5-2008

CHARACTERIZATION OF SOLUBLE MATRIX FROM MOLLUSCAN SHELL WITH AN EMPHASIS ON TWO MAJOR PHOSPHOPROTEINS FROM THE EASTERN OYSTER, Crassostrea virginica

Mary Johnstone
Clemson University, mbj@clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Part of the Zoology Commons

Recommended Citation
Johnstone, Mary, "CHARACTERIZATION OF SOLUBLE MATRIX FROM MOLLUSCAN SHELL WITH AN EMPHASIS ON TWO MAJOR PHOSPHOPROTEINS FROM THE EASTERN OYSTER, Crassostrea virginica" (2008). All Dissertations. 222.
https://tigerprints.clemson.edu/all_dissertations/222

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
CHARACTERIZATION OF SOLUBLE MATRIX FROM MOLLUSCAN SHELL WITH AN EMPHASIS ON TWO MAJOR PHOSPHOPROTEINS FROM THE EASTERN OYSTER, *Crassostrea virginica*

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Biological Sciences

by
Mary Oliver Johnstone
May 2008

Accepted by:
Dr. AP (Hap) Wheeler, Committee Chair
Dr. Jim Colacino
Dr. Karen Burnett
Dr. Richard Montanucci
ABSTRACT

The soluble organic matrix (SM) extracted from the shell of the Eastern oyster, *Crassostrea virginica* was analyzed to identify the constituent proteins and their structures. Two acidic proteins with estimated molecular weights of 48 and 55 kD were isolated and determined to have highly related structures. Both proteins are comprised primarily of aspartate, serine, and glycine which represent nearly 90% of the total amino acids. Studies indicate that both proteins are phosphorylated and have estimated isoelectric points in the range of 3 to 5. Evidence suggests that poly-aspartate and poly-serine domains are present in the 48 kD protein. Amino terminal sequences were determined to be D-E-A-D-A-G-D-A-Y-D-V-A-D-T, for the 48 kD protein, and S-K-G-X-E-P-D, for the 55 kD protein where, X may be phosphoserine. Immuno-analyses using antibodies generated to the 48 and 55 kD proteins revealed that they share many common epitopes. Additional proteins were identified in SM by these antibodies indicating that the 48 and 55 kD proteins are members of a prominent acidic protein family. The acidic character of the 48 and 55 kD proteins, and their presence in the shell and extra-pallial fluid supports the hypothesis that they function in shell formation.

The SM from 14 bivalve species was compared in order to determine whether or not specific matrix protein components correlate with shell microstructure or mineralogy. While the compositions from all species were enriched in acidic amino acids, SM derived from calcite shell, including prismatic and foliated microstructures, was determined to be more acidic than aragonite. A prominent phosphoprotein class was identified in folia SM
by electrophoresis and appears unique to species having this microstructure. Electrophoretic protein patterns of aragonite SM were more variable and may reflect its earlier appearance in the geologic record compared to calcitic shell structures. ELISA experiments using an antibody generated to the SM of the Eastern oyster demonstrated highest cross-reactivity among related species, irrespective of shell structure. Western analysis shows broad reactivity to component proteins among all species. Apart from the folia microstructure, there was no particular protein pattern that correlated with a specific shell structure, rather SM protein patterns were determined to be species specific.
DEDICATION

To my father, who was my first biology teacher, and to my mother, who, among many roles, acted as zookeeper; and, to my children, Cary and Mason, who have known me only as a student.
ACKNOWLEDGMENTS

I would like to thank my major advisor, Dr. A.P. Wheeler for giving me the opportunity and support to pursue my research interests in this fascinating field of study. I am especially grateful for his guidance, encouragement, great patience, and friendship throughout my years at Clemson. I would also like to thank my committee, Dr. K. Burnett at the College of Charleston, Dr. J. Colacino, and Dr. R. Montanucci for their advice and review of this manuscript, and especially for their encouragement and friendship.

I would also like to thank my lab cohorts, Dr. Andrew Mount, Neeraj Gohad, and Josh Mount for their support and advice while working through this process. I benefit greatly from their exceptional creativity and exuberance for understanding the natural world. I am grateful to Mark Stokes for his expertise in formatting this manuscript.

Very special thanks go to my beloved friends and family, for their love, encouragement, and support at many levels during this long process, especially in the last months. It is because of all of you that this endeavor has been realized. I especially want to thank my husband, David, for his long sacrifice and immeasurable support over the years, and to my children, Cary and Mason, for giving me perspective and for being my cheerleaders, even in the midst of having family time taken from them when I needed to write.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>TITLE PAGE</strong>..................................................................................</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td><strong>ABSTRACT</strong>.....................................................................................</td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td><strong>DEDICATION</strong>..................................................................................</td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td><strong>ACKNOWLEDGMENTS</strong>...........................................................................</td>
<td>v</td>
</tr>
<tr>
<td></td>
<td><strong>LIST OF TABLES</strong>...........................................................................</td>
<td>viii</td>
</tr>
<tr>
<td></td>
<td><strong>LIST OF FIGURES</strong>...........................................................................</td>
<td>ix</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><strong>MAIN INTRODUCTION</strong>.......................................................................</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>**ISOLATION AND CHARACTERIZATION OF ORGANIC MATRIX PHOSPHOPROTEINS DERIVED FROM THE SHELL OF THE EASTERN OYSTER, <strong><em>CRASSOSTREA VIRGINICA</em></strong> ..................................................................</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Introduction........................................................................................</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Materials and Methods........................................................................</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Results..............................................................................................</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Discussion.......................................................................................</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Summary............................................................................................</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Introduction........................................................................................</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Materials and Methods........................................................................</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>Results..............................................................................................</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Discussion.......................................................................................</td>
<td>146</td>
</tr>
<tr>
<td>Summary</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>4 CORRELATION OF ORGANIC MATRIX COMPOSITION TO SHELL MICROSTRUCTURE AND MINERALOGY</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>168</td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>Discussion</td>
<td>234</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td>252</td>
<td></td>
</tr>
<tr>
<td>5 FINAL DISCUSSION</td>
<td>255</td>
<td></td>
</tr>
<tr>
<td>APPENDICES</td>
<td>273</td>
<td></td>
</tr>
<tr>
<td>REFERENCES</td>
<td>279</td>
<td></td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Chemical Analyses of the SM Protein Fraction</td>
<td>25</td>
</tr>
<tr>
<td>2.2 Amino Acid Composition of Matrix Fractions and Gel Purified Matrix Proteins</td>
<td>53</td>
</tr>
<tr>
<td>2.3 Comparison of Proteins and Protein Fractions by Difference Index</td>
<td>54</td>
</tr>
<tr>
<td>2.4 Amino Acid Compositions of 48 and 55 kD Proteins Before and After Dialysis on Immobilon P®</td>
<td>56</td>
</tr>
<tr>
<td>3.1 Amino-terminal Sequences(^{(1)}) for the 48 and 55 kD Proteins</td>
<td>128</td>
</tr>
<tr>
<td>4.1 Shell Microstructure and Mineralogy of Selected Bivalve Species (Carter, 1990)</td>
<td>173</td>
</tr>
<tr>
<td>4.2 TAXONOMY OF SELECTED BIVALVE SPECIES (Carter, 1990)</td>
<td>174</td>
</tr>
<tr>
<td>4.3 Percent BCA Weight to Dry Weight of SM from Selected Species</td>
<td>184</td>
</tr>
<tr>
<td>4.4 Amino Acid Compositions of Soluble Matrix (SM) From Selected Bivalve Species</td>
<td>187</td>
</tr>
<tr>
<td>4.5 Comparison of Amino Acid Compositions Between Whole Shell and Nacreous Shell Layers of Two Bivalves</td>
<td>190</td>
</tr>
<tr>
<td>4.6 Chemical Analyses of SM Protein Fractions Relative to Shell Microstructure and Mineralogy from Selected Bivalve Species</td>
<td>204</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>SDS-PAGE analysis of SM and sIM.</td>
<td>28</td>
</tr>
<tr>
<td>2.2</td>
<td>Preparative electrophoresis of whole SM.</td>
<td>31</td>
</tr>
<tr>
<td>2.3</td>
<td>Comparison of whole soluble matrix (SM) protein patterns by</td>
<td></td>
</tr>
<tr>
<td></td>
<td>differential staining.</td>
<td>33</td>
</tr>
<tr>
<td>2.4</td>
<td>SDS-PAGE analysis of gel purified protein demonstrating bleed off of</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lower molecular weight bands following freeze-thaw.</td>
<td>35</td>
</tr>
<tr>
<td>2.5</td>
<td>Isoelectric focusing of SM and SM dephosphorylated with alkaline</td>
<td></td>
</tr>
<tr>
<td></td>
<td>phosphatase (SM-AP) with pH range of 3 to 7.</td>
<td>37</td>
</tr>
<tr>
<td>2.6</td>
<td>Two-dimensional analysis comparison of SM and SM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dephosphorylated with alkaline phosphatase (SM-AP) after Stains-all staining.</td>
<td>40</td>
</tr>
<tr>
<td>2.7</td>
<td>Isoelectric focusing of matrix proteins on gel supplemented with CHAPS with</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a pH 3 to 10 range.</td>
<td>43</td>
</tr>
<tr>
<td>2.8</td>
<td>Two-dimensional analysis comparison of SM and SM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dephosphorylated with alkaline phosphatase (SM-AP).</td>
<td>47</td>
</tr>
<tr>
<td>2.9</td>
<td>Two-dimensional analysis comparison of SM and SM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dephosphorylated with alkaline phosphatase (SM-AP) following silver staining.</td>
<td>51</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.10 Glycoprotein identification compared to non-glycosylated proteins of SM</td>
<td>62</td>
</tr>
<tr>
<td>3.1 Enzyme-linked immuno-sorbent assay demonstrating specificity and cross-reactivity of the chicken anti-48 kD SM antibody with selected proteins</td>
<td>96</td>
</tr>
<tr>
<td>3.2 Enzyme-linked immuno-sorbent assays demonstrating reactivity and cross-reactivity of the rabbit anti-48kD, 55kD, SM and IM antibodies to their protein antigens and other phosphoproteins</td>
<td>101</td>
</tr>
<tr>
<td>3.3 Western analyses of anti-matrix antibodies to selected proteins</td>
<td>104</td>
</tr>
<tr>
<td>3.4 Enzyme-linked immuno-sorbent assay demonstrating the effect of dephosphorylation of SM on reactivity of the anti-55 kD antibody</td>
<td>107</td>
</tr>
<tr>
<td>3.5 Western analyses of dephosphorylated matrix protein following alkaline phosphatase and base treatment</td>
<td>109</td>
</tr>
<tr>
<td>3.6 SDS-PAGE analyses of the 48kD protein following treatment with chymotrypsin and endoproteinase Glu-C (V8 protease)</td>
<td>113</td>
</tr>
<tr>
<td>3.7 SDS-PAGE analyses of 48 kD and dephosphorylated 48 kD protein following treatment with trypsin</td>
<td>117</td>
</tr>
<tr>
<td>3.8 SDS-PAGE analysis of cleavage products generated from treating SM protein with pepsin and papain</td>
<td>121</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.9</td>
<td>SDS-PAGE analysis of cleavage products generated from treating 48 kD SM protein with endoproteinase Asp-N and Proteinase K.</td>
</tr>
<tr>
<td>Figure 3.10</td>
<td>SDS-PAGE analysis of SM following a time course treatment with formic acid.</td>
</tr>
<tr>
<td>Figure 3.11</td>
<td>SDS-PAGE analysis of SM and 48 kD protein following treatment with CnBr.</td>
</tr>
<tr>
<td>Figure 3.12</td>
<td>Enzyme-linked immunosorbent assay comparing reactivities of the anti-55 kD antibody to soluble shell matrix protein (SM) to those of tissue extracts.</td>
</tr>
<tr>
<td>Figure 3.13</td>
<td>Western analysis of mantle, EPF and hemolymph extracts compared to shell SM.</td>
</tr>
<tr>
<td>Figure 3.14</td>
<td>SDS-PAGE of extra-pallial fluid (EPF), hemolymph, and mantle extract compared to shell SM.</td>
</tr>
<tr>
<td>Figure 3.15</td>
<td>SDS-PAGE analysis of sIM extracts from the three principal shell layers and their reactivity to the anti-55 kD antibody as demonstrated by Western blot analysis.</td>
</tr>
<tr>
<td>3.16</td>
<td>Enzyme-linked immuno-sorbent assay demonstrating reactivity of the anti 55 kD antibody to sIM matrix extracts from the three principal shell layers.</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Table showing the various varieties of shell microstructures based on their structural arrangement (Column 1) and how they occur most commonly in major shell layers of Bivalvia (Column 2).</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>A single valve from the selected bivalve species.</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>A molecular phylogeny of the bivalve mollusks.</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Amino acid compositions represented as rose diagrams of all species.</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Rose diagrams of all represented amino acids for selected species.</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Difference indexes (SΔN) of SM compositions among species.</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>Bar plot of total phosphate content determined for SM of selected bivalves.</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Bar plot of carbohydrate content estimated for SM from selected bivalves.</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>SDS-PAGE analyses and differential staining of soluble organic matrix proteins from selected bivalve species.</td>
</tr>
<tr>
<td>4.10</td>
<td>SDS-PAGE analysis of matrix proteins extracted from whole shell compared to matrix protein from the inner nacreous shell layers of <em>A. rigida</em> and <em>P. fucata</em>.</td>
</tr>
<tr>
<td>4.11</td>
<td>Glycoprotein identification compared to non-glycosylated proteins of soluble matrix from selected bivalve species using gel blot analysis.</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 4.12  Enzyme linked immunosorbent assay (ELISA) demonstrating cross-reactivity of the anti-SM antibody made against C. virginica to SM of selected species.</td>
<td>228</td>
</tr>
<tr>
<td>Figure 4.13  Bar plot showing % cross-reactivity of SM among species compared to C. virginica.</td>
<td>230</td>
</tr>
<tr>
<td>Figure 4.14  Western Blot showing cross-reactivity of the anti-SM antibody made to SM from C. virginica with SM from selected species.</td>
<td>231</td>
</tr>
<tr>
<td>Figure A  Summary of Protein Bands Identified in SM from Selected Bivalve Species.</td>
<td>274</td>
</tr>
<tr>
<td>Figure B  Enzyme linked immunosorbent assay (ELISA) cross-reactivity of the anti-48 kD antibody to SM of selected species.</td>
<td>278</td>
</tr>
</tbody>
</table>
CHAPTER ONE

MAIN INTRODUCTION

The ability of organisms to create mineralized hard parts such as bone, teeth, and shell is through a process called biomineralization. The remarkable array of organisms that carry out this process is as diverse as the structures they produce, including members from all five Kingdoms that have utilized more than 60 different minerals as their raw material. Of these minerals, calcium carbonate and silica by far compose the bulk of mineral formed biotically. Calcium carbonate occurs in most invertebrate species and on an even more massive scale in marine algae, coccolithophorids. Calcium phosphate constitutes only a fraction of the total mineral biomass. Because this mineral composes vertebrate teeth and bone it has merited special consideration (Lowenstam and Weiner, 1989).

The key feature that distinguishes biomineral from abiotic mineral is that it is a composite containing both mineral and an organic phase called the organic matrix. In molluscan shell, calcium carbonate most often exists as two crystalline polymorphs calcite or aragonite, which are distinguished by the arrangement of their lattice ions (Lowenstam and Weiner, 1989). This is also referred to as mineralogy. The matrix intimately associates with the mineral both surrounding and permeating each crystal unit. Consequently, biologically formed mineral often take on crystal shapes different from their abiotic homologs (Weiner and Addadi, 1997). For example, inorganically formed
crystals of calcite have a signature rhombohedral shape whereas aragonite forms as smaller, needle-like crystals under conditions of normal temperature and pressure. In molluscan shell, each can form an array of crystal shapes such as long, columnar prisms or flattened laminar shaped crystals (Fuchigami and Sasaki, 2005; Carter, 1980).

Individual crystals within mineral composites formed by organisms are deposited in regular and predictable arrangements. In bivalve shell, calcite and aragonite form more than eleven distinct yet highly conserved microstructures, each defined by distinct crystal shapes with extraordinarily ordered arrangements (Carter, 1980). Most shells are multi-layered and are often comprised of both calcite and aragonite mineralogies and two or more microstructural arrangements. This multi-layered architecture and the presence of organic matrix impart mechanical strength to biominerals that is far superior to those formed abiotically (Currey, 1989; Smith et al., 1999; Daw, 2004). It is this attribute, considered with the diversity and regularity of structures formed by organisms that warrants study of how the organic matrix controls mineral formation.

Shell formation is a sequential and highly regulated event that takes place at the shell margin in intimate association with the mantle organ (Simkiss and Wilbur, 1989). Mineralization occurs in a compartment defined by the mantle epithelium, the inner shell surface, and the periostracum, a scleratinized protein sheet secreted by the mantle which attaches to the outside of the shell, sealing the space between the shell and mantle from the external environment (Checa, 2000). This space is termed the extra-pallial cavity.
(EPC) and is filled with a gelatinous extrapallial fluid (EPF) which creates a liquid zone between the shell and the mantle epithelium. It is within this medium that organic matrix and mineral interact to form the shell.

EPF is supersaturated with the ions of the mineral being deposited (Crenshaw, 1972a). Sources of calcium come from the aqueous environment and from stores within the tissues of the oyster (Watabe et al., 1976). The shell itself may provide an additional source of lattice ions as the animal undergoes anaerobic respiration in which the shell dissolves, providing buffering ions to counteract the build-up of metabolic acid. Calcium may pass from the hemolymph (blood) into the EPF by diffusion via paracellular pathways (Neff, 1973; Hudson, 1992). CO₂ accumulated during aerobic metabolism passes readily across cell membranes and through the action of carbonic anhydrase is converted to CO₃ in the EPC (Wheeler, 1992). A matrix protein derived from the nacreous shell layer of the pearl oyster was found to contain a functional carbonic anhydrase domain suggesting that this conversion may occur within the secreted matrix at the mineralization front (Miyamoto et al., 1996; 2005).

Another source of shell mineral may come from mineral containing cells. Vesicles identified in specialized cells of the mantle epithelium are reported to contain mineral laden granules (Neff, 1972; Myers, 1999). These granules are believed to contain amorphous calcium carbonate (non-crystalline) which may be transported to the site of mineralization (Addadi et al., 2006). Granular blood cells (hemocytes) are
proposed to transport mineral to the site of shell formation (Mount et al., 2004). Because these cells are refractive, the mineral inside is crystalline and is proposed to be the source of nascent crystals for the initiation of the shell forming process.

Organic matrix components are secreted by cells in the mantle epithelium into the EPF in a highly orchestrated manner (Jolly et al., 2004; Jackson et al., 2006; Myers et al., 2007). Many of these components are characteristically acidic and include glycoproteins, phosphoproteins, peptides, carbohydrates, and lipids (Lowenstam and Weiner, 1989). It is this cocktail that is generally believed to influence shell mineralogy and microstructure. How these components effect shell structure might be explained in the context of their chemical character, where and when they are secreted from the mantle epithelium and where they localize within the shell mineral.

A general model for the sequence of shell formation starts with the secretion of the periostracum from the periostracal groove located in the mantle organ. The mantle margin elaborates into three lobes; an outer lobe which lies adjacent to the shell surface, a middle lobe, and an inner lobe, which flanks the pallium (Galstoff, 1964). The periostracal groove lies at the intersecting base of the outer and middle mantle lobes. This region is a well-defined collection of specialized cells that secrete the periostracal material which is extruded as a highly cross-linked polymer sheet (Degens et al., 1967). This material is largely proteinaceous and in many cases contains chitin (Lowenstam and
Weiner, 1989). The periostracum is the primary ground substance onto which the shell layers are subsequently deposited (Checa, 2000).

The second event believed to occur is the assembly of a portion of the matrix secreted by the outer lobe epithelium. These proteins are more hydrophobic and are often enriched in the amino acids alanine, glycine, and proline and contain cysteine to varying degrees (Zhang and Zhang, 2006). Accordingly, the protein monomers are thought to assemble through a combination hydrophobic, covalent, and electrostatic interactions. The resulting assemblage is believed to form a scaffold on the surface of the periostracal sheet and provides a three dimensional framework within which mineral deposits and grows (Checa et al., 2005). This protein localizes around each crystal unit and is often referred to as inter-prismatic matrix (Watabe, 1965; Weiner and Traub, 1984). When extracted from the shell, this protein remains insoluble in water following shell dissolution and consequently belongs to the insoluble matrix (IM) class of proteins (Wheeler, 1992). These proteins are often called “structural” matrix.

A third event that occurs during shell formation is the release of a more acidic class of proteins. Studies have shown that the IM scaffold is in fact a composite material containing a hydrophobic core which is coated with acidic matrix (Kawaguchi and Watabe, 1993; Marin et al., 2007). Some of these proteins are water soluble following shell dissolution and others require the use of anionic detergents such as sodium dodecyl sulfate (SDS) and chaotropic agents such as urea to disrupt its association with the IM
scaffold (Mount, 1999; Crenshaw, 1990). These proteins are classified as the soluble matrix (SM) class and presumed to initiate crystal growth through anionic domains either by localizing calcium and carbonate ions (Weiner and Hood, 1975; Greenfield et al., 1984) or by binding nascent crystals (Wheeler, 1992; Mount et al., 2004). The evidence for both mechanisms will be discussed further below.

An additional secretion of matrix protein, also from the highly anionic SM class, interacts with the growing crystals inside the IM framework. These proteins are occluded within each crystal unit, dividing it into “nano” units which collectively diffract as a single crystal (Weiner and Traub, 1984). These proteins are easily extracted from the mineral in water using mild acid or the calcium chelating agent, ethylenediamine tetraacetic acid (EDTA). As a group, SM proteins are often referred to as “control” molecules because of their ability to interact with crystal lattice ions and crystal surfaces (Gotliv et al., 2003; De Yoreo and Dove, 2004).

Finally, an organic “membrane” forms on the surface of each mineral layer in the shell (Addadi and Weiner, 2006; Lee and Choi, 2007; and unpublished observations). This material remains insoluble following shell dissolution in EDTA and is termed “interlamellar” matrix as described in earlier studies (Watabe, 1965). In shell nacre, this matrix is reported to be enriched in glycine and alanine and chemically comparable to spider silk (Addadi and Weiner, 2006). *In vivo*, these researchers show that this matrix is water absorbent and forms a gel in between successive mineral layers of the shell. In
vitro, extracts containing this matrix fraction inhibit crystal formation (Cohen, 2003). The occurrence of this organic membrane between distinct shell layers (Lee and Choi, 2007; unpublished observations) suggests it may hold some of the chemical cues for determining the mineralogy and microstructure of the crystals deposited.

Of course, the sequence of the events as described above are given only to understand how the proteins might interact with mineral as shell layer formation occurs and do not necessarily reflect their order of occurrence in vivo. Shell formation is a dynamic process and potentially involves the secretion of hundreds of protein components at specific times (Jackson et al., 2006). It is not known when these components are secreted during the process much less how they interact with each other to form the organic framework that supports mineral growth.

Some progress has been made towards understanding where in the mantle epithelium some matrix component mRNA’s are expressed in relation to shell structure. Proteins expressed near the mantle edge localize within the outer layer and are thought to be important for the construction of the corresponding mineralogy and microstructure. Proteins secreted dorsally along the epithelium towards the mantle center occur in the inner shell layer and are presumed to influence its mineral structure. Proteins expressed along the entire epithelium are thought to be important for the construction of both layers (Sudo et al., 1997; Jolly et al., 2004; Takeuchi and Endo, 2006; reviewed in Zhang and Zhang, 2006).
Our understanding of the mechanism by which matrix might influence shell mineralogy and microstructure comes from studies of the structures of shell matrix extracts and matrix isolates and their effect on mineralization in vitro. Early on, it was established that shell SM is dominated by aspartic acid (Asp), serine (Ser), glycine (Gly) and glutamic acid (Glu) to various degrees (Travis et al., 1967). Using dilute acid hydrolysis, a treatment specific for cleaving peptide bonds between Asps, it was suggested that Asp-X-Asp structures were present, with X being either Gly or Ser (Weiner and Hood, 1975; Weiner, 1983). It was proposed that tandem repeats of this arrangement would structurally match the spacing between Ca$^{2+}$ ions in the crystal lattice and thereby provide a template for calcite or aragonite nucleation. Then by altering the microenvironment such as ionic strength, the level of supersaturation of calcium carbonate, or the amount of Mg$^{2+}$, aragonite could be favored to form on the matrix surface over calcite. It was later proposed that sulfated carbohydrate groups present on the Asp-rich proteins could optimize a stereospecific coordination between the Ca$^{2+}$ and highly ordered carboxyl groups defined by the specific arrangement of Asp residues. This structural matching between lattice ions and the protein structure is referred to as epitaxy and is thought to be the mechanism that determines mineralogy and the orientation of the growing crystal (Addadi et al., 1987; Addadi and Weiner, 1989; Addadi et al., 2006).

Support for ordered nucleation gained momentum when several seminal studies demonstrated that SM fractions from either calcite or aragonite shell induced the
corresponding mineralogies from solution (Belcher et al., 1996; Falini et al., 1996; Feng et al., 2000; Thompson et al, 2000). This activity occurred with SM in solution (Belcher et al., 1996), or when the SM was immobilized on an insoluble substrate resembling IM (Falini et al., 1996). In some experiments, calcite was induced by its corresponding SM even in the presence of competing Mg$^{2+}$ ions. Collectively, these studies have been construed as compelling evidence that crystal mineralogy is determined by a specific matrix protein structure.

However, it is important to note that in all of the aforementioned studies, mineralization experiments were carried out in supersaturated calcium carbonate solutions and little to no sodium and were not representative of physiological conditions. In fact, studies of SM from oyster shell have shown that SM binds calcium at low ionic strength, but when salt is increased to physiological ionic strength, Ca$^{2+}$ binding is eliminated (Rusenko, 1988; Mount, 1999). Similar findings have been reported for SM from clam, *M. mercenaria* (Greenfield and Crenshaw, 1984). It may be that in the previously mentioned studies, SM binds nascent crystals formed in solution rather than specifically nucleating the mineral through ordered Ca$^{2+}$.

Along these lines, Wheeler and colleagues proposed an alternative to matrix acting in controlled nucleation (Wheeler, 1992). These authors observed that oyster SM in solution was a potent inhibitor of crystallization. The high affinity of SM for nascent crystal surfaces was insensitive to ionic strength (Sikes and Wheeler, 1988). When
immobilized on a surface, these same proteins were able to initiate crystal growth (Wheeler and Sikes, 1989). Thus, SM was proposed to bind and stabilize nascent crystals presumed to form in the EPF medium. Mount et al., (2004) identified refractive granular blood cells (hemocytes) as a source of nascent crystals for initiating mineralization. These cells appear to supply crystalline calcite mineral to the site of shell growth.

It is widely accepted that crystal binding by SM is a mechanism by which crystal morphology is affected. Unlike ordered nucleation, this interaction does not require a specific amino acid structure, rather proteins with an abundance of acidic and other anionic groups repetitively arranged can induce secondary structures such as \(\beta\)-sheets, helices and random coils which are known to react favorably with mineral surfaces (Wheeler, 1992; Lowenstam and Weiner, 1989). Many of the matrix protein structures determined so far have within them acidic domains that are presumed to bind \(\text{Ca}^{2+}\) or interact with mineral surfaces (Zhang and Zhang 2006). Several SM components have demonstrated crystal shaping abilities \textit{in vitro} (Michenfelder et al., 2003; Susuki et al., 2004; Kim et al., 2004; Marin et al., 2005,).

Furthermore, studies have shown that SM adsorbs to specific crystal surfaces in solution, limiting growth in one direction while allowing growth to proceed in another (Addadi and Weiner, 1985; Berman et al., 1988; Michenfelder et al., 2003). This preferential binding may be a mechanism by which at least some of the boundaries of a growing crystal are defined. Checa et al., (2005) suggests that crystal shape is defined by
the IM framework into which the crystals grow. A recent review of the literature maintains that crystal shape orientation and mineralogy are likely influenced by both classes of proteins working concurrently (Addadi et al., 2006).

Despite the recent progress in identifying component matrix proteins from gene sequences (Zhang and Zhang, 2006), an understanding of how these components work individually and collectively to interact with lattice ions and crystal surfaces is still unclear. Putative functions have been assigned to some components based on conserved structural motifs present in their structures. As examples, MS131 (Sudo et al., 1997) and Prismalin-14 (Susuki et al., 2004) are two proteins expressed in the mantle edge of *P. fucata* and are presumed to localize in the outer calcite shell layer. Both of these proteins have acidic domains which could interact with mineral. However, their structures are dominated by Gly-rich regions and hydrophobic amino acids which are proposed to facilitate interactions with other proteins. Consequently, both proteins are thought to function as framework proteins. On the other hand, Asp-rich and Aspein, two proteins identified in the outer prismatic shell layer of *A. rigida* and *P. fucata* respectively, are dominated by Asp-runs and Asp-X-X-Asp domains but lack prominent hydrophobic regions. These proteins are presumed to function primarily by binding mineral surfaces or lattice ions (Tsukamoto et al., 2004; Gotliv et al., 2005).

Any of these proposed activities for gene or shell isolates remain speculative as none have been demonstrated in vivo. Moreover, components deduced from gene
sequence lack critical information about post-translational processing that is extensive in shell proteins and thought to be crucial for activity. For example, glycosylation may be important for binding Ca$^{2+}$ (Crenshaw and Ristedt, 1976; Addadi and Weiner, 1989) and phosphorylation has been shown to be important for regulating crystal growth (Borbas et el., 1991). Enzymatic degradation of matrix proteins has been reported as a mechanism by which protein activity is altered during specific phases of mineralization (Robinson et al., 1995).

To fully comprehend matrix mediated shell formation, it must be approached at the gene level and post-secretion using molecular and biochemical techniques. The focus of this study was to chemically characterize two major phosphoprotein components previously identified in the SM from the inner shell layer (folia) of the eastern oyster, *Crassostrea virginica* (Myers et al., 1996; Myers, 1999). These proteins were estimated to be 48 and 55 kD and are the major acidic constituents of a previously identified reverse phase HPLC fraction of SM (RP-1) which was demonstrated to have crystal binding activities (Rusenko, 1988; Wheeler et al., 1988b). Using a variety of chemical analyses including amino acid analyses, Edman sequencing, immuno detection, and SDS-PAGE techniques, the structures of the 48 and 55 kD were partially elucidated. Particular attention was paid to phosphorylation, glycosylation and how these proteins appeared in the shell layers and fluids (EPF and hemolymph) which gave some indication of their post-secretory forms. Using similar techniques, a comparative study was conducted on SM from specific shell layers of 14 bivalve species to investigate if particular SM
characteristics or components correlated with a particular shell microstructure or mineralogy.
CHAPTER TWO

ISOLATION AND CHARACTERIZATION OF ORGANIC MATRIX
PHOSPHOPROTEINS DERIVED FROM THE SHELL OF THE EASTERN OYSTER,
CRASSOSTREA VIRGINICA

Introduction

The study of biologically controlled mineralization requires an understanding of the macromolecules associated with mineral, and the molecular details of how they interact with each other and the mineral phase to affect mineral formation. Accordingly, the identification and characterization of component matrix proteins has long been a primary focus of research in vertebrate and invertebrate systems. For a variety of reasons, this endeavor has been challenging. In molluscan shell, a contributing factor is that matrix complexity and composition of component proteins is highly variable among species (Grégoire, 1972). For example, matrix derived from oyster shell has a molecular weight range spanning several thousand to millions of Daltons (Ruenko et al., 1991; Myers, 1999; Mount, 1999) while the quahog, Mercenaria mercenaria, is reported to have only a single protein (Crenshaw, 1972b). Even within single species, investigators often do not agree on the organic content, chemical composition or the number of component proteins (Krampitz et al., 1983; Myers, 1999; Kawaguchi and Watabe, 1993).

In part, matrix variability is attributed to differences in extraction methods and biochemical analysis (Wheeler and Sikes, 1984; Rusenko, 1988). Furthermore, shell layers
of different microstructure found within the same shell have been shown to contain
different amino acid compositions (Grégoire, 1972; Samata, 1988) as well as component
proteins (Weiner, 1983, Fritz et al., 1994; Sudo et al., 1997; Kawaguchi and Watabe,
1993). Consequently, a true assessment of component proteins is dependant on the
accuracy of shell layer separation prior to extraction. To complicate the matter further,
environmental influences, such as temperature and salinity, have been reported to affect
matrix compositions of individuals of the same species (Grégoire, 1972).

Historically, isolating and characterizing component matrix proteins from mineral
using standard biochemical techniques have not been straightforward due to their unusual
chemical properties (Lowenstam and Weiner, 1989; Mann, 2001). In oyster,
distinguishing discrete matrix proteins following gel electrophoresis is difficult because
their high anionic character renders them unable to react appreciably with common
proteins stains, such as Coomassie blue or silver nitrate (Myers, 1996). Moreover, these
proteins are extensively phosphorylated (Wheeler et al., 1991) and fibrous in structure so
that accurate assessment of size and purity is problematic. For similar reasons, other
methods of purification, including gel chromatography and ion exchange, have been
unsuccessful at resolving single protein species from the whole matrix extracts (Rusenko,
1988). And, like many mineral derived matrices, a significant portion of the matrix
assembles into high molecular weight polymers that are intractable to solvent, thus
difficult to analyze (Rusenko, 1988; Wheeler et al., 1991; Mount, 1999).
Because the complexity of shell derived matrix cannot be deciphered entirely using biochemical analyses, researchers have turned their attention to using molecular cloning and related techniques to deduce amino acid structures of the component proteins. While these methods have been successful in identifying a growing number of matrix gene products (reviewed in Zhang and Zhang 2006), they do not reveal information about post-translational processing, which is typically extensive in mineral derived proteins and thought to be critical for function (Lowenstam and Weiner, 1989).

In vertebrate and invertebrate systems, protein phosphorylation and glycosylation have been shown to profoundly affect protein function. The water soluble matrix (SM) from oyster is reported to contain approximately 15% phosphate by weight (Rsuenko, 1988; Borbas, 1991). In vitro, phosphorylated SM was shown to be a more potent inhibitor of crystal growth than dephosphorylated SM (Borbas, 1991). A similar finding has been reported for dentin phosphoproteins (phosphophoryn) which are chemically and functionally similar to oyster shell SM (Butler, 1987; He et al., 2005). Osteopontin and osteonectin are found in the mineralized and non-mineralized tissues of vertebrates. However, in mineralized tissues, these proteins were to have a marked increase in their extent of mineralization compared with those identified in non-mineralized tissues (Boskey, 1998). In vivo, the bones of mice deficient in phosphorylated matrix proteins showed differences in crystal size and a deterioration of mechanical strength compared with control animals (Camacho et al., 1995).
Compared to phosphate, carbohydrate content in oyster SM is low (Rusenko, 1988); however, it is a prominent constituent of the shells of many molluscan species (Samata, 1990; Samata et al., 1980). In nacre formation, sulfates are thought to bind Ca2+ ions and facilitate crystal nucleation (Crenshaw and Ristedt, 1976; Addadi and Weiner, 1989). In crab carapace formation, specific glycoproteins correlate with the onset of mineralization (Shafer et al., 1995).

Shell proteins are processed further post-secretion in the area of mineralization. Some of these assemble into water insoluble aggregates which form the framework matrix (IM) (Simkiss and Wilbur, 1989). Recently, small proteins ranging from 4 kD to 10 kD have been isolated from shell matrix and have been shown to inhibit crystal growth (Zhang et al., 2006; Mann et al., 2007). These proteins could result from proteolytic cleavage of a parent protein. For example, AP7 and AP24 were isolated from the nacre of *H. rufescens* and show some sequence similarity with Lustrin A (Wustman et al., 2004), a much larger protein (65 kD) also found in the nacre of this species. These proteins were shown to inhibit crystal growth at specific sites on the crystal surface (Michenfelder et al., 2003; Kim et al., 2004). It is possible that AP7 and AP24 derive from Lustrin A.

It is crucial to continue biochemical characterization of shell matrix extracts to identify post-translational changes that might occur to the component proteins from the point of secretion to the point they are incorporated into the shell, and to evaluate the how
these modifications may be important for function in the mineralization process. To this end, the focus of this chapter is to continue characterization of the organic matrix derived from the eastern oyster, *Crassostrea virginica* with an emphasis on the water soluble (SM), anionic class of proteins. These proteins have previously been shown to have mineral regulating properties (Wheeler et al., 1981; 1988b; Wheeler and Sikes, 1989). Particular attention is placed on characterizing matrix heterogeneity and post–translational modifications.

**Materials and Methods**

**Extraction and Purification of Matrix Proteins**

Specimens of the Eastern oyster, *Crassostrea virginica*, were collected from oyster beds located in Sixty Bass Creek of North Inlet estuary, Georgetown, South Carolina. The animals were shucked from the shells and the shells were scrubbed with a wire brush under tap water to remove sediment and foreign encrustations. The outer shell layers, including the periostracum (outer most non-mineralized layer) and the prismatic layer (outer most mineralized layer) were removed with a high speed rotary cutting tool. The remaining shell was broken into small pieces with a hammer. Shell pieces composed of foliated mineral were selected and ground into powder by short bursts in an electric coffee grinder. Protein was extracted from the mineral phase by dissolving 25 g of foliated shell in 750 ml of 17% ethylenediamine tetraacetic acid (EDTA), pH 8. The
resulting matrix suspension was centrifuged at 27,000 x g for 30 min to pellet the insoluble matrix (IM). The supernatant containing the soluble matrix (SM) was decanted, concentrated to approximately 50 ml, then dialyzed against 1000 ml of 10 mM NaCl using a Millipore Minitan tangential flow filtration apparatus with a molecular weight exclusion limit of 10kD. The resultant dialysate was further dialyzed against distilled water and lyophilized. The IM pellet was washed three times with 10 mM NaCl and further extracted with 8 M urea and 2% SDS. This soluble extract (sIM) was dialyzed against 10mM NaCl 2X followed by two washes in distilled water and finally lyophilized. Lyophilized protein extracts were stored in a dessicator at -20 °C. Working sIM stock solutions were made by reconstituting the protein in distilled water to a concentration of 5 mg ml⁻¹ based on total BCA weight. Working SM stock solutions were made by reconstituting the protein in distilled water to a concentration of 5 mg ml⁻¹ based on total weight of the dried material.

SM proteins were purified using preparative gel electrophoresis. Approximately 5 mg of dry weight whole SM was loaded onto full length Tris-Glycine-SDS 12% polyacrylamide gels (as described by Laemmli (1970) and modified by Butler et al. (1981)) and run at 25 mA constant current at 4°C. Protein bands were visualized by staining the gel with 0.3 M CuCl₂, and bands of interest were excised from the gel, de-stained for approximately 30 min in three washes of 50 mM EDTA, pH 8, and eluted using an electro-elution module (Bio Rad) according to manufacture’s directions. Following electro-elution, protein was dialyzed against distilled water using Microcon concentrators.
(Millipore) or dialysis cassettes (PIERCE) both having a MWCO of 10 to 12K. Purified protein was lyophilized and stored at -20 °C.

Protein Estimation

When possible, protein was quantified by measuring the dry weight of the extracted material. In cases where very small amounts of material were obtained, protein was quantified using the BCA protein assay (PIERCE).

Analytical gel electrophoresis of soluble and insoluble matrix

Whole SM and sIM were resolved on gradient (4 to 20%) or 12% Tris-Glycine ready gels (Novex) either in the presence of sodium dodecyl sulfate (SDS) or under non-denaturing conditions. Protein was visualized using a variety of stains including, CuCl₂, 1-ethyl-2- [3-(1-ethynaphtho [1,2-d]thiazolin-2-yldene)-2-methylpropenyl]naphtho[1,2-d]thiazolium bromide (Stains-all) and silver nitrate. Gels stained with CuCl₂ were incubated for 5 to 10 min in 0.3 M CuCl₂ and then transferred to distilled water. Stains-all staining was done according to methods described by Campbell et al. (1983) as modified by Myers et al. (1996). Silver staining was conducted by method of (Morrisey, 1981) and was additionally used as a secondary stain after Stains-all staining. This “double staining” technique was found to amplify protein that was Stains-all positive especially the minor constituents whose concentrations were below the sensitivity for detection by Stains-all alone (Tsay and Veis, 1985; Myers et al., 1996). This technique
also had the added advantage of identifying proteins not reactive to Stains-all which appeared white against the light brown background. In this way, most of the protein components could be visualized simultaneously.

**Amino acid analysis**

Amino acid analysis was carried out by re-suspending 10ug of dry weight protein in 6 N HCL containing 2% phenol (N= total number of amino acid residues or mole percent). Hydrolysis was carried out *in vacuo* at 110 ° C for 24 h. Following hydrolysis, samples were dried in a Speed Vac apparatus, re-suspended in 150 ul sample buffer, and analyzed on a Beckman 6300 Autoanalyzer using standard ion exchange resins and ninhydrin detection. To determine whether the amino acid compositions were representative of related or different proteins, a difference index (SAN) was computed using the following equation (Cornish-Bowden, 1983): 

\[ S\Delta N = \frac{1}{2} \Sigma (n_{iA} - n_{iB})^2 \]

where \( n_{iA} \) or \( n_{iB} \) represents the quantity of the \( i^{th} \) type of amino acid contained in protein A or B.

This manipulation assumes that the proteins compared are similar in size and number of amino acids, a condition that is obviously too stringent for multi-component extracts. Still, its application here does give a basis for comparison among matrix extracts which contain a similar range of protein sizes. The values obtained with the equation are applied to the following criteria. If \( S\Delta N < 0.42 \) N (N= 100 when using mole percent), then there is a 95% certainty that the proteins are related. If \( 0.42 \) N < \( S\Delta N < 0.93 \) N, then
the proteins are weakly related. If $S\Delta N > 0.93 N$, then there is a 95% certainty that the proteins are unrelated.

**Phosphate Analysis**

Total phosphate was determined by the method of Eisenreich et al. (1975) as modified by Swift (1985). Generally, the reaction results in the conversion of phosphorus to the orthophosphate form using a persulfate-sulfuric acid digestion. Blue color is produced by the formation of a reduced molybdoantimonylphosphoric acid complex and is detected at 340 nm. The range of standards employed was 0.07 to 5.0 ug/ml total phosphate.

