

Primary Structure of Myostracal Prism Soluble Protein (MPSP) in Oyster Shell, *Crassostrea gigas*

Seung Woo Lee,¹ Young Moon Kim,¹ Hong Seok Choi,² Jai Myung Yang,² and Cheong Song Choi^{1,3}

Soluble protein (MPSP, myostracal prism soluble protein) obtained from myostracum in oyster shell (*Crassostrea gigas*) was characterized using biochemical and molecular biological techniques. From an analysis of secondary protein structure, it was shown that β -structure was predominant in MPSP. And via *in vitro* assays, the relation of MPSP to biomineral phase and morphology was studied. SDS-PAGE revealed one major protein band of 20 kDa. An amino acid sequence of 160 amino acids was deduced for myostracum by characterization of the complementary DNA encoding the protein. The deduced protein was composed of a high proportion of Gly and Asp, typifying a calcium-binding protein for shell formation, and a relatively high proportion of Val, Ala and Ile, typifying an adhesive protein. In contrast to prevailing expectations, (Gly-Asp)*n*-type sequence motifs exist in MPSP, demanding a revision of previous theories of protein-mineral interactions. The cDNA sequence of myostracum is elucidated for the first time.

KEY WORDS: Myostracum; soluble protein; β -structure; aragonite; oyster shell.

1. INTRODUCTION

The adult shell of the oyster, *Crassostrea gigas*, is composed mainly of calcite. Only two small, distinct, well-defined areas of the shell, the ligament and myostracum, are composed of aragonite (Stenzel, 1963). The place of attachment of the adductor muscle, the muscle scar, is the most conspicuous area on the interior surface of the valves of *C. gigas*. The muscle scar, which is formed at the site of attachment of muscle to the shell, is the exposed point of myostracum. Myostracum consists of prismatic ultrastructures (Fig. 1). As the shell grows, the area of muscle attachment advances leaving behind a 'trail' of myostracum. This is then covered by inner shell layers (Lowenstam and Weiner,

1989). In general, the mollusk shell is mainly composed of two layers, a prismatic and a laminated or nacreous layer. Both layers are in the forms of calcium carbonate crystal; however, in many cases, the prismatic layer is calcite and the laminated layer aragonite (Sarikaya *et al.*, 1995).

Calcium carbonate is the most widespread mineral in invertebrate calcified tissues. The mechanical properties of these tissues, such as strength, hardness, shape and solubility, depend on their organic matrix. Thus, mollusk shells are interesting models for the study of biomineralization processes. In such regulated processes, organic matrices, secreted from the mantle epithelia, have been suggested to play a critical role (Watabe and Wilbur, 1960): the major components of soluble organic matrices

¹ Department of Chemical & Biomolecular Engineering, Sogang University, Seoul, 121-742, Korea.

² Department of Life Science, Sogang University, Seoul, 121-742, Korea.

³ To whom correspondence should be addressed. E-mail: cschoi@sogang.ac.kr

Abbreviations: MPSP, Myostracal Prism Soluble Protein; CD, Circular Dichroism; FT-IR, Fourier Transform Infra-Red; SDS PAGE, Sodium Dodecyl Sulfate Poly-Acrylamide Gel Electrophoresis; RT-PCR, Reverse Transcriptase-Polymerase Chain Reaction; Val, Valine; Gly, Glycine; Asp, Aspartate; Glu, Glutamate; Ala, Alanine; Ile, Isoleucine.

are acid-rich calcium-binding proteins (Cariolou and Morese, 1988; Wheeler and Sikes, 1984). Falini *et al.* (1996) and Belcher *et al.* (1996) reported that soluble proteins, extracted from the nacreous shell layers, induced aragonite formation *in vitro*. In the main, previous studies on soluble proteins of bivalves have tended to concentrate on nacre and pearl.

There is both a sharp contrast and similarity between myostracum of oyster and nacre of abalone: myostracum is composed of prismatic layers, while nacre is of laminated layers; polymorphism of both layers consists of aragonite. However, at present, no soluble protein from myostracum by contrast with that from nacre has been reported. Thus, in this study, MPSP, which is known as a determinant of polymorphism and morphology, was the subject of investigation.

