Supporting Information for

Bioinspired Silicification Reveals Structural Detail

in Self-Assembled Peptide Cages

Johanna M. Galloway,*[†] Laura Senior,[‡] Jordan M. Fletcher,[†] Joseph L. Beesley,[†,‡] Lorna R. Hodgson,[‡] Robert L. Harniman,[†] Judith M. Mantell,[‡,§] Jennifer Coombs,[‡,□] Guto G. Rhys,[†] Wei-Feng Xue,[⊥] Majid Mosayebi,[‖,#] Noah Linden,[#] Tanniemola B. Liverpool,[‖,#] Paul Curnow,[‡,‖] Paul Verkade,[‡,§,‖] and Derek N. Woolfson*[†,‡,‖]

† School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK.

‡ School of Biochemistry, University of Bristol, Biomedical Sciences Building, University Walk, Bristol, BS8 1TD, UK.

§ Wolfson Bioimaging Facility, University of Bristol, Biomedical Sciences Building, University Walk, Bristol, BS8 1TD, UK.

□ Bristol Centre for Functional Nanomaterials, NSQI, Tyndall Avenue, University of Bristol, Bristol, BS8 1FD, UK.

⊥ School of Biosciences, Stacy Building, University of Kent, Canterbury, CT2 7NJ, UK.

‖ BrisSynBio, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol, BS8 1TQ, UK.

School of Mathematics, University Walk, University of Bristol, Bristol BS8 1TW, United Kingdom.

*Address correspondence to either d.n.woolfson@bristol.ac.uk or johanna.galloway@bristol.ac.uk.

CONTENTS

EXTENDED MATERIALS AND METHODS

Materials

Rink amide ChemMatrix™ resin for peptide synthesis was purchased from PCAS Biomatrix Inc. (St-Jean-sur-Richelieu, Canada). Fmoc L-amino acids and hydroxybenzotriazole (HoBt) were obtained from ACTG Bioproducts (Hessle, UK). Dowex® 50WX4 hydrogen form 100 – 200 mesh ion exchange resin ((C$_{10}$H$_{12}$·C$_{10}$H$_{10}$·C$_8$H$_8$)$_x$), 2,2'-dipyridyldisulfide (DPDS) and reagent grade sodium silicate ((NaOH)$_x$(Na$_2$SiO$_3$)$_y$·$_z$H$_2$O) were purchased from Sigma Aldrich (Gillingham, UK). All other reagents were obtained from Fisher Scientific (Loughborough, UK). All aqueous buffers and solutions were made using ultrapure water (18.2 MΩ^{-cm} at 25 °C) and filtered using a 0.22 µm PTFE sterile syringe filter purchased from Insight Biotechnology (London, UK) before use. Aluminum SEM stubs and carbon tape dots were purchased from Agar Scientific (Stanstead, UK). Formvar and carbon coated copper 200 mesh standard or Gilder Finder TEM grids were purchased from EMS, (Hatfield, PA, USA). SCANASYST-AIR-HR cantilevers (Bruker, CA, USA) were used for PeakForce (PF-AFM), and Scout-Beta (NuNano, Bristol, UK) for tapping mode (TM-AFM).

Peptide Synthesis and Assembly

Homotrimeric and heterodimeric coiled-coil peptides based on the <u>S</u>elf-<u>A</u>ssembled peptide ca<u>GE</u> (SAGE) design in Fletcher *et al*. (2013)[1] were synthesized using solid-phase peptide synthesis[2] (see Table S1 – S3) on Rink amide resin on a Liberty™ microwave peptide synthesizer (0.1 mM scale). Fmoc protected amino acids were coupled (4.5 eq. hydroxybenzotriazole (HOBt), 5 eq. N,N'-diisopropylcarbodiimide (DIC) in 7 mL dimethylformamide (DMF), 25 °C, 2 minutes then 50 °C, 25 W microwave irradiation, 5 minutes[*]) washed (5x 7 mL (DMF)), deprotected (20% (v/v) morpholine in DMF, 75 °C, 20 W microwave irradiation, 5 minutes), and washed (5x 7 mL DMF) before adding the next Fmoc protected amino acid.

Synthesized peptides were acetylated on resin by incubating with 3 eq. acetic anhydride and 4.5 eq. N,N-diisopropylethlamine (DIPEA) in 7 mL DMF for 30 minutes at room temperature. Carboxyfluorescein was coupled to an orthogonally protected lysine residue (Fmoc-lys(alloc)-OH) at the C terminus of the homotrimer. The lysine was selectively deprotected (1 eq. Pd(PPh$_3$)$_4$, 40 eq. phenylsilane in degassed dichloromethane (DCM), 25 °C, 30 minutes), and the carboxyfluorescein coupled to the exposed amine through amide bond formation chemistry (5 eq. carboxyfluorescein, 4.5 eq. HOBt, 5 eq. DIC in DMF, 25 °C, 30 minutes).

---

[*] Arginine was double coupled: 2x (45 minutes at 25 °C, 2 minutes at 75 °C + 30 W microwave irradiation).

Table S1. Diagrams, naming and abbreviations for the coiled-coil peptide sequences and modules used in this and previous work.

diagram	name	abbreviation in Fletcher *et al.* (2013)[1] style	shortened systematic abbreviation
	parent homotrimer	CC-Tri3	3
	4X residues on N terminus of homotrimer	X4-CC-Tri3	$^{X4}3$
	4X residues on C terminus of homotrimer	CC-Tri3-X4	3^{X4}
	dimer A	CC-Di-A	A
	dimer B	CC-Di-B	B
	parent hub A	CC-Tri3-CC-Di-A	3-A
	parent hub B	CC-Tri3-CC-Di-B	3-B
	4X residues on N terminus of homotrimer in hub A	X4-CC-Tri3-CC-Di-A	$^{X4}3$-A
	4X residues on C terminus of homotrimer in hub B	CC-Tri3-X4-CC-Di-B	3^{X4}-B

Table S2. Self-Assembled peptide caGE (SAGE) decorations and naming of assembled structures.[*] Percentages of components shown are as a proportion of the total SAGE hub peptide in the assembly.

SAGE name	SAGE long name	hub A component	hub B component
parent	parent	50 % 3-A CC-Tri3-CC-Di-A	50 % 3-B CC-Tri3-CC-Di-B
SAGE-E4	100% C-terminal E4	50% 3^{E4}-A CC-Tri3-E4-CC-Di-A	50% 3^{E4}-B CC-Tri3-E4-CC-Di-B
E4-SAGE	100% N-terminal E4	50% $^{E4}3$-A E4-CC- Tri3-CC-Di-A	50% $^{E4}3$-B E4-CC-Tri3-CC-Di-B
50% SAGE-K4	50% C-terminal K4	50 % 3-A CC-Tri3-CC-Di-A	50% 3^{K4}-B CC-Tri3-K4-CC-Di-B
SAGE-K4	100% C-terminal K4	50% 3^{K4}-A CC-Tri3-K4-CC-Di-A	50% 3^{K4}-B CC-Tri3-K4-CC-Di-B
50% K4-SAGE	50% N-terminal K4	50 % 3-A CC-Tri3-CC-Di-A	50% $^{K4}3$-B K4-CC- Tri3-CC-Di-B
K4-SAGE	100% N-terminal K4	50% $^{K4}3$-A K4-CC-Tri3-CC-Di-A	50% $^{K4}3$-B K4-CC-Tri3-CC-Di-B
50% SAGE-R4	50% C-terminal R4	50 % 3-A CC-Tri3-CC-Di-A	50% 3^{R4}-B CC-Tri3-R4-CC-Di-B
SAGE-R4	100% C-terminal R4	50% 3^{R4}-A CC-Tri3-R4-CC-Di-A	50% 3^{R4}-B CC-Tri3-R4-CC-Di-B
50% R4-SAGE	50% N-terminal R4	50 % 3-A CC-Tri3-CC-Di-A	50% $^{R4}3$-B R4-CC-Tri3-CC-Di-B
R4-SAGE	100% N-terminal R4	50% $^{R4}3$-A R4-CC-Tri3-CC-Di-A	50% $^{R4}3$-B R4-CC-Tri3-CC-Di-B
R4-SAGE-fl	Fluorescent N-terminal R4	50% $^{R4}3$-A R4-CC-Tri3-CC-Di-A	47.5% $^{R4}3$-B + 2.5% 3^{fl}-B 47.5% R4-CC-Tri3-CC-Di-B + 2.5% CC-Tri3-fl-CC-Di-B

---

[*] See Table S1 for details on module component naming. See Table S3 for details on the coiled-coil component amino acid sequences.

Table S3. Coiled-coil and control peptide sequences used in this work. Molecular weight and extinction coefficient are predicted based on the primary amino acid sequences using http://pepcalc.com (07/18/2016, Innovagen).

name	long name of peptide	molecular weight (Da)	extinction coefficient ε
	amino acid sequence of peptide		
A	CC-Di-A	2656.9	5690
	Ac-GEIAALEKENAALECEIAALEQGWG-Am		
B	CC-Di-B	2757.4	6970
	Ac-GKIAALKKKNAALKCKIAALKQGYW-Am		
3	CC-Tri3	2630.1	1280
	Ac-GEIAAIKKEIAAIKCEIAAIKQGYG-Am		
E43	E4-CC-Tri3	3260.7	1280
	Ac-EEEEGGGEIAAIKKEIAAIKCEIAAIKQGYG-Am		
3E4	CC-Tri3-E4	3260.7	1280
	Ac-GEIAAIKKEIAAIKCEIAAIKQGYGGGEEEE-Am		
K43	K4-CC-Tri3	3256.9	1280
	Ac-KKKKGGGEIAAIKKEIAAIKCEIAAIKQGYG-Am		
3K4	CC-Tri3-K4	3256.9	1280
	Ac-GEIAAIKKEIAAIKCEIAAIKQGYGGGKKKK-Am		
R43	R4-CC-Tri3	3369.0	1280
	Ac-RRRRGGGEIAAIKKEIAAIKCEIAAIKQGYG-Am		
3R4	CC-Tri3-R4	3369.0	1280
	Ac-GEIAAIKKEIAAIKCEIAAIKQGYGGGRRRR-Am		
3fl	CC-Tri3-fl	3287.9	5430
	Ac-GEIAAIKKEIAAIKCEIAAIKQGYGGG(K*Cb)G-Am		
GYGK4	Control K4 peptide	849.0	1280
	Ac-GYGKKKK-Am		
GYGR4	Control R4 peptide	961.1	1280
	Ac-GYGRRRR-Am		

Peptides were cleaved from the resin by incubation for 3 hours with a mixture of trifluoroacetic acid (TFA, 94% (v/v)), 1,2-ethanedithiol (EDT, 2% (v/v)), triisopropylsilane (TIPS, 2% (v/v)) and water (H_2O, 2% (v/v)) for 3 hours. The resin was removed by filtration, excess solvent evaporated and the crude peptide precipitated by adding 50 mL of ice cold (4 °C) diethyl ether (Et_2O). The peptides were pelleted by centrifugation (4000 xg, 10 minutes, 4 °C), the Et_2O poured off, and the pellet dissolved in 5 mL of an aqueous solution of 50% (v/v) acetonitrile (MeCN).

Cleaved peptides were purified using reverse phase – high pressure liquid chromatography (RP-HPLC) on a Kromatek (semi-micro, 5 µm, 100 Å, 10 mm ID x 150 mm L) C18 column using buffers A (0.1% (v/v) TFA in water) and B (0.1% (v/v) TFA in MeCN). ≈10 mg of peptide was dissolved in 3 mL of 20% buffer B and injected onto the column equilibrated with 20% buffer B. A linear gradient between 20% and 80% buffer B was applied at 3 mL minute^{-1} over 30 minutes. Monitoring of absorbance at a wavelength of $\lambda = 280$ nm and 220 nm was used to identify peptide peaks for collection. Purified peptides were analyzed by matrix-assisted laser desorption / ionization – time of flight mass spectrometry (MALDI-TOF, >95% purity) and analytical RP-HPLC, and these data normalized (Figure S1 and Table S4).

