ASPECTS OF THE BIOCHEMISTRY OF THE ORGANIC MATRIX OF EXTANT AND FOSSIL MOLLUSKS

Thesis by

Stephen Weiner

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1977

(Submitted October 18, 1976)

ACKNOWLEDGEMENTS

I am most thankful to Professor Lowenstam and Professor Hood for their tremendous help and encouragement throughout all stages of this project. The combination of Professors Lowenstam's and Hood's individual and overlapping interests in this study, have given me a unique, and well appreciated opportunity to study the biochemistry of mollusk organic matrices. Without their insights, experience and perspective, this project would not have been possible. Above all, I have enjoyed and personally benefitted greatly from our interactions and for this I am truly grateful.

I am also thankful to all the members of Professor Hood's laboratory who have helped and advised me constantly. In particular, I would like to acknowledge the help of Vincent Farnsworth and Bruce Black. Larry Mortin and Barbara Taborek also have worked with me on this project in the last year and have been of great help to me. Dr. Bill Richardson, the University of Alabama, performed the phosphate analysis.

This project has been funded in parts by NSF grant DES 75-08659 to Professor Lowenstam and NSF grant PCM 71-00770 to Professor Hood.

Finally, I would like to thank my wife, Nomi, for her constant support and encouragement throughout these years, and my father, who always encouraged me to pursue an academic career and always stood behind me in all my endeavours.

TABLE OF CONTENTS

Section Page

LIST OF FIGURES

LIST OF TABLES

Table Page

ABSTRACT

The biochemistry of the organic matrix of mollusks is investigated to improve our understanding of the function of this material in shell formation and to explore the possibility of using fossil organic matrix components to study aspects of molecular evolution.

The soluble fractions of the organic matrices of nine species of mollusks representing the three major classes of the phylum have been investigated. However, the organic matrix of the clam, Mercenaria mercenaria, has been studied in greatest detail. It is composed of protein with covalently bound carbohydrate. The protein-carbohydrate linkage is probably through serine and/or threonine residues. The organic matrix components vary greatly in their protein-carbohydrate proportions, amino acid compositions and hydrodynamic size.

The discrete molecular weight components of the nine species of mollusks investigated exhibit great heterogeneity above the species level. A particular repeating amino acid sequence in which every second amino acid is an aspartic acid separated by either glycine or serine, is present in all mollusks examined to date. It is suggested that this amino acid sequence is a potential template for crystal nucleation.

The organic matrix of an 80 million year old clam,

Scabrotrigonia thoracica, still contains the repeating aspartic acid sequence and discrete molecular weight components, indicating that this material is unusually well preserved, and therefore could be of use in future studies of molecular evolution based on material derived from the fossil record. The glycine content of the organic matrices of eight of these fossil shells is a sensitive indicator of early diagenesis. The alloisoleucine/isoleucine ratios of the shell protein ranged from the equilibrium value to extremely low values. The potential uses of fossil organic matrices in evolution studies are discussed.

I. INTRODUCTION

Mollusk shells are composed of bioinorganic and bioorganic constituents. The periostracum and the organic matrix comprise the bioorganic constituents. The periostracum covers the outer surface of the shell, whereas the organic matrix is located within the shell. A study of the biochemistry of the organic matrix was undertaken in order to contribute to our understanding of the function of this material in shell formation. Organic matrices of fossil shells were also studied. As the organic matrix is protected to some degree from the environment by the bioinorganic shell constitutents, it is relatively well preserved in the fossil record, and is thus of great interest in the study of evolution.

The organic matrix and the study of biomineralization

The biomineralization mechanism(s) employed by organisms is, as yet, poorly understood. Observations of the sequence of events occurring during shell formation of mollusks indicate that mantle cells first extrude an extracellular material and only subsequently minerals are observed to form and grow in the proximity of this material (Wada, 1961; Bevelander & Nakahara, 1968). The extracellular material is eventually encapsulated by the growing shell and constitutes the so-called "organic matrix." This sequence of events strongly suggests that the organic

matrix performs some function during shell formation. The extant organic matrix was analyzed in this study by the following approaches:

(a) A partial characterization of the chemical and molecular nature of the organic matrix.

(b) A comparison of aspects of the organic matrix from species within different classes of the phylum Mollusca in order to analyze the common, and possibly functionally important properties of this material.

(c) The development of a method whereby organic matrix components are located on a broken surface of the shell. Application of this technique in the future could provide important information about the function of this material based on the locations of specific components in the shell.

Fossil organic matrices and the study of evolution

A comparison of organic matrix components, especially protein, from fossils of different ages could provide a unique opportunity to study the evolution of these components directly at the molecular level. Changes detected in the organic matrix may be correlated to any morphologic, mineralogic or trace element content changes in the shell which occurred over the same period of time. Thus the study of fossil organic matrices could provide an opportunity to compare classic evolutionary change based on morphology with molecular evolutionary change based on structural differences of the matrix components. The following

approaches to this study were used:

(a) An investigation to determine whether fossil organic matrices (and in particular the protein components) are obtainable in a sufficiently well preserved state to be of use for further study.

(b) An examination of some of the factors which are indicative of the excellent state of preservation of the organic matrix, in order to be able to locate other sources of such material.

Future studies of fossil organic matrices will include the development of suitable techniques to discriminate diagenetic from evolutionary changes and to eventually compare homologous organic matrix components from fossils of different ages.

Many of the approaches to studying both the extant and fossil organic matrices outlined above have not been tried before. In fact very little is understood about the chemistry of mollusk organic matrices, and even less about the molecular organization of this material. However, without this knowledge progress is limited. Thus the major portion of this study is concerned with improving our understanding of the biochemistry of the organic matrix. Many of the experiments are in the formative stages and the sections which follow describe the status of the different aspects of this work, placed in the perspective of published information, if any, on the subject. Appended are papers

which have been published or are in the process of being published on completed aspects of this study.

II. EXTANT MOLLUSK ORGANIC MATRICES

Mollusk organic matrices are composed of two major constituents, protein and carbohydrate (Grégoire et al,1955; Beedham, 1958). Lipid has been detected in some species (Beedham, 1958). The protein components are enriched in acidic and polar amino acids (Grégoire, 1955; Piez, 1961; Hare, 1963) and the carbohydrates are predominantly sulfated mucopolysaccharides (Simkiss, 1965) with minor amounts of neutral (Ravindranath & Rajeswari Ravindranath, 1974), and amino sugars (Degens et al., 1967).

II. A. Previous Work on the Biochemistry of Extant Mollusk Organic Matrices

Two primary areas of investigation of mollusk organic matrices have been undertaken in the past, viz. the study of crystal-organic matrix relationships using electron microscopy and the comparative study of amino acid compositions of the protein constituents of the organic matrix. Information about the molecular organization of this material is confined to only a few relatively recent papers.

Organic matrix-crystal relationships of mollusk shells were first investigated using the electron microscope by Grégoire (1949; 1955). The organic matrix after ethylenediaminetetraacetic acid (EDTA) decalcification was seen to possess characteristic reticulate patterns peculiar

to each of the major mollusk classes (Grégoire, 1957). The organic matrix was observed to surround crystalline units in a situation analogous to cement surrounding bricks in a wall. The existence of organic matrix within the crystalline units is as yet not fully substantiated. Some investigators have observed what appears to be organic matrix within the crystalline units (Watabe, 1963; Travis, 1970) whereas others conclude that the material seen represents, at most, trapped organic matter and water (Towe & Hamilton, 1968; Towe & Thompson, 1972). Transmission and scanning electron microscopy of shell mantle sections have produced diverse interpretations of the crystal-organic matrix relationships during the formation of new shell. Bevelander and Nakahara (1968) observed crystals forming in the pallial fluid and not necessarily associated with the polymerized organic matrix. Erben and Watabe (1974) however, observed crystals forming in close association with the polymerized organic matrix. These observations contrast the two extreme roles postulated for the organic matrix viz. a passive role in which the organic matrix acts as a compartment in which crystal nucleation and growth occurs, or an active role in which the organic matrix itself induces crystal formation (see Towe, 1972 for a review).

Most of the analyses of the chemical composition of the organic matrix reported in the literature are of the amino acid composition of the protein components (Grégoire etal 1955;

Piez, 1961; Hare, 1963). The amino acid compositions of organic matrix proteins are known to vary specifically between species (Grégoire et al, 1955; Hare, 1963; Degens et al., 1967) shell layers and between organic matrices from aragonitic and calcitic shell layers from the same species (Hare, 1963). Variations in the amino acid compositions of organic matrices from the same species living in different ecological environments have also been observed (Hare, 1962; Degens et al., 1967).

The organic matrix of mollusks when decalcified by EDTA or HCL can be separated into two fractions, viz. the soluble and insoluble fractions (Travis et al., 1967; Voss-Foucart et al., 1969; Meenakashi et al., 1971). Differences in amino acid compositions have been reported between the two fractions (Travis et al., 1967; Meenakashi et al., 1971). Watabe (1963) studied thin sections of shells decalcified by EDTA using the scanning electron microscope and concluded that the soluble fraction was located within the crystal units. Crenshaw (1972) arrived at a similar conclusion after finding that destruction of protein by sodium hypochlorite oxidation of powdered shell only destroyed 7% of the soluble fraction, whereas 99% of the insoluble fraction was removed. This indicated that the soluble fraction was protected by the crystals. However, Iwata (1975) noted that organic matrix fixed with a chromium sulfate solution (Sundström, 1968) prior to

decalcification (when examined under the transmission electron microscope) is smooth and free of the reticulate pattern observed by Grégoire etal,1955. He concluded that at least some of the soluble fraction filled the holes formed by the insoluble fraction. This indicated that soluble fraction material is not only located within crystals, but also forms part of the organic matrix surrounding the crystals.

The relationship between the carbohydrate and protein components is not well understood. The identification of both of these components in the soluble fraction of the organic matrix of the clam Mercenaria mercenaria suggests that the soluble fraction is a glycoprotein (Crenshaw, 1972).

Amino acid composition differences between fractions of the organic matrix indicate that more than one protein component is present (Hare, 1963; Degens et al., 1967; Meenakashi et al., 1971). This interpretation is supported by the observation of Degens et al. (1967) that an acid soluble fraction of M. mercenaria when chromatographed on a gel filtration column yielded fractions ranging in molecular weight from 20,000 to 80,000. However polyacrylamide gel electrophoresis of the soluble organic matrix indicates that only one component is present in the fraction (Voss-Foucart, 1968; Crenshaw, 1972).

Calcium has been shown to bind specifically to the soluble organic matrix fraction of M. mercenaria (Crenshaw, 1972). This binding is influenced by the conformation of the protein component. As the characteristics of calcium binding are very similar to that of calcium binding to porcine rib cartilage, Crenshaw (1972) suggests that the calcium is bound to the ester sulfates from two adjacent chains. Recent reviews (Wilbur, 1972; Degens, 1976) of mollusk organic matrices contain more details of published information.

These studies of the organic matrix provide little information about the complexity, organization or function of this material. They do, however, indicate that some properties of the organic matrix are specific at the species level, while others are common to most organic matrices.

II. B. Present Study of the Biochemistry of Extant Mollusk Organic Matrices

This project has attempted to emphasize the study of the structural organization of the organic matrix and how it pertains to shell formation. The strategy adopted involves the extensive characterization of the organic matrix of one particular species of mollusk viz. the clam Mercenaria mercenaria. In cases where a particular property has been well defined in this species, comparative

studies involving nine different molluskan species representing the three major classes have been undertaken. To date only the EDTA soluble fraction of molluskan organic matrices has been studied in detail. Emphasis has been placed on the protein components, as opposed to the carbohydrate, as it is thought that the protein most likely performs important functions in shell formation.

a. The organic matrix of Mercenaria mercenaria

a.1. Preparation of the organic matrix. The procedure for preparing the organic matrix is outlined in Table 1. The preparation procedure yields three separate fractions viz. insoluble, soluble G-25A and soluble G-25B. The weights of each fraction per gram of shell are shown in Table 2. Weighed aliquots of these fractions were hydrolysed in redistilled 6N HCl under vacuum after flushing twice with nitrogen for 20 hours at 108°C. The amino acid compositions were analysed on a Durrum 500 amino acid analyzer (Table 2). The amino acid compositions of each of the three fractions are different indicating that they contain different collections of proteins. The proportions of protein in the fractions range from 97.5% to 2.5%.

The influence of the grain size of the shell fragments being decalcified, and the influence of extended periods of immersion of the decalcified organic matrix in the EDTA solution were examined over a period of seven weeks. These

Table 1 Preparation Procedure of the Organic Matrix*

Physical cleaning of shell (removal of periostracum and hinge areas)

↓

Dipped in 5.8% NH_4OH, washed, sonicated briefly, washed, dried.

↓

Crushed and sieved. Only the fraction larger than 710 μm used.

↓

5 to 10g aliquots dialysed vs 10% EDTA buffered with phosphate to pH 7.0 and containing sodium azide.

↓

Decalcification occurs within 4 to 7 days. Dialysed vs. redistilled water. Lyophilized.

↓

Soluble and insoluble fractions separated by centrifugation at about 3000 r.p.m. for 5 minutes

↓

Pellet, washed 3 times with water, lyophilized. Insoluble Fraction

Supernatent desalted on Sephadex G-25 in water.

↓

Excluded fraction Lyophilized. Soluble G-25A Fraction

Included fraction dialysed exhaustively against water. Lyophilized. Soluble G-25B Fraction

*For more details see Appendix p. 82.

Table 2 Characteristics of the Organic Matrix Fractions of M. mercenaria*

	Insoluble	Soluble	
Mole %		G-25A	G-25B
ASP + ASN	17.90	25.25	16.24
THR	5.19	4.30	5.08
SER	7.26	10.21	17.82
GLU + GLN	7.82	6.66	12.29
PRO	12.43	8.46	7.59
GLY	9.28	12.14	12.88
ALA	6.10	4.68	7.59
CYS	1.64	1.76	trace
VAL	3.04	2.46	3.05
MET	1.80	1.36	0.57
ILE	2.35	2.16	2.03
LEU	3.84	3.40	3.57
TYR	3.65	4.10	2.43
PHE	4.58	3.78	2.81
HIS#			
LYS	6.19	4.98	5.39
ARG	5.10	4.31	2.62
mg organic matrix/g shell	1.97	1.21	0.21
Proportion of + protein in fraction	97.5	70.0	2.5

Histidine was not determined quantitatively, but is usually about 1% of the protein.

* Obtained from shell fragments greater than 710 μm in size and decalcified in less than one week.

+ Calculated from moles of amino acid per unit weight assuming 25% of the "dry" weight to be water.

parameters, unless intentionally monitored, can vary from one preparation to another depending upon the extent of crushing and the amount of shell being decalcified. If a large amount of shell is being decalcified, then organic matrix released in the early stages will remain in the EDTA solution for a long period of time (weeks). Aliquots of ground shell (5g) of grain size either larger or smaller than 710 μm were decalcified by dialysis against EDTA and were then left in the EDTA solution. Every second week a bag of each grain size was removed and the organic matrix prepared as outlined in Table 1. The amino acid compositions of the aliquots removed immediately after decalcification was complete, are compared to those immersed for seven weeks in the EDTA solution (Table 3). Seven weeks of immersion in the EDTA solution significantly affects the amino acid compositions of all three fractions. The insoluble fraction is least affected. The aspartic acid content of the soluble G-25A fractions increases after extended EDTA immersion, whereas the amino acid contents of the rest of the protein remain relatively constant. The soluble G-25B fractions change markedly during the seven week period. The amount of protein in the insoluble and soluble G-25A fractions was significantly reduced after seven weeks. The grain size of the shell fragments decalcified, or otherwise identically treated samples, influences the amino acid composition of the protein.

Table 3 Effect of Shell Grain Size and Extended Periods of Immersion in EDTA on the Amino Acid Composition of M. mercenaria Organic Matrix

Fraction	Insoluble				Soluble G-25A				Soluble G-25B			
Grain Size	< 710 μm		> 710 μm		< 710 μm		> 710 μm		< 710 μm		> 710 μm	
Immersion Period (weeks)	T=0	T=7	T=0	T=7	T=0	T=7	T=0	T=7	T=0	T=7	T=0	T=7
Mole %												
ASP+ASN	16.04	15.58	17.90	18.29	21.84	24.52+	25.25	26.48	18.03	10.08	16.24	10.63
THR	5.73	5.61	5.19	5.20	5.30	4.79	4.30	4.26	4.99	5.48	5.08	5.43
SER	6.75	6.73	7.26	7.27	9.81	10.09	10.21	10.70	16.30	20.21	17.82	25.04
GLU+GLN	7.87	7.84	7.82	8.02	7.46	6.77	6.66	6.91	10.09	14.01	12.29	13.04
PRO	12.72	12.66	12.43	12.59	8.92	8.30	8.46	8.07	5.15	5.27	5.62	4.11
GLY	9.58	10.65	9.28	9.18	11.49	11.93	12.14	12.66	13.79	10.69	12.88	12.48
ALA	6.43	6.30	6.10	6.18	5.26	4.87	4.68	4.15	7.10	8.92	7.59	9.30
CYS	1.81	1.64	1.64	1.72	1.86	1.45	1.76	1.27	tr.	tr.	tr.	tr.
VAL	3.20	3.22	3.04	3.19	2.89	2.61	2.46	2.24	2.63	3.15	3.05	2.98
MET	1.90	1.89	1.80	1.74	1.56	1.37	1.36	1.33	0.47	0.78	0.57	0.56
ILE	2.53	2.55	2.35	2.41	2.26	2.32	2.16	2.29	2.16	2.81	2.03	2.09
LEU	4.22	4.17	3.84	3.75	3.52	3.61	3.40	3.54	3.79	4.49	3.57	3.62
TYR	4.38	4.45	3.65	4.17	4.27	3.93	4.10	3.62	2.42	2.30	2.43	1.97
PHE	5.22	5.24	4.58	4.71	3.76	4.19	3.78	4.06	4.42	3.89	2.81	3.38
HIS*												
LYS	6.56	6.23	6.19	6.13	5.62	5.19	4.98	4.70	6.20	6.23	5.39	3.54
ARG	5.06	5.24	5.10	5.38	4.19	4.06	4.31	4.15	2.47	1.98	2.62	1.81
% protein‡	79.1	76.4	97.4	86.6	73.6	54.1	70.0	49.7	n.d.	n.d.	n.d.	n.d

‡ Calculated from the moles of amino acids per unit weight assuming 25% of the "dry" weight to be water.

* Histidine was not determined quantitatively, but is usually about 1% of the protein.

+ Differences greater than 1% (mole %) are underlined.

Different grain size fractions behave differently with respect to their amino acid compositions during extended periods of immersion in the EDTA solution. During the seven week period a fraction which contains a relatively small proportion of protein which is low in aspartic acid seems to have been mobilized from the soluble G-25A and possibly the insoluble fraction, to the G-25B fraction. Some of the mobilized fraction might have passed through the dialysis bag as well. The physical grinding of the shell fragments to a fine grain size appears to facilitate the movement of this fraction. Thus the grain size of the shell fragments and the period of immersion of the decalcified organic matrix in the EDTA solution change the protein constituents of the different fractions.

In order to reduce the influence of grinding and immersion in the EDTA solution, all preparations of the organic matrix of M. mercenaria studied were obtained from shell fragments larger than 710 μm and were immersed in the EDTA for less than a week.

a.2. Insoluble fraction. This fraction quantitively comprises the major fraction of the organic matrix of M. mercenaria and is composed of close to 100% protein (Table 2). Its insolubility, however, makes it difficult to study and hence very little is known about this fraction.

High ionic strength solutions such as 4M guanidine buffered to pH 5.8 with acetate are capable of at least

partially dissolving the insoluble fraction. Once dissolved the material is soluble in water. Table 4 lists the amino acid compositions of material solubilized from the insoluble fraction after one and seventeen days of extraction in the guanidine buffer. The extract after one day has a very different amino acid composition from the 17 day extract. The amino acid composition of the insoluble fraction remaining after 17 days is surprisingly similar to that of the original insoluble fraction despite the fact that about 30% of the material has been solubilized. It appears that different protein constituents are extracted from those which remain insoluble, indicating that the insoluble fraction is composed of a heterogeneous collection of proteins.

a.3. Soluble G-25A fraction. This fraction comprises the major portion (about 86% by weight) of the soluble organic matrix and is composed of about 70% protein (Table 2.). The properties of this fraction have been investigated in the most detail.

