A shell matrix protein of *Pinctada mazatlanica* produces nacre platelets in vitro

Crisalejandra Rivera-Perez1, Iliana Alejandra Flores-Sánchez2, Josafat Jehu Ojeda Ramírez de Areyano3, Delia Irene Rojas Posadas3 & Norma Y. Hernández-Saavedra3

Nacre is the main component of the pearl oyster shells and it is synthesized by specialized soluble and insoluble shell matrix proteins. Insoluble proteins from the decalcification of the shell are the less studied proteins due to the technical problems to isolate them from the organic matrix. In this study, an insoluble shell matrix protein from *Pinctada mazatlanica*, pearlin (Pmaz-pearlin), was successfully cloned from the mantle tissue, and the native protein isolated from the shell was functionally characterized. The full coding sequence of Pmaz-pearlin mRNA consists of 423 base pairs, which encode to a 16.3 kDa pearlin. Analysis of the deduced amino acid sequence revealed that Pmaz-pearlin contained four acidic regions, an NG repeat domain, and Cys conserved residues, the latter potentially forms four disulfide bridges which might stabilize the protein structure. The isolated protein from the shell is a glycoprotein of ~16.74 kDa which can produce aragonite and calcite crystals in vitro. Our results show that Pmaz-pearlin is a well-conserved protein involved in nacre layer growth, which produces calcite crystals in the presence of CaCl2, aragonite crystal polymorphs with a hexagonal structure in the presence of MgCl2, and needle-like crystal structure polymorphs in the presence of CaCO3. The identity of the crystals was confirmed using RAMAN analyses.

The Mollusk shell is composed of aragonite and calcite, crystal polymorphs of calcium carbonate. The outer prismatic layer of the shell is formed by calcite and the inner nacreous layer is composed of aragonite. These structures are made by specialized proteins known as shell matrix proteins (SMPs), which are synthesized by the epithelial cells from the mantle of mollusks and released into the extrapallial space, between the mantle and the shell, where they perform the crystal nucleation, crystal growth and crystal regulation. The SMPs have been classified according to their solubility after shell decalcification with EDTA or acetic acid solutions as soluble acidic matrix (Asp-rich proteins) and insoluble framework matrix proteins (Gly and Ala-rich proteins), the latter is mostly composed by chitin and silk.

Soluble SMPs are known to determine the mineralogical and crystallographic properties of the shell, while insoluble proteins create the necessary microenvironments for crystal growth and supply a surface for specific molecular reorganization, the framework for the shell formation. Several insoluble framework proteins had been described from the nacreous and the prismatic layer e.g. N14/N16/pearlin, Pif80, Pif97, MSI proteins (7, 31 and 60), Fam20c, N25, nacrein, shematrin, silkmapin. Most of these proteins share specific domains which have been related to functions such as inhibitors of growth and precipitation of calcite and/or inducers of aragonite nucleation. Proteins such as MSI60 acts as structural support for crystal nucleation and growth due to the self-assembly of the protein forming a fiber-like structure, other structural proteins includes hic31 and hic52, both containing polyglycine blocks of (Gly)n and structure similar to collagen type I, alpha 1 and alpha 2. Pearlin is an insoluble SMPs that has drawn attention since is involved in the shell nacre formation and the pearl of oysters. Pearlin belongs to a family protein of low molecular weight, which includes pearlin, N14 and N16 proteins. They differ only by a few amino acids and exhibit moderately acid and basic isoelectric point. The members of this family have a molecular weight of 13.6 to 16.0 kDa, however, N14 can dimerize forming a 28 kDa protein. The amino acid sequences of the members of this family are mainly composed of...
Gly and Tyr. Also, they exhibit an NG repeat domain which differs in length and has been hypothesized to be involved in the regulation of the crystal growth. Moreover, they possess four short acidic domains, as well as disulfide bridges at conserved positions.

