A crayfish molar tooth protein with putative mineralized exoskeletal chitinous matrix c properties

Jenny Tynyakov\textsuperscript{1,2}, Shmuel Bentov\textsuperscript{1,2}, Shai Abehsera\textsuperscript{1}, Galit Yehezkel\textsuperscript{1}, Ziv Roth\textsuperscript{1}, Isam Khalaila\textsuperscript{3}, Simy Weil\textsuperscript{1}, Amir Berman\textsuperscript{3}, Inbar Plaschkes\textsuperscript{2}, Moshe Tom\textsuperscript{4}, Eliahu D. Aflalo\textsuperscript{1,2} and Amir Sagi\textsuperscript{1,2}

\textsuperscript{1}Department of Life Sciences, \textsuperscript{2}National Institute for Biotechnology in the Negev, \textsuperscript{3}Department of Biotechnology Engineering, Ben-Gurion University, P O. Box 653, Beer-Sheva 84105, Israel, \textsuperscript{4}Israel Oceanographic and Limnological Research, Haifa 8511911, Israel

\textsuperscript{1}To whom correspondence should be addressed. E-mail: sagia@bgu.ac.il

Keywords: Biomineralization, crustacean, RNAi, calcium, hydroxyapatite, chitin.
Abstract

Some crustaceans possess exoskeletons that are reinforced with calcium carbonate. In the crayfish *Cherax quadricarinatus*, the molar tooth, which is part of the mandibular exoskeleton, contains an unusual crystalline enamel-like apatite layer. As this layer resembles vertebrate enamel in composition and function, it offers an interesting example of convergent evolution. Unlike other parts of the crayfish exoskeleton, which is periodically shed and regenerated during the molt cycle, molar mineral deposition takes place during the pre-molt stage. The molar mineral composition transforms continuously from fluorapatite through amorphous calcium phosphate to amorphous calcium carbonate and is mounted on chitin. The process of crayfish molar formation is entirely extracellular and presumably controlled by proteins, lipids, polysaccharides, low-molecular weight molecules and calcium salts. We have identified a novel molar protein termed Cq-M15 from *C. quadricarinatus* and cloned its transcript from the molar-forming epithelium. Its transcript and differential expression were confirmed by a next generation sequencing library. The predicted acidic pI of Cq-M15 suggests its possible involvement in mineral arrangement. *Cq-M15* is expressed in several exoskeletal tissues at pre-molt and its silencing is lethal. Like other arthropod cuticular proteins, Cq-M15 possesses a chitin-binding Rebers-Riddiford domain, with a recombinant version of the protein found to bind chitin. Cq-M15 was also found to interact with calcium ions in a concentration dependent manner. This latter property might make Cq-M15 useful for bone and dental regenerative efforts. We suggest that, in molar, this protein might be involved in calcium phosphate and/or carbonate precipitation.

Introduction

Biomineralization, the biogenic formation of mineral deposits, plays a central role in the formation and hardening of exoskeletons and occurs in almost all eukaryal phyla, although the morphology and extent of biomineralization vary widely across species (Simkiss, 1975). Invertebrates deposit various mineral compositions, such as iron oxide (Lowenstam, 1967; Weaver et al., 2010), carbonate apatite (Currey et al., 1982; Lowenstam and Weiner, 1985, 1992), silica (van der Wal, 1989; Miller et al., 1990; Michels et al., 2012), protodolomite (Ma et al., 2008) amorphous calcium phosphate (Greenaway, 1985; Vijayan and Diwan, 1996; Levi-Kalisman et al., 2002; Luquet and Marin, 2004; Neues et al., 2007; Kunkel and Jercinovic, 2013; Huber et al., 2014) and crystalline hydroxyapatite (Weaver et al., 2012), to increase hardness of exoskeletal working surfaces. In vertebrates, enamel serves this purpose (Lucas, 2004). The mandibles of the arthropod *Cherax quadricarinatus* (Bentov et al., 2012) contain an unusual crystalline enamel-like apatite layer which resembles vertebrate enamel in both composition and function (Bentov et al., 2012). The appearance of an enamel-like apatite
layer in a crayfish molar and the similarities found in its mechanical properties to mammalian teeth suggesting a possible case of convergent evolution as previously described in other species (Lowenstam and Weiner, 1992; Bentov et al., 2012; Weaver et al., 2012). The apatite layer is presented in the form of fluorapatite (FAP) and is mounted on chitin reinforced with amorphous calcium carbonate (ACC) (Bentov et al., 2012). The mineral composition of the molar continuously transforms from the FAP to amorphous calcium phosphate (ACP) and from ACP to ACC (Bentov et al., 2012).

Unlike vertebrate teeth, the complex functional tooth structure of the crayfish develops quickly, and is shed and regenerated with each molt cycle (Shechter et al., 2008a; Tynyakov et al., 2015). The molt cycle in crustaceans can be divided into four major stages, whereas a new exoskeleton is assembled during pre-molt and post-molt and sclerotized and calcified during post-molt (Shechter et al., 2008b). During pre-molt the new epicuticle and exocuticle is build beneath the old cuticle (Roer and Dillaman, 1984), while these layers do not calcify until post-molt (Travis, 1963; Travis and Friberg, 1963). Thus the mineral deposition of epic- and exocuticle begins at post-molt together with endocuticle synthesis (Drach, 1939; Travis, 1963; Travis and Friberg, 1963; Travis, 1965) Unlike the exoskeleton, the crayfish molar is assembled and mineralized during pre-molt (Tynyakov et al., 2015). Similarly to crustacean cuticle, the process of crayfish molar formation is entirely extracellular and presumably controlled by proteins, lipids, polysaccharides, low-molecular weight molecules and calcium salts (Willis, 1987; Lowenstam and Weiner, 1989; Horst and Freeman, 1993; Luquet, 2012).

As in vertebrates (Fujisawa and Tamura, 2012), skeletal proteins also play vital roles in mineral formation in invertebrates (Glazer and Sagi, 2012; Luquet, 2012). Previous in vitro experiments identified crustacean proteins thought to be associated with calcium carbonate calcification (Nagasawa, 2011). In contrast to vertebrates, the mineral in crustaceans is precipitated on chitin which forms their exoskeletal structural components. As such, many exoskeletal proteins contain chitin-binding domains (Tabor, 1986), such as chitin-binding domain 2 (Tellam et al., 1999) and the Rebers–Riddiford (R-R) consensus sequence (Karouzou et al., 2007), and/or interact with calcium carbonate. Three known forms of the R-R consensus sequences, RR-1, RR-2, and RR-3, have all been recognized in proteins (Andersen, 2000) and shown experimentally to bind chitin (Rebers and Riddiford, 1988). However, still remains a lack of knowledge about the R-R sequence-containing proteins interactions with chitin and minerals and their localizations in the exoskeleton.

In vertebrates, different proteins have been suggested to play important roles in the mineralization of bones and teeth (George and Veis, 2008). In vivo studies suggested that some of these proteins may influence enamel plasticity (deformation behavior) and morphology, as well as the mineral content of teeth (Robinson et al., 2004; Gruenbaum-
Cohen et al., 2009). Accordingly, mineralization experiments revealed that recombinant versions of these proteins contribute to the regulation of the structure and configuration of mineral particles in vitro (Deshpande et al., 2010; Fang et al., 2011).

