Functional prioritization and hydrogel regulation phenomena created by a combinatorial pearl-associated 2-protein biomineralization model system.

Gaurav Jain,[1†] Martin Pendola,[1†] Yu-Chieh Huang,[2] Jose Juan Colas,[3] Denis Gebauer,[2] Steven Johnson,[4] and John Spencer Evans[1†]*

[1]Laboratory for Chemical Physics, Center for Skeletal and Craniofacial Biology, New York University, 345 E. 24th Street, NY, NY, 10010 USA.

[2]Department of Chemistry, Physical Chemistry, Universität Konstanz, Universitätstrasse 10, Konstanz D-78457, Germany.

[3]Department of Physics, University of York, Heslington, York, United Kingdom.

[4]Department of Electronics, University of York, Heslington, York, United Kingdom.

*To whom correspondence should be addressed: John Spencer Evans, Laboratory for Chemical Physics, Division of Basic Sciences and Center for Skeletal and Craniofacial Medicine, New York University College of Dentistry, 345 E. 24th Street, New York, NY, 10010. Tel.: (212) 998-9605; Fax: (212) 995-4087. Email: jse1@nyu.edu.

ABBREVIATIONS

rPFMG1 = recombinant *Pinctada fucata* mantle gene 1; rPFMG2 = recombinant *Pinctada fucata* mantle gene 2; QCM-D = quartz crystal microbalance with dissipation monitoring; PNC = prenucleation cluster; ACC = amorphous calcium carbonate; FIB = focused ion beam sectioning. FSC-A = forward scattered component; SSC-A = side scattered component; αCHCA = α-cyano-4-hydroxycinnamic acid ; SA = sinapinic acid; TEV = 27 kDa catalytic domain of the Nuclear Inclusion "a" (NIa) protein encoded by the tobacco etch virus; Trx = thioredoxin; AP7 = aragonite protein 7 of the nacre layer of the Pacific red abalone, Haliotis rufescens.

ABSTRACT

In the nacre or aragonitic layer of the oyster pearl there exists a 12-member proteome which regulates both the early stages of nucleation and nano-to-mesoscale assembly of nacre tablets and calcitic crystals from mineral nanoparticle precursors. Several approaches have been developed to understand protein-associated mechanisms of pearl nacre formation, yet we still lack insight into how protein ensembles or proteomes manage nucleation and crystal growth. To provide additional insights we have created a proportionally-defined combinatorial model consisting of two pearl nacre-associated proteins, PFMG1 and PFMG2 (shell oyster pearl nacre, *P. fucata*) whose individual *in vitro* mineralization functionalities are distinct from one another. Using SEM, AFM, Ca(II) potentiometric titrations and QCM-D quantitative analyses, we find that at 1:1 molar ratios, rPFMG2 and rPFMG1 co-aggregate in specific molecular ratios to form hybrid hydrogels that affect both the early and later stages of *in vitro* calcium carbonate nucleation. Within these hybrid hydrogels rPFMG2 plays a role in defining protein co-aggregation and hydrogel dimension, whereas rPFMG1 defines participation in non-classical nucleation processes, and both proteins exhibit synergy with regard to surface and subsurface modifications to existing crystals. The interactions between both proteins is enhanced by Ca(II) ions and may involve Ca(II)-induced conformational events within the EF-hand rPFMG1 protein, as well as putative interactions between the EF-hand domain of rPFMG1 and the calponin-like domain of rPFMG2. Thus, the pearl-associated PFMG1 and PFMG2 proteins interact and exhibit mineralization functionalities in specific ways, which may be relevant for pearl formation.

Keywords: Pearl oyster, pearl, nacre, biomineralization, PFMG2, PFMG1, amorphous calcium carbonate, pre-nucleation clusters, QCM-D, AFM, mini-proteome, EF-hand, calponin-like

INTRODUCTION

The formation of the nacre pearl in marine invertebrates represents an on-demand production of mineralization in response to an irritant or parasite threat to the mantle organ. Pearl is a composite of nanograin calcium carbonate aragonite (nacre) and the protein-based organic matrix known as conchiolin.[1] In some cases, there is an admixture of another calcium carbonate polymorph, calcite.[1] As a composite, the pearl is tougher than pure aragonite and the layered nature of the structure contributes to the luster and strength of the pearl.[2-8] At the molecular level the pearl formation process is complex and requires the coordination of extracellular mineralization events with the synthesis and assembly of a molecular protein framework.[9-14] The result is a highly ordered biomaterial with mesoscale bulk properties (fracture resistance, color, lustrous appearance) that originate from nanoscale components and their individual properties.[2-8]

Although this nano-to-mesoscale aragonite assembly process in the pearl is not well understood, studies have shown that protein families or proteomes play an important role in biomineralization and nano-to-mesoscale assembly of shell nacre in the mollusk.[9-20] It is known that individual nacre-associated proteins possess the capabilities of nanoparticle assembly and modifications that coincide with *in situ* features that are important for the formation of mesocrystal aragonite tablets.[21-30] In the case of the Japanese pearl oyster (Pinctada fucata), the pearl formation process is mediated by a 12-member proteome known as PFMG (Pinctada Fucata Mantle Gene), which also play a role in mollusk shell formation.[15-17] One of these proteins, PFMG1,[13,15-17] has been the subject of *in vitro* investigations that explored the mineralization and protein aggregation processes that are relevant to pearl formation.[21] Other members, such as PFMG4, have been found to be enhancers of cell response and differentiation,[31] and it is known that PFMG2, PFMG6, and PFMG8 are highly expressed or up-regulated at various

developmental stages of shell formation.[32] However, these studies have only focused on the function of individual proteins and have ignored the basic fact that protein "collectives" must be involved in managing the pearl biomineralization process. Thus if we are to derive a better understanding of the pearl biomineralization process and nano-to-mesoscale assembly, we need to study the function of PFMG proteins in combination relative to each individual protein.

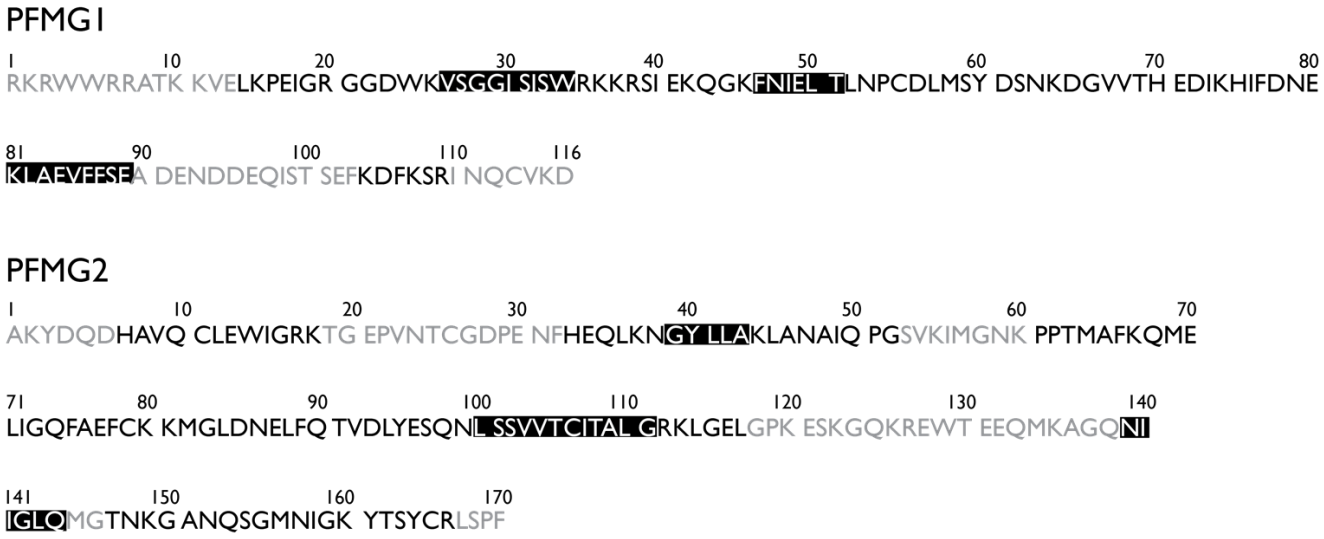

Figure 1 Primary amino acid sequences of Pinctada fucata pearl-associated PFMG1 (UniProt accession number Q3YL59) and PFMG2 (GenBank accession number AAZ76256.1) proteins, with membrane leader sequences deleted. Predicted intrinsically disordered (DISOPRED, IUP algorithms, gray) and amyloid-like cross-beta strand aggregation (AGGRESCAN, ZIBBER DB algorithms, black highlighting) sequences are shown. For PFMG1, putative EF-hand domain is located at D61-E102. For PFMG2, putative calponin domain is located at A1-G159.

Recently, a pair of unrelated nacre proteins, AP7 (*Haliotis rufescens*, Pacific red abalone shell nacre protein) and PFMG1 (*Pinctada fucata*) were investigated in defined molar ratios to better understand if the combination of two unrelated nacre proteins from different proteomes affects *in vitro* functionalities.[29] Surprisingly, this study revealed that both proteins in a 1:1 molar ratio exhibit synergistic mineralization activities and specific aggregation affinities under defined conditions.[29] Thus, a 1:1 molar mixture represents an interesting reference point for future combinatorial studies. We now take this concept one step further and under identical experimental conditions investigate two members

of the *P. fucata* PFMG proteome, PFMG1[13,15-17,21] and PFMG2,[15,32] both individually and in a 1:1 molar mixture. The choice of PFMG1 and PFMG2 for these combinatorial experiments was founded on the fact that PFMG1 contains an EF-hand Ca(II) binding domain,[13,21] and, PFMG2 contains an calponin-like domain that interacts with EF-hand domains in the presence of Ca(II).[15,16] Thus, putative interactions were suspected. As described in this report, we generated recombinant forms of PFMG1 (rPFMG1) and PFMG2 (rPFMG2) and studied each individual protein and a 1:1 molar mixture in parallel mineralization and aggregation experiments. What we found is quite intriguing: both proteins combine in specific molecular ratios to form hybrid supramolecular hydrogels[30] that synergistically modify the surfaces and create intracrystalline nanoporosities within calcite crystals. However, with regard to early events in non-classical nucleation,[33-36] no synergistic effects are observed for these hybrid hydrogels and rPFMG1 functionality predominates instead, with minimal participation from rPFMG2. Similarly, rPFMG2 appears to limit the overall dimensions and internal structure of the hybrid protein hydrogel particles and films in the presence and absence of Ca(II), and the interaction between rPFMG1 and rPFMG2 increases in the presence of Ca(II). Bioinformatics modeling reveals the basis for this observed interaction: both proteins contain folded domains that are interactive (PFMG1, pseudo-EF hand calmodulin-like domain; PFMG2, calponin-like domain), particularly in the presence of Ca(II). We conclude that within hybrid hydrogels rPFMG2 plays a role in defining interprotein complexation and hydrogel formation, whereas rPFMG1 defines participation in non-classical nucleation processes, and both proteins cooperate in surface and subsurface modification to existing crystals. These complex and unique relationships may play an important role in pearl nacre formation.