Isolation, cDNA cloning, and characterization of pearlin proteins have been limited to two oyster species, Pinctada fucata and Pinctada margaritifera. The pearlin transcript is highly expressed in the dorsal zone of the mantle. The resulting protein is a monomeric glycoprotein of 13.6 to 15.0 kDa, composed by a high proportion of Gly, Tyr, Cys, Asn, Asp and Arg and with calcium-binding properties. These proteins contain in their modular structure four acidic rich regions, ten conserved cysteine residues, a putative casein kinase II phosphorylation site (TDDD) and NG domain, the latter differ significantly in length between them, 10 and 35 residues respectively. These proteins can induce aragonite crystallization and establish inside the interlamellar matrix that separates layers, however, in vitro studies shown that high protein concentration (above 10 µg mL⁻¹) acts as an inhibitor of the precipitation of calcium carbonate, forming crystals of smaller size. Also, the modulation of the crystal growth has been hypothesized to be due to the presence of cofactors and/or the intrinsic properties of the protein, such as posttranslational modifications (e.g. phosphorylation, glycosylation, etc.).

Although biomineralization in mollusk has been studied in several species, the basic mechanism responsible for the inner nacre formation remains unknown. Thus, it is important to characterize the proteins related to this structure. The present study describes the isolation and characterization of the pearlin transcript from the mantle and the pearlin native protein from the shell of the pearl oyster Pinctada mazatlanica. The latter was functionally characterized. The results obtained from this study are beneficial for further studies to obtain a comprehensive understanding of the nacre formation during the biomineralization processes.

Results

Sequence analysis of cDNA pearlin from the mantle of P. mazatlanica. The full-length Pmaz-pearlin was composed of a 726-bp including 5′-UTR and 3′-UTR (292 bp), and an open reading frame (ORF) of 432 bp encoding 140 amino acids (Fig. 1). The Pmaz-pearlin has a calculated molecular mass of 16.3 kDa and a theoretical pl of 5.16 before any post-translational modification (PTM). Removal of the signal peptide sequence (residues 1–26) resulted in a theoretical molecular weight of 13.6 kDa and a theoretical pl of 4.69. It possesses high proportions of Asn (13.2%), Gly (14.9%), and Tyr (13.2%), which accounted for 41.3% of the total amino acid residues (Supplementary Table 1). Analysis of the Pmaz-pearlin sequence revealed the presence of several characteristic domains of the low molecular weight protein family from the shell matrix, including four acidic rich regions (AR1, residues 41–50; AR2, residues 68–73; AR3, residues 89–91 and AR4, residues 132–139), ten conserved cysteine residues (Cys18, Cys30, Cys36, Cys54, Cys55, Cys57, Cys66, Cys77, Cys85, and Cys86) and an NG domain (N92-G111) (Fig. 2A).

The Pmaz-pearlin sequence showed 98.57% similarity to the Pearlin from Pinctada margaritifera, 92.14% to N14 from Pinctada maxima, and 70–74.4% to pearlin, N14, and N16 from Pinctada fucata (Supplementary Table 2). All the protein sequences contained a signal peptide, four aspartic residues (AR1-4), ten conserved Cys residues and the NG domain, however, the length differ among them, being pearlin from P. mazatlanica, P.
Figure 2. Comparison of the deduced amino acid sequence of pearlin and related proteins from pearl oyster species. (A) Alignments were made with Clustal Omega (version 1.2.4). Heparin-binding motif (HB motif) and acidic regions (AR motif) are indicated. Signal peptide is underlined, and putative disulfide bridges are shown by connected dotted lines between cysteine amino acid residues. (B) Neighbor-Joining phylogenetic tree of pearlin from *P. mazatlanica* and related proteins. Protein sequences, pearlin from *P. mazatlanica* (GenBank no. MT746149), *P. margaritifera* (GenBank no. ABG24165.3), *P. fucata* (GenBank no. BAA75626.1), N14 from *P. maxima* (BAA90539.1), and N14.1 (GenBank no. BAA83734.1), N14.2 (GenBank no. BAA83734.1), N14.3 (GenBank no. BAA83735.1), N14.4 (GenBank no. BAA83736.1), N14.5 (GenBank no. BAA83737.1), N14.7 (GenBank no. BAA83739.1), N16-6 (GenBank no. BAN84255.1) and N16-7 (GenBank no. BAN67811) from *P. fucata*. Asterisk shows identical or conserved residues in all sequences in the alignment; colons are conservative substitutions; periods are semi-conservative substitutions.
margaritifera and N14 from P. maxima the longest from all the sequences analyzed, with 20 amino acid residues (Fig. 2A). Also, only N14.1 to N14.4 contained a heparin-binding motif (HB motif) composed by HKK residues.