**Dephosphorylation of SM proteins**

Whole SM was dephosphorylated by incubating 1 mg of dry weight protein with 25 units of insoluble alkaline phosphatase (Sigma) in a reaction buffer containing 1 mM ZnCl$_2$, 1 mM MgCl$_2$, and 50 mM glycine pH 10.4. The reaction was incubated at 37°C for 24 h with continuous rocking. After incubation, enzyme was removed from the reaction suspension by centrifugation. The supernatant was collected and dialyzed in dialysis cassettes (10K MWCO) against three changes of dH$_2$O to remove liberated phosphate. Finally, the samples were lyophilized, weighed, and total residual phosphate was determined by methods described above. Control experiments were run in parallel omitting the alkaline phosphatase.
Carbohydrate Analysis

Carbohydrate content was analyzed by two methods. Weight estimation was obtained using the Glycoprotein Carbohydrate Estimation Kit (PIERCE). Glycoprotein is oxidized with sodium meta-periodate. The aldehyde formed in this reaction reacts with a proprietary detection reagent to form a colored (purple) product. Absorbance is read at 550 nm and is proportional to the percentage of carbohydrate component in the protein. A standard curve is generated using proteins of known total carbohydrate amount.

Additionally, the DIG Glycan/Protein Double labeling kit (Boehringer Mannheim) was used to qualitatively label protein and glycoproteins simultaneously from blots following separation on SDS-PAGE gels. This method also uses sodium meta-periodate to oxidize sugar hydroxyl groups to form aldehydes. A hapten (DIG-Hydrazine) covalently interacts with the aldehydes which are then detected with an anti-DIG antibody conjugated to horseradish peroxidase. Development is carried out in TETON (4-Triethylenetrioxy-1-naphthol) substrate. Although staining of protein and glycoprotein can be carried out simultaneously, it was carried out sequentially in order to maximize the sensitivity for identifying glycoproteins. Once blots identifying glycoproteins were photographed, the blots were stained to identify non-glycosylated proteins so that both protein types were visualized at the same time.
Results

Extraction and quantification of matrix protein

Extraction of matrix protein from foliated shell was carried out according to the method of Mount (1999), which resulted in two distinct protein fractions. The water insoluble matrix (IM) fraction constitutes the bulk of matrix recovered and is dense and brown in color following lyophilization. The water soluble matrix (SM) fraction protein is comparatively light in weight and color. An additional protein fraction was extracted from the IM fraction using urea but the quantity extracted from this treatment was comparatively low. Over-all, the total amount of protein within the shell is approximately 1% protein to mineral which is consistent with previous studies (Mount, 1999; Rusenko, 1988). SM constitutes approximately 10 to 20% of the total protein. The IM fraction makes up the bulk of the protein and comprises the remaining 60 to 70% of the total protein extracted. About 2% of the IM can be solubilized with urea and SDS; this fraction is termed solubilized IM (sIM).

Protein was best quantified by measuring the dry weight of the SM extract instead of methods that utilize biuret chemistry, such as BCA (PIERCE). Protein estimation by this method relies mainly on the presence of cysteine, cystine, tryptophan, or tyrosine. Because these residues are not present in SM or IM protein in any appreciable amount (Rusenko, 1991; Mount, 1999), color formation (used as a signal for protein) is not indicative of the true quantity of protein. Consequently, BCA weight estimations for
soluble matrix protein fractions are significantly under-estimated. Over the course of this study, a ratio of soluble protein estimated by BCA to dry weight was determined to be only 20 to 25% (Table 2.1). For this reason, when BCA weights were used, it was assumed that the actual amount of protein was approximately four-fold greater.

**Table 2.1 Chemical Analyses of the SM Protein Fraction**

<table>
<thead>
<tr>
<th>Chemical Analysis¹</th>
<th>SM</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCA Concentration</td>
<td>243.9 ± 34.1 ug/ml</td>
</tr>
<tr>
<td>Phosphate</td>
<td>151.5 ± 3.9 ug/mg</td>
</tr>
<tr>
<td>Carbohydrate</td>
<td>38.8 ± 13.5 ug/mg</td>
</tr>
</tbody>
</table>

¹ All analyses are based on 1.0 mg of lyophilized weight and are reported as averages ± standard deviations (N=3)

An interesting property of SM that bears mentioning is its tendency to precipitate from solution during dialysis, despite its hydrophilic nature. Precipitation also occurred when lyophilized protein was reconstituted in water and subjected to multiple freeze thaw cycles. Precipitated protein was difficult to dissolve even when urea was added to the extract. At this time it is not clear if this observation is an artifact solely due to freeze-thaw cycles (a process known to precipitate some proteins) and the process of dialysis (where proteins can precipitate when dialyzed against buffers with significantly lower ionic strengths), or if some component in the crude SM protein fraction causes this
association. In any case, the observed self-association of SM \textit{in vitro} may be relevant to its function in shell formation.

**Analytical gel electrophoresis**

Previous studies demonstrated that Stains-all, a cationic stain, was best suited for visualizing the acidic shell matrix proteins following gel electrophoresis compared to other stains including Coomassie blue and silver nitrate (Myers et al., 1996). Proteins that are highly anionic stain a deep purple-blue color while proteins such as BSA stain red-pink. Red staining proteins, however, tend to fade quickly once the gel is exposed to light, unlike proteins that stain blue which remain stable upon exposure to light. This stain has also been used to identify highly phosphorylated proteins (Campbell et al, 1983). Considering that phosphoproteins historically have been implicated as key components in the mineralizing process in both vertebrate and invertebrates (Krampitz et al., 1983; Weiner and Traub, 1984; Lowenstam and Weiner, 1989; Veis et al., 1991), and specifically have been shown to bind to crystals and inhibit crystallization \textit{in vitro} (Termine et al., 1980, Fujisawa et al., 1987 Wheeler et al., 1991), the Stains-all stain conveniently provides a method for identifying potentially functional proteins contained within the whole matrix extract.

When SM and solubilized sIM are resolved on a Tris-Glycine-SDS 12\% polyacrylamide mini-gel, two prominent stains-all positive bands consistently appear in both
extracts (Figure 2.1 gel of sIM and SM). The SM extract reveals a major, intensely stained band at approximately 48 kD and another minor band at 55 kD. This result was verified quantitatively using densitometry and BCA quantification where it was found that the 48 kD band is present in two to three times greater quantity than the 55 kD protein in SM (data not shown). Gel analysis of the sIM extract reveals the same two molecular weight protein bands; however, the pattern is reversed where the 55 kD band is the major component, and the 48 kD is the minor band. This pattern is consistent with that reported by Mount (1999). Notably, band staining intensity was variable from prep to prep where, occasionally, the 55 kD in SM was barely visible following stains-all staining. Likewise, the 48 kD band sometimes appeared to be absent or very faint in sIM. Only when gels “missing” the 48 kD or 55 kD bands were back-stained with silver nitrate, did the two bands appear in both SM and sIM preps. The fact that variable staining intensity was observed for different preparations indicated that some preps were more efficiently extracted than others. For example, variability in gel pattern and staining intensity was observed between preps that were hand ground as opposed to mechanically ground prior to extraction.
Considered together, three inferences can be made from the above observations.

First, the 48 and 55 kD bands identified in the SM and sIM are likely equivalent.

Evidence supporting this idea is that their relative quantities and staining intensities are variable and seem to correlate with how rigorously the matrix is extracted from the
mineral. That is, the 48 kD is readily solubilized from the mineral and apparently, more of the 55 kD protein comes with it depending on the efficiency of extraction. Second, the 55 kD protein has an exceptional affinity for the IM. This is evidenced by the need for urea to release it from the IM following extraction. Finally, the fact that some portion of the 55 kD protein always extracts with the 48 kD protein despite the fact that urea is required to extract most of the 55 kD protein, may indicate that there is a tight association between the 48 and 55 kD proteins. The apparent affinities of the 48 kD with the 55 kD, and the 55 kD with IM may be indicative of true protein interactions \textit{in vivo} and may be relevant to their functions in shell formation.

In addition to the 48 and 55 kD phosphoprotein bands, a broad molecular weight continuum of stains-all positive bands is visible from approximately 6 kD to 50 kD. Also, the wells of the lanes containing SM, and especially sIM, consistently and intensely stain indicating the presence of extremely high molecular weight proteins that are unable to penetrate the gel. The presence of high molecular weight polymeric proteins in SM have been reported previously (Rusenko, 1988) and the fact that they are resistant to disassociation by SDS, β-mercaptoethanol, and heat prior to electrophoresis indicates they may be covalently cross-linked.

When whole SM was resolved on a full length (20 cm) preparative gel and stained with 0.3 M CuCl$_2$, a complex pattern of protein bands was revealed spanning a molecular weight distribution from 9 kD to 200 kD (Figure 2.2). Because this staining method is a
“negative” stain and actually stains the gel matrix around protein bands, theoretically, all protein components within the range of detectability should appear. Besides the major and minor phosphoprotein bands at 48 and 55 kD, several other prominent bands are visible. Two relatively broad bands appear at 10 and 66 kD, and another medium width band appears at 34 kD. Several small bands are visible at 50 and 60 kD. Other minor bands are barely visible between 20 and 30 kD and between 98 and 120 kD. This pattern results only when large amounts of SM extract are run. The appearance of a prominent 10 kD band varies from prep to prep and may be an artifact in the form of degradation products created during the extraction process. In addition, a prominent band migrates with the gel front and contains small proteins (< 1 kD). This gel pattern illustrates well the heterogeneity of the SM.
Figure 2.2 Preparative electrophoresis of whole SM.
Approximately 3 mg of dry weight protein was loaded into a single lane of a full length (approximately 20 cm) 12% SDS-PAGE gel and stained with 0.3 M CuCl₂ for 10 m. Following staining, the 48 and 55 kD bands were excised from the gel and electro-eluted. Mark 12® wide range protein standards (Invitrogen) were run in the reference lane and are labeled in kilodaltons.

In Figure 2.3, whole SM was resolved on a full length preparative 12% Tris-glycine-SDS gel and the protein was visualized by four different staining techniques. By staining sections of the same gel with each stain, the heterogeneity of the SM extract and to some extent, chemical characteristics of individual protein bands was illustrated simultaneously by their ability to interact with various stains. Section A was stained with
CuCl₂ and reveals a similar pattern to Figure 2.2. Most notably, the 48, 60, and 66 kD bands appear as does the 55, 34, and 10 kD bands. The absence of bands in Section A when compared to Figure 2.2 is due to the lower protein load used for this experiment which was necessary to avoid the appearance of overload when other stains were used. Section B was stained with Stains-all and, as expected, the only prominent bands that interacted with this stain were the 48 and 55 kD bands verifying their unique anionic character. Also, a continuum of minor bands was revealed spanning molecular weights of 6 kD to 48 kD. The added resolution of the full-length gel is evident here as this region, which typically appears as a smear on mini-gels, resolves a multitude of relatively discrete but minor bands. Section C, which was initially stained with Stains-all then back-stained with silver nitrate, again identified the prominent 48 and 55 kD proteins, and also identified the major bands at 60 and 66 kD. In Section D, where only silver stain was used to distinguish between anionic and more basic proteins, these same bands (60 and 66 kD) were identified, as well as a minor constituent at 50 kD, which was not identified by the other stains. These protein bands presumably have a more balanced amino acid composition compared to those of the stain-all positive bands, which are highly anionic. The 48 kD band stained lightly with silver due to the high protein load. Also noteworthy is the lack of staining by silver nitrate of the smear of proteins that appear following Stains-all staining indicating that most of the heterogeneity is found within the anionic class of proteins. Over-all, these results illustrate the complexity of the soluble matrix which is composed of a mixture of anionic proteins, including phosphoproteins, as well as proteins having a more basic character.
Figure 2.3 Comparison of whole soluble matrix (SM) protein patterns by differential staining.

Three mg of dry weight SM was loaded onto a full length (20 cm), 12% Tris-glycine gel, cut into four strips, and each strip was stained with one of the following stains. Mark 12 (Invitrogen) broad range molecular weight standards were run in the reference lane and are labeled in kilodaltons.

A - Copper chloride
B - Stains-all
C - Stains-all followed by silver nitrate
D - Silver nitrate

One factor found to contribute to matrix heterogeneity is the effect of freeze-thaw cycles on protein stocks once reconstituted. This was observed for gel purified proteins
where the more cycles protein stocks experienced correlated to an increase in band variability on gels. Shown in Figure 2.4, purified bands were subjected to a second round of electrophoresis after successive freeze-thaw cycles and resulted in “bleed off” of lower molecular weight proteins. When the 48 and 55 kD proteins were re-run immediately after electro-elution, no bleed off was evident. Evidently, the proteins undergo some structural change that alters their mobility (maybe loss of phosphate or carbohydrate groups), or, other proteins may co-migrate with the 48 and 55 kD bands that apparently resolve as a result of handling.
Figure 2.4 SDS-PAGE analysis of gel purified protein demonstrating bleed off of lower molecular weight bands following freeze-thaw.

The 48 and 55 kD protein bands were electro-eluted and re-run on a 12% Tris-glycine gel. Approximately 1ug of protein was loaded based on BCA weight estimates. The gel was stained with Stains-all.

Lane 1 - 55 kD protein run immediately after electro-elution
Lane 2 - 48 kD protein run immediately after electro-elution
Lane 3 - 48 kD re-run after 2 freeze-thaw cycles.
Lane 4 - 55 kD re-run after 2 freeze-thaw cycles.
IEF and 2D gel analysis

In an effort to further resolve the extent and origin of protein heterogeneity within the SM, and to determine if the 48 and 55 kD phosphoprotein bands were comprised of a single discrete protein, SM was subjected to IEF (isoelectric focusing) and 2-D gel analysis. To determine if varying degrees of phosphorylation contribute to heterogeneity observed for SM and the 48 and 55 kD proteins, SM was dephosphorylated with alkaline phosphatase (SM-AP), and run in parallel experiments with untreated SM. Considering the high anionic nature of SM proteins, studies were initially carried out using vertical pre-cast IEF gels pH 3 to 7 with a pI performance range of 3.0 to 6.5. Although gel patterns under these conditions were quite variable and the protein bands were distorted, it was determined that at least in part, protein heterogeneity identified by SDS-PAGE was in fact due to varying amounts of phosphate contained on the major matrix phosphoproteins (Figure 2.5). This finding was evidenced by the substantial decrease in the pI and number of Stains-all positive bands resulting from dephosphorylated SM compared to untreated SM. Whole SM generated three major stains-all positive bands with estimated pI’s of 5.0, 4.5, and 4.4. A less intensely stained, broad band appears in the pI range of 3.8. Conversely, dephosphorylated SM resolved only one prominent band with a pI of 4.3 and broad lighter bands 4.5 and 3.8. Phosvitin was included as a phosphoprotein control to observe how phosphoproteins generally behave when subjected to IEF. Staining continued with the dye front for both SM-AP and phosvitin indicating that the performance range of the gel on the lower pI range was exceeded.
Still, the SM contains a remarkable class of proteins with exceptionally high anionic character with pI’s approaching 2 and below. The fact that dephosphorylated SM resulted in fewer (lighter) bands as well as a shift down in pI suggests that some protein heterogeneity is indeed caused by varying degrees of phosphate.

Figure 2.5 Isoelectric focusing of SM and SM dephosphorylated with alkaline phosphatase (SM-AP) with pH range of 3 to 7.
Protein samples were loaded as dry weights. A single IEF gel was run at 100V for 1h, 200V for 2h, 500V for 30min, and cut in two pieces. Phosvitin (Sigma), a heavily phosphorylated protein, was used as a control. Gel piece “A” was stained with Stains-all. Gel piece “B” was stained with Coomassie brilliant blue 250. The pl ladder for IEF standards with range 3.6 to 6.6 is labeled.
Lanes a - IEF standard mix pI range 4.45 to 9.5 (Bio-Rad)
Lane b - IEF standard mix pI range 3.6 to 6.6 (Sigma)
Lane 1 - SM 25 ug
Lane 2 - SM-AP 25 ug
Lane 3 - phosvitin 5 ug
Following IEF, both SM and SM-AP were subjected to 2D analysis on 4 to 20% gradient gels. Only two prominent stains-all positive spots were discernable at 45 and 42 kD (Figure 2.6). It was speculated that the 45 kD spot was analogous to the 48 kD band identified by SDS-PAGE (verified below). However, the conspicuous presence of a 42 kD spot not normally observed by SDS-PAGE gels following Stains-all staining, and the absence of a 55 kD spot was unexpected and may indicate a shift in mobility of the 55 kD band or a “bleed-of” product from the 45 kD protein due to the IEF and 2D process (Figure 2.4). Still, both spots appeared to be discrete suggesting that they may be homogenous. Also, because no other phosphoprotein spots appeared, the molecular weight continuum of bands identified by SDS-PAGE (apparent as a smear here) may be largely artifact.
Figure 2.6  Two-dimensional analysis comparison of SM and SM dephosphorylated with alkaline phosphatase (SM-AP) after Stains-all staining.

Lanes from IEF gels (pI range 3 to 7) were excised and applied to a preparative well of 4 to 20% Tris-Glycine gradient gels. Arrows, drawn perpendicular to the top of gel A, define the boundaries of the pH range. Twenty-five ug of dry weight protein was run on the first dimension gel. Mark 12 molecular weight standards (Invitrogen) were run in the reference lane but are not visible following Stains-all staining.

A - SM
B - SM-AP

Because IEF results using the pH 3 to 7 system were often distorted, and the banding pattern highly variable, similar experiments were run using gels cast with an electrically neutral (zwitterionic) detergent, 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS) with a broader pH range of 3 to 10. This system was chosen because the band distortion and pattern variability that occurred using the pH 3 to 7 IEF gels was thought to be attributed to premature protein precipitation as the proteins and electrolytes moved towards their respective isoelectric points. By including CHAPS in the gels and in the sample buffer, it was anticipated that the protein would remain soluble until it reached its isoelectric point. Under these conditions, the gel patterns of SM protein were consistent and the band distortion was completely eliminated. However, the pI values of the SM and phosvitin control were significantly higher than those obtained with the pH 3 to 7 IEF gels. The considerable increase in pI values for the matrix proteins and phosvitin obtained when CHAPS was included in the media suggested that this additive, although non-ionic, may have affected the proteins causing them to focus in a higher than normal pI range. To address this question, the same proteins were focused on a pre-cast pH 3 to 10 IEF gel (Invitrogen) under the same
conditions excluding CHAPS. Although the bands were highly diffuse, the SM proteins and phosvitin focused in the same pI range as the gels with CHAPS included. This experiment clearly demonstrated the enhanced band resolution imparted by the CHAPS and the increased pI values were apparently due to the broader range of the gel. Because banding patterns were highly consistent and all the components resolved within the range of the standards, this system was best for exploring the questions concerning protein heterogeneity; however, pI values were assumed to be over-estimated by at least 2 fold.

Shown in Figure 2.7A are the protein patterns of the 48 kD (lane 1), SM (lane 2), and SM-AP (lane 3) on pH 3 to 10 CHAPS IEF gels following Stains-all staining. Apparent in the inset, the 48 kD resolved a prominent doublet with pI’s 7.2 and 7.1 and another prominent band at pI 6.8. Minor bands were identified at pI 7.4, 6.6, and 6.4. Notably, the SM resolved a nearly identical pattern to the 48 kD with only slight differences in pI values. The most visible difference was the low quantity of the pI 6.6 and 6.4 bands in the SM which is likely attributed to the fact that whole SM is comprised of a more complex mixture of proteins so that the same bands found in the 48 kD sample are in higher quantity relative to the whole SM protein. Still, a band with pI 6.9 appears to be more prominent in the SM.
Figure 2.7 Isoelectric focusing of matrix proteins on gel supplemented with CHAPS with a pH 3 to 10 range.

Prior to focusing, the 48 kD protein was gel purified from SM. SM protein was dephosphorylated using alkaline phosphatase (SM-AP). Protein was loaded based on dry weight except for the 48 kD protein, which was loaded using a BCA weight estimate. A single IEF gel was run at 100V for 1 h, 200V for 2 h, 500V for 30 min. and cut in two pieces. A- gel stained with Stains-all. B- gel stained with Coomassie brilliant blue 250.

Lane assignments are identical and are as follows. Standard pI ladders are labeled.
Lane a - IEF standards pI 4.45 to 9.5 (Bio-Rad)
Lane b - IEF standards pI 3.6 to 6.6 (Sigma)
Lane 1 - 48 kD protein 0.5 ug
Lane 2 - SM 25ug
Lane 3 - SM-AP 25ug

Inset shows enlarged section from gel slice A. Arrows indicate the bands for which pI’s were determined and are as follows:

<table>
<thead>
<tr>
<th>Lane 1 - 48 kD</th>
<th>Lane 2 - SM</th>
<th>Lane 3 - SM-AP</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.4</td>
<td>7.5</td>
<td>7.1</td>
</tr>
<tr>
<td>7.2</td>
<td>7.2</td>
<td>6.8</td>
</tr>
<tr>
<td>7.1</td>
<td>7.1</td>
<td>6.5</td>
</tr>
<tr>
<td>6.8</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>6.6</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When SM is compared to SM-AP, a marked difference in protein pattern resulted. While the SM resolved a continuum of bands spanning a pI range of 7.5 to 6.8 with a prominent doublet band appearing at pI 7.2 and 7.1 and additional bands at pI 6.8 and 6.9, the SM-AP resolves only two major bands at pI 6.8 and 6.5. It appears that the SM doublet (7.1 and 7.2) tightens after dephosphorylation resulting in a downward shift of the doublet to a single large band with pI 6.8 in SM-AP. Therefore, a shift of approximately 0.4 pI units occurred to this protein following dephosphorylation.

Similarly, assuming the 6.8 band in SM is equivalent to the 6.5 band in SM-AP, a downward shift by 0.3 units resulted from removing phosphate from this protein. Thus
the net negative charge of these proteins is increased when phosphate is removed and the doublet pattern tightened indicating that phosphorylation contributes to band complexity. Moreover, computer analysis of staining density revealed that the number of bands that focused in the SM-AP decreased when compared to those of untreated SM and is further evidence that differences in phosphorylation contribute to over-all matrix heterogeneity. In Figure 2.7B, a duplicate IEF gel was stained with Coomassie blue to identify other protein components not reactive with Stain-all. No additional bands stained and only faint staining was observed for the pI 6.8 bands in all lanes and the pI 6.9 bands in the SM and SM-AP. The prominent Stains-all positive doublet bands at pI 7.2 and 7.1 do not stain with Coomassie blue because of their anionic character.

Following IEF, the lanes resolving 48 kD, SM, and SM-AP were further analyzed on second dimension 4 to 20% gradient gel and stained (Figure 2.8). Here again, only two stains-all positive spots consistently appear on 2D gels for both SM and SM-AP despite heterogeneity revealed by IEF. This pattern occurred 100% of the time, regardless of the pH range of the gel or addition of additives in the first dimension (IEF). In figure 2.8A, SM resolves two stains-all positive spots at 45 and 42 kD with estimated pI’s of approximately 7. The same spots appear when SM-AP is resolved (Fig. 2.8 C); however, as expected, the spots had shifted slightly to a lower pI of 6.6. Also, the molecular weights shifted from 45.2 kD to 42.6 kD for the prominent spot, and from 41.9 kD to 38.6 kD for the smaller spot (see inset). The pI shift from 7.0 to 6.6 is a difference of 0.4 units which is consistent with shifts of the pI 7 bands observed in IEF, pH range 3
to 7 (Fig 2.7). Also, the change in molecular weight for each spot by about 3 to 4 kD illustrates well the effect of removing phosphate groups from protein which allows the protein to move more freely through the gel. Consequently, the true molecular weights for matrix phosphoproteins are likely lower than reported.
Figure 2.8 Two-dimensional analysis comparison of SM and SM dephosphorylated with alkaline phosphatase (SM-AP).
Lanes from gels IEF (pI range 3 to 10) were excised and applied to preparative wells of 4 to 20% Tris-Glycine gradient gels. Arrows drawn perpendicular to the top of gel A defines the boundaries of the pH range. Twenty-five micro-grams of dry weight protein was run on the first dimension gel. SeeBlue® molecular weight standards (Invitrogen) were run in the reference well and are labeled in kilodaltons.

A - SM stained with Stains-all
B - Gel in “A” back-stained with silver nitrate
C - SM-AP stained with Stains-all
D - Gel in “C” back-stained with silver nitrate

Inset shows section of gels A (SM) and C (SM-AP) superimposed. Note the decrease in pI and molecular weight of the Stains-all positive 45 and 42 kD spots resulting from dephosphorylation. A pI shift of 0.4 units was estimated as well as a molecular weight shift of approximately 3 to 4 kD for each spot.

In another experiment, purified 48 kD protein was resolved by 2D electrophoresis and verified that the 45 kD spot was, indeed, equivalent to the 48 kD protein identified by SDS-PAGE (data not shown). Also, it was determined that the doublet band appearing in the SM and 48 kD IEF gels at about pI 7 resolves as the major spot appearing at 45 kD on the 2D gels. Although the spot appears to be discrete, the potential for very closely related protein isoforms is possible given the broadness of the spot and that it appears as a tightly spaced doublet on IEF gels. Unresolved questions include why band heterogeneity observed following IEF did not result in multiple (corresponding) spots in the second dimension. Also, the absence of a spot corresponding to the 55 kD protein identified by SDS-PAGE is puzzling. Several explanations are possible. The 55 kD protein may migrate more quickly following such manipulations resulting in its migrating to a lower molecular weight- in this case to 42 kD or it could co-migrate with the 45 kD
spot. It is also possible that the 55 kD protein did not resolve under the 2D experimental conditions. If this is the case, then the 42 kD spot may be equivalent to a 38 kD Stains-all positive protein occasionally visible in SDS-PAGE that for unknown reasons resolves consistently under conditions of IEF and 2D electrophoresis; or it could represent “bleed off” from the 45 kD spot (refer to figure 2.4).

Following Stains-all staining, gels in Figure 2.8 A and C were back-stained with silver nitrate to identify any additional spots that were not detectable with Stains-all alone. Three additional spots appeared on the gel resolving SM (Figure 2.8 B) and one additional spot appeared on SM-AP (Figure 2.8 D). These same spots were most apparent when gels were stained with silver nitrate alone (Figure 2.9 A and B). In this figure, a complex pattern is evident on both SM and SM-AP gels. From this data, it appears that a 35 kD band has two isoforms (pI 6.2 and 5.3), a 20 kD band appears to have four isoforms (pI 8.6, 7.8, 6.1, and 5.8), and a 17 and 10 kD band resolves at least two isoforms each with pI’s of 5.8 and 6.1. The molecular weights of these proteins are fairly consistent with those identified by CuCl2 staining in Figure 2.3 where two large bands were identified at 34 and 10 kD, and another minor band at 20 kD. The absence of a shift in their position in the gels following dephosphorylation (Figure 2.9 B), and there inability to stain with stains-all suggests that these spots are not phosphoproteins or highly anionic proteins. Furthermore, micro-heterogeneity within this presumably more basic protein class seemed to be greater in the untreated SM revealed by the broader spots. It may be that the conditions of dephosphorylation, at pH 10, (rather than the
action of the enzyme because these proteins are not phosphoproteins) may have reduced micro-heterogeneity through base hydrolysis. Or it may be that staining is reduced by the alkaline conditions causing the spots to stain less prominently. An opposite affect occurred to the anionic 45 and 42 kD bands as these proteins stain better with silver nitrate following dephosphorylation which is likely due to the removal of phosphate.
Figure 2.9  Two-dimensional analysis comparison of SM and SM dephosphorylated with alkaline phosphatase (SM-AP) following silver staining.
Lanes containing 25 ug dry weight protein from IEF (pI range 3 to 10) were applied to preparative wells of 4 to 20% Tris-Glycine gradient gels. Arrows drawn perpendicular to the top of the gel defines the pI range boundaries. SeeBlue® molecular weight standards (Invitrogen) were run in the reference well and are labeled in kilodaltons.
A - SM  
B - SM-AP

A puzzling observation is the absence of the high molecular weight proteins (the 60 and 66kD silver positive proteins identified by SDS-PAGE) on 2D gels stained with silver. This considered with the absence of the 55 kD band on stains-all stained 2D gels raised the question of how various manipulations affect protein mobility and illustrates the difficulty in identifying component proteins from shell derived matrix using biochemical analyses. Still, the data clearly illustrate micro-heterogeneity within single molecular weight proteins, especially in the more basic class. Differences in phosphate content contribute, in part, to heterogeneity in the anionic class of proteins.

Amino Acid Analysis

The amino acid compositions of gel purified 48 and 55 kD proteins, as well as whole extracts of SM and IM (not to be confused with sIM, a urea soluble extract from IM), were analyzed. At a glance, the most striking feature in Table 2.2 is that all the proteins analyzed are composed largely of three amino acids, Asp, Ser, and Gly, each in a ratio of roughly 30%. This is particularly interesting in light of the matrix complexity revealed in the above electrophoresis studies. Moreover, the 48 and 55 kD bands are gel
purified bands from the SM, yet still highly resemble whole extracts of SM and IM. The remaining 10% of the compositions are distinguished essentially by varying degrees of hydrophobic residues. For example, IM has a slightly higher amount of hydrophobic groups and a slightly lower Asp content, which is reflected in a smaller C/HP ratio for IM compared to SM. The marginal distinction between these two protein classes is verified by their low SΔN index (0.38) reported in Table 2.3, indicating they are related.
Table 2.2 Amino Acid Composition of Matrix Fractions and Gel Purified Matrix Proteins

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>IM</th>
<th>SM</th>
<th>48 kDa</th>
<th>55 kD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>22.88</td>
<td>29.65</td>
<td>31.90</td>
<td>31.69</td>
</tr>
<tr>
<td>Thr</td>
<td>1.28</td>
<td>1.73</td>
<td>0.84</td>
<td>0.92</td>
</tr>
<tr>
<td>Ser</td>
<td>25.80</td>
<td>23.21</td>
<td>19.42</td>
<td>23.64</td>
</tr>
<tr>
<td>Glx</td>
<td>7.03</td>
<td>6.85</td>
<td>7.04</td>
<td>6.44</td>
</tr>
<tr>
<td>Pro</td>
<td>1.32</td>
<td>1.63</td>
<td>0.98</td>
<td>1.93</td>
</tr>
<tr>
<td>Gly</td>
<td>30.22</td>
<td>25.52</td>
<td>32.84</td>
<td>29.48</td>
</tr>
<tr>
<td>Ala</td>
<td>3.40</td>
<td>3.05</td>
<td>1.64</td>
<td>1.45</td>
</tr>
<tr>
<td>Val</td>
<td>1.54</td>
<td>1.40</td>
<td>1.15</td>
<td>0.35</td>
</tr>
<tr>
<td>Met</td>
<td>0.19</td>
<td>0.20</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Ile</td>
<td>0.46</td>
<td>0.39</td>
<td>0.23</td>
<td>0.19</td>
</tr>
<tr>
<td>Leu</td>
<td>1.28</td>
<td>0.95</td>
<td>0.55</td>
<td>0.59</td>
</tr>
<tr>
<td>Tyr</td>
<td>2.08</td>
<td>3.08</td>
<td>2.59</td>
<td>2.51</td>
</tr>
<tr>
<td>Phe</td>
<td>0.32</td>
<td>0.36</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>His</td>
<td>0.50</td>
<td>0.55</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Lys</td>
<td>1.61</td>
<td>1.12</td>
<td>0.35</td>
<td>0.81</td>
</tr>
<tr>
<td>Arg</td>
<td>0.88</td>
<td>0.52</td>
<td>0.48</td>
<td>0.00</td>
</tr>
</tbody>
</table>

| C/HP ratio² | 3.87 | 4.85 | 8.74 | 8.63 |

¹ Values reported as mole percent
² Charged residues (asp, glu, lys, arg, his) / Hydrophobic residues (pro, ala, val, met, ile, leu, phe)
Table 2.3 Comparison of Proteins and Protein Fractions by Difference Index

<table>
<thead>
<tr>
<th></th>
<th>SM</th>
<th>48 kD</th>
<th>55 kDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM</td>
<td>0.38</td>
<td>0.68</td>
<td>0.46</td>
</tr>
<tr>
<td>SM</td>
<td></td>
<td>0.39</td>
<td>0.13</td>
</tr>
<tr>
<td>48 kD</td>
<td></td>
<td></td>
<td>0.16</td>
</tr>
</tbody>
</table>

Numerical values ($S\Delta N$) determined by the equation $S\Delta N = 1/2 \sum (N_{iA} - N_{iB})^2$, where $N_{iA}$ and $N_{iB}$ is the quantity (mole percent) of the $i^{th}$ type of amino acid contained in protein A or B. If $S\Delta N < 0.42 \times N$ ($N = 100$ when using mole percent), then there is a 95% certainty that the proteins are related. If $0.42 < S\Delta N < 0.93 \times N$, then the proteins are weakly related. If $S\Delta N > 0.93 \times N$ then there is a 95% certainty that the proteins are unrelated.

Similarly, the 48 and 55 kD proteins are highly related having a $S\Delta N$ index of 0.16 and C/HP ratios of 8.74 and 8.63 respectively. The higher C/HP ratios of these proteins compared to whole SM and IM, and the high Ser content which carries the potential for phosphorylation, especially illustrates the remarkable anionic potential for the component proteins of SM.

Another observation that stands out when comparing protein difference indices is the slightly higher identity of the 55 kD protein to both the IM and SM. This suggests that the 55 kD protein has more chemical character in common to both SM and IM protein classes. The 48 kD protein appears to have more chemical character in common with SM than to IM. The 55 kD shows a higher identity with IM than does the 48 kD, and even borders the criteria for being “related” to IM having an index of 0.46. Still,
both proteins fall into the “weakly related” category when compared to IM and suggests that some part of IM is distinguished from the 48 and 55 kD proteins.

The higher identity of the 55 kD with the IM becomes more compelling when considered with the composition data reported in Table 2.4. In this experiment, samples were dialyzed briefly against water on Immobilon P prior to hydrolysis. This step was carried out to remove residual salt that had not been effectively removed from the samples following electro-elution. Interestingly, this treatment resulted in significantly different amino acid compositions for the 48 and 55 kD proteins and relatively poor yields. Only 29% of the 48 kD protein was recovered after dialysis demonstrating that this protein has a very low affinity for this membrane type, which adsorbs protein largely through hydrophobic interactions. A similar observation was made previously when western blot and Edman sequencing analyses of the 48 kD protein were carried out on PVDF, a nylon membrane which also adsorbs protein through hydrophobic interaction (Discussed in Chapter 3).
Table 2.4 Amino Acid Compositions of 48 and 55 kD Proteins Before and After Dialysis on Immobilon P®

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>48 kD</th>
<th>(dialyzed)$^2$</th>
<th>55 kDa</th>
<th>(dialyzed)$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>48 kD</td>
<td>55 kDa</td>
<td></td>
<td>55 kDa</td>
</tr>
<tr>
<td>Asp</td>
<td>31.90</td>
<td>11.18</td>
<td>31.69</td>
<td>20.82</td>
</tr>
<tr>
<td>Thr</td>
<td>0.84</td>
<td>5.41</td>
<td>0.92</td>
<td>3.64</td>
</tr>
<tr>
<td>Ser</td>
<td>19.42</td>
<td>10.82</td>
<td>23.64</td>
<td>13.51</td>
</tr>
<tr>
<td>Glu</td>
<td>7.04</td>
<td>11.72</td>
<td>6.44</td>
<td>9.73</td>
</tr>
<tr>
<td>Pro</td>
<td>0.98</td>
<td>7.22</td>
<td>1.93</td>
<td>3.60</td>
</tr>
<tr>
<td>Gly</td>
<td>32.84</td>
<td>21.60</td>
<td>29.48</td>
<td>25.22</td>
</tr>
<tr>
<td>Ala</td>
<td>1.64</td>
<td>5.19</td>
<td>1.45</td>
<td>3.23</td>
</tr>
<tr>
<td>Val</td>
<td>1.15</td>
<td>4.44</td>
<td>0.35</td>
<td>3.03</td>
</tr>
<tr>
<td>Met</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Ile</td>
<td>0.23</td>
<td>3.04</td>
<td>0.19</td>
<td>2.15</td>
</tr>
<tr>
<td>Leu</td>
<td>0.55</td>
<td>7.52</td>
<td>0.59</td>
<td>5.12</td>
</tr>
<tr>
<td>Tyr</td>
<td>2.59</td>
<td>1.91</td>
<td>2.51</td>
<td>2.21</td>
</tr>
<tr>
<td>Phe</td>
<td>0.00</td>
<td>1.59</td>
<td>0.00</td>
<td>0.98</td>
</tr>
<tr>
<td>His</td>
<td>0.00</td>
<td>1.17</td>
<td>0.00</td>
<td>0.70</td>
</tr>
<tr>
<td>Lys</td>
<td>0.35</td>
<td>3.63</td>
<td>0.81</td>
<td>3.31</td>
</tr>
<tr>
<td>Arg</td>
<td>0.48</td>
<td>3.55</td>
<td>0.00</td>
<td>2.76</td>
</tr>
<tr>
<td>C/HP ratio</td>
<td>8.74</td>
<td>1.08</td>
<td>8.63</td>
<td>2.06</td>
</tr>
<tr>
<td>% recovery$^3$</td>
<td>-</td>
<td>29%</td>
<td>-</td>
<td>61%</td>
</tr>
</tbody>
</table>

$^1$ Values as mole percent
$^2$ Prior to analysis, protein samples were applied to Immobilon P to eliminate residual glycine
$^3$ Percent protein recovered following dialysis on Immobilon P membrane
Conversely, the 55 kD protein demonstrated significantly greater adsorption to the membrane as 61% of the total protein was recovered. This marked difference in adsorption to the hydrophobic membrane may be relevant to function. Considering that the 55 kD shares more identity with the IM (in terms of the presence of hydrophobic groups) and that the 55 kD adsorbs better to the membrane, this observation could reflect its position *in vivo* where it may interact directly with IM (which is relatively more hydrophobic). In other words, this protein may be structurally optimized to interact with hydrophobic substrates through hydrophobic domains, despite its anionic character, thus supporting a role for its presence on the surface of IM.

Most striking from this experiment is the difference in compositions obtained when comparing the dialyzed samples with untreated samples. Both showed a marked decrease in Asp and Ser. This effect was most pronounced for the 48 kD protein. In addition, the amount of hydrophobic residues increased so that the C/HP ratios for the 48 and 55 kD proteins dropped from 8.74 and 8.63 to 1.08 and 2.06 respectively. Considered with the recovery data, two explanations are possible. First, it is possible that more than one protein co-migrates with the 48 and 55 kD bands. In the case of the 48 kD band, the protein analyzed from the blot is 29% of the total protein within that band and is significantly more hydrophobic. The 55 kD protein retained on the membrane represents over half of the total protein within the band and is also more hydrophobic. Additional evidence supporting the presence of protein isoforms with similar molecular weights comes from electrophoresis data demonstrating bleed-off of other lower
molecular weight bands that were re-run on gels (Figure 2.4). Furthermore, analytical IEF demonstrated that protein heterogeneity is due only in part to varying degrees of phosphorylation and therefore supports the potential for protein isoforms. Admittedly, 2D data is less convincing of band heterogeneity although the broadness of the spots may indicate closely related isoforms.

On the other hand, the failure of 2D analysis to identify clearly micro-heterogeneity within the stains-all positive spots suggests that the 48 and 55 kD bands may be discrete and leaves open the possibility that the difference in compositions could simply be attributed to degradation from handling prior to and during dialysis on the Immobilon P. The observation that additional bands bleed off of parent bands following freeze thaw cycles can also be used to argue for degradation occurring to a single protein. In this case, the differences in compositions for both proteins following dialysis is attributed to the more hydrophobic regions adsorbing to the membrane and the anionic regions elute into the dialysate. Therefore, it is tempting to suggest that these proteins are in part comprised of acidic domains separate from more hydrophobic domains, a protein arrangement common to structural proteins, and is being observed in many matrix proteins derived from mineral as more sequence information becomes available. An alternating and repetitive bipolar arrangement creates a linearized region, such as helical or pleated sheet structures, which maximize surface area for interacting with other proteins or surfaces. In enamel formation, Fincham et al. (1991) proposed that this property allows amelogenin monomers to align and aggregate in such a way as to adsorb
onto specific crystal surfaces promoting c-axial growth. Also, the loss of phosphate and other groups may affect the adsorption of the protein to the membrane which may explain the different affinities of and within the 48 and 55 kD proteins following dialysis.

Also potentially contributing to micro-heterogeneity within the 48 and 55 kD bands is minor contamination of protein fragments that are present in the SM. In addition to the 48 and 55 kD proteins, the SM resolves a molecular weight continuum of protein fragments spanning nearly the entire length of the gel. Upon staining, this continuum of protein fragments appears as a background smear and indicates that there is non-specific fragmentation that occurs to the matrix proteins. This degradation could have occurred during the extraction procedure and very likely to some extent, within the shell itself. Considering that the shell is not living tissue and that shells of mollusks live in aqueous environments, some loss of protein integrity is not surprising, especially for the older regions of the shell. So, knowing that this background protein exists; protein fragments that co-migrate with the 48 and 55 kD bands consequently would be excised and eluted together. Still, the contribution of these components is minor according to results of IEF and 2D analysis which resolve trace to no protein contamination within the 48 kD or 55 kD proteins.

Over-all, it is clear that these proteins and their putative isoforms are highly related and the fact that protein compositions of the 48 and 55 kD component proteins
highly resembles those of the whole SM and IM suggests that this type of protein (highly anionic and repetitive) is the dominant protein class within both the SM and IM.

