THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 17, pp. 12831–12839, April 23, 2010

A Protein Involved in the Assembly of an Extracellular Calcium Storage Matrix*

Received for publication, September 29, 2009, and in revised form, February 6, 2010 Published, JBC Papers in Press, February 11, 2010, DOI 10.1074/jbc.M109.071068

Lilah Glazer[‡§], Assaf Shechter[‡§], Moshe Tom[¶], Yana Yudkovski[¶], Simy Weil[‡], Eliahu David Aflalo[‡§], Ramachandra Reddy Pamuru[‖], Isam Khalaila[**], Shmuel Bentov[‡], Amir Berman[§], and Amir Sagi[‡§1]**

*From the [‡]Department of Life Sciences, [§]National Institute for Biotechnology in the Negev, and **Department of Biotechnology Engineering, Ben-Gurion University, P. O. Box 653, Beer-Sheva 84105, Israel, [¶]Israel Oceanographic and Limnological Research, Haifa 31096, Israel, and the [‖]Department of Biochemistry, Yogi Vemana University, Vemanapuram, Kadapa 516003, Andhra Pradesh, India*

Gastroliths, the calcium storage organs of crustaceans, consist of chitin-protein-mineral complexes in which the mineral component is stabilized amorphous calcium carbonate. To date, only three proteins, GAP 65, gastrolith matrix protein (GAMP), and orchestin, have been identified in gastroliths. Here, we report a novel protein, GAP 10, isolated from the gastrolith of the crayfish *Cherax quadricarinatus* and specifically expressed in its gastrolith disc. The encoding gene was cloned by partial sequencing of the protein extracted from the gastrolith matrix. Based on an assembled microarray cDNA chip, *GAP 10* transcripts were found to be highly (12-fold) up-regulated in premolt gastrolith disc and significantly down-regulated in the hypodermis at the same molt stage. The deduced protein sequence of GAP 10 lacks chitin-binding domains and does not show homology to known proteins in the GenBank™ data base. It does, however, have an amino acid composition that has similarity to proteins extracted from invertebrate and ascidian-calcified extracellular matrices. The GAP 10 sequence contains a predicted signal peptide and predicted phosphorylation sites. In addition, the protein is phosphorylated and exhibits calcium-binding ability. Repeated daily injections of *GAP 10* double strand RNA to premolt *C. quadricarinatus* resulted in a prolonged premolt stage and in the development of gastroliths with irregularly rough surfaces. These findings suggest that GAP 10 may be involved in the assembly of the gastrolith chitin-protein-mineral complex, particularly in the deposition of amorphous calcium carbonate.

Biomineralization in crustaceans has been and continues to be a subject of intensive study due to the key role of calcium metabolism in the periodic shedding of the exoskeleton during molting and its subsequent replacement. It is known that several crustacean species form temporary extracellular calcium carbonate storage deposits during the premolt stage (1). In freshwater crayfish, including our model organism the red claw crayfish *Cherax quadricarinatus*, these deposits include a pair of disc-like structures, known as gastroliths, that are located on each side of the stomach wall (2, 3). Gastroliths are also found in other decapods, including lobsters and several species of land crabs (1).

Gastrolith formation takes place in the gastrolith pouch, a cavity formed between the columnar epithelium of the gastrolith disc and the cardiac stomach wall. The main functions of the gastrolith disc epithelium are the transport of hemolymph calcium to the gastrolith and the synthesis of the gastrolith organic matrix (4). Like the exoskeleton, gastroliths consist of a chitinous organic matrix in which calcium carbonate is deposited (1, 5). In most crustacean species, the calcium carbonate of the exoskeleton exists in the form of one of two polymorphs, either the stable crystalline calcite (1, 6, 7) or the less stable amorphous calcium carbonate (ACC)[2] (8–12). We have reported previously that in *C. quadricarinatus* the major form of calcium present is ACC, not only in the gastroliths but also in the exoskeleton (12). It is likely that stabilization of otherwise unstable ACC requires unique conditions and the mediation of specialized macromolecules, such as proteins or peptides, within the extracellular matrix.

To date, only three transient storage organ proteins have been identified in crustaceans as follows: orchestin, gastrolith matrix protein (GAMP), and GAP 65. Orchestin is a calcium-binding phosphoprotein that was isolated from the organic matrix of the calcium storage organ of the terrestrial crustacean *Orchestia cavimana* (13). GAMP, which was isolated from the gastrolith of the crayfish *Procambarus clarkia* (14, 15), is believed to be a chitin-binding protein that inhibits calcium carbonate precipitation (16). GAP 65, which was extracted from late premolt gastroliths of *C. quadricarinatus*, is a calcium-binding glycoprotein, predicted to bind chitin. *In vivo* silencing of *GAP 65* resulted in the formation of gastroliths with morphological deformities and sparsely packed ACC spherules, suggesting that the protein has a dual function, *i.e.* both in the formation of the extracellular matrix and in the calcification process of the gastrolith (17). All three proteins have acidic pI values, and all are expressed in the premolt calcium stor-

* This work was supported in part by Israel Science Foundation Grant 1080/05, by German Israel Foundation Grant 950-9.5/2007, and by a grant from Amorphical Ltd.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) GQ231313.

[1] To whom correspondence should be addressed: Ben-Gurion University, P. O. Box 653, Beer-Sheva 84105, Israel. Tel.: 972-8-6461364; Fax: 972-8-6479062; E-mail: sagia@bgu.ac.il.

[2] The abbreviations used are: ACC, amorphous calcium carbonate; GAMP, gastrolith matrix protein; GAP, gastrolith protein; OPIM, optical phase interference microscopy; SH, Swedish height; MMI, molt mineralization index; GO, gene ontology; EST, expressed sequence tag; RT, reverse transcription; dsRNA, double strand RNA; MS, mass spectrometry.

age forming tissue, whereas GAMP (16, 18) and GAP 65 are also expressed in the hypodermis. These and additional studies performed on proteins involved in calcified chitinous matrices suggest the relevance of traits such as phosphorylation, calcium binding, acidity, and the ability to bind chitin to their performance.

In parallel to the studies of proteins isolated from chitinous matrices in crustaceans, a great deal of attention has been addressed to such proteins isolated from insect cuticles. To date, these studies have yielded a number of ubiquitous simple repeat sequences (19), *i.e.* the AAP(A/V) repeat and several glycine-rich regions such as GG*X* and/or GA repeats, which are also found in crustacean cuticular proteins (see the cuticleDB website). These repeats, and others such as A_n can also be found in spider silk proteins, where they are believed to be critical for the high tensile strength of spider silk (20).

Here, we report the identification and characterization of a novel protein obtained from the gastrolith matrix of premolt *C. quadricarinatus* and its encoding transcript. *In vivo* silencing of the transcript resulted in a prolongation of the premolt stage and the development of amorphous mineral-containing extracellular matrices with significant surface irregularities.

