Biochemical and Biophysical Research Communications **269,** 213–218 (2000)
doi:10.1006/bbrc.2000.2274, available online at http://www.idealibrary.com on IDEAL®

Molecular Mechanism of the Nacreous Layer Formation in *Pinctada maxima*

Makiko Kono,* Nakanobu Hayashi,† and Tetsuro Samata*,[1]

**Laboratory for Cell Biology, Department of Environmental Health, Azabu University, 1-17-71 Fuchinobe, Sagamihara-shi, Kanagawa, 229-8501, Japan; and †Omgen Inc., Oyaguchi Kitamachi 90-4, Itabashi-ku, Tokyo, 173-0002, Japan*

Received November 15, 1999

We have cloned the cDNAs that encode two kinds of molluscan shell matrix proteins, namely N66 and N14, in the nacreous layer of *Pinctada maxima*. N66 is composed of carbonic anhydrase-like and repeat domains, as described for nacrein (1) in the pearls of *P. fucata*. N14 is homologous to N16, recently found in the nacreous layer of *P. fucata* (2) and is characterized by high proportions of Gly, Tyr, and Asn together with NG repeat sequences. The molecular weights of these proteins were estimated as 59,814 and 13,734 Da, respectively. Structural differences were clearly indicated in the alignment and length of the repeat sequences of the sets of the homogeneous proteins (N66/nacrein and N14/N16). The longer repeat sequences of N66 and N14 may be responsible for *P. maxima*'s excellent property of calcification. The *in vitro* crystallization experiments revealed that the mixture of N66 and N14 could induce platy aragonite layers highly similar to the nacreous layer, once adsorbed onto the membrane of the water-insoluble matrix.

The organic matrix components figure in most schemes describing the regulation of biomineral formation. The organic matrix of molluscan shell is one of the best studied of all calcium carbonate biominerals and considerable information has been accumulated regarding the structure of these matrices (3–7), each of which is generally classified as either a water-soluble matrix (WSM) or a water-insoluble matrix (WISM). Recently, several genes encoding the matrix components have been isolated and their deduced amino acid sequences have been clarified (1, 2, 8–10). Since Crenshaw isolated Ca-binding glycoprotein from the shell of *Mercenaria mercenaria* (11), the Ca-binding property has been assigned to the WSM of many kinds of molluscan species (12–14). The other investigations regarding the function of organic matrices come from *in vitro* studies in which matrix activities are measured (15–20). These activities include matrix abilities related to crystal nucleation, crystal growth and termination of crystal growth. However, the precise function of the organic matrix remains unclear even *in vitro*. It seems relevant to investigate which structural properties of each matrix component are correlated to the regulation of each step of shell formation.

Pinctada maxima, a close relative of the Japanese pearl oyster *P. fucata*, has a large thick nacreous layer and is cultured in south Asian countries as mother-of-pearl. Such an excellent ability for calcification must be correlated to the structural properties of the organic matrix components in the nacreous layer of *P. maxima*.

Here we report the complete amino acid sequences of the two matrix components newly isolated from the nacreous layer of *P. maxima*, and compare these to the homogeneous components already reported from *P. fucata* (1, 2). This is the first comparative biochemical study of the homogeneous matrix protein from different molluscan species. The results indicate how phylogenetic and functional differences are reflected in the protein structure.

MATERIALS AND METHODS

cDNA cloning. Specimens of *P. maxima* cultured at the Institute of Tasaki Marine Biology (Tokushima, Japan) were used for RNA extraction. Total RNA was isolated from the dorsal region of the mantle tissue using the acid-guanidium-phenol-chroloform (AGPC) method (21). Poly(A)$^+$ RNA was fractionated by an Oligotex dT30 mRNA Purification Kit (Takara). Approximately 5 μg of mRNA was used to construct the cDNA using a Time Saver cDNA Synthesis Kit (Amersham Pharmacia Biotech.). The cDNA was then inserted into the Lambda gt10 vector and packaged with a Lambda DNA Packaging System (Amersham Pharmacia Biotech.).

For the screening of the cDNA library, two kinds of degenerate oligonucleotide probes were designed based on the amino acid sequences of N-terminal regions of N66 and N14. They were [5′-GCNWSNATGCAYMGNCAYGAYCAY-3′] and [5′-TAYATGGAYGGNMGNCARTGGMGNTAYCC-3′], corresponding to the amino acid

The complete cDNA and deduced amino acid sequences of N66 and N14 reported in this paper have been deposited in the EMBL/GenBank Accession Nos. AB032612 and AB032613, respectively.

[1] To whom correspondence should be addressed. Fax: 81-42-769-1938. E-mail: samata@azabu-u.ac.jp.

