

Skeletal matrix proteins of invertebrate animals: Comparative analysis of their amino acid sequences

Authors: SARASHINA, ISAO, and ENDO, KAZUYOSHI

Source: Paleontological Research, 10(4): 311-336

Published By: The Palaeontological Society of Japan

URL: https://doi.org/10.2517/prpsj.10.311

The BioOne Digital Library (<u>https://bioone.org/</u>) provides worldwide distribution for more than 580 journals and eBooks from BioOne's community of over 150 nonprofit societies, research institutions, and university presses in the biological, ecological, and environmental sciences. The BioOne Digital Library encompasses the flagship aggregation BioOne Complete (<u>https://bioone.org/subscribe</u>), the BioOne Complete Archive (<u>https://bioone.org/archive</u>), and the BioOne eBooks program offerings ESA eBook Collection (<u>https://bioone.org/esa-ebooks</u>) and CSIRO Publishing BioSelect Collection (<u>https://bioone.org/csiro-ebooks</u>).

Your use of this PDF, the BioOne Digital Library, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at <u>www.bioone.org/terms-of-use</u>.

Usage of BioOne Digital Library content is strictly limited to personal, educational, and non-commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne is an innovative nonprofit that sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Skeletal matrix proteins of invertebrate animals: Comparative analysis of their amino acid sequences

ISAO SARASHINA AND KAZUYOSHI ENDO

Doctoral Program in Earth Evolution Sciences, Graduate School of Life and Environmental Science, University of Tsukuba, Tsukuba, Japan (sarasina@arsia.geo.tsukuba.ac.jp, endo@arsia.geo.tsukuba.ac.jp)

Received July 10, 2006; Revised manuscript accepted September 21, 2006

Abstract. The genetic bases of skeletogenesis are expected to shed light on the origins of metazoan biomineralization. Here we review aspects of genetic machineries of invertebrate skeletogenesis, including regulatory genes involved in biomineralization as well, and with an enumerative reference to the genes encoding skeletal matrix proteins. The complete primary structure has been determined for a total of 77 skeletal matrix proteins in invertebrates representing five animal phyla. Presence of repeated sequences and prevalence of acidic proteins stand as common features among those proteins. Similarities are interpreted as convergence because these proteins are not similar at the primary structure level. C-type lectin-like domains are shared by the calcium carbonate skeletal matrix proteins of molluscs and deuterostomes. However, the important sites for carbohydrate binding are not conserved between these two groups. Several arthropod skeletal matrix proteins have the Rebers-Riddiford consensus sequence which is characteristic of noncalcified cuticular proteins of arthropods, indicating that these skeletal matrix proteins were recruited from the non-calcified cuticular proteins after arthropods diverged from other metazoan groups. Dermatopontin, a molluscan shell matrix protein, is also inferred to represent a cooption for biomineralization after molluscs diverged from other metazoan groups based on the molecular phylogemetic analysis. Those findings support the premise that the genetic machineries of biomineralization evolved independently many times after the divergence of metazoan phyla, and that some common genes that served for other functions have been coopted for biomineralization in various lineages.

Key words: acidic proteins, biomineralization, C-type lectin, repeated sequences, cooption, skeletal matrix proteins

Introduction

After the appearance of organisms which produced distinctly larger trace fossils toward the end of the Neoproterozoic and before the emergence of the relatively abundant fauna producing small shelly fossils that marks the beginning of the Cambrian explosion, metazoans variously adapted to a range of distinctive habitats and evolved to exhibit many disparate bodyplans (Conway Morris, 1998; Valentine, 2004). The occurrence of massive biominerals in metazoans may be one of the many novelties of their evolution which appeared in that period. Nevertheless, the apparent absence of biominerals in the Ediacaran fauna recognized at the close of the Neoproterozoic and nearly simultaneous occurrence of biominerals in cyanobacteria (in spite of some Proterozoic cyanobacteria having had mineralized sheaths), algae, foraminifers and radiolarians (Runnegar and Bengtson, 1990) seem to call for specific explanations for metazoan

biomineralization, and many researchers have tried to explain it. The explanations chiefly concern physicochemical environmental factors (e.g., O₂, CO₂, Ca, P, temperature), ecological relationships (e.g., predation, primary production) and other adaptive advantages (e.g., biomechanics, detoxification) (Simkiss, 1977; Vermeij, 1989; Runnegar and Bengtson, 1990; Bengtson, 1994; Nedin, 1999; Valentine, 2004; Cohen, 2005). While it is logical for us to search for triggers that brought about the simultaneous occurrences of biominerals, the mechanisms that enabled biomineral formation, or the underlying machineries of biomineralization, including its genetic bases, would also be important and must shed light on this problem.

So far as metazoans, our knowledge of biomineralization mechanisms in invertebrates has been limited compared to that in vertebrates. Thanks to the propagation of molecular biological techniques, however, data from invertebrates have been accumulated in recent years. It is thus probably worthwhile to consider the origins of invertebrate biominerals through the aspect of mechanisms at the molecular level. Initially we refer to the finding that implanted molluscan nacres induce osteogenesis in vertebrates. Then we review the gene regulatory networks involved in biomineralization, including vertebrates for comparison. Finally, we review matrix proteins in invertebrate skeletons which have been thought to play very important roles in biomineralization. We enumerate the invertebrate skeletal matrix proteins of which the complete primary structure has been clarified to date, and discuss the underlying mechanisms and the origins of invertebrate biomineralizations, referring also to osteogenic activity of molluscan nacre and the gene regulatory networks.

Abbreviations used in this article

BMP: bone morphogenetic proteins, cDNA: complementary DNA, CRD: carbohydrate recognition domain, CTLD: C-type lectin-like domains, ELISA: enzyme linked immunosorbent assay, EST: expressed sequence tag, IGFBP: insulin-like growth factor binding protein, PMC: primary mesenchyme cell, SDS-PAGE: sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Three-letter and one-letter codes of the common amino acids

Ala (A): alanine, Arg (R): arginine, Asn (N): asparagines, Asp (D): aspartic acid, Cys (C): cysteine, Gln (Q): glutamine, Glu (E): glutamic acid, Gly (G): glycine, His (H): histidine, Ile (I): isoleucine, Leu (L): leucine, Lys (K): lysine, Met (M): methionine, Phe (F): phenylalanine, Pro (P): proline, Ser (S): serine, Thr (T): threonine, Trp (W): tryptophan, Tyr (Y): tyrosine, Val (V): valine.

Osteogenic activity of molluscan nacre

Bobbio (1972) exhumed Mayan skulls in Honduras. He discovered that the ancient Mayans used nacre as a dental implant, which showed perfect osteointegration. Thus Bobbio rediscovered an osteogenic property of nacre. Atlan *et al.* (1997) took nacre, mother-of-pearl, from shells of the giant oyster *Pinctada maxima*, and implanted powdered-nacre into human bones, aiming at reducing bone loss. The implanted nacre was accepted showing no immunological response. Moreover, osteoblasts, bone-forming cells, were activated, and new bone was formed. The osteogenic activity of nacre obtained from *P. maxima* has also been shown by *in vivo* experiments in rats (Liao *et al.*, 1997), in sheep (Delattre *et al.*, 1997; Lamghari *et al.*, 1999, 2001b; Milet *et al.*, 2004) and in rabbits (Lamghari *et al.*, 2001a). When Dupoirieux *et al.* (1994) implanted coral skeletons into rat bones, half were accepted, but a complete bone repair was not achieved. This observation suggests that coral skeletons have weaker osteogenic activity than molluscan nacre. This may reflect evolutionary relationships among them, because mammals are evolutionarily more distantly related to corals than to molluscs (Westbroek and Marin, 1998). *In vitro* experiments using nacre chips of *P. maxima* and human osteoblasts suggested osteogenic activity of nacre. Induction of mineralization occurred preferentially in the osteoblasts attached to nacre chips (Lopez *et al.*, 1992; Silve *et al.*, 1992).

In order to confirm the hypothesis that the nacre contains diffusible signal molecules which are able to stimulate mammalian cells involved in bone formation, in vitro experiments using water-soluble matrix molecules from P. maxima nacre and mammalian cell cultures have been performed. These experiments showed that the water-soluble matrix is also effective in inducing differentiation or activation of rat osteoblasts (Mouriès et al., 2002), mouse pre-osteoblasts (Milet et al., 2004), human fibroblasts, the cells inducing bone formation (Almeida et al., 2000, 2001; Mouriès et al., 2002; Milet et al., 2004), and rat bone marrow stromal cells, a pool of bone precursor cells (Lamghari et al., 1999; Mouriès et al., 2002; Milet et al., 2004). The water-soluble matrix from P. maxima nacre was also revealed to increase the amount of Bcl-2 in rat osteoblasts, a protein which inhibits programmed cell death (Korsmeyer, 1992), consequently prolonging the life of mature osteoblasts (Mouriès et al., 2002). These observations in all implied that, although the bone and the nacre from P. maxima are not homologous morphologically, some parts of the complex machinery involved in their formation may be homologous (Westbroek and Marin, 1998).

The nacre from the Swedish freshwater pearl mussel *Margaritifera* sp. also promoted rat bone formation *in vivo*. In this case, however, no active osteoblasts were detected in direct contact with nacre at early stages. Therefore, the osteogenic activity of *Margaritifera* nacre could not be confirmed (Liao *et al.*, 2000, 2002).

Gene regulatory networks involved in metazoan biomineralization

Vertebrates

Gene regulatory mechanisms involved in biomineralization have been best studied in vertebrates. Osteoblast differentiation and bone formation are known to be regulated by many local factors. Among these, bone morphogenetic proteins (BMPs) are one of the most potent factors. The bone-inducing ability of BMPs was first reported in the 1960s (Urist, 1965; Urist and Strates, 1971), but it was very difficult to isolate BMPs as a single protein from the decalcified bone matrix until the late 1980s. Molecular cloning of human BMP-2 and -4, and the purification and sequencing of bovine BMP-3 were performed in the late 1980s (Wozney et al., 1988; Luyten et al., 1989; Wozney, 1992). Consequently it has been clarified that BMPs are signal molecules belonging to the transforming growth factor β (TGF- β) superfamily, and play critical roles in bone induction. To date, around 20 BMP members have been identified (Shen et al., 2004) and BMP-3 was revealed to be an antagonist to other osteogenic BMPs (Daluiski et al., 2001). Three signal molecules, Smad1 (Hoodless et al., 1996), 5 (Nishimura et al., 1998) and 8 (Chen et al., 1997) are revealed to be the immediate downstream molecules of BMP receptors and play a central role in BMP signal transduction (Chen et al., 2004), and other signal molecules, Sonic and Indian hedgehogs are involved in osteoblast differentiation by interacting with BMPs (Yamaguchi et al., 2000). Some transcription factors such as core-binding factor a-1 (Cbfa1), also known as Aml3 or Runx2, Osterix (Osx), which is a downstream molecule of Cbfa1, and the Activator protein-1 (AP-1) family including Fra-1 and Δ FosB are also known to play a major role in the differentiation and maturation of osteoblasts. Cbfa1 or Osx null mice showed a complete lack of ossification (Komori et al., 1997; Nakashima et al., 2002), and overexpression of Fra-1 and ΔFosB increased bone mass (Jochum et al., 2000; Sabatakos et al., 2000). The three transcription factors having a paired-type homeodomain, Cart1 (Zao et al., 1994), Alx3 (ten Berge et al., 1998), and Alx4 (Qu et al., 1997), regulate skeletons of the face, neck and limbs (Beverdam and Meijlink, 2001). Although Alx3 single mutant mice did not show obvious abnormalities, Alx3/Alx4 double mutant mice showed more severe craniofacial abnormalities than Alx4 single mutant mice (Beverdam et al., 2001). Cart1/Alx4 double mutant mice exhibited severe craniofacial and limb abnormalities (Qu et al., 1999). In humans, mutations in Alx4 caused skull ossification defects (Wu et al., 2000; Mavrogiannis et al., 2001). Furthermore, two signaling molecule, parathyroid hormone related peptide (PTHrP) and fibroblast growth factors (FGFs) are known to be critically involved in endochondrial bone formation, and three transcription factors of the Sox family, Sox9, L-Sox5 and Sox6 have essential roles in

chondrocyte differentiation (de Crombrugghe et al., 2001).

Invertebrates

As to gene regulatory mechanisms involved in biomineralization in invertebrates, several components of the gene networks and their regulatory relationships have been revealed to some extent. The identified components or the component candidates of the gene networks were transcription factors such as engrailed (Moshel et al., 1998; Jacobs et al., 2000; Wanninger and Haszprunar, 2001; Nederbragt et al., 2002), Hox1, Hox4 (Hinman et al., 2003), and signal molecules such as decapentaplegic (Nederbragt et al., 2002) in molluscs, and transcription factors such as engrailed (Lowe and Wray, 1997), ets1 (Kurokawa et al., 1999), Alx1 (Ettensohn et al., 2003), and the possibly transmembrane protein P16 (Cheers et al., 2005) in echinoderms. SM50, one of the skeletal matrix protein genes of sea urchins, is located directly downstream of ets1 (Kurokawa et al., 1999). P16 is downstream of Alx1 (Cheers et al., 2005) in the primary mesenchyme cell (PMC) gene regulatory network in sea urchins (Cheers et al., 2005). In the micromere PMC gene regulatory networks in sea urchins, genes upstream of Alx1 are Repressor X and Pmar1 (Ettensohn et al., 2003).

The gene regulatory networks in some metazoan phyla have the similar components in common, but it may not necessarily support evolutionary links of the biomineralization systems between them. engrailed (en) is considered to be at least indirectly involved in skeletogenesis in both molluscs and echinoderms. en is expressed at the boundary around the embryonic shell of molluscs (Nederbragt et al., 2002) and at the boundaries between newly forming skeletal ossicles in brittle stars (Lowe and Wray, 1997). en expression patterns in brittle stars are not a common feature among echinoderms and thus en may have been recruited for skeletogenic function within echinoderms (Lowe and Wray, 1997) independently from molluscs. decapentaplegic (dpp) is an ortholog of BMP2/4 which is also involved in osteogenesis in vertebrates. dpp-BMP2/4 was discovered from the mantle of the mollusc Pinctada fucata (Matsushiro and Miyashita, 2004). In the case of another mollusc Patella vulgata, however, Dpp-BMP2/4 is probably employed for setting up the boundary around shells rather than inducing shell formation (Nederbragt et al., 2002).

Members of the Cart1/Alx3/Alx4 subfamily are involved in skeletogenesis in both echinoderms (Alx1) and vertebrates (Cart1, Alx3, Alx4), possibly indicating the common origin of a part of their biomineralization systems. Most of the components, however, are different between echinoderms and vertebrates, and common features cannot be found between the regulatory relationships of the genes involved in their biomineralization. These arguments seem to harmonize with the independent origins of biomineralization in many metazoan phyla, and we could perhaps explain the common employment of Alx subfamily members in echinoderms and vertebrates as independent gain-offunction events. Except for vertebrates and echinoderms, however, there is very little information on the regulatory gene cascades involved in metazoan biomineralization, and it is premature to draw definitive conclusions.

Skeletal matrix proteins in invertebrates

Porifera

Silica is an amorphous form of hydrated SiO₂, an abundant element in the earth's crust. There are three classes of sponges, two of which, the Hexactinellida and the Demospongiae, produce silicified spicules, while the other one, the Calcarea, produces calcium carbonate spicules. Incidentally, the Calcarea is considered closer not to either or both the Hexactinellida and the Demospongiae but to the other metazoan phyla (Kruse *et al.*, 1998). These spicules, supporting the organisms and defending them against predation, are produced in membrane-enclosed vesicles in specialized cells, sclerocytes (Perry, 2003). After reaching a crucial size, the spicules are extruded from the cells and their growth proceeds in the extracellular space (Müller *et al.*, 2005).

Shimizu et al. (1998) isolated three very similar proteins of 27, 28 and 29 kDa in size from the spicules of Tethya aurantia belonging to the Demospongiae. They dubbed the three proteins, Silicate α , Silicate β and Silicate γ , respectively, and the complete primary structure of Silicate α (and Silicate β later on) was deduced from the full-length cDNA. Silicatein α is rich in serine, glycine, alanine and tyrosine, and has a high degree of sequence similarity to cathepsin L, a papain-like cysteine protease. The catalytic cysteine residues, however, at the active site of the cystein protease is replaced by serine in Silicatein α , and Silicateins did not display esterase activity when tested with synthetic chromogenic substrates (Shimizu et al. 1998). Recombinant Silicatein α , however, catalyzed the polymerization of silica at moderate temperature and pH in vitro (Cha et al., 1999), suggesting that Silicate α has an enzymatic function.

The axial filaments in the spicules of the common Mediterranean sponge *Petrosia ficiformis* (Demospongiae) contained two water-insoluble proteins of 30 and 23 kDa (Pozzolini et al., 2004). They showed this 23 kDa protein to be a homolog of Silicateins (Funayama et al. (2005) called it "Pf Silicatein"). Another homolog of Silicateins (Ef Silicatein) was also identified from the EST library of a freshwater sponge Ephydatia fluviatilis belonging to Demospongiae (Funayama et al., 2005). Recently a cluster of four genes consisting of ankyrin repeat gene, silicate in β (encoding SILICAb_SUBDO), tumor necrosis factor receptor-associated factor and a protein kinase was isolated from the S. domuncula genome, and the four genes including *silicatein* β were revealed to be expressed synchronously (Schröder et al., 2005). This observation suggests that a coordinated expression of physically linked genes is essential for the synthesis of the spicules.