The aim of this study was to identify and characterize a novel protein from the crustacean molar by means of transcript tissue expression pattern, transcript’s in vivo silencing, mineral precipitation and chitin binding assays with a recombinant protein. While in vivo silencing of the transcript was lethal, a recombinant version of the protein was found to bind chitin and interact in calcium carbonate and hydroxyapatite precipitation kinetics in a concentration-dependent manner in vitro.

Results

Identification, sequencing, bioinformatics analysis and biophysical characterization of Cq-M15. Initially efforts were directed at defining the protein content of C. quadricarinatus molars. Separation of EGTA-extracted molar proteins by two-dimensional polyacrylamide gel electrophoresis (2D PAGE) revealed the presence of at least twelve distinct spots. These spots were excised from the gel and subjected to mass spectrometry for peptide de novo sequencing. Degenerate primers corresponding to the obtained peptide sequences were constructed. Then cDNA fragments were amplified and eventually full transcripts were obtained. While screening the protein profile of the molar tooth, a protein termed Cq-M15 was identified as a spot of ~17 kDa and pI of ~5.4 (Fig. S1, circled spot, modified) (Tynyakov et al., 2015). Its deduced 740 bp transcript includes a 5’ UTR of 18 bp, an ORF of 465 bp encoding a deduced 155 amino acid protein, and a 3’ UTR of 257 bp (Fig. 1). The first 15 amino acids were predicted to be a signal peptide. A chitin-binding R-R domain was suggested for spanning residues 53-113. This domain is assigned to the RR-1 family, members of which have been found in soft and flexible cuticles (Kim et al., 2005). The predicted pI of Cq-M15 is 5.2, and approximately 5.8% of the total amino acids of the protein correspond to possible phosphorylation sites. Amino acid composition analysis of Cq-M15, moreover, revealed an abundance of non-polar, aliphatic amino acids, as well as polar but uncharged amino acids (Table 1).

Cq-M15 resembles crustacean cuticular proteins expressed in mineralized cuticles. Following the assembly of its sequence, the similarity of the putative Cq-M15 protein to other proteins in protein databases was assessed, revealing Cq-M15 to be similar to cuticular proteins found in crustaceans and insects (Table 2). Evolutionarily, the above proteins were devided into two clades (Fig. 2). Clade i includes Cq-M15 and cuticular proteins from the mineralized parts of crustacean cuticles and clade ii encompasses cuticular proteins expressed...
in the crustacean arthrodial cuticles (non-calcified and flexible cuticles), and cuticular proteins expressed in insects. From the evolutionary perspective, Cq-M15 is thus much closer to crustacean proteins from mineralized cuticles than to arthrodial and insect cuticles.

**Cq-M15 transcription significantly differs as a function of tissue and molt stage.** As suggested by its similarity to exoskeletal proteins from mineralized crustacean cuticles, and since the molar tooth is being built during pre-molt (Tynyakov et al., 2015), the specific expression of Cq-M15 was tested in a number of target tissues at different molting stages. RT-PCR revealed that Cq-M15 is transcribed at the mid- and late pre-molt stages in exoskeletal elements forming tissues, such as the cuticle, molar, basal segment, maxillae and gastrolith, as well as in non-exoskeletal tissue, such as the muscle, as compared to the transcription seen during inter-molt and post-molt (Fig. 3A). In most non-exoskeletal tissues, including testes, no transcription was detected. RNA sequencing of hypodermis, gastrolith and molar-forming tissue by Illumina (560 million reads), and additional 454 and Sanger sequencing (287,900 reads) was followed by *de novo* assembly, yielding 56,490 contigs as described previously (Tynyakov et al., 2015). Based on the above database a differential expression analysis of the Cq-M15 transcript showed significant increase (p <0.05) in normalized read count of early- and late- pre-molt compared to the post- and inter-molt stages at the molar-forming epithelium samples as described in Fig. 3B. This was confirmed by Real-time RT-PCR which was next performed to evaluate differences in the relative transcription levels of Cq-M15 between cuticle- and molar-forming tissues at different molt stages, with the level of transcript in testicular tissue serving as a control (Fig. 3C). The results show how Cq-M15 transcription significantly differs as a function of tissue and molt stage (Kruskal-Wallis chi-squared value = 79.3018, df = 14, p < 0.001). Multiple pair-wise comparisons using the Wilcoxon rank sum test revealed that at mid-pre-molt, Cq-M15 transcription in the molar-forming tissue was significantly higher than in cuticle-forming tissue (p < 0.05), while such transcription in testis was significantly lower than in the latter (p < 0.001). In late pre-molt, testicular tissue transcription of Cq-M15 was significantly lower than in cuticle- and molar-forming tissues (p < 0.001), while there was no significant difference between the latter two (W = 60, p = 0.09). Overall, Cq-M15 transcription in cuticle- and molar-forming tissues at the late pre-molt stage is highest, relative to other molt stages (p<0.05). On the other hand, Cq-M15 was almost not transcribed in the late post-molt and inter-molt stages (p < 0.05).

**In vivo injection of Cq-M15 dsRNA causes a high mortality rate.** Since the expression of Cq-M15 was significantly increased in relevant molt stages and locations related to exoskeleton and teeth assembly, a study of Cq-M15 function *in vivo* through an RNAi loss of function approach was attempted by Cq-M15 dsRNA administration. First, the Cq-M15
dsRNA sequence was compared with all known transcripts listed in *C. quadricarinatus* databases, in order to rule out non-specific silencing of other non-*Cq-M15* transcripts in the crayfish (Yudkovski et al., 2007; Yudkovski et al., 2010); no similar sequences were detected. The effectiveness of the RNAi-mediated gene knock-down was evaluated by qPCR (Fig. 4A). Examination of the relative quantification (RQ) values obtained revealed significant differences between the experimental and the control groups (Kruskal-Wallis chi-squared value = 24.6676, df = 5, p < 0.001). When multiple pair-wise comparisons were conducted using the Wilcoxon rank sum test, a significantly lower level of *Cq-M15* transcription of the experimental group, as compared to the control, was noted in the cuticle- and molar-forming tissues, as well as in testis (p<0.05).  

Once the evaluation of silencing effect of *Cq-M15* dsRNA injection had been established, a long-term silencing experiment addressing *Cq-M15* functionality was performed. This approach revealed a high mortality rate in the *Cq-M15* dsRNA-injected group (Fig. 4B). As such, the experiment was terminated after the death of the first seven experimental subjects, with no molt cycle differences noted between the groups. Since *Cq-M15* was found to transcribe at a pre-molt stage, the molt cycle was induced to initiate Cq-M15 transcription. Since the old exoskeleton doesn’t undergo any changes during the newly induced molt-cycle, the effects of silencing of *Cq-M15* could be visible only at the newly formed exoskeleton. None of the experimental group crayfish built a new exoskeleton, including the molar. Thus we could not perform any microscopic or spectral analysis of newly constructed molar morphology and composition. The rate of survival of the experimental group was 30% versus 100% for the control group (Fig. 4B), reflecting a significant difference starting from as early as the third day of the experiment (Wilcoxon rank sum test, W=2.5, p < 0.05).  