0006-291X/00 $35.00

sequences (ASMHRHDH) and (YMDGRQWRYP) of N66 and [5′-TAYCARMGNTGYWSNM-GNTA-3′] and [5′-TGGTAYTGYTGGAT-HCCNTA-3′], corresponding to (YQRCSRY) and (WYCWIPY) of N14. Positive clones obtained by screening were separated and subcloned into the plasmid vector with pGEM-T (Promega) and sequenced by the cycle sequence method using an automated DNA sequencer Prism 377 (PE Biosystems). The cDNA fragments that contained the 5′-end sequence were isolated by the 5′-RACE method using a Marathon cDNA Amplification Kit (Clontech), and the PCR products were subcloned and sequenced by the method described above.

Purification and aminoterminal analysis of the organic matrix components. We mechanically separated the inner nacreous layer of *P. maxima* from the outer prismatic layer and crushed it to powder. The powdered shell was decalcified with 10% (w/v) EDTA-4Na (pH 7.6) containing 0.01% (w/v) Na-azide, followed by dialysis against distilled water. The dialyzed solution was concentrated by evaporation and again dialyzed with a Centriprep-3 dialyzer (Grace Japan). Following this, WSM and WISM were separated by centrifugation at 15,000 × g for 20 min. The components in the WSM were analyzed with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using 10–20% (w/v) ready-made gel, Pagel NPG 10–20 (Atto). To determine the aminoterminal sequences, the bands of the WSM were blotted onto a polyvinylidene difluoride filter (PVDF Filter, Millipore) using a dry blotting system (Nippon Eido), cut from the filter and subjected to an automated amino acid sequencer LF3000 (Beckman).

RT-PCR analysis. Reverse transcription PCR (RT-PCR) was carried out using the degenerate oligonucleotide primers described below, and mRNAs in the dorsal region of the mantle and the mantle edge served as the templates. The primers were as follows: [5′-ATGTGGAGAATGACGACGCTTC-3′] and [5′-CTCTGCGTTG-ATACCACTGCT-3′], each of which matched the sense sequence of nucleotide (nt) 1–21 and antisense sequence of nt 1790–1810 of the clone coding for N66 and [5′-TCAGTGGTAACAACACAGAGT-3′] and [5′-TAATCAATAAAATATTGAAGG-3′], matching of the sense sequence of nt 22–42 and the antisense sequence of nt 763–783 of the clone coding for N14.

Crystallization experiment. The crystallization experiment was carried out according to the method of Samata *et al.* (2), using an aragonitic crystallization solution.

RESULTS

SDS-PAGE revealed the presence of two major components with molecular weights of approx. 66 kDa (N66) and approx. 14 kDa (N14) in the WSM of the nacreous layer of *P. maxima.* A minor band was also visible at the position of approx. 28 kDa. Amino acid sequence analysis showed that these components had different N-terminal sequences of ASMHRHDHYM-DGRQWRYP for N66 and AYQRCSRYWYCWIPY for N14 and 28 kDa component.