EXPERIMENTAL PROCEDURES

Sample preparation. The expression, preparation and purification of recombinant PFMG1 (rPFMG1, MW = 13608 Da) was performed as described previously (Figure S2, Supporting Information).[21,29] For the expression of recombinant PFMG2 (rPFMG2), we followed a cloning and expression protocol similar to that reported for rPFMG1[21,29] and this was performed by GenScript USA (Piscataway, NJ, USA; http://www.genscript.com/) using their proprietary OptimumGene system and recombinant expression systems. The synthetic DNA was created from the complete PFMG2 sequence (GenBank accession number AAZ76256.1) minus the membrane leader sequence (residues 1-20)(Figure S1, Supporting Information). To this DNA sequence a hybrid thioredoxin (Trx)-poly-$(His)_6$ tag - TEV protease cleavage site sequence was incorporated at the N-terminus.[26,27] Upon cleavage with TEV and removal of the solubility tag, this creates a protein that is 171 AA in length, with G1 as the first residue (remnant from TEV protease site) and F171 as the C-terminal residue (Figure S1, Supporting Information). This hybrid sequence was subcloned into expression vector E4. The recombinant plasmid was transiently transfected into 100 mL suspension of HEK 293 cell culture, which were then grown in 10 L serum-free media, collected and lysed at day 5 post-transfection. Cell pellets were resuspended in phosphate buffered saline (PBS) and sonicated for 2 minutes to reduce solution viscosity. The solution was then centrifuged and the target protein was captured from the cell lysate using HiTrap chelating Ni HP column and eluted with 200 mM imidazole buffer.

Following Ni elution, Trx-poly-$(His)_6$ tag removal was accomplished via TEV digestion and subsequent affinity capture of the TEV enzyme and removal of the Trx-His_6 fragment using HiTrap

chelating Ni HP column. The protein was then subjected to ultrafiltration using Amicon Ultra 10 kDa MWCO (Millipore Corporation, USA) ultrafilters to remove any additional impurities. The final protein purity was determined to be 93% (4%-20% gradient SDS-PAGE/Coomassie Blue staining, reducing conditions Figure S3, Supporting Information), with a yield of 15 mg of protein per 10 L of culture medium. Using a Bruker Daltonics MALDI-TOF and α-cyano-4-hydroxycinnamic acid (α CHCA) we determined the apparent MW of this purified material to be 18867.6 Da using (Figure S3, Supporting Information), which is close to the hypothetical MW of 18869.6 Da. Purified protein stock aliquots were stored in 50 mM Tris-HCl/10% v/v glycerol (pH 8.0) at -80 ºC until needed. For subsequent experimentation, rPFMG1 and rPFMG2 samples were created by exchanging and concentrating appropriate volumes of stock solution into unbuffered deionized distilled water (UDDW) or other appropriate buffers using Amicon Ultra 0.5, 3 kDa MWCO.

MALDI-TOF MS analysis of individual and rPFMG1, rPFMG2 mixtures. For rPFMG2 and the 1:1 rPFMG1 : rPFMG2 mixture, 5 µL of each protein stock solution (20 µM) were mixed with an equal volume of α-CHCA matrix solution (10 mg/mL, ThermoFisher Scientific, USA), to create a 10 µL sample. In the case of the rPFMG1 sample, 5 µL of protein solution was mixed with 5 uL of sinapinic acid (SA, 10 mg/mL).[21] 1.5 µL of the each protein+matrix sample was placed in a clean ground stainless steel target, and allowed to dry at room temperature. Once dried, the rPFMG2 and 1:1 rPFMG1:rPFMG2 samples were covered with 1.5 µL of the αCHCA matrix and the rPFMG1 samples was covered with 1.5 µL of the SA matrix and the drying process was repeated. Measurements were calibrated with Sigma Aldrich Standards insulin (5,730 Da) and aldolase (39,212 Da). Samples were analyzed using a Linear Mode method in a range of 10-95 kDa, 1.00 m/Z resolution, with 500 shots/sec, using a Bruker UltrafleXtreme MALDI-TOF/TOF instrument.[21,26] Calibration of spectra was made using a Cubic

Enhanced method. Data acquisition performed using the software FlexControl 3.0 (Bruker Daltonics, USA) and data processed with FlexAnalysis 3.0 (Bruker Daltonics, USA) and mMass 5.5.0.

In vitro micro-mineralization assays. Calcite-specific mineralization microassays were conducted by mixing equal volumes of 20 mM $CaCl_2$*$2H_2O$ (pH 5.5) and 20 mM $NaHCO_3$ / Na_2CO_3 buffer (pH 9.75 in Milli-Q Type I ultrapure water 0.22 μm filtered) to a final volume of 500 μL in sealed polypropylene tubes and incubating at room temperature for 1 hr.[22-29] The final pH of the reaction mixture was measured and found to be approximately 8.0 - 8.2.[22-29] Individual aliquots of rPFMG2 and rPFMG1 stock solutions were added to the calcium solution prior to the beginning of the reaction, with final assay concentrations of each protein to be 20 μM. In addition to individual protein-containing assays, we also created a scenario involving a 1:1 mole stoichiometry mixture (i.e., 20 μM rPFMG1 : 20 μM rPFMG2).

Mineral and protein deposits formed during the assay were captured on 5 x 5 mm Si wafer chips (Ted Pella, Inc.) that were placed at the bottoms of the vials. Upon completion of the mineralization assay period, the Si wafers were rinsed thoroughly with calcium carbonate saturated methanol and dried overnight at room temperature prior to analysis.[22-29] Imaging of the Si wafers extracted from the mineralization assays was performed using a Merlin (Carl Zeiss) field emission SEM (FESEM) using either an Everhart-Thornley type secondary electron detector (SE2) at an accelerating voltage of 3 kV and a probe current of 100 pA. Prior to analysis, samples were coated with iridium (4nm layer) using a Cressington 208HR sputter coater with thickness controller. X-ray microanalysis of the iridium coated Si wafers was performed using an Oxford Instruments EDS with integrated INCA software attached to the Merlin FESEM. To perform the analysis, samples were lowered to a working distance of 8.5 mm and the acceleration voltage and probe current were increased to 5-7 kV and 1 nA, respectively. Areas of interests were scanned for 400 seconds each and the data exported using the same software.

Focused ion beam sectioning of crystals. Using a Zeiss Auriga Small Dual-Beam FIB-SEM, imaging of internal crystal morphology was performed on crystals retrieved from 1:1 rPFMG1 : rPFMG2 assays. For these analyses all samples were first coated with 4 nm iridium prior to SEM imaging, then coated with 50 nm of Au prior to performing FIB.[22-24,28,29] A 30 kV 120 pA gallium ion beam was oriented perpendicular to the sample by tilting the sample stage to 54˚ and utilized to mill 15 nm serial cross-sections. SEM images of cross-sectioned surfaces were then obtained using a 2.0 kV 600 pA electron beam and a secondary electron detector at a working distance of 5.0 mm. Images of surfaces containing electron beam damage were created for comparison to images of undamaged surfaces but were not used for the purposes of discussion in this publication. Images were taken shortly after cross sectioning to limit the exposure of the uncoated surfaces to the electron beam.[22-24,28,29]

Calcium potentiometric titrations. The quantitative potentiometric titration assay is based on *in-situ* Ca(II)-ion-selective potentiometric measurements at constant pH levels.[33-36] A computer-controlled automatic titration instrument manufactured by Metrohm is utilized for the experiments. The setup includes two Titrando devices (Titrando 809_1 and Titando 809_2) controlling three dosing devices (800 Dosino) for dosing $CaCl_2$ (10 mM), NaOH (10 mM) and HCl (10 mM), respectively. Calcium potential and pH values are monitored by utilizing one Ca(II)-ion-selective electrode (Metrohm No. 6.0508.110) and one pH-electrode (Metrohm No. 6.0256.100), respectively. During a titration experiment, $CaCl_2$ (10 mM) is continually titrated into carbonate buffer (10 mM) at a constant rate of 20 µl / min and the pH is kept constant at pH 8.5 by automatic counter-titration of NaOH and HCl. Here, HCl titration is required for balancing the out-diffusion of CO_2, which becomes significant below pH 9.00. In the presence of 50 nM and 500 nM rPFMG proteins and 1:1 molar mixture (Table S2, S3, Supporting Information), the potentiometric titrations are performed in 20 mL and 7 mL carbonate buffer at pH 8.5, respectively. As

the pre-nucleation slopes and nucleation times depend on the volume, the quantitative data can only be compared to corresponding reference experiments.[33-36]

Imaging and flow cytometry of individual and hybrid hydrogel particles. For detection of mesoscale protein hydrogel particles, 5 μL of a 10 μM solution of either rPFMG1, rPFMG2, or 1:1 rPFMG1:rPFMG2 in 10 mM HEPES, pH 8.0 was placed on a clean glass slide and imaged using bright field microscopy (100x lens, Nikon DS-U3 Light Microscope).[30] For flow cytometry, final concentrations of 10 μM rPFMG1, rPFMG2, and 1:1 rPFMG1:rPFMG2 in 10 mM HEPES pH 8.0 and 10 mM HEPES / 10 mM $CaCl_2$, pH 8.0, 150 μL final volume, were created and tested using a BD LSRFortessa multi-parameter cell analyzer (BD Biosciences, USA).[30] Each solution was tested in a continuous flow rate of 25 μL/min using four laser excitation lines: 405nm, 488nm, 561nm, and 640nm to register two light-scattering parameters (FSC-A and SSC-A) and the number of events for each sample. Data was collected using the BD FACS DiVa (BD Biosciences, USA) software designed for the instrument. Files were later processed using FlowJo software (TreeStar, OR, USA).[30]

AFM imaging of protein hydrogels We investigated the morphological and quantitative characteristics of rPFMG1, rPFMG2, and 1:1 rPFMG1:rPFMG2 protein assemblies and nanoparticles captured onto mica substrates.[21,22,26,27,29] The apo or Ca(II) free forms were imaged in 10 mM HEPES buffer (pH 8.0) at a protein concentration of 20 μM for both rPFMG1 and rPFGM2 individually and as a 1:1 molar ratio. In addition, we also imaged these samples in 10 mM HEPES, 10 mM $CaCl_2$, pH 8.0, to mimic conditions similar to those found in our mineralization assays.[21,22,26,27,29] AFM experiments were conducted in the temperature range of 27-31 °C using the MFP-3D-BIO AFM instrument operating in tapping mode in a buffer solution. Olympus BL-AC40TS-C2 rectangular-shaped, gold-reflex coated silicon-nitride tips with a spring constant of approximately 0.09 N/m with a resonance frequency of 110 kHz were used to achieve the best image quality were used for imaging. All samples were aliquoted onto a freshly stripped

surface of mica (0.9 mm thick, Ted Pella, Inc.) and incubated for a period of 20 min at ambient temperature prior to measurement.[21,22,26,27,29] Gwyddion software was implemented for image processing, noise filtering, and analysis, including the calculation of R_q, i.e., the surface roughness of the imaging surface.[21,22,26,27,29]

