Collagen intrafibrillar mineralisation as a result of the balance between osmotic equilibrium and electroneutrality

Li-na Niu[1]*, Sang Eun Jee[2]*, Kai Jiao[1], Lige Tonggu[3], Mo Li[3], Liguo Wang[3], Yao-dong Yang[4], Ji-hong Bian[4], Lorenzo Breschi[5], Seung Soon Jang[2], Ji-hua Chen[1], David H. Pashley[6], Franklin R. Tay[6]

**These authors contributed equally to this work*

Supplementary Information

S1 Detailed materials and methods

1. Chemicals

Poly(allylamine) hydrochloride (PAH, Mw: 15 kDa) or poly-L-aspartic acid-sodium salt (PAsp, Mw: 27 kDa) were used to stabilise a supersaturated calcium phosphate (CaP) mineralisation medium. Other chemicals used were $CaCl_2$ $2H_2O$, K_2HPO_4, bovine serum albumin (BSA, Mw: 66 kDa), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, dimethyl sulfoxide (DMSO, Mw: 78 Da), spermine (Mw: 202.34 Da), citric acid (Mw: 192.12 Da), 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), N-hydroxysulfosuccinimide sodium salt (sulfo-NHS), 2,4,6-trinitrobenzenesulfonic acid (TNBS), methylene blue and fluorescamine. These chemicals were all purchased from MilliporeSigma (St. Louis, MO, USA) unless noted otherwise.

2. Preparation and characterisation of mineralising medium

2.1 Turbidity assessment

The polycationic polymer PAH was used to stabilise a supersaturated calcium phosphate (CaP) mineralisation medium. Different amounts (1.25, 2.5, 5, 10, 20 and 40 mg) of PAH were added to 50 mL of 9 mM $CaCl_2$ $2H_2O$ solution and then mixed with an equal volume of 4.2 mM K_2HPO_4 solution. To maintain the pH of the mineralisation solution at 7.4, calcium and potassium solutions were made in 10 mM HEPES buffer (with 15 mM NaCl). The final PAH-stabilised amorphous calcium phosphate (PAH-ACP) solutions contained 4.5 mM Ca^{2+}, 2.1 mM HPO_4^{2-} and 12.5, 25, 50, 100, 200 or 400 μg/mL of PAH. Turbidity assessment was employed for indirect evaluation of the stability of the different versions of PAH-ACP solutions. Optical density profiles of the PAH-ACP solutions were performed at 0, 1, 12, 24, 48, 72, 96 and 120 h using a spectrophotometer (Synergy HT, BioTek Instruments, Winooski, VT, USA) at 405 nm. The mineralisation solution containing no polymer served as the negative control and deionised water was used as the baseline.

2.2 Preparation of PAH-ACP and PAsp-ACP

The PAH concentration (200 μg/mL) that enabled the mineralisation solution to remain clear without precipitation or colloid formation for at least 5 days was used to prepare PAH-ACP solution for intrafibrillar mineralisation in different collagen models.

Using the same method, 75 μg/mL of poly-L-aspartic acid-sodium salt was used to stabilise the CaP mineralisation medium to prepare PAsp-ACP.

2.3 Particle size distribution and zeta potential measurement of PAH-ACP and PAsp-ACP

Size distribution of the nanoparticles present in the PAH-ACP solution and PAsp-ACP solution was conducted using dynamic light scattering (DLS). In DLS, the distribution of diffusion coefficients is determined through measurement and correlation of the statistical fluctuations in the light scattered from a system of particle diffusing under the influence of Brownian motion. Analysis was performed with Zetasizer Nano ZS (Malvern Instruments, Westborough, MA, USA) using a 4 mW He–Ne laser (633 nm) and a backscatter detector at a 173 ° measuring angle. The instrument recorded the intensity auto-correlation function, which was transformed into volume functions to obtain size information.

Zeta potential analyses were performed using the DelsaNano C Zeta Potential Analyzer (Beckman Coulter, Inc., Brea, CA, USA). Analysis of PAH-ACP and PAsp-ACP suspended in HEPES buffer was performed in duplicates using a flow cell at 25 ℃. The zeta potential of charged particles in solution was determined by the electrophoretic light scattering method.

2.4 Cryogenic transmission electron microscopy (Cryo-TEM) of PAH-ACP

Quantifoil Jena R2/2/gold grids with holey-carbon support film and 2 μm hole size (Electron Microscopy Sciences, Hatfield, PA, USA) were surface plasma-treated using a rotary pumped carbon coating system (EMS150R S, Electron Microscopy Sciences) to increase their wettability and used within 30 min after plasma treatment. Fifty μL drops of CaP mineralisation medium containing 200 μg/mL PAH (prepared immediately or aged for different time periods) were placed over the plasma-treated grids. The liquid film on each grid was blotted for 2 sec to ~1000 Å thick and vitrified by rapidly plunged into melting liquid ethane at liquid nitrogen temperature, using a custom-made semi-automated compressed air-driven cryo-plunger with foot switch control. Cooled by liquid nitrogen, the ethane was maintained well below its boiling point and thus served as an excellent cryogen[S1]. The vitrified PAH-ACP coated grids were stored in cryo grid boxes (Ted Pella, Inc., Redding, CA, USA) and stored in liquid nitrogen overnight to allow slow evaporation of remnant frozen ethane. The grids were transferred to a Gatan 626 cryo-transfer holder and kept under -170 ℃ all the time. Images were acquired using a Tecnai G2 F20 microscope (FEI, Hillsboro, OR, USA) at 200 kV, using 50,000X magnification and -3 μm defocus. The electron dose for each exposure was around 20 e/Å^2 and images were recorded on an Eagle 4k x 4k CCD camera (FEI).

3. Mineralisation of reconstituted single-layer collagen

3.1 Reconstitution of single layer collagen

Rat tail tendon fascicles enveloped by their peritendinea were retrieved in iso-osmotic saline (0.9% NaCl) from the tails of mature (70 day-old) Wistar rats. Five mg of the rat tail tendon collagen was dissolved in 50 mL of 0.1 M acetic acid (pH 3.0) at 4 ℃ overnight. A single-layer of type I collagen fibrils was reconstituted over Ni TEM grids by neutralising the 0.1 mg/mL collagen stock solution with ammonia vapour for 4 h. The neutralised collagen solution was left to gel by incubation at 37 ℃ for 3 days. Cross-linking of the reconstituted collagen fibrils was performed by floating the collagen-coated grids upside down in 0.3M EDC/0.06M NHS solution for 4 h. Thereafter, the collagen-coated grids were rinsed with Milli-Q water and air-dried.

3.2 Cryogenic TEM and electron tomography of mineralised single-layer collagen

Quantifoil Jena R2/2/gold grids with holey-carbon support film and 2 μm hole size were used for reconstituting collagen fibrils, using the method described in the Section 3.1. Collagen-coated grids were mineralised by floating the grids upside-down over 50 μL droplets of PAH-ACP solution inside a 100% humidity chamber for 72 h. The grids were retrieved from the PAH-ACP solution without further staining. Ten nanometer gold particles were added as fiducial markers[S2]. Grids were vitrified in the manner described in Section 2.4. The cryogenic specimens were imaged at liquid nitrogen temperature (below -170 ℃) in the FEI Tecnai G2 F20 TEM equipped with a FEI 4k x 4k Eagle (CCD) camera. Images were taken at a magnification of 25,000×, giving a pixel size of 4.4 Å in the specimen.

For electron tomography, specimens were tilted in 2 ° steps from approximately 60 ° to −60° using the UCSF TOMO software package[S3]. The total electron dosage for the entire tilt series was estimated at ~100 e/Å2 per tilt series. Tomograms were reconstructed using the weighed back-projection method in the IMOD package[S4]. Images were binned twice and Gaussian-filtered to 0.2 before visualisation using the ImageJ software (NIH, Bethesda, MD, USA). Additional segmentation and visualisation of the 3-D volume were performed using Amira 5.3.3. (Visage Imaging Inc., Andover, MA, USA) and Matlab R2014a and 3D Volume Visualization package 1.1 (The MathWorks, Inc., Natick, MA, USA).

3.3 Conventional TEM of mineralised single layer collagen

Reconstituted collagen-coated, nickel TEM grids were floated upside-down over 50 μL droplets of PAH-ACP solution inside a 100% humidity chamber for 12-72 h. Then mineralised collagen were examined without staining, or stained with 1% uranyl acetate (MilliporeSigma) for 30 sec prior to TEM examination. Examination of the grids was conducted using the JEM-1230 TEM (JEOL, Tokyo, Japan) at 110 kV. Selected area electron diffraction was performed to identify the crystallinity of the infiltrated minerals. Images were recorded using an UltraScan 1000 4k x 4k CCD camera (Gatan, Inc., Pleasanton, CA, USA).

To correlate intrafibrillar minerals with specific locations along type I collagen fibrils, periodic electron density patterns that appeared along stained fibrils were examined at two specific stages of mineralisation, after intrafibrillar infiltration with ACP, and after transformation of intrafibrillar ACP into apatite, following previously published stained band designations[S5]. With the use of ImageJ software, a densitometric plot of the bands was generated from negative (grey scale-inverted) TEM images of one D period of selected mineralised collagen fibrils.

4. Effect of surface charge modification on the mineralisation of collagen

4.1 Crosslinking PAH to carboxyl groups on collagen fibrils (adding positive charge)

Because methylation of collagen via esterification of existing carboxylic acid group[S6] produced denatured fibrils that were devoid of cross banding patterns when examined by TEM (not shown), cationic collagen was produced by adding positive charges instead of eliminating existing negative charges[S7]. Accordingly, 25 mg collagen sponges (Ace Surgical Supply Co., Inc, MA, USA) were incubated in 10 mL of 0.1 M 2-morpholinoethane sulfonic acid solution (MES, pH 5.5) containing 2.5 mM of PAH for 1 h at room temperature. Then, 10 mL MES solution containing 40 mM of EDC was added. After incubation for 24 h at room temperature, the collagen sponges were washed with Milli-Q water (10 min each for 6 times), lyophilised and used as the cationic collagen model. Similar procedures were used to prepare cationic collagen in the reconstituted single-layer collagen model, with lyophilisation replaced with air-drying of the TEM grids.

4.2 Crosslinking PAsp to amine groups on collagen fibrils (adding negative charge)

Because preparation of anionic collagen via acetylation[S8] or selective hydrolysis[S9] also produced denatured fibrils without cross banding at the TEM level, anionic collagen was producing by adding negative charges instead of eliminating existing positive charges[S10]. One hundred and twenty five milligram of PAsp was dissolved in 25 mL of 0.1 M MES buffer (pH 6.0) and mixed with 10 mg of EDC and 27.5 mg of Sulfo-NHS for 15 min for NHS-ester

activation. The pH of the mixture was increased to 7.3 using concentrated phosphate-buffered saline (PBS, 0.1 M sodium phosphate, 0.15 M NaCl, pH 7.2-7.5) and 25 mg of collagen sponges was added. The reaction was allowed to proceed for 2 h at room temperature. The collagen sponges were then washed with Milli-Q water (10 min x 6 times), lyophilised and used as the anionic collagen model. Similar procedures were used to prepare anionic collagen in the reconstituted single-layer collagen model, with lyophilisation replaced with air-drying of the TEM grids.

4.3 Surface zeta potential measurement of modified collagen sponges

Zeta potential of collagen sponges before and after cationic or anionic modification (N = 3) was performed with the DelsaNano C Zeta Potential Analyser (Beckman Coulter, Inc., Indianapolis IN, USA) equipped with a flat surface cell for measuring solid zeta potential. The flat surface cell was used to hold flat sponges, which were rehydrated with 0.01 M NaCl aqueous solution at 25 ℃. The zeta potential measurements were conducted through dynamic light scattering of charged probe particles under influences from both electro-osmosis and electrophoresis. Polystyrene latex particles (Mw: 300,000 g/mol) with a diameter of 520 nm coated with hydroxypropyl cellulose were used as mobility-monitoring probe particles. The true electro-osmotic flow due to solid surface zeta potential was obtained by subtracting the true electrophoretic mobility from the total apparent electro-osmotic flow of the charged probe particles. This experiment was conducted in triplicate.

4.4 Determination of surface amine and carboxyl groups in modified collagen sponges

4.4.1 2,4,6-trinitrobenzenesulfonic acid (TNBS) assay

The amount of free amine groups present after crosslinking of PAH to carboxylic acid groups on collagen fibrils was determined with TNBS assay[S11]. Briefly, 1 mL of a 4 wt% $NaHCO_3$ solution (pH 9.0) and 1 mL of a freshly-prepared 0.5 wt% TNBS solution in distilled water were added to 4.5 mg of collagen sample (unmodified collagen sponge control, cationic collagen sponge or anionic collagen sponge). After reaction for 2 h at 40 ℃, 3 mL of 6 N HCl was added and the temperature was kept at 60 ℃ for 90 min to dissolve the collagen. The resulting solution was diluted with 5 mL of demineralised-water and the absorbance was measured at 345 nm. The control was prepared by applying the same procedures, except that HCl was added before the addition of TNBS. The primary amine group content of the collagen specimens was determined using a calibration curve correlating the absorbance with known glycine concentrations. The amine group content was expressed as millimoles of amine per gramme of collagen. The experiment was performed in sextuplicate.

4.4.2 Methylene blue sorption

Methylene blue sorption isotherms were used to determine the concentration of carboxylic acid groups in the cross-linked collagen specimens, based on the method described in ASTM D 1926-00[S12]. Unmodified collagen sponges, cationic collagen sponges and anionic collagen sponges were used for testing. A 100 mL of methylene blue-buffer solution (pH 8.0) containing 0.2 mM methylene blue, 0.625 mM diethylbarbituric acid (barbital) and 0.4 mM sodium hydroxide in distilled water were first prepared. Three specimens were weighed for each type of collagen sponge, one estimated to give 50 % exhaustion of the dye solution, one 10 to 15% lower than 50%, and one 10 to 15% higher than 50%. The specimens were immersed in 5

mL of the methylene blue-buffer solution, agitated overnight and centrifuged. One mL of the supernatant was mixed with 1 mL of 0.12 N HCl and further diluted to 10 mL volume with distilled water. The absorbance of the solutions was read at 620 nm using a spectrophotometer. The concentration of methylene blue in the supernatant derived from each collagen specimen was calculated from a pre-determined calibration curve collaborating known methylene blue dye concentrations with absorbance values. The relation between the specimen size in each of the three collagen groups and methylene blue concentration was obtained by plotting the specimen size against the concentration of methylene blue in the supernatant liquid. From this plot, the weight of the specimen that resulted in 50 % exhaustion of the dye solution was determined for each of the three collagen groups.

The specimen size that resulted in 50 % exhaustion of 5 mL of 0.2 mM methylene blue-buffer solution has utilised 0.0005 mM of methylene blue in ion exchange with the collagen carboxyl groups. Hence, the number of millimoles of carboxylic acid per gramme of collagen (M) is: M = (0.0005/W), where W = weight specimen resulting in 50 % exhaustion of 5 mL of 0.2 mM methylene blue solution. The experiment was performed in sextuplicate.

4.5 Attenuated total reflection–Fourier transform infrared spectroscopy (ATR-FTIR) of modified collagen sponges

Infrared spectra of the cationic and anionic collagen sponges, PAH, PAsp and un-modified collagen sponges were collected using a Nicolet 6700 FT-IR spectrophotometer (Thermo Scientific, Waltham, MA, USA) equipped with an ATR setup. Infrared spectra between 4,000-400 cm^{-1} at 4 cm^{-1} resolution were collected using 32 scans. The spectrum of cationic collagen sponge was normalised against that of unmodified collagen sponge along the –OH peak (~3270 cm^{-1}). The spectrum of anionic collagen sponge was normalised against that of unmodified collagen sponge along the collagen amide I peak (~1635 cm^{-1}).

4.6 Mineralisation of modified collagen

Cationic, anionic or non-modified collagen sponges were incubated in PAH-ACP or PAsp-ACP for 1 day or 7 days and then fixed in 2% glutaraldehyde, post-fixed in 1% osmium tetroxide, dehydrated in an ascending ethanol series (50-100%), immersed in propylene oxide and embedded in epoxy resin for TEM examination without further staining. Cationic and anionic single layer collagen were incubated in PAH-ACP or PAsp-ACP for 3 days, stained with 1% uranyl acetate and examined with TEM.