Only one positive clone was isolated from the cDNA library by each screening procedure using an oligonucleotide probe based on the N-terminal sequence of N66 and N14, respectively. The DNA sequence analysis of these two clones showed that they were 1811 base pairs (bp) (C1) and 814 bp (C2) long. Their complete nucleotide sequences together with the deduced amino acid sequences are shown in Figs. 1 and 2. C1 revealed an open reading flame encoding 568 amino acids and C2 encoding 140 amino acids, both with the translation initiation codon ATG and the nucleotides corresponding to the signal peptides (nt 1–66 in C1 and nt 82–156 in C2). The stop codon was TAA for C1 and TAG for C2,

```
ATGTGGAGAATGACGACGCTTCTTCACTTGACTGCTCTGCTTGTTCTGATTCCATTATGT   60
[ M  W  R  M  T  T  L  L  H  L  T  A  L  L  V  L  I  P  L  C    20
CATTGCGCCCTCCATGCACAGGCATGACCATTATATGGACATGGATCAAACCTACCCTAAT  120
  H  C] A  S  M  H  R  H  D  H(11) Y  M  D  M  D  Q  T  Y  P(20) N   40
GGATTGGGATACTGTGAACCTTCAGGTGAAAGCAGCTGTAAAGCCGGATTTAGCTACAAT  180
  G  L  G  Y  C  E  P  S  G  E  S  S  C  K  A  G  F  S  Y  N    60
AGAGACATATGCCAAGGTCCGTATCATTGGCACACTATATCTAGTTGCTATAAGGCATGT  240
  R  D  I  C  Q  G  P  Y  H  W  H  T  I  S  S  C  Y  K  A  C    80
GGACATAAAAGGAGACAATCACCAATCAACATTTGGTCACATAAAGCTGTATTCTTACCT  300
  G  H  K  R  R  Q  S  P  I  N  I  W  S  H  K  A  V  F  L  P   100
TATCTGCCAAGACTGAAATTCAAGCCACATATGAAGTCATTGGATACGGACGTGACAAAT  360
  Y  L  P  R  L  K  F  K  P  H  M  K  S  L  D  T  D  V  T  N   120
CACCAAAATCGTGCCCCTGAATTCGAGCCGGAGGACGGAGATAAGCTTCATGTGAAACTA  420
  H  Q  N  R  A  P  E  F  E  P  E  D  G  D  K  L  H  V  K  L   140
AAGAATCTTGTTGATGGACATTATAAATTTCACAATCTCCATATTCACAACGGCAAAAGT  480
  K  N  L  V  D  G  H  Y  K  F  H  N  L  H  I  H  N  G  K  S   160
AGACGAAAGGGCTCGGAACACAGCGTGAACAGACATTTTACGCCCATGGAGGCTCATTTG  540
  R  R  K  G  S  E  H  S  V  N  R  H  F  T  P  M  E  A  H  L   180
GTGTTCCATCATGATGATAAAAAGGAAATCAAACCTCCAAGGGTTAAGTTAGGGGGAGTG  600
  V  F  H  H  D  D  K  K  E  I  K  P  P  R  V  K  L  G  G  V   200
TACGCTGGTCGTAACAAATTTGTTGTCGTTGGAGTCTTTCTAGAGGTGGGTGATGAAGGA  660
  Y  A  G  R  N  K  F  V  V  V  G  V  F  L  E  V  G  D  E  G   220
TACGGTGATGAACCGGACGACGATGAATGTAAGCGCATATTAAAGGGTCATTGCGAGAAC  720
  Y  G  D  E  P  D  D  D  E  C  K  R  I  L  K  G  H  C  E  N   240
AATGGGGACAATGGTAACAACTGTGATAACGGCAACAATGGTAACAACGACAACAATGGT  780
  N  G  D  N  G  N  N  C  D  N  G  N  N  G  N  N  D  N  N  G   260
AACAACGGAAACAATGGTAATGGTAACAACGGTTATAACGGTAATAACGGTGACAATGGA  840
  N  N  G  N  N  G  N  G  N  N  G  Y  N  G  N  N  G  D  N  G   280
AACAATGGCAATGGTAATGGTAACAACGGTTATAACGGTAATAACGGTTACAATGGCAAC  900
  N  N  G  N  G  N  G  N  N  G  Y  N  G  N  N  G  Y  N  G  N   300
AACGGAAACAATGGTAATGGTAACAATGACAATAATGGTAACGATAACAACGGAAATAAC  960
  N  G  N  N  G  N  G  N  N  D  N  N  G  N  D  N  N  G  N  N   320
GGTGGCAATGGTAACAACGGAAACAATGGTAATGGTAACAATGGAAATAATGGTAATGGT 1020
  G  G  N  G  N  N  G  N  N  G  N  G  N  N  G  N  N  G  N  G   340
AATAACGGAAATAACGGTGGCAATGGCAACAACGGAAACAATGGTAATAGTAACAACGGA 1080
  N  N  G  N  N  G  G  N  G  N  N  G  N  N  G  N  S  N  N  G   360
AATAATGGTAATGGTAACAACGGAAATAACGGTGGCAATGGCAACAACGGAAACAATGGT 1140
  N  N  G  N  G  N  N  G  N  N  G  G  N  G  N  N  G  N  N  G   380
AATGGTAACAATGAAAATAATGGTAACGGTAGTAATGGTAACAATGGTGGAAACGGCAAC 1200
  N  G  N  N  E  N  N  G  N  G  S  N  G  N  N  G  G  N  G  N   400
AATGGTAATAACGGTGATAACGGTAATGGCGACAATGGTTATAACGGTGATAATGGTAAC 1260
  N  G  N  N  G  D  N  G  N  G  D  N  G  Y  N  G  D  N  G  N   420
AGTGACGGGCGACTCAGACGCTGGGATTTGGCAAATGTCCGACGCATGCACGCCGAGCGA 1320
  S  D  G  R  L  R  R  W  D  L  A  N  V  R  R  M  H  A  E  R   440
TATCACTTTAGCGGAGGATGTATCGTCAAAAAAGCTAAACGCCTCAGCAGGATTCTTGAA 1380
  Y  H  F  S  G  G  C  I  V  K  K  A  K  R  L  S  R  I  L  E   460
TGCGCATATAGACACAAAAAAGTCAGAGAATTCAAAAGGAATGGAGAAGAAAAAGGTCTT 1440
  C  A  Y  R  H  K  K  V  R  E  F  K  R  N  G  E  E  K  G  L   480
GATGTTGATATTACACCGGAAATGGTTTTACCGCCAATGAAATACAGACATTACTATACT  1500
  D  V  D  I  T  P  E  M  V  L  P  P  M  K  Y  R  H  Y  Y  T   500
TATGAAGGATCTTTGACAACCCCTCCTTGCAATGAGACCGTCCTTTGGGTTGTTGAAAAA 1560
  Y  E  G  S  L  T  T  P  P  C  N  E  T  V  L  W  V  V  E  K   520
TGCCACGTGCAAGTATCCAGAAGGGTGCTTGATGCATTGCGGAACGTCGAAGGATATGAA 1620
  C  H  V  Q  V  S  R  R  V  L  D  A  L  R  N  V  E  G  Y  E   540
GATGGTACCACGCTGAGCAAGTATGGAACCAGACGTCCCACACAAAGAAACAAGCATCCT 1680
  D  G  T  T  L  S  K  Y  G  T  R  R  P  T  Q  R  N  K  H  P   560
CTACGTGTGTACAAAAACTCCATATAATGATCATGGCGAGAGAATGACGACGCTTCTTCA  1740
  L  R  V  Y  K  N  S  I  *                                     568
CTTGACTGCTCTGCCTCCTCCCCCCACCCCCCCCCGGCCATATGGCCACTCTGCGTTGAT  1800
ACCACTGCTT                                                     1811
```