QCM-D Interaction Studies. Quartz crystal microbalance with dissipation monitoring (QCM-D) experiments were performed using a Q-sense E4 system from Biolin Scientific.[29] Gold-coated AT-cut quartz sensors were used (QSX 301, Biolin Scientific), for which the fundamental frequency was 4.95 MHz ± 50 kHz. The sensors were initially subjected to a 10 min cleaning step in an UV−ozone cleaner, with the active side facing upward. The sensors were then placed successively in solutions of Hellmanex III (2%) and ultrapure Milli-Q water (twice) and sonicated in each bath for 10 min, with the active side facing upward in all instances. The sensors were then dried with N_2 gas and replaced in the UV−ozone cleaner for 30 min. Finally, the sensors were left to soak in 100% ethanol for approximately 30 min and dried with N_2 before installation in the flow modules.[29] The QCM-D flow chambers were flushed with ultrapure Milli-Q water before each measurement (for 60-80 min), until a stable baseline was established (< 0.5 Hz shift over 10 min). The flow module temperature was maintained at 16 °C throughout, and the flow rate was kept constant at 20 μL/min. The sensor surfaces were then functionalized with poly(L-lysine) by incubation with a 0.01% aqueous solution of poly(L-lysine) until the surface was saturated (around 20−30 min). The poly(L-lysine) acts as a biocompatible support for nonspecific immobilization of proteins. Next, we introduced a 2.5 μM PFMG2 solution in HEPES buffer (pH 8, 10 mM) into the QCM-D flow chambers for approximately 45 min. This was followed by the introduction of a 2.5 μM solution of PFMG1 dissolved either in HEPES buffer or in a 10 mM $CaCl_2$ solution in HEPES buffer (pH 8, 10 mM) again for approximately 45 min. Here, we expect that pre-aggregation of the flow-introduced protein occurs in the presence of Ca(II) prior to interactions with the adsorbed protein. All

QCM-D experiments were run in duplicate and found to be reproducible (Figures S5, S6, Supporting Information).

Bioinformatics. To determine the location of disordered sequence regions[37] within the PFMG1 and PFMG2 sequences, we employed the DISOPRED3[38] and IUP_PRED[39] prediction algorithms using default parameters. Subsequently, we utilized AGGRESCAN[40] and ZIPPER DB[41] with default parameters to globally identify putative cross-beta strand sequence regions which exhibit association propensities (Figure 1). To determine a hypothetical global structure of PFMG1 and PFMG2, we utilized the DISOclust (v1.1) - IntFOLD2 integrated protein structure and function prediction server (University of Reading, UK, using default parameters), which provides tertiary structure prediction/3D sequence homology modeling of protein sequences that contain folded and unfolded sequence elements.[42]

RESULTS

rPFMG1 and rPFMG2 jointly form unique nanotextured orthogonally arranged calcite crystals featuring intracrystalline nanochambers. As in our first nacre protein pair study,[29] we initiated our study of rPFMG1 and rPFMG2 with an examination of the individual and combined effects of these proteins on *in vitro* calcium carbonate crystal growth (60 min duration). Based upon previous combinatorial studies, it was recognized that the most interesting results were obtained using a 1:1 molar ratio of the two proteins,[29] and thus this ratio became the main focus of the present study. As shown in Figure 2, the 1:1 combination of the two proteins led to surface and directional effects on calcite crystal growth that reflect the simultaneous participation of both proteins (Figure 2). Note that these results do not reflect the definitive functionalities of either protein, which can only be ascertained using in situ studies. Individually, we note that rPFMG1 induces the formation of highly nanotextured crystals or crystal clusters, whereas rPFMG2 promotes new directional growth with evidence of nanotexturing at the

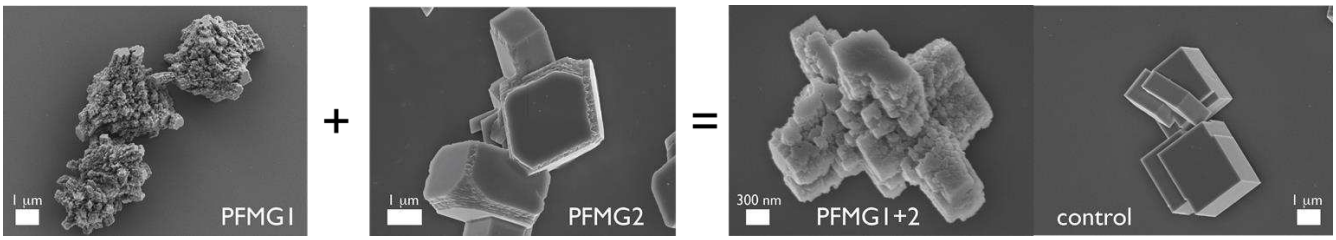

Figure 2. SEM images of Si-wafer collected calcite crystals generated in the presence of rPFMG1, rPFMG2, 1:1 rPFMG1:rPFMG2 and protein-deficient control assays (control).

junction of crystal surfaces. When both proteins are present at a 1:1 molar ratio, we observe the formation of orthogonally arranged calcite crystals (as verified by microRaman spectroscopy, Figure S4, Supporting Information) that feature nanotextured surfaces (Figure 2), i.e., the nucleation and crystal growth processes now simultaneously incorporates the individual functional properties of both rPFMG1 and rPFMG2, similar to what we have reported in previous combinatorial nacre protein mineralization studies.[29]

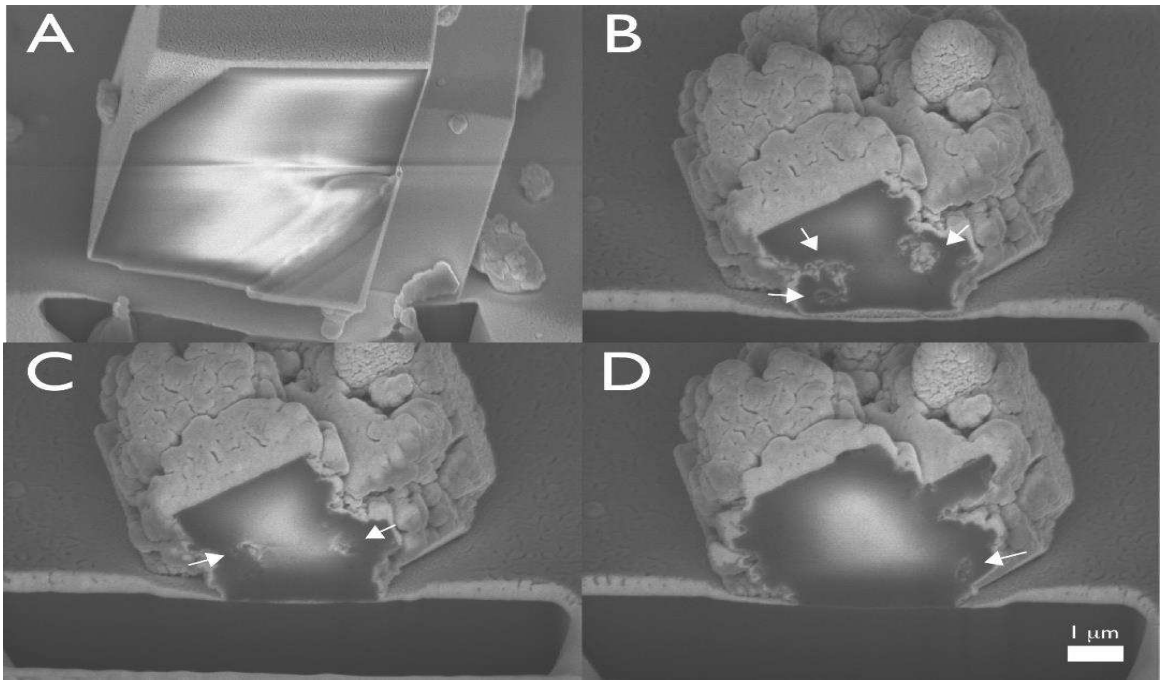

Figure 3. SEM images of FIB-sectioned calcite crystals. (A) Typical protein-deficient assay calcite crystal (60 minutes assay time), which features minimal porosities. (B,C,D) Representative sequential sections of calcite crystals generated in assays containing 1:1 rPFMG1 : rPFMG2 (20 µM each), where one can clearly denote multiple subsurface void or pore regions (arrows).

The 1:1 scenario generated very interesting results, and given that the nanotextured surfaces of the crystals might entrap protein aggregates or hydrogels and incorporate these within the mineral overgrowth phase,[22-24,28,29] we focused on these crystals for subsequent subsurface analysis. Using FIB-SEM we sectioned these crystals and discovered the presence of intracrystalline nanoporosities (Figure 3). The majority of these nanoporosities appear to be peripherally located within these hybrid crystals, indicating that the potential mechanism for intracrystalline nanoporosity formation in this scenario involved the initial deposition of both protein phases onto exposed crystal surfaces, followed by mineral overgrowth and entrapment of these phases within the nucleating crystals.[22-24,28,29] This same mechanism of intracrystalline nanoporosity formation has also been observed in nacre protein crystal growth studies in vitro[22-24,28,29] and will be discussed in more detail later on. Thus, with regard to later stages of crystal growth, rPFMG1 and rPFMG2 functionalities are both manifested and exhibit overlap, leading to changes in crystal morphology, the creation of nanotextured surfaces, and the introduction of intracrystalline nanoporosities.

