Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 304 (2003) 11–17
BBRC
www.elsevier.com/locate/ybbrc

Molecular cloning of a cDNA encoding a soluble protein in the coral exoskeleton

Isao Fukuda,[a] Syuuichi Ooki,[a] Toshihiro Fujita,[a] Emi Murayama,[b] Hiromichi Nagasawa,[b] Yeishin Isa,[c] and Toshiki Watanabe[a,*]

[a] *Division of Marine Biosciences, Ocean Research Institute, The University of Tokyo, 1-15-1 Minamidai, Nakano, Tokyo 168-8639, Japan*
[b] *Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo 113-8657, Japan*
[c] *Department of Chemistry, Biology and Marine Sciences, Faculty of Science, University of the Ryukyus, 1-Senbaru, Nishihara-cho, Okinawa 903-0213, Japan*

Received 10 March 2003

Abstract

Organic substances were extracted from the calcified exoskeleton of the reef coral *Galaxea fascicularis*. In an SDS–PAGE analysis of the extract, a protein with an apparent molecular mass of 53 kDa was detected as well as two other weaker bands. A Ca^{2+} overlay analysis failed to find a Ca^{2+}-binding protein in the extract. Periodic acid Schiff staining indicated that the 53 kDa protein was glycosylated. A cDNA containing the entire open reading frame for this protein was obtained. Analysis of the deduced protein sequence suggests that the protein, named galaxin, is synthesized as a precursor consisting of a signal peptide, a propeptide sequence, and a mature protein of 298 amino acids. Galaxin exhibits a novel amino acid sequence which is characterized by a tandem repeat structure. *Galaxin* transcripts were detected in the adult coral, but not in planktonic larvae.

Keywords: Calcification; Calcium carbonate; Scleractinia; Hermatypic coral; Exoskeleton; Glycoprotein; Planula

Skeletal structures that are mineralized with inorganic calcium crystals are found in a wide range of animals including both vertebrates and invertebrates [1,2]. In vertebrates, endoskeletal structures are found that are mineralized with hydroxyapatite, whereas in invertebrates exoskeletal structures mineralized with calcium carbonate crystals (usually aragonite or calcite) are common. Prior to the calcification of skeletal structures, cells involved in skeletogenesis synthesize organic matrices and secrete them into the extracellular space where calcification takes place. In corals, for instance, Isa [3] showed using electron microscopy that the initial phase of calcification occurred in close association with organic substances secreted from the calicoblastic ectoderm. Thus, it has been suggested that macromolecules in the matrices may regulate aspects of calcification such as nucleation (formation of miniscule seeds of crystals) and determination of the polymorph (e.g., aragonite/calcite), shape or size of crystals [4–7].

Researchers have isolated and characterized components of organic matrices, especially proteins, in diverse animal species including both vertebrates and invertebrates, in search for key molecules that play important roles in skeletogenesis. In the vertebrate bone, some of the non-collagenous proteins have been suggested to regulate calcification. Bone sialoprotein, for example, is implicated in inducing nucleation of hydroxyapatite crystals [8], whereas osteopontin, osteocalcin, and matrix Gla-protein have been suggested to negatively regulate calcification [9–11]. A number of matrix proteins in the calcified exoskeleton have also been characterized for the amino acid sequence in invertebrates such as crustaceans [12–20] and mollusks [21–29]. Some of these proteins have also been characterized for the function [19,21,25,26].

Biochemical studies on the exoskeleton of scleractinian corals demonstrated the presence of proteins and

* Corresponding author. Fax: +81-3-5351-6534.
E-mail address: toshi@ori.u-tokyo.ac.jp (T. Watanabe).

0006-291X/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0006-291X(03)00527-8

the amino acid composition of the matrix has been determined in a number of species [30–33]. However, no reports have been published to date on the primary structure or function of individual proteins. This article carries the first report on cDNA cloning and deduction of the primary structure of a matrix protein extracted from the calcified exoskeleton of a coral.

