2000

Further characterization of the large subunit of the major embryonic Artemia franciscana cysteine protease.

Ervin. Pullumbi

University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation


https://scholar.uwindsor.ca/etd/2175

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000 ext. 3208.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
FURTHER CHARACTERIZATION OF THE LARGE SUBUNIT OF THE MAJOR EMBRYONIC ARTEMIA FRANCISCANA CYSTEINE PROTEASE

By:

Ervin Pullumbi

A Thesis
Submitted to the College of Graduate Studies and Research through the Department of Biological Sciences
In Partial Fulfilment of the Requirements for the Degree of Master of Science at the University of Windsor

Windsor, Ontario, Canada

1999
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-52640-2
ABSTRACT

The encysted embryos of *Artemia franciscana* contain a potentially unique cysteine protease consisting of a cathepsin L-like protease and a tightly associated protein which has been denoted as the large subunit. The molecular masses of these proteins are 28.5 kD for the small subunit and 31.5 kD for the large subunit. Whereas the small subunit has been characterized in this laboratory as a cathepsin L-like cysteine protease, the large subunit has not been studied.

The dimeric cysteine protease from *Artemia* encysted embryos was purified to apparent homogeneity using a multi-step method involving gel filtration, anion exchange and affinity chromatography. Five isoforms were separated by fast protein liquid chromatography. The protease subunits from each isoform were isolated using reverse phase high performance liquid chromatography.

Protease Lys C and cyanogen bromide treatments were used to generate fragments for N-terminal amino acid sequencing of the large subunit from the major isoform of *Artemia* cysteine protease. One of the peptides generated by the cyanogen bromide treatment had a molecular weight of about 25 kD and was submitted for N-terminal amino acid sequencing. The first 22 amino acids of the peptide were identified to have a high homology with several cell adhesion molecules, such as mouse osteoblast-specific factor 2 (mOSF-2), *Drosophila* fasciclin I (dFAS-I), chick beta transforming growth factor induced protein
(cβIG-H3), and the 32 kD hyaline layer component from the sea urchin embryo (HLC-32).

Chromatography of the *Artemia* cysteine protease on Mono S at pH 5.0 led to the purification of the subunits in the native state. During the purification on Mono S, the dissociation of the subunits appeared to occur by an irreversible process which is dependent on both low pH and cysteine protease activity. The fact that the catalytic and non-catalytic subunits do not re-associate after separation (at pH 5.0) supports previous findings which show that the majority of this protease is localized in the cytoplasm or the extracellular matrix of the encysted embryo and not in acid vesicles or lysosomes.

Based on molecular and biochemical data, several possible roles for the large subunit of *Artemia* cysteine protease are proposed in this thesis. The large subunit of the major embryonic *Artemia* cysteine protease seems to be a unique protein which stabilizes the cathepsin L-like protease subunit at neutral pH and high temperature, conditions which often accompany dormant embryos in nature. As well, the 31.5 kD subunit may play a role in protease docking to specific sites in the embryo, and facilitate the role of the cathepsin L-like catalytic subunit in yolk degradation and modifications of the extracellular matrix throughout the embryo during early development.
AKNOWLEDGEMENT

I wish to thank very much Dr. A. H. Warner for providing me the opportunity to work in his laboratory as a Master’s student. I have appreciated very much all the training and his advice during this period.

I wish also to thank Dr. D. A. Cotter of the Department of Biological Sciences and Dr. B. Mutus of the Department of Chemistry and Biochemistry for their participation as graduate committee members, and for reviewing my thesis.

Many thanks are also due to the graduate students and the technical support staff of this department who helped me with various aspects of this thesis.
TABLE OF CONTENTS

ABSTRACT........................................................................................................iv
ACKNOWLEDGEMENTS..................................................................................vi
TABLE OF CONTENTS....................................................................................vii
LIST OF FIGURES...........................................................................................xi
LIST OF TABLES...............................................................................................xii

1. INTRODUCTION.........................................................................................1

1.1. Proteases.................................................................................................1

1.1.1. The serine proteases (E.C. 3.4.21.).........................................................5
1.1.2. The cysteine proteases (E.C. 3.4.22.)......................................................6
1.1.3. The aspartic proteases (E.C. 3.4.23.)......................................................9
1.1.4. Metalloproteases (E.C. 3.4.24.).............................................................10

1.2. Invertebrate proteases.............................................................................10

1.3. The proteases of *Artemia franciscana*..................................................12

1.4. Putative role of the major embryonic *Artemia* cysteine protease, a cathepsin L-like protease.................................................................15

1.5. Natural protease inhibitors as regulators of proteolytic activity............16

1.6. Protease inhibitors in *Artemia*...............................................................17

1.7. Objectives...............................................................................................18

2. MATERIALS AND METHODS.................................................................19

2.1. Materials...............................................................................................19

2.2. Methods................................................................................................20

2.2.1. Isolation of the major cysteine protease from *Artemia* cysts...........20
2.2.2. Cysteine protease assay.......................................................................24
     a) General cysteine protease assay..........................................................24
     b) Activity of the catalytic subunit alone or in the presence of the large subunit.........................................................25
2.2.3. Protein assays ................................................................. 25
2.2.4. The cyanogen bromide treatment of the Artemia large subunit .......... 26
2.2.5. Gel electrophoresis with amino group protection .......................... 26
2.2.6. Attempts to obtain amino acid sequence from the large subunit of Artemia CP ............................................................................. 27
2.2.7. Gel filtration of proteins from subcellular fractions of Artemia cysts ... 28
   a) Preparation of Artemia subcellular fractions ............................... 28
   b) Gel Filtration of the post-mitochondrial fraction (PMF) ............... 28
   c) Gel filtration of the myochoondrial/lysosomal lysate from Artemia embryos ................................................................. 29
   d) Gel filtration of CP after treatment to dissociate subunits .......... 30
2.2.8. Immunoaffinity purification of cysteine protease subunit-specific antibodies ................................................................. 30
2.2.9. Immunodetection of cysteine protease subunits by Western blotting ... 31
2.2.10. Elution of the purified Artemia cysteine protease from the Con A column .......................................................... 32
2.2.11. Fractionation of native cysteine protease subunits by FPLC on a Mono S column ................................................................. 33

3. RESULTS ........................................................................................................ 34

3.1. Isolation of the large subunit of isoform three (CP-3L) of the cathepsin L-like cysteine protease ......................................................... 35

3.1.1. Purification of the major Artemia cysteine protease ...................... 35
3.1.2. Separation of the cysteine protease isoforms ............................... 38
3.1.3. Separation of the large and small subunit of Artemia cysteine protease by HPLC ................................................................. 38

3.2. Amino acid analysis of peptides generated by CNBr and Lys C treatment of CP-3L protein ........................................................................ 41

3.2.1. Fragmentation of CP-3L and separation of the resulting peptides by SDS-PAGE and HPLC ............................................................. 41
3.2.2. Protein database search for homologues having the highest local similarity with the peptides sequenced ................................. 46

3.3. Isolation of Artemia embryo CP-3 subunits under native conditions .................................................................................... 48

3.3.1. Elution of Artemia CP from a concanavalin A column using glucose .... 50
3.3.2. Elution of Artemia CP from a concanavalin A Sepharose column with galactose ................................................................. 51
3.3.3. Gel filtration of the Artemia cysteine protease under acid conditions .... 55
3.3.4. Gel filtration of the Artemia cysteine protease under neutral pH following acid treatment at pH 5 ........................................... 58
3.3.5. Gel filtration of the Artemia cysteine protease after treatment at pH 5 in the presence of a reversible inhibitor .................................. 59

3.4. Purification of Artemia cysteine protease subunits on a Mono S column by fast protein liquid chromatography ........ 61

3.4.1. Separation of the CP subunits from a crude preparation under native conditions by FPLC ........................................ 61
3.4.2. Separation of the large and small subunits of the Artemia CP under native conditions by FPLC .................................. 65
3.4.3. Effect of the large subunit of the Artemia cysteine protease on activity of the catalytic (small) subunit ........................... 68

3.5. Characterization of the CP in fractions from Artemia embryos .................................................. 70

3.5.1. Gel filtration of the lysosomal fraction .................................................. 71
3.5.2. Gel filtration of the aqueous cytoplasmic fraction ............................... 72

4. DISCUSSION ................................................................. 75

4.1. Purification of the Artemia CP .................................................. 76

4.2. The partial sequence of the large subunit of the Artemia CP shows homology with some cell adhesion proteins ....... 79

4.3. Attempts to dissociate the Artemia CP subunits in native conditions ................................................. 80

4.3.1. Glucose gradient elution of the dimeric CP from a Con A column .... 81
4.3.2. Galactose elution of the dimeric CP followed by glucose elution from a Con A column ................................................. 81
4.3.3. Dissociation of the Artemia dimeric CP subunits at pH 5 ............. 82

4.4. The Artemia CP subunits do not re-associate after separation at low pH ..................................................... 82

4.5. The potential mechanism of dissociation ........................................ 83

4.6. Activity of the catalytic subunit of the Artemia CP in the presence or absence of the large subunit ................. 84
4.7. The putative roles of the large subunit of the embryonic
Artemia cysteine protease ..................................................85

4.7.1. Stabilization of the catalytic subunit..................................85
4.7.2. Protecting the catalytic subunit of the Artemia CP from high
temperatures........................................................................86
4.7.3. Targeting of the cysteine protease to the extracellular matrix....87

APPENDIX.....................................................................................90
1. MS profile of the Lys C-generated peptides from the Artemia CP-3L...90
2. HPLC profile of the Lys C-generated peptides from the Artemia CP-3L...91
3. Results of gel filtration column calibrations..................................92

REFERENCES...............................................................................93

VITA AUCTORIS...........................................................................102
LIST OF FIGURES

Fig. 1. Subunit composition of the cathepsin L-like cysteine protease of *Artemia franciscana* .......................................................... 37

Fig. 2. Purification of the cathepsin L-like cysteine protease CP-3 from *Artemia* embryos ............................................................... 39

Fig. 3. Separation of the *Artemia* cysteine protease subunits using HPLC ...................................................................................... 42

Fig. 4. Electrophoresis of the large subunit of *Artemia* cysteine protease third isoform (CP-3) ............................................................. 43

Fig. 5. Electrophoretic separation of CP-3L treated with CNBr ............ 47

Fig. 6. Elution of CP from a Con A Sepharose column using a glucose gradient .................................................................................. 52

Fig. 7. *Artemia* cysteine protease elution from a ConA Sepharose column with galactose followed by glucose ....................................... 53

Fig. 8. Gel filtration of *Artemia* embryo cysteine protease under acid and neutral pH ........................................................................... 56

Fig. 9. Gel filtration of the pH 5-treated *Artemia* cysteine protease at neutral pH conditions .............................................................. 60

Fig. 10. Gel filtration of *Artemia* cysteine protease after treatment at pH 5 in the presence of a reversible cysteine protease inhibitor .......... 62

Fig. 11. Separation of *Artemia* CP subunits by chromatography on a Mono S column ................................................................. 63

Fig. 12. Purification of the CP subunits under native conditions .......... 66

Fig. 13. Kinetics of the catalytic (small) subunit of the *Artemia* CP-3 in the presence and absence of the large CP-3 subunit at different pH ........................................................................................................... 69

Fig. 14. Gel filtration of a CP sample from lysosomal fraction ............. 71

Fig. 15. Gel filtration of a CP preparation from aqueous cytoplasm ...... 73
LIST OF TABLES

Table 1. Summary of the purification steps in the isolation of the 
Artemia cathepsin L-like cysteine protease...............................36

Table 2. Amino acid sequences of selected peptide fragments 
obtained from CNBr and endoproteinase Lys C 
treatment of CP-3L..........................................................45

Table 3. Alignment of the CNBr-generated fragment from Artemia 
CP-3L with various proteins showing high homology...................49
1. INTRODUCTION

1.1. Proteases

Proteases are enzymes that are able to catalyze the cleavage of peptide bonds between natural amino acids in proteins. Enzymes in general, and proteases in particular, have an essential role for the overall cellular functions, because they mediate virtually every biochemical reaction. The cell is able to regulate its many processes by a coordinated control of synthesis and degradation of each component, especially proteins. Protein degradation is a complex and highly regulated activity, which is carried out by several cellular proteolytic enzyme systems. In understanding the structure and functions of proteases, the evolutionary perspective is very useful. Proteases have been analyzed, and their differences explained by amino acid sequence divergence or convergence (Polgar, 1989). Besides the satisfaction of scientific curiosity, proteases are used to solve practical problems, such as those related to health, agriculture or industrial processes. But this goal can be met only with the support of fundamental research.

The classification of proteases has changed over the last several years from a system based initially on substrate specificity, to one based on catalytic mechanism and amino acid sequence similarities. The reason for this change is that as more proteases are identified, it is almost impossible to compare them based on classical methods alone because of the difficulty in purifying sufficient amounts of protease and the need to carry out many tests. On the other hand, it appears to be easier to determine the amino acid sequence of proteases by cDNA
cloning. However, it is not possible today to predict all properties of proteins (three dimensional structure) based on primary sequence alone. Therefore, a combination of amino acid sequence with catalytic mechanism data and three dimensional structure (when available) seems to be the current trend in protease characterization (Barrett et al., 1998). It is believed that proteases with high homology in their amino acid sequence imply a common ancestral origin, from which the entire family might have arisen by divergent evolution. Proteins that have a similar function (and three dimensional structure) but don’t share a high degree of homology are supposed to have arisen by convergent evolution (Barrett et al., 1998). Based on many studies it appears that all proteases have their origin in some common ancestral molecule(s), and have evolved to become more specialized to meet their catabolic role through repeated gene duplications, resulting in paralogous proteins (Berti et al., 1995; Neurath. 1984; Polgar, 1989; Ward et al., 1997). Because the three dimensional structure is essential for the function of macromolecules, many proteases have been crystallized and their structures determined. These results have revealed many structure-function relationships; however many details about protease function remain unknown. In general, proteases are synthesized as zymogens (inactive precursors) to prevent uncontrolled degradation. Only when they reach their target (localization), do they become activated typically by limited proteolysis, and their pro-peptide is removed. Their activation mechanisms are diverse, and they can be enzymatic or non-enzymatic. The pH is very important in these conversions which can be intra-molecular (auto-catalytic) or intermolecular (Menard et al., 1998). A
The common role for the pro-peptide is to sterically block the active site, preventing the binding of substrates; therefore, pro-peptides sometimes play the role of inhibitors (Khan & James, 1998). Furthermore, pro-peptides are important for folding, stability or sorting (Ellis, 1998; Groves et al., 1998). Another advantage of cloning is that it makes protease purification easier by providing sufficient amounts of pure protein for further biochemical characterization. Proteins obtained by cloning in bacterial cells lack glycosylation but this can be overcome by cloning in eukaryotic cells.

