and cathepsin S has been purified from bovine lymph nodes (Turnsek et al., 1975), and rabbit spleen (Kirschke et al., 1986; Maciewicz and Etherington, 1988).

Cathepsins B and L are similar in several properties (Okitani et al., 1988). In investigations of proteolytic degradation of the extracellular matrix, cathepsins B and L have been shown to be capable of cleaving acid-insoluble collagen and to play an important role in tumor invasion (Kirschke et al., 1982; Tryggvason et al., 1987). Both enzymes can cleave the synthetic substrate, Z-Phe-Arg-NMec (Barrett, 1980; Kirschke, 1982). However, cathepsin B can be distinguished from cathepsin L by its ability to cleave Z-Arg-Arg-NMec which is resistant to cathepsins H and L (Mason et al., 1984). Meanwhile, cathepsin L has higher endopeptidase activity than cathepsin B (Barrett and Kirschke, 1981). Cathepsin L has been reported to be the most active cysteine proteinase in degrading various protein substrates, such as azocasein, casein, collagen, and myofibrillar proteins (Okitani et al., 1980; Seymour et al., 1994).

Cathepsin H is an endoaminopeptidase which has endo- and aminopeptidase activity in the same order of magnitude (Kirschke, 1977). It differs from other lysosomal cysteine proteinases in its capacity to hydrolyse aminopeptidase as well as endopeptidase substrates. It can catalyze the hydrolysis of unblocked amino acid derivatives such as L-Arg-NMec which cathepsin B or L cannot. However, cathepsin H is unable to degrade collagen and Z-Phe-Arg-NMec (Kirschke et al., 1980; Mason et al., 1984).

Cathepsin S has some similarities to cathepsin L (Kirschke et al., 1989). For some time it was thought that differences between these two enzymes might be ascribed to species and/or tissue variations in the primary structure (Ritonja et al., 1991). Its bond specificity in the cleavage of the oxidized-insulin B chain is partly similar to that of cathepsin L, and it has the same specificity to the synthetic substrates as cathepsin L (Bromme et al., 1989; Kirschke et al., 1989). But comparison of the primary sequences of bovine cathepsin S with that of bovine cathepsin L have clearly indicated that the enzymes are structurally different (Ritonja et al., 1991). While cathepsin L has been reported to be glycoprotein, neither a glycosylated residue nor a potential glycosylation site was found in bovine cathepsin S (Mason et al., 1984; Ritonja et al., 1991). Even though cathepsin S was found to hydrolyze proteins at a similar rate to cathepsin L at pH below 7.0, cathepsin S retained the most activity at pH 7.0-7.5 (Bromme et al., 1989; Kirschke et al., 1989). Within this latter range of pH, cathepsin L was found to be completely inactive. In addition, cathepsin S has different kinetic properties toward some inhibitors and some synthetic substrates other than Z-Phe-Arg-NMec (Bromme et al., 1989).

Tissue Distribution of Cathepsins B, L, and S

The expression levels of rat cathepsins B, L, and S in different tissues were investigated by Northern blot analysis (Petanceska and Devi, 1993). It was found that the expression of cathepsins B and L was totally different from the expression of cathepsin S. Cathepsin L was expressed in the highest level in the kidney, comparable level in the spleen, brain, liver, and the lowest level in the lung. Cathepsin B was

expressed in high level in the kidney and low level in the heart and is fairly uniform in all other organs. Cathepsins B and L are distinguished from cathepsin S by their expression in all tissues, despite some variations found in the levels. Cathepsin S expression varied dramatically between tissues and was at much lower levels than cathepsins B and L. Cathepsin S was expressed at the highest level in the ileum followed by the spleen and brain. The expression levels of cathepsin S in the liver, kidney and jejunum were 6-9 fold lower than that in the ileum. This is in general agreement with the Western blot result of bovine cathepsin S distribution in some organs, except that cathepsin S was not detected in bovine liver (Kirschke et al., 1989).

Sequence Homology and Molecular Structure of Cysteine Proteinases

Cathepsins B, H, L, and S belong to the group of closely related proteins of the papain superfamily (Ritonja et al., 1991). Papain is a plant thiol proteinase which has been purified from the papaya fruit and studied extensively. Its amino acid sequence has been determined, and its molecular structure has been refined to a resolution of 1.65 Å (Kamphuis et al., 1984). The primary structure of cathepsins B, H, L, and S from several species has been published (Segundo et al., 1985; Ishisoh et al., 1987; Wiederanders et al., 1991; Lafuse et al., 1995). The comparison of the amino acid sequences of cysteine proteinases demonstrates a striking homology among their sequences. The amino acids in the active sites Cys^{25}, His^{159}, Asn^{175} and Trp^{177}, and the sequence around them are highly conserved in all cysteine proteinases. The N-terminal region (residues 1-70 in papain) and C-terminal region (residues 118-212 in papain) display the highest sequence homologies, whereas the lowest sequence homologies are

observed in the middle region (residues 71-117 in papain), a segment where most insertions/deletions are observed (Dufour, 1988). The highest sequence homology is found between cathepsins L and S, and cathepsins L and S have a higher homology with cathepsin H than with cathepsin B (Petanceska and Devi, 1992; Shi et al., 1992).

Based on the high similarity in the sequences and striking identity in the catalytic sites between papain and cathepsins, cathepsins B, H, L, and S are presumed to have similar three-dimensional structure and catalytic mechanisms to papain (Katunama and Kominami, 1983; Takio et al., 1983; Dufour, 1988). Like papain, cathepsins can be arbitrarily divided into three regions: N-terminal, C-terminal, and central regions (Takio et al., 1983). The N-terminal region contains the highly conservative sequence C-G-S-C-W-A-F, where Cys^{25} is the active-site residue. The C-terminal region contains Y-W-I-V-K-N-S-W, where Asn^{175} and Trp^{177} are the active-site residues. According to X-ray data of papain, the active-site Cys^{25} is nearby the active-site His^{159}; Asn^{175} is hydrogen-bonded to His^{159}, and Trp^{177} shields this bond from the solvent (Dufour, 1988). Fig.1 gives an outline of the structure of papain superfamily, showing some of the key conserved residues involved in the active sites (North et al.,1990).

Processing of Procathepsins

Many lysosomal proteins are synthesized as propeptides containing either N-terminal or C-terminal extensions that are removed during or soon after delivery to lysosomes (Tao et al., 1994). The precursor of rat liver cathepsin B encoded by cDNA clones appears to be extended at both ends and its C-terminal extension is a six amino acid peptide (Segundo et al., 1985). It has been proposed that all lysosomal enzymes

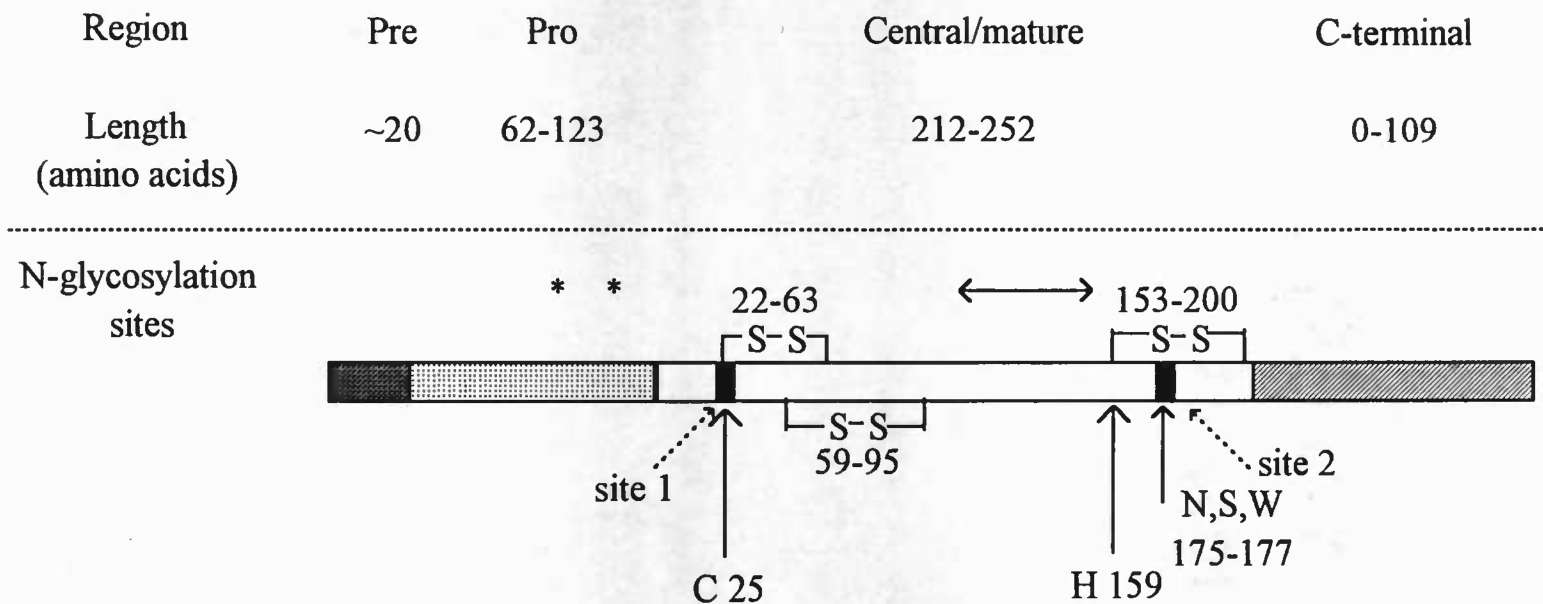

Fig. 1. The primary structure of a cysteine proteinase. The diagram gives an outline of the structure of a member of the papain superfamily, showing some of the key conserved residues involved in the active site (numbers refer to papain residues). ⟷ represents the region in mature enzyme containing known N-glycosylation sites; * represents putative N-glycosylation sites in some proregions. Site 1 and Site 2 refer to conserved sequences often used to design oligonucleotides (North et al.,1990).

may be synthesized with a transient C-terminal sequence that could function as a sorting sequence, being cleaved after completion of targeting (Erickson and Blobel, 1983). The cDNA of mouse cysteine proteinase, which is murine preprocathepsin L (Joseph, 1988), has an N-terminal extension of a 113 amino acid peptide. The first 17 amino acids constitute a signal peptide cleaved co-translationally after Ala-(-97). The following 96 amino acids are the pro-region between the signal sequence and the conserved amino terminus of mature cysteine proteinases (Portnoy et al., 1986), which render the enzyme inactive (Mason et al., 1987, 1989). For cathepsin L and other cathepsins, the propeptide may play a role in intracellular trafficking and regulation of the enzymatic activity of these proteins (McIntyre and Erickson, 1991).

Cathepsins B and L purified from mammalian liver are composed of a heavy and light chain (Kirschke et al., 1977; Mason et al., 1984, 1986; Mason, 1986). These two chains may be formed by limited proteolytic cleavage in the C-terminal part of a single polypeptide chain (Katunuma and Kominami, 1983; Wada et al., 1987). The postulated cleavage site is found on the surface when the primary structure of cathepsin B is fitted to the tertiary structure of papain (Katunuma and Kominami, 1983). The existence of single-chain of mammalian muscle cathepsins and integrated cathepsin cDNAs supports this hypothesis (Chan et al., 1986; Gal and Gottesman, 1988; Okitani et al., 1988).

The processing involved in the conversion of procathepsin L to cathepsin L has been investigated *in vitro* on dextran sulfate and *in vivo* by a pulse-chase method. Pulse-chase experiments with different types of cells suggest that the enzyme is translated in a prepro-form, processed into the proenzyme in the rough endoplasmic reticulum, and localized in lysosomes as the active enzyme (Kominami et al.,1988; Nishimura et al.,

1988). The *in vitro* study demonstrated that cathepsin L is processed autocatalytically by multi-step and the four amino acid residues at the C-terminal in the propeptide function to prevent the activation of processed cathepsin L (Ishidoh and Kominami, 1994).

Polymerase Chain Reaction (PCR)

PCR is used to amplify a segment of DNA that lies between two regions of known sequence by repeated cycles of denaturation, priming and extension (Old and Primrose, 1989). It is so sensitive that a single DNA molecule can be amplified, and single-copy genes are routinely extracted out of complex mixtures of genomic sequences (Innis et al., 1990). Since the announcement of PCR for DNA amplification (Saiki et al., 1985), numbers of modifications, improvements, and new applications of PCR technology have been made. For example, reverse transcription PCR (RT-PCR) has been used for sensitive detection and sequence analysis of RNA starting with extremely small amounts of template (Kawasaki et al., 1990). Rapid amplification of cDNA ends PCR (RACE-PCR) allows to amplify copies of the region between a known region in the transcript and the unknown 3' or 5' end (Loh, 1989). Marathon^{TM} PCR for cDNA amplification provides a fast and easy way to clone full-length cDNAs without cDNA library screening (Chenchik et al., 1996). PCR has been extensively applied in gene cloning, probe labeling, DNA sequencing, and diagnosis of human genetic diseases.

cDNA Cloning of Cathepsins

To date, many cathepsin cDNA genes have been cloned from the cDNA library by hybridization with probes. By using the synthetic degenerate oligonucleotides as a probe deduced from a short chain of amino acids of cathepsins, the first cathepsin B cDNA gene was isolated from the rat liver cDNA library (Segundo et al., 1985), and then from the human liver cDNA library (Fong et al., 1986).

The cDNA sequences from closely related genes and/or species were used as probes to screen the cDNA library for cathepsins. A cDNA clone for rat cathepsin L was identified with two oligonucleotide probes based on the amino acid sequence of rat cathepsin H (Ishidoh et al., 1987). By using a nick-translated 950-bp rat cathepsin B cDNA fragment as a probe and washing under the reduced-stringency condition, human and mouse preprocathepsin B cDNAs were cloned (Chan et al., 1986). Similarly, a cDNA for human pro-cathepsin L was purified with a mouse cDNA fragment as a probe (Gal and Gottesman, 1988).

