A new matrix protein family related to the nacreous layer formation of *Pinctada fucata*

Tetsuro Samata[a,*], Nakanobu Hayashi[b], Makiko Kono[a], Kyoko Hasegawa[c], Chie Horita[d], Syusaku Akera[d]

[a] *Laboratory of Cell Biology, Faculty of Environmental Health, Azabu University, 1-17-71 Fuchinobe, Sagamihara 229-8501, Japan*
[b] *Omgen Inc., Itabashi, Tokyo 173-0031, Japan*
[c] *New Japan Alfa Inc., Atsugi, Kanagawa 243-0811, Japan*
[d] *Tasaki Institute for Marine Biological Research, Hiwasa, Tokushima 779-23, Japan*

Received 2 August 1999; received in revised form 2 October 1999

Abstract We have isolated a new matrix protein family (N16) which is specific to the nacreous layer of the Japanese pearl oyster, *Pinctada fucata*, and have cloned and characterized the cDNAs coding for the components. Analysis of the deduced amino acid sequence revealed that N16 showed no definitive homology with other proteins. The in vitro studies of the crystallization clarified that N16 induced aragonite crystals when fixed on the substrate but inhibited crystal formation without it. The aragonite crystals showed platy morphology different from those formed inorganically, and long intervals of incubation resulted in crystalline layers highly similar to the nacreous layer.

Key words: Biomineralization; Nacreous layer; Organic matrix; cDNA cloning; Crystallization experiment

1. Introduction

Crenshaw [1] first identified a water soluble matrix (WSM) containing a Ca-binding glycoprotein as the major component of the shell of *Mercenaria mercenaria*. Subsequent studies have focused on fractionating the WSM into different components [2–4], and have obtained much information on the structure and function of the organic matrix components related to shell formation. Although the organic matrix has been considered to control the polymorphism and the morphology of calcium carbonate crystals, the precise processes remain to be clarified. Wilbur and Watabe [5] suggested that the water insoluble organic matrix (WISM) was involved in the polymorphic control. In contrast, Falini et al. [6] showed that the WSM, when combined with chitin and fibroin could determine the polymorphism. Belcher et al. [7] demonstrated that the WSM was sufficient to control the polymorphism of calcium carbonate crystals without the WISM sheet.

Several genes encoding matrix components in molluscan shells have recently been isolated and their deduced amino acid sequences have been clarified [8–10]. Miyamoto et al. [8] found a polyanionic nacrein with two domain structures of CA and Gly-Xaa-Asn (Xaa = Asp, Asn or Glu) repeat domain in the nacreous layer of *Pinctada fucata*. Subsequently, Gly-Ala rich (MSI 60) and Gly-Glu rich (MSI 31) framework proteins were isolated from *P. fucata* and found to be specific for the nacreous and prismatic layers, respectively [9]. Most recently, a lustrin A showing a modular structure and multifunctionality was identified from the nacreous layer of *Haliotis rufescens* [10].

Genus *Pinctada* belongs to one of the unique families of bivalve mollusca, Pteriidae, whose shells are composed of two polymorphs of calcium carbonate, i.e. calcite in the prismatic layer and aragonite in the nacreous layer, and therefore is a suitable model for investigating how molluscs control the polymorphism. We here present the complete amino acid sequence of the proteins that are specific for the nacreous layer and its role related to the formation of the aragonite crystals. This is the first report on the structure and function of the molluscan shell matrix protein.

2. Materials and methods

2.1. Animals

Samples for the cDNA cloning and protein analysis were derived from Japanese pearl oyster, *P. fucata martenshii*, cultured at Tasaki Institute for Marine Biological Research (Tokushima, Japan).

2.2. Extraction of the organic matrix components

The nacreous layer of *P. fucata* was decalcified with 10% EDTA-4Na (pH 7.6). After dialysis against DW, the WSM and WISM were separated by centrifugation at $15\,000 \times g$ for 15 min. The precipitated WISM was further extracted using dilute alkali solution of NH_4OH (pH 8.5), and the alkali soluble organic matrix (ALSM) was obtained by centrifugation at $15\,000 \times g$ for 20 min.

