RECENT ADVANCES IN STUDIES ON NACREOUS LAYER BIOMINERALIZATION. MOLECULAR AND CELLULAR ASPECTS

TETSURO SAMATA(1)

Keywords: Cell culture; Gene analysis; in vitro crystallization experiment; Nacreous layer; Organic matrix.

ABSTRACT

Recent advances in biotechnology promoted evolution in the general study of life sciences, and more specifically, in the study of biomineralization. Here I review information from recently developed studies on molecular and cellular aspects of molluscan shell nacreous layer formation. First, I summarize the primary and tertiary structures of the organic matrix (OM) proteins recently identified on the basis of cDNA cloning and classify them into several groups. Further discussion is focused on the possible function of the OM proteins involved in the regulation of CaCO3 polymorphism by in vitro examination of crystallization. In addition, several new findings from in vitro and in vivo studies of gene expression and cell culture are discussed in relation to biomineralization. In this manner, a current overview of the process of molluscan shell nacreous layer formation on the molecular level is presented.

INTRODUCTION

Hard tissue formation by organisms is commonly regulated by many kinds of macromolecules constituting an OM. The OM of molluscan shells is one of the best studied of all CaCO3 biomineral structures, and considerable information has been accumulated regarding its morphology, structure, and histochemical properties (Wada, 1964; Grégoire, 1972; Weiner & Hood, 1975; Weiner, 1979; Nakahara et al., 1980, 1982; Samata, 1988). An OM is roughly classified into a soluble matrix (SM) and an insoluble matrix (ISM) based on its solubility in solutions. Some confusion has arisen over the names of OM-fractions based on their solubility in solutions. Some confusion has arisen over the names of OM-fractions based on their solubility in solutions and no definitive nomenclature has yet been established. The general types of soluble proteins referred to in the field of biomineralization are: 1) proteins soluble in the solution used for decalcification such as EDTA and dilute acid; 2) soluble protein in water used for dialysis after decalcification, such as water-soluble protein; and 3) solubilized proteins incorporated into insoluble proteins by using dense urea solution or detergents. It is worth noting that it is not clear whether the various methods of decalcification release the same molecules. In fact, soluble matrix in Mili-Q water without any demineralization process contained components with completely different amino acid compositions with

(1) Laboratory of Cell Biology, Faculty of Environmental Health, Azabu University 1-17-71 Fuchinobe, Sagamihara 229-8501 Japan e-mail: samata@azabu-u.ac.jp
those soluble in EDTA (Pereira-Mouries, in press). In this paper, fractions that remain in the supernatant after centrifugation of the components corresponding to 1) and 2) are called as the soluble matrix (SM), and those that remain in the precipitate are called the insoluble matrix (ISM)

From the viewpoint of morphological designation, OMs have also been divided into several classes. These include the interlamellar matrix, a fibrous sheet that lies between the brick-wall of crystals, and the intercrystalline matrix surrounding each crystal, typically seen in the nacreous layer (Watabe, 1965). The crystal-facing layers of these two matrices have been suggested to be an envelope in which crystal nucleus formation and growth might occur (Nakahara, 1979).

Much work has been carried out to fractionate the SM into different components (Krampitz et al., 1976; Samata & Krampitz, 1982; Weiner, 1983). Since Crenshaw isolated a Ca-binding glycoprotein from the shell of Mercenaria mercenaria (Crenshaw, 1972), Ca-binding property has been assigned to the SM of many molluscan species (Samata & Krampitz, 1982; Cariolou & Morse, 1988; Samata, 1990) and some of the Ca-binding proteins have been purified to homogeneity and subjected to amino acid sequence (SQ) analysis of their N-terminal regions (Marxen & Becker, 1997; Matsushiro et al., 2003).

Hitherto, several genes encoding the matrix components in molluscan shells have been isolated and their deduced amino acid SQs clarified (Miyamoto et al., 1996; Sudo et al., 1997; Shen et al., 1997; Samata et al., 1999; Kono et al., 2000; Mann et al., 2000; Marín et al., 2000; Miyashita et al., 2000; Sarashina & Endo, 2001; Weiss et al., 2001). However, information available is confined to the nacreous, prismatic and foliated layers while information in the other shell layers is lacking on the gene level.

Other data were obtained through in vitro studies of crystallization in which matrix activities related to crystal formation were measured (Wheeler, et al., 1981, 1987; Greenfield, et al., 1984; Albeck et al., 1996; Belcher et al., 1996; Falini et al., 1996; Thompson et al., 2000). Other experiments have also been conducted specifically to investigate the nacreous layer, which has a specific geometry, in order to gain an understanding of the effect of OM on shell formation. However, results of various experiments were remarkably different depending on the methods used. Therefore, the precise function of OM remains still unclear even in vitro.

Little is known about the cellular aspects of molluscan shell formation. Studies have conducted structural and histochemical analyses of the mantle epithelium tissue (Tsujii, 1976; Zylstra et al., 1978), which is responsible for shell formation, and determination of distribution of enzymes such as alkaline phosphatase and carbonic anhydrase (CA) in the same tissue sample with a focus on their function in shell formation (Chun & Saleuddin 1974; Boer & Wittereen 1980). In addition, involvement of mammalian insulin and insulin-like peptides in shell formation of Helisoma duryi was discussed (Saleuddin et al., 1992). Several new data about the nature of the epithelial cells have been accumulated from primary tissue culture of the mantle epithelium (Machii 1985; Awaji 1991, 1997; Awaji & Suzuki, 1998). Additional reports related to OM secretion and crystal formation during cell culture (Machii, 1985; Machii & Wada, 1989; Samata et al., 1994) may indicate a new approach to the study of biomineralization.

The present report summarizes the results of the recent researches and attempts to identify possible roles of the OM components in molluscan shell nacreous layer biomineralization.

