The purification and characterization of an extrapallial fluid protein from the mollusc, Mytilus edulis

Stephen Joseph Hattan

University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/dissertation

Recommended Citation
https://scholars.unh.edu/dissertation/1890
The purification and characterization of an extrapallial fluid protein from the mollusc, Mytilus edulis

Abstract

Shell deposition within the bivalve mollusc Mytilus edulis is regulated by the macromolecules of the extrapallial fluid (EP fluid). These mineralization regulating macromolecules are thought to be responsible for the nucleation, growth regulation and growth cessation of the CaCO$_3$ mineral crystals that will inevitably constitute $\geq$95% of the mature shell.

This dissertation presents the results of the purification and characterization of the major EP fluid protein of the bivalve Mytilus edulis. The major EP fluid protein was determined to comprise 56% of the total fluid protein and to consist of 7.25% by weight carbohydrate. The protein was purified through a series of chromatographic procedures and, once pure, was shown to have a molecular weight of 28,300 amu and an isoelectric point of 4.85.

N-Terminal amino acid sequence analysis showed the primary protein structure to be, Asn-Pro-Asp-Asp-X-X-Asp-Asp-Ala-Pro-Ile-Val-Glu-X-X-Asp- (where X = an unknown amino acid residue). This amino acid sequence shows a repeat sequence similar to the theorized - (Asp-X) - (X = Gly) shell organic matrix repeat sequence proposed to act as a template for CaCO$_3$ crystal growth. Additionally, the EP fluid protein shows similarities with mollusc shell soluble organic matrix proteins in its amino acid composition with glutamic and aspartic acid residues being the two most abundant amino acid residues.

Lastly, the EP fluid protein was shown to bind calcium and in so doing self-assemble into higher order protomers. Protomers composed of as many as 32 subunits were identified. The ability to self-assembly is one of the most intriguing properties ascribed to organic matrix of molluscan shells. The results of this dissertation suggest that the major extrapallial fluid protein may be a building block of the organic matrix of the shell. The proposed model of the EP fluid protein protomer formation may give insight into how matrix assemble takes place.

Keywords

Chemistry, Analytical, Chemistry, Biochemistry, Biology, Animal Physiology

This dissertation is available at University of New Hampshire Scholars' Repository: https://scholars.unh.edu/dissertation/1890
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. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

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.
THE PURIFICATION AND CHARACTERIZATION OF AN EXTRAPALLIAL FLUID PROTEIN FROM THE MOLLUSC, MYTILUS EDULIS

BY

STEPHEN JOSEPH HATTAN
BA., Saint Anselm College, 1988

DISSERTATION

Submitted to the University of New Hampshire in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in

Chemistry

May, 1996
This dissertation has been examined and approved.

N. Dennis Chasteen
Dissertation Director, N. D. Chasteen
Professor of Chemistry

Sterling A. Tomellini,
Associate Professor of Chemistry

W. Rudolf Seitz,
Professor of Chemistry

Edward H. Wong,
Professor of Chemistry

Thomas M. Laue
Professor of Biochemistry

26 January, 1986
Date
DEDICATION

This dissertation is dedicated to my parents, Mary E. Hattan and Dr. Richard F. Hattan, for all of their love and support throughout my life.
ACKNOWLEDGEMENTS

I offer sincere thanks to my dissertation advisor, Professor N. D. Chasteen for his support and guidance throughout my graduate study. Dr. Chasteen's work ethic and professionalism are qualities which I will not forget and always admire. I thank Dr. Thomas Laue for his assistance with the ultracentrifuge experiments. I thank Dr. Paul Matsudaira for his assistance with the protein sequence experiments. I thank Dr. Robert Trimble for his help with the protein carbohydrate analysis and I thank Dr. John Osterhout for his assistance with the circular dichroism experiments.

I thank the friends that I have made in the UNH community with special thanks to John Grady for his valuable help and support.

I would also like to thank my parents, all of my brothers and sisters and my loving wife Kathleen for all of their love and encouragement throughout the years.
TABLE OF CONTENTS

DEDICATION .................................................. iii
ACKNOWLEDGEMENTS ................................. iv
LIST OF FIGURES ........................................ vii
LIST OF TABLES ........................................... x
ABSTRACT .................................................. xi

CHAPTER

I. Introduction .............................................. 1
   Mytilus Edulis and Its Shell ......................... 3
   The Organic Matrix ..................................... 6
   Soluble Organic Matrix ............................... 7
   Insoluble Organic Matrix ............................ 11
   Mechanisms of Mineral Formation ................. 12
   Mineral Formation by Epitaxy ..................... 13
   Mineral Formation by Ionotropy ................... 15
   The Extrapallial Fluid ............................... 18

II. Materials and Methods ............................... 21
   Mussel Source .......................................... 21
   Protein Buffer Solution ............................. 21
   Fluid Extraction ........................................ 22
   Centrifugation ......................................... 22
   Sample Storage ........................................ 23
   Protein Salt Precipitation ......................... 23
   Cation Exchange Chromatography ................. 25
   Anion Exchange Chromatography ................ 26
   Size Exclusion Chromatography .................. 27
   Electrophoresis ....................................... 28
   Protein Concentration Assay ....................... 29

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral Sugar Assay</td>
<td>29</td>
</tr>
<tr>
<td>Ultrafiltration</td>
<td>30</td>
</tr>
<tr>
<td>Protein Lyophilization</td>
<td>30</td>
</tr>
<tr>
<td>Mass Spectrometry</td>
<td>31</td>
</tr>
<tr>
<td>Amino Acid Analysis</td>
<td>31</td>
</tr>
<tr>
<td>N-Terminal Amino Acid Sequence</td>
<td>32</td>
</tr>
<tr>
<td>Ultracentrifugation</td>
<td>33</td>
</tr>
<tr>
<td>Infrared Spectroscopy</td>
<td>34</td>
</tr>
<tr>
<td>EDX Analysis</td>
<td>34</td>
</tr>
<tr>
<td>$^{31}$P Nuclear Magnetic Resonance</td>
<td>35</td>
</tr>
<tr>
<td>Protein Total Phosphate Assay</td>
<td>35</td>
</tr>
<tr>
<td>Sialic Acid Assay</td>
<td>36</td>
</tr>
<tr>
<td>Electron Paramagnetic Resonance</td>
<td>36</td>
</tr>
<tr>
<td>Atomic Absorption</td>
<td>37</td>
</tr>
<tr>
<td>Circular Dichroism</td>
<td>38</td>
</tr>
</tbody>
</table>

### III. Results
- Raw EP Fluid Protein and Carbohydrate Analysis...39
- Extrapallial Fluid Protein Distribution....44
- Protein Purification...47
- Molecular Weight Determination...53
- Isoelectric Focusing...58
- Determination of Protein Molar Absorptivity...60
- Amino Acid Analysis...64
- Identification of the Protein of a Glycoprotein...68
- $\text{N- vs. -O- Linked Carbohydrate Side Chains}$...71
- N-Terminal Amino Acid Sequence...77
- Protein Interaction with Calcium...79
- Sedimentation Velocity in the Presence of Calcium...83
- Circular Dichroism...91
- Investigation into Possible Calcium Binding Moieties...95
- Sulphate Determination by EDX Analysis...96
- Phosphate Analysis by Phosphate Assay and $^{31}$P NMR...99
- Sialic Acid Determination...101
- Copper Determination...102

### IV Discussion...105
LIST OF FIGURES

Figure:

1.1: Diagram of the shell of *Mytilus Edulis* ................................................. 5
1.2: Electron micrograph of the nacreous shell layer (x 17,500) ................. 9
1.3: Illustration of the cross section of the organic matrix ....................... 12
1.4: Illustration of calcium binding by the -[Asp-X]- amino acid sequence ...................................................................................................... 14
1.5: Illustration of calcium binding as it occurs in the ionotropic theory of shell formation ................................................................. 16
3.1: Plot of EP fluid [protein] and [carbohydrate] vs. month of extraction ......................... 40
3.2: 15% native PAGE gel of raw EP fluid samples ................................. 44
3.3: 15% SDS-PAGE gel of raw EP fluid samples ................................. 45
3.4: 15% native PAGE gel of raw fluid taken from individual animals . 45
3.5: Densitometric scan of a native PAGE gel ...................................... 46
3.6: Sephadex CM-50 cation exchange chromatogram ............................. 50
3.7: Hi-Trap Q anion exchange chromatogram ...................................... 50
3.8: Sepharose 6B-CL size exclusion chromatogram ............................... 51
3.9: 10% native PAGE gel showing the major EP fluid protein at different stages of purity ........................................................................ 51
3.10: 15% SDS-PAGE gel used to estimate EP fluid protein molecular weight ...................................................................................... 56

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
3.11: Plot of Response Factor vs. log Molecular Weight for SDS-PAGE protein standards .......................................................... 57
3.12: Time of flight mass spectrum of purified EP fluid protein ............ 57
3.13: Plot of Response Factor vs. Isoelectric point for IEF protein standards .......................................................................................... 61
3.15: UV-Vis absorption spectrum of purified EP fluid protein .............. 63
3.16: Plot of Absorbance vs. Concentration for several EP fluid protein samples .......................................................................................... 63
3.17: UV-Vis absorption spectrum of purified EP fluid protein .............. 64
3.17a: Pie chart displaying the EP fluid protein amino acid breakdown ... 65
3.18: 15% native gel identifying EP fluid protein as a glycoprotein........ 70
3.19: Composite plot of carbohydrate assay calibration curves ............. 72
3.20: Composite plot of protein assay calibration curves ....................... 72
3.21: Absorption scan of EP fluid after digestion in NaOH.................... 75
3.22: Difference spectrum of NaOH digest - H2O Control .................... 75
3.23: 15% native gel of ENDO H treated/untreated EP fluid protein .... 76
3.24: 10% native PAGE gels of EP fluid protein in the presence and absence of calcium................................................................. 82
3.25: Electrophoresis of EP fluid on a large stacking gel brought to 10 mM Ca2+ ......................................................................................... 82
3.26: Ultracentrifugation sedimentation velocity plot g(s) vs. s* for an 0.9 mg/ml protein sample with varying amounts of Ca2+ .......... 84
3.27: Multiple Gaussian fit to sedimentation velocity data for a 0.9 mg/ml protein sample analyzed in 10 mM Ca2+ ................................. 84
3.28: Ultracentrifugation sedimentation velocity plot g(s) vs. s* for four EP fluid protein samples run in the absence of Ca\textsuperscript{2+} ................. 86

3.29: Ultracentrifugation sedimentation velocity plot g(s) vs. s* of EP fluid protein samples in the presence and absence of Ca\textsuperscript{2+} .................. 89

3.30: Ultracentrifugation sedimentation velocity plot g(s) vs. s* displaying an increase in peak width at half height for the higher order protomers ................................................. 89

3.31: CD spectra of the EP fluid protein in the presence and absence of Ca\textsuperscript{2+} ................................................................. 94

3.32: CD spectra of peptides of known secondary structure ............... 94

3.33: FTIR spectrum of the EP fluid protein in the SO\textsubscript{4}\textsuperscript{2-}, PO\textsubscript{4}\textsuperscript{3-} absorption region .............................................. 97

3.34: EDX analysis of a lyophilized EP fluid protein sample .................. 98

3.35: 31P NMR spectrum of an 8 mg/ml EP fluid protein sample .......... 100

3.36: Composite plot of sialic acid standards used in sialic acid assay . 101

3.37: EPR spectrum of a 1 mM frozen EP fluid protein solution .......... 104

3.38: Atomic absorption Cu standard addition calibration curve used in the quantitation of Cu/protein stoichiometry .......................... 104

4.1: 5-25% SDS-PAGE gel of SOM protein from 9 different molluscs . 109

4.2: Cartoon of protomer assembly ..................................................... 115

4.3: Ultracentrifugation sedimentation velocity plot g(s) vs. s* of a 0.9 mg/ml EP fluid protein sample in 10 mM Ca\textsuperscript{2+} buffer ..................... 119

4.4: Ultracentrifugation sedimentation velocity plot g(s) vs. s* of a 0.9 mg/ml EP fluid protein sample in 10 mM Ca\textsuperscript{2+} buffer incubated for 24 hours .......................................................... 119

4.5: Cartoon of protomer assembly ..................................................... 124
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Extrapallial fluid protein concentration by month</td>
<td>41</td>
</tr>
<tr>
<td>3.2</td>
<td>Extrapallial fluid carbohydrate concentration by month</td>
<td>41</td>
</tr>
<tr>
<td>3.3</td>
<td>Extrapallial fluid protein concentration by individual animal</td>
<td>42</td>
</tr>
<tr>
<td>3.4</td>
<td>Data of the efficiency of the protein purification procedure</td>
<td>52</td>
</tr>
<tr>
<td>3.5</td>
<td>Molecular weight data for SDS-PAGE standards</td>
<td>56</td>
</tr>
<tr>
<td>3.6</td>
<td>EP fluid protein amino acid analysis data</td>
<td>67</td>
</tr>
<tr>
<td>3.7</td>
<td>Comparison of EP fluid amino acid analysis data against that of SOM proteins</td>
<td>68</td>
</tr>
<tr>
<td>3.8</td>
<td>Data of the carbohydrate/protein weight percent</td>
<td>70</td>
</tr>
</tbody>
</table>
Shell deposition within the bivalve mollusc *Mytilus edulis* is regulated by the macromolecules of the extrapallial fluid (EP fluid). These mineralization regulating macromolecules are thought to be responsible for the nucleation, growth regulation and growth cessation of the CaCO₃ mineral crystals that will inevitably constitute ≥ 95% of the mature shell.

This dissertation presents the results of the purification and characterization of the major EP fluid protein of the bivalve *Mytilus edulis*. The major EP fluid protein was determined to comprise 56% of the total fluid protein and to consist of 7.25% by weight carbohydrate. The protein was purified through a series of chromatographic procedures and, once pure, was shown to have a molecular weight of 28,300 amu and an isoelectric point of 4.85.

N-Terminal amino acid sequence analysis showed the primary protein structure to be, Asn-Pro-Asp-Asp-X-X-Asp-Asp-X-X-Asp-Ala-Pro-Ile-Val-Glu-
X-X-Asp- (where X = an unknown amino acid residue). This amino acid sequence shows a repeat sequence similar to the theorized -[Asp-X]- (X = Gly) shell organic matrix repeat sequence proposed to act as a template for CaCO$_3$ crystal growth. Additionally, the EP fluid protein shows similarities with mollusc shell soluble organic matrix proteins in its amino acid composition with glutamic and aspartic acid residues being the two most abundant amino acid residues.

Lastly, the EP fluid protein was shown to bind calcium and in so doing self-assemble into higher order protomers. Protomers composed of as many as 32 subunits were identified. The ability to self-assembly is one of the most intriguing properties ascribed to organic matrix of molluscan shells. The results of this dissertation suggest that the major extrapallial fluid protein may be a building block of the organic matrix of the shell. The proposed model of the EP fluid protein protomer formation may give insight into how matrix assemble takes place.
CHAPTER I

Introduction

The term biomineralization refers to the widespread phenomenon of biological regulation of inorganic mineral deposition. All of life's kingdoms contain species which biomineralize and these species are spread through 55 of the classifying phyla. To date, science has recorded some 60 different types of biogenic minerals and it is generally agreed that numerous other types remain as yet undiscovered. In nature, biological mineral production takes place on such a massive scale that it has a substantial influence on the chemistry of the ocean waters, the composition of the ocean sediments and the composition of sedimentary rock. Reef building species generate mineral deposits on such a scale as to permanently alter the Earth's geography. Just as impressive as the numbers and types of minerals and mineralizing species is the diversity of biological functions which mineral deposits are responsible for. The thin shell of the Mollusc Nautilus allows it to withstand the tremendous pressure of the ocean depths, while at the same time, the hammer, anvil, and stirrup bones within
the middle ear enable us to hear. The degree of control which biology has over mineralization events also shows a high degree of variability. While some events appear to be no more than biologically-induced amorphous mineral precipitation, others appear regulated in virtually every detail.\(^8,9\)

Despite many differences, there are certain aspects which biomineralization processes have in common.\(^10-13\) Generally, organisms regulate mineralization via the use of organic macromolecules\(^14-20\) (proteins, polysaccharides). In numerous controlled-mineralization events these macromolecules appear responsible for the nucleation, growth regulation, and growth cessation of the mineral crystals involved. The environment from which mineralization takes place is almost always unique and isolated from its surrounding. The content of the mineralization "Mother Liquor" is biologically regulated and supersaturated with the ions of the mineral being deposited.

This doctoral thesis is an investigation into the biomineralization process within molluscs. The specific research topic is the extrapallial fluid of the bivalve mollusc, *Mytilus Edulis*. The extrapallial fluid is the medium which fills a cavity located between the animal viscera and the external shell. The location and contents of the extrapallial fluid has led investigators to believe that it plays a vital role in the dynamic life of the shell.\(^21-24\) However, despite being highly significance in regards to shell mineralization, the EP fluid has received little detailed study.
Mytilus Edulis and Its Shell:

*Mytilus Edulis* is the common blue mussel which inhabits the intertidal zones of coastal New England. *Mytilus Edulis* is a molluscan species of the class Bivalvia. This class of mollusks is named for the bilateral symmetry of the external shell halves; a distinguishing feature possessed by all class members. *Mytilus Edulis* is a sedentary, ciliary feeder which typically lives its life in a colony, firmly attached to the rocks of shallow coastal waters. Our interest in *Mytilus Edulis* concerns its ability to biomineralize. Within this animal, the ability to biomineralize manifests itself in shell formation. Therefore, our investigation will begin with a brief description of the elements involved in shell formation as well with a description of the shell itself.

The organ most responsible for orchestrating shell formation is the mantle. The mantle is thin membrane consisting of two epithelial layers. It is the animal’s outer-most visceral organ and acts as a barrier between the other soft tissues and the external shell. The mantle is attached to the shell along the entire shell margin at a point called the pallial line. The result of this attachment is the creation of an effective cavity between the mantle and the inner shell surface. This cavity, termed the extrapallial cavity, is filled by the extrapallial fluid. The biological macromolecules involved in shell mineralization are synthesized by cells within the mantle and then secreted into the extrapallial fluid. It is thought that the medium of the
extrapallial fluid is the environment from which biological regulation of shell formation is carried out.\textsuperscript{31-33}

As a whole, the shells of mollusc's are 95 to 99.9\% by weight CaCO\textsubscript{3} with the residual mass being composed of biological macromolecules.\textsuperscript{34-35} Although the shell is made up largely of CaCO\textsubscript{3}, mineral distribution within the shell is not homogeneous. The shell of \textit{Mytilus Edulis} is composed of three spatially distinct layers: the periostracum, the prismatic region and the nacre (Figure 1.1). These layers are all continuous, ultrastructurally unique, and reside one on top of the other along the long axis of the shell. The inner two layers -- the prismatic region and the nacre -- are calcified. The outer periostracum layer is not.