A silicase identified from *S. domuncula* is known to be a silica-catabolizing enzyme (Schröder *et al.*, 2003), which shares the highest sequence similarity to the family of carbonic anhydrases. Silicase is able to degrade silica to form free silicic acid, and possibly contribute to the metabolism of silica deposition.

While Krasko *et al.* (2000) found that silicates upregulate SILICAa_SUBDO which is an *S. domuncula* ortholog of Silicatein- α , Müller *et al.* (2005) showed that selenium, a trace element essential for metazoans, up-regulates two other genes, and increased spicule formation using an *in vitro* cell culture system of *S. domuncula* belonging to the Demospongiae.

One of these two genes showed the closest similarity to the human selenoprotein M. The biological role of selenoprotein M in higher metazoan phyla is not known in detail, but it has previously been shown that selenocysteins, a 21st amino acid in ribosomemediated protein synthesis (Stadtman, 1996), in selenoproteins participates in redox reactions, together with a close cystein partner (Hatfield and Gladyshev, 2002). This constellation of selenocystein and cystein residues also exist in S. domuncula selenoprotein M, suggesting that this protein functions as an enzyme (Müller et al., 2005). Another up-regulated protein termed spicule-associated protein showed no significant sequence similarity to any protein in the database. This protein has six highly similar repeats comprising 20 amino acid residues, in which hydrophobic amino acids are surrounded by polar amino acids, expected to form a tight association with membranes. The immunohistological data showed that selenoprotein M, spicule-associated protein and Silicatein are all present in the axial filaments and at the surface of the spicules, suggesting that they are associated with spicule formation (Müller et al., 2005).

Collagen is the major structural protein in general. In sponges, spicules are embedded in an organic matrix containing collagen (Müller *et al.*, 2005). Some collagen-like proteins are identified in sponges (Exposito *et al.*, 1991; Krasko *et al.*, 2000), and one of them, COL1_SUBDO identified from *S. domuncula* (Schröder *et al.*, 2000), was revealed to be upregulated by silicates as well as by Silicateins. Thus COL1_SUBDO is likely to be involved in the forma-

Cnidaria

Reef corals belonging to the Scleractinia accrete hard skeletons leading to biomineralization products of spectacular size, while soft corals belonging to the Octocorallia and the Antipatheria have spicules supporting the organisms (Cohen and McConnaughey, 2003). Scleractinian skeletons are nearly exclusively composed of aragonite, and spicules of octocorals are composed of calcite (Lowenstam and Weiner, 1989). The spicules contain a proteinaceous axial filament.

tion of the collagenous sheath required for the func-

tional skeletons (Krasko et al., 2000).

Galaxin was identified from the soluble organic matrix in the exoskeleton of the scleractinian coral Galaxea fascicularis, which is the only coral skeletal matrix protein of which the complete primary structure has been elucidated (Fukuda et al., 2003). Galaxin transcripts were detected in the adult coral, but not in planktonic larvae. Galaxin has two potential N-linked glycosylation sites and a tandem repeat structure where a sequence of 27-31 amino acid residues repeats ten times. Each repeat unit contains a dicystein sequence. An in vitro study using a synthetic peptide suggests that Cys forms an intrarepeat cyclocystine loop (Liff and Zimmerman, 1998) and two terminal Cys may contribute to cross-linking to other molecules. The abundance of cystein residues raises the possibility that the protein is highly cross-linked to form a macromolecular network. Galaxin has no acidic domain and did not exhibit a significant Ca²⁺binding ability when examined using the Ca²⁺-overlay analysis (Maruyama et al., 1984; Fukuda et al., 2003). These observations, however, do not indicate that Ca²⁺-binding protein is not important nor present in the coral biomineralization systems (Isa and Okazaki, 1987). Partial amino acid sequences containing many aspartic acids have been reported from both exoskeletons of a scleractinian coral (Puverel et al., 2005) and spicules of an octocorallian coral (Rahman and Isa, 2005). Ca²⁺-binding proteins were also detected from hard tissues of both scleractinian and octocorallian corals (Watanabe et al., 2003; Rahman and Isa, 2005). On the other hand, Ca²⁺-binding phospholipids were

detected in scleractinian corals, and they might be important for scleractinian biomineralization, such as serving as seeding sites for calcium carbonate depositions (Isa and Okazaki, 1987).

Arthropoda

Most crustaceans possess a calcified outer skeleton, the carapace, which they replace completely and periodically with moulting cycles to permit growth. Thus they need a large amount of calcium during the premoult period. They have developed several calcium storage strategies, and during the period they absorb calcium from diverse storage structures such as gastroliths or midgut caeca as well as from food or water (Lowenstam and Weiner, 1989; Testenière et al., 2002). In crayfish, exoskeleton and gastrolith consist of calcite and amorphous calcium carbonate, respectively (Takagi et al., 2000). In general, exoskeletons of crustaceans consist of calcium carbonate under a crystalline state and/or an amorphous form, whereas calcium storage structures are always in an amorphous form (Luquet and Marin, 2004).

Andersen and coworkers extracted a soluble matrix protein from the calcified exoskeleton of the northern red shrimp Pandalus borealis and determined its complete amino acid sequence by combined mass spectrometry and Edman degradation (Jacobsen et al., 1994). This is the first report of the complete primary structure of a matrix protein identified from crustacean hard tissues. This protein, H1a, is rich in hydrophobic amino acids such as proline, alanine and valine (Jacobsen et al., 1994). On the other hand, Suzuki et al. (2001) extracted an insoluble matrix protein, Pb CP-12.7, from the same origin, the calcified exoskeleton of Pandalus borealis, and its complete primary structure was determined by direct sequencing. The sequence is rich in hydrophobic amino acids, and includes a repeated motif, X-A-G-X-X-P-Y. This protein is insoluble in water but soluble in methanol, and shows chitin-binding ability using chromatography with a chitin-packed column. Because P. borealis is a small shrimp and a batch of 50-100 shrimps was needed to extract proteins, they then studied bigger decapods, the American lobster Homarus americanus (Kragh et al., 1997; Nousiainen et al., 1998) and the rock crab Cancer pagurus (Andersen, 1999). They extracted 13 matrix proteins from the calcified exoskeleton of H. americanus; these proteins are numbered according to their molecular mass in kilodaltons and the number is preceded by the letters HaCP (H. americanus cuticular protein). Similarly, the 12 matrix proteins identified from the calcified exoskeleton of C. pagurus are indicated by CpCP (C. pagurus cuticular

protein), and the five matrix proteins from the arthrodial membranes, flexible regions of the exoskeleton, are indicated by CpAMP (C. pagurus arthrodial membrane protein). A common 18 amino acid motif contains three highly conserved glycine residues shared by HaCP11.6, CpCP12.46, CpCP12.43, CpCP11.58 (these four proteins contain four copies of the motif), HaCP6.3, HaCP5.9, HaCP5.6, HaCP4.6, HaCP4.5, HaCP4.4, CpCP4.98, CpCP4.66, CpCP4.63, CpCP4.59 and CpCP4.34 (these 11 proteins contain two copies). The fact that this 18 amino acid motif is found in 15 out of the 25 proteins derived from the calcified exoskeletons examined, and that this motif is not found in any other proteins derived from the arthrodial membranes examined nor in any of the cuticular proteins of insects suggest that it is involved in the calcification process (Andersen, 1999). The Rebers-Riddiford consensus sequence, which is a chitinbinding domain and found in many cuticular proteins of insects and crustaceans, is present in CpCP11.14, and a truncated version of the consensus sequence is found in HaCP18.8 and CpCP5.75 (Nousiainen et al., 1998; Andersen, 1999). The fact that the consensus sequence is found only in three out of the 25 proteins in the calcified exoskeleton examined and that the consensus sequence is mainly found in proteins derived from flexible cuticles (Andersen, 1999) suggests that the consensus sequence is not likely to be involved in the calcification process.

Orchestin is a water-soluble non-glycosylated acidic matrix protein extracted from calcium carbonate concretions in the midgut caeca of the land crustacean Orchestia cavimana (Testenière et al., 2002). The amino acid sequences are rich in acidic residues such as aspartic acid (16.7%) and glutamic acid (13.0%), and several tandem and periodic repeats such as E-S-R/E-E-E-P-R-K-L, D-D-S-R-E, S-D-E-S, S-D-E, S-R-E, S-D and D-S are recognized (Testenière et al., 2002). Orchestin was shown to have a calcium-binding ability using the Ca²⁺-overlay analysis. The Ca²⁺-binding occurred only via the phosphoserine residues, although tyrosine residues are also phosphorylated (Hecker et al., 2003). Both the transcripts and translation products are expressed in the calcium storage organ during the premoult periods (Testenière et al., 2002). The protein is synthesized also during the postmoult period as a component of the organic matrix of the calcium spherules, by which calcium resorbed from the storage organ is translocated (Hecker et al., 2004). Thus Orchestin is probably a key molecule in two transitory calcified mineralization processes in the calcium storage organ and the calcium spherules.

GAMP (Gastrolith Matrix Protein) is an insoluble

matrix protein extracted from gastroliths of the crayfish Procambarus clarkii (Ishii et al., 1996, 1998). The amino acid sequence is rich in alanine and glycine residues and has no potential N-linked glycosylation site. The sequence includes a repeated motif, Q-Q-A-A-P-A/T, and shows some sequence similarity to involucrin, a human keratinocyte protein, which is crosslinked to the membrane proteins (Tsutsui et al., 1999). GAMP was revealed to inhibit calcium carbonate precipitation in *in vitro* experiments, by monitoring the pH values of supersaturated calcium carbonate solutions, but Ca²⁺-overlay analysis indicated that GAMP has almost no Ca²⁺-binding ability. The transcripts are expressed in the gastrolith disc, in which a pair of gastroliths is formed, during the premoult period (Tsutsui et al., 1999). The immunolocalization signals of the translation products are also observed in the gastrolith disc as well as in the gastroliths and the exocuticle in the exoskeleton. The distribution of GAMP immunoreactivity roughly corresponds to that of calcium carbonate (Takagi et al., 2000).

Calcification-associated peptide-1 (CAP-1) and CAP-2, a similar protein to CAP-1, were isolated from the exoskeleton of the crayfish Procambarus clarkii, the same species from which GAMP had been identified (Inoue et al., 2001, 2003, 2004). In CAP-1, one of the serine residues was identified to be phosphorylated (Inoue et al., 2001). Both CAP-1 and CAP-2 have the Rebers-Riddiford consensus sequence, a chitin-binding domain, and actually exhibited chitinbinding ability when examined using the method of Folders *et al.* (2000). Both proteins also showed inhibitory activity on calcium carbonate precipitation in an in vitro dose-dependent anti-calcification assay, in which the formation of calcium carbonate precipitates was monitored by the turbidity of the solution. The transcripts of the two proteins are both expressed in the epidermal tissues during the postmoult period, where and when the calcification take place. Taking these results together, the authors surmised that CAP-1 and CAP-2 might serve as a connecting peptide between chitin and calcium carbonate in the cuticle and bind to the surface of calcium carbonate crystals so as to function as a nucleator or a regulator of calcification (Inoue et al., 2001, 2003, 2004).

In the prawn *Penaeus japonicus*, four cDNA species were identified using the differential display technique: those transcripts are expressed in the epidermal cells underlying the exoskeleton specifically during the postmoulting periods. All the four cDNA species, *DD4*, *DD5*, *DD9A*, *DD9B*, were sequenced to reveal the deduced amino acid sequences (Endo et al., 2000; Watanabe et al., 2000; Ikeya et al., 2001). The amino

acid sequence of DD4, which was renamed Crustocalcin (Endo et al., 2004), contained a glutamic acid-rich region and the Rebers-Riddiford consensus sequence, and exhibited sequence similarity to the Drosophila Ca²⁺-binding protein Calphotin. Ca²⁺-overlay analysis indicated clearly that Crustocalcin/DD4 has a Ca²⁺-binding ability (Endo et al., 2000) and partial segments of Crustocalcin/DD4 induced calcium carbonate crystallization in vitro (Endo et al., 2004). The amino acid sequence of DD5 consists of tandem repeats of a 93-98 amino acid sequence that is repeated 13 times. This repeated motif contains the Rebers-Riddiford consensus sequence, possibly contributing to cross-linking of chitin fibers. DD5 shows a sequence similarity to HaCP18.8, an exoskeletal protein of the American lobster H. americanus (Ikeya et al., 2001). Based on the spatio-temporal expression patterns of the four transcripts and the sequence similarities to crustacean and insect cuticular proteins exhibited by the four proteins, the authors interpreted that all the four proteins, Crustocalcin/DD4, DD5, DD9A and DD9B are the protein components of the exoskeleton, suggesting that they play a role in the calcification of the crustacean exoskeleton (Endo et al., 2000; Watanabe et al., 2000; Ikeya et al., 2001). However the authors also pointed out that two of them, DD9A and DD9B, may not be involved in calcification (Watanabe et al., 2000), because they exhibit sequence similarities with proteins extracted from arthrodial membranes, uncalcified flexible cuticles, of H. americanus (Andersen, 1998) and C. pagurus (Andersen, 1998).

Two genes, CsCP (*Callinectes sapidus* cuticular protein)8.2 and CsCP8.5, were identified from the EST database of the blue crab *Callinectes sapidus*, derived from the mRNA differentially expressed in the hypodermis. The two proteins encoded by these genes, containing partial Rebers-Riddiford consensus sequences, may be homologous to CAP-1, CAP-2, DD5 (Wynn and Shafer, 2005) and possibly also to HaCP18.8.

Mollusca

Most molluscan species have shells, one of the most dominant biominerals after corals in the metazoans. Molluscan shells consist of aragonite or a composite of aragonite and calcite, two of the polymorphs of calcium carbonate. The shells have been thought of as remarkable examples of a matrix-mediated biomineralization (Watabe and Wilbur, 1960; Lowenstam, 1981; Lowenstam and Weiner, 1989; Belcher *et al.*, 1996; Falini *et al.*, 1996), and unusually acidic molecules in the shell matrix have been predicted to play important roles in molluscan biomineralization (Hare, 1963; Weiner and Hood, 1975; Mitterer, 1978; Weiner, 1979, 1983; Runnegar, 1984; Weiner and Addadi, 1991; Wheeler, 1992; Albeck *et al.*, 1993; Gotliv *et al.*, 2003).

Chronologically speaking, the fifth protein identified from molluscan shells, MSP-1, is the first unusually acidic protein to have been sequenced (Sarashina and Endo, 1998; 2001). MSP-1 is a soluble shell matrix glycoprotein extracted from the calcitic foliated layer of the scallop *Patinopecten yessoensis*. The protein is rich in serine, glycine and aspartic acid. MSP-1 has a characteristic modular structure containing a basic domain close to the N-terminus and a large sequence unit repeated four times; each unit comprises three domains of SG, D and K domains. The basic and SG domains of MSP-1 are somewhat similar to the basic and the glycine- and serine-rich domains in Lustrin A (Shen et al., 1997) as described below. The K domains in MSP-1 have putative basic cleavage motifs (Barr, 1991). The D domain, rich in aspartic acid and containing Asp-Gly-Ser-Asp and Asp-Ser-Asp motifs, is supposed to represent Ca^{2+} -binding sites. Borbas *et al.* (1991) reported that phosphate groups were found in organic matrix in calcitic foliated shell structures in higher amounts than in other structures. MSP-1 contain by far the largest number of putative phosphorylation sites (247 sites) of all the skeletal matrix proteins of metazoans, suggesting that MSP-1 is heavily phosphorylated and more acidic than theoretically predicted from the deduced amino acid sequence alone (pI = 3.2). Because many potential phosphorylation sites and N-linked glycosylation sites are found in the D domains, these post-translational modifications are probably involved in the binding of calcium ions. MSP-1 shows sequence similarity to phosphophoryn, and the SG domain of MSP-1 shows similarity to the glycine- and serine-rich domains in Lustrin A (Shen et al., 1997). Both the similarities, however, may result from the biased amino acid compositions. Thus the evolutionary relationships among these proteins remain obscure. The gene encoding a second unusually acidic shell matrix protein, Aspein, was identified from cDNA constructed from the mantle of P. fucata (Tsukamoto et al., 2004). Aspein is more acidic than MSP-1, containing a high proportion of Asp (60.4%) and the predicted isoelectric point is 1.45; this is probably the most acidic of all known proteins. Aspein has a relatively neutral N-terminal region followed by an unusually acidic region containing 58 polyaspartate blocks ranging from 2 to 10 residues in size punctuated by Ser-Gly dipeptides. The Ser in the dipeptides can be phosphorylated, but any putative Nor O-linked glycosylation sites were not found. The transcript of Aspein is expressed only in the edge of the mantle corresponding to the prismatic layer (Takeuchi and Endo, 2005). Seven cDNA clones encoding the members of an unusually acidic shell matrix protein family, Asprich, were identified from a cDNA library constructed from the mantle of the bivalve Atrina rigida (Gotliv et al., 2005). Because the amino acid sequences of the signal peptides of Aspein and Asprich are highly similar, the two proteins are obviously homologous. Asprich has a short basic domain in the N-terminal region followed by the unusually acidic region. The unusually acidic region contains the putative calcium-binding domain, D-D-D-S-E-D-D-D-D-D-D, forming a specific pocket motif identified by alignment to Calsequestrin, a Ca²⁺-binding protein in skeletal and cardiac muscles of vertebrates (Yano and Zarain-Herzberg, 1994). Asprich also contains four D-E-A-D motifs, which are thought to bind ATP under the presence of Mg^{2+} in DEAD box RNA helicase (Pause and Sonenberg, 1992), suggesting that Asprich are capable of binding ATP through Mg²⁺. Antibodies raised against a synthetic peptide of a part of an acidic region of Asprich reacted with the matrix components of the calcitic prismatic layer, but not with those of the aragonitic nacreous layer. Interestingly, all the three unusually acidic proteins, MSP-1, Aspein and Asprich, are present only in the calcitic shell layers, not in the aragonitic shell layer (Gotliv et al., 2005).