1. Analysis of the structure of OM at the molecular level

Following the discovery of the genes encoding a major component in the nacreous layer of Pinctada fucata by Miyamoto et al. (1996), the primary structures of about 10 proteins were determined using the technique of cDNA cloning. Significant aspects of these molecules will be summarized in the following sections.

1-1) Primary structures of the OM proteins

The following nomenclatures of the protein groups are still tentative, and they should be redefined after enough data of the primary structures of the molluscan OM components for grouping them can be obtained.
1-1-1) Proteins in ISM

1-1-1a) Proteins in ISM in the nacreous layer

Three proteins (MSI60, N16, and N14) reported so far are classified into this group. MSI60 (Sudo et al., 1997), isolated from the nacreous layer of *P. fucata*, was found to be composed of 738 amino acids. N16 family (Samata et al., 1999), isolated from the nacreous layer of the same species was composed of 129-131 amino acids. N14 (Kono et al., 2000), isolated from the nacreous layer of *P. maxima*, was found to be composed of 140 amino acids. Perlin, reported on by Miyashita et al. (2000), is homologous with a component of the N16 family.

MSI60, the predominant constituent of the insoluble protein of the nacreous layer, contains 11 poly Ala blocks and two Ala-rich domains, which are located between poly Gly-rich regions, and shows high homology with drag line silks. MSI60 is one of the fundamental components constituting so-called "conchiolin" (Frémy, 1855).

N16 and N14 are specific to the nacreous layer of *P. fucata* and *P. maxima* which can be isolated by the treatment of dilute alkali from the main part of ISM. RT-PCR analysis of the N14 gene clearly indicated that the expression of this gene was enriched in the inner part of the mantle epithelium, which is involved in nacreous layer formation (Kono et al., 2000) (Fig.1). They are characterized by high proportions of Cys, Tyr, Gly and Asn, together with a specific repeat of Asn and Gly (NG-repeat SQ), which had high homology with NG-repeat domain of nacrein and N66. Acidic amino acids are present in high concentrations in four acidic regions. Moreover, the contents of aromatic amino acids are also high. BLAST searches revealed limited homology of N16 with other known proteins, the only exception being partial SQ similarity with high Gly/Tyr proteins (HGTs) in the rabbit hair keratin family (Fratini et al., 1993). At least 6 isoforms of N16 (named N16 family) have been identified (Samata et al., 1999).

N16 or N14 may form a complex structure with MSI60, as evidenced by the high affinity of these two molecules (Suganaga, 2003), and thus may form ISM membrane together with some other components such as chitin and unidentified protein components.
The possible function of these molecules will be discussed in a later section.

1-1-1b) Protein contained in ISM in the prismatic layer

MSI31 (Sudo et al., 1997), isolated from the prismatic layer of _P. fucata_ and composed of 334 amino acids, is classified into this group. It may be the predominant constituent of the insoluble framework protein in the prismatic layer. It is characterized in the N-terminal half by 10 poly Gly blocks composed of 3–5 Gly residues and in the C-terminal half by a large acidic region including six consecutive ESEEDX (X is an any amino acid) motifs. No homology of this molecule with any other known proteins has ever been determined.

Even though MSI60 and MSI31 constitute ISM in the nacreous and prismatic layers of _P. fucata_, there was no available information about the maturing process of these molecules to ISM after expression from the genes.

1-1-2) CA-like proteins

Two proteins (nacrein and N66) are classified into this group. Nacrein (Miyamoto et al., 1996), isolated from the nacreous and prismatic layers of _P. fucata_ (Samata & Kono, 2000), was found to be composed of 447 amino acids. N66 (Kono et al., 2000), isolated from the same shell layers of _P. maxima_, was found to be composed of 568 amino acids.

These proteins contain two functional domains: one is a carbonic anhydrase-like domain (CA domain), which is split into two sub-domains with insertion of a repeat domain rich in Asn and Gly (NG-repeat domain). The alignment and length of the NG-repeat SQ is species-dependent.

CA is one of the most important enzymes for molluscan shell formation catalyzing the interconversion between CO$_2$ and HCO$_3$$. The highest CA-activity measured in the gills, followed by that in the mantle epithelial tissue (Kawai, 1954). Recent analysis (Suganaga, 2003) of RT-PCR using primers designed based on the SQ of human CAII clarified the presence of a CA-like component in the adductor muscle of _P. fucata_. It had a high homology with nacrein, not only in terms of the CA domain, but also the NG-repeat domain. This finding implies that the so-called "molluscan CA" may be a molecule classified into the CA-like proteins, distributed not only in the mantle epithelium, but also in the gill or adductor muscle, indicating that nacrein itself may not be specifically responsible for shell formation. Furthermore, the fact that nacrein and N66 have been found in both the aragonitic nacreous and the calcitic prismatic layers suggests that the specific involvement of these molecules in the nucleation of aragonite crystal is doubtful.

These molecules may play a role as a kind of the components related to CO$_3$- concentration, and it is possible that they are universally distributed in molluscan shells. Consequently, search for this kind of protein in the shell layers except for the nacreous and prismatic layers must be one of an important subject of cDNA research in future.

1-1-3) Asp-rich protein

MSP-1 (Sarashina & Endo, 2001), isolated from the foliated layer of _Patinoprecten yessoensis_ and composed of 840 amino acids, is classified into this group.

This component is specific for the foliated layer of _P. yessoensis_. High similarity of amino acid composition between MSP-1 and the unpurified SM in oyster shell (Wheeler et al., 1988) implies the possible presence of MSP-1 or its relatives in the latter. MSP-1 includes a high proportion of Ser, Gly, and Asp, with a repeated unit comprised of 3 modules: the SG, D, and K domains. The D domain is rich in potential N-glycosylation and phosphorylation sites.