FIG. 1. The nucleotide sequence of the 1.8 kb cDNA clone and deduced amino acid sequence of N66. Brackets enclose the signal peptide sequences. The underlines show the amino acid sequences used for the design of the two kinds of the degenerate oligonucleotide probes. The asterisk represents the termination codon.

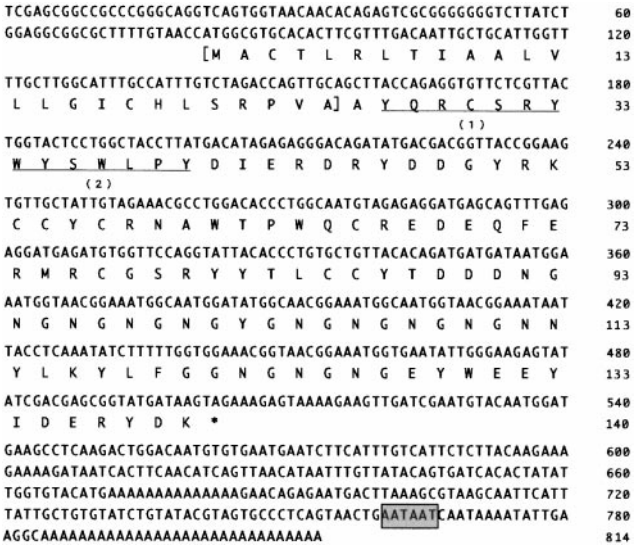

FIG. 2. The nucleotide sequence of the 814 bp cDNA clone and deduced amino acid sequence of N14. The brackets, underlines, and asterisk are used as described in the legend to Fig. 1. A putative polyadenylation signal is boxed in dark shadow.

which was located 179 nucleotides upstream of the putative polyadenylation signal (AATAAT).

Based on the deduced amino acid sequence, N66 was characterized as having high proportions of Asn and Gly containing a specific repeat which was composed mainly of Asn and Gly at the center of the molecule between residues 240–420. N14 was rich in Gly, Tyr and Asn with a specific repeat of NG between residues 92–113 (with 2 exceptions of Tyr instead of Asn and Asn instead of Gly) and 122–127. The calculated molecular weights of N66 and N14, before any post-translational modifications, were 59,814 Da and 13,734 Da, respectively.

A search of the Swissprot protein data bank revealed high homology between N66 and nacrein (39% identity for the repeat region and 72% for the remainder), isolated from pearl of *P. fucata* (1) and between N14 and N16 (71% identity) isolated from the nacreous layer of *P. fucata* (2). Moreover, N66 showed homology with the carbonic anhydrase (CA) family, exhibiting greatest identity with human CAVII (33% identity between residue 56–215, and 38% between 486–555). On the contrary, N14 had homology among short sequences with the cell wall protein, QID3 (22) and rabbit hair keratin, Glycine/Tyrosine-rich protein (23).

The result of the amino acid sequence comparison between N66 and nacrein is shown in Fig. 3 and schematically in Fig. 4A. It is summarized as follows: Regarding the CA-like domain, (1) The most striking identity of 97.2% was observed among the active site of CAII, including a perfect match of 3 zinc-binding sites; and (2) The degree of conservation of the amino acid residues varied according to the position of the peptide chain. Regarding the repeat domain, (1) The length of the domain in N66 was almost twice that of nacrein (see Fig. 4A); (2) Although the repeat domains of both components were characterized with high proportions of Gly and Asn, a slight difference was recognized in the alignment of the two amino acids between them; and (3) The N-terminal amino acids were identical but the C-terminals were different i.e., Ala for N66 and Lys for nacrein.