Each homotrimer peptide was joined back-to-back to either heterodimer A (CC-Di-A) or B (CC-Di-B) *via* an unsymmetric disulfide bond using 2,2'-dipyridyldisulfide (DPDS) mediated thiol activation.[3] 2 – 5 mg of heterodimer was dissolved in 2 – 5 mL water (Di-A) or phosphate buffered saline (Di-B, PBS, 137 mM NaCl, 2.7 mM KCl, 8.02 mM dibasic Na_2HPO_4, 1.98 mM monobasic KH_2PO_4, pH 7.4). 10 equivalents of DPDS was dissolved in 1 mL methanol (MeOH) and added to the peptide solution. After ≈1 hour, unreacted DPDS was removed by extracting with 3 x 30 mL washes Et_2O. The peptide fraction was freeze dried, purified using RP-HPLC and analyzed by MALDI-TOF (>95% purity) and analytical RP-HPLC (Figure S1 and Table S4). After freeze drying, the activated heterodimer and an equimolar amount of a homotrimer was dissolved at ≈1 mg mL^{-1} in water (Di-A) or PBS (Di-B). The activated heterodimer was mixed with the homotrimer, and over 3 – 16 hours they are linked back-to-back with an unsymmetric disulfide bond, which was confirmed MALDI-TOF (>95% purity) and analytical RP-HPLC (Figure S1 and Table S4) after RP-HPLC purification.

Circular Dichroism (CD) Spectroscopy

Circular dichroism (CD) spectra were recorded using a JASCO J-810 spectropolarimeter with a Peltier temperature controller. A 50 µM peptide solution was prepared in potassium phosphate buffer (PI, 10 mM potassium phosphate (8.02 mM dibasic K_2HPO_4, 1.98 mM monobasic KH_2PO_4), pH 7.4) with a 5x excess (250 µM) of tris(2-carboxyethyl)phosphine hydrochloride (TCEP) to prevent disulfide bond formation. As measurements were recorded in a quartz cuvette, it was not possible to add silicic acid, so 25 mM NaCl was added as a proxy to maintain a similar ionic strength to that found during mineralization. Peptide concentration was determined using a NanoDrop 2000 UV-Vis spectrophotometer (see Table S3 for extinction coefficients). Peptide solutions were cooled to 5 °C in a 1 mm path length quartz cuvette. Samples were heated at 40 °C hour^{-1} between 5 °C and 90 °C, and cooled back to 5 °C. Absorbance at a wavelength of $\lambda = 222$ nm was recorded every 1 °C, and full CD spectra between $\lambda = 260 - 190$ nm were recorded every 5 °C, and a post-melt scan was also recorded at 5 °C. These data were normalized to give a mean residue ellipticity (MRE, θ in deg cm^2 dmol^{-1} res^{-1}) using Equation S1:[4]

$$\theta = \frac{\theta_{obs}}{10 \times c \times l \times n} \tag{S1}$$

Where θ_{obs} is the observed ellipticity in degrees, c is the concentration in moles, l is the path length in cm and n is the number of amide bonds in the peptide. The measured α helical fraction of the peptide was calculated using Equation S2:[5,6]

$$Fraction\ \alpha\text{-helix}\ (\%) = \frac{100 \times (\theta_{222} - \theta_{coil})}{-42500 \times (1 - (3/n) - \theta_{coil}} \tag{S2}$$

Where θ_{222} is the MRE at a wavelength of λ = 222 nm, θ_{coil} is 415 in deg cm² dmol⁻¹ res⁻¹ at 5 °C, and n is the number of amide bonds in the peptide. Melting transition midpoint temperature (T_M) was calculated from the *x*-intercept of the second derivative plot of MRE at a wavelength of 222 nm against temperature.

Analytical Ultracentrifugation (AUC)

AUC sedimentation velocity (SV) experiments were done using Beckman Optima XL-A and XL-I analytical ultracentrifuges and an An-60 rotor at 20 °C. Peptides were dissolved at 325 mM (A₂₃₇ ≈1.0 for 1.2 cm path length) in PI buffer with a 5x excess of TCEP (1.625 mM) and 25 mM NaCl. 310 µL or 410 µL peptide solution was loaded into 1 channel of a 2 channel Epon or aluminum SV centerpiece respectively, with quartz windows or sapphire windows respectively. 320µL or 420 µL of blank respectively (PI buffer with a 5x excess of TCEP (1.625 mM) and 25 mM NaCl) was loaded into the other channel. Samples were spun at 60,000 rpm, with absorbance scans taken across 5.8 cm – 7.3 cm (total 120 scans). These data were fitted using SEDFIT[7] to a continuous c(s) distribution model at a 95% confidence level. For this, the baseline, bottom, frictional coefficient (f/f₀), and systematic time-invariant and radial-invariant noise were fitted, and the partial specific volume ($\bar{v}$) for the peptide, buffer density and viscosity were calculated using SEDNTERP.[8] (see Table S6).

Sedimentation equilibrium data were collected at 20 °C in a Beckman Optima XL-A or XL-I analytical ultracentrifuge using an An-50 rotor. Peptides were dissolved at 325 mM (A₂₈₀ ≈ 0.5 for 1.2 cm path length) in PI buffer with a 5x excess of TCEP (1.625 mM) and 25 mM NaCl. 3x 110 µL peptide solution and 3x 120 µL blank were loaded into an SE Epon 6 channel centerpiece with quartz windows and a 1.2 cm path length. The samples were centrifuged at 3 krpm intervals between speeds of 30,000 – 48,000 rpm. The partial specific volume ($\bar{v}$) for the peptide, buffer density and viscosity were calculated using SEDNTERP[8] (see Supplementary Table S6) These data fitted with a single, ideal species model using Ultrascan II,[9] with 95% confidence limits on the fitted mass determined *via* Monte Carlo analysis.

Silica Coating SAGEs

To determine the concentration of silicon in the as purchased sodium silicate solution, a range of dilutions were filtered, and the concentration of silicon measured using a Jobin Yvon Horiba Ultima 2 sequential

inductively coupled plasma atomic emission spectrometer (ICP-AES). Orthosilicic acid (referred to throughout as silicic acid) was freshly prepared just prior to use as per Alexander (1953).[10] 4 mL of a 720 mM sodium silicate solution (540 µL of the as purchased solution (~5.3 M silicon) was diluted in 3.46 mL water) with 4 mL of 10 mM sulfuric acid in the presence of 1.5 mg Dowex® 50WX4 resin. This was mixed for 10 minutes at 20 °C. 3.3 mL of the acidified solution was added to 6.7 mL water and filtered to make a 120 mM stock of silicic acid.

In a standard mineralization, SAGE peptides (at between 2 and 25 µM) were assembled in PI buffer for 1 hour at 20 °C. Whether using undecorated parent, decorated trimers, or a mixture of decorated SAGE peptides, a 1:1 ratio of total hub A to total hub B was always present to ensure stoichiometric assembly of the heterodimers. 500 µL of assembled SAGE peptide was then added to 500 µL PI buffer containing between 0 – 36 mM silicic acid from 120 mM silicic acid stock solution. This was incubated at 20 °C for 24 hours. 0.5 µL of 2-(4-pyridyl)-5-((4-(2-dimethylaminoethylaminocarbamoyl)methoxy)phenyl)oxazole (PDMPO, LysoTracker®)[11] was added to 100 µL of a mineralization reaction for detection of silica polymerization. In the absence of silica, PDPMO is excited at a wavelength of $\lambda = 388$ nm, and emits at 450 and 525 nm. The emission at 450 nm is quenched by the polymerization of silicic acid to form silica. Fluorescence spectra were recorded at 3 nm intervals between 400 – 600 nm on a CLARIOstar microplate reader (BMG Labtech, Ortenberg, DE). Samples were washed by centrifugation (17,000 xg, 5 minutes), removal of the supernatant, and resuspension in 1 mL water 3 times, and stored in ≈100 µL water.

Scanning Electron Microscopy (SEM)
Samples for SEM were prepared by resuspending 5 – 10 µL of washed sample using a pipette. This was dropped onto cleaned silicon wafer (1 mL water, 1 mL ethanol, 1 mL water, dried with N_2 stream) fixed to carbon tape on an SEM stub and dried. A 5 – 8 nm film of gold-palladium was sputter coated onto unmineralized samples using an Emtech 575X sputter coater. Sample stubs were mounted in a FEI Quanta 200 FEG-SEM and imaged at a working distance of ≈8.5 mm, and an accelerating voltage of 5 keV.

Transmission Electron Microscopy (TEM)
Samples for TEM were prepared by resuspending 5 – 10 µL of washed sample using a pipette, and then dropped onto a 10 nm formvar and 1 nm carbon coated copper 200 mesh TEM grids (FCF200-Cu, EMS, PA, USA) or Gilder Finder grids (FCF200F2-Cu, EMS, PA, USA) and dried. As a control, unmineralized SAGEs were dried on grids and negatively stained using 1% uranyl acetate (5 µL, 30 seconds). The stain was wicked off using filter paper, and the grid washed using 5 µL water wicked off with filter paper 3x and then dried. Sample grids were imaged on a Tecnai 12 – FEI 120 kV BioTwin Spirit TEM fitted with a tungsten filament and at an accelerating voltage of 120 keV. 15 nm gold fiducials were added to grids for tilt series acquisition. Tilt series between angles of ≈70° and ≈-70° for tomographic reconstructions were

collected using a Tecnai 20 – FEI 200 kV Twin Lens scanning transmission electron microscope (STEM) fitted with a LaB$_6$ filament at an accelerating voltage of 200 keV. Images were collected using an FEI Eagle 4k x 4k CCD camera. Energy dispersive X-ray (EDX) spectra were recorded using a JEOL 2100F STEM with a LaB$_6$ filament at an accelerating voltage of 200 keV, and images were collected using a Gatan Orius 11 megapixel 832 camera. X-rays were detected using 80 mm^2 AZtec detector, and processed and quantified using AZtecTEM software.

The grainsize of SAGEs was recorded using representative TEM images and ImageJ[12,13] to measure the length and width of 200 – 400 particles for each sample. These data were then binned into 25 nm bins and fitted with a single or double Gaussian fit in Origin 2015 64 bit, which uses Equation S3:

$$y = y_0 + \frac{A}{w\sqrt{\pi/2}}e^{-2\frac{(x-x_c)^2}{w^2}}$$

(S3)

Where y_0 is the offset, x_c is the peak center, w is the width and A is the amplitude of the peak. Grainsizes quoted are the average, and the error is one standard deviation of these data.

Tilt series were aligned using the 15 nm gold fiducials and reconstructed using eTomo in the IMOD[14,15] software package (University of Colorado, Boulder, US). 3D reconstructions of tomograms were rendered in Amira 3D (FEI Software). Movie files of the tomographic reconstruction software outputs were compiled in ImageJ[12] and compressed using HandBrake.

Correlative Light and Electron Microscopy (CLEM)[16,17]

Finder grids of fluorescent SAGEs were imaged in the bright-field and fluorescence channels using a 63x water immersion lens (numerical aperture 1.2) on a Leica DMI4000 B inverted epifluorescence microscope. The same areas of each grid was then imaged using TEM (see above). Fluorescence image data was overlain with TEM image data using the TurboReg Registration plug-in in Fiji[18] to produce correlative light and electron microscopy (CLEM) images. For this, the source image is a fluorescent & bright-field overlay, the target is the corresponding TEM image. Accurate and affine are selected, and 3 landmark positions that occur in both images are marked. Manual alignment is selected, which aligns the two images in rotation and scale. The fluorescence data is then overlain onto the TEM image to create a CLEM overlay.

Permeability Testing

Fluorescent SAGEs were coated using 12 or 24 mM silicic acid as above. These samples were resuspended in 100 µL PI buffer, and 10 µL transferred to a fluorescence 96 well plate. As a control, unmineralized fluorescent SAGEs were assembled in PI buffer for 1 hour, and then 40 uL was added to the 96 well plate. These samples were made up to a total volume of 50 µL with PI buffer (control, no additive) or PI buffer containing fluorescence quenching, reducing or proteolyzing additives. The quenching agent used was

10 µL 0.4% (w/v) Trypan Blue. The reducing agent was 10 µL 5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP). A protease (trypsin, 5 µL 1 mg mL^{-1}) was incubated with the SAGEs at 37 °C. After treatment, samples were imaged on a widefield fluorescence microscope (details above in CLEM section), by sandwiching 5 µL of incubated sample between a glass slide and a cover slip before mounting onto the imaging stage. The same illumination and detector settings were maintained so that the level of fluorescence could be compared semi-quantitatively between samples. In all cases, images containing any fluorescence were recorded, which shows the maximum fluorescence from the samples in all of the conditions tested.