MSP-1 shows the highest SQ similarity to phosphophoryn, a major component of non-collageneous proteins in dentin (Gu et al., 1998). In addition, the SG-domain of MSP-1 shows similarity to the GS-domain of Lustrin A, implying a functional similarity between these two proteins. Sarashina & Endo (2001) assumed that the rigid arrangement of COO groups of Asp in the D domain might be specific for inhibiting or orienting the nucleation of calcite.
1-1-4) Proteins with amino acid SQs similar to those of known proteins

1-1-4a) Mucoperlin (Marin et al., 2000)

Mucoperlin, isolated from the nacreous layer of *Pinna nobilis* and composed of 636 amino acids, may be specific to the nacreous layer. It has many of the structural features characteristic of mucins, such as a core with tandem repeats rich in Ser and Thr (Turner et al., 1995), which provide O-glycosylation sites, and rich in Pro, which accounts for its rigid rod-like structure.

The immunohistochemical observation that mucoperlin surrounds each aragonite crystal of the nacreous layer might indicate that its primary role is to terminate crystal growth and combine the crystals. This function may be facilitated by mucoperlin's highly glycosylated nature, which may enable it to act as some kind of glue.

1-1-4b) Perlucin (Mann et al., 2000)

Perlucin, isolated from the nacreous layer of *Haliotis laevigata* and composed of 155 amino acids, had amino acid SQ similar to the C-type carbohydrate-recognition domains of asialoglycoprotein receptors of various organisms. It was estimated to be a functional C-type lectin with broad carbohydrate-binding specificity.

1-1-4c) Perlustrin (Weiss et al. 2001)

Perlustrin, isolated from the nacreous layer of *Haliotis laevigata* and composed of 82 amino acids, had SQ similarity to the N-terminal domain of mammalian insulin-like growth factor binding proteins (IGFBPs). In fact, it showed a high capacity of binding human IGFs.

1-1-5) Proteins with modular structure

Lustrin A (Shen et al., 1997), isolated from the nacreous layer of *Haliotis rufescens* and composed of 1428 amino acids, revealed a highly modular structure with Gly-, Ser-, Cys- and Pro-rich domains. The alternating Cys-rich and Pro-rich module design is seen in frustulins (Kröger et al., 1996), which are found in the silificed cell wall of diatoms. This modular structure of Lustrin A suggests that it has multiple functions such as 1) a framework structural factor interacting with other molecules; 2) a factor protecting the organic matrix against degradation and 3) conforming elasticity to the shell. MSP-1 is the other protein with repeated modular structure.

The aforementioned proteins have very different characters in their structures, because gene analyses have been carried out independently on genes encoding the main OM proteins of the unrelated species. For the next step of gene analysis, the search for the homologous genes between taxonomically separated groups of molluscs is required. Further analysis of gene expression and determination of the whole gene structures containing the promoter regions of the already identified genes should be carried out.

2. Which moieties in OM are responsible for shell formation?

As it has been described for many molluscan shells, the protein moieties are considered the first candidate for shell formation. In addition, sugar components are thought to be important for shell formation (Albeck et al., 1996). Studies have focused on chitin at an initial stage of research (Peters, 1972). Chitin has been thought to be present in the core of the interlamellar matrix (ISM) of the nacreous layer (Nakahara, 1979), composing a tight structural relationship between protein sheet and aragonite.
crystal, allowing for a close geometric match between the lattice spacing of aragonite and the atomic spacing of OM (Weiner & Traub, 1981). Levi-Kalisman et al. (2001) found that β-chitin is one of the constituents of ISM of the nacreous layer of *Atrina* arranged at higher level of organization. Based on an *in vitro* crystallization experiment, Albeck et al. (1996) suggested that polysaccharide moieties of glycoprotein might be a controlling factor of calcite growth. Prior to that, Wada (1980) had demonstrated that the crystal nucleus formation might take place in specific organic granules composed of sulfated glycoprotein, which would provide an active surface to serve as a template for nucleation. Mucoperlin (Marin et al. 2000), which might be heavily glycosylated, was speculated to surround the nacreous tablets and terminate crystal growth. Daufin (2003) reported difference in sugar-protein ratios between the SM isolated from the calcitic prismatic layer of *Pinna nobilis* and that of *Pinctada margaritifera*.

Despite the common presence of sugar moieties as glycoproteins in molluscan shells, their function remains speculative due to the paucity of available information. Consequently, the precise identification of sugar components and their functional analysis is necessary for the next step of the experiment.

Data on the lipid components in OM of molluscan shells has not yet been accumulated. Samata & Ogura (1997) separated the lipid component from the ISM of the nacreous layer of *P. fucata* and clarified that it was mainly composed of monoelaidic acid, a kind of oleic acid. On the contrary, phospholipid, which was reported in the skeleton of hard corals (Isa & Okazaki, 1987) found to be absent in *P. fucata*.

3. Protein structure and shell formation

3-1) Primary structure and shell formation

Prior to studies at the molecular level, much of the evidence regarding the primary structure of the molluscan OM components was considered to indicate the presence of an Asp-rich domain, based on findings of Asx-rich compositions in OM (Weiner & Hood, 1975, Runneger, 1984). It is important to note that amino acid compositional analysis does not distinguish between Asp and Asn.
Several hypotheses have been proposed to describe the process of shell formation. One hypothesis that has been advanced takes into account the possible effect of the high Asp content on the OM components of molluscan shell (Weiner 1979, Weiner & Traub 1984, Weiner & Addadi 1991). In particular, an amino acid SQ of D-X-D (X is a neutral amino acid) along the β-sheet framework was thought to be able to create the binding configuration for Ca\(^{2+}\) at the OM interface (Mann, 1988). Nacrein has a NG-repeat domain containing a G-X-N (X=D, N, G) repeat SQ, and was thought by Miyamoto et al. (1996) to be the major Asp-rich Ca-binding protein of P. fucata. However, within this domain, Asn residues are more common than Asp residues. Moreover, the NG-repeat domain of N66, another type of the CA-like proteins, which was isolated from P. maxima, a close relative of P. fucata, contains only 8 Asp residues against 120 Asn, showing 93% of the result of calculating as Asx by amino acid compositional analysis was Asn. The data suggests that the CA-like proteins cannot be representative molecules of Asp model contrary to the opinion of Miyamoto et al. (1996).