EXPERIMENTAL PROCEDURES

Animals and Molt—C. quadricarinatus males were grown in artificial ponds at Ben-Gurion University of the Negev, Beer-Sheva, Israel, under the conditions described in Shechter *et al.* (12). Intermolt crayfish were held in individual cages and endocrinologically induced to enter premolt through removal of the X organ-sinus gland complex. The progression of the molt cycle was monitored daily by measuring gastrolith molt mineralization index (MMI), as described by Shechter *et al.* (11). For all dissection procedures, crayfish were placed on ice for 5–10 min until they were anesthetized.

Purification of Gastrolith Proteins—Gastroliths were dissected from endocrinologically induced premolt crayfish, cleaned from cellular matter, and ground to powder in liquid nitrogen. Gastrolith proteins were extracted and separated following the procedure detailed in Shechter *et al.* (17). Briefly, EGTA-extracted gastrolith proteins were separated on a DEAE column. Elution was performed with a 0–1 M NaCl gradient.

Two-dimensional Separation and Visualization—The fractions collected at 100–200 mM NaCl were separated in two dimensions according to Arnott *et al.* (21) with a few modifications. Briefly, for isoelectric focusing, Immobiline dry strips (11 cm, pH 3–10) were rehydrated and aligned on the isoelectric focusing tray. Samples were loaded adjacent to the anode, and voltage was applied for a total of 11.9 kV-h. Following isoelectric focusing, the gel strips were equilibrated in a buffer containing 6 M urea, 0.375 M Tris-HCl, pH 8.8, 2% SDS, 20% glycerol, and either 2% dithiothreitol for reduction or 8% iodoacetamide for alkylation. The strips were then mounted on a 15% SDS-PAGE with the Tris/glycine running buffer system according to Laemmli *et al.* (22). Bands were visualized with Coomassie Blue staining.

Mass Spectrometry—Reduction, alkylation and trypsinization steps were carried out according to Rosenfeld *et al.* (23). For peptide sequencing, extraction of the fragments from the gel, mass spectrometry, and data analysis were performed according to Shechter *et al.* (17). For validation of GAP 10 from the two-dimensional electrophoresis, the tryptic digest was separated on a 75-μm internal diameter fused silica column, packed with C-18 (New Objective, Woburn, MA) connected to an Eksigent nano-LC system (Eksigent, Dublin, CA), and the peptides were eluted with the following solutions: buffer A was composed of 2% acetonitrile, 0.1% formic acid, and buffer B was composed of 80% acetonitrile in 0.1% formic acid, in nano-pure water. A linear gradient of 20–65% of buffer B was created over 45 min. MS peptide analysis and tandem MS fragmentation were performed using the LTQ-Orbitrap (Thermo Fisher Scientific, San Jose, CA). The mass spectrometer was operated in the data-dependent mode to switch between MS and collision-induced dissociation tandem MS of the top six ions. The collision-induced dissociation fragmentation was performed at 35% collision energy and 30-ms activation time. Protein identification and validation were performed using Sequest and Mascot algorithms operated under Proteome Discoverer 1.0 (Thermo Fisher Scientific) using an unpublished data base containing the GAP 10 sequence.

Detection of Protein Phosphorylation—Proteins separated by two-dimensional electrophoresis were fixed on the gel overnight, washed, and stained with PhosDecor™ phosphoprotein stain (Sigma) according to the manufacturer's protocol. Visualization was performed using a 300-nm UV transluminator (Chemi-Genius Bio-imaging System, Syngene, Cambridge, UK).

Detection of Calcium-binding Proteins—Calcium overlay procedure was performed according to Maruyama *et al.* (24). Briefly, proteins separated by SDS-PAGE were transferred to a nitrocellulose membrane. The membrane was washed three times with incubation buffer (60 mM KCl, 5 mM MgCl, and 10 mM imidazole HCl, pH 6.8) followed by incubation in the same buffer containing 10 μCi/25 ml $^{45}CaCl_2$ for 10 min. The membrane was rinsed with distilled water for 5 min. Excess water was absorbed between two sheets of Whatman filter paper, and the dried membrane was exposed to a Kodak BioMax MS x-ray film at −80 °C for 7 days. For both the detection of protein phosphorylation and the calcium-binding assay, 5 μg of casein (Sigma) was used as positive control.

Sequencing of GAP 10 Transcript Using Degenerative Primers—Degenerative primer GAP 10 AF, 5′-GAY GGN GGN TGG GAR TTY AC-3′, was generated based on a partial fragment sequence of GAP 10 obtained by tandem MS. A fragment of the cDNA was amplified using this forward primer and oligo$(dT)_{18}$VN as a reverse primer. The fragment was cloned into the pGEM-T Easy vector (Promega, Corp., Madison, WI) and sequenced. 5′- and 3′-rapid amplifications of cDNA ends were carried out with the Clontech SMART™ rapid amplification of cDNA ends kit (Clontech), according to the manufacturer's protocol to obtain the entire open reading frame sequence of *GAP 10*.

The deduced amino acid sequence of the gene was obtained bioinformatically by the ExPASy Translate tool, available on line. Signal peptide and phosphorylation sites were predicted using the SignalP 3.0 and NetPhosK 1.0, respectively, CBS Prediction Servers.

Construction of a cDNA Library of the Gastrolith Disc Using Suppression Subtractive Hybridization—Total RNAs of the gastrolith disc and muscle tissue were extracted from premolt-induced males using EZ-RNA total RNA isolation kit (Biological Industries, Beit Haemek, Israel). cDNA was prepared from 1 μg of total RNA using the Super SMART PCR cDNA synthesis kit (Clontech). The cDNAs were then used to prepare a suppression subtractive hybridization library of the gastrolith disc with the PCR Select cDNA subtraction kit, according to the manufacturer's instructions (Clontech). Briefly, the cDNA from gastrolith disc was used as the tester, and the cDNA from muscle tissue was used as the driver. After two hybridization cycles, unhybridized cDNAs, representing mostly genes that were expressed in the tester but absent from the driver, were amplified by two PCRs. The primary (27 cycles) and secondary (20 cycles) PCRs were performed with the Advantage™ 2 PCR kit (Clontech), and the PCR products were cloned into the pGEM-T Easy vector (Promega) and electrophoretically transformed into competent bacteria. Clones containing the inserts were isolated and grown overnight. Plasmid DNA was purified (Qiagen Miniprep kit) and sequenced.

Construction of a C. quadricarinatus cDNA Microarray, RNA Hybridization, and Data Analysis—The cDNA microarray used here was performed as described previously by our group (25) but broadened by the addition of 1000 cDNAs cloned from premolt gastrolith disc suppression subtractive hybridization library in the framework of the present study. Sequences were annotated by using both BLASTX and BLASTN comparisons (26). When available, gene ontology (GO) annotations (27) were borrowed from the BLAST hits, thereby further assisting the appropriate gene annotation. Both sequence and GO annotations were obtained with the BLAST2GO software (28). Only BLAST hits with *e*-values $<10^{-5}$ and their derived GO terms were considered valid annotations.

Expressed sequence tags (ESTs) identified as up-regulated during premolt following hybridizations with mRNA from the gastrolith disc were sequenced. To enhance the quality of the selected ESTs, the obtained cDNA sequences were first stripped of low quality, vector, and primer sequences using Sequencher™ software (GeneCodes Corp., Ann Arbor, MI). Clustering, assembly, and annotation of the remaining sequences were performed as elaborated in Yudkovski *et al.*[3] Unannotated genes were searched for *GAP 10* sequence similarity using a local installation of tBLASTx algorithm (NCBI).