Further bioinformatics analysis of the Pmaz-pearlin sequence indicates that the protein contains several putative PTMs such as O-glycosylation (Ser and Thr) and N-glycosylation (Tyr), as well as eleven phosphorylation sites. The motif TDDD, a phosphorylation site found in other homologs proteins, displayed a conservative substitution in Pmaz-pearlin (D89E). Also, four putative disulfide bridges Cys30-Cys57, Cys54-Cys66, Cys55-Cys86, and Cys77-Cys85 were predicted for Pmaz-pearlin (Fig. 2A).

Phylogenetic analysis. Eleven homologs proteins were identified in other oyster species from the low weight molecular protein family. The phylogenetic analysis of those proteins including the deduced Pmaz-pearlin showed two main branches, the first branch contained two nodes with over 60% support, this branch was represented by all homologs protein from P. fucata (N14, N16, and pearlin), and the second branch with 100% support was represented by pearlins (P. margaritifera and P. mazatlanica), and one N14 from P. maxima (Fig. 2B).

Characterization of isolated pearlin protein from the shell. Shell matrix proteins were extracted from the shell of P. mazatlanica using acetic acid and the acetic acid-soluble proteins (ASM) and acetic acid-insoluble proteins (AIM) were separated. The proteins from ASM and AIM were separated by 12% SDS-PAGE and only a protein band of ~20 kDa was identified in the AIM (Fig. 3A). This protein band (~20 kDa) was found to have immunoreactivity with an antibody against Pmarg-pearlin (Fig. 3B,C), suggesting that this band corresponds to a pearlin protein.

The identified pearlin, named Pmaz-pearlin, was isolated from the AIM in a single step using preparative electrophoresis. The Pmaz-pearlin was eluted in 30 fractions from 150 fractions collected. Each purification procedure produced 30.2 µg of protein from 77 µg of the target protein from the AIM, leading a yield of 38.9% (Table 1). The isolated Pmaz-pearlin has a relative molecular mass of ~16.74 kDa (Fig. 3D–F), which differ significantly to the deduced molecular mass, ~13.6 kDa, this difference could be due to the presence of PTMs. According to PAS staining of the protein on 16% SDS-PAGE, Pmaz-pearlin possesses carbohydrates associated, however, the composition of the carbohydrates was not analyzed. The Pmaz-pearlin protein band analyzed by

Table 1. Purification of pearlin from the shell of Pinctada mazatlanica.

| Step                                      | Total protein (mg) | Target protein (mg) | Yield (%) | Purity (%) |
|-------------------------------------------|-------------------|---------------------|-----------|------------|
| Crude extract, AIM                        | 0.381             | 0.077               | 100.0     | 20.2       |
| Preparative electrophoresis and AMICON concentrate | 0.030             | 0.030               | 38.9      | 100.0      |