5. Penetrability of the collagen intrafibrillar milieu by molecules of different sizes

5.1 Accessibility of polymeric process-directing molecules to the intrafibrillar water

Modified size exclusion chromatography[S13,S14] that uses a column containing demineralised bone powder was employed to determine whether PAsp (Mw: 2.7×10^4 Da) and PAH (Mw: 1.5×10^4 Da), as well as small anionic and cationic molecules such as citric acid (Mw: 192.12 Da) and spermine (Mw: 202.34 Da) can access the intrafibrillar compartments of collagen fibrils. Bovine serum albumin (BSA, Mw: 6.6×10^4 Da) was used as the high molecular weight tracer control, while dimethyl sulfoxide (DMSO, Mw: 78 Da) was used as the low molecular weight tracer control.

Bovine femur cortical bone was obtained from a local abattoir and sectioned into 5 × 5 × 5 mm cubes under water cooling. The bone specimens were delipidated in a mixture of

chloroform and methanol (3:1 v/v), vigorously shaken for 3 days.Then the specimens were dehydrated in acetone for 20 min, immersed in liquid nitrogen for 15 min and pulverised into bone powder in a Retsch ball-milling machine (Model MM201, RetschInc, Newtown, PA, USA) for 4 sec at 30 Hz. The bone powder was passed through a series of stacked sieves. Powder with diameter between 425-600 μm was collected for further experiment. The bone powder was completely demineralised in 10-fold excess of Kristensen's solution (0.1 M formic acid buffered with sodium formate to a pH of 2.3; http://www.leicabiosystems.com/pathologyleaders/an-introduction-to-decalcification/) over three days at 4 ℃. The demineralised bone powder (4.0 g dry weight) was hydrated and packed into a 1 × 30 cm glass column to yield a volume of 23 mL. Packed bone powder within the column was equilibrated at room temperature with 10 mM HEPES buffer (pH 7.4) that contained 2 M NaCl to minimise nonspecific ionic interactions between test molecules and the demineralised bone powder[S15,S16]. The buffer (density: 1.076 g/ml) was used to elute the column at a flow rate of 14.4 mL/h using a precision syringe pump until the final effluent absorbance at 280 nm was less than 0.01, indicating that all extrafibrillar soluble protein has been rinsed from the column. Sample sizes were 1.0 mL and were collected in a fraction collector set to change every 4.17 min. The table below characterises the column packed with demineralised bovine bone powder.

Table I Characterisation of column packed with demineralised bovine bone powder

Total volume of packed column	23 mL
Determination of combined volume of extrafibrillar and intrafibrillar water	
Wet weight of column contents	57.692 g
Dry weight of column contents	39.298 g
Total weight of liquid in column	57.692 g - 39.298 g = 18.394 g
Volume of liquid in demineralised bone powder (i.e. **extrafibrillar and intrafibrillar water**)	18.394 g / (1.076 g/mL) = 17.09 mL
Determination of volume of intrafibrillar water	
Dry weight of demineralised bone powder	4.0 g
Wet weight of demineralised bone powder	10.628 g
Weight of liquid inside collagen	10.628 g - 4 g = 6.628 g
Volume of liquid inside collagen (i.e. **intrafibrillar water**)	6.628 g / (1.076 g/mL) = 6.06 mL

Tracer samples were dissolved in 0.5 mL of column buffer and contained approximately 2.5 mg of BSA (Mw: 6.6×10^4 Da), or 5.0 mg of PAsp (Mw: 2.7×10^4 Da), or 2.5 mg of PAH (Mw: 1.5×10^4 Da), or 5.0 mg of citric acid (Mw: 192.12 Da), or 5.0 mg of spermine (Mw: 202.34 Da), or 1 mg of DMSO (Mw: 78 Da), or 2.5 mg of PAH and 5.0 mg of spermine, or 5.0 mg of PAsp and 5.0 mg of citric acid. The volume of each fraction (1.01 ± 0.02 mL) was determined by weighing the fraction contents in a tared tube and divided by the density of the buffer (1.076 g/mL).

Elution of test tracers was determined as follows: BSA, absorbance at 280 nm; PAsp, and citric acid, at 228 nm; DMSO, absorbance at 220 nm; PAH and spermine, reacted with fluorescamine, and measured with Nanodrop 3300 fluorospectrometer (Thermo Scientific) at 470

nm.

5.2 Effects of inclusion of a short-chain polyamine (spermine) on collagen mineralisation with PAH-ACP

Collagen sponges (2 mm thick and 5 mm in diameter) were immersed in 5 mL of 200 μg/mL PAH-ACP solution, 200 μg/mL PAH-ACP + 0.3 mg/mL spermine solution, or 0.3 mg/mL spermine-CaP mineralisation solution at 37 ℃ for 1-7 days with the mineralising medium changed every 3 days. Specimens were then processed for TEM examination of the extent of intrafibrillar mineralisation of the collagen fibrils within the sponge leaflets.

5.3 Effects of inclusion of a short-chain tricarboxylic acid (citric acid) on collagen mineralisation with PAsp-ACP

Collagen sponges (2 mm thick and 5 mm in diameter) were immersed in 5 mL of 75 μg/mL PAsp-ACP solution, 75 μg/mL PAsp-ACP + 0.3 mg/mL citric acid solution, or 0.3 mg/mL citric acid-CaP mineralisation solution at 37 ℃ for 7 days with the mineralising medium changed every 3 days. Specimens were then processed for TEM examination of the extent of intrafibrillar mineralisation of the collagen fibrils within the sponge leaflets.

5.4 Effect of cationic or anionic polymer concentration on the osmolality of the polymer-stabilised ACP

The osmolality of the CaP solution containing 0, 50, 100, 200, 400, 600, 800, 1000, 1600, 2000 or 4000 μg/mL PAH or PAsp was measured with an osmometer (μOsmette, Precision Instruments Inc., Natick, MA, USA) and expressed as the number of milliosmoles per kilogramme of solvent (mOsm/Kg). The osmolality of PAH or PAsp aqueous solution with the same concentration was also measured for comparison.

6. Molecular dynamics (MD) simulations of the collagen molecules in presence of polyelectrolytes

Full atomistic molecular dynamics simulation of a high resolution hydrated collagen structure was constructed to investigate the transport of ions and water molecules from the intrafibrillar to the extrafibrillar region of collagen fibrils caused by Gibbs-Donnan equilibrium in the presence of polyanionic and polycationic process-directing agents, respectively.

6.1 Modelling collagen structure with intrafibrillar and extrafibrillar water

A collagen triple helix unit was constructed based on the X-ray crystallographic structures of type I collagen molecules of rat tail tendon reported by Orgel *et.al*. (PDB ID:3HR2)[S17]. Since this helical structure has low resolution and includes only α-carbons with sequence of *Rattus norvigecus*, full atomistic structures of collagen were constructed using homology modelling from human FASTA sequence, following the protocol reported by Nair *et al.*[S18] and Gautieri *et al*.[S19]. The FASTA sequences of α1(I) chain (entry number NP000079) and α2(I) chain (entry number NP000080) were from PubMed. Using Modeller 9.15 (https://salilab.org/modeller/download_installation.html)[S20], the obtained FASTA sequences were aligned on the crystallography-derived collagen structures; the structure with the lowest energy was chosen based on Discrete Optimised Protein Energy (DOPE) calculations. The resultant triple helix was 1.28 nm in diameter and 298.025 nm in length. Collagen fibrillar

structure was described using a 67 nm long simulation box. A long collagen fibril was produced by interacting the collagen triple helix units with their self-images using periodic boundary condition, as reported in previous MD simulation studies.

To obtain energetically stable collagen structures in their hydrated states, MD simulations of one unit cell was performed with 0.8 g water/g collagen[S21,S22], which contains primary hydration layers at its native states of collagen structures. The structures were energy-minimised using the steepest descent gradient method and equilibrated using constant number, volume and temperature (NVT) and then constant number, pressure and temperature (NPT) ensemble for 1.0 ns and 2.4 ns each until the energy, pressure and temperature reached equilibrium. The NPT simulation was then performed for 12 ns until the dimension of the structures converged. The converged dimensional specifications of the system was a = 40.23Å, b = 27.19 Å, c = 677.90 Å, $\alpha = 90°$, $\beta = 90°$, $\gamma = 105.58°$, which were similar to the dimensions reported by Orgel *et al*. based on crystallographic data (a = 39.97 Å, b = 26.95 Å, c = 677.90 Å, $\alpha = 89.24°$, $\beta = 94.59°$, $\gamma = 105.58°$). Slight dimensional differences could be caused by setting different α and β values for technical convenience to construct collagen fibrillar structures at the interface between the intrafibrillar space and the extrafibrillar space. To describe the interface, 3x1x1 cells of one equilibrated collagen cell were constructed based on the equilibrated single cell collagen structures, which contained five collagen triple helices in one cell and added water layer in the "a" direction, as shown in Fig.5a in the main text. To obtain equilibrated fibrillar structures with interface, the system was equilibrated without polyelectrolytes, using NVT and NPT equilibration for 1.6 ns and 2.0 ns, each with imposing restraint on the backbone atoms of the collagen molecules. After that, 40 ns of NPT simulation was performed to relax the collagen structures. The hydrated collagen structures at interface reached equilibrium at 20 ns. Hence, the structure at 20 ns was extracted to introduce polyelectrolytes.

The Groningen Machine for Chemical Simulations (GROMACS) molecular simulation packages (http://www.gromacs.org)[S23] were employed for performing all MD simulations. Interatomic interactions, collagen structure and water molecules were described by AMBER99sb-ILDN force field[S24] and simple pointy charge (SPC) water model[S25]. Newton's equations of motion were integrated using velocity-verlet algorithm with a 2 femtosecond time-step. Long range electrostatic interactions were described using PME (Particle Mesh Ewald) method. Pressure and temperature were controlled to 1 bar and 310 K, respectively, using the Berendsen thermostat and the Parrinello-Rahman barostat.

6.2 Modelling movement of ions and water molecules in the presence of polyelectrolytes

To investigate the effect of the presence of large polyelectrolyte molecules on the movement of ions and water molecules between the intrafibrillar and extrafibrillar spaces, bulky particles with electronic charges were placed in the centre of the extrafibrillar regions of the aforementioned collagen model, representing polycations (i.e. PAH) or polyanions (i.e. PAsp), as indicated in Fig.5a.

One particle was set to have $54e^+$ for the polycationic system ($54e^-$ for the polyanionic system), which was designated as 50% of the total collagen charges (+108e) of the system. The van der Waals radius of the particle was set to 1.5 nm. The distance between collagen extrafibrillar region and charged particle was 5 nm. Sodium ions and Cl ions were added to neutralise the system for MD simulations. 568 Na^+ and 1054 Cl^- were added to the polycationic system and 946 Na^+ and 676 Cl^- were added to the polyanionic system. The structures were energy-minimised and equilibrated with 1.6 ns NVT and 2.4 ns NPT equilibrium by restraining

the backbone atoms of the collagen structures. Then, 21.6 ns simulation was performed and the data was collected from the last 20 ns after 1.6 ns of equilibration. The changes in the number of Na^+, Cl^- and water molecules in the intrafibrillar regions and the changes in the number of bound water and hydrogen bonds between collagen and water in the 1st hydration layer of collagen structures were recorded as a function of simulation time.

It is technically impossible to measure internal pressure imposed on the collagen structures in equilibrium MD simulations due to pre-established constant pressure environment (the pressure of each atom was controlled to have a reference pressure of 1 bar using brostats in NPT molecular simulations). Hence, the structural variations of collagen molecules within the collagen structure were measured as indirect evidence of the existence of negative intrafibrillar pressure following the establishment of Gibbs-Donnan equilibrium in the presence of extrafibrillar polyelectrolytes. Root Mean Square Deviations (RMSD) of collagen molecules was first analysed between atomic position at time *t* and the initial positions of backbone atoms of collagen structures. Solvent Accessible Surface Area (SASA) was calculated to investigate the deformation of collagen based on the double cubic lattice method[S26]. To account for the free surface area for water molecules, a probe with a radius of 1.4Å was utilised, which corresponds to the hydrodynamic radius of a water molecule. Since changes in the approximation of the collagen molecules would also modify their hydrogen bond networks, the number of hydrogen bonds among the collagen molecules was also examined.

6.3 Modelling migration of CaP mineralisation precursors into contracted collagen structures in the presence of polyelectrolytes

To investigate the infiltration of CaP into the contracted collagen structures using molecular dynamics simulations, specific configurations of CaP mineralisation precursors and their populations should be clarified for describing interactions. However, the configuration of biological CaP is uncertain. Although Ca/P ratios and their structural configurations have been studied in amorphous calcium phosphate forms[S27,S28], little is known about the Ca/P ratio in the intrafibrillar regions of the collagen structures. Moreover, there is no robust forcefield to describe the divalent metal ions in a protein/water environment[S29,S30]. Thus, CaP precursors in the collagen structure is described by sole Ca ions[S31] or by the forcefield developed for hydroxyapatite[S32,S33]. Since the purpose of the present molecular simulation was to investigate qualitatively if re-swelling of the polyelectrolyte-induced contraction of the collagen fibrils would induce the intrafibrillar diffusion of CaP precursors, the model was simplified by assuming the Ca ions form CaP prenucleation clusters and treated the former as single particles with 135 Da and +2 charges. Amber forcefield[S34] for Ca ions were used to describe CaP prenucleation clusters with modification of ε (van der Waals interaction parameter) to describe larger dispersion interactions of CaP instead of Ca ions. 560 Ca ions were added to the entire system obtained from previous simulation of the polycationic system or the polyanionic system, and 1120 Cl ions were added to each system to neutralise the system. Molecular dynamics simulations were performed for 10 ns with the same conditions as previous simulations. After the simulation, the distribution of ions and water molecules surrounding the collagen molecules was analysed. The number of the ions and water molecules were counted at $r \sim r+\Delta r$, where r is the minimum distance from the collagen structures. The number of water/ions were counted Δr, which is every 0.05 nm, over 0~0.625 nm from the collagen molecules. The number of Ca ions, Na, Cl ions and water molecules closest to the collagen structures was counted and plotted.

Solvent Accessible Surface Area (SASA) was also calculated to investigate the structural variation of the collagen structure, following the introduction of Ca ions into the system.

7. Solution-induced dimensional changes in hydrated single layer collagen

Single layers of rat tail type I collagen were reconstituted as described in section 3.1 on freshly-cleaved muscovite mica (Grade V1, Ted Pella Inc., CA, USA) instead of TEM grids. Atomic force microscopy (AFM) imaging of the dried collagen was first conducted using silicon nitride cantilevers with integral tips (Applied Nanostructures, Santa Clara, CA, USA) on a Cypher S AFM (Asylum Research, Santa Barbara, CA, USA) with a 1-Hz scan rate. A closed fluid cell was employed so that the surfaces with adsorbed fibrils were totally immersed in 2 mL of solution for 1 h or 1 day before scanning. Prior to immersion in the different solutions, the adsorbed fibrils were rinsed with Milli-Q water for 6 times and 5 min/rinse. The solutions used were:

1. HEPES buffer (10 mM, pH 7.4) for 1 h
2. HEPES buffer containing 19.8 mM NaCl (which has the same ionic strength as the mineralising medium without polymer) for 1 h
3. HEPES buffer containing 200 μg/mL PAH for 1 h
4. HEPES buffer containing 200 μg/mL PAH and 19.8 mM NaCl for 1 h
5. HEPES buffer containing 200 μg/mL PAH and 19.8 mM NaCl for 1 day
6. HEPES buffer containing 200 μg/mL PAH, 4.5 mM $CaCl_2 \cdot 2H_2O$ and 2.1 mM K_2HPO_4 (PAH-ACP) for 1 h
7. HEPES buffer containing 200 μg/mL PAH, 4.5 mM $CaCl_2 \cdot 2H_2O$ and 2.1 mM K_2HPO_4 (PAH-ACP) for 1 day.

The acquired images (field size: 1 ×1 μm) were analysed with IGOR Pro 6.2 (WaveMetrics Inc., Lake Oswego, OR, USA) to obtain the diameter changes of the collagen fibrils in different solutions. This experiment was conducted in sextuplicate. For each repeat, images of the same collagen fibril in different solutions were collected. One-factor repeated measures analysis of variance was performed to detect significant differences among the various groups, after validation of the normality and homoscedasticity assumptions of the data set. Post-hoc multiple comparisons were conducted using the Holm-Sidak statistic. For all analyses, statistical significance was pre-set at $\alpha = 0.05$.

The aforementioned set of experiments was repeated by replacing PAH with PAsp.

