ZOOLOGICAL SCIENCE **16**: 619–628 (1999)

Cloning and Expression of a cDNA Encoding an Insoluble Matrix Protein in the Gastroliths of a Crayfish, *Procambarus clarkii*

Naoaki Tsutsui[1,2], Katsuaki Ishii[1], Yasuaki Takagi[3], Toshiki Watanabe[1], and Hiromichi Nagasawa[2]*

[1]*Department of Molecular Biology, Ocean Research Institute, The University of Tokyo, Minamidai 1-15-1, Nakano, Tokyo 164-0014, Japan,*
[2]*Department of Applied Biological Chemistry, The University of Tokyo, Yayoi 1-1-1, Bunkyo, Tokyo 113-8657, Japan and*
[3]*Otsuchi Marine Research Center, Ocean Research Institute, The University of Tokyo, Akahama 2-106-1, Otsuchi, Iwate 028-1102, Japan*

ABSTRACT—In the crayfish *Procambarus clarkii*, the gastroliths are formed as a paired structure in the stomach during the premolt period, and contain calcium carbonate and a small amount of an organic matrix. In this investigation, a cDNA encoding an insoluble matrix protein was isolated from *P. clarkii*. The open reading frame encoded 505 amino acid residues including two unique repeated sequences. The N-terminal half of the amino acid sequence, which included 10-amino-acid repeats, exhibited a high degree of similarity to that of involucrin, a protein synthesized in human keratinocytes. Northern blot analysis revealed that mRNA encoding the matrix protein is specifically expressed in the gastrolith discs during the premolt period in which the gastroliths formed. In the gastrolith discs, levels of expression of this mRNA correlated increases in weights of the gastroliths concomitant with their formation. Organ culture of the gastrolith discs suggested that expression of mRNA in the discs is induced by molting hormone, 20-hydroxyecdysone. These results reinforced the relationship between the matrix protein and formation of the gastroliths. Functional analysis showed that the protein inhibits calcium carbonate crystallization in a solution system, suggesting that the protein plays a role in the calcification of the gastroliths.

INTRODUCTION

Hard tissues, observed in many organisms, help maintain body structure, defend against enemies and store calcium. They often contain biominerals, which are mostly composed of inorganic compounds with a small amount of organic components which are considered to regulate formation of biominerals. In mammals, osteocalcin (Hauschka and Carr, 1982) and several proteins (Stetler-Stevenson and Veis, 1983; Gorski, 1992; George *et al*., 1993) have been identified as trace constituents contained in the calcium phosphate making up bones or teeth. In many invertebrates, the main component of such biomineral material is calcium carbonate, present in crystalline or amorphous form. Organic compounds contained in biominerals may control the formation of biomineral like in mammals, and influence crystal phase, affecting balances of substances such as aragonite or calcite (Belcher *et al*., 1996; Falini *et al*., 1996). A few matrix proteins have recently been characterized. cDNAs encoding matrix proteins in the spicules of the sea urchin have been isolated and their nucleotide sequences have been determined, thereby allowing the primary structure of the spicule matrix proteins to be deduced (Kato-Fukui *et al*., 1991). A soluble matrix protein, nacrein, has been characterized from the pearl oyster and found to have two functions, carbonic anhydrase activity and calcium-binding ability (Miyamoto *et al*., 1996). Two separate cDNA clones encoding framework proteins have been isolated and sequenced from the same animal (Sudo *et al*., 1997). A cDNA clone encoding Lustrin A, a matrix protein of the nacreous layer of abalone, have been isolated (Shen *et al*., 1997).

Crustaceans form skeletal structures by storing calcium carbonate within the cuticle comprising their exoskeletons. Some of the proteins which are contained in the exoskeleton have been studied. One of these is a matrix protein in the calcareous concretion of the terrestrial crustacean *Orchestia cavimana*, for which 14 amino acid residues at its N-terminus have been determined. This protein has been shown to possess calcium-binding ability (Luquet *et al*., 1996). In some crus-

* Corresponding author: Tel. +81-3-5841-5132;
FAX. +81-3-5841-8022.

taceans, calcium liberated from the exoskeleton during the premolt period is stored inside the body tissues. In the crayfish *Procambarus clarkii*, most of the liberated calcium is stored in a pair of gastroliths, a kind of biomineral, which form on a pair of gastrolith discs located at the fore of the stomach (Travis, 1960). Gastroliths are composed mainly of calcium carbonate and contain a small amount of organic materials including proteins and carbohydrates. Gastroliths continue to enlarge as the molt cycle proceeds. After molting, they are digested and the calcium carbonate in the gastroliths is used to harden the new exoskeleton. Gastrolith formation is known to be induced by the molting hormone, 20-hydroxyecdysone (Sonobe *et al.*, 1991; Ueno *et al.*, 1992). Histological examination has revealed that gastroliths form between the cuticle and epidermis of the gastrolith discs, suggesting that the epidermal tissue facing the gastroliths are responsible for gastrolith formation (Ueno, 1980). Based on the above evidence, we investigated gastrolith formation in consideration of calcification as well as transportation of calcium in synchronization with the molting cycle. We have thereby isolated and characterized a gastrolith matrix protein (GAMP) from the insoluble organic matrix of gastroliths (Ishii *et al.*, 1996, 1998). GAMP was thought to be a major protein contained in the gastroliths and to be tightly associated with chitin which is a main component of the insoluble matrix. It had a molecular mass of about 50.5 kDa and a blocked amino terminus. By enzymatic digestion and microsequencing, five partial amino acid sequences with a total of 225 amino acid residues have been identified. In the present study, we describe the cloning and sequencing of cDNA encoding GAMP, analysis of expression of its mRNA, and possible functioning of GAMP.

MATERIALS AND METHODS

Animals

Freshwater crayfish, *P. clarkii*, were purchased from a local dealer. They were maintained in indoor aquaria at 18–20°C and provided artificial pellets (MARUHA Co. Ltd., Tokyo) every 2 to 3 days. Males at the intermolt stage with a carapace width of 20–23 mm and body length of 80–90 mm were used for all experiments. In this study, crayfishes during premolt period are those animals which were artificially induced to reach premolting by bilateral eyestalk ablation (Sonobe *et al.*, 1991).

Isolation of RNA

Gastrolith discs were removed from crayfishes at 0, 2, 4, 6, 9, 12, 16, and 21 days after bilateral eyestalk ablation. In addition, the fan blade, abdominal muscle, intestine, and thorax connective tissue were dissected from crayfishes 0 and 16 days after eyestalk ablation. Total RNA was isolated from these organs using ISOGEN (Wako, Osaka), according to instructions of the manufacturer.

Construction of the cDNA library

Using Oligotex-dT30 super (Roche, Japan), poly(A)$^+$ RNA was prepared from total RNA which had been isolated from the anterior part of the stomach including the gastrolith discs of crayfish at 16 days after bilateral eyestalk ablation. The poly(A)$^+$ RNA was subjected to cDNA library construction using a TimeSaver® cDNA synthesis kit (Pharmacia Biotech). The synthesized cDNA was ligated to the λ ZAP-II vector (Stratagene). The ligation product was packaged using Gigapack® II Gold (Stratagene).