2.3. N-terminal amino acid determination

To determine the amino terminal sequence, a single band with molecular weight of 16 kDa obtained from SDS-PAGE of the ALSM was blotted onto a PVDF filter (Millipore) using a dry blotting system (Nippon Eido), cut off from the filter and applied to a Beckman LF3000 amino acid sequence analyser (Beckman).

2.4. Crystallization experiments

Two crystallization solutions were prepared, one which induces calcite (calcitic crystallization solution) and one which induces aragonite (aragonitic crystallization solution). The former was a supersaturated solution of $CaHCO_3$ prepared by purging a stirred aqueous suspension of $CaCO_3$ with CO_2 [11]. The latter was obtained by addition of 50 mM $MgCl_2$ to the former solution. Crystallization experiments were carried out by adding 0.1 to 50 μg of the matrix component to the crystallization solution with or without the substrate WISM membrane, which was fixed on a glass plate situated at the bottom of a 24-hole microplate at 18°C.

An Olympus BH2 differential interference contrast microscope (Olympus) and Rigaku PSPC-MDG 2000 microarea X-ray diffractometer (Rigaku) were used for morphological observation and identification of the induced crystals.

*Corresponding author. Fax: +81 (42) 769-1938.
E-mail: samata@azabu-u.ac.jp

PII: S0014-5793(99)01387-3

2.5. cDNA cloning

Mantle tissue was removed from the molluscan body and frozen in liquid nitrogen. We extracted RNA from the dorsal region of the mantle by the method of acid guanidium thiocyanate-phenol chloroform extraction [12], and poly(A)$^+$ RNA was fractionated in an Oligo(dT) column. The purified poly(A)$^+$ RNA was applied to construction of cDNA using a Time Saver cDNA Synthesis Kit (Amersham-Pharmacia). The cDNA was inserted into the γ gt 10 express vector and packaged with a γ DNA Packaging System (Amersham-Pharmacia). After the library was screened by probes as described below, two positive clones were isolated, and they were subcloned into the plasmid vector with pGEM-T (Promega) and sequenced by the cycle sequencing method using a PRISM 377 automated DNA sequencer (PE Biosystems). The cDNA fragments which contained a 5′-end sequence were isolated by the 5′-RACE method using a Marathon cDNA Amplification Kit (Clontech). The PCR products were subcloned and sequenced by the method described above.

RT-PCR was carried out using oligonucleotides described below as the primers and mRNAs in the dorsal region of the mantle and the mantle edge as the templates. The PCR products were subcloned and sequenced to search for the mutated genes.

2.6. Probes for cDNA library screening and primers for RT-PCR

Two degenerate oligonucleotides, 5′-GAYMGNMGNGAYAA-YGGNGGNAARAAR-3′ and 5′-GGNGGNAARAARAARTGYT-TYTTYTGY-3′, were synthesized for the library screening to correspond to a portion of the N-terminal sequence DRRDNGGKKK and GGKKKCFFC, determined for N16 by amino acid sequence analysis. Two oligonucleotide primers, 5′-TTCTGTGAGGCGGCGC-CTGGAATC-3′ corresponding to nucleotide (nt) 4–24 of one clone isolated from the cDNA library and 5′-GTTCATACTTTACTATA-TAATTTTACA-3′ matching the antisense sequence of nt 700–721 of the same clone were used for RT-PCR.

3. Results

We separated a protein component of 16 kDa length from the ALSM of the nacreous layer of *P. fucata* by SDS-PAGE and designated it N16. Amino acid sequence analysis of N16 showed the presence of at least two closely related sequences,

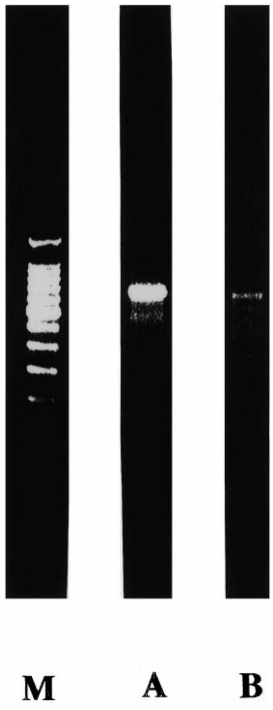

Fig. 1. RT-PCR of the mRNA in the dorsal region of the mantle epithelium. N16 mRNA was expressed at extremely high levels in the dorsal region but only in trace amounts in the mantle edge. Lane M, marker DNA; lane A, RT-PCR using N16 mRNA isolated from the dorsal region; lane B, RT-PCR using N16 mRNA isolated from the mantle edge.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1	M	K	C	T	L	R	W	T	I	T	A	L	V	L	L	G	I	C	H	L	A	R	P	A	Y	H	K	K	C	G	R	Y	S
2																																	
3		T																							F	R	T						