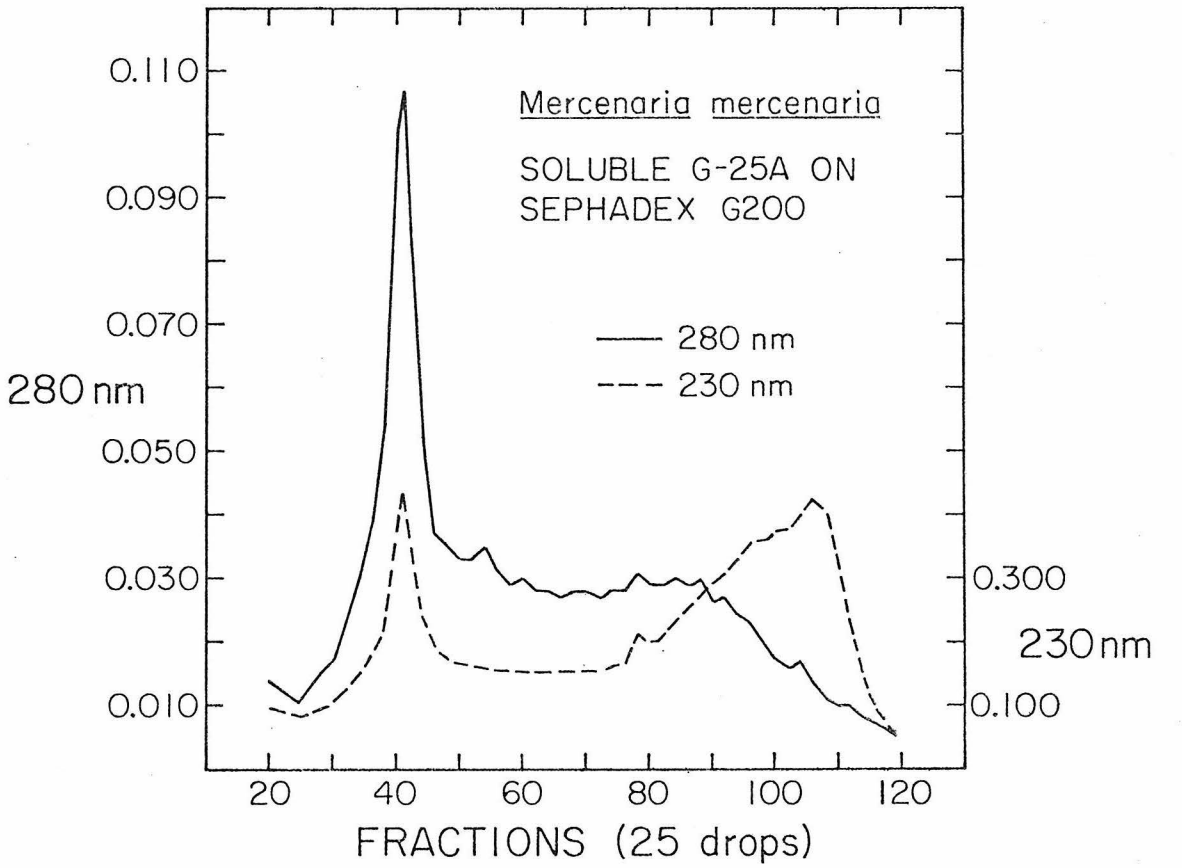
Chromatography on Sephadex G-200. The soluble G-25A fraction was chromatographed on a Sephadex G-200 column (1.5 x 90 cm) in 0.1M NH_4HCO_3 buffer (Fig. 1). As these components contain bound polysaccharides, no direct estimation of their molecular weights can be made. The material excluded on the column is probably greater than 200,000 daltons. The optical density of the eluted

Table 4 Amino Acid Composition of Material Solubilized in 4M Guanidine pH 5.8 (acetate) from M. mercenaria Insoluble Fraction

	Soluble Extracted Material		Remaining Insoluble Material	
Duration of Extraction (days)	T = 1	T = 17	T = 0	T = 17
Mole %				
ASP + ASN	21.55	25.31	16.04	16.98
THR	4.58	4.85	5.73	5.38
SER	10.47	10.37	6.75	6.86
GLU + GLN	8.23	6.74	7.87	7.75
PRO	8.43	9.31	12.72	13.94
GLY	16.82	13.48	9.58	9.32
ALA	6.53	5.68	6.43	6.20
CYS	1.65	1.93	1.81	3.28
VAL	3.41	3.26	3.20	3.85
MET	2.83	1.56	1.90	1.65
ILE	2.16	2.48	2.53	2.52
LEU	4.05	3.89	4.22	3.89
TYR	1.33	0.87	4.38	3.56
PHE	2.73	3.11	5.22	3.75
HIS	0.27	0.46	n.d.	0.53
LYS	2.13	3.18	6.56	6.47
ARG	2.80	3.52	5.06	4.88
Total weight of sample (mg)	n.d.+	n.d.	9.8	6.8

+ n.d. - not determined

Figure 1. The soluble G-25A fraction of M. mercenaria chromatographed on Sephadex G-200.


Mercenaria mercenaria
SOLUBLE G-25A ON
SEPHADEX G200
280 nm
230 nm
280 nm
230 nm
0.110
0.090
0.070
0.050
0.030
0.010
0.300
0.100
20
40
60
80
100
120
FRACTIONS (25 drops)

fractions was measured at 230 nm and 280 nm (Fig. 1). Carbohydrates and peptide bonds absorb relatively strongly at 230 nm, whereas only protein absorbs relatively strongly at 280 nm. The differences in the chromatograms at the above two wavelengths (Fig. 1) indicate that carbohydrate is more abundant among the constituents of smaller hydrodynamic size i.e., those that elute late on the column.* This conclusion was supported by estimating the proportions of protein present in the fractions eluted at different times from the Sephadex G-200 column. The proportion of protein decreases with decreasing hydrodynamic size from 60.9% to 0.61% (Table 5). The amino acid compositions of the fractions are also shown in Table 5. Significant differences in amino acid compositions (especially the aspartic acid content) are present, indicating that the protein constituents vary with increasing elution time on Sephadex G-200. As fraction VI contains a conspicuously different amino acid composition one possibility is that the differences observed in the other fractions might be a function of the amount of fraction VI present in each.

---

*The elution time of a component on a molecular sieve, such as Sephadex, is a function of molecular weight, and binding affinity for Sephadex. Thus the carbohydrate component might be responsible for altering the conformation and binding affinity thus causing a relatively high molecular weight component to elute late on the column.

Table 5 Amino Acid Composition Analyses of Pooled Fractions Obtained from Chromatographing 10 mg. of M. mercenaria Soluble G-25A Organic Matrix on Sephadex G-200

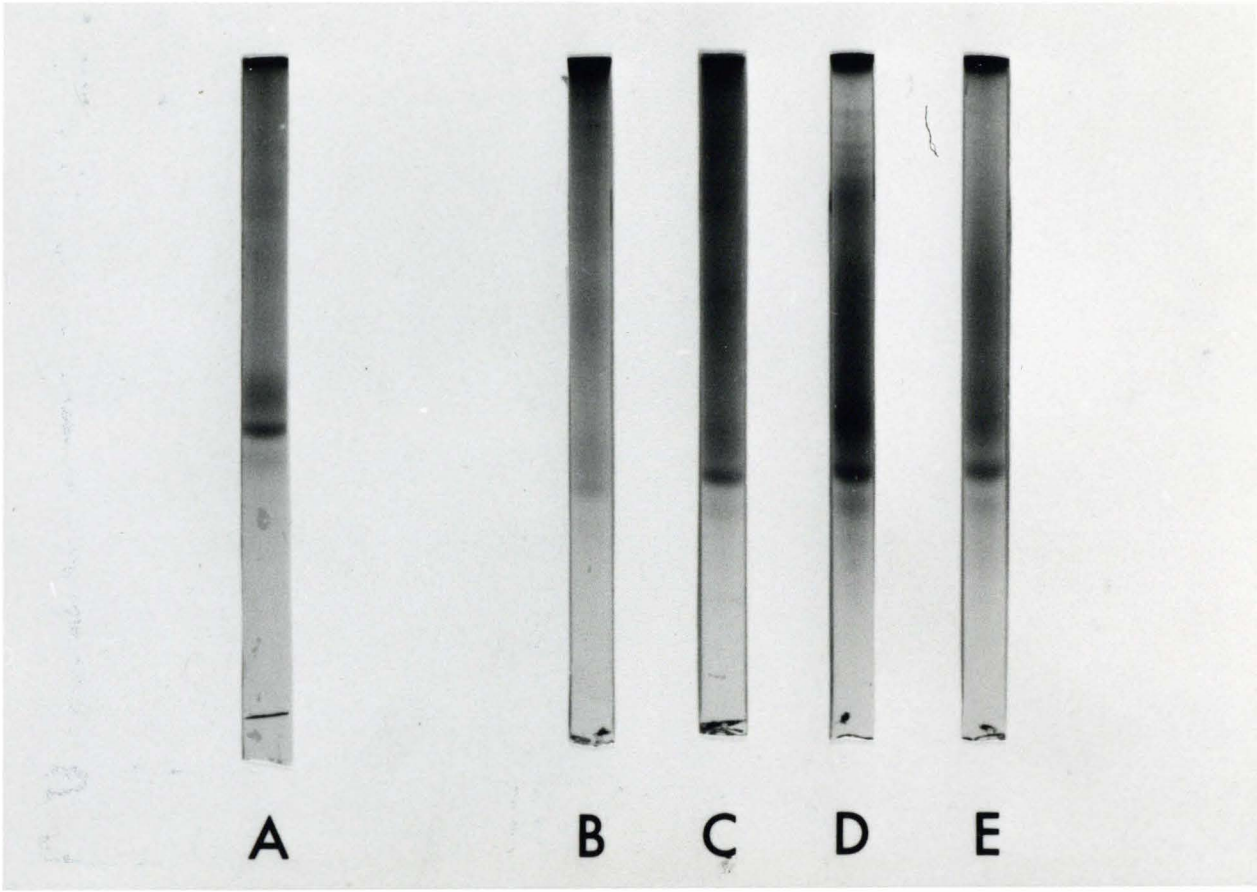
	I	II	III	IV	V	VI
Fractions Pooled (see Fig. 1)	(25-46)	(47-60)	(61-74)	(75-90)	(91-100)	(101-120)
Mole %						
ASP + ASN	23.69	29.51	26.31	21.03	18.61	13.95
THR	4.18	3.51	3.90	5.00	5.75	4.59
SER	9.47	10.84	10.43	10.17	9.26	7.99
GLU + GLN	6.61	5.42	6.43	8.63	8.58	8.66
PRO	12.02	8.90	8.80	7.05	7.41	8.44
GLY	11.54	14.33	14.11	13.36	12.80	16.88
ALA	5.29	4.05	4.27	5.14	5.20	6.81
CYS	2.28	1.91	2.31	3.21	4.32	2.48
VAL	2.99	2.65	2.98	3.76	4.91	6.14
MET	1.06	0.80	1.22	1.28	1.37	1.48
ILE	2.39	2.31	2.46	2.62	2.81	3.63
LEU	3.36	3.51	3.75	4.31	4.10	5.40
TYR	2.69	2.25	2.48	1.75	2.62	2.26
PHE	2.92	2.80	2.86	2.96	3.41	3.56
HIS	0.46	0.34	0.52	1.14	1.51	1.22
LYS	5.10	3.97	4.07	4.70	4.08	3.11
ARG	3.95	2.88	3.08	3.88	3.24	3.40
% Protein+	60.9	44.9	43.4	19.8	3.5	0.61

+ Calculated from the moles of amino acid per unit. weight assuming 25% of the "dry" weight to be water

Discrete molecular weight components. A small proportion of the soluble G-25A fraction is composed of discrete molecular weight components as shown (Fig. 2) by electrophoresis of this fraction on 10% polyacrylamide-urea-SDS gels (Swank & Munkres, 1971). A background of non-discrete molecular weight material is also seen on the gels. Aliquots of the fractions I, II, III+IV, V+VI obtained from the Sephadex G-200 column (see Fig. 1) were electrophoresed on the same gel system (Fig. 2). The major discrete molecular weight components are present in all four fractions, and are especially concentrated in fractions II and III+IV. Thus these discrete molecular weight bands are not chromatographing together on Sephadex G-200, but are associating with other soluble matrix components, and therefore eluting together with these components. The strongly dissociating conditions during electrophoresis allow these components to migrate together on polyacrylamide gels.

A repeating amino acid sequence with aspartic acid present every second residue. The soluble G-25A fraction was subjected to partial hydrolysis under conditions which cleave the protein on both sides of aspartic acid (Partridge & Davis, 1950). The products were analyzed on the Durrum 500 amino acid analyzer and were found to contain significant quantities of glycine and serine in addition to the anticipated aspartic acid, indicating that portions of the amino acid sequence are composed of an (Asp-Y)n type

Figure 2. 10% polyacrylamide-urea-SDS gels of M. mercenaria soluble G-25A fraction and the four pooled fractions of soluble G-25A material chromatographed on Sephadex G-200. A - soluble G-25A fraction; B - Sephadex G-200 fraction I; C - Sephadex G-200 fraction II; D - Sephadex G-200 fractions III+IV; E - Sephadex G-200 fractions V+VI.


A
B
C
D
E

sequence were Y is glycine or serine (see appended paper 1). The(Asp-Y)n type sequence was present in fractions I to IV obtained from the Sephadex G-200 column (Table 6) indicating that this sequence is not confined to one or only a few components in the soluble G-25A matrix, but probably constitutes part of the amino acid sequence of many components. Fractions V and VI are depleted in protein and there is difficulty interpreting the data (see discussion section II.B.a.4. following).

The protein-polysaccharide linkage. Carbohydrate-protein complexes can be classified according to the structure of the linkage region (Marshall, 1972). Thus characterization of the organic matrix linkage region can indicate the category of carbohydrate-protein complex to which the organic matrix belongs. In glycoproteins the carbohydrate linkage occurs through the asparagine of a specific tripeptide, -Asn-x-Ser/Thr (Marshall, 1972). In collagens, the carbohydrate is linked to hydroxylysine present in a Gly-X-Hyl-Gly peptide (Marshall, 1972). In proteoglycans the carbohydrate is linked to the protein through the serine or threonine of a -Ser-Gly- or Thr-Gly-peptide (Isemura and Ikenaka, 1975). These peptides at the linkage site are recognized by specific enzymes (glycosyl transferases) which initiate the linkage. The serine/threonine linkage of proteoglycans is labile at high pH (Anderson et al., 1963). In the presence of sodium

Table 6 Partial Acid Hydrolysis* of Aliquots obtained by Chromatographing M. mercenaria Soluble G-25A Fraction on Sephadex G-200

Sephadex G-200 Fraction	Yield %+			Minimum Proportion‡ $(Asp\text{-}y)_n$ Sequence %
	Asp	Gly	Ser	
I	58.6	20.8	15.6	7.8
II	63.1	24.3	19.9	11.3
III	63.6	22.1	17.9	10.0
IV	56.8	16.2	11.2	6.6
V	62.3	34.4	7.6	not valid‡
VI	55.0	70.4	15.0	not valid

* 48 hours of hydrolysis with 0.25M acetic acid at 108°C.

‡ Calculated from the moles released assuming an Asp-Y-Asp-Y . . . sequence divided by the total moles present in an equivalent hydrolyzed sample of organic matrix.

+ Yields are calculated from the amounts of the particular amino acids present in an equivalent, completely hydrolyzed sample of organic matrix.

‡ See discussion section II. B. a.4.

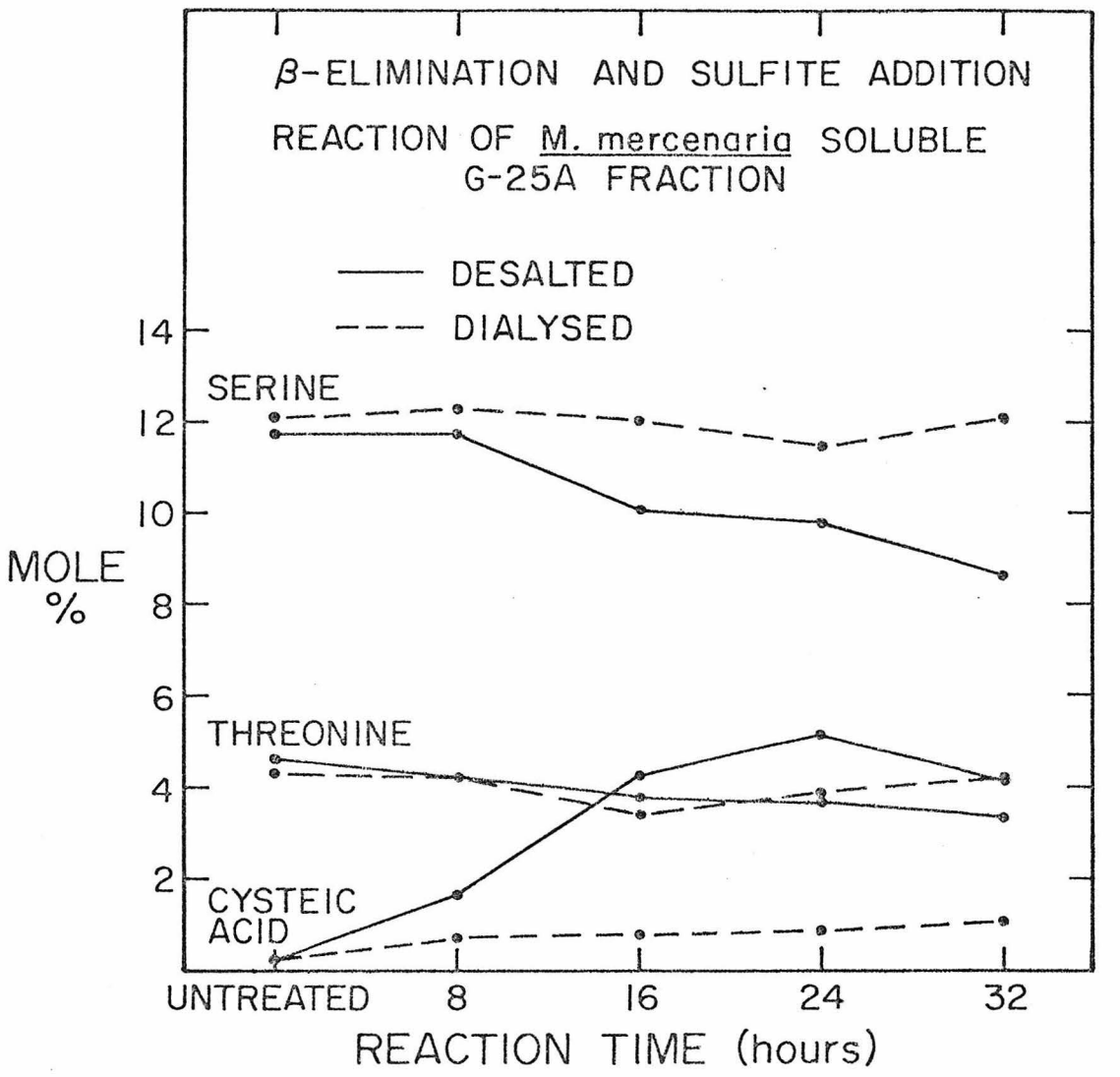
sulfite (0.2M) at pH 11.5, the protein-carbohydrate bond is cleaved by β-elimination with the addition of sulfite to the linked serine or threonine. The linked serine is converted to cysteic acid and the linked threonine to 2-amino-3-sulfonylbutyric acid (Simpson et al., 1972). Phosphate substituted hydroxyamino acids also participate in this reaction. Alkali-catalyzed peptide bond cleavage appears to be negligible (Simpson et al., 1972).

The soluble G-25A fraction of M. mercenaria was subjected to the β-elimination and sulfite addition conditions of Simpson et al. (1972) for 32 hours. The pH was checked every 8 hours and an aliquot was removed. The pH of the removed aliquot was then lowered to 7.0 with 6N HCl, and half the sample was desalted on Sephadex G-25 and half was dialyzed exhaustively against water. The two untreated samples were also desalted or dialyzed prior to hydrolysis. Figure 3 shows the changes in mole % of serine, threonine and cysteic acid of the completely hydrolyzed aliquots. The desalted samples decrease in serine and threonine, and increase in cysteic acid with increasing reaction time, indicating that about 25% of the hydroxyamino acids are substituted by carbohydrate and/or phosphate.* The dialyzed samples show no obvious trends

---

*The phosphate content of the soluble G-25A fraction is only 6 nanomoles/mg. Thus phosphate can account for less than 2.5% of the substituted hydroxyamino groups.

Figure 3. β-elimination and sulfite addition reaction of M. mercenaria soluble G-25A fraction.


β-ELIMINATION AND SULFITE ADDITION
REACTION OF M. mercenaria SOLUBLE G-25A FRACTION
DESALTED
DIALYSED
SERINE
THREONINE
CYSTEIC ACID
MOLE %
14
12
10
8
6
4
2
UNTREATED
8
16
24
32
REACTION TIME (hours)

in serine and threonine, and only a slight increase in cysteic acid with increasing reaction time. The latter might originate from the observed loss of cysteine during the reaction. Thus the protein moiety to which the carbohydrate was linked has passed through the dialysis bag. This fraction is still of sufficiently large hydrodynamic size, however, to be excluded on Sephadex G-25.

An aliquot of M. mercenaria soluble G-25A fraction was treated for 16 hours in 0.2M sodium sulfite at pH 11.5. After acidification to neutral pH the sample was dialyzed against two changes of 200 ml of water. The pooled 400 ml of dialysate was lyophilized and desalted on Sephadex G-25 to recover the components that passed through the dialyzed bag, but excluded on G-25. The remaining contents of the dialyzed bag were exhaustively dialysed against water. Amino acid compositions of the two aliquots were obtained (Table 7). The material that passed through the bag has an unusually low aspartic acid content. Thus the protein containing the major proportion of linked hydroxyamino groups has a different amino acid composition to that which is not linked and is of small hydrodynamic size. It is conceivable that this moiety arises as a result of peptide cleavage during the reaction, even though Simpson et al. (1972) thought that alkali-catalyzed peptide bond cleavage was negligible.