The calculation and statistical analysis of $\log_2$ expression ratios for each clone in a binary comparison were carried out using the LIMMA package (Linear Models for Microarray data Analysis (29)) as described in Yudkovski *et al.*[3] For each spot, LIMMA calculated an M value, $M = (\log_2(\text{Cy5})/\text{Cy3})$, where Cy5 and Cy3 are the normalized emission intensities of the spot and an emission intensity A value ($A = (\log_2(\text{Cy5})\cdot\text{Cy3})/2$). Mean $|M|$ values of >2 were plotted against mean A values of >9 in expression scatter plots.

Tissue-specific Expression of GAP 10—To identify tissue-specific *GAP 10* expression, RNA was extracted from the gastrolith disc, hypodermis, hepatopancreas, muscle, and sperm duct as described above. First strand cDNA was generated with oligo$(dT)_{18}$VN using expand RT (Roche Diagnostics). PCR was performed with the following primers: *GAP 10*-F 5′-CTTGGTGCTGGCCAGGTGGGAGGTGCTG-3′ and *GAP 10*-R 5′-CAGCCTCCGTCAGACTCC ACGCGCACAT-3′.

In Situ Hybridization—Gastrolith discs were dissected out from an induced premolt crayfish 5 days after induction through removal of the X organ-sinus gland. Fixation, paraplast embedding, sectioning, and the hybridization procedure were conducted as described by Ventura *et al.* (30). Digoxygenin-labeled oligonucleotides for antisense and sense probes corresponding to nucleotides 31–403 of *GAP 10* cDNA were synthesized using SP6 and T7 RNA polymerases, as described in the digoxygenin application manual (Roche Diagnostics). Hybridization was carried out as described previously by Shechter *et al.* (31), with the slight modification of adding 100 μg/ml tRNA to the hybridization solution.

Physicochemical Properties of GAP 10—The physicochemical properties of GAP 10, including molecular weight and amino acid composition, were deduced from its translated protein sequence using the ExPASy ProtoParam tool (32).

In Vivo Silencing of GAP 10 Transcripts during Premolt—Intermolt males (MMI = 0), each weighing 5–10 g, were divided to four groups. Ten animals were injected daily with both ecdysone (Sigma), 1 ng/μl (17), and dsRNA of *GAP 10*, 5 μg/g; eight were injected with ecdysone, 1 ng/μl; six were injected with both ecdysone and dsRNA of *CqVg*, 5 μg/g; and four were injected with ecdysone carrier only, *i.e.* 10% ethanol in diethyl pyrocarbonate-treated doubly distilled water. Injections were given into the sinus of the first abdominal segment. For each individual crayfish, the experiment was terminated at MMI = 0.1, at which time the animal was anesthetized, and its gastroliths were dissected out. Premolt duration was defined as the number of days until termination.

Optical Phase Interference Microscopy (OPIM) Analysis—OPIM was performed using the NewView 200 (Zygo Corp., Middlefield, CT) microscope, for analysis of gastrolith surface roughness. Four to eight different locations on the surface of each dissected gastrolith were analyzed. A cylindrical reference surface was fitted to and subtracted from each image. Swedish height (SH) was calculated as the height difference between two reference values on a topographic image; the upper reference was positioned at a depth that exposes 5% of the data and the lower reference at a depth that exposes 90% of the data.

Statistical Analysis—Premolt duration, SH, and relative transcript level data were expressed as means ± S.E. Statistical analysis for premolt duration and SH was performed using one-way analysis of variance, followed by Fisher's LSD multiple comparison test. For relative transcript levels, Mann-Whitney U test was used. p values < 0.05 were considered statistically significant.

Relative Quantitative Real Time RT-PCR—RNA was extracted as described above. First strand cDNA was generated by reverse transcription using random hexamer primers (Verso™ cDNA kit, Thermo Fisher Scientific). Relative quantification was performed with the following primers: qGAP_10_F, 5′-AAAAACACAGGTTCTGTTGCTCT-3′ and qGAP_10_R,

[3] Y. Yudkovski, L. Glazer, A. Shechter, A. R. Reinhardt, V. Chalifa-Caspi, A. Sagi, and M. Tom, submitted for publication.

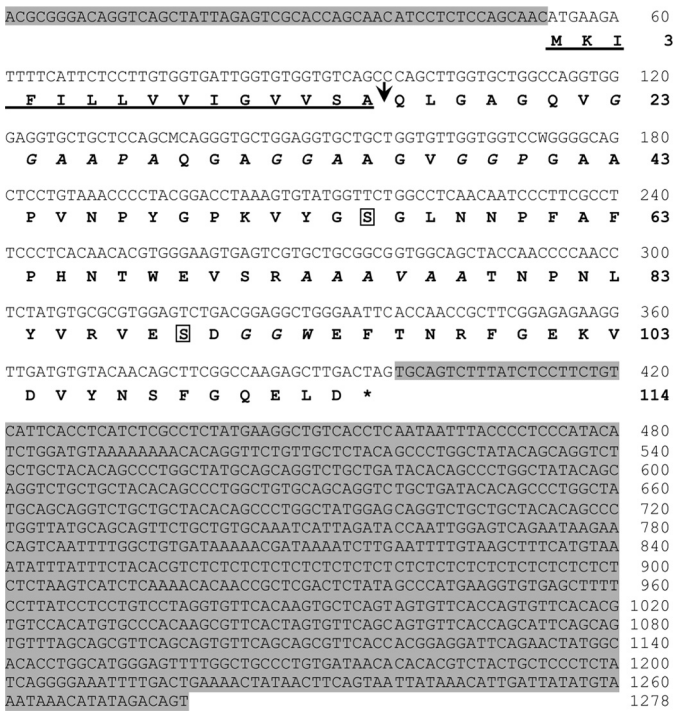

FIGURE 1. **Nucleotide sequence of *GAP 10* cDNA and deduced amino acids of its open reading frame.** The 5′- and 3′-untranslated regions are highlighted in *gray*. The putative signal peptide in the N terminus is *underlined*. *Arrowhead* indicates signal peptide cleavage site. Predicted phosphorylation sites are *boxed*. *Italic letters* indicate AAP(A/V), GG*X*, and A_n consensus sequences. The *asterisk* indicates a stop codon.

5′-TAGCCAGGGCTGTGTAGCA-3′. 18 S rRNA, used as the normalizing agent, was also evaluated by real time RT-PCR at similar conditions using the following primers: q18 S_F, 5′-CTGAGAAACGGCTACCACATC-3′, and q18 S_R, 5′-GCCGGGAGTGGGTAATTT-3′. Real time reactions were performed using Universal ProbeLibrary Probes 62 and 74, respectively (Roche Diagnostics), with an ABI Prism 7000 sequence detection system (Applied Biosystems, Foster City, CA).

Accession Numbers—GAP 10 sequence data has been deposited in the GenBank™ Data Bank with accession number GQ231313. The microarray platform was deposited in the Gene Expression Omnibus (GEO) data base of the NCBI (GEO accession number GPL8761), and the gastrolith disc and hypodermal dataset were assigned the accession number GSE16866.