	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
1	Y	C	W	I	P	Y	D	I	E	R	D	R	Y	D	N	G	D	K	K	C	C	F	C	R	Y	A	W	S	P	W	Q	C	N
2													R				G								N								K
3																																	K

	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
1	E	E	E	R	Y	E	W	L	R	C	G	M	R	F	Y	S	L	C	C	Y	T	D	D	D	N	G	N	G	N	G	N	G	N
2		D										H	K			Y	M												D				
3		D										H	K			Y	M																

	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131
1	G	N	G	L	N	Y	L	K	S	L	Y	G	G	Y	G	N	G	N	G	E	F	W	E	E	Y	I	D	E	R	Y	D	N
2	x	x		F																												K
3	x	x		F																												K

Fig. 2. Deduced amino acid sequence of the N16 protein family. At all positions lacking an amino acid, sequence identity is with the component 1. x indicates the absence of an amino acid. Shadowed amino acids correspond to the putative signal peptides. N16-1 and N16-2 were coded by the two clones isolated from the cDNA library, while N16-3 was coded by the clone prepared from the RT-PCR product.

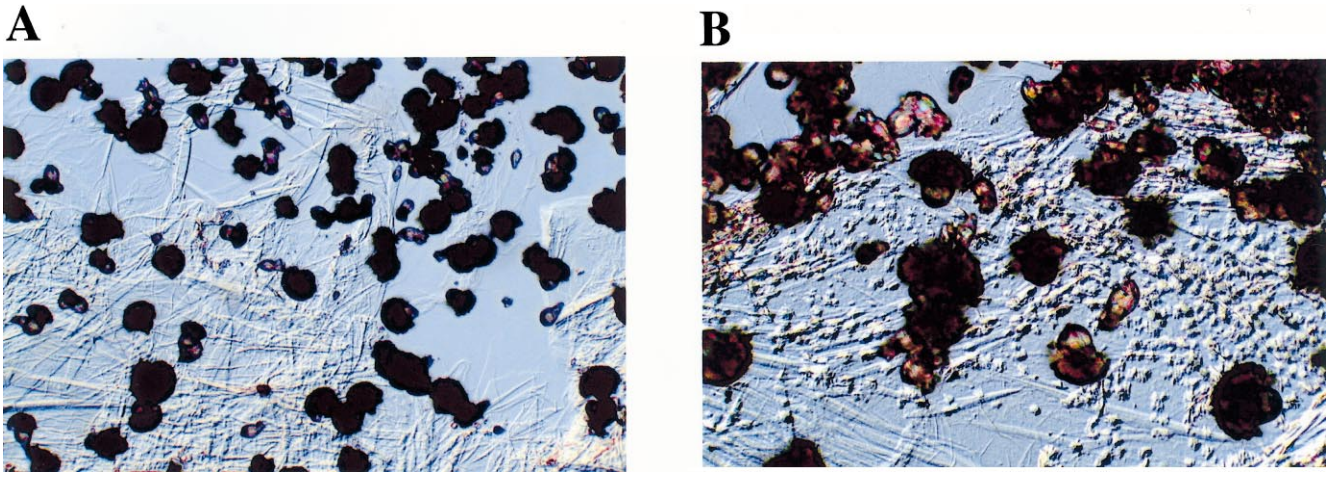

Fig. 3. Differential interference contrast microscopy images showing two types of aragonite crystals grown on the WISM membrane. A: Spherical crystals, grown on and around the membrane in the absence of N16 (×150). B: Mixture of spherical and platy crystals grown on the WISM membrane in the presence of 3 μg/ml N16 (×150). At low concentration of N16 below 5 μg/ml, the crystals formed showed a mixture of spherical and platy morphology, with the latter being located only on the membrane.

i.e. (1) AYHKKCGRYSYCWIPYDIERDRYDNGDKKKCFFC and (2) same sequence except for residue 2 (F), 3 (R), 4 (T), 23 (R) and 27 (G), both of which were different from nacrein found by Miyamoto et al. [8].