The effect of rPFMG1 + rPFMG2 on non-classical nucleation. Our focus now shifts to earlier *in vitro* mineralization events which involve the formation of pre-nucleation clusters (PNCs) and their transformation into amorphous calcium carbonate (ACC) , i.e., the non-classical nucleation scheme.[33-36] It has been established that nacre proteins modulate the early events in the non-classical nucleation pathway of calcium carbonates.[22,25-27,29] For this reason, the effects of rPFMG1 and rPFMG2 on $CaCO_3$ nucleation were investigated utilizing a quantitative potentiometric titration assay, which is based on *in-situ* Ca(II)-ion-selective potentiometric measurements at constant pH levels (Figure 4; Tables S2, S3, Supporting Information).[33] During the pre-nucleation regime, differences between detected and added Ca^{2+} ions evidence the binding of calcium ions indicating the formation of $CaCO_3$ pre-nucleation clusters (PNCs).[35] The slope of the pre-nucleation gradient is a measure for the stability of the PNCs, and the

timeline for their formation denotes their kinetics.[33,35] As shown in Figure 4A, in the presence of rPFMG1, there is significantly flatter pre-nucleation gradient in comparison with the reference titration curve at pH 8.5, showing that rPFMG1 increases the thermodynamic stability of PNCs at a protein concentration of 500 nM. In contrast, rPFMG2 does not significantly affect the stability of PNCs. Most importantly, the change of pre-nucleation gradient in the 1:1 protein mixture sample is similar to that in rPFMG1 sample, suggesting that there is no synergistic effect of rPFMG1 and rPFMG2 on PNC stability. Rather, rPFMG1 activity predominates. In the presence of rPFMG1, rPFMG2 and 1:1 protein mixtures, nucleation of a solid, which becomes apparent by the drop in free calcium (Figure 4A) and ion product (Figure 4B), is delayed, corresponding to a retarding factor of 1.36, 1.27 and 1.43, respectively, relative to the reference experiment (Tables S2, S3, Supporting Information). This result indicates that the 1:1 rPFMG1 : rPFMG2 mixture at best induces a slightly longer nucleation delay.[36] The solubility of initially precipitated phase (i.e., the threshold of the constant ion product after nucleation) for the individual rPFMG proteins or 1:1 protein mixtures are lower than reference scenario, suggesting that both rPFMG1, rPFMG2, and the 1:1 protein mixture promote the formation of less soluble calcium carbonate phases (Figure 4B; note that at 50 nM protein concentrations no significant effects are observed on PNCs stability, nucleation time or solubility of initially precipitated phase, Table S3, Supporting Information). In summary, rPFMG1 stabilizes PNCs, while rPFMG2 does not show any significant influence on thermodynamic stability of the PNCs. Individually, both proteins slightly inhibit the nucleation of ACC, whereas the initially formed phases are thermodynamically more stable (i.e., less soluble) than in the reference experiments. Thus, unlike the results observed for existing crystals at later stages of the nucleation process (Figures 2,3), the effects generated by the 1:1 protein mixtures on the early stages of nucleation show no synergistic tendencies like that we reported for AP7 - rPFMG1,[29] and are similar to those of rPFMG1 alone, indicating that within rPFMG1+rPFMG2 complexes it is the rPFMG1 protein

that largely manipulates the nucleation process, with minimal contributions from rPFMG2. This indicates that the surface modifications induced by the proteins on the crystals (Figure 2) indeed mostly develop during crystal growth rather than nucleation.

rPFMG1 + rPFMG2 assembly and hydrogelation processes are modulated by rPFMG2. As noted in previous individual[21,25-27,30] and combinatorial[29] nacre protein studies, these proteins are strong aggregators and will form hydrogels at pH 8, with the strongest aggregation propensity observed when

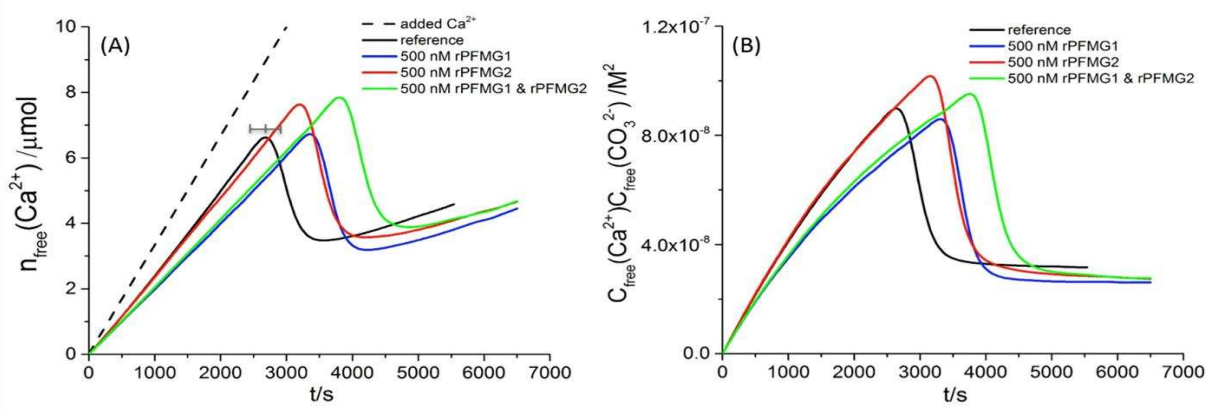

Figure 4. The development of (A) free Ca (II) ion concentration and (B) calcium carbonate ion product in potentiometric titrations in the absence and presence of 500 nM rPFMG proteins, as indicated, in 10 mM carbonate buffer at pH 8.5 as a function of time. In each plot the reference curve refers to experiments conducted in the absence of protein. Experiments were performed in duplicate (not shown), and the reproducibility was good (see Tables S2,S3). In (A), the gray bar shown at the maximum of the reference curve illustrates the average nucleation time (center) ± the maximum deviation observed in 4 individual reference experiments.

Ca(II) is present. This is certainly true for rPFMG1 as documented in published studies,[21,29] and, as shown in Figure S3, Supporting Information, MALDI-TOF-MS studies confirm that rPFMG2 is a strong aggregator with the unique capability of forming stable multimers[26] under the low pH conditions of the matrix mixture (i.e., dimer, trimer, tetramer, etc.), something which rPFMG1 does not exhibit (Figure S2). Thus, given that both protein sequences are aggregation-prone (Figure 1), with one being capable of multimeric complexation, the question arises as to how these traits manifest themselves when both proteins are present in a 1:1 molar ratio.

The answer to this question becomes evident when we examine this 1:1 mixture using MALDI-TOF-MS (Figure 5) and compare these findings to those obtained for the individual proteins (Figures S2, S3, Supporting Information). Here, we see the formation of detectable multimers;[26] however, the m/Z values do not correspond to those generated by rPFMG2 or rPFMG1 alone. Based upon the known m/Z values for both proteins (rPFMG1 = 13608 Da; rPFMG2 = 18867.6 Da), it appears that these multimers are in fact heterocomplexes that approximately correspond to rPFMG2 : rPFMG1

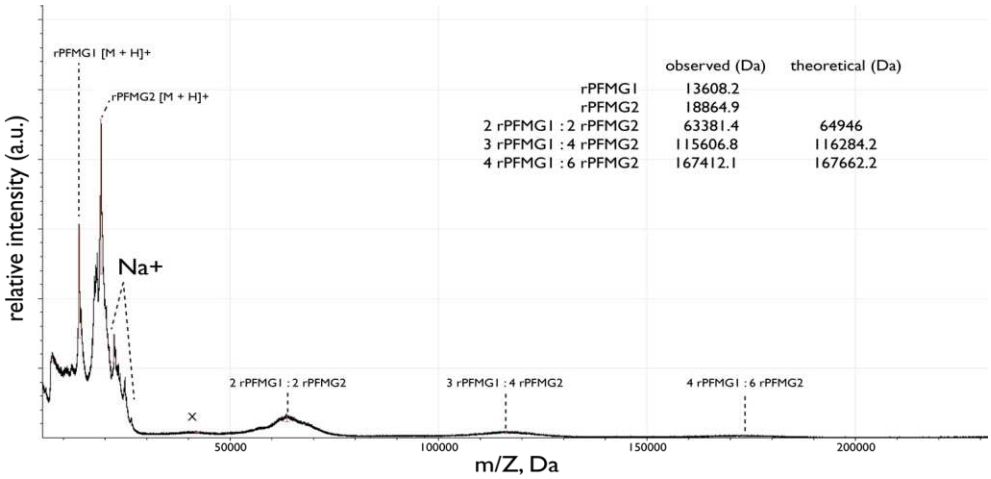

Figure 5. MALDI-TOF-MS spectra of 1:1 molar ratio rPFMG1:rPFMG2 20 μM sample in α-CHCA matrix. Assignment of adduct peaks, observed and theoretical protein ratios are provided in the figure.

ratios of 2:2, 4:3, and 6:4, ± 500-1500 Da. Thus, we have uncovered an interesting phenomenon: co-aggregation occurs between rPFMG1 and rPFMG2, with a defined complexation or aggregation ratio in which rPFMG2 predominates. Three things should be noted. First, MALDI-TOF-MS is qualitative and cannot discern protein-protein affinities, and thus at this stage we cannot comment on the binding strengths that each protein might manifests for its partner. Second, note that there may be the potential for even higher-ordered rPFMG2:rPFMG1 complexes that for various reasons cannot be detected by MALDI-TOF.[26] Finally, we note that matrix conditions are low pH or acidic, and thus the interactions between rPFMG1 and rPFMG2 may be fostered by charge neutralization of protein Asp, Glu carboxylate

residues. We will return to this topic later when we explore the effects of Ca(II) on rPFMG1 – rPFMG2 binding (see below).

We continued our investigations into the 1:1 rPFMG1:rPFMG2 aggregation process using light microscopy in combination with flow cytometry (Figure 6).[30] Here, we observe that individual rPFMG1 and rPFMG2 samples form translucent mesoscale hydrogel particles at pH 8, with evidence of internal porosities or void regions (Figure 6, top panel). Hydrogelation is a common phenomenon for

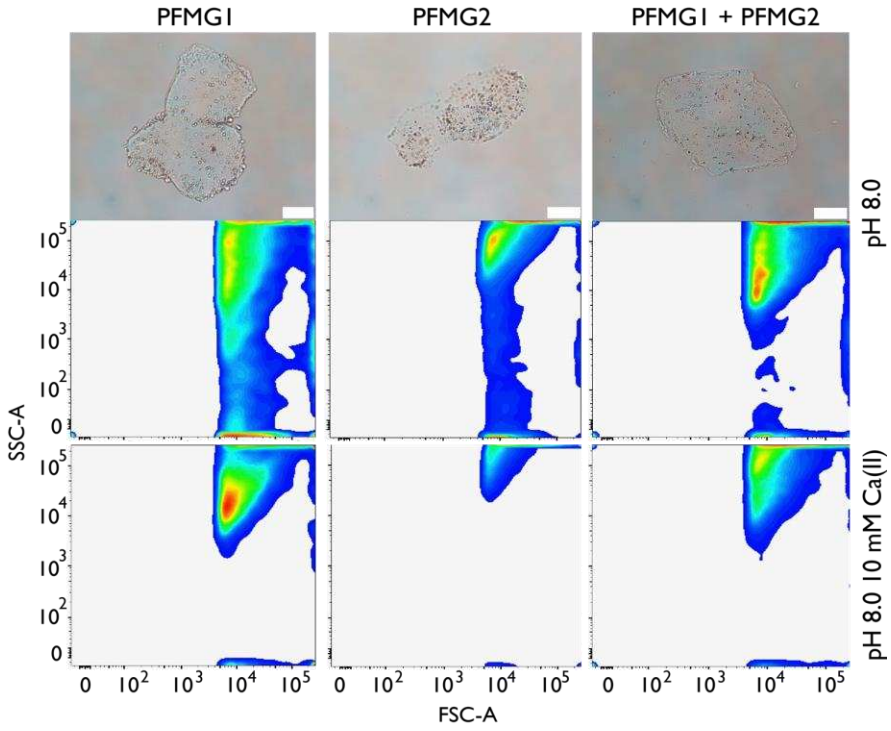

Figure 6. (TOP PANEL) Light microscopy images of nacre protein hydrogels, 10 mM HEPES, pH 8.0. Scalebars = 10 µm.(MIDDLE PANEL) Flow cytometry 2-D density plots of FSC as a function of SSC, Ca(II)-free, 10 mM HEPES, pH 8.0 (BOTTOM PANEL) Flow cytometry 2-D density plots of FSC as a function of SSC, 10 mM $CaCl_2$, 10 mM HEPES, pH 8.0 Forward scattered light (FSC, x-axis) determines particle size distributions; and side-scattered light (SSC, y-axis) measures refracted and reflected light that occurs at any interface within the particles where there is a change in refractive index (RI) that results from variations in particle granularity or internal structure.