The result of the amino acid sequence comparison between N14 and N16 is shown in Fig. 5 and schematically in Fig. 4B. It is summarized as follows: (1) A well comparable repeat sequence was found where the doublet Asn-Gly recurred 6 times in N16 and 12 times in

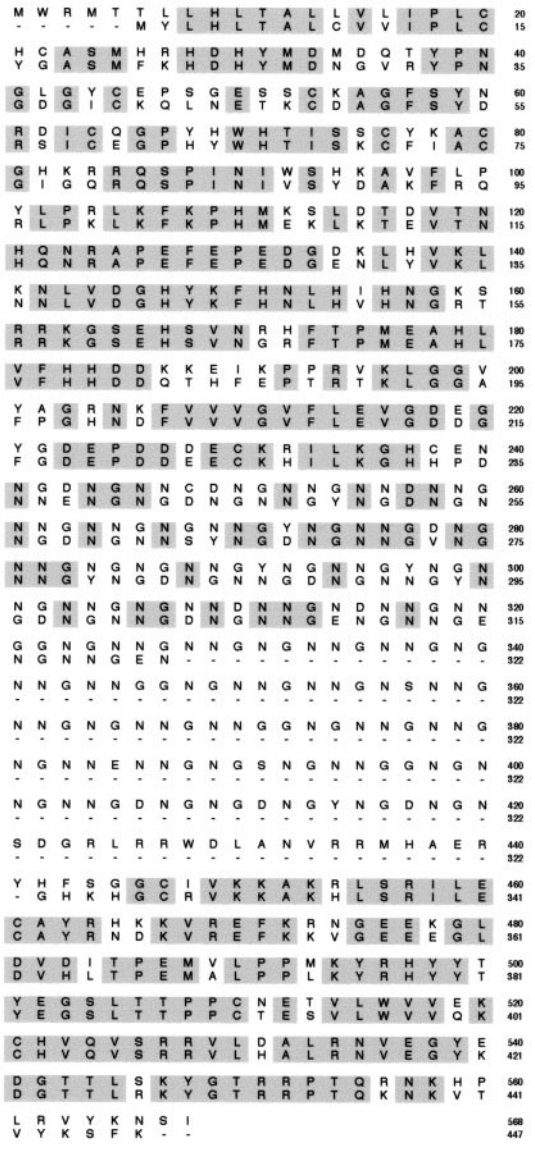

FIG. 3. Comparison of the deduced amino acid sequences of N66 (upper) and nacrein (lower). Dark shadowed backgrounds represent the identical amino acids. Alignment gaps in the sequences are represented by dotted lines.

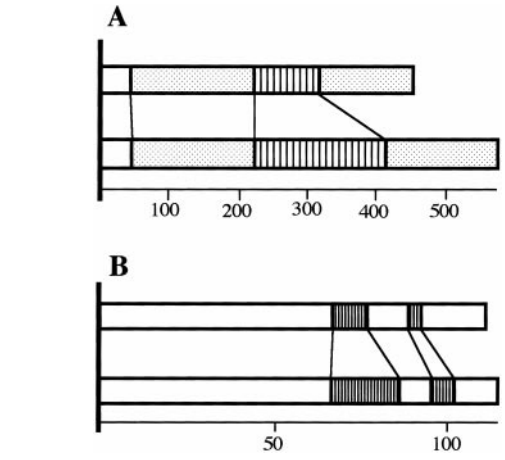

FIG. 4. (A) Schematic representation of the amino acid sequences of nacrein (upper) and N66 (lower). The column with vertical stripes corresponds to the repeat domain, the dotted column corresponds to the CA-like domain, and the blank column represents the remainder of the region. The numbers indicate amino acid residue positions. The alignment of the repeat domain in nacrein is (AB-A), (A = GDNGNN, B = G/SYN), whereas that in N66 is the irregular repeat of GNN and GN, in which the most frequent occurrence is GNNGNNGN. (B) Schematic representation of the amino acid sequences of N16 (upper) and N14 (lower). The column with the vertical stripes corresponds to the region with the NG repeat sequence and the blank column represents the remainder of the region. The numbers indicate amino acid residue positions.

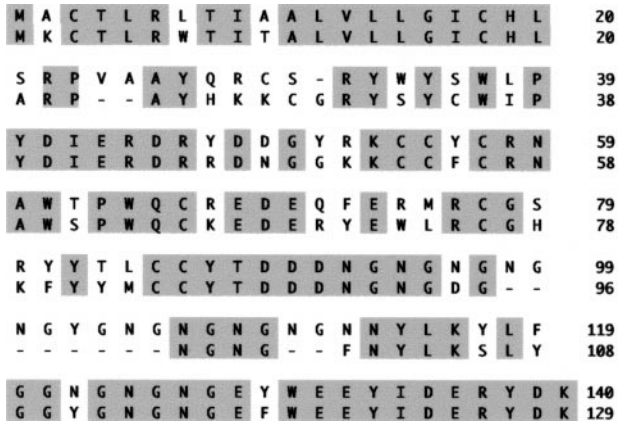

FIG. 5. Comparison of the deduced amino acid sequences of N14 and N16. Identical amino acids are boxed in dark shadow. Gaps are marked by dotted lines.