In the present study, two cDNA clones were isolated from the cDNA library, and one additional clone was prepared from the RT-PCR product. In addition, RT-PCR demonstrated that the N16 mRNA was expressed at extremely high levels in the dorsal region of the mantle, which region may be responsible for the nacreous layer formation, but only in trace amounts at the mantle edge, which region may be responsible for the prismatic layer formation (Fig. 1). These three clones revealed open reading frames encoding 129 and 131 amino acids, with the first 23 amino acids apparently comprising a signal peptide. From the deduced amino acid sequences, three protein components (N16-1, N16-2 and N16-3) were identified (Fig. 2), whose N-terminal sequences matched those determined by the amino acid sequence analysis of the 16 kDa component except for residues 53, 56 and 57. Although they shared a high sequence identity of about 89%, N16-1 component had no Lys but rather had Asn as the C-terminal amino acid. The common features among these components were high contents of Gly, Tyr, Asn and Cys. Acidic amino acids were present in high concentration in only four acidic regions. Moreover, the contents of aromatic amino acids were also high. On the other hand, Val was absent in all the components, and Met, His and Gln were contained in only one residue. Overall, N16 possessed a hydrophilic and acidic nature with a p*I* of 4.8 (N16-1) or 5.1 (N16-2 and N16-3), respectively. In addition, remarkable repeats of NG were found, constituting about 9 to 12% of all amino acid residues.

A search of the Swissprot protein database clarified that N16 had no definitive homology with known proteins except for a partial sequence identity with high glycine/tyrosine proteins (HGTPs) of rabbit hair [13].

The motif analysis revealed a phosphorylation site with a typical sequence of TDDD at Thr of the 87th residue. In addition, a sequence of x-BBB-xx-B-x (B = a basic amino acid, x = a non-basic amino acid) typically seen in the heparin-binding site was identified between residues 25 and 32 of the N16-1 and N16-2 components. With the exception of

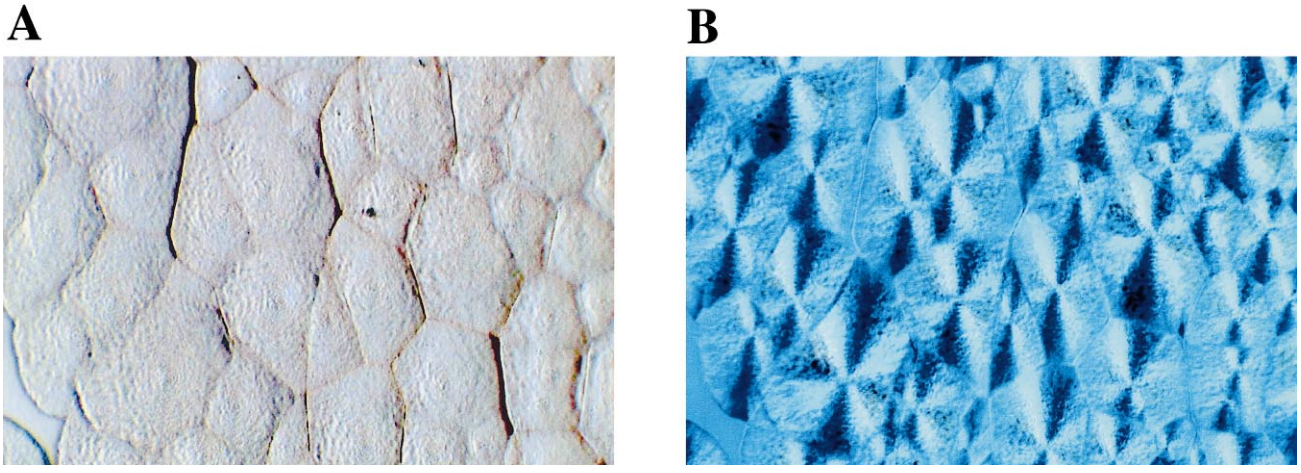

Fig. 4. Differential interference contrast microscopy images of crystal layers of platy aragonite developed on the WISM membrane in the presence of 10 μg/ml N16. A: Observation under normal light (×600). B: Observation under crossed nicols (×600).