Screening of the cDNA library

Two degenerate oligonucleotide primers, 2F and 2R, were designed based on the amino acid sequence of GAMP (Fig. 1). They were derived from the sequences of amino acid positions 3–8 and 30–35 of partial sequence 2 (Ishii *et al.*, 1998), respectively. PCR amplification was carried out in a 50 μl solution containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 0.02% gelatin, 400 μM dNTP, 3.2 μM of each primer 2F and 2R, and 2 U r*Taq* DNA polymerase (TaKaRa). After an initial denaturation at 94°C for 3 min, 36 cycles of PCR were carried out with the following program: denaturation at 94°C for 1 min, annealing at 45°C for 1 min and extension at 55°C for 3 min. Final extension was done at 55°C for 10 min. All PCR amplifications were performed on a PROGRAM TEMP CONTROL SYSTEM PC-700 (ASTEC). The product of the PCR was subcloned into the pCR™ II plasmid vector using a TA Cloning® Kit (Invitrogen). The *Eco*RI fragment of this plasmid was self-ligated using T4 DNA ligase to generate longer fragments, and labeled with digoxigenin (DIG)-labeled dUTP by using a DIG DNA Labeling and Detection Kit (Boehringer Mannheim). The cDNA library of the gastrolith discs was screened with this DIG-labeled DNA. Positive plaques were isolated in the second round of screening, and the recombinant pBluescript SK$^-$ phagemids were rescued from the bacteriophage clones according to instructions of the manufacturer (Stratagene). Dideoxy sequencing reactions were carried out using a Thermo Sequenase Cycle Sequencing Kit (Amersham), and electrophoresis and signal detection were done on an automatic sequencer SQ-3000 (HITACHI).

Rapid amplifications of 5' cDNA end (5' RACE)

Poly(A)$^+$ RNA (1 μg) isolated from the gastrolith discs of crayfish 16 days after bilateral eyestalk ablation was reverse-transcribed in a 20 μl solution containing 100 pM primer JR (nucleotides 921–940, Fig. 1), 500 μM dNTP, 10 mM dithiothreitol, 25 U RNase inhibitor (GIBCO BRL), and 200 U reverse transcriptase (GIBCO BRL) for 1 hr at 42°C. The reaction was stopped by the addition of 80 μl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). The resulting single-stranded cDNA was tailed with poly(dA) by a 3'-end labeling kit (Amersham). The PCR was carried out using an LA PCR Kit Ver.2 (TaKaRa) with primer NR, Adaptor, T17-Adaptor (500 nM each, Fig. 1), and the poly(dA)-tailed cDNA as a template. The initial denaturation at 94°C for 4 min, 3 PCR cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 3 min was followed by 37 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 3 min, and the final extension at 72°C for 15 min. The PCR product was subcloned and sequenced as described above. The same PCR was also performed using primer NR2 (Fig. 1) instead of primer NR and the poly(dA)-tailed cDNA as a template.

Northern blot analysis

Total RNA (10 μg/lane) obtained from the gastrolith discs, abdominal muscle, intestine, thorax connective tissue and fan blade at each intermolt and premolt period were run on a 1% formaldehyde-agarose gel, transferred to a Hybond-N$^+$ membrane (Amersham). A GAMP cDNA fragment (corresponding to nucleotides 1–271, Fig. 2) which had been labeled with [α-^{32}P]dCTP (Amersham) using Megaprime™ DNA labeling system (Amersham) was used as a probe. The membrane was hybridized with the radiolabeled probe (1×10^6 cpm/ml) at 42°C in 50% formamide, 6×SSC, 1×Denhard's solution, 0.4% SDS and 100 μg/ml calf thymus DNA for 14 hr. The membrane was washed with 0.1×SSC and 0.1% SDS at 65°C. RNA ladder (0.24–9.4 kb, GIBCO BRL) was used as the molecular weight marker.

A Northern blot of total RNA of the gastrolith discs of crayfish at 0, 2, 4, 6, 9, 12, 16, and 21 days after eyestalk ablation was hybridized with a part of the GAMP cDNA corresponding to nucleotides 1–271 (Fig. 2) as described above. Levels of hybridized transcripts were determined by a Fujix Bio-imaging analyzer BAS 1000 (Fuji Film).

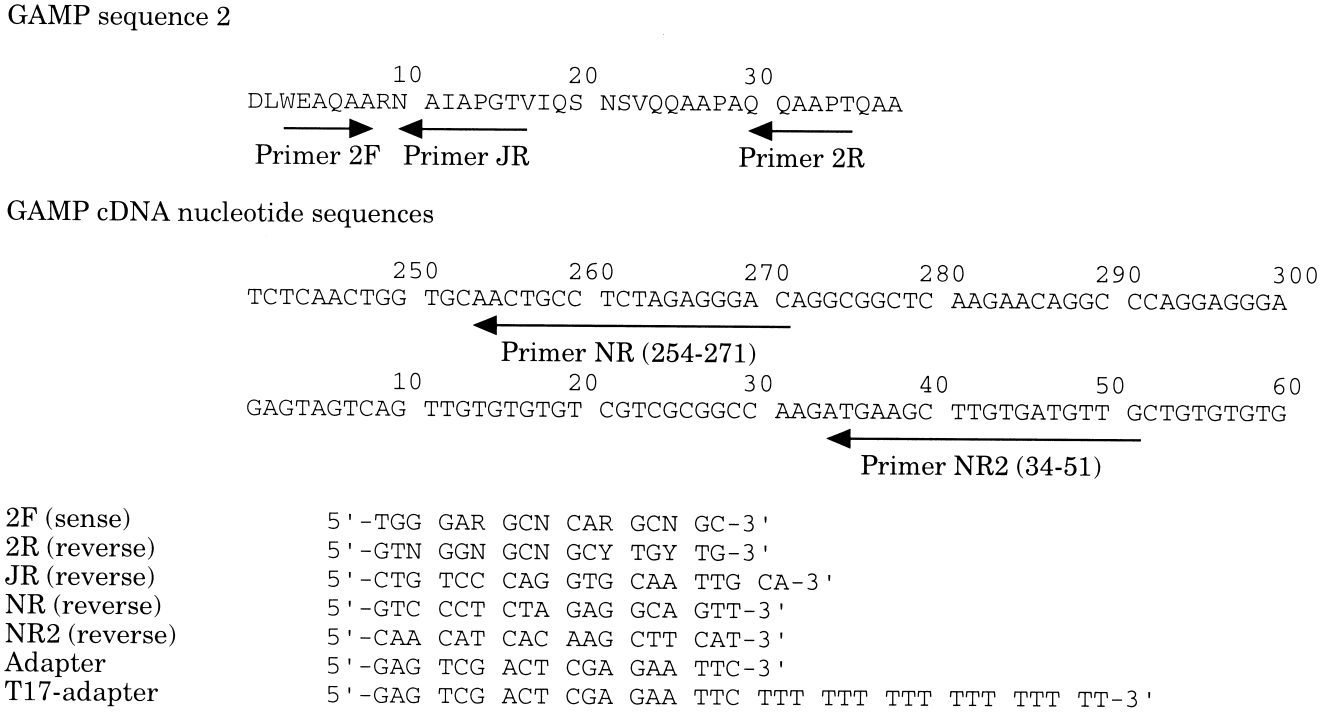

Fig 1. Positions and nucleotide sequences of oligonucleotide primers used to clone GAMP cDNA. Two degenerate primers, 2F and 2R, were used for PCR amplification, the products of which were used to screen a cDNA library of the gastrolith discs during the premolt period. The other primers were used for 5' RACE PCR amplifications. GAMP sequence 2 was obtained by protein sequencing (Ishii *et al.*, 1998). Nucleotide numbers of GAMP cDNA sequences are consistent with those of Fig. 2.