SI-2 Characterisation of poly(allylamine)-stabilised amorphous calcium phosphate (PAH-ACP) and poly(aspartic acid)-stabilised ACP

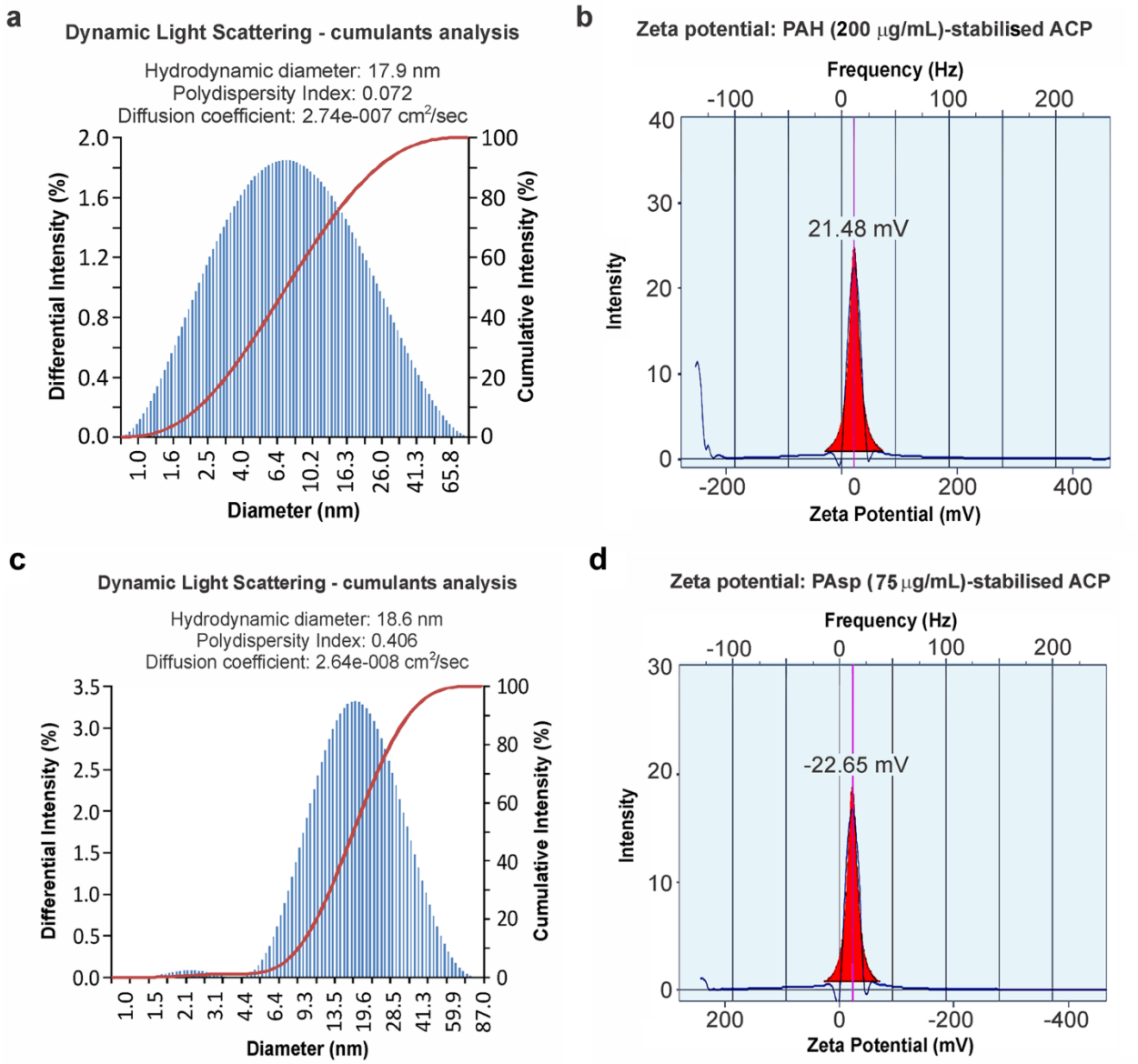


Figure SI-2 a) Dynamic light scattering and **b**) zeta potential measurement of ACP stabilised by 200 µg/mL PAH. The dimension and zeta potential of poly(aspartic acid)-stabilised ACP are included for comparison. **c**) Dynamic light scattering and **d**) zeta potential data of ACP stabilised by 75 µg/mL poly(aspartic acid) (Mw 27 kDa). Poly(aspartic acid)-stabilised ACP is dimensionally akin to PAH-ACP, but has negative zeta potential instead of positive zeta potential.

SI-3 Cryogenic TEM of vitrified PAH-stabilised CaP solution

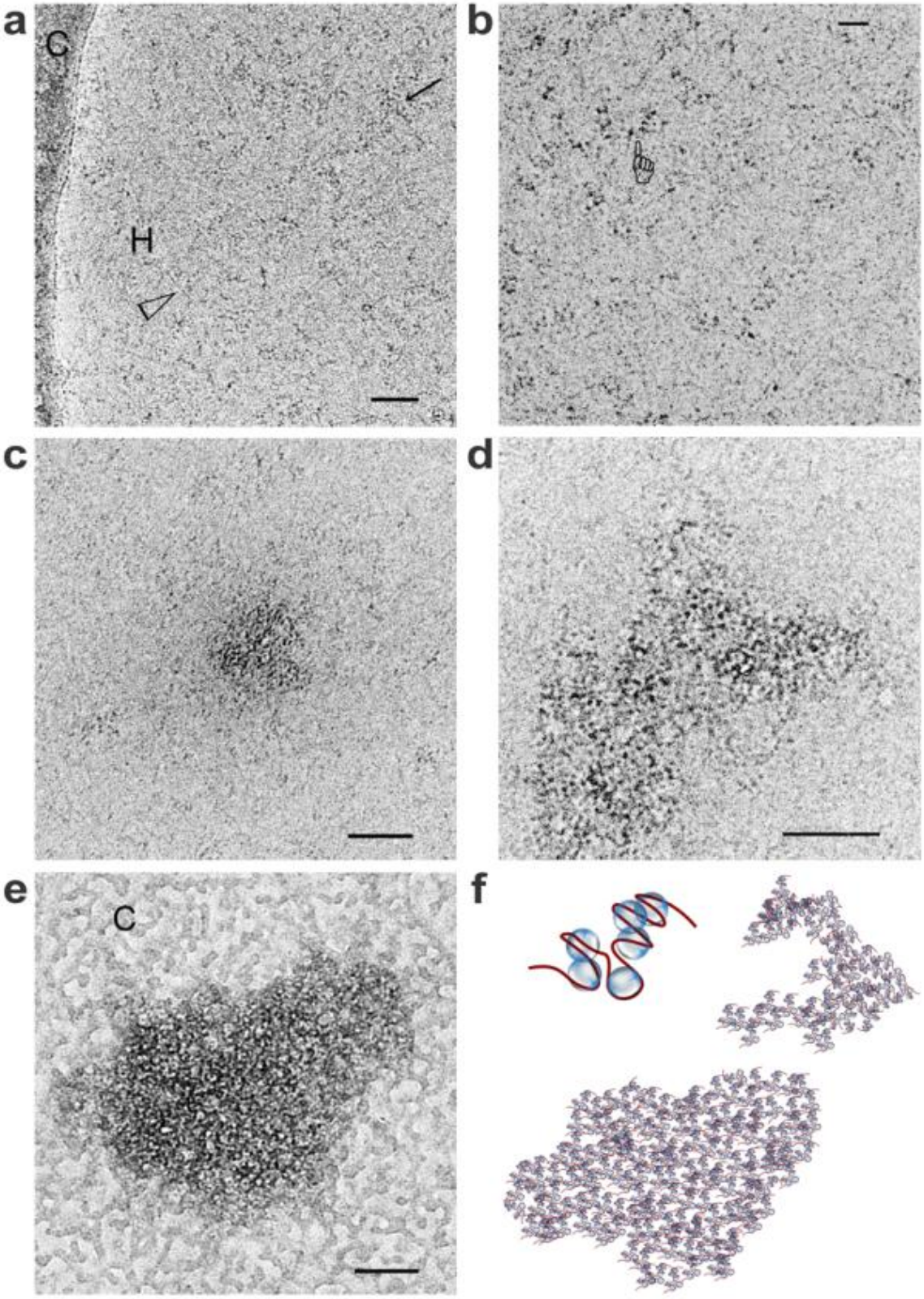


Figure SI-3 Unstained cryogenic TEM of vitrified PAH-stabilised CaP solution. **a**) 15 min post-preparation showing electron-dense prenucleation complexes (arrow) among filamentous structures (arrowhead). H: hole; C: carbon film; bar: 50 nm. **b**) High magnification of "**a**" showing association of the prenucleation complexes into ensembles (pointer) (bar: 20 nm). **c**) 24 h post-preparation showing densification of ensembles (bar: 50 nm). **d**) 3 days post-preparation, with 150-200 nm branched, CaP ensembles (bar: 50 nm). **e**) 3 days post-preparation. Carbon film region (C) showing further densification of loose CaP ensembles into ACP (bar: 50 nm). Contamination of crystalline ice on the carbon film resulted in a 'splotchy' appearance. **f**) Schematics of i) polymer-prenucleation cluster ensemble, ii) branched assembly, iii) ACP.

SI-4 Conventional transmission electron microscopy (TEM) reveals no difference in the morphology of PAH-stabilised CaP aggregates, compared to those observed by cryogenic TEM

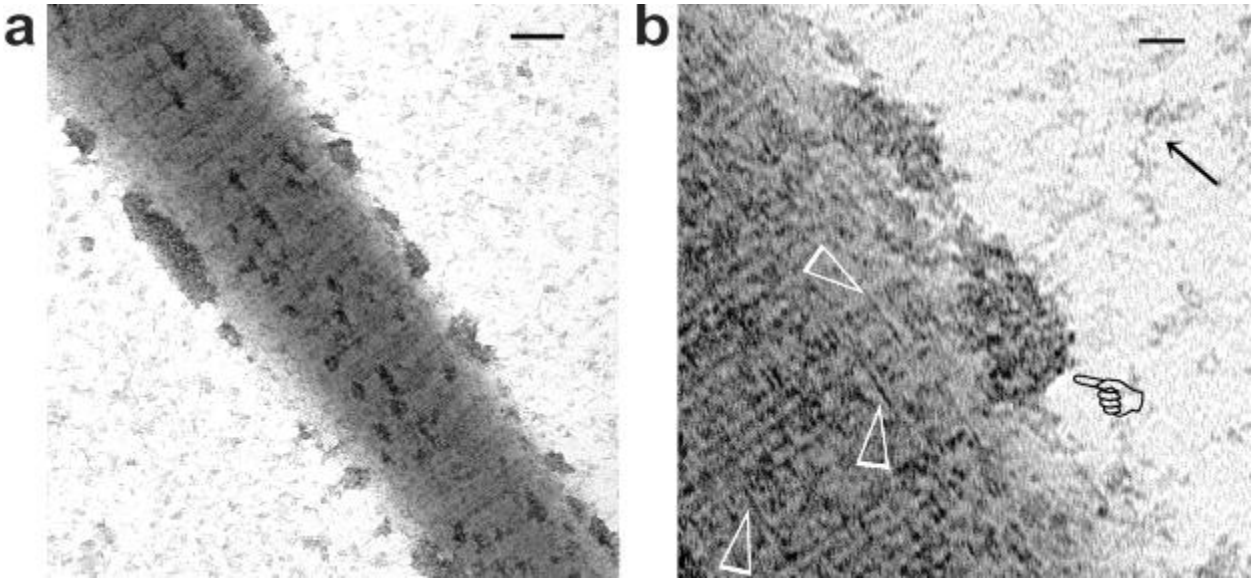
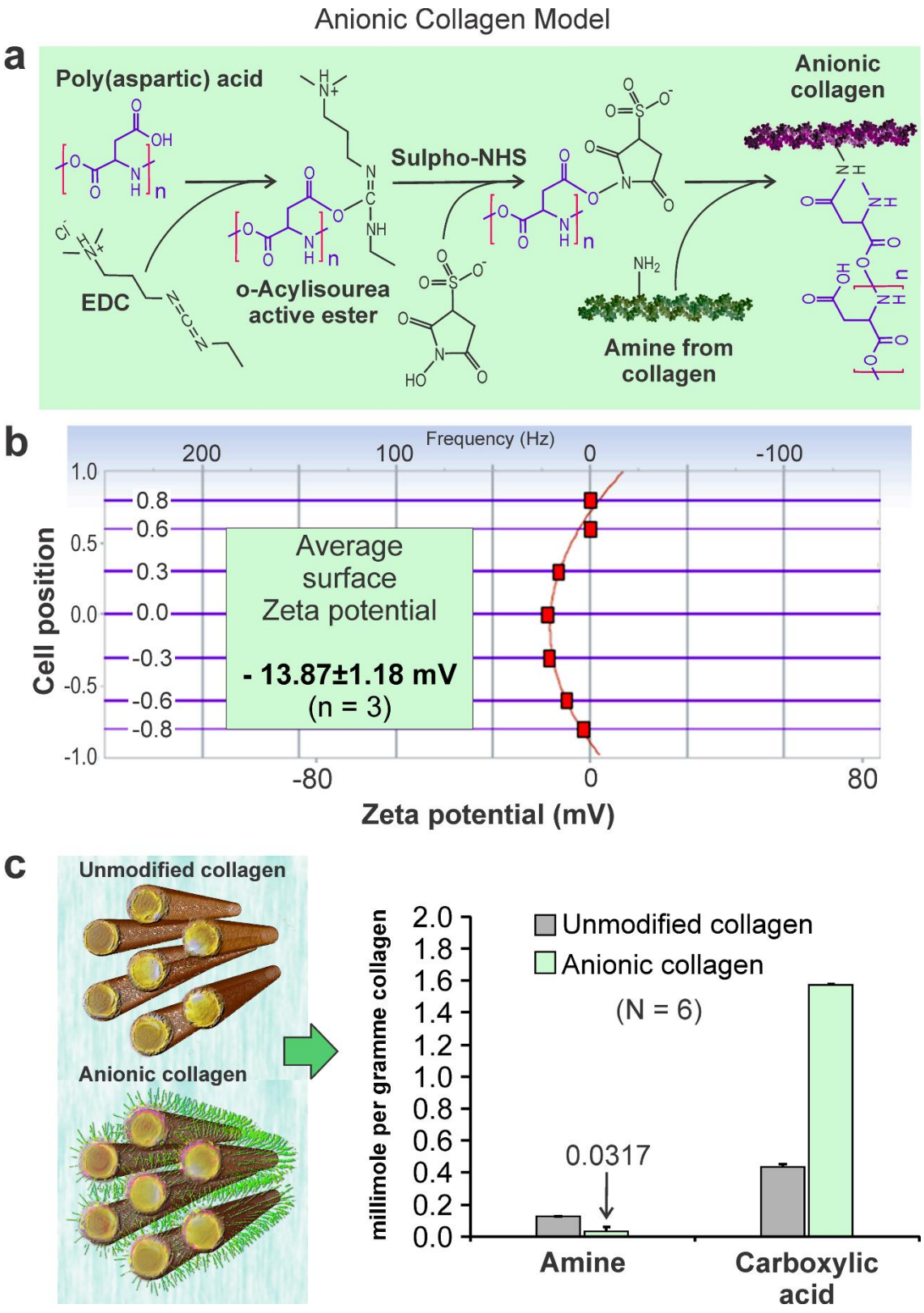


Figure SI-4 Uranyl acetate-stained conventional TEM images of the early stage of collagen mineralisation in 200 μg/mL PAH-stabilised CaP mineralising solution. **a**) Low magnification of the condensation of PAH-stabilised CaP aggregates on the surface of a collagen fibril (bar: 100 nm). These aggregates are similar in appearance to that shown in cryo-TEM (Figure SI-3). **b**) High magnification of the previous image showing PAH-stabilised CaP complexes (arrow) that condensed into PAH-stabilised ACP (pointer) on the surface of the partially-mineralised collagen fibril (bar: 20 nm). Open arrowheads: intrafibrillar apatite crystallites. Because CaP aggregates are well-stabilised by PAH and appear similar when examined after vitrification by cryogenic TEM, or after desiccation by conventional TEM, subsequent examination of collagen intrafibrillar mineralisation was conducted using conventional TEM.

