Extremely charged coral protein AGARP regulates calcium carbonate growth through liquid phase separation

Barbara P. Klepka[1], Agnieszka Michaś[1], Tomasz Wojciechowski[2], Anna Niedźwiecka[1,*]

[1] *Laboratory of Biological Physics, Institute of Physics, Polish Academy of Sciences, Aleja Lotnikow 32/46, PL-02668 Warsaw, Poland*

[2] *International Research Centre MagTop, Institute of Physics, Polish Academy of Sciences, Aleja Lotnikow 32/46, PL-02668 Warsaw, Poland*

* Corresponding author: annan@ifpan.edu.pl

Abstract

Biomineralization *via* the non-classical crystallization pathway is postulated to involve a transient liquid phase of calcium carbonate formed in the presence of polymers. In the context of coral biocalcification, these polymers may include coral acid-rich proteins (CARPs), secreted into the skeletal organic matrix. However, direct evidence for the existence of this liquid phase with proteins is lacking. Here we report the properties of the intrinsically disordered aspartic and glutamic acid-rich protein (AGARP), which is the first CARP cloned from one of the most studied scleractinian coral of the Great Barrier Reef, *Acropora millepora*, a model species for biomineralization and ecological studies. We show that AGARP can significantly influence early stages of $CaCO_3$ nucleation and crystal growth through liquid-liquid phase separation. Consequently, we introduce the concept of a biologically relevant crystallization precursor, a liquid protein-calcium condensate composed of CARP molecules and Ca^{2+} ions, that forms as a result of liquid phase separation in a viscous, crowded environment. Our work bridges the gap between the liquid phase separation and biomineralization research.

Introduction

Biomineralization mechanisms described within the framework of non-classical crystallization (NCC) theory involving putative liquid precursors are attracting increasing attention. Unlike the classical crystallization pathway, NCC assumes that crystal formation occurs through intermediate states such as prenucleation clusters (PNCs), amorphous precursors, or the liquid phases, rather than through simple atom or ion attachment[1]. The sequential emergence of these states and/or their coexistence in equilibrium, along with their transformations dependent on thermodynamic conditions, is still a matter of debate. The putative pre-nucleation stage, PNCs, comprise 1-3 nm clusters of ions that are both thermodynamically stable and highly hydrated[2–4]. Another hypothesized transient state, amorphous calcium carbonate (ACC), has been found in living biomineralizing organism – coral *Stylophora pistillata*. These ACC entities were postulated to form within the tissue and subsequently attach to the skeleton's surface, where they eventually undergo crystallization to aragonite ($CaCO_3$ polymorph)[5]. The transient liquid phase is postulated to be formed in the presence of polymers and is referred to as polymer-induced liquid precursors (PILPs)[6]. The PILP droplets were observed in the presence of poly(acrylic acid) by means of atomic force microscopy[7]. There is, however, no other direct experimental evidence of PILP. Conclusions about the prior existence of a polymer-induced liquid phase of calcium carbonate are made based on observations of the structural properties of the mineral phases at various stages of crystallization, without being able to prove the liquidity (e.g. [8,9]). Consequently, it was also speculated that PILPs are, in fact, ACC clusters assemblies driven by polymers[10].

In the context of coral skeleton formation, these polymers could be appointed to, among others, lipids, polysaccharides and, in particular, coral acid-rich proteins (CARPs) secreted to the extracellular medium at the interface between the coral tissue and the aragonite skeleton[11]. By far, only four CARPs have been cloned and partially characterized[12]. It was demonstrated that the CARPs from *S. pistillata* bind Ca^{2+} stoichiometrically and spontaneously precipitate aragonite *in vitro* in seawater[12]. CARP3, on the other hand, was shown to roughen the surface of $CaCO_3$ crystals *in vitro* and tune the polymorph selection depending on the Mg^{2+}:Ca^{2+} ratio[13]. Immunolocalization of CARPs revealed that those proteins are localized in the area adjoining the early mineralization zones and that their localization in the tissue corresponds to the pattern of calcium distribution[14]. Additionally, CARP1 has been found to sequester and concentrate Ca^{2+} ions in a non-biomineralizing organism *Nematostella vectensis*[15]. CARPs

may also interact with other proteinaceous components of the skeletal organic matrix (SOM) as predicted by interactome analysis[16].

CARPs, featuring low-complexity sequences enriched in disorder-promoting charged amino acid residues, such as aspartic and glutamic acid, suggest that they may be classified as intrinsically disordered proteins (IDPs). IDPs, unlike folded proteins, do not adopt stable three-dimensional structure. Instead, they exhibit high conformational flexibility, making them extremely sensitive to environmental conditions[17]. Some IDPs can also undergo liquid-liquid phase separation[18–20] and form liquid biomolecular condensates. It was speculated that IDPs are important for biomineralization due to their flexibility and ability to bind to different surfaces[21], but details of the biophysical mechanism of biomineralization remain elusive.

Here we report the properties of a novel CARP, the aspartic and glutamic acid-rich protein (AGARP), that was identified in the stony coral *Acropora millepora*[22], an organism whose genome has been fully sequenced[23] and is therefore a model species for the study of coral development,[24] biocalcification,[25] and climate changes[26]. We provide experimental evidence that the highly polyanionic AGARP is an IDP that can strongly affect both the early stages of $CaCO_3$ nucleation and further crystal growth through a liquid phase. In particular, we point to the possibility of the existence of a crystallization precursor that is more biologically relevant, *i.e.* liquid protein-calcium condensate (LPCC) composed of CARP molecules and associated Ca^{2+} ions, which is formed as a result of liquid-liquid phase separation (LLPS) in a highly viscous and crowded environment, corresponding to the properties of the coral extracellular SOM.

Materials and Methods

Reagents and buffers

All reagents used in this study were of analytical, spectroscopic or molecular biology grade: bromophenol blue (BPB), $CaCl_2 \cdot 2H_2O$, CHAPS, glycerin, glycine, imidazole, NaOH, sodium dodecyl sulfate (SDS), thiourea, Triton X-100, urea, guanidine hydrochloride (GdmCl) from Carl Roth, dithiothreitol (DTT), NaH_2PO_4, β-mercaptoethanol from Sigma-Aldrich Merck, isopropyl β-d-1-thiogalactopyranoside (IPTG) from GeneDireX, Tris(hydroxymethyl)aminomethane (Tris) from Merck Millipore, polyethylene glycol (PEG) 4k from Fluka, HCl from Chempur, NaCl from Stanlab, ampicillin from IBI Scientific, 96% ethanol from Pol-Aura. Proteins: chicken egg white lysozyme, bovine α-chymotrypsinogen

A, bovine serum albumin (BSA) with monomer fraction ≥ 97%, human serum albumin (HSA), horse spleen apoferritin were purchased from Sigma-Aldrich Merck.

Buffer composition: A (50 mM Tris/HCl, 150 mM NaCl, pH 8.00); B (25 mM Tris, 192 mM glycine, 0.1% SDS); C (50 mM Tris/H_2SO_4, 150 mM NaCl, pH 8.00); D (10 mM Tris/HCl, 100 mM NaCl, 5% glycerol, pH 8.00); E (50 mM NaH_2PO_4/Na_2HPO_4, pH 8.00); F (50 mM Tris/HCl, 150 mM NaCl, 0.5 mM EDTA, pH 8.00); equilibration buffer (6 M urea, 30% (w/v) glycerol, 2% (w/v) SDS in 0.05 M Tris/HCl buffer pH 8.8); rehydration buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 40 mM DTT, 0.2% (w/v) Bio-Lyte 3/10 Ampholyte (BioRad), 1% bromophenol blue); Laemmli buffer (62.5 mM Tris/HCl pH 6,8, 1% (w/v) SDS, 10% (v/v) glycerol, 0,05% (w/v) bromophenol blue, 12.5% (v/v) β-mercaptoethanol). All buffers and AGARP samples used in experiments were treated with Chelex 100 resin (BioRad) to remove residual bivalent ions. pH of buffers with the addition of GdmCl or $CaCl_2$ was ajusted individually.

Protein Expression and Purification

His_6-tagged SUMO-AGARP sequence (**Fig. S1**) was cloned into pET-15b vector between NdeI and XhoI restriction sites (**Fig. S2**); the coding sequence was purchased from Gene Universal. Rosetta2(DE3)pLysS *E. coli* cells were transformed by heat shock with the plasmid pET-15b encoding the His_6-SUMO-AGARP sequence. 100 μL of bacterial culture at a concentration of 10^4 CFU/μg in LB Broth (BioMaxima) was incubated with 100-150 ng of the corresponding plasmid for 5 min on ice, 1 min at 42 °C and for 1 min on ice. LB Broth was added to the *E. coli* culture to the final volume of 1 mL and the cell suspension was shaken for 1 h at 37 °C. 100 μL of the transformed cells were plated on the solid LB medium supplemented with 100 μg/mL ampicillin and incubated overnight at 37 °C. A single colony was used to inoculate a pre-culture that was grown overnight in 10 mL LB Broth with 100 μg/mL ampicillin while shaking with the frequency of 250 rpm on an SI-600R shaker (Lab Companion). Subsequently, 500 mL of the LB medium with 100 μg/mL ampicillin was inoculated with 5 mL of pre-culture and the cells were grown with shaking (as described above) at 37 °C until the OD_{600} reached ~ 0.6. Then the expression of His_6-SUMO-AGARP was induced by IPTG at the final concentration of 0.5 mM. The protein was expressed in *E. coli* culture grown at 15 °C for 4 h (**Fig. S3a**). The bacterial suspension was centrifuged and the pellet was stored overnight at -80 °C.

The pellet was resuspended to the final volume of 20 mL in buffer A supplemented with supernuclease (Sino Biological) (250 U/1 g cell suspension), one tablet of protease inhibitors

cOmplete Tablets, Mini, EDTA-free (Roche), and 2 mM DTT. The cell membranes were lysed by sonication for 4 min on ice in cycles of 30 s on, 30 s off at 50% of the maximal power in an ultrasonic homogenizer VCX130 (Sonics) and by Triton X-100 added to the sonicated suspension at a v/v ratio of 1:10. Dissolved proteins were separated from fragmented membranes by centrifugation at 44,400×*g* for 30 min at 4 °C. The supernatant was incubated for 1 hour with HIS-Select® Nickel Affinity Gel (Sigma), 300 μL of gel per 500 mL of His_6-SUMO-AGARP expression cell culture. The nickel affinity gel with adsorbed protein was washed 13 times with 2 mL of buffer A. The protein were eluted with 2 mL of 5, 10, 20, 50, 100, 300, 500 mM imidazole in buffer A, twice for each imidazole concentration. The purest fractions were selected and put together, dialyzed against buffer A without imidazole and incubated with SUMO Protease (Ulp1p, M1_K402del) (Kerafast) at a 200:1 ratio for 1 h at RT without stirring to cleave the His_6-SUMO tag from AGARP. To separate AGARP from His_6-SUMO, the protein solution was incubated again for 1 h with nickel affinity gel (~1 ml of gel per 15 mg of protein) at 4 °C with stirring. The pure untagged AGARP was found in the filtrate (**Fig. S3b**).

Size Exclusion Chromatography (SEC)

SEC was performed on a Superdex 200 Increase 10/300 GL column (Sigma) at 10 °C with a flow rate of 0.5 mL/min, an injection volume of 100 μL or 500 μL, and the detection wavelengths of 215 nm and 280 nm. The column was equilibrated with the appropriate buffer before each experiment. For the determination of Stokes radii (R_H) of AGARP forms, the setup was calibrated using standard proteins in buffer E: lysozyme (19 Å[27]), α-chymotrypsinogen A (22.4 Å[28]), BSA monomer (35.5 Å[28]) and dimer (43 Å[28]), HSA monomer (35.1 Å[29]), and apoferritin monomer (60.3 Å[30]) and dimer (88.8 Å[30]). For the studies of interactions with GdmCl, AGARP was incubated on ice in buffer D with a given concentration of GdmCl for 30 min prior to measurements, and the SEC calibration was carried out in buffer D without GdmCl. The exponential function was fitted to calibration data points: $R_H = (Y_0 - \text{Plateau})\cdot e^{-A\cdot K_{AV}} + \text{Plateau}$ (eq. 1), where: Y_0, Plateau, A were fitting parameters, and K_{AV} is the partition coefficient from SEC.

Protein labelling

AGARP, His_6-SUMO-AGARP, lysozyme, α-chymotrypsinogen A, HSA and apoferritin were labelled with AF488 NHS ester (Lumiprobe) according to manufacturer's instructions in 100

mM phosphate buffer pH 8.45. The unbound dye was removed from proteins with the use of SEC or Zeba Spin Desalting Columns 7K MWCO (Thermo Scientific).