Figure 3. Pearlin identification and purification from the acetic insoluble matrix of the shell of Pinctada mazatlanica. Protein profile of the acetic soluble (ASM) and insoluble (AIM) matrix analyzed by Coomassie Brilliant Blue (A) 12% SDS-PAGE, and (B) 16% SDS-PAGE, (C) Immunodetection of pearlin by a polyclonal antibody against pearlin by Western blot, (D) Purified P mazatlanica protein analyzed by 16% SDS-PAGE and stained with Coomassie Brilliant Blue, (E) Silver stain, and (F) periodic acid Schiff’s stain. MM: Molecular marker (kDa), CBB: Coomassie Brilliant Blue, PAS: periodic acid Schiff, ovalbumin protein was used as a positive control for Schiff’s stain. Gels and Blots were cropped for concise presentation, the full-length gels and blot are presented in the Supplementary Fig. 2.
LC–MS/MS produced four peptide sequences that matched with the deduced amino acid residues of the amplified pearlin transcript from the mantle of *P. mazatlanica* (Fig. 1, Table 2).

**In vitro crystallization of Pmaz-pearlin.** Three different salts, (1) NaHCO₃, CaCl₂, (2) NaHCO₃, MgCl₂, and (3) CaCO₃, were tested to evaluate the crystal growth in the presence of Pmaz-pearlin (Fig. 4). Control preparations, salts without protein, displayed small crystal structures of ~20 µm with rhombohedral structure (Fig. 4A–C). Preparations containing Pmaz-pearlin displayed structured crystals of >100 µm, which were significantly bigger than controls. Addition of Pmaz-pearlin to the saturated solution containing CaCl₂ formed structured crystals of 200 µm (Fig. 4D), and in the presence of MgCl₂, nacre platelets were observed with the typical hexagonal shape (Fig. 4E). Finally, Pmaz-pearlin produced large needle-like crystals (>300 µm) in the presence of CaCO₃ (Fig. 4F).

**Raman analysis.** The identity of the in vitro crystals produced by the pearlin from *P. mazatlanica* was confirmed by Raman spectroscopy. The spectra for the samples corresponding to Fig. 4D–F are shown in Fig. 5. Theoretical calculation of calcite and aragonite crystals using Raman spectroscopy have been widely studied. Calcite crystals contain two CaCO₃ units, for a total of ten atoms, and at Raman spectrum displays translator oscillations of CO₃ groups at 282.47 cm⁻¹, asymmetric bending at 712.48 cm⁻¹ (ν₄), and symmetric stretching of CO₃ groups (ν₁), confirming the calcite structure with space group *D*₆₃(*R*₃c) and parameters of a = 5.03 and c = 17.325 Å, respectively. Aragonite crystal, contains four CaCO₃ units, for a total of twenty atoms, and displays a Raman spectrum at 206.1 cm⁻¹, 702.595 cm⁻¹, and 1085.16 cm⁻¹ with space group *D*₁₆₂(∗Pnma) cell parameters of 5.008, 8.029, and 5.861 Å for the a, b, and c axes, respectively.

**Discussion.** The nacreous layer of pearl oysters is one of the major biominerals of biotechnological interest due to its fracture toughness. In bivalve mollusks, the nacre layer consists of a brick-and-mortar structure where aragonite platelets are arranged in parallel layers. Their construction is performed by specific shell matrix proteins (SMPs), which are known to be secreted from the edge region and the inner part of the mantle, also known as mantle edge and pallium, respectively. Once released to the extrapallial space, they form a framework that promotes the nucleation and deposition of aragonite platelets.
Several proteins have been recognized to be involved in nacre growth, they are classified by their domain composition, such as Pif, CA, VWA, LamG-, CBD2-, Kunitz-like, WAP, M-rich, D-rich, G-rich, Q-rich, A-rich, KU-like, chitin-binding domain, or RLCD proteins. Pearlins from *P. margaritifera* and *P. fucata*, are G-rich protein, which contains in their modular structure acidic regions, an NG domain, and Cys conserved residues which form disulfide bridges, including in *P. mazatlanica* in this study. The acidic rich regions, including the NG repeat domain, are hypothesized to induce secondary structures such as β-sheets, helices and random coils which are known to react favorably with mineral surfaces, such as Ca²⁺ or other mineral surfaces. Also, the length of the NG domain is associated with a stronger reaction with Ca²⁺ molecules, matrix components, and crystals. From all pearlins and homolog proteins analyzed, only pearlin from *P. margaritifera* and *P. mazatlanica* and the N14 from *P. maxima* have a 20 amino acid residues length of the NG repeat domain (Fig. 2A), which is two times longer to N14 and N16 proteins, suggesting that these proteins might present a better capability to form calcium carbonate crystals.