PCR amplification of cDNA fragments using degenerate oligonucleotides has been shown to be a powerful tool to isolate cDNA within a reasonably short period of time (Petanceska and Devi, 1992). The PCR product from the cDNA of interest can be used as a perfectly matched probe for that gene, so stringent hybridization conditions can be used to eliminate the spurious hybridization signals (Lee et al., 1990). PCR has been applied very successfully to clone serine and cysteine proteinase gene fragments from a number of parasites and mammals (Sakanari et al., 1989; North et al., 1990; Petanceska and Devi, 1992; Rosenthal and Nelson, 1992; Shi et al., 1992; Heussler and Dobbelaere, 1994; Mallinson et al., 1994). The degenerate oligonucleotides have been based upon

the two highly conserved regions around the active site residues Cys^{25} and Asn^{175}. The spacing of these conserved sequences in cysteine proteinase cDNAs (about 500 bp) makes them particularly suitable for amplification by PCR (North et al., 1990). The oligonucleotides containing all possible nucleotide combinations deduced from the amino acid sequences are more than 1000 fold. To minimize the degeneracy of the pool of oligonucleotides when human macrophage cathepsin S was cloned, the codon usage of known human and bovine cathepsin cDNAs for these sequences was employed (Shi et al., 1992). When cysteine proteinase genes were cloned from the parasites, liver fluke (*Fasciola hepatica*), trophozoite (*Plasmodium falciparum*), and protozoa (*Trichomonas vaginalis*), deoxyinosine was inserted when any of four deoxynucleotides was possible at the third position of a given codon, and RACE-PCR was used to amplify the 5' and 3' regions of the cDNAs (Rosenthal et al., 1992; Heussler et al.,1993; Mallinson et al., 1994).

Compared to the traditional labor-intensive, time-consuming cDNA-library screening method, PCR has proved to be a powerful tool for cloning full-length cysteine proteinase cDNAs. The full-length avian cathepsin B cDNA has been obtained from an osteoclast λgt11 cDNA library, by screening with PCR reaction using a gene-specific primer paired with a λ arm sequence primer (Dong et al., 1995).

Expression and Regulation of Procathepsin L

Cathepsin L has been found to be associated with a number of physiologic and pathologic processes including cell proliferation (Gottesman, 1978; Doherty et al., 1985), tumor invasion (Tryggvason et al., 1987), spermatogenesis (Zabludoff et al., 1990),

embryo implantation (Conliffe et al., 1995), and arthritis (Trabandt et al., 1990). The processing and regulation of its expression have been studied extensively (Kominami et al., 1988; Troen et al., 1991; Nishimura and Kato, 1992; Homma et al., 1994; Tao et al., 1994). The genomic genes of mouse and human cathepsin L have been cloned and localized on chromosomes (Troen et al., 1991; Chauhan et al., 1993), and their promoters, transcription sites and intron-exon splice junctions have been studied.

The recombinant procathepsin L has been expressed in *E.coli*, yeast, and mammalian cell systems for the studies of its function in the disease processes and the relationship between the structure and function (Smith and Gottesman, 1989; Nishimara and Kato, 1992; Tao et al., 1994). A cDNA clone encoding human cathepsin L was expressed at high levels in *E. coli.*, and an active enzyme with similar kinetic properties to the human liver enzyme was obtained after renaturation (Smith and Gottesman, 1989). Structure-function studies were performed by construction mutations in either the propeptide portion or the C-terminal light chain portion of the protein. It was shown that the propeptide is essential for proper enzyme folding and/or processing in the renaturation system, and the disulfide bond joining the light and heavy chains is essential for enzymatic activity (Smith and Gottesman, 1989). It was also proved that the proregion of cathepsin L is required for proper folding, stability, and endoplasmic reticulum exit *in vivo*, when the mouse cathepsin L cDNA and its mutants with the altered proregion were expressed in COS cells (Tao et al., 1994).

Chapter 3. CHARACTERIZATION OF PROTEINASE P-II FROM PACIFIC WHITING

Abstract

The N-terminal amino acid sequence of P-II, purified from Pacific whiting muscle and previously identified to be cathepsin L, is 61% identical with the sequences of chicken cathepsin L and 65% identical with cathepsins L and S from mammalian species, but only 39% identical with cathepsin B from rat and human. Amino acid composition analysis shows that P-II contains a considerably greater proportion of hydrophobic amino acids than cathepsin L from other species. Carbohydrate staining and monosaccharide composition analysis reveals no glycosylation in P-II. Three isozymatic forms of the enzyme were shown on non-denaturing PAGE gel stained for proteolytic activity. The autolysis of P-II, which was observed on SDS-PAGE in the absence of substrate, may contribute to its enzymatic stability.

Introduction

Pacific whiting (*Merluccius productus*) is the most abundant fishery resource off the northwest coast of the United States (Radtke, 1995). Its commercial utilization has been limited due to its soft texture and poor keeping quality associated with proteolysis of the muscle structure proteins (Konagaya and Aoki, 1981; Erickson et al., 1983). The protease activity in its muscle has been shown to correlate with the degree of the infection of its muscle tissue by the Myxosporean parasite, *Kudoa paniformis* (Patashnik et al., 1992; Toyohara et al., 1993; Seymour et al., 1994). At present, the majority of harvested Pacific whiting is used for the production of surimi.

Surimi is the raw material in seafood analogs and is prepared from minced and washed fish flesh. Myosin, the major component in surimi, has been shown to be rapidly degraded in muscle softening during cooking process (Chang-Lee et al., 1989; Wasson et al., 1992; Morrissey et al., 1993). Several heat-stable proteases have been shown to contribute to the thermal softening of muscle. These include alkaline proteases (Folco et al., 1984; Makinodan et al., 1987; Boye and Lanier, 1988; Stoknes et al., 1993), calpains (Jiang et al., 1991; Tsuchiya and Seki, 1991), and cathepsins B and L (Hara et al., 1988; Sherekar et al., 1988; Matsumiya et al., 1989; Yamashita and Konagaya, 1990a, c; Lee et al., 1993; An et al., 1994b). The enzymes were shown to be heat-stable and active at neutral pH. Among these, cathepsin L has been identified for its major role in muscle autolysis of Pacific whiting surimi (Masaki et al., 1993; An et al., 1994b; Seymour et al., 1994).

Two forms of proteinases (P-I and P-II), identified to be cathepsin L, have been purified from the sarcoplasmic fluid of parasitized Pacific whiting (Seymour et al., 1994). P-I was an inhibitor-complexed form which was composed of a single polypeptide of M_r 28,800 and other two low M_r components (An et al., 1995). P-II was the free form of enzyme which was also a single polypeptide of M_r 28,800 on SDS-PAGE gel. Acid-treated P-I without the two low M_r components had the same isoelectric point as P-II, which was 4.99 with a minor band at 5.02. While the previous work has revealed the biochemical properties of the enzymes, little has been known about its molecular properties. Therefore, the objective of this study was to characterize proteinase P-II of Pacific whiting with respect to molecular properties.

Materials and Methods

Materials

Pacific whiting were caught off the Oregon coast, stored in slurry ice and off-loaded at a local surimi processing plant. Whole fish were transferred to Oregon State University Seafood Lab, filleted, vacuum-packed and kept frozen at -20°C until used.

The electrophoresis unit and reagent grade, including acrylamide, bis-acrylamide, sodium dodecyl sulfate (SDS), Coomassie brilliant blue R-250, were purchased from Bio-Rad Laboratories, Hercules, CA. The reagents for carbohydrate staining were purchased from Sigma Chem. Co., St. Louis, MO., including dansyl hydrazine, periodic acid, dimethyl sulfoxide (DMSO), sodium metabisulfite, thymol, ovalbumin (Sigma A2512) and chymotrypsinogen (Sigma C4879). Sodium caseinate was purchased from U.S. Biochemical Corp., Cleveland, OH.

Enzyme purification

Proteinase P-II was purified from the sarcoplasmic fluid of Pacific whiting by the method of Seymour et al. (1994). Fraction P-II, free of inhibitors, was isolated by heat treatment and chromatography on butyl-Sepharose and DEAE. The fractions were pooled and concentrated by ultrafiltration using a Centri-Prep 10 cartridge (Amicon, W.R. Grace & Co.-Conn., Beverly, MA) and stored at −80°C until used.

Autolysis

The concentration of purified enzyme was analyzed by Lowry's assay (Lowry et al., 1951). P-II (24 μg) was mixed with an equal volume of McIlvaine buffer (0.2 M sodium phosphate, 0.1 M citrate and 1 mM sodium azide, pH 5.5), and incubated at 55°C for 5, 30, and 60 min. Samples were analyzed for degradation fragments on 15% SDS-PAGE gel by the method of Laemmli (1970). Low molecular weight standards (Pharmacia, Piscataway, NJ) included phosphorylase b (94,000), albumin (67,000), ovalbumin (43,000), carbonic anhydrase (30,000), trypsin inhibitor (20,000), and α-lactalbumin (14,400).

Native-polyacrylamide gel electrophoresis (PAGE) and activity staining

Duplicate PAGE minigels (7.3 x 8.3 x 0.75 cm) were loaded with 7.5 μg of enzyme and run by the method of Chrambach and Rodbard (1971). One gel was stained with Coomassie blue R-250, and the other was stained for proteolytic activity according to the method of García-Carreño et al. (1993). For activity staining, the gels were incubated in 50 ml of 2% sodium caseinate in 50 mM Tris-HCl buffer (pH 7.5) on ice for 1 hr and in 50 ml of 2% (w/v) sodium caseinate in McIlvaine buffer (pH 5.5) at 55°C for 45 min followed by staining with Coomassie blue R-250.

Carbohydrate staining and analysis

Glycosylation of the enzyme was tested by staining for carbohydrate on SDS-PAGE gels by thymol-sulfuric acid method (Racusen, 1979) and fluorescent method (Eckhardt et al., 1976). A glycoprotein, ovalbumin, and a non-glycoprotein,

chymotrypsinogen, were used as positive and negative controls, respectively. The monosaccharide composition of the purified enzyme was analyzed by Glyco, Inc. (Novato, California).

Amino acid composition analysis

Amino acid composition of purified enzyme was determined after acid hydrolysis in 6 N HCl containing 0.05% mercaptoethanol and 0.02% phenol at 115°C for 20 hrs. Serine was increased by 10% and threonine by 5% to compensate for destruction by acid. For determination of cysteine, samples were oxidized with performic acid prior to acid hydrolysis, and cysteine content was calculated from cysteic acid/alanine ratio. For determination of tryptophan, alkaline hydrolysis for 48 hrs at 135°C was carried out by the method of Hugli and Moore (1972). Amino acid analyses were performed on a Beckman amino acid analyzer model 6300 updated to a model 7300 using the sodium buffer system, column, and protocol recommended by the manufacturer. Data were collected and analyzed using Beckman's System Gold software.

N-terminal amino acid sequence

N-terminal sequence was determined by automated gas-phase Edman degradation using an Applied Biosystems 475A gas phase sequencer. Cysteine residue was determined after conversion to 4-vinylpyridine by the method of Andrews and Dixson (1987).

Result and Discussion

Autolysis

Cathepsin L has been shown to be much more stable in inactivated form than activated form during storage (Barrett and Kirschke, 1981). SDS-PAGE analysis of reactions revealed evidence of autolytic degradation (Fig. 2), when purified proteinase P-II was incubated under the optimum activity condition, 55°C and pH 5.5, in the absence of substrate. The protein band corresponding to the proteinase was noticeably smaller after 5 min and almost completely removed after 60 min. It was hypothesized that the loss of activity was mainly due to autolytic degradation (Barrett and Kirschke, 1981). It would seem likely that conditions which produced the highest rates of catalysis, pH 5.5-6.0 at 55°C, would also produce the highest rates of inactivation by autolysis. These results may not apply to situations where substrate is present.

Multi-isozymatic forms

While the purified proteinase P-II was a single band on SDS-PAGE gel (Seymour et al., 1994), it appeared as multiple bands on protein and activity staining of PAGE gel (Fig. 3). PAGE analysis showed two forms of proteinase P-II, corresponding to the major band and the minor band on PAGE gel stained with Coomassie blue R-250. This result is accordance with the previous report that Pacific whiting cathepsin L has pI point of 4.99 with a minor band at 5.02 (Seymour et al., 1994). Activity staining of PAGE gel revealed the third form of the enzyme, possibly due to the higher sensitivity. Kirschke et al. (1977) reported multiple forms of cathepsin L from rat liver lysosomes,

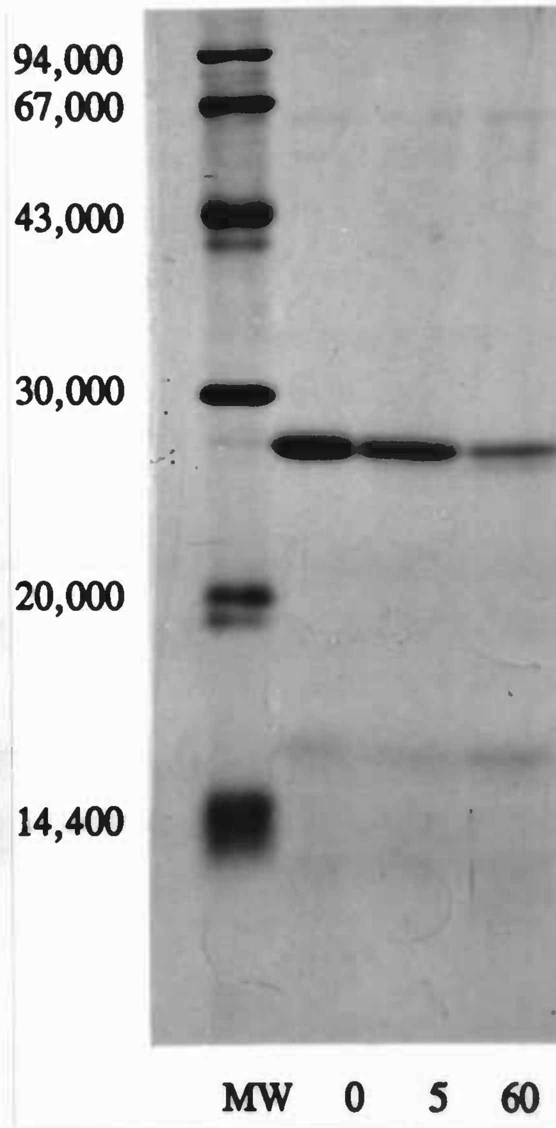

Fig. 2. The autolysis of P-II on SDS-PAGE in the absence of substrates. P-II (24μg) was loaded after incubation at 55°C and pH 5.5 for 0, 5, and 60 min.