SDS–polyacrylamide gel electrophoresis (SDS-PAGE)

After expression and purification, samples were incubated with Laemmli buffer at 95 °C for 10 min. The SDS-PAGE was run in 12% polyacrylamide gel in buffer B at 120 V until the BPB left the gel. The gels were stained with Bio-Safe™ Coomassie Stain (BioRad). BlueEasy Prestained Protein Marker (Nippon Genetics Co.) and Pierce™ Unstained Protein MW Marker (Thermo Fisher Scientific) were used for SDS-PAGE analysis of CARP expression and purification, respectively.

Isoelectric focusing (IEF) SDS-PAGE

Proteins His_6-SUMO-AGARP, AGARP and BSA (50 μg each) were extensively dialyzed against deionized water for 3 days (6 water changes) until there was no salt left and the proteins precipitated. Immobilized pH gradient (IPG) (pH 3–10) strip (ReadyStrip, BioRad) was loaded with proteins dissolved in 125 μl of rehydration buffer and passively rehydrated overnight at 25 °C. 1 h after addition of the protein solution, the strips were covered with mineral oil (BioRad) to prevent evaporation. Isoelectric focusing was carried out in PROTEAN i12 IEF Cell (BioRad) at 20 °C at the maximum current of 50 μA/gel at 250 V for 15 min (rapid ramp), 4,000 V for 1 h (gradual ramp) and until 40,000 Vh (rapid ramp) was reached. The strips were later equilibrated for 10 min with 2.5 mL of equilibration buffer with 1% DTT (w/v) and afterwards with 4% (w/v) iodoacetamide. Before the SDS-PAGE separation, the strips were rinsed in buffer B and covered with an overlay solution composed of 0.5% (w/v) agarose in buffer B with BPB.

Absorption and Fluorescence Spectroscopy

The absorption spectra of AGARP at 24.6 μM in buffer A were measured on Varian Cary 50 UV-Vis (Agilent Technologies) with a scanning speed of 120 nm/min, integration time of 0.5 s, and at 1 nm interval, against a reference buffer.

Emission spectra of AGARP at 0.8 μM in buffer A were measured on Fluorolog 3.11 (Horiba Jobin Yvon) in a thermostatted (20 ± 0.2 °C) semi-micro quartz cuvette with a magnetic stirrer, with path lengths for absorption and emission of 4 and 10 mm, respectively (119004F-10-40, Hellma), from 300 to 500 nm every 1 nm, with the excitation wavelength of 280 nm, integration time of 0.5 s, excitation and emission slits of 1 nm and 10 nm, respectively, in the S/R mode. The buffer spectrum was subtracted from the protein spectrum.

Circular Dichroism (CD) Spectroscopy

MOS-450/AF-CD spectrometer was used to acquire CD spectra. The source of light was a Xe lamp. The CD signal was collected from 190 to 260 nm, step 1 nm, integration time of 10 s at each point, in 4 replicates to exclude drift in time, then the spectrum was averaged. The measurements were carried out in a thermostatted quartz cuvette with the optical path length of 0.1 mm. The CD signal of the reference buffer was subtracted from each averaged AGARP spectrum. CD spectra for AGARP at 7 μM in buffer D containing 0, 0.5, 3.0, 5.0 M GdmCl were measured at 10.0 ± 0.5 °C to compare with SEC. The measurements for AGARP at 8.13, 8.05, 8.03 μM in buffer C containing 0, 1 μM, 10 mM $CaCl_2$, respectively, were performed at 20.0 ± 0.5 °C to compare with FCS. The final spectra were analyzed with BeStSel[31,32].

Laser scanning confocal microscopy imaging

Imaging of fluorescent samples was performed using an inverted laser scanning confocal microscope Axio Observer LSM 780 Zeiss, using MBS 488 beam splitter and BP 495-555 nm filter, excitation at 488 nm with argon ion laser. LLPS was observed using Plan-Apochromat 63x/1.40 oil immersion lens. The phases of $CaCO_3$ were visualized using C-Apochromat 40x/1.20 water immersion, air Plan-Apochromat 20x/0.8, or Plan-Apochromat 63x/1.40 oil immersion lenses.

Fluorescence correlation spectroscopy

The FCS measurements were performed on the same confocal microscope equipped with C-Apochromat 40x/1.2W corr M27 water immersion lens and a ConfoCor3 kit, essentially as decribed previously[33]. MBS 488 beam splitter and a BP 495-555 filter were used. Before each experiment, the samples were filtered through 0.22 μm membrane, and the 25 μL droplets were thermostated to RT of 24 °C for 3 minutes. The actual temperature inside the droplet was checked after every experiment with UT325 thermometer (UNI-T) equipped with a thermocouple with a diameter of 0.5 mm. The samples were excited at 488 nm with argon ion laser at 3% relative power. The setup: argon laser power, pinhole and objective ring, was calibrated on freely diffusing AF488 in pure water[34] and the structural parameter was determined separately for each microscope glass slide that was pre-passivated with BSA. The triplet state relaxation time was measured in the independent experiment series individually for every protein at a higher relative laser power (15%). The triplet state lifetime and diffusion time of AF488 in buffer E for a given glass slide were fixed parameters in a two-component model of 3D diffusion that was fitted to experimental autocorrelation data to account for

photophysics and a possible residual fraction of unbound dye. To determine R_H, a single measurement lasted 3 s and was repeated 50 times in a series. The series were repeated for 3 or 4 independent droplets at 25.0 ± 0.5 °C. Spikes in the count rate indicating the occurrence of oligomers or aggregates were excluded from further analysis. R_H values were determined from the Stokes-Einstein equation taking into account the buffer viscosity and temperature in the droplet, as described previously[33]. To contrast the hydrodynamic properties of AGARP relative to the folded proteins, their R_H were shown as a linear function of cubic root of molecular mass (M): $R_H = a \cdot M^{1/3}$ (eq. 2), and a as power function of the number of amino acid residues (N) in the protein chain from polymer diffusion theory[35]: $R_H = R_0 \cdot N^{\nu}$ (eq. 2), where a, R_0, and the critical exponent ν, were fitting parameters.

Interactions of AGARP at 0.2 μM with Ca^{2+} ions were studied at 20 ± 0.5 °C in buffer F. Microscope glass slides were cleaned with 96% ethanol. Unpassivated microscope slides were used to avoid interactions of Ca^{2+} with BSA. Measurements before and after the addition of Ca^{2+} to the final concentration of 5 mM were run for 6 s, in 10 repetitions in a series. The series were repeated 7 times with a 30 s interval. The experiments were repeated in 3 independent droplets. Lysozyme was used as a reference.

Scanning electron microscopy (SEM) and energy dispersive X-ray spectroscopy (EDS)

The phases that deposited on the microscope slide after FCS measurements with calcium were then visualized using SEM and their elemental composition was determined using EDS. The samples were coated with a 5 nm layer of gold using a Q150T Sputter coater (Quorum Technologies) and inspected under Carl Zeiss Auriga CrossBeam SEM-FIB equipped with QUANTAX 400 Bruker EDS XFlash 7100. SEM images were collected at 5 kV accelerating voltage (AV) using an InLens SE or SESI detector for Secondary Electron. EDS spectra we acquired at 5 kV and 1 kV AV.

Liquid-liquid phase separation experiments

Solution of AGARP labelled with AF488 was added to stock buffer A solutions containing appropriate amounts of PEG and $CaCl_2$ and mixed by gentle pipetting in a test tube. The 4 μL mixture was placed on the microscope slide with chambered coverglass (3 mm diameter, 1 mm depth, Grace Bio-Labs) and covered with another slide to prevent evaporation. The droplets or phases on the surface of the microscope slide and ~5 μm above it were visualized after ~3 h of incubation at 25 °C.

Droplets and AGARP-containing phases statistics

The percentage of area (%A) occupied by the AGARP-containing phases on the surface of the microscope slide and the diameter of droplets ~5 μm above the surface were determined from images using ImageJ[36]. The Hill function: $\%A = 100 \cdot c^n/(k^n + c^n)$ (eq. 4) turned out to be statistically the best best for analyzing the *%A* as a function of protein concentration *c*, where *n* is the apparent cooperativity coefficient, and *k* is a characteristic Ca^{2+} concentration at 50% of the observed process.

Fluorescence recovery after photobleaching and coalescence

To spot the early stages of droplets growth, liquidity and coalescence first the appropriate stock solution containing AGARP labelled with AF488 and PEG 4k was placed on the silanised microscope slide and later $CaCl_2$ stock solution was added to the final concentrations of 50 μM AGARP, 100 mM $CaCl_2$ and 10% PEG 4k. Immediately after, coalescence events were recorded by capturing images with 500 ms time interval using the laser confocal microscope. After ~5 min, 3 droplets were bleached with 100% laser power. The fluorescence recovery was observed for 52 s. The results were analyzed with FRAPAnalyzer 2.1.0 taking into account the contribution from background and bleaching occurring during the measurement (reference). Single exponential model was fitted, and the data were rescaled.

In vitro calcium carbonate precipitation

The crystallization experiment was carried out for 18 h at 25 °C on a unpassivated microscope slide using a Press-to-Seal silicone isolator (9 mm diameter, 0.5 mm deep, Invitrogen) in a 25 μL droplet containing 100 mM $CaCl_2$ and without or with 1 μM AF488-labelled AGARP and without or with 20% PEG 4k in buffer A with either 0.4 mg (12.5 mg/cm^3) or 2 mg (62.5 mg/cm^3) of solid $(NH_4)_2CO_3$ as a source of CO_2. A $CaCO_3$ precipitation was also carried out in the atmospheric CO_2. In the latter case the droplet was covered with cover slide after 1 h of incubation. The $CaCO_3$ particles grown in the presence of AF488-labelled AGARP were visualized with confocal microscope

Fractal analysis

ImageJ[36] and FracLac plugin[37] were used. The confocal images were rescaled to identical pixel dimensions and resized accordingly to obtain the same percentage of area occupied by the $CaCO_3$ phases. The brightness was automatically adjusted, and the threshold was set identically for all images to ensure visibility of the $CaCO_3$ phase contours. The holes within the phases were filled and the phase contour was outlined. The line was subsequently smoothened, sharpened, and smoothened again to avoid image quality-related artifacts. The

fractal analysis was performed using 100 positions for grids, power series for calculating the series of box sizes and default number of box sizes and parameters for smallest and largest box size. Only fractal dimensions determined with $r^2 \geq 0.999$ were considered in further statistical analysis.

Statistical analysis

Statistical analyses, non-linear regressions, and error analysis were performed according to propagation rules[38].

Bioinformatics Analysis

Using AGARP (UniProtKB B7W112) as a query sequence, the BLASTp and tBLASTn algorithm was used to find homologous proteins. The theoretical masses and pIs of AGARP and His_6-SUMO-AGARP without signal peptide were calculated with ProtParam[39]. The disorder tendency of AGARP was predicted with DISOPRED3[40] and DisMeta[41] that provides consensus results of seven disorder predictors[42–48]. Fraction of charged residues (FCR) and net charge per residue (NCPR) were calculated with CIDER[49] for sliding window size of 10. Servers CatGranule[50], PScore[51], ParSe[52] and FuzDrop[53,54] were used to evaluate the liquid-liquid phase separation (LLPS) propensity of AGARP. Composition profile of AGARP was obtained using Composition Profiler[55]. The reference repository was a set of ordered structures from the Protein Data Bank (PDB Select 25 dataset[56]). The charge-hydropathy plot[57] was generated by PONDR. Das-Pappu phase diagram[49] for AGARP was obtained by CIDER[49]. Secondary structure of AGARP were predicted by AlphaFold2[58] using the ColabFold platform[59] and only those predictions that had a per-residue confidence based on the predicted local distance difference test (pLDDT)[60] exceeding 70 were analysed. All regressions and graphs were prepared in OriginPro 2023 (OriginLab Corporation).

Results

Bioinformatics analysis of AGARP

AGARP is a unique protein among the CARPs, since both the BLASTp and tBLASTn search revealed only one homologous proteins to *A. millepora* AGARP, i.e. AGARP from a closely related *A. digitifera* species[23] (**Fig. S2**). Surprisingly, AGARP (**Fig. 1a**) does not show particular sequence similarities with other CARPs[12]. However, both AGARP and CARPs 1-

3[12] contain short 3-residue long blocks containing acidic residues: Glu-Asp-Glu, Asp-Glu-Glu, Asp-Glu-Asp that could presumably serve as Ca^{2+}- binding motifs.

According to the analysis by DisMeta[41], AGARP can be fully disordered (**Fig. 1b**) or can exhibit short-range order, as predicted by Disopred3[40] (**Fig. 1c**). AlphaFold2[58] predictions with pLDDT scores ≥ 70 indicating well modelled fragments suggest that AGARP may contain residual helical structures at the N- and C-termini (**Fig. 1d**). Those locally ordered fragments correlate with the depleted fraction of charged residues (**Fig. 1e**). Nonetheless, these results should be validated experimentally, since the choice of the disorder prediction server can influence the resulting disorder probabilities quite significantly[61].