This paper presents the structure of an unknown soluble protein of myostracum, as it relates to polymorphism and morphology of myostracum, in the hope of giving a deeper understanding of the structure and function of soluble proteins in calcium carbonate biomineralization.

2. MATERIALS AND METHODS

2.1. Separation of Myostracum from Adductor Muscle Scar

Shells of *C. gigas* (Namhae in Korea) were freshly collected, soaked in 5% NaOH, lightly scrubbed, and dried at room temperature. Myostracum under the muscle scar consists of prismatic layers from 20 to 30 μm (Fig. 1-b). Thus, the separation of pure myostracum without folia or chalky layer is very difficult. To obtain pure myostracum under the muscle scar, Feigl's solution (Feigl, 1961) that selectively stains aragonite was used. After the muscle scars were stained, they were finely ground and separated using an electric mill, cutting knife. To identify the color and mineralogy of separated particles, optical microscope (VL-11S) were used.

2.2. Preparation and Purification of Soluble Protein

The powder (25 g) of myostracum was extracted with 150 ml of 0.2 M EDTA (pH 8.0). Extraction was performed at 4°C with continuous stirring. After 2 days, the soluble extract was

obtained by centrifugation at $25,000 \times g$ for 20 min. The supernatant solution was diluted with an equal volume of DI water. The diluted solution was concentrated by ultrafiltration with a minimodule. The concentrated fraction was dialyzed against 1 l of DI water at 4°C for 3 days. The dialyzed fraction was concentrated to 20 ml and subjected to ethanol stored at -25°C for 1 week. The precipitate at -5°C with continuous stirring was dissolved in 3 ml of 50 mM $\text{NaHPO}_2 \cdot \text{H}_2\text{O}$ (pH 7.18) and then dialyzed against the DI water at 4°C for 3 days. The dialyzed solution was lyophilized.

2.3. CD (Circular Dichroism)

The purified soluble protein of myostracum was measured on a Jasco 720 spectro-polarimeter in a cell with a 0.1 cm path length. The samples were prepared at 0.1 mg/ml 0.02 mM $\text{NaHPO}_2 \cdot \text{H}_2\text{O}$ (pH 7.18) buffer. The sample was scanned from 200 to 250 nm at 25°C, using 1 nm bandwidth and a scan rate of 1 nm/s. Subtraction of the buffer contribution to background was performed, and the sample was purged with high purity N_2 gas to reduce far-UV spectral contributions arising from dissolved O_2 .

2.4. Crystal Growth

Calcium carbonate was grown by the slow diffusion (3 days) of $(\text{NH}_4)_2\text{CO}_3$ vapor into a cell culture dish containing 2 ml of 7.5 mM CaCl_2 in a desiccator. The effect of the soluble protein from myostracum on crystal morphology and polymorphism was examined by adding them in appropriate amounts ($2 \mu\text{m ml}^{-1}$) to CaCl_2 solutions. The mineralogy of crystal formed in the presence of protein was compared to that of crystals grown in parallel without protein.

2.5. FT-IR (Fourier Transform Infra-Red) Spectrometer Measurements

FT-IR spectrum was obtained using FTS-3000 (Bio-Rad) spectrophotometer under dry air at room temperature using KBr pellets. The resolution was 2 cm^{-1} and the system was purged with dry N_2 to reduce interfacing water vapor IR absorption.