Atomic Force Microscopy (AFM)

SiO$_2$-SAGEs were investigated both in tapping-mode (TM) and PeakForce (PF) feedback modes on atomic force microscopes (AFMs). Tapping-mode (TM-AFM) was conducted on a Multi-mode microscope with a Quadrexed Nanoscope III controller, utilizing Scout-Beta cantilevers with a nominal tip radius of 2 nm, resonant frequency of 350 kHz, and a spring constant of 42 N m^{-1}. PeakForce (PF-AFM) was conducted on a Multi-mode VIII microscope with Nanoscope V controller and a fast scan head unit in combination with SCANASYST-AIR-HR cantilevers of nominal tip radius 2 nm, and spring constant 0.4 N m^{-1}. In PeakForce imaging, force curves are taken by nonresonant low frequency oscillation of the cantilever above the sample, and the maximum interaction force between the cantilever tip and the sample is controlled as the feedback parameter. In this way, material properties such as Young's modulus, adhesion and dissipation of energy between the sample and the tip can be mapped simultaneous to the mapping of the surface topography. Data was analyzed using Nanoscope Analysis 1.5 software (Bruker, CA, US). Good agreement was found between maps of these surface properties and the more qualitative phase shift maps generated via tip-sample interaction forces in TM-AFM. To extract features present on the surfaces of individual particles, a least squares splines model describing a smooth, nonsymmetric convex surface were fitted to each of the scan-lines in the topological data of the particles using a script written in Matlab (Mathworks). The resulting geometric model describing the shape of the particles were subsequently subtracted from the original topology data. This generated a flattened topology map that contains detail of features on the surface of the particles.

Coarse-grained Computational Modelling of SAGE Assembly

Self-assembly of the SAGEs was investigated using a coarse-grained model designed on the level of a coiled-coil peptide as a single unit, parametrization and model design detailed in Mosayebi *et al.*[19]. Each coiled-coil in the model was a rigid body with Lennard-Jones (LJ) attractive patches on the surface to drive the formation of either heterodimeric or homotrimeric coiled-coils. A permanent bond connects each trimer component to a dimer component, as per the SAGE peptide hub design.
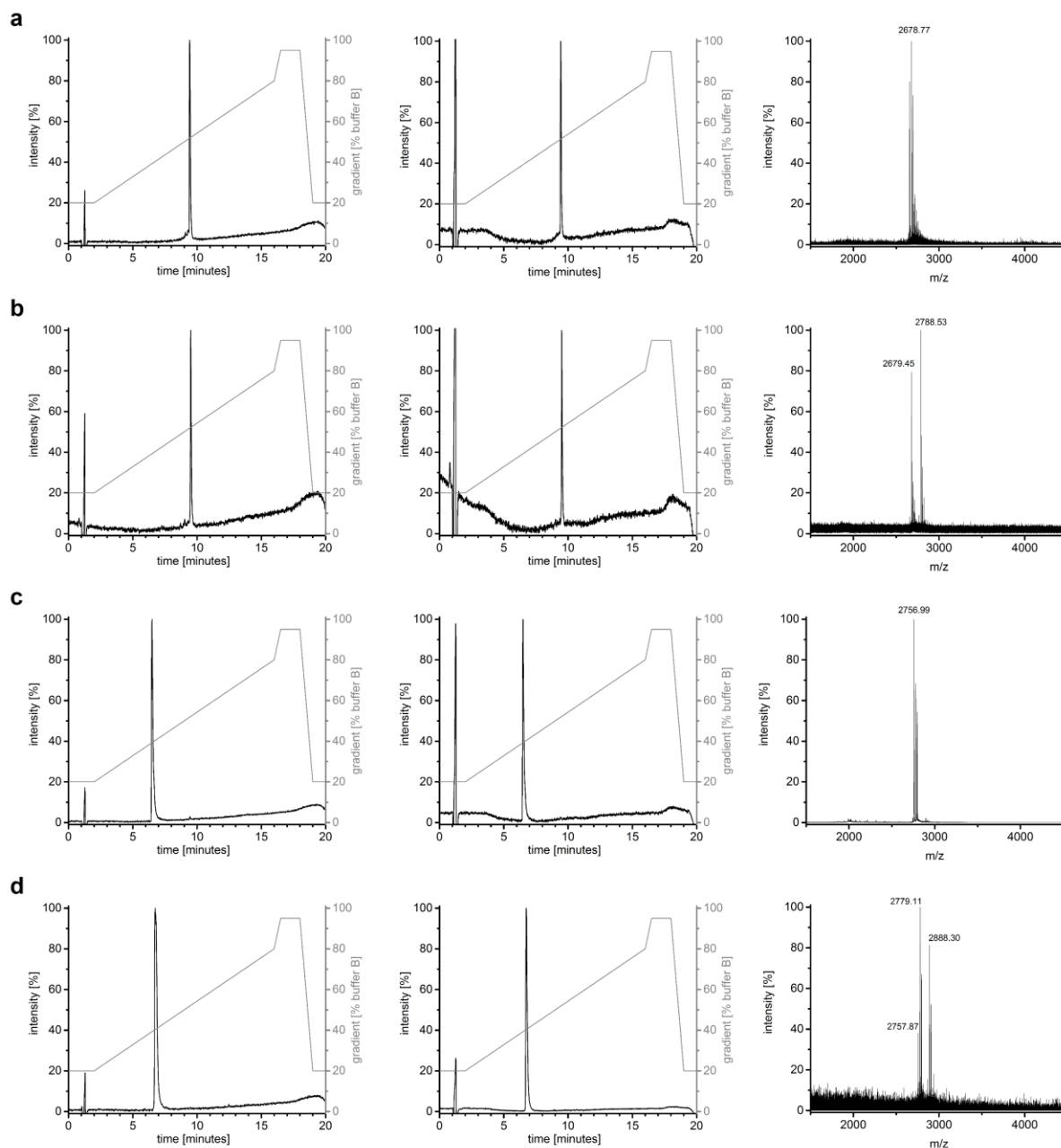
SUPPLEMENTARY FIGURES, TABLES AND COMMENTS

Figure S1. Analytical RP-HPLC and MALDI-TOF spectra for the peptides used in this study. Absorbance (shown in black) at (left) a wavelength of λ=220 nm, and at (center) λ=280 nm from analytical RP-HPLC (buffer gradient shown in grey); and (right) corresponding mass spectra (summarized in Table S4). Heterodimers: (a) CC-Di-A, (b) activated CC-Di-A, (c) CC-Di-B, and (d) activated CC-Di-B.

Notes on Peptide Characterization with MALDI-TOF and Analytical RP-HPLC

i.   Peptides were purified with RP-HPLC and their purity analyzed using analytical RP-HPLC to check that only one species was present in the sample, and MALDI-TOF to check their mass.

ii. For the activated peptides, the laser degraded some of the activated peptide to its unactivated form to give two peaks in the mass spectra. As the activated samples only had one peak in analytical

RP-HPLC, which had a retention time shifted from that of the unactivated peptide, it was assumed that they were a single, activated peptide species.

iii. The assembled hubs had between two and four peaks in the mass spectrometer as the laser degrades some of the assembled peptides to their component parts, and there was also a m/2 peak at half of the expected mass for the assembled hub peptides. As the assembled hubs only had one peak in analytical RP-HPLC, which had a retention time shifted from that of the component parts, it was assumed that they were a single peptide species.
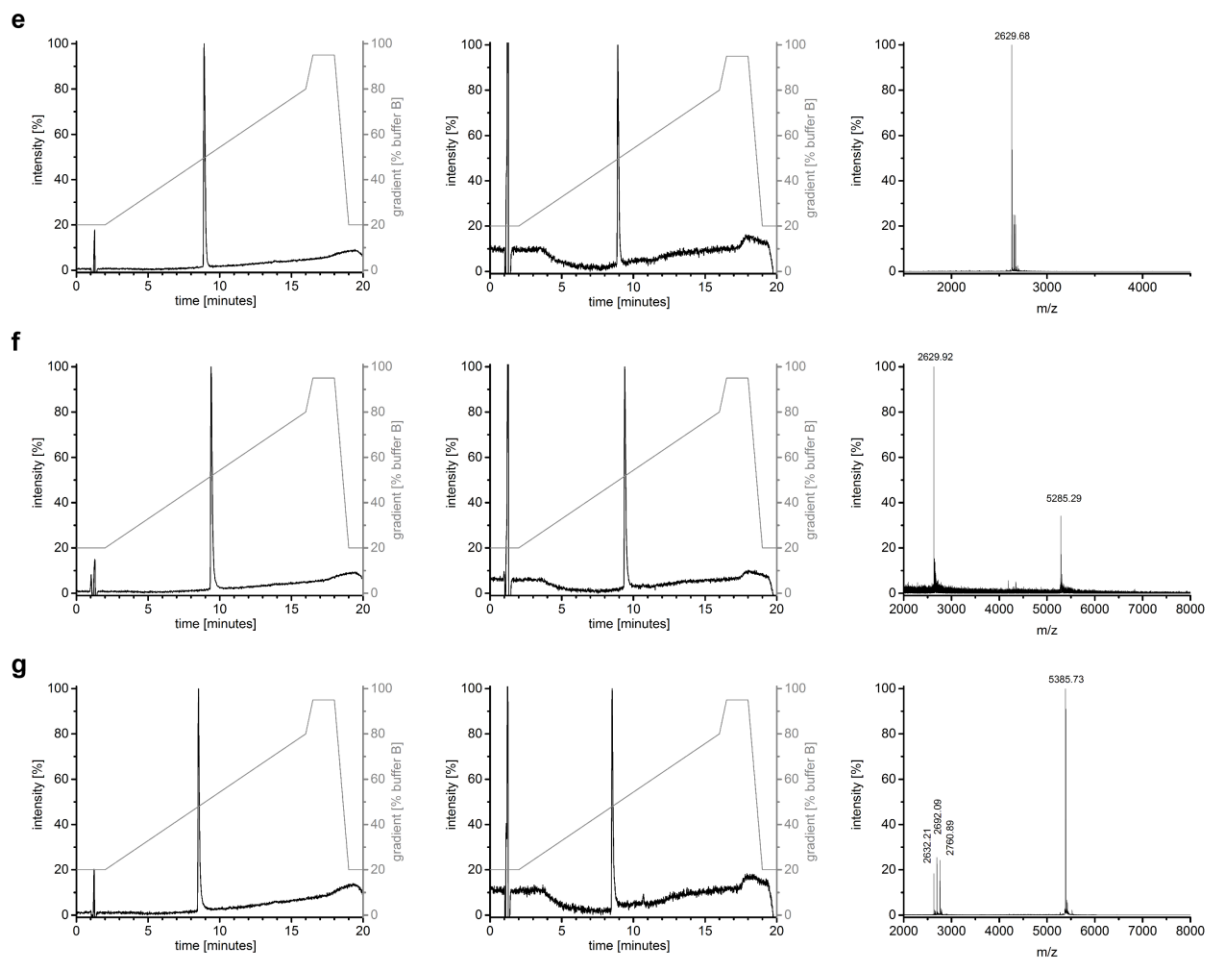