Contrary to the majority of IDPs, aspartic and glutamic acid are highly overrepresented, and glutamine is underrepresented in the AGARP sequence (**Fig. 1l**). On the charge-hydropathy plot[57] AGARP falls into the subset of disordered proteins (**Fig. 1m**).

AGARP is an extremely polyanionic protein with a net charge of –148 *e* per molecule and the linear charge density along the chain of –0.77 e/nm at pH 8.0, with FCR exceeding 0.5 for majority of the polypeptide chain as shown in **Fig. 1e**. Das-Pappu phase diagram[49] analysis (**Fig. 1n**) point to AGARP being a strong polyampholyte. However, as depicted in **Fig. 1f**, positive charges are consolidated in small patches within the AGARP sequence that is overall dominated by negative charge. This may indicate that AGARP, could in fact, exhibit properties of a strong polyelectrolyte.

Liquid-liquid phase separation propensity of AGARP

Four different servers used to evaluate the AGARP propensity to form droplets yielded puzzling results. Two short droplet-promoting regions, from 70 to 120 and from 390 to 440, have been identified in AGARP by catGranule[50] (**Fig. 1g**), a server which takes into account the protein RNA-binding features, disorder propensities, sequence composition (arginine-glycine and phenylalanine-glycine content) and polypeptide chain length. However, the average propensity score for the whole protein is relatively low and means that AGARP propensity is only by 1.14 standard deviation larger than the propensity to LLPS predicted for the yeast proteome. The AGARP overall LLPS propensity predicted with PScore[51], a server which algorithm is based on the frequency of π-π interactions, was 2.24 (**Fig. 1h**). This value does not exceed the confidence threshold for this predictor that is 4 standard deviations from the average PScore calculated for structures deposited in the PDB database[51], to be appointed as prone to LLPS. ParSe[52] identified a short, six-residue long fragment at the AGARP N-

terminus, overlapping with the droplet-promoting region predicted with catGranule[50] and a three-residue long fragment at the C-terminus, as prone to LLPS (**Fig. 1i**). ParSe[52] is a predictor that takes into account the β-turn tendency of the sequence and the strength of protein-solvent interactions resulting from the hydrodynamic size of the monomeric protein. The resulting scoring parameter r_{model}, indicating the ratio of β-turn propensity to polymer scaling exponent describing the protein size, was 1.73 for the whole AGARP sequence, while the r_{model} for phase separating proteins is expected to be larger than 2.1 ± 0.2 and therefore AGARP cannot be classified as droplet-promoting. The only server that predicts high probability of spontaneous LLPS of AGARP (p_{LLPS} = 0.9945) is FuzDrop[53,54] (**Fig. 1j**). FuzDrop[53,54] not only estimates the droplet-promoting tendency of the protein on the basis of the conformational entropy (the degree of disorder) of the chain in the free states but also on the basis of the binding entropy[53]. The calculated high LLPS tendency of AGARP may, therefore, arise from the large number of degrees of freedom of the chain, facilitating potential interactions with other binding partners. In case of CARPs such droplet-driving homotypic and heterotypic interactions may possibly occur with proteins[22,62], lipids[63], polysaccharides[64] or counterions, like calcium or magnesium[65], also present within the coral skeletal organic matrix (SOM)[66,67]. AGARP propensity to phase separate should be analyzed in the context of interactions with other components of the SOM, in particular that counterion binding can mediate multivalent homotypic interactions of CARPs.

Properties of AGARP

The low isoelectric points of AGARP (and its His_6-SUMO construct) were experimentally determined by 2D electrophoresis as 3.7 ± 0.2 and 3.9 ± 0.2, respectively (**Fig. 2a**), close to the calculated pI values of 3.94 and 4.09, respectively. The protein UV/VIS absorption spectrum with the maximum at 276 nm is mainly dominated by 10 side chains of Tyr and one Trp residue, with a visible contribution from 24 Phe residues at ~260 nm (**Fig. 2b**). The AGARP fluorescence spectrum displays the emission maximum at ~352 nm, which indicates that the Trp side chain is highly solvent-exposed.

The intrinsic disorder of both AGARP forms manifests itself in their hydrodynamic properties determined by SEC (**Fig. 2c, d**) and FCS (**Fig. 2e**). The Stokes radii of AGARP and His_6-SUMO-AGARP determined at 25 °C by FCS are 77 ± 5 and 81 ± 5 Å, and are higher than those obtained at 10 °C using SEC, 66 ± 4 and 74 ± 4 Å, respectively (**Fig. 2f**), pointing to the increase in R_H with temperature by about 9-17%. The AGARP forms are well beyond the

99 % CI of the Stokes-Einstein linear dependence of the R_H on the $M^{1/3}$ (R_H = (0.788 ± 0.014)·$M^{1/3}$) for folded proteins approximated as diffusing hypothetical spheres of homogenous density (**Fig. 2f**). In terms of universal scaling laws, the AGARP forms follow rather the power law for denatured proteins[68], since their hydrodynamic sizes are far from the power function for folded proteins determined under the conditions of this study (R_H = (4.3 ± 0.2) $N^{(0.312 \pm 0.007)}$) (**Fig. 2f, inset**). This indicates that AGARP behaves like ensemble of expanded conformations close to random coil.

To gain an insight into polyampholytic *vs.* polyelectrolytic nature of AGARP, its R_H was measured in the presence of increasing concentrations of ionic denaturant, GdmCl. As shown in **Fig. 2g**, at lower GdmCl concentrations, the R_H of AGARP decreases by ~25%, reaching a minimum at 0.5 M, while at higher GdmCl concentrations a gradual increase in R_H is observed. Similar phenomenon was also observed for other highly charged IDPs[69–71]. At low ionic strengths, the repulsion of negatively charged protein fragments is screened by the Gmd^+ counterion, allowing the protein chain to relax, while the subsequent expansion of the protein chain at higher ionic strengths can be attributed to the further nonspecific binding of the counterion[69].

Interestingly, those R_H changes are not accompanied by secondary structure rearrangements as revealed by CD (**Fig. 2h**). A negligible fraction of α-helices and no long-range secondary structures in the form of parallel β-sheets were identified in AGARP both in the absence and presence of GdmCl. Other forms of local order, like antiparallel β-sheets and β-turns that are indistinguishable from random coil conformations by CD[72] cover basically 100% of the spectrum. The CD results show that AGARP displays features of an extreme IDP even at 0.5 M GdmCl corresponding to the minimum of the R_H value. This implies that the compaction of the AGARP chain under ionic strength in the 0.5 M range, comparable to extracellular conditions in invertebrates[73], is an electrostatic collapse with no folding or disorder-to-order transition.

Interactions of AGARP with Ca^{2+}

The most biologically relevant counterion for CARPs is Ca^{2+}, that is a building block of coral $CaCO_3$ skeleton. For this reason, we investigated the interactions of AGARP with Ca^{2+} using FCS and SEM. The R_H of AGARP at 0.2 μM in Ca^{2+}-free conditions does not change within 95% CI during a ten-minute FCS measurement (**Fig. S8**). The addition of calcium ions to the protein solution shifts the individual autocorrelation curves of AGARP towards longer

diffusion times (*e.g.* after 3 and 12 min of incubation with Ca^{2+} in **Fig. 3a**) as a result of the rapid aggregation that can occur in real time during diffusion of AGARP molecules through the confocal volume in individual measurements. As show in **Fig. 3b**, this shift corresponds to formation of AGARP-containing aggregates with estimates of apparent hydrodynamic radius from ~100 to 10^4 Å. When lysozyme is used as a control, no larger protein-containing particles are visible in the FCS experiment.

After a 14-minute FCS measurement under ambient CO_2 conditions, AGARP triggers the formation of larger and clustered amorphous calcium carbonate (ACC) particles (**Fig. 3c**), which is contrast to $CaCO_3$ crystals of 200-300 nm that are deposited on the microscope glass slide both in the absence and presence of lysozyme, as revealed by the SEM. The crystals grown with lysozyme are slightly smaller. The surface of one of the largest AGARP-containing ACC aggregates (**Fig. 3e**) exhibits noticeable roughness originating from ~30 nm grains. The presence of AGARP in these structures was confirmed by EDS, as the peak corresponding to the Kα line of protein-derived nitrogen was detected at both 5 kV (**Fig. 3f, S11a**) and 10 kV accelerating voltage (**Fig. 3g, S11b**). Carbon Kα line originating from both AGARP and $CaCO_3$ was also detected at both AV. Strikingly, the calcium Kα line was only visible after excitation of the sample with the higher energy electron beam, which may indicate that AGARP is localized in the outermost layer of the ACC particles, while Ca^{2+} is located in deeper layers.

To gain insight into the AGARP conformation after interaction with Ca^{2+} upon $CaCO_3$ particles formation, CD was performed. We found that AGARP does not undergo significant chain rearrangements upon addition of 1 μM and 10 mM $CaCl_2$ but even complexed with Ca^{2+} remains disordered (**Fig. 3d**).

Ca^{2+}-driven liquid phase separation of AGARP

To investigate the interactions of AGARP with Ca^{2+} under conditions that more closely mimic the presence of other biopolymers in the extracellular skeletal organic matrix of corals, PEG was used as a viscous crowding agent. Contrary to former rapid precipitation of AGARP-containing ACC aggregates observed by FCS and SEM, under crowded and viscous conditions, AGARP undergoes binodal decomposition triggered by interactions with Ca^{2+} ions (**Fig. 4, 5, 6**). The liquid nature of the droplets has been demonstrated by observing multiple coalescence events (**Fig. 4a**) and the FRAP analysis (**Fig. 4b**) indicating 100% mobile fraction and the fluorescence recovery half-time of 11 ± 2 s (**Fig. 4c**). The LLPS

process was analyzed 5 μm above and at the surface of the microscope slide (**Fig. 5**), since once the droplets appear in the bulk by the binodal process, they coalesce rapidly and, when they become large enough, they sink by gravity and accumulate on the surface of the slide. This leads to the establishment of an equilibrium of three liquid phases: one dilute phase and two dense AGARP phases, the first being the dispersed fine droplet phase in the volume part and the second being the dominant phase on the slide surface. In the volume part, droplets (**Fig. 5a, 5b**) and individual aggregates (**Fig. 5b**) were observed. As shown in **Fig 5b,** the determinants for the formation of AGARP droplets are both the presence of the additional polymer, like PEG 4k, and the presence of calcium ions. Even at highest concentrations of AGARP (50 μM) and Ca^{2+} (100 mM) no droplets appear unless PEG is in the solution, and vice versa, even at the highest AGARRP and PEG 4k (20%) concentration no droplets emerged without Ca^{2+}. Surprisingly, the droplet diameter distribution depends the most strongly on the AGARP concentration (**Fig. 5c**), but are independent of the PEG and Ca^{2+} concentration. Conversely, the droplet population is strongly Ca^{2+}-, PEG-, and AGARP-dependent. For example, at 100 mM $CaCl_2$ and 50 μM AGARP the mean diameter of droplets are 2 ± 1, 1.9 ± 0.9 and 3 ± 1 μm for 7, 10 and 20% PEG. Similarly, at constant AGARP and PEG, the droplet sizes remain constant at increasing Ca^{2+} concentrations, *e.g.* at 5 μM AGARP and 20% PEG the diameters are 1.4 ± 0.5, 1.2 ± 0.5 and 1.4 ± 0.5 μm for 10, 50 and 100 mM Ca^{2+}. The only deciding factor that tunes the diameter of droplets and therefore the LLPS efficiency is the concentration of AGARP. For instance, at 20% PEG and 10 mM $CaCl_2$ the diameters increase gradually from 0.8 ± 0.3 μm at 1 μM AGARP to 3 ± 1 μm at 50 μM.

On the microscope slide, however, the AGARP-rich phase occupies more space as compared to the volume phase, and only at low AGARP concentrations (1 and 2 μM) separated droplets can be observed (as shown for 20% PEG 4k in **Fig. 5d**). At higher concentrations of AGARP (5-50 μM) in 20% PEG 4k, the pattern observed for the protein-rich phase is the result of multiple coalescence and droplet descent events occurring in the volume phase during the 2-3 hour equilibration period, although the pattern may seemingly resemble a spinodal decomposition. As shown in phase diagram in **Fig. 5e**, under conditions at which the dense liquid phase is not formed, single aggregates appear in the solution or both liquid and aggregated states once there is PEG. In PEG-free conditions, however, at highest AGARP concentration (50 μM) with either no or 100 mM $CaCl_2$ none of those phases are present. This observations together with previous analysis of LLSP phase diagram constructed based on images taken 5 um above the microscope slide indicate that the molecular crowing by a

polymeric molecule is the prerequisite for the liquid-solid or liquid-liquid phase separation to occur.