**Carbohydrate and Phosphate Analyses**

Phosphorylation and glycosylation of SM was estimated using several methods. Total phosphate was quantitatively determined on a weight basis by method of Swift (1992) and detects total phosphate in the range of 0.002 µg/ml to 1.1 µg/ml. Reported in Table 2.1, total phosphate was measured to be 155 µg/1 mg of dry weight SM; that is 15% of dry weight SM is phosphorylated, by composition mostly in the form of phosphoserine. These results are in agreement with those of Borbas et al. (1991), and Rusenko (1988) who reported similar phosphate concentrations for SM and fractions thereof. Based on the mole percent of serine (23.2) and threonine residues (1.7) reported for SM in Table 2.2, it is estimated that 70% of these residues are phosphorylated. Attempts to quantify phosphate for the 48 and 55 kD proteins were made but failed to produce a suitable signal for phosphate. This result may be attributed to the gel purification process, where the proteins are exposed to pH values approaching 9 for over 6 h. At this pH, phosphate groups are susceptible to base hydrolysis. Using alternative gel systems which operate near neutral pH may remedy this problem.
Carbohydrate was analyzed by two methods. First, carbohydrate content was estimated by comparing absorbance of sample proteins with those of known carbohydrate content. The second method is comparatively qualitative and identifies glycosylated proteins within a sample after it is separated on a gel and transferred to nitrocellulose. Glycoproteins and proteins are then identified simultaneously by immunoassay with a detection limit of 10 ng for glycoprotein and 50 ng for proteins. Although this method is not quantitative, it has the advantage of identifying discrete glycoproteins within a crude protein extract. Using the former technique, the carbohydrate content for whole SM was estimated to be between 3 to 5% on a dry weight basis (Table 2.1). Immuno-staining blots of SM for carbohydrate revealed the 48 kD band and a 34 kD band (also identified by CuCl₂ staining Figure 2.2) stained lightly compared to the control but was discernable above background (Figure 2.10). This indicates that these proteins are glycosylated to some extent. Also, a 98 kD band and several minor bands at 8, 14, and 30 kD lightly stain for carbohydrate. Interestingly, the 55 kD band was not visible above background which indicates that this protein has very little carbohydrate associated with it. When the same blot was sequentially developed for protein, the 48 kD band quickly changed brown while the 34 kD band remained blue. The time between when bands turn from blue to brown is a function of the amount of carbohydrate there is on a given protein. This is because over time, the staining reaction shifts in favor of protein staining. Therefore it was concluded that the 48 kD band is not as glycosylated as the 34 and 98 kD bands. The 8, 14, and 30 kD bands are also relatively sparsely glycosylated. Overall, SM does not
contain an abundance of discrete, heavily glycosylated proteins, especially when compared to the significantly higher phosphate content.

Figure 2.10 Glycoprotein identification compared to non-glycosylated proteins of SM.

Twenty-five micro-grams of dry weight protein was run on 4 to 20% Tris-glycine gradient gel and blotted onto nitrocellulose. In “A”, the blot was stained to identify glycosylated protein. In “B”, the blot was back-stained to identify non-glycosylated proteins relative to glycoproteins. The 48 kD protein in indicated by (→). Fetuin was run as a glycoprotein control; creatinase was run as a non-glycosylated protein control. The intensity of color is proportional to the amount of carbohydrate present on each protein.
Lane 1 - SeeBlue® protein standard (Invitrogen) labeled in kilodaltons
Lane 2 - SM
Lane 3 - Fetuin
Lane 4 - Creatinase
Discussion

The foliated (inner) shell layer of *Crassostrea virginica* contains an organic matrix composed of three distinct protein classes based on their solubility in water following demineralization in EDTA. The total organic matter extracted from oyster shell was approximately 1% of the shell weight. The IM, named for its insolubility in water, constitutes the bulk of the total organic matrix extracted and is about 70% of the matrix weight while the SM constitutes the remaining 30%. A urea and SDS solubilized fraction (sIM) is the least abundant protein class and constitutes only 1% of the total matrix. The relatively low amount of total organic matrix extracted from this species is consistent with previous studies (Rusenko, 1988; Kawaguchi and Watabe, 1993) and generally is typical of species that have foliated calcitic microstructure (Carter, 1980). Furthermore, the disproportionate amount of IM to SM, in this case about 5:1 is a trend observed often in both invertebrate and vertebrate systems (Lowenstam, 1989). In contrast, Kawaguchi and Watabe (1993) reported that SM is in higher proportion to IM in oyster shell following acid dissolution. This discrepancy illustrates how different extraction methods affect the amount of material recovered.

Differences in solubility alone suggest that the IM and SM protein fractions are fundamentally different and therefore perform distinct and separate functions. Specifically, the resistance of the bulk of the IM to solubilization by water and chaotropic agents, such as urea, suggests that this protein fraction is covalently cross-linked (Mount,
1999), a condition that is unique to this protein class and distinguishes it from the SM. Likewise, the release of additional protein (sIM) from the IM, using urea and SDS, suggests strong electrostatic interactions that are unique to this assemblage of proteins. In the shell, the IM protein class may be more resilient to degrading effects of the environment as indicated by the lower background staining of sIM on gels compared to SM and is further evidence of the strong interactions that occur within the IM class.

Paradoxically, despite the fundamental differences in solubility, SM is not easily distinguished from IM when their compositions are compared. Furthermore, electrophoretic analysis of component proteins patterns of SM and sIM are nearly identical and appear to vary mostly in their relative amounts. SM and sIM are, in part, composed of two prominent phosphoproteins with estimated molecular weights of 48 and 55 kD. The 48 kD band was found most abundantly in SM while the reverse is true for sIM, where the 55 kD band is the most prominent. (Mount, 1999). Considering that the 48 kD is water soluble and the 55 kD requires urea and SDS to be sufficiently extracted, this pattern illustrates the differences in affinity that each protein has for the IM protein and the mineral. That is, the prominence of the 48 kD band in the SM fraction confirms that this protein is readily released from the mineral upon EDTA dissolution, an observation that implies an intimate association of this protein to the mineral but not with the IM.
In contrast, because the 55 kD band is significantly less conspicuous in SM and requires urea to be effectively extracted suggests a very tight association of this protein with the IM. Furthermore, the relatively low level of the 55 kD protein present in the initial water soluble fraction (SM) infers a comparatively limited association of this protein to the mineral or the 48 kD protein. It is suspected that the 55 kD band, present as a minor phosphoprotein in SM, is the same protein apparent as a major protein component in the sIM fraction. This conclusion is supported by the fact its appearance in SM is enhanced depending on how well the protein is extracted and by their similar N-terminal sequences as discussed in Chapter 3. Likewise, the minor 48 kD protein, visible in the sIM fraction, corresponds to the 48 kD protein found in the SM fraction.

Interestingly, the amino acid compositions of the 48 and 55 kD proteins are strikingly similar, despite their obvious differences in molecular weight and solubility. Both proteins are rich in carboxyl groups contributed by a high aspartic acid and phosphorylated serine content, and as expected, are highly anionic as verified by their low pI values. The high serine content is similar to those reported for phosphophoryn, a class heavily phosphorylated proteins derived from tooth dentin that are purported to control hydroxapatite deposition (Fujisawa et al. 1987; Veis et al. 1991). Little negative charge is attributed to the presence of carbohydrate based on the low weight value estimated for SM; however, the 48 kD stained positive for glycosylation slightly in immunoassays. The 55 kD showed no staining demonstrating some distinction from the 48 kD protein. The presence or absence of such groups regarding function is interesting
in light of the recent discovery that matrix proteins from a variety of biominerals contain C-type lectin domains (reviewed in Mann et al., 2000). To date, only Perlucin, a C-type “lectin-like” protein derived from nacre of the abalone, has demonstrated carbohydrate binding (Mann et al., 2000) in vitro. Although speculative, such domains may promote specific interactions (such as assembly) between matrix glycoproteins and proteins containing lectin domains.

The similarities in composition suggest that the 48 and 55 kD proteins are highly related and likely originate from a single gene or are a part of a gene family. Gibson et al. (1991) found that related forms of amelogenins, a class of matrix proteins derived from tooth enamel, are produced from an alternatively spliced single mRNA. Matrix proteins derived from molluscan shell have demonstrated similar protein heterogeneity but an alternative splicing mechanism has not yet been identified. Perlucin has at least two variant forms distinguished by having varying in length and amounts of a repeat peptide sequences (Mann et al., 2000). A similar micro-heterogeneity was reported for lustrin A, a matrix protein derived from nacre of the abalone that contains a carbonic anhydrase domain (Shen et al., 1997). Most recently, families of highly related matrix proteins encoded by at least 8 distinct mRNAs have been found in the calcite layers of two bivalve species (Gotliv et al., 2005; Zhang et al., 2006; Yano et al., 2006).

The 48 and 55 kD proteins are major matrix components because their compositions highly resemble those of whole SM and IM. In fact, only when their...
compositions are statistically analyzed are subtle distinctions apparent. Over-all, the IM is distinct from the solubilized matrix proteins and contains slightly more hydrophobic groups. However, the 55 kD protein has a slightly higher identity to the IM than either the SM or the 48 kD protein. This result further supports a tight affinity of the 55 kD with the IM as residual 55 kD would contribute to the composition of IM resulting in a higher identity between the two. Also, the ability of the 55 kD protein to adsorb more effectively to a hydrophobic support suggests this protein has domains that more effectively interact with hydrophobic substrates and therefore may reflect its association with the IM in vivo. Accordingly, it is feasible that the 55 kD protein has hydrophobic domains that allow it to associate specifically with IM, while its more anionic domains interact with mineral surfaces. In bone formation, a class of small (Ms 25 to 28 kD), anionic matrix proteins with related structures, bind apatite and collagen effectively linking the mineral phase to the matrix. These proteins are distinguished by their relative affinities for hydroxyapatite which is presumed to correlate with variations in their structures. (Sodek et al., 1989). The SM, 48, and 55 kD proteins most certainly have common structural domains as their similarity indices identify them as highly related.

From these results, inferences can be made as to where each protein may be located within the mineral and, when considered with in vitro activities reported previously for shell matrix proteins, may provide some insight into their roles in shell formation. In early studies, microscopic examination of partially demineralized shell revealed that matrix localizes between (inter-crystalline) and is occluded within (intra-
crystalline) individual growing crystals (Watabe et al., 1965). Matrix in between crystals remained insoluble upon demineralization and largely intact leaving “imprints” resembling scaffolding (Nakahara et al., 1982). Proteins of this class, termed IM, tend to be polymeric and fibrous, and form thin, continuous sheets (Travis, 1965; Bevelander and Nakahara, 1980; Fritz et al., 1994). In fact, morphologically, shell IM has been likened to collagen and silk fibroin (Weiner and Addadi, 1997). In oyster, IM extract appears fibrous when observed microscopically (unpublished observation). Mount (1991) found that the cross-links present in IM are covalent and likely limited in number due to the exceptionally high water absorption capacity IM protein has on a weight to weight basis. This property has coined the term “gelling protein” to describe the IM (Mount, 1991; Sikes et al., 2000). Thus, the role of this protein class is assumed to be one of structure, where it is presumed to assemble prior to the onset of crystal formation, providing a “framework” or organizing mesh into (Nakahara, 1979) or onto (Erben and Watabe, 1974) which crystals grow.

In contrast, SM is thought to be mostly occluded within the crystals and associates with the immediate surface, as evidenced by its ready solubilization upon demineralization. Thus, it is presumed that this protein class interacts with crystals as they grow. This study and previous studies (Mount, 1999) demonstrate that the 48 kD SM phosphoprotein is easily released from the shell mineral and is therefore likely intra-crystalline matrix; it’s high anionic nature and the fact that it is poly-phosphorylated predicts that this protein component would interact well with the growing crystal
surfaces. The potential role of the 48 kD in mediating mineral growth in oyster shell comes from studies that demonstrate the ability of SM proteins to affect crystal morphology \textit{in vitro} by binding to specific crystal surfaces, allowing accelerated growth on faces that do not adsorb protein (Wheeler et al., 1981; 1988; Addadi and Weiner, 1985; Berman et al., 1988; Kono et al., 2000 Thompson et al., 2000). Moreover, a reverse phase HPLC fraction of SM termed RP-1 was previously shown to potently inhibit crystal growth in solution (Wheeler et al., 1981, 1988; Wheeler and Sikes, 1984) and \textit{in vivo} (Sikes and Wheeler, 1986). Myers (1999) identified the 48 kD protein as a major component of the RP-1 fraction.

Finally, a finding common to many systems is that some portion of the acidic matrix components are sandwiched between the surfaces of the crystals and the assembled IM (Lowenstam and Weiner, 1989). In this arrangement, the more acidic SM proteins essentially coat the relatively hydrophobic IM proteins often referred to as “core” proteins (Weiner et al, 1983; Bevelander and Nakahara, 1980; Watabe, 1965; Marin et al., 2005). In some mollusks, the interaction between the acidic macromolecules and the IM is apparently electrostatic yet quite strong as urea and similar agents are required for its release (Cariolou and Morse, 1988). Similarly, a variety of acidic proteins from bone and teeth such as phosphophoryn (Lee and Veis, 1980; Stetler-Stevenson and Veis, 1986) and osteonectin (Termine et al., 1981) bind tightly to collagen. In dentin, some phosphophoryn may be covalently bound to collagen because it remains even after repeated extractions (Lee and Veis, 1980).
The 55 kD protein fits well with the description of a “coating” protein. In this study and previous studies (Mount. 1999), the 55 kD protein is visible in the SM fraction but most abundantly in a urea extracted fraction (sIM). The fact that urea is required to liberate this protein indicates that it is tightly bound to the IM and, considered with its higher relatedness to IM, the 55 kD protein may coat the IM, potentially through common hydrophobic domains, ultimately localizing in the regions surrounding the crystal units as inter-crystalline matrix. This protein may be analogous to Lustrin A, a protein extracted from the IM of abalone shell nacre that is thought to act as an adhesive between aragonite tablets (Smith et al., 1999). Also, the apparently “residual” appearance of the 55 kD protein in the SM and the 48 kD in the sIM may be explained in that during crystal growth, the 48 kD protein would likely interact with the 55 kD protein at the interface between the mature crystal and the assembled IM, assuming that the 55 kD protein is located on the surface of the IM. Thus, some residual 48 kD in the sIM fraction (and 55 kD in the SM) following extraction would not be unanticipated considering the strong associations that are capable between the matrix proteins as demonstrated in this study.

In terms of function, the high anionic character of the 48 and 55 kD proteins and their respective affinity for shell mineral and IM suggests they play direct roles in shell formation. It is presumed those acidic macromolecules that are present on the assembled IM act to control mineral deposition in one of two ways: by acting as a nucleating sheet for initiating stereospecific crystal growth or by binding preformed crystals in an ordered manner. The first proposed function is born from studies where poly-anionic matrix
proteins were shown to promote crystal formation when immobilized on a solid support (Greenfield, 1984; Linde et al., 1989). More compelling are results from studies where proteins from shell layers, having different mineralogies, were reported to promote nucleation of the corresponding mineralogy (Falini et al., 1996; Feng et al., 2000). The significance of these findings, however, is overstated because experimental conditions used in these studies were not reflective of the ionic conditions of the mineralizing medium in vivo (Crenshaw, 1972a). In fact, when similar experiments were performed at physiological ionic strengths, neither SM (Sikes et al., 1998, 2000) or IM (Mount, 1999) is able to induce nucleation.

An alternative to the nucleation hypothesis is that assembled matrix binds preformed crystals (Wheeler et al., 1987). SM strongly inhibits crystal growth under physiological conditions, irrespective of competing calcium and suggests that SM has a high affinity for crystal nuclei. In fact, Mount (1999) demonstrated that intact IM (IM with associated sIM) binds nascent crystals tightly at intervals under physiological ionic strength. This finding considered with the recent discovery of seed crystals being supplied to the mineralization front by hemocytes (Mount et al., 2004) offers a convincing novel mechanism by which crystallization is promoted.

A model proposed by Wheeler et al. (1988), shows shell formation as a sequence of events where specific components are secreted in an ordered manner during crystal formation. In this model, IM initiates crystal formation by binding and stabilizing crystal
nuclei. Crystal shape is controlled by a secretion of SM, which adsorbs to specific crystal faces forcing growth on sides devoid of protein. Crystal growth is then completely inhibited by another secretion of SM which adsorbs non-specifically. Finally, pre-cursor IM proteins are secreted and assemble on the surface of the mature crystal in preparation for a new crystal to form. The characteristics of the 48 and 55 kD proteins and their relation to the IM and SM revealed in this study fit well into this model. Conceivably, the 55 kD protein may be secreted and adsorbed (through hydrophobic domains) onto the assembled matrix (IM) where it acts like a coating of “glue” and binds nascent crystals (perhaps of a particular mineralogy). The 48 kD protein may then be subsequently secreted and act mostly as a crystal growth inhibitor, affecting crystal growth rate and morphology. Obviously, any assignment for when a particular component might be deployed is premature. Also, to fully understand the role of any one matrix component, additional characteristics, revealed in this study, must be considered. These include the presence of high molecular weight proteins, the presence of prominent proteins not of the anionic protein class and the origin and cause of matrix protein heterogeneity.

In addition to the 48 and 55 kD proteins, a class of extremely high molecular weight anionic proteins of undetermined size is present in SM and especially sIM. Rusenko et al., (1991) reported a similar protein class in SM that exceeded the resolution of size exclusion chromatography. These proteins are unable to penetrate the gel of even the lowest percent polyacrylamide gels and, for the most part, are resistant to the denaturing conditions of SDS-PAGE and urea. The nature and origins of these macro-
molecular assemblages is currently unknown. However, in this study, it was observed that SM, following extraction, often formed precipitates that were resistant to dissolution. Precipitation did not occur in SM fractions resulting from size exclusion chromatography which suggests that the *crude* extract has components in it that promotes self-association.

Electrophoresis data shows that, in addition to a major anionic class of phosphoproteins, there are a number of other proteins, some of which are glycosylated. Molecules such as carbohydrates and lipids are capable of inducing self-association. Condensation reactions have been shown to occur between glycoproteins derived from the shell of *Mercenaria mercenaria* resulting in a partly insoluble, brown, macromolecular material (Collins et al., 1992). Wuthier (1975) identified a class of EGTA-extractable proteins which selectively bind Ca2+ in the presence of lipids forming protein-Ca- phospholipid complexes. Lipids have been shown to be tightly associated with the mineral of bone (Shapiro, 1970) and teeth (Shapiro, 1973) and are associated with the onset of mineral deposition (Boskey and Posner, 1977). In oyster shell formation, hemocytes deliver and deposit seed crystals to the mineralization front. These cells intimately associate with growing crystals, and appear to promote crystal remodeling (Mount et al., 2004). Thus, cell membrane components (including lipids) may be incorporated into the matrix as a by-product of cellular involvement.

Finally, proteins could interact through ion bridging if residual calcium remained in the SM extract, a mechanism common in nature which promotes spontaneous protein
assembly. In fact, a study of shell matrix derived from the bivalve, *Mytilus edulis*, suggests that some portion of the SM assembles through ionic bridging upon binding calcium, and may be the progenitors of IM (Hattan et al, 2001). An acid soluble matrix protein derived from nacre of the giant oyster, *Pinctada maxima* is reported to form 6 embarked oligomers through disulfide bridging (Bedouet et al., 2001).

Caspartin, an acid soluble protein isolated from the calcitic prismatic layer of the fan muscle *Pinna nobilis*, is a highly anionic with nearly 70% of its structure containing aspartic acid (Marin et al., 2005). This protein spontaneously oligomerizes in solution potentially through disulfide bridging. Because of the low Cysteine content, hydrostatic and hydrophobic interactions were not ruled out. This protein was immunolocalized within the prisms and forms a continuous “film” on both sides of the insoluble organic sheaths that surround each prism. In addition to a purported role in constraining crystal growth, these authors suggest that caspartin oligomers act to orient the growing crystals. Another mechanism of control attributed to oligomerization of acidic soluble matrix has been assigned to dentin matrix protein 1 (DMP-1). Assemblages of this protein were demonstrated to stabilize nascent calcium phosphate nuclei by forming ordered protein-mineral complexes, and thus proposed to promote bone and dentin mineralization at specific sites (He et al., 2005).

When oyster shell fragments are observed using atomic force microscopy, Sikes et. al. (1998; 2000) reported the presence of “foliar globules” on the surface of foliated
crystals that are clearly assemblages of matrix protein, the formation of which may be promoted by the presence of mineral salts. Given the high relatedness of the SM and IM proteins shown in this study (and Chapter 3), the idea that some portion of the SM, in particular the 55 kD protein may be the “building block” of IM is reasonable. In any case, the observation that SM self-assembles after extraction should be explored further considering protein assembly is a process concurrent with mineral formation.

SM also contains a molecular weight continuum of Stains-all positive bands that span nearly the length of the gel. Protein modifications are known to affect protein migration in gels and cause a smeared effect following staining. Protein degradation may be another source of this heterogeneity. Some degree of non-specific degradation likely occurs as a consequence of the extraction process, as demonstrated for the 48 and 55 kD bands, and some degradation is likely to have occurred within the shell itself. Considering that shell is not living tissue and that it is exposed to an aqueous environment, some loss of protein integrity would be expected, especially from older regions of the shell. An observation that supports this idea is that SM from younger specimens resulted in “cleaner” staining patterns on gels. Like bone, molluscan shell undergoes “remodeling” where mineral is purposely dissolved and re-mineralized (Lowenstam and Weiner 1989), a process that would seem to result in non-specific matrix protein degradation.
In bone and teeth, matrix heterogeneity is, in part, due to the directed degradation of a single “precursor” protein which produces several lower molecular weight proteins with similar compositions (Lee et al., 1983 and Uchiyama et al., 1984). In enamel, this processing is a purposeful and selective event where “parent” amelogenin matrix proteins are degraded and removed from developing enamel, in coordinated manner, prior to the onset of mineralization into mature enamel (Fincham et al., 1991; Robinson et al., 1995). Aoba and Moreno (1989), found this to be a specific proteolytic cleavage resulting in the loss of the anionic C-terminal domain causing the protein to loose affinity for the crystal surface. On the other hand, multiple, highly related proteins may represent discrete gene products resulting from differential RNA splicing as discussed above. Either case may explain the heterogeneity within the anionic class of proteins as well as the origins of the 48 and 55 kD bands.

Matrix heterogeneity was found to be caused, in part, by differences in phosphorylation of the component proteins- an observation well documented in documented in bone (Uchiyama et al, 1984) and dentin formation (Lee et al., 1983). Proof of this finding in oyster is the reduction in number of protein bands on IEF gels following de-phosphorylation as well as a significant shift in molecular weight on SDS-PAGE gels. Some degree of differential phosphorylation may be attributed to the environment as well as biochemical manipulation considering these groups are hydrolyzed in environments above pH 8 and would thus be artifact.
More compelling is the finding that dephosphorylation and phosphorylation of mineral matrix proteins are directed events where phosphate is removed or added, in an orchestrated manner, to change the activity of the protein. In vertebrates, alkaline phosphotase is present in the membranes of matrix secreting cells has been associated with mineralization in a variety of ways including acting as a hydrolase to deactivate inhibitors of apatite growth (Wuthier and Register, 1985). Osteopontin (OPN) and bone sialoprotein (BSP), matrix phosphoproteins associated with bone and teeth formation, are phosphorylated extra-cellularly by kinsases present on the surface of osteoblastic cells (Zhu et al., 2000). The extent to which these proteins are phosphorylated affects their activity and is reported to be the mechanism by which OPN achieves its multiple modes of function (Boskey, 1998).

Similarly, in studies of oyster shell formation, Borbas et al., (1991) found that a fraction of SM (RP-1) inhibits in vitro mineralization more effectively than dephosphorylated (RP-1). Furthermore, the dephosphorylated RP-1 more rapidly loses its inhibitory capacity during crystal growth assays. These studies suggest the possibility that at least some differential phosphorylation of oyster matrix has functional significance. In this case, perhaps the 48 kD protein undergoes dephosphorylation so that it is rendered inactive at specific times allowing rapid crystal growth on faces that are normally inhibited by the adsorption of this protein.
Amino acid analysis of the 48 and 55 kD proteins following adsorption to Immobilon revealed another level of heterogeneity within the bands themselves. Because this membrane depends largely on hydrophobic interactions to adsorb protein, it was presumed that the lack of these groups and the high content of anionic groups render this class of proteins unable to adequately interact with membranes of this type; thus, the low protein recoveries in this experiment were not unexpected. What was surprising, however, was the difference in affinity for the membrane exhibited by each protein, reflected in their recoveries, and the markedly different amino acid patterns that followed dialysis. Although some evidence supports that protein degradation (and to a lesser extent contamination from background protein) may contribute to band micro-heterogeneity, over-all, the data leans toward the possibility that the 48 kD and possibly the 55 kD bands contain protein isoforms. Perhaps two related proteins share the same molecular weight, one with a composition that is relatively less anionic and included more hydrophobic residues compared to another protein with a more anionic composition typically associated with mineral formation. The increased hydrophobicity of the former protein would allow it to bind to the membrane more effectively than a highly anionic protein. Isoelectric focusing of the 48 kD band supports band micro-heterogeneity where after eliminating heterogeneity due to phosphate, additional bands remained. Less convincing is 2-D analysis results which failed to resolve any additional discrete or prominent spots in the molecular weight range of 48 kD which should appear at different points along the pI gradient (positioned horizontally in the second dimension gel). Still, 2-D spots were broad and perhaps a tighter pI range would resolve this region more
effectively. In any case, if these isoforms exist, they are highly related as the difference in their surface charge is nominal.

Together, these studies support the possibility that some degree of matrix heterogeneity, with emphasis on the 48 and 55 kD phosphoproteins, is not artifact but represents an integral part of mineral deposition, where protein assembly, directed protein degradation (protein cleavage and/or dephosphorylation), or the expression of distinct related gene products are all strategies used to control the action of the matrix proteins. At this time, it is not clear which strategies are used during oyster shell formation.

Finally, in terms of matrix characterization, an unexplored and important line of study is revealed in the identification of a complex mixture of proteins that are not especially anionic. Several conspicuous proteins appear following CuCl₂ staining and silver nitrate staining, including variant forms identified by 2D electrophoresis. Because these proteins do not appear to be highly anionic, they would not be associated with functions typically purported for matrix proteins, such as crystal or ion binding. Still, their presence suggests they may be integral to the process and should be further investigated especially in light of the discovery that crystal bearing hemocyte cells are involved in shell formation (Mount et al., 2004). Such proteins may help recruit and organize cellular activity at the site of shell deposition.
Summary

In summary, oyster shell matrix is composed of a complex mixture of proteins. Amino acid analysis of whole SM extracts as well as those of the prominent 48 and 55 kD protein bands identified within SM supports the idea that a majority of the SM is highly anionic and phosphorylated. Heterogeneity within this protein class may be attributed to protein degradation that occurs diagenetically, as a result of biochemical manipulations or from unknown proteolytic processes. The loss of phosphate greatly contributes to heterogeneity within the anionic protein class. Although it is possible that some phosphate may be lost as a result of biochemical manipulation or diagenesis, the prospect that dephosphorylation may be a directed event is raised.

The 48 and 55 kD are highly related, both being made up primarily of serine, glycine, and aspartic acid. It is therefore very likely that they originate from the same gene through an alternative spicing mechanism or are part of a tightly related gene family. Moreover, this study suggests the possibility of additional protein isoforms, even within the 48 and 55 kD protein bands. Identifying gene sequences is key to ascertaining this form of heterogeneity.

Despite their exceptional similarity in composition and anionic character, the 48 and 55 kD proteins were distinguished in several ways. They vary slightly in amino acid compositions. The 55 kD shows a higher identity to sIM and more abundantly co-
extracts with sIM. This protein also adsorbs three times better to hydrophobic membrane supports than does the 48 kD protein. Finally, the 55 kD protein has no carbohydrate associated with it. In contrast, the 48 kD protein contains some carbohydrate moieties and is readily soluble indicating a more pervasive interaction with the mineral.

These differences suggest that despite their relatedness, the 48 and 55 kD proteins likely have distinct functions. The data suggests that in vivo, the 55 kD coats IM through common domains that promote a specific and tight association with assembled (IM) matrix while the 48 kD protein may act to modulate crystal habit through direct binding to crystal surfaces. Moreover, some portion of the 55 kD protein has an affinity for the 48 kD protein indicating that these proteins interact in some fashion in vivo. Still, in order to understand their individual and collective roles in shell formation, identifying the genes that encode the 48 and 55 kD proteins, the post-translational processing they undergo as they reach their mature forms and pinpointing time and location to which they are deployed during shell formation is crucial.

Future investigations should continue to focus on identifying the gene(s) for the 48 and 55 kD bands as well as the processing that occurs prior to their incorporation into the shell. This step will also facilitate understanding the macromolecular assemblages that occur within the SM and IM fractions of the matrix. However, equally important is to identify and characterize other minor protein constituents as these components may serve as key constituents in controlling mineral deposition in ways other than direct
contact with mineral surfaces or ion binding and may be more involved in protein interactions that help “set the stage” for controlled mineral deposition. These proteins are also excellent candidates for studies focused on the physiological aspects of shell formation.
CHAPTER THREE


Introduction

Molluscan shell is a multi-layer composite of inorganic mineral and organic macromolecules termed the organic matrix. Shell formation is a sequenced event which occurs at the shell margin in intimate association with the mantle organ. Epithelial cells along the outer mantle lobe actively secrete matrix components, among others, into the extra-pallial cavity (EPC), the region between the shell and the outer mantle lobe (Simkiss and Wilbur, 1989; Myers et al., 2007). There, matrix, crystal baring hemocytes (Mount et al., 2004) and inorganic ions interact to form matrix laden crystals which deposit as highly organized layers (Checa et al., 2005, 2006, 2007). These layers are defined in terms of the mineral type or mineralogy deposited and how the crystals are arranged, which is commonly referred to as microstructure (Carter, 1980). In bivalve shell, calcite and aragonite mineralogy occur most often. Among microstructures, prismatic (calcite), foliated (calcite), and nacre (aragonite) are most commonly studied.

The current model for shell growth states that the formation of the EPC is the first event to occur in a sequence of events. This space is sealed from the external environment by the periostracum, which is secreted from the periostracal groove at the
junction between the outer and inner mantle lobes (Simkiss and Wilbur, 1989). This highly cross-linked organic layer attaches to the shell, forming the outer most non-mineralized shell layer and provides the foundation for the subsequent deposition of the mineral layers (Checa, 2000). In oyster, the first mineral layer deposited is prismatic calcite followed by the deposition of foliated calcite, which forms the inner shell layer and bulk of the shell.

The formation of each layer is concomitant with a sequenced release of matrix proteins into the EPC fluid, which are believed to control every aspect of crystal deposition and growth (Simkiss and Wilbur, 1989; Sudo et al., 1997; Myers et al., 2007). Accordingly, the identification of these components and their protein structures has been the focus of many studies. Zhang and Zhang (2006) summarized the properties of more than 30 shell protein isolates, many of which are deduced from corresponding genes. These studies reveal that shell proteins have highly repetitive and modular structures. Gly- rich sequences, such as the GS domains of Lustrin A (nacre) and MSP-1 (foliated), GGY sequences found in MS131 and prismalin-14 (prismatic) are thought to impart elasticity to proteins and promote secondary structure optimal for “framework proteins” (insoluble matrix). Basic domains, often rich in arg and lys, and hydrophobic domains have also been reported and may promote interactions between macromolecules. Acidic domains of varying degrees are common to nearly all shell proteins. These domains are typically enriched with aspartic acid and are thought to chelate calcium (or magnesium) or directly interact with mineral surfaces. The most acidic matrix proteins are found in
calcitic shell. Among them, Aspein (prismatic) has the highest reported content of asp and contains (Asp)$_{2-10}$ repeats. Acidic proteins have also been identified in nacre. While there structures have not yet been determined, AP7, AP24 and AP8$_{\alpha\beta}$ are enriched in asp and were shown to interact with specific regions on crystal surfaces (Michenfelder et al., 2003; Fu et al., 2005).

Despite the recent advances in identifying matrix components and their gene sequences, this approach leaves out critical information regarding post-translational and post-secretory processing. Phosphorylation and glycosylation are known to be critical for the ability of shell matrix proteins to effect crystal formation and growth (Borbas et al., 1991; Matsushiro et al., 2003). Some matrix components assemble after secretion into the EPC and are thought to influence crystal orientation and shape (Matsushiro et al., 2003; Checa et al., 2005). More recently, matrix proteins have been shown to have additional cytokine properties (Almeida et al., 2000, 2001; Zhang et al., 2006) and likely act to regulate cells involved in mineralization including the mantle epithelium and hemocytes (Johnstone et al., 2007). As potential cytokines, matrix proteins may be further processed post-secretion to activate cell signaling capabilities (Bédouet et al., 2006; Van Lint and Libert, 2007).

These studies underscore the importance of continued studies using a two pronged approach where matrix protein genes and biochemical characterization of the expressed product are used to determine the complete structures of the protein. Knowing
in what form these proteins occur in the EPC compared to the shell is crucial for understanding matrix, mineral and cellular interactions that influence shell formation. To this end, the focus of this study was to determine the structures of shell derived 48 and 55 kD matrix phosphoproteins and to identify their presence in the mantle, EPC fluid and major shell layers.

Materials and Methods

Isolation of soluble (SM) and insoluble (IM) matrix protein

Specimens of the Eastern oyster, *Crassostrea virginica*, were obtained from oyster beds located in sixty bass creek of Winyah Bay, Georgetown South Carolina. The animals were shucked from the shells which were scrubbed with a wire brush under flowing tap water to remove sediment and foreign encrustations. The outer shell layers including the periostracum (outer most non-mineralized layer) and the prismatic layer (outer most mineralized layer) were removed with a high speed rotary cutting tool. The remaining shell was broken into small pieces with a hammer and shells pieces composed of foliated mineral were selected and ground into powder by short bursts in an electric coffee grinder. Twenty- five g of foliated shell was dissolved in 750 ml of 17% EDTA and centrifuged at 27,000 x g for 30 min to separate the water soluble fraction from the insoluble fraction. Soluble matrix (SM) contained within the water soluble fraction was concentrated to approximately 50 ml and then dialyzed against 1000 ml of 10 mM NaCl
using a Millipore Minitan tangential flow filtration apparatus with a molecular weight exclusion limit of 10 kD. The resultant dialysate was further dialyzed against distilled water and the final dialysate lyophilized. Soluble matrix was reconstituted in distilled water to a concentration of 5 mg ml\(^{-1}\) based on total weight of the dried material. The water insoluble fraction containing the insoluble matrix (IM) was washed with Tris buffered saline (TBS), pH 7.5 and extracted further in 2% SDS and 4 M urea. This suspension was centrifuged at 10,000 \(\times\) g for 5 min and the supernatant, containing solubized IM (sIM), was dialyzed against distilled water five times to remove salts. Protein was estimated using BCA protein assay kit (PIERCE).

To isolate sIM from specific shell layers, intact, individual shells (right valve) were placed in a 10% solution of EDTA, pH 8, and allowed to decalcify at room temperature over several days (Mount, 1999). As the shell decalcified, a transparent matrix gel like mold remained. Shell layers containing IM were collected by scraping the shell surface as it demineralized. sIM was extracted using a 2% SDS and 4% urea in Tris buffered saline, pH 7.5. Extracts were dialyzed as above and protein was estimated using BCA kit (PIERCE).

**Purification of 48 and 55 kD SM proteins**

Two mg of dry weight SM protein was loaded onto a 12% Tris-glycine SDS preparative gel for fractionation. Following electrophoresis, protein components were
identified by staining the gel with 0.3M CuCl$_2$. The 48 and 55 kD bands were excised from the gel, eluted with an electro-eluter apparatus (BioRad), and desalted by either dialysis using 10K MWCO dialysis cassettes (Pierce), or by using Centricon concentrators 10K MWCO (Amicon). The protein content of the resultant dialysate was estimated using BCA.

**Edman sequencing**

Prior to sequencing, approximately 200 ug of whole SM and IM protein were resolved on a 5 well, 4 to 20% Tris-glycine SDS gel and blotted onto Immobilon CD membrane (Millipore) and stained with a stain developed specifically for staining Immobilon CD according to manufactures directions. Protein bands of interest stained white against a purple background. These bands were excised from the blot, rinsed three times in dH$_2$O for 10 min each, air-dried and subjected to Edman sequencing.

**Enzymatic cleavage**

In an attempt to generate fragments for internal sequence analysis, several protease enzymes including, trypsin, chymotrypsin, V8 protease, pepsin, papain, protease K, and Asp-N (Boehringer Mannheim) were used. In preliminary, many of the enzymes were tested on whole SM to evaluate the effectiveness of the enzyme in cleaving the major 48 kD protein prior to treatment with gel purified 48 kD protein. Cleavage reactions containing trypsin, chymotrypsin, and V8 protease were carried out similarly
using a 1:10 ratio of enzyme to protein. When whole SM protein was tested, 25 ug of dry weight protein was reacted. In cases were purified 48 kD protein was tested, 1 ug of protein determined by BCA analysis was reacted. Bovine serum albumin was used as a positive control. Dry protein was reconstituted in 10 ul each of 8 M urea and 0.4 M NH₄ HCO₃. Two ul of 45 mM DTT was added and the reaction was incubated at 50 °C for 15 min. After the reaction was cooled to room temperature, 2 ul of 100mM IAA was added and the reaction was incubated at room temperature for 15 min. Finally, 20ul of dH₂O was added followed by the addition of the appropriate amount of enzyme in a 1ul volume. Negative control reactions containing each protein were treated as described above but excluded the addition of protease. The reactions were incubated for 24 h at 37 °C and were gently mixed by continuous rotation. Pepsin was reacted with BSA and SM protein in a buffer containing 10 mM HCL near pH 3. Papain was reacted with both proteins in a buffer containing 50 mM NaAcetate, 2M EDTA at pH 6. Both enzyme reactions were carried out in 20ul volumes used at a concentration of 1:50 ratio of enzyme to protein. Reactions were incubated overnight at room temperature with continuous rotation. All enzyme reactions were quenched by adding an equal volume of 2X Tris-tricine SDS sample buffer and heated to 90° C for 5min. The samples were resolved on a 10 to 20% Tris-tricine gel under denaturing conditions to evaluate cleavage of SM and BSA proteins.
Chemical cleavage

In an effort to obtain protein fragments suitable for sequencing, a time course study was undertaken to evaluate if limited proteolysis of the SM protein could be achieved by mild acid hydrolysis according to methods of Rusenko (1988). Twenty five ug of dry weight whole SM protein was reconstituted in 50 ul of 2% formic acid pH 2, and heated to 50 °C. Reactions were quenched at various time points over the course of 1 h by freezing. Three control reactions were run including whole SM that was not treated with acid or heat, SM that was heat treated only for 1 h, and SM that was treated with acid at room temperature for 1 h. Samples were frozen and lyophilized to dryness for analysis. Protein was also treated with CnBr in an effort to gain internal sequence data. In separate reactions, 25 ug of whole SM, and 1ug of purified 48 kD protein were reconstituted in 50 ul of 20 mg/ml CnBr in 70% TFA, and incubated at RT, in the dark, overnight. Two hundred ul of dH2O was added to each reaction and were then lyophilized (no heat). Samples were reconstituted again in 200ul of water and lyophilized to dryness (no heat). Control reactions were carried out identically except CnBr was not included. Potential proteolysis for both methods was determined by SDS-PAGE.

Dephosphorylation of the SM proteins

Whole SM was dephosphorylated using insoluble alkaline phosphotase as described in chapter I, and by base hydrolysis. For the later procedure, 2 mg of dry
protein was reconstituted in 0.5 M NaOH and incubated at room temperature for 24 h. The reaction was neutralized by the addition of HCl.

Production of antibodies to organic matrix protein

One hundred µg of gel purified 48 kDa SM protein was dissolved in 100 µl of Tris buffered saline (TBS) pH 7.5, mixed 1:2 with Freund’s complete adjuvant, and injected subcutaneously into white leghorn hens. A booster injection was performed at 10 days after the initial injection with an additional 100 µg of protein emulsified with an equal volume of Freud’s incomplete adjuvant. Eggs were collected two weeks after the booster injection and the yolk IgY was monitored by ELISA for specific antibody. Yolk IgY was isolated according to methods of Polson et al (1985) and final antibody extracts were reconstituted in TBS pH 7.5, to a protein concentration of approximately 15 mg/ml. Extracts exhibiting the highest specific activity were pooled and stored in 1 ml aliquots at -20 °C.

Additionally, antibodies were generated to 48 kD, 55 kD, whole SM and IM proteins in rabbit hosts. One hundred ug of each protein was reconstituted in 200ul of TBS, pH 7.5, and emulsified with an equal volume of complete Freunds adjuvant. The emulsion was injected subcutaneously into New Zealand White rabbits. Three weeks after the initial inoculation, the rabbits were boosted with an additional 100 ug of each
protein emulsified with an equal volume of Freunds incomplete adjuvant. Serum was harvested six weeks after the boost and stored in 1ml aliquots at -20 °C.

**Enzyme-Linked Immunosorbent Assay (ELISA)**

Selected proteins were dissolved in TBS, pH 7.5, to a concentration of 10 ug/ml. Fifty ul of each protein was applied to the wells of 96-well microtiter plates (Immulon) and allowed to adsorb overnight at 4 °C. Coat proteins were aspirated from the wells and the plate was washed gently with TBS, pH 7.5. Plates were blocked with 200 ul of 3% BSA in TBS, pH 7.5, for 3 h at room temperature. After blocking, plates were washed once with TBS, pH7.5. Yolk extract containing anti 48 kD antibody was serially diluted in 1% BSA in TBS, pH 7.5 (ELISA buffer) and 100 ul of each dilution was applied to the wells and incubated for 1.5 h at 37 °C. Plates were washed three times with TBS containing 0.05% Tween-20 (wash buffer) followed by the application of 100 µl of rabbit anti-chicken horseradish peroxidase conjugated antibody, diluted 1:35,000 in ELISA buffer. After 1 h incubation at 37 °C, plates were washed three times with wash buffer and color was developed at room temperature by the addition of 200 µl of o-phenylenediamine dihydrochloride (OPD). After 30 min, absorbance was measured at 450 nm wavelength.

ELISA’s using antibodies generated in rabbits were carried out under the same conditions as outlined above except a rabbit anti-goat antibody conjugated to alkaline
phosphatase was used to detect anti-SM antibody at a working dilution of 1:30,000. Color was developed using pNPP substrate (Sigma) and optical densities were determined after 30 min at 405 nm wavelength. After two injections of gel purified 48 kD protein, and 55 kD proteins and of whole soluble and insoluble matrix, the antibody titer and specificity to each protein was determined using ELISA and western blot analyses. For the ELISA test, color formation was measured at 405 nm and was plotted against the dilution factor of the sera collected following the initial challenge and one booster injection of the protein antigens.