Immunological quantification of ecdysteroids

Hemolymph collected from crayfish at 0, 2, 4, 6, 9, 12, 16, and 21 days after eyestalk ablation was subjected to radioimmunoassay (RIA) to determine hemolymph ecdysteroid titer according to Borst and O'Conner (1972) and Brown *et al.* (1993). An anti-20-hydroxyecdysone serum was provided from Dr. S. Takeda of the National Institute of Sericultural and Entomological Science, Ministry of Agriculture, Forestry, and Fisheries of Japan.

Organ culture of gastrolith discs

Crayfish carapace was washed with 70% ethanol, after which a pair of gastrolith discs including their surrounding stomach tissue were dissected. The gastrolith discs were rinsed with sterile culture medium which was prepared according to Sonobe *et al.* (1991) and separated from one another. One disc was cultured in 5 ml of the medium alone and the other was incubated in the medium containing 20-hydroxyecdysone (40 ng/ml). They were maintained in a Tissue Culture Flask (IWAKI GLASS) at 25°C for 20 hr. Using total RNA extracted from each cultured tissue, Northern blot analysis was carried out as described above.

Collection of gastroliths and purification of gastrolith matrix protein (GAMP)

A pair of gastroliths were taken out from crayfishes at 0, 2, 4, 6, 9, 12, 16, and 21 days after bilateral eyestalk ablation. GAMP was purified from gastroliths collected from the crayfish 12 and 16 days after bilateral eyestalk ablation according to methods described previously (Ishii *et al.*, 1998).

Endoproteinase Asp-N digestion of GAMP

GAMP (68 mg, 1.3 nmol) was digested with endoproteinase Asp-N (Boehringer Mannheim) and the digest was separated by reverse-phase HPLC as reported previously (Ishii *et al.*, 1998).

Mass spectral analyses

Mass spectra were measured on a time-of-flight (TOF) mass spectrometer (Voyager™Elite, Perceptive Biosystems) in the positive ion mode with sinapinic acid as the matrix.

Calcium binding test

GAMP and calmodulin were separated on a 12.5% SDS-PAGE (Laemmli, 1970) and transferred to a polyvinylidene difluoride membrane (ATTO). A $^{45}Ca^{2+}$ binding test was performed according to a standard protocol (Maruyama *et al.*, 1984) with $^{45}Ca^{2+}$ (DuPont New England Nuclear). Radioactivity was detected using a Fujix Bio-imaging analyzer BAS 1000 (Fuji Film).

Inhibition of calcium carbonate precipitation

The effect of GAMP on the rate of $CaCO_3$ precipitation was determined with supersaturated solutions prepared as described by Wheeler *et al.* (1981). $CaCO_3$ formation was followed by recording the pH decrease of a solution containing 1.5 ml of 20 mM $NaHCO_3$, pH 8.7, and 10 µl of sample, after 1.5 ml of 20 mM $CaCl_2$ was added. Changes of pH of the solutions were measured with a glass/calomel electrode pair.

RESULTS

Isolation and characterization of a GAMP cDNA

In order to isolate the cDNA clones encoding GAMP, a cDNA library was prepared from the gastrolith discs during the premolt period. Two degenerate primers, 2F and 2R (Fig. 1), were designed based on partial amino acid sequence 2 of GAMP (Ishii *et al.*, 1998). A PCR product of the expected size (98 bp) was obtained from the cDNA of the gastrolith discs