Next, we quantitatively investigated how the percentage of area (*%A*) of AGARP-rich droplets change with respect to protein, Ca^{2+} and PEG concentration (**Fig. 5f**, **S12**). Our main goal was to analyze the liquid phase formed at 20% PEG 4k, since under these conditions for each protein and Ca^{2+} concentration such phases were observable. For both 7% and 10% PEG 4k under certain AGARP/Ca^{2+} conditions both aggregates and liquid states are visible, therefore a reliable calculation of *%A* could be biased. As shown in **Fig. 5f**, at 20% PEG 4k with increasing AGARP concentration, under all tested $CaCl_2$ conditions, the *%A* of the liquid phase increases sigmoidally according to Hill's equation, indicating the cooperative nature of the process, which is consistent with the fact that the analyzed pattern results from coalescence and accumulation. Again, the presence of CARP turns out to be the main quantitative determinant of the droplets formation, with a characteristic AGARP concentration of 10.2 ± 0.3, 9.8 ± 0.3, and 7.5 ± 0.3 μM for 10, 50, and 100 mM Ca^{2+}, respectively. The LLPS process appears to be slightly affected by changes in calcium ion concentration in this range.

AGARP-controlled nucleation and growth of calcium carbonate

Having established that AGARP can stimulate the formation of either ACC clusters in simple water solutions or droplets under crowded viscous conditions upon interactions with Ca^{2+}, we investigated how these pre-crystallization states and the presence of polyanionic protein can tune the morphology of the emerging $CaCO_3$ phases. The starting point of crystallization is AGARP-Ca^{2+} complex in solution (**Fig. 6a**) or in liquid droplets (**Fig. 6h**) in PEG free or 20% PEG 4k conditions, respectively. At all CO_2 concentrations, AF488-labeled AGARP is always incorporated into $CaCO_3$ to a high extent, forming fluorescent structures (**Fig. 6b-g, j-n**). The $CaCO_3$ phases grown with AGARP in atmospheric CO_2 and in the vicinity of 12.5 mg/cm^3 $(NH_4)_2CO_3$ as source of CO_2 in no and 20% PEG 4k (**Fig. 6e, f, l, m**) have smoother and more rounded edges compared to no protein conditions (**Fig. 6r, w**). The effect of AGARP on changing the morphology of $CaCO_3$ is most pronounced once the $Ca^{2+} + CO_3^{2-} \rightarrow CaCO_3$ reaction rate is slowed down by decreasing the CO_2 concentration in the chamber to the natural environmental condition, since at high CO_2 concentration no significant difference in AGARP-containing (**Fig. 6g, n**) and AGARP-free (**Fig. 6s, x, z**) $CaCO_3$ morphology is visible. Fractal analysis (**Fig. 6aa**) of the most representative phases shown in **Fig. 6e-g, l-n**

revealed that AGARP alters the fractal dimensions of $CaCO_3$ phases to higher values as CO_2 levels decrease, indicating increased edge development per unit area of the crystal.

With increasing amount of CO_2 in the atmosphere of the crystallization chamber, with and without AGARP, the $CaCO_3$ phases become smaller but more abundant (**Fig. 6b-d**, **i-k**, **o-q**, **t-v**). This is a common phenomenon arising from the fact that at higher saturation the crystal growth is mostly dominated by crystal nucleation. The significant difference in the number of crystals is also observed in the presence of AGARP with (**Fig. 6i-k**) and without PEG (**Fig. 6b-d**). The number of crystals is higher when PEG is present in the crystallization solution. Additionally, in 20% PEG, the number of $CaCO_3$ phases is higher when AGARP is also present (**Fig. 6i-k, t-v**). This means that AGARP acts as a nucleator.

AGARP distribution in $CaCO_3$ phases and crystals

The specific AGARP distribution in $CaCO_3$ phases and crystals depends strictly on the reaction rate of $CaCO_3$ formation. Under lower reaction rates, AGARP accumulates selectively at the center and the edges of the $CaCO_3$ phases (**Fig. 7a, b, S14a**). It follows that AGARP is responsible for nucleation and controls the edge development of the growing crystal. Additionally, fibers with a cross-section diameter of 0.33 ± 0.08 μm are observed several micrometers above the microscope slide (**Fig. 7a, b, S15**). The arrangement of these fibers, extending from the crystal center towards the edge, likely reflects the protein mode of action. AGARP may first initialize the nucleation and subsequently spread outwards modifying the crystal edge and growth.

However, at a higher reaction rate of $CaCO_3$ formation, AGARP is more evenly distributed within the phases (**Fig. 7c, d, S14b**), further confirming that the controlling role of proteins in determining morphology is limited under these conditions. Nonetheless, the slight increase in fluorescence intensity at the center of the $CaCO_3$ phase in profile 1 (**Fig. 7d**) suggests that AGARP may still act as a nucleator even when the crystallization occurs more rapidly.

Discussion

CARPs are a group of proteins secreted to the coral extracellular matrix adjacent to the skeleton where they are postulated to be involved in the process of biomineralization[12,13,22]. To date, only four CARPs have been cloned and partially characterized[12,13]. In this work, we

have obtained and characterized AGARP, the fifth CARP, being the first from *A. millepora*, by studying its physicochemical properties and elucidating its potential involvement in regulation of the early stages of nucleation and control of the growth of $CaCO_3$ mineral phases.

Intrinsically disordered AGARP lacks a stable secondary and tertiary structure, a hydrophobic core, and has the hydrodynamic radius twice greater than that of a folded protein of the same molar mass. We showed that AGARP undergoes compaction at low GdmCl concentrations and further swelling at high GdmCl concentrations, a phenomenon that was also observed for other highly charged IDPs[69–71]. The high sensitivity of polyampholytes to the presence of salt[74] is manifested in AGARP interactions with Ca^{2+}, a building block of the $CaCO_3$ coral skeleton. We show that as a result of the electrostatic interactions with Ca^{2+} ions, AGARP induces the emergence of highly clustered, amorphous $CaCO_3$ particles, that could be appointed to ACC[75], with AGARP accumulated predominantly at the outer surface. The structure of the AGARP-induced ACC clusters resembles these reported recently as formed in the presence of negatively charged dsDNA after 2 h incubation with slow stirring, that may affect the final structure[10]. However, the ACC with AGARP are grown in minutes without stirring, which suggests that the effect of AGARP is stronger than dsDNA in mineral phase modulation. This results suggest that AGARP may be involved in the early stages of $CaCO_3$ precipitation in the form of ACC, which is further supported by the findings that SOMs, including proteins, accumulate in the centers of coral calcification[76] and that ACC was found to be one of the first precipitates formed in the pre-settled coral larva, as was shown for *S. pistillata*[77]. The protein-driven stabilization of ACC was previously observed only for the mixture of proteins of the coral skeleton organic matrix[78]. The AGARP-induced formation of ACC appears to be a determinant of the crystallization rate, as it has been shown that the crystal growth occurring *via* ACC attachment is more than 100 times faster than via the classical crystallization pathway[5].

In vivo, however, the crystallization processes in corals occurs in highly dense extracellular matrix packed with proteins[16,22,79], polysaccharides[80] and lipids[81]. Under such crowded conditions, introduced by PEG, AGARP undergoes LLPS in the presence of Ca^{2+} ions. This LLPS process could be described in terms of the percolation theory[82–84] where the AGARP molecules can form biomolecular condensates as a result of homotypic multivalent electrostatic interactions between acidic amino acid residues bound to Ca^{2+} counterions,

acting as stickers. In the studied system, below the percolation threshold, protein aggregates are formed, while above the threshold, an AGARP-rich liquid phase is formed.

The critical effect of AGARP concentration on binodal liquid phase decomposition in a viscous, crowded milieu, as in SOM, indicates the importance of the high multivalency of CARPs in regulating calcium carbonate nucleation and growth through homotypic and heterotypic Ca^{2+}-mediated interactions among CARP molecules. This may provide a clue to explaining how different coral species fine-tune their skeletal structure through different proportions of various CARPs, even though their sequences are not specific in terms of structure and binding sites in the traditional sense and cannot be treated in terms of homology and similarity analysis in the same way as simple folded proteins.

The involvement of liquid crystallization precursor, termed PILP[6], has been previously described in the literature. The nature of PILP is, however, qualitatively different from the observed liquid states in this work.

It has been postulated that PILP is the liquid form of $CaCO_3$ formed in the presence of a polymer[6], and the idea focuses on the interaction of Ca^{2+} with CO_3^{2-}, as if their encounter complex formed before polymer engagement. Here, the droplets are composed of AGARP molecules clustered together by Ca^{2+} ions, which is more biologically relevant, since the secretory CARPs meet Ca^{2+} ions during their biogenesis in ER, before reaching the extracellular environment, and thus the CARP-Ca^{2+} complex formation should be regarded as the first step in CARP-mediated biomineralization. Therefore, we introduce the term liquid protein-calcium condensate (LPCC) to describe the biomolecular condensate we observed as a biomineralization precursor.

The role of LPCC as a precursor of crystallization was confirmed by the crystallization experiments. We showed that when LPCCs formed prior to the introduction of CO_2 and the crystallization preceded slowly enough, AGARP tuned the morphology of the mineral phase. AGARP was found to smoothen and increase the development of the edges of $CaCO_3$ phases. A similar, but less pronounced effect of AGARP in determining the morphology of the mineral phase was observed when the crystallization precursors were ACC particles.

The subject concerning the role of different types of proteins in $CaCO_3$ crystal morphology determination has been previously investigated (eg. [12,13,85,86]) and dates back to 1872[87]. However, the role of CARPs was only explored in terms of polymorph selection strongly dependent on the concentration of Mg^{2+} ions[13] and its ability to spontaneously precipitate

aragonite[12]. In this work, we propose a comprehensive model for the non-classical crystallization pathway mediated by CARPs (**Fig. 8**). Our AGARP protein, similarly as other *A. millepora* CARPs[22], has a 16-amino acid long N-terminal signal peptide (**Fig. S16**) that directs the protein onto the secretory pathway[88] by targeting the nascent protein to the Sec translocon[89] located in endoplasmic reticulum membrane. The Sec translocon transports the translated protein to the lumen of ER, a compartment rich in Ca^{2+} ions[90], with Ca^{2+} concentrations ranging from 100 μM to 1 mM[91]. We postulate that at this stage interaction between CARPs and Ca^{2+} occurs and a complex is formed (1). From this stage forward, CARPs are found in a highly viscous environment. The ER lumen, has the viscosity nine times higher than water and three times higher than cytoplasm[92], estimated at approximately 6 cp at 37 °C. The lumen of Golgi apparatus and secretory vesicles through which CARPs are transported to the plasma membrane[93], are packed with proteins and other biomolecules[94] that increase the viscosity and osmolality of the inner solution. And finally, the presence of various biopolymers incorporated into the skeleton[22,79] may indicate that the extracellular matrix at the tissue-skeleton interface is a highly viscous environment, although the exact composition and viscosity is unknown. In this work, we show that the crowding leading to increased solution osmolality and viscosity raging from roughly ~ 5 cp (7% PEG 4k) to 11 cp (20% PEG 4k)[95] triggers the emergence of LPCC (2) which further coalesce forming larger scale LPCC (3). When the LPCCs are exposed to the environment rich in HCO_3^- ions produced from CO_2 by carbonic anhydrases, predominantly secreted to the EM adjacent to the skeleton[22,96], the crystallization process occurs (4). The formed $CaCO_3$ crystals have highly developed and smooth edges which match the characteristics of the native morphology of the skeleton[25,97]. This nucleation pathway involving CARPs is an extension of the proposed mineral formation model involving polystyrene sulphonate[98].

If the crystallization and precursor formation occurred in a low-viscosity environment close to that of water (~ 0.9 cp) in the absence of a polymeric crowder, the subsequent steps after the formation of the CARP-Ca^{2+} complex would be: (2') aggregation of CARPs, driven by electrostatic interactions, (3') formation of ACC and (4') crystallization, leading to the formation of crystals with residuals sharp edges. This pathway, however, would lead to the $CaCO_3$ morphologies that could potentially injure the soft tissues.

Our results deepen the understanding of the involvement of ACC-mediated crystallization pathway in the skeleton formation[5]. The formation of $CaCO_3$-based coral skeleton is an intricate process depending not only on CARPs but also other biopolymers and environmental

conditions. It is highly probable that various crystallization pathways cross, one of which being mediated by a CARP-containing liquid precursor.

All CARPs share a common feature, *i.e.* the negatively charged amino acid residues are highly overrepresented in their sequence[22,79], making them extremely acidic. However, the comparison of the results obtained for AGARP (**Fig. 6, 7**) with other proteins involved in biomineralization[86] shows that the distribution of the proteins within the $CaCO_3$ crystal, and thus their impact on nucleation and control of crystal growth, can be quite different. Our results may therefore suggest that the importance of the diversity of CARPs, which differ in the content and distribution of aromatic and hydrophobic residues, is not limited to providing a high net negative charge, but that the composition of CARPs in a given species may be one of the determinants of subtle features of skeletal morphology.