RESULTS

Identification and Sequencing of GAP 10 and Bioinformatic Analysis of the Deduced Protein—The protein fraction extracted from the gastrolith and eluted from a DEAE column with 100–200 mM NaCl revealed a prominent band, migrating at apparent molecular mass of ~11 kDa (data not shown). The transcript was successfully sequenced using tandem MS of trypsin-digested fragments, followed by nanospray QTof2-based degenerative primers, followed by specific primers for 5′- and 3′-rapid amplification of cDNA ends (Fig. 1). The 1278-bp transcript included a 5′-untranslated region of 53 bp, an open reading frame of 342 bp encoding a deduced 114 amino acid protein sequence, and a 3′-untranslated region of 880 bp. The first 15 amino acids included a signal peptide, as predicted by CBS Prediction Servers, with a predicted cleavage site between Ala^{-1} and Gln^{1}. The deduced mature protein has a calculated molecular mass of 10 kDa and was therefore termed GAP 10. According to a kinase-specific phosphorylation site prediction software, it is suggested that Ser^{40} and Ser^{74} are phosphorylated by protein kinase A.

TABLE 1
Physicochemical properties calculated for the GAP 10-deduced protein
Amino acid composition is categorized according to side chain properties. Percentage of amino acids corresponding to the total of each category is indicated in parentheses.

Amino acid class	Amino acid	GAP 10
Nonpolar aliphatic	Gly	18
	Ala	15
	Val	10
	Leu	4
	Ile	0
	Met	0
	Total	47 (47.5%)
Aromatic	Phe	5
	Tyr	4
	Trp	2
	Total	11 (11.1%)
Polar uncharged	Ser	4
	Pro	8
	Thr	3
	Cys	0
	Asn	8
	Gln	4
	Total	27 (27.2%)
Positively charged	Lys	2
	His	1
	Arg	3
	Total	6 (6%)
Negatively charged	Asp	3
	Glu	5
	Total	8 (8.1%)
Total amino acids	99	
Asx	11	
Glx	9	
Total	20 (20.1%)	

GAP 10 was not found to be similar to any protein or translated sequence in the GenBank™ data base. However, it does contain two known consensus sequences identified in arthropod cuticular proteins and in spider silk, *i.e.* one copy of the AAP(A/V) (residues 25–28) repeat and four copies of the glycine-rich GG*X* (residues 23–25, 32–34, 38–40, and 91–93) repeat, and an additional A_n motif starting at residue 73.

Table 1 presents the predicted physicochemical properties of the deduced protein. It has a relatively high percentage (43.5%) of the nonpolar aliphatic amino acids, Gly (18.2%), Ala (15.2%), and Val (10.1%) and also of the polar but uncharged amino acids Asn (8.1%) and Pro (8.1%).

Identification of GAP 10 as a Calcium-binding Phosphoprotein—The protein fraction extracted from the gastrolith and that contained GAP 10 was separated first by isoelectric focusing and then on SDS-PAGE, revealing several distinct proteins (Fig. 2*A*, *left*). A number of these proteins were found to be phosphorylated (Fig. 2*A*, *middle*) and to have calcium-binding capacity (Fig. 2*A*, *right*). Proteins that were both phosphorylated and calcium binding were prepared for liquid chromatography-MS. Two of them, migrating at apparent molecular

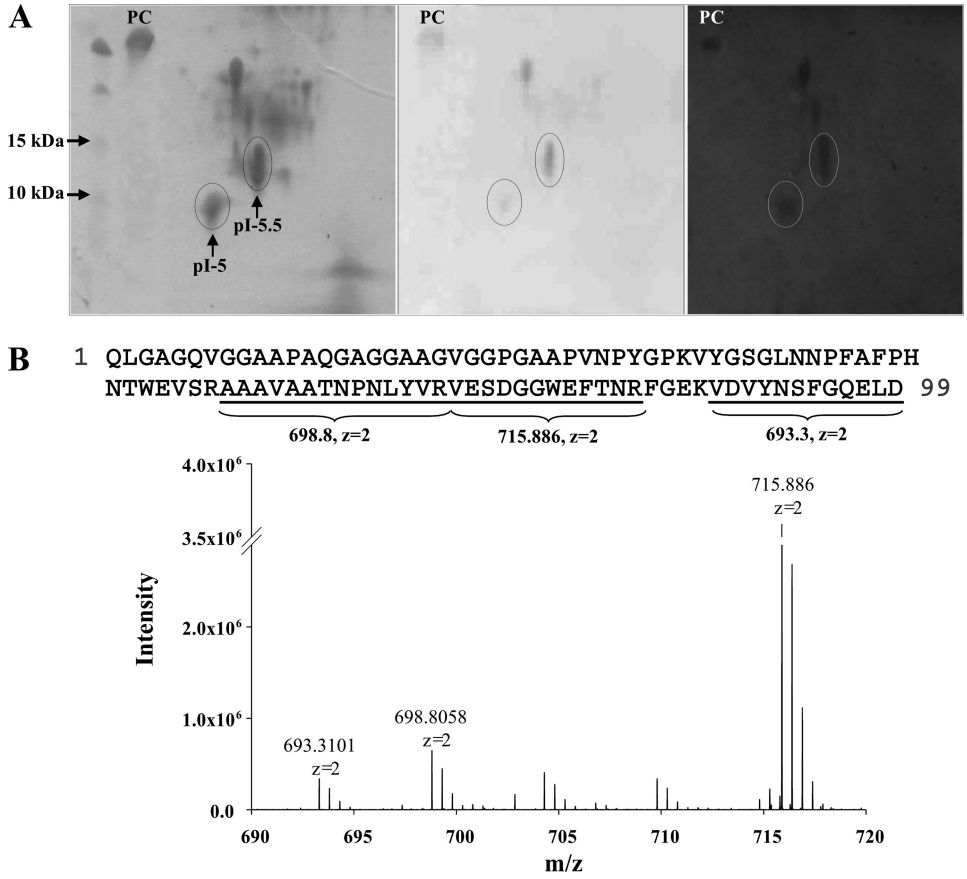

FIGURE 2. **Identification of GAP 10 and characterization of its phosphorylation and calcium-binding properties.** *A, left,* GAP 10-enriched fraction was separated first by isoelectric focusing and then on SDS-polyacrylamide gel. Molecular weight standard (*1st lane*) and casein (*2nd lane*) were added to the SDS-polyacrylamide gel. The gel was stained with Coomassie Blue. *Middle,* staining of the same gel as in *A* for phosphoproteins. *Right,* transfer of the same fraction as in *A* to a nitrocellulose membrane and incubation with $^{45}Ca^{2+}$. Spots identified as GAP 10 by liquid chromatography-MS are *circled*. Positive control (*PC*) is casein (5 μg). *B,* identification of GAP 10 peptides by liquid chromatography-MS. The identified peptides are *underlined* on the GAP 10 sequence, and the related peaks are indicated on the *spectrum*.

masses of ~9 and ~11 kDa (Fig. 2*A*, *circled*), were identified as GAP 10 (Fig. 2*B*), and their pI values were calculated as 5 and 5.5, respectively. Because GAP 10 has a mildly acidic pI, it is negatively charged at the pH of the gastrolith pouch, which is ~8.7 (12).