The periostracum is a thin layer of sclerotized protein\textsuperscript{36} which act as a sheath, covering the shell exterior. Its function is thought to be largely one of protection,\textsuperscript{37} guarding the underlying mineral layers against erosion from environmental forces. Additionally, the periostracum is thought to be the substrate upon which initial shell mineralization takes place.\textsuperscript{38}

As mentioned above, the prismatic and nacre regions of the shell are both mineralized. What makes these layers unique, is that the mineral morphology which constitutes each is different. The prismatic region is composed of CaCO\textsubscript{3} deposited as calcite crystals,\textsuperscript{38} while the nacre region has CaCO\textsubscript{3} deposited as aragonite crystals.\textsuperscript{40} The dividing line between these layers is sharp and these two mineralized shell fractions, both by
shells themselves and together, serve to illustrate the intricacy of the biomineralization process.

Most investigations into the processes of shell formation have focussed on the structure and composition of the shell. These investigations have yielded a good foundation of knowledge regarding shell ultrastructure as well as a general understanding of the forces which control mineral deposition. Results from some of this work are discussed below.
The Organic Matrix:

Molluscan shell deposition, along with several other biomineralization processes such as bone and tooth formation and egg shell formation is thought to be an organic matrix-mediated process. This means that the fundamental events surrounding mineral deposition, growth, and cessation are controlled by a network of biological macromolecules. This organic framework, called the organic matrix, is composed of protein, carbohydrate and glycoprotein. These are the macromolecules which are synthesized in the mantle and then secreted into the extrapallial fluid. Once secreted, these molecules are thought to self assemble into a matrix prior to mineral deposition. Once formed, the organic matrix is thought to be the substrate upon which mineral growth takes place. The matrix is believed to contain specific sites at which mineral crystal nucleation takes place. Atom spacing within the mineral phase, which dictates crystal morphology, as well as the direction of the crystal growth is mediated by the organic matrix. Additionally, cessation of crystal growth, which ultimately dictates crystal size and shape, is thought to be organic matrix-mediated. The organic matrix is credited for doing a lot of things. What is known about it?

Most studies of organic matrix material begins with shell demineralization. This is typically carried out by dialyzing the shell, or fragments thereof, in a solution saturated with EDTA. The mineral portion of the shell dissolves away, leaving the matrix macromolecules behind in
Regardless of species, what is usually found is that the organic matrix is constructed of two primary components. The fundamental distinction between the two fractions is that one component is water soluble and the other is water insoluble. This being the case, study of organic matrix material has been roughly categorized as pertaining to the Soluble Organic Matrix (SOM) or the Insoluble Organic Matrix (IOM). Both fractions have been the recipient of a fair amount of study and as such, an understanding of the functional roles that each play in mineral deposition has emerged.

**Soluble Organic Matrix:**

Soluble Organic Matrix is composed of protein and glycoprotein. The most noted characteristic of SOM protein, collectively, is that they are highly acidic. Amino acid studies on SOM proteins from certain species have reported molar ratios of as high as 40% aspartic acid. Despite reports of such a large quantity of a single amino acid, in general, the composition of SOM proteins is considered to be complex. It is not uncommon for SOM proteins to be reported as containing varying amounts of covalently-bound carbohydrate moieties (glycoprotein). Some of these sugar residues are reported to be sulfated. SOM protein may also contain phosphorylated serine and threonine amino acid residues. Much study has focused on their ability to bind calcium.
Present theory regarding mineral deposition attributes the acidic soluble matrix proteins with a diverse array of functions. Firstly, it is believed that these proteins contain specific nucleation sites responsible for controlling the timing and location of crystal formation. This hypothesis is based, in part, on electron microscopic images of mineralizing shells which show a high degree of precision in regard to sites at which new crystal form.\textsuperscript{58-60} Additionally, \textit{in vitro} experiments have been performed which show SOM material to enhance CaCO\textsubscript{3} crystal growth from a CaCO\textsubscript{3} supersaturated solution.\textsuperscript{61}

Responsibility for the regulation of growing crystals --post nucleation-- as well as cessation of crystal growth are also attributed to the SOM proteins. The extrapallial fluid is supersaturated with respect to CaCO\textsubscript{3}.\textsuperscript{62} Unrestricted crystal growth within such a medium ought to lead to crystals of assorted sizes, shapes and orientations. For \textit{Mytilus Edulis}, and most likely for all other shell forming species, this is not the case. The biological protocol for shell formation has a built-in mechanism(s) for the regulation of crystal morphology. Experimental evidence and observation, points to the acidic macromolecules of the soluble organic matrix as being responsible for fulfillment of this function.\textsuperscript{63,64} The microscopic image of a mineralized section of shell may be likened to a brick wall (Figure 1.2). The matrix material is the mortar, surrounding and dividing the bricks while simultaneously holding them together, and the individual CaCO\textsubscript{3} crystals are
representative of the bricks themselves. Such images suggest that the organic matrix assembles in such a way as to create numerous cavities. Crystal growth occurs within these cavities and it is the cavity bounds which dictate the ultimate size and shape of the mature crystal. Alternate mechanisms for growth regulation have also been suggested. In contrast to the in vitro experiments mentioned in the discussion of crystal nucleation, some in vitro experiments involving the addition of soluble matrix material to supersaturated CaCO₃ solution have
shown the matrix to inhibit the rate of CaCO₃ crystal formation.¹⁰⁻¹²
Inhibition of crystal growth is thought to occur via the formation of protein/crystal complexes which, once formed, inhibit further mineral deposition at the site of complexation. *In vivo*, this regulatory mechanism is thought to act as follows. Once nucleation and subsequent crystal growth have begun, proteins secreted into the mineralization medium bind to the growing crystals in a specific manner. Once complexed with protein, that given area or face of the crystal experiences no further mineral deposition. At the same time, regions which do not undergo protein binding grow unabated. By such a mode of operation, crystal growth is directed and ultimate crystal size and shape is determined. Cessation of growth, in the regions of crystal unaffected by protein binding, occurs when the mineralization front collides with an adjacent crystal. Such collisions are also thought to be responsible for the occlusion of the sculpting proteins into the shell.⁷⁰

Although evidence for both of the above mechanisms for crystal growth regulation exists, complete understanding of the true mechanism is far from complete. In all likelihood, there are numerous ways which nature has developed to deal with the need for crystal growth regulation. Regulatory mechanisms may differ across species lines or perhaps even differ as a function of the mineral morphology that a given mineralization event is trying to achieve.
Insoluble Organic Matrix:

Shell demineralization also yields a macromolecular component which is water insoluble. This insoluble portion of the shell matrix is thought to be structural in function.\textsuperscript{71,72} It is composed of both protein and carbohydrate, however, in comparison to the soluble matrix, the insoluble matrix components are hydrophobic in nature.\textsuperscript{73,74} Amino acid composition studies of ISM proteins show them to be rich in glycine and alanine. Also, by comparison, insoluble matrix proteins have higher amounts of aromatic amino acid residues and have a higher degree of crosslinking than do their soluble matrix counterparts. The structural polysaccharide chitin has been identified as a constituent of the insoluble matrix in several molluscan species.

The present picture of the organic matrix places the insoluble component at the core of the matrix structure (Figure 1.3). It is thought to act as the backbone or support element of the matrix. In addition to its structural function, the ISM is thought to be the substrate to which the soluble matrix components attach. Molluscan shells (sea shells) are as interesting from the viewpoint of material science as they are from the viewpoint of their chemistry/biochemistry.\textsuperscript{75} Shells are composed of simple, common materials, yet they are constructed in a way that makes them lightweight and strong. The insoluble matrix macromolecules undoubtedly contribute to the mechanical properties of the shell.
Mechanisms of Mineral Formation:

Having identified the primary constituents of the typical sea shell and having briefly outlined the theorized role that each constituent plays in the process of shell formation, a brief discussion of the leading theories of the events which lead to shell formation will be given.

The two leading theories of shell formation are formation by epitaxy and formation by ionotropy. To date, there is experimental evidence which supports both modes of shell formation, and as such, both have merit and deserve mention.
Mineral Formation By Epitaxy:

The basic premise of mineral growth by epitaxy is that the ultrastructure and orientation of crystal growth are determined exclusively by the substrate upon which the crystals grows.\(^{15}\) In shell formation, the organic matrix is that substrate. The organic matrix, more specifically the soluble organic matrix coating of the insoluble matrix framework, is seen as being a template for crystal growth. It is believed that inherent in the physical and chemical structure of the SOM is a pattern that sets up and dictates the crystal lattice structure of the developing mineral phase.

Initial studies based on the X-Ray diffraction patterns of organic matrices from 10 different decalcified molluscan shells, produced only one case where a general pattern of molecular orientation could be discerned.\(^{76}\) However, in another more recent examination utilizing electron diffraction to probe the matrix/mineral relationship, the nacreous layer of 4 partially decalcified molluscan shells showed, in all cases, that certain regions of the organic matrix had a preferred orientation.\(^{77}\) In addition, in two cases the spatial arrangement within the matrix was identified in the adjacent mineral phase. Such findings imply that localized regions of matrix may serve as a molecular template for crystal growth by epitaxy. These findings have also led to the hypothesis that these oriented sections of matrix may act as sites for crystal nucleation.

In support of the above findings, amino acid analysis data from
soluble matrix protein has revealed amino acid sequences which seem plausible candidates for the role of crystal template. The most noted sequence is the -[Asp-X]- sequence first proposed by Weiner and Hood. This sequence of alternating aspartic acid residues (the spacer (X) being Gly or Ala) is thought to possess the ability to bind calcium --via the carboxylic acid side chains of the Asp residues-- in such a way as to locally concentrate (crystal nucleation) and arrange the calcium ions in a pattern that will eventually dictate the lattice structure of the crystal that develops upon it (Figure 1.4). FTIR studies on soluble matrix protein indicate that, in

\[ \text{Asp} \ X \ \text{Asp} \]

\[ \text{Asp} \ X \ \text{Asp} \]

\[ \text{Ca} \]

\[ 0.95 \text{nm} \]

\[ 0.47 \text{nm} \]

\[ 0.69 \text{nm} \]

Figure 1.4 Illustration of the calcium liganded to two or three aspartic acid residues on the surface of a protein possessing β-sheet conformation and [-Asp-X-] primary structure. (Reference 1, p. 23)

the presence of Ca\(^{2+}\), the matrix protein adopts a beta sheet conformation. Such studies have concluded that it is not only the matrix
protein but rather the matrix protein in the presence of Ca$^{2+}$ which determines 3-dimensional protein structure and in turn, mineral morphology.

**Crystal Formation by I onotropy:**

Like the epitaxial model, the premier event in crystal formation by ionotropy is the binding of calcium by the organic matrix. The ionotropic model assumes, as does the epitaxial model, that the calcium binding sites are limited in number and strategically distributed across the matrix surface. However, in the ionotropic model, calcium binding sites do not act in concert to form a template for the crystal. Instead, the function of the calcium bind site is considered to be for crystal nucleation only. Once bound, the matrix/calcium complex is thought to seed crystal growth by attracting a layer counter ions (in EP fluid medium this counter ion would be CO$_3^{2-}$). This double layer of CO$_3^{2-}$ ions would in turn attract an additional layer of Ca$^{2+}$ ions. This snowballing effect of electrostatic interaction eventually leads to a highly localized concentrations of CaCO$_3$, which in turn results in crystal nucleation. The matrix does not act to set up the crystallographic lattice, in fact, initial mineralization is thought to be amorphous. There is evidence which supports amorphous deposition as the most energetically-favored mode of deposition in initial biomineralizational events. Also, deposits of amorphous CaCO$_3$ have been noted to occur in several species including mollusks. At some point after the initiation of...
crystal growth, a crystal phase transformation takes place wherein the crystal adopts its final morphological state. Transitions in mineral phase morphology have been noted to occur in several biomineralization processes.\textsuperscript{44}

![Diagram of ionotropic binding. The left drawing indicates the matrix binding of calcium. The right drawing indicates the formation of the carbonate and calcium double layer. (Reference 80)]

Regulation of crystal growth is achieved by the action of extrapallial fluid protein. The binding of EP Fluid protein to the growing crystal acts to inhibit further mineralization at the binding site.\textsuperscript{83,84} In doing so, EP fluid proteins are thought to sculpt the growing crystal into the desired shape. \textit{In vitro} experimental evidence has shown that, in CaCO\textsubscript{3}-saturated solution, SOM proteins fixed to a substrate may act to enhance CaCO\textsubscript{3} crystal
formation; however, when free in the solution these same SOM proteins act to inhibit CaCO$_3$ crystal formation. Such evidence has led to the hypothesis that when bound to the insoluble organic matrix, soluble matrix proteins act as effective CaCO$_3$ crystal nucleators and when free in solution the soluble matrix proteins act as crystal growth inhibitors. This dual mode of operation allows SOM protein to act as both crystal nucleator and crystal growth regulator.

In the ionotropic model of crystal formation, ester sulfate groups of protein-bound carbohydrate moieties are thought responsible for the binding of calcium. Experiments have been performed wherein shells are decalcified in solutions which fix glycoprotein. The result of such a procedure is the fixation of the SOM glycoprotein to the insoluble matrix substrate. Enzymatic investigation has shown these glycoproteins to be sulfated. Histological evidence has shown the fixed glycoproteins to be centers for calcium localization. Further studies have emphasized cooperation between sulfate and carboxylate groups as essential in crystal formation.

There are other theories which try to explain the controlling forces behind shell formation. One such view is the compartmentalized matrix theory. This theory holds that it is matrix cavities formed prior to mineral deposition which limit crystal size and shape. Another theory holds that it is a potential gradient which concentrates ions at the growing shell surface. Whatever the case, it is clear that mineral deposition is not
random. It is also clear from experimental evidence that any one of the above mentioned theories may hold clues to the \textit{in vivo} events of biomineralization. What is needed is further investigation. With this last thought in mind, we now discuss the topic of this doctoral work, the extrapallial fluid.

The Extrapallial Fluid:

As mentioned, the extrapallial fluid is the fluid medium which fills the extrapallial cavity. It is the medium within which matrix self assembly is thought to occur, as well as being the medium from which shell mineralization takes place. This being the case, the macromolecular components involved in shell formation, whether they are structural, crystal nucleators, growth regulators, ion transporters, etc..., no matter what their role in the mineralization process, they ought to reside, even if only transiently, in the extrapallial fluid. Qualitative analysis of extrapallial fluid has shown it to contain all the biomacromolecular materials which are found in the mature shell (protein, glycoprotein, carbohydrate).\textsuperscript{87,88} Despite being ascribed with a prominent role in shell formation, few studies, outside of this research laboratory,\textsuperscript{31-33} have focussed on the extrapallial fluid as a source for information on biomineralization.

As is thought to be the case for all biomineralization media, initial analyses of the extrapallial fluid have proved it to be unique in its
composition. Chemical analysis of the extrapallial fluid has shown the fluid to contain an assortment of different ionic components.\textsuperscript{89} The ionic makeup of the fluid has been shown to be different from that of the animal blood and the surrounding sea water.\textsuperscript{90} In addition, relative concentrations of the ionic species found within the extrapallial fluid have been shown to be regulated by the animal and not the environment.\textsuperscript{21} Studies on the fluid pH, [CO\textsubscript{2}], and [Ca\textsuperscript{2+}] as a function of shell opening and closure have shown the EP fluid to be a dynamic physiological fluid perhaps vital to pH regulation within the animal.\textsuperscript{91}

Studies on the EP fluid macromolecular composition have found it to heterogeneous in its protein content.\textsuperscript{31-33} EP fluid protein fractions are known to bind calcium.\textsuperscript{33} Studies have also shown that EP fluid protein of some species may act to inhibit CaCO\textsubscript{3} crystallization in much the same manner as protein derived from the shell matrix.\textsuperscript{24}

Studies in this laboratory have produced evidence that the extrapallial fluid of \textit{Mytilus Edulis} contains the enzyme tyrosinase.\textsuperscript{31,32} The tyrosinase family of enzymes are widespread in nature and accredited with a diverse array of functions.\textsuperscript{92} One of these functions is to act as an effective protein crosslinker. It is a tyrosinase enzyme that is thought to be responsible for protein cross linking in the periostracial layer of the \textit{Mytilus Edulis} shell.\textsuperscript{93} Perhaps it also acts to crosslink organic matrix proteins?

Given the significance of the role assigned to the EP fluid, along with
the results generated by the initial probes into its biochemical makeup and function, I undertook a research project to isolate and characterize the major protein component of the extrapallial fluid of *Mytilus Edulis*. The long term goal of this research is to gain further insight into to the shell formation process as it occurs in *Mytilus Edulis*, with hopes that this information may someday lead to a better understanding of biomineralization in general. It is evident from the foregoing that there are gaps in our knowledge of biomineralization, not only in the process as it occurs within *Mytilus Edulis*, but also, in the knowledge of all known biomineralizational processes which occur in Nature. Much time and effort have been put forth by numerous researchers in attempts to solve this puzzle of nature. However, as yet, a detailed understanding of the mechanisms which organisms utilize for mineralization in far from complete. I hope that the body of work presented in this dissertation will contribute toward the ultimate resolution of this mystery.
CHAPTER II

Materials and Methods

Mussel Source:

Mussels used for all experimental procedures were obtained from Great Eastern Mussel of Tenants Harbor, Maine. Mussels were purchased in 2 lb bags from Shaw's supermarket in Stratham, New Hampshire. The animals were kept on ice until the time of extrapallial fluid extraction and fluid was only taken from live animals.

Protein Buffer Solution:

The Protein Buffer Solution (PBS) used throughout the purification procedure was a 20 mM MOPS, pH 7.5, 0.1 M KCl, buffer solution. Prior to use at any stage of the protein purification procedure, the PBS was passed through a 1.5 x 25 cm column of Chelex 100 metal ion chelation resin (Sigma) and then filtered/degassed by suction filtration through a 0.45 μm pore Whatman filter.
**Fluid Extraction:**

The extrapallial fluid was removed from all mussels in the same fashion. Access into the shell cavity was gained by slicing the adductor muscle. The adductor muscle, which is responsible for the opening and closing of the bivalve shell, was cut with a scalpel and, once cut, the shell halves opened naturally. With the shell open, excess water was removed from the shell cavity by tilting the animal upright and letting the water drain onto a paper towel. The extrapallial fluid was then extracted with a 500 μl gas tight syringe. The syringe needle was placed bevel down on the shell margin and slid beneath the mantle, at the pallial line, into the extrapallial space. During the course of this procedure care was taken to disturb the mantle/shell attachment as little as possible. Having accessed the extrapallial cavity, the EP fluid was removed via the syringe. Once extracted, the EP fluid was placed in a polypropylene vial resting in an ice bath. Extrapallial fluid extraction was typically done with batches of 60-70 animals.

**Centrifugation:**

Immediately after extraction all EP fluid was subjected to centrifugation. Low speed centrifugation was performed in order to remove particulate matter (tissue, sand) residing in the EP fluid after extraction. The EP fluid was placed in polypropylene test tubes and spun in a tabletop
Adams Analytical Centrifuge at 3500 x g for five minutes. After centrifugation, the supernatant was retained and the pellet discarded. From each batch of centrifuged EP fluid, approximately two milliliters was reserved for raw fluid analysis. This fraction of fluid was placed in a polypropylene vial labeled "Raw Fluid" with the appropriate extraction date. All raw fluid samples (RF samples) were stored at 4 °C.