Materials and methods

Protein extraction and electrophoretic analyses. Colonies of *Galaxea fascicularis* were collected around Sesoko and Aka Islands, Okinawa, Japan. Apical parts of polyps covered by live tissues were collected, crushed into small pieces, and stirred in 1 N NaOH (2 h) to remove fleshy tissues. Skeletal pieces (10 g) were ground into fine powder with a mortar and pestle and stirred in 1 N NaOH (300 ml, 2 h) and subsequently 12.5% NaClO (300 ml, 2 h) to remove remaining soft tissues. The powder was then extensively washed in H_2O and decalcified in 0.5 M EDTA (300 ml, overnight). The decalcifying solution was centrifuged to remove insoluble materials and passed through two tandemly connected SEP-PAK C_{18} cartridges (Waters). The adsorbed macromolecules were eluted in 50% acetonitrile, concentrated by centrifugation under vacuum, and subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) or reverse-phase high performance liquid chromatography (HPLC). "Low molecular weight standard" (Amersham Pharmacia Biotech) and prestained molecular weight standard (Bio-Rad) were used in SDS–PAGE. The Ca^{2+} overlay analysis was carried out as described in [34].

Reverse-phase HPLC and protein sequencing. An ODP-50 column (4.6 × 250 mm, Shodex) was used with a 60-min linear gradient of 0–60% acetonitrile in 0.05% trifluoroacetic acid (TFA) at a flow rate of 1 ml/min. The eluate was collected in 60 fractions (each corresponding to 1 min of elution) and the protein content of the fractions was monitored using SDS–PAGE. Fractions containing the 53 kDa protein were combined, electrophoresed, and blotted onto PVDF membranes. The 53 kDa band, visualized by Coomassie brilliant blue (CBB) staining, was excised from a membrane and applied to a protein sequencer (model 476A, Applied Biosystems) in the pulsed-liquid mode. The 53 kDa protein was electro-eluted from acrylamide gels using a Maxyield-NP apparatus (model AE-6580, ATTO) and the eluate was concentrated by centrifugation in Ultrafree-MC (Millipore) and dried by centrifugation under vacuum. The protein was digested with *Staphylococcus aureus* V8 protease (Sigma) and the digest was applied to reverse-phase HPLC using a CAPCELL PAK C_{18} column (2.0 × 150 mm, Shiseido). Elution was performed with a 60 min linear gradient of 0–60% acetonitrile in 0.05% TFA (in the first round) or in 0.05% heptafluoroacetate (the second round) at a flow rate of 0.2 ml/min. Isolated peaks were applied to a sequencer as described above.

Preparation of total RNA and reverse transcription (RT)-PCR. Preparation of total RNA and synthesis of cDNA have been described [35]. To isolate a fragment of cDNA encoding the 53 kDa protein, a PCR was carried out with primers *D1* and *D2* (Fig. 3) using the following program: 30 s (5.5 min in the first cycle) at 94 °C, 30 s at 52 °C, and 1.5 min at 72 °C for 40 cycles. The PCR product was subcloned into the pGEM-T Easy vector (Promega) and sequenced. For PCR to study expression of *galaxin* and *LGfact*, the following program was used: 30 s (5.5 min only in the first cycle) at 95 °C, 30 s at 58 °C, and 40 s at 72 °C for 40 cycles. To detect the expression of *galaxin* and *LGfact*, the following primer sets were used: *F1* and *5′RACE2* (Fig. 3) and *actF* (5′-AAGCTCAAAGCAAACGTGGTA-3′) and *actR* (5′-AGGCTGG CTGGAACATGGC-3′), respectively.

Rapid amplification of cDNA ends. In the first round of 5′-RACE, cDNA was synthesized with the adult total RNA primed with *5′RACE1* (Fig. 3), and then poly(A)-tailed as described in [36]. The T_{17}*ADP* and *ADP* primers [36] were used as the sense primers and *5′RACE1* was used as the antisense primer in the following PCR: 1 min (5 min in the first cycle) at 94 °C, 1 min at 50 °C, and 3 min at 72 °C for 3 cycles, and then 1 min at 94 °C, 1 min at 55 °C, and 3 min at 72 °C for 37 cycles. The second round of 5′-RACE was carried out using a SMART RACE cDNA amplification kit (Clonetech) according to instructions from the manufacturer. *5′RACE2* (Fig. 3) was used as the antisense primer in the following PCR: 30 s (5.5 min only in the first cycle) at 94 °C, 30 s at 60 °C, and 3 min at 72 °C for 35 cycles. In 3′ RACE, cDNA was synthesized with the adult total RNA primed with T_{17}*ADP*. The *3′RACE* (Fig. 3) and *ADP* primers were used to amplify the cDNA in the following PCR: 30 s (5.5 min only in the first cycle) at 94 °C, 30 s at 54 °C, and 30 s at 72 °C for 40 cycles.