Microarray Analysis of GAP 10 Regulation in the Gastrolith Disc and Hypodermis during Premolt—Hybridizations of mRNA from the gastrolith disc of ecdysone-induced premolt animals *versus* intermolt control animals to the *C. quadricarinatus* cDNA microarray[3] revealed prominent up-regulation of *GAP 10* transcripts at premolt (Fig. 3*A*). *GAP 10* transcripts were up-regulated up to 65-fold, with an average of ~12-fold, and included ~12.7% of the total of up-regulated transcripts identified in this experiment. Hybridizations under the same conditions of mRNA from the hypodermal tissue of the same animals at the same molt stages revealed distinct down-regulation of *GAP 10* transcripts in this tissue, with an average of 7-fold, including ~9.2% of the total number of down-regulated transcripts identified in this experiment (Fig. 3*B*).

Tissue-specific Expression and Localization of GAP 10—Specific expression and localization of *GAP 10* were tested in premolt crayfish in a variety of target tissues by RT-PCR (Fig. 4*A*) and *in situ* hybridization (Fig. 4*B*). Both methods showed expression of *GAP 10* to be specific to the gastrolith disc during premolt. The transcript, showing a strong, specific signal using an antisense probe, was localized to the columnar epithelium of the gastrolith disc and was not detected in the adjacent muscle tissue (Fig. 4*B*). No signal was detected with the sense-strand probe.

In Vivo Silencing of GAP 10 Transcripts during Premolt—Repeated daily injections of ecdysone together with *GAP 10* dsRNA to intermolt animals resulted in an increase in premolt duration (the number of days until reaching MMI = 0.1, see under "Experimental Procedures") from an average of 10.1 days in the ecdysone-injected control group to 13.1 days in the group receiving ecdysone and *GAP 10* dsRNA, with the difference between the two groups being significant ($p < 0.05$) (Fig. 5). Another group was injected with ecdysone and *CqVg* dsRNA (a hepatopancreatic-specific gene found mostly in reproductive females) and served as a control for sequence-specific silencing. The average premolt duration in this group was 8.2 days, which is significantly lower than the value calculated for the ecdysone and *GAP 10* dsRNA-injected group ($p < 0.05$), but not significantly different from the one calculated for the ecdysone-injected control. It was also found that *GAP 10* silencing resulted in development of gastroliths with significant surface irregularities (Fig. 6*A*) as opposed to the fairly smooth gastroliths that developed in animals injected with ecdysone alone. For analysis of the gastrolith surface roughness, OPIM analysis was used (Fig. 6*B*). SH calculations on the topographic images obtained by OPIM revealed a significant difference ($p < 0.001$) between the ecdysone control group (4.4 ± 0.5 μm) and the ecdysone/*GAP 10* dsRNA-injected group (11.3 ± 1 μm) (Fig. 6*C*, *right*). *GAP 10* transcript levels in the ecdysone/*GAP 10* dsRNA-injected group were significantly lower than those found in the ecdysone control group ($p < 0.01$) (Fig. 6*C*, *left*).

DISCUSSION

In this study, a novel protein encoded by *GAP 10* was identified from the extracellular gastrolith matrix. Similarly to orchestin, GAMP, and GAP 65, GAP 10 has an acidic pI. The deduced protein sequence contains a predicted signal peptide, in accordance with the protein identification in the extracellular gastrolith matrix. Unlike GAP 65 (17), the premolt transcript expression of *GAP 10* is specific to the gastrolith disc, as

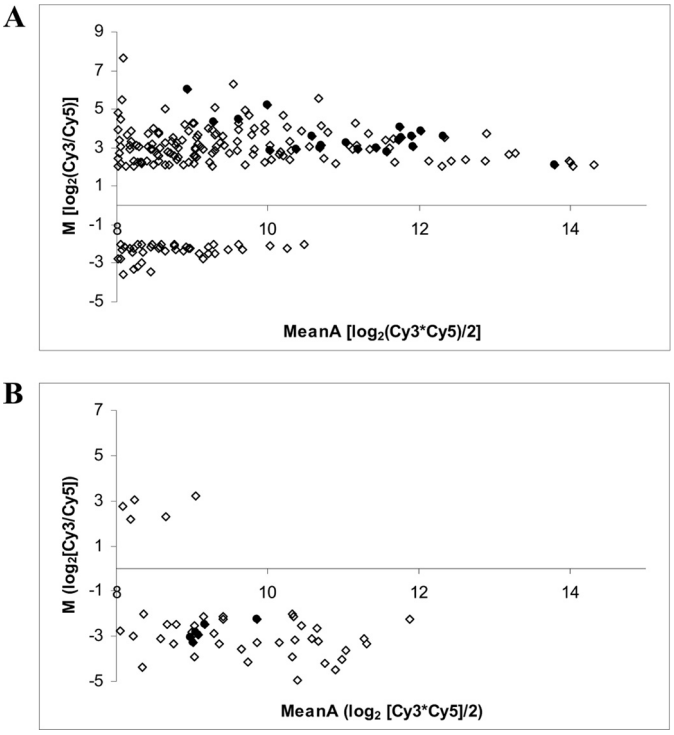

FIGURE 3. **Multigenic expression pattern of gastrolith disc (*A*) and hypodermis (*B*) in premolt *versus* intermolt crayfish.** Expression scatter plots of all the ESTs were identified as being differentially expressed between the treatment and the control. *M*, $\log_2$-fold change of normalized emission intensity between the treatment and the control filtered by $|M| > 2$; mean *A*, $\log_2$ of average signal intensity filtered by mean $A > 9$. Cy3 (premolt) and Cy5 (intermolt-control) are normalized microarray signals. *Empty diamonds* represent all differentially expressed ESTs. *Full circles* represent *GAP 10* ESTs.

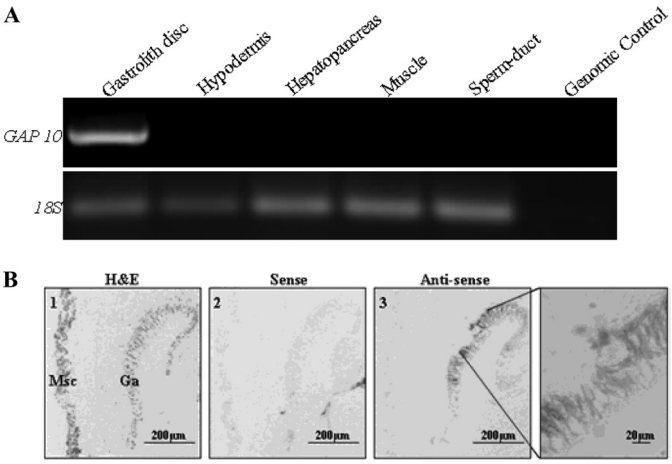

FIGURE 4. **Specific expression of *GAP 10* in premolt gastrolith disc as demonstrated by RT-PCR (*A*) and localization of *GAP 10* expression to the columnar epithelium forming the gastrolith by *in situ* hybridization (*B*).** Total RNA was extracted from the gastrolith disc, hypodermis, hepatopancreas, muscle, and sperm duct. Ribosomal 18 S unit was used to confirm RNA extraction. Genomic control with gastrolith disc RNA was used. *B*, *1st lane,* hematoxylin and eosin (*H&E*) staining. *2nd lane,* tissue probed with the negative control sense, *GAP 10* probe. *3rd lane,* tissue probed with the *GAP 10* antisense probe, with the far right image corresponding to an enlargement of a specific area. *Ga*, gastrolith disc; *Msc*, muscle tissue. *Scale bar*, 200 μm, except for in the *enlarged box*, where it represents 20 μm.