Sample Storage and Collection:
Throughout the course of this study all samples were stored in polypropylene vials (Fisher) at 4 °C. Small volume samples were stored in 1.5 ml polypropylene microfuge tubes (Fisher) and all chromatographic fraction collection was done with polyallomer test tubes (Seton). The practice of using non-glass containers was adopted in an attempt the limit the loss of protein material from adsorption onto glass surfaces.

Protein Salt Precipitation:
The remaining EP fluid was subjected to ammonium sulfate salt precipitation. This procedure was performed in order to remove and concentrate the proteinaceous material present within the fluid. The EP fluid was placed in Spectra/Pro 6000-8000 molecular weight cut off (MWCO) cellulose dialysis tubing and dialyzed against 1 L of PBS buffer brought to 85 % saturation (600 g/L) with ammonium sulfate (Sigma). The
(NH₄)₂SO₄/buffer solution was chilled to 4 °C prior to the dialysis procedure. Dialysis was performed for at least four hours, at 4 °C, under conditions of constant gentle stirring of the bulk buffer solution (magnetic stir bar). Once precipitation was complete, the precipitate within the dialysis bag was collected by centrifugation. Centrifugation was done in 1.5 ml polypropylene microfuge tubes at 12,500 x g in a Fisher Model 235 MicroCentrifuge. Centrifugation was performed for 10 minutes at 4 °C. After centrifugation the supernatant was drawn off and the protein precipitate dissolved in a minimum volume of PBS.

While research on the method for protein purification was in its initial stage of development, the supernatant from this procedure was returned to a dialysis bag and further exposed to the 85% (NH₄)₂SO₄ solution (up to 48 hours). This procedure was done to ensure that ample time had been allowed for the protein to salt out of solution. On no occasion did this extra dialysis step result in further precipitate formation. The supernatant from the salt precipitation step was also checked for the presence of protein by electrophoretic analysis and none was found. Therefore, in time, this procedure for checking the completeness of protein precipitation was discontinued and the supernatant from the salt precipitation step was simply discarded.
Equilibration in Protein Buffer Solution:

The precipitated/resolublized EP fluid protein was equilibrated in PBS via dialysis. The resolublized protein was placed in cellulose dialysis tubing, Spectra/Pro 6000 - 8000 MWCO, and dialyzed for several hours (usually overnight) against a liter of PBS to remove residual (NH₄)₂SO₄ remaining after the salting out process.

Cation Exchange Chromatography:

Cation exchange chromatography was performed on a 2 x 50 cm column filled with Sephadex CM -50 cation exchange resin (Sigma). The chromatographic resin was purchased as a dry powder. A literature estimate for the approximate volume of hydrated resin (1 g dry = 20 ml hydrated)⁸⁴ was used to determine the appropriate quantity of dry resin to use in column construction. The dry resin was allowed to swell and equilibrate in PBS for 72 hours prior to column packing. During this swelling/equilibration period the supernatant of the solution was twice decanted and replaced with fresh solution. PBS (as described above) was used as a mobile phase in all chromatographic procedures. Column flow was controlled by gravity. Strategic placement of the mobile phase reservoir in relation to the column inlet was used to maintain a flow rate of about 0.5 ml/min during periods of protein separation. Column effluent was monitored at 280 nm via a flow through detection cell and effluent fractions

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
were collected on an ISCO Retriever II fraction collector mounted with 16 x 102 mm polyallomer test tubes. 100 drops, or ~ 5 ml, was the typical fraction size collected. A salt gradient was not needed in order to achieve the separation of several EP fluid components. In order to inhibit bacterial growth during periods of non-use (more than a day or two) columns were filled with PBS containing 0.02% NaN₃.

Anion Exchange Chromatography:

Anion chromatography was performed using a Pharmacia Biotech HiTrap Q anion exchange column. Three 0.7 x 2.5 cm prepacked HiTrap Q columns were fastened together to make one 0.7 x 7.5 cm column. Protein buffer solution was used as a mobile phase. The column flow was controlled with a Gilson Minipuls II peristaltic pump and a flow rate of 1.5 ml/min was maintained during periods of protein separation. The column effluent was monitored at 280 nm via a flow through detection cell and effluent fractions were collected on an ISCO Retriever II fraction collector mounted with 16 x 102 mm polyallomer test tubes. 100 drops, or ~ 5 ml, was the typical fraction size collected. Protein elution was achieved with a PBS salt gradient of 0.1 - 0.4 M [KCl]. The salt gradient was created by connecting two Erlenmeyer flask reservoirs --one filled with PBS at starting [KCl] and one filled with PBS at final [KCl]-- with a piece of tygon tubing. Flow to the column was always siphoned from the low salt reservoir; in
addition, this low salt reservoir was maintained under conditions of constant, medium speed, stirring (magnetic stir bar). To complete each chromatographic run, the column was flushed with a 1M [KCl] PBS solution. Monitoring of the salt gradient was achieved by measuring the resistance of individual effluent fractions with a YSI Model 31 Conductivity Bridge. Resistance values of the individual effluent fractions were compared against the resistance values of PBS/[KCl] standards. This was done in order to determine the [KCl] concentration at which the various protein fractions eluted. In order to inhibit bacterial growth during periods of non-use (more than a day or two) columns were filled with PBS containing 0.02% NaN₃.

Size Exclusion Chromatography:

Size exclusion chromatography was performed on Sephadex 6B-CL size exclusion resin (Sigma). The resin was purchased preswollen; however, prior to column packing, the resin was allowed to equilibrate against the column mobile phase (PBS) for at least 24 hours (1 part resin/2 parts buffer). During this equilibration period the supernatant of the solution was twice decanted and replaced with fresh PBS. The column size was 1.5 cm inner diameter and 57 cm in length, creating a column bed of 100 cm³. The mobile phase was PBS and the flow was driven by gravity. A flow rate of about 15 ml/hr was maintained during protein separations. Column effluent was monitored at 280 nm via a flow through detection cell and effluent
fractions were collected on the ISCO Retriever II fraction collector mounted with 16 x 102 mm polyallomer test tubes. 100 drops, or ~ 5 ml, was the typical fraction size collected. Protein standards, (ferritin, albumin and cytochrome c) were used to assess column performance. During periods of non-use (more than a day or two) columns were filled with PBS containing 0.02% NaN₃.

Protein Electrophoresis:

All electrophoretic experiments were carried out on a Hoefer Mighty Small II electrophoresis unit. Electrophoretic reagent grade acrylamide, N,N'-methylene bis-acrylamide, lauryl sulfate sodium salt, TEMED and trizma base/ tris-HCl were purchased from Sigma. Baker analyzed reagent grade ammonium persulfate was purchased from J.T. Baker Chemical. Ampholines utilized in isoelectric focusing experiments (IEF) were purchased from Pharmacia. All the electrophoretic protocols (native PAGE, SDS PAGE, IEF) were taken from the Hoefer Scientific Instruments Manual 1992-1993. SDS PAGE molecular weight standards (serum albumin 66,200, ovalbumin 45,000, carbonic anhydrase 31,000, trypsin inhibitor 21,500, lysozyme 14,400) were purchased from Bio Rad and isoelectric focusing standards, carbonic anhydrase I (pl 6.6), carbonic anhydrase II (pl 5.9), trypsin inhibitor (pl 4.6), amylloglucosidase (pl 3.6), were purchased from Sigma. All gels were stained with brilliant coomassie blue R-250 protein staining solution as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
described in the Hoefer Manual. Glycoprotein specific gel staining was done by the Thymol/H$_2$SO$_4$ method of Rauchsen. All gels were 1.5 mm thick and electrophoresis experiments were conducted using a limiting current of 20 mA/gel with voltages typically ranging from 100 to 200 Volts. Gel scanning was performed on a Hoefer Scientific GS 300 transmittance/reflectance scanning densitometer interfaced to a MINC-23 computer (Digital Equipment Corp.).

Protein Concentration Assay:

The Bio Rad Protein Assay was used to estimate protein concentration the EP Fluid for various procedures throughout this thesis. The Bio Rad Assay is based on the Bradford Method of protein determination. The Bradford Method utilizes an absorbance shift in Coomassie Brilliant Blue G-250 (465 to 595nm) when the dye binds to protein. Bovine serum albumin was used as a standard and 5 point (0.2 - 1.0 mg/ml) standard calibration curves were constructed for protein assays. Both standard and protein samples were always run in duplicate. Absorbance measurements were performed in 1 cm methyl acrylate disposable cuvets (Fisher) on a Cary 219 UV-VIS spectrometer.

Neutral Sugar Assay:

For the determination/quantitation of neutral sugars moieties bound to
protein, the phenol-sulfuric acid colorimetric assay of Dubois\textsuperscript{98} was utilized. D-galactose served as a carbohydrate standard. Standard calibration curves ranging from 5 - 30 $\mu$g D-galactose in PBS were used to quantitate carbohydrate EP fluid samples. Absorption measurements (489 nm) were made in 1 cm quartz cuvets (Fisher) on the Cary 219 UV-VIS spectrometer.

**Ultrafiltration:**

Ultrafiltration was used to concentrate protein samples. Amicon 3, 10 and 50 ml cells, depending on sample volume, equipped with the correct size Amicon PM-10 (10,000 MW cutoff) ultrafiltration membrane were used for this purpose. Ultrafiltration was done at 4 °C under 20 psi of nitrogen pressure. Both ultrafiltrate and retentate were usually analyzed for protein content.

**Protein Lyophilization:**

Protein lyophilization was carried out after exhaustive dialysis (6000-8000 MWCO dialysis tubing) of EP fluid samples against either DI H$_2$O or a 5.0 mM NH$_4$HCO$_3$ solution at 4°C. Lyophilization was carried out in 1 x 10 cm pyrex test tubes containing 1 ml of protein solution. All samples for a given lyophilization experiment were frozen in a dry ice/acetone bath (-77 °C) and placed in a vacuum container until sublimated to dryness. The system vacuum was generated by a Sargent Welch DirecTorr vacuum
pump, model number 8811:

**Mass Spectrometry:**

The protein molecular weight was determined by matrix-assisted, laser desorption ionization, time-of-flight mass spectrometry on a Perceptive Biosystems, Voyager-Elite Biospectrometry Workstation. Instrument calibration was achieved using an insulin mass standard. Analysis of \( \text{NH}_4\text{HCO}_3 \) lyophilized EP fluid protein sample was performed using an accelerating voltage of 30,000 V, a pressure of \( 9.50 \times 10^{-8} \text{ torr} \), and a low mass cutoff of 500.0. The analysis was performed in the laboratory of Dr P. Matsudaira at the Whitehead Research Institute, Cambridge, Massachusetts.

**Amino Acid Analysis:**

Protein amino acid analysis was performed at AAA Laboratory, Mercer Island, Washington and the University of Michigan Protein and Carbohydrate Structure Facility, Ann Arbor, Michigan. The purified EP fluid samples submitted for analysis were dialyzed against a less concentrated than normal protein buffer solution, 5.0 mM MOPS, pH 7.5, 0.1 M KCl. Samples were hydrolyzed for 20 hours in 6 N HCl, 0.05% mercaptoethanol, 0.02% phenol, at 115°C. At the AAA laboratory, amino acid separation and quantitation were performed on a Beckman 7300 analyzer using system
Gold software and Beckman buffers. The procedure used was the ion-exchange method developed by Moore and Stein. At the University of Michigan laboratory, the analysis was performed on an Applied Biosystems 420H amino acid analyzer.

**N-Terminal Amino Acid Sequence:**

N-Terminal amino acid sequence analysis was performed at the Whitehead Research Institute, Cambridge, Massachusetts. EP fluid protein samples submitted for analysis were prepared by two methods. One sample of purified EP fluid protein was dialyzed against 0.5 mM NH₄HCO₃, pH 7.5 and then lyophilized. A second sample was prepared by dialysis of purified protein against a 10% (v/v) solution of acetic acid. Both samples were analyzed on a Applied Biosystems Inc., 475A Protein Sequencer, by the method of automated Edman degradation.

**UV-Vis Spectroscopy:**

All UV-Vis spectroscopy (except if otherwise noted) was conducted on a Cary 219 UV-Vis spectrometer. The cuvets used were 1 cm in length composed of either quartz or methyl methacrylate depending on the wavelength range of the analysis being performed. The instrument was always run in the double beam, auto gain mode, using a 1 nm slit width. Wavelength scans were conducted at 2 nm/s, except during the O-
glycosidic linkage experiment, for which scans were conducted at 1 nm/s.

All wavelength measurements/scans were made at room temperature (~22.5 °C), although the instrument sample cell temperature was not regulated.

Both sample and reference chambers were purged with a 5 psi flow of nitrogen gas.

**Ultracentrifugation:**

Sedimentation velocity experiments were performed on a Beckman XLA Analytical Centrifuge equipped with Rayleigh interference optics. Experiments were conducted at 20 °C using a four hole titanium rotor spinning at 40,000 rpm. The sample cells used were, two channel, charcoal filled, equipped with epon centerpieces and sapphire windows. Each cell was typically loaded with 480 µl of sample in one channel and an equal volume of sample buffer in the other. The buffer used in all sedimentation velocity analyses was 20 mM tris, pH 7.5, 0.1 M KCl; where mentioned, some samples were also analyzed with the addition of CaCl₂ to the tris buffer solution. The method of Stafford was used to obtain sedimentation coefficient distributions (g(S*)) from the time derivative of the concentration distributions (dc/dt). Protein partial specific volume (v 25 °C) was estimated by using the protein amino acid composition following the method of Cohn and Edsall. Values for partial specific volume of the individual amino acid residues were obtained from Laue.
the estimate of $v$ at 25 °C to the 20 °C value was also done in accordance with Laue.\textsuperscript{103} An estimate of solvent density was obtained by summing the density increments calculated for each solvent component according to the method of Svedberg and Peterson.\textsuperscript{103} Tabulated values for coefficient fitting parameters of the buffer components used in sedimentation velocity experiments were obtained from Laue.\textsuperscript{103}

Infra-Red Spectroscopy:

Infra-red spectroscopy was utilized for the identification of sulfate groups. A Nicolet MX-1 FT-IR Spectrometer was employed in order to identify the presence of sulfate moieties attached to either carbohydrate or protein. A lyophilized protein sample was combined with oven dried KBr in a 1/100, Protein/KBr, weight ratio. The salt/protein mixture was ground into a fine powder using a mortar and pestle. Once homogenized, a small fraction of the mixture was pressed into a clear window using a hand held pellet press. Output spectra were the result of a Fourier transform average of 128 individual scans. The scan range was 4000 to 400 cm\textsuperscript{-1}.

EDX Analysis:

Energy dispersion x-ray analysis (EDX) of EP fluid protein samples was performed on a Amray 1000 scanning electron microscope. Purified protein samples were prepared for analysis by exhaustive dialysis against DI
H₂O followed by lyophilization. Lyophilized fluid protein was given a carbon coating and analyzed with a 20 kV electron beam for 149 seconds. Instrument detection was capable of analyzing X-rays ranging in energy from 0 to 20 keV. This analysis range has the capability of detecting anything with an atomic mass greater than boron (B) and less than lead (Pb).

**Nuclear Magnetic Resonance:**

Phosphorus 31 NMR was employed to determine the presence/absence of protein bound phosphate groups. Two Spectrometers were used in the study; JEOL FX 90 FT-NMR and a Bruker AM 360 MHz superconducting NMR. Analyses were performed in 10 mm NMR tubes. Pulse parameters were selected based on values suggested for the analysis of phosphoprotein.¹⁰⁴ For the JEOL FX 90 FT-NMR, the following parameters were employed; Peak Width 38 μs, Sweep Width + 2000 Hz, Delay Time 0.5 s, Acquisition Time 1.86 s, Filter 1000, Flip Angle 70 degrees. A 0.25 mM phosphoric acid sample, analyzed separately, was used as an external reference standard.

**Total Phosphate Assay:**

Quantitative determination of protein bound phosphate was attempted using the assay of Buss and Stull.¹⁰⁵ The assay involved the conversion of protein bound phosphate into an inorganic phosphomolybdate compound.
The phosphomolybdate compound was then complexed to a dye (Malachite green), resulting in color production. Oven dried (3 hrs.) potassium phosphate (Fisher Scientific) was used to make standard solutions (4 - 16 nmol phosphate). Absorbance measurements (660 nm) were made with 1 cm methyl acrylate disposable cuvets (Fisher) on the Cary 219 UV-VIS spectrometer. Calibration standards, as well as EP fluid samples, were always run in duplicate.

**Sialic Acid Assay:**

The Periodate-thiobarbituric acid method of Aminoff was employed for the determination of sialic acid moieties. In this assay, sialic acid is oxidized by periodate (Sigma) resulting in the formation of a chromogen. A sodium arsenite (Sigma) solution is used to reduce excess periodate and color development takes place in an alkaline (pH = 9.0) thiobarbituric acid (Sigma) solution. Commercial N-acetylneuraminic acid (Sigma) was used as a standard. Standard solutions (5 - 40 μg) were used to construct a standard curve in all assay procedures. Absorbance measurements (549 nm) were made on the Cary 219 UV-VIS spectrometer.

**Electron Paramagnetic Resonance:**

Electron paramagnetic spectroscopy of EP fluid protein samples was performed on a homr built spectrometer composed of Bruker, Varian, and
Micro-Now components. Analyses were conducted at 77 K in a 4 mm EPR tube. Instrument settings were as follows; Scan Range 1200 G, Time Constant 0.3 s, Modulation Amplitude 5 G, Receiver Gain .003 V, Microwave Power 5 mW, Field Set 2950 G, Scan Time 500 seconds, Modulation Frequency 100 kHz, and Microwave Frequency 9.31 GHz.

Atomic Absorption:

Atomic absorption experiments were performed on a Thermal Jarrell Ash Smith-Hieftje-12 spectrometer equipped with a furnace atomizer. Elemental copper absorption was monitored at 324.7 nm. Protein samples were diluted so that a 10 µl injection volume would give an absorbance reading comparable to a copper concentration of approximately 10 - 50 µg/L. The sample injection volume was 10 µl. Once in the graphite furnace, the sample was dried at 150 °C for 10 s, then pyrolyzed for 15 s at 600 °C and for 10 s at 750 °C, last the sample was atomized at 2400 °C for 4 s. During the atomization time period, signal absorbance was measured. Copper quantitation in EP fluid protein samples was done by the method of standard addition. Standard addition calibration curves were constructed ranging from 0 - 100 µg/L copper. Standard solutions were made by the dilution of a 1,000 ppm copper atomic absorption standard (Fisher) in 2% nitric acid.
Circular Dichroism:

Circular dichroism (CD) experiments were done on a JASCO J700 Circular Dichroism Spectropolarimeter at the Rowland Institute of Science, Cambridge, Massachusetts. Instrument calibration was done on a daily basis using a 0.06% (+)-10-camphorsulfonic acid (CSA) solution, (Aldrich). The calibration procedure was done in accordance with the JASCO J700 manual. CD spectra of EP fluid protein (1.07 mg/ml) in PBS, EP fluid protein (1.07 mg/ml) in PBS brought to 10 mM Ca²⁺, as well as the buffers in the absence of protein (PBS and PBS 10 mM Ca²⁺) were obtained. In order to scan the protein samples in the far UV region (200 - 180 nm), a 0.05 mm sample cell was used. Scans were taken at room temperature and the sample chamber was continuously flushed with a 15 L/min flow of nitrogen. Instrument settings were as follows; Scan Range 260-180 nm, Scan Rate 1 nm/min, Wavelength Step 0.5 nm, Sensitivity 5 mdeg, Response Time 16 s.
CHAPTER III

Results

Raw EP Fluid Protein and Carbohydrate Analysis:

A typical batch extraction of mussels (60 - 70 animals) yielded 20 - 25 ml of EP fluid. Individual animals contained an average of 300 μl of fluid; however, a wide variation from practically nothing to in excess of 500 μl was observed.