Mason et al. (1984) also reported three main isoenzymatic forms of rabbit liver cathepsin L. Existing in multi-isozymatic forms might be one of the molecular properties of proteinase P-II.

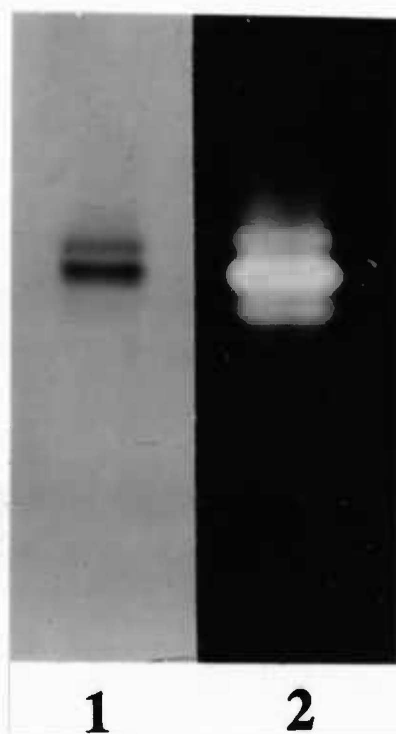

Fig. 3. The isozymatic forms of P-II on PAGE. Lane 1: PAGE gel stained directly with Coomassie blue R250; Lane 2: PAGE gel stained for activity.

Amino acid composition

Table 1 shows that Pacific whiting P-II has high hydrophobic amino acid content, including Ile, Leu, Phe, Pro, Trp, Tyr, Val. It contained much less lysine, and a considerably greater proportion of hydrophobic amino acids (42%) than comparable enzymes from other fish species. Cathepsin L from chicken and rat contained 36.1% and 37.0% hydrophobic residues, respectively (Towatari and Katunuma, 1983; Wada et al., 1987). Chum salmon cathepsin L (Yamashita and Konagaya, 1990a) and chum salmon cathepsin B (Yamashita and Konagaya, 1990b) contained 38 and 34% hydrophobic residues, respectively. The relatively high hydrophobicity may contribute to strong binding of cathepsin L to myofibrillar proteins during processing of Pacific whiting by hydrophobic interactions.

It was reported that cathepsin L was the major cysteine proteinase associated with myofibrillar proteins in surimi after extensive washing of muscle mince (An et al., 1994a). Myosin has been shown to contain high-affinity hydrophobic binding sites (Borejdo et al., 1983).

Glycosylation

Proteinase P-II showed no carbohydrates on SDS-PAGE gels stained by thymol-sulfuric acid and fluorescent methods. Monosaccharide analysis detected only glucose, which is usually from contaminating sources (Fig. 4). Both the carbohydrate staining and monosaccharide analysis data demonstrated that Pacific whiting P-II is not glycosylated. Kirschke et al. (1982), Mason et al. (1984) and Lee et al. (1993) reported that cathepsin L from rat liver, rabbit liver and mackerel (*Scomber australasicus*) muscle was not completely

Amino Acids	Amino acid residues	Integral numbers
Ala	21.8	22
Arg	5.49	5
Asx	28.74	29
Cys/2	6.59	7
Glx	24.86	25
Gly	26.19	26
His	6.8	7
Ile	17.8	18
Leu	17.2	17
Lys	8.1	8
Met	2.02	2
Phe	9.2	9
Pro	14.4	14
Ser	25.8	26
Thr	16.1	16
Trp	7.02	7
Tyr	21.0	21
Val	20.1	20
Total	279.2	279
Molecular weight	30,658	30,282

Table 1. Amino acid composition of Pacific whiting cathepsin P-II.

adsorbed by Con A-Sepharose, the specific activities of both glycosylated and unglycosylated forms of enzyme were very similar to each other. P-II from parasitized Pacific whiting muscle may be mainly unglycosylated.

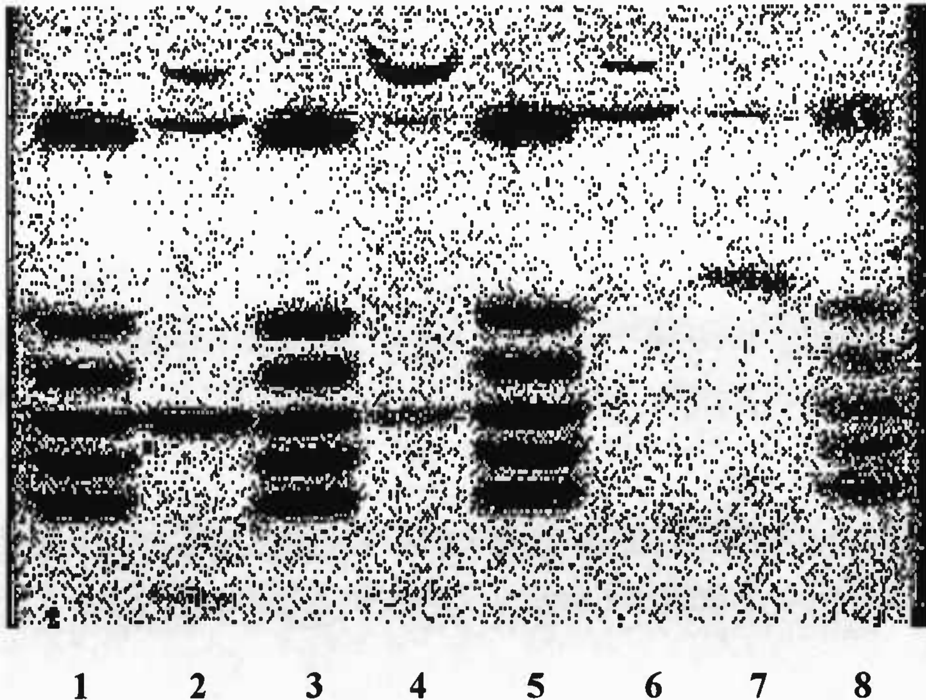

Fig. 4. Monosaccharide analysis of Pacific whiting P-II (Note: this picture was obtained form Glyco, Inc., Novato, CA). Lanes 1, 3, 5, and 8, monosaccharide ladder standards; Lane 2, amine hydrolysis; Lane 4, neutral hydrolysis; Lane 6, sialic acid hydrolysis; Lane 7, N-acetylneuraminic acid (NANA) standard.

N-terminal amino acid sequence

The N-terminal sequence was obtained for 23 of 25 amino acids, representing 10% of the total sequence. This sequence was aligned with the highly conserved N-terminal sequences of members of the cysteine proteinase superfamily (Table 2). The sequence was 65% identical with human and mouse cathepsin L (Mason et al., 1986), and 61% identical with chicken cathepsin L (Dufour et al., 1987). The sequence was also 65% identical with sequences of human cathepsin S, and 61% identical with papain (Mitchel et al., 1970), but only 39% identical with cathepsins B from rat liver (Takio et al., 1980) and human liver (Segundo et al., 1985). In the sequence near the active-site cysteine (residue 25), 9 out of 11 residues were identical with human cathepsin L, but only 5 of 13 with human cathepsin B. The presence of Cys^{22} indicates the possibility of a Cys^{22}-Cys^{63} disulfide bond existing in Pacific whiting proteinase P-II, as it does in papain (Mitchel et al., 1970).

					5					10							15							20					25
PW P-II	X	X	L	P	L	S	V	D	W	N	A	I	G	-	-	K	V	T	-	-	S	V	K	N	Q	G	K	C	G
Human CL		E	A	P	R	S	V	D	W	R	E	K	G	-	-	Y	V	T	-	-	P	V	K	N	Q	G	Q	C	G
Mouse CL	L	K	I	P	K	S	V	D	W	R	E	K	G	-	-	C	V	T	-	-	P	V	K	N	Q	G	Q	C	G
Human CS	R	I	L	P	D	S	V	D	W	R	E	K	G	-	-	C	V	T	-	-	E	V	K	Y	Q	G	S	C	G
Chicken CL			A	P	R	S	V	D	W	R	E	K	G	-	-	Y	V	T	-	-	P	V	K	D	Q	G	I	C	G
Papain			I	P	E	Y	V	D	W	R	Q	K	G	-	-	A	V	T	-	-	P	V	K	N	Q	G	S	C	G
Human CB			L	P	A	S	F	D	A	R	E	Q	W	P	Q	C	P	T	I	K	E	I	R	D	Q	G	S	C	G

Table 2. Comparison of N-terminal amino acid sequence of Pacific whiting P-II with other cysteine protease superfamily (human cathepsin L, mouse cathepsin L, human cathepsin S, chicken cathepsin L papain and human cathepsin B)

Chapter 4. PARTIAL CLONING OF CATHEPSIN L cDNA FROM RAINBOW TROUT

Abstract

The PCR product amplified from trout liver cDNA contained at least three different sequences, referred as to SFL2, SFL5, and SFL17. SFL5 was the partial cDNA sequence of trout cathepsin L (TCL) and was used to screen trout liver cDNA library. A 895-bp partial TCL has been identified, and its deduced amino acid sequence shared 83% identity with chicken cathepsin L. SFL2 might be the partial cDNA of a novel cathepsin L-related cysteine proteinase. It was 51% identical with Cysteine Proteinase 2 from *D. discoideum* and 44% identical with human cathepsin L at the amino acid level. SFL17 may be the partial sequence of cathepsin K. Its deduced amino acid sequence had 70% identity and about 89% similarity with rabbit and human cathepsin K over about 70% of mature enzyme.

Introduction

Cathepsin L is a lysosomal cysteine proteinase which belongs to the papain superfamily (Kirschke and Barrett, 1987). It exists extensively in many tissues and cells and plays an important role in metabolic and developmental processes, such as protein degradation (Barrett and Kirschke, 1981), bone resorption (Tezuka et al., 1994), sperm maturation (Zabludoff et al., 1990), embryo implantation (Conliffe et al., 1995), and hormone activation (Benjannet, 1991; Thomas et al., 1991).

Cathepsin L has been found to be involved in many diseases, and consequently more scientists are interested in the study of the function of cathepsin L. A cathepsin

L-like cysteine proteinase of *Plasmodium falciparum* trophozoites has been shown to be an essential malarial hemoglobinase, and it has become a potential target for antimalarial chemotherapy (Rosenthal and Nelson, 1992). High levels of cathepsin L expression have been found in many human tumors, and it has suggested that cathepsin L may play an important role in tumor invasion (Chauhan et al., 1993). The activation of cathepsins B and L in interleukin-6 transgenic mice has caused muscle atrophy (Tsujinaka et al., 1995) due to their high activity toward muscle structural proteins such as collagen and myosin (Kirschke, 1982).

Cathepsin L has been identified to play the major role in muscle softening of Pacific whiting (An et al., 1994b; Seymour et al., 1994), mackerel (Lee et al., 1993), and spawning salmon (Yamashita and Konagaya, 1990b, c). The activity of cathepsin L has been shown to be associated with the physiological state (Yamashita et al., 1990, 1991) and the degree of the infection by parasites (Konagaya et al., 1981; Erickson et al., 1983). Cloning the cathepsin L gene could be a first step towards understanding of the relationship of structure and function, regulation of expression in immune reaction and during developmental stages of the fish. While cathepsin L cDNA has been isolated from various sources, i.e., human (Gal and Gottesman, 1988; Joseph et al., 1988), rat (Ishidoh et al., 1987), mouse (Portnoy et al., 1986), lobster (Laycock et al., 1991), flesh fly (Homma et al., 1994), liver flute (Heussler and Dobbelaere, 1994), no cathepsins have been cloned from fish. The objective of this study was to clone and sequence cathepsin L cDNA from rainbow trout liver to elucidate the primary structure of cathepsin L from fish.

Materials and Methods

Materials

Rainbow trout (*Oncorhynchus mykiss*) liver cDNA and its cDNA library were kindly provided by Fugen Li at the Department of Food Science and Technology, Oregon State University. Taq DNA polymerase and other PCR core reagents were purchased from Perkin-Elmer Cetus, Foster, CA. Agarose was purchased from FMC, Rockland, Maine. The GENECLEAN II Kit was from BIO 101 Inc., La Jolla, CA. The DH5α competent cells and pUC19 plasmid were purchased from Gibco BRL Life Technologies, Gaithersburg, MD. EcoRI restriction enzyme, T4 DNA ligase, 5-brome-4-chloro-3-indolyl-β-D-galactoside (X-Gal), isopropyl-β-thiogalactopyranoside (IPTG) were purchased from Promega Corporation, Madison, WI. Digoxigen-3-O-methylcarbonyl-ϵ-aminocaproyl-[5-(3-aminoallyl)-2'-deoxy-uridine-5' triphosphate] tetralithium salt (DIG-11-dUTP), blocking reagent, anti-Digoxigenin-AP Fab fragments, disodium 3-(4-methoxyspiro{1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.1.1^{3,7}] decan}-4-yl) phenyl phosphate (CSPD) used for chemiluminescent labeling and detection were from Boehringer Mannheim, Indianapolis, IN. Nylon membranes were from Du Pont, Boston, MA. All other chemicals were purchased from Sigma Chem. Co., St. Louis, MO.

Designing primers for PCR amplification

The degenerate PCR 5' primer and 3' primer for amplifying partial cathepsin cDNA fragments from rainbow trout liver were designed based on the amino acid sequence of two highly conservative regions around C^{25} and N^{175} in the papain

superfamily. To reduce the degeneracy of the primers, the codon usage of cathepsin L cDNA from several species were referred (Figs. 5 and 6). An EcoRI site was attached to the 5' end of each primer. The sequence of 5' primer (sense) is: acc gaa ttc cag ggc (c/a)ag tg(c/t) gg(c/t) tcn tg(c/t) tgg, and the sequence of 3' primer (antisense) is: acc gaa ttc cc cca g(g/c)(a/t) gtt ctt nac na(g/a) cca.