several nacre- and sea urchin spicule-specific proteins,[21,22,25-27,29,30] and thus these two pearl-associated proteins apparently share this common trait. When present in a 1:1 molar ratio, we note that the resultant rPFMG1:rPFMG2 hydrogels are similar in appearance with regard to morphology and dimension of the hydrogels generated by the individual proteins (Figure 6, top panel), i.e., we do not note any particle

enlargement or other detectable changes. Using flow cytometry experiments, which allow us to move beyond simple light microscopic visualization and determine the particle size distributions (FSC-A, forward scattered component, x-axis) versus particle granularity or internal complexity (SSC-A, side-scattered component, y-axis) of the particles,[30] we found similar results (Figure 6, middle and lower panels). Here, at pH 8 in the absence of Ca(II) we see that there are differences in particle size distributions and internal complexities or structure for each protein. When we examine the 1:1 molar ratio sample under these conditions, we note that the FSC-A and SSC-A parameters have contributions from the individual proteins (Figure 6, middle panel) and this verifies that the 1:1 mixture generates *hybrid* hydrogel particles. These findings are also observed when Ca(II) is present: First, we note that the FSC-A and SSC-A parameters for rPFMG1 and rPFMG2 hydrogels are different, and, the 1:1 rPFMG1:rPFMG2 sample appears to share parameters that are similar to both rPFMG1 and rPFMG2 (middle panel, Figure 6). Second, we note that the introduction of Ca(II) leads to a significant change in the internal complexity or structure of the hydrogel particles individually formed by rPFMG1 and rPFMG2, and the 1:1 rPFMG1:rPFMG2 sample appears to demonstrate granularity traits shared by both proteins (bottom panel). Thus, individually rPFMG1 and rPFMG2 are Ca(II)-responsive hydrogelators and the 1:1 rPFMG1 and rPFMG2 mixture contains hybrid hydrogel particles at pH 8 in the presence and absence of Ca(II).

To examine the hydrogelation process on surfaces, we next studied the individual pearl proteins and their 1:1 mixture on freshly cleaved mica surfaces under different conditions using AFM tapping mode microscopy (Figure 7).[21,22,25-27,29] Here, we use AFM to determine the ultimate size ranges of the protein hydrogel particles that cannot be determined by methods such as MALDI-TOF-MS. What we find here is further evidence that supports both our MALDI-TOF-MS (Figure 5) and flow cytometry (Figure 6) datasets, namely, that rPFMG1 and rPFMG2 co-aggregate and form hybrid particles. If we

examine the first scenario [pH 8.0, no Ca(II)], it is clear that each individual protein forms particles and films on mica surfaces (Figure 7A). We note that the particle diameters for rPFMG1 are 1.5x greater than those of rPFMG2, whereas the particle heights are very similar for both proteins. Conversely, the film thicknesses of rPFMG2 are 5-10% greater than those of rPFMG1 (Figure 7B). Under these same conditions the particle dimensions and film thicknesses for the 1:1 rPFMG1 : rPFMG2 sample are smaller (by a factor of 2-3) than those of either rPFMG1 or rPFMG2 individually. This would suggest that when rPFMG2 combines with rPFMG1 (Figure 5) the overall dimensions and film thicknesses of the hybrid

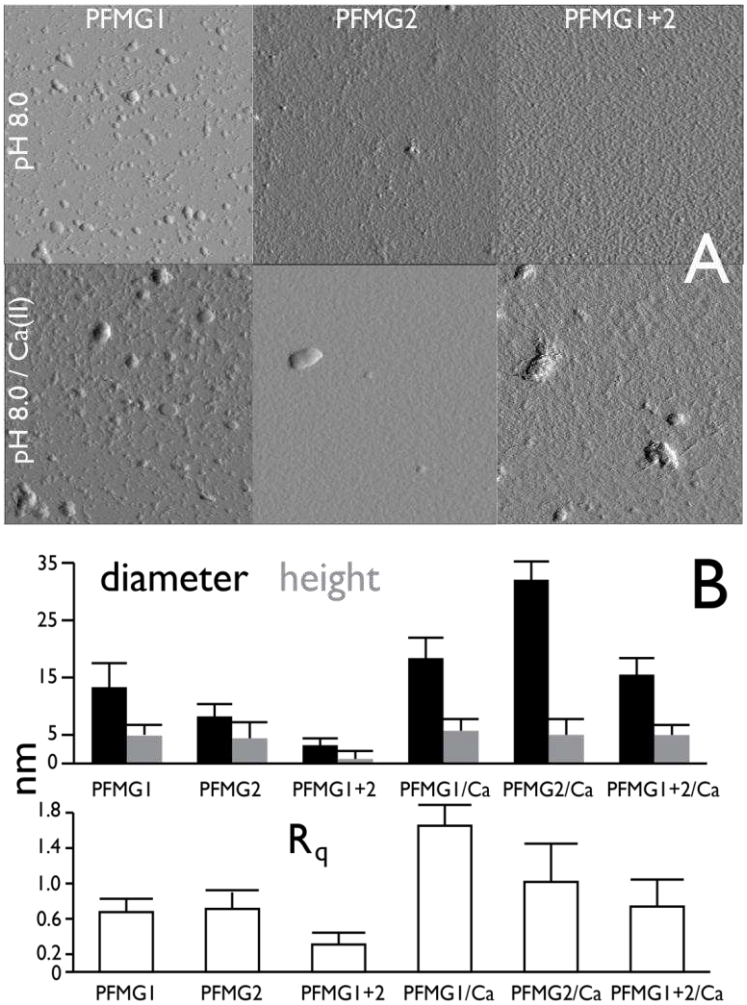

Figure 7. (A) Tapping mode AFM amplitude images of rPFMG1, rPFMG2, and 1:1 rPFMG1:rPFMG2 proteins samples in either 10 mM HEPES, pH 8.0, or 10 mM HEPES / 10 mM $CaCl_2$, pH 8.0. (B) Histogram plots of AFM-determined protein particle diameters, heights, and R_q (surface roughness)

particles are attenuated (Figure 7B). When we consider the second scenario [pH 8.0 with Ca(II)], we note increases in particle sizes and film thicknesses, with rPFMG2 > rPFMG1 in terms of particle diameters and film thicknesses (by 35%) and rPFMG2 ~ rPFMG1 in terms of particle heights. However, the 1:1 rPFMG1 : rPFMG2 sample still exhibits attenuation with regard to particle diameters and film thicknesses (Figure 7B) relative to the data obtained for the individual proteins in the presence of Ca(II). In summary, we find that rPFMG1 and rPFMG2 co-aggregate and form hydrogel particles and films on mica surfaces at pH 8.0 in the presence and absence of Ca(II), with evidence of dimensional attenuation, which we postulate is the result of rPFMG2-driven specific molecular combinations of each protein (Figure 5).

QCM-D studies of pearl protein adsorption reveal interesting molecular relationship between rPFMG1 and rPFMG2. From the foregoing, it appears that rPFMG1 and rPFMG2 form hybrid aggregates (Figure 5) or hydrogels (Figures 6,7), i.e., there are intermolecular interactions taking place between both nacre proteins under low ionic strength and mineralization conditions. To probe for these interactions we initiated a QCM-D study similar to the one utilized in earlier nacre combinatorial studies,[29] i.e., we started with one protein immobilized on the surface (rPFMG2) of an Au-coated QCM-D sensor placed in a flow module and introduced the second protein (rPFMG1) over time.[29,43] As we acknowledged in our earlier studies, this approach cannot fully replicate the 1:1 proportionality scenario we employed in our other experiments (Figures 2-7)[29] since we are immobilizing one protein onto the sensor first then introducing the second protein, rather than mixing them together in a predetermined ratio. Thus, we are merely measuring the interaction of the free protein component with the immobilized protein component.

The QCM-D experiments were conducted under two conditions: 10 mM HEPES, pH 8.0, and 10 mM HEPES / 10 mM $CaCl_2$, pH 8.0, to mimic mineralization conditions without interference from ACC or calcite precipitation (Figure 8A,B). Note that in the latter instance Ca(II) ions will promote

aggregation of either protein (Figures 6,7) and thus pre-aggregation of the flow-introduced protein component will occur prior to interaction with the surface-adsorbed protein component.[29,43] In these