The protein and carbohydrate concentration in untreated extrapallial fluid samples was determined throughout the course of a year. The study was done in order to generate data on the seasonal variation in the quantity of these macromolecules in the fluid. The Bio Rad protein assay and the phenol/H₂SO₄ neutral sugar assay were utilized for this purpose. Assays were performed on each new batch of EP fluid. Assay results are reported by month according to when the fluid extraction took place. Value for protein concentration are the cumulative results of 6-10 separate assays. The values reported for carbohydrate concentration are derived from only one or two separate carbohydrate assay(s). Graphical (Figure 3.1) and
tabular (Tables 3.1 and 3.2) displays of the data are given below. Table 3.3 shows the results of a protein assay performed on fluid samples taken from 10 individual *Mytilus Edulis* animals.

![Graph showing protein and carbohydrate concentration over months](filename: month.org)

**Figure 3.1** Extrapallial fluid protein/carbohydrate concentration plotted against the month in which fluid extraction took place. Inset is plot of [Protein] vs. [Carbohydrate] by month. (filename: month.org)
<table>
<thead>
<tr>
<th>Month</th>
<th>Pro. Conc. mg/ml</th>
<th>Std. Dev.</th>
<th>Rel. Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>4.79</td>
<td>0.61</td>
<td>12.7 %</td>
<td>6</td>
</tr>
<tr>
<td>February</td>
<td>6.53</td>
<td>0.58</td>
<td>8.09 %</td>
<td>6</td>
</tr>
<tr>
<td>April</td>
<td>4.58</td>
<td>0.68</td>
<td>14.8 %</td>
<td>10</td>
</tr>
<tr>
<td>May</td>
<td>2.41</td>
<td>0.21</td>
<td>8.71 %</td>
<td>6</td>
</tr>
<tr>
<td>June</td>
<td>2.68</td>
<td>0.40</td>
<td>14.9 %</td>
<td>8</td>
</tr>
<tr>
<td>July</td>
<td>3.51</td>
<td>0.21</td>
<td>5.98 %</td>
<td>6</td>
</tr>
<tr>
<td>October</td>
<td>2.83</td>
<td>0.41</td>
<td>14.3 %</td>
<td>6</td>
</tr>
<tr>
<td>December</td>
<td>6.98</td>
<td>0.87</td>
<td>12.4 %</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3.1 Extrapallial Fluid Protein Concentration (filename: month.org)

<table>
<thead>
<tr>
<th>Month</th>
<th>Carbohydrate Conc. mg/ml</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>3.0</td>
<td>2</td>
</tr>
<tr>
<td>February</td>
<td>3.5</td>
<td>2</td>
</tr>
<tr>
<td>April</td>
<td>4.2</td>
<td>1</td>
</tr>
<tr>
<td>May</td>
<td>2.1</td>
<td>2</td>
</tr>
<tr>
<td>June</td>
<td>3.4</td>
<td>1</td>
</tr>
<tr>
<td>July</td>
<td>3.6</td>
<td>1</td>
</tr>
<tr>
<td>October</td>
<td>4.2</td>
<td>2</td>
</tr>
<tr>
<td>December</td>
<td>6.2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3.2 Extrapallial Fluid Carbohydrate Concentration (filename: month.org)
Table 3.3 EP fluid Protein Concentration from 10 Animals (filename: month.org).

<table>
<thead>
<tr>
<th>ID #</th>
<th>[Protein] mg/ml</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.60</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4.10</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3.66</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4.82</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>4.88</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>4.96</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>5.24</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>5.20</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>5.14</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>3.75</td>
<td>1</td>
</tr>
</tbody>
</table>

Ave 4.58 mg/ml  
SD 0.68 mg/ml  
RSD 14.8 %

Results from the year-long study yielded an extrapallial fluid protein concentration of 4.29 mg/ml with a relative standard deviation of 40.8 %. The average extrapallial fluid carbohydrate concentration was 3.8 mg/ml with a relative standard deviation of 32 %. Figure 3.1 reveals some general trends in the EP fluid protein/carbohydrate data. The Figure 3.1 inset shows that EP fluid protein and carbohydrate concentrations are roughly correlated (R = 0.61). This trend indicates that the fluid protein/carbohydrate concentrations may be linked, perhaps via the presence of glycoproteins. Regardless of the connection between the macromolecules, with the
exception of months January and February, the fluid protein and carbohydrate levels are within experimental error of one another. This finding points to a consistency in the biological regulation of the fluid protein and carbohydrate levels. Generally speaking, it appears that EP fluid protein levels are higher during the northern hemisphere's winter months when compared against protein fluid levels assayed during the summer months. The average protein concentration for the months of December, January and February is 6.10 mg/ml compared with an average protein concentration of 2.86 mg/ml for months May, June and July. Before lending significance to this observation, several factors must be considered. Results from the protein assay of EP fluid taken from individual animals mimic the results of the monthly batch assays in regards to the standard deviation of the data (RSD ~ 12%). Imprecision in measurement of the bovine serum albumin assay calibration standard leads to an inherent 6.89% RSD in the reported concentration values. Therefore the remaining discrepancy in monthly fluid protein levels, stems from differences between individual animals, perhaps due to health or age, as well as from cycles related to seasonal or environmental variations. The extent to which each of these factors influences the protein/carbohydrate content of the EP fluid is unknown.
Extrapallial fluid Protein Distribution:

Initial investigation into the number and distribution of extrapallial fluid proteins was done by native and SDS-PAGE. Raw fluid samples taken from batch extracts as well as from individual animals were analyzed. The intent of the investigation was to determine whether the distribution and relative abundance of the various fluid proteins remained constant from batch to batch, month to month, and between individual animals. Figures 3.2, 3.3 and 3.4 are three electrophoretic gels which display the consistency of the EP fluid in its distribution of protein. All of the gels are stained with an indiscriminate protein stain containing Coomassie Brilliant Blue R-250 dye. The depth of staining at any given protein band represents the relative abundance of that protein. All of the gels clearly demonstrate that the EP fluid contains a dominant protein species (band which appears at the bottom of lanes 1-6 in Figure 3.2).

Figure 3.2 15% Native PAGE gel of fluid samples extracted in Oct, Aug, Jul
Lanes: 1) 20 µl Raw Fluid (RF) Oct 2) 10 µl RF Oct 3) 20 µl RF Aug 4) 10 µl RF Aug, 5) 20 µl RF Jul 6) 10 µl RF Jul 7) Protein Std's (from top to bottom), Transferrin (80,000); albumin (67,000); carbonic anhydrase (34,000); myoglobin (17,500).
Figure 3.3 15% SDS PAGE gel of raw fluid samples extracted in Oct, Aug, Jul
Lanes: 1) 5 µl Raw Fluid (RF) Oct 2) 10 µl RF Oct 3) 20 µl RF Oct 4) 5 µl RF Aug
5) 10 µl RF Aug 6) 20 µl RF Aug 7) 5 µl RF Jul 8) 10 µl RF Jul 9) 20 µl RF Jul
10) Protein Std’s (see table 3.5).

Figure 3.4 15% native PAGE gel of raw fluid samples from 8 animals.
Lanes: 1-8) 20 µl Raw fluid from 8 animals 9) Protein Std’s (same as Figure 3.2).
A gauge on the prominence of this protein was obtained by density scans of the individual gel lanes. Shown in Figure 3.5 is the scan of a native PAGE gel lane containing a raw fluid sample. The large peak in the density scan profile corresponds to the heavily stained protein band on the native gels in Figures 3.2 and 3.4. By integrating the scan peaks, and averaging the results of 3 separate scans, it was estimated that 56% ± 15% of the total extrapallial fluid protein migrates in this major band. This was the protein fraction targeted for further purification.
Protein Purification:

Purification of the EP fluid protein began with the salting out of all fluid proteins with 85% saturated (NH$_4$)$_2$SO$_4$ solution. The protein precipitate was resolublized in PBS and dialyzed against a solution of the same to remove residual (NH$_4$)$_2$SO$_4$. From this point, protein purification was accomplished via a series of chromatographic separations. A diagram outlining the purification procedure is shown on the following page. While developing this procedure, the protein fractions which resulted from chromatographic separations were analyzed by electrophoresis as a means of tracking the protein of interest. Discussed below are the results from the separation procedures used for protein purification. Details on the experimental procedures are given in the Material and Methods section.

Cation Exchange Chromatography

The first chromatographic procedure was cation exchange on Sephadex CM-50. This procedure yielded three different EP fluid fractions. Separation was accomplished without the aid of a salt gradient. The protein of interest was unretained and eluted in the void peak. Figure 3.6 is an example of a typical cation chromatogram.

Anion Exchange Chromatography

The second chromatographic procedure was anion exchange on
Protein Purification Flow Chart

Extraction
  ↓
Salt Precipitation
  ↓
Discard
  ↓
Precipitate
  ↓
Supernatant
  ↓
Dissolve in PBS
  ↓
Cation Exchange
  ↓
Discard
  ↓
Peaks Two & Three
  ↓
Void Peak
  ↓
Anion Exchange
  ↓
Discard
  ↓
Peak at 0.4 M KCl
  ↓
Everything Else
  ↓
Size Exclusion
  ↓
Purified Protein

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Pharmacia Hi-Trap Q. Protein elution required the use of a (0.1 - 0.4 M KCl) salt gradient. The protein of interest eluted at a [KCl] between 0.36 - 0.38. Figure 3.7 is an example of an anion exchange chromatogram. The peak containing the major EP fluid protein is labelled G.

**Size Exclusion Chromatography**

The final chromatographic step utilized size exclusion chromatography. The resin used was Sepharose 6B-CL. Figure 3.8 is a size exclusion chromatogram showing the separation profile for the EP fluid protein (labelled G). Following the fluid sample chromatogram is a chromatogram showing separation of three standard proteins. The standards are of known molecular weight and were used to estimate the unknown molecular weight of fluid proteins as well as serving as a gauge for column resolution.

Figure 3.9 shows a 10% native PAGE gel showing the major EP fluid protein band at the various steps of the purification procedure. Table 3.4 roughly displays the efficiency of the purification process by reporting the percent yield in milligrams of protein for the various purification steps.
Figure 3.6 Elution profile resulting for the Sephadex CM-50 cation exchange separation of resolublized (NH₄)₂SO₄ precipitated EP fluid protein (2 ml sample volume of ~6 mg/ml protein).

Figure 3.7 Typical elution profile of Hi Trap Q anion exchange separation of void peak from cation exchange separation (2 ml sample volume of ~2.7 mg/ml protein). The drawn in curve (red line) is a trace of the [KCl] (right axis) vs. Volume. After elution of peak G, the column is flushed with 1.0 M KCl in PBS.
Figure 3.8  Typical elution profile of Sepharose 6B size exclusion separation of peak G from anion exchange separation (1 ml sample volume of - 2 mg/ml protein). The three peaks to the right are standards, ferritin, albumin, ribonuclease.

Figure 3.9  10% Native gel showing the major EP Fluid protein in its various stages of purity. Lanes: 1) 20 µl Raw Fluid  2) Void peak off cation Ex. CM-50  3) G peak off anion Ex. Hi Trap Q  4) G peak off Size Exc. 6B
Step | Vol. added | mg protein added | Vol. collected | mg collected | mg added/mg collected | Efficiency
--- | --- | --- | --- | --- | --- | ---
Extraction | | | 24.0 ml | 53.5 mg | | |
Salt Precip | 24.0 ml | 53.5 mg | 7.00 ml | 45.9 mg | .857 | 85.7%
Cation Ex. | 2.00 ml | 13.1 mg | 4.00 ml | 10.7 mg | .810 | 81.0%
Anion Ex. | 2.00 ml | 5.30 mg | 3.60 ml | 1.94 mg | .366 | 36.6%
Size Ex. | 1.00 ml | 0.54 mg | 2.00 ml | 0.26 mg | .480 | 48.0%

| Table 3.4 | Data of the efficiency of the protein purification procedure. All protein values represent the total amount of protein present before or after a given purification step.

Based on the efficiencies displayed in Table 3.4, a batch purification process beginning with 53.5 mg of total protein, would produce 6.52 mg of purified protein. According to the densitometric gel scan, 56% of the total EP fluid protein is composed of the protein targeted by the purification procedure. The densitometric value allows for the estimation of (53.5 x 0.56) or 30.0 mg of the starting material being composed of the protein of interest. Therefore, 6.52 mg of purified protein, out of a possible 30.0 mg, results in ~ 22% yield. During all chromatographic procedures, it was common practice to only collect material which eluted at the center of the peak of interest. This practice contributed to the low overall protein yield. Protein volume and weight estimates, after any given purification step, were determined by the Bio-Rad protein assay after concentration by ultrafiltration. The measurements in Table 3.4 are approximate and represent only a single attempt to monitor the efficiency of the purification procedure. Therefore, at best, this study should be viewed as a rough
gauge on the efficiency of the purification procedure.

**Molecular Weight Determination:**

An estimate of the molecular weight of purified EP fluid protein was obtained by SDS-PAGE and time-of-flight mass spectrometry. For a given protein, the response factor (Rf) on an SDS-PAGE gel is a measure of the efficiency with which that protein migrates through the gel from the cathode (+) to the anode (-). The Rf is calculated by dividing the distance migrated by the protein in question by the distance migrated by the tracking dye. In all electrophoresis experiments discussed in this dissertation bromphenol blue was used as the tracking dye. Bromophenol blue is known to migrate through gels faster than most proteins. In addition, it is visible throughout the course of the experiment and thus serves as an indicator for completion of electrophoresis. When the tracking dye reached the bottom of the gel, electrophoresis was terminated. Sodium dodecyl sulfate (SDS) serves to denature the protein and to impose a net negative charge on the protein. The SDS induced negative charge insures that proteins will migrate from the cathode to anode. SDS also sets protein size as the dependent variable of migration efficiency. By comparing the Rf of a protein of unknown molecular weight against the Rf values obtained from protein standards of known molecular weight, an estimate of the unknown protein molecular weight can be obtained.
A plot of Rf vs. log (molecular weight) for several protein standards was used to construct a calibration curve from which the EP fluid protein MW was determined. Figure 3.10 shows a 15% SDS PAGE gel containing purified protein and MW standards. Table 3.4 gives the identity, molecular weight data, and gel Rf values for the standard proteins as well as the purified EP fluid protein. Figure 3.11 is a plot of the standard protein Rf vs. log MW used to estimate the EP fluid protein molecular weight. From Figure 3.11 a molecular weight of 37,300 ± 1,800 was determined.

SDS-PAGE experiments may also contain useful information about a protein's quaternary structure. Inclusion of a reducing agent, such as β-mercaptoethanol, into the sample preparation will likely result in the destruction of any inherent protein crosslinking disulfide bonds. The destruction of interprotein crosslinks will cause the protein to migrate through the gel as constituent subunits. Therefore, by analyzing a protein under reducing and nonreducing conditions, information about the number and size of protein subunits as well as information about the presence or absence of subunit crosslinking disulfide bonds may be obtained.

The EP fluid protein showed the same migrational pattern (that seen in Figure 3.10) under both reducing and nonreducing conditions. From this observation it may be concluded that the EP fluid protein is composed of a single type of subunit and that the protein contains no intersubunit disulfide bonds.
Matrix-assisted, laser desorption, ionization time-of-flight mass spectrometry was also used to determine the EP fluid protein's molecular weight. Figure 3.12 is the time of flight mass spectrum produced from the analysis of a lyophilized EP fluid protein sample. The peak centered at 28,340 mass units on the spectrum represents the anhydrous molecular mass of the ionized EP fluid protein \((M/1^+ )\). The peak centered at 14,200 mu represents the doubly ionized protein species \((M/2^+ )\). Mass values determined by this technique are thought to be accurate to within 5 mass units.

Figure 3.12 shows that the EP fluid protein peak in the mass spectrum is broad. The baseline width of the peak ranges from approximately 27,000 - 29,000 mass units. The peak breadth is indicative of a heterogeneous mixture of molecules and the peak profile suggests minor peaks at 27,000, 27,840 and 28,960 mu, in addition to the major 28,340 mu peak. As discussed later in this Chapter, the EP fluid protein is composed of ~ 7.25% carbohydrate (it is a glycoprotein). For glycoproteins, it is common to observe microheterogeneity in the extent of glycosylation. The mass of a single monosaccharide group averages 162 mu, and the addition or deletion of single polysaccharide chain may account for the observed mass heterogeneity.
Figure 3.10 15% SDS PAGE gel showing purified protein (Lanes 6-8) and protein standard molecular weight markers (Lanes 1&10 --see Table 3.5-- ).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Mol. Wt.</th>
<th>Log MW</th>
<th>Rf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum Albumin</td>
<td>66,200</td>
<td>4.821</td>
<td>0.139</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>45,000</td>
<td>4.653</td>
<td>0.254</td>
</tr>
<tr>
<td>Carbonic Anhydrase</td>
<td>31,000</td>
<td>4.491</td>
<td>0.416</td>
</tr>
<tr>
<td>Trypsin Inhibitor</td>
<td>21,500</td>
<td>4.332</td>
<td>0.607</td>
</tr>
<tr>
<td>Lysozyme</td>
<td>14,400</td>
<td>4.158</td>
<td>0.772</td>
</tr>
<tr>
<td>EP fluid Protein</td>
<td>-37,300</td>
<td>4.570</td>
<td>0.360</td>
</tr>
</tbody>
</table>
<pre><code>                             | from log MW | from graph | from gel |
</code></pre>

Table 3.5 Molecular weight data and Rf values used to construct the Rf vs. log MW standard curve (filename: sdsgei.org)
Figure 3.11 15% SDS PAGE Rf vs. log MW standard calibration curve

Figure 3.12 Time of flight mass spectrum of a purified EP fluid protein sample.
The presence of protein glycosylations may also be responsible for the discrepancy in the molecular weights as determined by the two methods, 37,300 vs. 28,340 mu. Protein glycosylations do not bind SDS and as a result will often lead to erroneously high MW estimates on SDS PAGE gels. In addition, a basic premise of MW determination by SDS-PAGE is that the protein behaves as a globular protein. Any deviation from the globular structure (if the protein is cylindrical or oblate) can lead to inaccuracies in the MW calculations. Due to the greater degree of accuracy inherent in molecular weight determination by mass spectrometry, an EP fluid protein molecular weight of 28,340 will be used in future discussions and calculations.