**rCq-M15 possesses chitin binding ability in vitro.** Recombinant Cq-M15 (rCq-M15) was used for *in vitro* studies. rCq-M15 was found to specifically bind to chitin in accordance with the prediction of a chitin-binding domain presence. Unbound rCq-M15 and bovine serum albumin (Shimizu et al., 1989) (which served as a control protein, lacking a chitin-binding domain) were recovered in the supernatant after consecutive washes of a chitin-containing solution with DDW (Fig. 5, lanes 1 and 4 correspondingly) and were also detected following further washings with weak salt solution (Fig. 5, lanes 2 and 5 correspondingly). Then, after boiling the washed chitin in 2% SDS with 20% 2-mercaptoethanol, only rCq-M15 was recovered (Fig. 5, lane 6). The later points to its strong binding to chitin.

**rCq-M15 undergoes a conformational transition in the presence of Ca^{2+}.** Evolutionary and sequence evidence point to the potential involvement of Cq-M15 in biomineralization,
specifically in the precipitation of Ca\(^{2+}\) ions and their carbonate and phosphate salts. Thus, the interactions of rCq-M15 with Ca\(^{2+}\) ions and its carbonate and phosphate salts found in the mandible and molar tooth were next tested in vitro. Initially, rCq-M15 was assessed by circular dichroism (CD). According to bioinformatics predictions, the first 96 residues of Cq-M15 contain a mixture of well-spaced loops and short beta-sheets, whereas the region spanning residues 120-140 is predicted to fold as an alpha helix. The CD curve obtained in calcium-free solution is characterized by minima at 213 and 216 nm (Fig. 6). In the presence of calcium, rCq-M15 undergoes a conformational transition, reflected by the appearance of prominent absorbance shoulders at 209, 215 and 218 nm.

**rCq-M15 interacts in the enhancement of calcium carbonate precipitation rate in vitro.** Finally, to the role of Cq-M15 in the biomineralization processes, in vitro calcium carbonate and phosphate precipitation assays were conducted with rCq-M15. No significant differences between calcium carbonate precipitation rates were noted in the absence of rCq-M15 (blank) and in its presence at concentrations of 57 µM and 230 µM. On the other hand, it was observed that at a concentration of 459 µM, rCq-M15 significantly interacted to promote calcium carbonate precipitation. At the end of the experimental period, the amount of deposited mineral presented by the absorbance was deemed to be concentration-dependent with the blank being the lowest and 459 µM rCq-M15 being the highest (Wilcoxon rank sum test, p <0.05) (Fig. 7).

**rCq-M15 interacts in apatite precipitation process in vitro.** With respect to calcium phosphate precipitation, rCq-M15 interacts in hydroxyapatite precipitation at 37°C in a concentration-dependent manner, compared to the blank (in the absence of rCq-M15), in which brushite (a hydroxyapatite precursor) (Park et al., 2002), was precipitated exclusively (Fig. 8A). At the initial concentration of 308 µM, rCq-M15 precipitated brushite with characteristic Raman peaks at 904 cm\(^{-1}\) and 990 cm\(^{-1}\) (Fig. 8A), while at a concentration of 615 µM, rCq-M15 interacted in precipitation of hydroxyapatite with characteristic Raman peak at 960 cm\(^{-1}\) (Fig. 8A), indicative of the presence of both compounds in the precipitate. In the presence of 923 µM and 1085 µM rCq-M15 (Fig. 8A), only the hydroxyapatite peak at 960 cm\(^{-1}\) was observed. To follow more realistic conditions of biomineralization in the crayfish, the above experiment was repeated for longer intervals and under ambient temperature. This experiment demonstrated that at a concentration of 923 µM, rCq-M15 seems to interact in the kinetics of crystallization by enhancing the level of octa-calcium phosphate (an additional precursor in the crystallization pathway of hydroxyapatite (Grynpas and Omelon, 2007)), with a characteristic peak at 985 cm\(^{-1}\) at ambient temperature, compared to the blank in which brushite, with characteristic peaks at 904 cm\(^{-1}\) and 990 cm\(^{-1}\), was
precipitated exclusively (Fig. 8B). In the presence of rCq-M15, octa-calcium phosphate remained stable during the two-week duration of the experiment (Fig. 8B), whereas in the blank, a mixture of brushite and octa-calcium phosphate gradually appeared (Fig. 8B). Fig. 9 presents the numerical values of intensity at 960 cm\(^{-1}\) (\(\nu_1\) vibration of PO\(_4\)) after subtraction of the intensity of the 1010 cm\(^{-1}\) (typical OCP peak). A significant higher apatite/OCP ratio is shown in the presence of 3.1 \(\mu\)M rCq-M15 compared to the control. No statistical difference in the apatite/OCP ratio was observed in the presence of 150 nM or 4.1 \(\mu\)M BSA.

**Discussion**

In the present study, a novel protein termed Cq-M15 was identified from the crayfish molar tooth and found to be expressed in several exoskeletal forming epithelia.

Characterization of the protein suggests its possible involvement in biomineralization. Indeed, this is the first report of an invertebrate protein suggested to interact in the precipitation of hydroxyapatite. The deduced protein sequence contains a predicted signal peptide, in accordance with the detection of the protein in the extracellular matrix. On the one side, the Cq-M15 has low percentages of acidic and putatively phosphorylated amino acids and possess RR1-type chitin-binding motif which are more compatible with the properties of structural cuticular proteins (Willis, 2010) and as such, the above properties are not in favor of an interaction with ions. Still, Cq-M15 could be enough acidic to possess the ability to interact in the mineral precipitation as was discovered with other hydrophilic and acidic molecules involved in the regulation of crystal nucleation and growth during biological mineralization in vertebrates (He et al., 2005; Gayathri et al., 2007), as well as in invertebrates (Glazer and Sagi, 2012; Luquet, 2012). Moreover, several crustacean proteins associated with calcification processes present amino acid composition profiles similar to Cq-M15 (Inoue et al., 2004; Glazer et al., 2010). In fact, such shared patterns of amino acid composition were also found in proteins from organic matrices involved in invertebrates calcium carbonate precipitation (Faircloth and Shafer, 2007; Gayathri et al., 2007; Marie et al., 2007; Inoue et al., 2008; Suzuki et al., 2013), as well as in protein extracted from vertebrate enamel (Robinson et al., 2005).