Figure S1 continued. Analytical RP-HPLC and MALDI-TOF spectra of parent (e) homotrimer (CC-Tri3), (f) parent hub A (CC-Tri3-CC-Di-A), and (g) parent hub B (CC-Tri3-CC-Di-B).
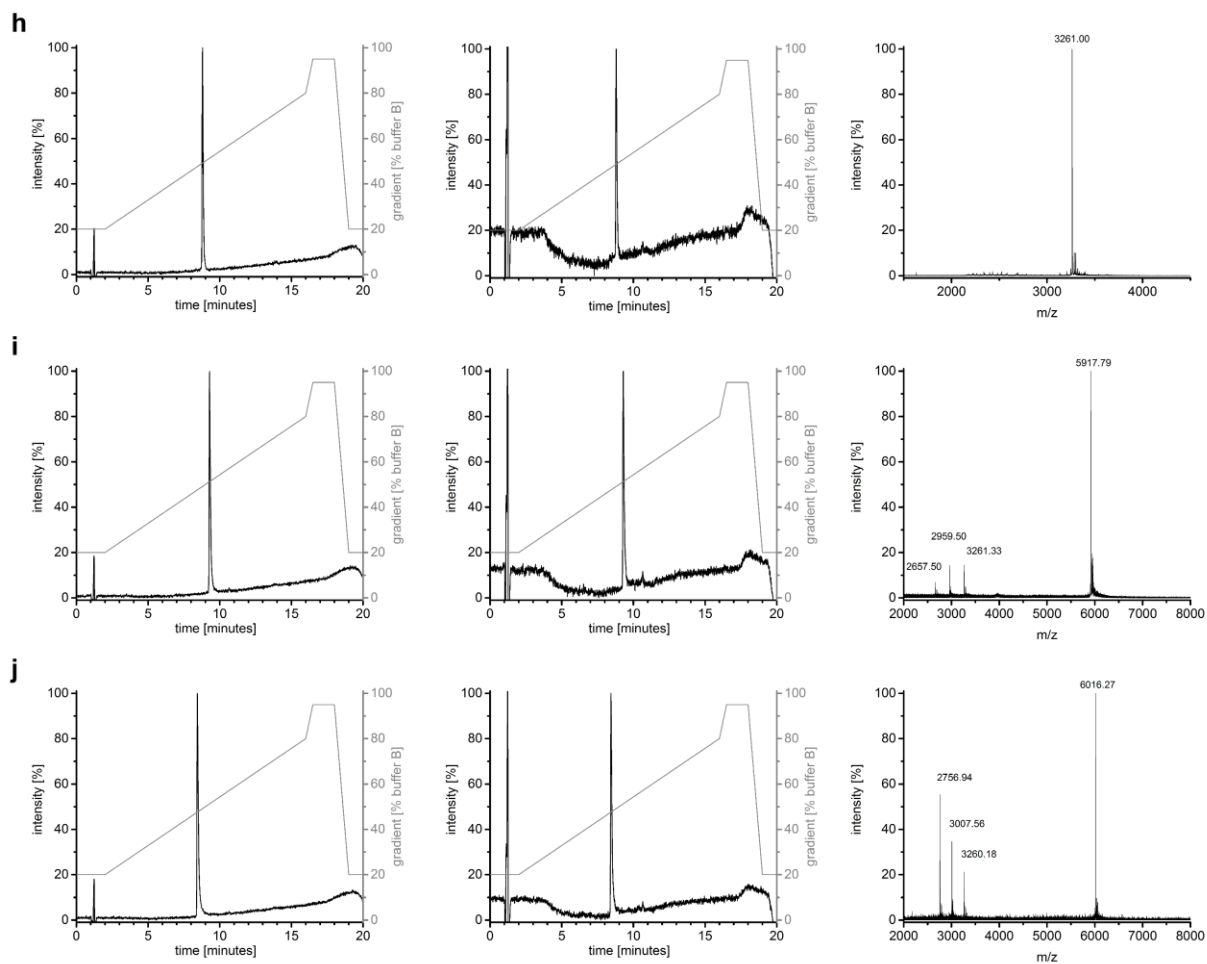
Figure S1 continued. Analytical RP-HPLC and MALDI-TOF spectra of negatively charged E4 decorated at the C terminus (h) 3^{E4} homotrimer (CC-Tri3-E4); and hubs (i) 3^{E4}-A (CC-Tri3-E4-CC-Di-A), and (j) 3^{E4}-B (CC-Tri3-E4-CC-Di-B).

Figure S1 continued. Analytical RP-HPLC and MALDI-TOF spectra of negatively charged E4 decorated at the N-terminus (k) ^E4^3 homotrimer (E4-CC-Tri3); and hubs (l) ^E4^3-A (E4-CC-Tri3-CC-Di-A), and (m) ^E4^3-B (E4-CC-Tri3-CC-Di-B).