Another important feature of SMPs is the presence of cysteine residues, which suggests that they form disulfide bonds. According to the bioinformatics analysis of the deduced protein sequence of Pmaz-pearlin, four putative disulfide bridges are formed, which seems to be conserved in all sequences analyzed (Fig. 2A). Several SMPs, soluble and insoluble, have disulfide bonds, such as P20, P60, PPP-10, and astring A, and P14, some of this disulfide bridges contribute to stabilize intermolecular subunits, such as P20 and P60 proteins. The presence of disulfide bonds in SMPs, including Pmaz-pearlin, might contribute to protect the protein from degradations during the synthesis of the aragonite platelets, as well as to provide rigidity to the structure, which might impact on their function.

Moreover, posttranslational modifications have been found to be extensive in SMPs and are thought to be crucial for activity. For example, glycosylation may be important for binding Ca²⁺ as well as surface recognition and polymorph selection; phosphorylation has been shown to be important for regulating crystal growth and sulfates are thought to bind Ca²⁺ ions and facilitate crystal nucleation. Deduced pearlin sequence of *P. mazatlanica* Pmaz-pearlin possess putative O- and N-glycosylations, which was corroborated by PAS staining, and potentially have similar effects than those previously described in other SMPs, however, the type of carbohydrate was not assessed, as well as the putative phosphorylation.

In vitro crystallization assay is the only assay to evaluate the activity of most of SMPs. Crystal carbonate polymorphs start with the growth of an amorphous crystal carbonate (ACC), a metastable hydrated phase, which then dehydrates to form calcite, aragonite or vaterite. It is known that the presence of ions such as Mg²⁺ or Sr²⁺ or even acidic residues such as aspartic acid or citric acid, affect the crystal polymorph; phosphorylation has been shown to be important for regulating crystal growth and sulfates are thought to bind Ca²⁺ ions and facilitate crystal nucleation. Deduced pearlin sequence of *P. mazatlanica* Pmaz-pearlin possess putative O- and N-glycosylations, which was corroborated by PAS staining, and potentially have similar effects than those previously described in other SMPs, however, the type of carbohydrate was not assessed, as well as the putative phosphorylation.

In vitro crystallization assay is the only assay to evaluate the activity of most of SMPs. Crystal carbonate polymorphs start with the growth of an amorphous crystal carbonate (ACC), a metastable hydrated phase, which then dehydrates to form calcite, aragonite or vaterite. It is known that the presence of ions such as Mg²⁺ or Sr²⁺ or even acidic residues such as aspartic acid or citric acid, affect the crystal polymorph. Magnesium ion is the most abundant ion in the shell, representing 2.55% of the total shell weight and it is known to promote aragonite crystals. Pmaz-pearlin in presence of MgCl₂ was able to produce aragonite platelets, with classic hexagonal shape, and calcite crystals in presence of CaCl₂, which were similar to published morphological studies of aragonite and calcite by SEM and previous identified crystals structure by Raman analysis, as well as those previously reported pearlins. However, Pmaz-pearlin in presence of CaCO₃ lead to needle-like crystals of aragonite, this behavior was also described for *Hyriopsis cumingii* when the whole extract was used for crystallization assay. The strong capability to produce aragonite platelets and needle-like crystals of aragonite might be enhanced by the intrinsic properties of the sequences (acidic regions and NG domain) and it’s PTMs (glycosylations).
Conclusions
The characterization of Pmaz-pearlin, contributes to the knowledge of the relatively small group of low molecular proteins involved in nacre growth. Pearlin from *P. mazatlanica* is a monomeric glycoprotein of ~16.74 kDa, which possesses the modular structure of the pearlins (AR, an NG domain, and disulfide bridges). The intrinsic properties of the sequences and the presence of PTMs might contribute to the capability to form aragonite platelets in the presence of MgCl₂ and CaCO₃.