SI-5 Anionic collagen model for studying intrafibrillar mineralisation by PAH-ACP



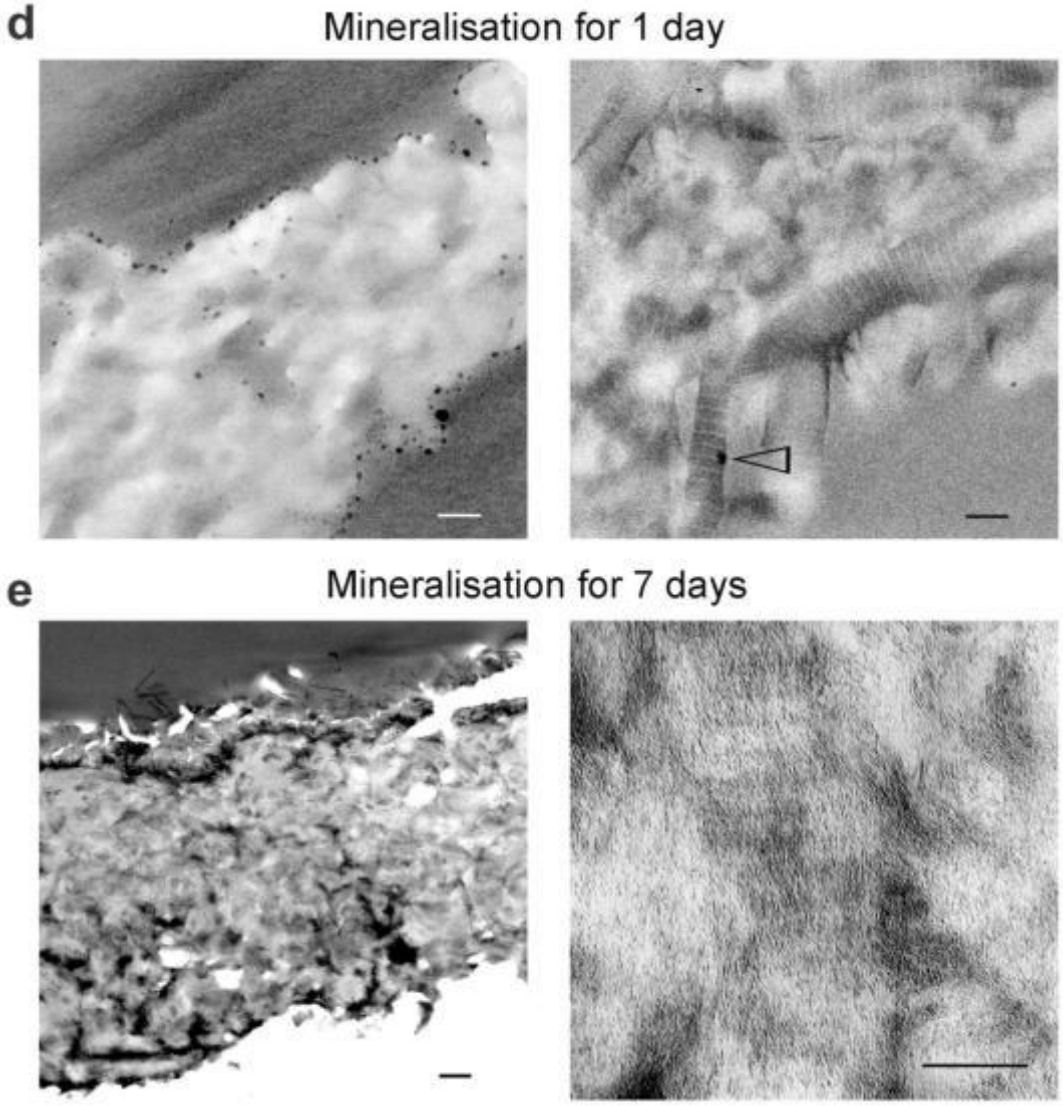


Figure SI-5 **a)** Method of conjugation of polyanions [e.g. poly(aspartic acid)] to collagen fibrils via amino acids containing $–NH_2$ side chains. Sulfo-NHS: N-hydroxysulfosuccinimide. **b**) Surface zeta potential of anionic collagen sponge. Zeta potential was determined from the measured negative electrophoretic mobility of the probe particles. **c**) Comparison of amine and carboxyl groups between anionic collagen and unmodified collagen. Decrease in amine groups is due to their conjugation with poly(aspartic acid) via the O=C−N-H linkage. **d**) TEM of collagen leaflets from anionic collagen mineralised for 1 day (Left – bar: 500 nm. Right – bar: 200 nm). Contrary to the belief that electrostatic attraction between opposite charges present on the collagen surface and PAH-ACP would enhance ACP infiltration into collagen fibrils, intrafibrillar mineralisation was not observed on the first day. The collagen surface was covered with electron-dense particles (arrowhead). **e**) TEM of collagen leaflets from anionic collagen sponges that were mineralised for 7 days (Left – bar: 500 nm. Right – bar: 200 nm). Mineralisation of anionic collagen by PAH-ACP was slow and inefficient compared with unmodified collagen. The anionic collagen fibrils were partially filled with fine intrafibrillar apatite that produced a faintly-banded appearance.

SI-6 Fourier transform infrared spectroscopy of cationic and anionic collagen

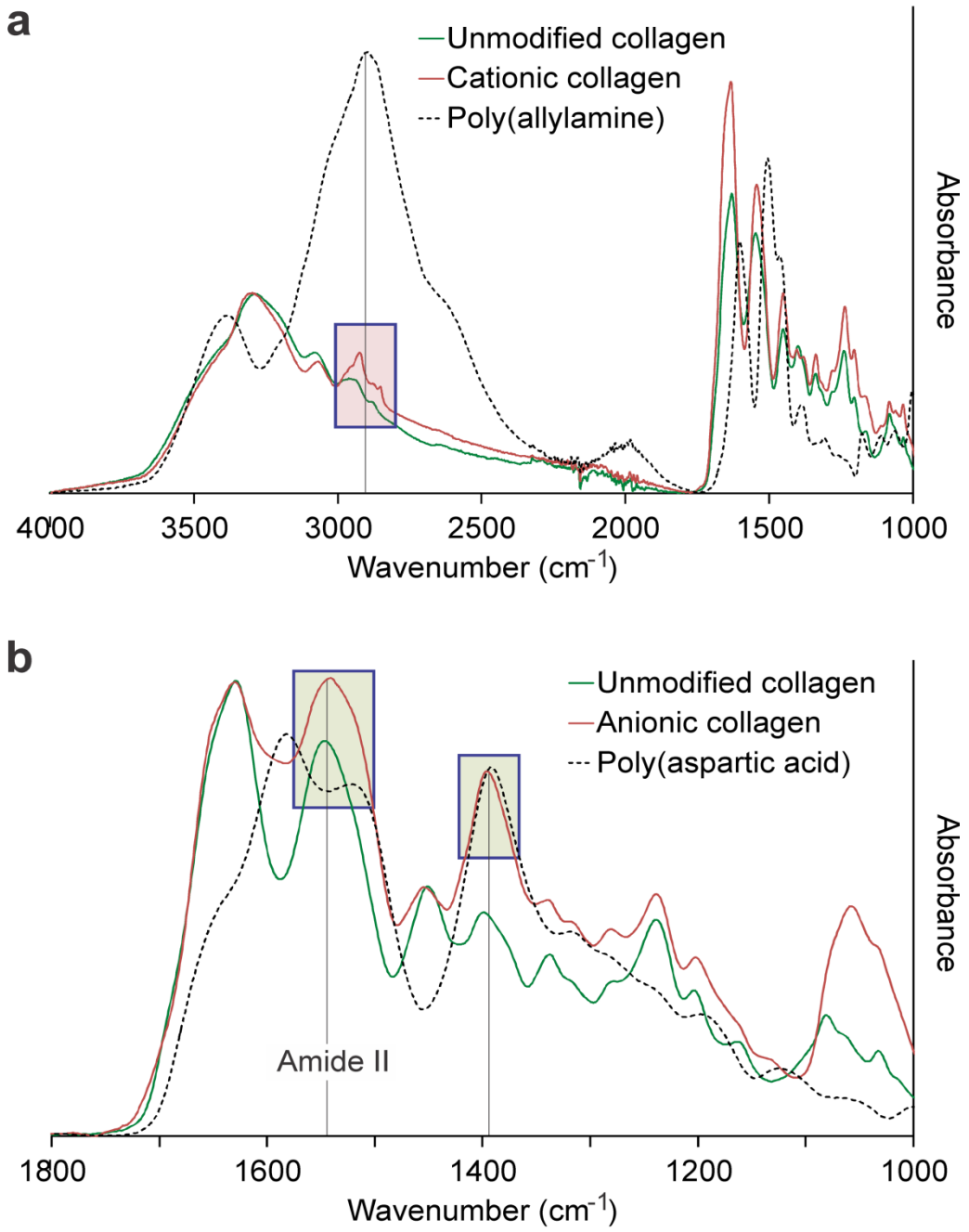


Figure SI-6 **a**) Cationic collagen: infrared spectra of PAH, unmodified and cationic collagen were used to ascertain that chemical modification has indeed occurred (box). **b**) Anionic collagen: infrared spectra of PAsp, unmodified and anionic collagen were used to validate the chemical modification process (boxes).

SI-7 PAH-ACP mineralisation of unmodified collagen (control) – 3-D collagen sponges

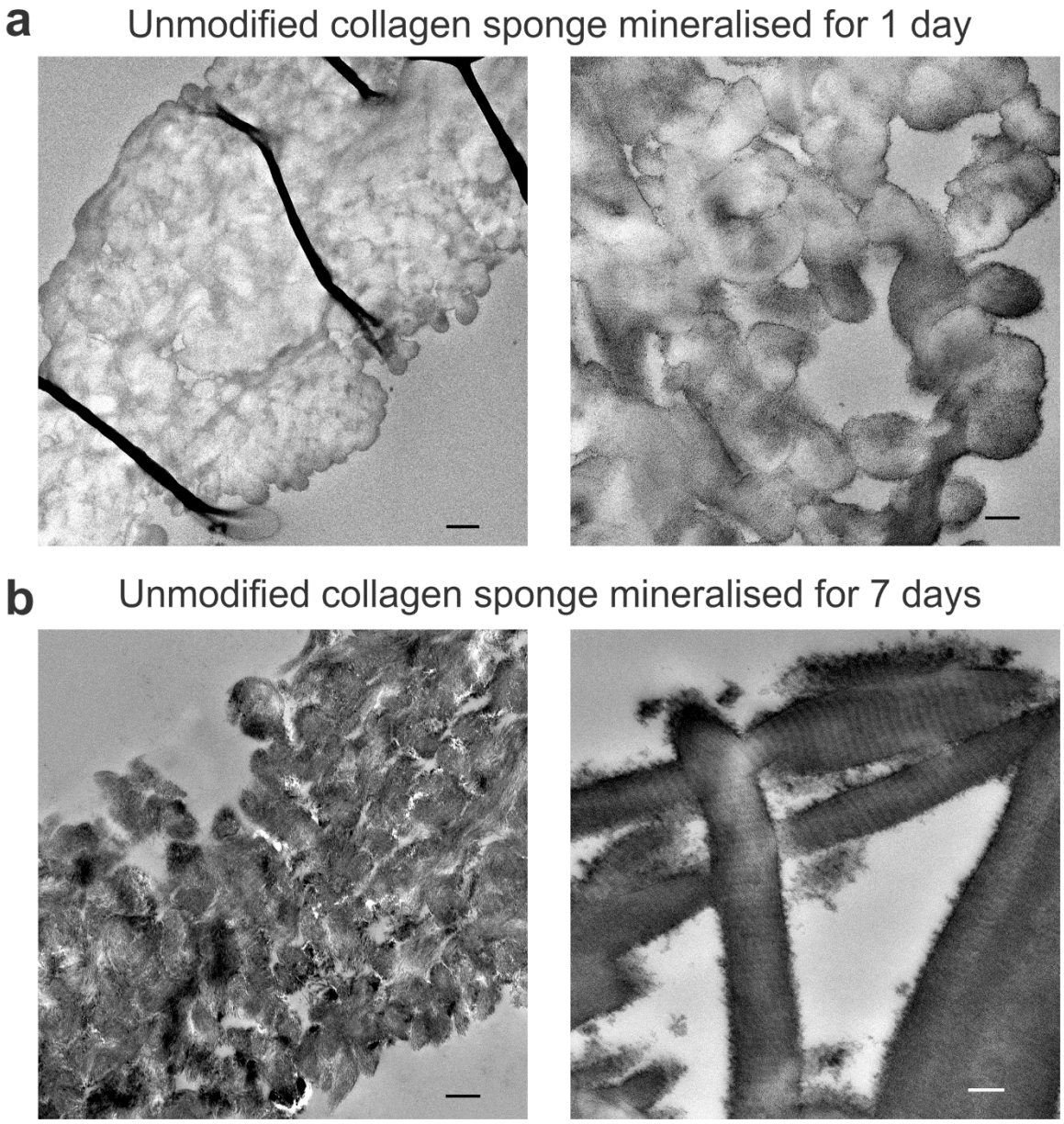


Figure SI-7 Unstained TEM images of leaflets from unmodified collagen sponges that were mineralised with PAH-ACP for a) one day and b) 7 days. Left column – low magnification images (bars: 500 nm). Right column – high magnification (bars: 200 nm).

SI-8 PAH-ACP mineralisation of cationic and anionic collagen – single layer collagen model

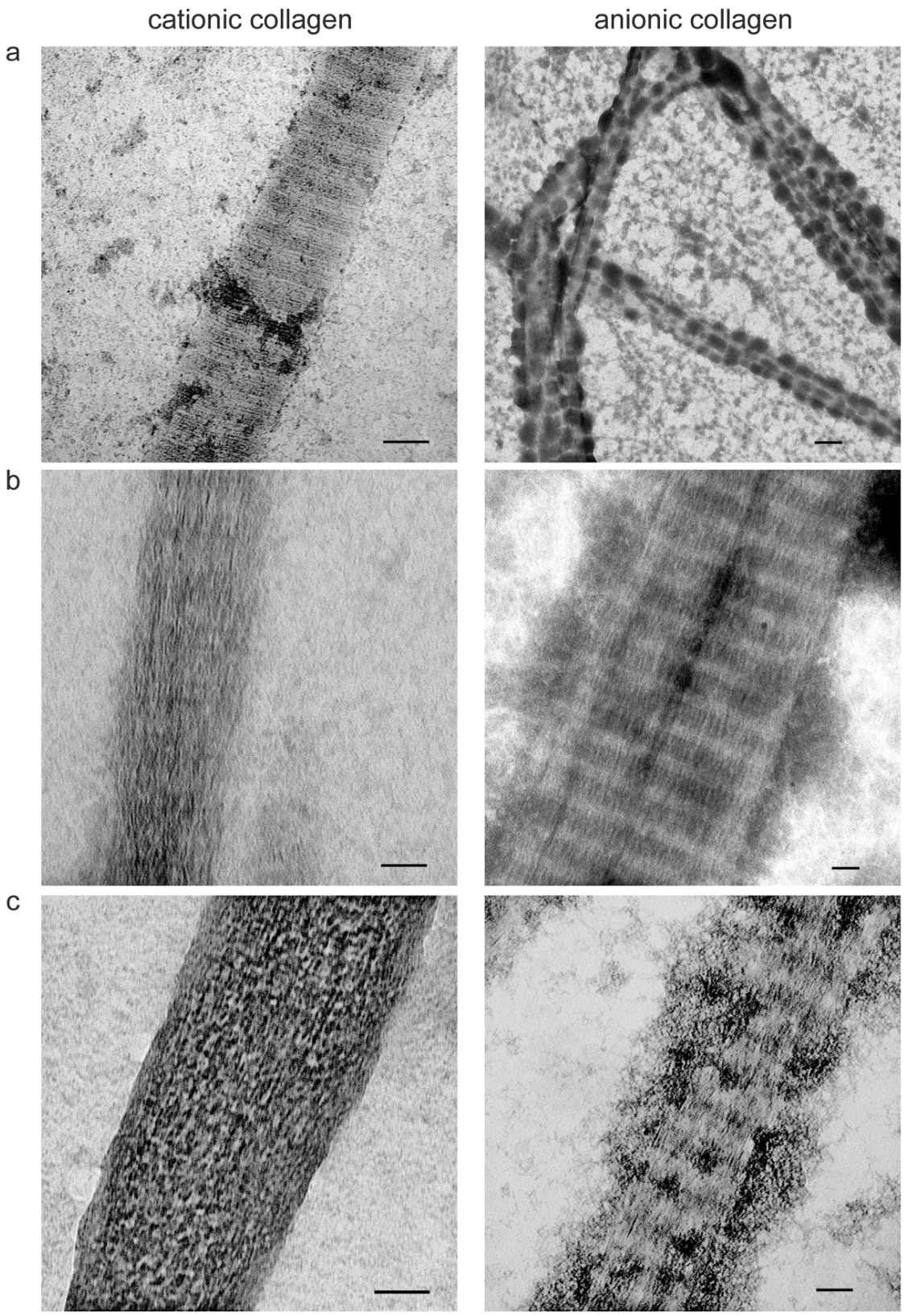


Figure SI-8 Uranyl acetate-stained TEM images of PAH-ACP mineralised reconstituted type I collagen fibrils that had been conjugated with PAH to render them cationic (left column), or conjugated with poly(aspartic acid) to render them anionic (right column), using the methods described section 4. For cationic collagen, the sequence of events is similar to unmodified collagen: **a**) infiltration of PAH-ACP with random sites of infiltration (bar: 100 nm), **b**) transformation of the intrafibrillar ACP to apatite (bar: 100 nm), eventually resulting in **c**) a mature, highly mineralised collagen fibril (bar: 50 nm). For anionic collagen, **a**) large, electron-dense particles are condensed on the fibril's surface (bar: 200 nm). They probably represent ACP coacervates following electrostatic interaction between the polycationic and polyanionic electrolytes. The latter is akin to the first step in layer-by-layer adsorption of polyelectrolyte multilayers[S35]. **b**) adsorption of these electron-dense particles to the collagen surface appears to slow down the infiltration of the CaP-PAH complexes (bar: 50 nm). **c**) in the mature stage, a mineralised fibril with less heavy intrafibrillar mineralisation is formed, while the surface coacervates eventually transformed into extrafibrillar apatite crystallites (bar: 50 nm). Information derived from the single-layer collagen models is complementary to the results obtained from 3-D models in the main text.