Acknowledgments

The work was supported by Polish National Science Centre grant no. 2016/22/E/NZ1/00656 to AN. The studies were performed in the NanoFun laboratories co-financed by ERDF within the Innovation Economy Operational Programme POIG.02.02.00-00-025/09. We thank Prof. Joanna Trylska for the access to the laboratory of circular dichroism at the Centre of New Technologies, University of Warsaw and Prof. Jarosław Stolarski for inspiring us with the topic of CARPs and encouraging us to undertake this research.

Data availability

All data necessary to interpret, verify and extend the research are included within the manuscript and Supplementary Information. Raw data available on request from the authors.

Contributions

A.N. conceived and designed the study, supervised and coordinated the project. A.N. and B.P.K. designed the experiments. A.M and B.P.K. cloned and expressed the proteins. T.W. and B.P.K acquired the SEM images and EDS spectra. B.P.K and A.N performed the experiments, analyzed and interpreted the data and wrote the manuscript with contributions from all authors. All authors revised the manuscript.

Ethics declarations

The authors declare no competing interests.

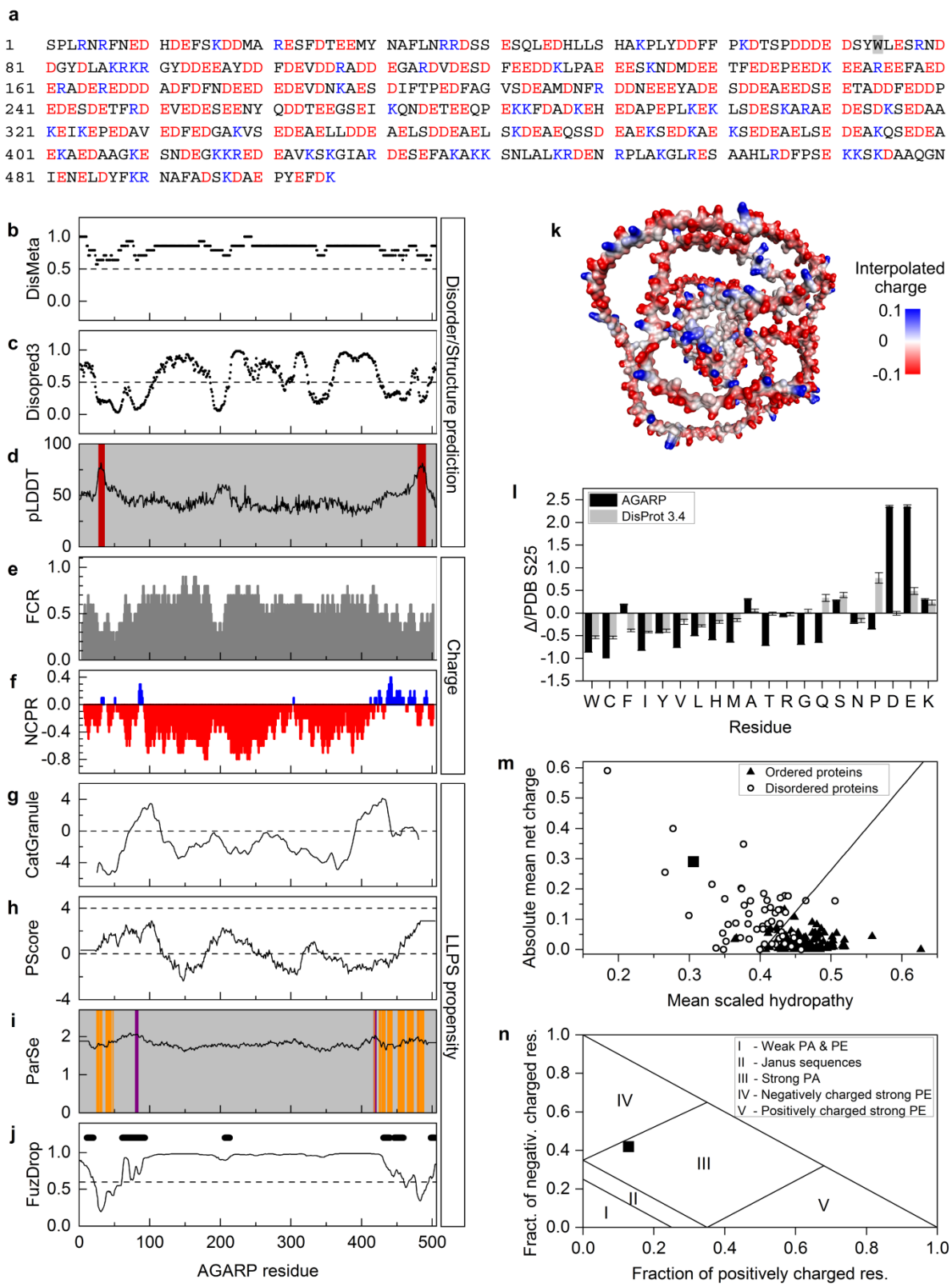

Fig. 1 Predicted properties of AGARP. a AGARP sequence without signal peptide with colored acidic (red) and basic (blue) residues and tryptophan highlighted in gray (Uniprot: B7W112). Bioinformatics analysis of AGAR; **b, c** disorder tendency[40–48]; **d** secondary

structure prediction[58,59]; **e**, **f** charge distribution[49]; **g** – **j** LLPS propensity[50–54]; **k** putative predicted conformation shown in the solvent-accessible surface representation and colored according to interpolated charge [58,59]; **l** composition profile[55]; **m** charge-hydropathy plot[57] and **n** Das-Pappu phase diagram[49]; Dashed lines denote the thresholds of respective methods of prediction. Values exceeding the threshold indicate that a region is (**b**, **c**) disordered or (**g**, **h**, **j**) prone to LLPS. α-helices predicted with pLDDT ≥ 70 in **d** are marked in dark red. Positive NCPR in **f** are marked with dark blue and negative – with dark red. In panel **i** – grey are regions intrinsically disordered but not prone to LLPS, orange are regions that can fold to a stable conformation, and violet are intrinsically disordered regions prone to undergo LLPS. Thick black lines in **j** mark aggregation hot-spots. The amino acids in **l** were ordered relative to the flexibility of the residue. The enrichment (positive values on y-axis) or the depletion (negative values on y-axis) of a certain residue was calculated with respect to the ordered structures from the Protein Data Bank (PDB Select 25 dataset[56]). In **n** PA stands for polyampholytes and PE for polyelectrolytes.

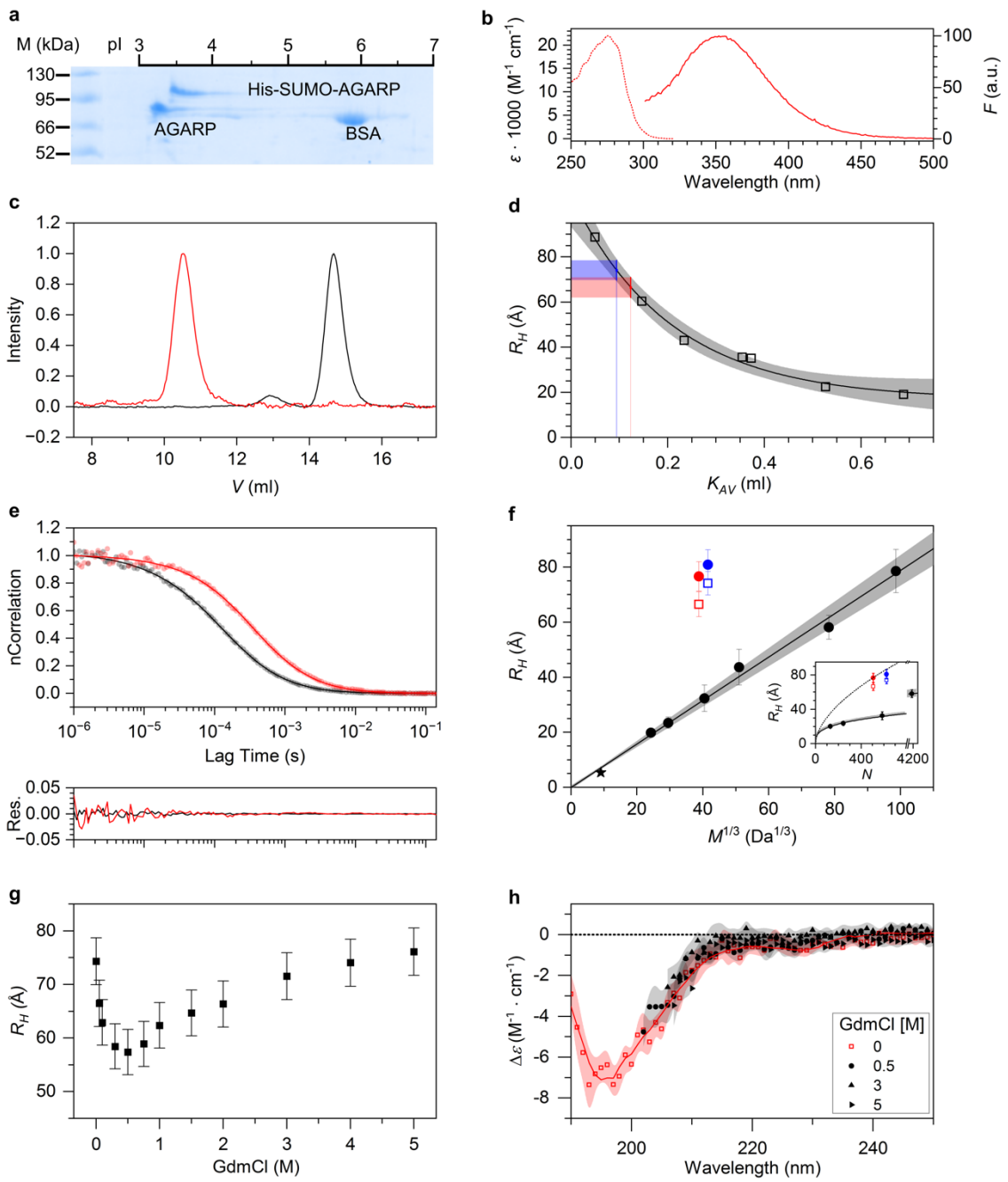

Fig. 2 Experimentally determined properties of AGARP. a 2D IEF SDS PAGE of AGARP and His$_6$-SUMO-AGARP. BSA was encompassed as a mass and pI standard; **b** Absorption (red dotted line) and emission (red solid line) spectrum of AGARP; **c** SEC profile of AGARP (red line) and HSA (black line); **d** Stokes radii of AGARP (red) and His$_6$-SUMO-AGARP (blue) determined by SEC at 10 °C with errors marked as shades colored correspondingly. Exponential function (black line) was fitted to data points determined for folded protein (open black squares); **e** Example normalized FCS autocorrelation curves for AGARP (red) and HSA (black) (dots, experimental points; lines, fitted models of diffusion) (upper panel) with corresponding fitting residuals (bottom panel); **f** Dependence of R_H on $M^{1/3}$

determined by FCS for folded proteins (black circles), AGARP (red circle) and His_6-SUMO-AGARP (blue circle) at 25 °C, and by SEC at 10 °C (hollow squares). Linear function was fitted to the values obtained for folded proteins and freely diffusing AF488 (black star). (inset) power function of R_H *vs.* number of protein residues (eq. 3) (solid black line) was fitted to experimental points for folded proteins. Shaded gray area, 99% confidence interval of the fit; dotted black line, power function previously determined for denatured proteins ($R_H = 2.21N^{0.57}$ Å)[68]; **g** Changes of AGARP R_H with increasing concentration of GdmCl as determined by SEC at 10 °C; **h** CD spectra of AGARP in the presence of different concentrations of GdmCl at 10 °C. Points indicate averaged experimental data with standard deviations marked as transparent shade colored correspondingly. Solid line is a fit performed by BeStSel[31,32].