Methods

**Biological material.** Mother pearl oysters, *Pinctada mazatlanica*, were provided by Perlas del Cortez S. de R.L. M.I. located at Bahia de la Paz B.C.S. Organisms were transported to the Molecular Genetics Laboratory at CIBNOR, and mantle tissue was dissected and stored at ~80 °C until used.

**Molecular characterization of Pearlin from mantle tissue of *P. mazatlanica*.** Total RNA extraction and RNA reverse transcription. Total RNA was extracted with TRIzol reagent (Invitrogen Life Technologies) according to the manufacturer’s instructions. Samples (100 mg) were homogenized using a glass pestle and two consecutive extractions of each sample were made. RNA concentration and purity were determined by spectrophotometry using a NanoDrop ND-2000 at 260/280 and 260/230 nm absorbance ratios (range, 1.9–2.0). The RNA integrity was assessed on a 1% (w/v) Synergel agarose gel. To ensure complete DNA absence, a direct PCR was performed using 1 µL (50 ng µL⁻¹) of each RNA preparation with 28S ribosomal (GenBank accession No. AY632555) specific primers, 28S_F (5'-GCAGGAAAAAGAAACTAAC-3') and 28S_R (5'-CCTCTAAGTGT TTCCAC-3'), a non-amplified control. After that, 1 µg of total RNA was used from each verified RNA sample for cDNA synthesis using the cloned AMV First-Strand cDNA Synthesis Reaction and oligo-dT primer. The cDNAs were stored at ~80 °C until use. Control reactions were performed without template or non-reverse transcribed RNA to determine the presence of DNA.

Pearlin cDNA sequencing. A set of primers that hybridize a conserved region of the homologous pearlin nucleotide sequence (*Pinctada margaritifera*, DQ665305; *Pinctada fucata*, AB020779; *Pinctada fucata*, AB023248-AB023253; *Pinctada fucata*, AB023067, and *Pinctada maxima*, AB032612) were designed. A 414-bp region of the mother pearl oyster pearlin cDNA was amplified by PCR. The reaction included 6 µL (GoTaq Green Master Mix, Promega), 1 µL (10 µM each) of each primer 80F (5’-CATGAMTGCGACACTTGGT-3’) and 482R (5’-CCGTYCTRCAGTRTACTC-3’), 30 ng of *P. mazatlanica* cDNA, and water to a final volume of 12 µL. PCR amplification was carried out for 5 min at 94 °C, followed by 5 cycles consisting 1 min at 94 °C, 1 min at 94 °C, 1 min at 50 °C, and 2 min at 72 °C. In the last cycle, the extension step at 72 °C lasted 10 min. PCR products were analyzed on 1% (w/v) Synergel agarose gels and visualized under UV light after staining with UView loading dye (Bio-Rad). The single band that was produced was sequenced in both directions, using the same set of primers.

The full-length sequence of pearl oyster pearlin transcript was obtained using a cDNA amplification kit (SMART RACE, Clontech Laboratories, Mountain View, CA). First-strand cDNA synthesis of 3’ and 5’-RACE was performed separately from 1 µg poly A + pearl oyster RNA obtained with a polyA Spin mRNA isolation kit (NEB England BioLabs, S1560). PCR amplification of both pearlin cDNA ends was performed using Universal primers included in the SMART Race kit and the specific primers 80F and 482R, using the 3’RACE and 5’RACE cDNA, respectively as template. Both PCR reactions produced a single PCR band that was used for ligation in a sequencing vector (pGEM-Teasy, Promega) and then cloned into *Escherichia coli* TOP10 cells (C4040, Thermo Fischer) following standard cloning methods. Plasmid DNA was isolated from three positive colonies using the alkaline lysis method and used for sequencing reactions.