Figure S1 continued. Analytical RP-HPLC and MALDI-TOF spectra of positively charged K4 decorated at the C-terminus (n) 3^{K4} homotrimer (CC-Tri3-K4); and from hubs (o) 3^{K4}-A (CC-Tri3-K4-CC-Di-A), and (p) 3^{K4}-B (CC-Tri3-K4-CC-Di-B).
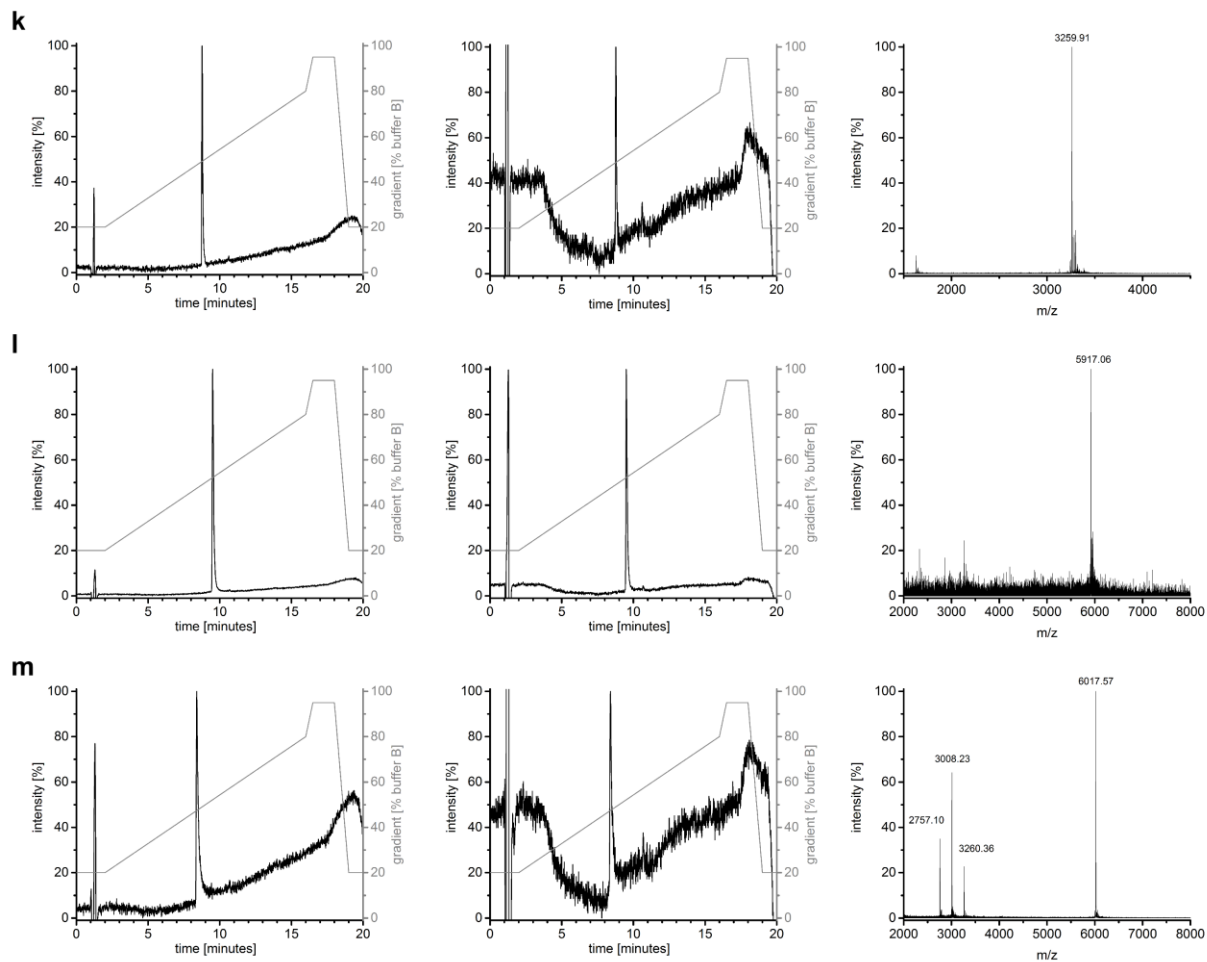
Figure S1 continued. Analytical RP-HPLC and MALDI-TOF spectra of positively charged K4 decorated at the N-terminus (q) [K4]3 homotrimer (K4-CC-Tri3); and hubs (r) [K4]3-A (K4-CC-Tri3-CC-Di-A), and (s) [K4]3-B (K4-CC-Tri3-CC-Di-B).
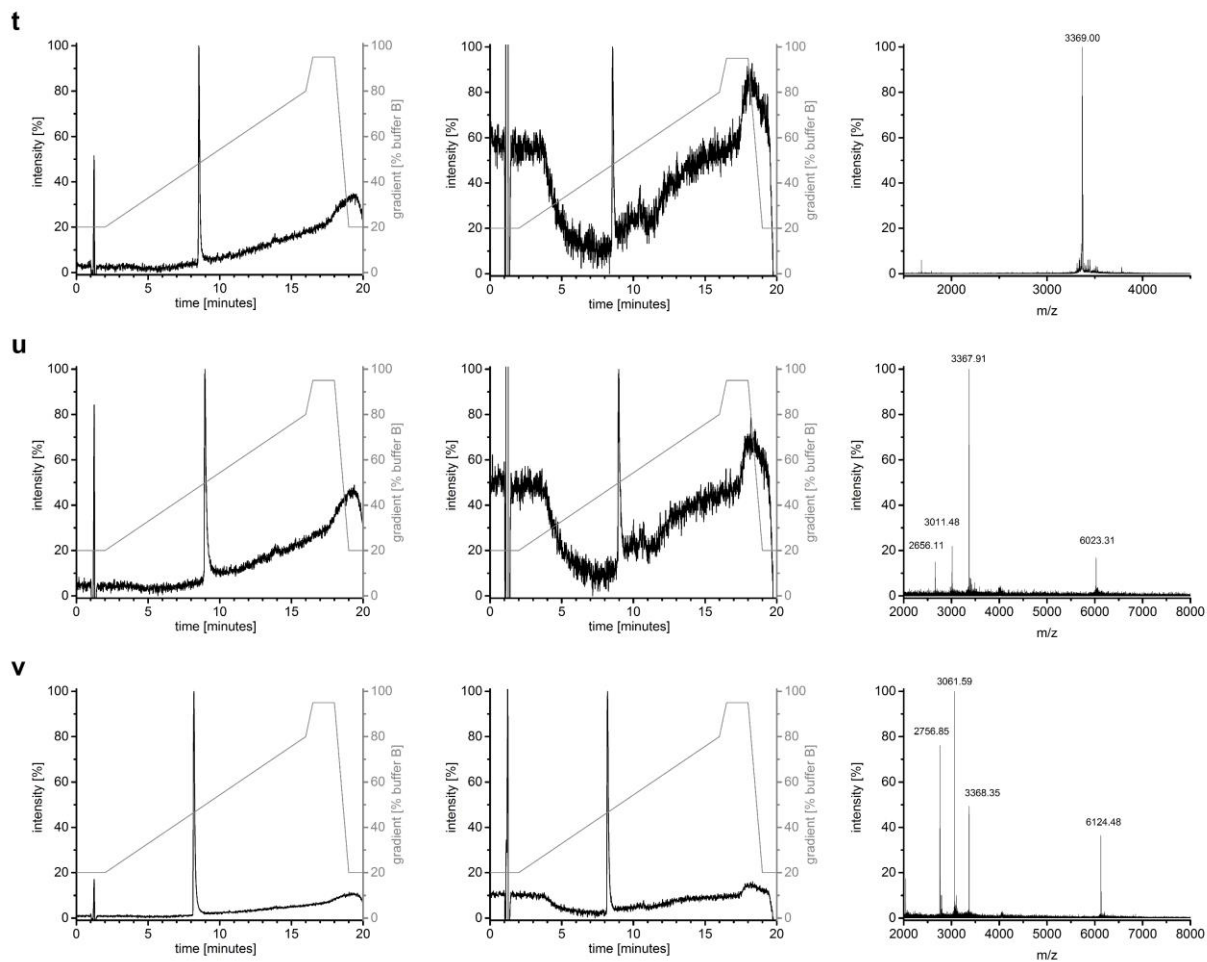
Figure S1 continued. Analytical RP-HPLC and MALDI-TOF spectra of positively charged R4 decorated at the C-terminus (t) 3^{R4} homotrimer (CC-Tri3-R4); and from hubs (u) 3^{R4}-A (CC-Tri3-R4-CC-Di-A), and (v) 3^{R4}-B (CC-Tri3-R4-CC-Di-B).
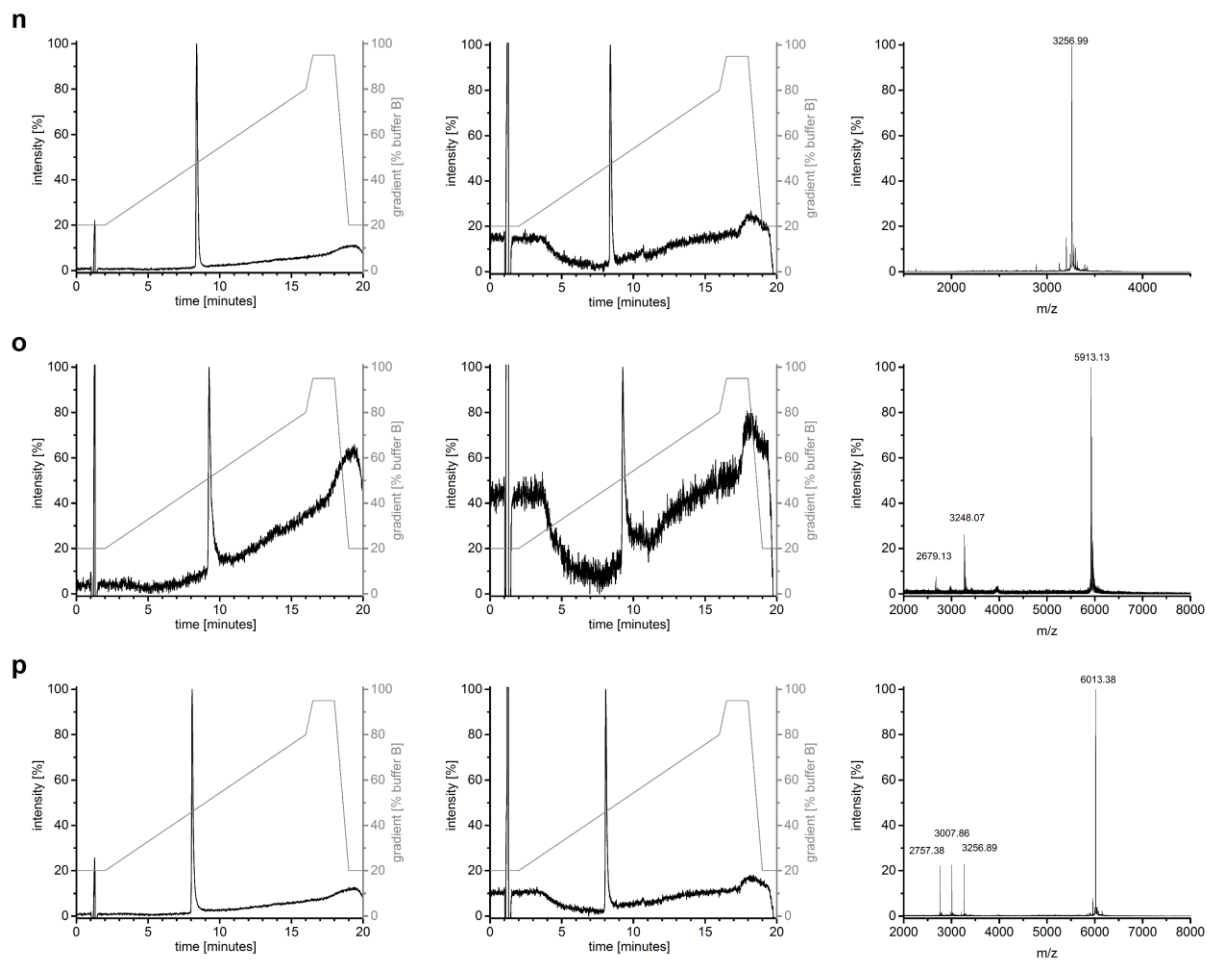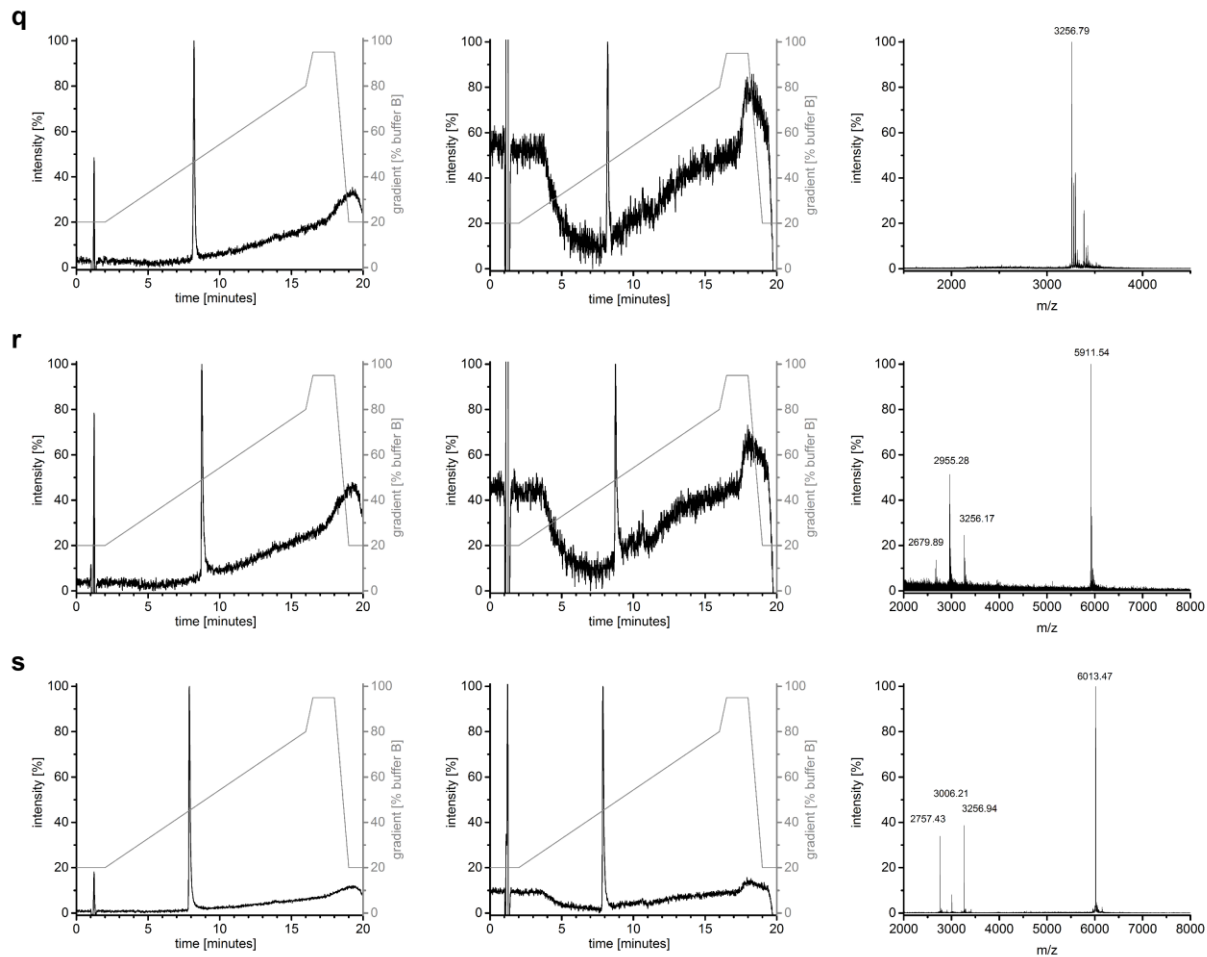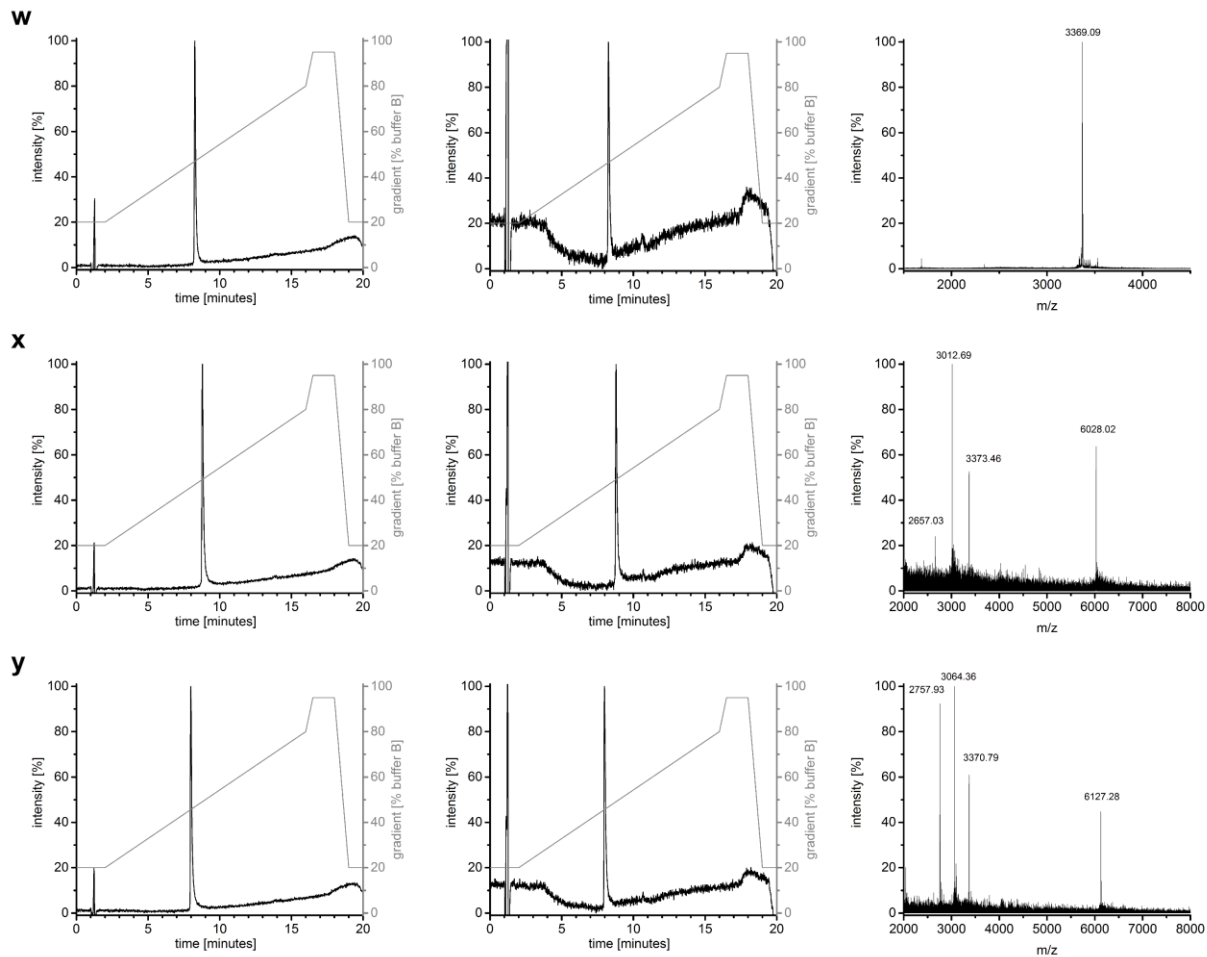
Figure S1 continued. Analytical RP-HPLC and MALDI-TOF spectra of positively charged R4 decorated at the N-terminus (w) [R4]3 homotrimer (R4-CC-Tri3); and hubs (x) [R4]3-A (R4-CC-Tri3-CC-Di-A), and (y) [R4]3-B (R4-CC-Tri3-CC-Di-B).