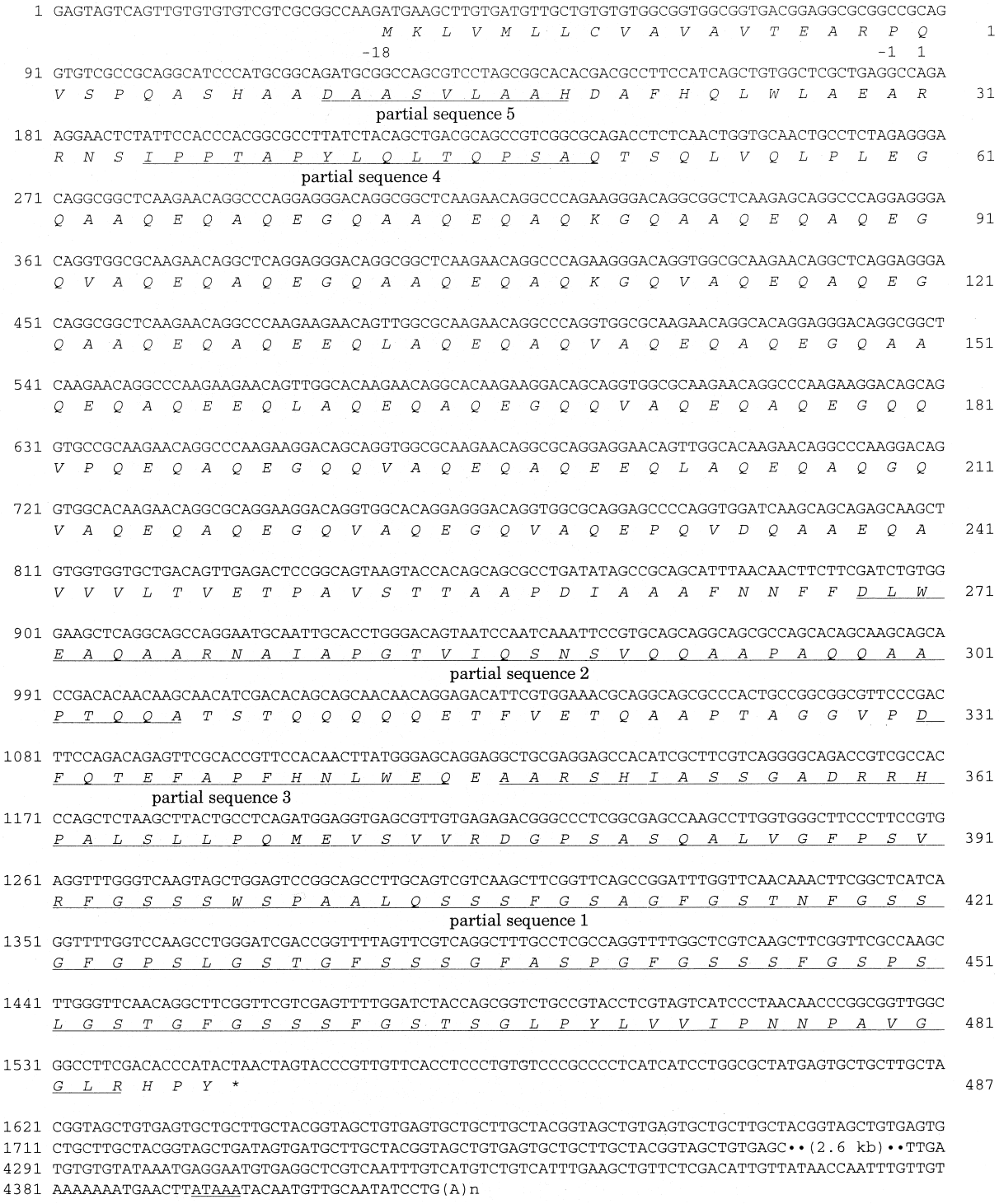

Fig. 2. Composite nucleotide sequence and deduced amino acid sequence for GAMP. The cDNA sequence was derived from a combination of a sequence of a clone (begins at 145) and a sequence obtained by 5' RACE. The deduced amino acid sequence which is shown below their respective codons in *italics* starts at nucleotide position 34 and the stop codon is indicated by an asterisk (*). Nucleotide and amino acid numbers are indicated on the left and right sides, respectively. Only part of 3' untranslated region is shown and the putative polyadenylation signal is underlined. This nucleotide sequence has been submitted to the GenBank™/EMBL Data Bank with accession number AB017017.

using these primers. Results of sequence analysis showed that the product was derived from a GAMP cDNA, because the deduced amino acid sequence was identical to a part of sequence 2 (Ishii *et al.*, 1998) of GAMP (data not shown). This PCR product was used as a probe to screen a cDNA library.

Among the isolated clones, that with the longest insert (4,275 bp excluding the poly(A)$^+$ tail) was analyzed for its nucle-

otide sequence (nucleotide position 145–4,419 in Fig. 2). In this cDNA clone, an open reading frame of 468 amino acids was found. Four out of the 5 partial amino acid sequences of GAMP (sequences 1, 2, 3, and 4 (Ishii *et al*., 1998)) were included in the deduced amino acid sequence (Fig. 2).

In the open reading frame, the first methionine was found in the partial sequence 1 (Fig. 2), suggesting that the analyzed clone was a partial one and that the 5' region including sequence 5 and the initial methionine was missing. To recover the missing part, 5' RACE was carried out, and an additional sequence of 144 bp (nucleotide position 1–144 in Fig. 2) was obtained. When this sequence was merged into the cDNA clone, the open reading frame was extended by 37 amino acid residues (amino acid position -18 – +19 in Fig. 2). The partial amino acid sequence 5 (Ishii *et al*., 1998) following an initial methionine was included in the added region (Fig. 2). Thus, all of the partial amino acid sequences of GAMP determined previously (Ishii *et al*., 1998) were included in the open reading frame. We, therefore, concluded that the analyzed cDNA encoded GAMP. GAMP mRNA contained 33 bp of a 5' untranslated region, 1,515 bp of an open reading frame, and 2,871 bp of a 3' untranslated region. A putative polyadenylation signal sequence ATAAA was found at the 3' untranslated region 20 nucleotides upstream from the poly(A) tail.

Table 1. Amino acid composition of GAMP

Amino acid	acid hydrolysis*	deduced from cDNA sequence*		
Gly	8.2	8.2		
Ala	16.1	17.9		
Val	6.5	6.2		
Leu	5.8	4.9		
Ile	2.4	1.2		
Met	0.0	0.2		
Phe	5.1	4.1		
Trp	—	0.8		
Pro	5.6	5.8		
Ser	8.9	9.5		
Thr	2.9	3.9		
Cys	0.0	0.0	Asn	1.9
Asx	2.9	3.5	Asp	1.6
Glx	31.8	29.5	Gln	20.5
Lys	1.4	0.4	Glu	9.0
His	0.0	1.4		
Arg	2.0	1.9		
Tyr	0.5	0.6		
Total	100.1	100.0		

* Amino acid compositions are represented as percent.
— : Tryptophan was not determined.

Determination of the N-terminal residue of GAMP

A fragment, designated D1, was obtained by digestion of GAMP with endoproteinase Asp-N (Ishii *et al*., 1998). D1 had a molecular mass of 977.48 Da as determined by TOF mass analysis and had possibly a blocked N-terminus, because it did not give any significant amino acid sequence by automated Edman degradation. The mass value was in complete agreement with the theoretical value for the decapeptide, pGlu-Val-Ser-Pro-Gln-Ala-Ser-His-Ala-Ala, providing that the N-terminal glutamine residue would be cyclized to a pyroglutamic acid residue. However, pyroglutamate aminopeptidase digestion of D1 did not provide clear results. In the post-source decay mass spectrum of D1, a series of b-type fragment ions were observed at *m*/*z* 211.26, 298.39, 395.70, 524.07, 594.69, 682.08, 818.50, and 889.44, indicating the presence of the sequence from Ser^{3} to Ala^{10}. A collision-induced dissociation mass spectrum of D1 clearly showed the presence of pGlu and Val as well as Ala, Ser, Pro, Gln, His. Thus, the amino acid sequence of D1 was established, and therefore it was determined that the N-terminal residue of GAMP was pGlu. The glutamine residue at position 1 of the deduced amino acid sequence was the second most likely site of cleavage calculated according to methods of Von Heijne (1986).

Characterization of the deduced amino acid sequence of GAMP

The mature protein had a calculated molecular mass of 50,854 Da. This value was close to the value of 50,501 measured on a TOF mass spectrometer (Ishii *et al*., 1998). There was no potential *N*-glycosylation site in the GAMP, which was consistent with the previous report (Ishii *et al*., 1998). The amino acid composition of the deduced sequence of GAMP was close to that obtained by amino acid analysis (Table I). In the deduced amino acid sequence of GAMP, two unique repeated sequences were found: one was Gly-Ser-Xaa-Ser/Gly-Phe (Xaa indicates any amino acid residue), which has been reported previously (Ishii *et al*., 1996), and the other was Gln-Val/Ala-Ala-Gln-Glu-Gln-Ala-Gln-Glu-Gly (Fig. 3). The BLAST algorithm (Altschul *et al*., 1990) identified that the latter repeated sequence consisting of 10 amino acid residues had some similarity to that of involucrin (Eckert and Green, 1986) (Fig. 3), a keratinocyte protein, which is cross-linked to membrane proteins by transglutaminase (Simon and Green, 1985; Thacher and Rice, 1985). Unlike GAMP, involucrin was composed mainly of repeats of 10 amino acid residues.