Nacrein is the first molluscan shell matrix protein of which the complete primary structure was revealed. Nacrein is an EDTA-soluble protein extracted from the nacreous layer of the pearl oyster *Pinctada fucata* (Miyamoto et al., 1996). Nacrein has two potential Nlinked glycosylation sites but the electrophoretic mobility of Nacrein did not change when treated by endopeptidase H and F. The protein was stained blue by Stains-all staining (Campbell et al., 1983), suggesting that the protein has a cation-binding ability. The protein exhibits a total of 26 tandem repeats of a G-X-N motif, where X is frequently D or N, a region suspected to be involved in calcium binding. The amino acid sequence of Nacrein shows the highest similarity to human carbonic anhydrase II, and Nacrein itself exhibits a carbonic anhydrase activity when examined using the CO₂-Veronal indicator method (Yang et al., 1985). Two homologs of Nacrein, N66 (Kono et al., 2000) and the Turbo marmoratus Nacrein (Miyamoto et al., 2003), were identified by extraction from the nacreous layer of Pinctada maxima, a closely related species to P. fucata, and by screening the cDNA library constructed from the mantle of the gastropod T. marmoratus, respectively. The transcript of N66 is expressed in both the dorsal region and the edge of

the mantle, suggesting that the protein is present in both the nacreous and the prismatic layer (Kono et al., 2000). Analysis by quantitative polymerase chain reaction also indicated that Nacrein transcript is present in both the dorsal and edge regions of the mantle in P. fucata (Takeuchi and Endo, 2005). In comparison with Nacrein, N66 has a longer repeat domain containing both G-X-N and G-N repeats. The Nacrein of T. marmoratus also has a longer repeat domain than that of P. fucata, but the repeat domain of the former is made up only of G-N repeats. These two repeat domains are far less acidic than that of P. fucata Nacrein, therefore they seem unlikely to have the calcium-binding ability. On the other hand, the carbonic anhydrase domain of the two homologs are conserved as in Nacrein of P. fucata.

Five insoluble, possibly "framework" proteins that provide a structure in which mineralization takes place, MSI60, MSI31 (Sudo et al., 1997), MSI7 (Zhang et al., 2003), N16/Pearlin (Samata et al., 1999; Miyashita et al., 2000), and Prismalin-14 (Suzuki et al., 2004) have been identified from the pearl oyster P. fucata, and a homolog of N16/Pearlin was also identified from another pearl oyster P. maxima (Kono et al., 2000). MSI60 was extracted from the nacreous layer, and the transcript of MSI60 was revealed to be expressed in the dorsal region of the mantle corresponding to the nacreous layer. MSI60 contains 11 polyalanine blocks of 9-13 residues conferring to the protein a similarity to spider-silk fibroins, and 39 polyglycine blocks of 3–15 residues possibly contribute to form β sheet conformation (Guerette et al., 1996). Polyaspartate blocks in both the N- and C-termini may bind calcium ions. The transcripts of MSI31 and MSI7 were isolated from cDNA library constructed from the mantle of P. fucata. MSI31 contains ten polyglycine blocks of 3-5 residues in the N-terminal half of the molecule. The C-terminal half is an acidic region including an X-S-E-E-D-Y motif tandemly repeated six times, where X is D or E, and Y is T or M, suggesting that this region binds calcium ions (Sudo et al., 1997). MSI7 shows a high sequence similarity to the N-terminal half of MSI31 and might be a truncated variant of MSI31. An in vitro experiment using the method of Weiss et al. (2000), monitoring the pH value dropping of CaCO₃-saturated solution containing the test material, suggests that MSI7 accelerates the precipitation of calcium carbonate. Interestingly, the transcript of MSI7 is expressed in both the dorsal region, which forms the nacreous layer, and the edge, which forms the prismatic layer, of the mantle (Zhang et al., 2003), whereas the transcript of MSI31 is expressed only in the edge of the mantle (Sudo et al.,

1997; Takeuchi and Endo, 2005). Taking account of the fact that the acidic X-S-E-D-Y repeated motif is present only in MSI31, while the polyglycine blocks are present in both MSI31 and MSI7, the acidic motif could be important to formation of the calcitic prismatic layer.

N16/Pearlin (Samata et al., 1999; Miyashita et al., 2000) was extracted from the nacreous layer of P. fucata and the gene encoding N14 (Kono et al., 2000) was identified from the cDNA library constructed from the dorsal region of the mantle of P. maxima. Their amino acid sequences, rich in Gly, Tyr and Asn, have four acidic regions, N-G repeats perhaps forming the glycine loops, Cys residues in the first two-thirds of the molecule and putative phosphorylation sites (Samata et al., 1999; Kono et al., 2000). Although N16/ Pearlin does not have a putative N-linked glycosylation site, periodic acid/Schiff (PAS) staining indicated that N16/Pearlin is glycosylated, suggesting that N16/ Pearlin is an O-linked glycosylated protein. Ca²⁺overlay analysis indicated that N16/Pearlin has a significant Ca²⁺-binding ability (Matsushiro et al., 2003). RT-PCR revealed that the transcripts of N16/Pearlin and N14 are expressed only in the dorsal region of the mantle. The crystallization experiment was carried out using CaCO₃-saturated solution containing Mg²⁺ and four types of test material, N16/Pearlin, N14, N66 and the mixture of N14 and N66; it indicated that, in all cases, the protein inhibited calcium carbonate precipitation in solution, but induced platy aragonitic tablets when adsorbed on insoluble matrix membranes (Samata et al., 1999; Kono et al., 2000).

Prismalin-14 was extracted from the prismatic layer, and the transcript was revealed to be expressed in the edge of the mantle corresponding to the prismatic layer. Prismalin-14 exhibited Ca²⁺-binding ability in the Ca²⁺-overlay analysis, and also showed inhibitory activity on calcium carbonate precipitation in an in vitro dose-dependent anti-calcification assay (Inoue et al., 2001). Thus acidic residues located in both the N- and C-termini as seen in MSI31 may bind calcium ions or the surface of calcium carbonate crystals. Glycine- and tyrosine-rich region in the C-terminal half of the molecule shows sequence similarity to keratin and may form the glycine loop which confers flexibility to the molecule (Steinert et al., 1984). A secondary structure prediction suggested that the region of P-I-Y-R repeats located in the N-terminus half forms β-strand conformation. Prismalin-14 is similar to MSI31 in that they are both water-insoluble, acidic and possibly responsible for the formation of the prismatic layer of P. fucata. Prismalin-14 shows no sequence similarity to MSI31 (Suzuki et al., 2004).

Four water-soluble matrix proteins have been extracted from the nacreous layer of abalones. Two of them, Perlucin and Perlustrin, identified from Haliotis laevigata, are neutral proteins exhibiting remarkable sequence similarities to vertebrate extracellular proteins. The other two, AP7 and AP24, identified from the red abalone H. rufescens, are moderately acidic proteins exhibiting no significant sequence similarities to known proteins. Perlucin (Weiss et al., 2000; Mann et al., 2000) is an N-linked glycosylated protein and has the C-type (Ca²⁺-dependent) carbohydrate-binding domain of the C-type lectins, followed by a C-terminal domain containing a motif of ten amino acid residues repeated twice. This C-terminal motif exhibits a significant sequence similarity to a repeated motif of P32 adhesin (Mann et al., 2000), a protein of Mycoplasma genitalium, which is involved in cytadherence and subsequent development of disease pathology (Reddy et al., 1995). Solid phase assays (Gabius et al., 1989) showed that Perlucin has a divalent metal iondependent ability to bind to glycoproteins containing galactose or mannose/glucose, indicating that Perlucin is a functional C-type lectin (Mann et al., 2000). Perlucin is different from other C-type lectin-like domain proteins isolated so far from other biominerals, because these proteins are not thought to have a carbohydrate-binding ability as described below. Perlucin promotes calcium carbonate precipitation (Weiss et al., 2000; Blank et al., 2003) as shown by the in vitro precipitation assay (Wheeler et al., 1981). The second protein, Perlustrin (Weiss et al., 2000; 2001), showed significant sequence similarity to the N-terminal domain of mammalian insulin-like growth factor binding proteins (IGFBPs). Surface plasmon resonance assays indicated that Perlustrin has notable binding affinity for human IGFs. Several IGFBPs are known to participate in bone metabolism in vertebrates, and IGFBP-5, the major bone IGFBP, is thought to bind to hydroxyapatite (Campbell and Andress, 1997). This is reminiscent of the fact that the water-soluble component of nacre stimulates bone formation (Westbroek and Marin, 1998; Lamghari et al., 1999; Almeida et al., 2000; Mouriès et al., 2002; Rousseau et al., 2003; Milet et al., 2004), and may support the notion that phylogenetically distant biomineralization systems, nacre and bone, contain components inherited from common ancestors (Weiss et al., 2001).

The other two abalone water-soluble proteins, AP7 and AP24, might form a dimer through hydrophobic interactions, because the ethanol eluant containing them could not be resolved into the two proteins by hydrophobic interaction chromatography, although they could be resolved by SDS PAGE (Michenfelder *et al.*, 2003). Synthetic N-termini of 30 amino acid residues of AP7 and AP24, which contain putative Ca^{2+} -binding sites, D-D and D-D-D-E-D, respectively, formed random coil-like structures confirmed by NMR spectroscopy (Wustman *et al.*, 2004). The crystal growth assay with calcite crystals and synthetic N-termini of AP7 and AP24 in CaCO₃-saturated solutions indicated that both AP7 and AP24 inhibited calcite growth, Michenfelder *et al.* (2003) thus inferred that these proteins contribute to the growth of aragonite. Electrophoretic mobility of AP24 changed when treated with N-glycosidase F, indicating that AP24 has N-linked carbohydrates (Michenfelder *et al.*, 2003).

Lustrin A is an insoluble, possibly a framework protein, identified from the nacreous layer of the abalone Haliotis rufescens. This protein is 1409 amino acid residues long excluding signal peptide (19 amino acids), which is the longest invertebrate skeletal matrix protein completely sequenced so far, and has a complex modular structure comparable to that of MSP-1 (Sarashina and Endo, 1998, 2001). Lustrin A is rich in serine, proline and glycine, and contains ten cystein-rich domains. Nine of them are interspersed by eight proline-rich domains and a glycine- and serine-rich domain lies between the two cystein-rich domains nearest to the C-terminus. They are followed by a basic domain and the C-terminal domain, and the latter is similar to protease inhibitors. The high sequence similarities among cystein-rich domains suggest that they undergo similar folding and perhaps contribute to protein-protein interactions. The density of proline residues of proline-rich domains are predicted to adopt an extended rod-like structure (Williamson, 1994), and the domains thus possibly contribute to separate the cystein-rich domains so that they can fold independently. The G-S and G-S-S-S repeats in the glycineand serine-rich domain may form rubber-like glycine loops (Steinert et al., 1991) and the domain may have elastic property. The basic and C-terminus domains may play a role in interacting with anionic molecules and protecting secreted matrix proteins.

Three completely sequenced shell matrix proteins have been identified from molluscan species other than pearl oysters and abalones. One is MSP-1 as described above, and the other two are Mucoperlin and Dermatopontin. Mucoperlin is a water-soluble acidic protein identified from a cDNA expression library constructed from the mantle tissues of the Mediterranean fan mussel *Pinna nobilis* (Marin *et al.*, 2000). Mucoperlin contains a tandem repeat region consisting of 13 repeated units of 31 amino acid residues rich in Ser and Pro, suggesting that this region has a rigid rod-like conformation (Williamson, 1994). Four putative N-linked glycosylation sites and 27 putative O-linked glycosylation sites are recognized in this region, suggesting that Mucoperlin may be heavily glycosylated. The acidic C-terminal region of Mucoperlin contains densely distributed putative phosphorylation sites and two putative sulfated Tyr sites (Huttner, 1987). Mucoperlin exhibits a certain sequence similarity to PGM, a pig gastric mucin. Mucoperlin also has structural features characteristic of mucins, such as Pro/Ser-rich repeats and putative O-linked serine residues. The immunodetection of Mucoperlin performed using ELISA and dot-blot analysis with the antiserum directed against recombinant non-glycosylated Mucoperlin indicated that Mucoperlin is located in the nacreous layer but not in the prismatic layer. In situ immunohistochemical localization assay using an antibody elicited against the recombinant Mucoperlin indicated that Mucoperlin is localized around polygonal nacre tablets. Mucoperlin may control the lateral extension of the tablet (Marin et al., 2000; Marin and Luauet, 2004).

Dermatopontin is a water-soluble neutral protein extracted from the aragonitic cross-lamellar layer of the freshwater snail Biomphalaria glabrata (Marxen and Becker, 1997; Marxen et al., 2003). Dermatopontin was revealed to have a pI of 7.4 by isoelectric focusing (IEF). Alcyan blue and Stains-all staining indicated that Dermatopontin is an acidic mucopolysaccharide and has no Ca2+-binding ability (Campbell et al., 1983; Butler et al., 1981). Dermatopontin contains a single N-glycosylation site, and that Asn in this site is modified by a short N-glycan was confirmed using tandem mass spectrometry (MS/MS). Dermatopontin is also called TRAMP (Tyrosine-Rich Acidic Matrix Protein), which is a component of extracellular matrix in vertebrates (Neame et al., 1989; Superti-Furga et al., 1993; Forbes et al., 1994) and invertebrates (Fujii et al., 1992; Schütze et al., 2001). Dermatopontin shows a widespread tissue distribution in mammals, including skin, skeletal muscle, heart, lung, kidney, cartilage and bone with possible functions in cell-matrix interactions and matrix assembly (Forbes et al., 1994). In mammals, they bind to decorin, and the Dermatopontin-decorin complex accelerates collagen fibrillogenesis (MacBeath et al., 1993) and modifies the behaviour of TGF-B (Okamoto et al., 1996, 1999). No collagen has been identified in the shell matrix of B. glabrata (Marxen and Becker, 1997), but the shell matrix contains a considerable amount of proteoglycans (Marxen et al., 1998). Thus, Dermatopontin may interact with shell proteoglycans to organize an organic scaffold for mineralization (Marxen et al., 2003). Dermatopontin is inferred to have been coopted for molluscan calcification after the divergence of molluscs and other metazoan phyla, possibly in much more recent times than the "Cambrian explosion" based on the data obtained from molecular phylogenetic analysis (Sarashina *et al.*, 2005).

Echinodermata

The mineral component of the spicules of sea urchins is magnesian calcite, the ratio of Mg^{2+}/Ca^{2+} being approximately 1/19 (Okazaki and Inoué, 1976). Amorphous calcium carbonate transforms into calcite during spicule growth (Beniash et al., 1997). The spicule is likely to be composed of a single crystal based on the results of X-ray diffraction analyses (Beniash et al., 1999). Some researchers, however, have argued that sea urchin spicules are built from a polycrystalline aggregate based on the results of X-ray diffraction analyses, chemical analyses and scanning electron microscopical observation (Okazaki and Inoué, 1976), and this possibility cannot be excluded because X-ray diffraction is not able to distinguish the assembly from a single crystal in some cases (Robach et al., 2005). The larval spicule also contains organic matrix of about 0.1-1% by weight (Seto et al., 2004; Okazaki and Inoué, 1976), which includes approximately four dozen matrix proteins at least (Killian and Wilt, 1996). In the case of Strongylocentrotus purpuratus, the complete amino acid sequences of eight spicule matrix proteins have been determined so far (Wilt, 1999, 2002; Wilt et al., 2003; Wilt, 2005).