**Isoelectric Focusing:**

Isoelectric focusing (IEF) was performed in order to determine the isoelectric point (pI) of the purified EP fluid protein. The protein pI corresponds to the pH value at which the protein has no net charge. The pI is determined by performing electrophoresis in the presence of a pH gradient. Proper construction of the pH gradient is monitored by the use of protein standards of known pI value. By comparing the focusing point of an unknown protein against the focusing points of protein standards, an estimate of the unknown protein's pI value may be obtained. For EP fluid protein IEF experiments, a pH gradient from pH 7 to pH 2.5 was
constructed. In order to achieve the pH gradient various amounts of gradient generating ampholines were used. The pH gradient in the IEF gel shown in Figure 3.11 was constructed using 550 µl pH 4-6, 200 µl pH 5-8, 150 µl pH 2-4 and 100 µl pH 7-9 ampholines. Carbonic anhydrase I (pI 6.6), carbonic anhydrase II (pI 5.9), trypsin inhibitor (pI 4.6) and amylglucosidase (pI 3.6) were used as focusing standards.

Figure 3.12 is a picture of an isoelectric focusing gel. From the gel it may be seen that the protein has numerous isoelectric points. This microheterogeneity in the isoelectric point is the result of microheterogeneity in the protein's surface charge. As mentioned, heterogeneity is a common feature of glycoproteins. Variance in surface charge may arise from heterogeneity in the number and/or type of sugar moieties attached to the protein core. In addition, there may be variation in the number and/or type of functionalities attached to the sugar residues (sialic acid, sulfate and phosphate groups). Microheterogeneity may also arise from protein related inconsistencies (partially filling of metal-binding sites, modified amino acid residues, protein sample degradation). To help ensure the observed EP fluid pI point heterogeneity was genuine, several IEF experiments were conducted. In each experiment the EP fluid protein focused as several bands. Depending on the resolution of the particular experiment, as many as six different bands could be distinguished. The EP fluid IEF value was determined by using the center, and most heavily stained, section of the
Figure 3.13  pl vs. Rf for isoelectric focusing protein standards. A pl of approximately 4.85 was determined for EP fluid protein.

Figure 3.14  Isoelectric focusing gel of EP fluid protein and IEF standards. 1), pl standards from top to bottom pl 6.6, 5.9, 4.6, 3.6  2) peak G anion Ex.  3) purified EP fluid protein  4) Close up of an EP fluid protein sample showing pl heterogeneity.
protein band stack as the migration distance in the pI response factor (Rf) calculation. Figure 3.13 is a plot of Rf vs. pI for the protein standards used in the focusing experiment. From this plot an EP fluid protein pI value of 4.85 was determined. The pI range over which the numerous protein bands focused was from pH 4.66 to 5.02.

Determination of Protein Molar Absorptivity:

Scans of the 400 - 250 nm UV spectral region of EP fluid protein samples consistently showed an absorbance maximum at 276 nm. Figure 3.15 shows a scan of the protein from 400 - 250 nm. Protein absorption in this region arises from the aromatic groups present on phenylalanine, tyrosine and tryptophan amino acid residues, and may be used for protein quantitation. By determining the protein concentration for several purified EP fluid samples and then measuring the absorption of these samples at 276 nm, a plot of absorbance vs. concentration was constructed. The protein concentration range covered was (0.1 - 1.0 mg/ml).

According to Beer’s law, for a given wavelength, absorption is directly proportional to concentration.

\[ A = ebc \]

A = Absorbance (A)  
b = pathlength (cm)  
c = concentration (mol/L)  
e = molar absorptivity (L/mol cm)
By fixing the cell pathlength (b), the absorption wavelength and thus the molar absorptivity (ε), if Beer's law is obeyed, a plot of absorbance vs. concentration should yield a straight line of slope (εb). By using a one-cm pathlength (b = 1), the slope obtained from the plot is rendered equal to the protein molar absorptivity at the wavelength used to make the absorption measurements. Figure 3.16 is a plot of absorbance at 276 nm vs. molar concentration for several EP fluid protein samples. From Figure 3.16 a molar absorptivity of 38,600 L/mol cm (at 276 nm) was obtained which is in the proper range for most proteins. The solvent used for absorption measurements was the protein buffer solution (PBS). PBS was also used as the double-beam reference for all measurements and to zero the instrument absorbance reading. Measurements were conducted at 25 °C in a 1 cm quartz cell.

Although 276 nm is the λ of maximum absorbance in the near-UV spectrum, it is by no means the λ maximum throughout the whole UV-Vis spectral region. Figure 3.17 shows a wavelength scan of the EP fluid protein from 820 - 190 nm. As shown in Figure 3.17, the λ maximum for the protein occurs at a wavelength of approximately 210 nm. This absorption arises from the backbone peptide chain of the protein molecule.¹¹²
Figure 3.15  UV-VIS absorbance spectrum of purified EP fluid protein. Sample was 0.87 mg/ml protein in PBS, pH 7.5. PBS was used as reference and to set 0 absorbance.

Figure 3.16  Absorbance vs. concentration for the determination of the protein molar absorptivity constant at 276 nm in PBS. (spectrum taken on Cary-5 spectrometer)
Amino Acid Analysis:

Amino acid analysis of the purified EP fluid protein was performed at two separate laboratories. On two occasions, AAA Laboratory of Mercer Island Washington USA was sent a purified protein sample, and on one occasion, the University of Michigan Protein Structure Facility, Ann Arbor Michigan USA was sent a purified protein sample. The individual and average results from these three analyses are shown in Table 3.6. The results are presented as percent molar composition for each amino acid residue (moles of amino residue/total moles for all amino acid residues), as well as in number of residues per protein based on a protein MW of 28,340 (last column). Only the University of Michigan analysis tested for the
presence of tryptophan. On one occasion, AAA laboratory analyzed for the presence of cysteine. Figure 3.17a is a pie chart displaying the percent composition of the fundamental amino acid groups. The pie chart was constructed from the average results of the three amino acid analyses.

Figure 3.17a EP fluid protein % molar composition of fundamental amino acid groups.

From the data a few generalization about the protein may be drawn. First, it may be said that the protein ought to be hydrophilic. Fifty seven percent of the molar amino acid composition is accounted for by charged or
polar amino acid residues with glutamic acid/glutamine, aspartic acid/asparagine and histidine being the amino acids present in the largest quantity. This finding comes as no surprise, considering the aqueous environment from which the protein originated, and the observation that it salted out of solution at a high (~70%) NH₄SO₄ salt concentration. By comparing the charged residues, it may be seen that there is a 6% excess of acidic residues over basic residues. Although the glutamic acid/glutamine, aspartic acid/asparagine ratios are not known, based on partial amino acid sequence data (vide infra) it appears that the residues are present in predominantly acidic form. The sequencing of 48 protein residues produced 11 aspartic acids to 3 asparagine residues (80% Asp), and 4 glutamic acid residues to 0 glutamines. This greater abundance of acidic residues results in an acidic protein with a pl value less than 7. The pl value for the protein was estimated to be 4.85 as determined by isoelectric focusing (Fig. 3.13).

The aliphatic residues are also well represented, accounting for 33% of the total residues. Although statistically this is not surprising (the aliphatic group is the largest of the amino acid groups), what is noteworthy is that glycine, the simplest amino acid, is not the dominant aliphatic residue. Instead, the more hydrophobic valine and alanine residues are present in greater quantity. The remaining aliphatics, leucine, isoleucine and glycine are present in approximately equal quantities.
Table 3.7 compares the amino acid breakdown of the EP fluid protein against that of soluble organic matrix proteins collected from other mineralizing species. Species from the phylum Foraminifera (protozoan), Echinodermata (sea urchin), and Mollusca (Nautilus), are represented in the table. In comparison to the matrix proteins, the EP fluid protein is deficient in its aspartic acid/asparagine and glycine residues, while being rich in histidine content. The ramifications of these differences are not clear at this time.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Average</th>
<th>AAA1</th>
<th>AAA2</th>
<th>UNich</th>
<th>sd</th>
<th>Res.Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Asx</td>
<td>11.92</td>
<td>11.00</td>
<td>13.63</td>
<td>11.11</td>
<td>1.49</td>
</tr>
<tr>
<td>2</td>
<td>Glx</td>
<td>14.12</td>
<td>12.15</td>
<td>14.43</td>
<td>15.77</td>
<td>1.83</td>
</tr>
<tr>
<td>3</td>
<td>His</td>
<td>11.58</td>
<td>11.27</td>
<td>14.08</td>
<td>9.38</td>
<td>2.36</td>
</tr>
<tr>
<td>4</td>
<td>Lys</td>
<td>5.59</td>
<td>5.79</td>
<td>6.22</td>
<td>4.77</td>
<td>0.74</td>
</tr>
<tr>
<td>5</td>
<td>Arg</td>
<td>2.30</td>
<td>2.06</td>
<td>1.35</td>
<td>3.50</td>
<td>1.10</td>
</tr>
<tr>
<td>6</td>
<td>Glu</td>
<td>4.79</td>
<td>4.65</td>
<td>5.20</td>
<td>4.51</td>
<td>0.37</td>
</tr>
<tr>
<td>7</td>
<td>Asp</td>
<td>2.31</td>
<td>2.31</td>
<td>2.11</td>
<td>2.50</td>
<td>0.19</td>
</tr>
<tr>
<td>8</td>
<td>Trp</td>
<td>1.16</td>
<td>1.16</td>
<td>0</td>
<td>2.83</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Tyr</td>
<td>5.64</td>
<td>5.92</td>
<td>5.57</td>
<td>5.43</td>
<td>0.25</td>
</tr>
<tr>
<td>10</td>
<td>Met</td>
<td>1.32</td>
<td>2.09</td>
<td>1.03</td>
<td>0.84</td>
<td>0.67</td>
</tr>
<tr>
<td>11</td>
<td>Ser</td>
<td>5.92</td>
<td>5.44</td>
<td>5.14</td>
<td>7.19</td>
<td>1.11</td>
</tr>
<tr>
<td>12</td>
<td>Cys</td>
<td>6.42</td>
<td>6.43</td>
<td>5.39</td>
<td>7.45</td>
<td>1.03</td>
</tr>
<tr>
<td>13</td>
<td>Ala</td>
<td>8.48</td>
<td>8.28</td>
<td>8.02</td>
<td>9.16</td>
<td>0.60</td>
</tr>
<tr>
<td>14</td>
<td>Val</td>
<td>7.38</td>
<td>8.02</td>
<td>7.59</td>
<td>5.94</td>
<td>1.10</td>
</tr>
<tr>
<td>15</td>
<td>Leu</td>
<td>6.27</td>
<td>5.73</td>
<td>7.43</td>
<td>5.65</td>
<td>1.01</td>
</tr>
<tr>
<td>16</td>
<td>Ile</td>
<td>5.31</td>
<td>5.67</td>
<td>6.50</td>
<td>3.77</td>
<td>1.40</td>
</tr>
<tr>
<td>17</td>
<td>Pro</td>
<td>3.14</td>
<td>3.20</td>
<td>3.17</td>
<td>3.04</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 3.6 Results of three amino acid analyses performed on purified EP fluid protein (filename: aaa.org).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 3.7 Comparison of the EP fluid amino acid breakdown against the amino acid breakdown of the organic matrix proteins from a protozoan, a sea urchin and another mollusc (filename: aaa.org).

Identification of the Protein as a Glycoprotein:

As mentioned in the Introduction, prior studies on the composition of both the extrapallial fluid and the shell reported the presence of glycoproteins. In order to determine the presence or absence of glycoprotein in the EP fluid of *Mytilus Edulis*, glycoprotein-specific staining of electrophoretic gels was employed. This investigation identified the major fluid protein --the protein targeted for purification and study-- as a glycoprotein. Figure 3.18 shows a 15% native PAGE gel which was loaded with duplicate samples of EP fluid protein, at the various stages of purification, and a duplicate loading of protein standards. When
electrophoresis was completed, the gel was sliced in half so that each of the gel halves contained one of the duplicate sample loadings. Next, the gel halves were stained by different methods. One half was stained indiscriminately for protein with Coomassie blue. The other gel half was stained with a thymol/H$_2$SO$_4$ solution, specific for the detection of glycoprotein. The left half of the gel (lanes 1 - 4), shows that the major EP fluid protein band develops the red/orange color indicative of glycoprotein. Of the standard proteins used in this experiment, cytochrome c, albumin, carbonic anhydrase, and transferrin (lanes 5 - 6), only the transferrin band stained in both gel halves. Transferrin is ~ 5% by weight carbohydrate. This analysis positively identifies the major protein component of the EP fluid as a glycoprotein.

The next question to be answered was: how much carbohydrate is bound to the protein? To answer this question, several purified EP fluid protein samples were quantitated for both protein and carbohydrate. The Bio Rad protein assay was used to determine protein concentration and the phenol/H$_2$SO$_4$ neutral sugar assay was used to determine carbohydrate concentration. Figures 3.19 and 3.20 show the composite standard plots of several individual assays. The values of mg carbohydrate/mg protein, obtained from the assays, were used to estimate the percent carbohydrate in the protein. Bovine serum albumin served as the standard in the protein assays and D-galactose was the standard in the carbohydrate assays.
Table 3.8 lists the results from 15 different sample preparations. Based on these results, it was determined that the EP fluid protein is composed of 7.25 % by weight carbohydrate.

Figure 3.18 15% native gel sliced in half and stained for glycoprotein (left) and indiscriminately for protein (right): Lanes: 1&10) peak G off anion Ex, 2&9) Raw fluid, 3&8) peak G off anion Ex, 4&7) peak G off size Ex, 5&6) protein Std's (same as figure 3.2)

| 1) 6.91% | 9) 8.12% |
| 2) 7.05% | 10) 6.43% |
| 3) 6.72% | 11) 5.91% |
| 4) 7.17% | 12) 7.33% |
| 5) 7.71% | 13) 6.09% |
| 6) 7.36% | 14) 5.82% |
| 7) 8.37% | 15) 7.63% |

| Ave | 7.25 % |
| SD  | 0.75   |
| RSD | 10.45 %|
| N   | 15     |

Table 3.8 Weight percent carbohydrate results from the analysis of 15 different EP fluid protein samples.
-N- vs -O- Linked Carbohydrate Side Chains:

The carbohydrate units of a glycoprotein are covalently linked to the peptide backbone via the side chain of certain amino acid residues. The carbohydrate/amino acid linkage may be either an N-linkage or an -O-linkage. An N-glycosidic linkage has the carbohydrate unit attached to the protein by the amide nitrogen of an asparagine residue. O-linkages have the carbohydrate attached via the oxygen of a serine, threonine, hydroxylysine or hydroxyproline residue. An easy way of distinguishing the type of carbohydrate attachment present on a given glycoprotein is to examine the alkaline stability of the carbohydrate/protein bond. Common O-glycosidic linkages (carbohydrate linked through serine or threonine) are readily cleaved when exposed to mildly basic conditions. The carbohydrate is released by a β-elimination mechanism; in the process, the serine and/or threonine residues are converted to α-aminopropenoic acid and α-amino-2-butenolic acid respectively. Both of these modified amino acid products absorb light at 240 nm and their detection may be used as evidence for O-glycosidic carbohydrate linkages.

A 0.3 mg amount of lyophilized EP fluid protein was dissolved and incubated in 0.1 M NaOH for 18 hours at room temperature. Simultaneously, a 0.3 mg amount of lyophilized protein was dissolved in DI H₂O and allowed to incubate for the same time period. After incubation, the NaOH digest was neutralized with 1.0 M HCl. Both samples were adjusted
Figure 3.19 Composite plot of carbohydrate standard calibration curve (filename: sugar.org).

Figure 3.20 Composite plot of protein standard calibration curve (filename: protein.org).
to give a 0.25 absorbance reading at 280 nm, when referenced against DI H₂O. A spectral scan (300 - 220 nm) of each sample was obtained. Figure 3.21 is a plot of absorbance vs wavelength for both samples as well as the H₂O reference. Figure 3.22 shows the difference spectrum (protein NaOH - protein H₂O). Figure 3.22 shows a strong absorption peak centered at 242 nm. This peak indicates the presence of the modified serine and/or threonine amino acid residues, and as such, indicates the presence of O-linked carbohydrate groups.

Investigation into the presence of N-linked high mannose carbohydrate moieties was conducted using the endo-β-N-acetylglucoaminidase H (Endo H) enzyme, and concanavalin A (Con A) affinity chromatography. Con A affinity chromatography resin specifically and reversibly binds α-linked mannose structures in N-linked glycoproteins. The ability of Con A to ubiquitously bind α-linked mannose has led to its broad and extensive use in glycoprotein purification procedures. However, when a sample of purified EP fluid protein was passed over a 0.5 x 5 cm chromatography column of Con A sepharose affinity resin (Pharmacia), the protein was not retained by the resin. This result suggests that the carbohydrate moieties are not composed of α-D-mannose.

The Endo H enzyme has the ability to hydrolyze the β1 - 4 bond between the GlcNAcβ1 - 4GlcHAc carbohydrate residues (GlcNAc = N-acetyl-D-glucosamine) on the inner core pentasaccharide of N-linked
Figure 3.21 Absorbance scans of NaOH digested EP fluid protein sample (---), DI H₂O/protein control (----) and DI reference (****). (spectra taken on Cary-5 spectrometer)

Figure 3.22 Difference spectrum produced by the subtraction of the DI H₂O/protein control sample from the NaOH protein digest.
glycans. To test for the effectiveness of EP fluid protein deglycosylation by Endo H, the procedure of Trimble and Maley was followed. Duplicate samples of 1 mg lyophilized purified EP fluid protein, and 1 mg of purified ovalbumin (Pharmacia) were dissolved in 50 mM citrate buffer, 1.2 mg/ml SDS, 0.1 M β-mercaptoethanol, and boiled for 90 seconds to denature the protein. Samples were then brought to 100 ug/ml Endo H and allowed to incubate for 22 hours at 37 °C. After incubation, all samples were analyzed by SDS-PAGE in order to determine if a shift in protein electrophoretic mobility (due to carbohydrate cleavage by Endo H) had occurred. Figure 3.23 shows a 15% SDS-PAGE gel of both ovalbumin and both EP fluid samples. As shown on the gel, the Endo H treated ovalbumin samples show a slight shift toward greater electrophoretic mobility (smaller MW), while the EP fluid samples show no shift in electrophoretic mobility. The average shift in MW (from calculated Rf values) for the ovalbumin samples is 3051 Daltons, which represents 6.67% of the total protein mass.

Although the shift in the MW of ovalbumin standard is high (ovalbumin is ~ 3.0% carbohydrate by weight), the fact that there is no observed shift between the Endo H treated and untreated EP fluid protein samples, indicates that the EP fluid does not contain Endo H accessible N-linked carbohydrate groups.