Results

Extraction of proteins from the skeleton and sequence analysis of the 53 kDa protein

Organic substances in the exoskeleton were solubilized in 0.5 M EDTA and proteins in the extract were analyzed using SDS–PAGE (Fig. 1). Three bands were reproducibly detected (Fig. 1A): a protein (53 kDa) was by far the most abundant, and the other two bands, at 45 and 20 kDa, appeared weaker. A Western blot was subjected to a Ca^{2+}-overlay analysis in which none of the CBB-stained bands exhibited significant Ca^{2+}-binding (Fig. 1B). Radioactivity was detected near the bottom of the autoradiogram. The corresponding area in a gel was not stained with CBB (data not shown), suggesting that the Ca^{2+}-binding is not due to a protein (see Discussion).

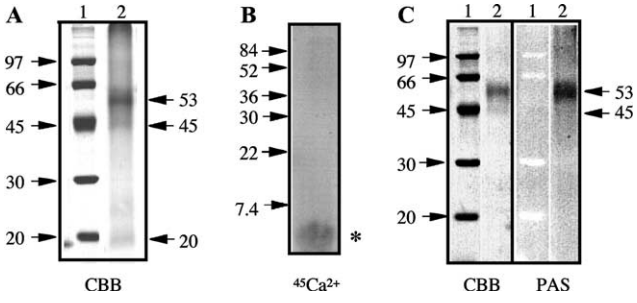


Fig. 1. Electrophoretic analyses of proteins extracted from the exoskeleton. (A) SDS–PAGE analysis of an eluate from C_{18} cartridges. Lane 1, molecular weight standards; lane 2, the eluate sample. The sample (derived from 10 g of skeleton) was electrophoresed on a 12% polyacrylamide gel and stained with CBB. The molecular masses of the standards and matrix proteins are indicated on the left and right, respectively (in kDa). (B) Ca^{2+} overlay analysis. An eluate (derived from 10 g of skeleton) was run on 15% polyacrylamide gel and transblotted. The molecular masses of the standards (in kDa) are indicated on the left. An asterisk indicates radioactive signals due to $^{45}Ca^{2+}$-binding. (C) Analysis of the combined fractions (44–52 min). Lanes 1, standards; lanes 2, combined fractions. The 53 kDa protein, as well as a faint band at 45 kDa, is detected with CBB staining (left panel). In PAS-staining of a Western blot, strong red/purple staining of the 53 kDa band, but not the 45 kDa band, is observed (right panel).

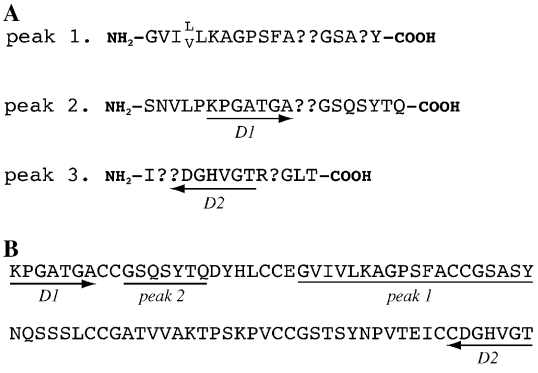


Fig. 2. Partial amino acid sequences of the 53 kDa protein. (A) Partial amino acid sequences of three fragments of the 53 kDa protein. Question marks indicate residues that could not be identified. Two degenerate oligonucleotide primers, D1 (sense) and D2 (antisense), were designed based on the amino acid sequences indicated with arrows. In designing D2, one of the unidentifiable residues was assumed to be cysteine. (B) Conceptual translation of an RT-PCR product. Sequences that are consistent with the peak 1 sequence and a part of the peak 2 sequence are underlined.

In an attempt to isolate the 53 kDa protein, extracts were subjected to reverse-phase HPLC. The protein did not form a discernible peak, but was eluted over a long period (44–52 min, data not shown). In these fractions, the 45 kDa protein was also detected (Fig. 1C). The 53 kDa protein stained positive in periodic acid Schiff (PAS) staining, indicating that it was glycosylated (Fig. 1C). In sequencing the N-terminus, signals were not detected in automated Edman degradation, suggesting the possibility of N-terminus blocking. The modification of the terminus was not likely to be formylation or acetylation, because treatments to remove formyl or acetyl groups did not result in detection of signals (data not shown).