Currently there are two systems for the classification of proteases. One system based on the catalytic mechanism, which is a "systematic" approach adopted by the Nomenclature Committee of IUBMB (Enzyme Nomenclature Recommendations, 1992), and the "evolutionary" system by Rawlings and Barrett (1993), the updated versions of which are found in Barrett et al. (1998) and in the MEROPS database (Rawlings & Barrett, 1999).

Many different properties can be used to classify proteases such as: site of function or localization (extracellular or intracellular: soluble or bound and/or compartmentalized), optimum activity pH (acid or alkaline), distribution in the living world (among viruses, bacteria, plants and animals), distribution in organism (system, tissue or cell), degree of proteolysis (limited or complete), position of cleavage ("endo" or "exo"), specificity of degradation (broad or selective), biological function (basic or specialized), role at different stages of development, regulation (expression or activity), physiological functions (normal or pathological), possible applications (health, agriculture, industry), size in kD
(small, large), number of subunits (monomeric or multimeric), other post translational modifications (glycosylation, acetylation, sulphation, phosphorylation), temperature optimum, inhibition and activation, and rate of turnover. Generally, intracellular proteolytic systems are discussed separately from the extracellular ones. The intracellular system includes the proteasome, calpains, caspases and the lysosomal enzymes. These systems have been reviewed by Solary et al. (1998) for mammals and by Mykles (1998) for invertebrates. Other proteases exist which have not yet been characterized or classified (Glas et al., 1998). For this reason, both classification systems have room for new entries. In the IUBMB system (1992), this is represented by E.C. 3.4.99, or by “Family U” for peptidases with uncertain position (Barrett et al., 1998). The evolutionary classification is based on families which are groups of peptidases showing homology at least in the catalytic region (Barrett et al., 1998). Families are denoted by a letter and a number. The letter represents the catalytic type (Serine, Cysteine, Aspartic, Metallo and Threonine) and the number is arbitrary. Several families can make a clan, which is denoted by two letters, the first indicating the catalytic type, the second is arbitrary. For example, when different families share a common three dimensional structure, but different catalytic mechanisms, the clan is denoted by P. for peptidase. The most recent protease classification by catalytic type was done by the Nomenclature Committee of The IUBMB in 1992, which recommended the use of the term peptidase for all subclasses of E.C. 3.4. instead of the widely used term protease. This classification is available online at the SwissPro database.
Peptidases (or proteases) can be further subdivided as "endo" or "exo" peptidases, respectively E.C. 3.4.11-3.4.19 and E.C. 3.4.21-3.4.99, according to whether the cleavage occurs within or from the ends of the amino acid sequence. Proteinase is a synonym for endopeptidase. The evolutionary system of Barrett et al. (1998) is a further expansion of the former system. This new system is more complex because the position of a protease in it depends not only on catalytic mechanism type but also on sequence and when available on 3D structure as well. Today, a lot of research is devoted to proteases because of their potential wide range of applications, especially in health-related areas. As the structure and function of proteases becomes better understood, more diseases will likely appear to be linked to them (Twining, 1994; Kato, 1999). Following is a short description of proteases according to the E.C. system.

1.1.1. The serine proteases (E.C. 3.4.21.1)

The mechanisms and specificity of the serine proteases have been reviewed by Barrett et al. (1998). These enzymes contain a catalytic serine residue which is part of the so-called catalytic triad (i.e., the ordered amino acids His-Asp-Ser) in their active site. The reaction proceeds through a tetrahedral transition state intermediate. The subclasses of this group are represented by chymotrypsin, which is secreted by the pancreas of mammals, and by subtilisin secreted from B. subtilis (Polgar, 1989). Most of these proteases are extracellular and have alkaline pH optima. In mammals, chymotrypsin, trypsin and elastin have digestive functions. They are secreted by the pancreas into the duodenum in
the form of zymogens where they are activated to degrade dietary (food) proteins (Hirschi et al., 1994). The conversion of the chymotrypsin zymogens to active enzymes was the earliest studied model of protease activation by limited proteolysis (Neurath, 1957). The serine protease specificity is normally determined by the residue at the P₁ position in the substrate (Schechter & Berger, 1967). Other physiological processes in mammals such as blood clotting, complement activation and blood pressure homeostasis require serine proteases (Polgar, 1989; Twining, 1994).

1.1.2. The cysteine proteases (E.C. 3.4.22.)

This group includes proteases found primarily in eukaryotic organisms. The most distinct and best studied member of this group is papain. The papain-like proteases are the most common among this group (Turk et al., 1997). The majority of the cysteine proteases require an acid or neutral pH for optimum activity, and in most of the cases they are found intracellularly. Their specificity is determined by the amino acid in the P₂ position of the substrate (Schechter & Berger, 1967). In this group are included many plant proteases, several mammalian lysosomal cathepsins (Barrett et al., 1998) and the calpains (Mykles, 1998). The three dimensional structure of the interleukin converting enzyme (ICE) revealed a novel type of structure among the cysteine proteases (Barrett et al., 1998). The catalytic reaction of all cysteine proteases proceeds through a covalent intermediate and involves a cysteine and a histidine residue.
Eukaryotic CP comprise enzymes from parasites, fungi, plants and animals. Many parasites affect the health of humans and animals by invading their tissues, causing considerable tissue damage (Rosenthal. 1999). One of the essential factors for the pathogenicity of these parasites is the secretion of cysteine proteases. Examples of these pathogens which are sources of secreted cysteine proteases are *Entamoeba* and *Fasciola* (Piacenza et al., 1999), *Leishmania* (Mottram et al., 1998), and *Trypanosoma, Plasmodium* and *Schisostoma* (Troeger et al., 1999). Considerable research has been conducted on inhibitors which can regulate the cysteine proteases of these parasites with the hope to be able to control them (Mc Kerrow et al., 1999).

Plant cysteine proteases are found mainly in the latex and in the fruits of plants. Their mechanism has been studied in detail especially for papain. However the precise biological function of these enzymes in plants is largely unknown (Storer & Menard, 1994; Turk et al., 1997). Plant proteases are actually the most used in practice, such as for clinical applications or industrial use as in the beer industry to prevent chill-haze and also for the tenderization of meat (Polgar, 1989).

Insect proteases have been studied for both a practical need in agriculture to use less toxic pesticides on food crops, and also for basic research on protease function. Insects which feed on plants depend on the breakdown of ingested protein for their normal life cycle (Jongsma & Bolter, 1997; Mochizuki, 1998). On the other hand, plants have developed a defense mechanism with a broad spectrum of compounds to prevent insects from eating them. Several species of
Coleoptera have been studied with the hope of constructing transgenic plants with specific inhibitors against its pest, but not against other organisms in the food chain (Hilder & Boulter, 1999). From the point of view of basic science as expressed by Maiz et al. (1995), "... insects have over decades served as classical biochemical models; they have provided important insights often applicable to mammalian systems". This is the reason why several insect cysteine protease genes have been cloned. For example cysteine protease genes have been cloned from the fruit fly D. melanogaster (Matsumoto et al., 1995; Tryselius & Hultmark, 1997) and the flesh fly Sarcophaga peregrina (Yamaguchi et al., 1999). Also, many other insect models have been used to study other aspects of protease function, such as their role in silk production. In the silk moth Bombyx mori, cysteine protease has been used as a model for zymogen activation, because the pro-CP accumulates in the eggs which is later activated by limited stepwise intra-molecular proteolysis. Other studies suggest that some insect cysteine proteases are involved in egg yolk protein degradation (Zhao et al., 1996). Among the arthropods that have been studied and found to contain cysteine proteases, are the American lobster, Homarus americanus (Laycock et al., 1991), and the Norway lobster, Nephrops norvegicus (Le Bouley et al., 1995). Results from these studies, show that these proteases are cathepsin L-like enzymes with high homology between each other; however, overall invertebrate cysteine proteases have been less studied than their counterparts in mammals (Maiz et al., 1995).
The prokaryotic CP are those of viruses and bacteria. The viral cysteine proteases play an essential role in the life cycle of the virus (herpesvirus, retrovirus, hepatitis C virus, and human rhinovirus) by cleaving the high molecular weight protein precursor (a polyprotein) to yield functional products such as structural proteins required for assembly or morphogenesis of the viral particle (Patick & Potts, 1998). Bacterial cysteine proteases are found in many types of bacteria, among which several are pathogenic. The cysteine proteases of pathogens are usually secreted out of the cell and contribute to host tissue destruction, disease or virulence. Notable examples of this group of cysteine proteases are streptopain from *Streptococcus*, clostripain from *Clostridium* and gingipain from *Porphyromonas gingivalis*, the latter a pathogen involved in tooth and gum disease (Kuramitsu, 1998; Okamoto *et al.*, 1995).

1.1.3. The aspartic proteases (E.C. 3.4.23.)

Most aspartic proteases belong to the pepsin family. The pepsin family includes digestive enzymes, as well as cathepsin D (which is a lysosomal enzyme), renin, and some fungal proteases (Barrett *et al.*, 1998). Another family of aspartic proteases comprises viral proteases, the most notable of which is the protease from the HIV virus, also called retropepsin. Crystal structure determinations have shown that these enzymes are “bi-lobed” molecules with the active site positioned between each lobe providing an aspartate residue. (Barrett *et al.*, 1998).
1.1.4. Metalloproteases (E.C. 3.4.24._)

This group may be one of the older classes of proteases. Members of this group are found in bacteria, fungi as well as in higher organisms. They differ widely in amino acid sequence and structure, but the majority of these proteases contain a catalytically active Zn$^{+2}$ atom, which can also be substituted by Co$^{+2}$ or Ni$^{+2}$ without affecting activity. This group contains matrix metalloproteases such as collagenase (MMP-1, or E.C. 3.4.24.7), gelatinases (MMP-2, synonym of gelatinase A, or E.C. 3.4.24.24, and MMP-9, or E.C. 3.4.24.35), stromelysins (MMP-3, or E.C. 3.4.24.17, and MMP-10, or E.C. 3.4.24.22), matrylysin or MMP-7 and MMP-23 (Barrett et al., 1998; Creemers et al., 1998).

1.2. Invertebrate proteases

Several cathepsin-like proteases have been found in invertebrate eggs, such as in Locusta, Drosophila, Periplaneta, Artemia, Bombyx and Ornithodorus, and trypsin-like proteases have been found in Artemia and Bombyx (reviewed by Yamamoto & Takahashi. 1993).

An invertebrate organism which is also well studied and contains several proteases is the slime mold, Dictostelium discoideum, where acid activation of latent proteases can be readily observed using zymography (North et al., 1996). The finding of Souza et al. (1997) are interesting because they show that cysteine proteases with different sugar modifications segregate to distinct compartments or vesicles.
Proteases in the gut of insects have been studied in order to be able to develop transgenic plants resistant to insects (Mochizuki, 1998). Parasite proteases and potential therapeutetic applications against them were reviewed recently (Coombs & Mottram, 1997). Research shows that parasitic worms are the most difficult to control by drugs compared to protozoan organisms. Several proteases have been sequenced from *D. melanogaster* tissues, including a cysteine protease with 67% homology with mammalian cathepsins L and H (Matsumoto *et al.*, 1995). Another cysteine protease in *Drosophila* is the CP-1, a cathepsin L-like enzyme expressed in a hemocyte cell line mbn-2 of the fly (Tryselius & Hultmark, 1997). However, few invertebrate proteases have been studied thoroughly for their role in development, and in this respect *Artemia* is one of the most studied systems (Warner, 1989).

Among those systems studied so far, several demonstrate an important role for proteases in early development in both invertebrates and vertebrates. Indirect proof for the role of a cathepsin L-like enzyme in the early development of *Xenopus* was suggested by the finding that the injection of a specific antibody for the protease stops further development (Myiata & Kuto, 1997).

The identification of proteases involved in the yolk degradation of invertebrate eggs has been related to the yolk composition and its proteins. Yolk and its proteins have been studied in insects (Kunkel & Nordin, 1985; Sappington & Raikhel, 1998) and in *Artemia* (De Chaffoy de Curcelles & Kondo, 1980). In *Blattella germanica* the vitellin processing enzyme is derived from a pre-protease of maternal origin (Giorgi *et al.*, 1997). In the sea urchin, the
hatching enzyme is a metalloprotease, and its gene is expressed at the 4-cell stage, which is also the earliest zygotic gene expressed in the embryo (Ghiglione et al., 1997). Another example is the cysteine protease from the Bombyx eggs which functions in the programmed yolk degradation (Yamamoto & Takahashi, 1993; Takahashi et al., 1996; 1997). Overall, these studies indicate that as a general rule invertebrate proteases carry out early roles in development and are essential for the continuity of the developmental program.

1.3. The proteases of Artemia franciscana

All biological processes are studied in model organisms which must offer some advantage over other systems in relation to the issue under examination. The organism used in this laboratory to study the role of cysteine proteases in development is the brine shrimp Artemia franciscana. Cathepsin L-like protease activity was found to be especially abundant in the dormant cyst, therefore it was thought to be developmentally important and was targeted for further characterization.

The brine shrimp Artemia franciscana, Kellog 1906, is a Crustacean of the order of Anostraca (Brusca & Brusca, 1990). Artemia is a widely used invertebrate model organism for several types of studies such as biochemical, cellular, developmental, physiological, and ecological as discussed in the three volume monograph by Persoone et al. (1980). The main reasons for the use of Artemia in biological studies are its commercial availability, simple storage in
the laboratory, short life cycle, and easy culturing giving almost synchronous populations of developing embryos.

Initial studies on *Artemia* proteases were conducted by Urbani *et al.* (1952) and later by Bellini (1957) who observed an increase of acid protease activity during development.

Nagainis and Warner (1979) and then Warner and Shridhar (1980) purified the enzyme to near homogeneity and characterized it as a cysteine protease. These studies characterized the enzyme as a heterodimeric protease with two subunits, one catalytic of 28.5 kDa, and a larger one of 31.5 kDa as determined by SDS-PAGE. The enzyme was also characterized based on substrate specificity, response to inhibitors, pH and temperature optima, and activity kinetics. Despite several studies on the characterization of the major protease from *Artemia* embryos some discrepancies still exist among various research groups (Warner, 1987). One of the factors which may have yielded conflicting results is that some properties of *Artemia* proteases depend on the population of cysts used, i.e., parthenogenetic or bisexual (Garesse *et al.*, 1980; Perona and Vallejo, 1982).

However, when the pattern of yolk protein hydrolysis *in vivo* is compared to that *in vitro*, along with immuno-histochemical localization studies, the results support the view that the *Artemia* cysteine protease has a role in yolk utilization (Warner & Shridhar, 1985; Warner *et al.*, 1995). Although the patterns of degradation *in vivo* and *in vitro* are very similar, they are not identical. Also the pattern of lipovitellin hydrolysis with a trypsin-like enzyme is very similar with
that of the CP (Ezquieta & Vallejo, 1985). These results suggest that more than one protease may be involved in yolk protein degradation in Artemia (Warner, 1989). Another unique fact about Artemia CP is its ability to specifically degrade the Artemia protein elongation factor EF-2 at pH 6 to 6.5, with little effect on other proteins, which are completely digested at a pH 3.5-4.5 (Yablonka-Reuveni and Warner, 1979). However, it is not known if this activity occurs also in vivo.