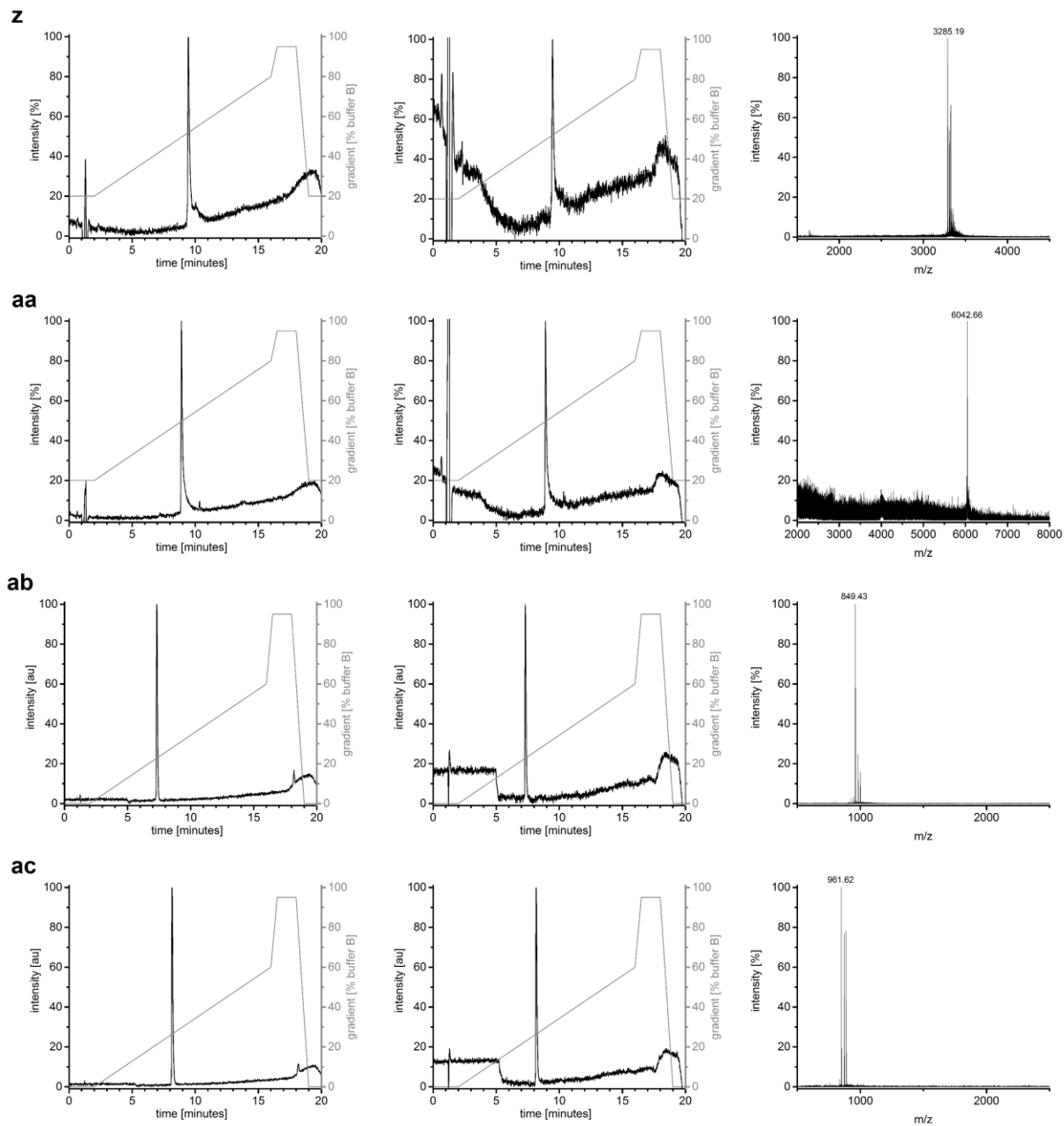**Figure S1 continued.** Analytical RP-HPLC and MALDI-TOF spectra of carboxyfluorescein C-terminally labelled (z) 3^{fl} homotrimer (CC-Tri3-fl), (aa) hub 3^{fl}-B (CC-Tri3-fl); and positively charged control peptides (ab) GYG^{K4} (GYGKKKK) and (ac) GYG^{R4} (GYGRRRR).

Table S4. Peptide mass recorded using MALDI-TOF, expected mass is proton (H$^+$ *i.e.* 1 mass unit more than the relative molecular mass) peak unless otherwise stated. For nomenclature, see Table S2, and for peptide sequences see Table S3. Spectra shown in Figure S1.

peptide name	long name of peptide	observed m/z	expected mass [Da]
A	CC-Di-A	2678.77*	2678.92*
activated A	activated CC-Di-A	2788.53*	2789.92*
B	CC-Di-B	2756.99	2758.39
activated B	activated CC-Di-B	2888.30*	2890.39*
3	CC-Tri3	2629.68	2631.11
3-A	CC-Tri3-CC-Di-A	5285.29	5288.03
3-B	CC-Tri3-CC-Di-B	5385.73	5388.50
3^{E4}	CC-Tri3-E4	3261.00	3261.67
3^{E4}-A	CC-Tri3-E4-CC-Di-A	5917.79	5918.59
3^{E4}-B	CC-Tri3-E4-CC-Di-B	6016.27	6019.06
^{E4}3	E4-CC-Tri3	3259.91	3261.67
^{E4}3-A	E4-CC-Tri3-CC-Di-A	5917.06	5918.59
^{E4}3-B	E4-CC-Tri3-CC-Di-B	6017.57	6019.06
3^{K4}	CC-Tri3-K4	3256.99	3257.91
3^{K4}-A	CC-Tri3-K4-CC-Di-A	5913.13	5914.83
3^{K4}-B	CC-Tri3-K4-CC-Di-B	6013.38	6015.30
^{K4}3	K4-CC-Tri3	3256.79	3257.91
^{K4}3-A	K4-CC-Tri3-CC-Di-A	5911.54	5914.83
^{K4}3-B	K4-CC-Tri3-CC-Di-B	6013.47	6015.30
3^{R4}	CC-Tri3-R4	3369.00	3369.96
3^{R4}-A	CC-Tri3-R4-CC-Di-A	6023.31	6026.88
3^{R4}-B	CC-Tri3-R4-CC-Di-B	6124.48	6127.35
^{R4}3	R4-CC-Tri3	3369.09	3369.96
^{R4}3-A	R4-CC-Tri3-CC-Di-A	6028.02	6026.88
^{R4}3-B	R4-CC-Tri3-CC-Di-B	6127.28	6127.35
3^{fl}	CC-Tri3-fl	3285.19	3287.87
3^{fl}-B	CC-Tri3-fl-CC-Di-A	6042.66	6046.26
GYG^{K4}	GYGKKKK	849.43	850.03
GYG^{R4}	GYGRRRR	961.62	962.09

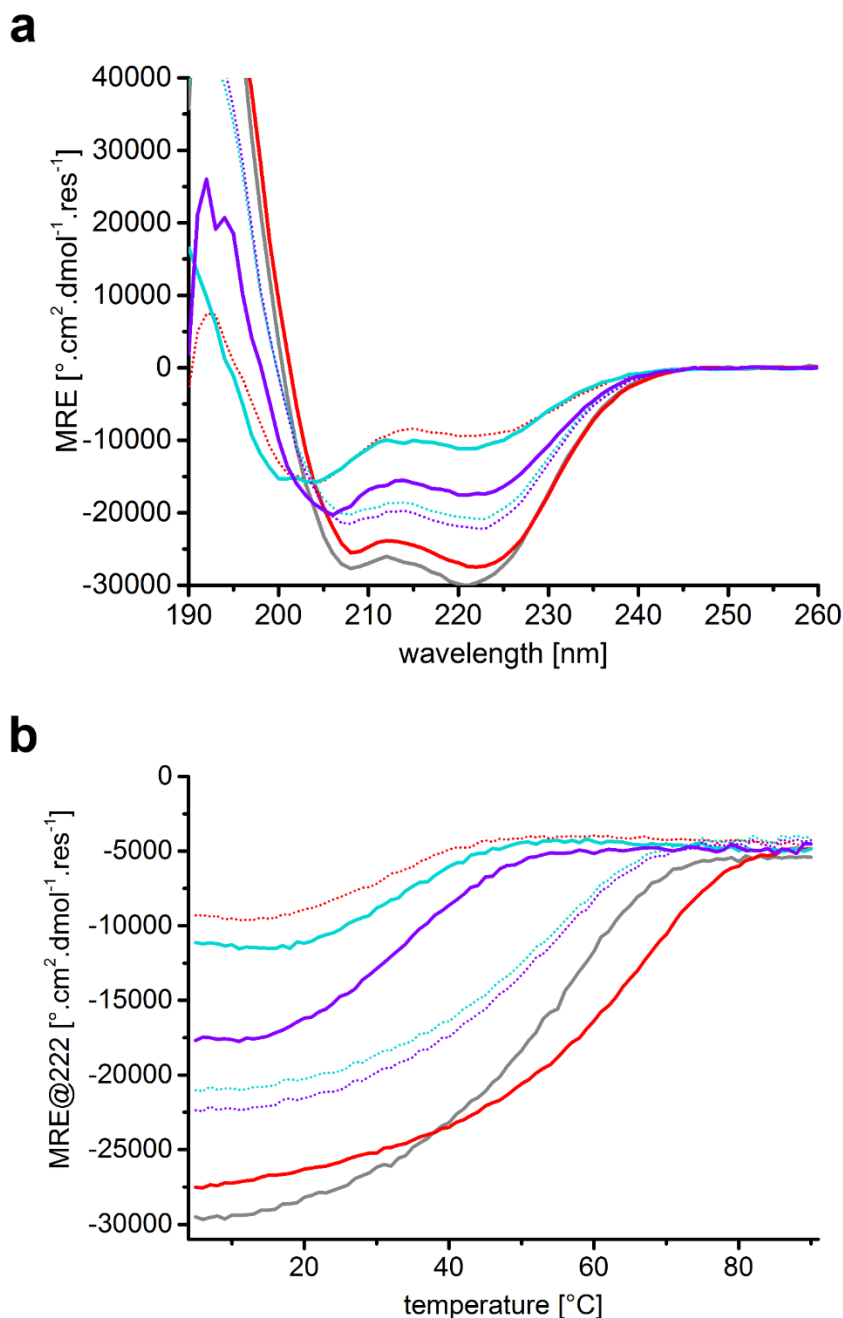* Na$^+$ peak mass (21 more than the H$^+$ peak)

Figure S2. Circular dichroism (CD) spectra and thermal unfolding curves for the homotrimeric coiled coils used as hub cores in this study. (a) CD spectra recorded at 5 °C between wavelengths of 190 – 260° (50 µM peptide in PI buffer with 250 µM TCEP and 25 mM NaCl). (b) The mean residue ellipticity (MRE) recorded at 222 nm as a function of temperature shows a sigmoidal thermal denaturation curve for each construct. Homotrimers shown: parent (CC-Tri3, **gray** solid line), $^{E4}3$ (E4-CC-Tri3, **red** solid line), 3^{E4} (CC-Tri3-E4, **red** dotted line), $^{K4}3$ (K4-CC-Tri3, **cyan** solid line), 3^{K4} (CC-Tri3-K4, **cyan** dotted line), $^{R4}3$ (R4-CC-Tri3, **violet** solid line), and 3^{R4} (CC-Tri3-R4, **violet** dotted line).

Table S5. Comparison of the proportion of α-helical residues predicted for each homotrimer sequence, which is based on the proportion of the sequence that should form a coiled-coil (Equation S2). The α-helicity as measured at 5 °C by CD spectroscopy. The experimental values were determined to the nearest 0.5% using the MRE at a wavelength of 222 nm at 5 °C. The α-helicities assume that for 100% α helix for a decorated homotrimer (31 amino acid chain), the MRE = -38400 cm^2 dmol^{-1} res^{-1} (25 amino acid parent MRE = -37400 cm^2 dmol^{-1} res^{-1}). The T$_M$ values were determined to the nearest 0.5 °C from the *x*-intercept of the second derivative of the MRE at a wavelength of 222 nm against temperature.

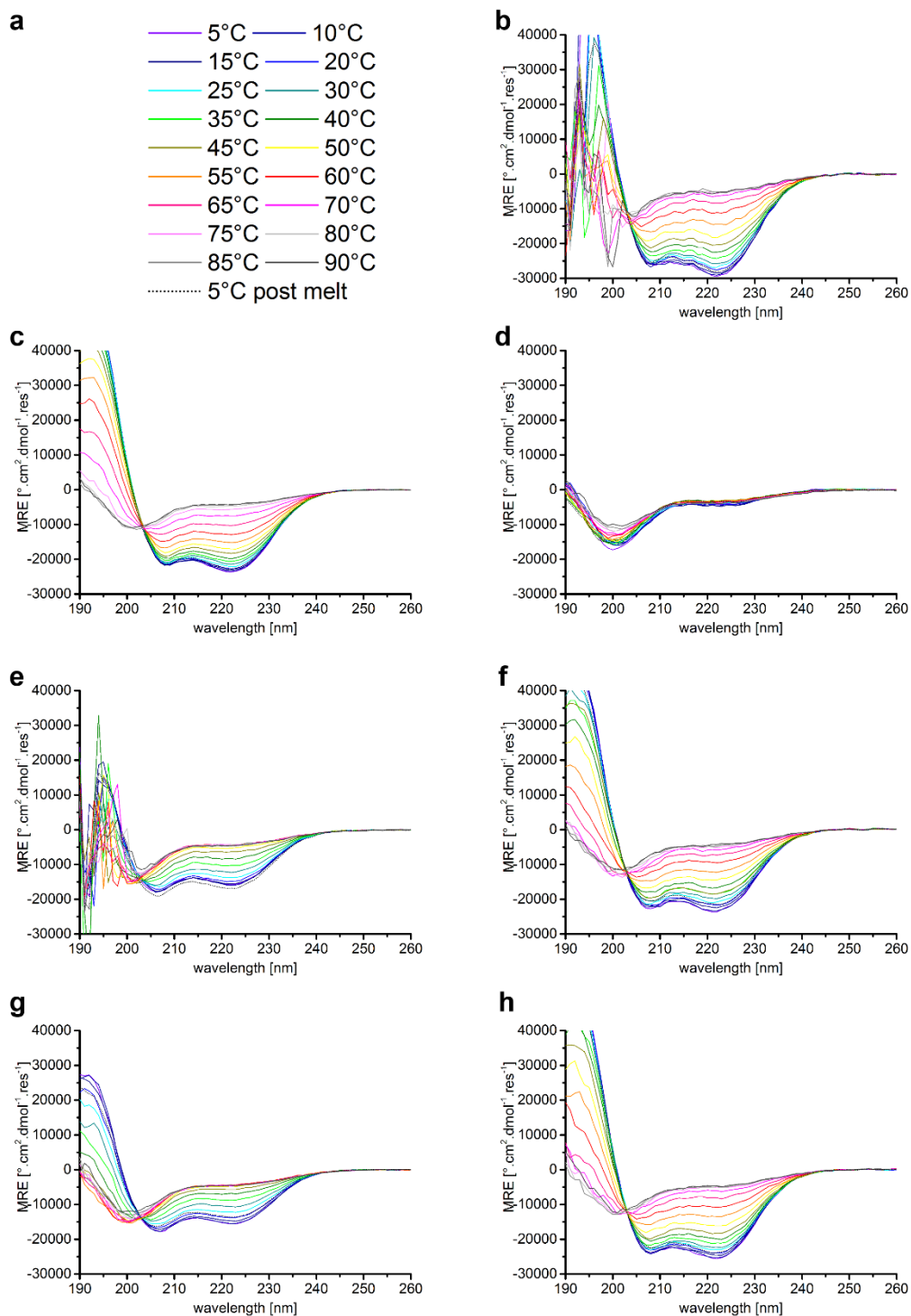|	homotrimer	predicted proportion of α-helical residues [%]	experimental α helicity at 5 °C [%]	melting transition midpoint, T$_M$ [°C]
3	CC-Tri3	84.0	81.5	56.0
E43	E4-CC-Tri3	67.5	72.0	65.5
3E4	CC-Tri3-E4	67.5	25.5	31.0
K43	K4-CC-Tri3	67.5	30.0	33.0
3K4	CC-Tri3-K4	67.5	58.0	54.0
R43	R4-CC-Tri3	67.5	46.0	34.0
3R4	CC-Tri3-R4	67.5	70.5	57.0