N14. Because of the longer NG-repeat, the total length of the N14 molecule was longer than the N16 molecule. (2) The most striking identity of 94.1% was observed among the acidic amino acids, resulting in a high conservation of the four acidic regions. An additional sequence match was determined at the sequence CCYTDDD, containing the phosphorylation site. (3) The Cys residue commonly present in the first half of both components was conserved fairly well (85.5% identity). (4) A sequence typically seen at the heparin-binding site of N16 was lacking in N14. (5) A slight difference was detected in the number of aromatic amino acids (particularly Tyr and Trp), which were contained in higher amounts in N14.

RT-PCR revealed that the N66 mRNA was expressed in both the dorsal region of the mantle and the mantle edge; the former may be responsible for the nacreous layer formation and the latter for the prismatic layer formation (8). In contrast, the N14 mRNA was expressed at extremely high levels only in the dorsal region of the mantle (Fig. 6).

A crystallization experiment was carried out using matrix components in each of the following three combinations: (1) WISM only, (2) N66 and N14 without WISM, (3) N66 and N14 with WISM as a substrate. The result was fundamentally the same with that obtained from *P. fucata* (2). N66 and N14 inhibited crystallization in the free state (without substrate membrane) but they could also induce platy aragonite crystals morphologically different from those formed spontaneously, once adsorbed onto the WISM membrane. Furthermore, addition of the mixture of N66 and N14 to the crystallization solution initiated more crystal formation compared with the separate addition of these two components. More than two days of incubation resulted in crystalline layers of flat aragonite highly similar to the nacreous layer of *P. maxima* (Figs. 7A and 7B).

DISCUSSION

As the calculated molecular weight of N14 was 13,734 Da, which is in agreement with that estimated by SDS-PAGE, the cDNA represents the full-length copy of the N14 mRNA coding regions, and no post-translational modifications can be assumed for the protein. In contrast, the calculated molecular weight of N66 was 59,814, which is slightly smaller than that estimated by SDS-PAGE as in the case of nacrein. Therefore, we cannot exclude the possibility of post-

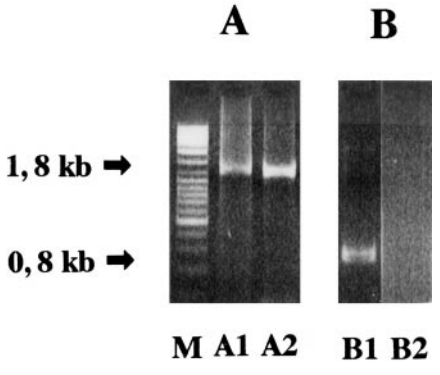

FIG. 6. (A) RT-PCR of N66 mRNA in the dorsal region of the mantle epithelium (A1) and in the mantle edge (A2). (B) RT-PCR of N14 mRNA in the dorsal region of the mantle epithelium (B1) and in the mantle edge (B2).