shown by RT-PCR and by *in situ* localization. Moreover, microarray hybridizations revealed that the transcript of *GAP 10* is notably highly up-regulated in the gastrolith disc during premolt, while being significantly down-regulated in the hypodermis, emphasizing its involvement in the construction process of the gastrolith matrix. The expression pattern of *GAP 10* is shared by the majority of the gastrolith transcripts, suggesting a mirror image pattern of regulation for the expression of structural and other genes in these two epithelial tissues during the molt cycle. Such a pattern coincides with the known opposite functions of these tissues, specifically with regard to calcification (12). During premolt, the hypodermis absorbs calcium from the cuticle, whereas the gastrolith disc deposits that same calcium into the gastrolith. The expression of *GAP 10* in the hypodermis of intermolt animals, as demonstrated by the microarray results, suggests a role for the protein during this particular molt stage.

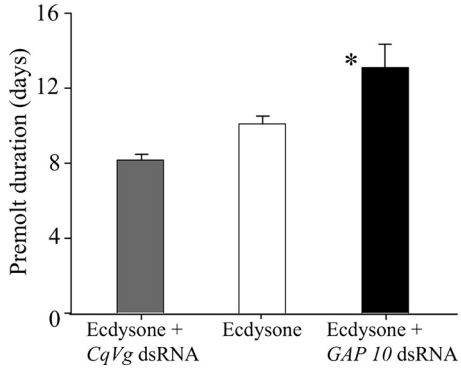

FIGURE 5. **Prolongation of premolt following *GAP 10* silencing.** Mean premolt duration (the number of days until experiment termination) was calculated for each experimental group injected with ecdysone ($n = 8$; *empty bars*) or ecdysone and GAP 10 dsRNA ($n = 10$; *black bars*) or ecdysone and CqVg dsRNA ($n = 6$; *gray bars*). *Asterisk* represents statistically significant difference of $p < 0.05$.

GAP 10 was not found to have significant similarity to any known protein in the GenBank™ data base. However, it does contain several known consensus sequences previously identified in arthropod extracellular structural proteins (19), including the AAP(A/V) repeat. This consensus sequence is a small, hydrophobic tetrapeptide, which occurs mainly in proteins of hard cuticles (33), in which water content is low and sclerotization is intense. It has been suggested that the importance of this element lies in its role in the optimal folding of these proteins (34), because the Pro residue induces bends in the peptide chain, and the hydrophobic residues stabilize the structure through hydrophobic interactions. Glycine-rich regions in cuticular proteins may be expected to have a high tendency for forming β-turns, where the turns will be stabilized by hydrogen bonds, and the side chains of the hydrophobic residues can interact hydrophobically. Both the AAP(A/V) consensus and the glycine-rich regions are thought to confer elasticity and high tensile strength on matrix proteins.

Amino acid composition analysis of GAP 10 revealed abundant nonpolar, aliphatic amino acids, Gly, Ala, and Val, and also polar but uncharged amino acids, Asn and Pro. An abundance of Ala and Gly is also characteristic of GAMP (16). Similar amino acid composition profiles are present in several crustacean cuticular proteins, such as CAP-1 and CAP-2, both identified from the exoskeleton of the crayfish *P. clarkii* (35, 36). These proteins are highly acidic and contain abundant Gly and Val. CAP-1 is also abundant in Ala. Both are calcification-asso-