Two genes encoding two spicule matrix proteins, SM50 (Benson et al., 1987; Sucov et al., 1987; Katoh-Fukui et al., 1991) and SM30, or pNG7 SM30 (George et al., 1991), have been identified from cDNA expression libraries constructed from the whole embryo of the purple sea urchin Strongylocentrotus purpuratus using antiserum directed against the total spicule matrix proteins. Within the embryo, both the SM50 and SM30 transcripts are expressed exclusively in the primary mesenchyme cells (PMCs). SM50 is a basic protein with a molecular mass of approximately 50 kDa, and has no putative N-linked glycosylation sites. The protein has a Pro-rich region and a repeated motif of 13 amino acid residues, Q-P-G-F/M/W-G-N/G-Q-P-G-V/M-G-G-R/Q. The N-terminal half of SM50 shows sequence similarity to C-type lectin. In the upstream region of the SM50 gene there exists an ets binding site that functions as a positive cis-regulatory element (Kurokawa et al., 1999). SM50 is expressed in both embryonic spicules and adult spines (Richardson et al., 1989). At the late gastrula stage, the expression is stronger in the ventrolateral clusters than dorsal chain in the PMC syncytium (Lee et al., 1999). In situ

immunohistochemical localization assay revealed that SM50 is localized around the spicule surface, on which spicule deposition occurs (Urry et al., 2000). LSM34, a homolog of SM50, was identified from the cDNA library constructed from mesenchymal and endoderm cell fractions of the painted sea urchin Lytechinus pictus using S. purpuratus cDNA encoding the spicule matrix protein as the probe (Livingston et al., 1991). This protein has a repeated motif of 7 amino acid residues, G-G-Q/R-Q-P-G-F, slightly different from that of SM50. Both repeated motifs of SM50 and LSM34 show sequence similarities to the repeated motifs of elastin and wheat gluten. The motifs have been considered to form a β -spiral supersecondary structure and confer elastic properties to the molecules (Venkatachalam and Urry, 1981; Miles et al., 1991). Peled-Kamar et al. (2002) showed that inhibition of LSM34 expression resulted in a block of spicule elongation. They injected antisense oligonucleotides directed against LSM34 transcript to the blastocoel of L. pictus embryos at the hatching blastula stage. Although small, possibly calcitic granules were recognized in the PMCs of antisense-treated embryos, those embryos had no apparent spicules. These results suggest that LSM34 is essential for spicule elongation, even if initial nucleation does not require LSM34. HSM41, a second homolog of SM50, was identified from the cDNA library constructed from the whole embryo of the Japanese sea urchin Hemicentrotus pulcherrimus screened with a cDNA clone of SM50 (Katoh-Fukui et al., 1992). HSM41 has a repeated motif of 13 amino acid residues, Q-P-G-F-G-N-Q-P-G-M/V-G-G-R/Q/N. This repeated motif is more similar to that of SM50 than to that of LSM34, obviously reflecting the evolutionary relationship of the three species.

SM30, or *pNG7* SM30, is an acidic matrix protein with a molecular mass of approximately 30 kDa (George et al., 1991) and possibly the most abundant spicule matrix protein judging from the signal intensities of ³⁵S-labeled proteins separated by twodimensional gel electrophoresis (Killian and Wilt, 1996). Southern blot analysis indicated that there are two to four copies of SM30 genes in the S. purpuratus haploid genome (Akasaka et al., 1994), while the SM50 gene occurs once per haploid genome (Sucov et al., 1987). Akasaka et al. (1994) isolated a genomic clone and found that at least two SM30 genes are tandemly arranged, designating them SM30-a and SM30-B. The former is completely and the latter is partially sequenced. The SM30 protein encoded by the pNG7clone previously reported by George et al. (1991), or pNG7 SM30, shows a high sequence similarity to the SM30- α protein. Most of the amino acid differences

between pNG7 SM30 and SM30-α, however, occur at the probably important sites conserved between SM30- α and other C-type lectins, suggesting that they may be not alleles but different forms of the SM30 proteins (Killian and Wilt, 1996). On the other hand, SDS-PAGE separations revealed that embryonic spicules contain SM30-A of 43 kDa and SM30-B of 46 kDa, and adult spines contain SM30-B and SM30-C of 49 kDa. Although SM30- α probably encodes SM30-A based on the expression patterns of the SM30- α transcript, we cannot be sure which SM30 gene encodes SM30-B or -C (Killian and Wilt, 1996). The observations on the electrophoretic mobilities of SM30-A, -B and -C treated with endoglycosydase F indicated that they are all N-linked glycosylated, harmonizing with the fact that each deduced amino acid sequence of pNG7 SM30 and SM30- α contains a putative N-linked glycosylation site (George et al., 1991; Akasaka et al., 1994). Both pNG7 SM30 and SM30-α proteins are rich in Pro and Ala, but do not have a conspicuous Pro-rich domain as in SM50. SM30-a exhibits high sequence similarity to the carbohydrate recognition domain (CRD) of C-type lectins, but lacks a pair of cysteine residues. The Ca²⁺-binding sites, which are conserved in the CRDs of other C-type lectines, are not conserved in SM30-a (Killian and Wilt, 1996). In situ immunohistochemical localization assay revealed that SM30 is localized within the spicule (Urry et al., 2000).

Five cDNAs, abundantly expressed in PMCs at the time of PMC ingression, were identified from the cDNA library constructed from differentiated micromere cultures of S. purpuratus (Harkey et al., 1988). The complete primary structure of one of them, designated PM27, was determined (Harkey et al., 1995). PM27 is possibly a nonglycosylated protein. PM27 has no potential N-linked glycosylation sites and, although it has potential O-linked glycosylation sites, they have very weak potential. The N-terminal domain consists of eight tandem repeats of P-G-M-G and five tandem repeats of Q-G, showing sequence similarity to several vertebrate fiber-forming proteins such as keratin, collagen, elastin and wheat glutinin. This N-terminal repeat domain is followed by the domain exhibiting sequence similarity to C-type lectin. The immunolocalization with the antibodies raised against recombinant PM27 indicated that PM27 accumulated at the growing surface of the spicule tips and disappeared from the mature mid-shaft region of the spicules. These results suggest that PM27 is involved in skeletal growth rather than maintenance or structural integrity of the spicules. In Western blot analyses, both antisera directed against the two different domains of PM27 detected two proteins in larvae and three proteins in adult tests. Since one is probably common to both larval and adult tests, at least four proteins sharing a similar structure are expressed during skeletogenesis (Harkey et al., 1995). The results of Southern and Northern blot analyses, however, are consistent with the conclusion that PM27 is a singlecopy gene (Harkey et al., 1988, 1995). Thus, the most plausible source of the multiple proteins is the posttranslational modifications. The SM37 gene was found in the genomic DNA fragment that contains the SM50 gene. The two genes are linked at a distance of about 12 kb (Lee et al., 1999). The SM37 gene included the same or highly similar cis-regulatory elements to that of the SM50 gene (Makabe et al., 1995; Lee et al., 1999). At the late gastrula stage, the expression of SM37 is stronger in the ventrolateral clusters than in the dorsal chain in the PMC syncytium. The results of both RT-PCR using the skeletogenic mesenchyme cell culture and the whole-mount in situ hybridization indicated that, in the embryos, the transcript of SM37 is expressed exclusively in skeletogenic mesenchyme cells, exactly as in SM50. The probe excess titration measurements (Killian and Wilt, 1989), which allow determination of the absolute number of transcripts, revealed that the transcripts of SM37 and SM50 are quantitatively almost identical through embryonic development (Lee et al., 1999), and these temporal schedules of SM37 and SM50 expression are different from SM30 (Guss et al., 1997). These results suggest that the two genes encoding SM37 and SM50 are regulated coordinately, and SM37 is very probably also a spicule matrix protein like SM50. SM37 has three putative N-linked glycosylation sites. Although the overall structure of SM37 is similar to SM50, the repeated motif of SM37, the consensus form of which is G-A/G-G-A/G-G-G-A-G-A-G-G-R-W-N-P-N-Q, is fairly different from SM50 except for the scaffold of Gly, Gln and Pro (Lee et al., 1999).

Zhu *et al.* (2001) carried out a large-scale analysis of mRNAs expressed in PMCs at the gastrula stage of *S. purpuratus.* Two species of mRNAs were identified to be PMC-specific using whole-mount in situ hybridization. Thereafter they reported the complete amino acid sequences of two putative spicule matrix proteins encoded by these two mRNAs, SpSM32 (originally named SM50-related protein) and SpC-lectin, and another new putative spicule matrix protein SpSM29, the transcript of which is also expressed only in the PMCs (Illies *et al.*, 2002).

SpSM29 and SpSM32 are basic proteins just like SM50, PM27 and SM37. SpSM29 has a Pro-rich repeat domain in the N-terminal side and a C-type lectin do-

main in the C-terminal side, similar to PM27, while SpSM32 has a C-type lectin domain in the N-terminal side and a Pro-rich repeat domain in the C-terminal side, similar to SM50 and SM37. The repeated motif of SpSM32 is P-X-Y, where X is usually Asn and Y is usually Gln. A region of approximately 200 nucleotides at the 5' end of SpSM32 mRNA, corresponding to the 5' untranslated region and the first 35 amino acid residues of SpSM32, is identical to the sequence of SM50 at the nucleotide level. One possibility is that SpSM32 and SM50 mRNAs share a common exon as the result of a cis-splicing. A single primary transcript including open reading frames of both SpSM32 and SM50 may undergo alternative cissplicing. The SM50 gene is known to be closely linked to the SpSM37 gene as described above, so it is possible that the SpSM32 gene is also a member of this complex (Illies et al., 2002).

SpC-lectin is acidic and has a C-type lectin domain but does not have a Pro-rich repeat domain, just like SM30 proteins. Although SpC-lectin is possibly a spicule matrix protein, it is not yet clear whether the SpC-lectin proteins interact directly with the mineral phase of the spicule (Illies *et al.*, 2002).

Prominent features of skeletal matrix proteins of invertebrates

Acidic nature

The complete amino acid sequences of 77 invertebrate skeletal matrix proteins have been published to date (Table 1, Figure 1). The information of the primary structure revealed some prominent features characteristic to skeletal matrix proteins. One of the most conspicuous is their acidic nature.

Aspartic acid (Asp) has long been known as a major component of skeletal matrix proteins, and has been regarded as important in biomineralization processes by virtue of its nature as an acid to interact with Ca^{2+} (Hare, 1963; Weiner and Hood, 1975; Mitterer, 1978; Weiner, 1979, 1983; Runnegar, 1984; Weiner and Addadi, 1991; Wheeler, 1992, Albeck et al., 1993; Gotliv et al., 2003). There were, however, some suspicions that the amino acids detected as aspartic acid in the amino acid composition analysis might in fact be present as their amide form, asparagine, in the matrix proteins (Young, 1971; Crenshaw, 1972, 1982). But characterization of the primary structure of the scallop shell protein MSP-1 (Sarashina and Endo, 1998, 2001) clearly demonstrated that there indeed exists a shell matrix protein rich in aspartate (Asp) residues (20 mol%). Based on the MSP-1 sequence, the gene sequence for even more Asp-rich shell matrix proteins

Aspein (60%) and Asprich (45%) were isolated (Tsukamoto et al., 2004; Gotliv et al., 2005). In fact, these three shell matrix proteins have the lowest values of isoelectric point (MSP-1, 3.2; Aspein, 1.5; Asprich, 2.5) among the 77 sequenced invertebrate skeletal matrix proteins. Among these 77 invertebrate skeletal matrix proteins, 43 are acidic (pI < 6.0), 20 neutral and 14 basic (pI > 8.0), suggesting that acidic skeletal matrix proteins are more common than basic ones. Only in echinoderms is the number of known basic skeletal matrix proteins (5) more than that of acidic ones (2). This fact is likely to reflect sampling bias rather than the whole nature of the spicule matrix proteins, because it is technically difficult to isolate unusually acidic proteins as a single protein (Wheeler et al., 1981; Gotliv et al., 2003). Two-dimensional gel electrophoreses of water-soluble spicule matrix proteins of the sea urchin S. purpuratus revealed that the majority of them are acidic (Killian and Wilt, 1996), supporting the interpretation that the basic nature of sequenced echinoderm spicule matrix proteins reflects the technical bias.

The acidic nature of skeletal matrix protein is mainly due to high content of acidic amino acid residues, Asp and Glu, although phosphorylation, sulfation and glycosylation may also be important. In 77 sequenced invertebrate skeletal matrix proteins, unusually acidic amino acid-rich (> 20%) proteins are Selenoprotein M (Porifera), CaCP5.75, CsCP8.2, CsCP8.5, Orchestin, Crustocalcin/DD4, DD9A, CAP-1, CAP-2 (Arthropoda), MSP-1, Aspein and Asprich (Mollusca). Among these, unusually Asp-rich (> 15%) proteins are CsCP8.2, CsCP8.5, Orchestin, CAP-1 (Arthropoda), MSP-1, Aspein and Asprich (Mollusca), and unusually Glu-rich (> 15%) proteins are Selenoprotein M (Porifera), Crustocalcin/DD4 and CAP-2 (Arthropoda). Moreover, Crustocalcin/DD4 contains a conspicuous Glu-rich region, while MSP-1, Aspein and Asprich have conspicuous Asp-rich regions. These facts suggest that, in arthropods, not only Asp but Glu is important for providing negative charges to skeletal matrix proteins.

Repeated structures

Repeated sequences are commonly found among skeletal matrix proteins. In fact, they are found in skeletal matrix proteins in all the five invertebrate phyla from which the sequence data have been obtained. Most repeated sequences are rather short, ten or less amino acids in length, such as D-S or S-D (Orchestin), G-N or N-G (N66, *Tm*Nacrein, N16/Pearlin), G-S or S-G (MSP-1, Lustrin A), Q-G (PM27), S-D-E, S-R-E (Orchestin), D-S-D (MSP-1), G-S-S (Lustrin

Table 1. List of matrix proteins in invertebrates referred to in this article. For all the proteins listed here, the complete primary structure has been determined. The source organisms, references and the DNA databank accession numbers are also listed. Numbers in parentheses are the accession numbers in the MIPS-Protein Sequence Database (Martinsried, Germany).

Porifera			
Silicatein a	Tethya aurantia	Shimizu et al., 1998	AF032117
Silicatein	Tethya aurantia	Funayama et al., 2005	AF098670
SILICAa_SUBDO	Suberites domuncula	Krasko et al., 2000	AJ272013
SILICAb_SUBDO	Suberites domuncula	Schröder et al., 2005	AJ784227
Ef Silicatein	Ephydatia fluviatilis	Funayama et al., 2005	AB219573
Pf Silicatein	Petrosia ficiformis	Pozzolini et al., 2004	AY158071
Silicase (SIA_SUBDO)	Suberites domuncula	Schröder et al., 2003	_
Selenoprotein M	Suberites domuncula	Müller et al., 2005	AJ875186
Spicule-associated protein	Suberites domuncula	Müller et al., 2005	AJ872182
COL1_SUBDO	Suberites domuncula	Schröder et al., 2000	AJ252241
Cnidaria			
Galaxin	Galaxea fascicularis	Fukuda et al., 2003	AB086183
Arthropoda	-		
H1a	Pandalus borealis	Jacobsen et al., 1994	_
Pb CP-12.7	Pandalus borealis	Suzuki et al., 2001	_
HaCP4.4	Homarus americanus	Nousiainen et al., 1998	(\$77929)
HaCP4.5	Homarus americanus	Nousiainen et al., 1998	(\$77931)
HaCP4.6	Homarus americanus	Kragh et al., 1997	Q7M4A1
HaCP5.6	Homarus americanus	Nousiainen et al., 1998	(\$77936)
HaCP5.9	Homarus americanus	Nousiainen et al., 1998	(\$77932)
HaCP6.3	Homarus americanus	Nousiainen et al., 1998	(\$77934)
HaCP9.3	Homarus americanus	Nousiainen et al., 1998	(\$77937)
HaCP11.6	Homarus americanus	Kragh et al., 1997	Q7M4E3
HaCP11.8	Homarus americanus	Nousiainen et al., 1998	(\$77925)
HaCP12.7	Homarus americanus	Nousiainen et al., 1998	(\$77926)
HaCP14.2	Homarus americanus	Nousiainen et al., 1998	(S77927)
HaCP18.8	Homarus americanus	Nousiainen et al., 1998	(\$77928)
HaCP20.2	Homarus americanus	Nousiainen et al., 1998	(S77930)
CpCP4.34	Cancer pagurus	Andersen, 1999	P81585
CpCP4.59	Cancer pagurus	Andersen, 1999	P81586
CpCP4.63	Cancer pagurus	Andersen, 1999	P81587
CpCP4.66	Cancer pagurus	Andersen, 1999	P81590
CpCP4.98	Cancer pagurus	Andersen, 1999	P81588
CpCP5.75	Cancer pagurus	Andersen, 1999	P81589
CpCP11.14	Cancer pagurus	Andersen, 1999	P81575
CpCP11.58	Cancer pagurus	Andersen, 1999	P81580
CpCP12.43	Cancer pagurus	Andersen, 1999	P81581
CpCP12.46	Cancer pagurus	Andersen, 1999	P81582
CpCP14.99	Cancer pagurus	Andersen, 1999	P81583
CpCP18.76	Cancer pagurus	Andersen, 1999	P81584
Orchestin	Orchestia cavimana	Testenière et al., 2002	AF124526
GAMP	Procambarus clarkii	Tsutsui et al. 1999	AB017017
Crustocalcin/DD4	Penaeus japonicus	Endo et al., 2004	AB049148
DD5	Penaeus japonicus	Ikeya et al., 2001	AB049147
DD9A	Penaeus japonicus	Watanabe et al., 2000	AB031223
DD9B	Penaeus japonicus	Watanabe et al., 2000	AB031224
CAP-1	Procambarus clarkii	Inoue et al., 2003	AB103035
CAP-2	Procambarus clarkii	Inoue et al., 2004	AB167814
CsCP8.2	Callinectes sapidus	Wynn and Shafer, 2005	AY752736
CsCP8.5	Callinectes sapidus	Wynn and Shafer, 2005	AY752735
Mollusca		<u>,</u>	
MSP-1	Patinopecten yessoensis	Sarashina and Endo, 2001	AB073617
Aspein	Pinctada fucata	Tsukamoto et al., 2004	AB094512
Asprich	Atrina rigida	Gotliv et al., 2005	AY660590
Nacrein	Pinctada fucata	Miyamoto et al., 1996	D83523
N66	Pinctada maxima	Kono et al., 2000	AB032613