In silico analyses. Sequence similarity searches were performed using the alignment tool BLAST. Pearlin homologs proteins were obtained from the National Center of Biotechnology Information (NCBI). An analysis degree of similarity among nucleotide sequences was performed using the ClustalW tool. Sequence similarity searches were performed using the alignment tool  BLAST. Pearlin homologs proteins were obtained from the National Center of Biotechnology Information (NCBI). An analysis degree of similarity among nucleotide sequences was performed using the ClustalW tool. A phylogenetic analysis of the sequences and the presence of PTMs might contribute to the capability to form aragonite platelets in the presence of MgCl₂ and CaCO₃.

**Native pearlin purification and characterization.** Shell matrix extraction. The organic matrix of prismatic and nacreous layers of *Pinctada mazatlanica* were crushed to a fine powder. The powdered matrices (20 g) were suspended in 100 mL cold acetic acid (4 °C, 10% v/v) and incubated 24 h with continuous stirring. Acetic acid-soluble matrix fractions (ASM) and acetic acid-insoluble matrix fractions (AIM) were separated by centrifugation at 13,000 × g for 20 min. The AIM was rinsed with distilled water and lyophilized. The ASM was dialyzed for another 24 h against cold acetic acid (4 °C, 1% v/v), afterward, the ASM was dialyzed for another 24 h against distilled water, and then lyophilized.
Sodium dodecyl sulfate–polyacrylamide gel electrophoresis. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to Laemmli. One milligram of lyophilized ASM and AIM of the shell from P. mazatlanica were mixed with 50 µL of 4× sample buffer (0.5 M Tris–HCl pH 6.8, 20% glycerol, 10% SDS, 10% β-mercaptoethanol and 0.05% bromophenol blue) respectively. Samples were boiled for 10 min and then centrifuged 1 min at 12,000×g. The supernatant (20 µL) was loaded into a 12% polyacrylamide gel. Broad range molecular weight standard (Bio-Rad, 1610317, California, USA) was loaded into the gel. Electrophoresis was run at 90-V at room temperature, using a vertical electrophoresis unit (Bio-Rad Protein II, California, USA). After electrophoresis, the gels were stained for 2 h with staining solution (0.05% (w/v) Coomassie Brilliant Blue R-250, 7% (v/v) acetic acid, 40% (v/v) methanol). Proteins were revealed by soaking for 2 h in distaining solution (7% (v/v) acetic acid and 40% (v/v) methanol). Gel proteins were analyzed using a gel imager (Chemi Doc XRS, Bio-Rad, California, USA).

Protein quantification. Quantification of pearlin present in the AIM was performed by pixel densitometry as described by Arroyo-Loranca. Briefly, a standard curve was performed using ovalbumin protein (0.25–7.0 µg/µL) and the linear equation obtained was used for the pearlin protein quantification, y = 10138x-4163.7 (Supplementary Fig. 1).

Immunodetection of pearlin by Western Blot. Polyclonal antibodies against recombinant pearlin from Pinctada margaritifera were used to detect the pearlin of the ASM and AIM from the shell of Pinctada mazatlanica. Protein samples (1 mg of lyophilized ASM and AIM) were separated using a 12% electrophoresis gel under reducing conditions as described before. Afterward, the proteins were electrophoretically transferred to a PVDF membrane with a semidyblotter (Thermo Fischer Scientific) at 18 V, 90 mA for 30 min. Then, the PVDF membrane was blocked with 5% skimmed milk in TNT buffer (10 mM Tris–HCl pH 8.0, 0.15 M NaCl, 0.05% Tween-20), incubated with the primary polyclonal antibody against pearlin (1:10,000), and followed with the second antibody anti-goat horseradish peroxidase conjugate (Santa Cruz, USA) at a 1:5000 dilution. Finally, the immunoblots were visualized in 3,3′-diaminobenzidine (DAB) solution.