Some trypsin-like activity has been found associated with the yolk platelets of Artemia (Ezquieta & Vallejo, 1985). Also some trypsin-like proteases have been found on the cyst envelope and some lysozyme-like activity has been found within the cysts. These results suggest that these proteases may have a role in immunity, by protecting the cyst from bacterial infections (Stabili et al., 1999). Besides the activated proteases, Artemia embryos contain a latent proteolytic activity with optimum at pH 7-8 at 40°C, with an unusual stability in 0.1 M HCl at 90°C for 30 minutes (Warner, 1989). This latent protease is activated when crude extracts are incubated for 30 minutes in pH 4.2-4.5 at 30°C. Its molecular weight (after activation) as evaluated by gel filtration is 5,000-10,000 Daltons.

If dormant cysts are incubated and allowed to hatch, an increase in alkaline protease activity appears in the cytosolic fraction about 20–24 hours after hatching of larvae (Osuna et al., 1977).

Four proteases, designed A, B, C and D have been purified and characterized from larvae (Olalla et al., 1978; Burillo et al., 1982). There is also
evidence that proteases A, B, C, and D found in larvae exist in *Artemia* adults (Sillero et al., 1980). At least two of these enzymes are serine proteases which play a role in molting (Warner and Matheson, 1998). Also, a protein called artemocyanin has been studied and thought to contain latent protease activity (Krissansen et al., 1983). Furthermore, two other SDS-stable proteases of *M*₄ 25 kD and 36 kD have been purified from adult *Artemia* homogenates (Krissansen et al., 1985).

1.4. Putative role of the major embryonic *Artemia* cysteine protease, a cathepsin L-like protease

The role of the major embryonic CP in *Artemia* has not been conclusively determined. However, there is evidence that the enzyme has a role in yolk protein utilization and in hatching of pre-nauplius larvae (De Chaffoy de Courcelles & Kondo, 1980; Perona et al., 1985; Warner & Shridhar, 1985; Warner et al., 1995) and also in molting (Warner & Matheson, 1998). Studies of other invertebrates support the role of these types of enzymes in yolk protein degradation *in vivo* (Fagotto, 1995; Yamamoto & Takahashi, 1993; Warner, 1989). However, in some studies using other models such as the cockroach *Blattella* (Liu & Nordin, 1998), the tick *Ornithodoros* (Fagotto, 1995), and the silk moth *Bombyx* (Takashi et al., 1997), cysteine proteases are present in the yolk. whereas in *Artemia* embryo the cysteine protease is not found inside the yolk platelet (Warner, 1989; Warner et al., 1995). The recent finding showing that cathepsin L mRNA is detected in unfertilized eggs of the flesh fly
Sarcophaga peregrina. is very interesting because it suggests that this protease is required in early embryogenesis (Yamaguchi et al., 1999).

1.5. Natural protease inhibitors as regulators of proteolytic activity

Many proteinase inhibitors have been identified in a variety of eukaryotic and prokaryotic organisms and characterized. The majority are proteins that range in size from 50 amino acids to 400 amino acids. Other small, non-protein inhibitors of proteases are also known to be produced in some microrganisms (Bode et al., 1992). The endogenous proteinaceous inhibitors are generally protease class-specific, except for the alpha-macroglobulin family. The most studied of these proteins are the serine protease inhibitors.

Recently, considerable progress has been made in the study of the family of cysteine protease inhibitors which includes cystatins, stefins, and kininogens (Hiltke et al., 1999).

As well, the family of metalloprotease inhibitors contains proteins such as the TIMP, tissue inhibitors of metalloproteases (Del Bigio & Seyoum, 1999) and the potato carboxidase inhibitor (PCI).

The aspartyl protease inhibitors are not well known or characterized but they appear to exist in several organisms (Sato et al., 1994; Baudys et al., 1991).

Among the protease inhibitors that have been identified, some are also found extracellularly. Most researchers believe that endogenous protease inhibitors are involved in the regulation of the proteolytic activity, either within
the same organism or across species. The role of protease inhibitors across species can be illustrated with some insects which feed on plants and have digestive proteases to utilize plant proteins. Alternatively, those plants that contain inhibitors to insect proteases are more able to resist excessive damage by their pests.

Throughout normal cell and tissue function, endogenous proteolytic activity is strictly controlled by endogenous inhibitors, pH and compartmentalization and when it happens that the balance is shifted in one direction or another, pathologic conditions may arise. Examples of pathologies related with disrupted balance between proteases and their endogenous inhibitors are inflammation, degenerative and immunological diseases, allergy and cancer (Nduwimana et al., 1995).

1.6. Protease inhibitors in *Artemia*

Several biochemical studies have demonstrated the presence of protease inhibitors at different developmental stages in *Artemia*. (Ezquieta & Vallejo, 1985; Nagainis and Warner, 1979). Homogenates of encysted embryos of *Artemia* show cysteine protease inhibitor activity (CPI) in the cytosolic fraction. Four cysteine protease inhibitors have been purified from the cytosolic fraction of *Artemia* embryos and denoted as CPI-1a, CPI-1b, CPI-2 and CPI-3. They have been biochemically characterized (Warner and Sonnenfeld-Karcz, 1992). In previous studies the developmental profile for the cytosolic inhibitors was determined (Warner, 1987). According to that study, the ratio of enzyme-to-
inhibitor activity was found to dramatically increase after hatching, suggesting that the role of the cysteine protease inhibitors may be to regulate proteolysis that may occur at the time of the cytosol acidification during anoxia (Utterback and Hand, 1987; Warner et al., 1997).

The yolk platelets contain lipovitellin which is supposed to inhibit the trypsin-like activity in *Artemia* by a mechanism similar to that of macroglobulins in mammals.

### 1.7. Objectives

The objectives of this thesis are to acquire new information about the *Artemia* cysteine protease large subunit to better understand previous data, such as the immuno-histochemical localization study and the potential role of this protease in development (Warner *et al.*, 1995). More specifically my objectives are: 1) to obtain partial amino acid sequence which can be used to design primers to help in sequence deduction by cDNA cloning; 2) to develop a method to dissociate the CP subunits without denaturing them; 3) to purify each of the CP subunits in the native form and compare the activities of the catalytic subunit alone or in the presence of the large subunit; and 4) to test whether dissociation occurs by a physical or an enzymatic process and whether association of the separated subunits can occur.
2. MATERIALS AND METHODS

2.1. Materials

The encysted embryos of *Artemia franciscana* used in these experiments were from Sanders Brine Shrimp Co., Ogden, Utah, lot No. 12715.

Grinding was carried out using an automated mortar and pestle from Torsion, Montreal, Quebec.

All chromatography supports including Sephadex G-25 (M), Sephadex G-150 (SF), Sephadex G-75 (SF), DEAE Sephadex A-50, Concanavalin A Sepharose. Mono Q and Mono S (FPLC®) were from Pharmacia Biotech AB, Uppsala, Sweden.

The C-18 (7 μ) reverse phase column was from Phenomenex, Torrance, California. All high performance liquid chromatography was carried out with an Altex chromatograph (Beckman Instruments, Fullerton, California) with a model 421 controller. Protein elution from FPLC or HPLC columns was monitored by a Beckman 20 μl analytical optical unit at 280 nm and registered on paper using a flat bed recorder (Kipp and Zonen, Holland).

The Bio-Rad Miniprotean system was used for gel electrophoresis with an LKB power unit, model 2197. The Trans-Blot unit from Bio-Rad was used for Western blotting. The nitrocellulose membrane used for Western blotting and the HRP-conjugated goat IgG were from Bio-Rad, Hercules, California.

A Sorvall RC-2B preparative centrifuge and a Beckman L5-65 ultracentrifuge were used to prepare cell-free extracts.
Membrane concentrators. YM-10 and C-10, were from Amicon.

Oakville, ON.

The protein determination kit (Bicinchoninic acid Assay) was from Pierce. Rockford, IL.

All other UV/VIS analyses were carried out using a spectrophotometer from Pharmacia BT. Cambridge, U.K.

The intensities of protein bands in Western blots were quantified using a scanner model CS-930 with recorder model DR-2 from Shimadzu Japan.

All reagents used were chemical grade or better.

2.2. Methods

2.2.1. Isolation of the major cysteine protease from Artemia cysts

The isolation of the major Artemia CP was carried out using a procedure based on the original method by Warner and Shridhar (1985), with some modifications, as described by Aiton (1997). All steps in the procedure were carried out at 4°C, unless stated otherwise.

First, cysts were hydrated overnight at 4°C in 25% sea water and after washing well, they were collected on a Buchner-type funnel with a coarse fritted disc (Kimax 600 ml/90C, from Kimble Glass Co.). Afterwards cysts were homogenized in 25 g batches with 7 ml of homogenization buffer A (50 mM Tris-Cl, pH 7.3, 5 mM KCl, 1 mM DTT, and 10 mM MgCl₂) using an automated mortar and pestle. After 15 minutes of homogenization, 115 ml of additional
homogenization buffer was added to the ground cysts and the suspension was stirred slowly for 15 minutes.

The homogenate was centrifuged at 10,400 g for 20 minutes and the supernatant from this step was filtered and centrifuged using a Ti60 rotor at 150,000 g for 2.5 hours. The supernatant from this step was filtered again and designated as the S-150 fraction.

Next, the S-150 fraction was adjusted to contain 25% ammonium sulfate and any insoluble material was allowed to precipitate overnight. The ammonium sulfate-treated fraction was centrifuged at 12,000 g for 30 minutes. The pellet was discarded and the supernatant was adjusted to contain 75% ammonium sulfate. Proteins were again allowed to settle overnight, then they were collected by centrifugation and the precipitate from this step (the 25-75% ammonium sulfate precipitate) was saved. Before gel filtration, the 25-75% ammonium sulfate precipitate was suspended in a minimal volume of column buffer 1 (15 mM K-PO₄, pH 6.8. 25 mM KCl, 10% glycerol) and cleared by centrifugation.

The soluble fraction was loaded on a G-25 (M) Sephadex column (3.7 x 40 cm) that had been previously equilibrated with column buffer 1, and eluted with the same buffer. One large fraction which could be visualized as an orange band was collected in a volume of about 100 ml.

This fraction was loaded directly onto a DEAE Sephadex A-50 column (4.2 x 40 cm) which had been previously equilibrated with column buffer 1. The column was then washed with the same buffer until the effluent had an absorbance of less than 0.02 at 280 nm. Bound proteins were eluted with a 25 to
750 mM KCl gradient in one liter of column buffer 1 which was pumped through the column at a flow rate of 1 ml/minute. Fractions of about 10 ml were collected and assayed for protein and cysteine protease activity.

Column fractions containing the majority of the activity were pooled and loaded onto a Concanavalin A (Con A) Sepharose column (1.7 x 21 cm) which had been equilibrated with column buffer 2 (15 mM K-PO₄, pH 6.8, 200 mM KCl). The column was washed until the effluent had an absorbance of less than 0.02 at 280 nm, and protein was eluted with column buffer 3 (15 mM K-PO₄, pH 6.8, 200 mM KCl, 1M glucose). Fractions of approximately 3 ml were collected and assayed for protein and cysteine protease activity.

The fractions containing the majority of the cysteine protease activity were pooled and concentrated by pressure filtration using a YM-10 membrane, and then by centricon C-10 to about one ml.

The cysteine protease preparation which eluted with glucose from the Con A column was concentrated then loaded onto a G-150 Sephadex (SF) column (1 x 80 cm) which had been equilibrated with column buffer 1. Fractions of about 1 ml were collected and each was assayed for protein content and cysteine protease activity. Those fractions containing the majority of the protease activity were pooled and concentrated by pressure filtration (membrane YM-10) and centricon (C-10) to about 1 ml.

About 4 mg protein was loaded on a Mono Q column (0.5 x 5 cm) which had been equilibrated with column buffer 1, and eluted with a 25 to 275 mM KCl gradient in column buffer 1, at a flow rate of 1 ml/min over 20', in order to
resolve the cysteine protease isoforms. Proteins eluting from the column were monitored at 280 nm for protein content, then each column fraction was analyzed for cysteine protease activity using the TNBS method.

Fractions containing the different isoforms were identified and pooled separately. The pooled samples were concentrated to a final volume of 0.25 ml using centricon units and re-equilibrated with column buffer 1.

The partially resolved isoforms were then re-chromatographed on the same column to enhance the purity of each isoform.

Initially, purification of the large subunit of the CP was performed under denaturing conditions using HPLC at room temperature. About 100 μg of the purified protease after the Mono Q column (CP isoform 3 or CP-3) was adjusted to a volume of 0.25 ml by addition of HPLC column equilibration buffer, then it was applied onto a C-18 reverse phase column (4.6 x 250 mm) which had been equilibrated with a buffer containing 12% acetonitrile and 0.1% trifluoracetic acid. Proteins were eluted with a 12-60% acetonitrile gradient containing 0.1% trifluoracetic acid at a flow rate of 1 ml/minute over 30 minutes. Protein elution was monitored at 280 nm.

Column fractions of 0.3 ml were collected and the fractions of interest were lyophylized in sterile 1.5 ml Eppendorf tubes of to remove the organic solvents. In this procedure, the Eppendorf tube containing the HPLC fractions was covered with two layers of parafilm which was pierced twice with a 26 gauge needle. The tubes were cooled to -70°C (until completely frozen), then transferred immediately to a lyophylizer and freeze-dried overnight.
2.2.2. Cysteine protease assay

a) General cysteine protease assay

Assays for cysteine protease activity were carried out according to Nagainis and Warner (1979), with some modifications. Most assays were conducted at pH 4 and at 40°C in a 200 μl reaction volume containing 1mM EDTA, 2.5 mg/ml protamine sulphate, 0.1 M sodium acetate, pH 4, and the required volume of sample.

All reaction vessels were incubated in a water bath at 40°C for 30 minutes or longer depending on the activity of the sample. The enzyme reactions were stopped by adding 950 μl stop buffer 1 (which contained 53 mM NaOH and 53 mM Na₂B₄O₇) to 50 μl of the reaction mix. The color reaction was carried out by adding 25 μl of 220 mM trinitrobenzene sulfonic acid (TNBS) to every tube and keeping them at room temperature for 15 minutes. The amino groups released during proteolysis combine with the TNBS to form a yellow colored complex with maximum absorbance at 420 nm. Colour development was stopped 15 minutes after the addition of the TNBS reagent with 2 ml of stop buffer 2 containing 0.1 M NaH₂PO₄ and 1.5 mM Na₂SO₃. Samples were mixed by vortexing and after 15 minutes their absorbance was measured in the spectrophotometer at 420 nm.