An aliquot of alkali treated M. mercenaria G-25A

Table 7 The Amino Acid Compositions of the Two Different Hydrodynamic Size Fractions of Alkali Cleaved *M. mercenaria* Soluble Organic Matrix

	Non-Dialyzable	Dialyzable, but Excluded on Sephadex G-25
Mole %		
Cysteic Acid	0.49	4.58
ASP + ASN	29.70	16.93
THR	3.87	4.47
SER	12.95	7.24
GLU + GLN	6.24	10.86
PRO	4.75	7.14
GLY	15.18	14.16
ALA	3.57	7.99
CYS	1.44	1.17
VAL	2.57	2.98
MET	0.73	1.28
ILE	2.01	1.92
LEU	2.89	3.73
TYR	3.84	3.41
PHE	2.47	2.88
HIS	0.66	0.96
LYS	3.82	4.26
ARG	2.79	4.05

fraction was electrophoresed on 10% polyacrylamide-urea-SDS gels (Swank & Munkres, 1971) together with an untreated sample (Fig. 4). All the discrete molecular weight bands are still present at approximately the same relative migration distances, indicating that these components do not have serine and/or theorine linked carbohydrate. In addition , no major redistribution of components has occurred as the Sephadex G-200 chromatogram of this material is similar to that of the untreated material.

Thus substituted hydroxyamino groups are present in the soluble G-25A fraction non-discrete molecular weight material. The serines and/or threonines linked to carbohydrate occur predominantly on a protein of small hydrodynamic size and low in aspartic acid content.

a.4. Soluble G-25B Fraction

This fraction is quantitatively the smallest fraction, and characteristically contains only about 2.5% protein (Table 2.). The remaining 97.5% is assumed to be carbohydrate with small amounts of salt that have not been removed by dialysis.

It is separated from the G-25A fraction by its small hydrodynamic size, in that it is included on Sephadex G-25. The amino acid composition of the soluble G-25B fraction is different from the other fractions (Table 2).

Figure 5 shows the elution profile of the soluble G-25B fraction on Sephadex G-200. Surprisingly, some of this

Figure 4. 10% polyacrylamide-urea-SDS gels of M. mercenaria soluble G-25A fraction and alkali treated aliquots of soluble G-25A material chromatographed on Sephadex G-200.

A - soluble G-25A fraction untreated; B - alkali treated material that excludes on Sephadex G-200; C - alkali treated material that includes on Sephadex G-200.

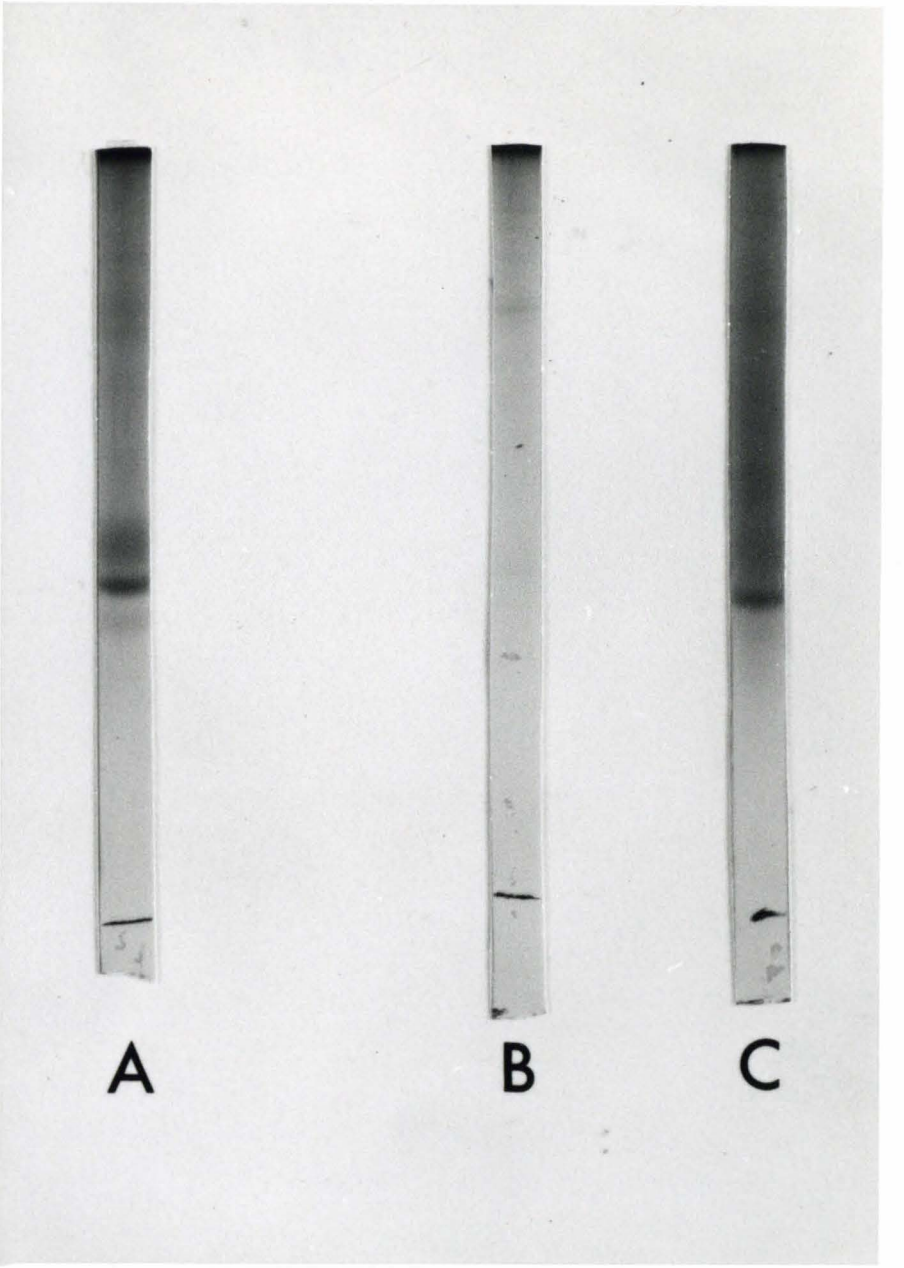
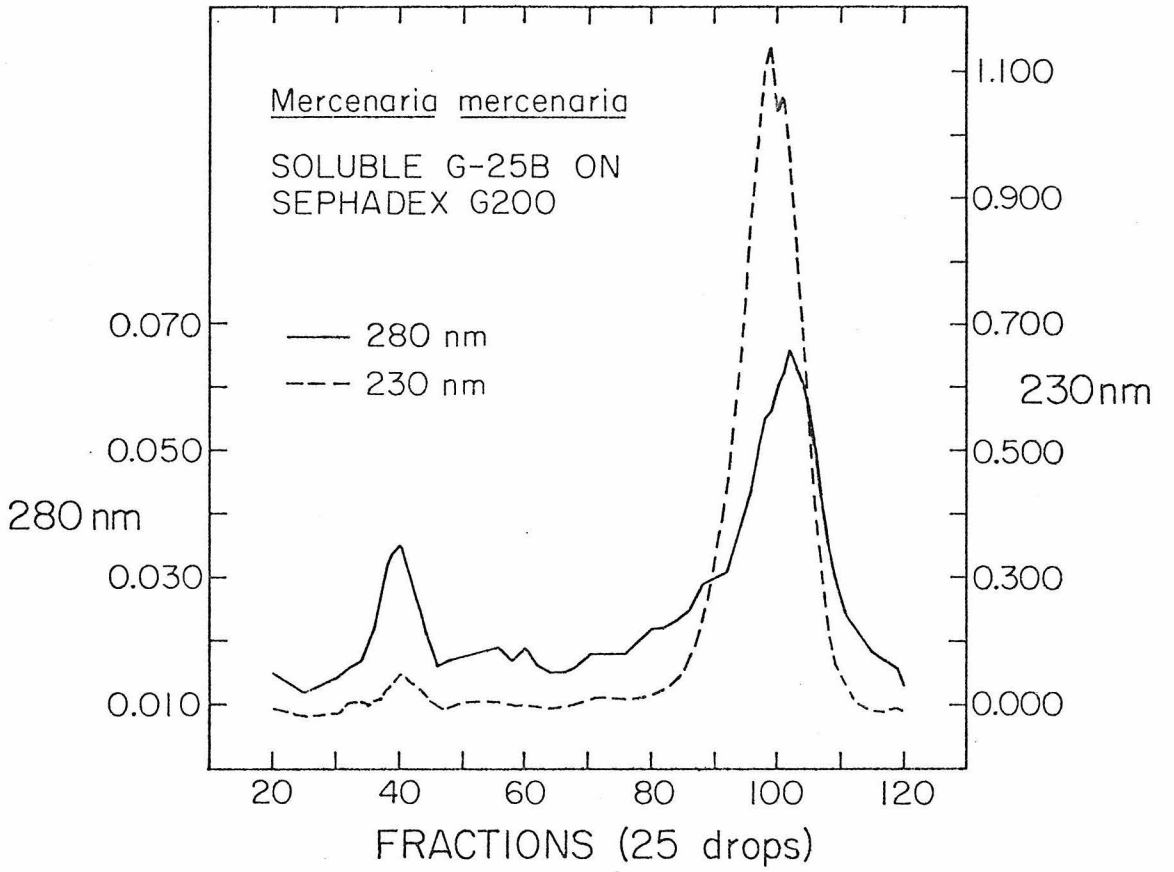

A
B
C

Figure 5. The soluble G-25B fraction of M. mercenaria chromatographed on Sephadex G-200.


Mercenaria mercenaria
SOLUBLE G-25B ON
SEPHADEX G200
280 nm
230 nm
280 nm
230nm
0.070
0.050
0.030
0.010
1.100
0.900
0.700
0.500
0.300
0.000
20
40
60
80
100
120
FRACTIONS (25 drops)

material excludes on Sephadex G-200, whereas it was included on Sephadex G-25. One possibility is that the excluded peak represents reaggregated material. The excluded fraction (the "a" peak) absorbs relatively strongly at 280 nm and contains about 8.1% protein whereas the included material (the "b" peak) absorbs strongly at 230 nm and contains about 1.3% protein (Fig. 5 and Table 8). The amino acid compositions of these two peaks is shown in Table 8. Both compositions are similar, except for the serine contents. It has been noted in other preparations of M. mercenaria G-25B material, that the serine contents can vary as much as 10 to 12 mole %. It is possible that part of the serine is present as a free amino acid, which is easily lost during the preparation procedure. It is of interest to note that the amino acid compositions of fractions V and VI obtained by chromatographing soluble G-25A material through Sephadex G-200 (Table 5) are very similar to the amino acid composition of the soluble G-25B material (Table 2), except for the low serine contents. These two fractions also contain very little protein. Thus the soluble G-25A and G-25B fractions appear to be end members of the same spectrum.

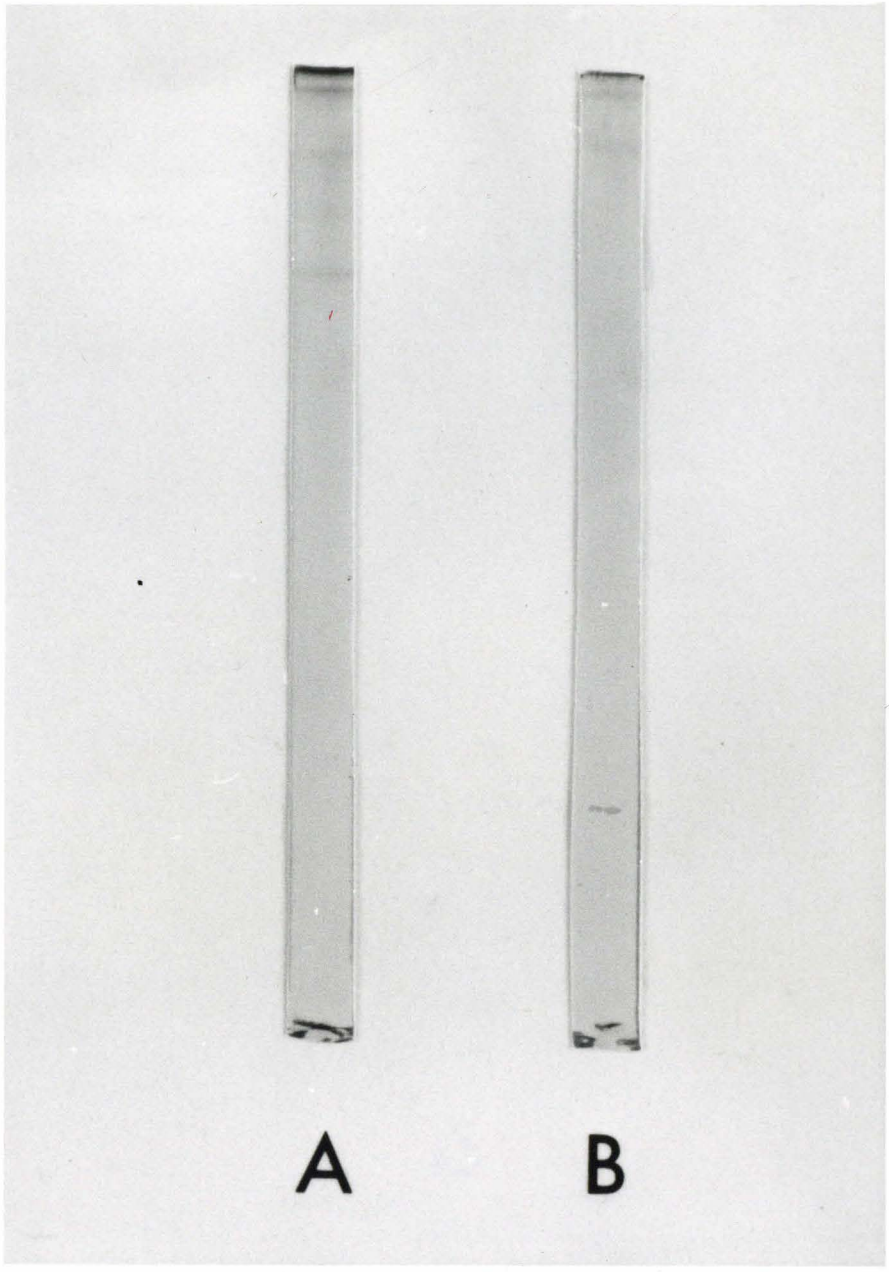
Gel electrophoresis of fractions "a" and "b" from the Sephadex G-200 column on 10% polyacrylamide-urea SDS gel (Swank & Munkres, 1971) shows faint discrete molecular weight bands (Fig. 6).

Table 8 Amino Acid Compositions of Fractions of M. mercenaria Soluble G-25B Chromatographed on Sephadex G-200

Fractions Pooled (see Fig. 5)	G-200a (30-50)	G-200b (70-120)
Mole %		
ASP + ASN	13.53	12.58
THR	5.40	4.50
SER	7.57	17.96
GLU + GLN	12.17	12.70
PRO	5.31	4.68
GLY	15.83	17.52
ALA	9.92	7.63
CYS	n.d.	n.d.
VAL	6.86	3.85
MET	0.95	0.97
ILE	4.39	2.24
LEU	6.33	3.20
TYR	0.68	1.68
PHE	3.48	2.44
HIS	0.33	0.95
LYS	3.57	4.76
ARG	3.70	2.34
% Protein[+]	8.1	1.3

[+] Calculated from moles of amino acid per unit weight assuming 25% of the "dry" weight to be water.

Figure 6. 10% polyacrylamide-urea-SDS gels of M. mercenaria soluble G-25B fraction chromatographed on Sephadex G-200. A - Sephadex G-200 peak "a"; B - Sephadex G-200 peak "b".


A
B

Partial acid hydrolysis of the G-200 "b" fraction results in extremely rapid release (in less than one hour) of large quantities of a number of amino acids other than aspartic acid (Fig. 7). The amount of aspartic acid released, even after 48 hours, is not sufficient to account for the large quantities of amino acids, such as serine, proline, glycine and alanine released (Table 9). These amino acids appear to be present as free amino acids bound to some substrate. Desalting an aliquot of G-200 "b" material on Sephadex G-25 in 0.1M NH_4HCO_2 after it was denatured (6M guanidine pH 5.8 for 24 hours) does not, however, release this "excess" amino acid containing material. It thus seems to be an integral part of the G-200 "b" fraction. The slow release of aspartic acid on the other hand, is comparable to the behavior of the soluble G-25A fraction under the same conditions and implies that the aspartic acid is present as part of a "normal" linear protein. These results indicate that the molecular organization of the soluble G-25B fraction is very unusual. It is composed mostly of carbohydrate with small amounts of peptides and possibly bound amino acids.

a.5. Conclusions

The soluble organic matrix of M. mercenaria contains three categories of components. The quantitatively largest fraction comprises the non-discrete molecular weight

Figure 7. Partial acid hydrolysis of M. mercenaria soluble G-25B fraction.

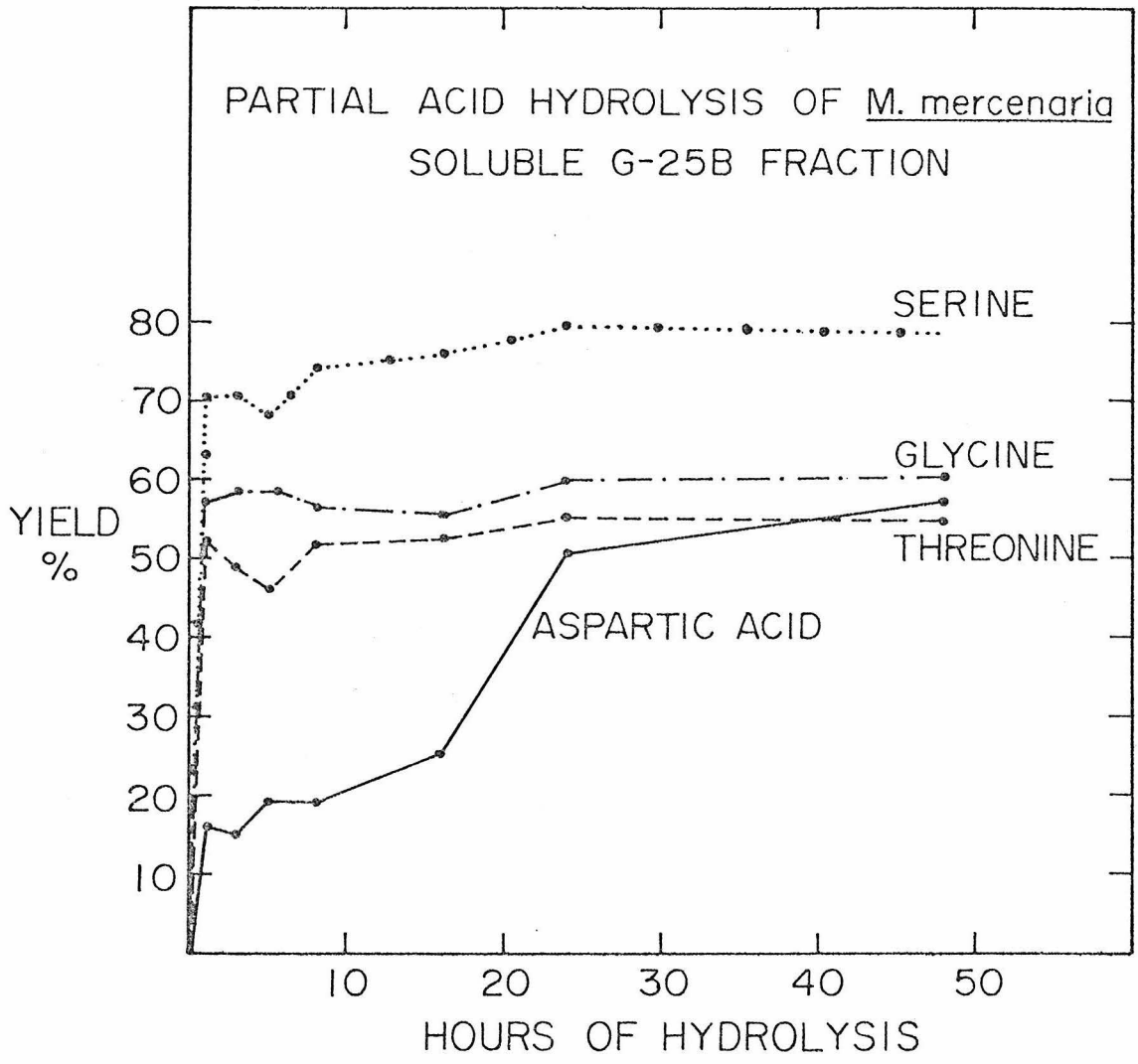

PARTIAL ACID HYDROLYSIS OF M. mercenaria
SOLUBLE G-25B FRACTION
SERINE
GLYCINE
THREONINE
ASPARTIC ACID
YIELD %
80
70
60
50
40
30
20
10
10
20
30
40
50
HOURS OF HYDROLYSIS

Table 9 Partial Acid Hydrolysis* of M. mercenaria G-25B Fraction Chromatographed on Sephadex G-200

		Yield						
		ASP	THR	SER	GLU	PRO	GLY	ALA
	%+	42.7	3.2	9.0	1.3	12.5	21.6	5.9
G-200a								
nanomoles released	‡	100.	2.9	11.8	2.7	11.5	60.0	10.1
	%	57.2	54.5	78.2	2.1	26.0	60.2	54.8
G-200b								
nanomoles released		100.	30.4	176.7	3.4	22.5	172.7	61.4

* 48 hours of hydrolysis with 0.25 M acetic acid at 108°C.