The expression of *Cq-M15* during the pre-molt stage coincides with molar tooth mineralization and the synthesis of the new cuticle during this particular molt stage (Tynyakov et al., 2015) and presumably indicates the involvement of Cq-M15 in exoskeleton assembly and specifically, in the early calcification of the molar tooth and other exoskeletal elements. Elevation of *Cq-M15* transcript expression in molar-forming epithelium at mid pre-molt coincides with the formation of the apatite part (Tynyakov et al., 2015) which could
implicate Cq-M15 interaction in calcium phosphate precipitation. Following the molt cycle progression, the highest Cq-M15 expression was detected at late pre-molt both in molar- and cuticle-forming epithelium which coincides with the exocuticle synthesis (Drach, 1939; Roer and Dillaman, 1984). The later could indicate the presence Cq-M15 in exocuticle and its interactions in calcium carbonate and ACC precipitation. We assume that in molar Cq-M15 is secreted by epithelial cells which extensions presumably formed an array of channels within the apatite layer (Bentov et al., 2012).

Cq-M15 has been found to transcribe in muscle as well. Previously it was published that crustacean muscle growth is associated with a molt cycle. Some studies demonstrated that muscle protein synthesis rates were elevated during the pre-molt (Mykles and Skinner, 1982; El Haj and Houlihan, 1987). For example, together with the protein synthesis elevation, claw muscles undergone atrophy (Skinner, 1965; Mykles and Skinner, 1982). This loss in muscle mass enabled the claw to be drawn through the narrow basi-ischium joint (Mykles and Skinner, 1990). It was beyond our scope to investigate relations between Cq-M15 expression, the muscle types and changes in muscles through the molt cycle. However, based on the known literature, we would suggest also possible involvement of Cq-M15 in atrophy and restoration of muscle proteins to facilitate withdrawal from the carapace during molting.

The newly identified Cq-M15 was found to share significant similarity to known proteins from arthropod cuticles and according to a, its sequence was more closely related to crustacean proteins from mineralized parts of the cuticle (Gosline, 1980; Inoue et al., 2008; Suzuki et al., 2013) and more distant from proteins found in crustacean non-mineralized arthrodial membranes (Gosline, 1980) and insect cuticles (Kim et al., 2005; Futahashi and Fujiwara, 2008; Werren et al., 2010). Such close homology further supports the involvement of Cq-M15 in biomineralization in the crayfish. Furthermore, in vivo silencing of Cq-M15 was lethal, suggesting its vital function during the molt cycle. While numerous gene silencing studies assessing the functionality of novel crustacean genes have been published previously (Ventura et al., 2009; Rosen et al., 2010; Ventura et al., 2012), to the best of our knowledge, the lethal effect of gene silencing has been previously reported in crustaceans only once (Christoffersen et al., 1989). We had shown earlier that silencing of particular transcripts in C. quadricarinatus affects the formation of the extracellular matrix and the calcification process (Shechter et al., 2008b; Glazer et al., 2010), as well as the duration of molt cycles (Pamuru et al., 2012; Tynyakov et al., 2015). However, the absence of C. quadricarinatus sequenced genome does not permit to completely rule out potential effects of non-specific small interfering RNAs derived from the Cq-M15 sequence acting within target cells, even though a complete sequence search was done prior to the silencing study.

Cq-M15 is predicted to contain the RR1-type chitin-binding motif. However, unlike most such proteins found in soft cuticular elements, i.e. insect cuticles (Rebers and Riddiford,
Cq-M15 was extracted from a hard exoskeletal organ, i.e. the molar tooth. This is thus only the second report of an RR1-type protein being present in a hard cuticle (Suderman et al., 2006). In accordance with the bioinformatics-based prediction, rCq-M15 chitin-binding ability suggests direct association with the extracellular chitinous cuticular matrix in different parts of the crayfish exoskeleton. In accordance with computer modeling predicting that beta-sheet folds of cuticular proteins form chitin-binding sites (Iconomidou et al., 2005), the ellipticity data obtained here confirmed the presence of beta-sheets and a helical structure in rCq-M15. In addition, CD analysis revealed that Cq-M15 undergoes a conformational transition in the presence of calcium, possibly to provide the proper environment for mineral nucleation.

In vitro mineralization studies under conditions that support the spontaneous formation of calcium phosphate crystals (Margolis et al., 2006) suggest that Cq-M15 interacts in the processes of calcium carbonate and calcium phosphate crystal nucleation. Amorphous calcium phosphate and/or octa-calcium phosphate have been suggested as corresponding to precursor phases of apatite crystals in vertebrates (Grynpas and Omelon, 2007; Beniash et al., 2009) and invertebrates (Lowenstam and Weiner, 1985; Bentov et al., 2012; Weaver et al., 2012). As the ability of rCq-M15 to interact in the initiation of hydroxyapatite nucleation depends on temperature, it seems that additional physiological factors regulating the transition to hydroxyapatite, yet to be discovered in the crayfish molar apatite, might also be involved in this transition.

In conclusion, the data presented here suggest that Cq-M15 possesses physical traits that would allow it to play a role in molar tooth and cuticle biomineralization. Since transcript silencing was lethal, we couldn’t demonstrate Cq-M15 involvement in ACP or AFP precipitation in vivo. Cq-M15 spatial and temporal expression and its interaction in the mineral precipitation in vitro point to possible interaction in calcium carbonate precipitation, including ACC precipitation. Thus, the above-discussed characteristics of Cq-M15 point to the involvement of the protein in the formation of the chitin-protein-mineral complex of the molar tooth and cuticle, through interactions with the chitinous matrix and calcium ions, and possibly with other proteins yet to be discovered. In addition, the level of the interaction depends on the amount of calcium ions in a concentration dependent manner. Since it is published everywhere that recombinant proteins which are produced in E. coli are lack of post-translational modification (Hannig and Makrides, 1998; Jonasson et al., 2002) we suppose that the Cq-M15 recombinant protein could possess a different tertiary and/or quaternary conformation compared to a native Cq-M15 protein in the in vitro medium. Moreover, the Cq-M15 recombinant protein interacts in hydroxyapatite precipitation, whereas
the molar is covered by fluoroapatite (Bentov et al., 2012), which are both molecular forms of apatite with a different in vitro conformation. Thus, this protein could be involved in calcium phosphate and/or carbonate precipitation. Further study of Cq-M15 properties and neighboring proteins will increase our knowledge on the involvement of proteins in exoskeletal calcified tissues and may prove instrumental for research in the fields of bone and teeth repair/regeneration and nano-structured materials.

Materials and methods

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 previously (Shechter et al., 2008a). Intermolts crayfish were held in individual cages and endocrinologically induced to enter pre-molt through removal of the X-organ-sinus gland (XO-SG) complex in the eyestalk or injection of ecdysone to the hemolymph (Shechter et al., 2005; Shechter et al., 2007). Gastrolith (the calcium storage organ) growth during the molt cycle was non invasively evaluated using X-ray in terms of the molt mineralization index (MMI) (gastrolith width/crayfish carapace length), which is used as an accurate molt stage marker (Shechter et al., 2007). For all dissection procedures, crayfish were placed on ice for 5-10 min until they were anesthetized.