A standard curve was generated using known quantities of alanine to determine the concentration of free amino groups released during proteolysis.
b) Activity of the catalytic subunit alone or in the presence of the large subunit

The purified subunits obtained from a Mono S column applying the method described in 2.2.12, were used to determine the activity of the small (catalytic) subunit (SSU) alone or in the presence of equimolar amounts of the large subunit (LSU) at pH values 4.0, 5.0 and 6.0. Activities were assayed according to Nagainis and Warner (1979). To determine cysteine protease activity in the presence or absence of the large subunit, reaction vessels of 200 µl were set up as described in the preceding section except that the reactions were carried out at pH 5 and pH 6 as well. Control reaction vessels contained water in place of the large subunit. At zero and other time intervals 45 µl was taken from the reaction and immediately transferred into 955 µl of stop buffer 1 and the TNBS assay was carried out as described above.

2.2.3. Protein assays

The protein concentration in samples from G-25 Sephadex, DEAE Sephadex, G-150 Sephadex (SF) and G-75 Sephadex (SF) columns was estimated by measuring the absorbance at 280 nm in the spectrophotometer, against bovine serum albumin at 1mg/ml as a standard.

Other protein determinations were performed based on the Pierce reagent (BCA), according to Smith et al. (1985). Bovine serum albumin was used as a protein standard and the assays were carried out according to the instructions of the manufacturer.
2.2.4. The cyanogen bromide treatment of the *Artemia* large subunit

Cyanogen bromide (CNBr) was used to generate peptide fragments for amino acid sequencing, because it specifically cleaves after methionines in the protein. The method was carried out according to Marco *et al.* (1998) with some modifications.

Two micrograms of the large subunit (protein) of the *Artemia* cysteine protease were dissolved in 40 µl of 70% formic acid in an Eppendorf tube, then 80 µg of CNBr in 70% formic acid were added and the tube was incubated overnight in the dark. Afterwards the CNBr-treated sample was diluted with distilled water up to 0.5 ml and the sample was freeze-dried. The same reaction was carried out with 2 µg of myoglobin as a positive control.

2.2.5. Gel electrophoresis with amino group protection

Gel electrophoresis was conducted according to the method of Laemmli (1970) on 7-18% polyacrylamide gradient gels (T%: 7-18, C%: 2.6). The CNBr-treated samples were electrophoresed on a 7-18% gradient SDS-PAGE according to the Bio-Rad instructions for protein sequencing using PVDF membranes (bulletin 1609). The gel was pre-run with 0.1 M thioglycolate in the upper buffer for 0.5 hours, then proteins were electrophoresed for 1.5 hours and next transferred to a PVDF membrane using a Trans-Blot unit (Towbin *et al.*, 1979). The transfer was carried out at 100 mA and 50 volts for 18 hours at room temperature. The PVDF membrane with the transferred proteins was soaked in
0.025% Coomassie blue R-250 in 40% methanol and afterwards destained for 5 minutes in a solution containing 50% methanol. Acetic acid was omitted from the stain and destaining solutions since it can cause acetylation of the N-terminus of proteins.

2.2.6. Attempts to obtain amino acid sequence from the large subunit of *Artemia* CP

All attempts at protein sequencing were done at the Eastern Quebec Peptide Sequencing Facility of The Hospital Center, University of Laval, Quebec by Dr. Sylvie Bourassa.

First the HPLC-purified large subunit isoform (CP-3L) was electrophoresed on a 7-18% (gradient) SDS-PAGE and stained with Commassie blue. Next, the gel slice containing the protein band was cut from the gel using a clean scalpel, transferred to an Eppendorf tube, then sent to the sequencing laboratory where in-gel digest was carried out with protease Lys C. The peptides generated from the Lys C treatment were eluted from the gel and about 10% of the sample was sent to Columbia University in New York for mass spectrometry (MS) analysis. The remaining 90% of the peptides were separated by reverse phase HPLC and five peptides were chosen to be sequenced using an automated Edman degradation protocol.

Another sample of about 2 micrograms of CP-3L was treated with CNBr and the digestion products were electrophoresed on a 7-18% gradient SDS-PAGE. After transfer of the resulting protein products to a PVDF membrane,
proteins on the membrane were stained with Coomassie blue and the membrane destained avoiding acetic acid. A protein band at about 25 kD representing the largest CNBr fragment was cut from the membrane and sent for sequencing.

2.2.7. Gel filtration of proteins from subcellular fractions of *Artemia* cysts

a) Preparation of *Artemia* subcellular fractions

Encysted embryos of *Artemia* were hydrated in 25% sea water for several hours, then collected on a cloth filter, blotted to remove excess of water and weighed. One gram of hydrated cysts was homogenized in 5 ml of a buffer containing 150 mM sorbitol, 70 mM potassium gluconate, 5 mM KH$_2$PO$_4$ and 35 mM Hepes, pH 6.8 with a glass hand homogenizer (Lurex).

The homogenate was centrifuged at 1.630 g for 10 minutes, and the supernatant resulting from this step was centrifuged again at 15,000 g for 30 minutes. The pellet from this step which represented the mitochondrial/lysosomal fraction was designated M+L fraction, and the supernatant or the aqueous cytoplasm was designated as PMF (or post mitochondrial fraction).

b) Gel Filtration of the post-mitochondrial fraction (PMF)

The PMF representing the aqueous cytoplasm of *Artemia* encysted embryos was concentrated to less than 0.5 ml using a centicon C-10 concentrator and then loaded onto a G-150 Sephadex (SF) column (1x 80 cm)
which had been previously equilibrated with column buffer 1 and calibrated using various molecular weight standards.

Fractions of about 1 ml were collected from the column and protease activity eluting from the column was assayed using the TNBS method. The protein content was determined by measuring the absorbance at 280 nm. The protease activity was assayed in the presence and absence of the fluoromethyl ketone inhibitor (Z-Phe-Ala-CH₂F) and also with pepstatin using the TNBS method (see above) to ensure that the proteolytic activity detected was only due to the cysteine protease.

c) Gel filtration of the myochondrial/lysosomal lysate from *Artemia* embryos

The mitochondrial/lysosomal (M+L) fraction from 5 g of cysts was first washed with the homogenization buffer to remove any cytoplasm. then it was lysed with the same buffer containing 0.05% Tween 20 and centrifuged to pellet any precipitate.

The soluble material from this step was applied to a calibrated G-150 Sephadex (SF) column and eluted with column buffer 1. Fractions of about one ml were collected and assayed for protein content by measuring the absorbance at 280 nm. and for protease activity using the TNBS method (see above). The M+L fraction was also analyzed for the presence of the two subunits of the *Artemia* cysteine protease by Western blotting.
d) Gel filtration of CP after treatment to dissociate subunits

A sample of concentrated post DEAE crude concentrate containing about 5000 mEU was first equilibrated in a buffer containing 20 mM sodium acetate and 1 mM EDTA at pH 5 and allowed to stay overnight on ice. After clearing the suspension by centrifugation at 10,000 g for 30 minutes the sample was loaded on a calibrated G-75 Sephadex (SF) column (1x 55 cm) which had been equilibrated with the same buffer. Protein was eluted with the same buffer and samples of about 1 ml were collected from the column and assayed for cysteine protease activity using the TNBS method to determine the fraction volume and molecular weight of the cysteine protease.

2.2.8. Immunoaffinity purification of cysteine protease subunit-specific antibodies

The purification of *Artemia* cysteine protease antibodies, specific to each subunit of the CP was done according to the method described by Lu (1991) and Lu and Warner (1991), starting with the total cysteine protease antiserum.

First, purified *Artemia* cysteine protease was electrophoresed on a 7-18% gradient SDS-PAGE and subsequently electroblotted on a 0.2 μm (pore size) nitrocellulose (NC) membrane at 100 mA and 50 volts for 18 hours. Next, the NC membrane was soaked in distilled water for 5 minutes and stained with Ponceau S for two minutes to visualize the protein bands.

The bands on the membrane corresponding to the large and small subunits of the CP were cut out and placed in separate small Petri dishes, with 5
ml of TTBS (10 mM Tris-Cl, pH 7.4, 140 mM NaCl, 0.1% Tween 20) containing 9% skimmed milk and incubated for 1-2 hours at room temperature with gentle shaking to block nonspecific sites. Afterwards, the NC strips were washed three times, 10 minutes each with TTBS and once with TBS (same as TTBS but without Tween 20) for five minutes.

Subsequently, the NC strips were incubated with a 1:10,000 dilution of crude rabbit CP antiserum for 2 hours at room temperature with gentle shaking. After washing unbound antibody from the strips, each strip (containing the large and the small subunit) of the CP was placed in separate Eppendorf tubes and eluted with three washes (0.35 ml each) of elution buffer containing 5 mM glycine, 0.5 M NaCl, pH 2.3 (HCl), 0.5% Tween 20 and 100 µg/ml BSA. After elution, the pH of the eluate was adjusted to neutrality with a 0.5M K₂HPO₄ solution. The eluted antibodies were concentrated with a centricon C-10 and stored at 4°C.

2.2.9. Immunodetection of cysteine protease subunits by Western blotting

The immunodetection of CP on Western blots was carried out using the immunoaffinity purified antibodies described above. After separation of the proteins on a 7-18% gradient gel SDS-PAGE using 125 V and 30 mA, proteins were electro-transferred onto a nitrocellulose membrane according to the original method of Towbin et al. (1979). The membrane was then soaked in water to remove the transfer solution and stained with Ponceau S to visualize any proteins transferred from the gel including the molecular weight markers. All subsequent
reactions were carried out at room temperature. Next the membrane was placed in the TTBS solution containing 9% skimmed milk and incubated with gentle shaking for 1 hour.

After washing the membrane in TTBS three times for 10 minutes each and 5 minutes in TBS, the membrane was incubated with a 1:500 dilution of the purified antibody in TTBS for 2 hours, with gentle shaking. The membrane was washed as described above and incubated in TTBS containing a 1: 3,000 dilution of HRP-conjugated goat IgG with gentle shaking for 1 hour. The membrane was washed as described above, then placed in 20 ml of 0.05 M Tris(HCl), pH 7.4. containing 5 mg diaminobenzidine (DAB) and 70 μl of 3% H₂O₂. After a few minutes, and the appearance of yellow/brown bands, the reaction was stopped by washing the membrane in distilled water.

2.2.10. Elution of the purified Artemia cysteine protease from the Con A column

A Con A column (0.5 x 3 cm) was washed and equilibrated with a buffer containing 15 mM potassium phosphate at pH 6.8 and 0.3 M KCl. A sample of purified CP after the Mono Q column, containing 3 mg protein and 5000 mEU, was loaded on the column and the column was washed with 5 ml of wash buffer. Fractions of 1 ml were collected and after the absorbance at 280 nm had dropped below 0.02, the elution was continued with the equilibration buffer to which glucose was added, to form a 0-1 M gradient. Samples of 1 ml were
collected and both protein content and cysteine protease activity were
determined as described above.

Aliquots from each fraction were also run on a 7-18% SDS-PAGE, and
the gel was stained with Commassie blue. The above procedure was repeated
using galactose in the elution buffer in place of glucose.

2.2.11. Fractionation of native cysteine protease subunits by FPLC on a
Mono S column

Separation of the cysteine protease subunits by FPLC was carried out at
room temperature. A concentrated post-DEAE sample with about 2000 mEU
was loaded on a Mono S column which had been previously equilibrated with a
buffer containing 20 mM sodium acetate, pH 5 and 1 mM EDTA. Protein was
eluted with a gradient of 20 mM to 700 mM NaCl in the same buffer using a
flowrate of 1 ml/minute over a period of 20 minutes. Fractions of 1 ml were
collected and assayed for cysteine protease activity.

Aliquots from each fraction were electrophoresed on a 7-18% gradient
SDS-PAGE and electroblotted onto a nitrocellulose membrane, which was
subsequently stained with Ponceau S to visualize protein bands then
immunostained as described above.

After determining the elution position of each of the cysteine protease
subunits with a crude protein preparation, the same conditions were used to
resolve the CP subunits starting with purified, dimeric cysteine protease.
3. RESULTS

An overview. In this thesis research experiments were conducted to characterize the large subunit (LSU) of the *Artemia* cathepsin L-like cysteine protease (CP) which is a component of the primary protein hydrolase in embryos and young larvae of *Artemia franciscana*. To achieve the above objective, the following approach was followed:

- First, the cysteine protease was isolated from encysted embryos and purified to homogeneity using a multi-step chromatographic scheme. Several isoforms of the CP were resolved by FPLC on a Mono Q column, and one of them (CP-3) was chosen for the isolation of the large subunit by HPLC.

- Amino acid analysis of fragments generated by endoproteinase Lys C or CNBr digest of CP-3L was carried out to obtain sequence data of the protein.

- Isolation of the CP subunits from the third isoform (CP-3) was conducted under native or non-denaturing conditions to test whether re-association of the large and small subunits of CP-3 could occur, after dissociation at pH 5.

- Purification of the CP subunits under native conditions was carried out to compare the influence of the large subunit on protease activity at pH 4, 5 and 6. As well, experiments were conducted to shed light on the potential mechanism of CP subunit dissociation.

- Gel filtration analyses of the aqueous cytoplasm and lysosomal fractions were carried out at pH 6.8 followed by electrophoresis and Western blotting to determine whether either subunit of the cysteine protease exists as a monomer in the *Artemia* embryo.
3.1. Isolation of the large subunit of isoform three (CP-3L) of the cathepsin L-like cysteine protease

3.1.1. Purification of the major *Artemia* cysteine protease

Purification of the major embryonic CP of *Artemia franciscana*, a cathepsin L-like cysteine protease, was carried out as described in Materials and Methods according to Warner and Shridhar (1985) as modified by Aiton (1997). Several batches of dormant cysts were needed to obtain a sufficient amount of pure protease for analyses. The results in Table 1 summarize the purification scheme used to purify the protease. Because of the many steps involved during purification, the total activity recovery was 10-15% of the original activity in the post-ribosomal fraction (S-150). The purity of the protease was confirmed by gel electrophoresis as shown in Fig. 1. A 7-18% polyacrylamide gradient was chosen because it had been determined to give the best resolution of the CP subunits (A.Warner, personal communication). After purification of the CP and staining with Coomassie blue, it was clear that the *Artemia* embryo CP is composed of two subunits, a small subunit (SSU) of 28.5 kD and a larger subunit (LSU) of 31.5 kD. The only visible bands on the polyacrylamide gel. In previous studies both of these subunits were found to be present after all purification steps and to resist separation under various conditions used in the purification of the CP activity.
Table 1. Summary of the purification steps in the isolation of the *Artemia* cathepsin L-like cysteine protease

<table>
<thead>
<tr>
<th>Purification step</th>
<th>Protein (mg)</th>
<th>Total Activity (EU)</th>
<th>Specific Activity (Eu/mg protein)</th>
<th>Activity Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-150 ultracentrifgation</td>
<td>5,000</td>
<td>450</td>
<td>0.09</td>
<td>100</td>
</tr>
<tr>
<td>25-75% ammonium sulfate</td>
<td>4,500</td>
<td>441</td>
<td>0.09</td>
<td>98</td>
</tr>
<tr>
<td>G-25 Sephadex</td>
<td>3,800</td>
<td>432</td>
<td>0.11</td>
<td>96</td>
</tr>
<tr>
<td>DEAE A-50 Sephadex</td>
<td>1,900</td>
<td>391</td>
<td>0.20</td>
<td>87</td>
</tr>
<tr>
<td>ConA Sepharose</td>
<td>53</td>
<td>315</td>
<td>5.90</td>
<td>70</td>
</tr>
<tr>
<td>G-150 Sephadex (first)</td>
<td>20.8</td>
<td>250</td>
<td>12</td>
<td>56</td>
</tr>
<tr>
<td>G-150 Sephadex (second)</td>
<td>10.8</td>
<td>157</td>
<td>14.5</td>
<td>35</td>
</tr>
<tr>
<td>Mono Q</td>
<td>4.3</td>
<td>67</td>
<td>15.5</td>
<td>15</td>
</tr>
</tbody>
</table>

* The values are averages from several batches of *Artemia* cysts and based on 100 g dry weight as starting material.
Fig. 1. Subunit composition of the cathepsin L-like cysteine protease of *Artemia franciscana*

About 1 µg of purified cysteine protease of *Artemia* embryos after chromatography on a Mono Q column (FPLC) was electrophoresed on a 7-18% (gradient) SDS-PAGE and stained with Coomassie blue.

