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:

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

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

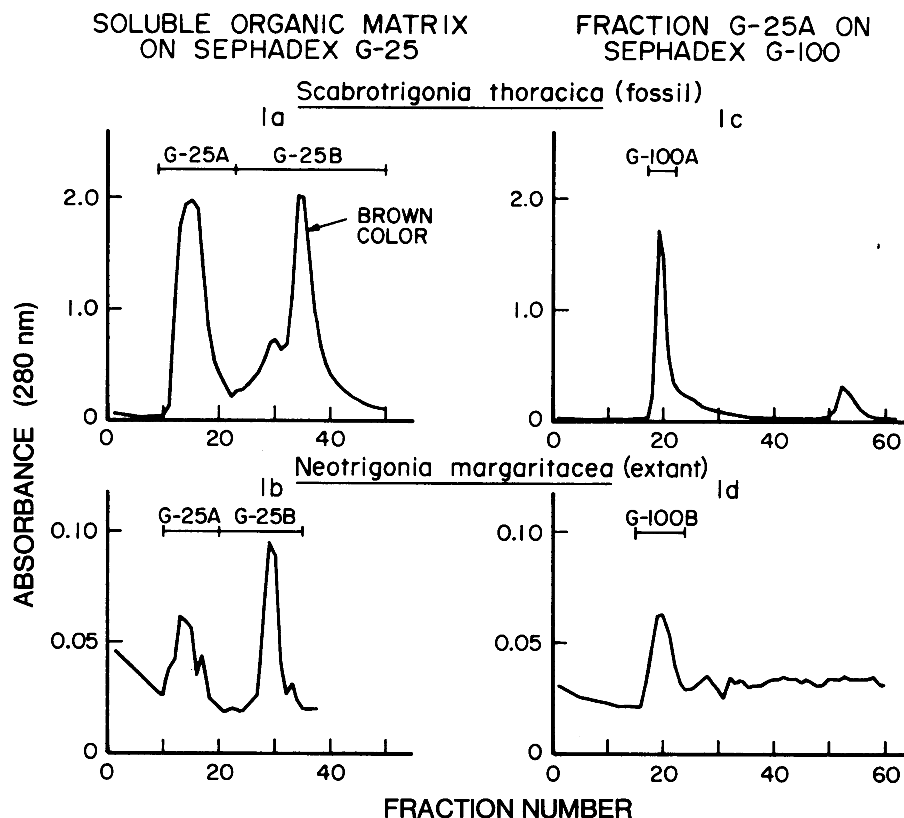

FIG. 1. Chromatograms of the soluble organic matrix fractions of *S. thoracica* and *N. margaritacea* on Sephadex G-25 and G-100.

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 (-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 (%)* Asp	Yield (%)* Gly	Yield (%)* Ser	Minimum† proportion Asp-Y-Asp-Y sequence (%)	Proportions of 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.

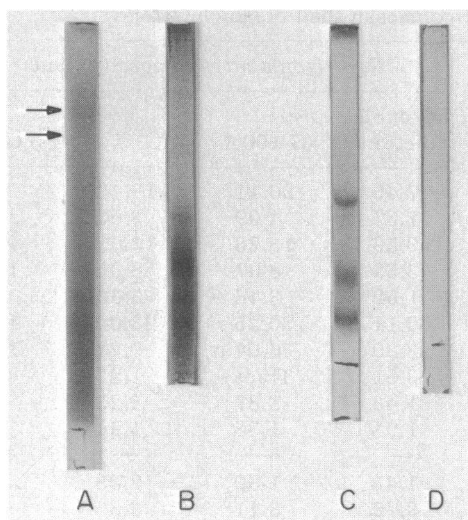

FIG. 2. Polyacrylamide gels of the fraction of fossil *S. thoracica* excluded from Sephadex G-100 (A); the fraction of *S. thoracica* included in Sephadex (B); the fraction of *N. margaritacea* excluded in Sephadex G-25 (C); and the fraction of *N. margaritacea* included in Sephadex G-25 (D). A and C 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.

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,

§ 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.

¶ 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.