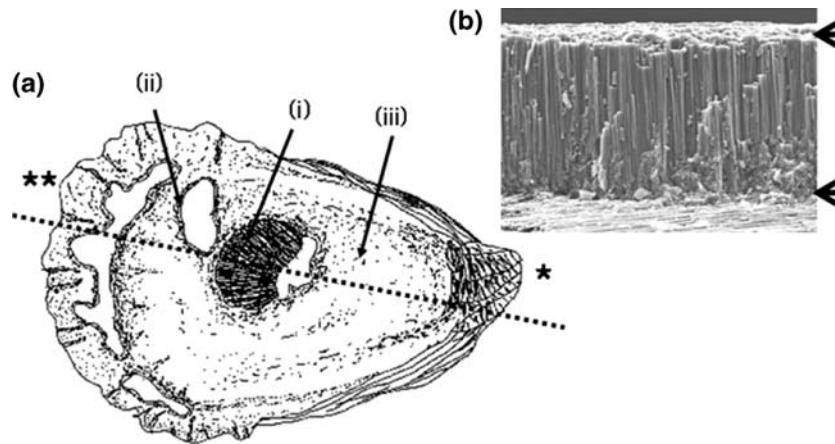


Fig. 1. Diagrammatic representation of oyster shell and scanning electron image of myostracum under the muscle scar. (a) left valve of shell [(i) muscle scar, (ii) chalky layer, and (iii) folia], a star represents anterior (umbo) and double star posterior. (b) Scanning electron microscope of fractured cross-section directional myostracum (black arrow) following as black dot line (Fig. 1-a). HFW = 25 μ m.

2.6. SDS PAGE (Sodium Dodecyl Sulfate Poly-Acrylamide Gel Electrophoresis)

The extracted soluble protein (1 mg per each well) was separated by SDS PAGE, using slab gels of 1.5 mm thickness containing 12% polyacrylamide. After electrophoresis, gels were stained by Coomassie Brilliant Blue R.

2.7. N-Terminal Sequence Determination

Following separation by SDS PAGE, the proteins were electroblotted onto polyvinylidene difluoride membrane in Caps buffer (10 mM, pH 11) containing methanol (12 vol% solution). N-terminal amino acid sequence analysis of the immobilized protein samples was used by Edman degradation using an automated protein sequencer (Perkin-Elmer Applied Biosystems). The N-terminal amino acid sequence, 5'-T-A-D-G-D-D-(D,S)-D-3', was determined at least twice reproduced by different SDS PAGE.

2.8. RNA Purification and RT-PCR (Reverse Transcriptase-Polymerase Chain Reaction)

RNA was extracted from the mantle tissue of a single specimen of *C. gigas* using RNeasy midi kit (Qiagen) following the manufacture's instruction. The RNA (5 mg) was applied as a template for reverse transcription to prepare complementary DNA (cDNA) in a 50- μ l reaction, primed with an oligo-

dT primer. RT (reverse transcriptase) was performed to synthesize cDNA as a template for PCR using the cDNA synthesis kit (M-MLV version (TaKaRa)) following the manufacture's instruction. Forward primer (5'-ACRGCY GAYGGY GAYG-3') for first PCR were designed by the nucleotide sequence determined by N-terminal amino acid sequence (5'-TADGDD-3'), as up to six three-base groups (codons) encoded the same amino acid and oligo-dT (18 nucleotides) is used as reverse primer.

PCR mixture in a final volume of 50 μ l is consisted of 10 mM Tris-Cl (pH 8.3), 50 mM $MgCl_2$, 0.4 mM dNTP mixture, 20 pmole of primers. A Cetus DNA Thermal Cycler (Perkin-Elmer) was employed with an initial step of 94°C for 5 min, then 35 cycles at 94°C for 30 s, 52.6°C for 45 s, 72°C for 1 min, followed by a final extension step of 72°C for 7 min. Second PCR was performed as described above using the forward primer (5'-ACGGCCGATGGTGACG-3') and the reverse primer (5'-GACGGGCTTGCCATC-3') which were designed from the sequencing data.

2.9. Amplification and Sequencing of cDNA 39-End

The PCR product was eluted using gel extraction kit (Qiagen), and ligated with pGEM T easy vector system (Promega) to amplify in *E. coli*. The product of ligation was sequenced by the chain termination method using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and an automated DNA sequencer (Perkin-Elmer Applied

Biosystems) primed with T7 promoter primer. The result of sequencing was translated by translation tool served by 'Justbio' website (<http://www.justbio.com>) and the secondary protein structure was analyzed by Protean (DNASTAR Inc.).

3. RESULTS

Shells of *C. gigas* (Namhae in Korea) were freshly collected, soaked in 5% NaOH, lightly scrubbed and dried at room temperature. Myostracum under the muscle scar consists of prismatic layers of 20–30 μm thick (Fig. 1b); thus, the separation of pure myostracum without folia or chalky layer is very difficult. To obtain pure myostracum under the muscle scar, Feigl's solution (Feigl, 1961) was used, which selectively stains aragonite. After muscle scars were stained, they were finely ground and separated using an electric mill, cutting knife. An optical microscope (VL-11S) was used to identify the color and mineralogy of the separated particles.