Pearlin purification by preparative SDS-PAGE. The identified pearlin from the AIM was isolated using a preparative polyacrylamide gel electrophoresis according to the Mini-Prep Cell Manual (Bio-Rad, model 491 Prep Cell), using 20 mg of lyophilized AIM according to Arroyo-Loranca. Proteins containing the pearlin protein were pooled, concentrated using an Amicon-Ultra filter (10,000 MW cut-off) and desalted using 30 mM Tris–HCl buffer pH 8.8. Isolated protein was stored at −20 °C until use.

Mass spectrometry analysis. Pearlin protein (2 µg) was separated by 12% SDS-PAGE, stained with Coomassie Brilliant Blue, and the band manually cut and extracted for the gel. The gel band was digested with trypsin and processed at the Laboratorio Universitario de Proteómica, UNAM.

Posttranslational modification analysis. Putative glycosylation of pearlin (2 µg) was determined by 12% SDS-PAGE polyacrylamide gels. Saccharide moieties were identified with Periodic Acid-Schiff Stain (PAS) (Sigma-Aldrich, S3133, St. Louis MO, USA).

Crystallization in vitro assay of pearlin. Growth of calcium carbonate crystals (calcite and/or aragonite) in the presence of purified pearlin was performed in vitro. Three different salt solutions were tested according to Weiss and Hillner, solution A containing 40 mM CaCl2 pH 8.2, 100 mM NaHCO3, solution B containing 40 mM MgCl2, pH 8.2, 100 mM NaHCO3, and solution C containing 100 mM CaCO3. All solutions were prepared using molecular biology grade reagents. Two micrograms of protein (10 µL) were mixed with 35 µl of each solution respectively, controls without protein were included. Each mixture was incubated over a sterile coverslip inside a six-well microplate and sealed with parafilm at 4 °C for 30 days. Each experiment was performed by triplicate. The morphology of the crystal was determined by Scanning Electron Microscopy (SEM) by triplicate at the Electronic Microscopy Laboratory at Centro de Investigaciones Biológicas del Noroeste S.C. (CIBNOR), México.

Raman. The identity of the in vitro crystals was obtained using Raman spectroscopy using an InVia micro-Raman spectrometer, using a 532 nm laser (YAG laser) as the excitation source at 100 mW power with a spot size of 4 µm2. The crystals were scanned from 100 to 1200 by triplicate for the specific identification. For all the measurements, the slits were set at 200 µm and a 100 x objective was used. The analysis was carried out at Laboratorio Nacional de Investigaciones en Nanociencias y Nanotecnología (LINAN)-IPICYT.

References
1. Wilt, F. H. Developmental biology meets materials science: Morphogenesis of biomineralized structures. Dev. Biol. 280, 15–25 (2005).
Acknowledgments

The authors are grateful to Dr. Daniel Saulnier from IFREMER for the antibodies against Pinctada margaritifera donated to this work; also, to Ariel Cruz, head of the Electronic Microscopy at CIBNOR for his technical support with SEM images. José Luis Rodríguez López and Beatriz A. Rivera for their technical support with Raman experiments. I.A.F.-S. and C.R.P. analyzed data. N.Y.H.S. and C.R.P. wrote the paper and prepared all figures and tables. All authors reviewed the manuscript.

Author contributions

N.Y.H.S. and C.R.P conceived and designed the experiments. I.A.F.-S., C.R.P., D.I.R.P. and J.J.O.R.A. performed experiments. I.A.F.-S. and C.R.P. analyzed data. N.Y.H.S. and C.R.P wrote the paper and prepared all figures and tables. All authors reviewed the manuscript.
Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-77320-7.
Correspondence and requests for materials should be addressed to N.Y.H.-S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020