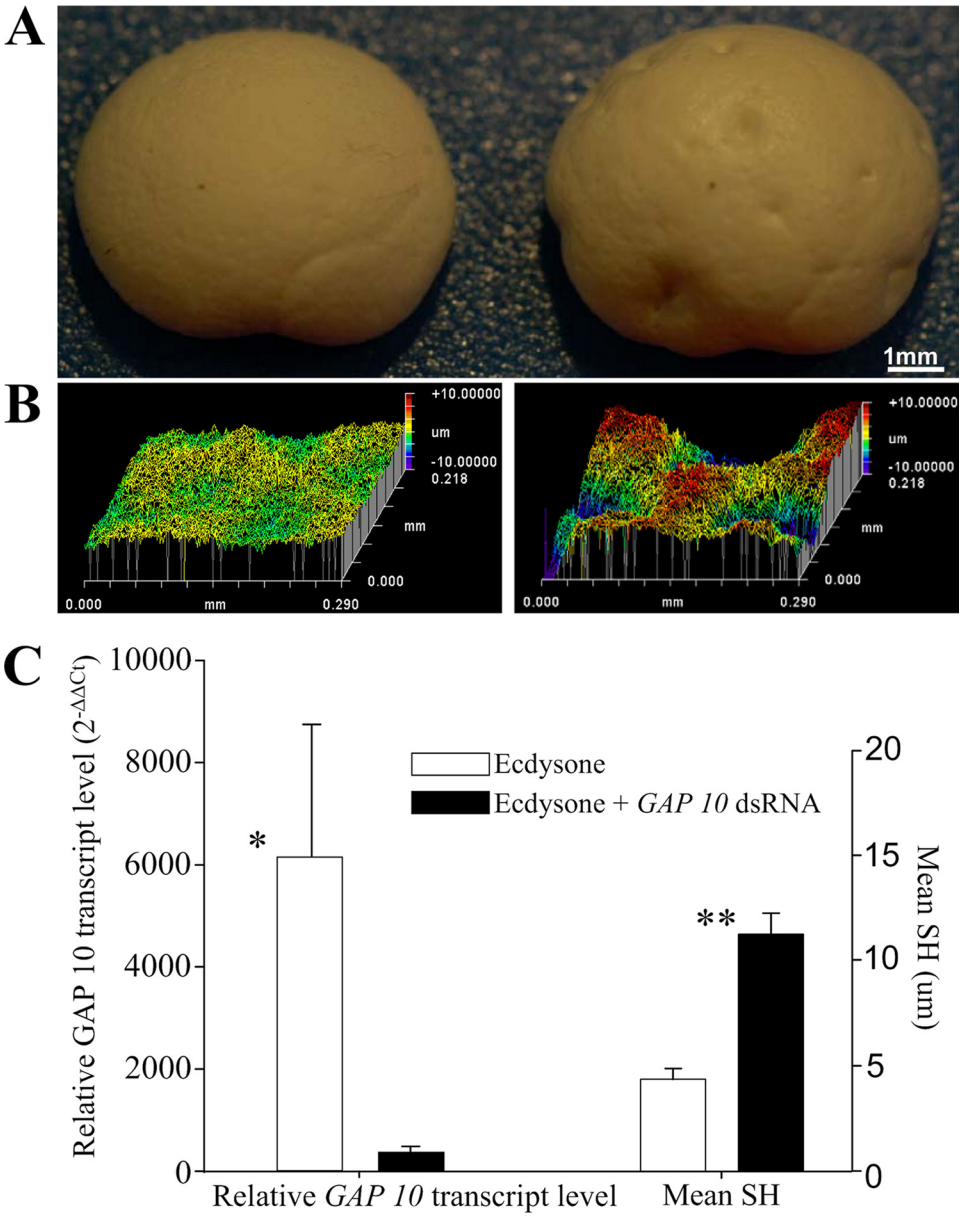

FIGURE 6. **Visualization and OPIM measurement of gastrolith surface irregularities and relative transcript levels of *GAP 10* in the gastrolith disc following *GAP 10* silencing.** *A, top view* of representative gastroliths, dissected from crayfish injected with ecdysone (*left*) or ecdysone and *GAP 10* dsRNA (*right*). *B,* topographic images of representative surface samples from each group. *C,* real time RT-PCR relative quantification of *GAP 10* transcript levels in the gastrolith disc as calculated for each experimental group injected with ecdysone ($n = 5$; *empty bars*) or ecdysone and GAP 10 dsRNA ($n = 6$; *black bars*) and mean SH as calculated for the two groups ($n = 3$ and $n = 4$, respectively). *Asterisk* represents statistically significant difference of $p < 0.01$. *Two asterisk* represent statistically significant difference of $p < 0.001$.

Both GAMP and GAP 65 are believed to bind chitin; GAMP could be extracted from a chitin-rich insoluble gastrolith matrix only by using aggressive conditions, trypsin digestion and 6 M urea, which suggest a tight association between the protein and the matrix (14). In GAP 65, the amino acid sequence contains a chitin-binding domain 2, also known as a peritrophin-A domain (17). In arthropods, there are two known chitin-binding domains, chitin-binding domain 2 (46) and the Rebers-Riddiford motif (1, 47, 48), both of which have also been reported in crustaceans. GAP 10 was extracted from the EGTA-soluble fraction of the gastrolith, and the deduced protein sequence does not have predicted chitin-binding domains either of the chitin-binding domain 2 or of the Rebers-Riddiford type. This lack of a chitin-binding domain suggests roles for GAP 10 that do not require a direct association with the chitinous matrix, such as participation in calcium carbonate precipitation.

GAP 65 was shown both to have calcium-binding ability and to play a role in the calcification process of the gastrolith by stabilization of synthetic ACC (17). GAMP inhibits calcium carbonate precipitation *in vitro* (16), and orchestin has calcium-binding ability dependent on phosphorylation of Ser (49). CAP-1, identified from the exoskeleton of *P. clarkii* (35), is an inhibitor of calcium carbonate precipitation, a property found to be associated with the phosphorylation of a specific Ser residue (50). GAP 10 was found to have calcium-binding ability and to be phosphorylated, with two predicted phosphorylation sites at Ser residues. The importance of protein phosphorylation for the inhibition of calcium carbonate precipitation has also been shown for the acidic, soluble organic matrix protein RP-1, isolated from the shells of the Antarctic scallop *Adamussium colbecki* (51); dephosphorylation of this protein resulted in almost complete loss of its inhibitory activity. This property has yet to be tested for GAP 10.

The *in vivo* silencing of *GAP 10* was followed by a considerable delay in premolt duration and the development of gastroliths with significant surface irregularities. Similarly, it was previously found that injection of *GAP 65* dsRNA into *C. quadricarinatus* resulted in surface irregularities in addition to

ciated peptide (CAP) proteins and are calcium-binding entities that inhibit calcium carbonate precipitation. Two other highly similar genes, expressed in the epithelial cells of the tail fan of the penaeid prawn *Penaeus japonicus,* are *DD9A* and *DD9B* (37). Both DD9-deduced proteins have acidic pI values, and their amino acid compositions are very similar to that of GAP 10, with regard to the richness in Gly, Ala, Val, and Pro. Moreover, this pattern of amino acid composition can be found in organic matrices and several proteins extracted from calcium carbonate precipitations of other invertebrate groups, such as echinoderms (38–40), mollusks (41–44), and also ascidians (45). The most abundant amino acids are Asx, Glx, and Gly, and in some matrices also Ser, Ala, and Pro.

much deeper structural deformities of the gastroliths (17). Scanning electron microscopy of fractures of gastroliths from *GAP 10*-silenced animals showed no apparent evidence for such penetrating structural deformities (data not shown).

Administration of ecdysone has been shown previously to induce molt in *C. quadricarinatus* (11, 35, 36). The typical induced premolt duration is 10–14 days until ecdysis, and the peak observed MMI is 0.125–0.145, which is reached 1–2 days before the molt event. In this study, the injections were terminated at MMI = 0.1 so as to prevent molting, and premolt duration was calculated up to that point. Therefore, the ~10-day mean premolt duration of the ecdysone-induced, nonsilenced animals was within the expected range, whereas the ~13-day mean premolt duration of the silenced animals implies a significant delay, which together with the structural effects suggests a multifaceted role for GAP 10.

We suggest that GAP 10 plays a crucial role in gastrolith formation, because depletion of the protein secreted into the matrix, following transcript silencing, significantly prolonged premolt and was manifested in irregularities appearing exclusively on the surface of the gastroliths, representing the most recently deposited layers. However, it should be noted that the lack of a sequenced genome for decapod crustaceans makes it impossible to reliably predict potential effects of small interfering RNAs within target cells on additional transcripts.

The above-discussed characteristics of GAP 10 suggest significant involvement of the protein in the formation of the chitin-protein-mineral complex of the gastrolith, especially with regard to the deposition of calcium carbonate. Because there is no indication for a possible association of GAP 10 with the chitinous matrix, protein-protein interactions between GAP 10 and a chitin-binding protein could be hypothesized. However, further studies are needed to unravel the role of this protein in the context of the gastrolith model used here, as a relatively simple model for the more complex biomineralization processes occurring throughout the animal kingdom.

Acknowledgments—We thank Shahar Shimshon and Roni Nisan for technical assistance, Juergen Jopp for the OPIM analysis, and Inez Mureinik for styling the manuscript. Animals for the study were supplied by Yossi Ben of Beer Tzofar and by Ayana Benet Perlberg of The Ministry of Agriculture and Rural Development, Department of Fisheries and Aquaculture, Dor Agriculture Center.