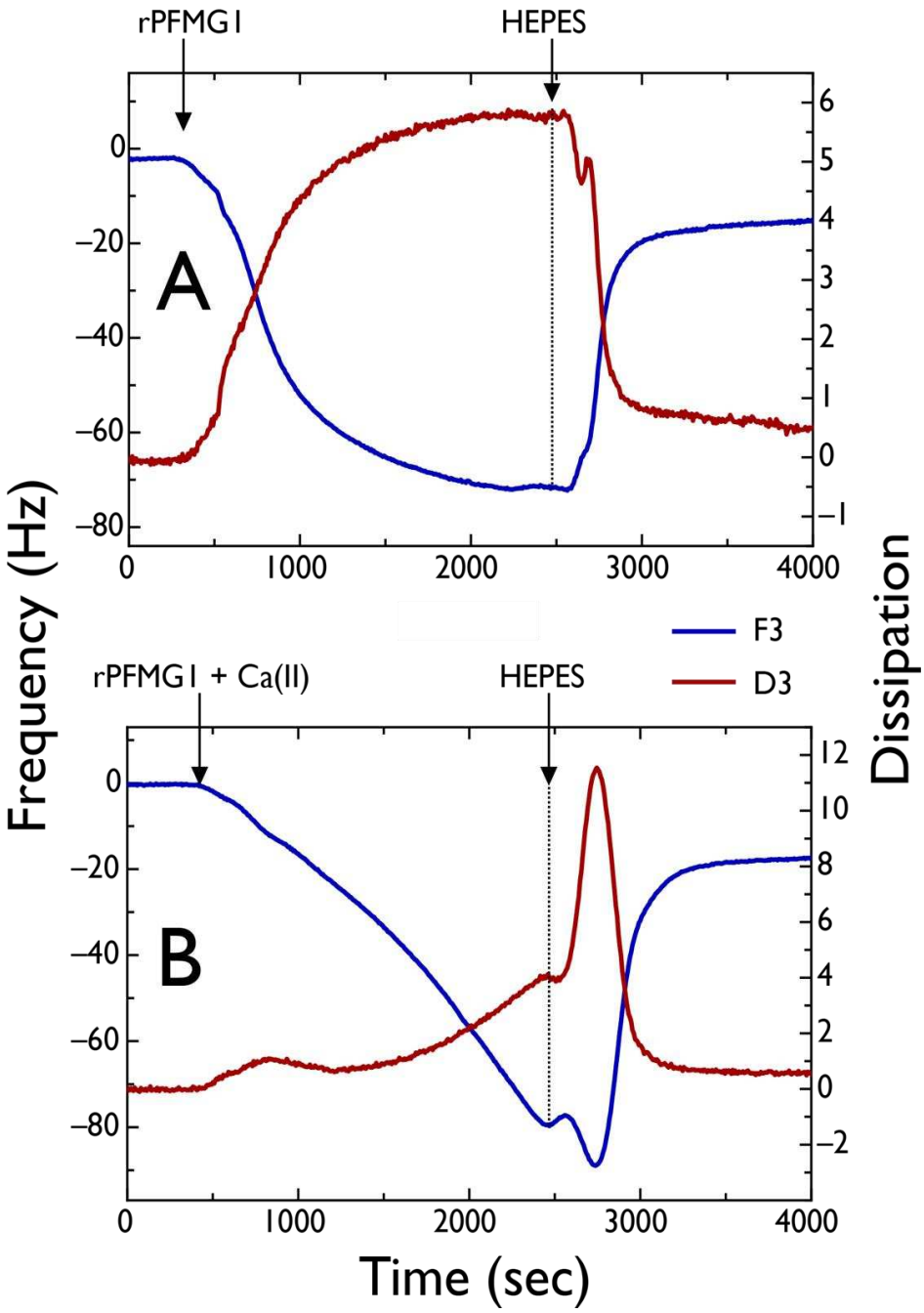

Figure 8. QCM-D experiments of immobilized rPFMG2 exposed to 2.5 μM rPFMG1 under the following conditions: (A) rPFMG1 in 10 mM HEPES, pH 8.0; (B) rPFMG1 in 10 mM HEPES, 10 mM $CaCl_2$, pH 8.0. Plots show the 3rd harmonic frequency (F3) and dissipation (D3) observed under each scenario. The time-dependent introduction of rPFMG1 alone or rPFMG1+Ca(II) and HEPES washing solutions are noted on the plots by arrows and extended dashed lines. These experiments were duplicated (See Supporting Information, Figures S5, S6) and found to be reproducible.

studies, poly(L-lysine)-coated Au QCM-D sensors were utilized to provide a biocompatible surface for initial rPFMG2 protein adsorption [note that both rPFMG1 and rPFMG2 contain Asp, Glu residues, Figure 1, and thus should be electrostatically compatible with poly(L-lysine) coatings]. In each of the following scenarios we created a layer of adsorbed rPFMG2 protein on poly(L-lysine) using a 2.5 μM protein solution. In the first scenario we introduce rPFMG1 protein (2.5 μM) in buffered media (arrow) to this immobilized rPFMG2 layer. As we see in Figure 8A, in the absence of Ca(II) we observe a standard Langmuir isotherm upon association of rPFMG1 to surface immobilized rPFMG2. When returning to running buffer alone (HEPES, 2nd arrow) we see a significant level of rPFMG1 dissociation from the rPFMG2 layer. The dissipation (which reflects viscoelasticity) also increases dramatically (following the same exponential behavior) before again dropping sharply upon washing in HEPES. These results were found to be reproducible (Figure S5, Supporting Information) and indicate that both proteins non-covalently interact with one another.

In the second scenario, 2.5μM rPFMG1 in a Ca(II) HEPES buffer is introduced to the immobilized rPFMG2 layer (Figure 8B, 1st arrow). Here, the presence of Ca(II) leads to more protein mass deposited compared to the HEPES-only scenario (Figure 8A), i.e. there is greater binding between rPFMG1 and rPFMG2 with Ca(II) ion present. This effect is repeatable (Figure S6, Supporting Information) and does not follow the standard Langmuir isotherm (see Figure 8A). Note also that in the presence of Ca(II) the binding kinetics are much slower - over the time scale of the experiment we do not reach the point where the rPFMG2 surface becomes saturated (blue frequency curve, Figure 8B). We also do not see the same large shift in dissipation as we did in the absence of Ca(II). With the introduction of HEPES buffer alone (2nd arrow) we flush the Ca(II) out of the flowcell and note the dissociation of rPFMG1 from the rPFMG2 layer (Figure 8B), similar to what we saw in the absence of Ca(II) (Figure 8A). Recall that under low pH or conditions of Asp, Glu carboxylate charge neutralization, we observed

complexation and specific binding ratios between these two proteins (Figure 5). From these results, our interpretations are as follows: a) at pH 8.0 rPFMG1 and rPFMG2 largely interact and the rPFMG1 just builds up on top of the rPFMG2 layer leading to a rapid and large increase in mass and a highly viscoelastic layer (Figure 8A). b) However, Ca(II) ions mediate the interactions between the two proteins (Figure 8B) leading to a much slower increase in mass and a more rigid rPFMG1 adlayer. Upon rinsing, the Ca(II) ions dissociate quickly so the rPFMG2 adsorbed layer now resembles that seen without the Ca(II) (i.e. highly viscoelastic - hence the peak in Figure 8B) after which the rPFMG1 desorbs. Given that rPFMG1 contains a Ca(II)-responsive pseudo-EF hand domain,[21] it is plausible that this change in adsorption kinetics is due to conformational changes occurring in rPFMG1, which, in turn, may affect how rPFMG1 self-aggregates and/or interacts with rPFMG2. Alternatively, based upon our MALDI-TOF-MS findings (Figure 5), charge screening or neutralization of protein Asp, Glu carboxylate groups [either by H+ or Ca(II)] may also be the driving force for rPFMG1 – rPFMG2 interactions.

The molecular basis for rPFMG1 – rPFMG2 interactions. Based upon the foregoing, we now understand that rPFMG1 and rPFMG2 are specifically interactive (Figures 5, 8) in the presence of Ca(II). Using predictive folding bioinformatics (DISOclust INTFOLD),[42] we generated putative structures of both proteins to allow us to visualize each structure and determine regions which may allow PFMG1 – PFMG2 to interact. As shown in Figures 1 and 9, both proteins possess a combination of intrinsically disordered regions and folded domains. Specifically, PFMG1 has a significant disordered N-terminal region and a disordered A90-D116 region hat overlaps with the pseudo-EF-hand helical domain located at D61-E102 (C-terminal). This EF-hand domain exhibits homology (~50%) to the crystal structures of Troponin C (4tncA2)[44] and Calmodulin-like domain protein kinase isoform 3 (3kheB2)[45](confidence P value = 3.139 E-3; global model quality score = 0.5452). Interestingly, PFMG2 possesses short regions of intrinsic

disorder at the C-terminus (list residues) and a folded helical-loop domain (A1-G159) that exhibits homology (~51%) with the crystal structures of the calponin domain of the Ras GTPase activating-

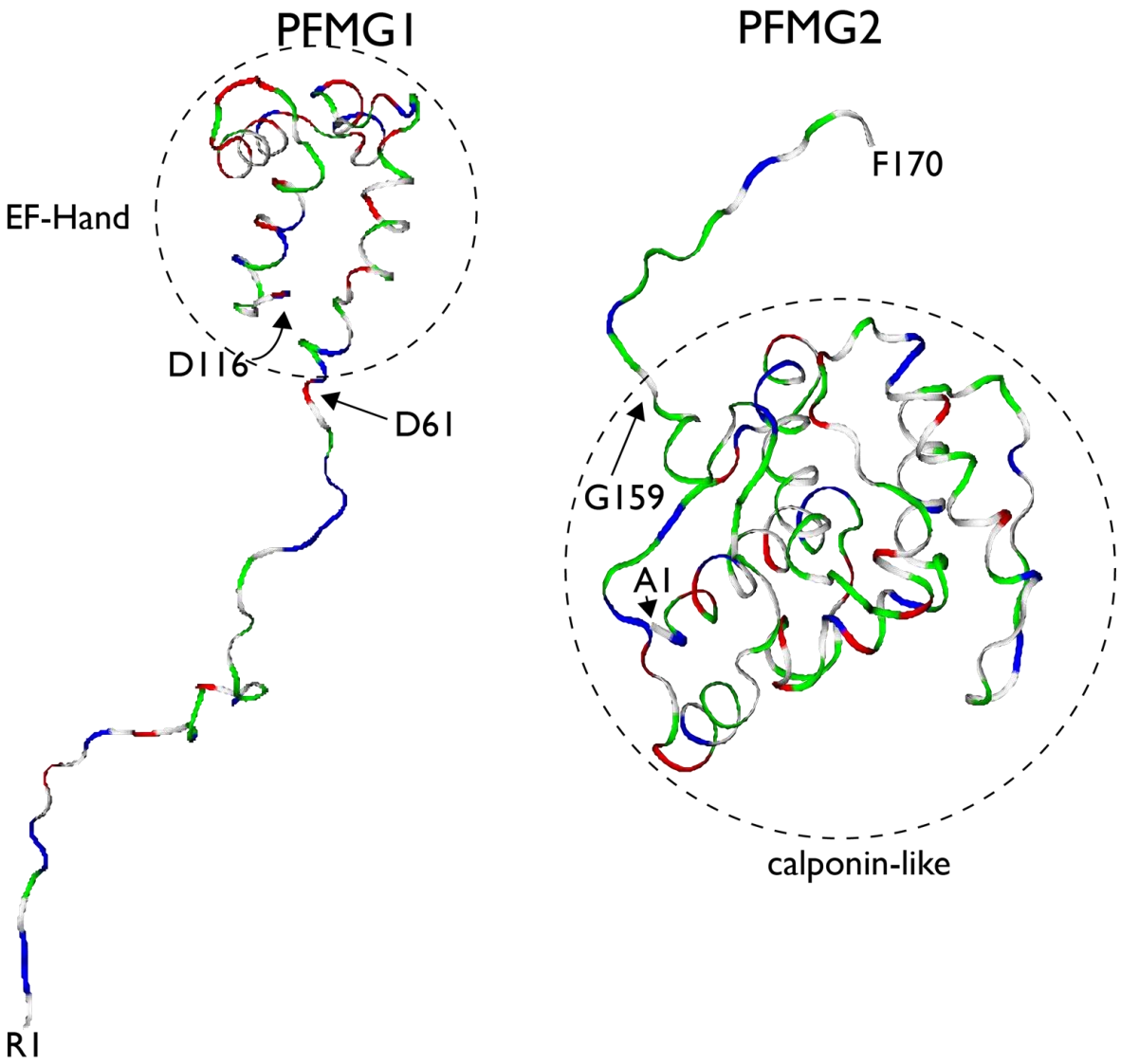

Figure 9. Multiple views of DISOclust IntFOLD2[42] predicted 3D modeling predicted structure of PFMG1 and PFMG2 proteins in ribbon representation (backbone only). The EF-Hand and calponin-like domains are encircled and domain boundary amino acids are identified. Color designations of backbone regions: red = anionic; blue = cationic; green = polar; white = hydrophobic.

like proteins 1p2xA (yeast)[46] and 3i6xA (human)[47](confidence P value = 4.256 E-4; global model quality score = 0.7368). This domain is notable for its dimerization interactions with the EF-hand domain of Ca(II)-bound calmodulin.[46,47] Thus, this explains why rPFMG2, with an alpha-helical calponin-like

domain, interacts with the pseudo-EF hand rPFMG1 (Figure 5, 8), particularly when Ca(II) is present (Figure 8). Moreover, the presence of unstable intrinsically disordered regions, coupled with aggregation-prone amyloid-like domains (Figure 1), may also enhance the interactiveness of rPFMG1 and rPFMG2 and promote the formation of higher-ordered protein complexes (Figure 5, Figure S3, Supporting Information) leading to the creation of hydrogel particles (Figures 6,7), as noted for other disordered, aggregation-prone nacre proteins.[21,22,25-27,29]