Next, digests of the protein with endoproteinase V8 were subjected to reverse-phase HPLC. Partial amino acid sequences could be determined in two of the collected peaks (peaks 1 and 2 in Fig. 2A). Another peak, which apparently contained more than one peptide, was applied to a second round of HPLC (data not shown), and sequenced (peak 3 in Fig. 2A).

Fig. 3. Nucleotide sequence of a cDNA coding for galaxin (GenBank/EMBL/DDBJ Accession No.: AB086183). The following four sequences were merged to generate the sequence of 1671 bp: (1) T^{1}–A^{579} (derived from the second round of 5′-RACE), T^{580}–C^{758} (the first round of 5′-RACE), T^{759}–T^{969} (the original PCR product), and T^{970}–T^{1671} (the 3′-RACE). Conceptual translation of the ORF (A^{76}–A^{1104}) is shown below the nucleotide sequence in the one letter representation of amino acids. The positions of nucleotides and amino acids are indicated on the left. The putative signal peptide is underlined and the protease target sequence (RYRR) is boxed. Sequences corresponding to the PCR primers are overlined; F or R in parentheses indicate whether the primer is a sense primer (F) or an antisense (R) primer. Two potential N-glycosylation sites are circled.

cDNA cloning

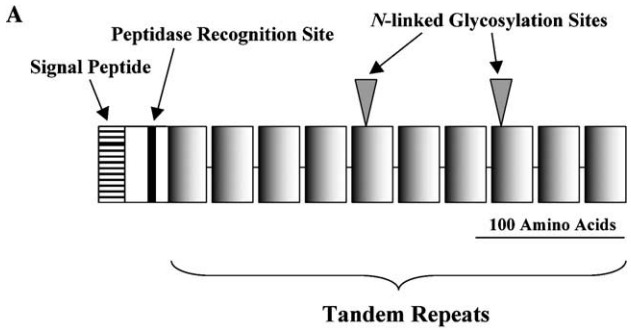
With degenerate oligonucleotide primers *D1* and *D2* (Figs. 2A and 3), a cDNA fragment of about 250 bp was amplified, which contained an open reading frame (ORF) for 70 amino acids (Fig. 2B). The translated sequence included the peak 1 sequence and a C-terminal part of the peak 2 sequence (Fig. 2B), indicating that the product was a part of a cDNA encoding the 53 kDa protein. Additional cDNA sequences were isolated using 3′ RACE and two rounds of 5′ RACE. By merging the sequences, a sequence of 1671 bp (excluding the poly(A) sequence) was obtained which contained a putative initiator codon, an upstream stop codon, a stop codon, and a variant (ATTAAA) of the typical polyadenylation signal 14 nucleotides upstream from the poly(A) sequence (Fig. 3). Thus, an ORF for 343 amino acids was revealed (Fig. 3). The deduced protein was named "galaxin" after the genus name of the coral.

To address the question whether the composite cDNA sequence was chimeric, RT-PCR was performed using primers derived from the 5′ and 3′ untranslated sequences. Two cloned PCR products had the same ORF sequences as in Fig. 3 except for substitution of two and three nucleotides, resulting in substitution of one and two amino acids, respectively (data not shown). These minor differences are likely to be due to polymorphism in the coral population, so we conclude that the ORF sequence was derived from a single transcript species.

In the deduced sequence, two potential N-glycosylation sites were found (Fig. 3). A computer-aided analysis [37] predicted a signal peptide (Met^{1}–Phe^{23}). Presence of a recognition site (RYRR) for dibasic processing endoproteases suggests another cleavage between Arg^{45} and Gln^{46}. Thus, the mature protein may consist of 298 amino acids, with the calculated molecular weight and isoelectric point of 29,890 and 4.57, respectively.