SI-9 PAsp-ACP mineralisation of cationic and anionic collagen – 3-D collagen sponges and single-layer collagen model

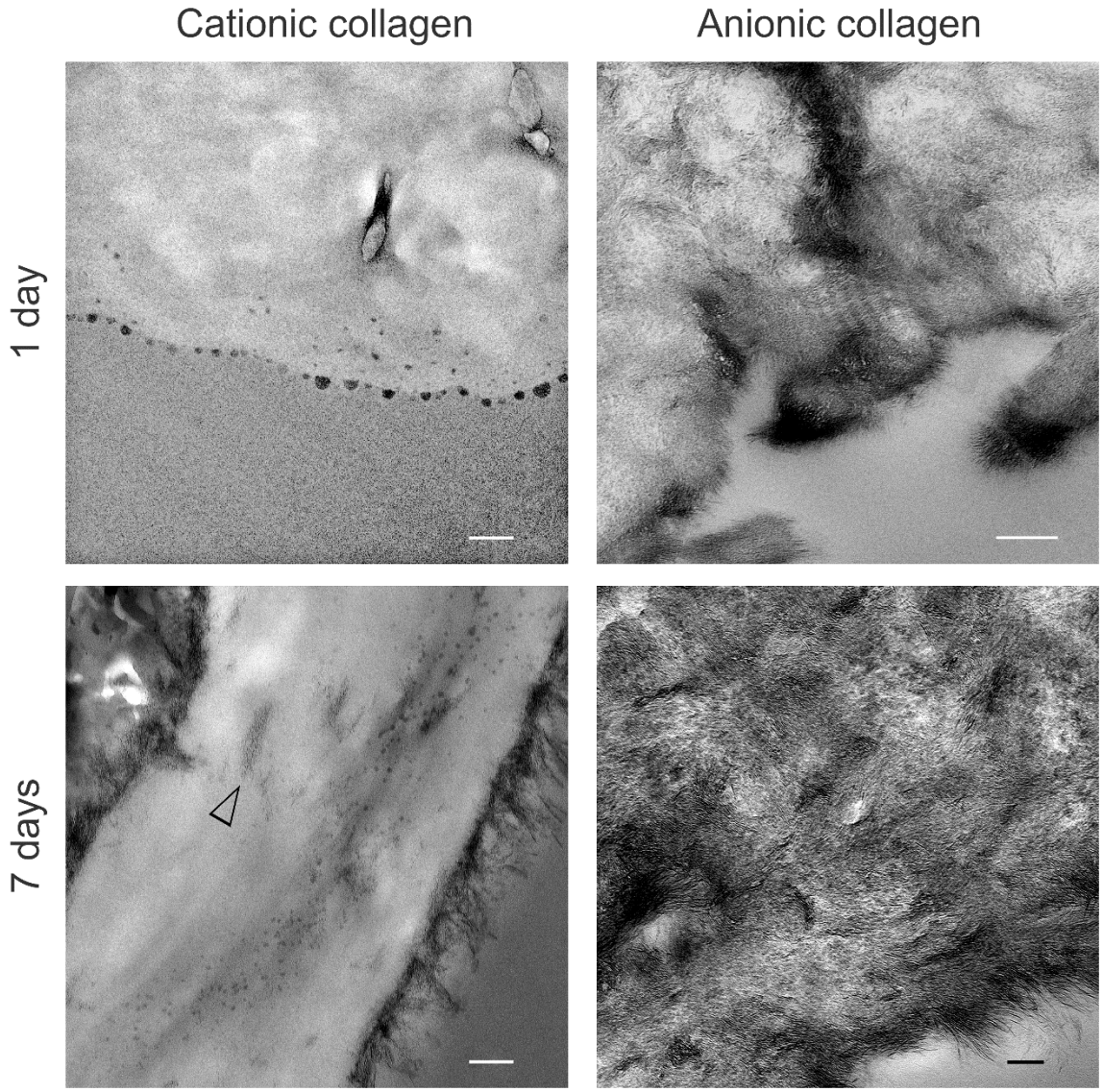


Figure SI-9A PAsp-ACP mineralisation of cationic (left column) and anionic (right column) collagen sponges. Left column: unstained TEM of collagen leaflets from cationic collagen mineralised with PAsp-ACP for 1 day and 7 days (bar: 500 nm for both images). Similar to what was found in the anionic collagen mineralised with PAH-ACP, the surface of cationic collagen is covered with electron-dense particles and intrafibrillar mineralisation is not observed on the first day. After incubation for 7 days, most of the apatite crystallites are identified on the surface of the collagen leaflets and intrafibrillar apatite deposition can hardly be observed (open arrowhead). Right column: unstained TEM of collagen leaflets from anionic collagen mineralised with PAsp-ACP for 1 day (bar: 200 nm) and 7 days (bar: 100 nm). After incubation for 1 day, extensive intrafibrillar mineralisation is observed and the banded appearance of mineralised collagen can be identified. After incubation for 7 days, the collagen fibrils are heavily mineralised and the banded appearance of mineralised collagen can no longer be observed.

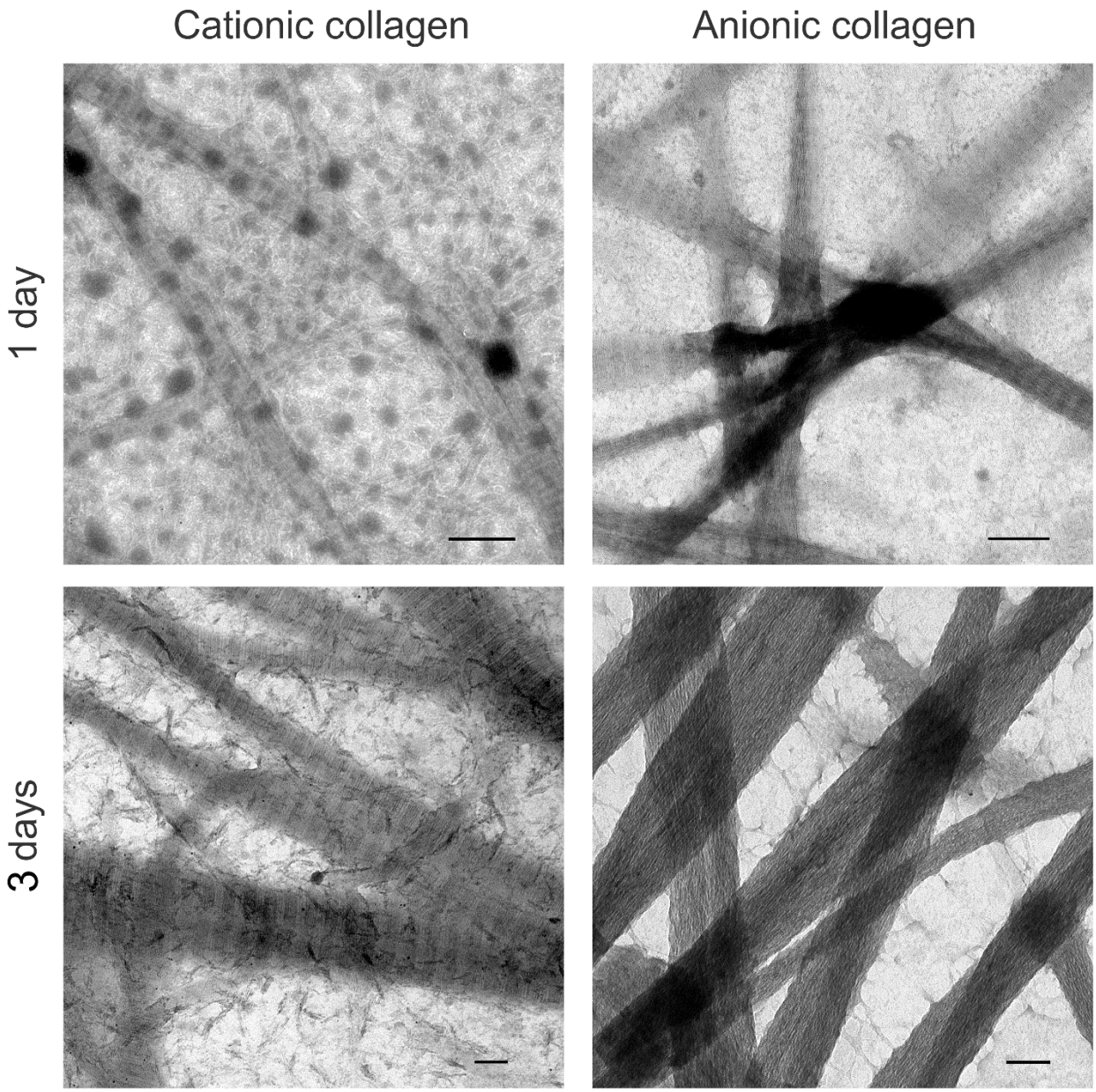


Figure SI-9B PAsp-ACP mineralisation of cationic (left column) and anionic (right column) collagen using a single-layer collagen model. Left column: Uranyl acetate-stained TEM images of cationic collagen fibrils mineralised with PAsp-ACP for 1 day (bar: 500 nm) and 3 days (bar: 100 nm). After incubation for 1 day, large, electron-dense aggregates can be identified. They probably represent coalesced, ACP following electrostatic interaction between the polycationic and polyanionic electrolytes. After incubation for 3 days, the apatite depositions are mostly extrafibrillar. Right column: Uranyl acetate-stained TEM images of anionic collagen fibrils mineralised with PAsp-ACP for 1 day (bar: 200 nm) and 7 days (bar: 100 nm). On the first day, intrafibrillar mineralisation, which initiates from the cross point of several collagen fibrils, has already commenced. On the 3rd day, most of the collagen fibrils are heavily mineralised with the intrafibrillar apatite crystallites.

SI-10 The effects of inclusion of a low molecular weight tricarboxylic acid on collagen mineralisation with poly(aspartic acid)-stabilised ACP

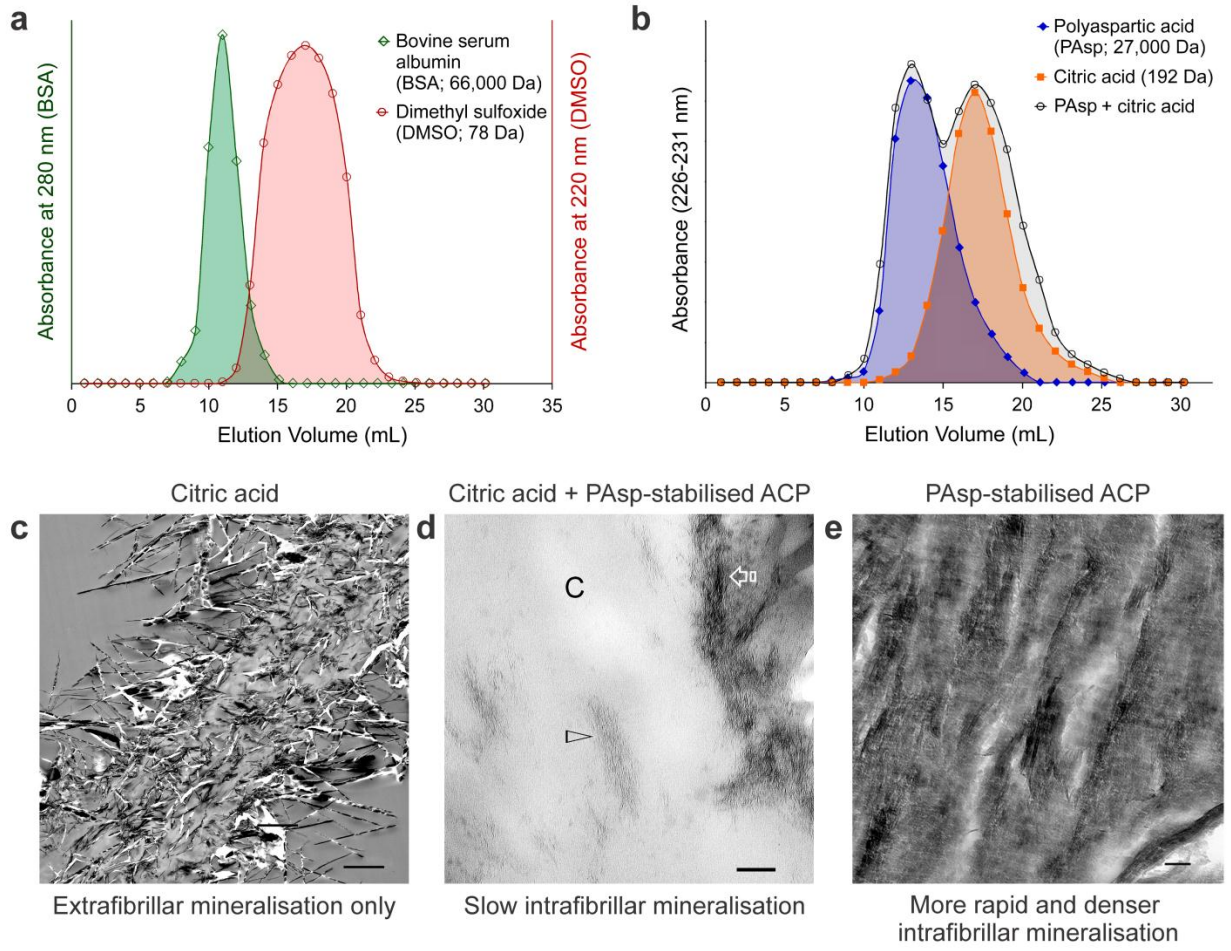


Figure SI-10 **a**) Modified size exclusion chromatography showing the elution profiles of dimethyl sulphoxide (DMSO, low molecular weight control) and bovine serum albumin (BSA, high molecular weight control) through a size exclusion column containing demineralised bone power as the stationary phase[S13,S14]. A small molecule such as DMSO (Mw 78 Da) that can penetrate every region of the stationary phase pore system "sees" a total liquid volume equal to the sum of the interparticle pore volume and the collagen intrafibrillar volume. This small molecule has more volume to traverse and will elute late. Conversely, a very large molecule such as BSA (Mw 66 kDa) that is too large to enter the collagen intrafibrillar spaces "sees" only the extrafibrillar volume and will elute earlier when this volume of mobile phase has passed through the column. **b**) Elution profiles of poly(aspartic acid) (long-chain polycarboxylic acid, Mw 27 kDa), citric acid (short chain tricarboxylic acid, Mw 192 Da), and poly(aspartic acid)/citric acid mixture. The mixture can be separated into two elution peaks that correspond well to those of the component molecules. By using this principle, molecules of different sizes will elute through the demineralised collagen stationary phase at different rates. This results in the separation of a solution of molecules based on their sizes. Subtraction of the smaller elution volume of BSA from that of DMSO gives the experimental intrafibrillar volume that cannot be reached by BSA.