Lane 1: Molecular weight standards

Lane 2: The purified cysteine protease subunits:

- Large subunit (LSU), 31.5 kD
- Small subunit (SSU), 28.5 kD
3.1.2. Separation of the cysteine protease isoforms

Researchers in our laboratory previously found that *Artemia* embryos contain several isoforms of the major cysteine protease which can be separated by FPLC on a Mono Q column using a KCl gradient (A. Warner, personal communication; A. Aiton, 1997). Therefore, the major isoforms of the CP were separated on a Mono Q column (0.5 x 5 cm) as shown in Fig. 2. Two cycles of fractionation on Mono Q were necessary to isolate individual isoforms.

Purification of the third isoform (CP-3) is shown in Fig. 2B. During the first chromatographic run, each activity peak was pooled separately as shown in Fig. 2A. Since it was previously established that the second peak of protease activity (see Fig. 2A) represented the major isoform (denoted as CP-3), the fractions containing CP-3 were pooled, concentrated and re-run on Mono Q to enrich for the third isoform (A. Warner, personal communications).

The results of these chromatographic steps yielded isoform CP-3. The same procedure was followed to isolate the other isoforms (CP-1, CP-2, CP-4 and CP-5) for use in other analyses (Warner et al., 1999).

3.1.3. Separation of the large and small subunit of *Artemia* cysteine protease by HPLC

From previous studies in this laboratory, it was determined that the *Artemia* cysteine protease subunits could be fractionated by HPLC under denaturing conditions (A. Warner unpublished observations, and A. Aiton, 1997).

Therefore, subunit separation was carried out using a C-18 reverse phase column
Fig. 2. Purification of the cathepsin L-like cysteine protease CP-3 from *Artemia* embryos

A. The partially purified cathepsin L-like CP from the second G-150 SF column was chromatographed on a Mono Q column (0.5 x 5 cm) equilibrated with column buffer 1. Proteins were eluted from the column with a 25-275 mM KCl gradient in column buffer 1 at a flow rate of 1 ml/min over 30 minutes and 0.33 ml fractions were collected. Total protein was monitored by recording the absorbance at A₂₈₀ nm (----). The protease activity (→→) was determined using the TNBS assay (A₄₂₀ nm). The various cathepsin L-like isoforms were numbered as indicated.

B. Re-chromatography of cathepsin L-like isoform 3 (CP-3) in panel A on a Mono Q column to resolve it from isoforms 1 & 2 and 4-6. The contents of all peak tubes were pooled, concentrated and stored at -15°C for further analyses.
as described in the Materials and Methods. Under these conditions the large subunit (LSU) and the small subunit (SSU) of the cysteine protease eluted at about 19 and 27.5 minutes, respectively. The elution characteristics of the LSU and SSU of the third isoform of *Artemia* cysteine protease (CP-3) are shown in Fig. 3. The other isoforms (CP-1, CP-2, CP-3, CP-4 and CP-5) showed the same elution profile from the HPLC column (data not shown), but because of its abundance the third isoform was selected for further analysis.

The LSU of CP-3 was concentrated by lyophilization and a small amount of the protein was electrophoresed on a 7-18% (gradient) SDS-PAGE gel together with native (purified) CP-3 as a positive control to confirm the identity and the purity of the protein. The gel was stained with Coomassie blue, destained and then scanned. These results are given in Fig. 4.

### 3.2. Amino acid analysis of peptides generated by CNBr and Lys C treatment of CP-3L protein

#### 3.2.1. Fragmentation of CP-3L and separation of the resulting peptides by SDS-PAGE and HPLC

Previous attempts to sequence CP-3L indicated that the N-terminal amino acid was blocked (A. Warner, personal communication). Thus two procedures were used to generate peptide fragments for N-terminal amino acid analysis. One treatment used the protease Lys C to generate peptides, and the second used CNBr to cleave the protein at methionine residues. These treatments are described in Materials and Methods.
Fig. 3. Separation of the *Artemia* cysteine protease subunits using HPLC

A sample of the cathepsin L-like cysteine protease isoform 3 (CP-3) protein from the Mono Q column (see Fig. 2B) was adjusted to contain 12% acetonitrile and 0.1% trifluoracetic acid (TFA) and applied to a RP-HPLC column (C-18, 4.6 x 250 mm) which had been equilibrated with the same solution. After washing the column with equilibration buffer (4 ml), the cysteine protease subunits were eluted with a linear gradient (12-60%) of acetonitrile in 0.1% TFA at a flow rate of 1 ml/min, over 30 minutes. Protein eluting from the column was monitored at 280 nm (—). The small subunit (SSU) of the protease eluted at 19 minutes, while the large subunit (LSU) of the protease eluted at 27.5 minutes. The contents of each peak were pooled and lyophilized to remove the organic solvents.
Fig. 4. Electrophoresis of the large subunit of *Artemia* cysteine protease third isoform (CP-3)

A sample containing about 1μg of the lyophylized LSU of CP-3 protein (CP-3L) from the HPLC column (Fig. 3) was electrophoresed on a 7-18% (gradient) SDS-PAGE gel and stained with Coomassie blue.

Lane 1: Purified CP-3 protein as a control

Lane 2: Large subunit of the *Artemia* CP-3 (CP-3L)
The peptides that eluted from the gel after the Lys C treatment were collected and 10% of the sample was sent to the protein laboratory of Columbia University for mass spectrometric analysis. The MS profile generated in this analysis is shown in Appendix 1. A search of the MS-related protein databases did not yield a match with Lys C peptides generated from known proteins.

Given that the MS analysis provided no clues to the identity of the entire polypeptide (i.e., CP-3L), the Lys C derived peptides (about 90% of the initial sample) were fractionated by reverse phase HPLC and five of the 18-20 peaks were sequenced. The HPLC profile of the Lys C digest of CP-3L is given in Appendix 2. Peaks no. 5, 8, 9 and 11 were submitted for amino acid sequencing, the results of which are given in Table 2. Since the Lys C generated peptides were not suitable for the design of nucleotide primers we used another approach to obtain useful amino acid sequence.

Cyanogen bromide which cleaves proteins at the carboxyl side of methionine was chosen to generate fragments of the CP-3L which could be sequenced from the N-terminus. Given that most proteins have only a few methionine residues, this treatment allowed us to obtain a large, unblocked peptide for sequencing with a high degree of confidence as each CNBr-generated peptide would not have a blocked N-terminus. About 2 μg of CP-3L were treated with CNBr, and the products of the reaction were separated by electrophoresis on a 7-18% (gradient) SDS-PAGE. Proteins in the gel were transferred to a PVDF membrane which was stained with Coomassie blue in the absence of acetic acid to avoid blocking the N-terminal amino acid. Treatment of CP-3L
Table 2. Amino acid sequences of selected peptide fragments obtained from CNBr and endoproteinase Lys C treatment of CP-3L.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Peptide sequenced</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lys C digest</td>
<td>Peak 5</td>
<td>(K)LIPLYTK</td>
</tr>
<tr>
<td></td>
<td>Peak 8</td>
<td>(K)YGEEK</td>
</tr>
<tr>
<td></td>
<td>Peak 9</td>
<td>(K)SLIFSIK</td>
</tr>
<tr>
<td></td>
<td>Peak 11</td>
<td>(K)FTLFASTK</td>
</tr>
<tr>
<td>CNBr digest</td>
<td>25 kD fragment</td>
<td>(M)ANVIDHEGKFTLFAPTNEAFKR</td>
</tr>
</tbody>
</table>

1 The origin of these peptides is shown in Appendix 2

2 Single letter code was used to identify amino acids in peptides
with CNBr generated an intense protein band of about 25 kD, two minor bands of about 16 and 6.5 kD, and a band at 31 kD which could represent either an unmodified CP-3L or the large subunit lacking the N-terminal methionine residue. These results are shown in Fig. 5. The band of about 25 kD was cut from the membrane and sent for amino acid sequencing because it was more homogenous on the gel and also on the PVDF membrane than the other lower molecular weight bands. The results of sequencing the first 22 amino acids of the 25 kD protein are given in Table 2 together with amino acid sequence of four peptides generated by protease Lys C treatment of CP-3L.

3.2.2. Protein database search for homologues having the highest local similarity with the peptides sequenced

The amino acid sequence information generated from the large CNBr fragment of CP-3L was compared to the amino acid sequences in the protein database. The longest fragment (22 amino acids) was aligned to several proteins with an acceptable similarity. The non redundant GenBank database was searched with BLASTP (1.4.11) for local alignment. Among the proteins with the highest-scoring segment pairs were: chick transforming growth factor induced protein (cβIG-H3), mouse bone ossification factors, (mOSF), fasciclin I from Drosophila. (FAS I) and the 32 kD protein component of the hyaline layer (HLC-32) of the sea urchin embryo. Alignments of the 22 amino acid CNBr-generated fragment from the large subunit of Artemia (CP-3L) with each of the best matching proteins revealed that Artemia CP-3L has a conserved domain
Fig. 5. Electrophoretic separation of CP-3L treated with CNBr

About 2 micrograms of CP-3L were treated with CNBr and the digestion products electrophoresed on a 7-18% gradient SDS-PAGE. After separation, proteins were transferred to a PVDF membrane, which was subsequently stained with Coomassie blue to identify any bands that resulted from the treatment, and the membrane scanned as shown here.

Lane 1: Molecular weight standards (Bio-Rad 161-0317)

Lane 2: CP-3L

Lane 3: CNBr treated CP-3L. The 25 kD band indicated by the arrow was sent for amino acid sequencing.
consisting of 13 amino acids found in several proteins including some cell adhesion proteins. These proteins are composed of similar repeats. The two ends of each repeat have been denoted as H1 (H, for highly conserved) at the N-terminal and H2 toward the carboxylic end (Kawamoto et al., 1998). The Artemia sequence of 24 amino acids from CP-3L has the highest local homology with the H1 sequences which are given in Table 3 as repeats 1-4. The results in Table 3 also show that the conserved domain in Artemia CP-3L is nearly identical with the H1 domain in the mouse osteoblast-specific factor 2 (mOSF-2), fasciclin I from D.melanogaster (dFAS I), chick transforming growth factor induced protein (cβIG-H3) and the 32 kD hyaline layer component from the sea urchin embryo (HLC-32).

3.3. Isolation of Artemia embryo CP-3 subunits under native conditions

As the partial amino acid sequence of CP-3L did not give enough clues to the function of the large subunit (LSU), other chromatographic procedures were conducted to obtain the subunits in the native form, starting with pure (dimeric) CP-3.

Previously, it was demonstrated that the large subunit of the Artemia cysteine protease has N-linked glycosylation, while the small subunit has O-linked sugar residues (Aiton, 1997). It was thought that the two subunits of the CP might associate through the O-linked sugar, which in some other cases is responsible for subunit binding (Maiz et al., 1995). It was known from previous
Table 3. Alignment of the CNBr-generated fragment from *Artemia* CP-3L with various proteins showing high homology.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Description</th>
<th>Position</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Artemia</em> CP-3L</td>
<td></td>
<td>(2-21)</td>
<td>ANVIDHEGKFTLFAPTNNEAF</td>
</tr>
<tr>
<td>mOSF-2</td>
<td>Repeat 1</td>
<td>(125-144)</td>
<td>RDEIEGKGSYTFAPSNEAW</td>
</tr>
<tr>
<td></td>
<td>Repeat 2</td>
<td>(262-281)</td>
<td>LESLGRDGHTFPTLAPTNNEAF</td>
</tr>
<tr>
<td></td>
<td>Repeat 3</td>
<td>(397-416)</td>
<td>ASSLKDGEYTLAAPVNNAP</td>
</tr>
<tr>
<td></td>
<td>Repeat 4</td>
<td>(527-546)</td>
<td>KDLLTQPGDWTFLAPNTNDAP</td>
</tr>
<tr>
<td>dFAS-I</td>
<td>Repeat 1</td>
<td>(45-64)</td>
<td>ANSTLSLRSTTIFPTNEAF</td>
</tr>
<tr>
<td></td>
<td>Repeat 2</td>
<td>(192-211)</td>
<td>ESVYDAAGQHFLVPDEGF</td>
</tr>
<tr>
<td></td>
<td>Repeat 3</td>
<td>(348-367)</td>
<td>LDDINSLTEVTILAPSNEAW</td>
</tr>
<tr>
<td></td>
<td>Repeat 4</td>
<td>(496-515)</td>
<td>LNN---TQRRTFYPFPRDGW</td>
</tr>
<tr>
<td>cβIG-H3</td>
<td>Repeat 1</td>
<td>(121-140)</td>
<td>RPEIEGPQTFTIFAPSNEAW</td>
</tr>
<tr>
<td></td>
<td>Repeat 2</td>
<td>(258-277)</td>
<td>NSLLESEGGYTLAPTNNEAF</td>
</tr>
<tr>
<td></td>
<td>Repeat 3</td>
<td>(393-412)</td>
<td>SSHLTGESRTILAPVNEVF</td>
</tr>
<tr>
<td></td>
<td>Repeat 4</td>
<td>(520-539)</td>
<td>MENLNRPGTFTVPFNAP</td>
</tr>
<tr>
<td>HLC-32</td>
<td>Repeat 1</td>
<td>(112-131)</td>
<td>VLRDSDLSSVTCMPSDKAV</td>
</tr>
<tr>
<td></td>
<td>Repeat 2</td>
<td>(252-271)</td>
<td>EDEVRNSDPITTLVPTNAF</td>
</tr>
<tr>
<td>Consensus</td>
<td></td>
<td></td>
<td>G TLFA P NEAF</td>
</tr>
</tbody>
</table>

1 The following proteins were compared: mOSF-2, mouse osteoblast-specific factor 2; dFAS-I, *Drosophila* fasciclin I; cβIG-H3, chick beta transforming growth factor induced protein; HLC-32, 32 kD hyaline layer component from the sea urchin embryo.