PCR amplification of cathepsin cDNA fragments

The PCR reaction was performed in a final volume of 50 μl containing 1 μl (50 ng/ml) cDNA template, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 1 μM each primer, 0.002% (w/v) gelatin, and 1.25 units of the AmpliTaq polymerase. The reactions were carried out in a thermocycler (Ericomp, San Diego, CA) for 30 cycles of denaturation (94°C, 50 sec), annealing (48°C, 2 min), and extension (70°C, 1.5 min). The products were held at 70°C for 5 min after the reaction.

Separation of PCR products

The PCR products were fractionated on a 1.3% low melting-temperature agarose gel by electrophoresis under 100 V on ice. The gels were stained with 10 μl ethidium bromide solution (10 mg/ml) in 100 ml deionized distilled water (ddH_2O) for 30 min, and destained with two changes of ddH_2O. The gels were observed under UV light at 254 nm.

Conservative amino acid sequences	Q	G	Q	C	G	S	C	W
Flesh fly CL	cag	ggc	cag	tgc	ggt	tca	tgc	tgg
Shrimp CL	cag		cag	tgc	ggc	tcc	tgc	tgg
Lobster CL	cag	gga	cag	tgt	ggc	tcc	tgc	tgg
Human CL	cag	ggt	cag	tgt	ggt	tct	tgt	tgg
Rat CL	cag	ggc	cag	tgt	ggt	tct	tgc	tgg
Mouse CL	cag	ggc	cag	tgc	ggg	tct	tgt	tgg
Liver fluke CL	cag	gga	cag	tgt	ggt	tcc	tgt	tgg
Blood fluke CL	cag	ggt	cag	tgt	gga	tca	tgt	tgg

Fig. 5. The codon usage of conservative amino acid sequences QGQCGSC^{25}W (papain numbering) in flesh fly cathepsin L, shrimp cathepsin L, lobster cathepsin L, human cathepsin L, rat cathepsin L, mouse cathepsin L, liver fluke cathepsin L, blood fluke cathepsin L.

Conservative amino acid sequences	W	L	V	K	N	S	W	G
Flesh fly CL	tgg	ttg	gtt	aag	aac	tct	tgg	gg
Shrimp CL	tgg	cgt	gtc	aag	aac	tcg	tgg	
Lobster CL	tgg	ctc	gtc	aag	aac	tct	tgg	gg
Human CL	tgg	ctg	gtg	aag	aac	agc	tgg	gg
Rat CL	tgg	ctt	gtc	aaa	aac	agc	tgg	gg
Mouse CL	tgg	ctt	gtc	aag	aac	agc	tgg	gg
Liver fluke CL	tgg	att	gtg	aaa	aat	agt	tgg	gg
Blood fluke CL	tgg	ctt	att	aaa	aat	agt	tgg	gg

Fig. 6. The codon usage of conservative amino acid sequences WLVKN^{175}SWG (papain numbering) in flesh fly cathepsin L, shrimp cathepsin L, lobster cathepsin L, human cathepsin L, rat cathepsin L, mouse cathepsin L, liver fluke cathepsin L, blood fluke cathepsin L.

Recovery of 500-bp PCR fragment from the agarose gel

The 500-bp PCR product was excised from the agarose gel and recovered by using the GENECLEANII kit according to the manufacturer's instruction. The DNA was eluted into 20 μl ddH_2O.

Isolation of pUC19 plasmid from *E.coli*

The pUC19 plasmid DNA was extracted by large-scale alkali method (Sambrook et al., 1989) from 100 ml *E. coli* overnight culture. The nucleic acid pellet was dried and dissolved in 1 ml TE (pH 8.0). The extracted plasmid DNA was digested with 0.25 mg/ml RNase A at 37°C for 1 hr, and extracted with saturated phenol, phenol:chloroform:isoamyl alcohol 25:24:1, and chloroform. The supernatant was precipitated with 1 volume of isopropanol. The pellet was washed with 70% ethanol, dried and dissolved in 500 μl ddH_2O.

Preparation of EcoRI-digested fragment for ligation reaction

About 22 μg pUC19 plasmid DNA was digested with 30 units EcoRI in 30 μl solution, and about 8 μg of 500-bp PCR fragment was treated with 40 units EcoRI in 40 μl solution at 37°C for 2.5 hrs. DNA concentration was determined spectrophotometerically by the method of Sambrook et al. (1989). The digested fragments were recovered from the agarose gel as previously described.

Subcloning the PCR product into pUC19

About 2 μg EcoRI-digested PCR fragment and about 1.2 μg EcoRI pUC19 fragment were ligated with 9 units of T4 DNA ligase in 25 μl solution at 16°C overnight. The reaction solution (15 μl) was loaded onto 1% agarose gel to verify the ligation reaction. The reaction solution (2 μl) was transformed into *E. coli* competent cell according to the manufacturer's instruction. The transformation solution (80 μl) was spread on Luria-Bertani medium (LB) plates with 100 μg/ml ampicillin, 5 μg/ml X-gal, and 40 μg/ml IPTG. After incubation at 37°C for 16 hrs, 26 white colonies were picked up and transferred into 2 ml LB liquid solution containing 100 μg/ml ampicillin.

The transformant plasmid DNA was extracted from the overnight culture by the mini-prep boiling lysis method (Sambrook et al., 1989) and digested with EcoRI. Seventeen clones were found to have the 500-bp insertion. The plasmid DNA from these clones were extracted from 3 ml culture by the mini-prep alkali lysis method (Sambrook et al., 1989) and purified with GENECLEANII kit by the modified method (after NEW WASH, the GLASSMILK/DNA complex was washed with 80% ethanol twice, and then eluted with 30 μl ddH_2O). Among the 17 clones obtained, five were sequenced by dye-primer method at the Gene Center Lab, Oregon State University. The deduced amino acid sequences from the DNA sequences were put into SwissProt data base in BLAST program for protein homology search.

Nonradioactive probe labeling

DIG-11-dUTP was incorporated into the 500-bp partial cathepsin L cDNA by PCR method under the above described condition except 52°C annealing temperature and

66 μM DIG-11-dUTP, 134 μM dTTP and 200 μM dATP, dGTP, and dCTP each in a 100 μl final volume. The labeled products were purified either by ethanol precipitation or agarose gel extracted using the GENECLEAN II kit.

Detection of labeling efficiency

The purified probe was serially diluted and 1 μl of each diluted solution was applied to nylon membranes. The DNA was bound to the membrane by UV cross linking for 3 min under 254 nm UV light. The chemiluminescent detection reaction with CSPD as a substrate was performed according to the procedure provided by the manufacturer. The membrane was exposed to X-ray film for 50 min and developed with developer for 5 min, rinsed with water for 30 sec, and fixed for 10 min in fixer.

cDNA library screening

E.coli strain LE392 was used to prepare plating cells according to the procedure of Sambrook et al. (1989). cDNA λgt11 phage solution (2.5 μl) was mixed with 60 μl phage buffer and 250 μl plating cell and incubated at 37°C for 20 min. Melted LB top agarose (6 ml) was added to the mixture and poured to a prewarmed dry 150 mm plate. The plates were incubated at 37°C for 10-12 hrs and chilled for at least one hour before plate lifting.

The phage DNA was transferred to nylon membranes in duplicate, denatured, and neutralized as described by Sambrook et al. (1989). DNA was immobilized on the membrane by UV cross linking under 254 nm UV light for 5 min and then baked for 30 min at 80°C. The membranes were prehybridized in 50% formamide, 2% blocking

reagent, 5 x SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0), 0.1% N-lauroylsarcosine, and 0.02% SDS hybridization buffer at 37°C for about 3 hrs. The gel-purified probe was denatured by heating at 100°C for 10 min and chilled on ice water for 10 min. The membranes were hybridized at 37°C for 12 hrs in hybridization buffer with 10 ~ 100 ng/ml labeled probe. The membranes were washed 2 times for 5 min at room temperature with at least 50 ml of 2 x SSC and 0.1% SDS per 100 cm^2 filter and 2 times for 15 min at 68°C with 0.1 x SSC and 0.1% SDS. The hybridization signals were detected as described in section "Detection of labeling efficiency". Positive clones were stored in 1 ml SM buffer with 1 drop of chloroform at 4°C as described in Sambrook et al. (1989). They were diluted serially and plated on 90 mm plates and secondarily screened under the similar condition, except that hybridization was performed at 42°C, and the probe was the direct PCR product which was not further purified. Single colonies from each strain were picked up and stored in 1 ml SM buffer with a drop of chloroform at 4°C.

Preparation of phage lysate

One single colony of LE392 was transferred into 5 ml of LB medium containing 0.2% maltose and 10 mM $MgSO_4$ and cultured overnight at 37°C with constant shaking. The overnight culture (100 μl) was added to a tube containing 40-80 μl of the phage plug elute and incubated at 37°C for 20 min. The infected culture was transferred into 20 ml of prewarmed (37°C) LB medium supplemented with 10 mM $MgSO_4$ and shaken at 200 rpm for 7 hrs. If lysis did not occur, 30 ml of prewarmed LB with $MgSO_4$ was added into the culture and shaken for an additional 3 hrs at 300 rpm. Chloroform (250 μl) was

added and the culture was shaken for 25 min. The lysate was centrifuged at 8000 x g for 10 min, and the supernatant was transferred to a sterile tube.

Extraction of phage DNA from lysate and PCR amplification of the insertion

The phage DNA was extracted from 25 ml of lysate as described (Sambrook et al., 1989). The DNA was dissolved into 40 μl TE (pH 8.0) buffer and 20 μl of phage DNA was digested with EcoRI. Phage DNA (1 μl) was amplified by 30 cycles of PCR reaction with the λgt11 forward and reverse sequence primers. The DNA was denatured at 94°C for 5 min followed by the 30 cycles of denaturation (94°C, 50 sec), annealing (55°C, 1.5 min), and extension (72°C, 1.5 min), and held at 70°C for 5 min. The PCR product was digested with EcoRI, subcloned into pUC19 and subjected to sequencing.

Result and Discussion

Generation of cathepsin cDNA fragments by PCR

The alignment of the amino acid sequence of a number of cysteine proteinases showed that the regions around active site residues C^{25} and N^{175} were highly conserved among animal cathepsins as well as among the plant and fungal cysteine proteinases (Dufour, 1988). The regions, $CGSC^{25}WAF$ and $YWLVKN^{175}SW$ were usually chosen for designing the primers (North et al., 1990; Shi et al., 1992; Petanceska et al., 1992). To minimize the degeneracy of the oligonucleotide primers for PCR, several considerations were taken. Firstly, the regions of $QGQCGSC^{25}W$ (N-terminal) and $WLVKN^{175}SWG$ (C-terminal) were chosen for designing the primers. Secondly, the

codon usage of cathepsin L from several species was compared. Among the available cathepsin L cDNA sequences from mammalian, parasite and anthropod, the degenerate primers for cloning trout cathepsin L (TCL) cDNA were designed mainly based on the codon usage of mammalian cathepsin L. Third, almost all of the possible nucleotide sequences in the 3' region were included in the primer and fewer possible sequences in the 5' region were included. The sequence of the 3' end of each primer is more important than that of the 5' end for specific binding between the primer and the template during PCR. Fourth, an EcoRI digestion site was attached to each primer in order to conveniently subclone the PCR fragment. Three extra nucleotides, ACC, were added to the 5' end of each EcoRI site to protect the ends. Both 33-mer 5' primer with 64-fold degeneracy and 32-mer 3' antisense primer with 128-fold degeneracy were synthesized and used for PCR. The PCR product showed mainly one band, which was about 500-bp long (Fig. 7). The 500-bp fragment was subcloned into pUC19 and screened by blue-white selection. Among the 26 white colonies screened, 17 were found to have the 500-bp insertion (Fig. 8). Five clones, referred to as SFL2, SFL5, SFL7, SFL10 and SFL17, were chosen and sequenced, and 3 different sequences were obtained (Figs. 9, 10, and 11). SFL2, 7, and 10 contained the same insertion.

SFL5 is the partial sequence of trout liver cathepsin L

The deduced amino acid sequence of subclone SFL5 has 82% identity with chicken cathepsin L and was labeled as a probe for trout liver cDNA library screening.

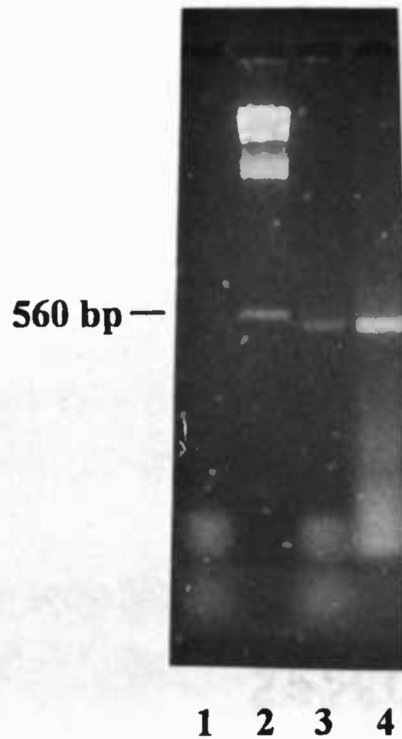

Fig. 7. PCR amplification of cysteine cathepsins cDNA with degenerate primers. Lane 1, negative control (without cDNA template); Lane 2, molecular weight standard (λ/HindIII); Lane 3, first amplification from cDNA; Lane 4, second amplification from the first PCR product.

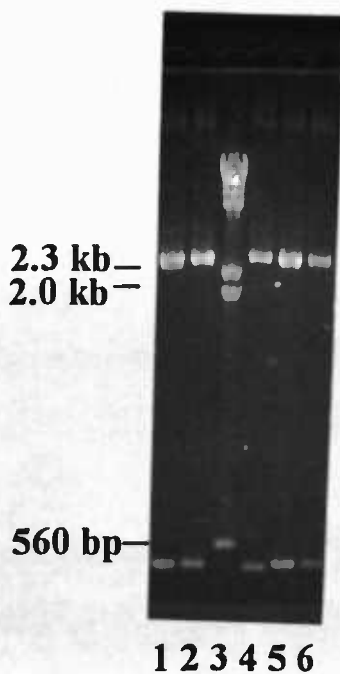

Fig. 8. EcoRI-digested plasmid DNAs of 5 white clones showing the 500-bp insertion. lanes 1, 2, 4, 5, 6, EcoRI-digested plasmid DNAs of 5 white clones; Lane 3, molecular weight standard (λ/HindIII).