**Western Blot analysis**

Selected proteins were resolved on 4 to 20% poly-acrylamide gradient gels (PAGE) in the presence of sodium dodecyl sulfate (SDS) prior to transfer onto nitrocellulose. The transfer was conducted for 1 h at 100 V in buffer containing 25 mM Tris, 192 mM glycine, 20% v/v methanol, pH 8.3 using a mini-transfer unit (Invitrogen). Following transfer, membranes were rinsed once with TBS, pH 7.5, then placed into 50 ml 3% BSA and 1% goat serum in TBS (blocking buffer) and allowed to incubate overnight at 4 °C. After blocking, membranes were incubated in 50 ml of one of the following anti-oyster matrix antibodies diluted in 3% BSA in TBS (western buffer) to the stated ratios: anti-55 and 48 kD antibodies were diluted 1:800 and anti-SM and IM antibodies were diluted 1:1000. After 2 h incubation in the selected anti-oyster antibody, membranes were washed three times in 100 ml of 1% BSA and 0.05% Tween-20 in TBS...
(wash buffer) for 5 min each. The membrane was incubated in 50 ml of alkaline-phosphotase conjugated goat anti-rabbit IgG (Sigma) diluted to 1:30,000 in western buffer for 1.5 h, and washed three times for 5 min each in wash buffer. A final rinse was carried out in TBS and color was developed with BCIP/NBT (Sigma). Development was stopped by repeated washing in distilled water over 10 min. All incubations and washes were carried out with continuous gentle shaking at room temperature.

**Results**

**Characterization of antibodies generated against matrix protein in chickens.**

Initially, antibodies to the 48 kD protein were generated in chickens. After two injections of the protein antigen, the antibody specificity and titer of the yolk extract was monitored by ELISA. Each egg clutch (4 to 5 eggs) was monitored individually. Extracts exhibiting the highest antibody titers were pooled. Pooled yolk extract was diluted serially and tested for reactivity to the 48 kD protein (protein antigen) and to other phosphoproteins including a crude and pure extract of phosphophoryn designated PP-Ca^{2+} and PP-4 respectively, and phosvitin. Myoglobin was used as a control. The color generated upon the addition of the peroxidase substrate was measured at 450nm and plotted vs. the dilution of the antibody preparation (Figure 3.1). While the antibody reacted well with the 48 kD antigen protein, an equally high cross-reactivity was observed for pure phosphophoryn and to a lesser extent, crude phosphophoryn. Notably,
the antibody reacted best with phosvitin. Cross-reactivity to phosphophoryn was not surprising given that it is a mineral derived protein with a chemical composition comparable to oyster shell matrix proteins (Rusenko, 1988). The substantial reactivity to phosvitin, a yolk derived protein, is likely due to a long run of phosphoserine (30mer) present near the C-terminal region of this protein (Byrne et al., 1984). Cross-reactivity with this protein suggests that poly-serine domains, in particular, are antigenic and are likely present in the 48 kD protein. Over-all, the 48 kD protein was not an especially potent antigen in chickens as evidenced by its low titer (1:7680). Control experiments were conducted using yolk extracts collected before challenge, each of which resulted in absorbances near or at background levels (data not shown).
Figure 3.1 Enzyme-linked immuno-sorbent assay demonstrating specificity and cross-reactivity of the chicken anti-48 kD SM antibody with selected proteins. The 48 kD SM protein was gel purified from whole SM. PP-4 is HPLC purified phosphophoryn from rat incisor. PP-Ca is calcium precipitated phosphophoryn from rat incisor (Rahima and Veis, 1988). Phosvitin is a chicken egg yolk phospho-glycoprotein (Sigma). Myoglobin was included as a control (Sigma). Wells were coated with 50 μl of 10 μg/ml protein solutions. Absorbance at 450 nm was plotted against the yolk extract dilution. This antibody shows significant cross-reactivity to the phosphoproteins, phosvitin, and phosphophoryn indicating that these proteins have epitope regions in common with the 48 kD protein.

Western analyses were inconclusive in identifying which bands within SM react to the anti-48 kD antibody, due to the fact that SM proteins were found to have poor affinity for PVDF membrane. However, phosvitin reacted strongly with anti-chicken
antibody on Westerns. In subsequent studies, it was discovered that nitrocellulose was most effective for blot applications involving oyster matrix protein. Because the antibody titer produced by chickens was low and the activity of the antibody preparations quickly deteriorated once extracts were reconstituted, efforts to generate matrix protein specific antibodies were redirected to using rabbits as the host animal. All studies described henceforth utilized antibodies generated in rabbits.

**Characterization of antibodies generated against matrix protein in rabbits**

Rabbit serum containing antibodies generated against whole SM and IM extracts, and gel purified 48 and 55 kD SM proteins were tested for reactivity to their antigens, and cross-reactivity to other phosphoproteins including phosvitin and phosphophoryn. In identical ELISA formats (Figure 3.2), shell matrix protein elicited a strong immune response in rabbit hosts. The anti 48 and 55 kD sera titered past the 1:30720 dilution (Figure 3.2, A and B), while anti-SM and IM sera titered past the 122,880 dilution (Figure 3.2, C and D). The curves generated for each of the antisera tested were nearly identical and indicates that the 48 and 55 kD proteins, and proteins within whole SM and IM extracts share common epitopes. Thus, shell matrix is comprised, in part, of a highly related suite of proteins. Notably, the anti-55 kD antibody consistently reacted most strongly with the 48 kD coat protein. Unlike antibodies made in chickens, none of the rabbit antibodies cross-reacted with phosphophoryn or phosvitin demonstrating that all four antibodies are specific for shell organic matrix and therefore better suited for
immunological procedures. Myoglobin (control) was also negative. Control experiments were conducted using sera collected before challenge, each of which resulted in absorbances near or at background levels (data not shown).
Figure 3.2 Enzyme-linked immuno-sorbent assays demonstrating reactivity and cross-reactivity of the rabbit anti-48kD, 55kD, SM and IM antibodies to their protein antigens and other phosphoproteins.
SM is the whole water soluble protein fraction that results following shell dissolution in EDTA. The 48 and 55kD proteins are gel purified from whole SM. Phosphophoryn is HPLC purified from rat incisor and graciously provided by Dr. Art Veis (Rahima and Veis, 1988). Phosvitin is a chicken egg yolk phosphoprotein (Sigma). Myoglobin was included as a control (Sigma). Absorbance at 405 nm was plotted against the serum dilution. Wells were coated with 50ul of 10ug/ml protein solutions.

Western blot analyses further demonstrates the high relatedness of component matrix proteins within the SM and sLM extracts. Illustrated in Figure 3.3 (A and B), anti-SM and IM antibodies reacted similarly in duplicate blots. Both antibodies broadly react with whole SM (Figure 3.3, Lane 1) and sLM extracts (Figure 3.3, Lane 4), and resulted in similar banding patterns. Anti-SM reacts highest with SM extract and shows specific activity, above background, for several bands including the 48 and 55 kD proteins (Figure 3.3, Lane 1). The anti-IM antibody recognizes several bands in sLM (Figure 3.3, Lane 4) including the 48 and 55 kD bands. Clearly, the antibodies similar reactivities indicate that SM and sLM have many proteins in common but appear to differ in quantity based on their relative band intensities. For example, a band detected at 60 kD appears in both SM and sLM but appears to be more prevalent in SM.
**Figure 3.3 Western analyses of anti-matrix antibodies to selected proteins.**
Proteins were resolved on 12% SDS-PAGE gels. Weights were determined by BCA. SM and sIM were loaded at slightly higher concentrations because they are crude fractions. Background staining was slightly elevated due to the high protein loads and is most obvious in the westerns immunostained with the anti-SM (A) and anti-IM antibodies (B). Lane “a” contains broad range protein standards as labeled (Bio-Rad).

- Lane a - Broad range biotinylated protein standards (Bio-rad)
- Lane 1 - hole SM 85 ug
- Lane 2 - Phosphophoryn 70 ug
- Lane 3 - Phosvitin 70 ug
- Lane 4 - sIM 85 ug
- Lane 5 - BSA 70 ug

The anti-55 and 48 kD antibodies produced nearly identical banding patterns on duplicate blots and show enhanced specificity to five bands in SM including the 48 and 55 kD proteins (Figure 3.3, C and D, Lane 1). Much of the background reactivity apparent with the anti-SM and anti-sIM antibodies is not apparent here. Moreover, these antibodies identified differences between SM and sIM extracts, mostly in the levels of proteins that are present. Also, as demonstrated in ELISA, when compared to the anti 48 kD blot (Figure 3.3, C), the anti-55 kD antibody (Figure 3.3, D) exhibited greater specific activity to the 48 and 55 kD proteins within the SM (Figure 3.3, Lane 1) and sIM (Figure 3.3, Lane 4) and the lowest background staining (compare Figure 3.3, Lane 1 in C and D). The fact that the 55 kD antigen produced an antibody that reacted best with the 48 kD protein suggests that the epitopes, shared by both proteins, are in greatest amount in the 48 kD protein. Consequently, the anti-55 kD antibody was most often used to identify the major SM proteins in experiments discussed herein, and further demonstrates that the 48 and 55 kD proteins, in particular, have highly related structures.
Curiously, the anti-55 kD antibody (which was made to 55 kD protein purified from SM) does not react as strongly as expected with the 55 kD protein in sIM (Figure 3.3, Lane 4, C and D) despite its prominence in this extract (as shown in chapter 2). Although this relative lack of reactivity suggests that the 55 kD proteins of the SM and sIM may be distinct proteins, the collective biochemical data presented here, and in Chapter 2 (same N-terminus), suggest that the proteins are identical, or at the very least highly related isoforms, which in theory should react similarly in immunoassays. Therefore, the inability for the anti-55 kD antibody to recognize well the 55 kD protein in sIM is likely a technical problem, stemming from its chemical character as an insoluble protein, which results in poor transfer efficiently or its epitopes are not available to antibody.

All Western analyses confirm that the anti 48 kD, 55 kD, SM and IM antibodies are specific to shell matrix, as there was no cross-reactivity with phosphophoryn (Figure 3.3, Lane 2) or phosvitin (Figure 3.3, Lane 3) above that observed for the control protein (Figure 3.3, Lane 5), which was negative.

Reactivity of the anti 48 kD antibody to dephosphorylated SM

Oyster SM proteins are heavily modified proteins, mostly in the form of phosphoserine (Rusenko, 1991 and chapter 1) Because post-translational modifications can themselves be antigenic, it was important to determine if the antibodies generated in
rabbits were, in fact, made to epitopes found within the amino acid structure- a point critical for studies using cross-reactivity to establish protein relatedness as well as for screening expression libraries made in bacteria, which do not carry out post-translational modifications.

Accordingly, antibodies were tested against matrix protein that was dephosphorylated by alkaline phosphotase (AP) or dilute base. In the ELISA shown in Figure 3.4, SM was recognized by the anti-55 kD antibody, despite the removal of phosphate with AP, while, antibody reactivity to base treated SM was significantly reduced. Western analyses (Figure 3.5, A) confirm this result and demonstrate that the 48 and 55 kD bands in SM (Figure 3.5, Lane 1) and sIM (Figure 3.5, Lane 2) remain active to both the anti-55 and anti 48 kD antibodies following AP treatment. In contrast, when SM is treated with base (Figure 3.5, B), immuno-staining was diminished in the 48 and 55 kD regions of the blot (Figure 3.5, Lanes 2 and 4 respectively). Similar results were obtained when the anti- 48 kD, SM and IM antibodies were used in duplicate experiments (data not shown). Myoglobin was included as a control and was negative.
Figure 3.4  Enzyme-linked immuno-sorbent assay demonstrating the effect of dephosphorylation of SM on reactivity of the anti-55 kD antibody.
Two methods of dephosphorylation, alkaline phosphotase (SM 254-AP) and dilute base treatment (SM 255 - base treated), are compared. SM 254 and SM 255 are control proteins which were exposed to experimental conditions but without the addition of enzyme or base. Myoglobin was included as a negative control.
Figure 3.5 Western analyses of dephosphorylated matrix protein following alkaline phosphatase and base treatment.
Samples were resolved on 12% SDS-PAGE gels and transferred to nitrocellulose. Forty ug (dry weight) of protein was loaded per lane.

A- Alkaline phosphatase treated SM compared to control SM. Protein was detected with the anti-55 and anti 48 kD as indicated. The 48 and 55 kD proteins are labeled. There is no apparent change in antibody reactivity to matrix proteins, including the 48 and 55 kD proteins, following dephosphorylation.
Lane a - Broad range biotinylated protein standards (Bio-Rad)
Lane 1 - SM- AP treated
Lane 2 - SM control

B- Base treated SM and sIM compared to control protein. Protein was detected using the anti 55 kD antibody. Antibody reactivity was diminished in the 48 and 55 kD protein regions of both the SM and sIM, while some higher molecular weight proteins that are not phosphoproteins, appear to be more resilient to this treatment.
Lane a - Broad range biotinylated protein standards (Bio-Rad)
Lane 1 - SM control
Lane 2 - SM base treated
Lane 3 - sIM control
Lane 4 -sIM base treated

C- SDS-PAGE analysis of base treated protein following silver stain. Base treatment caused a loss of protein structure for the 48 and 55kD proteins as evidenced by their disappearance following staining. The 48 and 55kD bands (marked with →) are very light in control SM because these proteins don’t react well with silver nitrate. However, the same result was observed when a duplicate gel was stained with Stains-all except there were no bands visible in either of the base treated samples indicating a loss of phosphate. The gel was back-stained with silver nitrate which resulted in heavy background staining but no additional bands were evident (data not shown).
Lane a - Broad range pre-stained protein standards (Bio-Rad)
Lane 1 - SM control – 40 ug (dry weight)
Lane 2 - SM- base treated – 40 ug (BCA weight)
Lane 3 - sIM- base treated – 40 ug (BCA weight)
The loss of reactivity of the antibodies to base treated SM and IM was due to the loss of protein structure containing the epitopes. This conclusion is supported by SDS-PAGE analysis of the SM and IM proteins following dilute base treatment which shows that the 48 and 55 kD phosphoprotein bands, in particular, are base labile (Figure 3.5, C). In lane 1, the 48 and 55 kD proteins are visible before treatment but disappear when exposed to dilute base (Figure 3.5, Lanes 2 and 3) and indicates that the proteins appear to have been effectively dephosphorylated following base hydrolysis. Two higher molecular weight proteins, assumed to lack poly-ser domains because they are not Stains-all positive, remained intact.

Over-all, alkaline phosphatase was the most effective means of removing phosphate from SM while maintaining the structural integrity of the proteins. Because antibody reactivity was unaffected when reacted with AP treated matrix, it was concluded that antibodies made in rabbit hosts were generated primarily to the amino acid structure of the 48 and 55 kD proteins, and not to phosphate groups.

**Enzymatic Cleavage**

Enzymatic and chemical cleavage was used in an effort to gain internal sequence data of the 48 kD protein, as well as a sense of the arrangement of its domains. Susceptibility to chemical and enzymatic cleavage gives limited data on the presence of specific groups and some logical assessment of where those groups may be located.
within the peptide chain based on the size and number of cleavage products that result. For example, because $\alpha$-chymotrypsin hydrolyzes carboxyl-terminal peptide bonds of aromatic amino acids and trypsin hydrolyzes the carboxyl-terminal peptide bond of lysine or arginine residues (Allen, 1981), few cleavages should result given the scarcity of these residues in whole SM and the 48 and 55 kD proteins. Large peptide fragments would result if these residues were located in the middle of the protein, as opposed to their location towards the ends of the peptide chain, which would produce small fragments. And depending on the substrate residues’ proximity to the N- or C- terminus, a slight or even undetectable change in molecular weight of the parent protein would occur.

V8 protease cleaves aspartic and glutamic acid residues, and considering that the former constitutes approximately 30% of the protein, this enzyme should generate many cleavage products of various sizes depending on the arrangement of these residues within the peptide chain. A caveat in using this protease is that -asp-asp- and -glu-glu- bonds are refractory to cleavage (Croft, 1980). Consequently, if aspartic acid residues exist as runs within the peptide chain the enzyme may be ineffective at generating fragments. Moreover, if aspartic residues are interspersed throughout the peptide, the cleavage products may be too small to isolate and sequence. Accordingly, high percent (10 to 20%) gradient Tris-tricine poly-acrylamide gels were used to evaluate cleavage because of their superior resolution of small proteins.
In this study, trypsin, chymotrypsin, and V8 proteases failed to fragment the 48 kD protein despite the presence of substrate amino acids susceptible to these enzymes (reported in Chapter 2). Shown in Figure 3.6, there was no difference in banding patterns of protein treated with chymotrypsin (Figure 3.6, A Lane 2), and control protein (Figure 3.6, A Lane 3). A duplicate gel was back-stained with silver nitrate to identify possible cleavage products that were not detectable with Stains-all alone (Figure 3.6, B Lanes 2 & 3). Again, there was no difference in banding patterns except for a band at approximately 6kD, which corresponds to the chymotrypsin band in the control (Figure 3.6, B Lane 1). Likewise, V8 protease failed to fragment the 48 kD protein as evidenced by the similar banding patterns of treated (Figure 3.6, A Lane 5) and untreated protein (Figure 3.6, A Lane 6). Back-staining with silver nitrate identified only bands originating from the V8 enzyme shown in the control (Figure 3.6, B Lane 4), but no bands resulting from the digestion of the 48 kD protein. The conditions of hydrolysis were not in question as the control protein, BSA, was successfully fragmented in both cases (C). The resistance of the 48 kD protein to proteolysis by V8 protease is consistent with previous studies where fractions of SM were refractory to this enzyme (Rusenko, 1988) and strongly suggests asp, to a large extent, is arranged as runs in this protein.
Figure 3.6 SDS-PAGE analyses of the 48kD protein following treatment with chymotrypsin and endoproteinase Glu-C (V8 protease).
In separate reactions, 5 ug (BCA weight) of dry 48 kD protein was reconstituted in 10 ul each of 8 M urea and 0.4 M NH₄ HCO₃. Next, 2 ul of 45 mM DTT was added and the reaction was incubated at 50°C for 15 min. After the reaction was cooled to room temperature, 2 ul of 100 mM IAA was added and the reaction was incubated at room temperature for 15 min. Finally, 20 ul of dH₂O was added followed by the addition of 1 ul of enzyme, so that the enzyme: protein ratio was 1:10. BSA was included as a control. All control samples were treated identically except either enzyme or protein was omitted. Reactions were incubated for 24 h at 37 °C and resolved on 10% to 20% Tris-tricine gels along with SeeBlue protein standard (Lanes a). Cleavage products from BSA were visualized by staining with Coomassie blue 250 while Stains-all and silver nitrate was used to visualize matrix protein products.

A- Digestion of 48kD SM protein; products visualized with Stains-all
Lane 1 - chymotrypsin control (no protein)
Lane 2 - 48kD with chymotrypsin
Lane 3 - 48kD control
Lane 4 - V8 control (no protein)
Lane 5 - 48kD with V8
Lane 6 - 48kD control
*Note the 48 phosphoprotein band has a higher estimated weight of 76.9kD when calculated using SeeBlue protein standards on 10 to 20% Tris-tricine gels.

B- Same gel as shown in A following secondary silver staining. The 48 kD protein appears to be resistant to cleavage by chymotrypsin as demonstrated by identical protein patterns in lanes 2 and 3. A band corresponding to chymotrypsin (lane 1) is also evident in lane 2). Although at first glance, V8 protease appeared to cleave the 48kD band, after silver staining, the protein pattern exhibited in lane B4 (V8 alone) demonstrates this not to be the case.

C- Digestion of BSA; products were visualized by Coomassie brilliant blue 250.
Lane 1 - chymotrypsin control (no protein)
Lane 2 - BSA with chymotrypsin
Lane 3 - BSA control
Lane 4 - V8 control (no protein)
Lane 5 - BSA with V8
Lane 6 - BSA control

The presence of phosphoserine has been shown to inhibit various endopeptidases from hydrolyzing the highly phosphorylated protein, phosvitin, despite the presence of many potential cleavage sites (Taborsky, 1974). To test if post-translational
modifications such as phosphate or carbohydrate moieties were obstructing available substrate residues within the 48 kD protein structure, trypsin digests were carried out using dephosphorylated protein. An initial experiment was conducted to test if trypsin could digest 48 kD protein in its native state. Shown in Figure 3.7, there was no variation in gel banding patterns of treated (Figure 3.7, A Lane 2) compared to control protein (A lane 3) despite the presence of substrate groups. Dephosphorylation of the 48 kD protein resulted in a significant downward shift in molecular mobility (compare Figure 3.7, B Lanes 1 and 2). However, even after phosphate was removed, the 48 kD protein remained refractory to trypsin as there was no visible difference in banding patterns of dephosphorylated (Figure 3.7, B Lane 3), and control protein (Figure 3.7, B Lane 2). The control protein, BSA, successfully fragmented under these conditions (compare Figure 3.7, A Lanes 5 and 6; Figure 3.7, B Lanes 4 and 5).
Figure 3.7  SDS-PAGE analyses of 48 kD and dephosphorylated 48 kD protein following treatment with trypsin.

In separate reactions, 5 ug each (BCA determined weight) of 48 kD protein, AP-48 kD (alkaline phosphatase treated) and 25 ug of BSA was mixed 1:10 with enzyme and incubated in reaction buffer as stated in Figure 3.6. Negative control reactions containing each protein were treated identically, except either enzyme or protein was omitted. The reactions were incubated for 24 h at 37 °C and the products were resolved on 10 to 20% Tris-tricine gels and stained with Stains-all.

A-  48 kD protein treated with trypsin
Lane a - Mark 12 molecular weight standards (Novex) - stained with Coomassie blue
Lane 1 - trypsin control (no protein)
Lane 2 – 48 kD and trypsin
Lane 3 – 48 kD control
Lane 4 - SM (25 ug)
Lane 5 - BSA with trypsin
Lane 6 - BSA control
*Note the 48 phosphoprotein band has a higher estimated weight of 52.9 kD when calculated using Mark 12 standards on 10 to 20% Tris-tricine gels.

B-  Dephosphorylated 48 kD protein treated with trypsin. Note the molecular weight shift of the 48 kD protein after alkaline phosphatase treatment. Mark 12 standards did not stain with Stains-all.
Lane a - Mark 12 molecular weight standards (Novex)
Lane 1 – 48 kD control
Lane 2 - AP-48 kD
Lane 3 - AP-48 kD with trypsin
Lane 4 - BSA control
Lane 5 - BSA with trypsin

There was no visible shift in molecular weight or in banding pattern of the 48 kD following incubation with trypsin demonstrating that this protein is resistant to cleavage by this enzyme. Likewise, dephosphorylation of the 48 kD failed to increase the proteins susceptibility to trypsin also demonstrated in B. Back-staining with silver nitrate did not resolve further any minor cleavage products (data not shown).
The results of this experiment suggest that the phosphate groups do not hinder the ability of the trypsin to cleave at its substrate residues, therefore some other modification and/or higher level protein arrangement must be obscuring these sites. These results remained consistent even under conditions where the incubation time was lengthened, urea and other denaturants were included, and/or the amount of enzyme was increased (data not shown). On the other hand, in the case of chymotrypsin and trypsin, it is possible that the substrate residues are located on the extreme end(s) of the peptide chain which may not produce a detectable molecular weight shift, even if cleavage occurred.

A series of reactions were carried out with whole SM and enzymes known to indiscriminately cleave protein. If extensive modifications and/or protein folding are blocking cleavage sites across long stretches of the peptide, then perhaps these “less-specific” proteases would generate fragments, possibly of substantial size. Several enzymes of this nature were used including protease K, papain, and pepsin. In addition, Asp-N, an endo-peptidase which specifically cleaves the C-terminal side of aspartic acid, was tested.

All four enzymes successfully cleaved SM proteins. Shown in figure 3.8, papain completely digested the 48 and 55 kD proteins as neither band was visible and no discernable cleavage products were evident following treatment (Figure 3.8, A, Lanes 3 and 4). Pepsin proved to be less aggressive and produced several distinct Stains-all positive bands when compared to the control (Figure 3.8, A, Lanes 1 and 2), and worked
best when proteolysis time was limited. Incubation at 5 h and 10 h produced four distinct products (marked with black arrows) with estimated molecular weights of 61, 52, 25, and 15 kD (Figure 3.8, B, Lanes 3 and 4) compared to the control (Figure 3.8, Lane 2). A pink staining band (white arrow) appears at 37.8 kD and was determined to be pepsin because it is visible in control lanes stained with coomassie blue (Figure 3.8, A, Lane 6 and B, Lane 5). An additional cleavage product of SM may co-migrate with the pepsin band because it becomes more prominent following treatment (lanes 3 and 4) compared to the pepsin control lanes. Apparently, the 55 kD protein was affected by the buffer conditions required for pepsin (pH 2) because it does not stain well with stains-all in control lanes (Figure 3.8, A, Lane 1 and B, Lane 2).
Figure 3.8 SDS-PAGE analysis of cleavage products generated from treating SM protein with pepsin and papain.
In separate reactions, 25 ug of SM protein was mixed 1:50 with each enzyme. Digestion with pepsin was carried out in 10 mM HCL, pH 3 and digestion with papain was carried out in 50 mM Na Acetate and 2M EDTA, pH 6. Control samples were treated identically except enzyme or protein was omitted. BSA was included as a control and products were visualized by staining with Coomassie blue 250. Stains-all was used to visualize matrix protein cleavage products. Reactions were resolved on 10 to 20% Tris-tricine gels along with SeeBlue protein standard (Lanes a) (Sigma).

A- Reactions were incubated for 24 h at 37 °C  
Lanes 1 and 3 - SM control  
Lane 2 - SM with pepsin  
Lane 4 - SM with papain  
Lanes 5 and 7 - BSA control  
Lane 6 - BSA with pepsin  
Lane 8 - BSA with papain  
*Note the 48 phosphoprotein band has a higher estimated weight of 69.9 kD when calculated using SeeBlue protein standards on 10 to 20% Tris-tricine gels.

Papain appeared to completely digest SM protein. Pepsin, on the other hand, resulted in 4 to 5 discrete protein fragments and was therefore used in a subsequent time course experiment shown in B.

B- Pepsin time course - reactions were incubated for 5 h and 10 h at 37 °C. Cleaved products marked with arrows.  
Lanes 1 and 5 - pepsin without protein  
Lane 2 - SM control  
Lane 3 - SM with pepsin incubated 5h  
Lane 4 - SM with pepsin incubated 10h  
Lane 6 - BSA control  
Lane 7 - BSA with pepsin

Protease K and Asp-N were tested with gel purified 48 kD, both of which generated several large protein fragments (Figure 3.9). Protease K cleaved the 48 kD protein into four discernable fragments of 61, 50, 48, and 30kD (Figure 3.9, A, lane 3). This gel was back-stained with silver nitrate (Figure 3.9, B) but no additional bands were evident. Likewise, digestion of the 48 kD protein with Asp-N (Figure 3.9, C), generated
several prominent bands at 58.8, 49.0, and 38.0kD. Back-staining (Figure 3.9, D) failed to identify any additional products. Over-all, Prot K, Pepsin and Asp-N enzymes produced prominent, large protein fragments and are the best means to obtain internal sequence information for the SM phosphoproteins.
Figure 3.9 SDS-PAGE analysis of cleavage products generated from treating 48 kD SM protein with endoproteinase Asp-N and Proteinase K.
In separate reactions, five ug of 48 kD protein was mixed 1:20 with each enzyme in Tris-HCL buffer, pH 8 containing a final concentration of 1M urea. Reactions were incubated for 24 h at 37 °C. Control samples were treated identically except no enzyme was added. Reaction products were run on 10 to 20% Tris-tricine gels along with SeeBlue protein standard (Lanes a) (Novex).

A - Cleavage products (denoted by →) resulting from Proteinase K, visualized with Stains-all. BSA was used as a control protein and gel sections resolving the products were stained with Coomassie brilliant blue 250.
Lane 1 - BSA with Proteinase K
Lane 2 - BSA control
Lane 3 - 48kD SM with Proteinase K
Lane 4 - 48kD control

*Note the 48 phosphoprotein band has a higher estimated weight of 65.0 kD when calculated using SeeBlue protein standards on 10 to 20% Tris-tricine gels.

B - Lanes a, 3 and 4, shown in A, followed by silver nitrate staining.

C - Cleavage products (denoted by →) resulting from Asp-N, visualized with Stains-all.
Lane 1 – 48 kD SM with Asp-N
Lane 2 – 48 kD control

D - Same gel as shown in C followed by silver staining.

Notice that in either case, no additional bands were identified by secondary silver staining. Due to expense, Asp-N was not run by itself. Its reported molecular weight is 27 kD. Proteinase K was linked to agarose beads which was removed from the sample by centrifugation prior to electrophoresis.

Chemical Cleavage

Under mild conditions, formic acid preferentially cleaves aspartyl peptide bonds (Inglis, 1983). Given the prevalence of this amino acid in both the 48 and 55 kD proteins
this treatment should produce numerous cleavage products. Rusenko (1988) reported two fragments resulted when a fraction of SM was subjected to mild acid hydrolysis after 0.5 and 1.0 h, and that the release of aspartic acid from SM fractions is essentially complete after 3 h. Because the goal of this study was to obtain long fragments suitable for isolation and sequencing, a shorter time course of treatment was used. Shown in figure 3.10, over the course of one hour, no discernable molecular weight shift of either band, or appearance of cleavage products were visible following mild acid hydrolysis when treated SM (Figure 3.10, Lanes 5 to 9) was compared to controls (Figure 3.10, Lanes 1 and 2). Back-staining with silver nitrate did not reveal any additional products. A longer time period may produce cleavage products, however, efforts were refocused on methods that may be better for creating products more favorable for sequencing.
Figure 3.10  SDS-PAGE analysis of SM following a time course treatment with formic acid.
Twenty five ug of dry weight whole SM protein was reconstituted in 50 ul of 2% formic acid pH 2, and heated to 50 °C. Reactions were quenched at various time points over the course of 1 h by freezing. Three control reactions were run including whole SM that was not treated with acid or heat, SM that was heat treated for 1 h, and SM that was treated with acid at room temperature for 1 h. After freezing, all samples were lyophilized, run on 10 to 20% Tris-tricine, and stained with Stains-all and silver nitrate gels to evaluate proteolysis.

A - Lane a - SeeBlue broad range molecular weight standards (Novex)
   Lane 1 - SM control
   Lane 2 - SM heat only
   Lane 3 - SM acid only (no heat)
   Lanes 5 to 9 - SM incubated in acid at 50 °C for 5, 15, 30, 45 min and 1h

B - Same gel as shown in A following secondary silver staining in an effort to identify potential minor cleavage products that may not be visible with stains-all alone.

There was no visible shift in molecular weight or in banding patterns of SM following staining indicating the 48 and 55 kD bands are resistant to cleavage under these conditions.

Like many proteins, methionine is present only in trace amounts in the 48 kD protein. Cyanogen bromide (CNBr) cleaves at the c-terminal side of methionine residues (Abelson, 1990) and was used in an attempt to generate large protein fragments. Moreover, because a methionine residue was identified at position 20 of the 48 kD n-terminus (discussed below, Table 3.1), it was anticipated that sequencing of a cleavage product generated by CNBr would provide a continuation of that sequence data. Shown in figure 3.11, CNBr failed to fragment proteins within the SM or the 48 kD protein as there was no difference in banding patterns between control (Figure 3.11, Lanes 2 and 5) and treated proteins (Figure 3.11, Lanes3 and 6). Theoretically, a 20kD product should
have resulted, however, the lack of any cleavage products suggest that methionine may not be present or is not accessible to the chemical.

### Table 3.1 Amino-terminal Sequences\(^{(1)}\) for the 48 and 55 kD Proteins

<table>
<thead>
<tr>
<th>Protein/shell matrix fraction(^{(2)})</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>asp-glu-alad-alagly-aspvalala-aspthr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55kD/SM(^{(4)})</td>
<td>( )-S-K-G-X-E-P-D-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )-( )(4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.11  SDS-PAGE analysis of SM and 48 kD protein following treatment with CnBr

BSA was included as a positive control. Dried protein was re-suspended in 50 ul of 20 mg/ml solution of CnBr in 70% TFA and incubated at room temperature in the dark overnight. Control reactions were carried out under identical conditions without CnBr. Additional controls that were not exposed to TFA were also included. Samples were run on a 10 to 20% Tris-tricine gel and initially stained with Stains-all then back-stained with silver nitrate.

A- Stains-all stained gel

Lane a - Benchmark pre-stained protein standards (Invitrogen)
Lane b - Mark 12 protein standards (Invitrogen)
Lane 1 - SM no treatment in TFA
Lane 2 - SM control
Lane 3 - SM with CnBr
Lane 4 - 48kD no treatment in TFA
Lane 5 - 48kD control
Lane 6 - 48kD with CnBr
Lane 7 - BSA with CnBr
Lane 8 - BSA control

B- Same gel as above back-stained with Silver nitrate in an effort to identify potential products not identified with Stains-all.

N-terminal sequence

The N-terminal sequence of the 48 and 55 kD SM proteins and the 55 kD sIM protein was determined by Edman sequencing directly from blots following SDS-PAGE, or from eluted protein. Results and are shown in Table 3.1. Single and three letter amino acid identifiers are used to illustrate the amino acid order. An ‘X’ is used to depict unknown amino acids which are not identifiable due to modifications such as phosphorylation or glycosylation. Nearly 20 residues were determined for the 48 kD protein. The sequence is composed roughly of equal amounts of hydrophobic groups and
polar groups, most of which carry a negative charge at neutral pH. No serine appears in the N-terminus despite the high composition reported for this amino acid. Aspartic acid is present but does not occur as a run in this particular protein segment. Interestingly, asp is evenly spaced forming triplicate “blocks” (Asp-X-X), where one “X” is alanine. This arrangement predicts a sheet or coil structure. Position twenty is presumed to be methionine. Two phosphorylation sites were predicted at position 14 (Thr) and 9 (Tyr) using NetPhos.2.0 (Kreegipuu et al., 1999).

Limited n-terminal sequence was determined for the 55 kD protein derived from SM and sIM. These proteins were particularly difficult to sequence. Both sequences share identical structure starting at position 3 through position 9. Also, Tyr appears to be present at roughly the same position. Although sequence information is ambiguous at position 1, and especially following position 9, it appears that the 55 kD protein derived from the SM and sIM are the same protein or are at least highly related forms (consistent with chapter 1). Such difficulty in sequencing is commonly associated with regions that are highly modified or where runs are present (Sabsay et al., 1991). In this case, aspartic acid or serine runs are most plausible.

There is little resemblance in the amino acid arrangements between the 55 and 48 kD N-terminal, apart from sharing some amino acid constituents. In the 55 kD sequences, serine and proline appear as do several polar groups including tyrosine, which is embedded in a run and/or heavily modified region. A repeat of QK (glutamine-
uncharged and lysine-charged) appears at the end of the 55 kD SM, but notably, at the beginning of 55 kD sIM. Over-all, the occurrence of runs, heavy modifications, and a generally distinct structural arrangement of the 55 kD N-terminus distinguishes it from the 48 kD N-terminal region, despite their similar amino acid compositions.

Both sequences were searched against protein databanks (Yutaka et. al., 1998), but identified no significant homology with other known protein sequences. However, recently, repeat sequences of D-E-A-D were reported in “Aspein”, an aspartic acid rich protein derived from prismatic layer of the stiff pen shell, *Atrina rigida* (Gotliv et al., 2005).

**Antibody reactivity to mantle and fluid extracts**

To determine the presence of the 48 and 55 kD proteins within the mantle and fluid tissues, the anti-55 kD antibody was used to probe extracts of mantle, extra-pallial fluid (EPF), and hemolymph in ELISA and western analyses. In ELISA (Figure 3.12), the reactivity of the antibody to mantle homogenate exceeded the reactivity to whole SM and indicates matrix protein prominent in this tissue. SM was detected, to a lesser extent, in extracts of EPF and hemolymph which were equally reactive to the antibody and still well over control optical density (OD) values (myoglobin). Western analysis confirmed these results and identified discrete bands within the tissue extracts that correspond to the 48 and 55 kD proteins identified in shell SM. Shown in figure 3.13, both proteins are
detected in EPF (Figure 3.13, Lanes 3 to 5) and hemolymph (Figure 3.13, Lanes 6 to 8) but the 48 kD appears to be the most abundant. The pattern of immuno-reactive proteins of mantle extract (Figure 3.13, Lane 2) and shell SM (Figure 3.13, Lane 1) are strikingly similar and show that the 48 and 55 kD proteins are present in mantle tissue. When mantle extract is resolved by SDS-PAGE, The 48 kD band is visible when low protein loads are resolved (Fig. 3.14 A, lane 6). As observed in ELISA, mantle tissue is highly reactive to the antibody as only 1 ug of protein extract (BCA determined), and a 1:1000 dilution of the antibody was needed for optimal signal to noise. Moreover, the relatively few bands that appear on western blots demonstrate the antibodies’ considerable specificity for the major shell matrix proteins, given the complexity of the mantle extract revealed by SDS-PAGE (Figure 3.14, A Lanes 2 to 4)). Together, these data support that the 48kD and 55 kD proteins are constituents in the shell and mantle tissue and are present in the fluid extracts.
Figure 3.12  Enzyme-linked immunosorbent assay comparing reactivities of the anti-55 kD antibody to soluble shell matrix protein (SM) to those of tissue extracts. Extracts included: mantle tissue homogenate, hemolymph (blood), and extra-pallial fluid (EPF-collected from the extra-pallial cavity). Myoglobin was included as a negative control.
Several additional bands within shell and tissue extracts are moderately immuno-reactive and occur between 33 kD to just above 250 kD (Figure 3.13, Lanes 1 to 8). Bands identified at 33 kD (Figure 3.13, Lanes 3 to 8) and above 250 kD (Figure 3.13, Lanes 5 and 8) appear to be unique to hemolymph and EPF. The 33kD band is lightly immuno-reactive and a prominent constituent in these extracts as evidenced by its intense staining following SDS-PAGE (Figure 3.14, B, Lanes 3 to 9). Several immuno-reactive bands, besides the 48 and 55 kD bands appear between 48 and 120kD in all four extracts. In particular, a 117 kD band is highly immuno-reactive and is present in shell SM and mantle extracts (Figure 3.13, Lanes 1 and 2), and to a lesser extent in EPF (Figure 3.13, Lane 5) and hemolymph (Figure 3.13, Lane 8). Unlike the 48 and 55 kD protein bands, this band is not Stains-all positive and suggests that despite having epitopes in common, this protein, to some extent, is chemically distinct from the 48 and 55 kD phosphoproteins.
Figure 3.13 Western analysis of mantle, EPF and hemolymph extracts compared to shell SM.

Samples were resolved on a 4 to 20% Tris-glycine gradient gel and blotted onto nitrocellulose. Loads for hemolymph, mantle, and EPF extracts were based on BCA protein estimates. The 48 and 55 kDa proteins are indicated by (→).

Lane a - SeeBlue pre-stained protein standards (Novex)
Lane b - Broad range biotinylated protein standards (Bio-Rad)
Lane 1 - SM 35ug dry weight
Lane 2 - Mantle extract 5 ug
Lanes 3 to 5 - EPF with 2, 6, and 15 ug protein respectively
Lanes 6 to 8 - Hemolymph with 2, 6, and 15 ug protein respectively

Note the prominent immuno-reactive bands in the mantle homogenate that correspond to the 48 and 55 kD bands of the SM. These bands are immuno-reactive in both EPF and hemolymph but the 55 kD protein is less conspicuous.
Figure 3.14 SDS-PAGE of extra-pallial fluid (EPF), hemolymph, and mantle extract compared to shell SM.
Samples were resolved on 4 to 20% tris-glycine gels and stained with Stains-all. Except for SM, all protein extracts were determined by BCA protein assay (PIERCE).

A- Gel analysis of varying amounts of mantle extract compared to SM
Lane a - Pre-stained protein standard (Bio-rad)
Lane 1 - SM 50 ug
Lanes 2 to 4 - Mantle extract with loads of 150, 100 and 50 ug respectively
Lane 5 - SM 10 ug
Lane 6 - Mantle extract 2 ug

B- Gel analysis of varying amounts of hemolymph and EPF compared to SM.
Lane a - SeeBlue protein standard (Novex)
Lane 1 - SM 25 ug dry weight
Lane 2 - SM 50 ug dry weight
Lanes 3 to 6 - EPF with loads of 2, 6, 15 and 30 ug respectively
Lanes 7 to 9 - Hemolymph with loads of 2, 6 and 15 ug respectively

Overall, these results conclusively demonstrate the presence of the 48 and 55 kD matrix proteins in shell, mantle, EPF and hemolymph extracts. Based on band intensity, it appears that the 55 kD protein is more variable in quantity among these tissues and is most prominent in mantle extract, while the 48 kD band is consistently prominent in all four extracts. Other proteins were identified in shell, mantle and fluid extracts that share some structural similarities to the 48 and 55 kD proteins based on their immuno-reactivity to anti-55 kD antibody. Some of these proteins appear to be tissue specific while others are present in all four extracts, but vary in quantity.

Characterization of sIM proteins derived from shell layers

An important question to explore is whether different matrix proteins are used to construct the different shell layers. While SM is difficult to obtain from individual shell layers in oyster, sIM is easily extracted from each layer using a whole dissolution process (Mount, 1999). In this process, the entire shell is allowed to dissolve in a 10% solution of...
EDTA, pH 8. Water insoluble matrix (IM) from the outer shell layer, which is composed largely of prismatic calcite and some periostracum, and the inner shell layer, which is mostly foliated calcite, is scraped off sequentially, as a gelatinous layer, during shell dissolution. Newly formed shell at the shell margin is comprised primarily of periostracum, a scleratinized protein layer that is not mineralized, and a thin layer of prismatic calcite, which is deposited onto the periostracum. This periostracal layer is easily collected by breaking off the growing shell edges. Following dissolution, sIM is extracted from the scrapings and edge pieces using urea and SDS. Extracts were subjected to SDS-PAGE as well as Western and ELISA analyses to identify differences in component proteins within each shell layer.