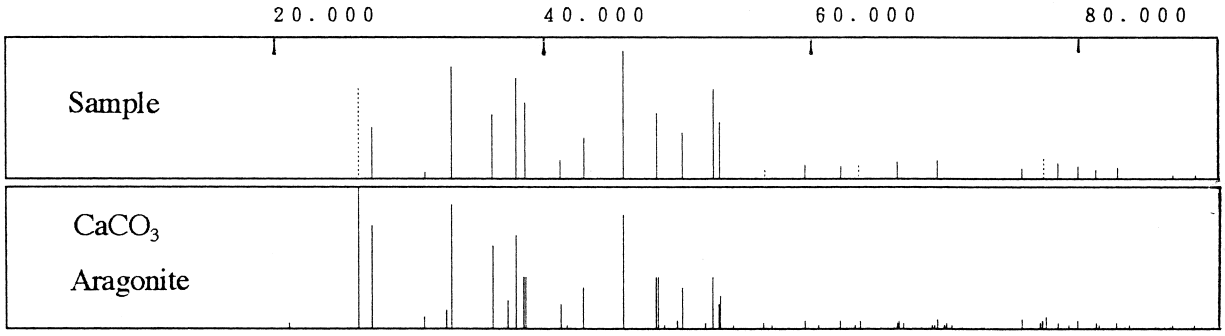

Fig. 5. X-ray diffractional analysis of the platy crystal layer developed on the WISM membrane in the presence of 10 μg/ml N16.

these sequences, N16 showed no sequence similarity to any of the known motifs, such as *N*-glycosylation, Ca^{-}-binding (EF hand protein), heavy metal-binding or cell adhesion.

Secondary structure analysis using the Chou and Fasman method [14] indicated that N16 contained a high degree of β-turns, some β-sheets and a few α-helices.

The crystallization experiment was carried out using matrix components in each of the following three combinations: (1) WISM only; (2) N16 without WISM; and (3) N16 with WISM as a substrate. For the last two combinations, N16 was mixed with the crystallization solution and added to the microplate. At first, in the absence of N16, needle-like or spherical aragonite crystals grew on and around the WISM membrane from the aragonitic crystallization solution, as is generally observed for a glass substrate (Fig. 3A). Crystals grown in vitro without a substrate in the presence of up to 1 μg/ml N16 showed needle-like or spherical morphology of aragonite that was characteristic of those formed inorganically. Increases of the concentration of N16 decreased the size and number of aragonite formed, and crystal formation was completely inhibited at addition of 10 μg/ml N16. In contrast, crystals grown in the presence of N16 with the substrate WISM displayed a very different morphology, forming oval to hexagonal platy aragonite crystals (Fig. 3B). At higher concentration of N16 (more than 10 μg) under the long intervals of incubation (more than 2 days) this resulted in aggregates of the tabular aragonite that coated the surface of the WISM membrane, and the crystals exhibited an architecture highly similar to those of the nacreous layer of *P. fucata* (Fig. 4A, B). X-ray diffractional analysis revealed that these were aragonite polycrystals (Fig. 5) that exhibited uniform extinction under cross-polarized light (Fig. 4B).

The experiment using the calcitic crystallization solution was carried out by adding N16 in the presence of the substrate WISM. Addition of N16 tended to result in decreases in the size and number of calcite crystals formed, indicating the inhibition of crystal formation in a concentration-dependent manner.

4. Discussion

Because N16 is insoluble in EDTA, it has generally been treated as an EDTA insoluble matrix and could not be separated from the WISM membrane. For the first time, we here isolated and purified it by using a dilute alkali solution.

Judging from the nucleotide structure, it is very likely that the cDNAs represent the full-length copy of the N16 mRNA coding region. Analysis of the deduced amino acid sequence revealed that N16 was a new protein family, which differed greatly from the protein components, recently isolated from molluscan shells [8–10]. N16 may be a kind of membrane protein-like component specific for the nacreous layer, and might be involved in both crystal formation and the formation of the WISM as microfibrillar matrix in a manner similar to FGTPs [13]. Support for the specific involvement of N16 in the nacreous layer formation comes from the result of our RT-PCR, which clarified that N16 mRNA was expressed only in the dorsal region of the mantle.