This method was used with the calculated intrafibrillar volume to yield the vol% of intrafibrillar water exposed to a test molecule (Table SI-10). **c**) Unstained TEM image of a leaflet in collagen sponge that was mineralised for 7 days with citric acid (0.3 mg/mL) and the supersaturated CaP solution without the use of poly(aspartic acid) to stabilise the CaP complexes. There was no intrafibrillar mineralisation. Only large, extrafibrillar needle-shaped crystallites can be observed along the surface of the leaflet, and smaller needle-shaped crystallites within the extrafibrillar spaces between the collagen fibrils within the leaflet (bar: 2 μm). **d**) Unstained TEM image of part of a leaflet in collagen sponge that was mineralised for 7 days with citric acid (0.3 mg/mL) added to CaP mineralising solution that was stabilised with 75 μg/mL poly(aspartic acid) (bar: 100 nm). A fine deposit of apatite deposits can be seen on the surface of the collagen leaflet (open arrow). Within the collagen leaflet (C), only a meagre amount of intrafibrillar mineralisation is observed (open arrowhead). **e**) Unstained TEM image of part of a leaflet in collagen sponge that was mineralised for 7 days with CaP mineralising solution that was stabilised with 75 μg/mL poly(aspartic acid) without addition of citric acid (bar: 100 nm). A heavy intrafibrillar deposition of apatite crystallites in a highly-ordered manner resulted in the recognition of banding patterns within the mineralised fibrils.

Test molecule	Mass (Da)	Elution volume (mL)	Volume separating test molecule from DMSO (mL)	Available intrafibrillar water (mL)*	Intrafibrillar water accessible by test molecule (%)
BSA (high molecular weight control)	66,000	11.01	**17.05**-11.01 = 6.04	**6.06 mL**	[(**6.06**-6.04/**6.06**)] x 100 = 0.33
PAsp	27,000	12.97	**17.05**–12.97 = 4.08		32.67
PAH	15,000	14.05	**17.05** –14.05 = 3.00		50.50
Spermine	202.34	16.98	**17.05** –16.98 = 0.07		98.64
Citric acid	192.12	17.01	**17.05**-17.01 = 0.04		99.34
PAsp + citric acid	NA	13.01 (PAsp) 17.00 (citric acid)	**17.05**-13.01 = 4.04 **17.05**-17.00 = 0.05		33.33 99.17
PAH + spermine	NA	14.00 (PAH) 17.04 (spermine)	**17.05**-14.00 = 3.05 **17.05**-17.04 = 0.01		49.67 99.83
DMSO (low molecular weight control)	78	**17.05**	**17.05**-17.05 = 0		[(**6.06**-0)/**6.06**] x 100 = 100

* data derived from Table I in Supplementary Information-1 (Detailed materials and methods), section 5.1

Table SI-10 Determination of the vol% of collagen intrafibrillar water that is accessible to test molecules.

SI-11 Osmolality and fluid partitioning inside and outside a collagen fibril

A. Osmolality measurements

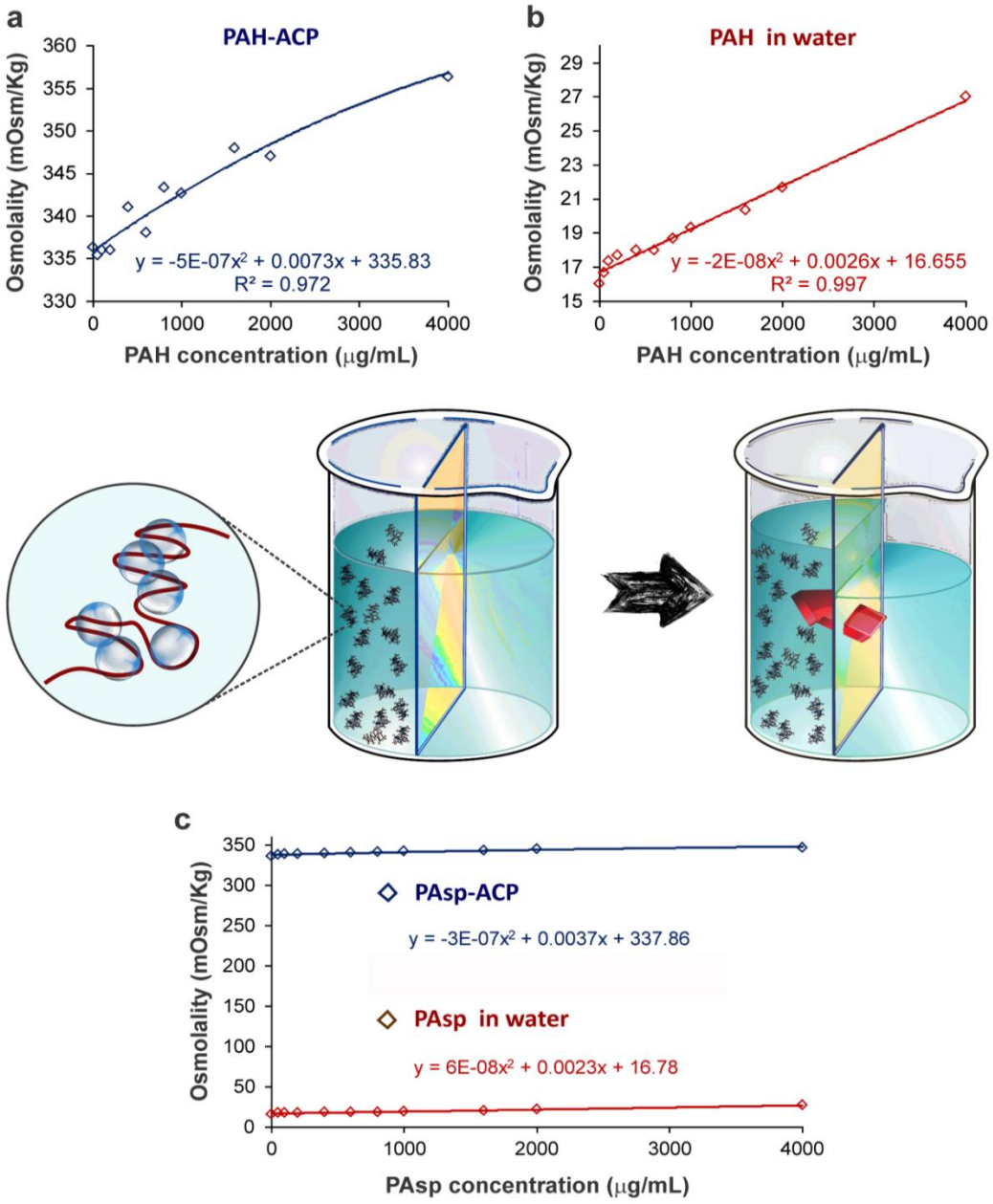


Figure SI-11A Effect of cationic or anionic polyelectrolyte concentration on the osmolality of the polyelectrolyte-stabilised ACP. Osmolality is defined as the number of osmoles per kilogramme of solvent (Osm/Kg). For biological solutions which are primarily water, this is not appreciably different from osmolarity (Osm/L). Osmolality is dependent on the number of particles present per unit mass regardless of their nature. Because of the low concentrations of solutes in physiological fluids, reporting osmolality in milliosmoles/kilogramme (mOsm/Kg) is

often more appropriate. For example, osmolarity of human plasma has a normal range of 285 mOsm/L. The relationship between the molarity of a solution and its osmolality depends on the number of particles derived from each molecule. For instance, glucose remains intact in solution and hence, a 20 mmol/L solution will have an osmolality of 20 mOsm/Kg. Conversely, ionic compounds such as KCl dissociates in aqueous solution, and may be conceptually considered as molecules each containing one potassium ion and one chloride ion. Thus, a 100 mmol/Kg KCl solution will generate an osmolality of 200 mOsm/Kg. Similarly, $MgCl_2$ dissociates into three moieties in solution, so that 200 mmol/L $MgCl_2$ will produce, after dissociation, an approximately 600 mOsm/Kg solution. Because the freezing point of water is depressed proportionally to the number of particles dissolved in the water, osmolality measurements in Figure S-12A were conducted using the principle of freezing point depression.

In "**a**", the osmolality of PAH-ACP increases quadratically with increasing PAH concentration. The osmolality of the solution is in the 300+ mOsm/kg range (higher than human plasma) because the mineralising solution used to prepare PAH-ACP was supersaturated to begin with. In "**b**", the osmolality of PAH in water also increases quadratically with increasing PAH concentration; however, the resultant osmolality is less than 1/10 of the PAH-ACP. Distilled water used to prepare the mineralising solution has a mean osmolality of 16 mOsm/Kg. For a PAH concentration of 200 μg/mL, the osmolality of PAH-ACP is 335.33 mOsm/Kg, while that of PAH in water is 16.67 mOsm/Kg (a 20-fold difference!). This means that any free PAH (i.e. not involved in stabilisation of CaP aggregates into ACP) that exists in a solution is unlikely to exert much oncotic pressure (i.e. colloidal osmotic pressure) in the presence of a selectively-permeable membrane. This is due to counterion condensation on the surface of the polyelectrolyte chains, which results in reduction of their surface charge density. In salt-free solutions of fully charged sodium poly(styrene sulfonate), for example, only 16% of all counterions contribute to the solution osmotic pressure[S36]. Likewise, only 24.5% of all counterions contribute to the solution osmotic pressure in DNA solutions[S37]. In the presence of a salt, the osmolality of the PAH-ACP solution is predominantly derived from osmotically-active counterions and salt ions[S38] (Na^+, Cl^- present in the original mineralising solution, with additional contribution from Ca^{2+} and HPO_4^- ions that may not have condensed into CaP aggregates). The same scenario applies to PAsp-stabilised ACP, as shown in "**c**", which combines the osmolality of PAsp-ACP and PAsp in water into a single chart.

B. Gibbs-Donnan equilibrium

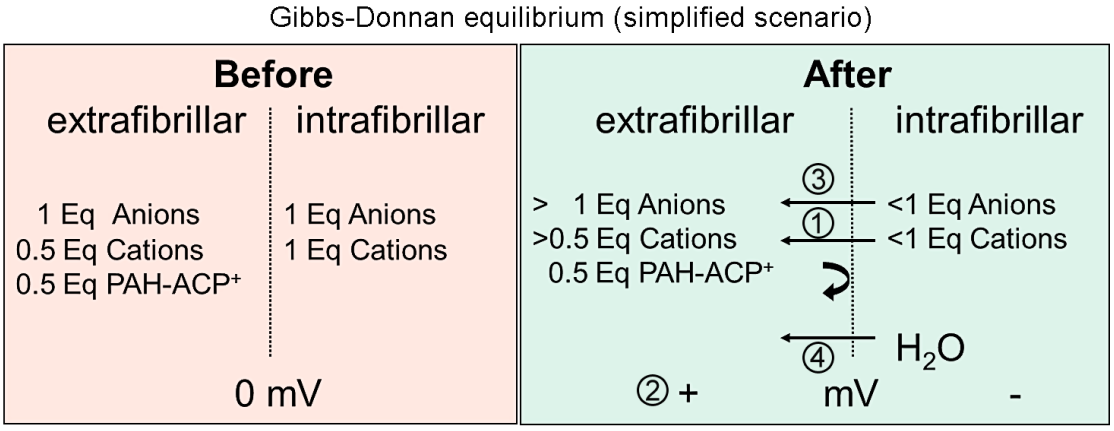


Figure SI-11B The capability of collagen fibrils to function as selectively-permeable membranes has previously been established from our size exclusion experiments (Figure 4 and Section SI-10). The long range interaction of PAH-ACP with collagen that causes ingress of fluid-like CaP aggregates into the intrafibrillar water compartments of collagen fibrils may be explained by the Gibbs-Donnan effect[S39], which involves the establishment of both electrical neutrality and osmotic equilibrium across a selectively-permeable membrane. Positively-charged electrolytes partially impermeable to the intrafibrillar fluid compartments of the collagen fibril will attract negatively-charged ions and repel positively-charged ions. The passive distribution of cations is altered to preserve electroneutrality in the extrafibrillar and intrafibrillar compartments (left). The diffusible cation concentration is lower in the extrafibrillar compartment containing PAH-ACP with positively zeta potential. Thus, cations such as Na^+ (counterions derived from the mineralising solution) and Ca^{2+} (free ions that have not aggregated into CaP clusters) will move out of the intrafibrillar compartment into the extrafibrillar compartment down its concentration gradient (right-1). The excess positive charges render the extrafibrillar compartment positive in terms of electrochemical potential (right-2). In order to establish electrochemical equilibrium across the selectively-permeable membrane, anions such as Cl^- and HPO_4^- will have to move out of the intrafibrillar compartment (right-3). The consequence of establishing Gibbs-Donnan equilibrium is that there are more osmotically active species in the extrafibrillar compartment than the intrafibrillar compartment at equilibrium. The high osmolality of the PAH-ACP solution is supported by data in Figure S-11A. As a result of the Gibbs-Donnan effect, the total osmolar concentration is higher in the extrafibrillar compartment. This extra osmotic force from diffusible ions is added to the (minor) osmotic forces exerted by the polycationic PAH molecules (Figure S-11A), resulting in water moving out of the intrafibrillar compartment to establish osmotic equilibrium (right-4). The net result of the Gibbs-Donnan effect is that more water moves into the extrafibrillar compartment than would be predicted on the basis of oncotic pressure (colloid osmotic pressure) of the PAH molecules alone. To balance the unstable negative pressure generated within the collagen fibril, fluid-like CaP aggregates (dissociated from the PAH by short range interaction with positive charges along the collagen fibril) will enter the intrafibrillar water compartment of the collagen fibril to maintain the intrafibrillar volume. A similar effect may be used to explain the ingress of CaP aggregates when collagen fibrils are immersed in PAsp-stabilised mineralising solution.

SI-12 Molecular dynamics (MD) simulations of collagen molecules in presence of polyelectrolytes

[Please refer to Figure 5 in the main text]

Hydrated collagen structure

The structure of equilibrated collagen (Fig.5a) was verified by analysing the hydration layers (Fig.5b) and charge density along the c axis (Fig.5c). Figure 5b shows the three hydration layers at minimum distance r from the surface of a collagen triple helix, which correspond to the tightly-bound, loosely-bound and free water molecules on the collagen structures. The 1st, 2nd and 3rd hydration layers were specified as (0-2.1 nm), (0.21-0.425nm) and (0.425-0.625 nm) from the surface of a collagen triple helix, based on curvature and inflection points. The 1st hydration layer, which corresponds to Ramachandran water bridges[S40], contains 0.06 g water/g collagen molecules and the summation of the three hydration layers is 0.80 g water/g collagen[S21, S22]. The collagen structure together with these three hydration regions surrounding each collagen helix are considered as intrafibrillar regions. The collagen structure has a dimension of approximately 12.06 nm, and consists of the fifteen collagen molecules arranged into five triple helices (~1.2 nm diameter) in the a axis. Other regions beyond those regions were classified as extrafibrillar regions, as shown in Fig.5d. Charge density along the c axis was extracted from the structures of 18-20 ns simulation region, and then the average value was plotted in Fig.5c. The a- and c-band regions showed positive charge densities as reported in other simulation and experimental structures[S41]. To investigate this structural variation specifically, Solvent Accessible Surface Area (SASA) was calculated based on the double cubic lattice method[S26]. To account for the free surface area of water molecules, a probe with a radius of 1.4 Å was utilised, which corresponds to the hydrodynamic radius of a water molecule. To verify that the collagen structure was fully relaxed and reached equilibrium before introducing polyelectrolytes, SASA of the collagen structures were plotted for 40 ns (Fig.SI-12A). The SASA of the collagen showed increasing tendency at initial simulation time and reached equilibrium at 20 ns.

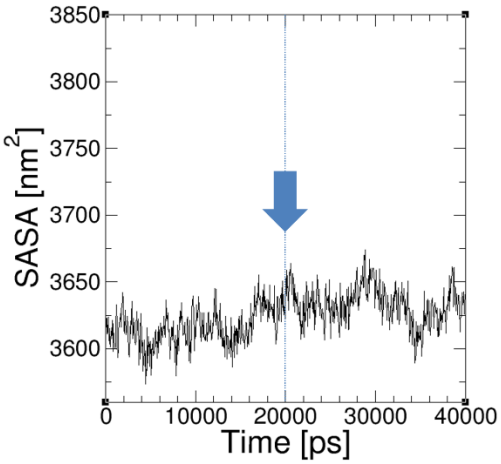


Figure SI-12A Solvent Accessible Surface Area (SASA) of collagen without polyelectrolytes. Arrow indicates the collagen structure was fully relaxed and reached equilibrium at 20 ns.