-

325

Turbo marmoratus Nacrein	Turbo marmoratus	Miyamoto et al., 2003	AB073680
MSI60	Pinctada fucata	Sudo et al., 1997	D86074
MSI31	Pinctada fucata	Sudo et al., 1997	D86073
MSI7	Pinctada fucata	Zhang et al., 2003	AF516712
N16/Pearlin	Pinctada fucata	Samata et al., 1999	AB023067
N14	Pinctada maxima	Kono et al., 2000	AB032612
Prismalin-14	Pinctada fucata	Suzuki et al., 2004	AB159512
Perlucin	Haliotis laevigata	Mann et al., 2000	S78774
Perlustrin	Haliotis laevigata	Weiss et al., 2001	P82595
AP7	Haliotis rufescens	Michenfelder et al., 2003	AF225916
AP24	Haliotis rufescens	Michenfelder et al., 2003	AF225915
Lustrin A	Haliotis rufescens	Shen et al., 1997	AF023459
Mucoperlin	Pinna noblis	Marin et al., 2000	AF145215
Dermatopontin	Biomphalaria glabrata	Marxen et al., 2003	P83553
Echinodermata			
SM50	Strongylocentrotus purpuratus	Katoh-Fukui et al., 1991	M16231
LSM34	Lytechinus pictus	Livingston et al., 1991	Q05904
HSM41	Hemicentrotus pulcherrimus	Katoh-Fukui et al., 1992	Q26264
<i>pNG</i> 7 SM30	Strongylocentrotus purpuratus	George et al., 1991	M63840
SM30-α	Strongylocentrotus purpuratus	Akasaka et al., 1994	Q26646
PM27	Strongylocentrotus purpuratus	Harkey et al., 1995	U18132
SM37	Strongylocentrotus purpuratus	Lee et al., 1999	AF068737
SpSM29	Strongylocentrotus purpuratus	Illies et al., 2002	AF519417
SpSM32	Strongylocentrotus purpuratus	Illies et al., 2002	AF519416
SpC-lectin	Strongylocentrotus purpuratus	Illies et al., 2002	AF519418

Table 1. Continued

A), G-X-N (Nacrein, N66), P-N-Q (SpSM32), S-D-E-S (Orchestin), D-G-S-D (MSP-1), P-I-Y-R (Prismalin-14), P-G-M-G (PM27), D-D-S-R-E (Orchestin), G-G-G-G-S (MSI60), E-E-D-M/T-E/S (MSI30), Q-Q-A-A-P-A/T (GAMP), X-S-E-E-D-Y (MSI30), X-A-G-X-X-P-Y (Pb CP-12.7), G-G-Q/R-Q-P-G-F (LSM34), E-S-R/E-E-P-R-K-L (Orchestin) and N/D-S-L-H-A-N-L-Q-Q-R (Perlucin). Some repeated sequences of skeletal matrix proteins are longer, ranging from 13 to 31 amino acids (a.a.) in length. They are recognized in SM50, HSM41 (13 a.a.), SM37 (17 a.a.), Spicule-associated protein (20 a.a.), Galaxin (27-31 a.a.) and Mucoperlin (31 a.a.). Gly, Ser, Glu, Asp, Gln, Pro and Asn residues occur frequently in these repeated motifs. Functions of the repeats have been proposed as binding to chitin (P-I-Y-R), binding to some macromolecules (repeats in Galaxin), binding to membrane (Q-Q-A-A-P-A/T, repeats in Spiculeassociated protein), flexible property (N-G, G-N, S-G, G-S, G-S-S, G-G-G-G-S, G-G-Q/R-Q-P-G-F, N/D-S-L-H-A-N-L-Q-Q-R and repeats in SM50 and HSM41), rod-like property (repeats in Mucoperlin) and Ca²⁺-binding in many repeated motifs which have acidic properties.

Extremely long repeated motifs are found in three skeletal matrix proteins, Lustrin A (75–88 a.a.), DD5 (93–98 a.a.) and MSP-1 (158-177). Lustrin A and MSP-1 are molluscan shell matrix proteins having a complex modular structure. The long repeated motifs

of MSP-1 and Lustrin A contain a Gly- and Ser-rich domain having an elastic property and a Pro-rich domain having an extended rod-like structure, respectively. They may contribute to separate the functionally important domains in the repeated motifs so they can fold independently and may bind to Ca^{2+} (in MSP-1) or other proteins (in Lustrin A). Each of the 13 repeated motifs in DD5 contains the Rebers-Riddiford sequence, a chitin-binding site, but does not contain the connecting domain such as the Gly-and Ser-rich domain in MSP-1 or the Pro-rich domain in Lustrin A. One DD5 protein molecule may be cross-linked to 13 chitin fibers to form a rigid and compact macromolecular network.

Although various functions for the repeated sequences have been proposed, the commonality of the short repeated structures among different animal phyla appear to suggest strong functional constraints such as interactions with the regular arrangement of ions in crystal lattices. On the contrary, the long repeated unit may contain domains which fold and function independently to the adjacent repeated units. Thus, the long repeats may not necessarily reflect the regular arrangement of ions in crystals.

C-type lectin-like domains

CRDs of C-type lectin are known to mediate sugarbinding under the presence of Ca^{2+} (Drickamer, 1999). Except for Tetranectin found in the cartilage of the



Figure 1.



Figure 1. Amino acid contents by percentage and isoelectric points of invertebrate skeletal matrix proteins the complete primary structure of which has been reported. The black bars on the left side represent percent aspartic acid residues. Dark grey and light grey bars represent glutamic acid residues and putative phosphorylation sites, respectively. Those three kinds of residues contribute to the acidic nature of proteins. Black bars on the right side represent basic residues including arginine, lysine and histidine residues.

shark, C-type lectin-like domains (CTLD) have been found in calcium carbonate skeletal matrix proteins, such as molluscan Perlucin, all ten echinodermal spicule matrix proteins as well as several vertebrate skeletal matrix proteins, Lithostathine (originally named Pancreatic stone protein; De Caro et al., 1987; Bertrand et al., 1996), Ovocleidin-17 (Hincke et al., 1995; Mann and Siedler, 1999), Ovocleidin-116 (Hinke et al., 1999), Ansocalcin (Lakshminarayanan et al., 2002, 2003), Struthiocalcin-1, Struthiocalcin-2 (Mann and Siedler, 2004), Dromaiocalcin-1, Dromaiocalcin-2, Rheacalcin-1, Rheacalcin-2 (Mann, 2004). Interestingly, CTLDs found in calcium carbonate skeletal matrix proteins in both vertebrates and echinoderms do not bind carbohydrates and do not retain the carbohydrate-binding amino acid residues which are characteristic of lectins. Lithostathine, a vertebrate pancreatic stone protein, is known to contain a putative Ca²⁺-binding site but it is not located within the CTLD (Bertrand et al., 1996). It is difficult to predict functions of these CTLDs in calcium carbonate skeletal matrix proteins, as CTLDs are known to exhibit diverse functions (Drickamer, 1999). Perlucin in molluscs, in contrast to the deuterostome calcium carbonate skeletal CTLD proteins, has a functional carbohydrate-binding ability, and may be involved in Ca²⁺-dependent glycoprotein-protein interaction within the skeletal matrix (Mann et al., 2000). Thus, functions of CTLDs found in skeletal matrix proteins are possibly different between deuterostomes and molluscs.

Tetranectin is a member of the family of mammalian C-type lectin, shared by the blood and the extracellular matrix. Neame *et al.* (1992) identified a homolog of Tetranectin from the reef shark cartilage, which is possibly capable of binding sugars. Tetranectin is also thought to play an important role during osteogenesis in mice (Wewer *et al.*, 1994), and its function is highly likely to be different from the calcium carbonate skeletal CTLD proteins.

Glycosylation

Glycosylation is a common modification of skeletal matrix proteins. General functions of protein glycosylation include protein folding and stability, receptor functioning, cell adhesion and signal transduction (Bhatia and Mukhopadhyay, 1999). Cho *et al.* (1996) proposed an oligosaccharide-mineral binding function based on the observation that oligosaccharide moieties are found in association with calcium carbonates of sea urchin spicules. In many cases, however, the significance of glycosylation in skeletal matrix proteins has not been clarified.

Implication and conclusion

From the morphological viewpoint, skeletons of many metazoan phyla appear to have evolved independently. However, because of the findings that the molluscan shell nacres have the ability to induce osteogenesis in vertebrates, metazoans belonging to distantly related phyla became suspected to have common underlying machineries of biomineralization, at least partly (Westbroek and Marin, 1998). Recently some skeletonization regulatory genes and their networks in invertebrates as well as vertebrates have been clarified. So far, the results seem to support the idea of independent origins of metazoan biomineralization, a premise which is also supported by the complete primary structure of skeletal matrix proteins as summarized above. Of course, currently we have very little information on molecules involved in biomineralization. We could not draw definitive conclusions and further information is awaited.

The repeated sequences are one of the common features of the skeletal matrix proteins. As described above, the amino acid compositions of the repeated sequences are biased and somewhat similar among metazoan phyla. But the amino acid sequences are different from each other. So this commonality of repeated sequences in metazoan skeletal matrix proteins may be a result of convergence. The acidic nature is also a major common feature of the skeletal matrix proteins. This common feature is also considered as convergence because significant sequence similarities are not observed among acidic skeletal proteins of different phyla. For instance, the unusually acidic shell matrix proteins of molluscs contain exclusively Asp as the acidic amino acid, whereas those of arthropods also contain Glu in addition to Asp. C-type lectin-like domains (CTLD) are found in some skeletal matrix proteins in both molluscs and deuterostomes, implying at first sight that this part of the underlying machineries has a common origin. The sites important for carbohydrate-binding, however, are not conserved between molluscs and deuterostomes, and thus the roles of CTLD in biomineralization are likely to be totally different between the two groups. These two groups are likely to have coopted the same gene independently after they diverged. Eight out of the known 37 arthropod skeletal matrix proteins have the Rebers-Riddiford consensus sequence which is found in non-calcified cuticular proteins in arthropods. This observation indicates that these arthropod biomineralization proteins have been coopted from the cuticular proteins after arthropods diverged from other metazoan groups. Dermatopontin, a molluscan shell matrix

protein, is also considered to have been coopted for biomineralization after molluscs diverged from other metazoan groups based on the molecular phylogenetic analysis.

The information obtained from regulatory genes involved in biomineralization, their networks and invertebrate skeletal matrix proteins, therefore, seems to support the idea that the machineries involved in metazoan biomineralization have evolved independently after the divergence of metazoan phyla. Although we find some homologous components among metazoan biomineralization machineries, they can be explained by independent cooptions of the common components. At least some metazoan groups coopted for biomineralization some common components which had served some other preexisting functions than biomineralization. This scenario is consistent with the observation that the molluscan shell nacre or matrices have the ability to induce osteogenesis in vertebrates. Marin et al. (1996) detected immunological crossreactivities between skeletal watersoluble matrices and mucous excretions of both a cnidarian and molluscs. This result seems to suggest that calcifying machineries of the two phyla derived, at least partly, from the common noncalcifying precursors. This observation also seems to harmonize with our scenario.

Furthermore, some researchers raise a possibility that metazoan biomineralization systems have been derived from a system that the last common ancestor of metazoans possessed. More specifically, Kirschvink and Hagadorn (2000) hypothesized that the ancestral eukaryotes inherited the ability to precipitate magnetite from magnetotactic bacteria during endosymbiotic events. The magnetite system would have to have been present in most of the metazoan phyla during the "Cambrian explosion", and could have provided a template for other biomineralization systems to evolve. At least so far, however, we have not found significant evidence to support this hypothesis.

References

- Akasaka, K., Frudakis, T. N., Killian, C. E., George, N. C., Yamasu, K., Khaner, O. and Wilt, F. H., 1994: Genomic organization of a gene encoding the spicule matrix protein SM30 in the sea urchin *Strongylocentrotus purpuratus*. *The Journal of Biological Chemistry*, vol. 269, p. 20592– 20598.
- Albeck, S., Aizenberg, J., Addadi, L. and Weiner, S. 1993: Interactions of various skeletal intracrystalline components with calcite crystals. *Journal of the American Chemical Society*, vol. 115, p. 11691–11697.
- Almeida, M.J., Milet, C., Peduzzi, J., Pereira, L., Haigle, J.,

Barthelemy, M. and Lopez, E., 2000: Effect of watersoluble matrix fraction extracted from the nacre of *Pinctada maxima* on the alkaline phosphatase activity of cultured fibroblasts. *The Journal of Experimental Zoology*, vol. 288, p. 327–334.

- Almeida, M. J., Pereira, L., Milet, C., Haigle, J., Barbosa, M. and Lopez, E., 2001: Comparative effects of nacre watersoluble matrix and dexamethasone on the alkaline phosphatase activity of MRC-5 fibroblasts. *Journal of Biomedical Materials Research*, vol. 57, p. 306–312.
- Andersen, S.O., 1998: Characterization of proteins from arthrodial membranes of the lobster, *Homarus americanus*. *Comparative Biochemistry and Physiology Part A*, vol. 121, p. 375–383.
- Andersen, S. O., 1999: Exoskeletal proteins from the crab, Cancer pagurus. Comparative Biochemistry and Physiology Part A, vol. 123, p. 203–211.
- Atlan, G., Balmain, N., Berland, S., Vidal, B. and Lopez, E., 1997: Reconstruction of human maxillary defects with nacre powder: histological evidence for bone regeneration. *Comptes Rendus de l'Académie des Sciences—Série III*, vol. 320, p. 253–258.
- Barr, P. J., 1991: Mammalian subtilisins: The long-thought dibasic processing endoprotease. *Cell*, vol. 66, p. 1–3.
- Belcher, A. M., Wu, X. H., Christensen, R. J., Hansma, P. K., Stucky, G. D. and Morse, D. E., 1996: Control of crystal phase switching and orientation by soluble mollusc-shell proteins. *Nature*, vol. 381, p. 56–58.
- Bengtson, S., 1994: The advent of animal skeletons. *In* Bengtson, S. *ed., Early Life on Earth: Nobel Symposium No. 84*, p. 412–425. Columbia University Press, New York.
- Beniash, E., Addadi, L. and Weiner, S., 1999: Cellular control over spicule formation in sea urchin embryos: A structural approach. *Journal of Structural Biology*, vol. 125, p. 50–62.
- Beniash, E., Aizenberg, J., Addadi, L. and Weiner, S., 1997: Amorphous calcium carbonate transforms into calcite during sea-urchin larval spicule growth. *Proceedings of the Royal Society of London. Series B*, vol. 264, p. 461–465.
- Benson, S., Sucov, H., Stephens, L., Davidson, E. and Wilt, F., 1987: A lineage-specific gene encoding a major matrix protein of the sea urchin embryo spicule. I. Authentication of the cloned gene and its developmental expression. *Devel*opmental Biology, vol. 120, p. 499–506.
- Bertrand, J. A., Pignol, D., Bernard, J. P., Verdier, J. M., Dagorn, J. C. and Fontecilla-Camps, J. C., 1996: Crystal structure of human lithostathine, the pancreatic inhibitor of stone formation. *The EMBO Journal*, vol. 15, p. 2678– 2684.
- Beverdam, A., Brouwer, A., Reijnen, M., Korving, J. and Meijlink, F., 2001: Severe nasal clefting and abnormal embryonic apoptosis in Alx3/Alx4 double mutant mice. *Development*, vol. 128, p. 3975–3986.
- Beverdam, A. and Meijlink, F., 2001: Expression patterns of group-I aristaless-related genes during craniofacial and limb development. *Mechanisms of Development*, vol. 107, p. 163–167.
- Bhatia, P. K. and Mukhopadhyay, A., 1999: Protein glycosylation: implication for *in vivo* functions and therapeutic applications. *Advances in Biochemica Engineering/ Biotechnology*, vol. 64, p. 155–201.
- Blank, S., Arnoldi, M., Khoshnavaz, S., Treccani, L., Kuntz, M., Mann, K., Grathwohl, G. and Fritz, M., 2003: The na-

cre protein perlucin nucleates growth of calcium carbonate crystals. *Journal of Microscopy*, vol. 212, p. 280–291.