Figure S3. CD spectra recorded for homotrimers at 5 °C intervals between 5 °C – 90 °C, (50 µM peptide, PI buffer with 25 mM NaCl and 250 µM TCEP). (a) Key for the temperatures at which the CD spectra were recorded. CD spectra for homotrimers from: (b) parent (CC-Tri3), (c) ^{E4}3 (E4-CC-Tri3), (d) 3^{E4} (CC-Tri3-E4), (e) ^{K4}3 (K4-CC-Tri3), (f) 3^{K4} (CC-Tri3-K4), (g) ^{R4}3 (R4-CC-Tri3), and (h) 3^{R4} (CC-Tri3-R4).
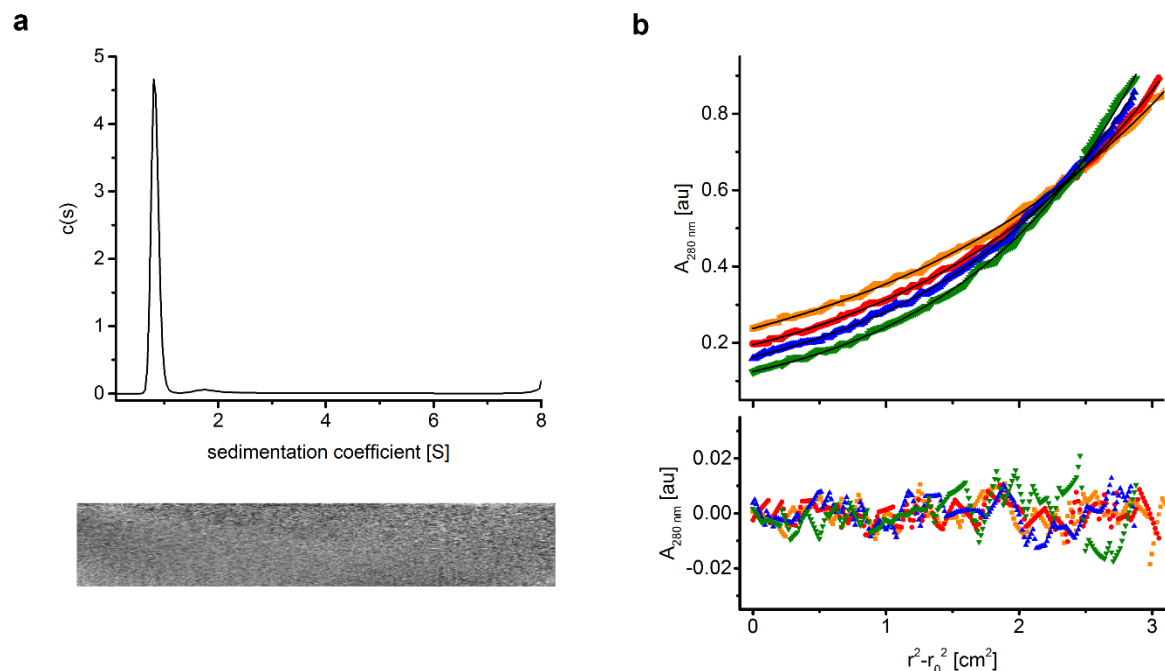
Figure S4. AUC data (upper) and residuals of fits (lower) for homotrimer peptides (325 µM peptide, 1.625 mM TCEP, PI buffer, 20 °C) used in this study. (left) Sedimentation velocity (SV) recorded at 60 krpm, fitted with a continuous c(s) distribution. (right) Sedimentation equilibrium (SE) plots show single ideal species fits (black lines), and data points recorded at 33 krpm (■), 36 krpm (●), 39 krpm (▲), and 42 krpm (▼). Details of masses determined and 95% confidence limits shown in Table S6.

(a) Parent homotrimer (CC-Tri3) ($\bar{v}$ = 0.76757) SV data returning s = 0.83 S, $s_{20,w}$ = 0.84 S, f/f₀ = 1.59 and a molecular weight of 9150 Da, consistent with a 3.5x monomer mass at 95% confidence level. SE fits showed a molecular weight of 7960 Da (3.0x monomer mass, 95% confidence limits 7920 – 7990 Da).
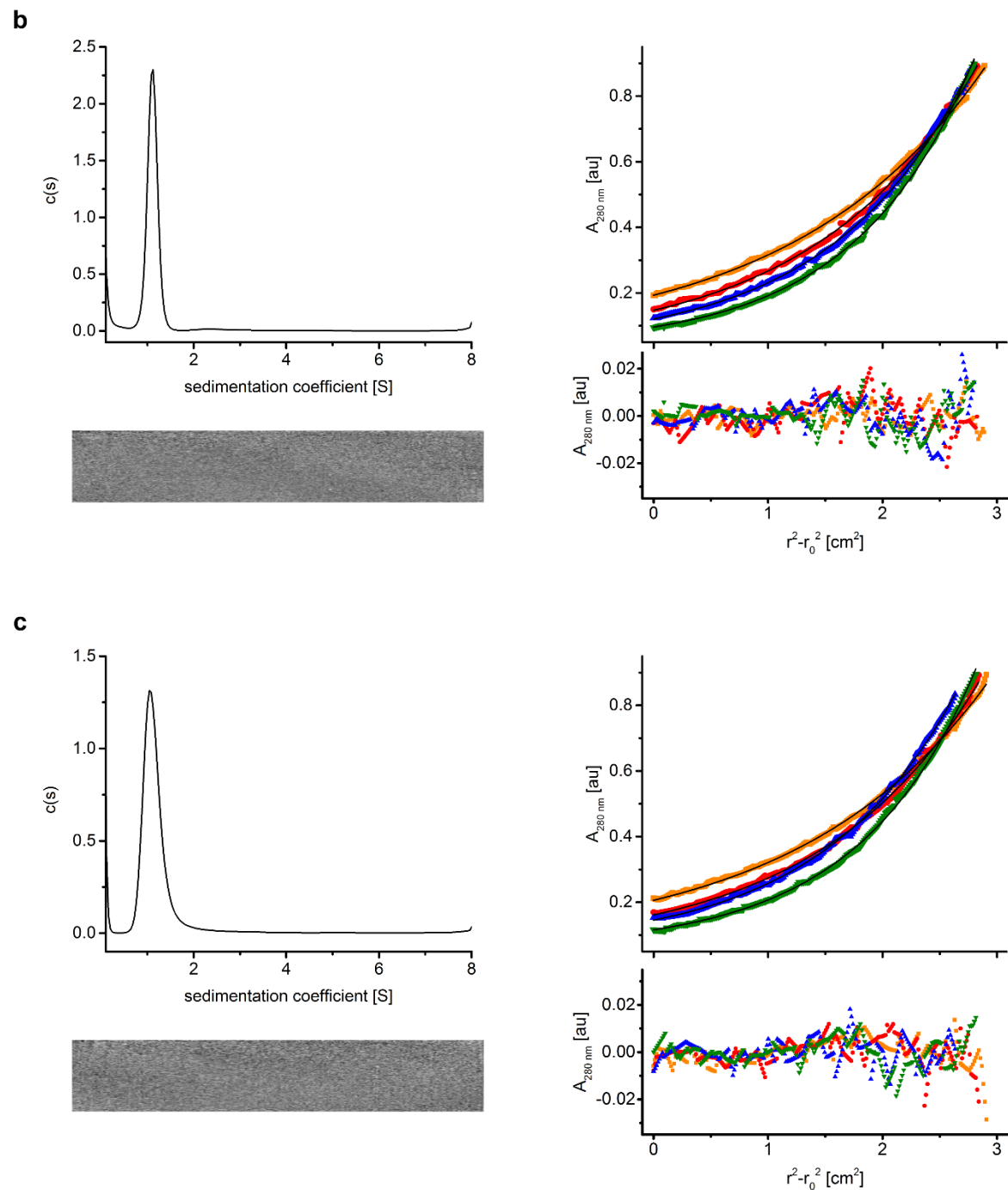
b

c

Figure S4 continued. Negatively charged E4 decorated homotrimer ($\overline{\nu}$ = 0.74996) (b) C-terminal E4 (E4-CC-Tri3) SV data returning s = 1.11 S, $s_{20,w}$ = 1.12 S, f/f$_0$ = 1.44 and a molecular weight of 10710 Da, consistent with a 3.3x monomer mass at 95% confidence level. SE fits showed 9710 Da (3.0x monomer mass, 95% confidence limits 9670 – 9720 Da). (c) N-terminal E4 (CC-Tri3-E4) SV data returning s = 1.15 S, $s_{20,w}$ = 1.16 S, f/f$_0$ = 1.41 and a molecular weight of 10910 Da, consistent with a 3.3x monomer mass at 95% confidence level. SE fits showed 9780 Da (3.0x monomer mass, 95% confidence limits 9730 – 9830 Da).
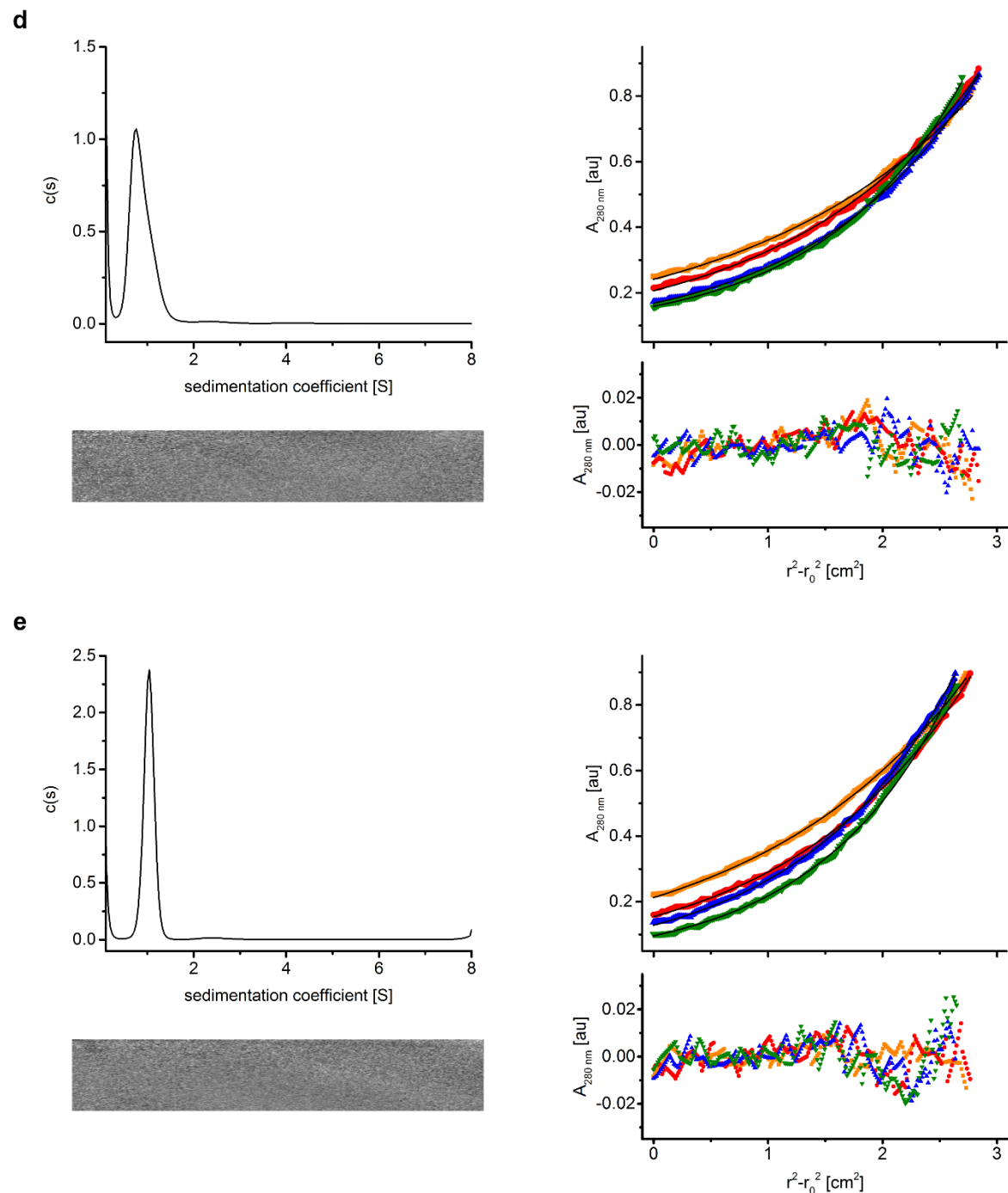
d