Transport of ions and water molecules through collagen fibrils

The number of Na ions (Fig. 5e), Cl ions (Fig.5f) and water molecules (Fig.5g) in the intrafibrillar region varied over simulation time, in the presence of polycations or polyanions in the extrafibrillar region. In both systems, the number of Na and Cl ions decreased as simulation time increased, which indicates diffusion of both ions from the intrafibrillar region to the extrafibrillar region. In the polycationic system, the number of Na ions in the intrafibrillar regions reduced from 207 to 196 (5.314% decrease) and Cl ions reduced from 328 to 297 (9.45% decrease). In the polyanionic system, the number of Na ions decreased from 273 to 223 (18.32% decrease) and Cl ions decreased from 362 to 308 (14.92% decrease). Movement of ions with positive and negative charges from the intrafibrillar region to the extrafibrillar region occurred irrespective of the charges present in the polyelectrolyte particles.

The water molecules in intrafibrillar regions also decreased in both systems (Fig.5g), following the migration of Na and Cl ions into the extrafibrillar region. The specific intrafibrillar regions from which the water molecules were removed were investigated by counting the number of water molecules in the gap and overlap regions of the collagen structures from 0 ns to 20 ns. In the anionic system, 69.43% of the diffused water molecules were derived from the gap region (c = 46.925 nm) while 30.58% came from the overlap region (c = 26.865nm). In the cationic system, 70.03% and 29.97% water molecules diffused from the gap and overlap regions, respectively. Most of the water molecules were removed from gap regions, where there are more loosely-bound and free water molecules filling the intrafibrillar spaces (Fig.5a). In the gap regions, water molecules were removed evenly from entire c axis instead of from the "a band" region alone. It has been reported that gap region is less resistant to external forces by experiments and simulations[S42]. Thus, in the presence of strong polyelectrolytes near the surface of collagen structures, most of the water molecules diffuse out of the gap regions.

To examine effect of migration of the monovalent ions and intrafibrillar water on the dimensional changes in collagen structures, Root Mean Square Deviations (RMSD) of collagen molecules was analysed. As shown in Fig.5h, RMSD of collagen backbones exhibited steep increases with increases in simulation time, indicating the collagen molecules undergo significantly large structural variations during osmosis. The SASA of collagen structures in the presence of polyelectrolytes was also investigated after equilibration (Fig. 5i). After ions and water molecules diffused out of the intrafibrillar water compartments into the extrafibrillar regions (Fig.5, e-g), the SASA of collagen molecules decreased from 3844.64 nm^2 to 3799.86 nm^2 (1.16% decrease) in the polyanionic system, and from 3783.54 nm^2 to 3746.11 nm^2 (0.99% decrease) in the polycationic system, over a period of 20 ns. The results indicate that the dimensional changes occurred due to contraction of the collagen structures. Specifically, sharp contraction occurred at 5 ns in the polyanionic system and at 11 ns in polycationic system (Fig.5i), while g water/g collagen in the intrafibrillar regions showed an identical decreasing tendency (Fig.5g). Combined with the observation that water molecules mainly diffuse from the gap region, it may be inferred that SASA mainly decreases in the gap regions where more spaces are initially filled by water molecules.

The stability of collagen structures in presence of osmosis

To investigate the effect of water diffusion from the intrafibrillar regions on the structural stability of the collagen structures, the amount of tightly-bound water molecules (in g water/g collagen) in the 1st hydration layer within 0.21 nm of the collagen triple helix were plotted as a

function of simulation time, as illustrated in Fig. SI-12B. The 1st hydration layer of tightly-bound water corresponds to the Ramachandran water bridge, which is the key for maintaining the structural stability of the collagen structure via interchain hydrogen bonds. It is well known that disruption of Ramachandran water bridges results in the loss of integrity of the collagen triple helix arrangement[S43]. In both the polycationic and polyanionic systems, g water/g collagen in the first tightly-bound water layer showed minimal changes over time. We further analysed hydrogen bond networks between collagen and water molecules, which is a critical factor in maintaining the mechanical stability of the collagen triple helix. Figure SI-12b shows that the number of hydrogen bonds remained stable over time, which has a similar tendency as the water content in 1st layer of tightly-bound water. From this observation, the stability of the collagen structure would be maintained in presence of the polyelectrolytes. Combining the observations on water distributions, we opine that the transport of water molecules occurs mainly in the loosely-bound and free water regions (i.e. the 2nd and 3rd hydration layers) in the spacing of the gap regions because water movement from the 1st layer of tightly-bound water would have disrupted the integrity of the collagen helical arrangement. Conversely, the 2nd and 3rd hydration layers would provide diffusion channels for movement of ions and water molecules out of the collagen fibrils.

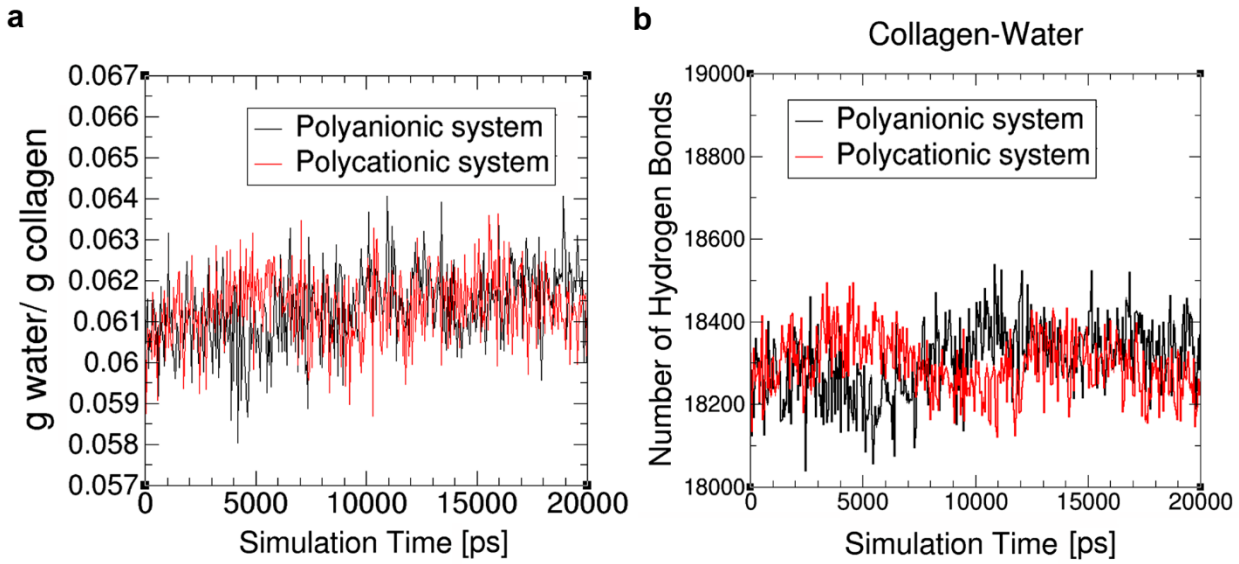


Figure SI-12B a) Variations of the 1st tightly-bound water layers over simulation time, and **b)** number of hydrogen bonds between collagen and water in the polyanionic (black) and polycationic (red) systems.

SI-13 Solution-induced dimensional changes in hydrated single-layer collagen

Solvent-induced dimensional changes of the fibrillary collagen matrices had been reported in 2001[S44]. The re-expansion of an air-dried collagen matrix and the shrinkage of moist, hydrated collagen matrix are influenced by polar solvents with different solubility parameters for H-bonding (δh). Water-free polar solvents of low hydrogen bonding ability do not re-expand dried, collapsed collagen fibrils and that they cause shrinkage of moist collagen matrices. Solvents with higher H-bonding capacities (methanol, ethanol, ethylene glycol, formamide, and water) re-expand the air-dried collagen fibrils in proportion to their solubility parameters for H-bonding (δh). Although that classic study provides strong support for the existence of water flow within a bulk fibrillar collagen scaffold under different solvation conditions, it is difficult to distinguish between the contribution proportion of intrafibrillar and extrafibrillar water flow to the total dimensional changes. A more recent study showed that stable atomic force microscopy (AFM) imaging of the same collagen fibril may be achieved in the presence of different liquids[S45]. In going from air to buffer, the collagen fibrils swell ~2-fold. In 100% ethanol, which is hygroscopic, the fibrillar diameters are reduced to values similar to those in air.

To validate the molecular dynamics simulation results, AFM was used to detect the effects of polyelectrolyte and polymer stabilised-ACP on the dimensional (diameter) changes of unmineralised collagen fibrils (Fig.SI-13A). Representative amplitude and height images of collagen fibrils under different conditions are shown in Fig.SI-13B. The AFM results indicate that the diameter of dried collagen fibril is about 204.86±17.35 nm, which is larger than those shown in TEM. This may be caused by the different hydrophilicity and charge between the surface of a TEM grid and a mica disc. When the collagen fibrils were soaked in HEPES buffer for about 1 h, the diameter of the collagen fibril was significantly increased to 319.71±29.88 nm ($P < 0.05$). Addition of more NaCl to the HEPES buffer, which resulted in same ionic strength as the mineralising medium without polyelectrolyte, did not induce further diameter change in the collagen fibril (309.43±30.61 nm) ($P > 0.05$). However, addition of PAH to HEPES buffer led to the shrinkage of the collagen fibrils (260.86±30.13 nm; $P < 0.05$). Addition of more NaCl to the polyelectrolyte-containing HEPES buffer (which resulted in same ionic strength as the polyelectrolyte-containing mineralising medium) does not induce further diameter change of the collagen fibril (254.29±23.42 nm; $P > 0.05$). There was no significant change in the diameter of the collagen fibrils after they were further incubated in HEPES buffer containing 200 μg/mL PAH and 19.8 mM NaCl for 1 day (253.43±25.26 nm; $P>0.05$). When the incubation buffer was changed to HEPES buffer containing 200 μg/mL PAH, 4.5 mM $CaCl_2$•$2H_2O$ and 2.1 mM K_2HPO_4 (PAH-ACP), the diameter increase of the collagen fibrils in the first hour (264.57±21.90 nm) was not statistically significant from those immersed in solutions without the ACP precursors ($P > 0.05$). The increase in collagen diameter was statistically significant ($P < 0.05$) after 1 day of incubation in PAH-ACP containing buffer (304.29±27.8 nm), which was not significantly different from those immersed in HEPES buffer alone or HEPES containing NaCl ($P > 0.05$).

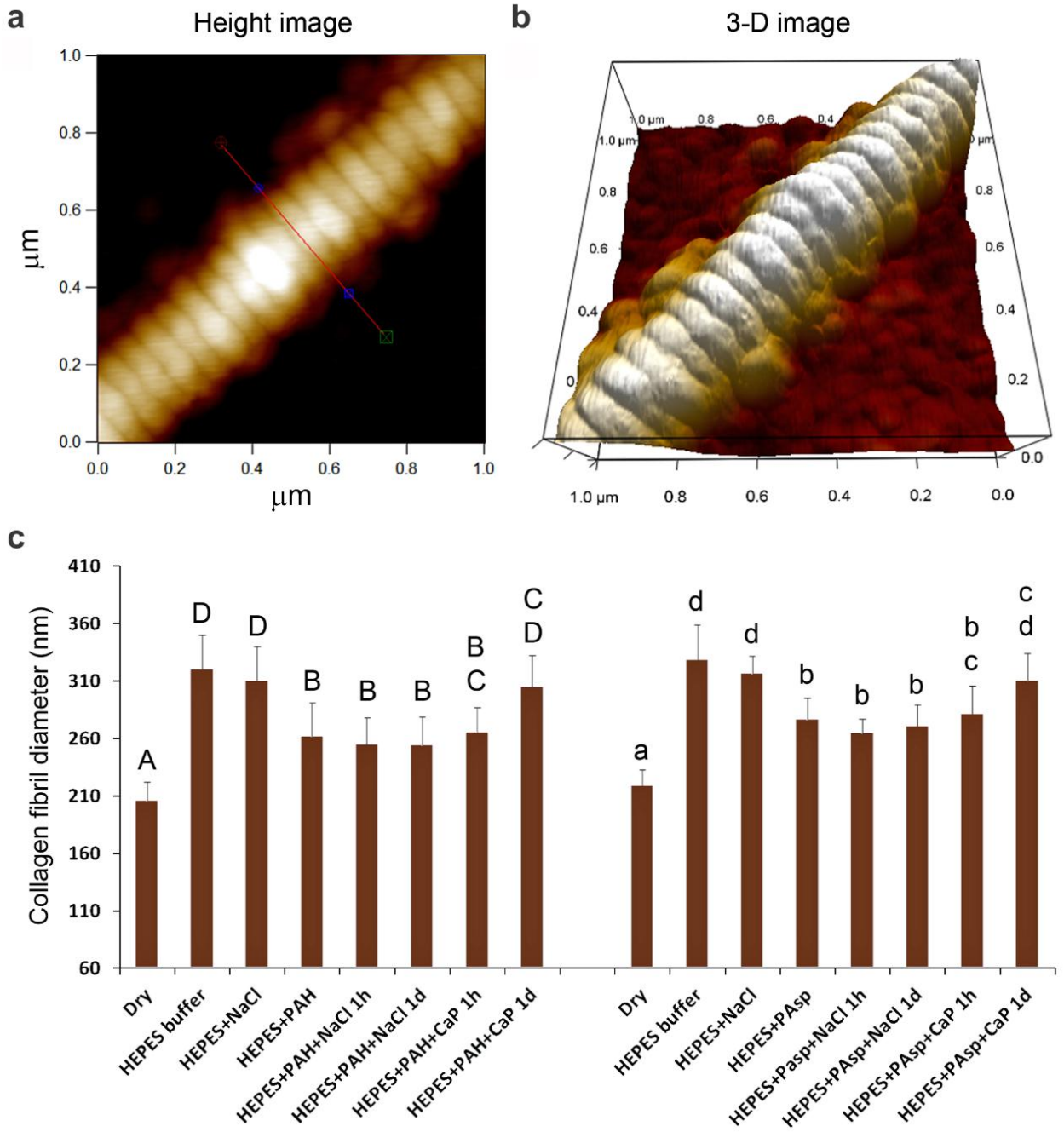


Figure SI-13A a) Representative AFM height image of a reconstituted collagen fibril immersed in HEPES buffer for 1 hour. **b)** Three-dimensional image of the surface topography of the same collagen fibril. **c)** Changes in diameter of the collagen fibrils after immersion in different solutions, as measured by the line section labelled in "**a**". The distance between the two blue markers indicates the diameter of the collagen fibril. For PAH (left), groups labelled with the same upper-case letters are not significantly different ($P > 0.05$). For PAsp (right), groups labelled with the same lower-case letters are not significantly different ($P > 0.05$).

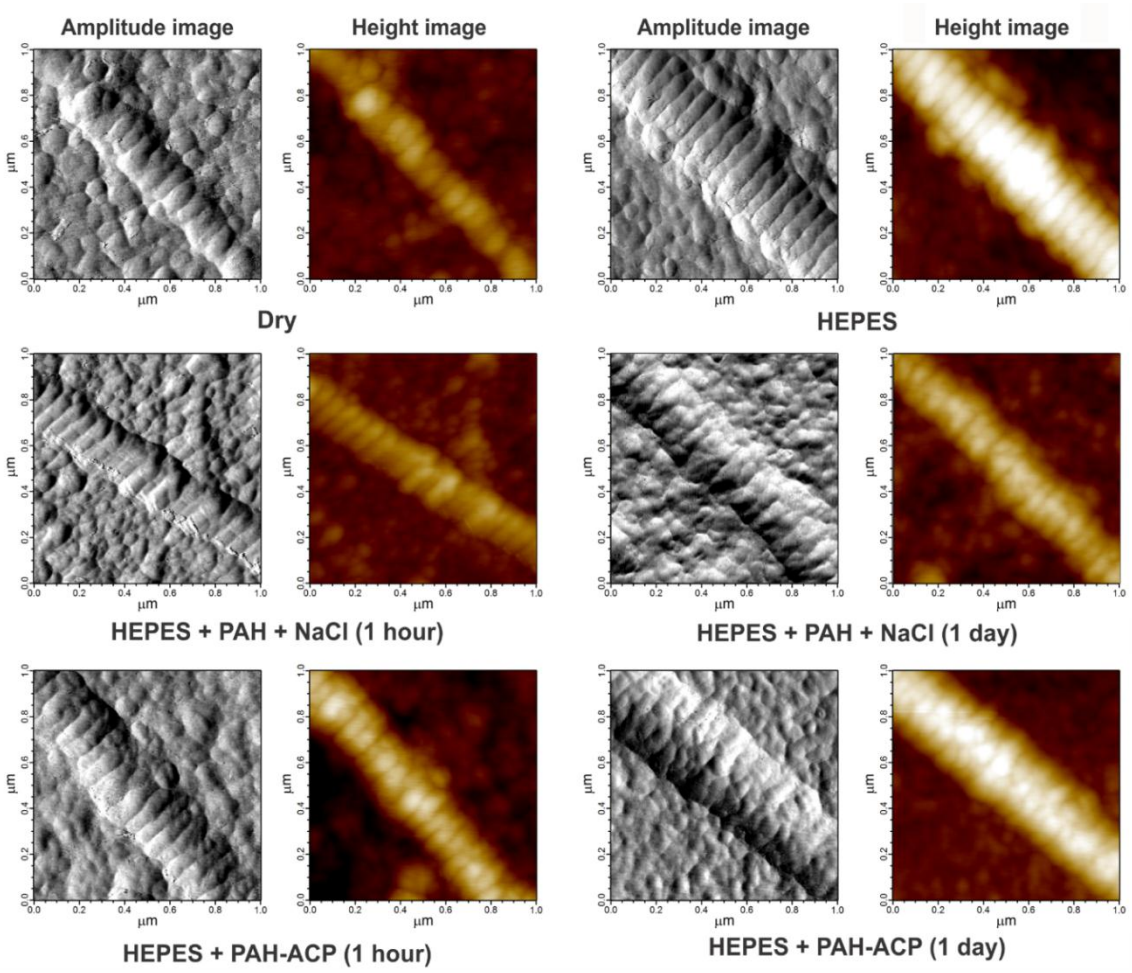


Figure SI-13B Representative amplitude and height images of collagen fibrils under different conditions (PAH groups).