REFERENCES

1. Luquet, G., and Marin, F. (2004) *Comp. Rend. Palev.* **3,** 515–534
2. Travis, D. F., and Friberg, U. (1963) *J. Ultrastruct. Res.* **59,** 285–301
3. Travis, D. F. (1960) *Biol. Bull.* **16,** 137–149
4. Ueno, M. (1980) *J. Exp. Zool.* **213,** 161–171
5. Roer, R., and Dillaman, R. (1984) *Am. Zool.* **24,** 893–909
6. Yano, I. (1975) *Bull. Jpn. Soc. Fish Oceanogr.* **41,** 1079–1082
7. Pratoomchat, B., Sawangwong, P., Guedes, R., Reis, Md., Mde, L., and Machado, J. (2002) *J. Exp. Zool.* **293,** 414–426
8. Sugawara, A., Nishimura, T., Yamamoto, Y., Inoue, H., Nagasawa, H., and Kato, T. (2006) *Angew. Chem. Int. Ed. Engl.* **45,** 2876–2879
9. Greenaway, P. (1985) *Biol. Rev. Camb. Philos. Soc.* **60,** 425–454
10. Addadi, L., Raz, S., and Weiner, S. (2003) *Adv. Mater.* **15,** 959–970
11. Shechter, A., Tom, M., Yudkovski, Y., Weil, S., Chang, S. A., Chang, E. S., Chalifa-Caspi, V., Berman, A., and Sagi, A. (2007) *J. Exp. Biol.* **210,** 3525–3537
12. Shechter, A., Berman, A., Singer, A., Freiman, A., Grinstein, M., Erez, J., Aflalo, E. D., and Sagi, A. (2008) *Biol. Bull.* **214,** 122–134
13. Testenière, O., Hecker, A., Le Gurun, S., Quennedey, B., Graf, F., and Luquet, G. (2002) *Biochem. J.* **361,** 327–335
14. Ishii, K., Yanagisawa, T., and Nagasawa, H. (1996) *Biosci. Biotechnol. Biochem.* **60,** 1479–1482
15. Ishii, K., Tsutsui, N., Watanabe, T., Yanagisawa, T., and Nagasawa, H. (1998) *Biosci. Biotechnol. Biochem.* **62,** 291–296
16. Tsutsui, N., Ishii, K., Takagi, Y., Watanabe, T., and Nagasawa, H. (1999) *Zool. Sci.* **16,** 619–628
17. Shechter, A., Glazer, L., Cheled, S., Mor, E., Weil, S., Berman, A., Bentov, S., Aflalo, E. D., Khalaila, I., and Sagi, A. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 7129–7134
18. Takagi, Y., Ishii, K., Ozaki, N., and Nagasawa, H. (2000) *Zool. Sci.* **17,** 179–184
19. Futahashi, R., Okamoto, S., Kawasaki, H., Zhong, Y. S., Iwanaga, M., Mita, K., and Fujiwara, H. (2008) *Insect Biochem. Mol. Biol.* **38,** 1138–1146
20. Gatesy, J., Hayashi, C., Motriuk, D., Woods, J., and Lewis, R. (2001) *Science* **291,** 2603–2605
21. Arnott, D., O'Connell, K. L., King, K. L., and Stults, J. T. (1998) *Anal. Biochem.* **258,** 1–18
22. Laemmli, U. K. (1970) *Nature* **227,** 680–685
23. Rosenfeld, J., Capdevielle, J., Guillemot, J. C., and Ferrara, P. (1992) *Anal. Biochem.* **203,** 173–179
24. Maruyama, K., Mikawa, T., and Ebashi, S. (1984) *J. Biochem.* **95,** 511–519
25. Yudkovski, Y., Shechter, A., Chalifa-Caspi, V., Auslander, M., Ophir, R., Dauphin-Villemant, C., Waterman, M., Sagi, A., and Tom, M. (2007) *Insect Mol. Biol.* **16,** 661–674
26. Kuballa, A. V., Merritt, D. J., and Elizur, A. (2007) *BMC Biol.* **5,** 45
27. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000) *Nat. Genet.* **25,** 25–29
28. Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005) *Bioinformatics* **21,** 3674–3676
29. Smyth, G. K. (2004) *Stat. Appl. Genet. Mol. Biol.* **3,**
30. Ventura, T., Manor, R., Aflalo, E. D., Weil, S., Raviv, S., Glazer, L., and Sagi, A. (2009) *Endocrinology* **150,** 1278–1286
31. Shechter, A., Aflalo, E. D., Davis, C., and Sagi, A. (2005) *Biol. Reprod.* **73,** 72–79
32. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) in *The Proteomics Protocols Handbook* (Walker, J. M., ed) pp. 571–607, Humana Press Inc., Totowa, NJ
33. Magkrioti, C. K., Spyropoulos, I. C., Iconomidou, V. A., Willis, J. H., and Hamodrakas, S. J. (2004) *BMC Bioinformatics* **5,** 138
34. Andersen, S. O., Højrup, P., and Roepstorff, P. (1995) *Insect Biochem. Mol. Biol.* **25,** 153–176
35. Inoue, H., Ozaki, N., and Nagasawa, H. (2001) *Biosci. Biotechnol. Biochem.* **65,** 1840–1848
36. Inoue, H., Ohira, T., Ozaki, N., and Nagasawa, H. (2004) *Biochem. Biophys. Res. Commun.* **318,** 649–654
37. Watanabe, T., Persson, P., Endo, H., and Kono, M. (2000) *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **125,** 127–136
38. Gayathri, S., Lakshminarayanan, R., Weaver, J. C., Morse, D. E., Kini, R. M., and Valiyaveettil, S. (2007) *Chemistry* **13,** 3262–3268
39. Benson, S. C., Benson, N. C., and Wilt, F. (1986) *J. Cell Biol.* **102,** 1878–1886
40. Raz, S., Hamilton, P. C., Wilt, F. H., Weiner, S., and Addadi, L. (2003) *Adv. Funct. Mater.* **13,** 480–486
41. Marie, B., Luquet, G., Pais De Barros, J. P., Guichard, N., Morel, S., Alcaraz, G., Bollache, L., and Marin, F. (2007) *FEBS J.* **274,** 2933–2945
42. Gotliv, B. A., Addadi, L., and Weiner, S. (2003) *ChemBioChem.* **4,** 522–529
43. Pereira-Mouriès, L., Almeida, M. J., Ribeiro, C., Peduzzi, J., Barthélemy, M., Milet, C., and Lopez, E. (2002) *Eur. J. Biochem.* **269,** 4994–5003
44. Zhang, Y., Xie, L., Meng, Q., Jiang, T., Pu, R., Chen, L., and Zhang, R. (2003) *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **135,** 565–573
45. Aizenberg, J., Lambert, G., Weiner, S., and Addadi, L. (2002) *J. Am. Chem.*

Soc. **124,** 32–39
46. Tellam, R. L., Wijffels, G., and Willadsen, P. (1999) *Insect Biochem. Mol. Biol.* **29,** 87–101
47. Karouzou, M. V., Spyropoulos, Y., Iconomidou, V. A., Cornman, R. S., Hamodrakas, S. J., and Willis, J. H. (2007) *Insect Biochem. Mol. Biol.* **37,** 754–760
48. Wynn, A., and Shafer, T. H. (2005) *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **141,** 294–306
49. Hecker, A., Testenière, O., Marin, F., and Luquet, G. (2003) *FEBS Lett.* **535,** 49–54
50. Inoue, H., Ohira, T., Ozaki, N., and Nagasawa, H. (2003) *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **136,** 755–765
51. Halloran, B. A., and Donachy, J. E. (1995) *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **111,** 221–231