Although both the Con A and the Endo H experiments cast doubt on the existence of N-linked carbohydrate units, the evidence from these
experiments does not conclusively prove their non-existence. What may be concluded from these experiments is that the EP fluid protein's carbohydrate moieties are not likely to be polysaccharides containing peripheral mannose residues (Con A experiment). In addition, if the EP fluid protein does contain N-linked carbohydrate, it is Endo H resistant. In order to conclusively show the non existence of N-linked carbohydrate, enzymatic digestion with other endoglycosidases (Endo G, Endo F, PNGnase A, PNGnase F) should be investigated.¹¹⁹

Monosaccharide analysis showed the carbohydrate component of the glycoprotein to be composed of predominantly galactosamine, although lesser amount of glucosamine, fucose, glucose and mannose were detected.

![Figure 3.23 15% SDS PAGE Gel of the Endo H treated and untreated EP fluid protein and Ovalbumin Standard. Lanes: 1 & 9) MW weight standards (same as Table 3.5), 2) ovalbumin, 3 & 4) Endo H treated ovalbumin, 5 & 7) EP fluid protein, 6 & 8) Endo H treated EP fluid protein.](image-url)
N-Terminal Amino Acid Sequence and Trypsin Digest:

N-terminal amino acid sequence analysis and protein digestion with the protease enzyme trypsin was performed in order to obtain a protein fingerprint. These tests were conducted with the hope that they would aid in the identification of the protein through searches of protein data bases containing such information.

The N-terminal sequences were obtained from purified EP fluid protein samples, prepared for the analysis by two different methods. Both protein samples were purified from the extrapallial fluid as described earlier in this Thesis. Once purified, one sample was dialyzed against a solution of 10% (volume) glacial acetic acid (Fisher). The other sample was dialyzed against 5 mM \( \text{NH}_4\text{HCO}_3 \), pH 7.5, and then lyophilized.

From the sample prepared by acetic acid dialysis, a sequence of 20 amino acid residues was obtained. From the lyophilized protein sample, a sequence of 11 residues was obtained. Both sequences are shown below using the three letter amino acid identifiers. All of the X's in the sequence data represent unknown amino acid residues. These positions are perhaps occupied by amino acids which have been modified --possibly by glycosylation-- in such a way as to make them nonidentifiable as one of the 20 basic amino acid residues or they may be occupied by less common amino acids. The elution position of an unknown peak on an amino acid analysis chromatogram suggests phosphoserine, hydroxyproline or
γ-carboxyglutamic acid as possible candidates.

Asn-Pro-Val-Asp-Asp-X-X-Asp-Asp-Ala-Pro-Ile-Val-Glu-X-X-Asp-
N-terminal sequence of the EP fluid protein sample prepared by acetic acid dialysis.

Examination of the above sequence shows that aspartic acid is the most abundant residue and appears in what seems to be a repetitive pattern (adjacent aspartic acids separated by a pair of unknown residues). This pattern is similar to the sequence proposed as a CaCO₃ crystal template in the crystal growth by epitaxy model of shell formation. The abundance of aspartic acid implies that this area of the protein is acidic and as such should have the ability to chelate metal ions. All that can be said with certainty about the -X- residues is that they have been modified. Whether or not the modifications are similar on all of the X residues is unknown. Based on their location in the sequence, it seems probable that the modified residues would act in concert with the Asp residues in the performance of a particular biological function.

Asn-X-Val-Asp-Asp-His-His-Asp-Asp-His-
N-terminal sequence of the EP fluid protein sample prepared by lyophilization.

The N-terminal analysis of a lyophilized sample showed slight differences in the sequence data obtained. The question marks (?) which appear above certain residues represent positions where a discrepancy in
the data exists. What is interesting is that all of these positions correspond to an unknown residue in one of the two sequences. It was made clear by the analyst (Dr. P. Matsudaira) that the amino acid yields at these positions were greatly reduced in comparison with the amino acid yields for sequence position where the residues match. The residues assigned to the question mark (?) positions were the amino acids present in the largest amount; however, the sequence data for these sites was not clean (numerous peaks) and the yields were low (indicating a high degree of modification). The reason that the assignments are shown is that these residues may represent the root amino acids at the modified position.

Both of the above sequences were entered into several protein data bases.\textsuperscript{121} None of the searches produced a protein containing a compatible sequence. Mass fragments produced by the trypsin digestion were also entered into a trypsin digest data base.\textsuperscript{122} This search also proved negative. Therefore, it is concluded that the EP fluid protein is probably the first protein of its kind to be studied at this level of detail.

Protein Interaction with Calcium:

Since the EP fluid proteins may be involved in shell formation, an examination of the protein behavior in the presence of calcium was undertaken. Changes in protein electrophoretic mobility in a calcium-doped gel was examined. An electrophoresis experiment was conducted wherein
all solutions utilized in the production and running of the gel were brought to 10 mM Ca$^{2+}$ by the addition of CaCl$_2$ salt. The experiment resulted in the conversion of the single major EP fluid protein band into several more slowly migrating protein bands. This transition appeared to be triggered by calcium. Figure 3.24 is a picture of two 10% native PAGE gels labelled A and B. Gel A is an example of the electrophoretic behavior of the EP fluid protein (lanes 2 - 8) in the absence of calcium. The main protein fraction is shown to migrate with the fastest moving standard myoglobin. Gel B shows the electrophoretic behavior, of the identical EP fluid samples, in the presence of 10 mM Ca$^{2+}$. Other than the addition of CaCl$_2$ to the gel B electrophoretic solutions, all other aspects of the experiments were identical. Figure 3.24 shows that the protein samples in gel B did not enter the 10% resolving gel; instead, the protein remained trapped at the stacking/resolving gel interface. The standards in gels A and B are shown to have migrated approximately equal distances. Gel B indicates that calcium is interacting with the EP fluid protein in such a way as to severely inhibit its electrophoretic mobility. Figure 3.25 is a picture of a 5% native PAGE gel, in effect a giant stacking gel, brought to 10 mM Ca$^{2+}$. The giant stacking gel demonstrates that, in the presence calcium, an apparently homogeneous, rapidly migrating protein band is transformed into several, distinguishable, more slowly migrating species.

In native PAGE, the electrophoretic mobility of a protein is a function
of both protein size and protein charge. In typical electrophoretic experiments, negative charge will enhance electrophoretic mobility because protein migration proceeds from the cathode (-) to anode (+). From isoelectric focusing experiments it was determined that the EP fluid protein had a pl point of 4.85 (acidic). Therefore, the pH at which native PAGE experiments are carried out (pH 8.55) will result in a protein with a net negative charge. Protein interaction with calcium might cause some of the protein's inherent negative charge to be neutralized. If the net negative charge decreases, then electrophoretic mobility ought to decrease as well.

If the stoichiometry of protein-calcium binding occurred in differential amounts, than differential effects on mobility ought to be observed. Variable amounts of protein-calcium complexation, resulting in differential surface charge neutralization, is one explanation of the observed fractionation, and reduced electrophoretic mobility, of EP fluid protein run on calcium doped native PAGE gels.

Another possible explanation is that calcium is causing protomer formation. The presence of calcium may be inducing the formation of multimers from monomeric protein species. If this were the case, then the observed decrease in electrophoretic mobility might be due to an increase in protein size and not the result of a decrease in negative surface charge. To help solve this riddle, sedimentation velocity analysis was employed.
Figure 3.24  Electrophoresis in the absence (gel A) and in the presence (gel B) of 10 mM Ca\textsuperscript{2+}. All other aspects of the gels are identical. Lanes: (for both gels) 1) Protein Std's (from top to bottom, Transferrin (80,000); albumin (67,000); carbonic anhydrase (34,000); myoglobin (17,500)) 2-8) EP fluid protein.

Figure 3.25  5% native gel—giant stacking gel—10 mM Ca\textsuperscript{2+}. Gel demonstrates the resolution of the slowly migrating protein bands formed in the presence of calcium.
Sedimentation Velocity in the Presence of Calcium:

Sedimentation velocity analysis measures the speed at which molecules sediment from solution as a result of an applied centrifugal force. The rate of sedimentation is dependent on the molecular size and shape, among other things, but is unrelated to inherent molecular charge. Therefore, if the EP fluid protein-calcium interaction was resulting in protein protomer formation, then when placed under an identical centrifugal force, the higher molecular weight protomers ought to sediment faster than their monomer component. Figures 3.26 and 3.27 demonstrate that the protein-calcium interaction indeed results in protomer formation. Figure 3.26 is a plot of distribution \( g(s) \) vs. apparent sedimentation velocity \( s^* \) for three sedimentation velocity cells containing protein at different calcium concentrations. The three cells contain equal amounts of protein (0.9 mg/ml) and varying amounts of calcium (1 - 100 mM). Regardless of the \( \text{Ca}^{2+} \) concentration several high molecular weight species are observed. Figure 3.26 is a look at the sedimentation velocity cell containing 9 mg/ml protein in 20 mM Tris, pH 7.5, brought to 10 mM \( \text{Ca}^{2+} \) by the addition of \( \text{CaCl}_2 \). This sedimentation velocity analysis profile was integrated with a multiple Gaussian curve fitting program (Origin 3.5). Figure 3.27 shows the result of the multiple Gaussian curve fit. This plot clearly demonstrates the presence of at least five, distinguishable, protomer species having \( s^* \) values of ~ 4.0, 10, 15, 19 and 22.
Figure 3.26  Distribution $g(s)$ vs. $s^*$, the apparent sedimentation coefficient, for sedimentation velocity analysis of samples containing 0.9 mg/ml of protein and varying (1 - 100 mM) amount of Ca$^{2+}$. Solvent for all cell was 20 mM Tris, pH 7.5, 0.1 M KCl. (file 062295v gs100.org)

Figure 3.27  Distribution, $g(s)$ vs. $s^*$, the apparent sedimentation coefficient, for 10 mM Ca$^{2+}$ sedimentation velocity cell the Figure 3.26 analysis (file 062295v gs100.org)
The results displayed in Figures 3.26 and 3.27 pose several questions. Is protomer formation entirely calcium dependent? Are other metals capable of producing a similar effect? What is the stoichiometry of protomer formation? Is the process reversible?

To answer the question of whether protomer formation is fully calcium dependent and/or mass action related, a sedimentation velocity experiment was conducted wherein protein solution of various concentrations (0.1 mg/ml - 1.0 mg/ml) were analyzed in the absence of calcium. The purpose of conducting the experiment was to investigate whether protomer formation showed a protein concentration dependency (does the system display any association by mass action?). Figure 3.28 shows the profile of species distribution g(s) vs. apparent sedimentation coefficient s* for four different concentrations of EP fluid protein as analyzed by sedimentation velocity ultracentrifugation. Figure 3.28 demonstrates that in the absence of calcium, the vast majority of the protein sediments with an s* ~ 4.0. Generally speaking, proteins which sediment at this rate have a molecular mass in the vicinity of 60,000 (bovine serum albumin, MW 66,500, s_{20,w} = 4.3; hemoglobin, MW 68,000, s_{20,w} = 4.3). Therefore, it is highly unlikely that the peak at s* ~ 4.0 represents the protein monomer of MW 28,340. An EP fluid protein dimer (MW = 56,680) or, depending on molecular shape, perhaps a protein trimer (MW 85,020), are more probable candidates for the s* ~ 4.0 peak.
Whatever the species responsible for the $s^* \sim 4.0$ peak, it does not decompose at low protein concentrations. This observation suggests that the $s^* \sim 4.0$ peak is stable and likely represents the apo-form (no metal) of the protein in solution.

Figure 3.28 Sedimentation velocity analysis. $g(s)$ vs $s^*$ for four different EP fluid protein concentration. Solvent is 20 mM Tris, pH 7.5, 0.1 M KCl. (file 061395v3 gs100.org)

Figure 3.28 also shows that the system displays a slight degree of association by mass action. As protein concentration increases, larger protomeric species are formed ($>s^* \sim 4.0$). The 0.10 and 0.25 mg/ml samples show no higher order protomers (nothing $> s^* \sim 4.0$). The 0.5 mg/ml sample shows a peak at $s^* \sim 10$, which accounts for $\sim 5.8\%$ of the material in the cell. The 1.0 mg/ml sample shows the same $s^* \sim 10$ peak.
accounting for 7.2% of the total cell material. This observation implies that higher order protomer formation (>\(s^* - 4.0\)) is not an entirely calcium dependent phenomena. Protomer formation may be considered, at least in part, a function of protein concentration. In Figure 3.28, the location of the protomer peak at \(s^* \approx 10\) suggests that the stoichiometric relationship between the \(s^* - 4.0\) species and the \(s^* - 10\) species is greater than 2 to 1 (more than two \(s^* - 4.0\) species are combining to create the larger \(s^* - 10\) species). However, without knowledge of the molecular weight or the frictional coefficient for either of these molecules, extraction of accurate information about the molecular size is not possible.

Figure 3.29 is an overlay of the sedimentation profile of a 1.0 mg/ml EP fluid protein sample analyzed in the absence of calcium, with a 0.9 mg/ml protein sample analyzed in buffer brought to 10 mm Ca\(^{2+}\). This plot demonstrates, that although protomer formation is not entirely calcium dependent, it is strongly influenced by calcium. The 10 mM Ca\(^{2+}\) sample shows that ~ 60% of the \(s^* - 4.0\) material is converted to larger protomers. Although it may be misleading to compare the sedimentation coefficients of molecules without knowing pertinent information such as molecular size, shape and frictional coefficient, the fastest sedimenting protomer in Figures 3.27 and 3.29 (\(s^* \approx 22\)) is sedimenting at a rate on par with proteins such as horse spleen apoferritin, (MW 465,000, \(s_{20,w}\) 17.8), squid hemocyanin, (MW 612,000, \(s_{20,w}\) 19.6), and bovine liver glutamate dehydrogenase, (MW
Comparison of Sedimentation Velocity Profile for ~ 1.0 mg/ml EPF Protein Sample w/ and w/o Ca$^{2+}$

Figure 3.29 Overlay of 1.0 mg/ml EP fluid sample no calcium, and 0.9 mg/ml EP fluid sample 10 mm Ca$^{2+}$. Experimental conditions are the same as stated in Figure 3.28 and 3.27 respectively (filename: 061395v3 gs100.org).

Figure 3.30 Sedimentation Velocity analysis displaying the increase in the peak width at half height for protomers with increasing molecular size (filename: 061395v3 gs100.org).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
1,015,000, $s_{20,w}$ 26.6). It is plausible that the $s^*$ ~ 22 protomer is composed of 20 monomer subunits or more.

Figure 3.30 displays an interesting trend common to the sedimentation profile of all the protein-calcium samples. As the protein protomers increase in size, distribution in the measurement of the apparent sedimentation coefficient ($s^*$) also increases (as the protomers get larger their respective peak width's at half height are seen to increase). This observation is counterintuitive. All other things being equal, as a molecule increases in size its rate of diffusion ought to decrease, resulting in narrower peak width.

According to the analysis by Stafford, the sedimentation radial diffusion constant is related to peak width at half height ($W_{1/2}$) by equation 3.1.

$$D = \frac{(W_{1/2} \cdot s^*)^2}{(2 \cdot t)}$$

(3.1)

Where:

- $W_{1/2}$ peak width at half height (in Svedberg units x $10^{-13}$)
- $s^*$ apparent sedimentation coefficient ($s^* x 10^{-13}$ s)
- $D$ radial diffusion constant (cm/s$^2$)
- $t$ time (seconds)
- $r$ radial distance between molecule and rotor center (cm)

The radial diffusion constant is related to the molecular weight by equation 3.2.

$$MW = \frac{(R \cdot T \cdot s^* / D \cdot (1 - \nu \rho))}{(1- \nu \rho)}$$

(3.2)

Where:

- $MW$ molecular weight
- $R$ gas constant (8.3144 x $10^9$)
- $T$ temperature (K)
- $s^*$ apparent sedimentation coefficient ($s^* x 10^{-13}$ s)
- $D$ radial diffusion constant (cm/s$^2$)
- $\nu$ protein partial specific volume (ml/g)
- $\rho$ solvent density (g/ml)
From equation 3.1, it may be concluded that $W_{1/2}$ is directly proportional to the radial diffusion constant. Equation 3.2 shows that the radial diffusion constant ($D$) is inversely proportional to molecular weight. Therefore, equations 3.1 and 3.2, taken together, indicate that as a molecule's molecular weight increases, its radial diffusion constant ought to decrease, and as a molecule's radial diffusion constant decreases, its peak width at half height ought to decrease as well. This is not the case for the EP fluid protein protomers.

Converse to theory, as the EP fluid protein protomers grow in size, their $W_{1/2}$ are shown to grow in magnitude. One explanation for this occurrence, is that the larger protomer peaks ($>s^* - 4.0$), and perhaps the $s^* - 4.0$ peak as well, do not represent homogeneous molecular species. Instead, these peaks represent families of protomers of approximately equal size showing a broad distribution in their sedimentation rates.

Sedimentation rates are highly dependent on molecular shape as well as molecular size. Increasing peak width with increasing molecular size, may be explained as representing an increase in the diversity of molecular shapes being formed (increasing protomer size is accompanied by an increase in the number and type of structural isomers being formed). The idea of structural heterogeneity in the assembly of the EP fluid protomers, is consistent with certain aspects of the monomer unit which have already been discussed. The monomer showed microheterogeneity in it's molecular
weight by mass spectrometry, isoelectric point, N-terminal sequence, as well as showing noticeable fine structure on the sedimentation velocity peaks (Fig. 3.28).

All of these findings taken together indicate that the major EP fluid protein component may not be a homogeneous species; but rather, the fluid protein is composed of a family of closely related, microheterogeneous forms. In the presence of calcium, the various protein forms are able to assemble into higher order protomers. The stoichiometry of protomer formation is finite, leading to species which are distinguishable and largely uniform in molecular size. However, although there is uniformity in the stoichiometry of assembly, there appears to be several different modes by which the proteins may combine (smaller protomers may combine to form larger protomers in different ways). The differential modes of assembly result in the formation various structural isomers. The structural isomers are largely homogeneous in mass but heterogeneous in 3-dimensional shape.

Circular Dichroism:

Circular dichroism (CD) is a measure of the ability of molecule to differentially absorb left and right circularly polarized light. A CD spectrum is a trace of this differential absorption plotted as a function of wavelength. By convention, CD absorbance is determined by subtracting the magnitude of absorption of right circularly polarized light (A_r) from that
of left circularly polarized light ($A_L$).

$$CD = A_L - A_R \quad (3.3)$$

For macromolecules, CD spectra arise from the asymmetric distribution of the molecule's chromophoric constituents. Circular dichroism has proven to be a useful technique for the empirical determination of protein secondary structure.\textsuperscript{127} Certain fundamental protein secondary structural motifs (α-helix, β-sheet, random coil) are known to give rise to characteristic CD absorbance profiles. By comparing the CD spectrum of a protein against the CD spectra of peptides or proteins of known secondary structure, insight into plausible secondary structures may be achieved.

The CD spectrum of the EP fluid protein was deemed to be of value for two reasons. For the reasons stated above, it was hoped that the CD spectrum would give insight into the protein's secondary structure. In addition, CD analysis was performed to qualitatively determine whether or not discernable structural changes accompanied calcium binding to the protein.