In contrast to the CA-like proteins, MSP-1 may be one of a model molecule of the Asp-rich protein. The Asp residues are arranged regularly in the D domains in core units of DGSD or DSD. In view of the regularity of the Asp in the D domains, Sarashina & Endo (2001) remarked that this molecule might be capable of specifically nucleating calcite.

Glu has not received very much attention, whereas Asp has been spotlighted, due to the fact that domains or SQs rich in Glu have been reported to be much less prevalent in molluscan OM. The Glu-rich components have been recognized as MS131, which exhibits a repeating motif of ESEEDX (X is any amino acid) and as N16, which contains acidic regions in which Glu was more common used than Asp. The reason why Asp is preferentially used as acidic amino acids in the OM components of molluscan shell is unknown.

In addition, attention must be paid to the site of phosphorylation as reported in Ostrea virginica (Wheeler, 1992) and P. vesseensis (Sarashina & Endo, 2001), whose shells are composed of the calcitic foliated layer. It is reasonable to assume that the negatively
charged phosphate side chain, which binds with Ser or Thr, may concentrate Ca\(^{2+}\) in a manner similar to that of phosphophorin (Gu et al., 1998) and participate in shell formation. Although phosphorylation sites can be commonly detected in the molluscan OM components by motif analysis, it remains unclear which of the residues actually bind phosphate.

3-2) Secondary and tertiary structures and shell formation

Unfortunately, there are very few essential studies that might provide insight into the secondary and tertiary structures of the OM components in molluscan shell. At least, it is clear, based on the primary structures identified until now, that the tertiary structure which is seen typically in the EF hand Ca-binding protein cannot be identified in the molluscan OM proteins.

The OM proteins in molluscan shells share the common structural feature of having β-sheet conformation. Infrared spectrum analysis showed that the ISM in the nacreous layer of Pelecypoda favored antiparallel β-sheet conformation, gastropoda a random coil or α-helix (Hotta, 1973). In addition, OMs in the crossed-lamellar layer were estimated to adopt in α-helix in Cypraea, Phalium, Strombus and Dosinia, in contrast to turns in Tridacna (Daufin & Denis, 2000). Weiner & Traub (1980) demonstrated the idea that Asp-rich OM proteins were specified by the antiparallel or parallel β-sheet conformation. Analysis of the secondary structure of N16 using Chou and Fasman's method (Chou & Fasman, 1978) indicated that the NG-repeat SQ was rich in turns (Samata et al., 1999). As shown in Fig.2, the tertiary structure of N16, which was estimated using a computer, contained a large area of Gly-loop (Samata unpublished data). The high proportion of aromatic amino acids supported this estimation. A similar Gly-loop structure was estimated to be present in the "GS" domain of Lustrin A.

3-3) Repeated SQs present in the OM proteins

Various types of the repeated SQs, which show difference in composition and the number of core units, have been reported in the OMs of molluscan shells. A high degree of conservation of the amino acid SQ comprising a core unit was characteristic to each molecule. Examples of the repeated SQ recognized in multiple molecules are the NG-repeat SQ in CA-like proteins and N16/N14, and the GS (or SG)-repeat SQ in Lustrin A and MSP-1. It is not clearly understood whether the tandem repeats of NG and GS (SG) or Gly-rich repeat relate to any function. However, Gly-rich SQs can easily form a Gly-loop structure, which may extend from the globular cores of these molecules to place this area in contact with other components, such as a protein (glycoprotein), sugar (chitin) or CaCO\(_3\) crystal. Sarashina & Endo (2001) anticipated that the repeated SQs in MSP-1 might have functional effects on crystal formation with only a small number of the molecules by extending and covering crystal surfaces. In addition, in the case of the NG-repeat SQ, the length of the repeated region matches the shell-size of the species examined (Kono et al., 2000), implying that this region is critical for shell formation. If it is assumed that a single unit in the repeated SQ may be consistent with a specific function, the molecules containing the repeated SQ may be more activated with an increase in the number of repeats.

4. Factors reflecting to the primary structure of the OM protein

4-1) Relationship between the primary structure of the OM protein and the shell microstructure

It has been pointed out that many factors such as shell microstructures, phylogenetic relationships, living environments, and individual variations of molluscs might reflect the composition of the OM components.

Molluscan shell microstructures are categorized into six to eight large groups, which are divided into several subgroups. The classification and terminology of the shell microstructures vary among reports (MacClintock, 1967; Taylor, et al., 1969; Carter 1980; Kobayashi 1980; Uozumi & Suzuki 1981; Suzuki et al., 1996).

Samata (1990) demonstrated that the amino acid composition of OM differed markedly in relation to the shell microstructures. The difference was most remarkable among three types of shell layers in particular: 1) the nacreous and prismatic, 2) the foliated and 3) the crossed lamellar and complex
layers. That result was supported by our recent analysis of the OM components isolated from various shell layers using SDS-PAGE (Samata, unpublished data), in which the total amount and degree of heterogeneity as well as the composition depended on the shell microstructure. For example, the nacreous, prismatic and foliated layers contain a large quantity of OM, which is composed of several different protein molecules in the case of the nacreous and prismatic layers. In the case of the foliated layer, the identical amino acid SQ in the N-terminal region can be detected among classes of various sizes of molecules. On the other hand, the other shell layers such as the crossed-lamellar layer contain considerably lower quantities of OM and a very small number of the OM components. According to Shimamoto (1993), who analyzed the amino acid compositions of OMs extracted from the homogeneous, composite prismatic, and crossed-lamellar layers of 17 Venerid bivalve species, the compositions were similar between those in the homogeneous and crossed-lamellar layers, and those in the composite prismatic layer were dissimilar.