Distribution of the GAMP mRNA

Tissue specific expression of the GAMP mRNA was analyzed by Northern hybridization. Total RNA was prepared from the gastrolith discs, abdominal muscle, intestine, thorax connective tissue, and fan blade during both intermolt and premolt periods. The RNA blot was hybridized with the ^{32}P-labeled GAMP cDNA. Hybridization signals were detected in the gastrolith discs only during the premolt period and, in contrast to expected results, in the fan blade only during intermolt period, but not in other tissues (Fig. 4). The probe hybridized to a single band of approximately 4.4 kb in the gastrolith discs and 3.2 kb in the fan blade. The size of the band in the gastrolith discs was in good agreement with that of the GAMP cDNA obtained by combination of screening of a cDNA library and 5' RACE method (4,419 bp). This result suggested that the GAMP cDNA was complete or nearly complete.

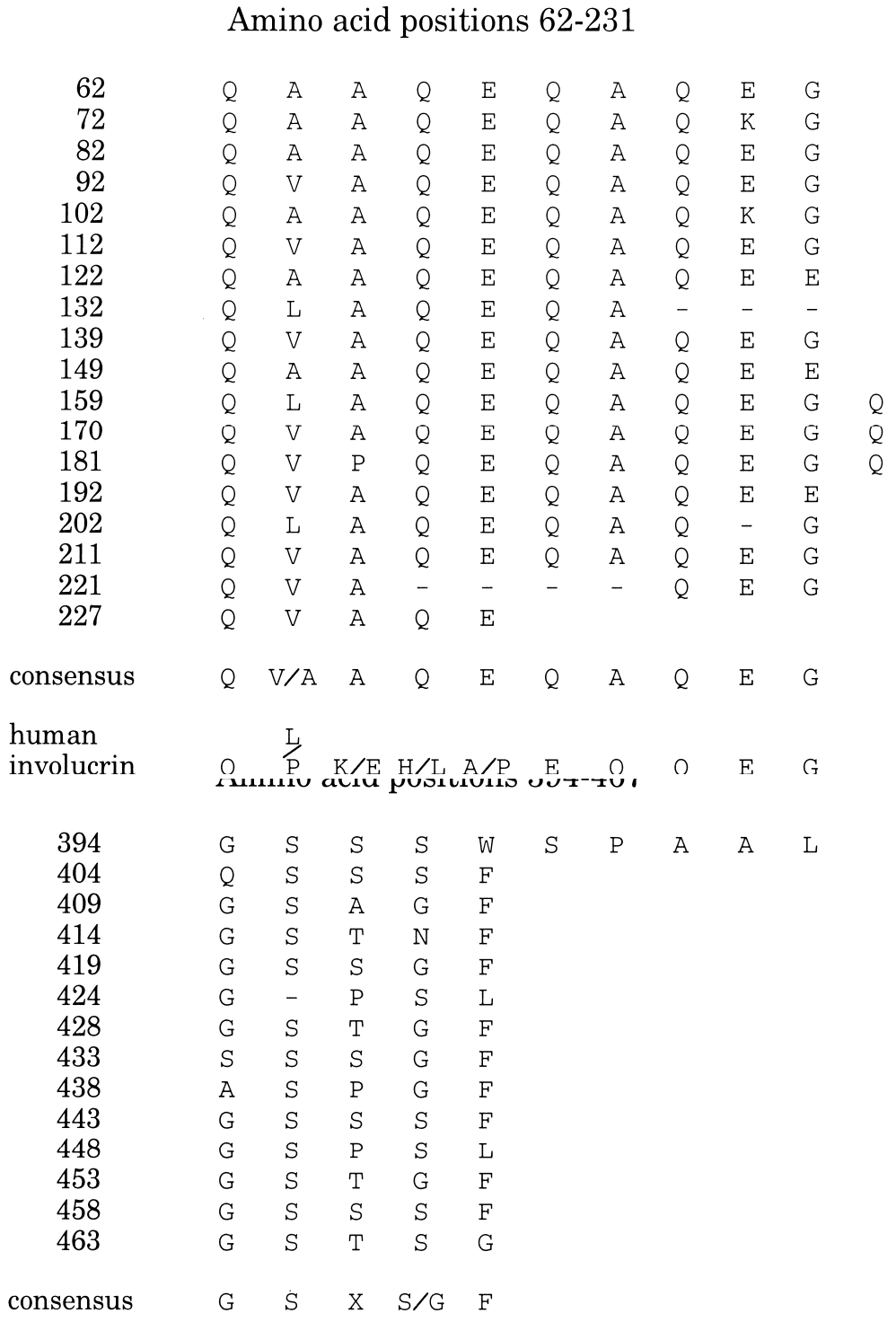

Fig. 3. Two repeated structures in the deduced GAMP sequence. The 10-amino-acid repeat found in amino acid positions 62–231 and the 5-amino-acid repeat in positions 394–467 are presented with the respective consensus sequence. Numbers on the left represent amino acid positions of the leftmost residues of each row. Consensus sequence of the 10-amino-acid repeats in amino acid positions 153–541 of human involucrin (Eckert and Green, 1986) is also shown.

Changes in relative levels of the GAMP mRNA in the gastrolith disc

Changes in levels of GAMP transcript in the gastrolith disc during premolt period were examined by Northern hybridization. The same cDNA fragment was used to probe a Northern blot containing total RNA from the gastrolith discs of crayfish at 0, 2, 4, 6, 9, 12, 16, and 21 days after bilateral eyestalk ablation. As a result, a single band approximately 4.4 kb long was detected in Day-4, -6, -9, -12, -16, and -21 crayfish. The levels of hybridization changed significantly

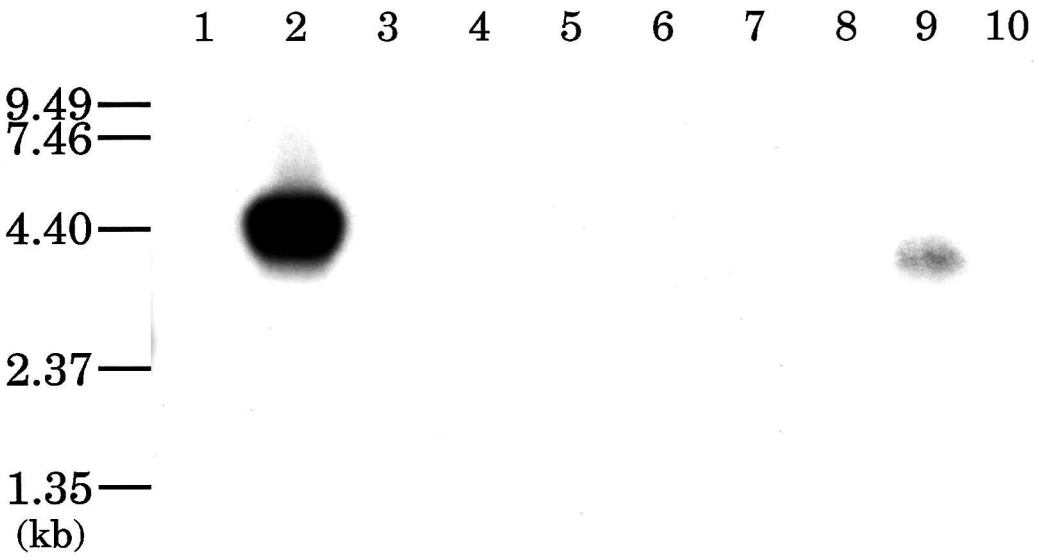

Fig. 4. Northern blot analysis of tissue specific expression of the GAMP mRNA. Total RNA (10 μg/lane) from the gastrolith discs during the intermolt period (lane 1), gastrolith discs during the premolt period (lane 2), abdominal muscle during the intermolt period (lane 3), abdominal muscle during the premolt period (lane 4), intestine during the intermolt period (lane 5), intestine during the premolt period (lane 6), thorax connective tissue during the intermolt period (lane 7), thorax connective tissue during the premolt period (lane 8), fan blade during the intermolt period (lane 9), and fan blade during the premolt period (lane 10) were blotted onto a nylon membrane and probed with GAMP cDNA which corresponded to nucleotides 1-271. Positions of the five molecular weight markers (9,490, 7,460, 4,400, 2,370 and 1,350 nucleotides) are indicated.