The shell of the juvenile and adult oyster consists mainly of calcite, except the ligament and myostracum, which are aragonitic layers, possibly as a continuation of folia of the valves. The muscle scar is the surface of the myostracum support for attachment of the adductor muscle (Fig. 1a) and is composed of regular and simple prisms (Fig. 1b). The myostracum in an anterior direction is deeply embedded in the folia and chalky layer and, thus, produces a steep layered slope that extends to the umbo (Carriker, 1996). Myostracal prisms are present in the entire myostracum zone, and in *C. gigas* attains a thickness of 25–30 μm (Fig. 1b).

The soluble protein extracted from myostracum was analyzed by SDS-PAGE (Fig. 2a), revealing one major protein band of 20 kDa when stained with Coomassie Brilliant Blue. In PCR (polymerase chain reaction), using a primer pair, only one component (500 bp) was amplified (Fig. 2b).

To obtain preliminary structural information on the polypeptide, the secondary structure preferences for MPSP in solution were quantitatively assessed using far-UV CD spectrometry (Fig. 3). As shown in Fig. 3a, the CD spectra of MPSP exhibits a broad negative band centered from 205 to 215 nm, which is usually associated with folding polypeptide structures. This result indicates that the MPSP does not possess a significant number of random coils, but adopts a β -structure conformation in solution. Turn is characterized by an intense nega-

tive band at 205 nm due to π - π^* transition and another high intensity band near 212 nm due to n - π^* transition (Johnson, 1985). The CD result differs from the structural features observed in other calcium carbonate mineral-binding sequences (Gerbaud *et al.*, 2000; Wustman *et al.*, 2002). The fractions are calculated as 18.4% α -helix, 41.8% β -structure, 27.8% turn and 12% random coil. Molecular ellipticity is expressed in $\text{deg cm}^2 \text{dmol}^{-1}$ per mole peptide.

The morphology of synthetic crystal used by the soluble protein obtained from myostracum consists of needle-like layers typically less than 80 nm (Fig. 4-a). Taylor *et al.* (1995) reported that the characteristic peak of aragonite is at 857 and 1082 cm^{-1} . As shown in Fig. 4-b, the synthetic crystal is identified as aragonite.

The complete nucleotide sequence of MPSP shows 480-bp cDNA encoding (Fig. 5). The nucleotide sequence revealed an open reading frame of 160 amino acids with translation. The deduced amino acid sequence contained a high proportion of valine (9.5%), glycine (7.6%) and acidic residues (Asp 19.1% and Glu 6.3%), in agreement with the bulk composition. The iso-electric point of the deduced amino acid sequence of MPSP is 3.9, as expected for this Asp-rich protein. The calculated molecular weight for MPSP, assuming no posttranslational modifications, is 18 kDa. The α -helix and β -structure of MPSP were predicted by analyses of the secondary structure, using the method of Chou and Fasman (1978).

4. DISCUSSION

It has been reported that soluble proteins play a key role of determining crystal polymorphism during shell formation (Belcher *et al.* 1996; Choi and Kim, 2000; Falini *et al.*, 1996). In an experiment on crystal growth of soluble protein obtained from myostracum in the shell of the Pacific oyster (*C. gigas*), the polymorph of the synthetic crystal is aragonite and is similar in shape to the myostracal prism (Fig. 4). Thus, it can be found that the soluble protein from myostracum is key factor to determine polymorphism and morphology of myostracum.

A β -structure is the major component of structural proteins, forming a stereochemical model for the controlled nucleation of calcium carbonate (Addadi and Weiner, 1985). In other words, a