No component appears to be universal in its distribution in molluscan shells. This result enables us to assume that the primary structure of OM component, with the singular exception of the CA-like proteins, which are found in both the nacreous and prismatic layers, primarily depends on the shell microstructure.

4-2) Relationship between the primary structure of the OM protein and the phylogeny of molluscs

Very few studies have investigated the amino acid SQ of the homogeneous OM protein, which might provide clear insight into the relationship between
phylogenetic characteristics and OM structure. Several researchers have reported that any difference found in amino acid composition of OM between shell layer types has been smaller than the variations found in the composition between taxonomic groups (Grégoire, 1972; Crenshaw, 1982).

Samata (unpublished data) has recently carried out a comparative analysis of the primary structure of three kinds of homologous proteins among five species from the genus *Pinctada*. The samples examined were as follows: 1) *P. fucata martenshii* (the so-called "Japanese pearl oyster") cultured in Japan, 2) the same species cultured in China (the so-called "Chinese pearl oyster"), 3) *P. fucata fucata* (the so-called "native Japanese pearl oyster"), 4) *P. margaritifera* (the so-called "black lip pearl oyster") and 5) *P. maxima* (the so-called "white pearl oyster"). Three genes, encoding nacrein, N16 and MSI60 were identified from these samples, respectively and the deduced amino acid SQs were compared. The overall highest homology among the species analyzed was recognized in the amino acid SQ of MSI60, and this finding implies that MSI60 may provide strictly ordered substrate for the components of crystal nucleation or be involved directly in nucleation by combination with another molecules.

With regard to nacrein and N66, the amino acid SQs of the NG-repeat domain were much less conservative than those of the CA domain. Asn and Gly residues were arranged regularly in NG-repeat domain in a core unit of NGNNG and YNG, which repeated 3 or 4 times in a species-dependent manner. A clear difference was recognized in the NG-repeat domain of *P. maxima*, characterized by a core unit of NNGNNGNG, which repeated more than 20 times and was much longer than the domain in *P. fucata* (Fig.4). Mutation was recognized in the CA domain, showing many replacements in amino acid residues in *P. maxima* as compared with those in the 3 specimens of *P. fucata*; in the latter group, only a few replacements in the amino acid residues were recognized. A similar type of mutation was recognized between the NG-repeat and the remaining regions of N16 and N14. Consequently, discussion must be focused on the mutation type seen between the NG-repeat and the other regions. In the former, the difference in frequency of the repetition, and in the latter, replacement of each amino acid residue, were characteristic, implying that mutation might have differently occurred between the SQs of NG-repeat region and CA or non-NG-repeat region.

Consensus SQs were located in a specific region of the CA-like proteins, which was the active site of CA, and of N16 and N14, which was rich in acidic amino acids, sulfation, and phosphorylation sites regardless of the species examined. The results correspond to the idea that the SQ of the functionally important region in homogeneous protein molecules can be conserved regardless of the phylogenetic relationship (Dicerson, 1971).

There is no definitive idea whether the length of the NG-repeat region in the CA-like proteins and N16/N14 is related to their capacity for shell formation or to changes in accompanying shell growth. A clear understanding may be obtained after analysis of the genes that encode the NG-repeat region in these proteins, using samples of the various stages of molluscan ontogeny.
4-3) Relationship between the primary structure of the OM protein and living environment of mollusks

Degens et al. (1967) demonstrated that living environment, particularly temperature affected amino acid compositions of the OM in some kinds of molluscs. To understand the relationship of them more precisely, attention must be paid to the primary structures of the specific OM components isolated from taxonomically closely related species whose living environments differ. Until now, no definitive data has been published in this regard.

Mori (in press) clarified the primary structure of the SM and ISM components in the nacreous layer of *Hyriopsis shlegeli*, a species of freshwater bivalve, and compared them with those of the marine species. One SM and ISM components had high SQ similarity with N16 an MSI60, while SQ of another SM component, which had similar molecular weight with nacrein was clearly different from latter. Whether that result reflects a corresponding difference in living environment or a phylogenetic relationship is unclear.

5. Possible function of the OM components

5-1) *In vitro* experiment of CaCO$_3$ crystallization

Many studies have attempted to elucidate that molluscan shell formation occurs under mechanisms of strict biological control. Examinations of crystallization have used several types of assays, which differ from those used to examine the composition of crystallizing solution, and the method used to monitor crystal formation.

With regard to the biomineralization of molluscs, Watabe & Wilbur (1960) and Wilber & Watabe (1963) showed that OM might control the CaCO$_3$-crystal polymorphism based on the results of *in vitro* and *in vivo* experiments. Several decades later, Belcher et al. (1996) demonstrated that the polyanionic proteins extracted from aragonitic nacreous layer of abalone shell were sufficient to induce aragonite on base crystal of calcite, indicating that the SM could exert greater activity than the ISM over control of crystal polymorphism. In contrast, Falini et al. (1996) published findings that macromolecules extracted from the aragonitic shell layers induced aragonite formation *in vitro* when they were first pre-adsorbed on a substrate organic sheet, suggesting that these macromolecules were responsible for the precipitation of aragonite.

It can be understood that aragonite commonly appears in marine molluscan shells because of the high Mg$^{2+}$ content in their extrapallial fluid (Kitano, 1962, 1990). On the contrary, common appearance of aragonitic shells in freshwater species, which contain much lower content of Mg$^{2+}$ in their extrapallial fluid than those of the marine species, may require mechanism of regulation of CaCO$_3$ polymorphism, such as the presence of the specific component capable for aragonite nucleation. Consequently, attention should be paid to experiments using the OM components in calcitic shell layers of the marine molluscs to search for the components involved in calcite nucleation as opposed to those in aragonitic shell layers of the freshwater molluscs to search for those involved in aragonite nucleation.