SDS-PAGE shows that the 48 scleratinized and 55 kDa bands are present in all three shell layers (Figure 3.15, A) but differ in their relative amounts. For comparison, whole SM and whole sIM (extracts from whole ground shell) were included and show that the 48 kD protein dominates in SM, while the 55 kD protein is dominant in sIM- a pattern typical for these extracts (Mount and chapter 2). When matrix from isolated shell layers are resolved, the 55 kDa protein is more prominent in the inner shell layer (Figure 3.15, Lane 3) compared to the 48 kDa protein. Both proteins occur to a lesser extent in Outer (Figure 3.15, Lane 4) and Edge (Figure 3.15, Lane 5) layers and appear to be in equal amounts based on staining intensity. While the inner shell layer is more complex than the Outer and Edge layers, most of the bands visible in Outer and Edge are also
present in the Inner layer. An exception is a prominent 12 kD band that appears in Edge and, to a lesser extent, in the Inner layer but is absent in the Outer layer.
Western analyses also identified similar banding patterns between shell layers (Figure 3.15, B). Positive staining proteins are assumed to be structurally similar to the 48 and 55 kD SM proteins. Like SDS-PAGE results, protein patterns of all the extracts tested (Figure 3.5, Lanes 1 to 5) are similar but some bands vary in their intensity. This result can be interpreted as differences in the quantity of protein between extracts or differences in reactivity with the anti-55 kD antibody which indicates structural variation among proteins. For example, the Outer layer (Figure 3.5, Lane 4) is “missing” 48 kD
band which was identified by Stain-all on SDS-PAGE (Figure 3.5, A Lane 4). However, a band slightly higher at 49 kD is immuno-reactive and may be an isoform of the 48 kD. Additional evidence for slight changes in structure is that the 48 kD protein, identified in Edge by Stains-all and antibody staining, migrates slightly lower in this extract (compare Figure 3.15, A Lanes 4 and 5; and in Figure 3.15, B compare Lane 5 with lanes 1 to 3).

Western analyses also identified bands that may be unique to specific shell layers. For example, an immuno-reactive doublet at 78 kD appears in the Inner layer but is not apparent in the Outer layer.

An observed inconsistency is that the anti-55 kD antibody recognizes the 55 kD band in SM (Figure 3.15, B Lane 1), but does not recognize the 55 kD band in any of the sIM extracts, including the whole sIM extract where it is particularly abundant (compare Figure 3.16, A Lane 2 with Figure 3.16, B Lane 2). In this blot (Figure 3.15, B); the 55 kD region of the sIM extracts appears as a negative band (Figure 3.15, B Lanes 2 to 5) which may indicate that the protein does not adhere well to the membrane under the conditions used for this experiment. It is also possible that the antibody, which was made against 55 kD protein derived from the SM, does not recognize 55 kD derived from sIM. Because these proteins have similar compositions, this lack of reactivity is not likely due to disparate epitopes, rather some other structural difference that may hinder the antibody from reacting with protein epitopes. In any case, the different behavior of the 55 kD from SM and the 55 kD in sIM in western experiments suggests that these proteins are different in some way and is further evidence that the matrix is composed of closely
related but distinct protein isoforms. This observation underscores the importance of identifying corresponding genes and the processing which occurs, pre and post translationally.
Figure 3.16  Enzyme-linked immuno-sorbent assay demonstrating reactivity of the anti 55 kD antibody to sIM matrix extracts from the three principal shell layers. Matrix protein was solubilized from IM isolated from the outer, inner, and edge (shell margin) shell layers according to Mount, 1999. SM and sIM are matrix extracts from whole shell. Myoglobin was included as a control.
With respect to matrix protein patterns of shell layers, a final and more general observation is that starting from Edge (Figure 3.15, A and B Lane 5), which is the first shell layer deposited, and moving toward the Inner shell layer (Figure 3.15, A and B Lanes 4 and 3), a progression of expression of proteins is apparent. For example, Western analysis shows that the Edge extract contains proteins similar to those found in the Inner layer as well as SM, then as the Outer layer is deposited, the expression pattern changes. This result is confirmed by ELISA (Figure 3.16) which shows that whole sIM, Inner sIM, and SM have more proteins in common, followed by edge extract and then Outer which is well above background, but less immuno-reactive. Therefore, proteins similar to the 48 and 55 kD from SM are most prevalent in the active growing edge and inner shell layer.

Discussion

The aim of this study was twofold- to characterize the structures of the major 48 and 55 kD phosphoproteins identified in the organic matrix, and to identify and compare these proteins within the shell, mantle, and EPF. Antibodies made to matrix proteins were used in conjunction with conventional sequencing efforts to gain information about the structures of the 48 and 55 kD proteins. Anti- matrix antibodies used to determine the relation among matrix components as well as their occurrence within the major shell layers and areas of mineralization. Efforts first focused on generating and characterizing anti-matrix antibodies.
Interestingly, chicken and rabbit hosts produced anti-matrix antibodies with distinct reactivity’s. While all the antibodies reacted well with their matrix protein antigens, the chicken anti 48 kD antibody cross-reacted considerably with phosphophoryn and phosvitin, two proteins known for their extraordinary content and arrangement of phosphoserine. Phosphophoryn is major matrix component of vertebrate teeth and bone and comprises a family of acidic proteins with reported compositions comparable to the 48 and 55 kD SM proteins (Chapter 2) where 50 to 90% of the protein is composed of serine and aspartic acid residues (George, et al., 1993, 1996, 1998). Serine, most of which is phosphorylated, is arranged as short runs interspersed with aspartic acid which are abundant throughout the phosphophoryn structures (George et al., 1998). Phosvitins, a class of phosphoproteins found in the eggs of oviparous vertebrates, have multiple runs of phosphoserine that occur in remarkable lengths up to n=14 (Maurizio et al., 2004; Byrne et al., 1984). The high reactivity of the anti 48 kD antibody with these proteins suggests that the primary epitopes recognized are phosphoserine regions. More importantly, this result provides evidence that the 48 kD protein contains phosphoserine runs.

Poly-phosphoserine domains are thought to be critically important in the mechanisms of crystal interaction, ion accumulation, and binding to collagen matrix in vertebrate teeth and bone formation (Veis et al., 1991, Dahl and Weis 2003). For example, specific motifs within the phosphophoryn structure, including triplet (DSS)n and doublet (SD)n repeats, form domains with distinct ion and crystal binding properties.
which are dependant on the presence of phosphate (George et al., 1996; He et al., 2005). Similarly, a fraction of oyster shell SM (RP-1) is a more effective inhibitor of in vitro mineralization than dephosphorylated RP-1 (Borbas et al., 1991). The finding that the 48 kD phosphoprotein (and likely the 55 kD protein) shares some structural epitopes with phosphophoryn suggests they may have similar domains and therefore may perform analogous functions during mineral formation. And, despite their distant relationship, some domains of shell matrix phosphoproteins and those derived from vertebrate tooth and bone may be conserved. The recent discovery that shell derived Aspein shares DNA sequence homology with phosphophoryn and bone sialoprotein further supports the idea that functional domains of mineral derived matrix proteins may be conserved between phyla (Tsukamoto et al., 2004).

Antibodies generated in rabbits did not cross-react with phosvitin or phosphophoryn and therefore were considered more specific for shell organic matrix. The epitopes most recognized by the anti 48 and 55 kD antibodies were apparently to unique regions of matrix protein and not to phosphorylated regions given that antibody reactivity to SM was unaffected after phosphate groups were removed. These antibodies were best suited for studies looking at relatedness among shell matrix components, as well as identifying matrix proteins within the tissues and shell.

Western analyses using the anti-SM and sIM antibodies shows that SM (water soluble), and sIM (urea and SDS soluble) are highly related and complex classes of
proteins. This finding is consistent with results here (Chapter 2), and previous studies where compositions of SM and IM extracts were found to be nearly indistinguishable except for an increased presence of basic amino acids in IM (Rusenko, 1988; Kawaguchi and Watabe, 1993; Mount, 1999). In this study, immuno-staining patterns of SM and sIM components were shown to be nearly identical. In fact, the only visible difference between these two protein classes is that their component proteins vary in intensity which indicates that SM and sIM have variable amounts of similar proteins.

The high relation among shell matrix is most evident in studies using the anti-48 and 55 kD antibodies. In chapter 2, the amino acid compositions of the 48 and 55 kD were reported to be nearly identical, except for a slightly higher mol% of hydrophobic groups in the 55 kD protein. Here, the anti 48 and 55 kD antibodies reacted similarly in both ELISA and Western analyses, where absorbance plots and blot patterns were essentially the same. Together, these results strongly suggest that the 48 and 55 kD proteins have similar structures. In fact, the 55 kD antigen produced an antibody that reacted best with the 48 kD protein suggesting that the epitopes shared by both proteins are in greatest amount in the 48 kD protein. These antibodies also reacted well with several other proteins within the SM and sIM not detected with Stains-all following SDS-PAGE, and indicates the presence of another class of proteins, distinct from the phosphoprotein class, yet still related.
Despite the high relation between the 48 and 55 kD proteins, they are distinct in their N-terminal regions. Sequence determined by Edman chemistry, revealed that in this region, the 48 and 55 kD proteins varied in the specific groups present, their arrangements, and the extent they are modified. The 48 kD protein contains a repeat pattern of “blocks” (Asp-X-X), where one X is alanine. This arrangement predicts a random coil or β-sheet secondary structure. Both motifs are commonly predicted for mineral derived proteins and thought to be ideal for protein-crystal interactions as well as protein assembly and high capacity ion binding (Addadi and Weiner, 1985, 1989; Wilt, Choi and Kim, 2000; Asakura et al., 2006). Notably, no serine is present this segment of the n-terminus. The DEAD structural motif has recently been identified in Asprich proteins where it is arranged as repeats and is purported to have Mg$^{2+}$ binding capabilities (Gotliv et al., 2005). The 55 kD protein has different groups including proline and lysine and has highly modified regions, presumed to be poly-phosphoserine, or Asx runs. Both motifs are refractory to Edman chemistry.

It is important to note that the N-terminal sequences reported here are thought to represent the true N-terminal regions of these proteins and not the N-termini of cleaved products, even though protein cleavage is a mechanism known to occur to mineral derived matrix proteins (Lowenstam and Weiner, 1989). Evidence supporting that the 48 and 55 kD proteins are not cleaved is that they appear to be the same size in the mantle and shell SM which suggests they are in a mature, monomeric form immediately after secretion (discussed further below).
Some general structure estimations of the 48 kD protein were extrapolated from how a variety of chemical and enzymatic treatments affected its structure. First, the failure of several common enzymes to cleave the protein despite the presence of substrate groups suggests a structure that is “protected” by bulky side groups and folded in such a manner rendering it essentially impervious to many common proteases. Even Protease K and pepsin, which are enzymes that hydrolyze peptide bonds indiscriminately, produced a limited number of fairly large protein fragments from the 48 kD protein while control protein was virtually fragmented completely. Also lending to its robust structure is its homogenous composition. Because roughly 60% of the protein is serine (much of which is phosphorylated) and aspartic acid (Chapter 1), it would seem that at least some of these groups would be arranged as runs. Some enzymes are unable to act efficiently at substrate sites flanked by stretches of a single amino acid such as aspartic acid. V8 protease is such an enzyme (Inglis, 1983). The fact that 48 kD protein is completely resistant to this enzyme is strong evidence for the presence of (-asp-)n sequences and is consistent with earlier studies where fractions of SM were suggested to have poly-aspartate runs (Rusenko et al., 1991). Also, only 4 or 5 products resulted from Asp-N endoprotease despite having numerous substrate groups. The fact that most of the resulting fragments are only slightly smaller than the parent protein suggests that poly-asp runs likely exist near the middle of the protein backbone. These poly-asp runs are not easily cleaved because of their repetitive nature. The relatively sparse presence of asp on the N-terminus would be more susceptible to cleavage and which would explain the small change in molecular weight of the resulting products.
Asp-rich domains acid runs have been identified in many mineral derived proteins (Zhang and Zhang, 2006). In shell formation, their significance is most often envisioned as crystal growth inhibitors through the specific adsorption to crystal surfaces (Addadi and Weiner 1985; Aizenberg et al., 1994; Gotliv et al., 2003) and binding lattice ions thereby controlling the polymorph deposited (Belcher et al., 1996; Falini et al., 1996). The recently discovered acidic “AP8” proteins derived from nacre are composed of approximately 35 mol% Asx and 40% Gly and are therefore comparable to the 48 and 55 kD proteins. These proteins were shown to alter the morphology of growing crystals by preferentially adsorbing to specific areas on the crystal faces (Fu et al., 2005).

Aspein, reported to be the most acidic shell protein known, is proposed to function by sequestering either Ca$^{2+}$ or Mg$^{2+}$ via long runs of Asx which act to control the local relative concentrations of these ions thereby favoring the formation of prismatic calcite over aragonite (Tsukamoto et al., 2004). However, because hemocyte cells have been shown to supply calcite seed crystals for calcitic folia growth in oyster (Mount et al., 2004), an extra-cellular sequestering function, as a nucleating mechanism, would not be required in this case. Rather Aspein and similar proteins may act primarily as an ion sink for supporting crystal growth. In any case, because this study supports that the 48 and 55 kD proteins have poly-asp domains, it is reasonable to speculate that they have similar activities.
Over-all, this study shows that numerous related proteins occur in SM and sIM. Specifically, the 48 and 55 kD proteins from SM share nearly identical composition and share common epitopes which indicates their structures are similar in many regions with the exception of the N-terminus. Anti-matrix antibodies identified additional components which were not visible with Stains-all. These components are likely less phosphorylated and but share domains with the 48 and 55 kD to varying degrees. Together, these proteins constitute a major acidic class and bring into question its origin. One possibility is that the 48 and 55 kD are expressed through a single gene that is alternatively spliced, a mechanism known to generate related forms of matrix proteins in teeth (phosphophoryns) (George et al., 1998). Also possible is that these and other related proteins are members of a multi-gene family. Recently, a family of glycine rich matrix proteins called the “shematrins” has been identified in association with the prismatic shell layer formation in *P. fucata*. These proteins share many structural domains but are encoded by distinct cDNA’s (Zhang et al. 2006; Yano, et. al., 2006).

A second phase of this study involved using the anti-55 kD antibody to examine matrix protein levels and patterns within the tissues and shell layers, and to detect possible structural changes that may have occurred. Mineral derived proteins are heavily post-translationally modified and include phosphorylation, glycosylation, protein cleavage, and assembly (Lowenstam and Weiner, 1989). In these studies, particular attention was given to molecular weight shifts and changes in immuno-reactivity of the
48 and 55 kD proteins within the tissues and shell layers, which may indicate the occurrence of structural change or the presence of protein isoforms.

Both the 48 and 55 kD proteins are present in the mantle tissue and extra-pallial fluid extracts where the 48 kD appears to be more abundant. The mantle is the organ responsible for making and secreting matrix components into the extra-pallial fluid (EPF), which lies between the mantle tissue and the shell (Simkiss and Wilbur, 1989). This fluid is composed of inorganic ions (Crenshaw, 1972a), and a variety of organic macromolecules (Wilbur and Bernhardt, 1984; Mount, 1991; Nair and Robinson, 1998; Hattan et al., 2001). It is within this medium that matrix is thought to interact with itself and with crystal surfaces influencing the growth of the mineral deposited. The presence of the 48 and 55 kD proteins in the EPF supports that they are important components in the shell formation process.

The 48 and 55 kD proteins may undergo different processing en route to the mineralization front. The 48 kD protein appears largely unaltered from the time just prior to its secretion, to the time it is incorporated into the shell. This deduction is based on the observation that the migration and immuno-reactivity of the 48 kD band from the mantle and EPF is comparable to that identified in shell SM. It may be that this protein obtains its mature form inside mantle cells, following modification in the Golgi, and remains in this form through its transport to the mineralization front and incorporation into the mineral phase. The 55 kD SM protein, to some degree, also appears to mature
intracellularly as indicated by its identical size in mantle homogenate compared to shell matrix. However, an interesting difference is that this protein is not nearly as prevalent in tissues outside the mantle. It seems to “disappear” from the EPF and is relatively minor in the SM. This result is evidence that some portion of the 55 kD SM protein undergoes further processing inside the EPF soon after it is secreted by mantle cells. One possibility is that it self associates or incorporates into an existing assembled matrix which is an idea congruent with chapter 1, and the common observation that some SM associates tightly with IM (Crenshaw, 1990; Dauphin, 2002). In any case, these results support that despite their high relation, the 48 and 55 kD proteins have distinct functions.

The presence of the 48 kD protein and, to a much lesser extent, the 55 kD protein in the hemolymph raises questions regarding its function in shell formation and the possibility of other functions not directly associated with mineral deposition. Its presence in this fluid is consistent with the identification of matrix producing cells on the pallial side of the mantle (Myers et al., 2007). Proteins identified in the blood and EPF of *M. mercenaria*, two of which were common to both fluids, sequester calcium at low affinity and are thought support calcium exchange between the blood and EPF (Nair and Robinson, 1998). A histidine and aspartic acid rich protein identified in the EPF of *M. edulis* has been associated with shell formation (Hattan et al., 2001). This protein has recently been sequenced and was found to share homology with a heavy metal binding protein found in hemolymph and therefore proposed to carry out multiple functions as a calcium transport protein, a matrix protein, and as a detoxifier of heavy metals (Yan et
al., 2005). With these studies in mind, the presence of the 48 and 55 kD proteins in oyster hemolymph supports their potential for multi-faceted functions.

The identification of additional proteins structurally related to the 48 and 55 kD proteins that occur in shell and hemolymph raises the question of how these, too, function in both tissues and how their activities are related to mineral deposition. Although it is possible that proteins from fluid tissues may become “trapped” within the matrix as shell forms, and their presence an artifact as has been suggested by Lowenstam and Weiner (1989), new evidence shows that the process of shell building is tightly linked to the cellular activity of hemocytes (Mount et al., 2004) and so common proteins between the shell and hemolymph would be expected, especially those that help direct hemocytes from the hemolymph to the site of mineral deposition and support their activity there.

The last tissue examined was the individual layers of the shell to see if a correlation exists between matrix protein patterns and the formation of the two major mineral layers in oyster shell- the outer calcitic prismatic layer and the inner calcitic foliated layer. A common belief is that IM proteins act with SM components to direct crystal shape, size and the polymorph deposited (Lowenstam and Weiner, 1981; Addadi and Weiner, 2006), resulting in one of five types of conserved microstructures that occur in shell (Carter, 1980). A sound hypothesis deduced from this premise is that the matrix protein composition, as well as specific protein structures, somehow influence which microstructure and mineralogy, is deposited. That is to say, the particular shape and type
of crystal that forms may be dictated by the unique protein-protein and protein-crystal interactions specific to the matrix proteins of a shell layer having a particular mineralogy or microstructure. The origin of this theory comes from early comparative studies where matrices from distinct shell layers, within and between species, were chemically compared and found to be different (Travis et al., 1967; Keith et al., 1993; Dauphin and Denis, 2000; Lee and Choi, 2007). Few studies have examined the relatedness among components and where they appear in the shell.

In this study, shell layer IM matrices were isolated by a whole shell dissolution process that essentially allowed examination of matrix components beginning from the onset of prismatic layer formation (periostracum collected at the shell edge), followed by the prismatic layer (outer), and finally from the foliated layer (inner), which forms the bulk of oyster shell. The results show that sIM from all three extracts have distinct protein patterns varying mostly in the intensity of similar bands. The component proteins are related in structure to the 48 and 55 kD phosphoproteins, and a few appear to be unique to a particular shell layer. Notably, there is evidence that the 48 kD protein varies slightly in structure among shell layers as Western analyses identified a 48 kD band that differed slightly in molecular weight in all three extracts. The fact that this protein is phosphorylated, and the extent of matrix protein phosphorylation has been shown to alter adsorption to crystal surfaces in a variety of systems, it is tempting to suggest that differentially phosphorylated isoforms of the 48 kD protein influence which microstructure is constructed.
This study provides evidence for two mechanisms by which matrix protein influences the construction of a particular microstructure. The first is timing. The data supports that “layer specific” proteins, as well as proteins common to more than one layer, are secreted at precise times as the different shell layers form. Western and ELISA show a general timed appearance of proteins related to the 48 and 55 kD which are more prevalent in the edge (periostracum) and inner (folia) layers and suggests the outer (prismatic layer) matrix has a distinct component composition. More specifically, a 78 kD doublet appears in the inner shell layer but is absent from outer layer. Therefore, this protein seems necessary for the construction of folia, but may not be critical for prism formation. A 12 kD protein identified by Stains-all is prominent in the edge but does not appear in the outer layer and is barely apparent in the inner layer. This protein must be important for the initiation of the prismatic layer, but is not used as a major constituent during bulk prism formation.

Recently, the cDNA’s encoding several matrix proteins from the pearl oyster, *P. fucata* have been shown to localize in specific regions in the mantle, which correspond to the inner nacre (aragonite) layer and the outer prismatic layer (Sudo et al., 1997; Takeuchi and Endo, 2006). Among these proteins, Nacrein, a water soluble protein thought to catalyze the production of carbonate, is found in both shell layers, while Aspein, a water soluble matrix protein and MS131 and Prismalin-14, both water insoluble, are purportedly specific to the prismatic layer. MS160 and N16 were associated exclusively with the nacre layer (Tsukamoto et al., 2004; Takeuchi and Endo,
2006). These studies illustrate well a correlation between specific matrix proteins being secreted at particular times as the different shell layers are deposited. Moreover, Takeuchi (2006) proposed that the expression of proteins specific for each shell layer are “dually coordinated” and their genes may be regulated by a common regulatory factor. To date, all recent studies comparing matrix protein components between shell layers have been carried out with species that have prismatic calcite outer layers and nacre aragonite inner layers. This study is unique because it involves a single mineralogy, calcite. And so, it may be that differences in matrix patterns noted here are more related to control of crystal shape, growth, and orientation than with initiation of a specific polymorph.

The second mechanism involves varying the levels of particular matrix proteins that are deployed to the mineralization front as the different layers are formed. As examples, this study shows that Stains-all positive forms of the 48 and 55 kD are present in both mineral layers but their levels appear to be regulated. Stains-all and immuno-staining show that the 48 kD protein is significantly more abundant in SM from the inner folia layer which suggests it plays a primary role, as a soluble protein, in forming this layer. Lower levels of the 48 kD protein are detected in all three sIM extracts by Stains-all, however, immuno-staining shows that of these three extracts, the 48 kD, or a closely related form, is most prominent in the periostracal extract. Thus, it appears that an increased amount of the 48 kD protein is important for the onset of prismatic calcite formation. Stains-all indicates that the 55 kD protein as most prevalent in the sIM from
folia compared to SM (folia) or prismatic and periostracal sIM extracts. Therefore, an increased level of the 55 kD protein correlates with the construction of foliated calcite, as a protein that associates with IM.

Some recent studies support that levels of matrix proteins may influence microstructure. Nacrein appears in both calcitic and nacreous layers but is expressed more abundantly in the mantle regions corresponding to prismatic layer formation (Takeuchi and Endo, 2006). In the otoliths of teleost fish, a water soluble acidic protein called “starmaker” was identified as a protein that strongly influenced mineralization patterns, and interestingly shares homology with dentin proteins. Alteration of the expression of this protein had a profound affect on the morphology on the developing otolith which resulted in the formation of elaborate star structures instead of the normally round smooth flat appearance. Moreover, it was found that otolith morphology was strongly influenced by expression levels of this protein (Sollner et al., 2003). Accordingly, it is possible that the formation of prismatic and foliated calcite layers in shell may be, in part, influenced by the levels of matrix proteins.

Together, these strategies, supported by this study and previous studies, reveal the highly orchestrated and specific nature of matrix protein deployment to the mineralization front as the different shell layers are formed. In this study, there was no major protein component identified as unique to either mineral layer, rather the proteins involved in making foliated and prismatic calcite are structurally related and may
influence mineral deposition and growth mostly through a regulated release of identical or protein isoforms. The idea of mineralogy and microstructure being determined by a specific suite of matrix components is examined further in the following chapter.

Summary

The results of this study provide insight into the general structure of 48 and 55 kD proteins, their relation with other acidic matrix components and where proteins of this class localize within the shell, mantle, EPC and hemolymph.

The 48 and 55 kD proteins are highly related phosphoproteins which occur in both the foliated and prismatic shell layers. While their N-terminal sequences are distinct, immuno-reactivity data indicates that they are structurally similar and probably originate from a common gene or gene family. These proteins also occur in the mantle, hemolymph, and EPC fluid extracts. The 48 kD does not appear to change structurally among these extracts. The 55 kD protein is less prominent in EPC fluid and may indicate that this protein undergoes a structural change following its secretion into the EPC.

Western and protein cleavage studies suggest poly-ser and poly-asp domains are present in the 48 kD protein. These domains very likely occur in the 55 kD protein based its high relation with the 48 kD protein. The resistance of the 48 kD protein to cleavage by V8 protease, and the fact that few fragments of substantial size were generated by
Asp-N protease indicates that poly-asp domains are prevalent and likely occur near the N and C terminal regions. The cross-reactivity of the anti 48 and 55 kD antibodies with other known phosphoproteins indicates the presence of poly-ser domains. Furthermore, the cross-reactivity with phosphophoryn, a vertebrate tooth derived protein, suggests that proteins from mineralized tissues share common domains across phyla. The conserved region is likely in the phosphoserine domain. Collectively, this data and amino acid compositions reported in Chapter 2 indicate that the structures of the acidic class of proteins from oyster shell folia are highly repetitive and modular.

The water solubilized matrix (SM) and the SDS/urea solubilized matrix (sIM) contain mostly similar proteins which are members of the acidic class. The anti 48 and 55 kD antibodies identified a similar pattern of protein bands in the SM and sIM fractions. These proteins share epitopes with the 48 and 55 kD proteins but are not as acidic as indicated by their lack of Stain-all staining. This difference may be attributable to lack of phosphorylation.

Finally, this study shows that SM from the foliated layer and sIM from the three shell layers are distinguished mostly by having variable amounts of similar proteins including the 48 and 55 kD proteins. The 48 and 55 kD proteins appear to be involved in the formation of both shell layers, but are most abundant in the foliated layer. These observations are evidence that the timing and levels of matrix proteins are important during layer formation and may influence the mineral type and microstructure deposited.
CHAPTER FOUR

CORRELATION OF ORGANIC MATRIX COMPOSITION TO SHELL MICROSTRUCTURE AND MINERALOGY

Introduction

Shell formation is a sequenced event which occurs at the shell margin in intimate association with the mantle organ (Simkiss and Wilbur, 1989), hemocytes cells (Mount et al., 2004) and a coordinated secretion of organic matrix components from specialized cells in the mantle epithelium (Myers et al., 2007; Takeuchi and Endo, 2006; Jolly et al., 2004). The organic matrix is presumed direct the deposition, orientation, and growth of the mineral phase (Watabe and Wilbur, 1960). The resulting structure is most often multilayered, each of which are composed of highly ordered crystalline aggregates with exceptional mechanical properties that far exceed those of their inorganically formed counterparts (Daw, 2004). The strength of biominerals is attributed to the organic matrix which surrounds and is occluded within each crystal unit, and appears between aggregate crystal sheets (Addadi et al., 2006).

The mineral layers of bivalves have been described in great detail first by Boggild (1930), followed by more recent descriptions (Taylor et al., 1969, Carter, 1979, Carter, 1990). From these studies, shell layers are defined according to the type of mineral deposited, called mineralogy or polymorph, and the arrangement of the crystals within the layer, which is referred to as the microstructure or texture. With few exceptions,
molluscan shell is composed of calcium carbonate which, in bivalves, occurs most commonly in the form of aragonite and calcite. Carter (1980) describes 8 types of arrangements of crystals that form the basic units of shell mineral (Figure 4.1, Column 1), and categorized these types into five microstructural groups that commonly occur in bivalves (Figure 4.1, Column 2). Using this classification and the collective observations of numerous researchers, an exhaustive report summarizing the shell structures for Bivalvia was created (Carter, 1990).
<table>
<thead>
<tr>
<th>Microstructure groups and their principal varieties</th>
<th>Microstructure categories and their constituent microstructures</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Prismatic</td>
<td>I. Aragonitic prismatic</td>
</tr>
<tr>
<td>A. Simple prismatic</td>
<td>II. Calcitic prismatic</td>
</tr>
<tr>
<td>B. Fibrous prismatic</td>
<td>III. Nacreous (aragonitic)</td>
</tr>
<tr>
<td>C. Spherulitic prismatic</td>
<td>IV. Porcelainous (aragonitic)</td>
</tr>
<tr>
<td>D. Composite prismatic</td>
<td>A. Aragonitic crossed lamellar</td>
</tr>
<tr>
<td>II. Spherulitic</td>
<td>B. Aragonitic crossed acicular</td>
</tr>
<tr>
<td>III. Laminar</td>
<td>C. Aragonitic complex crossed lamellar</td>
</tr>
<tr>
<td>A. Nacreous</td>
<td>D. Aragonitic crossed-matted/lineated</td>
</tr>
<tr>
<td>B. Regularly foliated</td>
<td>E. Aragonitic homogeneous</td>
</tr>
<tr>
<td>IV. Crossed</td>
<td>V. Foliated (calcitic)</td>
</tr>
<tr>
<td>A. Crossed lamellar</td>
<td>A. Regularly foliated</td>
</tr>
<tr>
<td>B. Crossed acicular</td>
<td>B. Calcitic crossed lamellar</td>
</tr>
<tr>
<td>C. Complex crossed lamellar</td>
<td>( = crossed foliated)</td>
</tr>
<tr>
<td>D. Crossed-matted/lineated</td>
<td>C. Calcitic complex crossed lamellar</td>
</tr>
<tr>
<td>V. Homogeneous</td>
<td>( = complex crossed foliated)</td>
</tr>
<tr>
<td>A. Homogeneous s.s.</td>
<td></td>
</tr>
<tr>
<td>B. Graular</td>
<td></td>
</tr>
<tr>
<td>VI. Isolated spicules or spikes</td>
<td></td>
</tr>
<tr>
<td>VII. Isolated crystal morphotypes</td>
<td></td>
</tr>
</tbody>
</table>

Microstructures rarely encountered in the Bivalvia

| I. Spherulitic                                    |
| II. Isolated spicules or spikes                   |
| III. Isolated crystal morphotypes                 |

Calcitic prismatic  Regularly foliated  Crossed foliated

Nacre (sheet nacre)  Aragonitic homogeneous  Aragonitic crossed lamellar
Figure 4.1 Table showing the various varieties of shell microstructures based on their structural arrangement (Column 1) and how they occur most commonly in major shell layers of Bivalvia (Column 2).

SEM micrographs of the microstructures represented in this study are labeled accordingly. Table and micrographs were reproduced from “Environmental and Biological Controls of Bivalve Shell Mineralogy and Microstructure”, by J. G. Carter in Skeletal Growth of Aquatic Organisms, eds. D.C. Rhoads and R.A. Lutz. 1980. pp 69 to 113; also see, Appendix 2, pp 645 to 673.

In many bivalves, the first mineral layer deposited is prismatic microstructure which is composed of large, adjacently positioned columnar prisms that are deposited and grow with their long axis perpendicular to the shell surface (Simkiss and Wilbur, 1989). Prismatic calcite crystals have been referred to as “mega crystals” because of their remarkable 50um - 100um diameters (Susuki and Uozumi, 1981). Prisms are surrounded by a thick interprismatic matrix envelope or sheath which are subsequently filled in with mineral and are presumed to strongly influence prism morphology. Its high organic content make prismatic shell exceptionally tough and ideal as an outer protective mineral layer (Checa, 2005).

The inner shell layer is most often composed of nacreous, foliated, or porcelaneous microstructure, which is deposited in multiple layers onto the outer prismatic layer and causes shell thickening (Simkiss and Wilbur, 1989). Nacre and foliated microstructures are composed of flattened, regularly shaped crystal units which grow with their long axes parallel to the plane of the shell and coalesce to form sheets or lamina, (Carter, 1980). Homogenous structure consists of irregularly shaped but similarly sized crystal units with no regular arrangement. Of these microstructures, folia
is the most highly organized with its crystals arranged in a coherent crystallographic orientation (Checa et al., 2007).

Amid the diversity of species that comprise Mollusca, the fact that many, including Bivalvia, construct repeatedly these microstructures and mineralologies indicates that the process of shell formation is conserved. Given that shell matrix is purported to control the deposition and arrangement of the crystal units, it would seem that the components themselves would be conserved as well. Numerous analyses of the water soluble portion of shell matrix (SM) have been reported for a variety of molluscan species (Travis, 1967 and reviewed in Keith et al., 1993). Collectively, these studies established the ubiquity of the amino acids Asp, Ser, Gly, and Glu which are thought to be critical for conferring “crystal shaping” and ion binding activities to SM proteins (Weiner and Addadi, 1991; Gotliv et al, 2003; De Yoreo and Dove, 2004).

Even more compelling, SM proteins derived from nacre and calcitic shell layers have been shown to promote the formation of crystals in vitro with corresponding mineralogies (Falini et al., 1996; Belcher et al., 1996; Kono et al., 2000). While these studies provide convincing evidence for layer specific proteins determining the mineralogy and microstructure deposited, each was conducted using a single species.

A limited number of studies have compared matrix components among species having similar shell structures (Dauphin and Dennis, 2000,) and species with different
shell structures (Keith et al., 1992; Dauphin, 2001). The available data is restricted to matrix derived from aragonite shell or from mixed shell matrix extracts representing more than one mineralogy and microstructure. Furthermore, comparing data between studies is problematic due to the difficulty in resolving reproducibly, discrete protein components (Weiner et al., 1977; Samata, 1990).

The aim of this study was to further explore a possible correlation between SM component characteristics and shell structure by using a comparative approach. The SM from a range of bivalve species, representing three prominent microstructures and varying degrees of phylogenetic relation, were chemically analyzed and compared. SM comparisons between individuals that are distantly related with similar shell microstructures, and species more closely related with dissimilar shell microstructures are especially important to identifying microstructure specific proteins that are conserved across species.

Materials and Methods

Selection of animals and explanation of phylogenetic relatedness

Thirteen species of bivalves, having various shell structures and degrees of relation, were selected to investigate if a particular matrix protein composition promotes the formation of a specific microstructure and/or mineralogy. A single valve of each species is shown in Figure 4.2. With the exception of two, the shells are composed of
two major mineral layers each with distinct microstructures and mineralogies. These
layers are described in Table 4.1 and the specific layers subjected to study are indicated.
Species were classified according to Carter (1990) and are based on historical and recent
morphological studies (Table 4.2). A molecular phylogeny of bivalves, constructed from
rDNA sequence data derived from the 18S subunit, was included for additional reference
(Adamkewicz et al., 1997). Of the 32 bivalve species examined in that study, seven of
the species are represented here (Figure 4.3).
Figure 4.2  A single valve from the selected bivalve species.
Specimens of C. virginica and M. mercenaria were collected from 60 Bass Creek in Winyah Bay, Georgetown, SC. Specimens of C. gigas were given by Paul Montagena from the University of Texas, Galveston. Specimens of O. edulis were a gift from Dr. Andy Brand from the Isle of Man, Britain. Specimens of A. simplex, P. magellanicus, A. iradians, A. rigida, and D. robustum were purchased from Gulf Specimen Marine Lab. Specimens of E. complanata were given by Dr. Arnie Eversole, and collected from Lake Hartwell in Clemson. Specimens of M. edulis were purchased from a commercial fishery. Specimens of P. fucata were a gift from Dr. Akira Machii, Mie-Ken, Japan. Specimens of P. gibbosa were collected from Sanibel Island, FL., by Dr. Andy Mount.
### Table 4.1 Shell Microstructure and Mineralogy of Selected Bivalve Species (Carter, 1990).

<table>
<thead>
<tr>
<th>Species</th>
<th>Shell layer and relative thickness</th>
<th>Mineralogy/ Microstructure¹</th>
<th>Shell layer analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. virginica</td>
<td>outer- very thin</td>
<td>calcite/prismatic</td>
<td>Inner</td>
</tr>
<tr>
<td></td>
<td>inner- very thick, bulk of shell</td>
<td>calcite/foliated</td>
<td></td>
</tr>
<tr>
<td>C. gigas</td>
<td>outer -very thin</td>
<td>calcite/prismatic</td>
<td>Inner</td>
</tr>
<tr>
<td></td>
<td>inner- very thick, bulk of shell</td>
<td>calcite/foliated</td>
<td></td>
</tr>
<tr>
<td>O. edulis</td>
<td>outer- very thin</td>
<td>calcite/prismatic</td>
<td>Inner</td>
</tr>
<tr>
<td></td>
<td>inner- very thick, bulk of shell</td>
<td>calcite/foliated</td>
<td></td>
</tr>
<tr>
<td>P. gibbosa</td>
<td>outer- thick, bulk of shell</td>
<td>calcite/foliated</td>
<td>whole shell</td>
</tr>
<tr>
<td></td>
<td>inner- thin</td>
<td>aragonite /porcelaneous</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(crossed lamellar)</td>
<td></td>
</tr>
<tr>
<td>A. simplex</td>
<td>outer- thick, bulk of shell</td>
<td>calcite/foliated</td>
<td>Inner</td>
</tr>
<tr>
<td></td>
<td>inner- very thin</td>
<td>aragonite/porcelaneous</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>lamellar</td>
<td></td>
</tr>
<tr>
<td>A. irradians</td>
<td>outer- very thin</td>
<td>calcite/prismatic</td>
<td>Inner</td>
</tr>
<tr>
<td></td>
<td>inner- thick, bulk of shell</td>
<td>calcite/foliated</td>
<td></td>
</tr>
<tr>
<td>P. magellanicus</td>
<td>whole shell</td>
<td>calcite/foliated</td>
<td>whole shell</td>
</tr>
<tr>
<td>P. fucata</td>
<td>outer- thick</td>
<td>calcite/prismatic</td>
<td>whole shell</td>
</tr>
<tr>
<td></td>
<td>inner- thin</td>
<td>aragonite/ nacreous</td>
<td>inner</td>
</tr>
<tr>
<td>A. rigida</td>
<td>outer- thick</td>
<td>calcite/prismatic</td>
<td>whole shell</td>
</tr>
<tr>
<td></td>
<td>inner- thin</td>
<td>aragonite/ nacreous</td>
<td>inner</td>
</tr>
<tr>
<td>M. edulis</td>
<td>outer –thick</td>
<td>calcite/prismatic</td>
<td>inner</td>
</tr>
<tr>
<td></td>
<td>inner - thin</td>
<td>aragonite/ nacreous</td>
<td></td>
</tr>
<tr>
<td>M. mercenaria</td>
<td>outer- thin</td>
<td>aragonite/prismatic</td>
<td>inner</td>
</tr>
<tr>
<td></td>
<td>inner- very thick bulk of shell</td>
<td>aragonite/ porcelaneous</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(homogeneous)</td>
<td></td>
</tr>
<tr>
<td>D. robustum</td>
<td>whole shell</td>
<td>aragonite/porcelaneous</td>
<td>whole shell</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(crossed lamellar)</td>
<td></td>
</tr>
<tr>
<td>E. complanata</td>
<td>outer- thin</td>
<td>aragonite/prismatic</td>
<td>inner</td>
</tr>
<tr>
<td></td>
<td>inner- thick, bulk of shell</td>
<td>aragonite/ nacreous</td>
<td></td>
</tr>
</tbody>
</table>

¹ Foliated (calcitic) mineral most often occurs as a mixture of the arrangements described by Carter (1980) and shown in column 1. These arrangements are not often differentiated in the literature. In the above species having foliated shell mineral, regularly foliated is the dominant arrangement. Porcelaneous (aragonitic) mineral arrangement types are indicated in ( ).
Table 4.2 TAXONOMY OF SELECTED BIVALVE SPECIES (Carter, 1990).

**SUBCLASS PTERIOMORPHIA**

**Order OSTREOIDA**

Superfamily Plicatuloidea
Family Plicatulidae

* Plicatula gibbosa (Cat's paw)

Superfamily Ostreoida
Family Ostreidae
Subfamily Crassostreinae

* Crassostrea virginica (Eastern oyster)
* Crassostrea gigas (Pacific oyster)

Subfamily Ostreinae

* Ostrea edulis (European oyster)

**Order PECTINOIDA**

Superfamily Pectinoidea
Family Pectinidae

* Argopecten irradians (Bay scallop)
* Placopecten magellanicus (Sea scallop)

Superfamily Anomioidea
Family Anomiidae
Subfamily Anomiinae

* Anomia simplex (Jingle shells)
Order **PTERIOIDA**
Suborder PTERIINA
Superfamily Pterioidea
Family Pteriidae

* Pinctada *fucata* (Pearl oyster)

Suborder PINNINA
Superfamily Pinnoidea
Family Pinnidae

* Atrina *rigida* (Stiff pen shell)

Order **MYTILOIDA**
Superfamily Mytiloidea
Family Mytilidae
Subfamily Mytilinae

* Mytilus *edulis* (Blue mussel)

SUBCLASS PALAEOHETERODONTA
Order UNIONOIDA
Family Unionidae

* Elliptio *complanata* (Eastern complanata)

SUBCLASS HETERODONTA
Order VENEROIDA
Superfamily Veneracea
Family Veneridae

* Mercenaria *mercenaria* (Quahog)

Superfamily Cardiacea
Family Cardiidae

* Dinocardium *robustum* (Giant heart cockle)
A caveat of this study is that the constructed phylogenies, based on either morphological or DNA similarities, do not often agree and even conflict among studies using similar methods of classification. However, as more DNA sequences become known, the congruence between studies using both classification methods is growing (Giribet and Wheeler, 2002). Another point to consider is that calcite microstructures
appeared more recently than aragonite structures. Consequently, species having foliated calcite microstructure are, in many cases, more closely related than species with aragonitic microstructures which span a significantly longer evolutionary period (Carter, 1980). Thus, a correlation between shell matrix composition and shell structure may not be visible against the background of phylogenetic relation; that is, matrix similarities between genera may result from being closely related and little to do with having a common shell microstructure or mineralogy. Even so, considering the wealth of studies demonstrating the effect of shell matrix proteins on mineralization, a comparative investigation of matrices from bivalves of various shell structure will provide a legitimate data set to be used in conjunction with DNA sequences of matrix proteins as they become available. Together, these data will further our understanding of the evolutionary relationships of shell proteins, their chemical characteristics, and how these characteristics may influence crystal size, shape, and orientation.

**Soluble Matrix Extraction and Quantification**

Bivalve specimens were shucked from their shells and were scrubbed with a wire brush under flowing tap water to remove sediment and foreign encrustations. The inner and outer shell layers of *P. fucata* and *A. rigida* were easily separated by striking the shells on a hard surface. In other cases, the outer or inner layers were abraded away with a small rotary tool leaving the remaining layer for study. Shell fragments from a single layer were ground to a fine powder following a few short (1 sec) bursts in a coffee
grinder. Twenty-five grams of shell was dissolved in 750 ml of 17% EDTA. The resulting suspensions were centrifuged at 27,000 x g for 30 min to separate water soluble matrix (SM) from insoluble matrix (IM). Soluble fractions containing SM were concentrated to approximately 50 ml and then dialyzed against 1000 ml of 10 mM NaCl using a Millipore Minitan tangential flow filtration apparatus with a molecular weight exclusion limit of 10 kD. The resultant dialysate was further dialyzed against distilled water and lyophilized. The insoluble pellets containing IM were lyophilized and weighed. SM was reconstituted in distilled water to a concentration of 5 mg ml⁻¹, based on total weight of the dried material. Additionally, protein was estimated chemically using bicinchoninic acid (BCA) (PIERCE). Because SM proteins generally lack the amino acids that react with Biuret chemistry, BCA estimates were at least two fold lower than dry weight measurements.