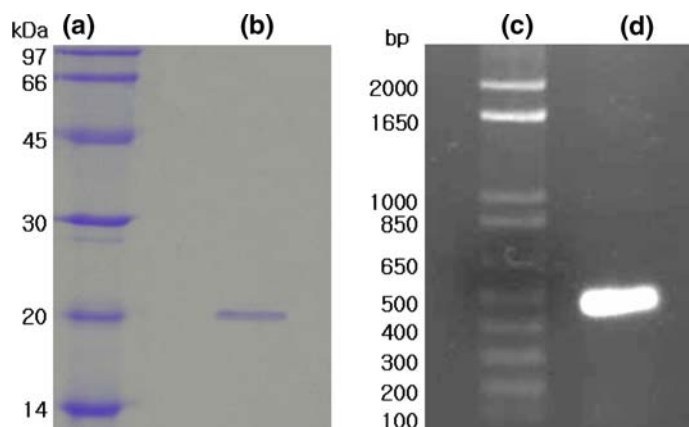


Fig. 2. Electrophoretic pattern of MPSP by SDS-PAGE. (a) marker and (b) the purified soluble protein of myostracum. The major peak fraction was separated on an SDS-PAGE by Coomassie brilliant blue staining. The sample was prepared by Material and Methods (Separation and purification of soluble protein); Electrophoretic pattern of PCR product. (c) molecular weight standards and (d) PCR product. A detected signal of about 500 bp is indicated.

β -structure, which adopts regular repeating negative charges, could bind calcium ions. Also, it has been shown that polypeptide could have a dramatic effect on aragonite crystal morphology and the effect is conformation-dependent (Weiner and Traub, 1984). Therefore, it is necessary to understand the relationship/association between the secondary structures of proteins and carbonate ions to definitively understand the shell-forming process. The deduced amino acid sequence revealed the secondary protein structures of MPSP, in which a highly conserved unit is repeated by α -helix, β -sheet and turn (Fig. 3). Repeating sequence motifs, such as the (Asp-X) n -type (Runneger, 1984; Weiner and Hood, 1975) and (Asp) n -type (Wheeler, 1992), have been predicted for molluscan soluble matrix proteins, combined with their functions. It has been postulated that Asp-rich domains in shell proteins play a role as templates on which epitaxial growth of the mineral phase takes place (Weiner and Hood, 1975).

It shows in Fig. 5, MPSP is rich in Asp, which is consistent with previous research on soluble shell proteins. The above results show that most Asp, which is known to play a role in the calcium-binding domain of MPSP, exists in the α -helix and turn domains rather than in the β -structure domain (Chou and Fasman, 1978; data not shown). This means that the association not only between calcium ions and the β -structure, but also the α -helix and turn domains, can act as a variable in the shell-forming process.

Glycine is an effective secondary structure-breaking amino acid in the α -helix and β -sheet (Chou and Fasman, 1978). It has been demonstrated that, in the case of *C. gigas*, MPSP has an open conformation, which agrees with the high content of glycine. Meanwhile, it has been reported that each concentration of valine, alanine and isoleucine was relatively greater than the existing level of soluble protein in oyster (Fig. 4) (Wheeler, 1992). Burzio

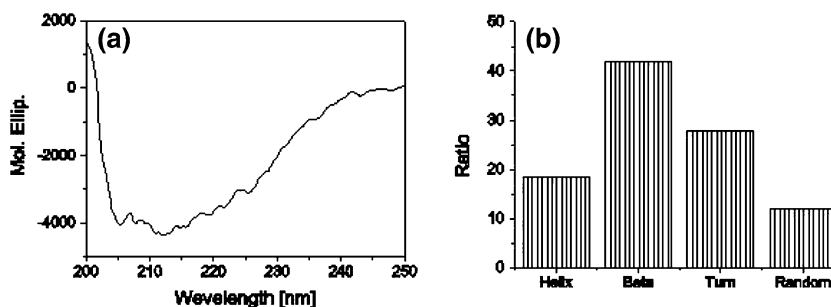


Fig. 3. CD (a) spectra and (b) fractions of secondary structure of soluble protein of myostracum.