Purification of molar proteins. Inter-molt stage animals were endocrinologically induced as above and 14 days later, their molar teeth were dissected at late pre-molt. Molars were pulverized in liquid nitrogen using a mortar and pestle and proteins extracted as previously described (O’Brien et al., 1991; Tynyakov et al., 2015). In brief, protein extract was obtained by 5 M guanidine thiocyanate, 0.5 M EGTA, 5 mM Pipes, pH 7, (1: 10, w: v), at room temperature for 5 h. The extract was centrifuged at 16,000 X g for 5 min in an Eppendorf microfuge, and the supernatant was dialyzed against 10 mM ammonium bicarbonate, pH 7.8 overnight at 4°C and concentrated using a Vivaspin 20 (MWCO 10,000; Vivascience).

Two-dimensional separation and visualization. The molar protein extract was separated in two dimensions as described previously (Arnott et al., 1998) with a few modifications. Briefly, for isoelectric focusing, the immobiline dry strip (11 cm, pH 3–10) was rehydrated and aligned on an isoelectric focusing tray. The molar proteins extract was loaded adjacent to the anode, and voltage was applied for a total of 11.9 kV-h. Following isoelectric focusing, the gel strip was equilibrated in a buffer containing 6 M urea, 0.375 M Tris-HCl, pH 8.8, 2% SDS, 20% glycerol and incubated in 2% dithiothreitol for reduction following 8% iodoacetamide for alkylation. Then the strip was mounted on a 15% SDS-PAGE with the
Tris/Glycine running buffer system as described previously (Lucas, 2004). Spots were visualized with Coomassie Blue staining and the corresponding gel pieces were excised for mass spectrometry.

**Mass spectrometry.** After undergoing two dimensional gel separation and visualization, proteins from molar extracts were trypsin (Gold Trypsin, Promega) digested and the peptides were purified on a C-18 column and dissolved in 0.1% formic acid. Nanoliquid chromatography and mass-spectrometry analysis was performed as described previously (Glazer et al., 2010) using a 75-μm internal diameter fused silica column, packed with C18 (NewObjective, Woburn, MA, USA) connected to an Eksigent nano-LC system (Eksigent, Dublin, CA, USA). Mass spectra were acquired using an LTQ-Orbitrap XL (Thermo Fisher Scientific, San Jose, CA, USA) and analyzed using the PEAKS Mass Spectrometry Software, including a *de novo* peptide sequencing approach. Further M15 protein validation was performed by the Sequest and Mascot algorithms operated under Proteome Discoverer 2.0 (Thermo Fisher Scientific) using an NCBI data base containing the newly identified M15 sequence.

**Transcript sequencing.** Degenerative primers Cq-M15 dF (5'-CAY TTY ACN GCN YTI WSI GC-3') and Cq-M15 dR (5'-CGN CGN ATR GGY GGYGGY GG-3') were generated based on the sequence of a partial fragment of *Cq-M15* obtained by tandem MS. A fragment of the cDNA was amplified using Cq-M15 dF as a forward primer and Cq-M15 dR as a reverse primer. The PCR conditions using the REDTaq ReadyMix™ PCR Reaction Mix (Sigma) were: 94 °C for 3min, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 3 min, followed by final elongation step of 72 °C for 10 min. The fragment was cloned into the pGEM-T Easy vector (Promega, Madison, WI) and sequenced. 5'- and 3'-rapid amplifications of cDNA ends were carried out with the SMART rapid amplification of cDNA ends kit (Clontech), according to the manufacturer’s protocol to obtain the entire sequence of Cq-M15.

**Accession Number** – *Cq-M15* sequence data has been deposited in the GenBank™ Data Bank with accession number KF739426.

**Bioinformatic Analysis.** The deduced amino acid sequence of the gene was obtained bioinformatically by the ExPASy Translate tool (Artimo et al., 2012). Signal peptide and phosphorylation sites were predicted using the SignalP 4.0 (Petersen et al., 2011) and NetPhos 2.0 (He et al., 2005) respectively. Identification of putative domains in the deduced protein sequence was predicted by SMART (Schultz et al., 1998; Suderman et al., 2006) and
Motif Scan (Suzuki et al., 2013) and its type – by a CuticleDB server (Karouzou et al., 2007). Physico-chemical parameters of the Cq-M15 protein sequence including molecular weight and amino acid composition, were predicted using ProtParam tool (Fujisawa and Tamura, 2012). Secondary structure predictions were done by PSIPRED (Futahashi and Fujiwara, 2008).

**Phylogenetic analysis.** A search for homologous sequences to Cq-M15 (www.ncbi.nlm.nih.gov) was performed using BLAST and their phylogenetic analysis among arthropod sequences was performed by the neighbor-joining method (Saitou and Nei, 1987) using MEGA5 software (Andersen, 2000). Evolutionary distances were computed using the Poisson correction method (Lu et al., 2007) and are expressed as amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the analysis.

**Temporal and spatial transcription.** To identify tissue specific temporal and spatial expression of *Cq-M15*, RNA was extracted from the forming epithelia of cuticle, mandible, molar, and basal segment tissues, as well as from the hepatopancreas, muscle and testis, using an EZ-RNA Total RNA Isolation Kit (Biological Industries, Beit Haemek, Israel), according to the manufacturer's protocol. First strand cDNA was generated with oligo (dT) 22 VN using expand RT (Roche Diagnostics). Cq-M15 was amplified by means of PCR with REDTaq ReadyMix™ PCR Reaction Mix (Sigma; one cycle at 94 °C –1 min; 30 cycles at 94°C –1 min, 55°C –2 min, 72°C –3 min; one cycle at 72°C –10 min) with the Cq-M15F (5'-CAACATGAAGTCTATTGTAATGTTGGC-3') and Cq-M15R (5'-TGTTTCAAATAAAGTATATTGAACC -3') primers. 18 S rRNA was evaluated as described previously (Glazer et al., 2010).

**In silico confirmation of Cq-M15 sequence and transcription pattern.** Reference *C. quadricarinatus* transcriptome was constructed from sequencing of four sets of samples as described previously by us (Tynyakov et al., 2015): (1) RNA extracted from hypodermis, gastrolith and molar-forming epithelium sequenced paired-end 100 bp by Illumina HiSeq2000 (USA). (2) RNA extracted from molar and gastrolith forming epithelium from different molt stages (intermolt, early pre-molt, late pre-molt and post-molt), barcoded and sequenced single-end 50 bp by Illumina HiSeq2000 (USA). (3) RNA sequenced by 454 technology, as described by Glazer et al. (Glazer et al., 2013). (4) RNA sequenced “single pass” by Sanger technology ("Expressed Sequence Tags") as described by Yudkovsky et al. (Yudkovski et al., 2010). All sequences (also called "reads") were de novo assembled with CLC Genomics Workbench 6.51 (CLC Bio, Aarhus, Denmark) using default parameters. The resulting contigs were submitted to a second assembly by CAP3, yielding a collection of putative
transcripts regarded as the "Reference *C. quadricarinatus* transcriptome". The 50 bp sequences from molar and gastrolith samples in different molting stages as described above, were aligned to the Reference *C. quadricarinatus* transcriptome using STAR 2.3. For estimation of transcript expression levels, the number of sequences aligned to each transcript in each stage was counted, and further normalized to the total number of sequences in that stage. A statistical test for assessing the differences in expression among stages was carried out using DESeq (Anders and Huber, 2010). The Cq-M15 sequence was searched against the reference transcriptome using BLAST, and the transcript with the highest score and lowest e-value was regarded as the Cq-M15 transcript in different molt stages: inter-molt (pool of animals, n=1), early pre-molt (pool of animals, n=1), late pre-molt (two single animals and one pool, n=3) and post-molt (two single animals, n=2).