‡ nanomoles released are normalized to 100 nanomoles of aspartic acid released.

+ Yields (%) are calculated from the amounts of the particular amino acids present in an equivalent, completely hydrolyzed sample of soluble organic matrix.

components of the G-25A material, which are high in aspartic acid, substituted hydroxyamino acids, and are composed of a relatively large proportion of protein. The second category comprise components with discrete molecular weights and few, if any, substituted hydroxyamino groups. The third category consist of components with very low amounts of protein, low in aspartic acid and of small hydrodynamic size. Although these groups probably overlap, their differentiation provides a convenient means of categorization for further investigation.

b. Comparative aspects of mollusk organic matrices

Three properties of mollusk organic matrices observed previously in M. mercenaria have been studied in different species representing the major molluskan classes. Two of the properties exhibit great diversity among the species studied, whereas the third shows that some aspects of the molecular nature of mollusk organic matrices, are shared among all the species examined.

b.1. Organic matrix content and the relative proportions of the constituent fractions vary greatly among mollusk shells. Hare and Abelson (1965) report that the organic matrix content of mollusk shells ranges from about 0.01% to 5% of the dry weight of the shell. The amount of organic matrix correlates with the type of shell structure. Furthermore the distribution of the organic matrix between its constituent fractions varies greatly.

Table 10 lists the amounts of insoluble and soluble G-25A fractions obtained by decalcifying different mollusk shells representing the three major classes. The soluble G-25B fraction data are not included as this fraction is present in only very small quantities. Despite the poor reproducibility (see footnote Table 10), it is clear that the proportions of the different fractions vary greatly, both absolutely and relative to each other. Although in almost all cases the insoluble fraction contains more material than the soluble G-25A fraction, the gastropods S. gigas and S. costata contain about ten times more soluble material than insoluble material. Thus, at this level of characterization, mollusks exhibit great diversity with respect to their organic matrix contents, and the relative proportions of organic matrix fractions.

b.2. The presence of discrete molecular weight components of mollusk organic matrices indicates that the organic matrix is heterogeneous. Estimates of the degree of heterogeneity of the soluble fraction of the organic matrix range from one component(Voss-Foucart , 1968; Crenshaw, 1972) to "at least" a few components (Hare, 1963; Degens et al., 1967; Meenakashi et al., 1971). An investigation of the discrete molecular weight components of organic matrices by polyacrylamide gel electrophoresis indicates that numerous discrete molecular weight components are usually present in mollusk organic matrices (see appended paper 2).

Table 10 The Relative Proportions of the Insoluble and Soluble G-25A Fractions in Various Mollusk Organic Matrices (mg/g dry shell weight)+

	Insoluble	Soluble G-25A
Cephalopoda		
Nautilus pompilius	44.0, 42.3, 38.9, 30.1	2.7, 7.7, 2.3, 3.2
Gastropoda		
Strombus gigas	0.02	0.73
Strombus costata	0.08, 0.02, 0.26	0.82, 0.93
Lottia gigantea	20.1	2.9
Bivalvia		
Mercenaria mercenaria	2.04, 1.99, 1.95, 1.88	1.0, 1.21, 1.46, 1.18
Neotrigonia margaritacea	1.02	0.14
Mytilus edulis	8.1	1.7
Mytilus californianus	8.5	2.4
Crassostrea irredescens	4.0, 3.0	0.86, 1.2
Crassostrea gigas	2.0	1.2

+The grain size of fragments decalafied and length of time immersed in EDTA were not controlled (except for *M. mercenaria*). These factors could account for the poor reproducibility observed.

The assemblage of discrete molecular weight components is most similar for species of the same genus. The presence of these components is consistent with the supposition that the soluble organic matrix fraction is capable of performing diverse functions. (NB. The discrete molecular weight components of mollusk organic matrices are the subject of appended paper 2).

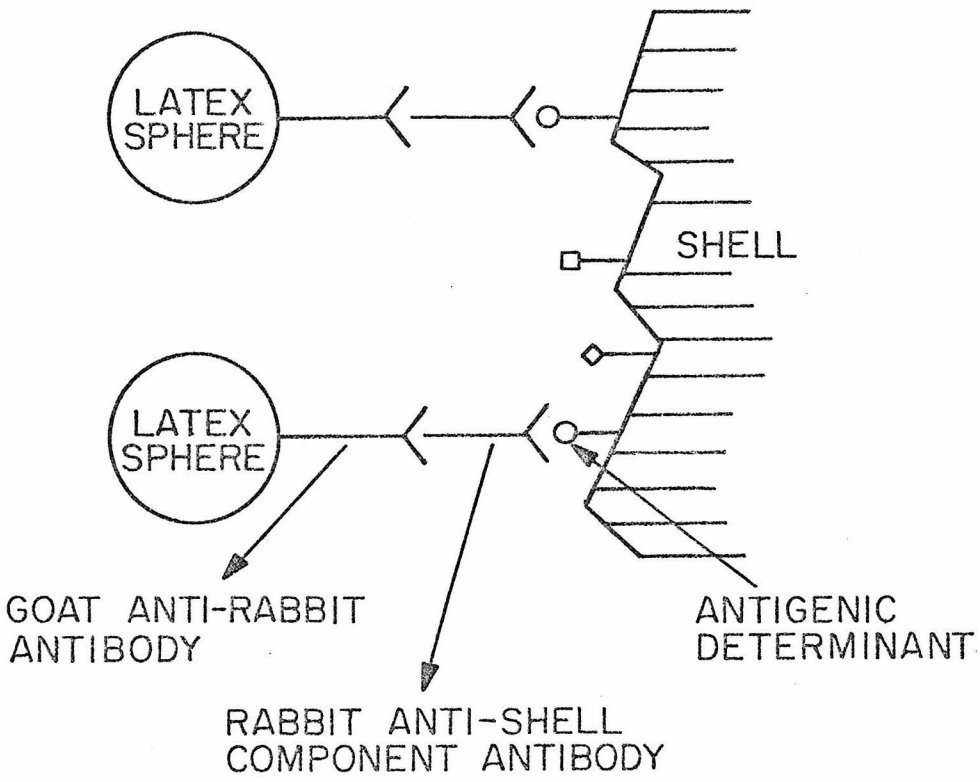
b.3. A regularly repeating sequence of aspartic acid separated by either glycine or serine is present in the protein of the soluble fraction of mollusk organic matrices. Partial acid hydrolysis of proteins cleaves the protein on both sides of aspartic acid (Partridge & Davis, 1950). Application of this technique to studying the soluble fraction of different mollusk organic matrices resulted in not only aspartic acid being produced as a free amino acid, but also glycine and/or serine. This indicates that a portion of the protein is composed of a repeating sequence in which aspartic acid is separated by either glycine or serine (see appended paper 1). The regularly spaced negatively charged aspartic acid might act as a template upon which mineralization of the shell occurs. Thus, in this respect, all mollusk organic matrices examined to date exhibit a common property, viz. a repeating $(Asp\text{-}Y)_n$ in type sequence. (NB. The $(Asp\text{-}Y)_n$ type sequence identified in the soluble protein fraction of mollusk organic matrices is the subject of appended paper 1).

Although few in number, these comparative studies do indicate that some organic matrix properties are sensitive to variations between mollusks at the species or genus level, whereas others are present in all mollusks examined representing the major classes. These observations support the notion that the organic matrix functions in a complex and diverse manner.

c. The distribution of matrix components in the shell could provide an indication of the nature of the function(s) the component performs. A method which maps the location of matrix components on a broken surface of the shell is being developed. The principle of the method involves the use of antibodies generated in a rabbit immunized with a specific shell component. The antibodies are used to locate the component exposed on the broken shell surface. The specific antibodies are then visualized by latex spheres which attach to the "backs" of the specific "shell" antibody via a second antibody bridge (Fig. 8). This method was developed for mapping cell surface components by Molday et al. (1974).

Antibodies have been raised successfully against organic matrix components. Latex spheres with bound goat anti-rabbit antibodies have been shown to bind only specifically to the shell surface. Furthermore latex spheres have been bound through a "shell antibody bridge" to the shell surface, whereas the preimmunization serum

Figure 8. Schematic representation of the binding of latex spheres to shell organic matrix components.


LATEX SPHERE
SHELL
LATEX SPHERE
GOAT ANTI-RABBIT ANTIBODY
ANTIGENIC DETERMINANT
RABBIT ANTI-SHELL COMPONENT ANTIBODY

from the same rabbit only binds a few spheres non-specifically. At present attempts are being made to prepare antibodies which are specific for only one shell component in order to be able to map out the location of that component in the shell. This approach could provide another independent means of determining the type of functions that the organic matrix performs, based on specific distributions of matrix components in the shell itself. This technique could be applied to other calcified tissues such as bones and teeth.

d. Similarities between mollusk organic matrices and calcifying matrices from other phyla might be indicative of a common mechanism of biomineralization. Bones, cartilage and teeth, the major calcifying tissues of vertebrates, have been studied in some detail. A brief comparison of mollusk organic matrix properties with these calcifying matrices follows. The major component of vertebrate bone matrix is collagen, which is not present in the mollusk organic matrix.* The role of collagen in bone formation has, however, in recent years been minimized (e.g. see a review by Leaver et al., 1975) and more attention is being given to the so-called "non-collagenous" components. These non-collagenous components comprise assorted glycoproteins,

---

*Travis et al. (1967) observed collagen in mollusk shells. This observation has not been substantiated by others.

chondroitin sulfate, peptides, albumin, "structural glycoproteins" and lipids (Leaver et al., 1975). A characteristic of the protein moieties of these non-collagenous components is their enrichment in the acidic and polar amino acids such as aspartic acid, glutamic acid, serine, glycine and alanine. These are the same amino acids which are enriched in mollusk organic matrix proteins. Sulfated mucopolysaccharides have been identified in the mollusk organic matrix (Simkiss, 1965) and are also present in bone matrix. The susceptibility of both the mollusk and vertebrate matrices to the β- elimination reaction indicates the presence of carbohydrate and/or phosphate linked to serine and/or threonine residues (Anderson et al., 1963) and section II.B.a.3) in both matrices. One particular phosphoprotein from bovine dentin, identified by Veis and Perry (1967) is known to be present at the calcification front of rat incisors (Weinstock & Leblond, 1972) and to induce the formation of hydroxylapatite during in vitro synthesis of amorphous calcium phosphate (Nawrot et al., 1976). Almost 80% of this phosphoprotein is composed of the amino acids, aspartic acid, serine and phosphoserine. The amino acid sequence of this protein is at present being investigated in collaboration with Arthur Veis in an effort to determine whether or not the repeating aspartic acid sequence present in soluble mollusk organic matrices (section II.B.b.3) is also present in this protein.

This approach has the potential of highlighting those aspects of calcifying tissues which are fundamental to calcifying systems. The onset of calcification is well documented in the fossil record as the worldwide appearance of preserved calcified tissues. This occurred during a long period beginning at the base of the Cambrian (Lowenstam, 1963) about 650 million years ago. The factors responsible for initiating calcification in many different phyla are not understood. An important clue to aid in understanding this problem, is whether most organisms employ basically the same mechanism of calcification, or different organisms use different means of calcifying.

e. Conclusions

Although characterization of the chemical and molecular nature of this material is the basis for any further studies of the organic matrix, it does not necessarily mean that an improvement in our understanding of the functions of the organic matrix will inevitably follow. However, based on our present knowledge of mollusk organic matrices certain conclusions can be made about the nature of the functions that this material might perform. The heterogeneity of all fractions of the organic matrix indicates that this material is at least capable of performing diverse functions. The observed variations show that mollusk organic matrices are different even at the species level. On the other hand, the identification of a

particular repeating amino acid sequence in all mollusk organic matrices examined, shows that these matrices share similar functions. The possibility that this sequence does indeed function as a template for crystal nucleation is an intriguing one, but proof awaits further investigation.

It is hoped that the different approaches to this problem outlined above, will aid in improving our understanding of the function of the organic matrix in shell formation.

III. FOSSIL MOLLUSK ORGANIC MATRICES

The study of evolution as inferred from the fossil record has been primarily based on a comparison of morphologic traits, which are generally interpreted in terms of adaptive strategies. As only the mineralized hard parts of the organism are usually preserved, there are few alternatives to supplement this approach. One possibility is to study the biominerals themselves. Although only limited information is available, trends in the changes of mineralization products utilized by a particular taxonomic group in geologic time have been observed (Lowenstam, 1963; 1974). Another approach is to study the major trace constituents of the biominerals. In the case of carbonate skeletons, definite trends in the strontium and magnesium concentrations relative to calcium do occur (Chave, 1954; Lowenstam, 1954; Dodd, 1967). Some of these trends, e.g. the Sr/Ca ratio in the higher mollusks show a tendency to exclude Sr relative to Ca over geologic time, which is thought to be related to the development of a mechanism of biochemical discrimination by these organism (Lowenstam, 1963; Hallam & Price, 1966; 1968). One facet of the preserved hard parts which has not yet been fully utilized for evolution studies is the organic matrix.

III. A. Previous Work on Fossil Organic Matrices

The unique location of organic matrices, i.e., within the crystalline framework of the inorganic mineral, results in this material being relatively well protected from the environment. For this reason, fossil organic matrices have been studied from many different locations and ages. Abelson (1954) was the first to identify fossil amino acids in shells as old as 25 million years. Abelson (1955) reported that the protein in fossil Merceneria merceneria shells was almost completely degraded by the Miocene, with only free amino acids remaining. Many studies have since reported that despite a drastic drop in concentration of amino acids per unit shell weight, the mole percent amino acid composition of the fossil organic matrix, even as old as the Devonian, still bears a striking resemblance to the closest living relatives (Voss-Foucart & Grégoire, 1971; Akiyama, 1971; Hare & Mitterer, 1967, Matter et al., 1970). Many studies indicate that the amino acids are often still present as polypeptides (Degens & Love, 1965; Grandjean et al., 1964; Jope, 1967; Brieteaux-Grégoire et al., 1968, Akiyama, 1971; Grégoire & Voss-Foucart, 1970; Wyckoff, 1972). Transmission electron micrographs of fossil organic matrices show that characteristic three-dimensional structures can be preserved (Grégoire, 1959; Towe & Urbanek, 1974). A recent immuno-diffusion study provides further evidence for the preservation of

structural similarities between fossil cephalopods and their closest living relatives (de Jong et al., 1974). A book by Wyckoff (1972) contains more details on fossil organic matrices. For the purposes of using fossil organic matrices to study evolution, it was necessary to ascertain whether or not material that is sufficiently well preserved to be of practical use for further examination, existed.

III. B. Present Study of Fossil Organic Matrices

a. An occurrence of unusually well preserved fossil organic matrices (see appended paper 3). The organic matrices of some Upper Cretaceous bivalves (Scabrotrigonia thoracica) from Coon Creek, Tennessee were investigated. Both the inorganic and organic constituents of the fossil shell were compared to shells of an extant representative of the superfamily Trigoniacea, Neotrigonia margaritacea.

Both Trigoniacean shells are 100% aragonite. Their ultrastructures, and Mg and Sr contents are very similar. The oxygen and carbon isotopic composition of the fossil shells indicates that little or no exchange with ground water has occurred. The organic matrices of 8 shells were examined and found to be in various states of preservation (see III.B.b), even though the inorganic shell components indicated more or less uniformly good preservation. The organic matrices of all shells were almost totally soluble in water and contained components which were excluded and included on Sephadex G-25. The best preserved fossil

organic matrix was shown to still contain the original repeating aspartic acid sequence identified in extant mollusks (appended paper 1). In fact the proportions of glycine and serine separating the aspartic acid residues were almost identical to those of the extant N. margaritacea indicating that this sequence has been generally conserved in the last 80 million years. Both the Sephadex G-25 excluded and included fractions contained discrete molecular weight bands when electrophoresed on polyacrylamide gels.

It was therefore concluded that the state of preservation of this material is sufficiently good to warrant further investigation into the possibility of using this material to study certain aspects of evolution.

b. Criteria for evaluating well preserved fossil organic matrices. As the success of any future studies in this field will depend upon the availability of well preserved fossil shells, it is useful to determine parammeters of the shell which are sensitive to even small diagenetic change, but are still easily monitored.

Eight shells of Scabrotrigonia thoracica, all obtained within a few meters of each other at Coon Creek, Tennessee, were examined. The ultrastructure, mineralogy, trace element content and stable isotopic composition of some of the shells, are at present being investigated. The organic matrices of the eight shells have been investigated.

Table 11 lists the amino acid compositions of the soluble G-25A and G-25B fractions and the amount of protein per unit shell weight of each shell. Almost all the shells contained negligible amounts of insoluble organic material. X-ray analysis of some of the insoluble fractions showed that graphite was a major constituent. It should be noted that the amino acid compositions vary considerably between equivalent fractions from the different specimens despite the fact that they were obtained within a few meters of each other. The range of glycine contents, however, is far greater than the fluctuations observed in each of the other amino acids. Figure 9 is a plot of glycine/alanine (alanine is relatively constant in all analyses) against the ratio of the epimers alloisoleucine/isoleucine. In general the alloisoleucine/isoleucine ratio increases as the glycine/alanine ratio increases. A number of the shells have relatively high gly/ala, and negligible amounts of alloisoleucine. The glycine content is apparently a very sensitive indicator of diagenetic change in this group of shells. In fact, these shells, which in many other respects are identical, can be differentiated by their glycine contents. As the glycine content is easily determined on small amounts of sample, it could prove to be a valuable indicator of early diagenetic change.

The soluble G-25B fractions, in seven of eight shells examined, showed lower glycine/alanine and alloiseleucine/

Table 11 Amino Acid Compositions of Upper Cretaceous S. thoracica Organic Matrices

	1		2		3		4	
	G-25A	G-25B	G-25A	G-25B	G-25A	G-25B	G-25A	G-25B
Mole %								
ASP + ASN	11.82	11.77	11.90	9.91	10.57	10.03	9.81	8.93
THR	6.11	6.36	6.13	5.83	5.96	5.73	3.31	4.54
SER	7.00	8.19	6.91	8.56	6.52	7.60	5.46	12.27
GLU + GLN	9.35	9.36	9.98	12.04	9.56	10.24	5.20	11.00
PRO	6.29	5.85	4.78	5.05	8.44	4.30	4.35	3.58
GLY	12.70	11.92	13.83	10.51	12.71	12.78	38.40	26.78
ALA	10.70	10.97	11.48	10.32	11.36	11.89	5.59	8.50
CYS*	0.88	1.24	1.50	1.48	1.57	1.65	tr.	1.08
VAL	5.47	4.90	5.84	5.05	6.07	6.06	3.31	3.39
MET	1.41	1.32	1.43	1.06	1.46	0.99	0.97	0.85
ALLO ILE	0.24	0.07	0.36	0.05	0.34	0.17	2.73	1.23
ILE	3.41	3.22	3.42	3.33	3.37	2.53	2.01	2.27
LEU	6.88	7.38	7.06	8.56	6.64	7.71	3.18	4.39
TYR	2.70	1.90	2.85	2.69	2.69	1.98	1.17	1.65
PHE	3.82	3.29	2.71	3.47	3.04	3.52	2.08	1.31
HIS	1.59	1.83	1.35	1.53	1.57	1.98	3.57	1.42
LYS	4.59	5.56	4.13	5.00	4.16	5.51	4.35	3.66
ARG	5.06	4.90	4.35	5.56	3.94	5.29	4.48	3.16
ALLO ILE/ILE	0.07	0.02	0.10	0.02	0.08	0.07	1.43	0.54
ug protein per g. shell weight	6	6	3	6.5	5.5	5.0	n.d.	n.d.

* The accuracy of cysteine values is very low.