2 Each protein has multiple repeat units which are shown here for comparison. The consensus sequence is given at the bottom.
work that the majority of the *Artemia* CP binds to Con A and can be eluted with glucose (Aiton, 1997). Because Con A normally binds high mannose structures which mostly are associated with N-linked sugars it was expected that the CP was binding to Con A through the LSU which carries N-glycosylation (Bittiger and Schnebli, 1976).

Experiments were carried out to determine whether the CP-3 subunits could be eluted differentially from a Con A column using a glucose gradient or different concentrations of galactose.

One experiment showed that dissociation of the CP-3 subunits could be achieved under native conditions at pH 5.0. Those observations suggested that each subunit could be purified in their native state. Also an experiment was conducted to elucidate the mechanism of CP subunit dissociation.

### 3.3.1. Elution of *Artemia* CP from a concanavalin A column using glucose

To determine whether the large subunit of the *Artemia* CP could be purified by exploiting differential binding of the subunits to concanavalin A, and the potential dissociation caused by the eluting agent, elution of the column with a glucose gradient was carried out as described in the Materials and Methods. According to the original purification method, the Con A affinity chromatography was carried out with a single step elution (Aiton, 1997). It was not known what would be the subunit composition of the CP in each of the fractions if a gradient elution was tried.
About 2 mg of a partially purified CP protein preparation with 5000 mEU was loaded on a Con A column (0.5 x 3 cm) which had been previously equilibrated with a buffer containing 15 mM K-PO₄, pH 6.8, and 0.3 M KCl. After washing the column with starting buffer to remove unbound protein, elution with a 0-1 M glucose gradient in the same buffer was started and fractions of 1 ml were collected. The protein and CP activity profiles for each column fraction are shown in Fig. 6. The elution profile shows that a considerable amount of protein eluted in the wash with little CP activity, whereas a broad peak of CP activity began to elute at about 0.2 M glucose. The recovery of the activity was about 95%. Also, SDS-PAGE electrophoresis and Western blotting of each column fraction indicated that both subunits are present in the column effluent with the same intensity, suggesting a high affinity of the two subunits for one another (data not shown).

3.3.2. Elution of Artemia CP from a concanavalin A column with galactose

About 2 mg of partially purified Artemia embryo cysteine protease with 5000 mEU was loaded on a Con A column (0.5 x 3 cm) which had previously been equilibrated with a buffer containing 15 mM K-PO₄, pH 6.8, and 0.3 M KCl. After washing the column with the equilibration buffer, the bound proteins were eluted with 1M galactose and then with 1M glucose. The cysteine protease elution profile is given in Fig. 7A.

Each column fraction, was assayed for the CP subunits by electrophoresis on a 7-18% gradient SDS-PAGE and Coomassie blue staining using 20 μl
Fig. 6. Elution of CP from a Con A Sepharose column using a glucose gradient

About 2 mg of partially purified *Artemia* cysteine protease containing 5000 mEU were loaded on a Con A column (0.5 x 3 cm) which had been previously equilibrated with a buffer containing 15 mM K-PO₄, pH 6.8 and 0.3 M KCl. The column was washed with the equilibration buffer to remove unbound protein, then the bound proteins were eluted with a 0-1M glucose gradient in the same buffer. Fractions of about 1 ml were collected and assayed for protein content (○—○) by measuring absorbance at 280 nm, and cysteine protease activity (●—●) using the TNBS method (absorbance at 420 nm). The glucose gradient was started at an elution volume of 5 ml (→), in a total volume of 15 ml.
Fig. 7. *Artemia* cysteine protease elution from a ConA Sepharose column with galactose followed by glucose

A. About 2 mg of a partially purified total cysteine protease preparation containing 5000 mEU of activity was loaded on a Con A column as described in Materials and Methods. The column was washed with the equilibration buffer, then proteins were eluted with 5 ml of 1 M galactose, then with 10 ml of 1M glucose. Fractions of about 1 ml were collected and assayed for cysteine protease activity (A420) using the TNBS method.

B. Aliquots of 20 μl were taken from selected column fractions and electrophoresed on a 7-18% gradient SDS-PAGE. The gel was stained with Coomassie blue and scanned.

Lane 1: Molecular weight standards

Lane 2: Column wash (column volumes 1-5 ml combined, panel A)

Lane 3-8: Galactose elution (elution volumes 6-11 ml)

Lane 9: Glucose elution (elution volumes after 11 ml were combined and concentrated to 1 ml)
aliquots from each column fraction. The results in Fig. 7B show the cysteine protease subunit composition for fractions from the Con A column (Fig. 7A). While some CP does not bind to Con A, and is washed through the column, both subunits appear in all fractions eluted by galactose and glucose, indicating that the subunits have a high affinity for each other.

3.3.3. Gel filtration of the *Artemia* cysteine protease under acid conditions

Since galactose did not dissociate the *Artemia* CP subunits or promote their differential elution from Con A at a neutral pH, I investigated the influence of mild acidic conditions on the dissociation of the CP subunits. A calibrated G-75 Sephadex (SF) column (1 x 55 cm) was equilibrated with a buffer containing 20 mM sodium acetate and 1 mM EDTA at pH 5. A sample of *Artemia* embryo cysteine protease (0.5 ml) after the DEAE Sephadex purification step (see Table 1) containing about 25 mg protein and 5000 mEU was allowed to equilibrate overnight in the same buffer at pH 5 and 4°C. After centrifugation to remove all insoluble material, the acid treated CP was applied to the G-75 Sephadex column. Proteins were eluted from the column with equilibration buffer and the activity profile was determined with the TNBS method. The results in Fig. 8A show that CP activity elutes from the column between 22 and 31 ml with a peak at 28 ml corresponding to a molecular weight of 30 kD based on the calibration of the column with several molecular weight markers (see Appendix 3).
Fig. 8. Gel filtration of *Artemia* embryo cysteine protease under acid and neutral pH

A. A sample of 0.5 ml of post DEAE crude concentrate (in column buffer 1 at pH 6.8) containing about 25 mg protein and 5000 mEU was equilibrated at 4°C overnight in a buffer containing 20 mM sodium acetate and 1 mM EDTA at pH 5. Precipitated proteins were removed by centrifugation. The clear supernatant was loaded onto a calibrated G-75 Sephadex (SF) column which had been previously equilibrated with the same buffer at pH 5. Fractions of about 1 ml were collected and assayed for protein at 280 nm (○—○) and cysteine protease activity (●—●) using the TNBS method. The CP activity eluted at a molecular weight of about 30 kD, based on the previous calibration of the column with known molecular weight standards (see Appendix 3).

B. The same column was loaded with an equal amount of untreated sample in column buffer 1 at pH 6.8 and run as described in the Materials and Methods section. Fractions of about 1 ml were collected and assayed for protein content (○—○) and cysteine protease activity (●—●) using the TNBS method. The CP activity eluted at a molecular weight of about 60 kD (see Appendix 3).
As a control, an untreated sample of CP (25 mg protein and 5000 mEU of activity) was applied to the calibrated G-75 Sephadex column described above and eluted with column buffer 1 at pH 6.8. The protein and activity elution profiles for this chromatographic step are given in Fig. 8B. These results show that the untreated Artemia cysteine protease activity elutes from the column with maximum activity at 18 ml, corresponding to a molecular weight of about 60 kD (see Appendix 3).

Collectively, the results in Fig. 8A and 8B, show that incubation of the cysteine protease in acetate buffer at pH 5, lowers the molecular weight of the protease activity from 60 kD to about 30 kD as would be expected if the CP subunits dissociated under these conditions.

3.3.4. Gel filtration of the Artemia cysteine protease under neutral pH following acid treatment at pH 5

To test whether the the cysteine protease subunits could reassociate once they had been separated by incubation in acidic conditions (at pH 5), a calibrated G-75 Sephadex (SF) column (1 x 55 cm) was equilibrated with column buffer 1 at pH 6.8 and then loaded with 0.5 ml of crude Artemia CP after the DEAE Sephadex purification step containing about 5 mg protein and 1000 mEU which was previously treated with 20 mM acetate buffer at pH 5 for one hour at 4°C. After centrifugation to clarify the acid-treated sample, the pH of the CP preparation was adjusted to 6.8 with 0.5 M K₂HPO₄. Column fractions of about 1 ml were collected and assayed for cysteine protease activity using the TNBS
method and protein content was determined by measuring the absorbance of each fraction at 280 nm. The protein and protease activity profiles are given in Fig. 9. These results show that the acid-treated cysteine protease elutes at a volume of about 28 ml corresponding to a molecular weight of about 30 kD (see Appendix 3) indicating that the CP subunits which dissociate can not readily re-associate by raising the pH to near neutrality.

3.3.5. Gel filtration of the *Artemia* cysteine protease after treatment at pH 5 in the presence of a reversible inhibitor

The observation that dissociation of the *Artemia* cysteine protease subunits at pH 5 appears to be an irreversible process suggested that the mechanism is not only physical but may require autocatalytic activity as well. To test if this hypothesis is correct, the dissociation of the cysteine protease subunits was carried out at pH 5 in the presence of dithiodipyrindine (PDS), a reversible cysteine protease inhibitor. A sample of crude *Artemia* CP from the DEAE Sephadex step containing about 20 mg protein and 4000 mEU was equilibrated overnight in 50 mM sodium acetate at pH 5 containing 10 µM dithiodipyrindine (PDS). After centrifugation to remove any insoluble material, the sample was readjusted to pH 6.8 and 0.5 ml was loaded on a calibrated G-75 Sephadex (SF) column (1x 55 cm) equilibrated with column buffer 1 (pH 6.8). Proteins were eluted with the same buffer and fractions of 1 ml were analyzed for protein content and cysteine protease activity. The cysteine protease assay contained 40µM of DTT in each reaction vessel to reverse the inhibition by
Fig. 9. Gel filtration of the pH 5-treated *Artemia* cysteine protease at neutral pH conditions

A sample of crude cysteine protease preparation after the DEAE Sephadex step containing about 5 mg protein and 1000 mEU was adjusted to contain 20 mM sodium acetate, pH 5, then incubated for 1 hour in an ice bath. After centrifugation to remove insoluble proteins, the pH of the sample was readjusted to 6.8. and after concentration to 0.5 ml, it was applied to a calibrated G-75 Sephadex (SF) column (1 x 55 cm) equilibrated at pH 6.8. Proteins were eluted with the same buffer, and fractions of 1 ml were collected for the assay of protein content (○—○) as absorbance at 280 nm, and cysteine protease activity (●—●) by the TNBS method (absorbance at 420 nm).
dithiodipyridine (PDS). The protein and activity profiles for this experiment are
given in Fig. 10. These results show two peaks of cysteine protease activity, one
eluting at 18 ml, corresponding to a Mr of about 60 kD and the second peak
eluting at 28 ml, corresponding to 30 kD. These results suggest that the
dissociation mechanism is dependent on both pH and cysteine protease activity.

3.4. Purification of *Artemia* cysteine protease subunits on a Mono
S column by fast protein liquid chromatography

3.4.1. Separation of the CP subunits from a crude preparation under native
conditions by FPLC

Attempts to separate the CP subunits were based on the finding which
suggested that the *Artemia* cysteine protease subunits dissociate by incubation at
pH 5 (see 3.3.3), and also on the assumption that the catalytic (small) subunit of
*Artemia* CP should strongly bind to a cationic exchanger at pH 5 as all the other
characterized cathepsin L-like proteases (Mason *et al.*, 1984).

Initially, it was necessary to determine the elution positions of the
subunits from a Mono S column. A cysteine protease sample after the DEAE
Sephadex purification step (about 5 mg protein and 1000 mEU) was loaded on a
Mono S column (0.5 x 5 cm) which had been previously equilibrated with a 20
mM sodium acetate buffer at pH 5. Proteins were eluted using a gradient of 20
mM to 700 mM NaCl in the same buffer using a flowrate of 1 ml/min over 20
minutes. Fractions of 1 ml were collected and assayed for CP activity and protein
content as described in the Materials and Methods. The results in Fig.11A show
Fig. 10. Gel filtration of *Artemia* cysteine protease after treatment at pH 5 in the presence of a reversible cysteine protease inhibitor

A sample of post DEAE Sephadex purified cysteine protease from *Artemia* embryos containing about 20 mg protein and 4000 mEU was adjusted to contain 0.1M sodium acetate, pH 5 containing 10 μM dithiodipyridine. After one hour incubation at 4°C the sample was centrifuged to remove any insoluble material and the pH was readjusted to 6.8. The soluble sample was loaded on a calibrated G-75 Sephadex (SF) column (1x 55 cm) equilibrated with column buffer 1 at pH 6.8. Fractions of 1 ml were collected and assayed for protein (o–o) and cysteine protease activity (••••) in the presence of 40 μM DTT to reverse the inhibition by the dithiodipyridine inhibitor.
Fig. 11. Separation of *Artemia* CP subunits by chromatography on a Mono S column

A. Partially purified cysteine protease after the DEAE Sephadex containing 5 mg protein and 1000 mEU was adjusted to contain 20mM sodium acetate, 1 mM EDTA, pH 5 and was allowed to stay on ice for one hour. Precipitated proteins were removed by centrifugation and the clear soluble fraction was loaded onto a Mono S column (0.5 x 5 cm) equilibrated with 20mM sodium acetate, 1 mM EDTA, pH 5. Proteins were eluted as described in the Material and Methods section. Fractions of about 1 ml were collected and assayed for protein content (---) and cysteine protease activity (●●●). About 80% of the CP activity applied to the column was recovered in the “activity peak”.

B. Aliquots (20 μl) from each fraction were electrophoresed on a 7-18% gradient gel, electroblotted on a nitrocellulose membrane then probed with the purified antibody specific against the large and small subunits of the cysteine protease. The numbers above each lane on top of the NC membrane scan indicate the elution volume from the Mono S column shown in panel A.
that the protein and CP activity elutes at a volume of about 11 ml. Aliquots from each fraction eluting from the Mono S column were electrophoresed on a 7-18% gradient SDS-PAGE and then electroblotted onto a NC membrane. In Fig.11B is shown a scan of the immunostained membrane for each column fraction in Fig. 11A. These results show that the small and the large subunits of the CP could be readily separated by cation exchange chromatography at pH 5, and indicate the elution position from Mono S of each subunit of the *Artemia* embryo cysteine protease. This information was used to isolate the pure native CP subunits starting with a purified preparation in the dimeric form (60 kD).