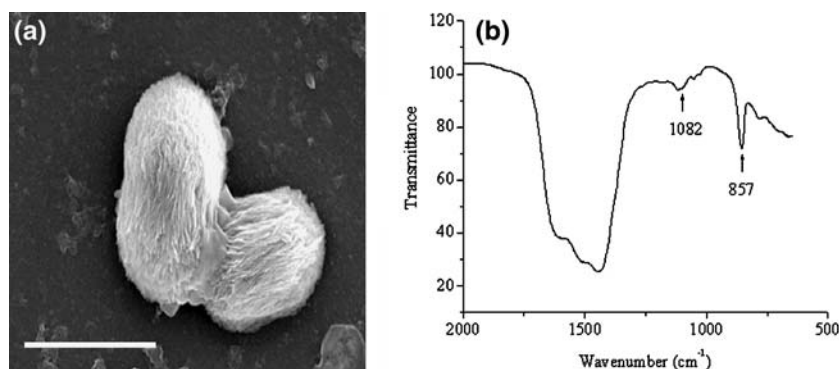


Fig. 4. (a) Synthetic crystal obtained from soluble protein of myostracum (scale bar: 10 μm) and (b) FT-IR spectra of synthetic crystal.

et al. (2000) reported that the major peptides of adhesive protein in *Aulacomya ater* contain seven amino acids corresponding to the consensus sequence AGYGGXK, whereas the X residue in position 6 was either valine, leucine or isoleucine, the carboxyl terminal residue was either lysine or hydroxylysine.

Therefore, it can be deduced that the substantial levels of the amino acids, valine and isoleucine, are associated with the functional characteristics of myostracum in direct contact with the adduct muscle.

Also, a noticeable point is that the repetitive structure of Gly and Asp (Fig. 4, underline) exist within MPSP.

The GD repeats are very similar to the acidic Gly-Xaa-Asn (Xaa = Asp, Asn or Glu) of nacrein (Miyamoto *et al.*, 1996). They reported that the Gly-Xaa-Asn is correlated with calcium-binding protein. Moreover, Gly-Xaa-Yaa (Yaa = any amino acid) repeats in the collagenous domain are also found in a structural protein of the inner ear, which

produces calcium carbonate crystals (Davis *et al.*, 1995). Possibly, the GD repeats could contribute to crystal growth by forming three-dimensional lattices or mesh works, like collagen molecules.

To find known proteins that have similar sequences to MPSP, a homology search was carried out against all protein sequences stored in the NCBI databank. There are no exact similarities with previous soluble proteins, including aspein of *Atrina rigida* (Gotliv *et al.*, 2005). However, the domains around the GD repeats of MPSP showed the highest sequence similarity with the Gly-Asp repeats of aspein, a soluble protein of the pearl oyster, *Pinctada fucata* (Tsukamoto *et al.*, 2004).

Although the soluble acidic matrix protein (MPSP) of oyster shell is believed to play an important role in the bio-fabrication of myostracum, its complete function and domain structures have not been characterized. Elucidation of the exact function in relation to the structure of MPSP and other acidic shell proteins awaits further investigation.

```

1  ACG GCC GAT GGT GAC GAT GAC GAT GAC GAA GTA GAT GCT ACC CTC TAT GAG
   T A D G D D D D E V D A T L Y E
52  AAT ATC CTC GGT AGG GAT TCT GTA CAC GAT GAT AGG CAC GAA TCC ACG ATC
   N I L G R D S V H D D I H E S T I
103 GGA GTC GAT GAT GCC GAT GGC GAT GAC ATC AAC GCC GAT GAT GAC GAT CGC
   G V D D A D G D D I N A D D D D R
154 ATC GAA GGC TCG CCC ATC GAG GCC GTC GAG ATC CGC TCG GTA CTC ACC TGT
   I E G S P I E A V E I R S V L T C
205 GAG TCT ATG AAG GGT GAT TGC GCT AAG TGC TAC GGC CGT AAC CTC GCT ACG
   E S M K G D C A K C Y G R N L A T
256 AAC ATC CTC GTC CAG AAG GGT GAT GTC GTC GGG GAT GTC ATC GCT GCT GAA
   N I L V Q K G D V V G D V I A A E
307 TCT ATC GGT GAG CCT GGT ACA CAG CTG ACG CTC AAT ACC TTC CAC GTC GGG
   S I G E P G T Q L T L N T F H V G
358 GGG GAT GAT TCG AAC GTC GCT ACG CGC AAT AGT GTC GCT GCT AAG TAC GAT
   G D D S N V A T R N S V A A K Y D
409 GGT ATC GTC ACC TTC GAT GAC CTG CGT GTC GGC GAA GTG GGT TCC GAA GAT
   G I V T F D D L R V G E V G S E D
460 GGT GAT GAT GGC AAG CCC GTC TAA
   G D D G K P V *

```

Fig. 5. Partial cDNA sequence of MPSP and the deduced amino-acid sequence. Numbers on the left indicate positions of the nucleotides in the MPSP cDNA sequence.