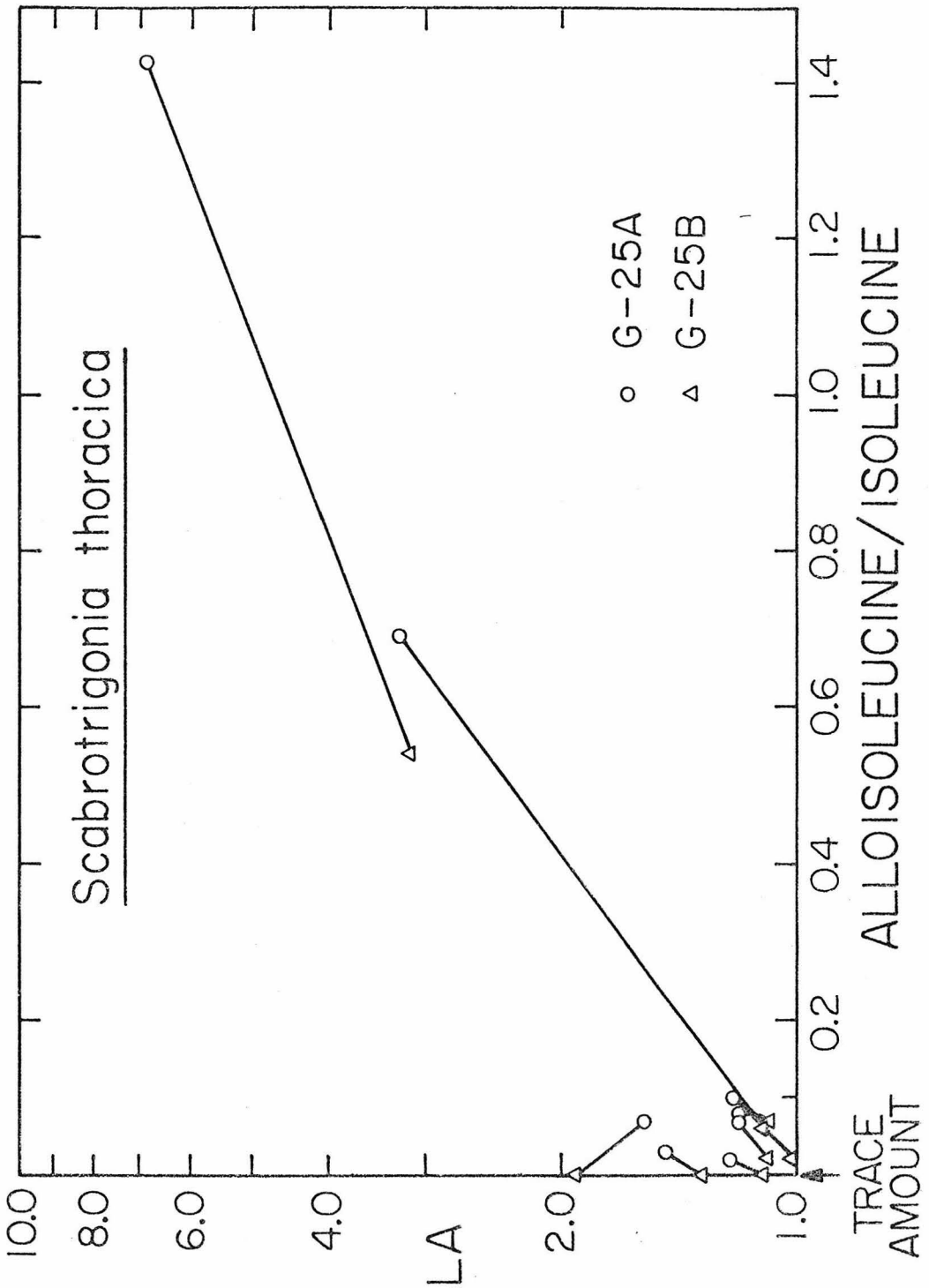
(Table 11--(Cont.)

Table 11--(Cont.)

	5		6		7		8	
	G-25A	G-25B	G-25A	G-25B	G-25A	G-25B	G-25A	G-25B
Mole %								
ASP + ASN	9.83	10.88	10.49	9.53	11.35	9.10	10.88	8.03
THR	3.91	5.67	5.24	5.32	4.75	4.62	6.45	4.79
SER	5.70	5.88	5.94	7.40	6.97	9.34	6.24	15.51
GLU + GLN	10.39	12.89	10.99	10.77	11.20	11.81	9.32	13.77
PRO	4.02	4.69	4.25	4.64	5.31	4.81	4.43	2.54
GLY	29.83	12.33	14.60	13.66	17.40	16.11	19.03	18.65
ALA	9.27	11.19	11.98	12.47	11.82	12.16	12.19	9.63
CYS*	2.79	0.57	tr.	tr.	tr.	tr.	tr.	tr.
VAL	3.80	5.72	5.81	5.62	5.83	5.23	6.60	4.09
MET	1.12	1.34	1.91	1.70	1.54	1.53	1.66	1.00
ALLO ILE	1.56	0.15	0.07	tr.	0.09	tr.	0.25	tr.
ILE	2.23	3.82	3.83	3.40	3.24	3.19	3.52	2.59
LEU	4.02	7.37	7.65	7.06	6.85	6.58	6.90	4.74
TYR	2.91	2.37	2.41	1.96	1.57	1.48	1.56	1.60
PHE	2.12	3.46	3.90	3.40	3.30	2.98	3.83	2.34
HIS	1.34	1.96	1.56	2.30	1.08	1.72	0.81	3.04
LYS	2.57	4.64	4.75	6.43	4.04	5.26	3.02	4.79
ARG	2.57	5.05	4.61	4.34	3.64	4.08	3.32	2.89
ALLO ILE/ILE	0.69	0.06	0.02	Very low	0.03	Very low	0.07	Very low
ug protein per g. shell weight	n.d.	n.d.	4.5	7.5	5.0	6.5	4.5	8.5

* The accuracy of cysteine values is very low.

Figure 9. Glycine/alanine vs. alloisoleucine/isoleucine ratios of the organic matrices of the fossil bivalve, S. thoracica.


Scabrotrigonia thoracica
GLY/ALA
10.0
8.0
6.0
4.0
2.0
1.0
TRACE AMOUNT
0.2
0.4
0.6
0.8
1.0
1.2
1.4
ALLOISOLEUCINE/ISOLEUCINE
G-25A
G-25B

isoleucine ratios than the G-25A fractions from the same shell. Thus if these criteria are indicative of state of preservation, then the G-25B fractions are better preserved than the G-25A fractions.

The G-25B fractions invariably contain more protein per shell weight than the G-25A fractions (Table 11). As only dialyzed material is being analyzed, all these components are of relatively high molecular weight. Material that is non-dialyzable,but is included on Sephadex G-25, must have a small hydrodynamic size or high affinity for Sephadex. This implies that properties which result in the G-25B components of the fossil organic matrices being included in Sephadex are still preserved. Thus the G-25B fraction is selected for by this property and is therefore expected to be relatively well preserved compared to the G-25A fraction which is "diluted" by diagenetically altered material of high molecular weight.

The range of alloisoleucine/isoleucine ratios (from "trace" to 1.4 the equilibrium ratio (Hare & Mitterer, 1967)) in organic matrices from shells of the same species, obtained within a few meters of each other, raises serious questions about the use of amino acid racemization and epimerization for stratigraphic or dating purposes. As equilibration between the epimers, alloisoleucine and isoleucine, generally occurs within only hundreds of thousands of years (Hare & Mitterer, 1969; Bada, 1972) the

very low values obtained in most of the shells studied here is unusual. An anhydrous environment of preservation of these fractions within the shell could account for the almost total absence of epimerization in some of the shells, as laboratory simulation experiments show that water is necessary for the epimerization and racemization of amino acids (Hare & Mitterer, 1969). Thus amino acid epimerization and racemization rates are dependent, not only on time, temperature, pH etc., but on the conditions of preservation of each individual shell.

c. Development of suitable techniques for studying fossil proteins. In order to compare the amino acid sequences of two fossil proteins, it is necessary to be able to first isolate homologous proteins. The heterogeneity of the material combined with a diagenetic overprint, makes this a formidable problem. The use of immunological techniques to isolate serologically homologous components is currently being investigated. This approach has the inherent advantage of selecting not only a collection of specific components, but also components that still have preserved antigenic determinants. It is hoped that a combination of serologic and physical criteria will be sufficient to isolate homologous components.

The comparison of amino acid sequences can be made indirectly using "fingerprint" peptide maps of radioactively labelled components, or directly by using microsequencing

techniques on a sequenator (Silver & Hood, 1975). The former approach has been successfully used on extant M. mercenaria components, whereas the latter approach has not been attempted.

d. Conclusions

Verification of the availability of well preserved fossil shell protein is the first important step towards the eventual utilization of this material for improving our understanding of evolution. It is only in the last few years that techniques suitable for analyzing microgram quantities of protein have been developed. Application of these techniques, together with the combined use of immunological and physical methods of isolating and characterizing the protein components, promises to make the study of fossil proteins an interesting field of research.

IV. GENERAL CONCLUSIONS

The Organic Matrix of *Mercenaria mercenaria*

1. The grain size of the shell fragments decalcified, and the period of immersion of the decalcified organic matrix in the EDTA solution change the protein constituents of the organic matrix fractions.
2. The insoluble fraction is composed of a heterogeneous collection of proteins.
3. The proportion of protein of the soluble G-25A fraction decreases with decreasing hydrodynamic size or increasing affinity to Sephadex from about 60% to 0.5% of the dry weight of the sample.
4. The nature of the protein constituents of the soluble G-25A fraction changes with decreasing hydrodynamic size.
5. Portions of the proteins of relatively large hydrodynamic size comprise a repeating amino acid sequence in which aspartic acid residues are separated by glycine and serine.
6. A small proportion of the soluble G-25A fraction is composed of discrete molecular weight components which are strongly associated with the non-discrete molecular weight components.
7. Substituted hydroxyamino groups are present in the soluble G-25A fraction non-discrete molecular weight material. The linked serines and threonines occur

predominantly on a protein of small hydrodynamic size and low in aspartic acid content.

8. The soluble G-25B fraction is composed mostly of carbohydrate with small amounts of protein, some of which is of discrete molecular weight.

Comparative Aspects of Mollusk Organic Matrices

(representing the major classes)

1. Mollusks exhibit great diversity with respect to their organic matrix contents and the relative proportions of organic matrix fractions.

2. A regularly repeating sequence of aspartic acid separated by either glycine or serine is present in the soluble fraction of all mollusk organic matrices studied.

3. The discrete molecular weight components of the soluble G-25A fractions of mollusks are very diverse above the species level.

Fossil Mollusk Organic Matrices

1. The organic matrix of an Upper Cretaceous bivalve still contains preserved repeating aspartic acid sequence and discrete molecular weight components.

2. The glycine content of fossil organic matrices of Upper Cretaceous bivalves appears to be a very sensitive indicator of early diagenetic change.

3. The states of preservation of the organic matrices of

specimens of upper Cretaceous bivalves show large variations with respect to their amino acid compositions and alloisoleucine/isoleucine ratios despite the fact that they were all located within a few meters of each other.

V. REFERENCES

Abelson, P. H., 1954. Organic constituents of fossils. Carnegie Inst. Wash. Yearbook, 53, 97-101.

Abelson, P. H., 1955. Organic constituents of fossils. Carnegie Inst. Wash. Yearbook, 54, 107-109.

Akiyama, M., 1971. Amino acid composition of fossil scallop shell proteins and non-proteins. Biomineralization, 3, 65-70.

Anderson, B., Hoffman, P. and Meyer, K., 1963. A serine-linked peptide of chondroitin sulfate. Biochim. Biophys. Acta., 74, 309-311.

Bada, J. L., 1972. The dating of fossil bones using the racemization of isoleucine. Earth. Planet. Sci. Letters, 15, 223-231.

Beedham, G. E., 1958. Observations on the non-calcareous component of the shell of the Lamellibranchia. Quat. J. Microscop. Science, 99, Part 3, 341-357.

Bevelander, G. and Nakahara, H., 1968. An electron microscope study of the formation of the nacreous layer in the shell of certain bivalve mollusks. Calc. Tiss. Res., 3, 84-92.

Bricteau-Grégoire, S., Florkin, M. and Grégoire, C., 1968. Prism conchiolin of modern or fossil molluscan shells. Comp. Biochem. Physiol., 24, 567-572.

Chave, K. E., 1954. Aspects of the biogeochemistry of magnesium: 1. Calcareous marine organisms. Jour. Geol., 62, 266-283.

Crenshaw, M. A., 1972. The soluble matrix from Mercenaria mercenaria shell. Biomineralization Research Reports, 6, 6-11.

Degens, E. T. and Love, S., 1965. Comparative studies of amino acids in shell structures of Gyraulus trochiformis Stahl, from the Tertiary of Steinheim, Germany. Nature, Lond., 205, 876-878.

Degens, E. T., 1976. Molecular mechanisms on carbonate, phosphate and silica deposition in the living cell. Topics in Current Chemistry, 64, 1-112.

Degens, E. T., Spencer, D. W. and Parker, R. H., 1967. Paleobiochemistry of molluscan shell proteins. Comp. Biochem. Physiol., 20, 553-579.

de Jong, E. W., Westbroek, P., Westbroek, J. F. and Bruining, J. W., 1974. Preservation of antigenic properties of macromolecules over 70 Myr. Nature, 252, 63-64.

Dodd, J. Robert, 1967. Magnesium and strontium in calcareous skeletons: a review. Jour. of Paleo., 41, 1313-1329.

Erben, H. K. and Watabe, N., 1974. Crystal formation and growth in bivalve nacre. Nature, 248, 128-130.

Grandjean, J., Grégoire, C., and Lutts, A., 1964. Mineral components and remnants of organic structures in shells of fossil molluscs. Bull. Acad. Belg. Cl. Sci., Ser. 5, 50, 562-595.

Grégoire, C., 1957. Topography of the organic components in the mother-of-pearl. Jour. Biophys. Biochem. Cytology, 3, 1-36.

Grégoire, C., 1959. Conchiolin remnants in mother-of pearl from fossil Cephalopoda. Nature, 184, 1157-1158.

Grégoire, C., Duchateau, G. and Florkin, M., 1949. Examen au microscope electronique de la pellicule préracree et de la nacre décalcifee de l'Anodonte. Arch. intern. Physiol., 57, pp. 121.

Grégoire, C., Duchateau, G., and Florkin, M., 1955. La trame protidique des nacres et des perles. Annls. Inst. Oceanogr., 31, 1-36.

Grégoire, C. and Voss-Foucart, M. F., 1970. Proteins in shells of fossil Cephalopods (Nautiloids and Ammonoids) and experimental simulation of their alterations. Arc. Int. Physiol. Bioch., 78, 191-203.

Hallam, A. and Price, N. B., 1966. Strontium contents of recent and fossil aragonitic cephalopod shells. Nature, 212, 25-27.

Hallam, A. and Price, N. B., 1968. Further notes of the strontium contents of unaltered fossil Cephalopod shells. Geol. Mag., 105, 52-55.

Hare, P. E., 1962. The amino acid composition of the organic matrix of some recent and fossil shells of some West Coast species of Mytilus. Ph.D. Dissertation, California Institute of Technology, Pasadena.

Hare, P. E., 1963. Amino acids in the proteins from aragonite and calcite in the shells of Mytilus californianus. Science, 139, 216-217.

Hare, P. E. and Abelson, P. H., 1965. Amino acid composition of some calcified proteins. Carnegie Inst. Wash. Yearbook, 64, 223-232.

Hare, P. E. and Mitterer, R. M., 1967. Non-protein amino acids in fossil shells. Carnegie Inst. Wash. Yearbook, 65, 362-364.

Hare, P. E. and Mitterer, R. M., 1969. Laboratory simulation of amino acid diagenesis in fossils. Carnegie Inst. Wash. Yearbook, 67, 205-208.

Isemura, M. and Ikenaka, T., 1975. β-elimination and sulfite addition reaction of chondroitin sulfate peptidoglycan and the peptide structure of the linkage region. Biochim. Biophys. Acta., 404, 11-21.

Iwata, K., 1975. Ultrastructure of the conchiolin matrices in molluscan nacreous layers. Jour. Fac. Sco.

Hokkaido Univ. Ser. IV, 17, 173-229.

Jope, M., 1967. The protein of Brachipod shell-IV. Shell protein from fossil inarticulates: amino acid composition and disc electrophoresis of fossil articulate shell protein. Comp. Biochem. Physiol., 30, 225-232.

Leaver, A. G., Triffitt, J. T. and Holbrook, I. B., 1975. Newer knowledge of non-collagenous protein in dentin and cortical bone matrix. Clin. Orthog., 110, 269-92.

Lowenstam, H. A., 1954. Factors affecting the calcite-aragonite ratios in carbonate secreting marine organisms. Jour. Geol., 62, 284-322.

Lowenstam, H. A., 1963. Biologic problems relating to the composition and diagenesis of sediments. In: T. W. Donnelly (ed.), The Earth Sciences, pp. 137-195. Univ. Chicago Press.

Lowenstam, H. A., 1974. Impact of life on chemical and physical processes. In: E. D. Goldberg (ed.), The Sea, 5, pp. 715-796. J. Wiley and Sons, New York and London.

Marshall, R. D., 1972. Glycoproteins. Ann. Rev. Biochem., 41, 673-694.

Matter, P., Davidson, F. D. and Wyckoff, R. W. G., 1970. The microstructure and composition of some Pliocene fossils. Comp. Biochem. Physiol., 35, 291-298.

Meenakashi, V. R., Hare, P. E. and Wilbur, K. M., 1971. Amino acids of the organic matrix of Neogastropod shells. Comp. Biochem. Physiol., 40B, 1037-1043.

Molday, R. S., Dreyer, W. J., Rembaum, A. and Yen, S. P. S., 1974. New immunolatex spheres: visual markers of antigens on lymphocytes for scanning electron microscopy. J. Cell. Biol., 64, 75-88.

Nawrot, C. F., Campbell, D. J., Schroeder, J. K. and Van Valkenberg, M., 1976. Dental phosphoprotein-induced formation of hydroxylapalite during in vitro synthesis of amorphous calcium phosphate. Biochemistry, 15, 3445-3449.

Piez, K. A., 1961. Amino acid composition of some calcified proteins. Science, 134, 841-842.

Ravindranath, M. H. and Rajeswari Ravindranath, J. H., 1974. The chemical nature of the shell of molluscs: I. Prismatic and nacreous layers of a bivalve Lamellidans marginalis (Unionidae). Acta Histochem., 48, 26-41.

Silver, J. and Hood, L., 1975. An approach to amino acid sequence analysis of transplantation antigens. Nature, 256, 63.

Simkiss, K., 1965. The organic matrix of the oyster shell. Comp. Biochem. Physiol., 16, 427-435.

Simpson, D. L., Hranisavljevic, J. and Davidson, E. A., 1972. β Elimination and sulfite addition as a means of localization and identification of substituted seryl and threoryl residues in proteins and proteoglycans. Biochemistry, 11, 1849-1856.

Sundström, B., 1968. Histological decalcification using aqueous solutions of basic chromium (III) sulphate. Odont. Revy., 19, 3-19.

Swank, R. T. and Munkres, K. D., 1971. Molecular weight analysis of oligopeptides by electrophoresis in polyacrylamide gel with sodium dedecyl sulfate. Anal. Biochem., 39, 462-477.

Towe, K. M., 1972. Invertebrate shell structure and the organic matrix concept. Biomin. Res. Reports, 4, 2-14.

Towe, K. M. and Hamilton, G. H., 1968. Ultrastructure and inferred calcification of the mature and developing nacre in bivalve mollusks. Calc. Tiss. Res., 1, 306-318.

Towe, K. M. and Thompson, G. R., 1972. The structure of some bivalve shell carbonates prepared by ion-beam thinning. Calc. Tiss. Res., 10, 38-48.

Towe, K. M. and Urbanek, A., 1974. Fossil organic material: a unique fibril ultrastructure in Silurian graptolites. Eighth International Congress on Electron Microscopy, Canberra, II, 694-695.

Travis, D. F., 1970. The comparative ultrastructure and organization of five calcified tissues. In: Biological Calcification (ed. H. Schraer), Appleton Century Crofts, New York, 462 pp.

Travis, D. F., Francois, C. J., Banar, C. L. and Glimcher, M. J., 1967. Comparitive studies at the organic matrices of invertebrate mineralized tissues. J. Ultrastructure Res., 18, 519-550.

Veis, A. and Perry, A., 1967. The phosphoprotein of the dentin matrix. Biochemistry, 6, 2409-2416.

Voss-Foucart, M. F., 1968. Essais de solubilisation et de fractionnement d'une conchioline (nacre murale de Nautilus pompilius (Mollusque, Cephalopode). Comp. Biochem. Physiol., 26, 877-886.

Voss-Foucart, M. F. and Grégoire, C., 1971. Biochemical composition and submicroscopic structure of matrices of nacreous conchiolin in fossil Cephalopods (Nautiloids and Ammonoids). Bull. Inst. Sci. Nat. Belg., 47, 1-39.

Wada, K., 1961. Crystal growth of molluscan shell. Bull. Natl. Pearl. Res. Lab., 7, 703-785.

Watabe, N., 1963. Decalcification of thin sections for electron microscope studies of crystal-matrix relationships in mollusc shells. J. Cell Biology, 18, 701-703.

Weinstock, M. and Leblond, C. P., 1973. Radioautographic visualization of the deposition of a phosphoprotein at the mineralization front in the dentin of the rat incisor. J. Cell. Biol., 56, 838-845.

Wilbur, K. M., 1972. Shell formation in mollusks. Chemical Zoology, 103-145.

Wyckoff, R. W. G., 1972. The Biochemistry of Animal Fossils. Williams and Wilkins Company, Baltimore, 152 pp.