Figure 3.31 shows the CD spectra of EP fluid protein in the presence and absence of calcium. The figure clearly demonstrates a significant change in the CD spectrum of the protein when it is allowed to interact with calcium. What's more, when compared against the CD spectra produced by
common protein secondary structures (Figure 3.32), it may be inferred, that in the absence of calcium the EP fluid protein produces a CD spectrum which resembles that of a random coil (low degree of order). Conversely, in the presence of calcium, the EP fluid protein CD spectrum more closely resembles the spectra produced by peptides of more defined structure (α-helix and/or β-sheet). The increase in (+) CD absorbance at 192 nm, as well as the increase in (-) CD absorbance 210 - 230 nm region give this claim validity. The magnitude of the structural rearrangement, and a possible quantitation of the secondary structural motifs involved, may be obtained by the semi-empirical method of singular-value decomposition (SVD). The software necessary to do the spectral deconvolution by the SVD method has only recently been acquired. Therefore, results of SVD analysis are not available at this time. If the SVD results become available prior to the publication of this thesis, they will be reported in an addendum to the main body of the thesis.
Figure 3.31 CD spectra of a 1.07 mg/ml protein solution 20 mm MOPS pH 7.5, 0.1 M KCl in the absence (----), and in the presence (solid line) of 10 mm Ca²⁺.

Figure 3.32 The CD for various secondary structures: α-helix (---), antiparallel β-sheet (-- --), and random coil (×××××).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Investigation of Possible Calcium-Binding Moieties:

With the discovery of the EP fluid protein/calcium interaction, a search for possible calcium-binding moieties was started. As mentioned in the Introduction, sulfate and phosphate functionalities have been identified on soluble matrix proteins of some shell-forming species. What's more, it is carbohydrate-bound sulfate groups that have been suggested as being sites of crystal nucleation in the ionotropy theory. Sialic acid residues on glycoproteins are known to contribute to charge microheterogeneity and, in addition, sialic acid residues are known to bind calcium. Knowing that sulfate, phosphate, and sialic acid functionalities were all possible calcium binding groups, these three functional groups were targeted for identification.

The first experiment performed was Fourier Transform infrared spectroscopy (FTIR). A protein FTIR spectrum was obtained in order to determine if absorption bands typical of $\text{SO}_4^-$ and $\text{PO}_4^-$ residues could be detected. Several milliliters of purified protein (in PBS) was dialyzed against several changes of distilled water and then lyophilized. Lyophilized protein was crushed into a KBr pellet and analyzed by FTIR.

The infrared spectra of tetravalent sulfur compounds containing a sulfur-oxygen linkage show strong absorptivities in the 1000-1250 cm$^{-1}$ region of the IR spectrum. The absorption is due to the strong $\text{S}=\text{O}$ stretching vibration. The intensity of these bands is proportional to
concentration and band absorptivities are in the vicinity of 300 L/mol cm. Specific ester sulfate absorption bands (RO-SO$_3^-$) are known to occur between the ranges of 1220 - 1315 cm$^{-1}$ and 1050 - 1140 cm$^{-1}$.

Phosphate compounds show absorbance in this spectral region as well. The P = O stretching vibration gives rise to a strong band with specific sub ranges for the (R-PO$_3^{2-}$) group being in the 970 - 1125 cm$^{-1}$ and 960 - 1000 cm$^{-1}$ region of the spectrum. The other absorber in this spectral region which is present in proteinaceous material is the C-N bond.

Figure 3.33 shows the protein FTIR spectrum of the sulfate/phosphate region. Several bands appear in the 1300 - 1040 cm$^{-1}$ region. The two large bands centered at 1652 cm$^{-1}$ and 1529 cm$^{-1}$ are due to C=O stretching and to N-H deformation respectively. In comparing the EP fluid protein spectrum against that of purified proteins, albumin and ovalbumin (a glycoprotein), as well as carbohydrates glucose and galactose, several of the bands in the fluid protein spectrum were found to be unique (1040, 1081, 1218, 1229 cm$^{-1}$). The FTIR experiment was viewed as a positive result in search for sulphate and/or phosphate protein bound moieties. With this positive result, more exacting methods for the determination of the presence or absence of these functional groups were sought.

Sulfate Determination by EDX Analysis:

The next examination performed was energy dispersive x-ray analysis
Figure 3.33 FTIR spectrum showing the sulphate/phosphate absorption region of EP fluid protein

by scanning electron microscopy. A solid sample of purified EP fluid protein (lyophilized as described in FTIR section), was analyzed under an Amray 1000 scanning electron microscope. The bombardment and excitation of a protein sample with the electron beam of the electron microscope, is used to induce the emission of x-rays by the elements that make up of the sample. The detection and evaluation of the emitted x rays may be used to positively identify the elemental composition of the protein sample. Figure 3.34 is a printout of the results for the EDX analysis of the EP fluid protein sample. Figure 3.34 shows that in addition to carbon, oxygen and nitrogen,
(the primary constituents of protein) the only other element detected is sulfur.

Although EDX results are only marginally quantitative, the analysis does indicated an abundance of sulfur in the protein sample. From the amino acid analysis it may be estimated that the protein contains 2 methionine residues and 3 cysteine residues per protein molecule. Contribution from the methionine and cysteine residues leads to an EP fluid protein sulfur content of approximately 6 parts per thousand. The detection limit for the EDX analysis is in the part per thousand range. The magnitude of the EDX sulfur signal, coupled with the quantity of sulfur containing amino acids makes it difficult to determine if the methionine and cysteine residues are fully responsible for the EDX signal. It is possible that some of the EDX signal is arising from sulfated carbohydrate residues.

Figure 3.34  EDX analysis of a lyophilized EP fluid protein sample. The results show an abundance of sulfur present in the protein.
Phosphate Analysis by Phosphate Assay and $^{31}$P NMR:

In order to analyze for the presence of protein bound phosphate, the phosphomolybdate assay of Buss and Stull$^{105}$ was employed. The assay operates by converting protein bound phosphate into inorganic phosphate. Inorganic phosphate is then allowed to react with molybdenum to form a phosphomolybdate complex. The phosphomolybdate binds to the dye malachite green creating a chromophore, the absorbance of which (660 nm) may be used for phosphate quantitation. Although this method is reported to have the ability to accurately detect phosphate in quantities as low as 0.2 nmol, there was difficulty experienced in the attainment of precise measurements for the phosphate standards. In addition, EP fluid protein samples showed a high degree of variation in their analytical results. One thing that was consistent was that EP fluid protein phosphate estimates were always extremely low, often negative, and --at best-- were at the threshold limit of detection for method. Therefore, an alternate means of phosphate measurement was sought.

Phosphate NMR was used in an attempt to identify protein bound phosphate groups. Unfortunately, like its phosphomolybdate assay predecessor, the analytical results obtained by NMR were inconsistent, and --at best-- challenged the limits of detection for the method. Figure 3.35 shows the $^{31}$P NMR spectrum of an 8 mg/ml protein sample. The spectrum does show a single phosphorus peak with a chemical shift of 1.53 ppm, as
compared to a $\text{H}_3\text{PO}_4$ acid external standard. When calcium was added to this protein sample, the phosphorus peak shifts to 1.39 ppm. The fact that the phosphorus peak showed a slight change in chemical shift may indicate that it is associating with the calcium. However, when two other protein samples of equal concentration were analyzed, no phosphorus signal was observed.

Due to the large quantity of protein which analysis by NMR demands, and the difficulty experienced in obtaining precision in sample measurements, this method was abandoned.

![Figure 3.35 NMR spectrum of an 8 mg/ml EP fluid protein sample. Sample is in PBS, pH 7.5.](image)
Sialic Acid Determination:

Determination of the presence or absence of sialic acid functionalities on the EP fluid protein was achieved by the method of Aminoff. Four purified EP fluid protein samples and two raw fluid samples were analyzed for sialic acid. All samples tested negative for the presence of sialic acid. Figure 3.36 shows a composite plot of standard curves produced from two separate sialic acid assays. The results of the EP fluid protein analyses indicate that there are no sialic acid moieties present on the EP fluid protein.

Figure 3.36 Composite plot of the sialic acid standards used in two separate sialic acid assays.
Determination of Copper:

Prior studies on EP fluid protein, have produced results which indicate that some EP fluid proteins are copper containing metalloproteins. To examine whether or not the protein of this thesis work was a copper-containing protein, the protein was analyzed by electron paramagnetic resonance spectroscopy (EPR). Figure 3.37 shows the EPR spectrum of a 1 mM sample of purified EP fluid protein. The spectrum shows a signal characteristic of Cu$^{2+}$ have values of $g_{\perp} = 2.038$ and $g_{||} = 2.216$ as measured against a Varian strong pitch standard. The hyperfine structure seen on the $g_{\perp}$ absorption peak is suggestive of nitrogen ligation. The magnitude of splitting between the hyperfine lines is $-15$ G. The observed EPR indicates that the protein is indeed a copper metalloprotein. However, as was the case with some of the other studies discussed in this thesis, examinations into the presence of copper in the protein were mired with inconsistency.

Several attempts were made to determine the Cu-protein stoichiometric ratio. Protein concentrations were determined by protein assay (Bio Rad), and Cu was quantitated by atomic absorption (AA). Figure 3.38 shows a typical Cu standard addition calibration curve, used for the quantitation of Cu in protein samples. Several different protein preparations were analyzed for Cu content and the Cu-protein stoichiometries obtained varied from a high of 1.28 Cu/protein to a low of 0.38 Cu/protein.
Activity for the Cu-containing enzyme tyrosinase has been reported in previous studies on EP fluid protein material.\textsuperscript{31,32} Two assays designed to indicate tyrosinase activity (one which measured O\textsubscript{2} consumption with an O\textsubscript{2} electrode,\textsuperscript{132} and a color-ometric assay which measured the production of dopaquinone),\textsuperscript{133} were performed on purified EP fluid samples. Both produced negative results for tyrosinase activity.

An assay for lysyl oxidase activity was also attempted.\textsuperscript{134} Lysyl oxidase is the copper containing enzyme involved in the production of interprotein crosslinks in vertebrate connective tissue.\textsuperscript{135} Assay of the purified EP fluid protein proved negative for lysyl oxidase activity; however, lysyl oxidase activity was noted in assays of raw EP fluid. Due to time constraints, the fluid component responsible for the lysyl oxidase activity was not pursued.

In conclusion, the observed variability in the Cu-protein stoichiometry, as well as the role that copper atom plays in protein function are not understood at this time.
Figure 3.37 EPR spectrum of a frozen 1mM purified EP fluid sample.

Figure 3.38 Atomic absorption Cu standard addition calibration curve used for the quantitation of Cu in protein.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
CHAPTER IV

Discussion

In this section we will summarize the information gathered on the Extrapallial fluid of *Mytilus Edulis*, and put forth a hypothesis on the functional role of the major EP fluid protein which is the focus of this study.

Some interesting things may be said about the EP fluid itself. Although not new information, this study reinforces previous work on the EP fluid\(^{31-33}\) in finding that it is rich in macromolecular content (~ 4.00 mg/ml in both protein and carbohydrate). Throughout the course of a year, variation in the fluid protein concentration was noted (Fig. 3.1), with higher protein levels being reported for the winter months. Studies have shown that the rate of shell deposition is greater during the summer months than during the winter.\(^{21}\) Greater involvement in shell deposition, perhaps leading to greater incorporation or adhesion of the protein to the surface of the growing shell, may explain the lower fluid protein levels. Another possibility is that the cold winter temperatures necessitate a higher fluid protein concentration to combat fluid freezing. Proteins which function specifically as antifreeze are known to exist\(^{136}\) and the intertidal habitat of *Mytilus Edulis*...
does allow it to be exposed to the extreme variations in temperature inherent to New England.

EP fluid carbohydrate levels correlate with fluid protein levels. That the correlation is the result of the presence of glycoproteins in the EP fluid, is partly correct. Analysis has shown the major EP fluid protein to be a glycoprotein; composed of ~ 7.25% carbohydrate. However, this small amount of protein-bound carbohydrate accounts for only ~ 5.0% of the total carbohydrate within the fluid. Therefore, whatever the biological mechanism(s) which operate to control fluid protein and carbohydrate content, this study suggests that they may operate in accord. Both protein and carbohydrate are components of the organic matrix. Beyond determining the EP fluid carbohydrate concentration, the carbohydrate component was not studied in depth.

Although variation in fluid protein levels was noted, the inventory of fluid proteins remained consistent from month to month, year to year, throughout the four year course of this study. Electrophoretic gels displayed a fluid protein distribution which remained regular, and consistently showed a single, dominant, protein band (~56% of total). Inventories performed on soluble organic matrix (SOM) proteins have reported numbers which vary from a single component\textsuperscript{137} to several discrete protein components in the supernatant of EDTA dissolved shells.\textsuperscript{138} However, regardless of whether the SOM is homogeneous or heterogeneous
in protein content, there does appear to be a limited number of prevalent protein species. Figure 4.1 shows an SDS gel of the SOM material from nine different species representing the three major classes of mollusca. Lane 14 of Figure 4.1 shows that the SOM from *Mytilus Edulis* contains two prominent protein bands with molecular weights of approximately 18,000 and 29,000. Although a direct comparison of Figure 4.1 with the gels of this study is of questionable validity (the study represented in Figure 4.1 was done some time ago (1977) under significantly different experimental conditions) it is apparent that an EP fluid protein/SOM protein comparative study ought to be informative. In order to make a valid comparison of *Mytilus Edulis* EP fluid and SOM protein, and perhaps answer more definitively questions about the ultimate destination and biological function of EP fluid proteins, a thorough, contiguous examination of EP fluid and SOM proteins needs to be performed.

This study demonstrates that the major EP fluid protein is a glycoprotein; what’s more, the major fluid protein was the only fluid protein positively identified as such, the other fluid proteins staining negative for carbohydrate. The major protein was shown to be microheterogeneous in mass, with molecules ranging in size from 27,000 to 28,960 mu (Fig. 3.12), as well as microheterogeneous in pi’s ranging from 4.66 - 5.02 (Fig. 3.13). The glycosidic connection between protein and sugar was shown to be partially, if not fully, an -O- glycosidic linkage. N-acetyl-D-galactosamine
Figure 4.1 SDS PAGE 5-25% acrylamide gradient gel of the soluble fractions of mollusc organic matrices. Lanes: 1) Standard Proteins (from top to bottom) A, bovine serum albumin (MW 65,000); B, ovalbumin (MW 45,000); C, inactivated trypsin (MW 23,800); D, lysozyme (MW 14,000); 2) Nautilus pompilius; 3) Lottia gigantea; 4) Strombus costata; 5) S. gigas; 6) Crassostrea iridescent; 7) C. iridescent (non-myostracal layers); 8) C. gigas (non-myostracal layers); 9) Mercenaria mercenaria; 10) Mytilus californianus (inner prismatic layer); 11) M. californianus (outer prismatic and nacreous layers); 12) M. Californianus (outer prismatic layer); 13) M. californianus (all shell layers excluding myostracum); 14) Mytilus Edulis (all shell layers excluding myostracum); 15) Protein Standards (same as 1). (Reference 133)

was determined to be the primary monosaccharide. IR and EDX analyses suggest that the carbohydrate residue may contain sulfate functionalities (Fig. 3.33, 3.34). Carbohydrate was determined to constitute ~ 7.25% of total protein mass (Table 3.8).

Even for the most studied of mineralized tissues, with the exception of human bone, a complete inventory of the constituents of the organic matrix is not available, let alone sufficient detail on the functions which the various constituents perform before, during and after mineralization.
However, when comparing what is known about the macromolecules of the organic matrix with the information presented here on the major EP fluid protein component, several similarities are noted. It is generally accepted that the soluble organic matrix (SOM) is composed of a heterogeneous mixture of proteins, with the major constituent of material thus far studied being reported as acidic glycoproteins.\textsuperscript{1,12,15,43} The carbohydrate moieties of these glycoproteins have received little study. However, a study of SOM glycoproteins from the clam \textit{Mercenaria mercenaria} suggests the existence of glycosubstituted serine residues (O-links).\textsuperscript{139} The claim was based on the ability to functionalize serine residues via a $\beta$-elimination reaction, implying that the serine residues have an ester linked moiety assumed to be carbohydrate. Although the study was not a conclusive nor a thorough investigation of the carbohydrate, the results are consistent with those reported here for the EP fluid protein-carbohydrate linkage. The extent of glycosylation and the type(s) of monosaccharide units present on SOM glycoprotein has also received little detailed study; however, of the examinations which have been performed, several report the presence of sulphated carbohydrate moieties. In two studies the sulphated residues were identified as hexosamines.\textsuperscript{140} These sulphated carbohydrate residues are thought to be important in the binding of calcium to the matrix.\textsuperscript{80} Reports of the presence of sulphated hexosamines are in line with carbohydrate residues of the EP fluid protein. N-sulphated-galactosamines
are known to exist on several extracellular calcium-binding proteins, such as clotting factor proteins. In addition, there are other -O- linked glycoproteins known to play structural or support roles elsewhere in nature.

Another, common feature of the SOM acidic glycoproteins is that they are rich in aspartic acid and glutamic acid. Whether these amino acids are present in the acidic or amine forms seems to be species dependent, with reports ranging from the residues being fully aminated to being fully acidic. At first glance, comparison of the EP fluid protein amino acid composition against that of soluble organic matrix proteins from other mineralizing animals (Table 3.7), does not appear to show a good resemblance. Although aspartic acid and glutamic acid are the most abundant residues in the EP fluid protein, they do not occur in as large a quantity as that noted for SOM proteins. Based on EP fluid protein amino acid sequence data, both aspartic acid and glutamic acid seem to be present in predominantly acidic form; however some asparagine residues were found. Something that must be noted is that the data on soluble matrix proteins presented in Table 3.7 is representative of the amino acid composition of an undetermined number of matrix proteins. The analysis of the EP fluid protein represents the amino acid composition of a more homogeneous protein sample. Therefore, direct comparison of the data must be done with caution. Also, as is apparent is Figure 4.1, different
animals are likely to have matrices composed of an undetermined number of different proteins. Therefore, in order to confidently compare the EP fluid protein's amino acid composition to SOM protein, data on the amino acid composition of the purified SOM proteins from the shell of *Mytilus Edulis* need to be obtained.