The results of the crystallization experiments showed that while N16 inhibited the crystallization in the free state, it could also – when adsorbed on the WISM membrane – induce platy aragonite specific to the nacreous layer. A similar observation was reported by Greenfield et al. [15]. On the other hand, N16 was not involved in the regulation of the $CaCO_3$ crystal polymorphism. This result is opposed to those of Falini et al. [6] and Belcher et al. [7], who demonstrated the involvement of the WSM components in the control of polymorphism with or without substrate. This might be due in part to the different conditions used in these crystallization experiments.

Although N16 had the capacity for Ca-binding (data not shown), it contained no motif that was typical of a Ca-binding protein, and thus the determination of the Ca-binding site remains only speculative. The possible candidates are acidic amino acids distributed among four acidic domains. In addition, we cannot exclude the possibility of the Ca-binding of heparin or heparan sulfide through their sulfate side chains in spite of the negative reaction of sugars by lectin-peroxidase test in N16 (data not shown), since the sulfate residues might interrupt the reaction in some cases. We also focused on the NG repeat regions which can form the structural motif known as the glycine loop, and this flexible structure might play some important roles in the reaction with other molecules, such as Ca^{2+}. The repeat regions may also serve for adsorption of N16 to other matrix components or aragonite crystals, thus controlling the crystal formation in a manner similar to that demonstrated by Wheeler [16].

Plans to further elucidate the molecular mechanism of the nacreous layer formation by crystallization experiments using the recombinant proteins expressed from the N16 genes are currently underway in our laboratory.

Acknowledgements: We are grateful to N. Watabe for his critical reading of the manuscript. We also thank H. Hirano, M. Sato, N. Fujiwara, K. Wada, N. Wada and A. Machii for their valuable discussions. 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. Further information of the DNA sequences can be obtained at the DDBJ/EMBL/GenBank nucleotide databases under accession number AB023067.

References

[1] Crenshaw, M.A. (1972) Biomineralization 6, 6–11.
[2] Weiner, S. (1979) Calcif. Tissue Int. 29, 163–167.
[3] Cariolou, M.A. and Morse, D.E. (1988) J. Comp. Phys. B 157, 717–729.
[4] Samata, T. (1990) The Veliger 33, 191–201.
[5] Wilbur, K.M. and Watabe, N. (1963) Ann. N.Y. Acad. Sci. 109, 82–112.
[6] Falini, G., Albeck, S., Weiner, S. and Addadi, L. (1996) Science 271, 67–69.
[7] Belcher, A.M., Wu, X.H., Christensen, R.J., Hansma, P.K., Stucky, G.D. and Morse, D.E. (1996) Nature 381, 56–58.
[8] Miyamoto, H., Miyashita, T., Okushima, M., Nakano, S., Morita, T. and Matsushiro, A. (1996) Proc. Natl. Acad. Sci. USA 93, 9657–9660.
[9] Sudo, S., Fujikawa, T., Nagakura, K., Tanaka, M., Nakashima, K. and Takahashi, T. (1997) Nature 387, 563–564.
[10] Shen, X., Belcher, A.M., Hansma, P.K., Stucky, G.D. and Morse, D.E. (1997) J. Biol. Chem. 272, 32472–32481.
[11] Oomori, T. and Kitano, Y. (1991) in: The Mechanism and Phylogeny of Mineralization in Biological Systems (Suga, S. and Nakahara, H., Eds.), pp. 501–505, Springer-Verlag, Tokyo.
[12] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[13] Fratini, A., Powell, B.C. and Rogers, G.E. (1993) J. Biol. Chem. 268, 4511–4518.
[14] Chou, P.Y. and Fasman, G.D. (1978) Annu. Rev. Biochem. 47, 251–276.
[15] Greenfield, E.M., Wilson, D.T. and Crenshaw, M.A. (1984) Am. Zool. 24, 925–932.
[16] Wheeler, A.P. (1992) in: Hard Tissue Mineralization and Demineralization (Suga, S. and Watabe, N., Eds.), pp. 171–187, Springer-Verlag, Tokyo.