In our latest study (Samata et al., 1999), we used an assay system containing two kinds of crystallizing solution with different compositions. One is a supersaturated CaCO$_3$ solution to induce calcite (calcitic crystallizing solution) and the other is the same solution but with enough Mg$^{2+}$ ions to induce aragonite (aragonitic crystallizing solution). Additionally, we added the purified OM components extracted from the nacreous layer of *P. fucata* (nacrein, N16 and ISM, respectively) as additives.

In every preparation, the characteristic efficiency of aragonite nucleation could not be observed for these three OM components in either the isolated state or with any combination of them. Two SM components were proved to have activity when being present in free state may inhibit crystal formation, but the same molecule, when attached to ISM, may regulate the form and size of aragonite crystals (Fig.5). The inhibitory reaction may be caused by binding of the SM components to all sides of the crystal nucleus, and the regulatory reaction by adsorption of them to the specific site of the crystal nucleus. It is assumed that the molecule-crystal interface might be mediated by the Gly-loop structure, which relates to the NG-repeat SQ and largely extended from the surface of nacrein.
and N16 molecules. We undertook further experiments to clarify this process, and are now attempting a binding assay of these two molecules to crystal surfaces, using fluorescence dye. An additional effect of the promotion of crystal growth was associated with the activity of nacrein based on the fact that increased efficiency of crystal growth was observed in the presence of nacrein as an additive in both the aragonitic and calcitic crystallizing solutions (Samata et al., in press). This effect may be due to the enzymatic activity of its CA domain.

One of a commonly used strategy of the in vitro crystallization experiment is adaptation of the supersaturated CaCO₃ solution with the OM components. However, crystal nucleus formation can occur without OM in such a solution, and therefore it is difficult to analyze the successive process of ion concentration and crystal nucleation under control of the OM components. In response to the abovementioned circumstances, an experiment can be designed to use a crystallizing solution, containing ions of Ca²⁺, HCO₃⁻, and SO₄²⁻ with/without Mg²⁺ under the similar concentration with extrapallial fluid of marine species (Wada & Fujinuki, 1976, Wada, 1988). As it is in the condition of under-saturation about CaCO₃, no crystal formation occurs without the OM components in our preliminary experiment. This approach will yield information useful in establishing the function of the OM components involved in the crystal nucleation.

One remaining problem related to the experiment using extracted OM is that we cannot guarantee the complete absence of undissolved CaCO₃ crystals in the sample even after most careful treatments to remove the mineral phase. Only tiny contaminated crystals may cause trouble in the result of this kind of experiment, because a new nucleus can be formed from any pre-existing crystals. Consequently, most of the results of in vitro crystallization experiments, using the OM components extracted from shell as additive or fixative, containing our own, should be prudently analyzed in particular when discussion is concerned to crystal nucleation. Repeated experiments should be applied to avoid decisive mistakes. Furthermore, to solve this problem, synthesized materials or recombinant proteins can be taken into consideration as the additives or fixatives to the system of crystallization. One of the former materials is a simplified surface molecular-assemblage that mimics biological membrane such as Langmuir monolayers (Landau et al., 1985, Mann, et al., 1991) or self-assembled monolayers (SAMs) (Bunker et al., 1994). By using SAMs of ω-terminated alkanethiols, micropatterned on metal films, Aizenberg (in press) demonstrated the face-selective nucleation of calcite crystals. The other type of material is synthesized peptides with homologous amino acid SQs with those of the OM proteins (Wheeler et al., 1987). In spite of the problem that the activities of these components are not always equal to those of the OM components, which discrepancy is due to a difference in their tertiary structures with those of the OM components, this new approach may give us a new insight into the process of crystal nucleation.

The time of nucleation is analogous to the induction time, which can be measured as the change in the composition of the crystallizing solution, which, in this system, can be detected after the first appearance of crystals. Therefore, the experimental system should be improved by using not only an optical measurement method, but also a chemical measurement method, such as the one used by Wheeler et al. (1981).

5-2) Control of the CaCO₃ polymorphism

CaCO₃ crystal of molluscan shell occurs as aragonite or calcite microstructura depending manner, and in some cases both crystal types exist for one species. Several researchers have suggested close involvement of the OM components in control of crystal type (Belcher et al., 1996, Falini et al., 1996, Thompson et al., 2000).

Presence of the OM components responsible specifically for nucleation of aragonite or calcite might be explained by the notion that the regular spacing of carboxyl side chains of Asp is a close match to that of with Ca²⁺ in CaCO₃ crystal lattices, and therefore controls the crystal polymorphism. A recent experiment (Sekiguchi, in press) using two kinds of the SM components as additives in the presence of fixative ISM isolated from the freshwater bivalve H. shlegeli in the calcitic crystallizing solution resulted in formation of platy crystals morphologically similar to those induced by combination of SM and ISM of marine
species in aragonitic crystallizing solution. Because of the small size and very thin morphology of these platy crystals, the accurate identification of the crystal form was not complete, even though a presumptive analysis using microarea XDR clarified the presence of aragonite together with calcite and vaterite.

An observation presented in another report indicated that nucleus formation is favored when a matrix molecule binds to the surface of a nascent crystal nucleus, decreasing the surface energy and lowering the energy available for nucleus formation (Wheeler, 1992). Wada et al. (1993) reported that the aragonite nucleus, once formed, may undergo a solid-to-solid transformation into a more stable calcite nucleus by the adsorption of the OM components onto the crystal in the case of the calcitic shell formation of the marine species. Based on this idea, a factor for calcite nucleation is not required, and the calcite nucleus results in the formation of high-magnesium calcite. At this point, an exact explanation as to why selectivity between these two crystals is a common feature of the biomineralized system is still required.