**Amino acid analysis**

Amino acid analysis was carried out by re-suspending 10 ug of dry weight protein in 6 N HCL containing 2% phenol. Hydrolysis was carried out at 110 °C under N₂ for 24 h. Following hydrolysis, samples were dried in a Speed Vac apparatus, re-suspended in 150 ul sample buffer, and analyzed on a Beckman 6300 Autoanalyzer using (post-column) ninhydrin detection and standard ion exchange resins.
Phosphate Analysis

Total phosphate was determined by method of Eisenreich et al. (1975) as modified by Swift (1985). Generally, the reaction results in the conversion of phosphorus to the orthophosphate form using a persulfate-sulfuric acid digestion. Blue color is produced by the formation of a reduced molydoantimonylphosphoric acid complex and is detected at 340 nm. Samples were tested against a standard solution with a dilution range of 0.07 ug/ml to 5.0 ug/ml total phosphate.

Carbohydrate Analysis

Carbohydrate content was analyzed by two methods. Weight estimates were obtained using the Glycoprotein Carbohydrate Estimation Kit (PIERCE). Generally, glycoprotein is oxidized with sodium meta-periodate. The aldehyde formed in this reaction reacts with a proprietary detection reagent to form a colored (purple) product. Absorbance is read at 550 nm and is proportional to the percentage of carbohydrate component in the protein. A standard curve is generated using proteins of known total carbohydrate amount so that percent carbohydrate for unknown samples is interpolated from this curve. Additionally, the DIG Glycan/Protein Double labeling kit (Boehringer Mannheim) was used to qualitatively immuno label protein and glycoproteins simultaneously from blots following separation on SDS-PAGE gels. This method uses sodium meta-periodate to oxidize sugar hydroxyl groups to form aldehydes. A hapten (DIG-Hydrazine) covalently interacts with the aldehydes which are then detected with an
anti-DIG antibody conjugated to horseradish peroxidase. Development is carried out in POD substrate. Although staining of protein and glycoprotein can be carried out simultaneously, it was carried out sequentially in order to maximize the sensitivity for identifying glycoproteins. Once blots identifying glycoproteins were photographed, the blots were stained to identify non-glycosylated proteins so that both protein types were visualized at the same time.

**Analytical gel electrophoresis of soluble matrix**

SM from each species was resolved on gradient (4 to 20%) or 12% Tris-Glycine ready gels (Novex) in the presence of sodium dodecyl sulfate (SDS). Protein was visualized using a variety of stains including, Coomassie Brilliant blue R-250, 1-ethyl-2-[3-(1-ethyl)naphtho[1,2-d]thiazolin-2-ylidene)-2-methylpropenyl]naphtho[1,2-d]thiazolium bromide (Stains-all) and silver nitrate. Gels were stained in 0.1% Coomassie blue in 10% acetic acid and 50% methanol. Stain-all staining was done according to methods described by Campbell et al. (1983) as modified by Myers et al. (1996). Silver staining was conducted by method of (Morrisey, 1981) and was additionally used as a secondary stain after Stains-all staining. This “double staining” technique was found to amplify the detection of anionic proteins and revealed minor constituents whose concentrations were below the sensitivity for Stains-all alone (Myers et al., 1996). This technique also had the added advantage of identifying proteins not reactive to Stains-all which appeared white.
against the light brown background so that most of the protein components could be visualized simultaneously.

**ELISA**

Selected proteins were dissolved in TBS, pH 7.5, to a concentration of 10 ug/ml. Fifty milliliters of each protein was applied to the wells of 96-well microtiter plates (Immulon) and allowed to adsorb overnight at 4 °C. Coat proteins were aspirated from the wells and the plate was washed gently with TBS, pH 7.5. Plates were blocked with 200 ul of 3% BSA and 1% goat serum in TBS, pH 7.5, for 3 h at room temperature. After blocking, plates were washed 1x with TBS, pH7.5. Rabbit anti-SM antibody was serially diluted in 1% BSA in TBS, pH 7.5 (ELISA buffer) and 100 ul of each dilution was applied to the wells and incubated for 1.5 h at 37 °C. Plates were washed three times with TBS containing 0.05% Tween-20 (wash buffer) followed by the application of 100 µl of goat anti-rabbit antibody conjugated to alkaline phosphatase (Sigma) at a working dilution of 1:30,000 in ELISA buffer. Color was developed using pNPP substrate (Sigma) and optical densities were determined after 30 min at the 405 nm wavelength.

**Western Blot**

SM from each species was resolved on 4 to 20% poly-acrylamide gradient gels (PAGE) in the presence of sodium dodecyl sulfate (SDS), and then transferred onto
nitrocellulose. The transfer was conducted for 1 h at 100 V in buffer containing 25 mM Tris, 192 mM glycine, 20% v/v methanol, pH 8.3, using a mini-transfer unit (Invitrogen). Following transfer, membranes were rinsed once with TBS, pH 7.5, then placed into 50 ml 3% BSA and 1% goat serum in TBS (blocking buffer) and allowed to incubate overnight at 4 °C. After blocking, membranes were incubated in 50 ml of anti-SM oyster antibody diluted 1:1000 in 3% BSA in TBS (western buffer) for 2 h. Membranes were washed three times in 100 ml of 1% BSA and 0.05% Tween-20 in TBS (wash buffer) for 5 min each. The membrane was incubated in 50 ml of alkaline-phosphatase conjugated goat anti-rabbit IgG (Sigma) diluted to 1:30,000 in western buffer for 1.5 h, and washed three times for 5 min each in wash buffer. A final rinse was carried out in TBS and color was developed with BCIP/NBT (Sigma). Development was stopped by repeated washing in distilled water over 10 min. All incubations and washes were carried out with continuous gentle shaking at room temperature.

Results

Protein extraction and quantification

Matrix protein was extracted from 25 g of dry weight shell. Protein yield improved significantly when hand (course) ground shell fragments were ground further to a fine powder, following a few short (1 s) bursts in a coffee grinder. This was especially the case for aragonitic shell as this mineral type was most difficult to dissolve. The
amount of SM extracted from the shells was variable from prep to prep, especially among aragonite shells. For this reason, a definitive SM weight/shell weight is not reported here. Still, some consistent trends were observed. Approximately 50% more SM is recovered from foliated (calcite) shell than either nacre or porcelaneous shell (aragonite). Conversely, IM recoveries from aragonitic shell are upwards of 50% greater than those recovered from calcitic shells. This indicates that aragonite matrices have more protein associations that are not water soluble and are fundamentally different from calcite derived SM.

Differences in sensitivity to BCA are also an indicator of chemical differences between the matrices. Protein estimation by this method relies mainly on the presence of cysteine, cystine, tryptophan, or tyrosine and the availability to these groups. A comparison of BCA weights to dry weights of total SM is reported in Table 4.3. At most, 50% of the dry weight SM is sensitive to BCA chemistry. Generally, matrix derived from foliated shell is more sensitive to BCA chemistry compared to aragonitic mineralogies with the exception of the scallop species. *M. mercenaria* is also an exception as nearly 50% of dry weight protein was BCA sensitive.
Table 4.3  Percent BCA Weight to Dry Weight of SM from Selected Species

<table>
<thead>
<tr>
<th>Species</th>
<th>BCA weight</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. virginica</em></td>
<td>38.3</td>
</tr>
<tr>
<td><em>C. gigas</em></td>
<td>35.4</td>
</tr>
<tr>
<td><em>O. edulis</em></td>
<td>37.2</td>
</tr>
<tr>
<td><em>P. gibosa</em></td>
<td>31.7</td>
</tr>
<tr>
<td><em>A. simplex</em></td>
<td>44.0</td>
</tr>
<tr>
<td><em>A. irradians</em></td>
<td>18.8</td>
</tr>
<tr>
<td><em>P. magellanicus</em></td>
<td>15.8</td>
</tr>
<tr>
<td><em>P. fucata</em></td>
<td>21.6</td>
</tr>
<tr>
<td><em>A. rigida</em></td>
<td>8.7</td>
</tr>
<tr>
<td><em>M. edulis</em></td>
<td>11.2</td>
</tr>
<tr>
<td><em>E. complanata</em></td>
<td>19.1</td>
</tr>
<tr>
<td><em>M. mercenaria</em></td>
<td>48.1</td>
</tr>
<tr>
<td><em>D. robustum</em></td>
<td>19.2</td>
</tr>
</tbody>
</table>

1 Species are grouped according to shell microstructure and mineralogy
SM from the first group comes from foliated calcite
SM from the middle group comes from nacre aragonite
SM from the last group comes from porcelaneous aragonite

2 Values are percents of BCA weights to dry weights (ug)

A noteworthy observation is that following extraction and dialysis, SM appeared
to re-associate in solution, forming precipitates that were resistant to dissolution. All of
the matrices tested exhibited re-association to some degree; however, it was especially
noticeable in species having an aragonite shell. This phenomenon was most pronounced
in species with pigmented SM including *P. gibosa, D. robustum,* and *P. fucata.*
Although “salting out” of proteins is a common occurrence, this case is interesting because SM proteins should be exceptionally water soluble given their hydrophilic and highly anionic compositions (reported below). Precipitates that were treated with urea and resolved on gels resulted in identical banding patterns as SM.

The occurrence of SM precipitates is also interesting in light of the shell formation process. Some portion of the matrix assembles into polymers presumably before or shortly after it is secreted into the extra-pallial cavity (Bevelander and Nakahara, 1980; Myers et al., 2007). Whether the observed SM precipitates are artifact, or reflect an association that occurs in vivo is unclear. In any case, re-association is a product of the unique chemistries of the matrix components. This observation, considered with the differences in the amounts of SM and IM recovered and their sensitivity to BCA, demonstrate the unique chemistries of calcite and aragonite shell matrix.

**Amino Acid analysis**

Amino acid analysis of SM from each species is reported in Table 4.4. At the surface, the compositions are strikingly similar, regardless of shell structure or genetic relation. The most outstanding feature is the high aspartic acid content which range from 20 mole% in *M. mercenaria* to over 50 mole% in *P. gibbosa* and *D. robustum*. Also
prominent are glycine, serine, and glutamine. This data indicates that all matrices are acidic which is typical of many mineral derived proteins (Lowenstam and Weiner, 1989).
Table 4.4 Amino Acid Compositions of Soluble Matrix (SM) From Selected Bivalve Species

<table>
<thead>
<tr>
<th>Genus Species</th>
<th>C. virginica</th>
<th>C. gigas</th>
<th>O. edulis</th>
<th>P. gibbosa</th>
<th>A. simplex</th>
<th>A. irradians</th>
<th>P. magellanicus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mineralogy/ Microstructure</td>
<td>CF</td>
<td>CF</td>
<td>CF</td>
<td>CF</td>
<td>CF</td>
<td>CF</td>
<td>CF</td>
</tr>
<tr>
<td>Amino acid2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASP</td>
<td>35.77</td>
<td>32.16</td>
<td>35.57</td>
<td>52.88</td>
<td>18.84</td>
<td>39.05</td>
<td>33.04</td>
</tr>
<tr>
<td>THR</td>
<td>1.51</td>
<td>2.13</td>
<td>1.37</td>
<td>1.67</td>
<td>0.86</td>
<td>0.60</td>
<td>0.55</td>
</tr>
<tr>
<td>SER</td>
<td>18.61</td>
<td>24.77</td>
<td>10.55</td>
<td>7.08</td>
<td>26.70</td>
<td>23.20</td>
<td>27.58</td>
</tr>
<tr>
<td>GLU</td>
<td>6.30</td>
<td>6.17</td>
<td>11.56</td>
<td>3.39</td>
<td>4.00</td>
<td>4.35</td>
<td>3.84</td>
</tr>
<tr>
<td>PRO</td>
<td>2.39</td>
<td>2.34</td>
<td>3.35</td>
<td>1.04</td>
<td>0.46</td>
<td>1.23</td>
<td>0.75</td>
</tr>
<tr>
<td>GLY</td>
<td>23.23</td>
<td>20.19</td>
<td>15.52</td>
<td>15.05</td>
<td>35.43</td>
<td>21.03</td>
<td>22.37</td>
</tr>
<tr>
<td>ALA</td>
<td>1.53</td>
<td>0.51</td>
<td>1.49</td>
<td>2.14</td>
<td>4.12</td>
<td>4.84</td>
<td>4.24</td>
</tr>
<tr>
<td>VAL</td>
<td>1.13</td>
<td>1.15</td>
<td>1.05</td>
<td>1.03</td>
<td>1.35</td>
<td>0.56</td>
<td>0.48</td>
</tr>
<tr>
<td>MET</td>
<td>0.67</td>
<td>0.45</td>
<td>0.25</td>
<td>0.54</td>
<td>0.37</td>
<td>0.23</td>
<td>0.24</td>
</tr>
<tr>
<td>ILE</td>
<td>0.87</td>
<td>0.87</td>
<td>1.94</td>
<td>0.65</td>
<td>0.50</td>
<td>0.25</td>
<td>0.46</td>
</tr>
<tr>
<td>LEU</td>
<td>1.17</td>
<td>1.12</td>
<td>1.13</td>
<td>1.04</td>
<td>0.74</td>
<td>1.50</td>
<td>0.72</td>
</tr>
<tr>
<td>TYR</td>
<td>2.99</td>
<td>2.52</td>
<td>2.74</td>
<td>0.44</td>
<td>2.41</td>
<td>0.28</td>
<td>0.30</td>
</tr>
<tr>
<td>PHE</td>
<td>0.57</td>
<td>0.66</td>
<td>0.40</td>
<td>1.79</td>
<td>0.45</td>
<td>0.33</td>
<td>0.24</td>
</tr>
<tr>
<td>HIS</td>
<td>0.51</td>
<td>0.58</td>
<td>3.64</td>
<td>0.23</td>
<td>0.19</td>
<td>0.30</td>
<td>0.64</td>
</tr>
<tr>
<td>LYS</td>
<td>1.45</td>
<td>1.92</td>
<td>7.03</td>
<td>5.61</td>
<td>2.04</td>
<td>1.84</td>
<td>2.94</td>
</tr>
<tr>
<td>ARG</td>
<td>1.08</td>
<td>1.35</td>
<td>1.99</td>
<td>0.61</td>
<td>1.50</td>
<td>0.38</td>
<td>1.59</td>
</tr>
<tr>
<td>% PP3</td>
<td>77</td>
<td>70</td>
<td>45</td>
<td>13</td>
<td>70</td>
<td>76</td>
<td>77</td>
</tr>
<tr>
<td>C/HP ratio4</td>
<td>5.41</td>
<td>5.94</td>
<td>6.23</td>
<td>7.64</td>
<td>3.33</td>
<td>5.15</td>
<td>5.92</td>
</tr>
<tr>
<td>Hydrophobicity5</td>
<td>0.12</td>
<td>0.10</td>
<td>0.13</td>
<td>0.11</td>
<td>0.14</td>
<td>0.13</td>
<td>0.10</td>
</tr>
</tbody>
</table>

1 Categories defined as: CF= calcite foliated
   AL= aragonite lamellar
   AN= aragonite nacre
   AP= aragonite porcelaneous
   AH= aragonite homogenous

2 Values are reported as mole percent and are averages, where n=2

3 Estimation of the percent phosphorylation occurring on Ser and Thr residues calculated as (% Phosphate by weight) / (mole% Ser +Thr) x 100. This estimate does not account for phosphate from other organic matrix components such as phospholipids.

4 Charged residues (Asp, Glu, Lys, Arg, His) / hydrophobic residues (Pro, Ala, Val, Met, Ile, Leu, Phe)

5 Hydrophobicity is the sum of hydrophobic residues/hydrophilic (Lys, Arg, His, Asp, Glu, Ser, Thr, Tyr)
(Continued on next page)
Table 4.4 (Continued)

<table>
<thead>
<tr>
<th>Genus</th>
<th>Species</th>
<th>Mineralogy/Microstructure&lt;sup&gt;1&lt;/sup&gt;</th>
<th>P. fucata</th>
<th>A. rigida</th>
<th>M. edulis</th>
<th>M. mercenaria</th>
<th>D. robustum</th>
<th>E. complanata</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AN</td>
<td>AN</td>
<td>AN</td>
<td>AH</td>
<td>ACL</td>
<td>AN</td>
<td></td>
</tr>
<tr>
<td>Amino acid&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASP</td>
<td></td>
<td>40.80</td>
<td>16.70</td>
<td>19.92</td>
<td>20.21</td>
<td>53.54</td>
<td>43.33</td>
<td></td>
</tr>
<tr>
<td>THR</td>
<td></td>
<td>2.98</td>
<td>4.58</td>
<td>5.26</td>
<td>5.72</td>
<td>2.84</td>
<td>2.32</td>
<td></td>
</tr>
<tr>
<td>SER</td>
<td></td>
<td>7.96</td>
<td>5.92</td>
<td>8.91</td>
<td>8.09</td>
<td>3.32</td>
<td>6.69</td>
<td></td>
</tr>
<tr>
<td>GLU</td>
<td></td>
<td>7.56</td>
<td>7.10</td>
<td>8.22</td>
<td>7.20</td>
<td>5.11</td>
<td>3.96</td>
<td></td>
</tr>
<tr>
<td>PRO</td>
<td></td>
<td>4.25</td>
<td>5.99</td>
<td>6.96</td>
<td>11.83</td>
<td>6.86</td>
<td>4.07</td>
<td></td>
</tr>
<tr>
<td>GLY</td>
<td></td>
<td>5.91</td>
<td>13.89</td>
<td>13.04</td>
<td>12.69</td>
<td>6.41</td>
<td>14.39</td>
<td></td>
</tr>
<tr>
<td>ALA</td>
<td></td>
<td>5.59</td>
<td>8.91</td>
<td>5.13</td>
<td>7.92</td>
<td>3.27</td>
<td>4.03</td>
<td></td>
</tr>
<tr>
<td>VAL</td>
<td></td>
<td>2.70</td>
<td>3.25</td>
<td>4.12</td>
<td>3.12</td>
<td>2.47</td>
<td>2.25</td>
<td></td>
</tr>
<tr>
<td>MET</td>
<td></td>
<td>1.92</td>
<td>1.42</td>
<td>1.10</td>
<td>0.93</td>
<td>2.00</td>
<td>2.76</td>
<td></td>
</tr>
<tr>
<td>ILE</td>
<td></td>
<td>2.90</td>
<td>3.12</td>
<td>3.56</td>
<td>2.07</td>
<td>2.15</td>
<td>2.55</td>
<td></td>
</tr>
<tr>
<td>LEU</td>
<td></td>
<td>2.96</td>
<td>5.69</td>
<td>4.38</td>
<td>3.00</td>
<td>2.80</td>
<td>2.92</td>
<td></td>
</tr>
<tr>
<td>TYR</td>
<td></td>
<td>2.74</td>
<td>3.67</td>
<td>2.70</td>
<td>3.70</td>
<td>1.82</td>
<td>1.87</td>
<td></td>
</tr>
<tr>
<td>PHE</td>
<td></td>
<td>1.79</td>
<td>3.99</td>
<td>2.11</td>
<td>1.79</td>
<td>1.40</td>
<td>1.63</td>
<td></td>
</tr>
<tr>
<td>HIS</td>
<td></td>
<td>1.20</td>
<td>2.54</td>
<td>1.46</td>
<td>1.34</td>
<td>0.48</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>LYS</td>
<td></td>
<td>5.51</td>
<td>8.37</td>
<td>7.48</td>
<td>7.18</td>
<td>3.32</td>
<td>4.81</td>
<td></td>
</tr>
<tr>
<td>ARG</td>
<td></td>
<td>3.23</td>
<td>4.85</td>
<td>5.61</td>
<td>3.05</td>
<td>1.81</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td>% PP&lt;sup&gt;3&lt;/sup&gt;</td>
<td></td>
<td>19</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>C/HP ratio&lt;sup&gt;4&lt;/sup&gt;</td>
<td></td>
<td>2.64</td>
<td>1.22</td>
<td>1.56</td>
<td>1.27</td>
<td>3.07</td>
<td>2.70</td>
<td></td>
</tr>
<tr>
<td>Hydrophobicity&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
<td>0.31</td>
<td>0.60</td>
<td>0.46</td>
<td>0.54</td>
<td>0.29</td>
<td>0.31</td>
<td></td>
</tr>
</tbody>
</table>

<sup>1</sup> Categories defined as: CF= calcite foliated
AL= aragonite lamellar
AN= aragonite nacre
AP= aragonite porcelaneous
AH= aragonite homogenous

<sup>2</sup> Values are reported as mole percent and are averages, where n=2

<sup>3</sup> Estimation of the percent phosphorylation occurring on Ser and Thr residues calculated as (% Phosphate by weight) / (mole% Ser + Thr) x 100. This estimate does not account for phosphate from other organic matrix components such as phospholipids.

<sup>4</sup> Charged residues (Asp, Glu, Lys, Arg, His) / hydrophobic residues (Pro, Ala, Val, Met, Ile, Leu, Phe)

<sup>5</sup> Hydrophobicity is the sum of hydrophobic residues/hydrophilic (Lys, Arg, His, Asp, Glu, Ser, Thr, Tyr)
A deeper examination of the data reveals correlations between composition, mineralogy, and microstructure. Compared to calcite, aragonite derived matrix has a significantly higher amount of hydrophobic groups especially proline, alanine, leucine, and valine which is reflected in the C/HP and hydrophobicity values (Table 4.4). While all species contain high amounts of serine and glycine, these amino acids are especially prevalent in species having foliated/calcite microstructure. Compositions determined from different shell layers of two species, *P. fucata* and *A. rigida* (Table 4.5), show that whole shell extracts, which are predominantly derived from outer prismatic (calcite) shell layers, have serine and glycine levels more comparable to aragonite shell. This is further evidence that a prevalence of Ser and Gly is a hallmark of foliated calcite. Prismatic calcite has a higher C/HP ratio than the nacre layer but generally lower than those of foliated mineral. Because the charged groups determined are predominately Asp and Glu, the C/HP ratios indicate that calcite foliated matrix has the highest acidic character, followed by prismatic calcite and aragonite microstructures.
Table 4.5  Comparison of Amino Acid Compositions Between Whole Shell and Nacreous Shell Layers of Two Bivalves

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>P. fucata whole shell</th>
<th>P. fucata nacre only</th>
<th>A. rigida whole shell</th>
<th>A. rigida nacre only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>40.44</td>
<td>40.80</td>
<td>43.23</td>
<td>16.70</td>
</tr>
<tr>
<td>Thr</td>
<td>1.83</td>
<td>2.98</td>
<td>1.94</td>
<td>4.58</td>
</tr>
<tr>
<td>Ser</td>
<td>8.84</td>
<td>7.96</td>
<td>5.77</td>
<td>5.92</td>
</tr>
<tr>
<td>Glu</td>
<td>4.67</td>
<td>7.56</td>
<td>22.93</td>
<td>7.10</td>
</tr>
<tr>
<td>Pro</td>
<td>1.38</td>
<td>4.25</td>
<td>1.01</td>
<td>5.99</td>
</tr>
<tr>
<td>Gly</td>
<td>23.18</td>
<td>5.91</td>
<td>4.25</td>
<td>13.89</td>
</tr>
<tr>
<td>Ala</td>
<td>5.79</td>
<td>5.59</td>
<td>13.53</td>
<td>8.91</td>
</tr>
<tr>
<td>Val</td>
<td>2.32</td>
<td>2.70</td>
<td>2.38</td>
<td>3.25</td>
</tr>
<tr>
<td>Met</td>
<td>0.61</td>
<td>1.92</td>
<td>0.37</td>
<td>1.42</td>
</tr>
<tr>
<td>Ile</td>
<td>1.05</td>
<td>2.90</td>
<td>0.18</td>
<td>3.12</td>
</tr>
<tr>
<td>Leu</td>
<td>1.92</td>
<td>2.96</td>
<td>1.76</td>
<td>5.69</td>
</tr>
<tr>
<td>Tyr</td>
<td>1.35</td>
<td>2.74</td>
<td>0.59</td>
<td>3.67</td>
</tr>
<tr>
<td>Phe</td>
<td>1.09</td>
<td>1.79</td>
<td>0.43</td>
<td>3.99</td>
</tr>
<tr>
<td>His</td>
<td>1.24</td>
<td>1.20</td>
<td>0.29</td>
<td>2.54</td>
</tr>
<tr>
<td>Lys</td>
<td>2.58</td>
<td>5.51</td>
<td>0.99</td>
<td>8.37</td>
</tr>
<tr>
<td>Arg</td>
<td>1.66</td>
<td>3.23</td>
<td>0.33</td>
<td>4.85</td>
</tr>
</tbody>
</table>

C/HP ratio | 3.58 | 2.64 | 3.45 | 1.22 |
Hydrophobicity | 0.23 | 0.31 | 0.26 | 0.60 |

1 Values reported as mole percent
2 Whole shell includes the inner nacreous layer and the outer calcitic layer

Rose plots (Figure 4.4) of SM compositions graphically illustrate a correlation between matrix composition and shell mineralogy. In these plots, the most prevalent amino acids were considered individually (Ser, Gly, Asp, Glu), and the remaining in groups based on their chemical character. Plot shapes of calcitic shell structure (Figure
4.4, A) generally have thin “right” handed shapes, due to the dominance of glycine and serine compared to aragonite group (Figure 4.4, B), which resulted in “left” handed spikes due to the higher amounts of hydrophobic groups and the presence of charged/basic groups. Note too, that the whole shell matrix plot for *P. fucata* results in a “right” handed spike similar to those of foliated calcite group and indicates that matrices from foliated and prismatic structures are similar.
Figure 4.4  Amino acid compositions represented as rose diagrams of all species.
Amino acids were plotted in groups according to charge, hydrophobicity and the most prevalent single amino acids including, Asp, Glu, Ser and Gly. The “uncharged” group includes Thr and Tyr, the “charged/basic” group includes Lys, Arg and His, and the “hydrophobic” group includes Ala, Val, Leu, Ile, Met, Pro and Phe. Except for *O. edulis*, all calcitic microstructures resulted in “right-handed” shapes and aragonitic microstructures resulted in “left-handed” shapes. “Whole shell” plots of *P. fucata* and *A. Rigida* include compositions from both the calcitic/prismatic and aragonitic/nacreous layers of these species. Note the influence of the calcitic shell layers on the rose diagram shapes. Matrix from the calcitic shell layer of *P. fucata* results in a plot that resembles foliated calcite, while matrix from pure nacre resembles plot shapes similar to those of aragonite.

These plots do not resolve any clear correlation between amino acid composition and shell microstructure within the aragonite group. As examples, *A rigida*, *M. edulis* and *M. mercenaria* have strikingly similar and relatively broad shapes even though the later species has a porcelaneous microstructure. *P. fucata*, and *E. complanata* have nacre microstructures but their plot shapes closely resemble *D. robustum* which has porcelaneous.

On the other hand, the general plot shapes for the foliated group are more similar suggesting that their matrices are also similar, especially between closely related species such as Crassostrea. The nearly identical plot shapes that result among the scallops and the Crassostrea oysters suggest that shell matrix is conserved within these families. *O. edulis* is an exception and has a shape with characteristics of aragonite and calcite plot shapes. This observation may represent a divergence from the shell matrix proteins from Ostreidae, especially when considered with immuno-reactivity results discussed later.
By plotting the full spectrum of amino acids (Figure 4.5), the degree of phylogenic relation between species becomes considerably more apparent. All three oyster species have similar plot shapes that are distinct from the other calcite foliated species. Likewise, A. simplex resembles the shapes constructed from compositions of the scallop species; all three of these species share the Order Pectinoida. Even E. complanata, despite having nacre/aragonite shell, shows some shape similarity with the oyster species which agrees with their closer molecular phylogenetic relationship than say M. mercenaria (refer to Figure 4.3), and suggests that some matrix proteins are genetically conserved between calcite and aragonite microstructures. The resemblance of the plot shapes between P. fucata (whole shell which is mostly prismatic calcite) and oysters (folia), and the similarity between M. edulis (nacre) and M. mercenaria (porcelaneous) supports that some shell proteins are conserved between different microstructures of identical mineralogy. As a side note, these results demonstrate that using rose plot graphs to compare shell SM is a viable method to evaluate phylogenetic relationships with respect to shell proteins.
Figure 4.5  Rose diagrams of all represented amino acids for selected species.
Values were plotted on a logarithmic scale to offset the dominance of Asp, Ser, Gly, and Glu.

As a whole, rose plots illustrate strong matrix similarity between closely related individuals and within shells of identical mineralogy, but less evident is a correlation between matrix composition and microstructure, especially among species with aragonite mineralogy. These findings support that a particular matrix composition does influence the selection of mineralogy but de-emphasizes its influence in promoting the formation of a specific microstructure. Nonetheless, the plot shapes of foliated matrix are similar enough to leave open the possibility that a conserved set of proteins promotes this microstructure.

Amino acid compositions were also compared using the similarity index (Cornish Bowden 1983) (Figure 4.6, A). Although the criteria for assessing similarity between two compositions are too stringent for comparing shell matrix extracts, the values do reveal correlations between matrix, shell structure, and phylogeny similar to those deduced from rose plots. Generally, species with foliated shell are more related to C. virginica (Figure 4.6, B). After C. gigas, the scallop species are within 0.6 units of C. virginica, placing them ahead of more genetically related species including, O. edulis, and P. gibbosa. This supports that some matrix components are conserved in foliated microstructures. However, the appearance of E. complanata and P. fucata in the middle of this group, which are relatively close to Crassostrea (Adamkewicz et al., 1997), shows
the influence of phylogeny in matrix composition and de-emphasizes the importance of a specific amino acid structure in determining a particular shell structure.
<table>
<thead>
<tr>
<th></th>
<th>C. virginica</th>
<th>C. gigas</th>
<th>O. edulis</th>
<th>P. gibbosa</th>
<th>A. simplex</th>
<th>A. irradians</th>
<th>A. magellanicus</th>
<th>P. fucata</th>
<th>A. rigida</th>
<th>M. edulis</th>
<th>M. mercenaria</th>
<th>D. robustum</th>
<th>E. complanata</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. virginica</td>
<td>0.31</td>
<td>0.98</td>
<td>2.64</td>
<td>2.59</td>
<td>0.31</td>
<td>0.58</td>
<td>2.56</td>
<td>3.78</td>
<td>2.94</td>
<td>3.30</td>
<td>4.36</td>
<td>1.59</td>
<td></td>
</tr>
<tr>
<td>C. gigas</td>
<td>1.52</td>
<td>4.00</td>
<td>2.18</td>
<td>0.42</td>
<td>0.22</td>
<td>3.24</td>
<td>3.96</td>
<td>2.94</td>
<td>3.40</td>
<td>5.74</td>
<td>2.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O. edulis</td>
<td>2.04</td>
<td>5.25</td>
<td>1.60</td>
<td>2.26</td>
<td>1.06</td>
<td>2.52</td>
<td>1.75</td>
<td>2.09</td>
<td>2.75</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. gibbosa</td>
<td>9.92</td>
<td>2.57</td>
<td>4.42</td>
<td>1.26</td>
<td>7.05</td>
<td>6.19</td>
<td>6.43</td>
<td>0.74</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. simplex</td>
<td>3.18</td>
<td>1.90</td>
<td>9.10</td>
<td>5.14</td>
<td>4.91</td>
<td>5.44</td>
<td>13.27</td>
<td>7.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. irradians</td>
<td>0.31</td>
<td>2.65</td>
<td>4.88</td>
<td>4.03</td>
<td>4.36</td>
<td>4.39</td>
<td>1.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. magellanicus</td>
<td>3.96</td>
<td>4.66</td>
<td>3.86</td>
<td>4.31</td>
<td>6.63</td>
<td>3.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. fucata</td>
<td>3.72</td>
<td>3.14</td>
<td>3.50</td>
<td>1.09</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. rigida</td>
<td>0.48</td>
<td>0.92</td>
<td>7.63</td>
<td>3.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. edulis</td>
<td>0.23</td>
<td>6.33</td>
<td>3.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. mercenaria</td>
<td>6.27</td>
<td>3.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. robustum</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SAΔN <0.42=95% related
0.42< SΔN<0.93= weakly related
SΔN >0.93= unrelated
Collectively, analyses of amino acid compositions show no definitive correlation between matrix composition and shell microstructure in the aragonite group and a correlation was not easily evident against the backdrop of phylogenetic relation in the calcitic group. However, a few general traits do appear to correlate with shell structure. Aragonite shell matrix has more hydrophobic amino acids than calcite and is reflected in their lower C/HP values. Foliated calcite matrix has the highest content of serine and glycine compared to other microstructures, including prismatic calcite. A more detailed look at matrix components and their chemical characteristics was needed to evaluate the possibility of a true correlation between matrix and microstructure.

Phosphate and Carbohydrate Analyses

Total phosphate and carbohydrate values were determined for specific shell layers from each species and are summarized in Table 4.6. By comparing total phosphate values graphically (Figure 4.7), a stark contrast between shell structures is evident. Foliated calcite shell has as much as 100 times the amount of phosphate compared to nacre or porcelaneous aragonite shell. Moreover, the extent to which serine/threonine groups are phosphorylated is estimated to be upwards of 70%, a substantial difference to that estimated for aragonite microstructures which range from 2 to 19% (Table 4.6).
edulis is an exception and has a relatively low amount of total phosphate compared to other species with foliated shell. Still, it is estimated that nearly 50% of its serine/threonine residues are phosphorylated. The low phosphate amount determined for P. gibbosa correlates with the low Ser/Thr determined for this species (footnote).

Admittedly, degrees of phosphorylation by this calculation cannot be assumed for a specific SM component. However, when considered with SDS-PAGE following Stains-all staining, some idea of which components are more highly phosphorylated can be deduced. This is discussed further in the following section.
Table 4.6 Chemical Analyses of SM Protein Fractions Relative to Shell Microstructure and Mineralogy from Selected Bivalve Species

<table>
<thead>
<tr>
<th>Species</th>
<th>Shell layer</th>
<th>Mineralogy/ Microstructure</th>
<th>Phosphate $^1$</th>
<th>%PP $^2$</th>
<th>Carbohydrate $^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. virginica</em></td>
<td>inner</td>
<td>calcite/foliated</td>
<td>15.47±0.58</td>
<td>77</td>
<td>2.93</td>
</tr>
<tr>
<td><em>C. gigas</em></td>
<td>inner</td>
<td>calcite/foliated</td>
<td>16.43±0.32</td>
<td>70</td>
<td>2.54</td>
</tr>
<tr>
<td><em>O. edulis</em></td>
<td>inner</td>
<td>calcite/foliated</td>
<td>5.4±0.17</td>
<td>45</td>
<td>2.61</td>
</tr>
<tr>
<td><em>P. gibbosa</em></td>
<td>outer</td>
<td>calcite/foliated</td>
<td>1.13±0.21</td>
<td>13</td>
<td>2.72</td>
</tr>
<tr>
<td><em>A. simplex</em></td>
<td>outer</td>
<td>calcite/foliated</td>
<td>19.27±0.58</td>
<td>70</td>
<td>2.70</td>
</tr>
<tr>
<td><em>A. irradians</em></td>
<td>inner</td>
<td>calcite/foliated</td>
<td>18.00±0.10</td>
<td>76</td>
<td>2.85</td>
</tr>
<tr>
<td><em>P. magellanicus</em></td>
<td>whole shell</td>
<td>calcite/foliated</td>
<td>22.40±0.30</td>
<td>77</td>
<td>1.75</td>
</tr>
<tr>
<td><em>P. fucata</em></td>
<td>outer</td>
<td>calcite/prismatic aragonite/ nacreous</td>
<td>*3.60</td>
<td>*1.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>inner</td>
<td></td>
<td>2.11</td>
<td>19</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>whole shell</td>
<td></td>
<td>5.71±0.07</td>
<td></td>
<td>2.79</td>
</tr>
<tr>
<td><em>A. rigida</em></td>
<td>outer</td>
<td>calcite/prismatic aragonite/ nacreous</td>
<td>*1.53</td>
<td>*1.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>inner</td>
<td></td>
<td>0.92</td>
<td>9</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>whole shell</td>
<td></td>
<td>2.45±0.12</td>
<td></td>
<td>2.37</td>
</tr>
<tr>
<td><em>M. edulis</em></td>
<td>inner</td>
<td>aragonite/ nacreous</td>
<td>0.47±0.12</td>
<td>3</td>
<td>1.41</td>
</tr>
<tr>
<td><em>M. mercenaria</em></td>
<td>inner</td>
<td>aragonite/porcelaneous (homogeneous)</td>
<td>0.26±0.02</td>
<td>2</td>
<td>1.46</td>
</tr>
<tr>
<td><em>D. robustum</em></td>
<td>whole shell</td>
<td>aragonite/porcelaneous (crossed lamellar)</td>
<td>0.33±0.04</td>
<td>5</td>
<td>1.41</td>
</tr>
<tr>
<td><em>E. complanata</em></td>
<td>inner</td>
<td>aragonite/ nacreous</td>
<td>0.38±0.06</td>
<td>4</td>
<td>1.54</td>
</tr>
</tbody>
</table>

$^1$ Total phosphate was determined according to method in Swift, 1992. For each sample total phosphate was determined from a standard curve and reported as percent weight of phosphate to dry weight of SM protein. Values and standard deviations are averages where N=3. Values not reported with standard of deviations are averages where N=2.

$^2$ Estimation of the percent phosphorylation occurring on Ser or Thr residues reported in Table 4.4. Values are reported as: [% total phosphate / (mole% Ser + mole% Thr)] 100.

$^3$ Carbohydrate was determined using the Glycoprotein Carbohydrate Estimation Kit (Pierce). Percent carbohydrate for SM samples was estimated from a standard curve generated from a glycoprotein of known carbohydrate content. Values are reported as percent weight of total carbohydrate to dry weight of protein and are averages where N=2

*Values marked as such are theoretical values and are the differences between the empirical values determined for whole shell and nacre shell layers.
Figure 4.7 Bar plot of total phosphate content determined for SM of selected bivalves. Values are reported as percent weight of total phosphate to dry weight protein. Species are arranged by microstructure and mineralogy and to a degree by phylogenetic relation to *C. virginica*, according to Carter (1990).

The nacre of *P. fucata* and to a lesser extent that of *A. rigida* has slightly more phosphate compared to other aragonitic microstructures. Even so, when phosphate values are considered with the amounts of serine/threonine reported for these species (Table 4.4), only 19% is estimated to be phosphorylated in *P. fucata*, and even less (9%) for *A. rigida*. Thus, the degree of phosphorylation of matrix derived from aragonite microstructures is lower than that of foliated calcite.
A theoretical phosphate value was determined for the outer prismatic calcite layers of both species by subtracting the empirically determined values of nacre from those determined for whole shell (Table 4.6). By this extrapolation, prismatic shell has more phosphate than nacre. This estimate is offered with caution because the whole shell value reflects a mixture of the two matrices. However, evidence supporting this conclusion is that whole shell matrix stains more intensely with Stains-all following SDS-PAGE than nacre in both species (discussed further below). Still, on average, foliated calcite has about ten times more phosphate than prismatic calcite. Therefore, matrix from folia is enriched in highly phosphorylated proteins compared to prismatic calcite and aragonite shell structures.

A clear, but subtle difference is evident in the amounts of carbohydrate estimated for SM from each species. Those with foliated shells have slightly more total carbohydrate than those with nacre or porcelaneous aragonite shells (Figure 4.8). Carbohydrate content was estimated to be 2.79 % for the whole shell of *P. fucata*, and 1.40 % for the inner nacreous layer. Therefore, it can be extrapolated (with caution) that the prismatic calcite layer contains about 1.39 % total carbohydrate; a value comparable to aragonite microstructures. A similar conclusion can be drawn from values estimated for *A. rigida*. This suggests that either the extent of glycosylation or amounts of glycoprotein is most prevalent in the foliated microstructure.
Figure 4.8  Bar plot of carbohydrate content estimated for SM from selected bivalves. Species are arranged by microstructure and mineralogy and to a degree, by phylogenetic relation to *C. virginica* according to Carter (1990).

**SDS-PAGE analysis**

Using a variety of staining methods, component protein patterns of each species were compared. Stains-all, a cationic dye, has been shown to specifically stain phosphoproteins and other anionic proteins dark blue to purple, while most other proteins do not stain, or stain red but fade quickly upon exposure to light (Campbell et al., 1983). Coomassie blue and silver nitrate stain a broad range of proteins. Coomassie blue stains in part through hydrophobic interactions; consequently this stain does not react well with
highly acidic proteins such as those that occur in oyster shell (Rusenko et al., 1991; Myers, 1999). By using a combination of stains, a more complete picture of the protein components that comprise each matrix emerged as well as a general sense of their chemical character.

Whole matrix protein extracts from each species were resolved on duplicate gels and stained with either Stains-all, Coomassie brilliant blue or silver nitrate in an effort to distinguish phosphoproteins and other anionic proteins from those with more basic character (Figure 4.9). For this analysis, lane assignments were arranged according to microstructure/mineralogy and loosely to phylogenetic relation. Figure 4.9, Lanes 1 to 7 resolve shell matrices of species from the subclass Pteriomorphia, whose shells are composed mostly of foliated/calcite and include the oysters (C. virginica, C. gigas and O. edulis), members of the order Ostreoida, and members of the order Pectinoida, which include scallops (A. iradians; P. magellanicus), cat’s paw (P. gibbosa) and the jingle shell (A. simplex). Figure 4.9, Lanes 8 to 10 and 13 resolve shell matrices having nacreous/aragonite and include the pearl oyster (P. fucata) and stiff pin shell (A. rigida), which are members of the order Pterioida, and the Blue Mussel (M. edulis) from the order Mytiloida. Figure 4.9, Lane 13 resolves matrix from the Eastern complanata (E. complanata), a freshwater species from the Subclass Palaeoheterodonta and the order, Unionoida. Figure 4.9, Lanes 11 and 12 resolve shell matrices having porcelaneous aragonite microstructure/mineralogy and include the common quahog (M. mercenaria) and giant heart cockle (D. robustum), both of which belong to the subclass Heterodonta,
order Veneroida. A summary log of matrix protein bands identified for each species is found in Appendix A.
Figure 4.9  SDS-PAGE analyses and differential staining of soluble organic matrix proteins from selected bivalve species.
Samples were resolved on 4 to 20% Tris-glycine gels. All proteins were loaded in a quantity of 25 ug dry weight per well.