The protein sequence consists of tandem repeats

The bulk of the protein sequence consists of tandem repeats of a unit sequence of about 30 amino acids (Figs. 3 and 4). Two dicysteine sequences (Cys-Cys) are found at fixed positions in nine repeats. The most abundant residue in Gln^{46}–Ala^{343} is cysteine (13.4%), followed by small residues such as serine (11.1%), glycine (10.4%), and alanine (9.4%). Fractions of acidic amino acids (aspartate and glutamate) are not high (5.4% and 1.3%, respectively). The galaxin sequence was compared with protein sequences in databases (Protein All ver. 32t09) using the FASTA program [38], without finding a significant match. Proteins with the highest similarity scores included two hypothetical proteins in *Caenorhabditis elegans*, ultra-high sulfur keratin in mouse [39], and lacunin in the insect wing [40]. The similarity between galaxin and these proteins was limited to occurrence of dicysteine repeats. Dicysteine sequences have also been found in proteins isolated from the calcified molluscan exoskeleton such as lustrin A, N14, and N16/




```
B  56: -CGGAPFSTATHICCNGNAEPKTGSTPM---
   83: CCDSNSYDPLSQICCEGTVSHKATSPGAMPA
  114: CCASDGYDMSTQLCCNDNVMHKPTGPTALPG
  145: CCGDHSYDASVQLCCDSNVVPKMGSLSA---
  173: CCGPNSYDTNTTLCCDSNVAFVSGPQAQ---
  201: CCGSQGYDGATQLCCDSNVLPKPGATGA---
  229: CCGSQSYTQDTHLCCEGVIVLKAGPSFA---
  257: CCGSASYNQSSSLCCGATVVAKTPSKPV---
  285: CCGSTSYNPVTEICCDGHVGTRAGLTSPT--
  314: CCGGAVFDAATAKCCDGVPVFNVPSCAGLA-

Consensus sequence
     CCGS-SYD--T-LCCDGNV--K-GS--A---
```

Fig. 4. (A) A schematic representation of the primary structure of galaxin. (B) The tandem repeat structure in galaxin. The numbers on the left represent the amino acid positions and "–" indicates a gap. Below the alignment is shown a deduced consensus sequence in which only the amino acids that are observed in at least five sequences are shown and "–" represents an arbitrary amino acid.

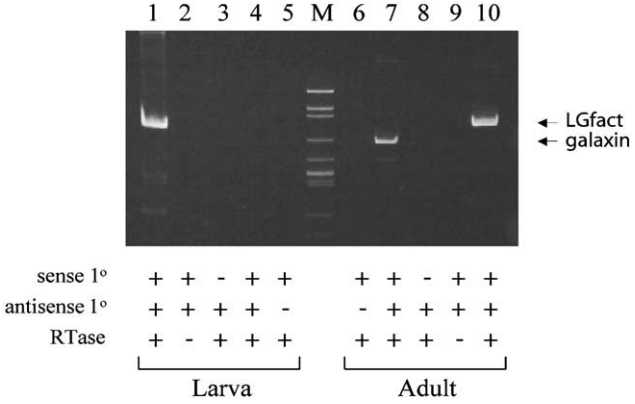


Fig. 5. Expression of *galaxin* in larvae and adult polyps. RT-PCR products were run on a polyacrylamide gel and stained with ethidium bromide. Lanes 1–5, larval cDNA templates; lane M, molecular weight standards; lanes 6–10, adult cDNA templates. Lanes 1 and 10 are products of PCR to detect *LGfact*. *Galaxin* transcripts were detected in adult (lane 7) but not in larvae (lane 4). Comparison of lanes 7 and 9 shows that the PCR product in lane 7 is derived not from genomic DNA but from cDNA.

pearlin [23,25–27], although galaxin did not exhibit significant similarities to these proteins.

Stage specificity of galaxin expression

Larvae (planulae) of scleractinian corals are planktonic and lack skeletal structures: corals initiate skeletogenesis after the settlement and metamorphosis. Thus, we presumed that the expression of *galaxin* was specific to the postmetamorphic stage. To address this issue, RNA samples prepared from larvae and adult polyps were subjected to RT-PCR to detect the expression of *galaxin*. As shown in Fig. 5, a PCR product of the expected size was detected only with the adult RNA, whereas the expression of *LGfact*, a coral actin [35], was observed in both.

Discussion

This paper reports isolation of a protein extracted from the skeletal organic matrix of a scleractinian coral and deduction of its amino acid sequence for the first time. The sequence analysis suggests that galaxin is synthesized as a precursor containing a signal peptide and a propeptide sequence. The prediction for a signal sequence is consistent with the possibility that the protein is secreted by the calicoblastic ectoderm. Although the role of the putative second processing is unknown, we may speculate that the propeptide is necessary for targeting the protein to specialized secretion vesicles, but not for the skeletogenic process after secretion.