6. Possible mechanism of molluscan shell formation

6-1) Possible mechanism of crystal nucleation

Crystal formation in a biological system is considered to involve 1) crystal nucleation and 2) crystal growth (Simkiss & Wilbur, 1989). Thus far, explanation of the initial stage of shell formation includes 1) homogeneous (spontaneous) nucleation and 2) heterogeneous nucleation. The first type of nucleation occurs in the condition of a pure supersaturated solution, in which the crystal nuclei are formed spontaneously. It may not occur in biological systems, because the solution in which crystals may be induced includes many kinds of organic materials (Simkiss, 1976). Consequently, heterogeneous nucleation is commonly achieved in biological systems by the varied involvement of the OM components. The process of heterogeneous nucleation has been explained in template hypothesis (epitaxial hypothesis) (Neuman & Neuman, 1958; Seifert, 1972; Weiner & Hood, 1975; Crenshaw & Ristedt, 1976; Degens, 1976; Wada, 1980) and

![Figure 6](image_url)

**Figure 6.**
SDS-PAGE electrophoretogram of the whole proteins expressed by E. coli. Arrow shows the position of the recombinant N16 protein. (min.); time after induction for gene expression.
compartment hypothesis (Bevelander & Nakahara, 1969, 1980). The former emphasizes the activity of the OM components as a template for crystal nucleation, and the latter regards the presence of an organic compartment that covers crystals to control their nucleation, as important morphology, size, and polymorphism.

Based on measurements of the ion products of calcium and carbonate compared among environmental water, blood, and extrapallial fluid, Wada (1988) showed that the value of the activity product \( \text{Ca} \times \text{CO}_3 \) in marine species was fairly low. The result implies that the site for crystal nucleation in marine bivalves may be under a condition of saturation or low super saturation with respect to \( \text{CaCO}_3 \). In such a case, the specific OM components are thought to be involved in ion-concentration prior to crystal nucleation. \( \text{Ca} \)-binding protein (or glycoprotein) and phosphorylated or sulfated protein (or glycoprotein) may act to concentrate \( \text{Ca}^{2+} \), although their action mechanisms may be different, while the \( \text{CA} \)-like proteins may be a molecule participated in concentrating \( \text{CO}_3^{2-} \) due to the activity of its \( \text{CA} \) domain. In the case of freshwater molluscs, extrapallial fluid may be under a condition of super-saturation with respect to \( \text{CaCO}_3 \) (Wada, 1988), implying the presence of a different mechanism of crystal nucleation from marine species.

Degens (1976) presented molluscan shell biomineralization based on template hypothesis. It explained that the structural framework (ISM) was covered with polyanionic soluble protein capable of nucleating crystals. This idea was further extended by Weiner (1975, 1979) that sequential arrangement of the carboxyl group of Asp may be the main constituent of the template.

Sulfate has been thought to be the other important component for crystal nucleation. The possibility of \( \text{Ca} \)-binding by sulfate esters in the shell of \( M. \ mercenaria \) was shown by Crenshaw (1972), and this finding was further supported by the fact that sulfate esters concentrated at the nucleation site by histochemical analysis (Crenshaw & Ristedt, 1976). In his work using a glass coverslip preparation, Wada (1980) confirmed the sulfur concentration at the nucleation site and proposed that the sulfur-binding glycoprotein is critical for crystal nucleation. Recent analysis of sugar moieties contained in the SM of the calcitic prismatic layer of two pteriomorphid bivalves, Daufin (2003) showed that sulfur is mainly associated with acidic sulfated sugars and not with amino acids, although the role of the sulfated sugars is still not well understood.

A hypothesis proposed by Addadi & Weiner (1989) took into account the possible effect of two cooperating factors, \( \beta \)-sheet carboxylate-rich protein and sulfated polysaccharide. According to them, the sulfate created a flux of \( \text{Ca}^{2+} \) towards the nucleation site, and carboxylates, regularly arranged in a protein \( \beta \)-sheet domain, provided the structural organization necessary for nucleation.

Based on the Cryo-TEM studies of the OM in the nacreous layer of \( Atrina \), Levi-Kalisman (2003) represented a scheme where the silk fibroin-like gel phase is located between the interlamellar sheets composed mainly of \( \beta \) chitin, which is arranged in highly ordered parallel arrays and may exist as a substrate for crystal nucleation. Meanwhile, acidic Asp-rich protein bound to the chitin sheet surface may be responsible for crystal nucleation. In vitro assay system using chitin fibers or chitosan films as substrate showed favorable nucleation of calcite and vaterite dependence on the additive acidic macromolecules (Zhang & Gonsalves, 1995, Kato & Amamiya, 1999).

Until now, information about these hypotheses was mainly indirect, whereas more direct information may be provided on crystal nucleation by the combination of analyses regarding the genes encoding the OM components, of observation of the microstructure of the shell layer and OM, and of in vitro and in vivo experiments of crystallization. In addition, Mann (1988) demonstrated that structural correspondence between the surface of the OM components and crystal nuclei implies a stereochemical matching. Therefore, structures of the OM-crystal interface should be analyzed in two dimensions. In this respect, analysis of the tertiary structures of the OM proteins may give elucidation the mechanism of crystal nucleation at the molecular level.
6-2) Possible mechanisms of crystal growth

Once formed, the crystal nucleus continues to add ions to crystal surfaces and the crystal face develops dislocation and passes through steps important for subsequent growth (Nielsen & Christoffersen, 1982). The process of crystal growth includes regulation of crystal morphology, which is controlled by the inhibitory activity of the specific OM components. The process of inhibition of crystal growth was experimentally clarified by using adsorption and crystal growth assays by using OM phosphoproteins and synthetic peptides in a report by Wheeler (1992). The result implies that the matrix phosphoproteins and polyaspartic acid simultaneously have high affinity for several sites on crystals, thus offering inhibition of crystal growth. In OM phosphoprotein, a lower acidic component was speculated to interact with a specific face of the crystals, regulating the crystal morphology while a highly acidic component was nonspecifically terminating crystal growth. However, it should be noted that the interactions of the OM components with crystals cannot be entirely modeled after those of small peptides due to differences in their tertiary structure with those of the OM components. Based on these experiments, Wheeler (1992) represented a hypothetical scheme related to the oyster shell formation. It contains stages of crystal initiation, and control and limitation of crystal growth. Crystal initiation is performed by ISM from the previous crystal layer, which has higher affinity for crystal nuclei than Ca$^{2+}$. Later two stages are regulated by SMs, which may adsorb onto the specific site (control) or whole sites (limitation) of growing crystals.