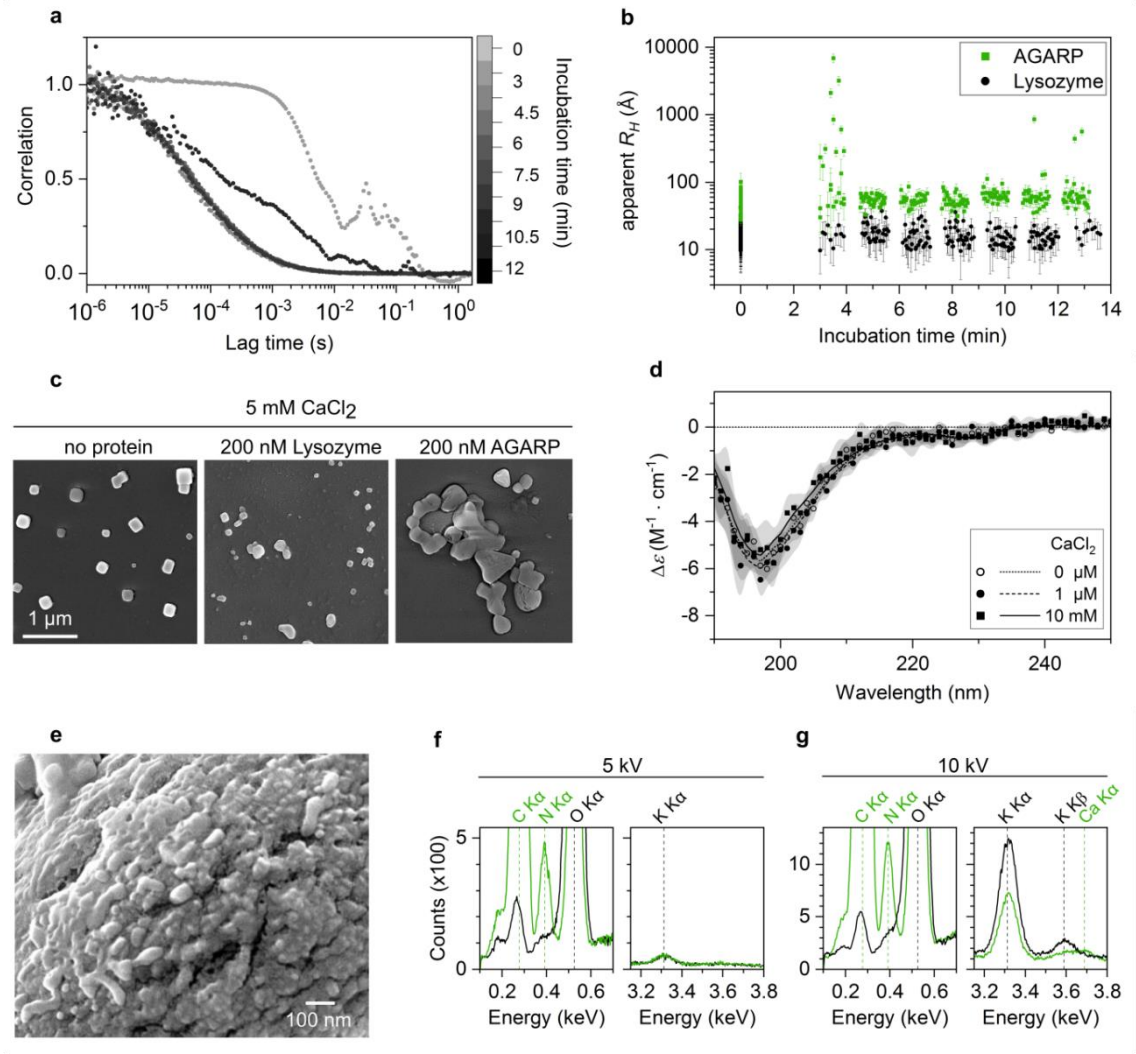

Fig. 3 AGARP-driven formation of ACC. a Normalized FCS autocorrelation curves obtained for 200 nM AGARP at subsequent time points after the addition of 5 mM $CaCl_2$; **b** (Apparent) R_H values of the phases containing AGARP and lysozyme (control) shown as a function of incubation time with Ca^{2+}; **c** SEM images of the phases deposited on the glass slide after the experiment; scale bar is 1 µm in all images; **d** Normalized circular dichroism spectra of AGARP in the absence and presence of $CaCl_2$. Points indicate averaged experimental data with standard deviations marked as transparent shade. Lines are fits performed by BeStSel[31,32]; **e** SEM image of the fragment of the phase formed after the FCS experiment analyzed by EDS at 5 kV AV; **f, g** EDS spectra obtained at (**f**) 5 kV and (**g**) 10 kV accelerating voltage. Green spectrum, $CaCO_3$ phase containing AGARP; black spectrum, glass slide (control); dashed lines, K lines of elements.

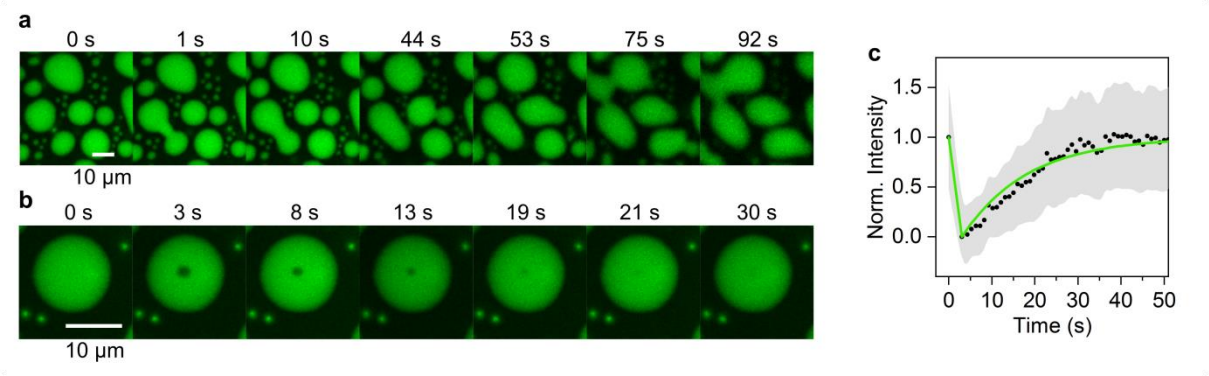

Fig. 4 Ca^{2+}-driven LLPS of AGARP. Liquid properties of the droplets formed after ~ 5 min at 50 µM AGARP, 100 mM $CaCl_2$ and 10% PEG 4k; **a** Fluorescence confocal microscopy images showing coalescence of droplets; **b** Fluorescence recovery after photobleaching, and **c** the corresponding FRAP curve. Shaded area, standard deviation of the measured intensities; green line, one component exponential function fitted in FRAPAnalyzer 2.1.0. Scale bar is 10 µm in all images.

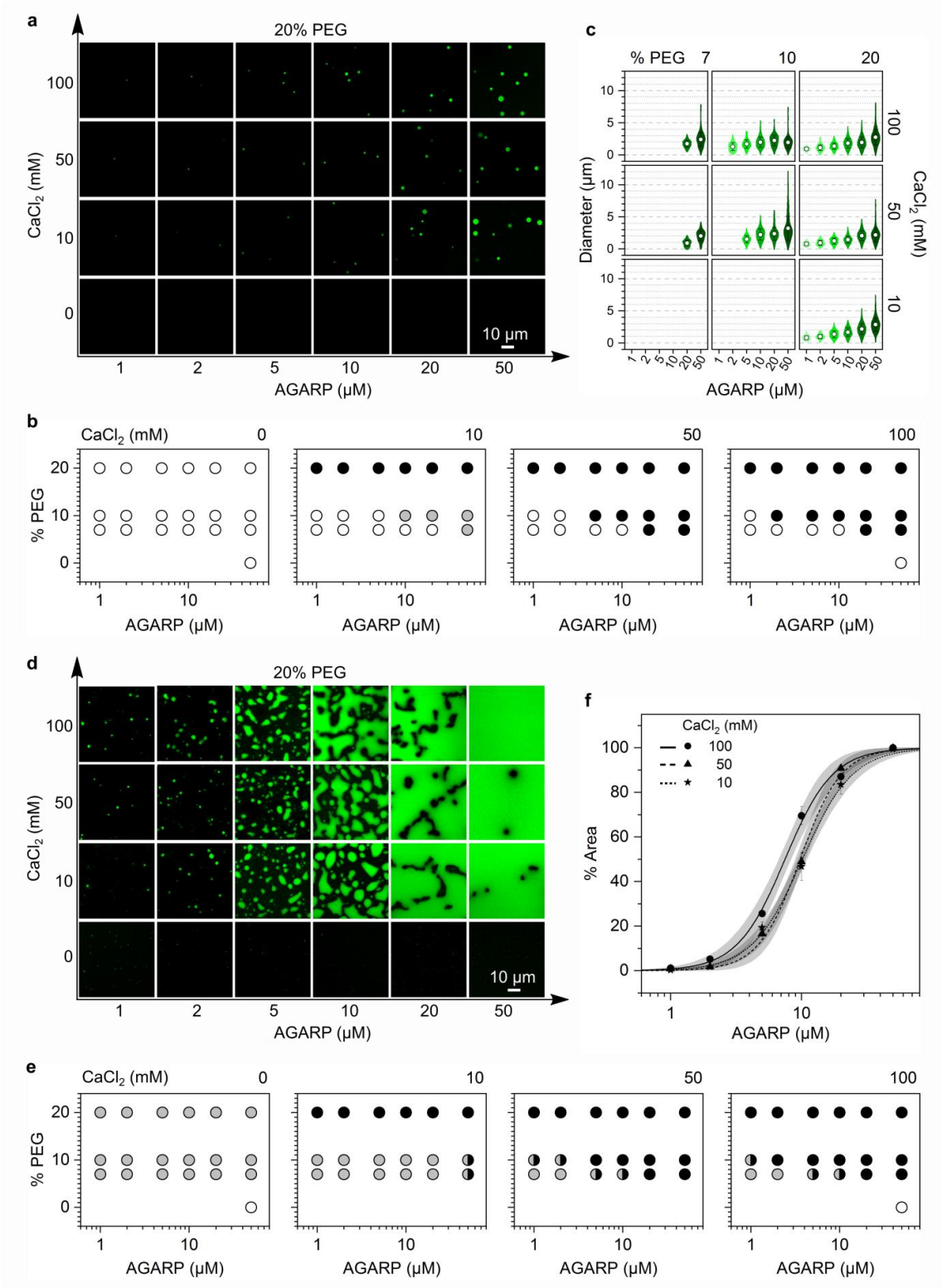

Fig. 5 Qualitative and quantitative analysis of Ca^{2+}-driven LLPS of AGARP. a Fluorescence confocal microscopy images of droplets formed in 20% PEG in the presence of varying concentrations of AGARP and $CaCl_2$; **b** Phase diagram of LLPS in the volume part (5 µm above the slide). Black, droplets; white, no droplets or aggregates; gray, aggregates; **c** Droplets diameter distribution. Increasing intensity of green indicates increasing

concentration of AGARP; **d** Fluorescence confocal microscopy images of AGARP-rich phases formed in 20% PEG in the presence of varying concentrations of AGARP and $CaCl_2$; **e** Phase diagram of LLPS observed at the surface of the slide. Black, droplets or phases formed as a results of droplets coalescence; white, no droplets or aggregates; gray, aggregates. gray/black; both droplets and aggregates; **f** Percentage of area occupied by AGARP-rich phase at 20% PEG 4k as a function of AGARP concentration. Eq. 4 was fitted to experimental data (lines). Scale bar is 10 μm in all images.