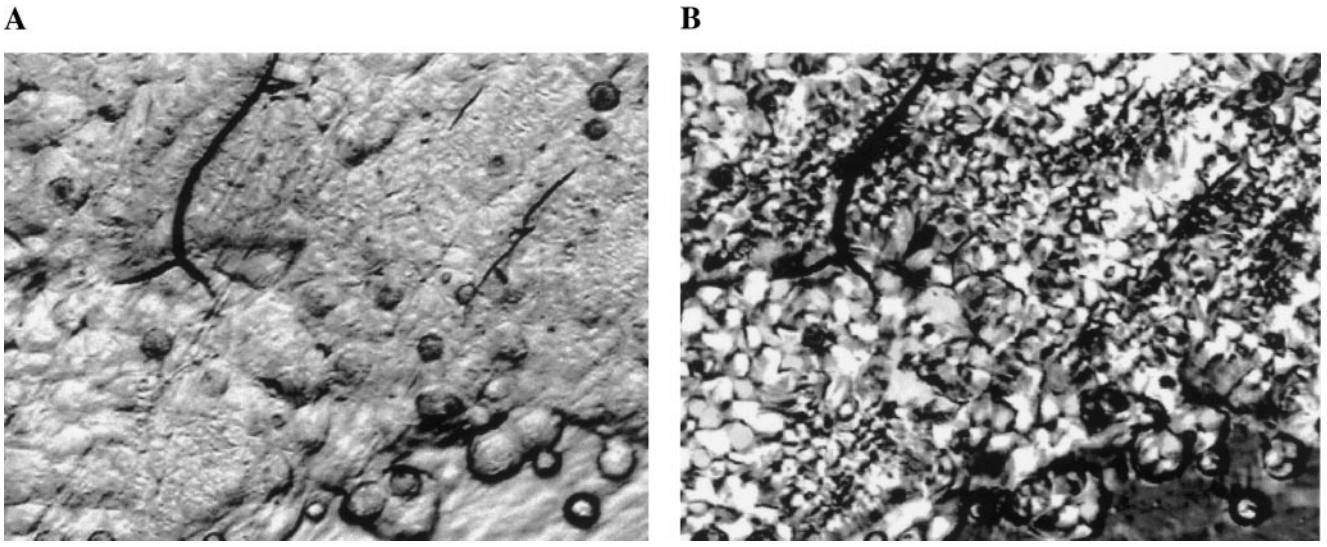

FIG. 7. Differential interference contrast microscopy images of the crystal layer of platy aragonite covering the WISM membrane in the presence of the mixture of each 2.5 μg/ml N66 and N14, similar to that in the nacreous layer. (A) Observation under the normal light (×600). (B) Observation under the crossed nicols (×600).

translational modifications of N66 such as sulfation by sulfated polysaccharides in spite of the negative reaction of sugars by lectin-peroxidase test, since sulfate residues might interrupt the reaction in some cases. An additional 28 kDa component was thought to be the dimer of N14 as it had the same N-terminal amino acid sequences with N14. In addition, N66 and N14 appear to be homogeneous proteins with nacrein and N16 in *P. fucata,* respectively.

As already found in nacrein of *P. fucata* (1), N66 was composed of two domains, one CA-like and another a repeat domain, which is inserted into the former. In contrast to nacrein, which had its highest homology with human CAII, N66 showed higher homology with human CAVII than CAII, although the meaning of this finding remains unclear. Judging from the structural analyses and crystallization experiments, we believe a multi-functional molecule N66 may serve as a producer of HCO_3^- through the CA-like domain and may also relate to crystal formation through the repeat region. We also believe that N14 may be involved in the formation of the membranous structure of the WISM as a microfibrillar matrix through the keratin-like sequence (23) and relate to crystal formation in a similar manner as N66 through the NG repeat sequence or acidic amino acids. However, how each of the organic matrix components i.e., WISM, N66 and N14, take part in each step of shell formation such as the nucleation of crystal (regulation of polymorphism), crystal growth (regulation of crystal morphology) and termination of crystal growth, is still open to speculation. It appears that N66 may be involved in the formation of both the nacreous and prismatic layers, while N14 seems to be specifically involved in the formation of the nacreous layer by the results of RT-PCR. In addition, the crystallization experiment clarified that N66 and N14 might react with each other to promote the nucleation of aragonite crystals. Elucidation of the precise process of this reaction is now underway in our laboratory.

Based on the sequence analysis of N16 in *P. fucata,* Samata *et al.* (2) have noted the essential involvement of the NG repeat region for the formation of the nacreous layer. A similar repeat domain composed of Gly-X-Asn (X = Asp, Asn or Glu) was reported from nacrein of *P. fucata* (1). We also found well comparable repeat sequences in N66 and N14 of *P. maxima,* except for the length of the sequences. Although a sequence characterized with the repeat of Asn and Gly has been reported from several proteins, such as the halocyanin (24), precursor of QID3, a kind of cell wall protein (22) and lipoprotein in *Mycoplasma penetrans* (25), it accounted for only a small portion of the whole residues of these proteins. Therefore, this is the first report of an unusually long repeat sequence, accounting for about 32% in N66 and 21% in N14 of the length of whole residues. The presence of such a long specific sequence leads us to believe that the sequence may be associated with some important role in calcification. Until now, no definite idea has been proposed about the function of such a repeat sequence. However, the aforementioned results at least imply that the regulation of crystal formation by this region requires Asn instead of Asp, in contrast to a commonly accepted central tenet of shell formation, that a regular arrangement of carboxyl groups of Asp results in a near-match of the spacing with Ca^{2+} in $CaCO_3$ crystal lattices. Samata *et al.* (2) proposed the possibility that the NG repeat can situate on the surface of N16, enabling this molecule to react with other molecules such as Ca^{2+} on the analogy of the Gly-loop, typically seen in a Gly-rich region. An addi-

tional idea is the involvement of the sequence in adsorption of the components onto the other matrix components or crystal surfaces, enabling the components to be involved in the regulation of crystal formation in a manner similar to that demonstrated by Wheeler (26). According to this theory, the longer repeat region may be related to the stronger reaction of the proteins with Ca^{2+}-molecules, matrix components and crystals, which would result in the excellent calcification ability of N14.