Another report on the effect of the OM components to crystal growth was obtained from the assay by optical microscopy, using atomic force microscope with or without the SM components, isolated from the calcitic shell of Haliotis rufescens (Walters et al., 1997). Atomic step edges on the crystal surface were altered in shape, which was consistent with the habit modification when OM was added.

Figure 7.
Various types of crystals developed over the surface of the cultured epithelial cells with the result of EDS analysis.
(This figure was reprinted from Sumata et al., (1994), under the permission of the authors.)
7. Several new studies on biomineralization

The previous section summarized the structure of the OM components involved in the molluscan shell formation and the results of in vitro experiments related to the function of these molecules. The next step of research must be focused on in vivo experiments on biomineralization, and several new technologies involving gene expression and tissue culture will be introduced in this section.

7-1) Expression of the genes involved in the formation of the nacreous layer

7-1-1) Expression of N16 gene by E. coli

Oyama (2002) have succeeded in an experiment on the expression of the gene encoding the N14 protein in the nacreous layer of _P. maxima_ after transforming the N14 gene amplified by PCR to M15 E. coli by using the pQE-70 expression vector. The positive clones were induced to express recombinant N14 protein, which had the same molecular weight and amino acid composition as N14 isolated from shell (Fig.6). This result implies that the N14 gene is functional, excluding the possibility of a pseudogene of it. The crystallization experiment using recombinant N14 protein as additive resulted in an equal effect on crystal formation with lesser amount of it than that using N14 extracted from shell.

7-1-2) Seasonal and daily change in expression of genes involved in the nacreous layer formation

In the field of Japanese pearl culturing, it has been noticed for a long time that the aragonite crystals of the nacreous layer show seasonal changes in form, size, and density (Wada, 1961, 1972, 1999), leaving research on molecular level for later. As the first step, we designed an experiment to trace the daily sequential expression of the correspondent genes as follows: the epithelial tissue was peeled from the cultured pearl oyster every 3 hours for 3 days in the summer and winter seasons, respectively. After isolating total RNA, RT-PCR and real-time PCR was carried out to determine the expression of the genes. The primers for PCR analysis were designed on the basis of the SQ information of nacrein, N16 and MSI60 genes that the genes encoding them showed completely the same cycle of expression, which was closely related with daily cycle of tidal wave (Samata, unpublished data). The same experimental setup to investigate the expression of them in samples collected in winter is now underway.

In another experiment, Moura _et al._ (2000, in press) reported the monthly variation of the total weight and composition of organic and inorganic components contained in the body fluid and extrapallial fluid of freshwater bivalve, _Anodonta cygnea_, which relates to the cycle of the nacreous layer formation.

7-2) Research on primary culture of the mantle epithelial tissue and the cells responsible for shell formation

Tissue culture, a new technique practiced in shell and pearl formation studies, is of great use. Long term culture of mantle epithelial tissue has been conducted by Machii (1985) by using the Japanese pearl oyster _Pinctada fucata_. This line of research was advanced by improvements of the culture medium, including appropriately balanced solution, fundamental medium adjusted to the amino acid composition of the body fluid, and physiological saline suited to the species, and a method for the decontamination of animals. Subsequently, Machii & Wada (1989) reported evidence of the secretion of an organic substance and of crystals using an explants culture of mantle epithelial tissue from _P. fucata_. Microcrystals with similar morphology were found by Samata _et al._ (1994) on the surface of a black secretion on cultured mantle epithelial tissue of the same species (Fig.7). Microcrystals of various sizes and shapes contained large amounts of Ca and S.

In addition, Sud _et al._, (2001) analyzed the effect of the water-soluble OM in the nacreous layer of _P. maxima_ for the regulation of cell activity using mantle epithelial cells of _Haliotis tuberculata_, and showed that OM reduced the global viability of the cells in a dose-dependent manner.

The next step towards the successful cultivation of mantle outer epithelial cells will be to develop a method for collecting sufficient amounts of cells of high purity and viability. Awaji (1991, 1997) reported
a method to isolate pure outer epithelial cells of *Pinctada fucata* by enzymatic digestion treatment. The isolated cells were mostly epidermal ones, and DNA synthesis was found to occur in migrated cells during cell culture (Awaji & Suzuki, 1998).

7-3) *In vivo* experiments of shell formation

Several attempts have been made to determine the initial stage of molluscan shell formation by the experiments of transplant, shell regeneration and glass coverslip preparation. Wada (1980) represented the detailed explanation of these experiments, and these methods should be further extended in observing and analysing the direct process of shell formation.

8. Issues remaining to be clarified

1) Different shell layers with distinct crystal types are directly related to different cell types and, therefore, a specific mantle epithelial region controls crystal formation. Experimental studies that might provide insight into the relationship between cell type and crystal formation are still necessary.

2) Genes involved in shell formation express proteins (OM precursor proteins) after post-translational modification are secreted from the mantle epithelial cells. After release, these proteins may undergo physical and chemical changes, and subsequently mature into the various OM components. These aspects of proteins need closer examination.

3) The analysis of the OM proteins (genes) responsible for shell formation is still in its early stage, when the nomenclature of newly found protein is, as expected, problematic. Indeed, some confusion has arisen over similar names being given to completely different proteins or different names being given to closely related proteins. The nomenclature should be redefined after enough data from the primary structures of the molluscan OM proteins or grouping them can be obtained.

ACKNOWLEDGEMENTS:

I am grateful to Dr. Koji Wada, Dr. Katsuhiko Wada, Dr. Akira Machii and Dr. Makiko Kono for their valuable discussions and suggestions.

REFERENCES


(Received: June, 10, 2003. Accepted: October, 05, 2003)