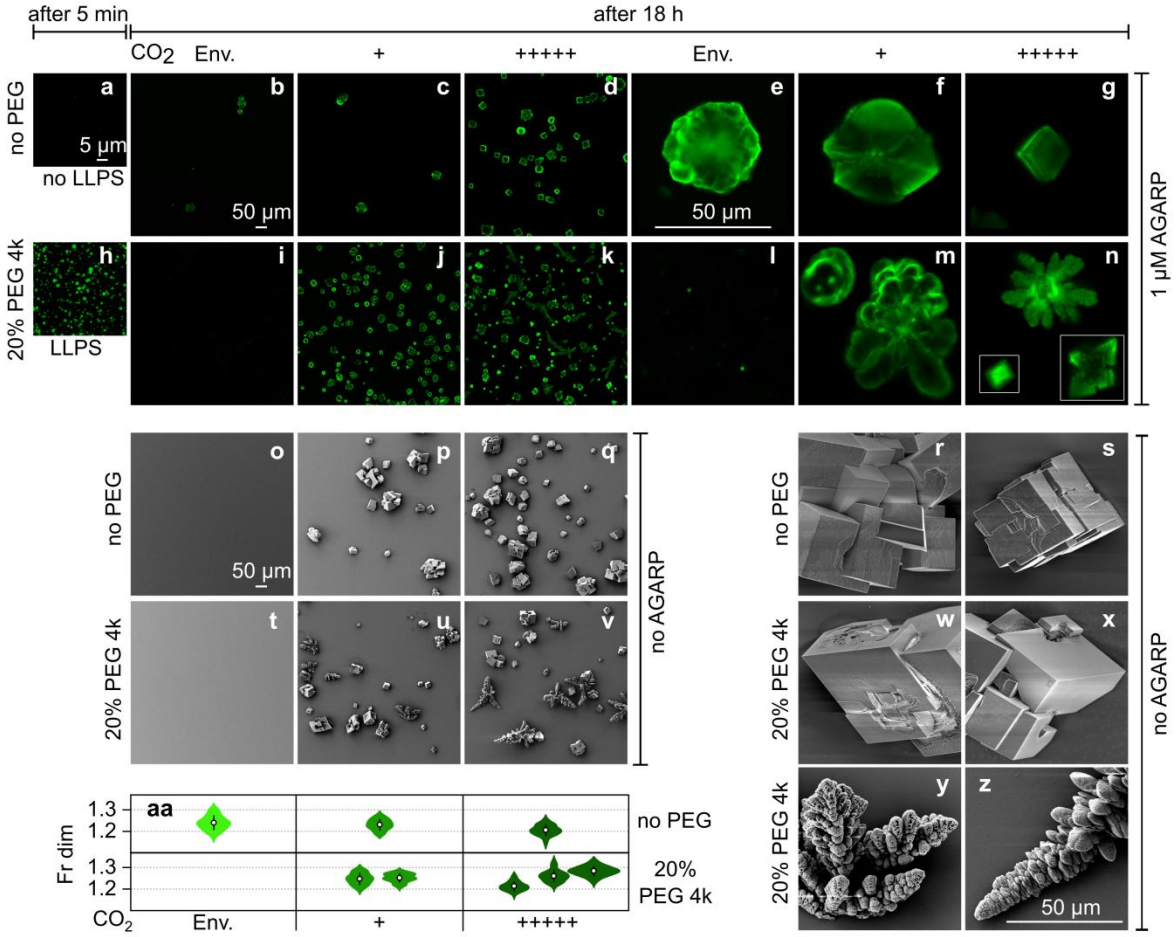

Fig. 6 $CaCO_3$ phases and crystals grown in the presence and absence of 1 µM AGARP with and without 20% PEG 4k after 18 h of crystallization at increasing concentrations of CO_2. a, h Fluorescence confocal microscopy images of phases formed in the presence of 1 µM AGARP prior to crystallization; $CaCO_3$ phases formed at increasing concentrations of CO_2 (**b**, **c**, **d**, **i**, **j**, **k,** fluorescence confocal microscopy images) in the presence and (**o**, **p**, **q**, **t**, **u**, **v**, SEM images) absence of AGARP found in the 708.5 x 708.5 µm area of observation and characteristic phases grown in the given conditions (**e, f, g, l, m, n** and **r, s, w, x, y, z**, respectively). Env., environmental source of CO_2; + and +++++, 12.5 mg/cm^3 and 62.5 mg/cm^3 $(NH_4)_2CO_3$ as a source of CO_2; **aa** Fractal analysis of the phases shown in panels **e, f, g, m, n** performed using ImageJ[36] and FracLac plugin[37]. Hollow black circle, mean; black line, standard deviation. Scale bar is 50 µm in all images.

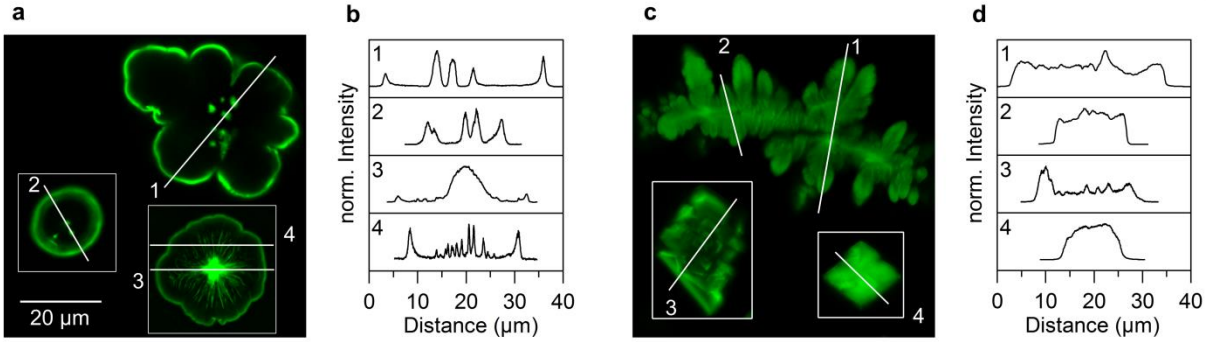

Fig. 7 AGARP distribution in $CaCO_3$ phases and crystals grown in the presence of 20% PEG 4k after 18 h of crystallization. Fluorescence confocal microscopy images of AGARP-rich $CaCO_3$ phases grown in the presence of (**a)** 12.5 mg/cm^3 and **(c)** 62.5 mg/cm^3 $(NH_4)_2CO_3$ as a source of CO_2 with profile lines (white) shown in **b** and **d**, respectively. Scale bar is 20 µm in all images.

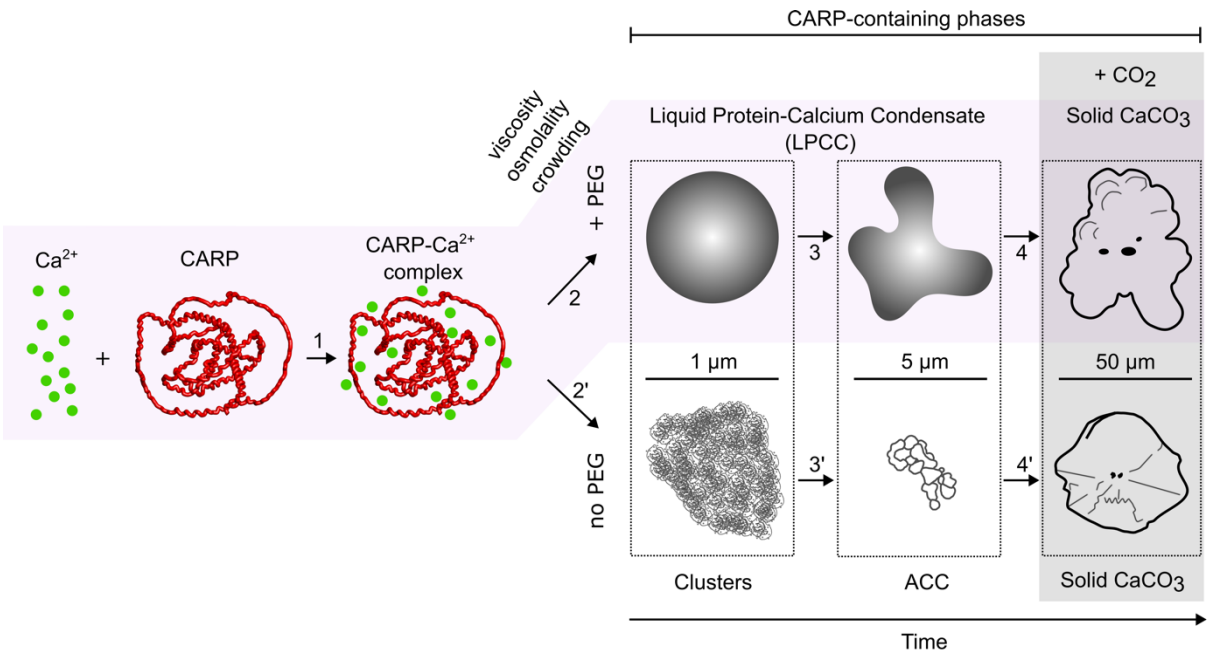

Fig. 8 Model of the CARP-mediated non-classical crystallization pathway.

When coral acid-rich proteins (CARPs) are exposed to the Ca^{2+}-rich milieu in ER, CARP-Ca^{2+} binding occurs (1). In highly viscous and crowded environment of SOM (violet area), LPCCs are formed via liquid-liquid phase separation (2). LPCCs coalesce (3) forming higher-order liquid phases. When LPCCs are found in the environment rich in CO_2, crystallization occurs (4) and $CaCO_3$ crystals with smooth and well developed edges appear. In less crowded and viscous conditions, CARP-Ca^{2+} complexes aggregate (2') which is followed by the transformation to amorphous calcium carbonate (ACC) (3'). Once ACC is exposed to CO_2 atmosphere, crystallization takes place (4') and $CaCO_3$ crystals with residual sharp edges are formed.

References

1. Jehannin, M., Rao, A. & Cölfen, H. New Horizons of Nonclassical Crystallization. *J. Am. Chem. Soc.* **141**, 10120–10136 (2019).
2. Gebauer, D., Völkel, A. & Cölfen, H. Stable Prenucleation Calcium Carbonate Clusters. *Science* **322**, 1819–1822 (2008).
3. Pouget, E. M. *et al.* The Initial Stages of Template-Controlled CaCO3 Formation Revealed by Cryo-TEM. *Science* **323**, 1455–1458 (2009).
4. Avaro, J., Moon, E. M., Schulz, K. G. & Rose, A. L. Calcium Carbonate Prenucleation Cluster Pathway Observed via In Situ Small-Angle X-ray Scattering. *J. Phys. Chem. Lett.* **14**, 4517–4523 (2023).
5. Mass, T. *et al.* Amorphous calcium carbonate particles form coral skeletons. *Proceedings of the National Academy of Sciences* **114**, E7670–E7678 (2017).
6. Gower, L. B. & Odom, D. J. Deposition of calcium carbonate films by a polymer-induced liquid-precursor (PILP) process. *Journal of Crystal Growth* **210**, 719–734 (2000).
7. Wolf, S. L. P., Caballero, L., Melo, F. & Cölfen, H. Gel-Like Calcium Carbonate Precursors Observed by *in situ* AFM. *Langmuir* **33**, 158–163 (2017).
8. Wolf, S. E. *et al.* Strong Stabilization of Amorphous Calcium Carbonate Emulsion by Ovalbumin: Gaining Insight into the Mechanism of 'Polymer-Induced Liquid Precursor' Processes. *J. Am. Chem. Soc.* **133**, 12642–12649 (2011).
9. Stifler, C. A., Killian, C. E. & Gilbert, P. U. P. A. Evidence for a Liquid Precursor to Biomineral Formation. *Crystal Growth & Design* **21**, 6635–6641 (2021).
10. Xu, Y. *et al.* Microscopic structure of the polymer-induced liquid precursor for calcium carbonate. *Nat Commun* **9**, 2582 (2018).
11. Allemand, D., Tambutté, É., Zoccola, D. & Tambutté, S. Coral Calcification, Cells to Reefs. in *Coral Reefs: An Ecosystem in Transition* (eds. Dubinsky, Z. & Stambler, N.) 119–150 (Springer Netherlands, Dordrecht, 2011). doi:10.1007/978-94-007-0114-4_9.
12. Mass, T. *et al.* Cloning and characterization of four novel coral acid-rich proteins that precipitate carbonates in vitro. *Curr Biol* **23**, 1126–1131 (2013).
13. Laipnik, R. *et al.* Coral acid rich protein selects vaterite polymorph in vitro. *Journal of Structural Biology* **209**, 107431 (2020).
14. Mass, T., Drake, J. L., Peters, E. C., Jiang, W. & Falkowski, P. G. Immunolocalization of skeletal matrix proteins in tissue and mineral of the coral Stylophora pistillata. *Proceedings of the National Academy of Sciences* **111**, 12728–12733 (2014).
15. Foster, B. *et al.* A novel *in vivo* system to study coral biomineralization in the starlet sea anemone, *Nematostella vectensis*. *iScience* **27**, 109131 (2024).
16. Mummadisetti, M. P., Drake, J. L. & Falkowski, P. G. The spatial network of skeletal proteins in a stony coral. *J R Soc Interface* **18**, 20200859 (2021).
17. Białobrzewski, M. K. *et al.* Diversity of hydrodynamic radii of intrinsically disordered proteins. *Eur Biophys J* **52**, 607–618 (2023).
18. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419–434 (2019).
19. Yoshizawa, T., Nozawa, R.-S., Jia, T. Z., Saio, T. & Mori, E. Biological phase separation: cell biology meets biophysics. *Biophys Rev* **12**, 519–539 (2020).
20. Uversky, V. N. Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. *Current Opinion in Structural Biology* **44**, 18–30 (2017).
21. Boskey, A. L. & Villarreal-Ramirez, E. Intrinsically disordered proteins and biomineralization. *Matrix Biology* **52–54**, 43–59 (2016).
22. Ramos-Silva, P. *et al.* The skeletal proteome of the coral Acropora millepora: the