3.4.2. Separation of the large and small subunits of the *Artemia* CP under native conditions by FPLC

By applying the same conditions as those described in Fig 11A, but starting with a pure CP-3 preparation containing approximately 70μg protein and 1000 mEU, it was possible to isolate individual subunits in their native state as shown in Fig.12. the results in Fig. 12A show the elution profile from the Mono S column which is similar with that shown in Fig. 11. The results in Fig.12B give the results of SDS-PAGE and Western blotting of concentrated column fractions showing the identity of these proteins by immunostaining with antibodies specific for each CP subunit. In Fig.12C is shown a duplicate of the SDS-PAGE gel stained with Coommassie blue to identify the molecular size of the major protein peaks in Fig. 12A. As shown in Fig. 12 (B, C) no other protein bands are visible in the gel. Also the immunostaining with purified antibodies
Fig. 12. Purification of the CP subunits under native conditions

A. A purified sample of cysteine protease (pH 6.8) containing 70 μg protein and 1000 mEU was used to isolate the subunits separately. The sample was chromatographed on a Mono S column (0.5 x 5 cm) as described in the Materials and Methods. Fractions of 1 ml were collected and assayed for CP activity (○—○) using the TNBS method (A 420 nm). The protein profile (—) was monitored by absorbance at 280 nm, and on paper. About 30 μg of each subunit was recovered along with about 100 mEU of activity.

B. Scan of a nitrocellulose membrane probed with the subunit-specific antibodies. Lane 1. purified cysteine protease as a positive control. Lane 2 and 3. column fractions corresponding to elution volumes of 10 and 11 ml respectively in panel A. Lane 4. sample from fraction eluting at a volume of 19 ml in panel A.

C. Scan of a Coomassie blue stained gel from column fractions corresponding to the protein peaks. Lane 1. molecular weight markers, Lane 2. purified cysteine protease showing both subunits as a control, Lane 3. protein eluting at 11 ml. and Lane 4. protein eluting at 19 ml.
against the large and the small subunit demonstrates clearly that the large and small subunits of the *Artemia* CP can be resolved by cation-exchange chromatography on a Mono S column. About 30 μg protein of each subunit were recovered from the Mono S column which represented 86% of the dimeric CP protein loaded onto the column. However, the total cysteine protease activity recovered in the purified SSU fraction was only 10% of the initial starting material, for reasons which are not clear.

### 3.4.3. Effect of the large subunit of the *Artemia* cysteine protease on activity of the catalytic (small) subunit

In order to assess the potential role of the large subunit on the activity of the *Artemia* cysteine protease, the catalytic (small) subunit from CP-3 was assayed for activity either alone or in the presence of the large subunit at three pH values (4.0, 5.0 and 6.0). The results of this experiment are given in Fig. 13. A slight decrease in activity was noted at pH 4 and pH 6 when the large subunit of the CP was added to the reaction vessel. However, as the amounts of purified subunits were limited these experiments could not be repeated at varying concentrations of each subunit. With this limited set of results, it appears that the large subunit has little influence on the activity of the catalytic subunit at the three pH values tested.

However the results shown in Fig. 13 provide only an indication about the kinetics of the CP and more assays would be required to characterize the
Fig. 13. Kinetics of the catalytic (small) subunit of the *Artemia* CP-3 in the presence and absence of the large CP-3 subunit at different pH.

Equal amounts of purified small subunits (5 μg protein, 16 mM) were combined and the CP activity assayed using the TNBS method, compared to reaction vessels containing the catalytic subunit only. The CP activity was measured at pH 4 (panel A), pH 5 (panel B) and pH 6 (panel C). Cysteine protease activity in the presence of the catalytic subunit alone (○—○), and in combination with equal amounts of the large subunit (●—●).
activity of the catalytic subunit and the potential role of the large subunit on the CP activity.

3.5. Characterization of the CP in fractions from *Artemia* embryos

Previous experiments showed that CP activity in the cytosolic fraction of *Artemia* encysted embryos elutes from a gel filtration column (G-150 Sephadex) at a molecular weight of about 90 kD (Warner. 1997). In contrast, CP from the lysosomal fraction of *Artemia* embryos eluted from the gel filtration column as a protein with a molecular weight of about 60 kD (A. Warner, personal communication). In order to determine whether the subunit composition of the CP in these two embryo fractions differ, samples from the aqueous cytoplasm and the lysosomal/mitochondrial fractions were separated on a calibrated G-150 Sephadex (SF) column (1x 80 cm) and analysed for each subunit as described in Materials ans Methods.

3.5.1. Gel filtration of the lysosomal fraction

To analyse the lysosomal CP, a sample of mitochondrial/lysosomal lysate containing 5 mg protein with about 1000 mEU of CP activity was loaded on the same G-150 Sephadex (SF) column used to fractionate the cytosolic fraction (control) and eluted with column buffer 1. The protein and activity elution profiles are given in Fig. 14. It can be seen that unlike cytosol preparations of *Artemia* embryos representing the cytoplasm, where the activity elutes at a
Fig. 14. Gel filtration of a CP sample from lysosomal fraction

A lysosomal CP preparation obtained as described in Materials and Methods containing about 5 mg and 500 mEU of CP activity was loaded on a calibrated G-150 Sephadex (SF) column (1 x 80 cm) which was previously equilibrated with column buffer 1. Fractions of 1 ml were collected as they eluted from the column. Protein concentration (○—○) was determined at 280 nm and the cysteine protease activity (●—●) was assayed with the TNBS method.
molecular weight of about 90 kD (see Fig. 15), the CP activity from the crude L/M fraction elutes at a molecular weight of about 60 kD. Western blotting showed the same staining pattern indicating an almost equal concentration of the subunits in the CP activity peak region (data not shown).

3.5.2. Gel filtration of the aqueous cytoplasmic fraction

When a sample of the aqueous cytoplasm fraction containing 30 mg protein with 3000 mEU of CP activity was loaded onto a G-150 Sephadex (SF) column (1 x 80 cm) which had been calibrated and equilibrated with column buffer 1 and proteins were eluted with the same buffer, the results shown in Fig. 15A were obtained. The activity peak elutes at a molecular weight of about 90 kD (see Appendix 3 for calibration). The protease activity was also assayed in the presence of pepstatin (5 μg/ml) and the fluromethyl ketone inhibitor Z-Phe-Ala-CH₂F (10 μg/ml), to make sure that most of the activity is due to the cysteine protease (data not shown). The elution of the CP activity at a molecular weight of about 90 kD indicates that the catalytic subunit of CP should be complexed with other proteins. The data in Fig. 15B show the results of Western blotting of selected column fractions in Fig. 15A separated by SDS-PAGE. It can be seen from this figure that both subunits of the *Artemia* CP give an almost equal staining in the column fractions corresponding to the CP activity. Traces of the small subunit were found in fractions corresponding to molecular weights greater than 90 kD and these were thought to be due to the association of the small subunit with potential protein complexes.
**Fig. 15. Gel filtration of a CP preparation from aqueous cytoplasm**

A. A CP preparation from the aqueous cytoplasm containing 30 mg protein and about 3000 mEU units was loaded on a G-150 Sephadex (SF) column (1x 80 cm) previously equilibrated with column buffer 1. Fractions of 1 ml were collected and assayed for protein at 280 nm (○○○) and cysteine protease activity (●●●) using the TNBS method.

B. Aliquots of selected fractions from the column in panel A were electrophoresed on a 7-18% gradient SDS-PAGE gel. Proteins were electrotransferred from the gel onto NC membrane and immunostained using the antibodies specific to the both large and small subunits. Lanes from left to right are: CP, a crude CP preparation as a positive control; numbers 20-32 represent fraction elution volumes shown in panel A.
4. DISCUSSION

Several experiments and procedures were conducted during this project to characterize the large subunit of embryonic Artemia cysteine protease in order to:

- obtain partial amino acid sequence of the large subunit of the dimeric cysteine protease (CP) to design primers for PCR and cDNA cloning experiments,
- develop a method to dissociate and purify the CP subunits under native conditions, (i.e., without destroying the activity of the catalytic subunit),
- determine if re-association of the subunits can occur after dissociation at pH 5.0, followed by pH readjustment to neutrality,
- assess whether embryos contain a relative excess of the large subunit protein,
- purify both (native) subunits after dissociation.
- compare the activity of the catalytic subunit alone and in the presence of the large subunit and,
- probe the dissociation mechanism by testing whether a catalytic step is essential for the separation of subunits to occur.

Previous research carried out in this laboratory has characterized the small subunit of the Artemia CP as a cathepsin L-like cysteine protease with O-linked glycosylation (Aiton, 1997). Results from cell fractionation and immuno-histochemical analysis suggest that this protease, unlike the classical cathepsin L which is normally found inside lysosomes, is not found exclusively
in lysosomes but is distributed throughout the embryo (Warner and Shridhar, 1985; Warner et al., 1995). However it should be noted that *Artemia* cysteine protease, unlike many other cathepsin L proteases, is found in an undifferentiated embryonic stage, which is not identical to mammalian tissues where cathepsin L was first described (Bowers, 1998). Also, at each step in the *Artemia* cathepsin L-like cysteine protease purification protocol, the protease has always been found tightly associated with another protein of about 31.5 kD, which has been designated as the large subunit. This protein has not been observed to be associated with cathepsin L in other biological systems. Several studies have shown that this enzyme is very important for the continuity of development in *Artemia* embryos and is the key player in yolk degradation and larval hatching and molting (Warner et al., 1995; Warner and Matheson, 1998).

Yet, although the catalytic subunit of the *Artemia* CP has been quite well characterized biochemically in research carried out in this laboratory, and its cDNA cloned and sequenced, the nature and function of the protein associated with the catalytic subunit is not known.

**4.1. Purification of the *Artemia* CP**

Initially, the *Artemia* embryo cysteine protease was purified to homogeneity using a multi-step protocol involving gel filtration, ion exchange and affinity chromatography according to the method followed by Aiton (1997). This was a modification of a previous method by Warner and Shridhar (1985). A close monitoring of every step in the purification of the main *Artemia* cysteine
protease activity during this project, and results from previous work in our laboratory (A. Warner, personal communication) indicated that the cathepsin L-like protease is tightly associated with a protein of about 31.5 kD, which was designated as the large subunit. The purification protocol used in this project was followed because a recent immuno-histochemical localization study showed that the *Artemia* CP is found mainly in the cytoplasmic and extra-embryonic regions which are thought to have a pH close to neutrality (Warner *et al.*, 1995). Based on that information and on previous differential centrifugation studies using several homogenization buffers (Warner and Shridhar, 1985), it appeared that the majority of the CP is present in the soluble fraction of the embryo.

One study characterized the major embryonic *Artemia* cysteine protease as a cathepsin B-like enzyme, and suggested that the localization of this enzyme is in the lysosomes (Perona *et al.*, 1987). According to that study, the cysteine protease of *Artemia* embryos is present exclusively in the lysosomes. However, the results from our laboratory are not in agreement with these conclusions. It is possible that a fraction of the *Artemia* CP is present in lysosomes without the large subunit (Butler, 1999), but we find the majority of the cysteine protease to be present in the soluble fraction and composed of two subunits.

Transient changes during cyst re-hydration can explain the presence of a dimeric CP in *Artemia* embryos if the analogy with the swelling of mammalian hepatic cells is valid (Waldegger *et al.*, 1997). According to this view, cell swelling leads to alkalization of the lysosomes, and cell shrinkage brings about acidification of the lysosomes. During cysts hydration (before homogenization),
swelling could lead to alkalization and disruption of lysosomes. This may explain both the presence of CP in the cytoplasm and also the fact that almost all the CP is isolated as a dimer (see section 3.3.4) in contrast to what would be the case if the CP were in acidic lysosomes.

Many methods for the purification of cathepsin L from various mammalian tissues have been published and compared (Coetzer et al., 1995). In the majority of these studies, researchers start with liver tissue which is relatively easy to homogenize with mild mechanical force. Among the early steps of protease purification, either the lysosomal fraction is isolated first and purification continues from there, or many proteins are precipitated from the crude homogenate by a low pH treatment, and the supernatant is incubated at pH 4.2 for 4 hours at 37°C to activate the protease (Mason et al., 1984). These conditions are different from the protocol used to purify the embryonic Artemia CP where the pH was maintained at 6.8-7.2 throughout the protocol.

An unexpected finding was the fact that sheep cathepsin L can form proteolytically active complexes with stefin B which have enhanced activity and are stable under physiological conditions (Coetzer et al., 1995). This can explain the ability of secreted cysteine protease to degrade extracellular matrix proteins under neutral conditions. The fact that the optimal stability of the free cathepsin L is in the pH range 4-6 (Mason et al., 1984), whereas cathepsin L associated with stefin B can function at a more neutral pH, is an example of modification of the cathepsin L properties by an associated protein or subunit. The fact that proteolytic activity is retained after the association with an inhibitor indicates
that the binding between the stefin B and cathepsin L in this case is not through the active site. Therefore, cathepsin L may associate with other proteins not only through the active site cleft, but also through other domains in the protein.

Clearly, until the amino acid sequence and the three dimensional structure of the dimeric cysteine protease is known, a conclusive answer to the importance of the cysteine protease-associated protein may not be possible.

4.2. The partial sequence of the large subunit of the *Artemia* CP shows homology with some cell adhesion proteins

The longest fragment of the large subunit of *Artemia* cysteine protease sequenced in the current study is a 22 amino acids region from the N-terminus of a 25 kD peptide generated by CNBr treatment of the CP-3L protein. Local homology search indicated that the peptide has a conserved domain consisting of 13 amino acids found in several proteins with a high degree of similarity including the chick transforming growth factor induced protein (cβIG-H3), mouse bone ossification factors, (mOSF), fasciclin I from *Drosophila*, (FAS I) and the 32 kD protein component of the hyaline layer (HLC-32) of the sea urchin embryo extra-cellular matrix. All of these proteins function as either cell adhesion molecules or as signalling molecules (Lipke, 1996). These proteins are composed of similar repeats of about 150 amino acids each. The two ends of each repeat have been denoted as H1 (H, for highly conserved) at the N-terminal and H2 toward the carboxylic end (Kawamoto et al., 1997). However, the high local homology observed does not necessarily mean that the LSU of *Artemia* CP
(CP-3L) has also high global homology with the above mentioned proteins. Potential function(s) of the *Artemia* CP-associated protein will be discussed later.

### 4.3. Attempts to dissociate the *Artemia* CP subunits in native conditions

While the two subunits could be readily separated under denaturing conditions, several experiments were conducted to find a way to dissociate them under native conditions. The observation that some "free catalytic subunit" (SSU) was eluting from a Con A column during the wash suggested that the eluting agent (glucose) could differentially elute/dissociate the CP subunits. The LSU was found to contain N-like sugars, whereas the SSU had only O-linked glycosylation (Aiton, 1997). Because Con A normally binds high mannose structures which mostly are associated with N-linked sugars, it was thought that the binding of the CP in the Con A column occurred through the N-glycan of the LSU (Bittiger and Schnebli, 1976). Also, it was hypothesized that the two subunits might associate through the O-linked sugar of the SSU which in some other cases is responsible for subunit binding (Maiz *et al.*, 1995). According to this hypothesis the LSU was thought to be a lectin-like protein which could selectively immobilize and target the cathepsin L-like embryonic protease to the extracellular matrix during early development (Lipke, 1996).
4.3.1. Glucose gradient elution of the dimeric CP from a Con A column

According to the method followed for the purification of the CP, the Con A affinity purification was carried out in a single-step elution with 1 M glucose (Aiton, 1997). It was not known what would be the subunit composition of each fraction if a gradient of glucose was used to elute the protease. Unlike what was expected, the results indicated that there was no observable dissociation or differential elution of the subunits in the fractions eluted with a glucose gradient (see Fig. 6). The apparently "free" SSU observed in the wash fractions of Con A column was probably non-glycosylated SSU as suggested by Mason et al. (1984), or it may be the free-form of the protease which was dissociated according to the equilibrium constant at pH 6.8.