APPENDIX

DETAILED DESCRIPTION OF THE PREPARATION PROCEDURE OF THE ORGANIC MATRIX

Five to 10 g aliquots of the cleaned and crushed shell are placed in 45 mm dialysis bags (obtained from VWR Scientific division of Univar). 20-40 ml of the buffered EDTA dialysate is used to suspend the ground shell in the bag. Each bag is dialyzed against about 1 l of 10% (w/v) EDTA buffered to pH 7.0 with sodium phosphate (0.05 - 0.1 M) containing about 0.01% (w/v) sodium azide. The dialysate is stirred at room temperature and the contents of the bag are agitated manually once a day. Decalcification occurs within 4 to 7 days. The salt is partially removed by dialyzing against three changes, each of 6 l, of distilled deionized water. The contents of the bag are lyophilized, resuspended in 12 ml of distilled deionized water and centrifuged at 3170 g (3000 rpm on a clinical centrifuge) for about 5 min. The supernatant is lyophilized and the pellet is washed 3 times with distilled deionized water. The lyophilized supernatant is desalted on a Sephadex G-25 column (30 cm x 1.25 cm I.D.) in water. 50 drop (3.4 ml) fractions are collected and their absorbtion at 280 nm measured. The excluded and included fractions are pooled separately and lyophilized. The remaining salt in the included fraction is removed by dialyzing (3-4 changes of 6 l each) against distilled deionized water and then lyophilized.

VI. APPENDED PAPER 1

Soluble Protein of the Organic Matrix of Mollusk Shells: A Potential Template for Shell Formation.

Published in Science, 190, 987-989, 1975.

Authors: S. Weiner and L. Hood.

Reprinted from
5 December 1975, Volume 190, pp. 987-989

SCIENCE

Soluble Protein of the Organic Matrix of Mollusk Shells: A Potential Template for Shell Formation

Stephen Weiner and Leroy Hood

Soluble Protein of the Organic Matrix of Mollusk Shells: A Potential Template for Shell Formation

Abstract. *A significant proportion of the soluble protein of the organic matrix of mollusk shells is composed of a repeating sequence of aspartic acid separated by either glycine or serine. This regularly spaced, negatively charged aspartic acid may function as a template upon which mineralization occurs.*

Mollusk shells are generally composed of calcium carbonate crystals enclosed in an organic matrix. Chemical analyses indicate that the organic matrix is primarily a glycoprotein, characteristically containing large proportions of acidic amino acids (*1*) and acid mucopolysaccharides (*2*). Lipids are also found in some organic matrices (*3*). The organization of these components in the matrix is not understood.

The organic matrix is observed to form prior to mineralization (*4*) and probably is involved in the formation of the shell (*5*). Two different functions have been ascribed to the matrix during mineralization (*6*). One theory envisages the organic matrix as a template which controls crystal growth (*7*), while the other proposes that the matrix functions as a compartment in which crystals grow (*4*). We report here a partial characterization of the protein component of the organic matrix. This information should be useful for obtaining a better understanding of the role of organic material in the process of mineralization.

Our strategy in this study was dictated by the observation that the organic matrix extracted from these shells contained from 15 to 43 mole percent of aspartic acid (Table 1). Accordingly, a mild acid hydrolysis procedure was employed that preferentially cleaves on both sides of aspartic acid (*8*). This procedure released significant quantities of additional free amino acids from all shell proteins, indicating that portions of the shell proteins are made up of simple repeating sequences with alternating aspartic acid residues (Fig. 1).

Shells of *Crassostrea virginica* and *Mercenaria mercenaria* (freshly collected) and *Crassostrea irredescens* and *Nautilus pompilius* (air-dried and stored) were physically cleaned and the periostraca removed. The shell portions and their mineralogy used for analysis are identified in Table 2. The shells were decalcified by dialysis at room temperature against phosphate-buffered 8 percent ethylenediaminetetraacetic acid, *p*H 6.9, containing 0.1 percent sodium azide. After complete decalcification the organic matrix was dialyzed against distilled water. The soluble and insoluble fractions were separated by centrifugation. The insoluble fraction was washed repeatedly with distilled water and then lyophilized. The soluble fraction was lyophilized and then desalted on a Sephadex G-25 column in distilled water in order to remove any adherent free amino acids. The observation that the soluble fraction of the organic matrix of *C. virginica* is partially excluded on Sephadex G-200 and included on Sepharose 4B suggests that it has a maximum molecular weight of the order of 1 million. The yields of soluble and insoluble fractions per unit of shell weight, given in Table 1, compare favorably with yields reported by Hare and Abelson (*5*). Weighed portions of the soluble fraction containing norleucine as an internal standard were hydrolyzed in vacuo with 0.5 ml of redistilled 6*N* HCl at 108°C for 20 hours. The hydrolysis products were analyzed on a Durrum 500 amino acid analyzer, and the amino acid compositions are given in Table 1. Shell proteins were also treated with 1.0 ml of 0.25*M* acetic acid in vacuo at 108°C for 48 hours. The resulting freed amino acids were analyzed directly on the Durrum 500 amino acid analyzer.

Cleavage on both sides of aspartic acid resulted in the release of significant quantities of glycine and serine in addition to the expected aspartic acid [see Table 2 and

Table 1. Amino acid compositions of the soluble fraction of the shell organic matrix and the proportions of soluble and insoluble organic matrix per unit shell weight.

Quantity	*Crassostrea virginica*			*Crassostrea irredescens*		*Mercenaria mercenaria*	*Nautilus pompilius*
	Foliated and chalky layers, calcite	Foliated layers, calcite	Chalky layers, calcite	Foliated and chalky layers, calcite	Adductor myostracum, aragonite	shell layers minus myostracum, aragonite	shell layers minus myostracum, aragonite
Amino acid composition of soluble fraction (mole %)							
Aspartic acid + asparagine	32.53	36.75	33.22	42.96	15.10	25.60	26.13
Threonine	2.12	1.34	1.49	0.85	4.48	3.84	4.82
Serine	21.54	12.00	27.14	11.85	8.43	11.06	7.94
Glutamic acid + glutamine	5.25	7.08	5.31	6.60	13.70	7.16	6.59
Proline	2.78	2.43	1.53	0.70	6.43	8.11	4.58
Glycine	24.67	27.27	25.70	26.02	15.89	12.28	23.58
Alanine	1.47	1.52	1.15	2.80	4.85	4.78	4.44
Cysteine						2.08	
Valine	1.05	0.94		1.36	2.65	2.38	1.52
Methionine	0.33		0.29	0.64	1.64	1.22	0.57
Isoleucine	0.91	1.14	0.39	0.69	4.53	2.16	1.72
Leucine	0.89	0.96	0.43	0.98	6.45	3.23	2.21
Tyrosine	3.20	3.19	2.04	1.44	2.55	3.88	6.43
Phenylalanine				0.68	3.38	2.24	2.06
Histidine	0.46	0.87	0.14	0.44	2.26	0.99	2.12
Lysine	1.90	2.67	1.09	1.73	5.04	4.74	2.84
Arginine	0.90	1.77		0.87	2.54	4.18	2.38
Soluble and insoluble nondialyzable organic matter in shell (%)*	0.33	0.58	0.91	0.45	0.60	0.33 ± 0.04	4.05 ± 0.6
Soluble nondialyzable organic matter in shell (%)	0.12	0.29	0.85	0.11	0.07	0.09 ± 0.01	0.27 ± 0.05
Proportion of protein in soluble fraction (%)†	42.70	45.0	19.4	63.6	11.3	83.6	14.7

*Represents the proportions of soluble and insoluble organic matrix after decalcification and exhaustive dialysis against water. †Calculated from moles of amino acid per unit weight soluble fraction obtained on complete hydrolysis assuming 25 percent of the "dry" weight to be water.

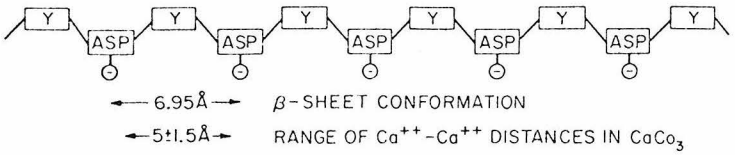


Fig. 1. Schematic illustration of the $(Asp\text{-}Y)_n$ sequence, where Y represents serine or glycine. The distance from one aspartic acid residue to the next in the β-sheet conformation is 6.95 Å. The Ca-Ca distance in the unit cells of aragonite and calcite ranges from about 3.0 to 6.5 Å.

Table 2. Proportions of free amino acids released after 48 hours of hydrolysis with 0.25*M* acetic acid at 108°C. After 48 hours the maximum yield of aspartic acid (Asp) was obtained. Yields are calculated from the amounts of the particular amino acids present in an equivalent, completely hydrolyzed sample of soluble organic matrix. The yields of gylcine (Gly) and serine (Ser) are significantly greater than yields obtained as a result of random cleavage (*9*). Values in the last column represent the moles released assuming an Asp-Y-Asp-Y-Asp-Y. . . sequence divided by the total moles present in an equivalent, completely hydrolyzed sample of soluble organic matrix.

Species and shell parts analyzed	Mineral	Yield (%)			Minimum proportion Asp-Y-Asp-Y sequence (%)
		Asp	Gly	Ser	
Crassostrea virginica					
Foliated and chalky layers	Calcite	82.1	28.8	20.0	29.9
Foliated layers	Calcite	58.9	25.7	13.0	17.8
Chalky layers	Calcite	89.9	33.3	27.7	37.8
Crassostrea irredescens					
Foliated and chalky layers	Calcite	80.2	36.8	28.8	26.4
Adductor myostracum	Aragonite	65.4	14.7	17.1	8.3
Mercenaria mercenaria					
Shell layers minus myostracum	Aragonite	65.0	26.5	18.3	13.3
Nautilus pompilius					
Shell layers minus myostracum	Aragonite	68.0	26.7	10.0	14.3

(*9*)]. Threonine, glutamic acid, and alanine are released in such small quantities that quantitation is not reliable.

The release of serine and glycine together with aspartic acid implies that sequences of the type $(Asp\text{-}Y)_n$, where Y is predominantly serine or glycine, are found in the protein of all the organic matrices examined (Fig. 1). The minimum proportion of $(Asp\text{-}Y)_n$ sequence is calculated, assuming one continuous Asp-Y-Asp-Y-Asp-Y. . . sequence, from the moles of glycine and serine released on partial acid hydrolysis compared to the total number of moles of amino acids present in a completely hydrolyzed equivalent portion of soluble organic matrix. The results are shown in Table 2. In all the species examined the $(Asp\text{-}Y)_n$ sequence comprises a significant part of the soluble organic matrix protein. In no case could it account for more than about 55 percent of the protein present. These results do not, however, distinguish between long continuous or short discontinuous sequences of the $(Asp\text{-}Y)_n$ type. Furthermore, it is not known whether such sequences comprise a majority of certain shell polypeptides or whether they are interspersed among a heterogeneous collection of proteins. This sequence, because of its repeating nature and the predominance of glycine and serine in the repeating unit, is similar to that of other structural proteins, such as collagen and silk.

This characterization of the protein components does not reveal any consistent differences between the calcitic shells of the oyster species and the aragonitic shells of *Mercenaria* and *Nautilus*. However, the protein fractions of the separated shell layers of *C. virginica* contain different amounts of $(Asp\text{-}Y)_n$ sequence and the proportions of Gly and Ser released are also different.

Negatively charged aspartic acid residues are thus found to be present as approximately every second residue, forming a significant portion of the organic matrix proteins (Fig. 1). It is possible that this sequence, comprising regular repeating negative charges, could bind Ca^{2+} ions and thus perform an important function in mineralization. The distance from one aspartic acid residue to the next, if the protein is fully extended, is 7.27 Å (*10*). Most of the Ca^{2+}-Ca^{2+} distances in the crystal lattices of aragonite and calcite range from about 3.0 to 6.5 Å (*11*). As folding will reduce the aspartate-aspartate distances, the spatial requirements for interacting with crystals of calcium carbonate are consistent with an $(Asp\text{-}Y)_n$-type sequence, provided the charges are located on one side of the protein only. This latter provision could be accounted for if the $(Asp\text{-}Y)_n$-type sequence adopted a β-sheet conformation. Infrared spectroscopy indicates that at least part of the protein is present in the β-sheet conformation (*12*). Furthermore, when aspartic acid, glycine, and serine occur in proteins, more than 50 percent of the time they are most likely to be in the β-sheet conformation (*13*).

It is interesting to note that one of the calcium binding sites of carp and hake myogen has a sequence remarkably similar to the organic matrix protein, namely, Gly-Asp-Ser-Asp-Gly-Asp-Gly-Val-Asp-Glu (*14*). Although the mechanisms of binding are probably different, the known association of calcium with this type of sequence lends support to the hypothesis that the $(Asp\text{-}Y)_n$-type sequence itself binds calcium.

The $(Asp\text{-}Y)_n$-type sequence is shown to be present in the organic matrices from the five mollusks examined. This repeating sequence clearly plays an important functional role in the organic matrix and may function as a template for mineralization.

STEPHEN WEINER
Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena 91125

LEROY HOOD
Division of Biology, California Institute of Technology

References and Notes

1. C. Gregoire, G. Duchateau, M. Florkin, *Ann. Inst. Oceanogr. (Paris)* 31, 1 (1955); K. A. Pier, *Science* 134, 841 (1961); P. E. Hare, *ibid.* 139, 216 (1963).
2. K. Simkiss, *Comp. Biochem. Physiol.* 16, 427 (1963); R. H. Hedley, *Proc. Zool. Soc. London* 130, 569 (1958).
3. M. Jope, in *Comprehensive Biochemistry*, M. Florkin and E. H. Stotz, Eds. (Elsevier, Amsterdam, 1971), vol. 26, part C, p. 749.
4. G. Bevelender and H. Nakahara, *Calcif. Tissue Res.* 3, 84 (1969).
5. P. E. Hare and P. H. Abelson, *Carnegie Inst. Washington Yearb.* 64, 223 (1965).
6. K. M. Towe, *Biomineral. Res. Rep.* 4, 1 (1972).
7. M. J. Glimcher, in *Calcification in Biological Systems*, R. C. Sognnaes, Ed. (AAAS, Washington, D.C., 1960), p. 421; K. M. Wilbur, in *Physiology of Molluscs*, K. M. Wilbur and C. M. Yonge, Eds. (Academic Press, London, 1964), vol. 1, p. 243.
8. S. M. Partridge and H. F. Davis, *Nature (London)* 165, 62 (1950); J. Schultz, H. Allison, M. Grice, *Biochemistry* 1, 694 (1962); V. M. Ingram, *Methods Enzymol.* 6, 225 (1963); *ibid.*, p. 831.
9. Four control proteins, egg-white lysozyme, yeast glyceraldehyde-3-phosphate dehydrogenase, bovine ribonuclease, and bovine insulin A, were partially hydrolyzed under the designated experimental conditions in order to determine the degree of nonspecific cleavage. The average mole percentages of glycine and serine released nonspecifically are 4.4 ± 1.5 and 3.1 ± 1.7, respectively, despite the fact that they are present in different proportions in the original proteins. These values are well below those reported for mollusk shell proteins in Table 2. It was also noted that the mole percentage of aspartic acid released after partial acid hydroly-

sis of ribonuclease A, glyceraldehyde-3-phosphate dehydrogenase, and lysozyme corresponds to ± 10 percent of the known mole percentage of aspartic acid in these proteins. For insulin A, however, which contains only asparagine, 70.4 percent of the asparagine was released as aspartic acid, indicating that considerable deamidation had taken place. Thus, the mole percentage of aspartic acid released from shell proteins should be regarded only as an approximation of the amount of aspartic acid present relative to asparagine.

10. L. Pauling, R. B. Carey, H. R. Branson, *Proc. Natl. Acad. Sci. U.S.A.* 37, 205 (1951).
11. F. Lippman, *Sedimentary Carbonate Minerals* (Springer-Verlag, New York, 1973), pp. 6 66.
12. S. Hotta, *Chikyu Kagaku* 23, 133 (1969).
13. P. N. Lewis, F. A. Momany, H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.* 68, 2293 (1971).
14. C. E. Nockolds, R. H. Kretzinger, C. J. Coffee, R. A. Bradshaw, *ibid.*, 69, 581 (1972).
15. We are grateful to H. A. Lowenstam for invaluable assistance. We thank J. Erez, Woods Hole Oceanographic Institute, for the *M. mercenaria* samples and N. Haven, Hopkins Marine Station, for the *N. pompilius* sample. Supported by NSF grant GA36824 x 1.

24 June 1975; revised 25 August 1975

APPENDED PAPER 2

Discrete Molecular Weight Components of Mollusk Shell Organic Matrices.

Submitted for publication August, 1976.

Authors: S. Weiner, H. A. Lowenstam and L. Hood.

DISCRETE MOLECULAR WEIGHT COMPONENTS OF MOLLUSK SHELL ORGANIC MATRICES

S. Weiner[1], H. A. Lowenstam[1] and L. Hood[2]

[1]Division of Geological and Planetary Sciences

[2]Division of Biology

California Institute of Technology

Pasadena, California 91125

Contribution No. 2797, Division of Geological and Planetary Sciences,

California Institute of Technology, Pasadena, California 91125

Abstract

1. Discrete molecular weight components are present in the soluble fraction of mollusk organic matrices.
2. The number of components, relative migration distances and concentrations vary greatly in the specimens analyzed.
3. The relative migration distances of discrete molecular weight components are most similar for species of the same genus.
4. The soluble fraction of mollusk organic matrices is heterogeneous and presumably capable of performing diverse functions.

INTRODUCTION

A knowledge of the degree of heterogeneity of the organic matrix of mollusk shells is important for understanding the molecular organization of this material. Furthermore, the number of components present in the organic matrix may be indicative of the complexity of functions which this material is capable of performing. The specific functions of the mollusk organic matrix are not known, although it has been postulated that the organic matrix is in some way involved in shell formation (Hare and Abelson, 1965; Towe, 1972; Weiner and Hood, 1975). Thus an investigation of the discrete molecular weight components was made in order to improve our understanding of the degree of heterogeneity of mollusk organic matrices.

The organic matrix of mollusk shells comprises approximately 0.01 percent to 5 percent by weight of the shell (Hare and Abelson, 1965). Protein and polysaccharide are the two major constituents. The proteins are generally enriched in the polar and acidic amino acids (Grégoire, 1955; Piez, 1961; Hare, 1963; Degens et al., 1967) and the carbohydrate fraction is known to contain sulfated mucopolysaccharides (Simkiss, 1965), neutral sugars (Ravindranath and Ravindranath, 1974) and amino sugars (Degens et al., 1967).

Estimations of the degree of heterogeneity of mollusk

organic matrices have been made using a number of techniques. Variations in the amino acid compositions of organic matrices from different species of mollusks have been interpreted to indicate that in certain cases more than one protein component is present (Hare, 1963; Degens et al., 1967, Meenakashi et al., 1971). This interpretation is supported by the observation of Degens et al. (1967) that one fraction of the organic matrix of Mercenaria mercenaria has a wide range of molecular weight as determined by gel filtration. However, polyacrylamide gel electrophoresis of the soluble fractions of Nautilus pompilius (Voss-Foucart, 1968) and Mercenaria mercenaria (Crenshaw, 1972) indicated that only one discrete component is present in these fractions at levels that can be detected. In order to improve our understanding of the degree of heterogeneity of mollusk organic matrices, the soluble fractions from nine different mollusk species, representing the three major classes, were electrophoresed on polyacrylamide gradient slab gels.

MATERIALS AND METHODS

Material

The following specimens were used in this study: Bivalvia - Crassostrea gigas (Morro Bay, California); Crassostrea irredescens (Mazatlan, Mexico); Mytilus

californianus (total shell sample) and Mytilus edulis (Los Angeles, California); Mytilus californianus (separate shell layer samples, Punta Cabra, Mexico); Mercenaria mercenaria (Woods Hole, Massachusetts). Gastropoda - Lottia gigantea (Punta Cabra, Mexico); Strombus gigas and Strombus costata (Bermuda). Cephalopoda - Nautilus pompilius (Philippines). All the specimens were collected fresh and air-dried. The shells were physically cleaned and the periostraca removed. Myostracal shell layers and hinge areas of the bivalves were excluded from analysis, except for C. irredescens where the myostracum was analyzed separately. The total shell of L. gigantea, including the myostracum, was used, whereas only the outer lip of the two species of Strombus and the outer margin of the body chamber (the peristome) of N. pompilius were analyzed. The outer prismatic, inner prismatic and a combination of the outer prismatic and nacreous shell layers of M. californianus were analyzed separately. As the prismatic layers are composed of calcite and the nacreous layer of aragonite, the quantities of aragonite in these three samples, analyzed by powder X-ray diffraction are proportional to the nacreous layer contents of the samples. The outer prismatic layer sample contains less than 2 percent aragonite, the inner prismatic layer sample contains 10 $\pm$ 5 percent aragonite and the mixture of outer prismatic and nacreous layers sample contains 45 $\pm$ 5 percent aragonite.

Preparation of organic matrix

The shells were briefly dipped in 5.8 percent ammonium hydroxide to remove contaminating superficial amino acids and peptides. They were then washed with redistilled water, sonicated for about a second and then rewashed. The shells were crushed and dialyzed against 10 percent EDTA* buffered with phosphate to pH 7.0 containing small amounts of sodium azide. After decalcification, the organic matrix was dialyzed against water, lyophilized, the insoluble fraction removed by centrifugation, and the soluble fraction desalted on Sephadex G-25 in water. The fraction excluded on Sephadex G-25 was used for analysis.

Gel electrophoresis

An exponential gradient SDS+ polyacrylamide slab gel was used. The starting solutions contained 5 percent and 25 percent acrylamide, 0.33 percent and 0.07 percent bis-acrylamide respectively and were buffered to pH 8.8 according to the method of Laemmli (1970). The stacking gel composed of 3 percent acrylamide, 0.4 percent bis-acrylamide at pH 6.8 (Laemmli, 1970). The samples were loaded in a reducing buffer after being denatured at 80°C for 1-2 min. The gel was run at 70 volts for 15 hr and stained in Coomassie Blue solution (Weber and Osbourne, 1969).