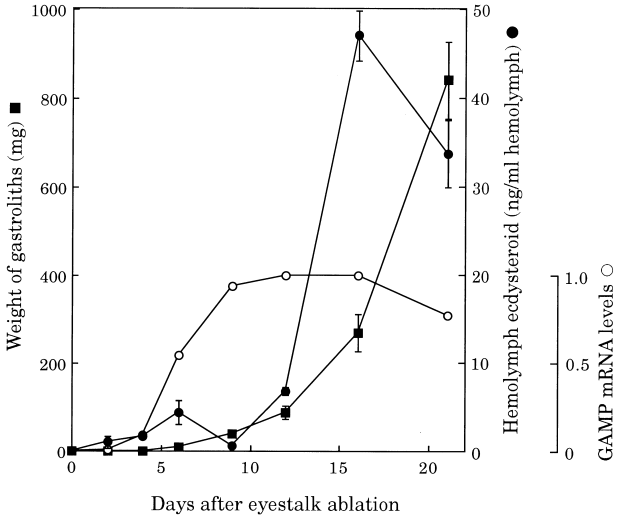

Fig. 5. Changes in relative expression levels of the GAMP mRNA in the gastrolith disc, weight of the gastroliths, and hemolymph ecdysteroid titers after eyestalk removal. Solid squares, weight of gastroliths; solid circles, ecdysteroids in the hemolymph; open circles, relative levels of the GAMP encoding mRNA.

after eyestalk ablation, the minimal level being observed 4 days after eyestalk ablation and the levels being increased with days (Fig. 5). Maximum levels were observed on days 12 and 16, whereas a significant decrease was observed 21 days after eyestalk removal, just prior to molting. To ensure equal loading of RNA samples, gels were stained with ethidium bromide before blotting; levels of ribosomal RNA thus visualized were almost equal (data not shown).

Results of changes in levels of GAMP mRNA in the gastrolith discs after eyestalk ablation, weight of gastroliths and hemolymph ecdysteroid titers in the same animals are shown in Fig. 5. The increase in the hemolymph ecdysteroid titer was observed at 2 days after eyestalk ablation and that was followed by the increase in GAMP mRNA and weight of gastroliths.

Induction of expression of GAMP mRNA by 20-hydroxyecdysone

Organ culture of gastrolith discs was carried out to examine whether the expression of GAMP was induced by ecdysteroids. One disc from a pair of gastrolith discs was placed into culture medium alone and the other was plated into the same medium containing 20-hydroxyecdysone and incubated at 25°C for 20 hr. Northern blot analysis using total RNA extracted from the cultured organs revealed that expression of GAMP mRNA was significantly enhanced in the gastrolith disc cultured with 20-hydroxyecdysone compared with that organ cultured without 20-hydroxyecdysone (Fig. 6).

Ca^{2+}-binding property of GAMP

As the first step to study the role of GAMP in the calcification of the gastroliths, we examined Ca^{2+} binding ability of GAMP according to a procedure developed by Maruyama *et al.* (1984). A Western blot of GAMP and calmodulin (a positive control) was immersed in a solution containing $^{45}Ca^{2+}$ and radioactivities in the protein bands were detected. Weak signals were observed at positions where GAMP was electrophoresed, but the intensity of the signals did not increase in proportion to loaded amounts of protein as compared with calmodulin, a calcium binding protein, whose intensity was

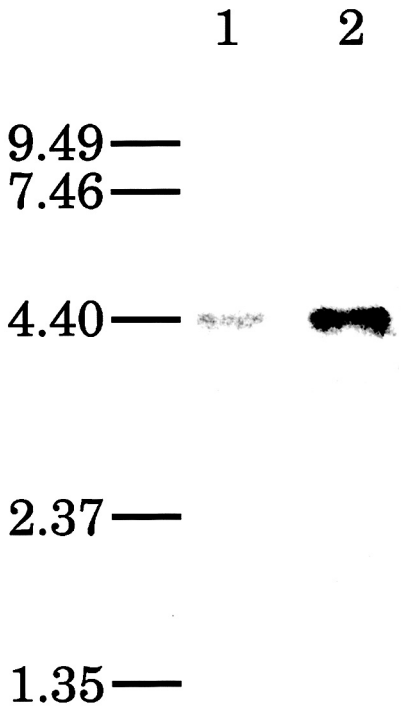

Fig. 6. Northern blot analysis of expression of GAMP mRNA from cultured gastrolith discs. Total RNA (15 μg/lane) from the gastrolith discs cultured without 20-hydroxyecdysone (lane 1) and cultured with 20-hydroxyecdysone (lane 2) were blotted onto a nylon membrane and probed with GAMP cDNA which corresponded to nucleotides 1–271. Positions of the five molecular weight markers (9,490, 7,460, 4,400, 2,370, and 1,350 nucleotides) are indicated.

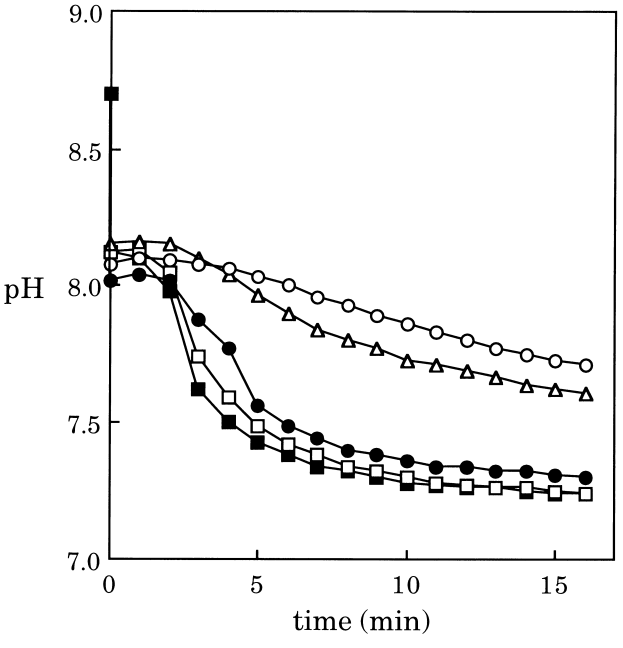

Fig.7. Effects of GAMP on CaCO3 precipitation. Changes of pH value of the solutions containing 1.5 ml of 20 mM $NaHCO_3$ after the addition of 1.5 ml of 20 mM $CaCl_2$ at time 0 are shown. Solid squares, control DW; solid circles, 5×10^{-7} M BSA; open squares, 5×10^{-9} M GAMP; open triangles, 5×10^{-8} M GAMP; open circles, 1×10^{-7} M GAMP (final concentration).

increased with the loaded amounts of protein. Thus, the signal may be attributed to non-specific biding. In conclusion, significant binding of Ca^{2+} to GAMP could not be detected.