- Bobbio, A., 1972: The first endosseous alloplastic implant in the history of man. *Bulletin of the History of Dentistry*, vol. 20, p. 1–6.
- Borbas, J. E., Wheeler, A. P. and Sikes, C. S., 1991: Molluscan shell matrix phosphoproteins: correlation of degree of phosphorylation to shell mineral microstructure and to in vitro regulation of mineralization. *The Journal of Experimental Zoology*, vol. 258, p. 1–13.
- Butler, W. T., Bhown, M., Dimuzio, M. T. and Linde, A., 1981: Noncollagenous proteins of dentin. Isolation and partial characterization of rat dentin proteins and proteoglycans using a three-step preparative method. *Collagen and Related Research: Clinical and Experimental*, vol. 1, p. 187–199.
- Campbell, P. G. and Andress, D. L., 1997: Insulin-like growth factor (IGF)-binding protein-5-(201–218) region regulates hydroxyapatite and IGF-I binding. *American Journal of Physiology: Endocrinology & Metabolism*, vol. 273, p. E1005–E1013.
- Campbell, K. P., MacLennan, D. H. and Jorgensen, A. O., 1983: Staining of the Ca²⁺-binding proteins, calsequestrin, calmodulin, troponin C, and S-100, with the cationic carbocyanine dye "Stains-all". *The Journal of Biological Chemistry*, vol. 258, p. 11267–11273.
- Cha, J. N., Shimizu, K., Zhou, Y., Christiansen, S. C., Chmelka, B. F., Stucky, G. D. and Morse, D. E., 1999: Silicatein filaments and subunits from a marine sponge direct the polymerization of silica and silicones in vitro. Proceedings of the National Academy of Sciences of the United States of America, vol. 96, p. 361–365.
- Cheers, M. S. and Ettensohn, C. A., 2005: P16 is an essential regulator of skeletogenesis in the sea urchin embryo. *Developmental Biology*, vol. 283, p. 384–396.
- Chen, D., Zhao, M. and Mundy, G. R., 2004: Bone morphogenetic proteins. *Growth Factors*, vol. 22, p. 233–241.
- Chen, Y., Bhushan, A. and Vale, W., 1997: Smad8 mediates the signaling of the receptor serine kinase. Proceedings of the National Academy of Sciences of the United States of America, vol. 94, p. 12938–12943.
- Cho, W. J., Partin, S. J. and Lennarz, J. W., 1996: A technique for detecting matrix proteins in the crystalline spicule of the sea urchin embryo. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, p. 1282–1286.
- Cohen, A. L. and McConnaughey, T. A., 2003: Geochemical perspectives on coral mineralization. In Dove, P. M., De Yoreo, J. J. and Weiner, S. eds., Biomineralization: Reviews in Mineralogy and Geochemistry, vol. 54, p. 217– 247. Mineralogical Society of America and the Geochemical Society, Washington.
- Cohen, B. R., 2005: Not armour, but biomechanics, ecological opportunity and increased fecundity as keys to the origin and expansion of the mineralized benthic metazoan fauna. *Biological Journal of the Linnean Society*, vol. 85, p. 483– 490.
- Conway Morris, S., 1998: *The Crucible of Creation: The Burgess Shale and the Rise of Animals*, 324 p. Oxford University Press, Oxford.
- Crenshaw, M.A., 1972: The soluble matrix from *Mercenaria mercenaria* shell. *Biomineralisation*, vol. 6, p. 6–11.
- Crenshaw, M. A., 1982: Mechanisms of normal biological mineralization of calcium carbonates. *In* Nancollas, G. H. *ed.*,

Biological Mineralization and Demineralization, p. 243–257. Springer-Verlag, Berlin, Heidelberg, New York.

- Daluiski, A., Engstrand, T., Bahamonde, M. E., Gamer, L. W., Agius, E., Stevenson, S. L., Cox, K., Rosen, V. and Lyons, K. M., 2001: Bone morphogenetic protein-3 is a negative regulator of bone density. *Nature Genetics*, vol. 27, p. 84– 88.
- De Caro, A. M., Bonicel, J. J., Rouimi, P., De Caro, J. D., Sarles, H. and Rovery, M., 1987: Complete amino acid sequence of an immunoreactive form of human pancreatic stone protein isolated from pancreatic juice. *European Journal of Biochemistry*, vol. 168, p. 201–207.
- De Crombrugghe, B., Lefebvre, V. and Nakashima, K., 2001: Regulatory mechanisms in the pathways of cartilage and bone formation. *Current Opinion in Cell Biology*, vol. 13, p. 721–727.
- Delattre, O., Catonne, Y., Berland, S., Borzeix, S. and Lopez, E., 1997: Use of mother of pearl as a bone substitute— Experimental study in sheep. *European Journal of Orthopaedic Surgery & Traumatology*, vol. 7, p. 143–147.
- Drickamer, K., 1999: C-type lectin-like domains. *Current Opinion in Structural Biology*, vol. 9, p. 585–590.
- Dupoirieux, L., Costes, V., Jammet, P. and Souyris, F., 1994: Experimental study on demineralized bone matrix (DBM) and coral as bone graft substitutes in maxillofacial surgery. *International Journal of Oral and Maxillofacial Surgery*, vol. 23, p. 395–398.
- Endo, H., Persson, P. and and Watanabe, T., 2000: Molecular cloning of the crustacean *DD4* cDNA encoding a Ca²⁺-binding protein. *Biochemical and Biophysical Research Communications*, vol. 276, p. 286–291.
- Endo, H., Takagi, Y., Ozaki, N., Kogure, T. and Watanabe, T., 2004: A crustacean Ca²⁺-binding protein with a glutamaterich sequence promotes CaCO₃ crystallization. *Biochemical Journal*, vol. 384, p. 159–167.
- Ettensohn, C. A., Illies, M. R., Oliveri, P. and De Jong, D. L., 2003: Alx1, a member of the Cart1/Alx3/Alx4 subfamily of Paired-class homeodomain proteins, is an essential component of the gene network controlling skeletogenic fate specification in the sea urchin embryo. *Development*, vol. 130, p. 2917–2928.
- Exposito, J.-Y., Le Guellec, D., Lu, Q. and Garrone, R., 1991: Short chain collagens in sponges are encoded by a family of closely related genes. *The Journal of Biological Chemistry*, vol. 266, p. 21923–21928.
- Falini, G., Albeck, S., Weiner, S. and Addadi, L., 1996: Control of aragonite or calcite polymorphism by mollusc shell macromolecules. *Science*, vol. 271, p. 67–69.
- Folders, J., Tommassen, J., van Loon, L.C. and Bitter, W., 2000: Identification of a chitin-binding protein secreted by *Pseudomonas aeruginosa. Journal of Bacteriology*, vol. 182, p. 1257–1263.
- Forbes, E.G., Cronshaw, A.D., MacBeath, J. R.E. and Hulmes, D.J.S., 1994: Tyrosine-rich acidic matrix protein (TRAMP) is a tyrosine-sulphated and widely distributed protein of the extracellular matrix. *FEBS Letters*, vol. 351, p. 433–436.
- Fujii, N., Minetti, C. A. S. A., Nakhasi, H. L., Chen, S.-W., Barbehenn, E., Nunes, P. H. and Nguyen, N. Y., 1992: Isolation, cDNA cloning, and characterization of an 18-kDa hemagglutinin and amebocyte aggregation factor from *Limulus polyphemus. The Journal of Biological Chemistry*, vol. 267, p. 22452–22459.

- Fukuda, I., Ooki, S., Fujita, T., Murayama, E., Nagasawa, H., Isa, Y. and Watanabe, T., 2003: Molecular cloning of a cDNA encoding a soluble protein in the coral exoskeleton. *Biochemical and Biophysical Research Communications*, vol. 304, p. 11–17.
- Funayama, N., Nakatsukasa, M., Kuraku, S., Takechi, K., Dohi, M., Iwabe, N., Miyata, T. and Agata, K., 2005: Isolation of *Ef silicatein* and *Ef lectin* as molecular markers for sclerocytes and cells involved in innate immunity in the freshwater sponge *Ephydatia fluviatilis*. *Zoological Science*, vol. 22, p. 1113–1122.
- Gabius, S., Hellmann, K.-P., Hellmann, T., Brinck, U. and Gabius, H.-J., 1989: Neoglycoenzymes: A versatile tool for lectin detection in solid-phase assays and glycohistochemistry. *Analytical Biochemistry*, vol, 182, p. 447–451.
- George, N. C., Killianb, C. E. and Wilt, F. H., 1991: Characterization and expression of a gene encoding a 30.6-kDa *Strongylocentrotus purpuratus* spicule matrix protein. *Developmental Biology*, vol. 147, p. 334–342.
- Gotliv, B.-A., Addadi, L. and Weiner, S., 2003: Mollusc shell acidic proteins: in search of individual functions. *Chem-BioChem*, vol. 4, p. 522–529.
- Gotliv, B.-A., Kessler, N., Sumerel, J. L., Morse, D. E., Tuross, N., Addadi, L. and Weiner, S., 2005: Asprich: A novel aspartic acid-rich protein family from the prismatic shell matrix of the bivalve *Atrina rigida*. *ChemBioChem*, vol. 6, p. 304–314.
- Guerette, P. A., Ginzinger, D. G., Weber, B. H. F. and Gosline, J. M., 1996: Silk properties determined by gland-specific expression of a spider fibroin gene family. *Science*, vol. 272, p. 112–115.
- Guss, K. A. and Ettensohn, C. A., 1997: Skeletal morphogenesis in the sea urchin embryo: regulation of primary mesenchyme gene expression and skeletal rod growth by ectoderm-derived cues. *Development*, vol. 124, p. 1899–1908.
- Hare, P. E., 1963: Amino acids in the proteins from aragonite and calcite in the shells of *Mytilus californianus*. *Science*, vol. 139, p. 216–217.
- Harkey, M. A., Klueg, K., Sheppard, P. and Raff, R. A., 1995: Structure, expression, and extracellular targeting of PM27, a skeletal protein associated specifically with growth of the sea urchin larval spicule. *Developmental Biology*, vol. 168, p. 549–566.
- Harkey, M. A., Whiteley, H. R. and Whiteley, A. H., 1988: Coordinate accumulation of five transcripts in the primary mesenchyme during skeletogenesis in the sea urchin embryo. *Developmental Biology*, vol. 125, p. 381–395.
- Hatfield, D. L. and Gladyshev, V. N., 2002: How selenium has altered our understanding of the genetic code. *Molecular Biology of the Cell*, vol. 22, p. 3565–76.
- Hecker, A., Testenière, O., Marin, F. and Luquet, G., 2003: Phosphorylation of serine residues is fundamental for the calcium-binding ability of Orchestin, a soluble matrix protein from crustacean calcium storage structures. *FEBS Letters*, vol. 535, p. 49–54.
- Hecker, A., Quennedey, B., Testenière, O., Quennedey, A., Graf, F. and Luquet, G., 2004: Orchestin, a calciumbinding phosphoprotein, is a matrix component of two successive transitory calcified biomineralizations cyclically elaborated by a terrestrial crustacean. *Journal of Structural Biology*, vol. 146, p. 310–324.
- Hincke, M.T., Gautron, J., Tsang, C.P., McKee, M.D. and Nys, Y., 1999: Molecular cloning and ultrastructural local-

ization of the core protein of an eggshell matrix proteoglycan, ovocleidin-116. *The Journal of Biological Chemistry*, vol. 274, p. 32915–32923.

- Hincke, M. T., Tsang, C. P., Courtney, M., Hill, V. and Narbaitz, R., 1995: Purification and immunochemistry of a soluble matrix protein of the chicken eggshell (ovocleidin 17). *Calcified Tissue International*, vol. 56, p. 578–583.
- Hinman, V.F., O'Brien, E.K., Richards, G.S. and Degnan, B.M., 2003: Expression of anterior Hox genes during larval development of the gastropod *Haliotis asinina*. Evolution and Development, vol. 5, p. 508–521.
- Hoodless, P. A., Haerry, T., Abdollah, S., Stapleton, M., O'Connor, M. B., Attisano, L. and Wrana, J. L., 1996: MADR1, a MAD-related protein that functions in BMP2 signaling pathways. *Cell*, vol. 85, p. 489–500.
- Huttner, W. B., 1987: Protein tyrosine sulfation. *Trends in Bio*chemical Sciences, vol. 12, p. 361–363.
- Ikeya, T., Persson, P., Kono, M. and Watanabe, T., 2001: The DD5 gene of the decapod crustacean Penaeus japonicus encodes a putative exoskeletal protein with a novel tandem repeat structure. Comparative Biochemistry and Physiology Part B, vol. 128, p. 379–388.
- Illies, M. R., Peeler, M. T., Dechtiaruk, A. M. and Ettensohn, C. A., 2002: Identification and developmental expression of new biomineralization proteins in the sea urchin *Strongylocentrotus purpuratus. Development Genes and Evolution*, vol. 212, p. 419–431.
- Inoue, H., Ohira, T., Ozaki, N. and Nagasawa, H., 2003: Cloning and expression of a cDNA encoding a matrix peptide associated with calcification in the exoskeleton of the crayfish. *Comparative Biochemistry and Physiology Part B*, vol. 136, p. 755–765.
- Inoue, H., Ohira, T., Ozaki, N. and Nagasawa, H., 2004: A novel calcium-binding peptide from the cuticle of the crayfish, *Procambarus clarkii. Biochemical and Biophysical Re*search Communications, vol. 318, p. 649–654.
- Inoue, H., Ozaki, N. and Nagasawa, H., 2001: Purification and structural determination of a phosphorylated peptide with anti-calcification and chitin-binding activities in the exoskeleton of the crayfish, *Procambarus clarkii. Bioscience Biotechnology and Biochemistry*, vol. 65, p. 1840–1848.
- Isa, Y. and Okazaki, M., 1987: Some observations on the Ca2+-binding phospholipid from scleractinian coral skeletons. *Comparative Biochemistry and Physiology Part B*, vol. 87, p. 507–512.
- Ishii, K., Tsutsui, N., Watanabe, T., Yanagisawa, T. and Nagasawa, H., 1998: Solubilization and chemical characterization of an insoluble matrix protein in the gastroliths of a crayfish, *Procambarus clarkii. Bioscience Biotechnology* and Biochemistry, vol. 62, p. 291–296.
- Ishii, K., Yanagisawa, T. and Nagasawa, H., 1996: Characterization of a matrix protein in the gastroliths of the crayfish *Procambarus clarkii. Bioscience Biotechnology and Biochemistry*, vol. 60, p. 1479–1482.
- Jacobs, D. K., Wray, C. G., Wedeen, C. J., Kostriken, R., De-Salle, R., Staton, J. L., Gates, R. D. and Lindberg, D. R., 2000: Molluscan engrailed expression, serial organization, and shell evolution. *Evolution and Development*, vol. 2, p. 340–347.
- Jacobsen, S. L., Andersen, S. O. and Højrup, P., 1994: Amino acid sequence determination of a protein purified from the shell of the shrimp, *Pandalus borealis. Comparative Biochemistry and Physiology Part B*, vol. 109, p. 209–217.

- Jochum, W., David, J. P., Elliott, C., Wutz, A., Plenk, H. Jr., Matsuo, K. and Wagner, E. F., 2000: Increased bone formation and osteosclerosis in mice overexpressing the transcription factor Fra-1. *Nature Medicine*, vol. 6, p. 980–984. Erratum, 2000: *Nature Medicine*, vol. 6, p. 1412.
- Katoh-Fukui, Y., Noce, T., Ueda, T., Fujiwara, Y., Hashimoto, N., Higashinakagawa, T., Killian, C. E., Livingston, B. T., Wilt, F. H., Benson, S. C., Sucov, H. M. and Davidson, E. H., 1991: The corrected structure of the SM50 spicule matrix protein of *Strongylocentrotus purpuratus*. *Developmental Biology*, vol. 145, p. 201–202.
- Katoh-Fukui, Y., Noce, T., Ueda, T., Fujiwara, Y., Hashimoto, N., Tanaka, S. and Higashinakagawa, T., 1992: Isolation and characterization of cDNA encoding a spicule matrix protein in *Hemicentrotus pulcherrimus* micromeres. *The International Journal of Developmental Biology*, vol. 36, p. 353–361.
- Killian, C.E. and Wilt, F.H., 1989: The accumulation and translation of a spicule matrix protein mRNA during sea urchin embryo development. *Developmental Biology*, vol. 133, p. 148–156.
- Killian, C. E. and Wilt, F. H., 1996: Characterization of the proteins comprising the integral matrix of *Strongylocentrotus purpuratus* embryonic spicules. *The Journal of Biological Chemistry*, vol. 271, p. 9150–9159.
- Kirschvink, J. L. and Hagadorn, J. W., 2000: A grand unified theory of biomineralization. In Bäuerlein, E. ed., The Biomineralization of Nano- and Micro-structures. P. 139–150. Wiley-VCH Verlag, Weinheim, Germany.
- Komori, T., Yagi, H., Nomura, S., Yamaguchi, A., Sasaki, K., Deguchi, K., Shimizu, Y., Bronson, R. T., Gao, Y. H., Inada, M., Sato, M., Okamoto, R., Kitamura, Y., Yoshiki, S. and Kishimoto, T., 1997: Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell*, vol. 89, p. 755– 764.
- Kono, M., Hayashi, N. and Samata, T., 2000: Molecular mechanism of the nacreous layer formation in *Pinctada maxima*. *Biochemical and Biophysical Research Communications*, vol. 269, p. 213–218.
- Korsmeyer, S. J., 1992: Bcl-2 initiates a new category of oncogenes: regulators of cell death. *Blood*, vol. 80, p. 879–886.
- Kragh, M., Molbak, L. and Andersen, S.O., 1997: Cuticular proteins from the lobster, *Homarus americanus. Comparative Biochemistry and Physiology Part B*, vol. 118, p. 147– 154.
- Krasko, A., Lorenz, B., Batel, R., Schroder, H. C., Muller, I. M. and Muller, W. E., 2000: Expression of silicatein and collagen genes in the marine sponge *Suberites domuncula* is controlled by silicate and myotrophin. *European Journal* of Biochemistry, vol. 267, p. 4878–87.
- Kruse, M., Leys, S. P., Muller, I. M. and Muller, W. E., 1998: Phylogenetic position of the Hexactinellida within the phylum Porifera based on the amino acid sequence of the protein kinase C from *Rhabdocalyptus dawsoni. Journal of Molecular Evolution*, vol. 46, p. 721–728.
- Kurokawa, D., Kitajima, T., Mitsunaga-Nakatsubo, K., Amemiya, S., Shimada, H. and Akasaka, K., 1999: HpEts, an ets-related transcription factor implicated in primary mesenchyme cell differentiation in the sea urchin embryo. *Mechanisms of Development*, vol. 80, p. 41–52.
- Lakshminarayanan, R., Kini, R. M. and Valiyaveettil, S., 2002: Investigation of the role of ansocalcin in the biominerali-

zation in goose eggshell matrix. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, p. 5155–5159.