The function of galaxin remains unknown, since significant Ca^{2+}-binding was not detected, and its sequence was novel. The galaxin sequence, however, showed two structural features that are observed in some hard tissue proteins: a tandem repeat structure [36,41] and occurrence of dicysteine repeats (Results). The galaxin sequence is rich in Cys residues, most of which (28 out of 30) occur in dicysteine repeats (Fig. 4). Finding of dicysteine sequences in skeletal matrix proteins in two separate phyla (Cnidaria and Mollusca) may suggest a role of these motifs in calcification or interaction with other components of the organic matrix. The abundance of Cys residues in galaxin may raise the possibility that the protein is highly cross-linked to form a macromolecular network. However, an in vitro study using a synthetic peptide containing two dicysteine sequences showed that Cys residues in dicysteine sequences had propensity to form a small intrarepeat cyclocystine loop [42]. Thus, only the two terminal Cys residues (Fig. 4) may be available for intermolecular disulfide bonds.

Amino acid composition of the organic matrix has been determined in a number of scleractinian species [30–33] (*G. fascicularis* was not included in these studies). The content of Asx (aspartate plus asparagine) was high in many species, ranging between 11.5% (*Fungia echinata*) [30] and 59.8% (*Flabellum alabastrum*) [33]. The fraction of Asx in mature galaxin is 9.7% or 11.2% excluding Cys and Trp. The latter value is comparable to the Asx content in some of the examined species such as *F. echinata* and *Pocillopora damicornis* (12.3%) [33].

Proteins with regions that are rich in acidic amino acids have been identified in calcified exoskeletons of invertebrates [19,20,24]. These acidic domains are implicated in regulation of calcification through interaction with Ca^{2+} or calcium carbonate crystals [4,5]. In the present study, no acidic domain was found in the galaxin sequence, and none of the extracted proteins exhibited significant Ca^{2+}-binding. These observations, however, may not necessarily argue that Ca^{2+}-binding is not present in the skeletal matrix of *G. fascicularis*, since the identified proteins may represent only a subset of matrix proteins of the skeleton. Proteins have been extracted from the skeleton of another scleractinian, *Tubastrea aurea*, and subjected to a Ca^{2+}-binding analysis using the same procedure as in the present study. Two CBB-stained bands exhibited Ca^{2+}-binding (K. China, Y. Isa, unpublished observation), raising the possibility that similar Ca^{2+}-binding proteins are present in *G. fascicularis*.

The Ca^{2+}-binding substance in the extract (Fig. 1B) is not likely to be a protein, based on the lack of CBB staining. The corresponding area of a gel was stained yellow with Stains-all (T. Watanabe, unpublished observation), suggesting that a lipid(s) may bind Ca^{2+} [43]. The presence of Ca^{2+}-binding phospholipids has been demonstrated in the exoskeleton of the scleractinian coral, *Acropora hebes* [44]. This observation may lend support to the presence of Ca^{2+}-binding lipids in the exoskeleton of *G. fascicularis*.

Acknowledgments

The work described in this paper was in part supported by a Grant-in-Aid (12NP0201) from the Ministry of Education, Science, Culture and Sports, Japan.

References

[1] K. Simkiss, K.M. Wilbur, Biomineralization: Cell Biology and Mineral Deposition, Academic Press, San Diego, 1989.
[2] R.J.P. Williams, The functional forms of biominerals, in: S. Mann, J. Webb, R.J.P. Williams (Eds.), Biomineralization—Chemical and Biological Perspectives, VCH Press, New York, 1989, pp. 1–61.
[3] Y. Isa, An electron microscope study on the mineralization of the skeleton of the staghorn coral, *Acropora hebes*, Mar. Biol. 93 (1986) 91–101.