1	AGC	TTT	GCC	ACC	ACT	GGA	GCA	GTA	GAG	GGC
1	S	F	A	T	T	G	A	V	E	G
31	GCT	CTC	TTC	CTG	AAA	TCA	GGG	TCC	CTC	CAG
11	A	L	F	L	K	S	G	S	L	Q
61	GTC	TTG	TCC	CAG	CAG	ATG	TTG	GTA	GAT	TGT
21	V	L	S	Q	Q	M	L	V	D	C
91	AGC	TGG	GGC	TTT	GGT	AAT	AAC	GGA	TGT	GAT
31	S	W	G	F	G	N	N	G	C	D
121	GGA	GGA	GAG	GAG	TGG	AGG	GCC	TAC	GAG	TGG
41	G	G	E	E	W	R	A	Y	E	W
151	ATC	ATG	AAA	CAC	GGA	GGC	ATC	GCC	ACT	ACG
51	I	M	K	H	G	G	I	A	T	T
181	GAG	ACC	TAT	GGT	TCC	TAC	ATG	GGC	ATG	AAC
61	E	T	Y	G	S	Y	M	G	M	N
211	GGC	CTG	TGC	CAC	CTC	AAC	ACT	TCT	CAG	CTG
71	G	L	C	H	L	N	T	S	Q	L
241	ACA	GCT	CGT	GTC	CAG	AGC	TAC	ACC	AAT	GTG
81	T	A	R	V	Q	S	Y	T	N	V
271	ACG	TCA	GGT	GAT	GCC	GAG	GCG	CTG	AAG	GTG
91	T	S	G	D	A	E	A	L	K	V
301	GCG	TTG	TTT	AAA	CAT	GGC	CCA	GTG	GCC	GTC
101	A	L	F	K	H	G	P	V	A	V
331	AGT	ATT	GAC	GCA	GGA	CAT	AGA	TCC	TTT	GTC
111	S	I	D	A	G	H	R	S	F	V
361	TTC	TAC	AGC	CAC	GGG	ATC	TAC	TAT	GAA	CCC
121	F	Y	S	H	G	I	Y	Y	E	P
391	AAA	TGT	GGA	AAC	ACA	ACT	GAC	TCC	CTG	GAC
131	K	C	G	N	T	T	D	S	L	D
421	CAC	GCG	GTG	CTT	GCA	GTG	GGT	TAT	GGT	GTC
141	H	A	V	L	A	V	G	Y	G	V
451	ATG									
151	M									

Fig. 9. Nucleotide sequence and deduced amino acid sequence of clone SFL2, 7, and 10.

1	GCG	TTC	AGC	ACC	ACC	GGG	GCC	ATG	GAG	GGC
1	A	F	S	T	T	G	A	M	E	G
31	CAG	CAG	TTC	AGG	AAG	ACT	GGC	AAG	CTG	GTG
11	Q	Q	F	R	K	T	G	K	L	V
61	TCT	CTG	AGT	GAA	CAG	AAC	CTG	GTG	GAC	TGC
21	S	L	S	E	Q	N	L	V	D	C
91	TCC	AGA	CCG	GAG	GGC	AAC	GAG	GGC	TGT	AAC
31	S	R	P	E	G	N	E	G	C	N
121	GGT	GGG	CTC	ATG	GAC	CAG	GCC	TTT	CAG	TAC
41	G	G	L	M	D	Q	A	F	Q	Y
151	ATC	CAG	GAC	AAT	GCC	GGC	CTG	GAC	ACA	GAA
51	I	Q	D	N	A	G	L	D	T	E
181	GAG	TCC	TAC	CCC	TAC	GTC	GGC	ACT	GAT	GAG
61	E	S	Y	P	Y	V	G	T	D	E
211	GAC	CCT	TGC	CAC	TAC	AAG	CCA	GAG	TTC	AGT
71	D	P	C	H	Y	K	P	E	F	S
241	GCT	GCC	AAC	GAG	ACT	GGC	TTT	GTG	GAC	ATC
81	A	A	N	E	T	G	F	V	D	I
271	CCC	AGT	GGC	AAG	GAG	CAT	GCT	ATG	ATG	AAG
91	P	S	G	K	E	H	A	M	M	K
301	GCT	GTG	GCT	GCA	GTC	GGT	CCT	GTC	TCT	GTT
101	A	V	A	A	V	G	P	V	S	V
331	GCC	ATC	GAT	GCC	GGC	CAC	GAG	TCC	TTT	CAG
111	A	I	D	A	G	H	E	S	F	Q
361	TTC	TAT	GAG	TCT	GGG	ATC	TAC	TAT	GAG	AAG
121	F	Y	E	S	G	I	Y	Y	E	K
391	GAG	TGC	AGC	AGT	GAG	GAG	TTG	GAC	CAT	GGT
131	E	C	S	S	E	E	L	D	H	G
421	GTT	CTA	GTG	GTG	GGA	TAT	GGT	TTT	GAA	GGA
141	V	L	V	V	G	Y	G	F	E	G
451	GAA	GAT	GTG	GAT	GGC	AAG	AAA	TAC	TGG	ATT
151	E	D	V	D	G	K	K	Y	W	I
481	GTC	AAG	AAC	AGC	TGG	AGT	GAG	AAA	TGG	GGA
161	V	K	N	S	W	S	E	K	W	G
511	GAC	AAA	GGC	TAT	ATC	TAC	ATG	GCC	AAA	GAC
171	D	K	G	Y	I	Y	M	A	K	D
541	AGG	AAG	AAC	CAC	TGT	GGT	ATC	GCC	ACG	GCA
181	R	K	N	H	C	G	I	A	T	A
571	TCC	AGC	TAC	CCA	CTG	GTC	TAG	TGTTTTAGAACCTG		
191	S	S	Y	P	L	V	End			

606 GATGTGATCTAGACTTTTTTTTTTTATGTTTGAAAAATGTTCATAAAGGG
656 CAATGATGACTACAGTGCTGCCTTATTAAGTCTGTGAAAGGCACTAGTGA
706 ATCCTGCTGCATGGGTTTTAAAGACCGTTTTAGGGAATCTATGAAATTCT
756 ACCTAGTTTGTTTTTATACTTGATTTGATATATATATAGTTGAGTCAAGT
806 AAATGTTACAGGTGTGACCTCTGTGTACATTCCTAGTTTACTGCTTTTAC
856 ACCTGTTCTGTAAAAATGTTGTACATATGAGCAAAATAAAC

Fig. 10. Partial cathepsin L cDNA sequence and its deduced amino acid sequence.

1	GCC	TTC	AGC	TCT	GCA	GGG	GCC	CTG	GAG	GGC
1	A	F	S	S	A	G	A	L	E	G
31	CAG	CTG	GCA	AGG	ACC	ACA	GGC	AAA	CTG	ATA
11	Q	L	A	R	T	T	G	K	L	I
61	GAC	CTC	AGC	CCA	CAG	AAC	CTG	GTG	GAC	TGT
21	D	L	S	P	Q	N	L	V	D	C
91	GTC	ACT	GAG	AAC	AAT	GGC	TGT	GGT	GGA	GGC
31	V	T	E	N	N	G	C	G	G	G
121	TAC	ATG	ACC	AAC	GCC	TTC	GAA	TAC	GTT	GAG
41	Y	M	T	N	A	F	E	Y	V	E
151	GAA	AAC	GGA	GGC	ATC	GAC	ACA	GAG	GAG	GCT
51	E	N	G	G	I	D	T	E	E	A
181	TAC	CCT	TAC	CTT	GGC	CAG	GAT	GAG	CAG	TGT
61	Y	P	Y	L	G	Q	D	E	Q	C
211	GTC	TAC	AAC	GCG	TCT	GGC	ATG	GGT	GCT	CAG
71	V	Y	N	A	S	G	M	G	A	Q
241	TGT	CGC	GGG	TTC	AAG	GAG	ATC	CCT	GAG	GGA
81	C	R	G	F	K	E	I	P	E	G
271	GAC	GAG	TGG	GCA	CTG	ACC	AAG	GCT	GTA	GTC
91	D	E	W	A	L	T	K	A	V	V
301	AAA	GTG	GGG	CCT	GTG	GCT	GTG	GGC	ATT	GAT
101	K	V	G	P	V	A	V	G	I	D
331	GCC	ACC	CTC	TCC	ACC	TTC	CAA	TTC	TAC	CAG
111	A	T	L	S	T	F	Q	F	Y	Q
361	AGA	GGC	GTG	TAC	TAC	GAC	CCC	AAC	TGC	AAC
121	R	G	V	Y	Y	D	P	N	C	N
391	AAG	GAT	GAC	ATC	AAC	CAC	GCC	GTG	CTT	GCA
131	K	D	D	I	N	H	A	V	L	A
421	GTG	GGC	TAT	GGA	CAA	ACT	GCC	AAG	GGT	GTG
141	V	G	Y	G	Q	T	A	K	G	V
451	AAA	TTC								
151	K	F								

Fig. 11. Nucleotide sequence and deduced amino acid sequence of clone SFL17.

About 6 x 10^5 plaques were screened, and 6 positives with 600 ~ 800 bp insertions were identified after secondary screening (Figs. 12 and 13). Two of the largest ones, termed as LA and LC, were sequenced and LC was found to contain a 300-bp-long overlapping sequence with SFL5. Combining the sequence of SFL5 and LC, a 895-bp nucleotide sequence composed of 307 bp of the 3' untranslated part and about 88% of the mature part of TCL cDNA was obtained (Fig. 10). This sequence contains active site His (TCL 139), the highly conserved region YWIVKNSW (TCL 158-165) in papain family, and one stop codon. It has also one putative N-glycosylation site (N-X-S/T) (TCL N 83). There is one AATAAA sequence in the end of the 3' untranslated region. Unlike rat cathepsin B cDNA which has a six-amino-acid-long transient C-terminal sequence (Segundo et al., 1985), no C-terminal extension region was found in TCL cDNA. The restriction map of this sequence is shown in Fig. 14.

The deduced amino acid sequence of the identified partial TCL cDNA is 83% identical, 91% similar to chicken cathepsin L and 73% identical, 86% similar to human cathepsin L (Table 3). The N-terminal (TCL 1-54) and the C-terminal (TCL 137-196) of partial TCL have about 92% identity with chicken cathepsin L, while the central part (55-136) has 71% identity with chicken cathepsin L. The missing part of mature TCL belongs to the highly conservative N-terminal region in papain superfamily and may have a very high identity with chicken cathepsin L. Because the N-terminal amino acid sequence of the cysteine proteinase P-II from Pacific whiting has only 65% identity with human cathepsin L and 61% identity with chicken cathepsin L, this enzyme might be a cathepsin L-related cysteine proteinase, which is expressed specifically under the parasite-infected condition. The sequences of cathepsin L-related genes share close similarity

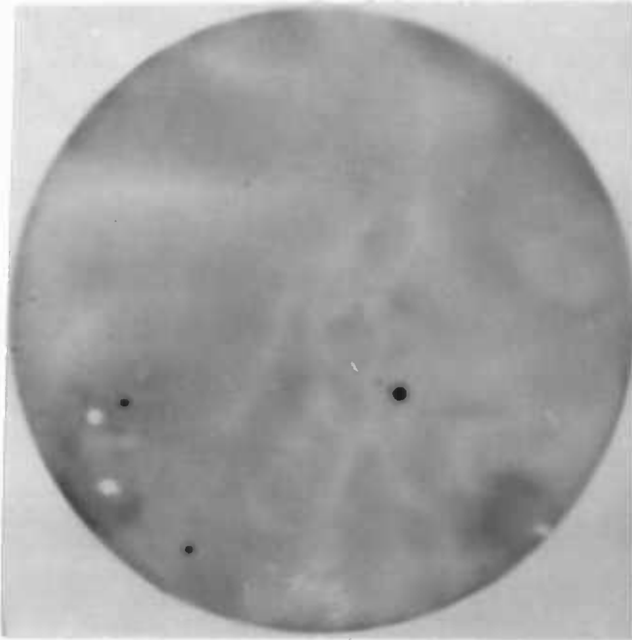

Fig. 12. Positive clone LA from secondary screening of cDNA library.

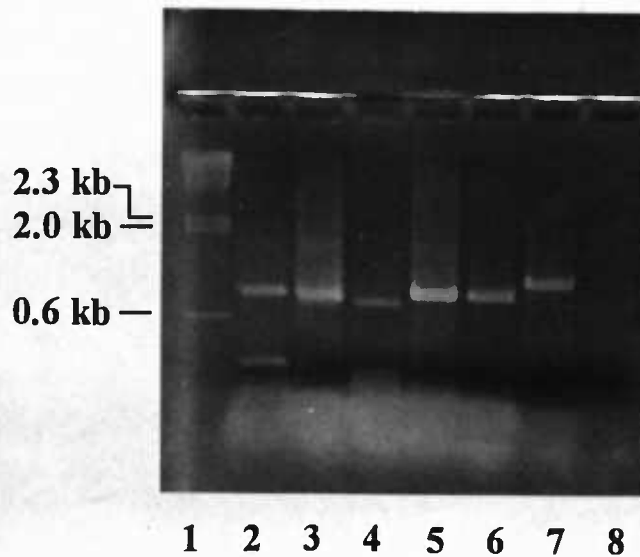

Fig. 13. PCR products from the six positive clones with λgt11 forward and reverse sequence primers. Lane 1, molecular weight standard (λ/HindIII); Lane 2, clone LF as template DNA; Lane 3, clone LE as template DNA; Lane 4, clone LD as template DNA; Lane 5, clone LC as template DNA; Lane 6, clone LB as template DNA; Lane7, clone LA as template DNA; Lane 8, negative control (without template DNA).