Inhibition of calcium carbonate precipitation by GAMP

In order to estimate the functioning of GAMP, the effects of GAMP on the inhibition of calcium carbonate precipitation was examined by monitoring the pH value of supersaturated calcium carbonate solution with time. Fig. 7 shows these results. When GAMP was added to a $NaHCO_3$ solution prior to the addition of $CaCl_2$ solution, the rate of decreasing pH value was suppressed significantly as compared with H_2O or bovine serum albumin as negative controls. GAMP did not show inhibitory effects at concentrations of 5×10^{-9} M, but weak inhibition was observed at concentrations of 5×10^{-8} M. The inhibitory effects of GAMP were enhanced at concentrations of 1×10^{-7} M.

DISCUSSION

We have been studying the molecular mechanisms of biomineralization through purification and characterization of proteins in organic matrices of mineralized tissues. In the present experiment, we have isolated and characterized cDNA encoding GAMP, which constitutes a significant proportion of the insoluble matrix proteins of the gastroliths of the crayfish. This is the first report detailing a complete amino acid sequence and describing the functioning of a purified insoluble matrix protein contained in invertebrate biomineral material. All experiments conducted thus far on insoluble matrix proteins have revealed only amino acid composition, apparent molecular weight on SDS-PAGE, and ability to regulate crystallization. There had been little evidence demonstrating the purity of such insoluble proteins.

From the deduced amino acid sequence of GAMP, it was seen that GAMP had two different sets of unique repeated sequences. All matrix proteins from invertebrate hard tissues studied thus far have repeated sequences, but sequences differed from one another (Kato-Fukui *et al.*, 1991; Miyamoto *et al.*, 1996; Shen *et al.*, 1997; Sudo *et al.*, 1997). These repeated sequences would be associated with crystal growth by binding of calcium ions or by formation of a particular structure appropriate or inappropriate for crystallization. Data base searches have detected a similarity of the repeated sequence composed of 10 amino acids of GAMP with involucrin (Fig. 3), which is a protein produced specifically in terminally differentiated keratinocytes. Involucrin forms an insoluble envelope with other cytosolic proteins through a γ-glutamyl-ε-lysine isopeptide bond by transglutaminase, which prevents solubilization of cross-linked proteins from denaturants and reductants. In the case of gastroliths, a main component in the insoluble matrix was chitin. GAMP appears to become insoluble as a result of associating with chitin. Each 10-amino-acid repeat possesses glutamine residues, which can act as a primary amine acceptor in involucrin (Simon and Green, 1988), at four conserved positions. However, it is difficult to

postulate that transglutaminase renders GAMP insoluble as in the case of involucrin, because GAMP could be solubilized in 1% SDS containing 10 mM DTT. GAMP has no cysteine residues, and therefore there is no possibility of direct binding of GAMP with chitin through disulfide bonds. Details regarding the manner of the binding are still unknown. A model for a chitin-protein complex in insect cuticle has been proposed (Blackwell and Weih, 1980), but the detailed structure connecting these two large molecules has not been identified.

Amino acid composition of GAMP calculated from the deduced amino acid sequence was in good agreement with the analytical data in our previous report (Ishii *et al.*, 1998). The data estimated the existence of a moiety rich in glutamine and/or glutamate residues in GAMP. Most of the two residues are contained in the 10-amino-acid repeats as mentioned above. This moiety, which was predicted to form α-helical structures according to Chou-Fasman's method (Chou and Fasman, 1978), was not recovered from various enzymatic digests by reverse-phase HPLC possibly because of low solubility or due to unusual tertiary structure, which could be responsible for association with chitin or regulation of crystal growth. On the other hand, Chou-Fasman's method predicted that the 5-amino-acid-repeats near to the C-terminus would be abundant in random coils and turns, which suggested that the repeats could be relatively flexible within the structure of GAMP.

GAMP showed no similarity to any proteins other than involucrin thus far identified, which presented difficulties in estimating its function. However, it is of interest that both proteins are produced in the epidermis. GAMP mRNA was observed in tissues which were calcified and during periods in which calcification took place. GAMP had a unique repeated sequence of 10 amino acid residues with each unit containing two glutamate residues. It was assumed that GAMP possesses calcium binding property, because some known calcium binding proteins have carboxyl groups of Asp and/or Glu side chains which are responsible for calcium binding (Hauschka and Carr, 1982). GAMP had almost no calcium ion binding ability under the conditions which were used in the present experiment. However, GAMP is expected to have a weak affinity to calcium ion because of its ability to inhibit crystallization of calcium carbonate.

In Northern blot analysis, GAMP mRNA was detected in the gastrolith discs during premolt period and also in the fan blade during the intermolt period. Gastroliths are formed on the gastrolith discs during the premolt period and calcification of the cuticle is observed during the postmolt and intermolt periods. Hence, these results explain that GAMP mRNA is expressed in the calcifying tissues. This is interesting from a point of view relating to not only the process of calcification itself, but also to the question of calcium transportation as a process synchronized with the molting cycle. As a result of detailed analyses of the gastrolith discs, the increase in the hemolymph ecdysteroid titer after transition to the premolt period, as induced by bilateral eyestalk ablation, paralleled the levels of GAMP mRNA and the increase in weight of the gastroliths with a delay of several days. These results strongly suggest that GAMP is responsible for the calcification of the epidermis as well as of the gastroliths. The existence of GAMP in both tissues was also supported by immunohistochemical studies using an anti-GAMP antiserum (unpublished data). The stage-specific and tissue-specific expression of GAMP under the control of molting hormone indicates that GAMP would be a good marker for not only calcification but also postembryonic development. The relationship between hemolymph ecdysteroid titer and weight of gastroliths has already been reported (Sonobe *et al.*, 1991; Keller and Schmid, 1979). The expression of GAMP mRNA induced by molting hormone was demonstrated more directly by organ culture of the gastrolith discs. Formation of gastroliths is an essential event in the molting process, and it is reasonable to consider that the event is induced by molting hormone. By histological examination, it has been suggested that many ecdysteroid-binding sites appear in the gastrolith discs during the premolt period (Ueno *et al.*, 1992), and some genes induced by ecdysteroids would encode proteins involved in formation of gastroliths such as GAMP. On the other hand, what regulates the expression of GAMP mRNA in the epidermis has not been known. It is possible to assume that molt-inhibiting hormone induces the expression of GAMP mRNA in the epidermis and ecdysteroids suppress it contrary to the case of the gastrolith discs. In Northern blot analysis, a signal observed in the fan blade was approximately 3.2 kb (Fig. 4), which was 1.2 kb shorter than that of the gastrolith discs. This difference in size would result from deletion in the 3' untranslated region or from alternative splicing which generates different transcripts in each tissue. Alternatively, the signal would be mRNA originated from another gene which was homologous to GAMP.