A- Gels were stained with Stains-all
B- Gels were stained with Coomassie brilliant blue 250
C- Gels were stained with silver nitrate.

Lane a - SeeBlue broad molecular weight protein standard (Novex)
Lane b - Mark 12 protein standard (Novex)

Lane 1 - C. virginica
Lane 2 - C. gigas
Lane 3 - O. edulis
Lane 4 - P. gibbosa
Lane 5 - A. simplex
Lane 6 - A. irradians
Lane 7 - P. magellanicus
Lane 8 - P. fucata
Lane 9 - A. rigida
Lane 10 - M. edulis
Lane 11 - M. mercenaria
Lane 12 - D. robustum
Lane 13 - E. complanata

A cursory examination of the gels show that SM proteins from foliated/calcite shell (Figure 4.9, A Lanes 1 to 7) stain much more profusely with Stains-all compared to Coomassie or silver stains (Figure 4.9, B and C Lanes 1 to 7). With the exception of P. gibbosa\(^1\), several prominent bands of comparable size are detected only by Stains-all which indicates they are highly acidic. Some of these bands, particularly those visible in

\(^1\) The lack of resolution and staining of proteins from P. gibbosa is likely due to the fact that matrix was extracted from shells that were collected post mortem. It is reasonable to assume that some degradation of protein could occur, this is especially the case for phosphoproteins as phosphate is susceptible to hydrolysis at pH over 8. Still, Plicatula shares the order Ostreoida with oysters and resolves a light SA positive band against a high background staining similar in size to other folia derived matrices. While its over-all composition is informative, identification of specific protein bands is limited.
O. edulis (Figure 4.9, A Lane 3), A. simplex (Figure 4.9, A Lane 5) and the scallops (Figure 4.9, A Lanes 6 and 7), are quite diffuse. Moreover, background staining appears as a smear spanning at least half the length of the gel lanes (Figure 4.9, A Lanes 1 to 7). These staining characteristics are commonly observed for shell proteins (Mann, 2001 and Marin et al., 2005). Considered with amino acid compositions and phosphate values reported herein (Table 4.4 and 4.6), these results support that SM from foliated calcite is comprised predominantly of acidic proteins and phosphoproteins are the dominant class.

The SM from aragonite shell generally does not stain well with Stains-all (Figure 4.9, B Lanes 8 and 9), but Coomassie and silver staining detect diverse protein patterns among the nacre and porcelaneous microstructures. This tendency is illustrated best in the nacre of M. edulis (compare Figure 4.9, A, B, and C, lane 10), and the porcelaneous M. mercenaria (Compare Figure 4.9 A, B and C, lane 11). Moreover, a majority of the proteins identified in aragonite SM are below 50 kD compared to the acidic proteins of folia, which are generally above 50 kD. This may indicate that proteins in aragonite SM are not as highly modified as those of folia which migrate less efficiently causing their weights be over-estimated (Mann, 2001). Also, the lower phosphate and carbohydrate determined for SM from the aragonitic group supports this conclusion (Table 4.6).

When considered with chemical compositions (Table 4.4 and 4.6), SDS-PAGE results illustrate that nacre and porcelaneous aragonite have more diverse mixture of proteins with moderate acidity and slightly more hydrophobic properties compared with
those from foliated calcite. In addition, matrix complexity is higher in aragonite SM and quite variable between species where the number of components detected ranged from 4 in some species to as many as 16 components in others (see Appendix A). This may reflect this group’s longer evolutionary history (Carter, 1980).

There are exceptions to the above observations. Among species with foliated/calcite shells, *O. edulis* is unique due to an exceptionally prominent 14 to 17 kD doublet band that is visible only by Coomassie blue stain (Figure 4.9, B Lane 3) and resembles a similarly sized band in *M. mercenaria* (Figure 4.9, B Lane 11). The nacre layers of *P. fucata* and *A. rigida*, appear to possess characteristics of both calcite/foliated and aragonitic groups as their SM components stain reasonably well with Stains-all and Coomassie blue (Figure 4.9, A and B respectively, Lanes 8 and 9). The fact that some of the components are sensitive to Stains-all and that total phosphate is estimated to be higher for these species compared to other aragonitic species (Table 4.6) suggests that *P. fucata* and *A. rigida* have phosphoprotein components.

Similar SM protein patterns are evident among related species. The three oyster species (Figure 4.9, A Lane 1 to 3) belonging to the family Ostreidea, each have two prominent Stains-all positive bands and are presumed to be phosphoproteins. In all three cases, the major band is the lower molecular weight protein and measures approximately

---

2 Molecular weights are approximate as the exact measurements will shift slightly from run to run and especially when using different gel systems (personal experience). For example, the 58kd stains-all positive band in *C. virginica* has been reported as a 55 kDa band in chapter 1 and 2 and previously (Mount,
48 kD in *C. virginica*, 52 kD in *C. gigas* and 57kDa in *O. edulis*. Likewise, the minor band is the higher molecular weight band and measures at 58, 56, and 66 kD respectively. Silver staining reveals protein similarities between the *Crassostrea* species where three prominent proteins of approximately 70, 60, and 55 kD differ only slightly in molecular weight and staining intensities (Figure 4.9, C Lane 1 & 2). Coomassie blue staining (Figure 4.9, B) identified in *C. gigas* two bands at approximately 69 and 15 kD; similar bands appear *C. virginica* but are barely visible. Overall, Proteins appear to be conserved within the *Crassostrea* genus.

Differential staining reveals distinct differences of SM components between the genera *Ostrea* and *Crassostrea*. *O. edulis*, as mentioned above, has a major doublet band at 14 and 17 kD that distinguishes it from the other oysters (Figure 4.9, B). Moreover, this species contains about half the amount of serine and about one third the amount of phosphate reported for *Crassostrea*. In Figure 4.9, C, *O. edulis* (Lane 3) appears to lack the three prominent silver staining proteins identified in *Crassostrea* (Figure 4.9, Lanes 1 and 2). Still, within the oyster family Ostreidae, the phosphoprotein class appears to be conserved.

---

1999, Myers, 1998). The difficulty in resolving highly modified proteins by electrophoresis and other chromatography methods are well documented and are, in part, due to the effect of the bulky side groups that cause the proteins behave erratically in methods involving chromatographic separation. Such modifications act by increasing surface charge and promoting the formation of structures that can over-estimate the actual size of the proteins during chromatographic separation.
SM from members of the Order Pectinoida (Figure 4.9, A Lanes 5 to 7) display similar banding patterns and staining characteristics and illustrates their close genetic relationship. Both scallops (family Pectinidae) and the jingle shell (family Anomiidae) resolve prominent, high molecular weight phosphoprotein bands. In *A. simplex* this band occurs at 172 kD and around 114 kD in both *A. iradians* and *P. magellanicus*. All three species share a series of relatively minor stains—all positive bands of varying intensities between 49 and 77 kD (Appendix A). In figures 4.9, B and C, some bands distinguish *A. simplex* from the scallops, including one identified by Coomassie stain occurs at 5.3 kD (Figure 4.9, B Lane 5), and one identified by silver stain occurs at 28 kD (Figure 4.9, C Lane 5). However, several bands that are similar in size but vary in staining intensity appear to be common to all three species.

Remarkably, the banding patterns of the scallops from the Argopecten and Placopecten genera are nearly identical (Figure 4.9, B and C). Components, identified by each of the stains, exhibit only slight shifts in migration and appear to differ mostly in their staining intensities. Only a single band detected at 41 kD (Figure 4.9, C Lane 6) appears unique to *A. iradians*. The remainder of the bands appears in both species but vary considerably in their intensity. For example, a band at 76 kD (Figure 4.9, C lane 6) stains lightly in *A. iradians*; a similar band at 77 kD (Figure 4.9, C Lane 7) stains darker in *P. magellanicus*. These same bands stain oppositely with Stains-all; that is, the 76 kD stains darkly compared to the lighter staining 77 kD band. This may indicate a difference in the extent to which the proteins are phosphorylated. Within the Pectinidae family, a
large part of the SM components appeared to be conserved including the phosphoprotein class.

Overall, SM from all of the species with foliated calcite shell which represent two phylogenetic orders possess one distinctly prominent phosphoprotein band. This suggests phosphoproteins are conserved over a broad range of species and correlate with foliated shell structure. SM components, not from the anionic class of proteins, appear to be conserved within families and to some degree between families. These proteins vary mostly in their staining intensities while their molecular weights are nearly identical.

SM among species having aragonite shell structure exhibits considerably more diverse protein patterns than those from folia. Compared to orders Ostreoida (oysters) and Pectinoida (scallops and jingle shell), there are less protein similarities within and between orders of the aragonite group. And, unlike foliated shell, there is no identifiable “hallmark” protein that stands out for any particular microstructure. For example, *M. mercenaria* and *D. robustum* belong to the order Veneroida, and both have porcelaneous aragonite shell structures. In Figure 4.9, B and C, Lane 11, *M. mercenaria* stained well with Coomassie blue and with silver nitrate revealing several prominent bands, some of which are lightly stains-all positive indicating that they are anionic, and possibly phosphorylated proteins (Table 4.6). *D. robustum* (Figure 4.9, Lane 12) did not stain with either Coomassie or Stains-all (Figure 4.9, B and C, Lane 12). Silver stain (Figure 4.9, C Lane 13) detected intense staining in the well region, but no discrete bands were
evident with any of the stains. This species has a remarkable content of aspartic acid; at nearly 50 mole %, it is the highest amount determined for any shell matrix. It was observed that SM from this species, in particular, formed aggregates that were difficult to solubilize prior to electrophoresis. This portion may contain a majority of the anionic protein components. The heavy staining near the well of the gel indicates the presence of exceptionally high molecular weight protein. So, despite their close relation and similar shell microstructures, matrix components from *M. mercenaria* and *D. robustum* are quite distinct.

The same can be said for nacre. A comparison of *P. fucata* (Figure 4.9, Lane 8), *A. rigida* (Figure 4.9, Lane 9) *M. edulis* (Figure 4.9, Lane 10) and *E. complanata* (Figure 4.9, Lane 13) in all three gels shows no obvious pattern among these identical shell structures. Each contain a distinct anionic protein detected by Stains-all (Figure 4.9, A), but their molecular weights are significantly different. For example, *M. edulis* and *E. complanata* occur at 99 and 15 kD respectively (Figure 4.9, A, Lane 10 and 13). *P. fucata* and *A. rigida* are more closely related and they do have a similarly sized acidic protein at around 54 and 59 kD respectively (Figure 4.9, A, Lane 8 and 9). However, this band is significantly more prevalent in *P. fucata* and occurs as doublet. In fact, *P. fucata* resolves a variety of acidic (Stains-all positive) proteins between 28 to 60 kD, not observed in any other aragonite SM.
Coomassie and silver stain detect the greatest number of unique SM proteins that occur in nacre which are summarized for each species in Appendix A. A major band that is present in both *P. fucata* (60 kD) and *A. rigida* (59 kD) is an exception and stains best with Coomassie (Figure 4.9, B, Lanes 8 and 9) and silver (Figure 4.9, C, Lanes 8 and 9). This particular protein may represent a conserved protein between these species which share the Order Pterioida. Over-all, SDS-PAGE shows that SM protein patterns among the nacre and porcelaneous aragonite microstructures are distinct.

For two species, it was possible to compare protein patterns between shell layers. For *P. fucata* and *A. rigida*, SM from the inner nacreous shell layers were compared with those from whole shell SM, which contains mostly prismatic calcite (Figure 4.10, A and B). In Figure 4.10, A, SM from the whole shell of both species (Figure 4.10, A Lanes 2 and 4) stains more deeply with Stains-all than SM from nacre (Figure 4.10, Lanes 3 and 5) and indicates that prismatic calcite has more acidic components. Some bands appear to be layer specific. For example, a 78 kD band in *A. rigida*, is prominent in whole shell (Figure 4.10, A Lane 4, indicated by a red arrow) but not apparent in nacre (Figure 4.10, A Lane 5). Differences between layers are less apparent in *P. fucata* but silver staining in the 30 to 40 kD range appears to be differ (compare Figure 4.10, B Lanes 2 and 3).
Figure 4.10  SDS-PAGE analysis of matrix proteins extracted from whole shell compared to matrix protein from the inner nacreous shell layers of *A. rigida* and *P. fucata*.

Samples were resolved on a 4 to 20% Tris-glycine gel, cut in two sections, and stained with Stains-all and silver nitrate. Extracts were loaded in a quantity of 25 µg dry weight per well.

A - Gel section stained with Stains-all  
Lane 1 - SeeBlue protein standard (Novex)  
Lane 2 - *P. fucata* nacre+calcite  
Lane 3 - *P. fucata* nacre  
Lane 4 - *A. rigida* nacre+calcite  
Lane 5 - *A. rigida* nacre

B - Same gel as shown in “A” stained with silver nitrate
Note that samples containing both shell layers stain more readily with Stains-all than lanes containing only nacre matrix which indicates that the calcite layer contains proteins with higher anionic character compared to those within nacre. In *P. fucata*, little difference in protein pattern is evident between the layers with Stains-all. *A. rigida* exhibits a different banding pattern of anionic proteins between layers, particularly a 78 kD protein indicated by (←). A 59 kD protein indicated by (↔) is present in both layers and is visible with Stains-all and silver staining. This protein is more prevalent in the calcitic layer (A lane 4). Silver stain reveals protein pattern differences between whole shell and nacre shell layers for both species; these proteins are presumed to be less anionic.

For the most part, however, the SM patterns of both layers are quite similar but the components vary markedly in staining intensity and color. This is especially evident in *P. fucata* following stains-all staining (Figure 4.10, A Lanes 2 and 3). In *A. rigida*, only hints of the prevalent proteins found in whole shell matrix are visible in nacre. For example, two lightly staining bands are evident at approximately 40 kD (Figure 4.10, A Lane 5). This same region stains significantly darker in lane 4. A stains-all positive band at 59 kD (Figure 4.10, A Lanes 4 and 5 indicated by a green arrow), occurs in significantly higher quantity in whole shell (Figure 4.10, A Lane 4) as indicated by its broad band width and intense staining. This same band appears as a “negative” band with silver nitrate (Figure 4.10, B Lane 4), but stains brown in nacre (Figure 4.10, B Lane 5). This staining variation is indicative of a difference in surface charge, as proteins with excessive anionic charge can appear as a negative band following silver staining (Rabilloud, 1990). So, in this case, the band is more highly phosphorylated (probably not glycosylation in this case; Figure 4.11), or there is simply more of it in the calcite layer than in the nacre layer. Similar staining differences are observed in *P. fucata*. For
example, two bands at 54 and 31 kD (both stain-all positive) appear as negative bands in whole shell (Figure 4.10, B Lane 2) but stain brown in nacre (Figure 4.10, Lane 3).

Figure 4.11 Glycoprotein identification compared to non-glycosylated proteins of soluble matrix from selected bivalve species using gel blot analysis
Twenty-five ug of dry weight protein from each matrix was run on 4 to 20% Tris-glycine gradient gels and blotted onto nitrocellulose. Blots were first stained to identify glycosylated protein. Blots were then back-stained to identify non-glycosylated proteins. Fetuin was run as a glycoprotein control and stains blue; creatinase was run as a non-glycosylated protein control and stains brown. The intensity of color is directly proportional to the amount of carbohydrate present on each protein. The 48 kD protein in C. virginica is indicated. A (←) indicates a possible conserved glycoprotein among aragonite species. Circled regions indicate a glycoprotein pattern among species.

A - Blot stained with anti-DIG/ TETON, an antibody/substrate identifying glycosylated proteins.
Lane A - SeeBlue protein standard (Novex)
Lane 1 - C. virginica
Lane 2 - C. gigas
Lane 3 - O. edulis
Lane 4 - P. gibbosa
Lane 5 - A. simplex
Lane 6 - A. irradians
Lane 7 - P. magellanicus
Lane 8 - Fetuin
Lane 9 - Creatinase
Lane 10 - P. fucata nacre + calcite
Lane 11 - A. rigida nacre + calcite
Lane 12 - P. fucata nacre
Lane 13 - A. rigida nacre
Lane 14 - M. edulis
Lane 15 - M. mercenaria
Lane 16 - D. robustum
Lane 17 - E. complanata

B - Blot in “A”, back-stained with Anti-Fluos/INT, an antibody/substrate identifying non-glycosylated proteins relative to glycosylated proteins.

Over-all, this study shows that in a single species, many SM proteins are common to prismatic calcite and nacre, but there appears to be higher levels and more extensive phosphorylation or glycosylation of these proteins in the prismatic calcite shell layer.

**Glycoprotein Identification**

Component glycoproteins of each matrix were identified from non-glycosylated proteins by staining gel blots with a glycoprotein specific stain followed by a general protein stain (Figure 4.11). Glycoproteins are identified first and appear blue, and then blots are back-stained to reveal other proteins which stain brown. Although this assay is generally qualitative, the staining intensity of identified glycoproteins correlates to the
amount of carbohydrate groups within that protein band. Bands that change from blue to brown have less carbohydrate than those that remain blue. Moreover, the rapidity with which this reaction occurs is an indicator of the extent of glycosylation. That is, the faster a blue band turns brown, the less carbohydrate is present.

In the Ostreidae family, both oysters of *Crassostrea* genus show broad glycoprotein staining and nearly identical banding patterns. Some of these components vary slightly in molecular weight but significantly in their staining intensities. For example, both species have similar glycoprotein bands at 48 kD (52 kD in *gigas*), 35 kD doublet (36 kD doublet in *gigas*), and 29 kD (28 kD in *gigas*); however, the 48 and 35 kD bands stain more intensely in *C. virginica* and the 28 kD protein stains more intensely in *C. gigas*. This suggests that between these species, similar proteins vary in the degree of glycosylation or in the amounts of the glycosylated protein. *C. virginica* has a unique glycoprotein at 105 kD and *C. gigas* has a heavily glycosylated protein at 69 kD (also Coomassie positive) not observed in *virginica*. This demonstrates that there are species specific glycoproteins between these closely related individuals. *O. edulis* shows a pattern quite distinct from the *Crassostrea* genus, including two moderate to heavy glycosylated proteins (also Coomassie positive) at 30 and 14 kD. No distinct glycoprotein band resolves for *P. gibbosa*. Also of note is that the glycoproteins identified in the oysters are not especially stains-all reactive, with the exception of the major stains-all positive bands at 48 kD in *C. virginica*, and 52 kD in *C. gigas*. None of the Stains-all positive bands in *O. edulis* stain for carbohydrate indicating that the
glycoprotein class is generally distinct from the phosphoprotein class in the order Ostreoida (Appendix A)

Within the scallop family Pectinidae, both species display similar protein patterns; however, the amount and extent of glycosylated protein varies substantially. *P. magellanicus* exhibits a broad and intense glycoprotein staining compared to *A. irradians* (Figure 4.11, A and B, Lanes 6 and 7). A common glycoprotein at 114 kD is significantly more prevalent in *P. magellanicus*. *A. irradians* has a unique and heavily stained doublet glycoprotein at 16 kD compared to a larger and lightly stained 25 kD glycoprotein in *P. magellanicus*. *A. simplex* is distinct from the scallops and contains heavily glycosylated proteins at 31(doublet), 29, and 10 kD. This species exhibits remarkable pattern similarity with oysters, and like oysters, the major glycoprotein bands do not coincide with the major proteins identified with Stains-all. Conversely, the major Stains-all positive bands identified in both scallops are glycoprotein positive, especially in *P. magellanicus*.

Overall, glycoproteins vary within and between the families of the orders Ostreoida and Pectinoida, both in the amounts of protein and extent of glycosylation. The anionic class of proteins in scallops are phosphorylated and glycosylated. In oysters, there seems to be a demarcation between glycoproteins and phosphoproteins. As far as shell structure, there is no component glycoprotein or pattern that is common to the foliated calcite group.
Despite the fact that total carbohydrate amounts estimated for aragonite matrices are slightly lower than those for calcite (see Figure 4.8), several prominent glycoproteins appear in each of the aragonite species, some of which are similar in size and staining intensities. For example, in Figure 4.11, A and B, *P. fucata* (Lane 12) and *A. rigida* (Lane 13), share a similar major glycoprotein of 61 and 59 kD, respectively. A similar band appears in *M. edulis* (60 kD), *M. mercenaria* (63 kD), *D. robustum* (63 kD, only visible in Figure 4.11, B, lane 16), and *E. complanata* (59 kD), but varies in staining intensity. Because this band is present in all of the species, which are distantly related, it may represent isoforms of a conserved matrix glycoprotein. However, it does not correlate with shell structure, as it is abundant in the prismatic calcite layer of *P. fucata* and *A. rigida*, which is discussed below.

*M. edulis*, *E. complanata*, and *M. mercenaria* stand out among species with aragonitic shell as having more discrete glycoproteins with a relative higher level of glycosylation. In fact, these species display remarkably similar glycoprotein banding patterns despite their distant relation. *M. edulis* contains a major glycoprotein (triplicate) at 21 kD and minor components at 60, 31, and 11 kD. *E. complanata* has a major glycoprotein band at 43 kD and minor components at 62, 30, and 17 kD. The intense staining of matrix from this species, in particular, was surprising given the low total carbohydrate estimation. *M. Mercenaria* has glycoproteins of 65, 17, and 13 kD. The similarity in pattern of these species is evidence that some glycoproteins may be conserved in species with aragonite shell.
Differences in glycoproteins were observed between the prismatic calcite and nacre shell layers. In *P. fucata*, the 61 kD glycoprotein (a major Coomassie positive band) is more heavily glycosylated or in higher amounts in the whole shell (Figure 4.11, A and B, Lane 10) compared to nacre (Lane 12). In *A. rigidida*, the opposite was observed where the major 59 kD glycoprotein (major Stains-all and Coomassie positive protein) is either more prevalent or more extensively glycosylated in nacre (compare Figure 4.11, Lanes 11 and 13). In this case, the extent of glycosylation of a similar protein and/or the amount is variable between microstructures.

There is evidence of conserved patterns of glycoproteins in several species including the approximately 60 kD protein as discussed above and another low molecular weight pattern (circled) in Figure 4.11. However, no specific glycoprotein observed in any of the species correlates with any particular shell mineralogy or microstructure. Another point worth mentioning is that additional bands were identified using this double staining method which were not detected with the conventional stains following SDS-PAGE. Although many of these proteins are minor components, it underscores the complexity of the matrix beyond the acidic class.

**ELISA and Western Analysis**

A poly-clonal antibody generated to SM from *Crassostrea virginica* (anti-SM), was used in ELISA and western analysis to ascertain the degree of relation among species
with respect to their phylogeny, shell matrix and shell structure. In these studies, the antibody is also considered as an “anti-folia” antibody and *C. virginica* serves as the out-group for comparing matrix/shell structure relationships. Results from indirect ELISA indicate that there is a strong conservation of matrix protein epitopes between closely related individuals. This is well illustrated among the oysters which show the strongest immunoreactivity to the anti-SM antibody (Figure 4.12). More significant, however, is that the degree of immuno-reactivity continued to generally parallel phylogenetic relation, despite shell microstructure or mineralogy. This is evidenced by the rankings of individuals according to their degree of cross-reactivity with the anti-SM antibody (Figure 4.13). In this plot, after oysters, the highest activities are observed for *M. edulis* and *P. fucata*, both nacre species but relatively close relatives to *Crassostrea* according to genetic phylogeny (Figure 4.3). Following these species are *P. magellanicus*, *A. simplex*, *A. irradians* (all foliated), *A. rigida*, and *E. complanata* (both nacre). The most distantly related species, *D. robustum* and *M. mercenaria* exhibit the lowest cross-reactivity. The lack of a strong reactivity with *P. gibbosa* may be attributed to the deterioration of matrix protein structure (as footnoted above).
Figure 4.12 Enzyme linked immunosorbent assay (ELISA) demonstrating cross-reactivity of the anti-SM antibody made against C. virginica to SM of selected species. Anti-SM serum was serially diluted and reacted with 10μg/well of SM protein from each species. Myoglobin was included as a negative control. Y-intercept values were determined by applying a logarithmic regression curve fit to the data.
Figure 4.13  Bar plot showing% cross-reactivity of SM among species compared to C. virginica

Shell microstructure and mineralogy are indicated. Cross-reactivity was calculated using Y-intercept values determined from ELISA plot with C. virginica value (1.321) representing 100%. M. edulis and P. fucata are cross-reactive with C. virginica despite having different shell structure, and may reflect their molecular phylogenetic relationship to Crassostrea (Adamkewicz et al., 1997).

Western blot data supports the findings of ELISA and identified component proteins that contribute to immunoreactivity. Shown in Figure 4.14, the anti-SM antibody recognizes a broad spectrum of matrix protein bands identified in C. virginica (Appendix A), especially the 48 kD Stains-all positive band and the 70 kD silver stain positive band (Figure 4.14, Lane 1). The antibody also recognizes several other bands including the 55 kDa phosphoprotein (see Appendix A). A nearly identical pattern of reactivity was observed for C. gigas (Figure 4.14, Lane 2), which indicates that many matrix components of the genus Crassostrea, including the phosphoprotein class, have conserved structures. In the lower molecular weight range, some species specific bands appeared including an 8 kD band in virginica, and a 28 kD band in gigas. O. edulis exhibited less broad reactivity compared to the genus Crassostrea, but the pattern of immuno-reactive bands was nearly identical (compare Figure 4.14, Lanes 1 to 3). Based on staining intensity, these common bands appear to vary mostly in abundance and may represent genetically conserved proteins among the Ostreidae family. The prominent Stains-all positive band (56 kD) identified in O.edulis is moderately immuno-reactive but the major Coomassie blue positive bands (14 and 17 kD) are not (Appendix A and Figure
This suggests that some of the major matrix proteins of *O. edulis* are structurally distinct from the *Crassostrea* oysters.

**Figure 4.14 Western Blot showing cross-reactivity of the anti-SM antibody made to SM from *C. virginica* with SM from selected species**

Twenty-five ug of dry weight protein from each species was resolved on 4 to 20% Tris-Glycine gels in the presence of SDS, and transferred to nitrocellulose for 1 h at 100 V in buffer containing 25 mM Tris, 192 mM glycine, 20% v/v methanol, pH 8.3 using a mini-transfer unit (Bio- Rad). Band molecular weights are slightly lower using the Biotinylated protein standards (see lane 1 note below).

Lane a - Biotinylated protein standard (Bio-Rad)
Lane 1 - *C. virginica* (black arrows indicate the 48 and 55 kD proteins)
Lane 2 - *C. gigas*
Lane 3 - *O. edulis*
Lane 4 - *P. gibbosa*
Lane 5 - *A. simplex*
Lane 6 - *A. irradians*
Lane 7 - *P. magellanicus*
Lane 8 - *P. fucata* nacre+calcite
Lane 9 - *A. rigida* nacre+calcite
Lane 10 - *M. edulis*
Lane 11 - *M. mercenaria*
Lane 12 - *D. robustum*
Lane 13 - *E. complanata*
The scallops of the Pectinidae family and *A. simplex* from the family Anomiidae are distantly related to the oysters but share similar shell structure (calcite/folia). Both scallops exhibited a similar pattern of immuno-reactivity with the anti-SM antibody (Figure 4.14, Lanes 6 and 7). Several discrete bands are medium to strongly reactive and are generally comparable in size and number to those of the *Crassostrea* genus (Figure 4.14, Lanes 1 to 3 and Appendix A). This suggests that some SM proteins may be conserved between the Pectinidae, Anomiidae, and Ostreidae families. *A. simplex* (Figure 4.14, Lane 5), exhibits a relatively light reactivity despite its higher ELISA value. Most important is that the major stains—all positive bands identified for these species including, 208 kD in *A. simplex*, 113 kD in *A. irradians*, and 114 kD in *P. magellanicus* (Figure 4.9, as summarized in Appendix A) show only light immuno-reactivity. Therefore, the major phosphoproteins are structurally distinct among the bivalve families with foliated shell.

Immuno-reactive bands occur in all species with aragonitic shell matrices (Figure 4.14, Lanes 8 to 13), many of which have molecular weights similar to those observed for calcitic shell (Figure 4.14, Lanes 1 to 7 and Appendix A). The strongest immuno-reactivity is observed in nacre where a broad spectrum of discrete proteins showed medium to high reactivity. These proteins are not confined to any one class based on their differential staining characteristics (Appendix A). Of the nacre shelled species, *P. fucata* and *M. edulis* exhibit the highest immunoreactivity and in consistent with ELISA. *P. fucata* (Figure 4.14, Lane 8) exhibits strong reactivity to a wide range of proteins.
Notably, the prominent Stains-all positive doublet bands identified at 50 and 54 kD are less immuno-reactive than the major Coomassie staining bands at 60, 36, and 28 kD (Appendix A). *M. edulis* also shows a broad range of distinct proteins that are strongly immuno-reactive and, like *P. fucata*, the major Stains-all positive band shows relatively weak cross-reactivity compared to those identified with other stains. In both of these species, proteins in the 90 to 200+ range were only visible by immuno-staining (Appendix A). The high reactivity of the anti-SM antibody with matrix components of *P. fucata* and *M. edulis* is likely a result of their relatively close relationship with *C. virginica*.

*A. rigida* immuno-stains less intensely than the *P. fucata* and *M. edulis*, however, the proteins identified correspond nearly perfectly with those identified for it’s close relative, *P. fucata*. The matrix of *E. complanata* exhibits strong reactivity to all the protein components identified by SDS-PAGE and especially to a 15 kD protein which is acidic and glycosylated (Figure 4.14, Lane 13 and Appendix A). Matrix proteins from porcelaneous aragonite shell were the least reactive to the anti-SM antibody. Both *M. mercenaria* (Figure 4.14, Lane 11) and *D. robustum* (Figure 4.14, Lane 12) show only a few weakly staining bands; however, similarly sized bands are evident in oysters, *A. irradians, P. fucata* and *M. edulis*. These bands may represent conserved shell matrix proteins.
It is important to note that the first ELISA experiments were conducted with anti 48 kD antibody made from *C. virginica* (Appendix B). Curiously, this antibody did not react well with other species, even species within Ostreoida and supports that the phosphoproteins among the species, though related, have distinct primary structure arrangements.

Over-all, similar protein patterns are evident among related species with foliated shell following immuno-staining. Aragonite patterns were more variable. There was no obvious correlation between immuno-reactive bands with shell structure.

**Discussion**

A question unanswered in studies of shell formation is whether shell mineralogy and microstructure are determined by a specific set of shell matrix proteins. This question stems from several commonalities among bivalve shell structures and their constituent matrix characteristics. Shell microstructures are conserved over a wide range of species and almost exclusively occur as either calcite or aragonite polymorphs of calcium carbonate (Lowenstam, 1981). In addition, most shell soluble matrix (SM) is unusually anionic and *in vitro* studies have demonstrated an ability to control many attributes of shell structure. When immobilized to a surface, SM proteins control the size, number, mineralogy, and orientation of the crystals that form (Falini et al., 1996; Mount, 1999; Samata et al., 1999; Feng et al., 2000). Acting from solution, SM adsorbs
to crystals affecting their rate of growth, shape, and mineralogy (Wheeler and Sikes, 1984; Belcher et al., 1996; Kono et al., 2000; Thompson et al., 2000; Michenfelder et al., 2003; Kim et al., 2004).

The mechanisms by which matrix carries out these activities remains unknown but a popular theory is that the interactions occur through ordered anionic domains repeated within the protein structure that interact specifically with mineral (Weiner and Hood, 1975; Weiner and Traub, 1984; Weiner and Addadi 1991). It is reasonable to speculate that these domains might be conserved and potentially microstructure specific.

**Mineralogy**

This study included seven species with predominantly calcite shell and six species with shells composed of aragonite. Analyses of their respective SM show that these mineralogies share some general matrix characteristics. All are enriched in glycine and anionic amino acids. Aspartic acid is the most prominent followed by serine and glutamine. The dominance of these amino acids is typical for shell derived matrix (Lowenstam, 1981, Travis et al., 1967). In addition, SM components resolved by SDS-PAGE generally fall into two main protein classes based on differential staining: a highly acidic class, identified by Stains-all, including proteins that are phosphorylated and to a lesser extent glycosylated in some species, and a moderate to low acidic class, identified with Coomassie blue and silver nitrate.
Nonetheless, a more detailed analysis of SM compositions and their constituents reveals that calcite and aragonite are distinct. Calcite SM, especially from foliated mineral, was found to have more acidic character compared to aragonite SM. Rose plots of amino acid groups (charged/basic, hydrophobic, and uncharged) and the most prevalent amino acids (Asp, Ser, and Gly) results in two basic, but distinct plot shapes characteristic of each mineralogy. “Right handed” plot shapes result from a dominance of acidic groups and were characteristic of calcite SM. Aragonite SM appeared as “left handed” shapes which reflect a higher hydrophobic and basic amino acid content. This chemical characteristic may explain, in part, the lower solubility of aragonite SM in water.

The higher anionic character of calcite matrix is further supported by the reported C/HP ratios (charged amino acids compared to hydrophobic groups) which are substantially higher than those calculated for aragonite SM. Moreover, phosphate levels are elevated in calcite SM, which contribute to the profuse staining of its constituents by stains-all. In fact, differential staining reveals that calcite SM has a higher proportion of highly acidic proteins compared to aragonite SM which does not stain well with stains-all.

These analyses support that at least, on a general level, matrix composition may influence the selection of calcite or aragonite mineralogy. In particular, acidic proteins appear to correlate with calcite mineralogy while more hydrophobic proteins associate
with aragonite. A recent review of sequences, mostly from gene isolates, reports that proteins derived from calcite prisms and folia have pI’s in the range of 1.45 to 3.9 compared with aragonite derived proteins which fall into the range of 4.9 to 8 pI (Zhang and Zhang, 2006).

Microstructure

SM from each species was compared with respect to their constituent microstructures. Two calcitic microstructures, including foliated and prismatic, and three aragonite microstructures, including nacre, and homogenous and crossed lamellar structures from the porcelaneous category (Figure 4.1), were represented. Among these, foliated SM is most distinct and represents the best evidence for a particular matrix influencing the deposition of a specific microstructure. Compared to prismatic calcite and aragonite microstructures, SM from folia is the most acidic as evidenced by its higher C/HP ratios and lower hydrophobicity values (range 0.10 to 0.14). It has the highest serine and glycine content and as much as 100 times more total phosphate than aragonite derived SM and, on average, 10 times more than prismatic calcite. Moreover, the extent to which Ser/Thr is phosphorylated is greater than 70 % for most species. Folia SM is also slightly richer in carbohydrate content than either prismatic calcite or aragonite microstructures.
Another hallmark of folia SM is the presence of a phosphoprotein class not observed in any other microstructure. Within related groups, the phosphoprotein pattern is highly conserved and considered with their broad background staining and intensity, these components are presumed to be highly phosphorylated. Amino acid compositions support that as a class, these proteins are major SM constituents. With that being said, these phosphoproteins are structurally distinct between families. They vary in size and number between groups, and ELISA and Western analyses show that the epitopes are not conserved to a high degree, except within the Crassostrea genus. In addition, these proteins vary in their extent of phosphorylation and glycosylation based on their staining intensities using agents that detect these groups. So, while folia derived phosphoproteins are chemically similar, the arrangement of their amino acids structures and modifications are distinct.

Compared to folia, SM from aragonite microstructures were more complex and highly variable between species. There was no correlation in protein patterns with any one microstructure. This is illustrated well in the distinct gel protein patterns of species with nacre shells including *P. fucata, A. rigida, M. edulis, and E. complanata*. Dauphin and Dennis (2000) report similar findings which show that SM from 6 species of mollusks with aragonitic crossed lamellar shell microstructure were highly varied in chromatographic elution profiles and protein/carbohydrate ratios.
The variability of aragonite SM may reflect its earlier appearance in the geologic record which is well before the appearance of foliated structure (Carter, 1980). Some theories suggest that microstructure evolution has progressed from aragonite prismatic and nacre, to porcelaneous, and finally to calcite foliated structures (Taylor, 1973; Carter and Tevesz, 1978a, b). The fact that all of the aragonite microstructures contain protein bands which are immuno-reactive with the anti-SM (folia) antibody indicates epitopes are conserved in shell SM proteins, irrespective of shell microstructure and mineralogy. This is illustrated well in *E. complanata* and *M. mercenaria*, where a good many of the major SM proteins are immuno-reactive. In the case of *P. fucata* and *M. edulis*, immuno-reactivity is observed in the acidic proteins. Because they are closer relatives of *Crassatrea* these proteins may represent divergent forms of an acidic shell protein. If bivalves are monophyletic (Winnepenninckx et al., 1994), and there is some debate on this issue (Adamkewicz et al., 1997), these observations support that shell proteins have divergently evolved but have maintained some protein domains even while their shell structures changed. Broad cross-reactivity among the major shell proteins is further evidence that shell structure is not defined by a well defined set of SM proteins.

The most compelling evidence against microstructure specific proteins is the finding that SM protein patterns between distinct shell layers in a single species are mostly similar. This was illustrated for two species, *P. fucata* and *A. rigida*, both of which have an outer prismatic calcite layer and an inner nacre aragonite layer. The proteins that appeared “layer specific” were relatively few, especially in *P. fucata*, and in
both species, the “major” proteins (defined according to band width and staining intensity) occurred in both layers. What varied between the two microstructures were the amounts of these similar proteins. Evidence also suggests some of these proteins differ in phosphorylation and glycosylation levels. The significance of these findings as determinants of shell structure is discussed further below.

Among species having similar shell structures, this study shows no evidence for a conserved set of “folia” or “nacre” specific protein per se; rather a variety of proteins, with chemically similar but distinct primary structures, contribute to the formation identical shell microstructures. In a single species, the fact that many of the major SM proteins occur in both shell layers further undermines the importance of “layer specific” proteins. These findings do not support the idea that crystal initiation, growth, and orientation require a stereo-chemical match between the mineral lattice and precise arrangements of amino acids within the peptide chain (Addadi and Weiner, 1989). For example, (Asp-Y)$_n$ domains were predicted to occur in SM proteins. This spacing was calculated to match the spacing of the Ca$^{2+}$ ions in a crystal lattice of aragonite and calcite crystals (Weiner and Hood, 1975). Altering the local ionic strength, especially Mg$^{2+}$ concentration, could favor the formation of aragonite over calcite (Simkiss et al., 1982). An (Asp-Gly-X-Gly-X-Gly)$_n$ domain was predicted to interact specifically with the lattice arrangement of a particular type of folia (Runnegar, 1984). So far, neither domain has been identified in the 20 plus proteins for which structures have been determined; in fact, the primary sequences are quite variable (Zhang and Zhang, 2006).
If primary protein structure is not the defining influence on shell structure, what other factors may promote the formation of a particular microstructure?

**Higher order structure**

*Secondary structure*

Secondary protein structures may influence shell structure. This study shows that SM from the shells of all the species tested are enriched in Asp, Ser, Gly, and Glu to varying degrees, regardless of mineralogy and microstructure. By virtue of their prominence, one would predict that these groups would arrange as repeated domains within a protein backbone. As gene sequences become available, it is now clear that shell proteins are, in fact, highly modular and repetitive (Zhang and Zhang, 2006). Some of the common modules include acidic domains, Gly-rich domains, basic domains, and hydrophobic domains. These structural elements vary in size, number, position in the backbone, and arrangement of their constituent amino acids, however, their secondary structures are conserved.

Gly-rich domains often form β-sheets or loops. The prismatic calcite derived protein MS131, contains a β-sheet region comprised of 10 (Gly)_{3.5} blocks (Sudo et al., 1997). Similar, but not identical, domains have been reported for prismilin-14 (Susuki et al., 2004) and the shematrin shell protein family (Yano et al., 2006). These gly-rich, β-
sheet motifs are thought to promote matrix interactions that contribute to the formation of the shell matrix (IM) framework, particularly of the prismatic calcite shell layers.

Loop structures have been predicted for the “GS” domain positioned near the C-terminus of the nacre derived protein, Lustrin A. These repetitive Gly, Ser regions are occasionally interspersed with aromatic amino acids (Shen et al., 1997). A similar domain was identified near the N-terminus of MSP-1, a protein derived from folia of the scallop, though it lacked hydrophobic groups (Sarashina and Endo, 1998; 2001). A loop structure was also predicted for the (GGG)$_3$ (GG)$_3$ domain located near N-terminus of MS17- a protein that occurs in both shell layers of $P. fucata$ (Zhang et al., 2003). These structures are similar to those found in keratins are presumed to impart elasticity and flexibility to the protein backbone (Shen et al., 1997).

Acidic domains occur in most of the shell proteins but are quite variable. In the nacreous derived nacrein, a single acidic domain (Gly-X-Asn), where X= either Asp, Asn, or Glu is repeated in the center of the protein and spans approximately 100aa of the total 417 amino acid structure (Miyamoto et al., 1996). Another protein from nacre, MS160, has a single large acidic domain near the N-terminus comprised of consecutive (ESEEDX)$_6$ motifs that are thought to arrange as B-sheets (Sudo et al., 1997).

Acidic domains are far more prevalent in calcitic microstructures and are dominated by Asp. MSP-1 from the foliated layer of the scallop contains 2 highly
conserved and repetitive (Asp-Gly-Ser-Asp) motifs of approximately 50aa in length (Sarashino and Endo, 1998; 2001). In *C. virginica*, the major 48 and 55 kD phosphoproteins isolated from folia are enriched in asp, ser, and gly (Chapter 2) and likely contain repeated poly-asp sequences (Chapter 3) and (Asp-Ser-Gly) sequences (Wheeler, 1992). Similar asp-rich sequences have been identified in prismatic calcite. Aspein, isolated from in *P. fucata*, contains (Asp)$_{2-10}$ regions with intervening SG dipetides. Asp constitutes 60% of the protein (Tsukamoto et al., 2004). The Asp-rich protein family has members containing several asp-runs up to 30 residues in length (Gotliv et al., 2005). Although secondary structures are not reported for these proteins, asp, ser, and gly arranged together often result in $\beta$ – structure (Choi and Kim, 2000). Poly-asp regions of calsequestrin which are similar to those found in asprich occur in $\alpha$ helical motifs (Wang et al., 1998). Both motifs are presumed to interact optimally with Ca$^{2+}$ ions and crystal surfaces.