The comparison of the amino acid sequence between the two kinds of the homogeneous proteins (N66 and nacrein, and N14 and N16) demonstrates that the consensus sequences are located at the presumably significant regions for the formation of tertiary structure and function of the components. They are the active site of CAII in N66/nacrein and the acidic amino acids, Cys and phosphorylation site in N14/N16. The results correspond to the idea that the sequence of the functionally important region in homogeneous protein molecules can be conserved regardless the phylogenetic relationship (27). Alternatively, phylogenetic relationship may be reflected in the variation of the amino acid sequence of the remaining regions. Degens *et al.* (28) demonstrated that matrix compositions vary from species to species, even within closely related groups, on the amino acid level. This is the first report relating the phylogenetic comparison of the amino acid sequences of the homogeneous proteins in the organic matrices. Further analysis of diverse species of the different phylogenetic relationship will propose important information for understanding the process of molecular evolution happened on organic matrix.

ACKNOWLEDGMENTS

We are grateful to Dr. N. Watabe, Dr. H. Hirano, Dr. M. Sato, Dr. N. Fujiwara, Dr. K. Wada, Dr. N. Wada, and Dr. A. Machii for their valuable discussions. We also thank Mrs. C. Horita and Mr. S. Akera for kindly providing the samples for cDNA cloning. This work was supported in part by a grant (Development of Fundamental Technologies for Effective Genetic Improvement of Aquatic Organisms) from the Ministry of Agriculture, Forestry and Fisheries of Japan.

REFERENCES

1. Miyamoto, H., Miyashita, T., Okushima, M., Nakano, S., Morita, T., and Matsushiro (1996) *Proc. Natl. Acad. Sci. USA* **93,** 9657–9660.
2. Samata, T., Hayashi, N., Kono, M., Hasegawa, K., Horita, C., and Akera, S. (1999) *FEBS Lett.* **462,** 225–229.
3. Weiner, S., and Hood, L. (1975) *Science* **190,** 987–989.
4. Weiner, S. (1979) *Calcif. Tissue Int.* **29,** 163–167.
5. Nakahara, H., Bevelander, G., and Kakei, M. (1982) *Venus* **41,** 3–46.
6. Samata, T. (1988) *Venus* **47–2,** 127–140.
7. Bowen, C., and Tang, H. (1996) *Comp. Biochem. Physiol.* **115A–4,** 269–275.
8. Sudo, S., Fujikawa, T., Nagakura, K., Ohkubo, T., Sakaguchi, K., Tanaka, M., Nakajima, K., and Takahashi, T. (1997) *Nature* **387,** 563–564.
9. Shen, X., Belcher, A. M., Hansma, P. K., Stucky, G. D., and Morse, D. E. (1997) *J. Biol. Chem.* **272,** 32472–32481.
10. Sarashina, I., and Endo, K. (1999) *Amer. Mineral.* **83,** 1510–1515.
11. Crenshaw, M. A. (1972) *Biomineralization* **6,** 6–11.
12. Samata, T., and Krampitz, G. (1982) *Malacologia* **22,** 225–233.
13. Cariolou, M. A., and Morse, D. E. (1988) *J. Comp. Phys. B* **157,** 717–729.
14. Samata, T. (1990) *The Veliger* **33,** 191–201.
15. Wheeler, A. P., George, J. W., and Evans, C. A. (1981) *Science* **212,** 1397–1398.
16. Wheeler, A. P., Rusenko, K. W., George, J. W., and Sikes, C. S. (1987) *Comp. Biochem. Physiol.* **87B–4,** 953–960.
17. Greenfield, E. M., Wilson, D. C., and Crenshaw, M. A. (1984) *Am. Zool.* **24,** 925–932.
18. Albeck, S., Weiner, S., and Addadi, L. (1996) *Chem. Eur. J.* **2–3,** 278–284.
19. Falini, G., Albeck, S., Weiner, S., and Addadi, L. (1996) *Science* **271,** 67–69.
20. Belcher, A. M., Wu, X. H., Christensen, R. J., Hansma, P. K., Stucky, G. D., and Morse, D. E. (1996) *Nature* **381,** 56–58.
21. Chomoczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
22. Lora, J. M., de la Cruz, J., Benitez, T., Llobell, A., and Pintor-Toro, J. A. (1994) *Mol. Gen. Genet.* **242**(4), 461–466.
23. Fratini, A., Powell, C. B., and Rogers, E. G. (1993) *J. Biol. Chem.* **268**(6), 45111–45118.
24. Matter, S., Scharf, B., Kent, S. B. H., Rodewald, K., Oesterhelt, D., and Engelhard, M. (1994) *J. Biol. Chem.* **269–21,** 14939–14945.
25. Ferris, S., Watson, H. L., Montagnier, N. L., and Blanchare, A. (1995) *FEMS Micriobiol. Lett.* **130,** 313–320.
26. Wheeler, A. P. (1992) in Hard Tissue Mineralization and Demineralization (Suga, S., and Watabe, N., Eds.), pp. 171–187, Springer-Verlag, Tokyo.
27. Dockerson, R. E. (1971) *J. Mol. Evolution* **1,** 26–34.
28. Degens, E. T., Spencer, D. W., and Parker, R. H. (1967) *Comp. Biochem. Physiol.* **20,** 553–579.