Microheterogeneity in the amino acid sequence of GAMP was observed in our previous report (Ishii *et al.*, 1998). There was a large difference between the deduced amino acid sequence and partial sequence 1; lack of the sequence Phe-Gly-Ser-Pro-Ser-Leu-Gly-Ser-Ala-Ala between residues 441 and 442 of the deduced amino acid sequence. It is supposed that the deletion of two units of 5-amino-acid repeats, and this type of deletion, such as deletion of a repeating unit, would exist, because some microheterogeneities were observed in a GAMP preparation from a single gastrolith. The cDNA clone obtained in this study might be one derived from similar clones. The band detected in a series of Northern blot analyses gave the appearance of a singular entity and was considerably sharp with an approximate size of 4.4 kb in the gastrolith discs, which suggested that the heterogeneities were sufficiently small in size.

ACKNOWLEDGEMENTS

We are grateful to Mr. M. Takamiya for the ecdysteroid radioimmunoassay. We also thank Dr. Marcy N. Wilder of Japan International Research Center for Agricultural Sciences for critical reading of this manuscript. This work was supported partly by a grant for COE projects (1995 and 1996) from the

Ocean Research Institute of the University of Tokyo, and partly by a grant from the Takeda Science Foundation (to H. N.).

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410

Belcher AM, Wu XH, Christensen RJ, Hansma PK, Stucky GD, Morse DE (1996) Control of crystal phase switching and orientation by soluble mollusc-shell proteins. Nature 381: 56–58

Blackwell J, Weih MA (1980) Structure of chitin-protein complexes: ovipositor of the ichneumon fly *Megarhyssa*. J Mol Biol 137: 49–60

Borst DW, O'Conner JD (1972) Arthropod molting hormone: radioimmune assay. Science 178: 418–419

Brown JJ, Kiuchi M, Kainoh Y, Takeda S (1993) In vitro release of ecdysteroids by an endoparasitoid, *Ascogaster reticulatus* Watanabe. J Insect Physiol 39: 229–234

Chou PY, Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol 47: 45–148

Eckert R, Green H (1986) Structure and evolution of the human involucrin gene. Cell 46: 583–589

Falini G, Albeck S, Weiner S, Addadi L (1996) Control of aragonite or calcite polymorphism by mollusk shell macromolecules. Science 271: 67–69

George A, Sabsay B, Simonian PA, Veis A (1993) Characterization of novel dentin matrix acidic phosphoprotein: Implications for induction of biomineralization. J Biol Chem 268: 12624–12630

Gorski JP (1992) Acidic phosphoproteins from bone matrix: a structural rationalization of their role in biomineralization. Calcif Tissue Int 50: 391–396

Hauschka PV, Carr SA (1982) Calcium-dependent alpha-helical structure in osteocalcin. Biochemistry 21: 2538–2547

Ishii K, Yanagisawa T, Nagasawa H (1996) Characterization of a matrix protein in the gastrolith of the crayfish *Procambarus clarkii*. Biosci Biotech Biochem 60: 1479–1482

Ishii K, Tsutsui N, Watanabe T, Yanagisawa T, Nagasawa H (1998) Solubilization and chemical characterization of an insoluble matrix protein in the gastroliths of the crayfish, *Procambarus clarkii*. Biosci Biotech Biochem 62: 291–296

Kato-Fukui Y, Noce T, Ueda T, Fujiwara Y, Hashimoto N, Higashinakagawa T, Killian CE, Livingston BT, Wilt FH, Scov HM, Davidson EH (1991) The corrected structure of SM50 spicule matrix protein of *Strongylocentrotus purpuratus*. Dev Biol 145: 201–202

Keller R, Schmid E, (1979) In vitro secretion of ecdysteroids by Y-organs and lack of secretion by mandibular organs of the crayfish following molt induction. J Comp Physiol 130: 347–353

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685

Luquet G, Testeniére O, Graf F (1996) Characterization and N-terminal sequencing of a calcium binding protein from the calcareous concretion organic matrix of the terrestrial crustacean *Orchestia cavimana*. Biochim Biophys Acta 1293: 272–276

Maruyama K, Mikawa T, Ebashi S (1984) Detection of calcium binding proteins by ^{45}Ca autoradiography on nitrocellulose membrane after sodium dodecyl sulfate gel electrophoresis. J Biochem 95: 511–519

Miyamoto H, Miyashita T, Okushima M, Nakano S, Takashi M, Matsushiro A (1996) A carbonic anhydrase from the nacreous layer in oyster peals. Proc Natl Acad Sci USA 93: 9657–9660

Shen X, Belcher AM, Hansma PK, Stucky GD, Morse DE (1997) Molecular cloning and characterization of lustrin A, a matrix protein from shell and pearl nacre of Haliotis rufescens. J Biol Chem 272: 32472–32481

Simon M, Green H (1985) Enzymatic cross-linking of involucrin and other proteins by keratinocyte particulates in vitro. Cell 40: 677–683

Simon M, Green H (1988) The glutamine residues reactive in transglutaminase-catalyzed cross-linking of Involucrin. J Biol Chem 263: 18093–18098

Sonobe H, Kamba M, Ohta K, Ikeda M, Naya Y (1991) In vitro secretion of ecdysteroids by Y-organs of crayfish *Procambarus clarkii*. Experientia 47: 948–952

Stetler-Stevenson WG, Veis A (1983) Bovine dentin phosphophorin: Composition and molecular weight. Biochemistry 22: 4326–4335

Sudo S, Fujikawa T, Nagakura T, Ohkubo T, Sakaguchi K, Tanaka M, Nakashima K, Takahashi T (1997) Structures of mollusc shell framework proteins. Nature 387: 563–564

Thacher SM, Rice RH (1985) Keratinocyte-specific transglutaminase of cultured human epidermal cells: Relation to cross-linked envelope formation and terminal differentiation. Cell 40: 685–695

Travis DF (1960) The deposition of skeletal structures in the crustacea. I. The histology of the gastrolith skeletal tissue complex and the gastrolith in the crayfish, *Orconectes* (*Cambarus*) *virilis* Hagen-Decapoda. Biol Bull 118: 137–149

Ueno M (1980) Calcium transport in crayfish gastrolith disc: Morphology of gastrolith disc and ultrahistochemical demonstration of calcium. J Exp Zool 213: 161–171

Ueno M, Bidmon HJ, Stumpf WE (1992) Ecdysteroid binding sites in gastrolith forming tissue and stomach during the molting cycle of crayfish *Procambarus clarkii*. Histochemistry 98: 1–6

Von Heijine G (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14: 4683–4691

Wheeler AP, George JW, Evans CA (1981) Control of calcium carbonate nucleation and crystal growth by soluble matrix of oyster shell. Science 212: 1397–1398

(Received March 15, 1999 / Accepted April 7, 1999)