REFERENCES

- Addadi, L., and Weiner, S. (1985). *Proc. Natl. Acad. Sci. USA* **82**: 4110–4114.
- Belcher, A. M., Wu, X. H., Christensen, R. J., Hansma, P. K., Stucky, G. D., and Morse, D. E. (1996). *Nature* **381**: 56–58.
- Burzio, L. A., Saez, C., Pardo, J., Waite, J. H., and Burzio, L. O. (2000). *Biochimica et Biophysica Acta* **1479**: 315–320.
- Cariolou, M. A., and Morse, D. E. (1988). *J. Comp. Physiol. B* **157**: 717–729.
- Carriker M. R. (1996). In: Kennedy, V. S., Newell, R., and Eble, I. E., (eds.), *The Eastern Oyster, Crassostrea virginica*, Maryland Sea Grant College, pp 75–93.
- Choi, C. S., and Kim, Y. M. (2000). *Biomaterials* **21**: 213–222.
- Chou, P. Y., and Fasman, G. D. (1978). *Adv. Enzymol.* **47**: 45–148.
- Davis, J. G., Oberholtzer, J. C., Burns, F. R., and Greene, M. L. (1995). *Science* **267**: 1031–1034.
- Falini, G., Albeck, S., Weiner, S., and Addadi, L. (1996). *Science* **271**: 67–69.

- Feigl F. (1961) Spot tests in inorganic analysis 5th edn. Elsevier Publishing Company, Maruzen Company.
- Gerbaud, V., Pignol, D., Loret, E., Bertrand, J. A., Berland, Y., Foontecilla-Camps, J. C., Canselier, J. P., Gabas, N., and Verdier, J. M. (2000). *J. Biol. Chem.* **275**: 1057–1064.
- Gotliv, B. A., Kessler, N., Sumeral, J. L., Morse, D. E., Tuross, N., Addadi, L., and Weiner, S. (2005). *ChemBioChem.* **6**: 304–314.
- Johnson W. C. Jr. (1985) *Methods Biochem. Anal.*, New York, vol. 31 pp 61–163.
- Lowenstam, H. A., and Weiner, S. (1989) *On Biomineralization*. New York: Oxford University Press 103–110.
- Miyamoto, H., Miyashita, T., Okushima, M., Nakano, S., Morita, T., and Matsushiro, A. (1996). *Proc. Natl. Acad. Sci. USA* **93**: 9657–9660.
- Runneger, B. (1984) *Alcheringa* **8**: 273–290.
- Sarikaya M., Liu J., Aksay I. A. (1995). In: Sarikaya M. and Aksay I. A. (eds), *Biomimetics design and processing of materials*. American Institute of Physics, AIP Press, pp 35–90.
- Stenzel, H. B. (1963) *Science* **142**: 232–233.
- Taylor, D. R., Crowther, R. S., Cozart, J. C., Sharrock, P., Wu, J., and Soloway, R. D. (1995). *Hepatology* **22**: 488–496.
- Tsukamoto, D., Sarashina, I., and Endo, K. (2004). *Biochem. Biophys. Res. Commun.* **320**: 1175–1180.
- Watabe, N., and Wilbur, K. M. (1960). *Nature (London)* **188**: 334.
- Weiner, S., and Hood, L. (1975). *Science* **190**: 987–989.
- Weiner, S., and Traub, W. (1984). *Philos. Trans. R. Soc. Lond. Ser. B* **304**: 425–434.
- Wheeler A. P. (1992). abcd. In: Suga S. and Watabe N. (eds), *Hard Tissue Mineralization and Demineralization*. New York, Springer-Verlag, pp 171–187.
- Wheeler, A. P., and Sikes, C. S. (1984). *Am. Zool.* **24**: 933–944.
- Wustman, B. A., Weaver, J., Morse, D. E., and Evans, J. S. (2002). *Langmuir* **19**: 9373–9381.