- Lakshminarayanan, R., Valiyaveettil, S., Rao, V.S. and Kini, R.M., 2003: Purification, characterization, and in vitro mineralization studies of a novel goose eggshell matrix protein, ansocalcin. *The Journal of Biological Chemistry*, vol. 278, p. 2928–2936.
- Lamghari, M., Almeida, M.J., Berland, S., Huet, H., Laurent, A., Milet, C. and Lopez, E., 1999: Stimulation of bone marrow cells and bone formation by nacre: *in vivo* and in vitro studies. *Bone*, vol. 25, suppl. 1, 91S–94S.
- Lamghari, M., Antonietti, P., Berland, S., Laurent, A. and Lopez, E., 2001: Arthrodesis of lumbar spine transverse processes using nacre in rabbit. *Journal of Bone and Mineral Research*, vol. 16, p. 2232–2237.
- Lamghari, M., Berland, S., Laurent, A., Huet, H. and Lopez, E., 2001: Bone reactions to nacre injected percutaneously into the vertebrae of sheep. *Biomaterials*, vol. 22, p. 555– 562.
- Lee, Y.-H., Britten, R.J. and Davidson, E.H., 1999: SM37, a skeletogenic gene of the sea urchin embryo linked to the SM50 gene. Development, Growth and Differentiation, vol. 41, p. 303–312.
- Liao, H., Brandsten, C., Lundmark, C., Wurtz, T. and Li, J., 1997: Responses of bone to titania-hydroxyapatite composite and nacreous implants: a preliminary comparison by in situ hybridization. *Journal of Materials Science: Materials in Medicine*, vol. 8, p. 823–827.
- Liao, H., Mutvei, H., Hammarstrom, L., Wurtz, T. and Li, J., 2002: Tissue responses to nacreous implants in rat femur: an in situ hybridization and histochemical study. *Biomaterials*, vol. 23, p. 2693–2701.
- Liao, H., Mutvei, H., Sjostrom, M., Hammarstrom, L. and Li, J., 2000: Tissue responses to natural aragonite (*Margaritifera* shell) implants *in vivo*. *Biomaterials*, vol. 21, p. 457– 468.
- Liff, M. I. and Zimmerman, M. N., 1998: NMR study of crosslinking by oxidation of four-cysteine polypeptide models of the elastic network phase of wool fibre. *Polymer International*, vol. 47, p. 375–385.
- Livingston, B. T., Shaw, R., Bailey, A. and Wilt, F., 1991: Characterization of a cDNA encoding a protein involved in formation of the skeleton during development of the sea urchin Lytechinus pictus. Developmental Biology, vol. 148, p. 473–480.
- Lopez, E., Vidal, B., Berland, S., Camprasse, S., Camprasse, G. and Silve, C., 1992: Demonstration of the capacity of nacre to induce bone formation by human osteoblasts maintained in vitro. *Tissue and Cell*, vol. 24, p. 667–679.
- Lowe, C. J. and Wray, G. A., 1997: Radical alterations in the roles of homeobox genes during echinoderm evolution. *Nature*, vol. 389, p. 718–721.
- Lowenstam, H. A., 1981: Minerals formed by organisms. *Science*, vol. 211, p. 1126–1131.
- Lowenstam, H. A. and Weiner, S., 1989: On biomineralization, 324p. Oxford University Press, Oxford.
- Luquet, G. and Marin, F., 2004: Biomineralisations in crustaceans: storage strategies. *Comptes Rendus Palevol*, vol. 3, p. 515–534.
- Luyten, F. P., Cunningham, N. S., Ma, S., Muthukumaran, N., Hammonds, R. G., Nevins, W. B., Woods, W. I. and Reddi, A. H., 1989; Purification and partial amino acid sequence

of osteogenin, a protein initiating bone differentiation. *The Journal of Biological Chemistry*, vol. 264, p. 13377–13380.

- MacBeath, J. R., Shackleton, D. R. and Hulmes, D. J., 1993: Tyrosine-rich acidic matrix protein (TRAMP) accelerates collagen fibril formation *in vitro*. *The Journal of Biological Chemistry*, vol. 268, p. 19826–19832.
- Makabe, K. W., Kirchhamer, C. V., Britten, R. J., and Davidson, E. H., 1995: *Cis*-regulatory control of the *SM50* gene, an early marker of skeletogenic lineage specification in the sea urchin embryo. *Development*, vol. 121, p. 1957–1970.
- Mann, K., 2004: Identification of the major proteins of the organic matrix of emu (*Dromaius novaehollandiae*) and rhea (*Rhea americana*) eggshell calcified layer. *British Poultry Science*, vol. 45, p. 483–490.
- Mann, K. and Siedler, F., 1999: The amino acid sequence of ovocleidin 17, a major protein of the avian eggshell calcified layer. *Biochemistry and Molecular Biology International*, vol. 47, p. 997–1007.
- Mann, K. and Siedler, F., 2004: Ostrich (*Struthio camelus*) eggshell matrix contains two different C-type lectin-like proteins. Isolation, amino acid sequence, and posttranslational modifications. *Biochimica et Biophysica Acta— Proteins and Proteomics*, vol. 1696, p. 41–50.
- Mann, K., Weiss, I. M., André, S., Gabius, H. J. and Fritz, M., 2000: The amino acid sequence of the abalone (*Haliotis laevigata*) nacre protein perlucin. *European Journal of Biochemistry*, vol. 267, p. 5257–5264.
- Marin, F., Corstjens, P., De Gaulejac, B., De Vrind-De Jong, E., and Westbroek, P., 2000: 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). *The Journal of Biological Chemistry*, vol. 275, p. 20667–20675.
- Marin, F. and Luquet, G., 2004: Molluscan shell proteins. Comptes Rendus Palevol, vol. 3, p. 469-492.
- Marin, F., Smith, M., Isa, Y., Muyzer, G. and Westbroek, P., 1996: Skeletal matrices, muci, and the origin of invertebrate calcification. *Proceedings of the National Academy* of Sciences of the United States of America, vol. 93, p. 1554–1559.
- Maruyama, K., Mikawa, T. and Ebashi, S., 1984: Detection of calcium binding proteins by ⁴⁵Ca autoradiography on nitrocellulose membrane after sodium dodecyl sulfate gel electrophoresis. *The Journal of Biochemistry*, vol. 95, p. 511–519.
- Marxen, J. C. and Becker, W., 1997: The organic shell matrix of the freshwater snail *Biomphalaria glabrata*. *Comparative Biochemistry and Physiology Part B*, vol. 118, p. 23–33.
- Marxen, J. C., Hammer, M., Gehrke, T. and Becker, W., 1998: Carbohydrates of the organic shell matrix and the shellforming tissue of the snail *Biomphalaria glabrata* (Say). *The Biological Bulletin*, vol. 194, p. 231–240.
- Marxen, J. C., Nimtz, M., Becker, W. and Mann, K., 2003: The major soluble 19.6 kDa protein of the freshwater snail *Biomphalaria glabrata* is an N-glycosylated dermatopontin. *Biochimica et Biophysica Acta—Proteins and Proteomics*, vol. 1650, p. 92–98.
- Matsushiro, A. and Miyashita, T., 2004: Evolution of hardtissue mineralization: comparison of the inner skeletal system and the outer shell system. *Journal of Bone and Mineral Metabolism*, vol. 22, p. 163–169.
- Matsushiro, A., Miyashita, T., Miyamoto, H., Morimoto, K., Tonomura, B., Tanaka, A. and Sato, K., 2003: Presence

of protein complex is prerequisite for aragonite crystallization in the nacreous layer. *Marine Biotechnology*, vol. 5, p. 37–44.

- Mavrogiannis, L. A., Antonopoulou, I., Baxova, A., Kutilek, S., Kim, C. A., Sugayama, S. M., Salamanca, A., Wall, S. A., Morriss-Kay, G. M. and Wilkie, A. O., 2001: Haploinsufficiency of the human homeobox gene *ALX4* causes skull ossification defects. *Nature Genetics*, vol. 27, p. 17–18.
- Michenfelder, M., Fu, G., Lawrence, C., Weaver, J. C., Wustman, B. A., Taranto, L., Evans, J. S. and Morse, D. E., 2003: Characterization of two molluscan crystalmodulating biomineralization proteins and identification of putative mineral binding domains. *Biopolymers*, vol. 70, p. 522–530. Erratum, 2004: *Biopolymers*, vol. 73, p. 291.
- Miles, M. J., Carr, H. J., McMaster, T. C., l'Anson, K. J., Belton, P. S., Morris, V. J., Field, J. M., Shewry, P. R. and Tatham, A. S., 1991: Scanning tunneling microscopy of a wheat seed storage protein reveals details of an unusual supersecondary structure. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, p. 68–71.
- Milet, C., Berland, S., Lamghari, M., Mouries, L., Jolly, C., Borzeix, S., Doumenc, D. and Lopez, E., 2004: Conservation of signal molecules involved in biomineralisation control in calcifying matrices of bone and shell. *Comptes Rendus Palevol*, vol. 3, p. 493–501.
- Mitterer, M. R., 1978: Amino acid composition and metal binding capability of the skeletal protein of corals. *Bulletin of Marine Science*, vol. 28, p. 173–180.
- Miyamoto, H., Miyashita, T., Okushima, M., Nakano, S., Morita, T. and Matsushiro, A., 1996: A carbonic anhydrase from the nacreous layer in oyster pearls. *Proceedings* of the National Academy of Sciences of the United States of America, vol. 93, p. 9657–9660.
- Miyamoto, H., Yano, M. and Miyashita, T., 2003: Similarities in the structure of nacrein, the shell-matrix protein, in a bivalve and a gastropod. *Journal of Molluscan Studies*, vol. 69, p. 87–89.
- Miyashita, T., Takagi, R., Okushima, M., Nakano, S., Miyamoto, H., Nishikawa, E. and Matsushiro, A., 2000: Complementary DNA cloning and characterization of pearlin, a new class of matrix protein in the nacreous layer of oyster pearls. *Marine Biotechnology*, vol. 2, p. 409–418.
- Moshel, S. M., Levine, M. and Collier, J. R., 1998: Shell differentiation and *engrailed* expression in the *Ilyanassa* embryo. *Development Genes and Evolution*, vol. 208, p. 135– 141.
- Mouriès, L. P., Almeida, M. J., Milet, C., Berland, S. and Lopez, E., 2002: Bioactivity of nacre water-soluble organic matrix from the bivalve mollusc *Pinctada maxima* in three mammalian cell types: fibroblasts, bone marrow stromal cells and osteoblasts. *Comparative Biochemistry and Physiology Part B*, vol. 132, p. 217–229.
- Müller, W. E. G., Borejko, A., Brandt, D., Osinga, R., Ushijima, Hamer, H. B., Krasko, A., Xupeng, C., Müller, I. M. and Schröder, H. C., 2005: Selenium affects biosilica formation in the demosponge *Suberites domuncula*. Effect on gene expression and spicule formation. *The FEBS Journal*, vol. 272, p. 3838–3852.
- Nakashima, K., Zhou, X., Kunkel, G., Zhang, Z., Deng, J. M., Behringer, R. R. and de Crombrugghe, B., 2002: The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. *Cell*, vol. 108, p. 17–29.

- Neame, P. J., Choi, H. U. and Rosenberg, L. C., 1989; The isolation and primary structure of a 22-kDa extracellular matrix protein from bovine skin. *The Journal of Biological Chemistry*, vol. 264, p. 5474–5479.
- Neame, P.J., Young, C.N. and Treep, J.T., 1992: Primary structure of a protein isolated from reef shark (*Carcharhinus springeri*) cartilage that is similar to the mammalian Ctype lectin homolog, tetranectin. *Protein Science*, vol. 1, p. 161–168.
- Nederbragt, A. J., van Loon, A. E. and Dictus, W. J., 2002: Expression of *Patella vulgate* orthologs of engrailed and dpp-BMP2/4 in adjacent domains during molluscan shell development suggests a conserved compartment. *Developmental Biology*, vol. 246, p. 341–355.
- Nedin, C., 1999: Anomalocaris predation on nonmineralized and mineralized trilobites. Geology, vol. 27, p. 987– 990.
- Nishimura, R., Kato, Y., Chen, D., Harris, S. E., Mundy, G. R. and Yoneda, T., 1998: Smad5 and DPC4 are key molecules in mediating BMP-2-induced osteoblastic differentiation of the pluripotent mesenchymal precursor cell line C2C12. *The Journal of Biological Chemistry*, vol. 273, p. 1872– 1879.
- Nousiainen, M., Rafn, K., Skou, L., Roepstorff, P. and Andersen, S. O., 1998: Characterization of exoskeletal proteins from the American lobster, *Homarus americanus. Comparative Biochemistry and Physiology Part B*, vol. 119, p. 189–199.
- Okamoto, O., Fujiwara, S., Abe, M. and Sato, Y., 1999: Dermatopontin interacts with transforming growth factor β and enhances its biological activity. *Biochemical Journal*, vol. 337, p. 537–541.
- Okamoto, O., Suzuki, Y., Kimura, S. and Shinkai, H., 1996; Extracellular matrix 22-kDa protein interacts with decorin core protein and is expressed in cutaneous fibrosis. *The Journal of Biochemistry*, vol. 119, p. 106–114.
- Okazaki, K. and Inoué, S., 1976: Crystal property of the sea larval sea urchin spicule. *Development Growth & Differentiation*, vol. 18, p. 413–434.
- Pause, A. and Sonenberg, N., 1992: Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *The EMBO Journal*, vol. 11, p. 2643–2645.
- Peled-Kamar, M., Hamilton, P. and Wilt, F. H., 2002: Spicule matrix protein LSM34 is essential for biomineralization of the sea urchin spicule. *Experimental Cell Research*, vol. 272, p. 56–61.
- Perry, C. C., 2003: Silicification: The processes by which organisms capture and mineralize silica. *In* Dove, P. M., De Yoreo, J. J. and Weiner, S. *eds.*, *Biomineralization: Reviews in Mineralogy and Geochemistry*, vol. 54, p. 291– 327. Mineralogical Society of America and the Geochemical Society, Washington.
- Pozzolini, M., Sturla, L., Cerrano, C., Bavestrello, G., Camardella, L., Parodi, A. M., Raheli, F., Benatti, U., Muller, W. E. and Giovine, M., 2004: Molecular cloning of silicatein gene from marine sponge *Petrosia ficiformis* (Porifera, Demospongiae) and development of primmorphs as a model for biosilicification studies. *Marine Biotechnology*, vol. 6, p. 594–603.
- Puverel, S., Tambutté, E., Pereira-Mouriès, L., Zoccola, D., Allemand, D. and Tambutté, S., 2005: Soluble organic matrix of two scleractinian corals: Partial and comparative analy-

sis. Comparative Biochemistry and Physiology Part B, vol. 141, p. 480-487.