**Cq-M15 silencing.** Cq-M15 and CqVg (*C. quadricarinatus* vitellogenin) vitelo dsRNAs were synthetized as described previously (Rosen et al., 2010). Cq-M15 and CqVg dsRNAs were synthesized as described previously (Glazer et al., 2010). For a preparatory short-term silencing experiment, inter-molt males (MMI= 0; 5-10 g) were injected with 5 µg/g of Cq-M15 dsRNA (n=5) or CqVg dsRNA (n=5). Four consecutive injections were given once a day. On the fifth day, the animals were dissected to isolate the cuticle- and molar-forming epithelia and testis. For long-term silencing, the experimental group (n=10) was injected daily with dsRNA of Cq-M15, while the control group (n=10) was injected with CqVg dsRNA. In both experiments, the daily injections included 1 ng/µl ecdysone (Sigma) (Glazer et al., 2010) to induce molting, applied into the sinus of the first abdominal segment.

**Quantitative real time RT-PCR.** RNA was extracted as described above from the testis and the forming epithelium of the molar and cuticle at the following molt stages: mid- and late pre-molt (n₁, n₂=9), early and late post-molt (n₃, n₄=4) and inter-molt (n₅=4). The same was done with experimental and control males (n₆, n₇=5) of the short-term preliminary in vivo Cq-M15-silencing experiment. First strand cDNA was generated by reverse transcription using random hexamer primers (VersoTMcDNA kit, Thermo Fisher Scientific). Relative quantification was performed with the following primers: Cq-M15 F (5’ATGTTGGCTCTGCTTGCTG -3’), Cq-M15R (5’- TGGTGCTCCATAGGTTGTGA -3’) and Universal Probe Library Probe #83 (Roche). Relative quantification of *M15* transcript levels was obtained using the following primers: Cq-M15qPCR_ (F5’ATGAATTACATCATCTCTGTGAGGAACC-3’), Cq-M15qPCR_R(5’CAATGAGCATAAACCAGCCCA-3’) and Probe TM FAM-CTTGATTITTTACTGCGGTTGACACACTGAGTGGAGCAGCAACC-3’, especially designed by TIB MOLBIOL Synthese labor (Roche). *Cq-18S* (accession no. AF235966), used as the
normalizing agent, was evaluated in all the experiments by real time RT-PCR in similar conditions (Glazer et al., 2010).

**Recombinant protein production.** Due to the difficulty of extraction from tissue samples and receiving of negligible amounts of Cq-M15 and, in order to elucidate Cq-M15 functions *in vitro*, rCq-M15 was produced and purified and verified through Western blot by GenScript (GenScript Inc., Piscataway, NJ). In brief, rCq-M15 was chemically synthesized, cloned into pUC57-Kan vector and expressed in *E. coli* BL21 (DE3) at 37°C with 200 rpm agitation. SDS-PAGE and Western blot were used to detect protein expression.

**Chitin-binding assay.** A chitin-binding assay was performed as described previously (Inoue et al., 2001). Briefly, a solution of rCq-M15 was incubated with chitin powder and further successively washed starting with DDW and then with 0.2 M NaCl and finally boiled for 10 min in 2% SDS, 20% (v/v) β-mercaptoethanol. The supernatant of each wash was electrophoresed on 15% SDS-PAGE. The bands were visualized by silver staining. BSA served as a negative control.

**Circular dichroism.** Circular dichroism was conducted with a J750 Spectropolarimeter (Jasco, Easton, STATE) equipped with a Pelletier device. Spectra in the absence and presence of calcium were measured at pH 7.4 using a 0.1 cm optical path Suprasil quartz cuvette (Hellma, Müllheim, Germany) in 400 µl of a 0.028 mM rCq-M15 prepared in 0.1 mM Hepes. Then, 12 µl calcium chloride solution (1 M) was added such that a molar ratio of rCq-M15 to Ca²⁺ ions of 1:1071 was obtained. Spectra were collected at a wavelength range of 190-260 nm at ambient temperature with bandwidth of 1 nm, scan speed of 10 nm min⁻¹ and a time constant of 4 seconds. Each spectrum represents an average of five successive scans over the 190–260 nm wavelength range.

**Mineral precipitation assays.** Different concentrations of rCq-M15 were tested in a calcium carbonate precipitation kinetics assay as described previously (Inoue et al., 2001), with slight modifications. In brief, 100 µl of 22 mM NaHCO₃, pH=8.7, were added to 20 µl rCq-M15 in DDW at different concentrations (57, 230 and 459 µM), to which 100 µl of 22 mM CaCl₂ were added. Addition of DDW instead of rCq-M15 served as a negative control (blank). Turbidity was measured at 570 nm with a spectrophotometer.

Apatite precipitation experiments were performed as described previously (Beniash et al., 2011), with slight modifications. 100 mM stock solutions of CaCl₂ were added to 20 µl rCq-M15 in DDW at different concentrations (308, 615, 923 and 1850 µM). Prior to being added
to the above solutions, rCq-M15 was dialyzed against DDW and concentrated to 3 mg/ml using a SpeedVac apparatus. 100 mM (NH₄)₂HPO₄ was added to yield final concentrations of 2.5 mM Ca²⁺ and 1.5 mM phosphate. 20 µl droplets of the solution were placed on quartz slides in a humidity chamber for incubation at 37°C overnight at 100% humidity. In addition, 20 µl droplets of rCq-M15 at 923 µM were placed on quartz slides in a humidity chamber for incubation at ambient temperature for 24 hours and for two weeks. After incubation, the quartz slides were dipped in DDW, air-dried and subjected to Raman spectroscopy. Raman spectra were acquired on a Jobin-Yvon LabRam HR 800 micro-Raman system equipped with liquid nitrogen-cooled detector as described previously (Bentov et al., 2012).