evolution of calcification by co-option and domain shuffling. *Mol Biol Evol* **30**, 2099–2112 (2013).
23. Ying, H. *et al.* The Whole-Genome Sequence of the Coral Acropora millepora. *Genome Biology and Evolution* **11**, 1374–1379 (2019).
24. Technau, U. & Steele, R. E. Evolutionary crossroads in developmental biology: Cnidaria. *Development* **138**, 1447–1458 (2011).
25. Ramos-Silva, P. *et al.* The Skeleton of the Staghorn Coral Acropora millepora: Molecular and Structural Characterization. *PLOS ONE* **9**, e97454 (2014).
26. Cohen, A. L. & Holcomb, M. Why Corals Care About Ocean Acidification: Uncovering the Mechanism. *Oceanography* **22**, 118–127 (2009).
27. Parmar, A. S. & Muschol, M. Hydration and Hydrodynamic Interactions of Lysozyme: Effects of Chaotropic versus Kosmotropic Ions. *Biophys J* **97**, 590–598 (2009).
28. Andrews, P. Estimation of molecular size and molecular weights of biological compounds by gel filtration. *Methods Biochem Anal* **18**, 1–53 (1970).
29. Armstrong, J. K., Wenby, R. B., Meiselman, H. J. & Fisher, T. C. The Hydrodynamic Radii of Macromolecules and Their Effect on Red Blood Cell Aggregation. *Biophysical Journal* **87**, 4259–4270 (2004).
30. Petsev, D. N., Thomas, B. R., Yau, S.-T. & Vekilov, P. G. Interactions and Aggregation of Apoferritin Molecules in Solution: Effects of Added Electrolytes. *Biophysical Journal* **78**, 2060–2069 (2000).
31. Micsonai, A. *et al.* BeStSel: webserver for secondary structure and fold prediction for protein CD spectroscopy. *Nucleic Acids Research* **50**, W90–W98 (2022).
32. Micsonai, A. *et al.* BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res* **46**, W315–W322 (2018).
33. Waszkiewicz, R. *et al.* Hydrodynamic Radii of Intrinsically Disordered Proteins: Fast Prediction by Minimum Dissipation Approximation and Experimental Validation. *J. Phys. Chem. Lett.* **15**, 5024–5033 (2024).
34. Petrásek, Z. & Schwille, P. Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. *Biophys J* **94**, 1437–1448 (2008).
35. Le Guillou, J. C. & Zinn-Justin, J. Critical Exponents for the n-Vector Model in Three Dimensions from Field Theory. *Phys. Rev. Lett.* **39**, 95–98 (1977).
36. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671–675 (2012).
37. Karperien, A. FracLac for ImageJ. (1999).
38. Zięba, A. *Analysis of Experimental Data in Science and Technology*. (Cambridge Scholars Publishing, 2023).
39. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook* 571–607 (2005) doi:10.1385/1-59259-890-0:571.
40. Jones, D. T. & Cozzetto, D. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* **31**, 857–863 (2015).
41. Huang, Y. J., Acton, T. B. & Montelione, G. T. DisMeta: a meta server for construct design and optimization. *Methods Mol Biol* **1091**, 3–16 (2014).
42. Linding, R. *et al.* Protein disorder prediction: implications for structural proteomics. *Structure* **11**, 1453–1459 (2003).
43. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. & Jones, D. T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* **337**, 635–645 (2004).
44. Cheng, J., Sweredoski, M. J. & Baldi, P. Accurate Prediction of Protein Disordered Regions by Mining Protein Structure Data. *Data Min Knowl Disc* **11**, 213–222 (2005).

45. Prilusky, J. *et al.* FoldIndex©: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* **21**, 3435–3438 (2005).
46. Linding, R., Russell, R. B., Neduva, V. & Gibson, T. J. GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res* **31**, 3701–3708 (2003).
47. Dosztányi, Z., Csizmók, V., Tompa, P. & Simon, I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J Mol Biol* **347**, 827–839 (2005).
48. Vucetic, S., Brown, C. J., Dunker, A. K. & Obradovic, Z. Flavors of protein disorder. *Proteins: Structure, Function, and Bioinformatics* **52**, 573–584 (2003).
49. Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. G. & Pappu, R. V. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys J* **112**, 16–21 (2017).
50. Bolognesi, B. *et al.* A Concentration-Dependent Liquid Phase Separation Can Cause Toxicity upon Increased Protein Expression. *Cell Rep* **16**, 222–231 (2016).
51. Vernon, R. M. *et al.* Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *eLife* **7**, e31486 (2018).
52. Paiz, E. A. *et al.* Beta turn propensity and a model polymer scaling exponent identify intrinsically disordered phase-separating proteins. *J Biol Chem* **297**, 101343 (2021).
53. Hardenberg, M., Horvath, A., Ambrus, V., Fuxreiter, M. & Vendruscolo, M. Widespread occurrence of the droplet state of proteins in the human proteome. *Proceedings of the National Academy of Sciences* **117**, 33254–33262 (2020).
54. Vendruscolo, M. & Fuxreiter, M. Sequence Determinants of the Aggregation of Proteins Within Condensates Generated by Liquid-liquid Phase Separation. *J Mol Biol* **434**, 167201 (2022).
55. Vacic, V., Uversky, V. N., Dunker, A. K. & Lonardi, S. Composition Profiler: a tool for discovery and visualization of amino acid composition differences. *BMC Bioinformatics* **8**, 211 (2007).
56. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242 (2000).
57. Uversky, V. N., Gillespie, J. R. & Fink, A. L. Why are 'natively unfolded' proteins unstructured under physiologic conditions? *Proteins* **41**, 415–427 (2000).
58. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
59. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* **19**, 679–682 (2022).
60. Mariani, V., Biasini, M., Barbato, A. & Schwede, T. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **29**, 2722 (2013).
61. Necci, M., Piovesan, D. & Tosatto, S. C. E. Critical assessment of protein intrinsic disorder prediction. *Nat Methods* **18**, 472–481 (2021).
62. Drake, J. L. *et al.* Proteomic analysis of skeletal organic matrix from the stony coral Stylophora pistillata. *Proceedings of the National Academy of Sciences* **110**, 3788–3793 (2013).
63. Farre, B., Cuif, J.-P. & Dauphin, Y. Occurrence and diversity of lipids in modern coral skeletons. *Zoology* **113**, 250–257 (2010).
64. Goldberg, W. M. Acid polysaccharides in the skeletal matrix and calicoblastic epithelium of the stony coral Mycetophyllia reesi. *Tissue and Cell* **33**, 376–387 (2001).
65. Pilson, M. E. Q. Major constituents of seawater. in *An Introduction to the Chemistry of the Sea* 67 (Cambridge University Press, New York, 2013).
66. Goffredo, S. *et al.* The Skeletal Organic Matrix from Mediterranean Coral Balanophyllia europaea Influences Calcium Carbonate Precipitation. *PLoS ONE* **6**, e22338

(2011).
67. Adamiano, A. *et al.* Analytical pyrolysis-based study on intra-skeletal organic matrices from Mediterranean corals. *Anal Bioanal Chem* **406**, 6021–6033 (2014).
68. Wilkins, D. K. *et al.* Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. *Biochemistry* **38**, 16424–16431 (1999).
69. Müller-Späth, S. *et al.* Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proceedings of the National Academy of Sciences of the United States of America* **107**, (2010).
70. Różycka, M. *et al.* Intrinsically Disordered and Pliable Starmaker-Like Protein from Medaka (Oryzias latipes) Controls the Formation of Calcium Carbonate Crystals. *PLoS One* **9**, e114308 (2014).
71. Poznar, M. *et al.* Structural properties of the intrinsically disordered, multiple calcium ion-binding otolith matrix macromolecule-64 (OMM-64). *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1865**, 1358–1371 (2017).
72. Micsonai, A. *et al.* Disordered–Ordered Protein Binary Classification by Circular Dichroism Spectroscopy. *Frontiers in Molecular Biosciences* **9**, (2022).
73. Transport of Ions and Small Molecules across Cell Membranes. in *Molecular cell biology* (ed. Lodish, H. F.) (W.H. Freeman, New York, 2003).
74. Higgs, P. G. & Joanny, J. Theory of polyampholyte solutions. *J. Chem. Phys.* **94**, 1543–1554 (1991).
75. Cartwright, J. H. E., Checa, A. G., Gale, J. D., Gebauer, D. & Sainz-Díaz, C. I. Calcium Carbonate Polyamorphism and Its Role in Biomineralization: How Many Amorphous Calcium Carbonates Are There? *Angewandte Chemie International Edition* **51**, 11960–11970 (2012).
76. Von Euw, S. *et al.* Biological control of aragonite formation in stony corals. *Science* **356**, 933–938 (2017).
77. Akiva, A. *et al.* Minerals in the pre-settled coral Stylophora pistillata crystallize via protein and ion changes. *Nat Commun* **9**, 1880 (2018).
78. Reggi, M. *et al.* Biomineralization in Mediterranean Corals: The Role of the Intraskeletal Organic Matrix. *Crystal Growth & Design* **14**, 4310–4320 (2014).
79. Drake, J. L. *et al.* Proteomic analysis of skeletal organic matrix from the stony coral Stylophora pistillata. *Proc Natl Acad Sci U S A* **110**, 3788–3793 (2013).
80. Naggi, A. *et al.* Structure and Function of Stony Coral Intraskeletal Polysaccharides. *ACS Omega* **3**, 2895–2901 (2018).
81. Farre, B., Cuif, J.-P. & Dauphin, Y. Occurrence and diversity of lipids in modern coral skeletons. *Zoology (Jena)* **113**, 250–257 (2010).
82. Kar, M. *et al.* Phase-separating RNA-binding proteins form heterogeneous distributions of clusters in subsaturated solutions. *Proceedings of the National Academy of Sciences* **119**, e2202222119 (2022).
83. Deniz, A. A. Percolation physics and density transition frameworks converge in biomolecular condensation. *Proceedings of the National Academy of Sciences* **119**, e2210177119 (2022).
84. Mittag, T. & Pappu, R. V. A conceptual framework for understanding phase separation and addressing open questions and challenges. *Molecular Cell* **82**, 2201–2214 (2022).
85. Wojtas, M., Wołcyrz, M., Ożyhar, A. & Dobryszycki, P. Phosphorylation of Intrinsically Disordered Starmaker Protein Increases Its Ability To Control the Formation of Calcium Carbonate Crystals. *Crystal Growth & Design* **12**, 158–168 (2012).
86. Różycka, M. *et al.* Lattice Shrinkage by Incorporation of Recombinant Starmaker-Like Protein within Bioinspired Calcium Carbonate Crystals. *Chemistry – A European*

Journal **25**, 12740–12750 (2019).
87. Harting, P. *Recherches de morphologie synthétique sur la production artificielle de quelques formations calcaires organiques*. (Van der Post, 1872).
88. Blobel, G. Intracellular protein topogenesis. *Proceedings of the National Academy of Sciences* **77**, 1496–1500 (1980).
89. Denks, K. *et al.* The Sec translocon mediated protein transport in prokaryotes and eukaryotes. *Mol Membr Biol* **31**, 58–84 (2014).
90. Daverkausen-Fischer, L. & Pröls, F. Regulation of calcium homeostasis and flux between the endoplasmic reticulum and the cytosol. *J Biol Chem* **298**, 102061 (2022).
91. Rossi, A. M. & Taylor, C. W. Reliable measurement of free Ca2+ concentrations in the ER lumen using Mag-Fluo-4. *Cell Calcium* **87**, 102188 (2020).
92. Dayel, M. J., Hom, E. F. Y. & Verkman, A. S. Diffusion of Green Fluorescent Protein in the Aqueous-Phase Lumen of Endoplasmic Reticulum. *Biophysical Journal* **76**, 2843–2851 (1999).
93. Viotti, C. ER to Golgi-Dependent Protein Secretion: The Conventional Pathway. in *Unconventional Protein Secretion: Methods and Protocols* (eds. Pompa, A. & De Marchis, F.) 3–29 (Springer, New York, NY, 2016). doi:10.1007/978-1-4939-3804-9_1.
94. *The Golgi Apparatus: State of the Art 110 Years after Camillo Golgi's Discovery*. (Springer, Wien, 2008).
95. Gonzalez-Tello, P., Camacho, F. & Blazquez, G. Density and Viscosity of Concentrated Aqueous Solutions of Polyethylene Glycol. *J. Chem. Eng. Data* **39**, 611–614 (1994).
96. Moya, A. *et al.* Carbonic Anhydrase in the Scleractinian Coral Stylophora pistillata: CHARACTERIZATION, LOCALIZATION, AND ROLE IN BIOMINERALIZATION *. *Journal of Biological Chemistry* **283**, 25475–25484 (2008).
97. Tambutté, E. *et al.* Observations of the tissue-skeleton interface in the scleractinian coral Stylophora pistillata. *Coral Reefs* **26**, 517–529 (2007).
98. Smeets, P. J. M., Cho, K. R., Kempen, R. G. E., Sommerdijk, N. A. J. M. & De Yoreo, J. J. Calcium carbonate nucleation driven by ion binding in a biomimetic matrix revealed by in situ electron microscopy. *Nature Mater* **14**, 394–399 (2015).