The results of dimensional changes of collagen fibril under different solvation conditions support the mineralisation mechanism proposed in the present study. When dried collagen fibril was soaked in HEPES buffer, the small size and strong H-bonding capacity, 42.3 $(J/cm^3)^{1/2}$, of water break the weak forces between collagen polypeptides and cause expansion of the collapsed collagen fibrils. Addition of a small amount of salt does not influence the diameter of the hydrated collagen fibrils, because the ions derived from the salt can move freely in both the intrafibrillar and extrafibrillar spaces. Because of the high molecular weight and large charge density of the polyelectrolyte (PAH or PAsp), and the semi-permeable feature of collagen fibril *per se*, addition of the polyelectrolyte to HEPES buffer causes free water to flow from the intrafibrillar space into the extrafibrillar space. This, in turn, results in reduction in the diameter of the collagen fibril. Because of the viscoelastic properties of hydrated collagen fibrils[S46,S47], the contracted fibrils put the remaining intrafibrillar water under tension (i.e. below ambient pressure), which, in turn, lowers the vapour pressure of the intrafibrillar compartments. The PAH-ACP or PAsp-ACP precursors in the vicinity of the collagen fibrils are drawn in from the extrafibrillar milieu, down a pressure gradient, to restore the intrafibrillar volume.

SI-14 MD simulations of infiltration of CaP mineralisation precursors into the intrafibrillar regions of the contracted collagen structures

[Please refer to Figure 6 in the main text]

To validate the hypothesis that CaP precursors infiltrate the intrafibrillar compartments to compensate for the contraction of collagen structures, infiltration of CaP particles into the contracted collagen fibrils was investigated using molecular dynamics simulations. CaP mineralisation precursors (simplified as Ca ions) with +2 charges were introduced to the contracted collagen structures induced by the polycationic or polyanionic electrolyte. Additional Cl ions were added to neutralise the system for molecular simulations. In both systems, in presence of polycationic or polyanionic electrolyte, the number of Na, Cl ions and water molecules decreased in the contracted collagen structures (Fig.6a-c) while the number of Ca ions in intrafibrillar regions of the contracted collagen structures increased over simulation time (Fig.6d). During Ca ions infiltration, SASA of the contracted collagen structure in presence of polycationic agents, which was 3746.11 nm^2 without Ca ions increased to 3762.21 nm^2 (0.43% increase), while SASA in presence of polyanionic agents increased from 3799.86 nm^2 to 3874.33 nm^2 (1.96 % increase) (Fig.SI-14a). The result infers that re-swelling of the contracted collagen structures causes the introduction of CaP precursors into the intrafibrillar regions. Structural swelling was significantly larger in polyanionic system than polycationic system thus the water molecules are diffusing into the intrafibrillar region in polyanionic system while they are diffusing out to extrafibrillar region in polycationic system (Fig.6c).

The number of Ca ions surrounding the collagen structures increased continuously mainly in the regions of the 2nd water layers (0.21-0.42 nm). Conversely, the number of Ca ions decreased in the regions of the 3rd water layer (0.42-0.625 nm) during the 10 ns simulation time in presence of the polyanionic or the polycationic electrolytes (Fig.SI-14, b and c). No Ca ions were found in the 1st water layer over the entire simulation time, the latter being occupied by double water bridges and the sidechains of the collagen molecules[S48]. The recent molecular dynamics simulation study performed by Xu *et al.*[S49] showed that charged amino acid side chains are oriented toward the gap zones of the collagen fibril. Such an orientation contributes significantly to collagen acting as a template for the nucleation of mineralisation precursors. Combining with previous observations, the present molecular simulation results indicate that re-swelling of the flexible collagen structures introduces CaP precursors into the side chains of the collagen structures, which would initiate collagen mineralisation without affecting the Ramachandran water bridges in the 1st bound water layers. There was little difference in the Ca ions distribution in the polyanionic system and the polycationic system (Fig.SI-14, b and c). The number of intrafibrillar Ca ions ($r < 0.625$nm) present during the 7.5-10 ns simulation time are similar in both the polyanionic (204.2157) and the polycationic systems (209.8039) (Fig.6d). Therefore, although oppositely-charged electrolytes affect the degree of swelling of collagen, the effect is mainly on water adoption rather than the infiltration of CaP precursors.

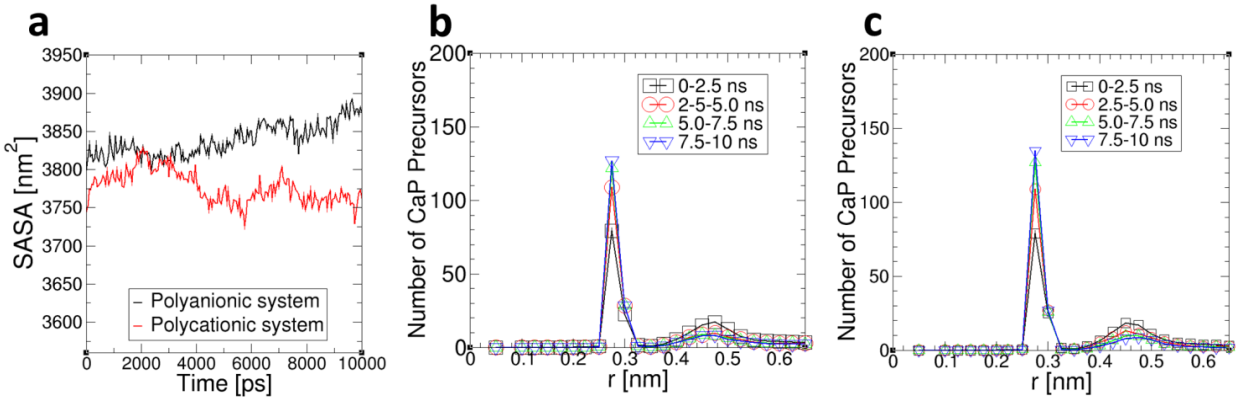


Figure SI-14 a) SASA of collagen structures when CaP precursors (simplified as Ca ions) are diffusing into the intrafibrillar regions in both systems. Distribution of Ca ions surrounding collagen structures at distance r during the four simulation periods (0-2.5 ns; 2.5-5.0 ns, 5.0-7.5 ns and 7.5-10 ns) in **b)** polyanionic system and **c)** polycationic system.

SI-15 References quoted in Supplementary Information

S1 Grassucci, R. A., Taylor, D. J. & Frank, J. Preparation of macromolecular complexes for cryo-electron microscopy. *Nat. Protoc.* **2,** 3239-3246 (2007).

S2 Ress, D., Harlow, M. L., Schwarz, M., Marshall, R. M. & McMahan, U. J. Automatic acquisition of fiducial markers and alignment of images in tilt series for electron tomography. *J. Electron. Microsc. (Tokyo)* **48,** 277-287 (1999).

S3 Zheng, S. Q. *et al*. UCSF tomography: an integrated software suite for real-time electron microscopic tomographic data collection, alignment, and reconstruction. *J. Struct. Biol.* **57,** 138-147 (2007).

S4 Kremer, J. R., Mastronarde, D. N. & McIntosh J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116,** 71-76 (1996).

S5 Chen, L., Jacquet, R., Lowder, E. & Landis, W. J. Refinement of collagen-mineral interaction: a possible role for osteocalcin in apatite crystal nucleation, growth and development. *Bone* **71,** 7-16 (2015).

S6 Wang, J. *et al*. Evaluation of collagen and methylated collagen as gene carriers. *Int. J. Pharm.* **279,** 115-126 (2004).

S7 Ma, L., Gao, C., Mao, Z., Zhou, J. & Shen, J. Enhanced biological stability of collagen porous scaffolds by using amino acids as novel cross-linking bridges. *Biomaterials* **25,** 2997-3004 (2004).

S8 Green, R. W., Ang, K. P. & Lam, L. C. Acetylation of collagen. *Biochem. J.* **54,** 181-187 (1953).

S9 Bet, M. R., Goissis, G. & Lacerda, C. A. Characterization of polyanionic collagen prepared by selective hydrolysis of asparagine and glutamine carboxyamide side chains. *Biomacromolecules* **2,** 1074-1079 (2001).

S10 Pereira, M. & Lai, E. P. Capillary electrophoresis for the characterization of quantum dots after non-selective or selective bioconjugation with antibodies for immunoassay. *J. Nanobiotechnology* **6,** 10 (2008).

S11 Grover, C. N. *et al*. Crosslinking and composition influence the surface properties, mechanical stiffness and cell reactivity of collagen-based films. *Acta. Biomater.* **8,** 3080-3090 (2012).

S12 ASTM Standard C33 (ASTM D1926-00). *Standard test methods for carboxyl content of cellulose*. (ASTM International, West Conshohocken, 2011).

S13 Toroian, D., Lim, J. E. & Price, P. A. The size exclusion characteristics of type I collagen: implications for the role of noncollagenous bone constituents in mineralization. *J. Biol. Chem.* **282,** 22437-22447 (2007).

S14 Takahashi, M. *et al*. The importance of size-exclusion characteristics of type I collagen in bonding to dentin matrices. *Acta. Biomater*. **9,** 9522-9528 (2013).

S15 Sinha, N. & Smith-Gill, S. J. Electrostatics in protein binding and function. *Curr. Protein Pept. Sci.* **3,** 601-614 (2002).

S16 Blackburn, R. S. *et al*. Sorption of chlorhexidine on cellulose: mechanism of binding and molecular recognition. *J. Phys. Chem. B* **111,** 8775-8784 (2007).

S17 Orgel, J. P., Irving, T. C., Miller, A. & Wess, T. J. Microfibrillar structure of type I collagen in situ. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 9001-9005 (2006).

S18 Nair, A. K., Gautieri, A. Chang, S-W. & Buehler, M. J. Molecular mechanics of mineralized collagen fibrils in bone. *Nat. Commun.* **4,** 1724 (2013).

S19 Gautieri, A., Vesentini, S., Redaelli, A. & Buehler, M. J. Hierarchical structure and nanomechanics of collagen microfibrils from the atomistic scale up. *Nano Lett*. **11,** 757-766 (2011).

S20 Fiser, A., Do, R. K. & Sali, A. Modeling of loops in protein structures. *Protein Sci.* **9,** 1753-1773 (2000).

S21 Fullerton, G. D. & Rahal, A. Collagen structure: The molecular source of the tendon magic angle effect. *J. Magn. Reson. Imaging* **25,** 345-361 (2007).

S22 Fullerton, G. D. & Amurao, M. R. Evidence that collagen and tendon have monolayer water coverage in the native state. *Cell Biol. Int.* **30,** 56-65 (2006).

S23 Van Der Spoel, D. *et al*. GROMACS: Fast, flexible, and free. J. *Comput. Chem*. **26,** 1701-1718 (2005).

S24 Lindorff-Larsen, K. *et al*. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78,** 1950-1958 (2010).

S25 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. **79,** 926 (1983).

S26 Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C. & Scharf, M. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem*. **16,** 273-284 (1995).

S27 Habraken, W. J. *et al.* Ion-association complexes unite classical and non-classical theories for the biomimetic nucleation of calcium phosphate. *Nat. Commun.* **4,** 1507 (2013).

S28 Eanes, E. D. Amorphous calcium phosphate. *Monogr. Oral Sci*. **18,** 130-147. (2001).

S29 Berneche, S. & Roux, B. Energetics of ion conduction through the K^+ channel. *Nature* **414,** 73-77 (2001).

S30 Allen, T. W., Andersen, O. S. & Roux, B. Ion permeation through a narrow channel: using gramicidin to ascertain all-atom molecular dynamics potential of mean force methodology and biomolecular force fields. *Biophys. J.* **90,** 3447-3468 (2006).

S31 Xu, Z. *et al.* Molecular mechanisms for intrafibrillar collagen mineralization in skeletal tissues. *Biomaterials* **39,** 59-66 (2015).

S32 Kawska, A., Hochrein, O., Brickmann, J., Kniep, R. & Zahn, D. The nucleation mechanism of fluorapatite–collagen composites: ion association and motif control by collagen proteins. *Angew. Chem. Int. Ed.* **47,** 4982-4985 (2008).

S33 Hauptmann, S., Dufner, H., Brickmann, J., Kast, S. M. & Berry, R. S. Potential energy function for apatites. *Phys. Chem. Chem. Phys.* **5,** 635-639 (2003).

S34 Li, P., Roberts, B. P., Chakravorty, D. K. & Merz, K. M. Jr. Rational design of particle mesh Ewald compatible Lennard-Jones parameters for +2 metal cations in explicit solvent. *J. Chem. Theory Comput.* **9,** 2733-2748 (2013).

S35 Ducher, G. Fuzzy nanoassemblies: Toward layered polymeric composites. *Science* **227,** 1232-1237 (1997).

S36 Dobrynin, A. V. & Rubinstein, M. Theory of polyelectrolytes in solutions and at surfaces. *Prog. Polym. Sci.* **30,** 1049-1118 (2005).

S37 Raspaud, E., da Conceicao, M. & Livolant, F. Do free DNA counterions control the osmotic pressure? *Phys. Rev. Lett.* **84,** 2533-2536 (2000).

S38 Carrillo, J-M. Y. & Dobrynin, A. V. Salt effect on osmotic pressure of polyelectrolyte solutions: simulation study. *Polymers* **6,** 1897-1913 (2014).

S39 Rose, B. D. *Clinical Physiology of Acid-Base and Electrolyte Disorders* Ch. 22 (McGraw-Hill, New York, 1994).

S40 Ramachandran, G.X., Chandrasekharan, R., Interchain hydrogen bonds via bound water molecules in the collagen triple helix, *Biopolymers*, **6,** 1649-1658 (1968).

S41 Nudelman, F. *et al*. The role of collagen in bone apatite formation in the presence of hydroxyapatite nucleation inhibitors. *Nat. Mater.* **9,** 1004-1009 (2010).

S42 Masic, A. *et al*. Osmotic pressure induced tensile forces in tendon collagen. *Nat. Commun*. **6,** 5942 (2015).

S43 Mogilner, I. G., Ruderman, G. & Grigera, J. R. Collagen stability, hydration and native state. *J. Mol. Graph. Model.* **21,** 209-213 (2002).

S44 Pashley, D. H. *et al*. Solvent-induced dimensional changes in EDTA-demineralized dentin matrix. *J. Biomed. Mater. Res*. **56,** 273-281 (2001).

S45 Grant, C. A., Brockwell, D. J., Radford, S. E. & Thomson, N. H. Tuning the elastic modulus of hydrated collagen fibrils. *Biophys. J*. **97,** 2985-2992 (2009).

S46 Chen, Z. L., Kahn, H., Ballarini, R. & Eppell, S. J. Viscoelastic properties of isolated collagen fibrils. *Biophys. J*. **100,** 2008-2015 (2011).

S47 Morin, C., Hellmich, C. & Henits, P. Fibrillar structure and elasticity of hydrating collagen: a quantitative multiscale approach. *J. Theor. Biol*. **317,** 384-393 (2013).

S48 Cameron, I. L., Short, N.J., Fullerton, G. D. Verification of simple hydration/dehydration methods to characterize multiple water compartments on tendon type 1 collagen. *Cell Biol. Int.* **31,** 531-539 (2007).

S49 Xu, Z. *et al.* Molecular mechanisms for intrafibrillar collagen mineralization in skeletal tissues. *Biomaterials* **39,** 59-66 (2015).