[4] L. Addadi, S. Weiner, Stereochemical and structural relations between macromolecules and crystals in biomineralization, in: S. Mann, J. Webb, R.J.P. Williams (Eds.), Biomineralization—Chemical and Biological Perspectives, VCH Press, New York, 1989, pp. 133–156.
[5] A.P. Wheeler, C.S. Sikes, Matrix–crystal interactions in $CaCO_3$ biomineralization, ibid, pp. 95–132.
[6] A.M. Belcher, X.H. Wu, R.J. Christensen, P.K. Hansma, G.D. Stucky, D.E. Morse, Control of crystal phase switching and orientation by soluble mollusk-shell proteins, Nature 381 (1996) 56–58.
[7] G. Falini, S. Albeck, S. Weiner, L. Addadi, Control of aragonite or calcite polymorphism by mollusk shell macromolecules, Science 271 (1996) 67–69.
[8] G.K. Hunter, H.A. Goldberg, Nucleation of hydroxyapatite by bone sialoprotein, Proc. Natl. Acad. Sci. USA 90 (1993) 8562–8565.
[9] G.K. Hunter, P.V. Hauschka, A.R. Poole, L.C. Rosenberg, H.A. Goldberg, Nucleation and inhibition of hydroxyapatite formation by mineralized tissue proteins, Biochem. J. 317 (1996) 59–64.
[10] P. Ducy, C. Desbois, B. Boyce, G. Pinero, B. Story, C. Dunstan, E. Smith, J. Bonadio, S. Goldstein, C. Gundberg, A. Bradley, G. Karsenty, Increased bone formation in osteocalcin-deficient mice, Nature 382 (1996) 448–452.
[11] G.B. Luo, P. Ducy, M.D. McKee, G.J. Pinero, E. Loyer, R.R. Behringer, G. Karsenty, Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein, Nature 386 (1997) 78–81.
[12] S.O. Andersen, Cuticular proteins from the shrimp, *Pandalus borealis*, Comp. Biochem. Physiol. 99B (1991) 453–458.
[13] S.O. Andersen, Exoskeletal proteins from the crab, *Cancer pagurus*, Comp. Biochem. Physiol. 123A (1999) 203–211.
[14] M. Kragh, L. Molbak, S.O. Andersen, Cuticular proteins from the lobster, *Homarus americanus*, Comp. Biochem. Physiol. 118B (1997) 147–154.
[15] S.S. Kumari, J.H. Willis, D.M. Skinner, Proteins of crustacean exoskeleton: IV. Partial amino acid sequences of exoskeletal proteins from the Bermuda land crab, *Gecarcinus lateralis*, and comparisons to certain insect proteins, J. Exp. Zool. 273 (1995) 389–400.
[16] M. Nousiainen, K. Rafn, L. Skou, P. Roepstorff, S.O. Andersen, Characterization of exoskeletal proteins from the American lobster, *Homarus americanus*, Comp. Biochem. Physiol. 119B (1998) 189–199.
[17] T. Watanabe, P. Persson, H. Endo, M. Kono, Molecular analysis of two genes, *DD9A* and *B*, which are expressed during the postmolt stage in the decapod crustacean *Penaeus japonicus*, Comp. Biochem. Physiol. 125B (2000) 127–136.
[18] H. Endo, P. Persson, T. Watanabe, Molecular cloning of the crustacean *DD4* cDNA encoding a Ca^{2+}-binding protein, Biochem. Biophys. Res. Commun. 276 (2000) 286–291.
[19] H. Inoue, N. Ozaki, H. Nagasawa, Purification and structural determination of a phosphorylated peptide with anti-calcification and chitin-binding activities in the exoskeleton of the crayfish, *Procambarus clarkii*, Biosci. Biotechnol. Biochem. 65 (2001) 1840–1848.
[20] H. Endo, P. Persson, T. Ikeya, Y. Takagi, T. Watanabe, Molecular analysis of genes expressed during the calcification of the exoskeleton of the decapod crustacean *Penaeus japonicus*, in: I. Kobayashi (Ed.), Biomineralization, Tokai University Press, Tokyo, 2003 (in press).
[21] H. Miyamoto, T. Miyashita, M. Okushima, S. Nakano, T. Morita, A. Matsushiro, A carbonic anhydrase from the nacreous layer in oyster pearls, Proc. Natl. Acad. Sci. USA 93 (1996) 9657–9660.
[22] S. Sudo, T. Fujikawa, T. Nagakura, T. Ohkubo, K. Sakaguchi, M. Tanaka, K. Nakashima, T. Takahashi, Structures of mollusc shell framework proteins, Nature 387 (1997) 563–564.
[23] X. Shen, A.M. Belcher, P.K. Hansma, G.D. Stucky, D.E. Morse, Molecular cloning and characterization of lustrin A, a matrix protein from shell and pearl nacre of *Haliotis rufescens*, J. Biol. Chem. 272 (1997) 32472–32481.
[24] I. Sarashina, K. Endo, Primary structure of a soluble matrix protein of scallop shell: implications for calcium carbonate biomineralization, Am. Mineral. 83 (1998) 1510–1515.