---

*Ethylenediaminetetraacetic acid

+Sodium dodecyl sulfate

RESULTS

Polyacrylamide gel electrophoresis of the soluble fractions of mollusk organic matrices shows that discrete molecular weight components are present (Fig. 1). The number, relative migration distances and quantities of protein present in these bands varies greatly in the organic matrices examined. No common band is present in all the organic matrix fractions analyzed as determined by relative migration distance.

As the components of the soluble fraction of mollusk organic matrices are thought to be glycoproteins (Crenshaw, 1972; Weiner and Hood, unpublished data), an estimation of their molecular weights compared to the molecular weight of the standard proteins, is not possible. The standard proteins do, however, provide a rough estimate of the molecular weight range of the shell components.

The detection of discrete molecular weight components is a function of the amount of protein present and the extent to which the particular component binds Coomassie Blue stain. The apparent absence of discrete molecular weight bands in, for example, the samples of C. irredescens (non-myostracal layers) and C. gigas (Columns 7 and 8, Fig. 1), could therefore be due to insufficient material or its inability to bind the stain, and does not necessarily imply that the shells of these species do not contain discrete molecular weight components.

The distribution of distinct bands on the gel could be affected if the components were proteolytically cleaved during sample preparation. In order to investigate this possibility, bovine serum albumin was incubated at 37°C at pH 8 for 24 hr with aliquots of the soluble fraction of M. mercenaria. No cleavage products were observed when these samples were electrophoresed on a 10 percent SDS polyacrylamide gel. Therefore, proteolytic activity, at least in this case, has probably not altered the distribution of discrete molecular weight components.

A study of the influence of increased periods of immersion of the ground shells in the decalcifying EDTA solution on the distribution of discrete molecular weight bands was undertaken. It showed that in the case of the soluble fraction of M. mercenaria a component that migrates faster than lysozyme was lost (presumably through the dialysis bag) after about three weeks of decalcification. Because the period of decalcification of the samples examined in Figure 1 varied from about one to three weeks depending upon the amount of shell material being decalcified, its grain size after crushing and the nature of the shell itself, portions of the low molecular weight components in the organic matrix fractions might have been lost.

The discrete molecular weight components account for only a small part of the protein content of the soluble

fraction. For example, discrete molecular weight bands of the soluble fraction of M. mercenaria electrophoresed on 10 percent SDS polyacrylamide gels (Fairbanks et al., 1971) were eluted from the gel with 0.01 percent SDS (Silver and Hood, 1975). The protein content, as estimated from an amino acid analysis, of the three most prominent bands accounted for approximately 10 percent of the total protein loaded on the gel (Weiner and Hood, unpublished data). The remaining protein fraction of M. mercenaria is composed of very high molecular weight components which do not enter the 5 percent polyacrylamide at the top of the gel, and lower molecular weight components which enter the gel and form a dense background of nondiscrete molecular weight material (column 9, Fig. 1). As the organic matrix samples of most of the species examined (Fig. 1) resemble M. mercenaria in this respect, it is thought that the protein content of the discrete molecular weight components constitutes only a small part of the soluble fraction protein.

The relative migration distances of the discrete molecular weight components are most similar for species of the same genus, e.g., M. californianus and M. edulis (columns 13 and 14, Fig. 1). C. gigas and C. irredescens (non-myostracal layers) are similar in that they both have barely distinguishable discrete molecular weight components (columns 7 and 8, Fig. 1). The relative migration

distances of the components of the gastropods, S. costata and S. gigas appear to be identical (columns 4 and 5, Fig. 1). Species from different higher taxonomic levels show no obvious similarities. This may be a function of the very small number of species examined.

Three different shell layers of M. californianus were investigated separately viz. the outer prismatic layer, a composite of the outer prismatic layer and nacreous layer, and the inner prismatic layer (columns 10, 11, and 12, Fig. 1). Although the basic assemblages of discrete molecular weight components are all very similar, each shell layer contains at least one unique facet, viz. an additional component, the absence of a component, or the presence of a particular component in relatively large concentrations. On the other hand the set of discrete molecular weight components of the myostracum of C. irredescens, is very different from the pattern observed for the organic matrices of the combined foliated and chalky layers of this oyster species (columns 6 and 7, Fig. 1).

DISCUSSION

The presence of numerous discrete molecular weight bands in the soluble fractions of the mollusk organic matrices examined, shows that this material is heterogeneous. In addition, most specimens are observed to contain both high molecular weight components that do not enter the

top of the gel and lower molecular weight components that enter the gel to form a dense background of nondiscrete molecular weight material. Therefore, the degree of heterogeneity of these organic matrix fractions is in all likelihood, much greater than the sum of the discrete molecular weight components observed. Thus, the heterogeneity of this material is consistent with the supposition that mollusk shell organic matrix is capable of performing diverse functions.

Species of the same genus show similar or identical distributions of distinct components on the gel. M. edulis and M. californianus (total shell) samples contain numerous discrete molecular weight components whose relative migration distances, although similar, show many distinct differences (Fig. 1). The oysters, C. gigas and C. irredescens (non-myostracal layers) are similar in that they both contain barely discernible bands. The pair of gastropod species, S. gigas and S. costata contain discrete molecular weight components which appear to have identical relative migration distances. The limited sampling precludes drawing any conclusions about comparative gel patterns at higher taxonomic levels. However, the striking differences in gel patterns observed in the specimens examined above the genus level, strongly suggests that there is significant diversity in the distribution of discrete molecular weight bands in the soluble fractions of

mollusk shell organic matrices. Further investigations might show whether or not this diversity in gel patterns is due to the relatively rapid evolution of these proteins. As the rate of evolution of a protein is related to the functional requirements to which the protein is subjected (Dickerson, 1972), this information could provide an indication of the functional requirements to which shell glycoproteins are subjected.

The factors which affect the assemblage of discrete molecular weight components present in the soluble fraction of mollusk organic matrices are, no doubt, numerous and complex. The following discussion elaborates on some of these factors.

Individual shell layers of a particular specimen contain unique assemblages of proteins. Amino acid composition differences of the shell layers of M. californianus have been observed (Hare, 1963). The distribution of discrete molecular weight components from the three shell layers of M. californianus also show distinct characteristics (Fig. 1). The three shell layers differ with respect to their mineralogy, microarchitecture and the location of the cells in the mantle which are responsible for their formation. The outer and inner prismatic shell layers are composed of calcite. They differ in that the outer prismatic layer is formed by cells at the mantle edge, whereas the inner prismatic layer is formed by mantle cells

located within the pallial line. The middle layer has a nacreous ultrastructure, is composed of aragonite and is formed by mantle cells also within the pallial line. Despite these differences, the basic assemblages of discrete molecular weight components are very similar (Fig. 1). On the other hand, a comparison of the adductor myostracal shell layer with the non-myostracal layers of C. irredescens shows marked differences (Fig. 1). The myostracum has a prismatic ultrastructure, is composed of aragonite (Lowenstam, 1964) and is formed by cells of the adductor muscle. The non-myostracal layers comprise a mixture of foliated and chalky calcite layers which are formed by mantle cells. Although these three factors, viz. ultrastructure, mineralogy and the type of cells which are responsible for the formation of the particular shell layer must surely be reflected in the overall composition of the organic matrix, their relationship to the assemblages of discrete molecular weight components in the individual shell layers analyzed, is not clear.

Other factors which might alter the assemblage of glycoproteins found in the organic matrix are the stage of development of the mollusk examined and the environment in which it grew. Variations in the amino acid compositions of mollusk organic matrices have been ascribed to changes in the environment. For example, changes in temperature and salinity seem to affect the amino acid compositions of

the organic matrix from a particular shell layer of M. californianus (Hare, 1962). Individuals of the same species from different environments are known to differ in the amino acid composition of their organic matrices (Degens et al., 1967). The effect of environment or stage of development on the assemblage of discrete molecular weight components, is not known. On the other hand the observation that assemblages of discrete molecular weight glycoproteins are most similar at the species level, suggests that the discrete molecular weight components might somehow be involved in determining the physical characteristics of the shell of one species which enables it to be distinguished from other species.

CONCLUSIONS

The presence of soluble discrete molecular weight components in mollusk organic matrices, shows that the organic matrix is heterogeneous. As the discrete molecular weight components form only a part of the total soluble fraction, the degree of heterogeneity is expected to be considerably greater than just the sum of distinct components observed on the gel. Thus, the heterogeneity of the organic matrix is consistent with the supposition that these shell glycoproteins are capable of performing diverse functions.

ACKNOWLEDGEMENTS

We thank N. Haven, Hopkins Marine Station, for the N. pompilius sample, J. Erez, Woods Hole Oceanographic Institution, for the M. mercenaria sample and B. Buchardt Larsen, Institute of Historical Geology, Copenhagen, for the X-ray diffraction analyses.

This study was supported in part by NSF Grant DES 75-08659 to Dr. Lowenstam and NSF Grant PCM 71-00770 to Dr. Hood.

REFERENCES

Crenshaw, M. A. (1972) The soluble matrix from Mercenaria mercenaria shell. Biomineralization Research Reports 6, 6-11.

Degens, E. T., Spencer, D. W. and Parker, R. H. (1967) Paleobiochemistry of molluscan shell proteins. Comparative Biochemistry and Physiology 20, 553-579.

Dickerson, R. E. (1972) The structure and history of an ancient protein. Scientific American 226, 58-72.

Fairbanks, G., Steck, T. L. and Wallach, D. F. H. (1971) Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. Biochemistry 10, 2606-2617.

Grégoire, C., Duchateau, G. and Florkin, M. (1955) La trame protidique des nacres et des perles. Annales de L'Institut Océanographique 31, 1-36.

Hare, P. E. (1962) The amino acid composition of some recent and fossil shells of some west coast species of Mytilus. Ph.D. Thesis, California Institute of Technology, p. 109.

Hare, P. E. (1963) Amino acids in the proteins from aragonite and calcite in the shells of Mytilus californianus. Science 139, 216-217.

Hare, P. E. and Abelson, P. H. (1965) Amino acid composition of some calcified proteins. Carnegie Institution of Washington Yearbook 64, 223-232.

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

Lowenstam, H. A. (1964) Coexisting calcites and aragonites from skeletal carbonates of marine organisms and their strontium and magnesium contents. In Recent Researches in the Field of Hydrosphere, Atmosphere and Nuclear Geochemistry. (Ed., Editorial Committee of Sugawara Festival Volume) Maruzen Co. Ltd., Tokyo, 373-404.

Meenakashi, V. R., Hare, P. E. and Wilbur, K. M. (1971) Amino acids of the organic matrix of Neogastropod shells. Comparative Biochemistry and Physiology 40B, 1037-1043.

Piez, K. A. (1961) Amino acid composition of some calcified proteins. Science 134, 841.

Ravindranath, M. H. and Rajeswari Ravindranath, J. H. (1974) The chemical nature of the shell of molluscs: I. Prismatic and nacreous layers of a bivalve Lamellidans marginalis (unionidae). Acta Histochimica 48, 26-41.

Silver, J. and Hood, L. (1975) An approach to amino acid sequence analysis of transplantation antigens. Nature 256, 63-64.

Simkiss, K. (1965) The organic matrix of the oyster shell. Comparative Biochemistry and Physiology 16, 427-435.

Towe, K. M. (1972) Invertebrate shell structure and the organic matrix concept. Biomineralization Research Reports 4, 2-14.

Voss-Foucart, M. F. (1968) Essais de solubilisation et de fractionnement d'une conchioline (nacre murale de Nautilus pompilius, Mollusque Cephalopode). Comparative Biochemistry and Physiology 26, 877-886.

Weber, K. and Osbourne, M. (1969) The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. Journal of Biological Chemistry 244, 4406-4412.

Weiner, S. and Hood, L. (1975) Soluble protein of the organic matrix of mollusk shells: a potential template for shell formation. Science 190, 987-989.

FIGURE LEGEND

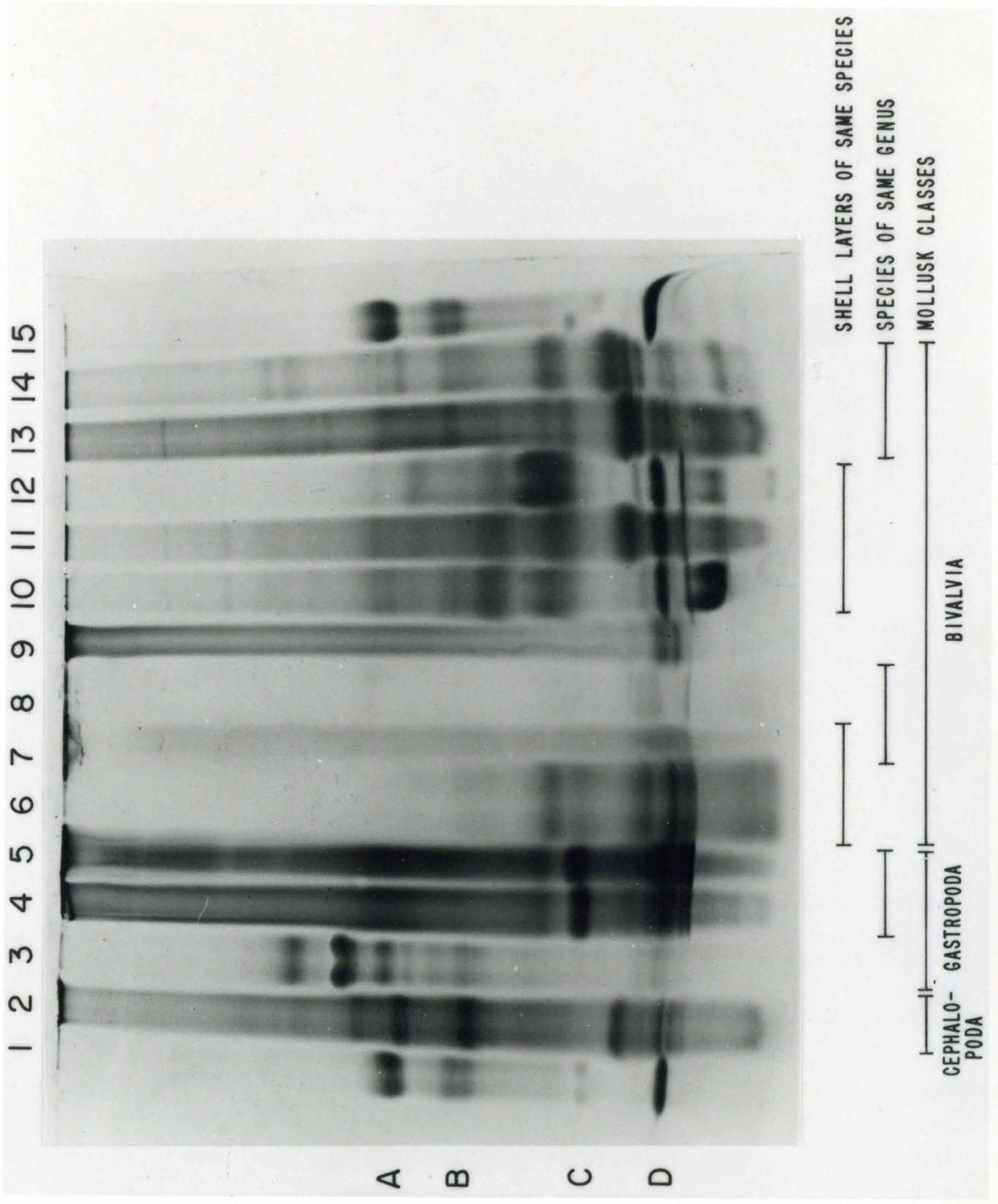
Figure 1. Gel electrophoresis of the soluble fractions of organic matrices on an SDS-polyacrylamide gradient slab gel. 1 - Standard proteins; 2 - N. pompilius; 3 - L. gigantea; 4 - S. costata; 5 - S. gigas; 6 - C. irredescens (myostracum only); 7 - C. irredescens (non-myostracal layers); 8 - C. gigas; 9 - M. mercenaria; 10 - M. californianus (inner prismatic layer); 11 - M. californianus (outer prismatic and nacreous layers); 12 - M. californianus (outer prismatic layer); 13 - M. californianus; 14 - M. edulis; 15 - Standard proteins.
Standard proteins are: A - Bovine serum albumin (65,000); B - ovalbumin (45,000); C - Inactivated trypsin (23,800); D - lysozyme (14,100).

APPENDED PAPER 3

Characterization of 80-Million Year Old Mollusk Shell Proteins.

Published in Proceedings of the National Academy of Sciences USA, 73, 2541-2545, 1976.

Authors: S. Weiner, H. A. Lowenstam and L. Hood.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
A
B
C
D
SHELL LAYERS OF SAME SPECIES
SPECIES OF SAME GENUS
MOLLUSK CLASSES
CEPHALO-PODA
GASTROPODA
BIVALVIA

Reprinted from
Proc. Natl. Acad. Sci. USA
Vol. 73, No. 8, pp. 2541–2545, August 1976
Geology

Characterization of 80-million-year-old mollusk shell proteins

(amino acid analyses/gel electrophoresis/shell organic matrix/Trigoniacea)

S. WEINER*, H. A. LOWENSTAM*, AND L. HOOD†

* Division of Geology and † Division of Biology, California Institute of Technology, Pasadena, Calif. 91125

Communicated by Leon T. Silver, May 21, 1976

ABSTRACT **Fossil glycoproteins of the soluble organic matrix are present in an 80-million-year-old mollusk shell from the Late Cretaceous Period. Discrete molecular weight components, as determined by gel electrophoresis, are preserved. The fossil organic matrix was compared with the organic matrix of a living representative species of the same superfamily. A particular repeating amino acid sequence, found in contemporary mollusk shell proteins, was identified in the fossil glycoproteins. The ultrastructure, mineralogy, and chemistry of the inorganic components of the fossil and contemporary shells provide information on the state of preservation of the fossil. The use of fossil shell proteins to further our understanding of molecular evolution is discussed.**

This paper reports on the chemical characterization of fossil proteins from the shell organic matrix of the mollusk, *Scabrotrigonia thoracica*, which lived during the Late Cretaceous Period, about 80 million years ago. We were particularly interested in the state of preservation of these fossil proteins in order to assess their potential value in the study of protein and shell evolution. In 1954 Abelson (1) first reported the presence of amino acids in fossil bones and shells as old as about 350 million years. The presence of amino acids in fossils has been reported subsequently by numerous investigators (2–6). Peptide bonds (7, 8) and even fossil organic material with preserved antigenic determinants have been reported (9). This report describes the presence of shell glycoproteins of discrete molecular weight in fossil *S. thoracica* that share a particular repeating amino acid sequence with their contemporary counterparts. In addition, the ultrastructure, chemistry, and mineralogy of the shell inorganic components provide information on the conditions under which the shell has been preserved.

MATERIALS AND METHODS

Choice of a Suitable Fossil Shell. Mollusk shells of the Late Cretaceous (Lower Maestrichtian) deposits at the Coon Creek type locality in southwestern Tennessee are well-known for their exceptional physical preservation (10, 11). Geochemical studies of the shell carbonate from a variety of species have shown that the original shell mineralogy is preserved and that strontium, magnesium, and $\delta^{18}O$ contents of the shells have escaped detectable diagenetic alterations (12–14). We have chosen to study shells from Coon Creek of the bivalve, *Scabrotrigonia thoracica*. The fact that these shells still retain their original pearly luster in the nacreous layer provides additional evidence for their exceptional state of preservation. Shells of an extant representative of the superfamily Trigoniacea, *Neotrigonia margaritacea*, from South Australia were chosen for comparison. The Late Cretaceous trigoniacean shells were recovered from the Coon Creek deposits about 2 m below the oxidized outcrop surface. The shells of the extant species were obtained from specimens collected alive. After removal of the soft parts, the shells were dried at room temperature.

Shell Analyses. The shell ultrastructures of *S. thoracica* and *N. margaritacea* were examined with an ETEC Autoscan scanning electron microscope. Shell mineralogy was determined by means of a Perkin Elmer 180 spectrophotometer using KBr pellets. Strontium and magnesium analyses were obtained with a MAC V automated electron microprobe. The oxygen and carbon isotopic compositions were obtained by the method of Epstein *et al.* (15). The results are expressed as:

$$\delta \text{ (parts per thousand)} = \frac{R_{\text{sample}} - R_{\text{standard}}}{R_{\text{standard}}} \times 1000$$

where R is the isotopic ratio $^{18}O/^{16}O$ or $^{13}C/^{12}C$ and the standard is the Chicago PDB-I standard.

Purification of the Organic Matrix. After mechanical cleaning and removal of periostracal remnants, the shells were briefly dipped in 5.8% (wt/vol) ammonium hydroxide to remove contaminating superficial amino acids and peptides. They were then washed with redistilled water, sonicated for about a second, and then rewashed. The myostracal shell layers and hinge areas were excluded from analysis. The shells were crushed and dialyzed against 10% EDTA (wt/vol) buffered with phosphate to pH 7.0 containing small amounts of sodium azide. After decalcification, the organic matrix was dialyzed against water and lyophilized. The insoluble fraction was removed by centrifugation, and the soluble fraction desalted on Sephadex G-25. The organic matrix of *S. thoracica* was almost completely soluble, whereas that of *N. margaritacea* was predominantly insoluble. The included Sephadex G-25 fractions were exhaustively redialyzed against water and lyophilized. Portions of the fraction excluded from Sephadex G-25 were run through a Sephadex G-100 column, equilibrated with redistilled water.