4.3.2. Galactose elution of the dimeric CP followed by glucose elution from a Con A column

Galactose elution was tried based on the assumption that the binding between the two CP subunits was mediated by the O-linked sugars of the small subunit. It was expected that elution of the small subunit would occur with a sugar mostly found associated with this type of glycan (i.e., galactose). The results of this experiment indicated that galactose does not dissociate the dimeric protease to yield a protein fraction enriched with the SSU (see Fig. 7). Therefore, neither glucose nor galactose could be used to differentially elute the cysteine protease subunits from the Con A-Sepharose column.
4.3.3. Dissociation of the *Artemia* dimeric CP subunits at pH 5

The fact that many lectins do not bind sugars at low pH, and the use of low pH in the purification of cathepsin L from mammalian sources, suggested that perhaps incubation of the *Artemia* embryonic cysteine protease preparation at a low pH (5.0) could allow separation of the subunits without denaturing them. This approach resulted in a method to dissociate the CP subunits under native conditions. The catalytic activity of the small subunit served as an indicator of the native state after the separation of the subunits. The dissociation of the subunits was found to be pH dependent, and observed after incubation in an acetate buffer at pH 5.0 for at least 15 minutes at 0-4°C. However, the rate of subunit dissociation at pH 5.0 is not known. Initially gel filtration at pH 5.0 indicated a shift in the CP activity from a molecular weight of about 60 kD for the dimeric CP to about 30 kD for the dissociated subunits. The CP assay was conducted using protease inhibitors to ensure that the cysteine protease activity was the result of the CP only. The gel filtration results were confirmed using FPLC on a Mono S column (at pH 5.0) to resolve the subunits of the dimeric CP, which due to similar molecular mass could not be resolved using gel filtration.

4.4. The *Artemia* CP subunits do not reassociate after separation at low pH

Another interesting result is that after dissociation, we were unable to restore the dimeric structure of the enzyme at neutral pH, which may have important developmental significance.
One implication of this result is the suggestion that the *Artemia* CP may not be found within acid vesicles (lysosomes). If the majority of the enzyme were in lysosomes, where the internal pH is expected to be about 5.0, the dimeric protease structure could not be maintained, no matter if the pH is re-adjusted during enzyme isolation. Therefore the isolation of the CP in association with the large subunit would not be possible. This is in agreement with the previous localization studies from our laboratory (Warner et al., 1985 and Warner et al., 1995). Also, pH measurements of the *Artemia* embryos, indicate that the pH of the cytoplasm does not drop much below 6.4 during prolonged anoxia and it may be even higher than 7.9 during development (Crowe et al., 1987).

### 4.5. The potential mechanism of dissociation

The observation that the CP subunits can not readily re-associate after separation by incubation at pH 5, suggested that the dissociation mechanism may not be exclusively a physical process. If the large subunit has lectin-like properties it should re-associate with the small subunit when the pH is readjusted above 6.5. However, in analogy with concanavalin A which needs calcium, manganese and magnesium ions for optimal binding, it may be that after acid dissociation these ions are no longer present at the right concentration to allow re-association of the subunits.

Subunit dissociation is not complete (at 4°C) if the dimeric form of the CP is incubated at pH 5.0 in the presence of dithiodipryridine (PDS), a reversible cysteine protease inhibitor. This observation suggests that proteolytic activity is
required for full dissociation. The mechanism of this pH-dependent catalytic activity could be better understood if the binding site between the subunits was known. Also the rates of dissociation and catalysis are not known, as well as the influence of temperature on the dissociation process. If the binding of the subunits is through the catalytic cleft of the SSU, then dissociation of the two subunits could proceed in analogy with the auto-catalytic activation, or removal of the pro-peptide at low pH. The auto-activation can be intra-molecular or inter-molecular (Khan and James, 1998). In contrast, if the subunits are not binding at/near the active site, and the large subunit is a lectin-like protein which binds the SSU through the O-linked sugars, then it may be supposed that the catalytic event is required to modify the configuration of the dimer in order to facilitate dissociation. This potential structural arrangement could explain the high stability of the dimer CP in analogy with immobilized enzymes.

4.6. Activity of the catalytic subunit of the Artemia CP in the presence or absence of the large subunit

The time course activity of the catalytic subunit alone and in the presence of the large subunit at three pH values (pH 4.0, 5.0 and 6.0) gave results which were not significantly different from one another. However, since there was insufficient amounts of purified subunits available for additional studies, kinetic parameters for the catalytic subunit alone and in the presence of the large subunit could not be determined. With sufficient amounts of purified subunits, a more rigorous approach would determine first the $K_m$ and $V_m$ for the
catalytic subunit alone, and then in the presence of the LSU, using the cathepsin L specific substrate (N-Cbz-Phe-Arg-4 methoxy-β-naphtylamide). Then it would have been possible to determine the $k_{cat}/K_m$ which provides a good measure of the substrate specificity and catalytic activity of the protease. Also, activity time courses obtained at several concentrations of the LSU could shed more light about potential (inhibitory) effects of the LSU on the catalytic activity of the small subunit (SSU) of *Artemia* CP.

4.7. The putative roles of the large subunit of the embryonic *Artemia* cysteine protease

4.7.1. Stabilization of the catalytic subunit

The recovery of the cysteine protease activity after the dissociation of the subunits depended on the protease preparation and its purity. While a high recovery of cysteine protease activity was observed (80%) after the separation of the subunit fractions of a crude CP preparation by FPLC (see Fig. 11), a much lower yield of activity was obtained (10%) after the FPLC separation of the CP on a Mono S column starting with the purified dimeric protease (see Fig. 12). This observation suggests that the stability of the catalytic subunit is improved by higher concentrations of other proteins. Therefore, one of the functions of the LSU may be to stabilize the cathepsin L-like protease (SSU) activity. Since the large subunit protein of the embryonic *Artemia* cysteine protease appears to be unique and not found associated with the CP in other systems, it may be supposed that its function permits the embryo to adapt to dormancy, anoxia or
anhydrobiosis, or any other harsh condition encountered during the life cycle of
*Artemia*. The LSU may stabilize the catalytic subunit during early development
when the cytosolic pH can reach 7.9 and then after acidification of the cytoplasm
to 6.3 (during anoxia). In the meantime the protease must be available for yolk
degradation as yolk proteins become mobilized and released into the cytoplasm
where proteolytic degradation takes place (Utterback and Hand, 1987; Warner
*et al*., 1997).

The classical (mammalian) cathepsin L has good stability in the pH range
4-6, but it is unstable at pH values higher than 6.5 (Mason *et al*., 1984).
Therefore, the LSU may modify the pH stability of the catalytic subunit when
they are associated, and this property may be lost after subunit dissociation.
However, if the SSU *in vivo*, has catalytic activity in the form associated with the
LSU at a higher pH, this property could extend the role of the CP during
development in *Artemia*.

4.7.2. Protecting the catalytic subunit of the *Artemia* CP from high
temperatures

Another possible role of the LSU is to protect the cathepsin L-like
subunit from high temperatures. Due to limited amounts of the purified subunits
it was not possible to compare the stability of the small subunit at different
temperatures. However, gel filtration of crude homogenates showed that the
*Artemia* CP eluted from the column as a complex with a molecular weight of
about 90 kD. Under natural conditions inside the cyst, perhaps the protease is
present in complexes with more than one protein (Warner et al., 1997). On the
other hand, the "protecting role" of the LSU may come into effect after removal
of the pro-peptide (activation). If this is the case, the pro CP should be found
alone in the embryo, and the association with the large subunit occurring
during/after the removal of the pro-peptide. However it is not known when and
where activation of the pro CP occurs in *Artemia* embryos.

4.7.3. Targeting of the cysteine protease to the extracellular matrix

A lot of research has been conducted to understand the mechanisms of
protein sorting to various sites in the cell. the role of sugar modifications of
proteins and the steps of organogenesis in many cell types. The results of these
studies have revealed that the mechanisms involved with the above processes are
highly diversified. On the other hand, it is not known what are the precise protein
targeting mechanisms and the processes leading to organogenesis in
undifferentiated cells during early development. The dormant embryo of *Artemia*
consists of about 4000 cell nuclei in a common cytoplasm where cell separation
has not yet been completed (Clegg and Conte, 1980). In some cell types such as
the haemopoietic cells, the whole lysosome may be secreted to the extracellular
matrix using the regulated secretory pathway (Griffiths, 1996). However,
whether this process occurs inside the common cytoplasm of *Artemia* dormant
embryos is not known.

With the limitation by the above consideration, it can be conceived that
another role of the LSU could be targeting of the catalytic subunit to the
extracellular matrix. The catalytic subunit has only O-linked glycosylation, and as such would not be secreted out of the cell, whereas the LSU which has N-linked glycosylation can in principle be secreted out of the cell. Therefore, the attachment of the SSU to the LSU would lead to the co-secretion of the dimeric protease into the extracellular matrix (Moore, 1999). According to this interpretation, it is assumed that removal of the propeptide happens in some acidic vesicle where both subunits would be present. The dimer with the secretion signal (N-glycosylation of the LSU) would then be transferred out of the cell via the regulated secretory pathway. However, this model would be valid only if the *Artemia* dormant embryo is similar with the mammalian cells which serve as models for protein secretion.

Another observation suggests that the LSU of the *Artemia* CP is found on the outer surface of the yolk platelets (A. Warner, personal communication. Warner et al., 1995). According to this observation there is some LSU protein observed in purified yolk platelets, while there is no trace of the SSU, suggesting that either the LSU is present on the outer surface of the yolk granule or that it can be present inside the yolk (Warner et al., 1995). If this hypothesis is valid, assembling of the CP subunits could occur on the outer surface of the yolk platelets. The homologous protein of the LSU, the HLC-32 from the sea urchin is found inside yolk platelets before fertilization, but in the extracellular matrix in the developing embryo (Mayne and Robinson, 1998). By analogy, the LSU of the *Artemia* CP could be stored in the yolk and secreted toward the outer yolk surface after fertilization. This is consistent with the view that *Artemia* yolk
proteins first become mobilized and then released from the yolk granule (Utterback and Hand. 1987), where they (potentially) become degraded by the cysteine protease (Warner et al., 1997).

From all the available data, it may be hypothesized that there are several populations of the CP in the Artemia encysted embryo. One (small) population is present in lysosomes while the majority of the cysteine protease is in the extracellular matrix and in the cytoplasm of embryos.

The high degree of local homology of the LSU with some cell adhesion molecules (i.e., HCL-32 from the sea urchin) which are known to be localized in the extra-cellular matrix, suggests that the large subunit of the cysteine protease plays a significant role in the control and/or function of the catalytic subunit at specific times in development when morphological differentiation requires considerable extracellular protein degradation. This view is consistent with the immuno-histochemical localization study which detected the presence of the cysteine protease in various regions of the extracellular matrix in newly hatched larvae of Artemia (Warner et al., 1995).
Appendix 1

MS profile of the Lys C-generated peptides from the *Artemia* CP-3L

The SDS-PAGE gel slice containing the 31kD subunit of CP was treated with protease Lys C. Peptides were eluted from the gel and collected. About 10% of the resulting peptides were submitted for mass spectrometric analysis. Below is the scan of the MS profile generated from this analysis.
Appendix 2

HPLC profile of the Lys C-generated peptides from the *Artemia* CP-3L

The SDS-PAGE gel slice containing the 31kD subunit of CP-3 as treated with protease Lys C. Peptides were eluted from the gel and collected. About 90% of the peptides were chromatographed by RP-HPLC. The figure shows a scan of the HPLC profile obtained at the protein sequencing laboratory of the University of Laval. The handwritten numbers correspond to the peptides collected and sequenced. The same numbers are used to identify the peptide sequences in Table 2.
Appendix 3
Results of gel filtration column calibrations

1. G-75 Sephadex (SF) column (1 x 55 cm):

<table>
<thead>
<tr>
<th>PROTEIN, Molecular weight</th>
<th>ELUTION VOLUME (ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue Dextran, 2000 kD</td>
<td>15</td>
</tr>
<tr>
<td>BSA, 67 kD</td>
<td>18</td>
</tr>
<tr>
<td>Ovalbumin, 45 kD</td>
<td>23</td>
</tr>
<tr>
<td>Cytochrome C, 13 kD</td>
<td>43</td>
</tr>
<tr>
<td>Phenol Red, 0 kD</td>
<td>46</td>
</tr>
</tbody>
</table>

2. G-150 Sephadex (SF) column (1 x 80 cm):

<table>
<thead>
<tr>
<th>PROTEIN</th>
<th>ELUTION VOLUME (ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue Dextran, 2000 kDa</td>
<td>24</td>
</tr>
<tr>
<td>Gamma-Globulin, 158 kDa</td>
<td>25</td>
</tr>
<tr>
<td>BSA, 67 kDa</td>
<td>32</td>
</tr>
<tr>
<td>Ovalbumin, 45 kDa</td>
<td>39</td>
</tr>
<tr>
<td>Myoglobin, 17 kDa</td>
<td>48</td>
</tr>
<tr>
<td>Cytochrome C, 13 kDa</td>
<td>50</td>
</tr>
<tr>
<td>Phenol Red, 0 kDa</td>
<td>65</td>
</tr>
</tbody>
</table>
REFERENCES


Everts, V., 1998. Gelatinase A (MMP-2) and cysteine proteinases are essential for the degradation of collagen in soft connective tissue. Matrix Biology, 17: 35-46.


VITA AUCTORIS

NAME: Ervin Pullumbi

BORN: December 19, 1964, Tirana, Albania.

EDUCATION: 1997-1999
Master of Science Program,
Department of Biological Sciences,
University of Windsor,
Windsor, Ontario, Canada.

1982-1987
Diploma “Biologist” at
Tirana University, Faculty of Natural Sciences
Tirana, Albania.

EMPLOYMENT: 1988-1996
Research Assistant,
Department of Microbiology,
Food Institute,
Tirana, Albania.

PUBLICATIONS:

ABSTRACTS: 1) Characterization of a novel heterodimeric cathepsin L-like protease from the brine shrimp *Artemia franciscana*.

2) Cloning and Characterization of a Novel Cathepsin L-like Cysteine Protease from the Brine Shrimp, *Artemia*.