[25] T. Samata, N. Hayashi, M. Kono, K. Hasegawa, C. Horita, S. Akera, A new matrix protein family related to the nacreous layer formation of *Pinctada fucata*, FEBS Lett. 462 (1999) 225–229.
[26] M. Kono, N. Hayashi, T. Samata, Molecular mechanism of the nacreous layer formation in *Pinctada maxima*, Biochem. Biophys. Res. Commun. 269 (2000) 213–218.
[27] T. Miyashita, R. Takagi, M. Okushima, S. Nakano, H. Miyamoto, E. Nishikawa, A. Matsushiro, Complementary DNA cloning and characterization of pearlin, a new class of matrix protein in the nacreous layer of oyster pearls, Mar. Biotechnol. 2 (2000) 409–418.
[28] F. Marin, P. Corstjens, B. de Gaulejac, E. de Vrind-De Jong, P. Westbroek, Mucins and molluscan calcification—molecular characterization of mucoperlin, a novel mucin-like protein from the nacreous shell layer of the fan mussel *Pinna nobilis* (Bivalvia, Pteriomorphia), J. Biol. Chem. 275 (2000) 20667–20675.
[29] I.M. Weiss, S. Kaufmann, K. Mann, M. Fritz, Purification and characterization of perlucin and perlustrin, two new proteins from the shell of the mollusc *Haliotis laevigata*, Biochem. Biophys. Res. Commun. 267 (2000) 17–21.
[30] S.D. Young, Organic material from scleractinian coral skeletons—I. Variation in composition between several species, Comp. Biochem. Physiol. 40B (1971) 113–120.
[31] R.M. Mitterer, Amino acid composition and metal binding capability of the skeletal protein of corals, Bull. Mar. Sci. 28 (1978) 173–180.
[32] B. Constanz, S. Weiner, Acidic macromolecules associated with the mineral phase of scleractinian coral skeletons, J. Exp. Zool. 248 (1988) 253–258.
[33] J.P. Cuif, Y. Dauphin, A. Freiwald, P. Gautret, H. Zibrowius, Biochemical markers of zooxanthellae symbiosis in soluble matrices of skeleton of 24 *Scleractinia* species, Comp. Biochem. Physiol. 123A (1999) 269–278.
[34] K. Maruyama, T. Mikawa, S. Ebashi, Detection of calcium-binding proteins by ^{45}Ca autoradiography on nitrocellulose membrane after sodium dodecyl–sulfate gel-electrophoresis, J. Biochem. 95 (1984) 511–519.
[35] I. Fukuda, S. Imagawa, K. Iwao, T. Horiguchi, T. Watanabe, Isolation of actin-encoding cDNAs from symbiotic corals, DNA Res. 9 (2002) 217–223.
[36] T. Ikeya, P. Persson, M. Kono, T. Watanabe, The *DD5* gene of the decapod crustacean *Penaeus japonicus* encodes a putative exoskeletal protein with a novel tandem repeat structure, Comp. Biochem. Physiol. 128B (2001) 379–388.
[37] H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne, Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites, Protein Eng. 10 (1997) 1–6.
[38] W.R. Pearson, D.J. Lipman, Improved tools for biological sequence comparison, Proc. Natl. Acad. Sci. USA 85 (1988) 2444–2448.
[39] L. Wood, M. Mills, N. Hatzenbuhler, G. Vogeli, Serine-rich ultra high sulfur protein gene expression in murine hair and skin during the hair cycle, J. Biol. Chem. 265 (1990) 21375–21380.
[40] J.B. Nardi, R. Martos, K.K. Walden, D.J. Lampe, H.M. Robertson, Expression of lacunin, a large multidomain extracellular matrix protein, accompanies morphogenesis of epithelial

monolayers in *Manduca sexta*, Insect Biochem. Mol. Biol. 29 (1999) 883–897.
[41] S.O. Anderson, P. Hojrup, P. Roepstorff, Insect cuticular proteins, Insect Biochem. Mol. Biol. 25 (1995) 153–176.
[42] M.I. Liff, M.N. Zimmerman, NMR study of crosslinking by oxidation of four-cysteine polypeptide models of the elastic network phase of wool fibre, Polym. Int. 47 (1998) 375–385.
[43] K.P. Campbell, D.H. MacLennan, A.O. Jorgensen, Staining of the Ca^{2+}-binding proteins, calsequestrin, calmodulin, troponin C, and S-100, with the cationic carbocyanite dye "Stains-all", J. Biol. Chem. 258 (1983) 11267–11273.
[44] Y. Isa, M. Okazaki, Some observations on the Ca^{2+}-binding phospholipid from scleractinian coral skeletons, Comp. Biochem. Physiol. 87B (1987) 507–512.