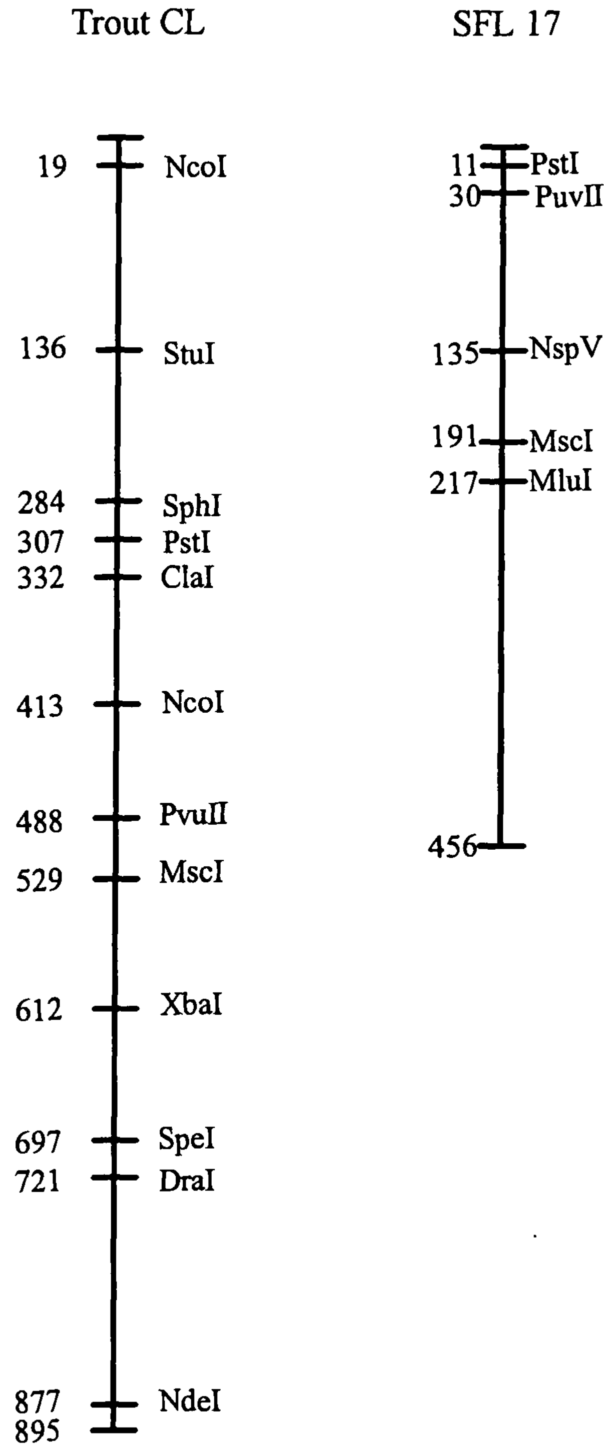

Fig. 14. The restriction maps of SFL 17 and partial trout CL cDNA sequences.

```
Trout CL       1   AFSTTGAMEGQQFRKTGKLVSLSEQNLVDCSRPEGNEGCNGGLMDQAFQY
Chicken CL    27   AFSTTGALEGQHFRKTGKLVSLSEQNLVDCSRPEGNQGCNGGLMDQAFQY
Human CL     140   AFSATGALEGQMFRKTGRLISLSEQNLVDCSGPQGNEGCNGGLMDYAFQY

Trout CL      51   IQDNAGLDTEESYPYVGTDEDPCHYKPEFSAANETGFVDIPSGKEHAMMK
Chicken CL    77   VQDNGGIDSEESYPYTAKDDEDCRYKAEYNAANDTGFVDIPQGHERALMK
Human CL     190   VQDNGGLDSEESYPYEAT-EESCKYNPKYSVANDTGFVDIPK-QEKALMK

Trout CL     101   AVAAVGPVSVAIDAGHESFQFYESGIYYEKECSSEELDHGVLVVGYGFEG
Chicken CL   127   AVASVGPVSVAIDAGHSSFQFYQSGIYYEPDCSSEDLDHGVLVVGYGFEG
Human CL     238   AVATVGPISVAIDAGHESFLFYKEGIYFEPDCSSEDMDHGVLVVGYGFES

Trout CL     151   EDVDGKKYWIVKNSWSEKWGDKGYIYMAKDRKNHCGIATASSYPLV
Chicken CL   177   ----GKKYWIVKNSWGEKWGDKGYIYMAKDRKNHCGIATAASYPLV
Human CL     288   TESDNNKYWLVKNSWGEEWGMGGYVKMAKDRRNHCGIASAASYPTV
```

Table 3. Comparison of amino acid sequences of partial trout cathepsin L and chicken, human cathepsin L .

with, but significant differences from cDNA sequences that have been reported for cathepsin L (Bryce et al., 1994). A novel family of cathepsin L-like sequences have been recently identified and localized on human chromosome 10q, which is distinct from the cathepsin L locus on chromosome 9 (Bryce et al., 1994). The transcripts corresponding to two cathepsin L-like genes were detected in human bladder and/or liver tissues (Bryce et al., 1994). It could be assumed that the multiple isozymatic forms of cathepsin L might be encoded by several cathepsin L-related genes. It has been suggested that cathepsin L is a housekeeping gene, and it is expressed in all tissues with some variations in the levels (Petanceska and Devi, 1992). Cathepsin L-related genes are only expressed in some tissues and at certain physiological and pathological stages. A novel cathepsin L-related protein has been found to be expressed in rat placenta, but not kidney or liver (Conliffe et al., 1995). It reaches its highest expression levels during late gestation. Cathepsins K and S also are closely related to cathepsin L and share high similarity with cathepsin L. For instance, human and bovine cathepsin S share 58.5% and 59.6% identity with chicken cathepsin L, respectively, at the amino acid level (Tezuka et al., 1994). Rabbit cathepsin K is 60.1% identical to chicken cathepsin L amino acid sequence (Tezuka et al., 1994). The expression of cathepsins K and S has been found to be tissue-specific by Northern blot analysis (Petanceska and Devi, 1992; Inaoka et al., 1995).

SFL2 is the partial sequence of a cysteine proteinase gene

The sequence analysis of the clones SFL2, 7, and 10 revealed that the insertion is a cysteine proteinase cDNA sequence, which has 51% identity and 65% similarity with

Cysteine Proteinase 2 from *Dictyostelium discoideum*. It also has 44% identity and 60% similarity with human cathepsin L (Table 4), and 42% identity and 58% similarity with chicken cathepsin L. The alignment of the deduced amino acid sequence of these clones with that of human cathepsin L showed it contained the active sites, H^{159}, G^{43}, C^{56}, and G^{166} (papain numbering) reported in all cysteine proteinases (Mitchel et al., 1970; Petanceska and Devi, 1992). E^{35}, which is highly conserved in papain, cathepsins L and H and plays an important role in an inter-domain electrostatic interaction (Dufour, 1988), was also found in SFL2. E^{50}, which plays a similar role to the E^{35} in plant cysteine proteinases and cathepsins L and H, was replaced by Q. Because cysteine proteinases have been shown to be highly conserved among different vertebrate species (Tezuka et al., 1994), the moderate similarity of SFL2 with human and chicken cathepsin L suggests that SFL2 might be the partial cDNA of a novel cysteine proteinase which is related to cathepsin L.

SFL17 is the partial sequence of cathepsin K

The 456-bp-long sequence of SFL17, which encodes 152 amino acids and presents about 70% of mature enzyme, was identified. This sequence contains active site H (SFL17 136), one putative N-glycosylation site (N-X-S/T) (SFL17 73), and single recognition site for restriction enzymes, MluI, MscI, NspV, PstI and PvuII (Fig. 14). The protein homolgy search showed that the amino acid sequence of SFL17 was 71% identical and 88% similar to rabbit cathepsin K; and 70% identical and 90% similar to human cathepsin K. Alignment of amino acid sequences of SFL17 with rabbit and human cathepsin K (Table 5) showed that the N-terminal part (SFL17 1-50) and the C-terminal

```
                      35      43     50    56
SFL2         1 SFATTGAVEGALFLKSGSLQVLSQQMLVDCSWGFGNNGCDGGEEWRAYEW
Human CL   140 AFSATGALEGQMFRKTGRLISLSEQNLVDCSGPQGNEGCNGGLMDYAFQY

SFL2        51 IMKHGGIATTETYGSYMGMNGLCHLNTSQLTARVQSYTNVTSGDAEALKV
Human CL   190 VQDNGGLDSEESYPYEATEESCKYNPKYSVANDTGFVDIPKQE--KALMK

                                                      159   166
SFL2       101 ALFKHGPVAVSIDAGHRSFVFYSHGIYYEPKCGNTTDSLDHAVLAVGYGV
Human CL   240 AVATVGPISVAIDAGHESFLFYKEGIYFEPDC--SSEDMDHGVLVVGYGF

SFL2       151 M
Human CL   286 E
```

Table 4. Alignment of amino acid sequences of SFL2 and human cathepsin L. The bold letters represent the identical amino acids.

```
SFL17          1  AFSSAGALEGQLARTTGKLIDLSPQNLVDCVTENNGCGGGYMTNAFEYVE
Rabbit CK    141  AFSSVGALEGQLKKKTGKLLNLSPQNLVDCVSENYGCGGGYMTNAFQYVQ
Human CK     141  AFSSVGALEGQLKKKTGKLLNLSPQNLVDCVSENDGCGGGYMTNAFQYVQ

SFL17         51  ENGGIDTEEAYPYLGQDEQCVYNASGMGAQCRGFKEIPEGDEWALTKAVV
Rabbit CK    191  RNRGIDSEDAYPYVGQDESCMYNPTGKAAKCRGYREIPEGNEKALKRAVA
Human CK     191  KNRGIDSEDAYPYVGQEESCMYNPTGKAAKCRGYREIPEGNEKALKRAVA

SFL17        101  KVGPVAVGIDATLSTFQFYQRGVYYDPNCNKDDINHAVLAVGYG
Rabbit CK    241  RVGPVSVAIDASLTSFQFYSKGVYYDENCSSDNVNHAVLAVGYG
Human CK     241  RVGPVSVAIDASLTSFQFYSKGVYYDESCNSDNLNHAVLAVGYG
```

Table 5. Alignment of amino acid sequence SFL17, rabbit cathepsin K, and human cathepsin K. The bold letters represent the identical amino acids in these three sequences.

part (SFL17 102-144) have much higher similarity with rabbit and human cathepsin K than the central region (SFL17 51-101), as was observed in other proteinases belonging to the papain superfamily (Dufour, 1988). Because the remaining unidentified sequence of SFL17 represents the highly conserved N-terminal and C-terminal parts of the mature enzyme, the above figures may underestimate the homology of SFL17 with rabbit and human cathepsin K. Based on the high similarity between SFL17 and rabbit and human cathepsin K, SFL17 was postulated to be the partial sequence of trout cathepsin K.

Cathepsin K cDNA was first found in rabbit and human osteoclasts (Tezuka et al., 1994; Inaoka et al., 1995). The deduced amino acid sequences of cathepsin K have a high similarity with cathepsins L and S, and encode a novel member of the cysteine proteinase family. Cathepsin K has been found to be expressed at high levels in osteoclasts and low levels in other tissues. It is speculated that cathepsin K is closely involved in human and rabbit bone resorption, even though the proteins encoded by cathepsin K cDNAs have not yet been identified or characterized.

It has been suggested that cathepsin K might be a cysteine proteinase closely related to cathepsin L with the high homology to cathepsin L at amino acid level (Tezuka et al., 1994). SFL17 shares 60% identity, 74% similarity with chicken cathepsin L and 54% identity, 71% similarity with bovine cathepsin S (Table 6). This homology is similar to rabbit and human cathepsin K. Rabbit cathepsin K showed 60.1% identity with mature chicken cathepsin L, 59.2% identity with mature bovine cathepsin S (Tezuka et al., 1994). Human cathepsin K was 48% identical to human cathepsin S and 42% identical to human cathepsin L over the entire deduced amino acid sequence (Inaoka et al., 1995). The partial deduced amino acid sequence of trout cathepsin K shares 51%

```
SFL17          1  AFSSAGALEGQLARTTGKLIDLSPQNLVDCVT-E--NNGCGGGYMTNAFE
Chicken CL    27  AFSTTGALEGQHFRKTGKLVSLSEQNLVDCSRPE-GNQGCNGGLMDQAFQ
Bovine CS     27  AFSAVGALEAQVKLKTGKLVSLSAQNLVDCSTAKYGNKGCNGGFMTEAFQ

SFL17         48  YVEENGGIDTEEAYPYLGQD-EQCVYNASGMGAQCRGFKEIPEGDEWALT
Chicken CL    76  YVQDNGGIDSEESYPYTAKDDEDCRYKAEYNAANDTGFVDIPQGHERALM
Bovine CS     75  YIIDNNGIDSEASYPYKAMD-GKCQYDVKNRAATCSRYIELPFGSEEALK

SFL17         97  KAVVKVGPVAVGIDATLSTFQFYQRGVYYDPNCNKDDINHAVLAVGYG
Chicken CL   126  KAVASVGPVSVAIDAGHSSFQFYQSGIYYEPDCSSEDLDHGVLVVGYG
Bovine CS    126  EAVANKGPVSVGIDASHSSFFLYKTGVYYDPSCTQN-VNHGVLVVGYG
```

Table 6. Comparison of amino acid sequences of SFL17, chicken cathepsin L and bovine cathepsin S. The bold letters represent the identical amino acids.

identity with the partial deduced amino acid sequence of TCL. The cloning of both cathepsins K and L in this study provides the strong evidence for Tezuka's proposal (1994) that cathepsin K is a novel cysteine proteinase which is closely related to but significantly different from cathepsin L. Further study will be necessary to understand the function of cathepsin K as well as the expression regulation of this enzyme in fish. Our study is the first report on the partial sequences of cysteine cathepsins in fish.