Qualitative evaluation of the abundance of apatite and its precursor octacalcium phosphate (OCP) was performed as described previously (Crane et al., 2006). Calcium phosphate precipitation was performed in solution by mixing 1M Calcium Chloride (CaCl₂) and 0.5M Sodium Phosphate (Na₂HPO₄, NaH₂PO₄). Both salts were used in different ratio to adjust the pH to 7.4) in 1.5 ml reaction volume, to a final concentration of 20 and 10 mM, respectively. The precipitation assay was performed in the presence of: 150 nM, 4.1 µM BSA, 620 nM and 3.1 µM rCq-M15. All the reactions were performed in two duplicates and were compared to control (without the addition of protein). Each tube was loaded first with CaCl₂, H₂O (adjust to the volume of 1.5 ml) and the tested protein in the final desired concentration. Later the (Na₂HPO₄, NaH₂PO₄) solution was added to start the precipitation in the solution. After 24 h of incubation, the samples were centrifuged at 10,000 rpm and the pellet was separated and washed with DDW and centrifuged again. The pellet was than smeared on Opal slides (for low background) and allowed to dry completely. Samples were analyzed using Raman and qualitative evaluation of the abundance of apatite and its precursor (octa calcium phosphate, OCP) was performed by subtracting the intensity of 1010 cm⁻¹ peak from the main PO₄ ν1 vibration of 960 cm⁻¹. Micro-Raman analyses of specimens were performed on a multichannel bench Renishaw InVia Reflex spectrometer (Renishaw InVia Raman, UK) coupled with a Peltier-cooled CCD detector. Excitation was provided by the 785 nm line of a diode laser. The samples were scanned from 200 to 1200 cm⁻¹ wavenumber shift at a spectral resolution of 2 cm⁻¹. The scattered light was analyzed by a spectrograph with holographic grating (600 g/mm), slit width 150 mm and opened confocal hole (1000 mm). The time of acquisition of a particular spectral window was 10 s.

Statistical analysis. Data are expressed as means ± SE. Non-parametric tests were used as follows: Statistical analysis for relative transcript levels among different tissues and molt stages, differences in survival rates after silencing with Cq-M15 dsRNA and changes in turbidity of calcium carbonate precipitation with rCq-M15 final concentrations, as well as for
qualitative evaluation of the apatite abundance, was performed using the Kruskal-Wallis rank sum test, followed by multiple pair-wise comparisons using the Wilcoxon rank sum test (with Bonferroni correction) for a two-sided test of significance. \( P \) values <0.05 were considered statistically significant.

**Acknowledgments.** This study was supported in part by grants from the Israel Science Foundation (ISF, Grants 102/09 and 613/13) and the National Institute for Biotechnology in the Negev (NIBN). We thank Leila Zeiri for technical assistance with Raman studies and to Vered Chalifa-Caspi for her professional support in NGS analysis. Animals for the study were supplied by the Department of Fisheries and Aquaculture, Dor Agriculture Center.

**Competing interests.** The authors declare no conflict of interest.

**Author contributions.** Author contributions: J.T. and A.S. designed the research; J.T., S.B. and S.W. performed the research; G.Y., Z.R., A.B. and I.K. contributed analytic tools; J.T., S.B., I.K., S.A., E.D.A., I.P., M.T. and A.S. analyzed the data; J.T. and A.S. wrote the paper.
References


Travis, D. F. (1965). The deposition of skeletal structures in the crustacea. 5. The histomorphological and histochemical changes associated with the development and


**Figure 1**: Nucleotide sequence of Cq-M15 cDNA and the deduced amino acids of the open reading frame. The 5' and 3' untranslated regions are highlighted in gray. The putative signal peptide at the N-terminus of the protein is underlined. The predicted chitin binding R&R domain is marked in bold. Predicted phosphorylation sites are boxed. The asterisk indicates a stop codon.
Figure 2: Phylogenetic analysis of Cq-M15 in comparison to cuticular proteins from crustaceans and insects. Phylogenetic tree based on the *C. quadricarinatus* Cq-M15 protein sequence and those of homologous proteins from selected invertebrates shows four distinct clades: one clade (i) comprising cuticular proteins from mineralized parts of cuticles and the second clade (ii) comprising cuticular proteins expressed in arthrodial cuticles of crustaceans and cuticular proteins expressed in insects. The numbers shown next to the branches indicate the percentage of replicate trees in which the associated taxa clustered together in the bootstrap test. The branch lengths correspond to evolutionary distances (expressed in the number of amino acid substitutions per site) used to infer the phylogenetic tree.
Figure 3: Spatial and temporal expression patterns of the *Cq-M15* transcript in vitro and in silico. A. Agarose gels showing RT-PCR products demonstrate spatial and temporal expression patterns of the Cq-M15 transcript in cuticle-, molar-, basal segment (B.S.), maxillae- and gastrolith-forming tissues, as well as in hepatopancreas (Hepato), testis and muscle. B. In silico transcriptomic analysis of Cq-M15 read counts from four different molt stages. Different letters above columns represent statistically significant differences (\( p < 0.05 \)). C. Real-time RT-PCR-based relative quantification of Cq-M15 transcript levels in the cuticle-, molar- and testis-forming tissues of crayfish injected with ecdysone. Error bar represents standard error. Different letters above columns represent statistically significant differences (\( p < 0.05 \)).
**Figure 4:** Relative quantification (RQ) of *Cq-M15* transcript levels. A. Levels of *Cq-M15* transcripts following in vivo dsRNA injections in male crayfish, as assessed by real-time RT-PCR following short-term silencing. The groups injected with ecdysone or either *dsCq-M15* or *dsCqVg*. Different letters represent significant difference and error bars represent standard error (p < 0.05 ± SE). B. Survival of *Cq-M15*-silenced *C. quadricarinatus* males. The experimental group injected with *Cq-M15* dsRNA is represented by a solid black line with black circles. The control group injected with *CqVg* dsRNA is represented by a solid black line with white circles. Both groups were injected with ecdysone to induce their molt cycle. Asterisk represents a significant difference (p < 0.05).
Figure 5: Chitin-binding ability of rCq-M15. Following incubation with chitin powder, equal amounts of rCq-M15 and BSA underwent three consecutive washes prior to SDS-PAGE. Lane 1 contains non-bound BSA washed with DDW, lane 2 contains BSA washed with 0.2 M NaCl and lane 3 contains BSA washed with SDS/2-mercaptoethanol. Lane 4 contains non-bound rCq-M15 washed with DDW, lane 5 contains rCq-M15 washed with 0.2 M NaCl and lane 6 contains rCq-M15 washed with SDS/2-mercaptoethanol.
Figure 6: Circular dichroism analysis of rCq-M15. The absorbance profiles of rCq-M15 in 0.1 mM Hepes (black) or in 0.1 mM Hepes with CaCl\(_2\) (protein: Ca 1:1071 molar ratio) (grey) are shown.
Figure 7: Calcium carbonate precipitation in the presence of rCq-M15. Changes in the turbidity of solution containing NaHCO$_3$ with different concentrations of rCq-M15 after addition of CaCl$_2$ are shown. Asterisk represents a significant difference (p < 0.05).
**Figure 8:** Calcium phosphate precipitation in the presence of rCq-M15. A. Raman spectra of the calcium phosphate phases that precipitate in the presence of different concentrations of rCq-M15 at 37°C. The different charts represent the following concentrations of rCq-M15: 308 µM, 615 µM, 923 µM and 1085 mM. The black line in each chart represents mineral precipitated (brushite) with no addition of protein, while the grey line represents mineral that precipitated in the presence of rCq-M15. The peak of the hydroxyapatite is at 960 cm⁻¹, while
that of brushite is at 990 cm$^{-1}$. An additional peak of brushite at 904 cm$^{-1}$ is typical for those phases that contain HPO$_4^{2-}$ groups due to the P–O–H$^+$ configuration (usually between 875 and 925 cm$^{-1}$). B. Raman spectra of the calcium phosphate phases that precipitate in the presence of rCq-M15 at a concentration of 923 µM at 24°C. The different charts represent the different days of the experiment, i.e. 24 h and seven days. The black line in each chart represents mineral precipitated with no addition of protein, while the grey line represents mineral that precipitated in the presence of 923 µM rCq-M15. The peak of brushite is at 990 cm$^{-1}$, octocalcium phosphate is at 985 cm$^{-1}$ and hydroxyapatite is at 960 cm$^{-1}$. An additional peak of brushite at 904 cm$^{-1}$ is typical for those phases that contain HPO$_4^{2-}$ groups due to the P–O–H$^+$ configuration (usually between 875 and 925 cm$^{-1}$).