- Qu, S., Tucker, S. C., Zhao, Q., de Crombrugghe, B. and Wisdom, R., 1999: Physical and genetic interactions between *Alx4* and *Cart1. Development*, vol. 126, p. 359–369.
- Qu, S., Li, L. and Wisdom, R., 1997: Alx-4: cDNA cloning and characterization of a novel paired-type homeodomain protein. *Gene*, vol. 203, p. 217–223.
- Rahman, M. A. and Isa, Y., 2005: Characterization of proteins from the matrix of spicules from the alcyonarian, *Lobophytum crassum. Journal of Experimental Marine Biology* and Ecology, vol. 321, p. 71–82.
- Reddy, S. P., Rasmussen, W. G. and Baseman, J. B., 1995: Molecular cloning and characterization of an adherencerelated operon of *Mycoplasma genitalium*. *Journal of Bacteriology*, vol. 177, p. 5943–5951.
- Richardson, W., Kitajima, Wilt, T. F. and Benson, S., 1989: Expression of an embryonic spicule matrix gene in calcified tissues of adult sea urchins. *Developmental Biology*, vol. 132, p. 266–269.
- Robach, J. S., Stock, S. R. and Veis, A., 2005: Transmission electron microscopy characterization of macromolecular domain cavities and microstructure of single-crystal calcite tooth plates of the sea urchin *Lytechinus variegates*. *Journal of Structural Biology*, vol. 151, p. 18–29.
- Rousseau, M., Pereira-Mouriès, L., Almeida, M. J., Milet, C. and Lopez, E., 2003: The water-soluble matrix fraction from the nacre of *Pinctada maxima* produces earlier mineralization of MC3T3-E1 mouse pre-osteoblasts. *Comparative Biochemistry and Physiology Part B*, vol. 135, p. 1–7.
- Runnegar, B., 1984: Crystallography of the foliated calcite shell layers of bivalve molluscs. *Alcheringa*, vol. 8, p. 273–290.
- Runnegar, B. and Bengtson, S., 1990: Origin of hard parts— Early skeletal fossils. *In* Briggs, D.E.G. and Crowther, P.R. eds., *Palaeobiology: A synthesis*, p. 24–29. Blackwell Scientific Publications, Oxford.
- Sabatakos, G., Sims, N. A., Chen, J., Aoki, K., Kelz, M. B., Amling, M., Bouali, Y., Mukhopadhyay, K., Ford, K., Nestler, E. J. and Baron, R., 2000: Overexpression of Delta-FosB transcription factor(s) increases bone formation and inhibits adipogenesis. *Nature Medicine*, vol. 6, p. 985–990.
- Samata, T., Hayashi, N., Kono, M., Hasegawa, K., Horita, C. and Akera, S., 1999: A new matrix protein family related to the nacreous layer formation of *Pinctada fucata*. *FEBS Letters*, vol. 462, p. 225–229.
- Sarashina, I. and Endo, K., 1998: Primary structure of a soluble matrix protein of scallop shell: Implications for calcium carbonate biomineralization. *American Mineralogist*, vol. 83, p. 1510–1515.
- Sarashina, I. and Endo, K. 2001: The complete primary structure of Molluscan Shell Protein 1 (MSP-1), an acidic glycoprotein in the shell matrix of the scallop *Patinopecten yessoensis. Marine Biotechnology*, vol. 3, p. 362–369.
- Sarashina, I., Yamaguchi, H., Haga, T., Iijima, M., Chiba, S. and Endo, K., 2005: Molecular evolution and functionally important structures of molluscan Dermatopontin: Implications for the origins of molluscan shell matrix proteins. *Journal of Molecular Evolution*, in press.
- Schröder, H. C., Krasko, A., Batel, R., Skorokhod, A., Pahler, S., Kruse, M., Müller, I. M. and Müller, W. E., 2000: Stimulation of protein (collagen) synthesis in sponge cells by a cardiac myotrophin-related molecule from *Suberites domuncula. The FASEB Journal*, vol. 14, p. 2022–31.

- Schröder, H. C., Krasko, A., Le Pennec, G., Adell, T., Wiens, M., Hassanein, H., Müller, I. M. and Müller, W. E., 2003: Silicase, an enzyme which degrades biogenous amorphous silica: contribution to the metabolism of silica deposition in the demosponge *Suberites domuncula. Progress in Molecular and Subcellular Biology*, vol. 33, p. 249–268.
- Schröder, H. C., Perovic-Ottstadt, S., Grebenjuk, V. A., Engel, S., Müller, I. M. and Müller, W. E., 2005: Biosilica formation in spicules of the sponge *Suberites domuncula*: synchronous expression of a gene cluster. *Genomics*, vol. 85, p. 666–678.
- Schütze, J., Skorokhod, A., Müller, I. M. and Müller, M. E., 2001: Molecular evolution of the metazoan extracellular matrix: cloning and expression of structural proteins from the demosponges *Suberites domuncula* and *Geodia cydonium. Journal of Molecular Evolution*, vol. 53, p. 402–415.
- Seto, J., Zhang, Y., Hamilton, P. and Wilt, F., 2004: The localization of occluded matrix proteins in calcareous spicules of sea urchin larvae. *Journal of Structural Biology*, vol. 148, p. 123–130.
- Shen, D., Zhao, M. and Mundy, G. R., 2004: Bone morphogenetic proteins. *Growth Factors*, vol. 22, p. 233–241.
- Shen, X., Belcher, A. M., Hansma, P. K., Stucky, G. D. and Morse, D. E., 1997: Molecular cloning and characterization of Lustrin A, a matrix protein from shell and pearl nacre of *Haliotis rufescens*. *The Journal of Biological Chemistry*, vol. 272, p. 32472–32481.
- Shimizu, K., Cha, J., Stucky, G.D. and Morse, D.E., 1998: Silicatein α: Cathepsin L-like protein in sponge biosilica. Proceedings of the National Academy of Sciences of the United States of America, vol. 95, p. 6234–6238.
- Silve, C., Lopez, E., Vidal, B., Smith, D.C., Camprasse, S., Camprasse, G. and Couly, G., 1992: Nacre initiates biomineralization by human osteoblasts maintained in vitro. *Calcified Tissue International*, vol. 51, p. 363–369.
- Simkiss, K., 1977: Biomineralization and detoxification. Calcified Tissue Research, vol. 24, p. 199–200.
- Stadtman, T.C., 1996: Selenocysteine. Annual Review of Biochemistry, vol. 65, p. 83–100.
- Steinert, P. M., Mack, J. W., Korge, B. P. and Gan, S.-Q., 1991: Glycine loops in proteins: their occurrence in certain intermediate filament chains, loricrins and single-stranded RNA binding proteins. *International Journal of Biological Macromolecules*, vol. 13, p. 130–139.
- Steinert, P. M., Parry, D. A. D., Racoosin, E. L., Idler, W. W., Steven, A. C., Trus, B. L. and Roop, D. R., 1984: The complete cDNA and deduced amino acid sequence of a type II mouse epidermal keratin of 60,000 Da: Analysis of sequence differences between type I and type II keratins. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 81, p. 5709–5713.
- Sucov, H. M., Benson, S., Robinson, J. J., Britten, R. J., Wilt, F. and Davidson, E. H., 1987: A lineage-specific gene encoding a major matrix protein of the sea urchin embryo spicule. II. Structure of the gene and derived sequence of the protein. *Developmental Biology*, vol. 120, p. 507–519.
- Sudo, S., Fujikawa, T., Nagakura, T., Ohkubo, T., Sakagushi, K., Tanaka, M., Nakashima, K. and Takahashi, T., 1997: Structures of mollusc shell framework proteins. *Nature*, vol. 387, p. 563–564.
- Superti-Furga, A., Rocchi, M., Schäfer, B. W. and Gitzelmann, R., 1993: Complementary DNA sequence and chromosomal mapping of a human proteoglycan-binding cell-

adhesion protein (Dermatopontin). *Genomics*, vol. 17, p. 463–467.

- Suzuki, M., Murayama, E., Inoue, H., Ozaki, N., Tohse, H., Kogure, T. and Nagasawa, H., 2004: Characterization of Prismalin-14, a novel matrix protein from the prismatic layer of the Japanese pearl oyster (*Pinctada fucata*). *Biochemical Journal*, vol. 382, p. 205–213.
- Suzuki, K., Okamori, M., Katsuzaki, H., Komiya, T. and Imai, K., 2001: Isolation and characterization of the novel lipophilic protein, Pb CP-12.7, from the shell of the pink shrimp, *Pandalus borealis. Bioscience, Biotechnology, and Biochemistry*, vol. 65, p. 1038–1044.
- Takagi, Y., Ishii, K., Ozaki, N. and Nagasawa, H., 2000: Immunolocalization of gastrolith matrix protein (GAMP) in the gastroliths and exoskeleton of crayfish, *Procambarus clarkii. Zoological Science*, vol. 17, p. 179–184.
- Takeuchi, T. and Endo, K., 2005: Biphasic and dually coordinated expression of the genes encoding major shell matrix proteins in the pearl oyster *Pinctada fucata*. *Marine Biotechnology*, in press.
- ten Berge, D., Brouwer, A., El Bahi, S., Guénet, J.-L., Robert, B. and Meijlink, F., 1998: Mouse *Alx3*: an aristaless-like homeobox gene expressed during embryogenesis in ectomesenchyme and lateral plate mesoderm. *Developmental Biology*, vol. 199, p. 11–25.
- Testenière, O., Hecker, A., Le Gurun, S., Quennedey, B., Graf, F. and Luquet, G., 2002: Characterization and spatiotemporal expression of *orchestin*, a gene encoding an ecdysone-inducible protein from a crustacean organic matrix. *Biochemical Journal*, vol. 361, p. 327–335.
- Tsukamoto, D., Sarashina, I. and Endo, K., 2004: Structure and expression of an unusually acidic matrix protein of pearl oyster shells. *Biochemical and Biophysical Research Communications*, vol. 320, p. 1175–1180.
- Tsutsui, N., Ishi, K., Takagi, Y., Watanabe, T. and Nagasawa, H., 1999: Cloning and expression of a cDNA encoding an insoluble matrix protein in the gastroliths of a crayfish, *Procambarus clarkii. Zoological Science*, vol. 16, p. 619– 628.
- Urist, M. R., 1965: Bone: formation by autoinduction. *Science*, vol. 150, p. 893–899.
- Urist, M. R. and Strates, B. S., 1971: Bone morphogenetic protein. *Journal of Dental Research*, vol. 50, 1392–1406.
- Urry, L. A., Hamilton, P. C., Killian, C. E. and Wilt, F. H., 2000: Expression of spicule matrix proteins in the sea urchin embryo during normal and experimentally altered spiculogenesis. *Developmental Biology*, vol. 225, p. 201–213.
- Valentine, J. W., 2004: On the Origin of Phyla, 614p. The University of Chicago Press, Chicago.
- Venkatachalam, C. M. and Urry, D. W., 1981: Development of a linear helical conformation from its cyclic correlate. β-Spiral model of the elastin poly(pentapeptide) (VPGVG)n. *Macromolecules*, vol. 14, p. 1225–1229.
- Vermeij, G.J., 1989: The origin of skeletons. *Palaios*, vol. 4, p. 585–589.
- Wanninger, A. and Haszprunar, G., 2001: The expression of an engrailed protein during embryonic shell formation of the tusk-shell, *Antalis entalis* (Mollusca, Scaphopoda). *Evolution and Development*, vol. 3, p. 312–321.
- Watabe, N. and Wilbur, K. M., 1960: Influence of the organic matrix on crystal type in molluscs. *Nature*, vol. 188, p. 334.
- Watanabe, T., Fukuda, I., China, K. and Isa, Y., 2003: Molecular analyses of protein components of the organic matrix in

the exoskeleton of two scleractinian coral species. *Comparative Biochemistry and Physiology Part B*, vol. 136, p. 767–774.

- Watanabe, T., Persson, P., Endo, H. and Kono, M., 2000: Molecular analysis of two genes, *DD9A* and *B*, which are expressed during the postmolt stage in the decapod crustacean *Penaeus japonicus*. *Comparative Biochemistry and Physiology Part B*, vol. 125, p. 127–136.
- Weiner, S., 1979: Aspartic acid-ruch proteins: major components of the soluble organic matrix of mollusc shells. *Calcified Tissue International*, vol. 29, p. 163–167.
- Weiner, S., 1983: Mollusc shell formation: Isolation of two organic matrix proteins associated with calcite deposition in the bivalve *Mytilus californianus*. *Biochemistry*, vol. 22, p. 4139–4145.
- Weiner, S. and Addadi, L., 1991: Acidic macromolecules of mineralized tissues: the controllers of crystal formation. *Trends in Biochemical Sciences*, vol. 16, p. 252–256.
- Weiner, S. and Hood, L., 1975: Soluble protein of the organic matrix of mollusc shells: a potential template for shell formation. *Science*, vol. 190, p. 987–989.
- Weiss, I. M., Göhring, W., Fritz, M. and Mann, K., 2001: Perlustrin, a *Haliotis laevigata* (abalone) nacre protein, is homologous to the insulin-like growth factor binding protein N-terminal module of vertebrates. *Biochemical and Biophysical Research Communications*, vol. 285, p. 244–249.
- Weiss, I. M., Kaufmann, S., Mann, K. and Fritz, M., 2000: Purification and characterization of Perlucin and Perlustrin, two new proteins from shell of the mollusc *Haliotis laevigata*. *Biochemical and Biophysical Research Communications*, vol. 267, p. 17–21.
- Westbroek, P. and Marin, F., 1998: A marriage of bone and nacre. *Nature*, vol. 392, p. 861–2.
- Wewer, U.M., Ibaraki, K., Schjorring, P., Durkin, M.E., Young, M.F. and Albrechtsen, R., 1994: A potential role for tetranectin in mineralization during osteogenesis. *The Journal of Cell Biology*, vol. 127, p. 1767–1775.
- Wheeler, A. P., George, J. W. and Evans, C. A., 1981: Control of calcium carbonate nucleation and crystal growth by soluble matrix of oyster shell. *Science*, vol. 212, p. 1397–1398.
- Wheeler, A. P., 1992: Phosphoproteins of oyster (*Crassostrea virginica*) shell organic matrix. *In* Suga, S. and Watabe, N. eds, Hard Tissue Mineralization and Demineralization, p. 171–187. Springer-Verlag, Tokyo.
- Williamson, M. P., 1994: The structure and function of prolinerich regions in proteins. *Biochemical Journal*, vol. 297, p. 249–260.
- Wilt, F. H., 1999: Matrix and mineral in the sea urchin larval skeleton. *Journal of Structural Biology*, vol. 126, p. 216– 226.
- Wilt, F. H., 2002: Biomineralization of the spicule of sea urchin embryos. *Zoological Science*, vol. 19, p. 253–261.
- Wilt, F. H., 2003: Developmental biology meets materials science: Morphogenesis of biomineralized structures. *Developmental Biology*, vol. 280, p. 15–25.

- Wilt, F. H., Killian, C. E. and Livingston, B. T., 2003: Development of calcareous skeletal elements in invertebrates. *Differentiation*, vol. 71, p. 237–250.
- Wozney, J. M., 1992: The bone morphogenetic protein family and osteogenesis. *Molecular Reproduction and Development*, vol. 32, p. 160–167.
- Wozney, J. M., Rosen, V., Celeste, A. J., Mitsock, L. M., Whitters, M. J., Kriz, R. W., Hewick, R. M. and Wang, E. A., 1988: Novel regulators of bone formation: molecular clones and activities. *Science*, vol. 242, p. 1528–1534.
- Wu, Y. Q., Badano, J. L., McCaskill, C., Vogel, H., Potocki, L. and Shaffer, L. G., 2000: Haploinsufficiency of *ALX4* as a potential cause of parietal foramina in the 11p11.2 contiguous gene-deletion syndrome. *The American Journal of Human Genetics*, vol. 67, p. 1327–1332.
- Wustman, B. A., Morse, D. E. and Evans, J. S., 2004: Structural characterization of the N-terminal mineral modification domains from the molluscan crystal-modulating biomineralization proteins, AP7 and AP24. *Biopolymers*, vol. 74, p. 363–376.
- Wynn, A. and Shafer, T. H., 2005: Four differentially expressed cDNAs in *Callinectes sapidus* containing the Rebers– Riddiford consensus sequence. *Comparative Biochemistry* and Physiology Part B, vol. 141, p. 294–306.
- Yamaguchi, A., Komori, T. and Suda, T., 2000: Regulation of osteoblast differentiation mediated by bone morphogenetic proteins, hedgehogs, and Cbfa1. *Endocrine Reviews*, vol. 21, p. 393–411.
- Yang, S. D., 1971: Organic material from scleractinian coral skeletons-I. Variation in composition between several species. *Comparative Biochemistry and Physiology Part B*, vol. 40, p. 113–120.
- Yang, S.-Y., Tsuzuki, M. and Miyachi, S., 1985: Carbonic anhydrase of *Chlamydomonas*: Purification and studies on its induction using antiserum against *Chlamydomonas* carbonic anthydrase. *Plant and Cell Physiology*, vol. 26, p. 25–34.
- Yano, K. and Zarain-Herzberg, A., 1994: Sarcoplasmic reticulum calsequestrins: structural and functional properties. *Molecular and Cellular Biochemistry*, vol. 135, p. 61–70.
- Zhang, Y., Xie, L., Meng, Q., Jiang, T., Pu, R., Chen, L. and Zhang, R., 2003: A novel matrix protein participating in the nacre framework formation of pearl oyster, *Pinctada fucata. Comparative Biochemistry and Physiology Part B*, vol. 135, p. 565–573.
- Zhao, G.-Q., Eberspaecher, H., Seldin, M. F. and de Crombrugghe, B., 1994: The gene for the homeodomaincontaining protein Cart-1 is expressed in cells that have a chondrogenic potential during embryonic development. *Mechanisms of Development*, vol. 48, p. 245–254.
- Zhu, X., Mahairas, G., Illies, M., Cameron, R. A., Davidson, E. H. and Ettensohn, C. A., 2001: A large-scale analysis of mRNAs expressed by primary mesenchyme cells of the sea urchin embryo. *Development*, vol. 128, p. 2615– 2627.