Chapter 5. SUMMARY

The N-terminal amino acid sequence of cysteine proteinase P-II, purified from Pacific whiting, was found to be 61% identical with the sequences of chicken cathepsin L and 65% identical with cathepsins L and S from mammalian species, but only 39% identical with cathepsin B from rat and human. Amino acid composition analysis showed that Pacific whiting P-II contains much less lysine, and a considerably greater proportion of hydrophobic amino acids than cathepsin L from other species. No glycosylation was found in Pacific whiting P-II by either carbohydrate staining or monosaccharide composition analysis. Three isozymatic forms of P-II were shown on non-denaturing PAGE gel, and the rapid autolytic degradation of the enzyme was observed on SDS-PAGE in the absence of substrate.

Rainbow trout liver cDNA was amplified by PCR with degenerate primers based on two highly conserved amino acid sequences in the papain superfamily. The 500-bp PCR product contained at least 3 different sequences. Clone SFL5 was the partial cDNA sequence of TCL and was labeled and used as a probe to screen the trout liver cDNA library. A 895-bp partial TCL sequence had been identified. One stop codon and a AATAAA sequence was found in the 307-bp 3' untranslated region. The deduced amino acid sequence of the partial TCL was 83% identical with chicken cathepsin L. Its N-terminal and C-terminal regions had about 92% identity with chicken cathepsin L. Based on the moderate identity of Pacific whiting P-II with several cathepsin L, the enzyme purified from parasitized Pacific whiting muscle is proposed to be encoded by a cathepsin L-related gene, which is expressed under the pathological condition. Clone SFL2 may be the partial cDNA of a novel cathepsin L-related cysteine proteinase. It was 51%

identical with Cysteine Proteinase 2 from *D. discoideum* and 44% identical with human cathepsin L at the deduced amino acid level. Clone SFL17 may be the partial sequence of cathepsin K, which is closely related to cathepsins L and S. The deduced amino acid sequence of SFL17 had 70% identity and about 89% similarity with rabbit and human cathepsin K over about 70% of mature enzyme.

BIBLIOGRAPHY

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J.D. (Eds.) 1994. Vesicular, and traffic in the secretory and endocytic pathways. Chap. 13 in *Molecular Biology of the Cell.* 3rd Ed., pp 599-654. Garland Publishing, Inc., New York & London.

An, H., Seymour, T.A., Wu, J.-W., Morrissey, M.T. 1994a. Assay system and characterization of Pacific whiting (*Merluccius productus*) protease. J. Food Sci. 59: 277-281.

An, H., Weerasinghe, V., Seymour, T.A., and Morrissey, M.T. 1994b. Degradation of Pacific whiting surimi proteins by cathepsins. *J. Food Sci.* 59: 1013-1017, 1033

An, H., Peters, M.Y. Seymour, T.A., and Morrissey, M.T. 1995. Isolation and activation of cathepsin L-inhibitor complex from Pacific whiting (*Merluccius productus*). J. Agric. Food Chem. 43: 327-330.

An, H., Wu, J.-W., Fan, X., Morrissey, M.T., and Seymour, T.A. 1996. Molecular and kinetic properties of cathepsin L from Pacific whiting (*Merluccius productus*). Submitted.

Andrews P.C. and Dixson, J.E. 1987. A procedure for *in situ* alkylation of cystine residues on glass fiber prior to protein microsequence analysis. Anal. Biochem. 161: 524-528.

Barrett, A.J. and Kirschke, H. 1981. Cathepsin B, cathepsin H and cathepsin L. Meth. Enzymol. 80: 535-561.

Benjannet, S., Rondeau, N., Day, R., Chretien, M., and Seidah, N.G. 1991. PC1 and PC2 are proprotein convertases capable of cleaving proopiomelanocortin at distinct pairs of basic residues. Proc. Natl. Acad. Sci. USA 88: 3564-3568.

Bond, J. S. and Butler, P. E. 1987. Intracellular proteases. Ann. Rev. Biochem. 56: 333-364.

Borejdo, J. 1983. Mapping of hydrophobic sites on myosin and its fragments. Biochem. 22: 1182-1187.

Boye, S.W. and Lanier, T.C. 1988. Effects of heat-stable alkaline protease activity of Atlantic menhaden (*Brevoorti tyrannus*) on surimi gels. J. Food Sci. 53: 1340-1342.

Bromme, D., Steinnert, A., Friebe, S., Fittkau, S., Wiederanders, B., and Kirschke, H. 1989. The specificity of bovine spleen cathepsin S, a comparison with rat liver cathepsin L and B. Biochem. J. 264: 475-481.

Bryce, S.D., Lindsay, S., Gladstone, A.J., Braithwaite, K., Chapman, C., Spurr, N.K., and Lunec, J. 1994. A novel family of cathepsin L-like (CTSLL) sequences on human chromosome 10q and related transcripts. Genomics 24: 568-576.

Chan, S.J., Segundo, B.S., McCormick, M.B., and Steiner, D.F. 1986. Nucleotide and predicted amino acid sequences of cloned human and mouse preprocathepsin B cDNAs. Proc. Natl. Acad. Sci. USA 83: 7721-7725.

Chang-Lee, M.V., Aguilar, R.P., Crawford, D.L., and Lampila, L.E. 1989. Proteolytic activity of surimi from Pacific whiting (*Merluccius productus*) and heat-set gel texture. J. Food Sci. 54: 1116-1119.

Chauhan, S.S., Popescu, N.C., Ray, D., Fleischmann, R., Gottesman, M.M., and Troen, B.R. 1993. Cloning, genomic organization, and chromosomal localization of human cathepsin L. J. Biol. Chem. 268: 1039-1045.

Chenchik, A., Moqadam, F., and Siebert, P.D. 1996. Marathon™ cDNA amplification: A new method for cloning full-length cDNAs. CLONTECH Laboratories, Inc., Palo Alto, CA.

Chrambach, A. and Rodbard, D. 1971. Polyacrylamide gel electrophoresis. Science 172: 440-451.

Conliffe, P.R., Ogilvie, S., Simmen, R.C.M., Michel, F.J., Saunders, P., and Shiverick, K.T. 1995. Cloning and expression of a rat placental cDNA encoding a novel cathepsin L-related protein. Mol. Reprod. Dev. 30: 285-292.

Devi, L. 1991. Peptide processing at monobasic sites. Chap. 7 in *Peptide Biosynthesis and Processing* Fricker, (L.D. ed). pp 175-198. CRC Press, Boca Raton, FL.

Doherty, P. J., Hua, L., Liau, G., Gal, S., Graham, D. E., Sobel, M., and Gottesman, M. M. 1985. Malignant transformation and tumor promoter treatment increase level of a transcript for a secreted glycoprotein. Mol. Cell. Biol. 5: 466-473.

Dong, S.S., Stransky, G.I., Whitaker, C.H., Jordan, S.E., Schlesinger, P.H., Edwards, J.C., and Blair, H.C. 1995. 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 1251: 69-73.

Dufour, E. 1988. Sequence homologies, hydrophobic profiles and secondary structures of cathepsins B, H, and L: comparison with papain and actinidin. Biochimi. 70: 1335-1342.

Eckhardt, A.E., Hayes, C.E., and Goldstein, I.J. 1976. A sensitive fluorescent method for the detection of glycoproteins in polyacrylamide gels. Anal. Biochem. 73: 192-197.

Erickson, A.H. and Blobel, G. 1983. Carboxyl-terminal proteolytic processing during biosynthesis of the lysosomal enzymes β-glucuronidase and cathepsin D. Biochem. 22: 5201-5205.

Erickson, M.C., Gordon, D.T., and Anglemier, A.F. 1983. Proteolytic activity in the sarcoplasmic fluids of parasized pacific whiting (*Merluccius productus*) and unparasitized True cod (*Gadus macrocephalus*). J. Food Sci. 48: 1315-1319.

Folco, E.J., Busconi, L., Martone, C., Trucco, R.E., and Sanchez, J.J. 1984. Action of two alkaline proteases and a trypsin inhibitor from white croaker skeletal muscle (*Micropogon poercularis*) in the degradation of myofibrillar proteins. FEBS Lett. 176: 215-219.

Fong, D., Calhoun, D.H., Hsieh, W.T., Lee, B., and Wells, R.D. 1986. Isolation of a cDNA clone for the human lysosomal proteinase cathepsin B. Proc. Natl. Acad. Sci. USA 83: 2909-2913.

Gal, S. and Gottesman, M.M. 1988. Isolation and sequence of a cDNA for human pro-(cathepsin L). Biochem. J. 253: 303-306.

García-Carreño, F.L., Dimes, L.E., and Harrd, N.F. 1993. Substrate-gel electrophoresis for composition and molecular weight of proteinases or proteinaceous proteinase inhibitors. Anal. Biochem. 214: 65-69.

Golde, T.E., Estus, S., Younkin, L.H., Selkoe, D.J., and Younkin, S.G. 1992. Processing of the amyloid protein precursor to potentially amyloidogenic derivatives. Science 255: 728-730.

Gottesman, M.M. 1978. Transformation-dependent secretion of a low molecular weight protein by murine fibroblasts. Proc. Natl. Acad. Sci. USA 75: 2767-2771.

Hara, K., Suzumatsu, A. and Ishihara, T. 1988. Purification and charaterization of cathepsin B from ordinary carp muscle. Nippon Suisan Gakkaishi 54: 1243-1252.

Heussler, V.T. and Dobbelaere, D.A.E. 1994. Cloning of a protease gene family of *Fasciola hepatica* by the polymerase chain reaction. Mol. Biochem. Parasitol. 64: 11-23.

Homma, K., Kurata, S., and Natori, S. 1994. Purification, characterization, and cDNA cloning of procathepsin L from the culture medium of NIH-Sape-4, and

embryonic cell line of *Sarcophaga peregrina* (Flesh Fly), and its involvement in the differentiation of imaginal discs. J. Biol. Chem. 269: 15258-15264.

Inaoka, T., Bilbe, G., Ishibashi, O., Tezuka, K., Kumegawa, M., and Kokubo, T. 1995. Molecular cloning of human cDNA for cathepsin K: novel cysteine preteinase predominantly expressed in bone. Biochem. Biophys. Res. Comm. 206: 89-96.

Innis, M.A.and Gelfand D.H. 1990. Optimization of PCRs. Chap. 1 in *PCR Protocols: A Guide to Methods and Applications.* (Innis M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. Eds) pp 3-12. Academic Press, Inc., San Diego, CA.

Ishidoh, K., Towatari, T., Imajoh, S., Kawasaki, H., Kominami, E., Katunuma, N., and Suzuki, K. 1987. Molecular cloning and sequencing of cDNA for rat cathepsin L. FEB 223: 69-73.

Ishidoh, K., and Kominami, E. 1994. Multi-step processing of procathepsin L *in vitro.* FEBS Lett. 352: 281-284.

Jiang, S.T., Wang, J.H., and Chen, C.S. 1991. Purification and some properties of calpain II from tilapia muscle (*Tilapia nilotica* X *Tilapia aurea*). J. Agric. Food Chem. 39: 237-241.

Joseph, L.J., Chang, L.C., Stamenkovich, D., and Sukhatme, V.P. 1988. Complete nucleotide and deduced amino acid sequences of human and murine preprocathepsin L. J. Clin. Invest. 81: 1621-1629.

Kamphuis, I.G., Kalk, K.H., Swarte, M.B.A., and Drenth, J. 1984. Structure of papain refined at 1.65Å resolution. J. Mol. Biol. 179: 233-256.

Katunuma, N. and Kominami, E. 1983. Structures and functions of lysosomal thiol proteinases and their endogenous inhibitor. Curr. Top. Cell. Regul. 22: 71-101.

Kawasaki, E.S. 1990. Amplification of RNA. Chap. 3 in *PCR Protocols: A Guide to Methods and Applications.* (Innis M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. Eds) pp 21-27. Academic Press, Inc., San Diego, CA.

Kirschke, H. 1977. Cathepsin H: an endoaminopeptidase. Acta. Biol. Med. Germ. 16: 1547-1548.

Kirschke, H., Langner, J., Wiederanders, B., Ansorge, S., and Bohley, P. 1977. Cathepsin L: a new proteinase from rat-liver lysosomes. Eur. J. Biochem. 74: 293-301.

Kirschke, H., Langner, J., Riemann, S., Wiederanders, B., Ansorge, S., and Bohley, P. 1980. Lysosomal cysteine proteinases. *Protein degradation in health and disease.* Ciba Foundation Symposium 75: 15-35.

Kirschke, H., Kembhari, A.A., Bohley, P., and Barrett, A.J. 1982. Action of rat liver cathepsin L on collagen and other substrates. Biochem. J. 201: 367-372.

Kirschke, H., Schmidt, I., and Wiederanders, B. 1986. Cathepsin S the cysteine proteinase from bovine lymphoid tissue is distinct from cathepsin L (EC 3.4.22.15). Biochem. J. 240: 455-459.

Kirschke, H. and Barrett, A.J. 1987. Chemistry of lysosomal proteases. Chap. 6 in *Lysosomes: their role in protein breakdown.* pp 193-238. Academic Press, London.

Kirschke, H., Wieseranders, B., Bromme, D. and Rinne, A. 1989. Cathepsin S from bovine spleen: purification, distribution, intracellular location and action on proteins. Biochem. J. 264: 467-473.

Kominami, E., Tsukahara, T., Hara, K., and Katunuma, N. 1988. Biosyntheses and processing of lysosomal cysteine proteinases in rat macrophages. FEBS Lett. 231: 225-228.

Konagaya, S. and Aoki, T. 1981. Jellied condition in Pacific hake in relation to Myxosporidian infection and protease activity. Bull. Tokai Reg. Fish. Res. Lab. 105: 1-16.

Koohmaraie, M., Schollmeyer, J.E., and Dutson, T.R. 1986. Effect of low-calcium-requiring calcium activated factor on myofibrils under varying pH and temperature conditions. J. Food Sci. 51: 28-32, 65.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Lafuse, W.P., Brown, D., Castle, L., and Zwilling, B.S. 1995. IFN-γ increases cathepsin H mRNA levels in mouse macrophages. J. Leukoc. Biol. 57: 663-669.