**Figure 9:** Numerical values evaluation of the apatite/OCP ratio in the presence of 620 nM or 3.1 µM rCq-M15. Boxes present the intensity at 960 cm$^{-1}$ ($\nu_1$ vibration of PO$_4^{3-}$) after the subtraction of the intensity at 1010 cm$^{-1}$ (typical OCP peak). Results are compared to the apatite/OCP ratio in the presence of 150 nM and 4.1 µM BSA and to a control (without any protein). The lines represent median values and the asterisk represents a significant difference (p < 0.05).
Table 1: Physicochemical properties calculated for the Cq-M15-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.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Cq-M15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
<td>15 (9.7%)</td>
</tr>
<tr>
<td>Ala</td>
<td>19 (12.3%)</td>
</tr>
<tr>
<td>Val</td>
<td>11 (7.1%)</td>
</tr>
<tr>
<td>Leu</td>
<td>10 (6.5%)</td>
</tr>
<tr>
<td>Ile</td>
<td>7 (4.5%)</td>
</tr>
<tr>
<td>Met</td>
<td>2 (1.3%)</td>
</tr>
<tr>
<td>Total</td>
<td>64 (41.4%)</td>
</tr>
<tr>
<td>Phe</td>
<td>8 (5.2%)</td>
</tr>
<tr>
<td>Tyr</td>
<td>4 (2.6%)</td>
</tr>
<tr>
<td>Trp</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>12 (7.8%)</td>
</tr>
<tr>
<td>Ser</td>
<td>13 (8.4%)</td>
</tr>
<tr>
<td>Pro</td>
<td>17 (11.0%)</td>
</tr>
<tr>
<td>Thr</td>
<td>4 (2.6%)</td>
</tr>
<tr>
<td>Cys</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Asn</td>
<td>1 (0.6%)</td>
</tr>
<tr>
<td>Gln</td>
<td>12 (7.8%)</td>
</tr>
<tr>
<td>Total</td>
<td>47 (30.4%)</td>
</tr>
<tr>
<td>Lys</td>
<td>4 (2.6%)</td>
</tr>
<tr>
<td>His</td>
<td>4 (2.6%)</td>
</tr>
<tr>
<td>Arg</td>
<td>7 (4.5%)</td>
</tr>
<tr>
<td>Total</td>
<td>15 (9.7%)</td>
</tr>
<tr>
<td>Asp</td>
<td>10 (6.5%)</td>
</tr>
<tr>
<td>Glu</td>
<td>6 (3.9%)</td>
</tr>
<tr>
<td>Total</td>
<td>16 (10.4%)</td>
</tr>
</tbody>
</table>
Table 2: The full-length Cq-M15 homology to other arthropod cuticular proteins.

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Animal</th>
<th>Homology (%) to Cq-M15 protein</th>
<th>NCBI Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcification associated soluble matrix protein 2 (Casp 2)</td>
<td><em>Procambarus clarkia</em></td>
<td>52%</td>
<td>BAF73806.1</td>
</tr>
<tr>
<td>Strongly Chitin-Binding Protein-1 (SCBP-1)</td>
<td><em>Procambarus clarkia</em></td>
<td>54%</td>
<td>BAM99303.1</td>
</tr>
<tr>
<td>Early Cuticle Protein 5 (ECP5)</td>
<td><em>Callinectes sapidus</em></td>
<td>58%</td>
<td>ADI59753.1</td>
</tr>
<tr>
<td>Early Cuticle Protein (ECP6)</td>
<td><em>Callinectes sapidus</em></td>
<td>54%</td>
<td>ADI59754.1</td>
</tr>
<tr>
<td>Early Cuticle Protein (ECP2)</td>
<td><em>Callinectes sapidus</em></td>
<td>50%</td>
<td>ADI59750.1</td>
</tr>
<tr>
<td>Early Cuticle Protein (ECP1)</td>
<td><em>Callinectes sapidus</em></td>
<td>47%</td>
<td>ADI59749.1</td>
</tr>
<tr>
<td>Early Cuticle Protein (ECP4)</td>
<td><em>Callinectes sapidus</em></td>
<td>47%</td>
<td>ADI59752.1</td>
</tr>
<tr>
<td>Early Cuticle Protein (ECP3)</td>
<td><em>Callinectes sapidus</em></td>
<td>46%</td>
<td>ADI59751.1</td>
</tr>
<tr>
<td>Calcified cuticle protein CP14.1</td>
<td><em>Callinectes sapidus</em></td>
<td>54%</td>
<td>ABB91676.1</td>
</tr>
<tr>
<td>Arthrodial cuticle protein16.5 (AMP16.5)</td>
<td><em>Callinectes sapidus</em></td>
<td>42%</td>
<td>ABB91677.1</td>
</tr>
<tr>
<td>Arthrodial cuticle protein16.3 (AMP16.3)</td>
<td><em>Callinectes sapidus</em></td>
<td>42%</td>
<td>ABC26005.1</td>
</tr>
<tr>
<td>Arthrodial cuticle protein8.1 (AMP8.1)</td>
<td><em>Callinectes sapidus</em></td>
<td>37%</td>
<td>AAV28476.1</td>
</tr>
<tr>
<td>Cuticular protein CPR3</td>
<td><em>Papilio xuthus</em></td>
<td>36%</td>
<td>BAG30726.1</td>
</tr>
<tr>
<td>CP RR-1 family member 40 precursor (CP RR-1 40)</td>
<td><em>Nasonia vitripennis</em></td>
<td>39%</td>
<td>NP_001161305.1</td>
</tr>
<tr>
<td>CP RR-1 family member 19 precursor (CP RR-1 19)</td>
<td><em>Nasonia vitripennis</em></td>
<td>46%</td>
<td>NP_001161334.1</td>
</tr>
<tr>
<td>CP RR-1 motif 6 precursor (CP RR-1 6)</td>
<td><em>Bombyx mori</em></td>
<td>48%</td>
<td>NP_001166742.1</td>
</tr>
<tr>
<td>CP RR-1 motif 8 (CP RR-1 8)</td>
<td><em>Antheraea yamamai</em></td>
<td>34%</td>
<td>AER27816.1</td>
</tr>
<tr>
<td>Pupal cuticle protein</td>
<td><em>Culex quinquefasciatus</em></td>
<td>46%</td>
<td>XP_001863934.1</td>
</tr>
</tbody>
</table>