Studies show that the secondary structures of proteins from different shell layers vary. In *C. gigas*, SM from the prismatic aragonite layer of the myostracum (the small adductor scar region) adopted mostly $\alpha$-helix (18%), $\beta$-sheet structures (41%) and turn structures (27%) and just 12% random coil. The composition from this layer was reported to contain 37% Asp, Ser, and Gly and 36% Glu, Ala, and Val. SM from folia formed slightly less $\beta$-structure (40%) and less $\alpha$ helical (13%) and turn (10%) structures compared to the myostracum but significantly more random coil (35%) structure was detected. The SM from this layer contained very few hydrophobic groups and 74% Asp,
Ser, and Gly (Lee and Choi, 2006). In a similar study of the same species, prismatic calcite was found to have the most random structure compared with either folia or the aragonite layers (Choi and Kim, 2000). Dauphin and Denis (2000) report that SM derived from the crossed lamellar aragonite layer of six species of bivalves and a gastropod occurred as similar amounts of β-sheet, a helix and turn secondary structures, even while their component proteins varied considerably in number, molecular weight, acidity and glycosylation. These studies demonstrate a correlation between microstructure and mineralogy with secondary protein structure.

Tertiary and Quaternary structure arrangements

Higher order structure may influence selection of mineralogy and microstructure. This study shows that the matrix extracts from aragonite and calcite shell were fundamentally different. Matrix from aragonite largely remained insoluble in water (IM) and the SM self-associated to a higher degree than calcite SM, forming aggregates that were refractory to dissolution. In addition, high molecular weight protein aggregates were more evident in the wells of aragonite SM extracts. Marin and colleagues (2005) report that the prism derived SM protein, caspartin, undergoes oligomerization following extraction. It is proposed that the polymerized form interacts with prism nuclei so that they grow in a preferred crystallographic orientation. A major SM protein (20 kD) found in *P. maxima* extracts as a polymer (120 kD) from nacre (Bedouet et al., 2001). The interactions that promote macromolecular assemblages of matrix are unknown, but
hydrophobic, ionic, and covalent mechanisms are often proposed (Crenshaw, 1980). In this study, aragonite matrix was generally less sensitive to BCA chemistry than calcite, despite having more of the reactive groups which are aromatic. This is indirect evidence that hydrophobic associations may be acting to obscure the reactive amino acids.

In the realm of understanding the mechanism of how higher order protein structure affects mineralization, studies of vertebrate enamel formation have been most instructive.

In a proposed model, amelogenin monomers are secreted by ameloblasts and assemble into spherical quaternary structures which are thought to promote growth of initial crystallites along the c-axis while inhibiting lateral growth through face specific adsorption to a,b surfaces. C-terminal regions of amelogenin are then cleaved by proteases which disrupt the adsorption of the “nanospheres” and promote growth in the lateral direction (Robinson et al., 1995). This mechanism has not been demonstrated for shell derived matrix; however, this study and those cited above lend support to the possibility that the differences in tertiary and quaternary macromolecular structures between aragonite and calcite shell matrix may be important in the selection of mineralogy and formation of their corresponding microstructures.
Deployment of SM components

Comparison of SM protein patterns between the shell layers of *A. rigida* and *P. fucata* support that during shell formation, the secretion of layer specific proteins and variable amounts of common proteins are factors that influence the microstructure deposited. Much attention has been given to protein differences between distinct shell layers, particularly those whose genes have been determined that exhibit a zonal mRNA expression pattern in the mantle. Protein expression that is confined to the outer lobe edge (near the shell margin) is presumed to act in the formation of prismatic calcite while those expressed dorsally on the outer lobe are presumed to act in nacre formation (Sudo et al., 1997; Takeuchi and Endo, 2006). Many of these studies suggest that the properties of these proteins directly influence which shell structure is deposited.

Sudo (1997) and colleagues were the first to identify a zonal expression pattern of layer specific matrix proteins in the mantle of *P. fucata*. The protein MS131 was confined to the outer edge of the mantle and directs the formation of prismatic calcite while the protein MS160 was expressed dorsally in the outer mantle, in close proximity with nacre, and presumed to promote the formation of nacre (Sudo et al., 1997). More recently, in the same species, the acidic shell proteins aspein (Tsukamoto et al., 2004) and Prismalin-14 (Susuki et al., 2004) were reported to be specific for the calcitic layer based on their expression in the outer edge of the mantle. Aspein (41 kD) is currently the most acidic matrix protein known and contains numerous (Asp)_{2-10} repeats. Both proteins

246
inhibit calcium precipitation and cause morphological changes in crystals grown *in vitro*. In *A. rigida*, a protein family collectively called Asprich includes at least 7 related proteins which were found to localize exclusively to the prismatic calcite layer (Gotliv et al., 2005).

This study shows that SM from *P. fucata* and *A. rigida* appears to have some layer specific proteins, especially those of the acidic class which are enriched in the calcitic prismatic shell layer. Some of these proteins may correspond to those in the aforementioned studies, which are thought to influence shell structure. But equally significant is the finding that SM protein patterns between the prismatic calcite and nacre shell layers are mostly similar, including the most prominent proteins. What vary most are the staining intensities of these common proteins. A similar pattern has been reported for caspartin, an acidic Ca$^{2+}$ binding protein found in the Fan mussel, *Pinna noblis*. This protein occurs in both the calcitic prismatic and nacreous shell layers, but is far more concentrated in the prismatic layer (Marin et al., 2005). What is interesting is that caspartin was shown to modify the morphology, size, and density of calcite crystals *in vitro*, in a dose dependent manner. These findings suggest that microstructure or mineralogy or both are driven, in part, by the levels of similar proteins deployed during shell layer formation. Additional support for this theory is found in studies of the starmaker protein which mediates otolith formation in fish. When the expression level of this protein was altered by functional knockout, the normally flat, disc shaped aragonite structure that forms during this process changed to resemble a star shape, and finally
appeared as pure calcite crystals when expression was completely inhibited. The degree of this metamorphosis was dependant on the amount of starmaker present (Sollner et al., 2003).

Changes in post-translational modifications

The removal or addition of phosphate and carbohydrates from matrix proteins have been shown to coincide with mineral deposition in a variety of vertebrate and invertebrate systems (Hoshi et al., 1999; Boskey et al., 1997; Shafer et al., 1995; Kabakoff et al., 1992). This study shows that foliated shell is unique among other microstructures, both in the amount of phosphorylated proteins in their SM and the extent to which they are phosphorylated. This characteristic places it in a category comparable with the phosphoryns (PP), the highly phosphorylated SM proteins derived from vertebrate tooth dentin (Veis, 1989). For PP, the degree of phosphorylation is critical for effecting crystal growth in dentin formation (He et al., 2005). Borbas (1991) demonstrated a similar effect using a fraction of oyster SM (RP-1) which contains the 48 and 55 kD phosphoproteins (Myers, 1999). In this study, the most extensively phosphorylated SM proteins were better inhibitors of crystal growth. Synthetic peptide analogs designed to mimic matrix control of crystal growth in vitro were found to promote formation of hydroxyapatite only when serine residues were phosphorylated; amorphous deposits formed when phosphoserine was replaced with serine (Hartgerink et al., 2001). These studies are
compelling evidence that altering phosphorylation is a mechanism that effects crystal growth.

Altering phosphate levels of matrix proteins can be carried out extra-cellularly, by the action of alkaline phosphatase or by translating different forms of the protein with more or less phosphorylation potential. The phosphophoryns occur as three variant forms of varying acidic groups and phosphorylation and suggests that at least the later mechanism may work in dentin formation (George et al., 1998). This study shows C. virginica has highly related phosphoproteins (Chapter 2 and 3) which could also represent differently phosphorylated proteins. Recently, mass spectral analysis identified a peptide in SM that matched an EST encoding alkaline phosphatase (unpublished data) and is evidence for the former mechanism. In any case, phosphorylation clearly correlates with folia and should be studied as a potential mechanism that influences the deposition this microstructure.

Glycoproteins, particularly glycoaminoglycans (GAGS), are a significant component of shell SM (Mann, 2001; Kawaguchi and Watabe, 1993). Their activity is often envisioned as providing anionic potential for interacting with mineral or as a “gelling” agent that increases the viscosity of the mineralizing medium (Myers, 1999). In studies of crab carapace formation, mucin-like glycoproteins identified in the newly formed, non-mineralized cuticle are de-glycosylated extra-cellularly prior to the onset of mineralization. It is proposed that removal of these groups changes the proteins activity
from an inhibitor to a promoter of crystal formation and growth (Coblentz et al., 1998). Two mucins have been identified in nacreous shell and have mineralizing properties (Marin et al., 2000; Mann et al., 2000; Blank et al., 2003).

This study shows that folia have about twice the amount of total carbohydrate than either prismatic calcite or aragonite microstructures. Thus, higher carbohydrate content may influence the deposition of this microstructure. However, this correlation could not be attributed to any one component as the patterns among folia were quite variable. Even similar bands in related species varied in their glycoprotein levels. On the other hand, all the aragonite microstructures resolved what appeared to be a conserved glycoprotein. This protein may be important for promoting aragonite mineral formation.

Also interesting was the identification of a set of similar low molecular weight glycoproteins. Because these were identified in folia and aragonite shells, they may represent conserved glycoproteins that function in the regulation of cells rather than a direct influence on mineral growth. The regulation of mineralization is a largely unexplored area but is garnering more attention as studies reveal the complexity of secreted mantle proteins that are spatially and temporally controlled (Jolly et al., 2004; Jackson et al., 2006). Some of these products no doubt influence the cells associated with mineralization including the mantle epithelium (Myers et al., 2007) and hemocytes.
Glycoproteins may influence shell structure by organizing and directing cell activity at the site of mineralization.

**Influence of minor protein components**

In this study it was observed that many “minor” components (defined by their breadth of staining) occur in all of the species analyzes. Some of these are only detected by immuno-staining with the anti-SM antibody and indicates they have common epitopes with oyster folia SM. Some are only visible after staining with anti-glycoprotein and anti-protein antibodies (see glycoprotein detection results section), which are far more sensitive than Stains-all, Coomassie or silver stain. Some of these proteins, particularly those of similar size and that appear in numerous species, may not be part of the “mineral interacting” class of proteins. Rather, these proteins may play support roles for cellular activity including the mantle epithelium. For example, alkaline phosphatase is a multimeric glycoprotein which occurs as subunits in the 50 kD range in vertebrates (Yokoto, 1978; Cyboron and Wuthier, 1981). This protein may de-phosphorylate proteins involved in shell mineralization, particularly those of the foliated layer (as noted above). Its presence may also indicate cell differentiation processes related to mineralization.

A range of cell regulating molecules are likely present in shell SM as a by-product of the cellular activity and regulation that occurs during the mineralization process.
(Mouries et al., 2002; Mount et al., 2004; Jolly et al., 2004; Simpson et al., 2006; Myers et al., 2007). The recent identification of several minor proteins from aragonite shell and in the mantle organ supports this contention. These include, bone morphogenic factor (BMP-2), a member of the TGF-β (transforming growth factor) protein family that induces bone and cartilage formation (Matsushiro and Miyashita, 2004), a dermatopontin (the vertebrate form binds TGF-β and promotes cell and protein aggregation) (Marxen and Becker, 2000; Marxen et al., 2003) and perlustrin, a member of the insulin-like growth factor binding protein family which was shown to bind IGF’s (Weiss et al., 2001). More recently, a small 10 kD protein named p10 was identified in the nacreous layer of *P. fucata*. This protein was able to induce and alter the growth of aragonite crystals, and promote cell differentiation (Zhang et al., 2006). How these relatively minor matrix constituents influence mineralization is not known but their potential to effect cells and mineral suggests they could play a crucial role in determining shell structure.

**Summary**

The SM from 14 bivalve species was compared in order to determine whether or not specific matrix protein components correlate with shell microstructure or mineralogy.

The amino acid compositions for all the species were found to be enriched in glycine and acidic amino acids, particularly aspartic acid. However, SM derived from calcite shell, including prismatic and foliated microstructures, was determined to be more
acidic than aragonite. This finding is supported by the staining characteristics of the component proteins following SDS-PAGE. Proteins from prismatic and especially foliated calcite stain profusely with the cationic stain, Stains-all, while the proteins from aragonite microstructures stain well Coomassie blue indicating a less acidic character.

Among the microstructures, foliated calcite is distinguished. Folia SM has the highest amounts of glycine (15 to 35 mole%) and serine (10 to 28 mole%) and up to 22% total phosphate by weight of material compared with aragonite microstructures, which were determined to have 6 to 15 mole% glycine, 3 to 9 mole% serine and than 1% total phosphate by weight. Folia SM also contains about 2 to 3% carbohydrate compared to aragonite microstructures which contain about 1 to 1.5% carbohydrate by weight.

Folia SM contains a highly phosphorylated protein class not observed in any other microstructure. The SDS-PAGE patterns of this class are comparable among all species, but are especially similar within families including the Ostreidae (oysters) and Pectinidae (scallops) families. Western blot analyses using an antibody generated to the SM of *Crassostrea virginica* indicate that the phosphoproteins between families have unique primary structures.

Electrophoretic protein patterns of aragonite SM, including nacre and porcelaneous microstructures, were more variable than calcite microstructures and may reflect its earlier appearance in the geologic record. ELISA experiments demonstrated
cross-reactivity among related species, irrespective of shell structure. For example, the
anti-SM oyster antibody cross-reacted with the nacre of *P. fucata* and *M. edulis*, more
intensely than with other folia shelled species. Western analysis shows broad reactivity
to component proteins among all species. Apart from the folia microstructure, there was
no particular protein pattern that correlated with a specific shell structure, rather SM
protein patterns were determined to be species specific.
CHAPTER FIVE

FINAL DISCUSSION

The first part of this study represents a continuation of research into the biochemical characteristics of shell organic matrix from the Eastern oyster, *C. virginica*, with the aim of identifying component proteins which may influence shell structure. In this species, extraction of foliated shell matrix results in three fractions; a water soluble (SM) fraction, an insoluble (IM) fraction and a urea and SDS soluble fraction of the IM (sIM). Previous studies revealed that these fractions are chemically similar and the SM contains up to 17% phosphate by weight (Rusenko, 1988, Borbas et al., 1991; Mount, 1999). In this study, two acidic phosphoprotein components of the SM were isolated and characterized. These proteins are reported to have estimated molecular weights of 48 and 55 kD. Previous studies identified these proteins as major constituents of a reverse phase HPLC fraction of SM termed (RP-1), which has demonstrated crystal binding activities *in vitro* (Rusenko et al., 1991; Wheeler et al., 1991; Myers, 1999). In the second part of this study, the SM from 14 bivalve species was compared to determine if a correlation between matrix structure and shell structure exists.

*Characteristics of the 48 and 55 kD phosphoproteins*

This study shows that the 48 and 55 kD are members of a dominant acidic protein class that occurs in foliated shell matrix. Nearly 85% of their compositions consist of aspartic acid, serine, and glycine in roughly in equimolar ratios. Because their individual
compositions do not deviate significantly from whole SM or IM fractions, and that they stain intensely with the cationic dye, Stains-all supports that the 48 and 55 kD proteins are phosphorylated and the most prevalent acidic components of this class.

The 48 kD protein is a major constituent of the intra-cellular matrix and presumed to localize around and within individual crystal units. This is evidenced by its ready solubilization in water and abundance in the SM fraction. Soluble shell proteins are often referred to as “crystal shapers” because of their ability to influence crystal morphology and growth rate through anionic domains (Mann, 1991).

The 55 kD is presumed to be a protein constituent of the inter-crystalline matrix and may coat the surface of IM and localize around individual crystals. The residual occurrence of this protein in SM is evidence of a more limited association with crystal surfaces while its relative abundance in the sIM fraction indicates a tight association with the IM. One possible mechanism of association between the 55 kD and IM is through hydrophobic interactions. This is supported by their similarity indexes, which shows that the 55 kD protein shares more identity with IM due to its lower ratio of charged to non-polar amino acids. Moreover, this protein adsorbed to hydrophobic supports (PVDF) more efficiently than the 48 kD protein. However, hydrophobic interactions are a relatively weak interaction and cannot explain fully the requirement for stringent agents to release the 55 kD from IM. It is likely that a combination of hydrophobic and strong ionic interactions is acting.
It is important to note that in terms of function, the 48 and 55 kD proteins represent a paradox in categorization as a “crystal shaper” or “adhesive” protein according to their amino acid compositions. The 48 kD actually has a higher content of hydrophobic groups such as Val, Ala and Ile which are groups typically equated with an adhesive function (Mann, 1991). However, the 55 kD exhibits more “adhesive” properties than the 48 kD. Both proteins are equally high in acidic groups which potentially act in crystal and calcium ion binding. Apparently, the ratio of charged (anionic and basic) to hydrophobic groups bears on their interactions with other proteins and mineral surfaces.

The 48 and 55 kD have highly related structures. Their amino acid compositions are nearly identical and antibody cross-reactivity demonstrates that they have multiple common epitopes. In fact, compared to the known shell isolates (reviewed in Zhang and Zhang, 2006), the 48 and 55 kD proteins from oyster have the highest amino acid relation and likely originate from a common gene that is alternatively spliced or are members of a multi-gene family. Recently, a gene family encoding at least 7 highly related matrix proteins was discovered in the prismatic layer of *P. fucata* (Yano et al., 2006; Zhang et al., 2006). Similarly, the Asp-rich protein found in the same species has at least seven highly related members with distinct cDNA’s (Gotliv et al., 2005). Both of these classes are thought to arise from alternative RNA splicing of a single gene. The high relation of the 48 and 55kD protein suggest they too, may come from a common gene.
Both the 48 and 55 kD proteins are phosphorylated. Although the degree of phosphorylation was not empirically determined for either protein, circumstantial evidence supports that a good deal of the serine is phosphorylated based on the prominence of these proteins in SM, the fact that SM contains 16% by weight phosphate, their staining characteristics, resistance to sequencing, and in the case of the 48 kD protein, loss of molecular weight following dephosphorylation.

Evidence supports that poly-serine domains exist in the 48 and 55 kD proteins. Anti 48 kD antibodies generated in chickens cross-reacted strongly with other proteins with known poly-serine domains including, phosvitin, a class of highly phosphorylated proteins found in vertebrate egg yolk (Byrne et al., 1984; Maurizio et al., 2004), and phosphophoryn, a matrix phosphoprotein associated with the formation of vertebrate tooth (George et al., 1998); this indicates that these proteins share structural similarity with the 48 kD protein and a common phosphoserine domain is likely a dominant epitope recognized (Rusenko et al., 1991; Maurizio et al., 2004).

Attempts to directly sequence the 48 and 55kD proteins resulted in limited success, especially for the 55 kD which was most difficult to sequence by Edman chemistry. Both proteins were refractory to cleavage by several common proteases. The 48 kD proteins absolute resistance to V8 protease and its limited digestion by Asp-N from the N and/or C- termini indicates that the 48 kD has poly-aspartic acid domains which are likely concentrated in the middle of the protein backbone. Poly-asp peptides
were previously identified in the RP-1 fraction of SM which is predominantly comprised of the 48 and 55 kD proteins (Rusenko et al., 1991; Myers et al., 1999). And, the fact that these proteins have similar structures, it is likely that the 55 kD, too, has poly-asp motifs. Both proteins may participate in calcium or mineral binding. Poly-asp runs have been identified in the protein aspein (Tsukamoto et al., 2004) and the Asp-rich family of proteins which derive from prismatic calcite (Gotliv, et al., 2005). Aspein is proposed to bind Ca2+ at high capacity and function as a supply protein in the area of mineralization. The 48 kD has been identified in association with calcium loaded mucous strands secreted by the shell facing mantle epithelium and may function in a similar capacity (Myers et al., 2007).

While the 48 and 55 kD proteins are structurally similar, several distinguishing characteristics were identified. First, the proteins have distinct N-terminal sequences. The 55 kD has proline and serine and possible Q-K-G repeats. The difficulty in sequencing this region is evidence for poly-asp or poly-ser domains and may also reflect a higher degree of post-translational modifications, perhaps in the form of phosphoserine or glycosylation. The N-terminus of the 48 kD was found to have repeat sequences of D-X-X-D, including a sequence D-E-A-D. The D-E-A-D motif has been identified in the Asp-rich proteins from *A. rigida*, and is highly conserved among the seven discrete proteins reported for this family (Gotliv et al., 2005). The authors note that this motif occurs in RNA helicases and is responsible for Mg2+ coordinated ATP hydrolysis. While speculative, they propose a similar role for Asp-rich.
A second distinguishing character is that the 48 kD protein is lightly glycosylated compared to the 55 kD protein, which has little to no detectable carbohydrate. The significance of this can only be speculated, however, “nucleating” proteins histochemically identified in nacre of the *Nautilus* appear in the center of each crystal, where the anionic carboxylate groups form a core surrounded by a ring of sulfate. These authors propose that the sulfate groups bind additional calcium which facilitates selective nucleation of aragonite (Crenshaw and Ristedt, 1976; Addadi et al., 2006). However, as nucleation by SM has not been proven (Wheeler et al., 1991; Mount, 1999, Mount et al., 2004), these same centers could bind nascent crystals. Another possibility is that the carbohydrate imparts a cell recognition/signaling capability to the 48 kD protein. This is particularly interesting given that hemocytes regulate this protein in response to shell notching (shell induction), and it is present in a variety of tissues indicating it may have multiple functions as a cytokine (Johnstone et al., 2007).

Finally, evidence supports that these proteins are processed differently during shell deposition. The 48 kD remains in a soluble form after it is secreted which is supported by its presence in the extrapallial fluid and its occurrence in solubilized matrix (SM). Furthermore, the molecular weight of the 48 kD is consistent in mantle tissue, shell and EPF extracts which suggests that it assumes a mature form, pre-secretion and retains that form during its incorporation into the mineral phase. The 48 kD protein does not appear to undergo proteolytic cleavage. The 55 kD protein is present in the mantle tissue extract but it is less evident in the EPF. This observation, considered with its tight
association with IM, supports that the 55 kD associates with an assembled matrix shortly after it is secreted by the mantle epithelium into the EPF.

Phosphorylation and Matrix Heterogeneity

A common attribute of shell matrix is its heterogeneity, which varies greatly among species and method of extraction. The source of this diversity remains a central question in studies of shell matrix. In this study, differential phosphorylation was found to be a contributing factor to matrix heterogeneity of SM derived from the folia of *C. virginica*. Treatment of SM extract with alkaline phosphatase significantly reduced band number and smearing in IEF and 2D analyses. This was especially evident in the 48 kD region of PAGE gels and partly explains the micro-heterogeneity observed for this protein.

In terms of function, the extent to which matrix protein is phosphorylated has been shown to significantly affect its ability to inhibit crystal growth (Borbas et al., 1991). Similar findings have been reported for vertebrate tooth and bone matrix phosphoproteins including phosphophoryn (PP) (He et al., 2005), and osteopontin (OPN) (Gericke et al., 2005), respectively. It is tempting to suggest that directed removal of phosphate is a mechanism that influences shell microstructure by altering the inhibitory action of the 48 kD protein during shell formation. Recent observations show that the 48 kD protein concentrates on the outer mantle lobe epithelium, both inside the cells and on
their apical surfaces (Johnstone et al., 2007 and unpublished observations). If this protein is de-phosphorylated during mineral formation, a concentration of alkaline phosphatase activity would be localized in the area of mineralization, perhaps in the plasma membrane of mineralizing cells.

**Comparison of SM from calcitic and aragonitic shell microstructures**

The SM from 13 bivalve species was analyzed to investigate if a specific set of matrix proteins correlated with a particular shell structure. A comparison of aragonite and calcite revealed that aragonite SM has more hydrophobic character which was reflected in amino acid analyses as well as staining characteristics following SDS-PAGE. Calcite SM contains more acidic components and has higher phosphate and carbohydrate content than aragonite SM. These data suggest some general matrix characteristics do correlate with mineralogy.

Analyses of component proteins revealed a less definitive correlation of matrix components with mineralogy or microstructure. Protein patterns among nacre and porcelaneous microstructures were especially variable in number, size, and staining characteristics. Less protein variability occurred in calcitic structures including folia and prismatic shell which are dominated by an acidic protein class. Folia SM is distinguished from all other microstructures by a “hallmark” phosphoprotein class that is similar among species but vary in size and primary structures according to Western analyses.
As a whole, these studies reveal that a variety of component matrix proteins erect similar shell microstructures. This finding agrees with studies reporting gene sequences for shell matrix proteins which to date, show only minor instances where primary structure is conserved among proteins derived from similar shell structures (reviewed in Zhang and Zhang, 2006). On the other hand, the secondary structures predicted for some of the deduced protein sequences illustrate that shell proteins, as a class, are highly modular, repetitive and have conserved of secondary structure motifs such as crystal/ion binding (random coil, B-sheet) hinge regions (glycine loops), for flexibility, and hydrophobic regions (B-sheets) for protein assembly. The proteins are distinguished by the number and length of these domains and may a factor that influences shell structure. For example, calcite mineral have especially acidic proteins with long anionic domains. This structure may facilitate interactions with lattice ions or mineral surfaces that is optimal for the formation of the relatively long crystals of the prismatic and folia microstructures (Checa, 2005; Checa, 2007).

It was concluded that primary structure may not be the determining factor of shell mineralogy or microstructure, rather conserved secondary and higher level structure may be more important for crystal-protein interactions that promote the formation of a specific shell structure. This contradicts the idea that SM proteins promote crystal mineralogy and orientation through ordered nucleation.
Comparison of shell layers

For three species, protein patterns of distinct shell layers were analyzed by SDS-PAGE and Western blot. Interestingly, most of the major protein bands between shell layers, irrespective of mineralogy or microstructure, were similar. The major difference was the abundance or the amount of phosphorylation and glycosylation of common proteins. This finding suggests that the levels of secreted matrix proteins and their post-translational processing influence the microstructure and mineralogy of the layer being deposited.

In *C. virginica*, the 48 and 55 kD are present in both the prismatic and foliated shell layers but their levels are regulated. The 48 kD protein appears to have two temporal spikes of secretion during layer formation. Its first appearance occurs on newly secreted periostracum and is therefore presumed important for the *onset* of prismatic layer formation. Its second and most prolific secretion occurs during folia formation. The lowest level of the 48 kD protein occurs in the prismatic layer and suggests its secretion is down regulated during formation of this layer, or it may be regionally secreted from the mantle epithelium. The 55 kD is secreted mostly during formation of the folia layer and becomes associated with the IM. Assuming that IM assembly precedes mineral deposition, this protein is likely secreted before the 48 kD. Also significant is the possibility that the 48 kD occurs as isoforms between shell layers which
was indicated by slight shifts in migration is SDS-PAGE. IEF and 2D analysis supports that these shifts may be attributed to differential phosphorylation.

In *P. fucata* and *A. rigida*, the protein pattern of SM from the outer prismatic calcite layer was similar to that of the inner nacre layer, despite having distinct mineralogy and microstructure. However, compared to nacre, calcite components stained more intensely with Stains-all suggesting that the proteins are more abundant or are more extensively phosphorylated/glycosylated in this layer. Differential phosphorylation and glycosylation of shell proteins may be a controlling factor in mineralization as discussed above.

In addition to regulated levels of similar proteins, some layer specific proteins were visible in the three species and support that the timed release of these proteins influence shell structure. In oyster, a 78 kD protein appeared specific to the folia layer and a 12 kD protein distinguished the periostracal layer. This protein is particularly interesting because it is highly acidic and appears to be secreted exclusively at the onset of prismatic layer formation and is not a major matrix constituent of either mineral layer. This protein may be similar to the nacre derived protein p10 which has mineralizing activity and induces cell differentiation in osteoblastic (bone) cells (Zhang et al., 2006). Some layer specific proteins were also identified in *P. fucata* and *A. rigida*. 
Differences in matrix components from shell layers have been reported previously, and are often interpreted as a primary factor in determining the mineralogy and microstructure of the layer from which it is derived (Mann, 2001). Support for this hypothesis is demonstrated in studies where matrix derived from the nacre and prismatic calcite shell layers promoted the formation of crystals having the corresponding mineralogy and shape (Falini et al. 1996; Belcher et al., 1996). However, this study shows that within a single species, SM components from distinct shell layers are more similar than they are different. Thus, similar proteins can erect shell layers with different mineralogy and microstructure.

The observation that the levels of similar these proteins change during shell layer formation offers a novel mechanism by which shell structure may be determined- that is regulated levels of particular proteins influence crystal growth. Support for this idea comes from studies that show the shape and texture of crystals grown in solution are altered by the addition of shell protein in a dose dependant manner (Marin et al., 2005).

The variety of matrix components that occur in identical microstructures among species and the finding that shell layers from a single species have nearly identical protein patterns despite having different shell structure calls into question conventional thinking that shell structure is determined by an epitaxial interaction between mineral and protein (Addadi and Weiner, 1985). This mechanism requires that there be a stereochemical match between the primary structure of the protein and the crystal lattice,
however, such stringency does not agree with the variety of proteins and structures that occur in shell SM.

Regulation of shell matrix protein and cells associated with shell formation

A significant conclusion drawn from this study is that the regulation of matrix proteins is an important mechanism in determining shell structure, maybe even more so than their individual chemical characteristics. With the growing number of identified matrix gene isolates, studies have shifted to understanding the regulation of shell formation as a cell driven process, mediated by the concerted actions of the mantle epithelium, immune system hemocytes, and shell matrix proteins.

A histological survey of the molluscan mantle reports that specialized epithelial cells secrete acidic matrix components near the mantle edge where new shell growth occurs (Myers et al., 2007). Gene expression patterns for some of these proteins show that they are secreted from discrete mantle zones which are presumed to correlate with the formation of specific shell layers (Jolly et al., 2004; Jackson et al., 2006; Sudo et al., 1997). Clearly, the release of matrix by the mantle epithelium is a highly orchestrated event that is temporally and spatially controlled by signaling pathways not yet discovered.
Hemocyte cells have only recently been identified as a fundamental component of shell formation. Crystal-bearing hemocytes, identified in oyster hemolymph following shell notching (shell induction), were shown to deposit nascent calcite crystals onto newly forming shell and is evidence that mineralization is initiated intra-cellularly (Mount et al., 2004). It stands to reason that mineralogy (calcite or aragonite) may be determined by hemocytes. The signals that cause hemocytes to form crystals and direct them to the area of shell deposition (chemokine) are currently unknown. It is interesting to speculate that mineralogy and microstructure is chemically cued. A promising area to look for the signals that regulate cell activities is the periostracal sheet and secondary organic “membrane” that support the mineral layers. Presumably, the former contains the signals for erecting prismatic calcite while the latter directs the formation of folia.

Recent evidence shows that shell SM, including the 48 kD protein, is expressed and regulated by hemocytes in response to shell notching (shell formation induction) (Johnstone et al. 2007). Furthermore, SM co-localizes with collagen inside hemocytes and is extruded as a complex from induced hemocytes (Patel, 2004); the 48 kD protein is one of the SM proteins released (Johnstone et al., 2007). Mount (2004) showed that hemocytes associate with a fibrous protein network on newly formed shell which has been preliminarily identified as Type 1 collagen (Johnstone unpublished data). These studies provide compelling evidence for direct involvement of the immune system in shell formation. More specifically, shell repair appears to elicit a cellular response
similar to those that occur during soft tissue repair (Franchini and Ottaviani, 2000). It is likely that biochemical signals associated with these processes are similar.

Other studies support that SM regulates cellular processes related to shell formation. Shell SM proteins derived from the nacre of two species, *Pinetada maxima* and *Pinetada fucata*, have been shown to have regulatory activities on mineralogenic cells (Almeida et al., 2000; 2001; Sud et al., 2001; Muries et al., 2002; Zhang et al., 2006). A gene isolate called PFMG1 was recently identified in mantle of *P. fucata* (Liu et al., 2007). This gene was highly expressed in the outer mantle epithelium and the expressed product was shown to affect crystal growth in solution in a manner similar to nacre SM. Expression was also identified in the adductor muscle. But more interestingly, this gene product upregulated expression of genes related to cell differentiation and growth in osteogenic cells and gene markers associated with mineralization including BMP-2, osterix and osteopontin. In oyster, the regulated occurrence of the 48 kD protein in hemocytes in response to shell repair, and its occurrence in blood serum and the epithelia of a variety of organ systems including the mantle, gill, heart, adductor muscle and vessel linings, supports that it is multi-functional (Johnstone et al., 2007) and should be evaluated further as a potential regulator of cell activity, particularly those associated with mineralization.

While the ability of mineral derived matrix to regulate cellular behavior is a relatively new observation in studies of shell formation, matrix proteins involved in
vertebrate tooth and bone formation are well known for their dual roles as mineral controllers and cytokines. A major matrix phosphoprotein of bone and teeth, osteopontin (OPN) acts to organize cellular activity during mineralization through RGD domains (Sodek et al., 2002; Bornstein, 2000). During enamel formation, amelogenin proteins are disrupted from the mineral surface by proteolytic cleavage and the resulting products have cell signaling activities (Robinson et al., 1995; Veis, 2003). During dentin formation, phosphophoryn (PP) has been shown to promote osteoblast maturation (Jadlowiec et al., 2004). OPN and PP in particular, are structurally and functionally comparable to oyster SM in terms of their roles in mineralization (Rusenko, 1988).

OPN may be the best model for exploring the role of oyster SM in regulating cellular activities related to shell formation, particularly those of the immune system. The occurrence of the 48 kD SM protein in a variety of organ systems including the mantle, gill, adductor muscle and blood serum, and its regulation by hemocytes suggests that it is multi-functional (Johnstone et al., 2007). Similarly, OPN is present in the epithelium of a wide range of tissues (Brown, 1992), and functions broadly within immune system processes. It is secreted by activated macrophages, T-lymphocytes and leukocytes, and it circulates in blood serum. OPN levels are elevated in most tissues during wound healing and inflammation (Giachelli, 2000; Gravallese, 2003). It binds to an array of integrins stimulating cell adhesion, migration (as a chemokine), signaling and survival mechanisms for a variety of mesenchymal, epithelial, and inflammatory cells (Zhu et al, 2004; Gravallese, 2003; Giachelli and Steitz, 2000; Giachelli et al., 1998;
OPN also acts as an anti-mineralizing agent to prevent ectopic calcification of soft tissues by physically inhibiting the growth of hydroxyapatite (Giachelli, 2001) and by regulating a cellular response that promotes the dissolution of ectopic deposits (Steitz et al, 2002).

The 48 kD protein is a good candidate for study as a potential regulator of cells associated with mineralization because of its presence in hemolymph serum, its association with hemocytes and its potential to exist as isoforms, distinguished in part by degrees of phosphorylation. SM proteins that are present at the growing edge, including the 48 and 10 kD phosphoproteins, are also good candidates for study and may act to recruit (chemokine), organize (cell adhesion) or promote cell differentiation, in addition to their presumed roles of controlling mineral growth. Recently, attention has been paid to the small SM peptides (less than 1 kD) that occur in nacre SM. These peptides are reported to contribute 60% of the total SM protein and proposed to act as cell signals (Bédouet et al., 2006). This study reveals a similar protein class occurs in oyster folia SM.

Finally, an effort should be made to identify proteins known to act in tissue repair and cell differentiation processes within the SM will help understand how these processes relate to shell formation. For example, regulatory proteins to look for may include TGF-β, and stimulating factors such as bone morphogenetic proteins (BMP’s), all of which would be highly conserved. In fact, cDNA’s encoding BMP have been isolated from the
mantle of *C. gigas* (Myamoto et al., 2002). Evidence for such proteins is the occurrence of similar proteins in shell SM, serum and mantle extract (shown here and reported by Johnstone et al., 2007). In addition, the occurrence of a multitude of minor components and prominent proteins not belonging to the acidic class in shell SM, some of which appear to be conserved among the species, may represent cell regulatory proteins.

To conclude, the collective findings of this and recent studies underscore the importance of understanding the regulation of matrix proteins and cells associated with shell formation, including the hemocytes and those of the mantle epithelium. Identifying the signaling pathways that control these cells and their products are as important to understanding shell formation as the matrix proteins, themselves, and they may well be the driving force behind the type of mineralogy and microstructure deposited during shell layer formation. An interesting prospect, stemming from the emerging parallels between shell, bone and tooth formation, is that some of these pathways are common among invertebrates and vertebrates. If demonstrated, the mineralization strategies utilized by these phyla are more conserved than previously thought. But most inspiring is that shell formation, and possibly biomineralization in animals, apparently represents a merging of two fundamental processes, mineralization and immunity. Continued research along these lines will undoubtedly reveal how the formation of hard parts evolved within the animal kingdom.
APPENDICES
### Appendix A

**Figure A  Summary of Protein Bands Identified in SM from Selected Bivalve Species**

<table>
<thead>
<tr>
<th>Species</th>
<th>MW</th>
<th>SA</th>
<th>CE</th>
<th>AgNO₃</th>
<th>DIG</th>
<th>Flour</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. virginica</td>
<td>105</td>
<td>◊</td>
<td>◊</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>72</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>64</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. gigas</td>
<td>69</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O. edulis</td>
<td>66</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. gibbus</td>
<td>103</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>67</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. simplex</td>
<td>208</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>172</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>77</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>63</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30(double)</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13(double)</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>◊</td>
<td>◊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.3</td>
<td>◊</td>
<td>◊</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>MW</td>
<td>SA</td>
<td>CB</td>
<td>AgNO₃</td>
<td>DIG</td>
<td>Flour</td>
<td>IR</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>----</td>
</tr>
<tr>
<td><em>A.</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>irradians</em></td>
<td>113</td>
<td>●</td>
<td>-stein</td>
<td>●</td>
<td>-</td>
<td>●</td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>91</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>76</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>○</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>41</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>15(doublet)</td>
<td>○</td>
<td></td>
<td>●</td>
<td>+</td>
<td>●</td>
<td>○</td>
</tr>
<tr>
<td><em>P.</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>magellanica</em></td>
<td>114</td>
<td>●</td>
<td>-stein</td>
<td>●</td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>77</td>
<td>○</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>72</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td><em>P.</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>fuscata</em></td>
<td>115</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>○</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>●</td>
<td>-stein</td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>○</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td><em>A.</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>rigida</em></td>
<td>133</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>59</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td><em>M.</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>salutis</em></td>
<td>200+</td>
<td>●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>117</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>99</td>
<td>●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>71</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>66</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>57</td>
<td>○</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>53</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>○</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td>○</td>
</tr>
<tr>
<td>Species</td>
<td>MW</td>
<td>SA</td>
<td>CB</td>
<td>AgNO₃</td>
<td>DIG</td>
<td>Flour</td>
<td>IR</td>
</tr>
<tr>
<td>------------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>----</td>
</tr>
<tr>
<td><em>M. edulis</em></td>
<td>200+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adults</td>
<td>117</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>●</td>
<td>○</td>
<td>+</td>
<td>○</td>
</tr>
<tr>
<td>71</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>+</td>
<td>○</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>-</td>
<td>○</td>
<td></td>
</tr>
<tr>
<td>21 (triplet)</td>
<td>●</td>
<td>○</td>
<td></td>
<td></td>
<td>-</td>
<td>○</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>●</td>
<td>○</td>
<td>●</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.5</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>-</td>
<td>○</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. mercenaria</em></td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>●</td>
<td>+</td>
</tr>
<tr>
<td>57</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>D. robustum</em></td>
<td>139</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. complanata</em></td>
<td>99</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>+</td>
<td>○</td>
</tr>
<tr>
<td>59</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Staining intensities are defined by pie symbols where ○ is light, ● is medium, and ●● is dark. For each stain, intensity symbols were assigned by comparing staining intensities of bands within a single sample and generally not across species.

SA = Stains-all stained bands  
CB = Coomassie blue stained bands  
AgNO$_3$ = Silver stained bands, (−stain) indicates bands that appear translucent  
DIG = protein stain  
Flour = glycoprotein stain, indicated by (+), or stained and (−) indicates no staining  
IR = immuno-reactivity to anti-SM antibody

*Note: B. complanata*, bands at 32, 29, and 26 compose a single broad band.
Appendix B

Figure B  Enzyme linked immunosorbent assay (ELISA) cross-reactivity of the anti-48 kD antibody to SM of selected species.
Anti-48 kD antibody was generated to gel eluted 48 kD phosphoprotein isolated from the SM of C. virginica. Serum containing the antibody was serially diluted and reacted with 10ug/well of SM protein from each species. Myoglobin was included as a negative control. The anti-48 kD generally did not cross-react with other species.
REFERENCES


Degens, ET; Spencer, DW; Parker, RW (1967) Paleobiochemistry of molluscan shell proteins. Comp. Biochem. Physiol. 20:553-579


285
Gibson, CW, Golub E; Ding WD, Shimokawa, H; Young, M; Termine, J; Rosenbloom, J. (1991). Identification of the leucine-rich amelogenin peptide (LRAP) as the translation product of an alternatively spliced transcript. Biochem Biophys Res Commun. Feb 14;174(3):1306-12


Inglis, A.S. 1983 “Cleavage at Aspartic Acid” Meth. Enz. 91:324-332


Keith, J; Stockwell, S; Ball, D; Remillard, K, Kaplan, D; Thannhauser, T; Sherwood, R. (1993) Comparative analysis of macromolecules in mollusk shells. Comp. Biochem. Physiol. 105B:487-496


Neff, J. M. (1972) Ultrastructure of the outer epithelium of the mantle in the clam Mercenaria mercenaria in relation to calcification of the shell. Tissue & Cell. 4: 591-600


Samata, T., (1990) Ca-binding glycoproteins in molluscan shells with different types of ultrastructure. Veliger 33 (2), 190–201


Travis, DF; Francois, CJ; Bonar, LC; Glimcher, MJ. (1967) Comparative studies of the organic matrices of invertebrate mineralized tissues. J. Ultrastructure Research. 18:519-550


Wustman BA, Morse DE, Evans JS (2004) Structural characterization of the N-terminal mineral modification domains from the molluscan crystal-modulating biomineralization proteins, AP7 and AP24. Biopolymers 74: 363–376


Yan Yin, Jing Huang, Michael L. Paine, Vernon N. Reinhold, and N. Dennis Chasteen. (2005) Structural Characterization of the Major Extrapallial Fluid Protein of the Mollusc Mytilus edulis: Implications for Function Biochemistry, 44 (31), 10720-10731


Yutaka Akiyama, Kentaro Onizuka, Tamotsu Noguchi, Makoto Ando (1998) Parallel Protein Information Analysis (PAPIA) system running on a 64-node PC Cluster. Proc. the 9th Genome Informatics Workshop (GIW'98), Universal Academy Press, pp.131-140