Laycock, M.V., Mackay, R.M., Fruscio, M.D., and Gallant, J.W. 1991. Molecular cloning of three cDNAs that encode cysteine proteinases in the digestive gland of the American lobster (*Homarus americanus*). FEBS 292: 115-120.

Lee, C. and Caskey, C.T. 1990. cDNA cloning using degenerate primers. Chap. 6 in *PCR Protocols: A Guide to Methods and Applications.* (Innis M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. Eds) pp46-53. Academic Press, Inc., San Diego, CA.

Lee, J.J., Chen, H.C., Jiang, S.T. 1993. Purification and characterization of proteinases identified as cathepsin L and L-like (58 kDa) proteinase from Mackerel (*Scomber australasicus*). Biosci. Biotech. Biochem. 57: 1470-1476.

Loh, E.Y., Elliott, J.F., Cwirla, S., Lanier, L.L., and Davis, M.M. 1989. Polymerase chain reaction with single-sided specificity: Analysis of T cell receptor δ chain. Science 243: 217-220.

Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. 1951. Protein measurement with Folin phenol reagent. J. Biol. Chem. 193: 256-275.

Maciewicz, R.A. and Etherington, D.J. 1988. A comparison of four cathepsins (B, L, N, and S) with collagenolytic activity from rabbit spleen. Biochem. J. 256: 433-440.

Makinodan, Y., Toyohara, H., and Ikeda, S. 1984. Comparison of muscle proteinase activity among fish species. Comp. Biochem. Physiol. 79B: 129-134.

Makinodan, Y., Yokoyama, Y., Konoshita, M., and Toyohara, H. 1987. Characterization of an alkaline proteinase of fish muscle. Comp. Biochem. Physiol. 87B: 1041-1046.

Mallinson, D.J., Lockwood, B.C., Coombs, G.H., and North, M.J. 1994. Identification and molecular cloning of four cysteine proteinase genes from the pathogenic protozoon *Trichomonas vaginalis*. Micorbiology 140: 2725-2735.

Masaki, T., Shimamukai, M., Miyauchi, Y., Ono, S., Tuchiya, T., Mastuda, T., Akazawa, H., and Soejima, M. 1993. Isolation and characterization of the protease responsible for jellification of Pacific hake muscle. Nippon Suisan Gakkaishi 59: 683-690.

Mason, R.W., Taylor, M.A.J., and Etherington, D.J. 1984. The purification and properties of cathepsin L from rabbit liver. Biochem. J. 217: 209-217.

Mason, R.W., Green, G.D.J., and Barrett, A.J. 1985. Human liver cathepsin L. Biochem. J. 226: 223-241.

Mason, R.W., Walker, J.E., and Northrop, F.D. 1986. The N-terminal amino acid sequences of the heavy and light chains of human cathepsin L. Biochem. J. 240: 373-377.

Mason, R.W. 1986. Species variants of cathepsin L and their immunological identification. Biochem. J. 240: 285-28.

Mason, R.W., Gal, S., and Gottesman, M.M. 1987. The identification of the major excreted protein (MEP) from a transformed mouse fibroblast cell line as a catalytically active precursor form of cathepsin L. Biochem. J. 248: 449-454.

Mason, R.W., Wilcox, D., Wilkstrom, P., and Shaw, E.N. 1989. The identification of active forms of cysteine proteinases in Kirsten-virus-transformed mouse fibroblasts by use of a specific radiolabelled inhibitor. Biochem. J. 257: 125-129.

Matsumiya, M., Mochizuki, A., and Otake, S. 1989. Purification and characterization of cathepsin B from ordinary muscle of common mackerel (*Scomber japonicus*). Nippon Suisan Gakkaishi 55: 2185-2190.

McIntyre, G.F. and Erickson, A.H. 1991. Procathepsin L and D are membrane-bound in acidic microsomal vesicles. J. Biol. Chem. 266: 15438-15445.

Mitchel, R.E.J., Chaiken, I.M., and Smith, E.L. 1970. The complete amino acid sequence of papain. J. Biol. Chem. 245: 3485-3492.

Morrissey, M.T., Wu, J.W., Lin, D., and An, H. 1993. Protease inhibitor effects on torsion measurements and autolysis of Pacific whiting surimi. J. Food Sci. 58: 1050-1054.

Nishimura, Y., Amano, J., Sato, H., Tsuji, H., and Kato, K. 1988. Biosynthesis of lysosomal cathepsin B and H in cultured rat hepatocytes. Arch. Biochem. Biophys. 262: 159-170.

Nishimura, Y. and Kato, K. 1992. Expression of mouse cathepsin L cDNA in *Saccharomyces cerevisiae*: Evidence that cathepsin L is sorted for targeting to yeast vacuole. Arch. Biochem. Biophys. 298: 318-324.

North, M.J., Mottram, J.C., and Coombs, G.H. 1990. Cysteine proteinases of parasitic protozoa. Parasitology Today. 6: 270-275.

Okitani, A., Matsukura, U., Kato, H., and Fujimaki, M. 1980. Purification and some properties of a myofibrillar protein-degrading protease, cathepsin L, from rabbit skeletal muscle. J. Biochem. 87: 1133-1143.

Okitani, A., Matsuishi, M., Matsumoto, T., Kamoshida, E., Sato, M., Matsukura, U., Watanabe., M., Kato, H., and Fujimaki, M. 1988. Purification and some properties of cathepsin B from rabbit skeletal muscle. Eur. J. Biochem. 171: 377-381.

Old, R.W. and Primrose, S.B. (Eds.) 1989. Nucleic acid probes and their applications. Chapter 15 in *Principles of gene manipulation; An introduction to genetic*

engineering. 4th Ed. pp 319-343. Oxford Blackwell Scientific Publications, London.

Patashnik, M., Groninger, H.S., Jr., Barnett, H., Kudo, G., and Koury, B. 1982. Pacific whiting, *Merluccius productus*: I. Abnormal muscle texture caused by Myxosporidian-induced proteolysis. Mar. Fish. Rev. 44(5): 1-12.

Petanceska, S. and Devi, L. 1992. Sequence analysis, tissue distribution, and expression of rat cathepsin S. J. Biol. Chem. 267: 26038-26043.

Portnoy, D.A., Erickson, A.H., Kochan, J., Ravet Alh, J.V., and Unkeless, J.C. 1986. Cloning and characterization of a mouse cysteine proteinase. J. Biol. Chem. 261: 14697-14703.

Racusen, D. 1979. Glycoprotein detection in polyacrylamide gel with thymol and sulfuric acid. Anal. Biochem. 99: 474-476.

Radtke, H. 1995. Windows on Pacific whiting: An economic success story for Oregon state fishing industry. Report for Oregon Coast Zone Management Association. Newport, OR.

Ritonja, A., Colic, A., Dolenc, I., Ogrinc, T., Podobnik, M., and Turk, V. 1991. The complete amino acid sequence of bovine cathepsin S and a partial sequence of bovine cathepsin L. FEBS. 283: 329-331.

Rosenthal, P.J. and Nelson, R.G. 1992. Isolation and charaterization of a cysteine proteinase gene of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 42: 55-62.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350-1354.

Sakanari, J.A., Staunton, C.E., Eakin, A.E., Craik, C.S., and McKerrow, J.H. 1989. Serine proteases from nematode and protozoan parasites: Isolation of sequence homologs using generic molecular probes. Proc. Natl. Acad. Sci. USA 86: 4863-4867.

Saneshige, S., Mano, H., Tezuka, K., Kakudo, S., Mori, Y., Honda, Y., Itabash, A., Yamada, T., Miyata, K., Hakeda, Y., Ishii, J., and Kumegawa, M. 1995. Retinoic acid directly stimulates osteoclastic bone resorption and gene expression of cathepsin K/OC-2. Biochem. J. 309: 721-724.

Segundo, B.S., Chan, S.J., and Steiner, D.F. 1985. Identification of cDNA clones encoding a precusors of rat liver cathepsin B. Proc. Natl. Acad. Sci. USA 82: 2320-2324.

Seymour, T.A., Morrissey, M.T., Peters, M.Y., An, H. 1994. Purification and characterization of Pacific whiting proteases. J. Agric. Food Chem. 42: 2421-2427.

Sherekar, S. V., Gore, M. S., and Ninjoor, V. 1988. Purification and characterization of cathepsin B from the skeletal muscle of fresh water fish, *Tilapia mossambica.* J. Food Sci. 53:1018-1023.

Shi, G.P., Munger, J.S., Meara, J.P., Rich, D.H., and Chapman, H.A. 1992. Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease. J. Biol. Chem. 267: 7258-7262.

Smith, S.M. and Gottesman, M.M. 1989. Activity and deletion analysis of recombinant human cathepsin L expressed in *Escherichia coli.* J. Biol. Chem. 264: 20487-20495.

Stoknes, E., Rustad, T., and Mohr, V. 1993. Comparative studies of the proteolytic activity of tissue extracts from cod (*Gadus morhua*) and herring (*Clupea harengus*). Comp. Biochem. Physiol. 106B: 613-619.

Takio, K., Towatari, T., Katunuma, N., and Titani, K. 1980. Primary structure study of rat liver cathepsin B: a striking resemblance to papain. Biochem. Biophys. Res. Comm. 97: 340-346.

Takio, K., Towatari, T., Katunuma, N., Teller, D.C., and Titani, K. 1983. Homology of amino acid sequences of rat liver cathepsins B and H with that of papain. Proc. Natl. Acad. Sci. USA 80: 3666-3670.

Tao, K., Stearns, N.A., Dong, J., Wu, Q. and Sahagian, G.G. 1994. The proregion of cathepsin L is required for proper folding, stability, and ER exit. Arch. Biochem. Biophys. 311: 19-27.

Tezuka, K., Tezuka, Y., Maejima, A., Sato, T., Nemoto, K., Kamioka, H., Hakeda, Y., and Kumegawa, M. 1994. Molecular cloning of a possible cysteine proteinase predominantly expressed in osteoclasts. J. Biol. Chem. 269: 1106-1109.

Thomas, L., Leduc, R., Thorne, B.A., Smeekens, S.P., Steiner, D.F., and Thomas, G. 1991. Kex2-like endoproteases PC2 and PC3 accurately cleave a model prohormone in mammalian cells: Evidence for a common core of neuroendocrine processing enzymes. Proc. Natl. Acad. Sci. USA 88: 5297-5301.

Toyohara, H., Kinoshita, M., Kimura, I., Satake, M., and Sakaguchi, M. 1993. Cathepsin L-like protease in Pacific hake muscle inected by Myxosporisian parasites. Nippon Suisan Gakkaishi 59: 1101-1102.

Trabandt, A., Aicher, W.K., Gay, R.E., Sukhatme, V.P., Nilson-Hamilton, M., Hamilton, R.T., McGhee, J.R., Fassbender, H.G., and Gay, S. 1990. Expression of the collagenolytic and Ras-induced cysteine proteinase cathepsin L and proliferation-associated oncogenes in synovial cells of MRL/l mice and patients with rheumatoid arthritis. Matrix 10: 349-361.

Troen, B.R., Chauhan, S.S., Ray, D., and Gottesman, M.M. 1991. Downstream sequences mediate induction of the mouse cathepsin L promoter by phorbol esters. Cell Growth Differ. 2: 23-31.

Tryggvason, K., Hoyhtya, M., and Salo, T. 1987. Proteolytic degradation of extracellular matrix in tumor invasion. Biochim. Biophys. Acta 907: 191-217.

Tsuchiya, H. and Seki, N. 1991. Action of calpain on α-actinin within and isolated from carp myofibrils. Nippon Suisan Gakkaishi 57: 1133-1139.

Tsujinaka, T., Ebisui, C., Fujita, J., Kishibuchi, M., Morimoto, T., Ogawa, A., Katsume, A., Ohsugi, Y., Kominami, E., and Monden, M. 1995. Muscle undergoes atrophy in assocation with increase of lysosomal cathepsin activity in interleukin-6 transgenic mouse. Biochem. Biophys. Res. Comm. 207: 168-174.

Turnsek, T., Kregar, I., and Lebez, D. 1975. Acid sulphydryl protease from calf lymph nodes. Biochem. Biophys. Acta 403: 514-520.

Wada, K., Takai, T., and Tanabe, T. 1987. Amino acid sequence of chicken liver cathepsin L. Eur. J. Biochem. 167:13-18.

Wasson, D.H., Babbitt, J.K., and French, J.S. 1992. Characterization of a heat stable protease from arrowtooth flounder, *Atheresthes stomias*. J. Aquat. Food Prod. Tech. 1:167-182.

Wiedernanders, B., Broemme, D., Kirschke, H., Kalkkinen, N., Rinne, A., Paquette, T., and Toothman, P. 1991. Primary structure of bovine cathepsin S comparison to cathepsin L, H, B and papain. FEBS. 286: 189-192.

Yamashita, M. and Konagaya, S. 1990a. Participation of cathepsin L into extensive softening of the muscle of chum salmon caught during spawing migration. Nippon Suisan Gakkaishi 56: 1271-1277.

Yamashita, M. and Konagaya, S. 1990b. Purification and characterization of cathepsin L from the white muscle of chum salmon (*Oncorhynchus keta*). Comp. Biochem. Physiol. 96B: 247-252.

Yamashita, M. and Konagaya, S. 1990c. Purification and characterization of cathepsin B from the white muscle of chum salmon (*Oncorhynchus keta*). Comp. Biochem. Physiol. 96B: 733-737.

Yamashita, M. and Konagaya, S. 1991a. Hydrolytic action of salmon cathepsins B and L to muscle structural proteins in respect of muscle softening. Nippon Suisan Gakkaishi 57: 1917-1922.

Yamashita, M. and Konagaya, S. 1991b. Immunohistochemical localization of cathepsins B and L in the white muscle of chum salmon (*Oncorhynchus keta*) in spawning migration: probable participation of phagocytes rich in cathepsins in extensive muscle softening of the mature salmon. J. Agric. Food Chem. 39: 1402-1405.

Zabludoff, S.D., Karzai, A.W., and Wright, W.W. 1990. Germ cell-sertoli cell interactions: the effect of testicular maturation on the synthesis of cyclic protein-2 by rat sertoli cells. Biol. Reprod. 43: 25-33.