Characterization of the Organic Matrix. The preservation of original amino acid sequence was determined by cleaving the protein on both sides of aspartic acid residues using mild acid hydrolysis and then analyzing the resultant products for free glycine and serine. A repeating $(\text{-Asp-Y-})_n$ sequence, where Y is glycine or serine, is present in the soluble organic matrix proteins of a number of extant mollusks (16). Cleavage on both sides of aspartic acid was obtained by hydrolyzing aliquots of the different molecular weight fractions in 1.0 ml of 0.25 M acetic acid under reduced pressure at 108° for 48 hr (16). Additional aliquots were totally hydrolyzed in 0.5 ml of redistilled 6 M HCl under reduced pressure at 108° for 20 hr. Portions of the fraction of *S. thoracica* excluded from Sephadex G-100 and of the fraction of *N. margaritacea* excluded from Sephadex G-25 were electrophoresed on 7.5% acrylamide-sodium dodecyl sulfate gels (17). The fractions of the living and fossil trigoniaceans included in Sephadex G-25 were electrophoresed on 10% (wt/wt) urea-acrylamide-sodium dodecyl sulfate gels (18).

RESULTS

The preservational state of inorganic and organic properties of eight valves was investigated. The best preserved valve is

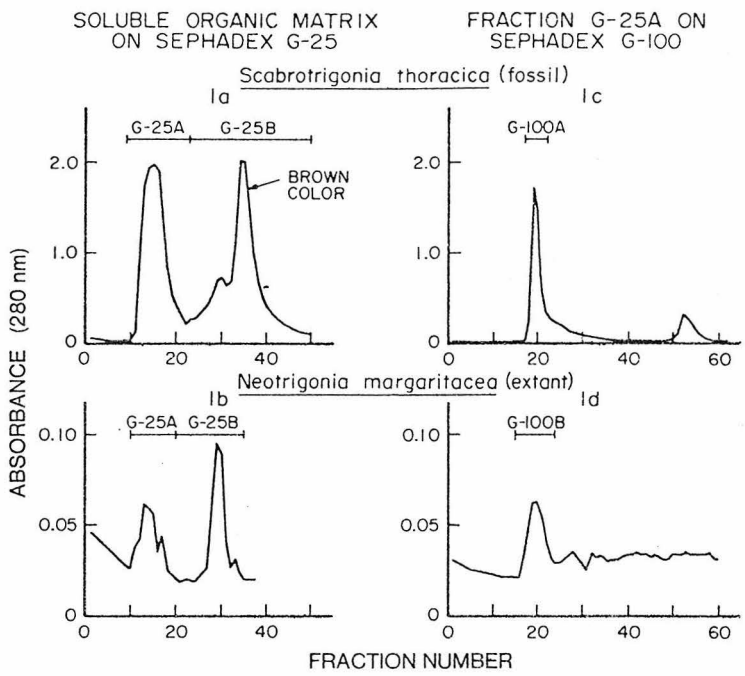


FIG. 1. Chromatograms of the soluble organic matrix fractions of *S. thoracica* and *N. margaritacea* on Sephadex G-25 and G-100.

described in this report. The inorganic and organic properties of the remaining valves are to be reported elsewhere.

Characterization of inorganic ultrastructure and composition

X-ray diffraction and infrared absorption spectra show that the bioinorganic fraction of the shells from the Recent and Late Cretaceous species is aragonite. Scanning electron micrographs show that the ultrastructures of the recent and fossil shell are basically alike in that they both have an outer prismatic layer and two underlying nacreous layers consisting of sheet nacre. Close inspection of the micrographs from the fossil shell established that there is no evidence of even trace amounts of secondary calcite development either as partial replacement of aragonitic sheet nacre or in the form of intercrystalline cement.

The average $MgCO_3$ content of the recent shells is 0.04% and of the late Cretaceous shell 0.06% by weight. The recent shells contain, on the average, 0.16% $SrCO_3$ and the fossil shell 0.23% by weight. The $\delta^{18}O$ and $\delta^{13}C$ contents of the fossil shell carbonate, relative to the PDB standard, are −0.56 and +1.60 parts per thousand, respectively.

Gel filtration of organic matrix

Fig. 1 shows the gel filtration patterns on Sephadex G-25 and G-100 of the soluble organic matrices of *S. thoracica* and *N. margaritacea*. The organic matrix of *S. thoracica*, after decalcification, is almost totally soluble, whereas the organic matrix of *N. margaritacea* is predominantly insoluble. Both trigoniacean organic matrices have high-molecular-weight fractions excluded on Sephadex G-100. *S. thoracica* has an additional low-molecular-weight component (Fig. 1c). The included fractions of *S. thoracica* and *N. margaritacea* on Sephadex G-25 are bimodal (Fig. 1a and b). The peak around fraction 30 is composed predominantly of undialyzed salt, whereas the second peak (around fraction 34) contains proteinaceous material. A proteinaceous fraction included in Sephadex G-25 is also present in living *Mercenaria mercenaria* organic matrix, which when electrophoresed on urea-sodium dodecyl sulfate-polyacrylamide gels, has discrete high-molecular-weight components. This observation shows that its chromatographic behavior on Sephadex G-25 is anomalous (unpublished data).

Amino acid compositions of organic matrix fractions

The uncorrected amino acid compositions of the fractions obtained from Sephadex G-25 and G-100 are shown in Table 1. The amino acid compositions of the different molecular weight fractions of *S. thoracica* are all basically similar. The amino acid compositions of the molecular weight fractions of *N. margaritacea* (Table 1) show distinct differences, indicating that these fractions are composed of different protein components. The amino acid compositions of the fossil and living trigoniacean organic matrices do not show any obvious similarities (Table 1).

The alloisoleucine/isoleucine ratios of all the molecular weight fractions of *S. thoracica* are very low (Table 1)[‡]. In general, equilibration between these two epimers occurs after

[‡] The extent of racemization of eight amino acids (Ala, Val, Leu, Pro, Asp, Phe, Glu, and Lys) of the fractions from one particular *S. thoracica* specimen included in and excluded from Sephadex G-25 was analyzed using a gas chromatographic technique (21). The analysis showed that little or no racemization has occurred (G. E. Pollock, personal communication). The alloisoleucine/isoleucine ratios of these fractions are also very low. The alloisoleucine/isoleucine ratios of the fractions from the remaining *S. thoracica* shells excluded from Sephadex G-25 ranges from less than 0.1 to 1.4, the equilibrium value (3).

Table 1. Amino acid compositions of extant and fossil trigoniacean shell organic matrices

Amino acid composition (mole %)	*Scabrotrigonia thoracica* (fossil)			*Neotrigonia margaritacea* (extant)			
	G-100A	G-25A	G-25B	Insoluble fraction	G-100A	G-25A	G-25B
Aspartic acid + asparagine	11.98	11.82	11.77	7.45	26.11	18.13	7.85
Threonine	6.40	6.11	6.36	1.37	4.09	3.53	4.09
Serine	7.33	7.00	8.19	10.59	13.23	13.09	16.68
Glutamic acid + glutamine	10.43	9.35	9.36	3.32	8.97	8.06	15.96
Proline	6.61	6.29	5.85	1.52	3.46	3.60	3.25
Glycine	11.25	12.70	11.92	32.11	20.25	23.02	18.11
Alanine	10.84	10.70	10.97	23.30	6.04	9.28	11.23
Cysteine*	Trace	0.88	1.24	0.81	Trace	1.22	0.97
Valine	5.83	5.47	4.90	1.43	1.87	2.23	3.57
Methionine	1.29	1.41	1.32	1.98	1.33	1.94	0.65
Alloisoleucine	0.15	0.24	0.07	—	—	—	—
Isoleucine	3.10	3.41	3.22	1.45	1.60	2.09	2.34
Leucine	7.33	6.88	7.38	3.72	3.11	3.53	4.02
Tyrosine	2.84	2.70	1.90	2.73	1.33	1.44	0.97
Phenylalanine	2.94	3.82	3.29	2.71	2.93	2.73	2.14
Histidine	1.65	1.59	1.83	0.44	0.71	0.72	2.60
Lysine	5.06	4.59	5.56	1.98	2.13	2.52	3.31
Arginine	5.01	5.06	4.90	3.11	2.84	2.88	2.27
Alloisoleucine/Isoleucine	0.05	0.07	0.02	—	—	—	—
Proportion of protein in shell (μg/g dry shell)	—	6	6	3280	—	96	40

* *S. thoracica* might contain aminobutyric acid which cochromatographs with cysteine.

only hundreds of thousands of years (19, 20). The low concentration of alloisoleucine indicates unusual preservational conditions. An anhydrous environment of preservation of these fractions within the shell could account for the almost total absence of epimerization, since laboratory simulation experiments show that water is necessary for the epimerization and racemization of amino acids (19).

The amount of nondialyzable soluble protein in the fossil shell is about one-eleventh that of the soluble protein fraction in the extant *N. margaritacea* shell. However, the protein content of the fossil shell is about 1/280 that of the total present in the *N. margaritacea* shell (Table 1).

Cleavage of organic matrix proteins on both sides of aspartic acid residues

Partial acid hydrolysis of the fractions of the organic matrices of *S. thoracica* and *N. margaritacea* excluded from Sephadex G-100 releases significant and very similar yields of aspartic acid, glycine, and serine (Table 2). The proportions of Asp:Gly:Ser released in the soluble fractions of the fossil and living trigoniaceans are also very similar. The release of small quantities of other amino acids is probably the result of nonspecific cleavage (16). This suggests that the repeating $(\text{-Asp-Y-})_n$ sequence is present in significant quantities in fossil as well as the contemporary proteins.

Gel electrophoresis of organic matrix components

The fraction of *S. thoracica* excluded from Sephadex G-100 and the fraction included in Sephadex G-25, when electrophoresed on polyacrylamide gels, show faint but distinguishable high-molecular-weight bands in addition to a heterogeneous collection of low-molecular-weight proteinaceous material (Fig. 2). The fractions of *N. margaritacea* excluded from and included in Sephadex G-25 do not show corresponding bands. Accordingly, discrete polypeptides of well-defined length are present in the fossil shell matrix.

DISCUSSION

The *S. thoracica* Shell Appears Well Preserved by Physical and Chemical Criteria. The preservation of the original microarchitecture and original mineralogy suggest that the Late Cretaceous shell is exceptionally well preserved. The compar-

Table 2. Proportions of the free amino acids released from the G-100 A fractions after 48 hr of hydrolysis with 0.25 M acetic acid at 108°

Species	Yield (%)*			Minimum† proportion Asp-Y-Asp-Y sequence (%)	Proportions of Asp:Gly:Ser
	Asp	Gly	Ser		
Scabrotrigonia thoracica (fossil)	51.3	14.5	13.7	5.2	6.3:1.7:1
Neotrigonia margaritacea (extant)	58.5	18.1	16.3	11.7	7.2:1.7:1

* Yields are calculated from the amounts of the particular amino acid present in an equivalent, completely hydrolyzed sample.

† Represents the moles released assuming an Asp-Y-Asp-Y-Asp . . . sequence divided by total moles present in an equivalent, completely hydrolyzed sample.

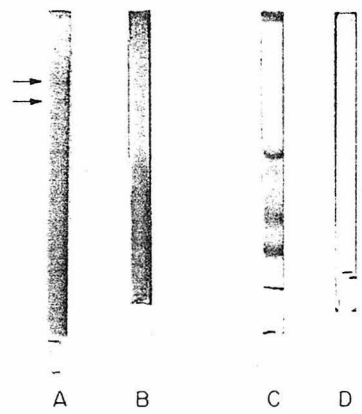


FIG. 2. Polyacrylamide gels of the fraction of fossil *S. thoracica* excluded from Sephadex G-100 (A); the fraction of *S. thoracica* included in Sephadex (B); the fraction of *N. margaritacea* excluded in Sephadex G-25 (C); and the fraction of *N. margaritacea* included in Sephadex G-25 (D). A and C are 7.5% sodium dodecyl sulfate-polyacrylamide gels. B and D are 20% urea-sodium dodecyl sulfate-polyacrylamide gels.

ison of the scanning electron micrographs of the fossil and contemporary shells indicates that the crystal fabrics are identical and that there is no indication of physical alteration of the fossil shell. The preservation of aragonite, one of the less stable polymorphs of calcium carbonate, is also indicative of unusual preservational conditions.

The Late Cretaceous trigoniacean shell studied here has its original or nearly original oxygen isotopic composition. Indeed, the samples having the most positive $\delta^{18}O$ values were shown to be isotopically free of diagenetic alterations or only altered to a very minor extent. The positive value of the carbon isotopic composition of the shell is also in agreement with this interpretation.

The oxygen isotopic composition of the fossil shell is 1.2% more positive than the average of the published values for three conspecific samples from the same collecting site (12). The previously analyzed shells were collected at or very near the surface of the Coon Creek exposure. Differences of the same magnitude have been found in the isotopic composition of other fossil suites between samples taken from exposed surfaces and from extensively sediment-shielded collecting sites (22). These data indicate that the fossil shell exposure to meteoric ground water was negligible throughout its postdepositional history.

The magnesium content of the fossil shell is higher than that of the contemporary shell. This is probably due to environmental temperature or species differences.[§] Since information on the factors affecting the uptake levels of strontium in extant bivalves is as yet incomplete, it is not possible to evaluate the significance of the strontium contents of the fossil shell.

[§] Magnesium contents of recent marine skeletal aragonites can differ between mollusk species. Within a species, however, a positive correlation with environmental temperature always exists (23, 24). The extant trigoniacean shells are from temperate water and the fossil shells are from subtropical water, as determined by the oxygen isotopic composition of the shell carbonate from Coon Creek fauna (12). Thus the observed differences in magnesium contents between recent and fossil shells appears to be primarily related to environmental temperature differences, possibly to species differences, but not diagenesis.

We conclude, therefore, that the mineralogy, microarchitecture, and magnesium, $\delta^{18}O$, and $\delta^{13}C$ contents of the fossil Trigoniacea shell are still preserved.

Discrete Glycoproteins with Preserved Regions of Amino Acid Sequence Are Present in the Fossil Shell. The presence of discrete fossil components of high molecular weight (greater than about 10,000) on polyacrylamide gels indicates that glycopeptides have been preserved in the fossil organic matrix (Fig. 2). Random hydrolytic or proteolytic cleavage would be expected to convert protein components to a series of polypeptides of many different sizes. Accordingly, these discrete protein bands probably represent the best preserved fraction of the fossil organic matrix.

In an earlier study on contemporary shell proteins, the use of a chemical technique to cleave aspartic acid from polypeptides demonstrated that a significant fraction of the shell glycoprotein was composed of a repeating sequence $(\text{-Asp-Y-})_n$ where Y was glycine or serine (16). When this same technique was used on fossil glycoproteins, it appeared that the $(\text{-Asp-Y-})_n$ sequence is still present (Table 2). Indeed, the proportions of glycine and serine released from the glycoproteins of *S. thoracica* and *N. margaritacea* are very similar.[¶] This similarity implies that contemporary and fossil glycoproteins have similar regions of the repeating $(\text{-Asp-Y-})_n$ sequence.

There are distinct differences in the fossil and contemporary glycoproteins that might be ascribed either to evolutionary or diagenetic changes. (*i*) Distinct amino acid compositional differences are noted (Table 1). In this regard, it should be noted that there is significant reduction of nondialyzable material in the fossil shell as compared to the living trigoniacean. This could result in a selected, but not representative, fraction of organic matrix being preserved. Of course, the preferential destruction of more labile peptides and amino acids, or the conversion of one amino acid to another (25), could also account for the observed differences. However, characterization at the amino acid composition level cannot, in this case, distinguish diagenetic from evolutionary changes. (*ii*) The fact that corresponding molecular weight bands are not found in the fossil and contemporary shell proteins suggests either that changes in the molecular weights of the components have occurred during evolution or that a nonrandom diagenetic change has converted larger glycoproteins into those seen on the acrylamide gels (*iii*) Finally, the minimum proportion of the $(\text{-Asp-Y-})_n$ sequence in the fossil glycoproteins is about half that of the contemporary glycoproteins (Table 2). Again, this difference could be an evolutionary or diagenetic change.

Fossil Shell Proteins May Contribute to Our Understanding of Molecular Evolution. The finding of fossil proteins with at least partially preserved amino acid sequence and discrete molecular weight components offers the opportunity to study fossil glycoproteins and to compare them with their contemporary counterparts. Thus it may be possible to compare proteins at various stages of evolution. The shell proteins are of particular interest because they may play a determinant role in shell morphology. Accordingly, it may be possible to compare the evolution of shell morphology with the evolution of shell proteins.

There are obvious difficulties with these proposed studies. First, one must determine the diagenetic changes that may occur with time in fossil proteins. In part this limitation can be circumvented by selecting unusually well-preserved material,

[¶] The proportions of glycine:serine released from other soluble organic matrix proteins of mollusk shells vary greatly from one species to another (unpublished data).

such as the Coon Creek fauna. Second, more should be known about the protein chemistry of contemporary shells. This additional information will be necessary in order to make homologous comparisons of contemporary and fossil proteins. Third, techniques will need to be devised for the isolation of discrete molecular weight components in quantities sufficient for more detailed chemical analysis. In spite of these limitations, the identification of unusually well preserved 80-million-year-old fossil proteins shows that the possibility exists for using this and other such well-preserved fossil proteins for studies in molecular evolution.

We thank S. Stevenson for the shells of *Neotrigonia margaritacea*, S. Savin and C. Emiliani for the oxygen and carbon isotopic determinations, and N. F. Sohl for informing us about the generic assignment of the Late Cretaceous species. F. Stehli participated in the recovery of fresh shells of *S. thoracica* and A. Chodos performed the electron probe analyses of the trace elements. G. E. Pollock analyzed the amino acid racemization states. This study was supported in part by NSF Grant DE575-08659 to H.A.L. and NSF Grant PCM71-00770 to L.H. This work is Contribution no. 2715, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, Calif. 91125.

1. Abelson, P. H. (1954) *Carnegie Inst. Washington Yearb.* 53, 97–101.
2. Grégoire, C., Duchâteau, G. & Florkin, M. (1955) *Ann. Inst. Oceanogr. (Paris)* 31, 1–36.
3. Hare, P. E. & Mitterer, R. M. (1967) *Carnegie Inst. Washington Yearb.* 65, 362–364.
4. Degens, E. T. & Love, S. (1965) *Nature* 205, 876–878.
5. Matter, P., III, Davidson, F. D. & Wyckoff, R. W. G. (1970) *Comp. Biochem. Physiol.* 35, 291–298.
6. Wyckoff, R. W. G. (1972) *The Biochemistry of Animal Fossils* (Scientechnica, Bristol, England), 152 pp.
7. Jope, M. (1967) *Comp. Biochem. Physiol.* 30, 225–232.
8. Grégoire, C. & Voss-Foucart, M. F. (1970) *Arch. Int. Physiol. Biochim.* 78, 191–203.
9. de Jong, E. W., Westbroek, P. & Westbroek, J. F. (1974) *Nature* 252, 63–64.
10. Wade, B. (1926) *U.S. Geol. Surv. Prof. Pap.* 137, 192 pp.
11. Sohl, N. F. (1960) *U.S. Geol. Surv. Prof. Pap.* 331-A, 151 pp.
12. Lowenstam, H. A. & Epstein, S. (1954) *J. Geol.* 62, 207–248.
13. Lowenstam, H. A. (1963) *The Earth Sciences: Problems and Progress in Current Research* (Univ. Chicago Press, Chicago), 137 pp.
14. Lowenstam, H. A. (1964) in *Isotopic and Cosmic Chemistry*, eds. Craig, H., Miller, S. L. & Wasserburg, G. J. (North Holland Publishing Co., Amsterdam), pp. 114–132.
15. Epstein, S., Buchsbaum, R., Lowenstam, H. A. & Urey, H. C. (1953) *Geol. Soc. Am. Bull.* 64, 1315–1326.
16. Weiner, S. & Hood, L. (1975) *Science* 190, 987–989.
17. Fairbanks, G., Steck, T. L. & Wallach, D. F. H. (1971) *Biochemistry* 10, 2606–2617.
18. Swank, R. T. & Munkres, K. D. (1971) *Anal. Biochem.* 39, 462–477.
19. Hare, P. E. & Mitterer, R. M. (1969) *Carnegie Inst. Washington Yearbk.* 67, 205–208.
20. Bada, J. L. (1972) *Earth Planet. Sci. Lett.* 15, 223–231.
21. Pollock, G. E., Oyama, V. I. & Johnson, R. D. (1965) *J. Gas Chromatogr.* 3, 174–176.
22. Lowenstam, H. A. (1961) *J. Geol.* 69, 241–260.
23. Chave, K E. (1954) *J. Geol.* 62, 587–599.
24. Dodd, J. R. (1967) *J. Paleontol.* 41, 1313–1329.
25. Vallentyne, J. R. (1964) *Geochim. Cosmochim. Acta* 28, 157–188.