DISCUSSION

This report follows up from our initial 2-nacre protein study with different proteomic origins: AP7 (intracrystalline, Pacific Red Abalone *Haliotis rufescens*) and rPFMG1 (pearl-associated, Japanese pearl oyster *Pinctada fucata*).[29] These unrelated proteins exhibited an unusual, unexpected synergistic effect at 1:1 molar ratio on both the early and later stages of calcium carbonate nucleation as well as the aggregation/hydrogelation process. In our present study, we focus on a different pair, PFMG1 and PFMG2, which are co-members of the same pearl PFMG proteome in *P. fucata*.[10,11,13,15,16] Here, we assessed how each individual protein and the 1:1 molar ratio affect calcium carbonate nucleation (Figures 2-4) and protein hydrogelation (Figures 5-8). To some extent the results obtained in the present study mirror those obtained in the 1:1 AP7 : rPFMG1 study,[29] namely, that the 1:1 rPFMG1 : rPFMG2 sample exhibit synergistic effects on later stages of calcium carbonate crystal growth (Figure 2) where we note overlapping morphological (rPFMG1) and directional growth (rPFMG2) changes in calcite. Further, we note that in both studies the 1:1 protein mixtures also create intracrystalline nanoporosities within these surface-modified crystals (Figure 3).[29] Presumably, both the surface and subsurface modifications to these crystals are due to the deposition and mineral engulfment of hybrid hydrogel particles created by

the protein mixtures (Figures 6,7), and the dimensions of these hydrogels are Ca(II)-dependent, as we noted in earlier 1:1 AP7 : rPFMG1 studies.[29]

However, other aspects of the 1:1 rPFMG1 : rPFMG2 nucleation and aggregation processes are quite different from those reported for 1:1 AP7 : rPFMG1, and our data suggests that there exists a special molecular and functional relationship between the pearl proteomic rPFMG1 and rPFMG2 members. Within the 1:1 rPFMG1 : rPFMG2 mixture we detected co-aggregates that possessed specific molecular combinations (rPFMG2 : rPFMG1 = 2:2; 4:3, and 6:4)(Figure 5). These combinations follow the relationship rPFMG2 > rPFMG1. Since rPFMG2 is capable of defined multimeric assembly (Figure S3) but rPFMG1 is not (Figure S2), we conclude that rPFMG2 incorporates rPFMG1 molecules into multimeric co-assemblies. This finding is echoed in the formation of hydrogels generated by the 1:1 rPFMG1 : rPFMG2 mixture, where the hydrogel particles do not increase in overall dimension as we witnessed in the 1:1 AP7 : rPFMG1 study. Rather, there appears to be a limitation imposed by rPFMG2 upon hybrid hydrogel particle and film dimensions (Figures 6,7) in the presence and absence of Ca(II). Functionally, although these hybrid PFMG complexes exhibit synergistic effects on the later stages of nucleation (Figures 2,3) they do not exhibit synergism during the earlier stages of nucleation that involve PNC and ACC formation (Figure 4; Supporting Information, Tables S2, S3). Rather, it appears that within the 1 : 1 rPFMG1 : rPFMG2 complexes it is the rPFMG1 protein that largely manipulates the non-classical nucleation process, with rPFMG2 offering minimal contributions. Thus, unlike the highly cooperative relationship between AP7 and rPFMG1,[29] rPFMG1 and rPFMG2 are functionally synergistic only within a specific scenario and the co-aggregation and hydrogelation processes are specific in scope.

Our studies also provide information regarding the non-covalent interactions between rPFMG1 and rPFMG2 (Figure 8). We know that rPFMG2 contains a calponin-like folded domain[46,47] that is interactive with EF-hand domains, such as the type found in rPFMG1,[21,44,45] (Figure 9). At this point in

time we do not know which residue-specific docking interactions occur between these two proteins. However, we do have clues as to what types of amino acid residues may be involved in rPFMG1 – rPFMG2 binding. For example, under low pH conditions, which would foster Asp and Glu sidechain carboxylate charge neutralization, specific complexation ratios of rPFMG1 and rPFMG2 form (Figure 2). Further, in the absence of Ca(II) at pH 8.0, both proteins interact with one another, but are easily displaced with simple washing techniques (Figure 8A). We conclude that the non-covalent interactions between both proteins can change based upon solution conditions and may involve Asp, Glu residues. This becomes even more evident once we introduce Ca(II) (Figure 8B), where we find that Ca(II) mediates the intermolecular interactions between rPFMG1 and rPFMG2 and intensifies the binding of these two proteins but slows the overall binding kinetics. Under these conditions a number of potential scenarios may explain this binding effect, such as non-specific Ca(II) induced charge neutralization of protein Asp, Glu carboxylate groups or Ca(II) coordination at Asp, Glu residues with concomitant protein conformational change (i.e. the EF hand domain of rPFMG1 binding to the calponin domain of rPFMG2, Figure 9).[44-47] Additional studies will be required to firmly establish the non-covalent interactions, specific domain associations, and affinities that drive rPFMG1 – rPFMG2 interactions and define complexation ratios.

This now leads us to some interesting hypotheses regarding the *in vitro* relationship between rPFMG1 and rPFMG2, and potentially provide insights into the *P. fucata* PFMG proteome itself. We believe that the specific co-aggregation of rPFMG2 with rPFMG1 (Figure 5) regulates the dimensions (Figures 6,7) and internal structure (Figure 6) of the hybrid hydrogels, with rPFMG2 acting as a "gatekeeper" of overall hydrogelation. These intermolecular interactions would be intensified in the presence of Ca(II)(Figures 6-8), and would involve a Ca(II)-induced conformational event in the pseudo-EF hand domain of the rPFMG1 protein (Figure 8)[21] and subsequent interaction with the calponin-like

domain of rPFMG2.[46,47] With regard to the early stages of non-classical nucleation, given the functional activity relationship rPFMG2 < rPFMG1 (Figure 4; Tables S2, S3, Supporting Information), the hydrogelator "gatekeeper" rPFMG2 contributes minimally to this process, with rPFMG1 functional activities being largely manifested in the hybrid complexes. But, this is not the case for later stages of the nucleation process (Figures 2,3), where the functional activities of both proteins appear relatively balanced and as a result both proteins within the hybrid hydrogels synergistically modify both the surface and the subsurface regions of calcite crystals. Thus, we must conclude that by combining two proteins with different functionalities in a specific molecular ratio in vitro, the resulting hydrogels are restricted in size and internal structure and selectively manifest certain functional characteristics (individual or synergistic) during different timepoints (early, Figure 4; late, Figures 2,3) in the mineralization process. Although we do not know the true molecular ratio or temporal expression of PFMG1 and PFGM2 expressed *in situ* within the pearl oyster,[11-13,15-17] from our experiments we can infer that if both proteins were co-expressed simultaneously or co-exist at a common timepoint during pearl formation, then PFMG1, PFMG2, and potentially other PFMG proteins would complex together in specific ratios to form a hybrid hydrogel phase (Figures 4,6,7). In turn, this hybrid hydrogel phase could selectively regulate the early (Figure 5) and later (Figure 2) stages of calcite and aragonite nucleation leading up to the formation of the composite pearl that possesses internal nanoporosities or protein hydrogel inclusions (Figure 3) that modulate the fracture resistance, strength, and appearance of the pearl itself. We intend to follow up on these hypotheses with further experimentation involving different members of the PFMG family using high throughput methodologies.

ASSOCIATED CONTENT

Supporting Information. Primary sequence of bacterial recombinant rPFMG2 (Figure S1); MALDI-TOF-MS of rPFMG1 (Fig S2); MALDI-TOF-MS and SDS-PAGE gel electrophoresis of rPFMG2 (Fig S3); Experimental procedures and MicroRaman spectra of mineral deposits generated from (-) and (+) 1:1 rPFMG1:rPFMG2 mineralization assays (Fig S4, Table S1); Potentiometric dataset obtained from rPFMG1, rPFMG2, and 1:1 rPFMG1 : rPFMG2 (Tables S2, S3). Full QCM-D dataset obtained for rPFMG1 – rPFMG2 interactions on Poly-L-Lys coated Au sensors in the presence and absence of Ca(II), pH 8.0 (Figure S5, S6).

AUTHOR INFORMATION

Corresponding Author

*To whom correspondence should be addressed: John Spencer Evans, Laboratory for Chemical Physics, Division of Basic Sciences and Center for Skeletal and Craniofacial Medicine, New York University College of Dentistry, 345 E. 24th Street, New York, NY, 10010. Tel.: (212) 998-9605; Fax: (212) 995-4087. Email: jse1@nyu.edu.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

Portions of this research (protein production and purification, assays, SEM, AFM) were supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Award DE-FG02-03ER46099 (JSE), and the QCM-D research was supported by grants from the EPSRC (EP/M028127/1) and the MRC (Discovery Grant MCPC15073)(SJ, JJC).

ACKNOWLEDGMENT

We thank the Molecular Cytology Core Facility at Memorial Sloan-Kettering Cancer Center New York for their help with the AFM imaging experiments. This paper represents Contribution Number 86 from the Laboratory for Chemical Physics, New York University.

REFERENCES

1 http://www.galleries.com/minerals/gemstone/pearl/pearl.htm

2 Wegst, U.G.K., Bai, H., Saiz, E., Tomsia, A.P., Ritchie, R.O. (2015) Bioinspired materials. *Nature Materials* 14, 23-36.

3 Studart, A.R. (2012) Towards high-performance bioinspired composites. *Adv. Materials* 24, 5024-5044.

4 Sun, J., Bhushan, B. (2012) Hierarchical structure and mechanical properties of nacre: A review. *RSC Adv*. 2, 7617-7632.

5 Zhang, G., Li, X. (2012) Uncovering aragonite nanoparticle self-assembly in nacre – A natural armor. *Crystal Growth and Design* 12, 4306-4310.

6 Li, X., Huang, Z. (2009) Unveiling the formation mechanism of pseudo-single-crystal aragonite platelets in nacre. *Phys. Rev. Lett.* 102, 075502-075506.

7 Zheng, G., Xu, J. (2013) From colloidal nanoparticles to a single crystal: New insights into the formation of nacre's aragonite tablets. *J. Struct. Biol.* 182, 36-43.

8 Xia, S., Wang, Z., Chen, H., Fu, W., Wang, J., Li, Z., Jiang, L. (2015) Nanoasperity: Structure origin of nacre-inspired nanocomposites. *ACS Nano* 9, 2167-2172.

9 Wang, X., Marimoto, K., Fuji, R., Liu, J., Li, L., Wang, P., Akaike, T., Want, Z. (2014) Pinctada fucata mantle gene 4 (PFMG4) from pearl oyster mantle enhances osteoblast differentiation. *Biosci. Biotech. Biochem.* 79, 558-565.