The same argument may be made for the comparison of the EP fluid protein N-terminal amino acid sequence data against sequence data of SOM proteins. The SOM sequence which has been given the most attention is the theorized sequence:

-Asp-X-Asp-X-Asp-X-Asp-X- \( (X = \text{Glycine or Serine}) \)

This amino acid arrangement is based on results collected from the amino acid analysis of SOM proteins from several mineralizing species after a 48 hour hydrolysis with 0.25 M acetic acid at 108 °C. This hydrolysis procedure is reported to cleave the peptide chain on either side of the aspartate residue. Based on the amino acid analysis % yield of aspartate, glycine and serine produced after the above hydrolysis procedure, the (-Asp-X-) repeat sequence was hypothesized.\(^{78}\) This particular sequence is thought to act as a template for CaCO\(_3\) crystal growth by epitaxy.\(^{78}\) While this exact sequence is not observed in the EP fluid protein data, a similar sequence does appear. As shown on page 78, the following stretch of
sequence was obtained from an EP fluid protein sample.

\[-\text{Asp-Asp-X-X-Asp-Asp-X-X-Asp-} (X=\text{unknown amino acid})\]

In the case of the SOM protein, the glycine residue is thought to provide the ideal distance for the complexation of calcium (Fig. 1.4). If this is true, then the same cannot be true for the EP fluid protein. The distance between the carboxyl groups on adjacent aspartic acid residues would be too short, while the gap created by adjacent unknown residues would be too large. Regardless of the differences in the spacing of the aspartate residues, the stretches of sequence are both acidic and as such have the potential to function as calcium chelators.\textsuperscript{139} A comparison of the EP fluid protein N-terminal amino acid sequence with those of protein sequence data bases\textsuperscript{121} produced no matches. The negative result of these data base searches indicates that the EP fluid protein is unique, with no other protein of similar sequence having yet been reported. When comparing the two above sequences, it should be kept in mind that the EP fluid sequence is the actual protein sequence as determined by an amino acid sequencer. The SOM sequence is a theorized sequence based on the results of an amino acid analysis. The fact that two relatively similar aspartic acid sequences are not present in sequence data banks, but seem to occur in the SOM proteins and now in the EP fluid protein, lends substantial validity to notion that the EP
fluid protein is a matrix protein. This idea is further justified by the behavior of the protein in the presence of calcium.

Data from ultracentrifugation, circular dichroism, and electrophoresis, (Figs. 3.29, 3.31 and 3.24) demonstrate that the protein undergoes major structural changes when allowed to interact with calcium. Sedimentation velocity ultracentrifugation reveals that in the presence of calcium, the EP fluid protein readily forms higher order protomers (Fig. 3.29). Protomeric formation followed, what appeared to be, a periodic stoichiometric assembly resulting in the production of a finite number of larger aggregates (Fig. 3.27). Results from the electrophoretic analysis of the protein, run on calcium-doped gels, support the results of the sedimentation velocity study. On native gels, a single protein band was observed to transform into several, distinguishable, more slowly migrating protein bands (Figs. 3.24, 3.25). The assembly of the protein into higher order protomers of discrete size was once again apparent. Electrophoretic analysis also showed that protomer formation was fully reversible. In addition, magnesium and manganese were also capable of inducing protomer formation (results not presented).

What was not immediately clear were the "tell tale" signs of the assembly mechanism's complexity. From inspection of the species distribution profiles of the sedimentation velocity experiment, one possible conclusion is that the products of protomer formation, while largely

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
homogeneous in size, are heterogeneous in form. The calcium enhanced protein aggregation may be resulting in the formation of protomers of similar size having a number of different arrangements (based on the analysis of peak width at half height each protomer appears to form as several structural isomers, Fig. 3.28). Hypothetically, if the protein is a structural protein eluted into the fluid to complex calcium, upon binding calcium the protein may be reticulating into a three dimensional network. If this were indeed the case, then the protomer network may be designed to assemble into structurally isomeric network fragments. Figure 4.2 is a two-dimensional cartoon of what may be happening. In the drawing each ellipse (labelled A) represents the native protein in the absence of calcium ($s^* \sim 4.0$). As mentioned in the Results section, based on the magnitude of the apparent sedimentation coefficient, this apo-protein species is likely to be a multimer of the 28,340 MW monomer. In the discussion that follows this $s^* \sim 4.0$ apo-protein will be referred to as A. The addition of calcium somehow enhances A to assemble into protomers. In the discussion that follows, these higher order protomers of A will be referred to as B, C, D, and E respectively, in order of increasing size. From experimental evidence it is apparent that only certain stoichiometric combinations of protein are stable; what's more, based on sedimentation velocity data it appears as though the first step in protomer formation results in the creation of
something larger than a dimer ($nA \neq B$, $n > 2$). The basis for this claim is
the large increase in the sedimentation coefficients upon going from $A$ ($s^* \approx 4.0$) to $B$ ($s^* \approx 10$). The reasons for this are unknown; and unfortunately, the exact stoichiometry of formation is also unknown. For the purposes of
Figure 4.2, $B$ was arbitrarily chosen to be a tetramer. For the purposes of
discussion, we assume that the three products shown are the only stable $B$
species formed. These tetrameric products will have the same mass and on
the native gels used in protomer analysis ought to migrate together ($Ca^{2+}$-
doped gels were only 5% acrylamide and thus the proteins experienced minimal molecular sieving. However, due to their different shapes, they will sediment at slightly different rates. Differential sedimentation rates will lead to the observed band broadening in sedimentation velocity analysis. The three B products may combine to form 6 different (C) octomers etc...

As the larger protomers form, the number of structural isomers increases. While the model in Figure 4.2 is likely to be a simplification of actual protomer formation, it does provide one reasonable explanation of the data.

Another explanation to the observed increase in peak width with increasing molecular weight is that the phenomenon is being caused by the disruption of the equilibrium between the various protomer species. From the centrifuge data it seems clear that species A is the building block from which the higher order species (B, C, D, E) are formed. In a calcium containing solution, the protomers of adjacent size are in equilibrium with one another. For example, in calcium containing solution protomer C would be in equilibrium with both protomer B and protomer D.

\[
\begin{align*}
B + B & \rightleftharpoons C & K_D &= [C]/[B]^n \\
C + B & \rightleftharpoons D & K_D &= [D]/[C][B]
\end{align*}
\]

Due to the differential sedimentation rates of the various protomers, as the sedimentation analysis progresses the various protomers become
separated from one another. The isolation of an individual protomer will cause an equilibrium imbalance between that given protomer and its size neighbors. Again, using protomer C as an example, from the centrifuge data it is observed that C sediments at a rate of ~ 15 s. This sedimentation rate is different from that of the two protomers with which C is in equilibrium (B, s ~ 10 and D, s ~ 19). As the analysis advances, C will become separated from both B and D. The separation will lead to a disruption in the equilibrium between the protomers. To counter the equilibrium imbalance, protomer C will breakdown in an attempt to reestablish equilibrium with B and well as build up to reestablish equilibrium with D. These occurrences will result in the conversion of molecules which began the analysis sedimenting at a rate of s ~ 13, into species which migrate at a rates of s ~ 10 and s ~ 19 respectively. Consequently, the drive toward reestablishment of equilibrium will result in the broadening of the sedimentation profiles of the individual protomers.

There are numerous, intriguing questions which remain unanswered. What is the stoichiometry of protomer formation? Why are only certain stoichiometric combinations stable? Why does there appear to be an upper limit in the size of the observed protomers? How does the protein bind calcium? Why is calcium so effective at promoting protomer formation?

Presently, there is insufficient data to begin speculation on the above posed questions. However, other topics may be addressed. The protein
interacts with calcium and in so doing is prompted to form protomers. How is this accomplished? Are calcium ions being shared by protein molecules? If so, is the joint binding of calcium acting to bridge the proteins together? Or, is the binding of calcium causing protein structural rearrangement, which results in the ability of the molecule to associate with other proteins in a stable manner (H-bonding, hydrophobic interactions)?

Referring back to Figure 3.26, it is observed that the species distribution profiles from the sedimentation velocity analysis of three 0.9 mg/ml protein samples run in buffer containing differential amounts of calcium are almost identical in appearance. The sample with the smallest calcium concentration (1.0 mM) has a calcium/protein subunit (28,340 MW) ratio of ~ 31.5. The sample with the largest calcium concentration has a calcium/protein subunit ratio of 3150. Figure 3.26 demonstrates that protomer formation is promoted by calcium. However, if calcium is involved in the chemical equilibrium at each successive step in the formation of higher order protomers then an environment with a higher calcium concentration ought to lead to the production of larger protomers. For a fixed protein concentration a sample with a higher calcium concentration should form a larger number of higher order protomers. Le Chatelier's principle ought to mandate that the reaction be driven to the right.
\[ nA + m\text{Ca}^{2+} \rightleftharpoons \text{Ca}_m A_n = B \]
\[ K_{A,B} = \frac{[B]/[\text{Ca}^{2+}]^m [A]^n}{[\text{Ca}^{2+}]^m [A]^n} \]

The same would be true for the conversion of B into C.

\[ qB + m\text{Ca}^{2+} \rightleftharpoons \text{Ca}_m B_q = C \]
\[ K_B = \frac{[C]/[\text{Ca}^{2+}]^m [B]^q}{[\text{Ca}^{2+}]^m [B]^q} \]

(etc... for C into D and D into E)

This phenomenon is not observed in EP fluid protomer formation. The similarity of the distribution profiles in Figure 3.26 indicates that Ca^{2+}-protein binding is already saturated at a 1 mM [Ca^{2+}]. If this is true, then how can the system be explained?

Figure 4.3 illustrates another interesting feature of protomer formation. In Figure 4.3, there is complete separation between peaks A and B, yet peaks B, C, D and E show a high degree of peak overlap. This observation suggests that the reactions involved in the establishment of equilibrium between peaks A and B may be kinetically slower than those involved in equilibrium between the higher order protomers (peaks B, C, D, E).

Figure 4.4 shows the distribution profile of a 0.6 mg/ml protein
sample that was bought to 10 mM Ca$^{2+}$ and allowed to equilibrate for 24 hours prior to sedimentation velocity analysis. Figure 4.4 shows that the A (s* ~ 4.0) peak has been essentially entirely converted into protomers, yet the distribution profile of the protomers, and the extent of overlap between the protomer peaks, remains similar to that observed in Figures 3.26 and 4.3. What seems to be occurring is that calcium is binding with the apo-protein A (s*~ 4.0) and upon doing so enhances the formation of the first protomer species B (s* ~ 10). For the 0.9 mg/ml [A], a 1.0 mM [Ca$^{2+}$] is sufficient to saturate the Ca$^{2+}$/protein binding sites and thus protomer formation is not enhanced by additional calcium (Figure 3.26). Once a sufficient number of the B protomers (s* ~ 10) have been formed, these species may interact with one another to form higher order protomers. This secondary protein-protein interaction does not depend on additional calcium. The equilibrium of formation/dissociation between B $\rightleftharpoons$ C, C $\rightleftharpoons$ D and D $\rightleftharpoons$ E, occurs at a faster rate than the reactions involved in establishing the Ca$^{2+}$ + A $\rightleftharpoons$ B equilibrium. The result of the differential equilibrium kinetics is that a baseline separation of A from the protomers B, C, D and E may be achieved but separation of the protomers themselves may not (Figure 4.3).

If calcium is not involved in the protomer-protomer interactions than what enables them to bind? Circular dichroism scans of the fluid protein, taken in calcium free and calcium containing (10 mM) buffer solutions

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 4.3 Sedimentation velocity distribution profile of a 0.9 mg/ml sample in 20 mM Tris buffer, pH 7.5, 0.1 M KCl, 100 mM Ca\(^{2+}\).

Figure 4.4 Sedimentation velocity distribution profile of a 0.6 mg/ml sample in 20 mM Tris buffer, pH 7.5, 0.1 M KCl, 10 mM Ca\(^{2+}\) incubated for 24 hours at 4 °C prior to analysis.
clearly show that the protein undergoes a significant structural change once calcium binding has occurred (Fig. 3.31). When compared against CD scans produced by peptides of known secondary structure (Fig. 3.32), it appears that the calcium-protein complex has a more defined secondary structural arrangement than does the calcium-free protein. Once the calcium-protein structural rearrangement has occurred, the positions of various protein functionalities may become fixed as the protein as a whole becomes less flexible. This greater rigidity may help establish a protein structure which is better able to interact (H-bonding, hydrophobic interactions) with other protein molecules and as such enhance protomer assembly.

According to the experimental observations of Gilbert and Gilbert\(^{143}\), who studied the reversible self-association of bovine β-lactoglobulin A, an \(n\) increase in the molecular weight of a molecule ought to be accompanied by an \(n^{2/3}\) increase in the molecule's sedimentation coefficient if the molecule is spherical in shape. For example, in the case of two monomers combining to form a dimer \(n = 2\). According to Gilbert and Gilbert the dimer ought to have a sedimentation coefficient of \(2^{2/3}\) or 1.58 times that of its constituent monomer.

Be applying the Gilbert and Gilbert formula to our system, estimates of the stoichiometries of the assembly for the various protomers may be obtained. In the case of \(A = B\), there is an observed increase in the
sedimentation coefficient from $s \sim 4.0$ to $s \sim 10$. This represents an increase in the sedimentation coefficient of $10/4$ or $2.5$ times. Using this value, we may back-calculate to determine the increase in molecular size. A $2.5$ times increase in sedimentation coefficient equals an $2.5^{3/2}$ or $4.0$ times increase in molecular size. Therefore, according to this model, 4 $A$ molecules (with the aid of Ca$^{2+}$) combine to form one $B$ molecule. The same calculation may be applied to the formation of the other protomers.

<table>
<thead>
<tr>
<th>Peak</th>
<th>Observed $s$</th>
<th>Increase in $s$</th>
<th>Calculated increase in size</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4.0</td>
<td>10/4.0 or 2.5</td>
<td>$2.5^{3/2} = 4.0 \div 4.0$ for $A_4 = B$</td>
</tr>
<tr>
<td>B</td>
<td>10</td>
<td>15/10 or 1.5</td>
<td>$1.5^{3/2} = 1.8 \div 2.0$ for $A_9 = C$</td>
</tr>
<tr>
<td>C</td>
<td>15</td>
<td>19/15 or 1.3</td>
<td>$1.3^{3/2} = 1.5 \div 1.5$ for $A_{12} = D$</td>
</tr>
<tr>
<td>D</td>
<td>19</td>
<td>22/19 or 1.2</td>
<td>$1.2^{3/2} = 1.3 \div 1.3$ for $A_{16} = E$</td>
</tr>
</tbody>
</table>

What these numbers indicate is that once the initial, calcium-dependent, protomer formation has occurred ($4A = B$), higher-order protomer formation occurs the successive addition of the $B$ protomer.

$$B + B \equiv C \quad K_D = [C] / [B]^2$$

$$C + B \equiv D \quad K_D = [D] / [C][B]$$

$$D + B \equiv E \quad K_D = [E] / [D][B]$$

Figure 4.5 is a cartoon of this assembly scheme where $A$ itself is a multimer ($n > 2$) of the 28,340 MW subunit.
Although comparison of Figures 4.3 and 4.4 implies the initial, calcium-dependent, protomer formation is a kinetically slower process than protomer-protomer interaction, it also suggests that the equilibrium constant of the former process is larger than those of the latter. Figure 4.4 shows the nearly complete conversion of 0.6 mg/ml of EP fluid apo-protein into it's protomeric form, when allowed to incubation in a 10 mM $[\text{Ca}^{2+}]$ buffer at 4
°C for 24 hours. In the simple case of a Ca/protein stoichiometry of 1 to 2, the initial calcium binding reaction would have an equilibrium constant of $pK_D \sim 9$.

If the protomer-protomer interactions are assumed to follow the equation, $B + B \rightleftharpoons B_2$ or $C, C + B \rightleftharpoons BC$ or $D$, etc., then from the peak areas in Figure 4.4 equilibrium constant of $\sim 4.5$ may be calculated for all of the protomer-protomer associations. Given below is a list of the equations and equilibrium expressions for the reactions mentioned in the preceding discussion.

EP fluid monomer = 28,340 μm

in solution $p28,340 = A$

\[
\begin{align*}
nA + mCa^{2+} & \rightleftharpoons B & pK_D &= \frac{[B]}{[A]^n[Ca^{2+}]^m} = 9.25 \\
B + B & \rightleftharpoons C & pK_D &= \frac{[C]}{[B]^2} = 4.55 \\
C + B & \rightleftharpoons D & pK_D &= \frac{[D]}{[C][B]} = 4.55 \\
D + B & \rightleftharpoons E & pK_D &= \frac{[E]}{[D][B]} = 4.33
\end{align*}
\]

The similar magnitude of the above calculated protomer-protomer dissociation constants suggests that the chemical interaction in all cases is likely to be the same.

The ability to self assemble is one of the most intriguing properties ascribed to the macromolecules of the organic matrix. The ability of newly-
secreted matrix material to assemble into a functioning, precisely-arranged network, capable of directing the construction of mineral formations with exacting detail is truly a wonder of nature. The observations described above may hold clues to the details of how the organic matrix is developed. Soluble glycoproteins are secreted into the EP fluid. Once there, these protein undergo calcium enhanced transformation which leads to the formation of a significantly larger Ca-protein multimer. Either as a prerequisite or as a result of calcium binding, the secondary structure of the glycoprotein adopts a more ordered conformation. This new structural arrangement enables the individual Ca-protein complexes to interact with one another. These Ca-protein-Ca-protein interactions lead to the formation of larger protomeric species. With the exception of the initial protomer formation (protomer B), the equilibrium of protomer association/dissociation is not influenced by calcium; however the structural stability of the system is highly calcium-dependent. In the absence of Ca$^{2+}$, protomers which may realistically be composed of $\geq 32$ subunits (protomer E), are observed to completely revert back into their constituent subunit (species A). Perhaps in vivo the protomers undergo covalent crosslinking in order to enhance their stability.

Previous work performed in this laboratory detected tyrosinase activity in the extrapallial fluid.$^{31,32}$ Tyrosinase is a copper-containing
metalloenzyme which is thought to be responsible for crosslinking the proteins of the periostracum, and has been suggested as a possible crosslinker for the insoluble organic matrix proteins. As mentioned in the Results section, the EP fluid protein studied here is also a copper-containing protein. Although assays for tyrosinase activity proved negative, the protein may be in a proenzyme state and, being removed from it's normal environment, has lost it's mode of activation. The existence of a protyrosinase molecule which is activated by enzymatic removal a protein fragment has been reported. A more thorough investigation into this possibility should be pursued.

To conclude, this investigation into the macromolecular content of the extrapallial fluid of *Mytilus Edulis*, has led to the purification and partial characterization of a structurally-unique glycoprotein (A). This glycoprotein has the ability to bind calcium and this calcium-binding enhances the formation of a Ca-protein complex (protomer B). Accompanying calcium binding, the Ca-protein complex appears to undergo significant secondary structural rearrangement. This structural rearrangement enhances the Ca-protein protomers ability to self associate into higher order protomers (protomers C, D and E). Protomer formation appears to follow a stoichiometric pattern and is fully reversible. Conclusive data on the true biological function of the protein has yet to be obtained; however, based on several of its characteristic features, some of its inherent similarities with
SOM protein, as well as its anatomical location within the animal itself, it is believed to be intimately involved in shell formation and may well be a precursor or building block to the soluble organic matrix.
BIBLIOGRAPHY


121. Data Bases of the National Center for Biotechnology Information (NCBI), Brookhaven Protein Data Bank, Swiss-Prot, Genpept, GenBank, Kabat Sequences

122. Trypsin Data base located at internet web site http://home.mcom.com/


125. Stafford diffusion constant equation personnel communication


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