10 Liu, J., Yang, D., Liu, S., Li, S., Xu, G., Zheng, G., Xie, L., Zhang, R. (2015) Microarray: A global analysis of biomineralization-related gene expression profiles during larval development in the pearl oyster. *BMC Genomics* 16, 325-340.

11 Liu, X., Li, J., Xiang, L., Sun, J., Zheng, G., Zhang, G., Wang, H., Xie, L., Zhang, R. (2012) The role of matrix proteins in the control of nacreous layer deposition during pearl formation. *Proc. R. Soc. B* 279, 1000-1007.

12 Gardner, L.D., Mills, D., Wiegand, A., Leavesley, D., Elizur, A. (2011) Spatial analysis of biomineralization associated gene expression from the mantle organ of the pearl oyster, Pinctada maxima. *BMC Genomics* 12, 455-470.

13 Liu, H.L., Liu, S.F., Ge, Y.J., Liu, J., Wang, X.Y., Xie, L.P., Zhang, R., Wang, Z. (2007) Identification and characterization of a biomineralization related gene PFMG1 highly expressed in the mantle of Pinctada fucata. *Biochemistry* 46, 844-851.

14 Su, J., Zhu, F., Zhang, G., Wang, H., Xie, L., Zhang, R. (2016) Transformation of amorphous calcium carbonate nanoparticles into aragonite controlled by ACCBP. *Cryst. Eng. Comm.* 18, 2125-2134.

15 Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X.,m Qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P.W.H., Wang, X., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, J., Peng, C., Meng, J., Yang, L., Liu, J., Wen, B., Zhang, N., Huang, Z., Zhu, Q., Feng, Y., Mount, A., Hedgecock, D., Xu, Z., Liu, Y., Domazet-Loso, T., Du., Y., Sun, X., Zhang, S., Liu, B., Cheng, P., Jiang, X., Li, J., Fan, D., Wang, W., Fu, W., Wang, T., Wang, B., Zhang, J., Peng, Z., Li, Y., Li, N., Wang, J., Chen, M., He, Y., Tan, F., Song, X., Zheng, Q., Huang, R., Yang, H., Du, X., Chen, L., Yang, M., Gaffney, P.M., Wang, S., Luo, L., She, Z., Ming, Y., Huang, W., Zhang, S., Huang, B., Zhang, Y., Qu, T., Ni, P., Miao, G., Wang, W., Zhang, S., Haung, B., Zhang, Y., Qu, T., Ni, P., Miao, G., Wang, J., Wang, Q., Steinberg, C.E.W., Wang, H., Li, N., Qian, L., Zhang, G., Li, Y., Yang, H., Liu, X., Wang, J., Yin, Y., Wang, J., (2012) The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* 490, 49-54.

16 Fang, D., Xu, G., Hu, Y., Pan, C., Xie, L., Zhang, R. (2011) Identification of genes directly involved in shell formation and their functions in pearl oyster, Pinctada fucata. *PLOS One 6*, 1-13.

17 Xiang, L., Su, J., Zheng, G., Liang, J., Zhang, G., Wang, H., Xie, L., Zhang, R. (2013) Patterns of expression in the matrix proteins responsible for nucleation and growth of aragonite crystals in flat pearls of Pinctada fucata. *PLOS One* 8, e66564, 1-10.

18 Jackson, D.J., McDougall, C., Woodcroft, B., Moase, P., Rose, R.A., Kube, M., Reinhart, R., Rokhsar, D.S., Montagnani, C., Joube, C., Piquemal, D., Degnan, B.M. (2010) Parallel evolution of nacre building gene sets in mollusks. *Mol. Biol. Evol.* 27, 591-608.

19 Marie, B., Joubert, C., Tayale, A., Zanella-Cleon, I., Belliard, C., Piquemal, D., Cochennec-Laureau, N., Marin, F., Gueguen, Y., Montagnani, C. (2012) Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell. *Proc. Natl. Acad. Sci USA* 109, 20986-20991.

20 Xiang, L., Su, J., Zheng, G., Liang, J., Zhang, G., Wang, H., Xie, L., Zhang, R. (2013) Patterns of expression in the matrix proteins responsible for nucleation and growth of aragonite crystals in flat pearls of Pinctada fucata. *PLOS ONE* 6, e66564, 1-9.

21 Perovic, I., Mandal, T., and Evans, J.S. (2013) A pseudo EF-hand pearl protein self-assembles to form protein complexes that amplify mineralization. *Biochemistry* 52, 5696-5703.

22 Chang, E.P., Evans, J.S. (2015) Pif97, a von Willebrand and Peritrophin biomineralization protein, organizes mineral nanoparticles and creates intracrystalline nanochambers. *Biochemistry* 54, 5348-5355.

23 Chang, E.P., Russ, J.A., Verch, A., Kroeger, R., Estroff, L.A., Evans, J.S. (2014) Engineering of crystal surfaces and subsurfaces by framework biomineralization protein phases. *Cryst.Eng. Commun.* 16, 7406-7409.

24 Chang, E.P., Russ, J.A., Verch, A., Kroeger, R., Estroff, L.A., Evans, J.S. (2014) Engineering of crystal surfaces and subsurfaces by an intracrystalline biomineralization protein. *Biochemistry* 53, 4317-4319.

25 Perovic, I., Chang, E.P., Lui, M., Rao, A., Cölfen, H., Evans, J.S. (2014) A framework nacre protein, n16.3, self-assembles to form protein oligomers that participate in the post-nucleation spatial organization of mineral deposits. *Biochemistry* 53, 2739-2748.

26 Chang, E.P., Perovic, I., Rao, A., Cölfen, H., Evans, J.S. (2016) Insect cell glycosylation and its impact on the functionality of a recombinant intracrystalline nacre protein, AP24. *Biochemistry* 55, 1024-1035.

27 Perovic, I., Chang, E.P., Verch, A., Rao, A., Cölfen, H., Kroeger, R., Evans, J.S. (2014) An oligomeric C-RING nacre protein influences pre-nucleation events and organizes mineral nanoparticles. *Biochemistry* 53, 7259-7268.

28 Chang, E.P., Williamson G., Evans, J.S. (2015) Focused ion beam tomography reveals the presence of micro-, meso-, and macroporous intracrystalline regions introduced into calcite crystals by the gastropod nacre protein AP7. *Crystal Growth and Design* 15, 1577-1582

29 Chang, E.P., Roncal-Herrero, T., Morgan, T., Dunn, K.E., Rao, A., Kunitake, J.A.M.R., Lui, S., Bilton, M., Estroff, L.A., Kroeger, R., Johnson, S., Coelfen, H., Evans, J.S. (2016) Synergistic biomineralization phenomena created by a nacre protein model system. *Biochemistry* 55, 2401-2410.

30 Perovic, I., Davidyants, A., Evans, J.S. (2016) Aragonite-associated mollusk shell protein aggregates to form mesoscale "smart" hydrogels. *ACS Omega* 1, 886-896.

31 Wang, X., Harimoto, K., Fuji, R., Liu, J., Li, L., Wang, P., Akaike, T., Wang, Z. (2015) Pinctada fucata mantle gene 4 (PFMG4) from pearl oyster mantle enhances osteoblast differentiation. *Biosci. Biotech. Biochem*. 79, 558-565.

32 Liu, J., Yang, D., Liu, S., Li, S., Xu, G., Zheng, G., Xie, L., Zhang, R. (2015) Microarray: a global analysis of biomineralization-related gene expression profiles during larval development in the pearl oyster, Pinctada fucata. *BMC Genomic* 16, 325-340.

33 Gebauer, D., Volkel, A., Cölfen, H. (2008) Stable prenucleation of calcium carbonate clusters. *Science* 322, 1819-1822.

34 Gebauer, D., Cölfen, H. (2011) Prenucleation clusters and non-classical nucleation. *Nano Today* 6, 564-584.

35 Kellermeier, M., Cölfen, H., Gebauer, D. (2013) Investigating the early stages of mineral precipitation by potentiometric titration and analytical ultracentrifugation. in Research Methods in Biomineralization Science. Ed. By De Yoreo, J.J. Book Series: *Methods in Enzymology* 532, 45-69.

36 Picker, A., Kellermeier, M., Seto, J., Gebauer, D., Cölfen, H, (2012) The multiple effects of amino acids on the early stages of calcium carbonate crystallization. *Zeitschrift für Kristallographie* 227, 744-753.

37 Evans, J.S. (2012) Identification of intrinsically disordered and aggregation - promoting sequences within the aragonite-associated nacre proteome. *Bioinformatics* 28, 3182-3185.

38 Kozlowski, L.P., Bujnicki, J.M. (2012) MetaDisorder: A meta-server for the prediction of intrinsic disorder in proteins. *BMC Bioinformatics* 13, 111-122.

39 Dosztányi, Z., Csizmók, V., Tompa, P., Simon, I. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, *Bioinformatics* 21, 3433-3434.

40 Conchillo-Sole, O., de Groot, N.S., Aviles, F.X., Vendrell, J., Daura, X., Ventura, S. (2007) AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics* 8, 65-82.

41 Goldschmidt, L., Teng, P.K., Riek, R., Eisenberg, D. (2010) The amylome, all proteins capable of forming amyloid-like fibrils. *Proc. Natl. Acad. Sci USA* 107, 3487-3492.

42 Roche, D. B., Buenavista, M. T., Tetchner, S. J., McGuffin, L. J. (2011) The IntFOLD server: an integrated web resource for protein fold recognition, 3D model quality assessment, intrinsic disorder prediction, domain prediction and ligand binding site prediction. *Nucleic Acids Res*. 39, W171-176.

43 Richter, R., Mukhopadhyay, A., Brisson, A. (2003) Pathways of lipid vesicle deposition on solid surfaces: A QCM-D and AFM study. *Biophys. J.* 85, 3035-3047.

44 Satyshur, K.A., Rao, S.T., Pyzalska, D., Drendel, W., Greaser, M., Sundaralingam, M. (1988). Refined structure of chicken skeletal muscle troponin C in the two-calcium state at 2-Å resolution. *J Biol Chem* 263, 1628-1647.

45 Wernimont, A., Amani, M., Qiu, W., Pizarro, J.C., Zrtz, J.D., Lin, Y.H., Lew, J, Hutchinson, A., Hui, R. (2011) Structures of parasitic CDPK domains point to a common mechanism of activation. *Proteins: Structure, Function, and Bioinformatics* 79, 803-820.

46 Wang, C.H., Balasubramanian, M.K., Dokland, T. (2004) Structure, crystal packing, and molecular dynamics of the calponin-homology domain of Schizosaccharomyces pombe Rng2. *Acta Crystallographica Section* D60, 1396-1403.

47 Liu, J., Kurella, V.B., LeCour, L., Vanagunas, T., Wothylake, D.K. (2016) The IQGAP1 N-terminus forms dimers, and the dimer interface is required for binding F-actin and calcium-bound calmodulin. *Biochemistry* 55, 6433-6444.

For Table of Contents Use Only