e

Figure S4 continued. Positively charged K4 decorated homotrimer ($\overline{v}$ = 0.74996) (d) C-terminal K4 (K4-CC-Tri3) SV data returning s = 0.87 S, $s_{20,w}$ = 0.90 S, f/f_0 = 1.57 and a molecular weight of 9070 Da, consistent with a 2.8x monomer mass at 95% confidence level. SE fits showed 8540 Da (2.6x monomer mass, 95% confidence limits 8480 – 8590 Da). (e) N-terminal K4 (CC-Tri3-K4) SV data returning s = 1.04 S, $s_{20,w}$ = 1.05 S, f/f_0 = 1.47 and a molecular weight of 10660 Da, consistent with a 3.3x monomer mass at 95% confidence level. SE fits showed 9040 Da (2.8x monomer mass, 95% confidence limits 8990 – 9090 Da).
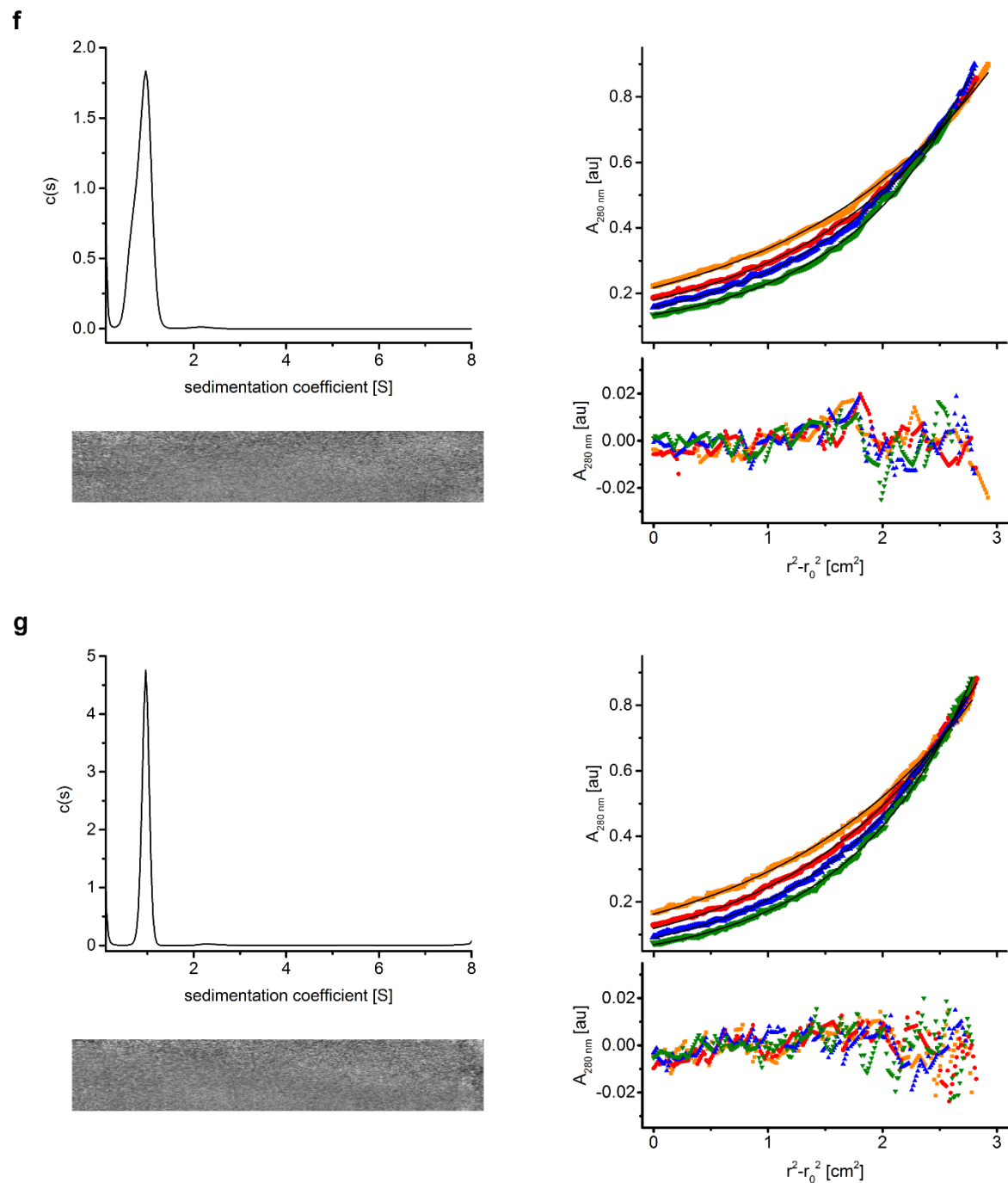
f

g

Figure S4 continued. Positively charged R4 decorated homotrimer ($\bar{v} = 0.74996$) (f) C-terminal R4 (R4-CC-Tri3) SV data returning s = 0.89 S, $s_{20,w}$ = 0.90 S, f/f$_0$ = 1.65 and a molecular weight of 9480 Da, consistent with a 2.8x monomer mass at 95% confidence level. SE fits showed 9180 Da (2.7x monomer mass, 95% confidence limits 9120 – 923 Da). (e) N-terminal R4 (CC-Tri3-R4) SV data returning s = 0.96 S, $s_{20,w}$ = 0.97 S, f/f$_0$ = 1.68 and a molecular weight of 10920 Da, consistent with a 3.2x monomer mass at 95% confidence level. SE fits showed 9270 Da (2.8x monomer mass, 95% confidence limits 9220 – 9320 Da).

Table S6. Summary of analytical ultracentrifugation (AUC) data used to determine the mass of each homotrimer used in this study. Masses are quoted to nearest 10 Da, sizes are quoted to nearest 0.1x monomer mass, $\bar{v}$ values were calculated using SEDNTERP.[8] Plots of data, fits and residuals are shown in Figure S4.

| homotrimer name	trimer mass [Da]	$\bar{v}$ [cm³g⁻¹]	mass SV [Da]	size SV [x mono]	mass SE [Da]	size SE [x mono]	95% confidence [Da]
parent CC-Tri3	7970	0.76757	9150	3.5	7960	3.0	7920– 7990
E43 E4-CC-Tri3	9780	0.74922	10710	3.3	9710	3.0	9670– 9720
3**E4** CC-Tri3-E4	9780	0.74992	10910	3.3	9780	3.0	9730– 9830
K43 K4-CC-Tri3	9770	0.75899	9070	2.8	8540	2.6	8480– 8590
3**K4** CC-Tri3-K4	9770	0.75899	10660	3.3	9040	2.8	8990– 9090
R43 R4-CC-Tri3	10110	0.74996	9480	2.8	9180	2.7	9120– 9230
3**R4** CC-Tri3-R4	10110	0.74996	9270	2.7	9270	2.8	9220– 9320
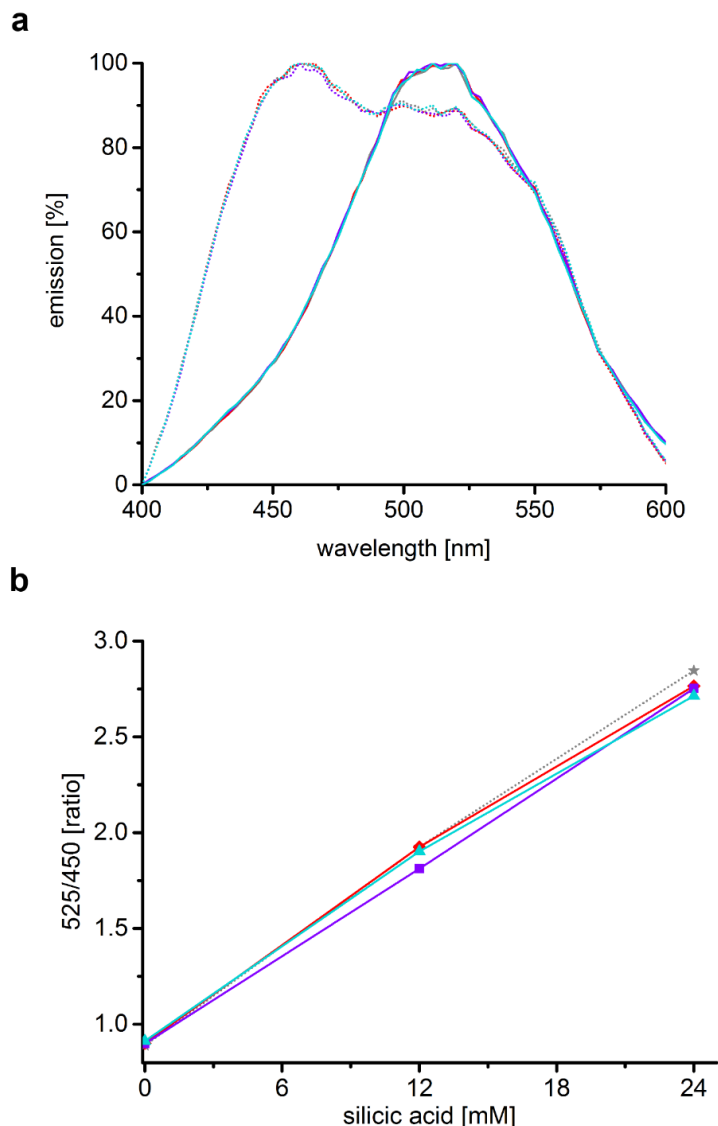
Figure S5. Fluorescence spectra from PDMPO in mineralization reactions recorded 24 hours after addition of silicic acid. Peptide assembly (2 µM peptide, assembled for 1 hour in PI buffer) followed by mineralization (24 mM silicic acid, 24 hours, 20 °C), with 0.5 µL PDMPO indicator in 100 µL reaction volume. (a) Fluorescence spectra recorded with no silicic acid (dotted lines, unquenched emission) and 24 mM silicic acid (solid lines, quenched at $\lambda = 450$ nm); parent SAGE (**grey**), E4-SAGE (**red**), K4-SAGE (**cyan**), and R4-SAGE (**violet**). (b) Ratio of emission at a wavelength of $\lambda = 525$ nm to 450 nm; no peptide (**grey** dotted line with ★ star), parent (**grey** solid line with ◀ left arrow), E4-SAGE (**red** solid line with ♦ diamond), K4-SAGE (**cyan** solid line with ▲ triangle), and R4-SAGE (**violet** solid line with ■ square). The PDMPO indicates that there is the same amount of silica precipitated in the presence or absence of peptide, as the emission at 450 nm is quenched to the same extent for each of the peptides tested.
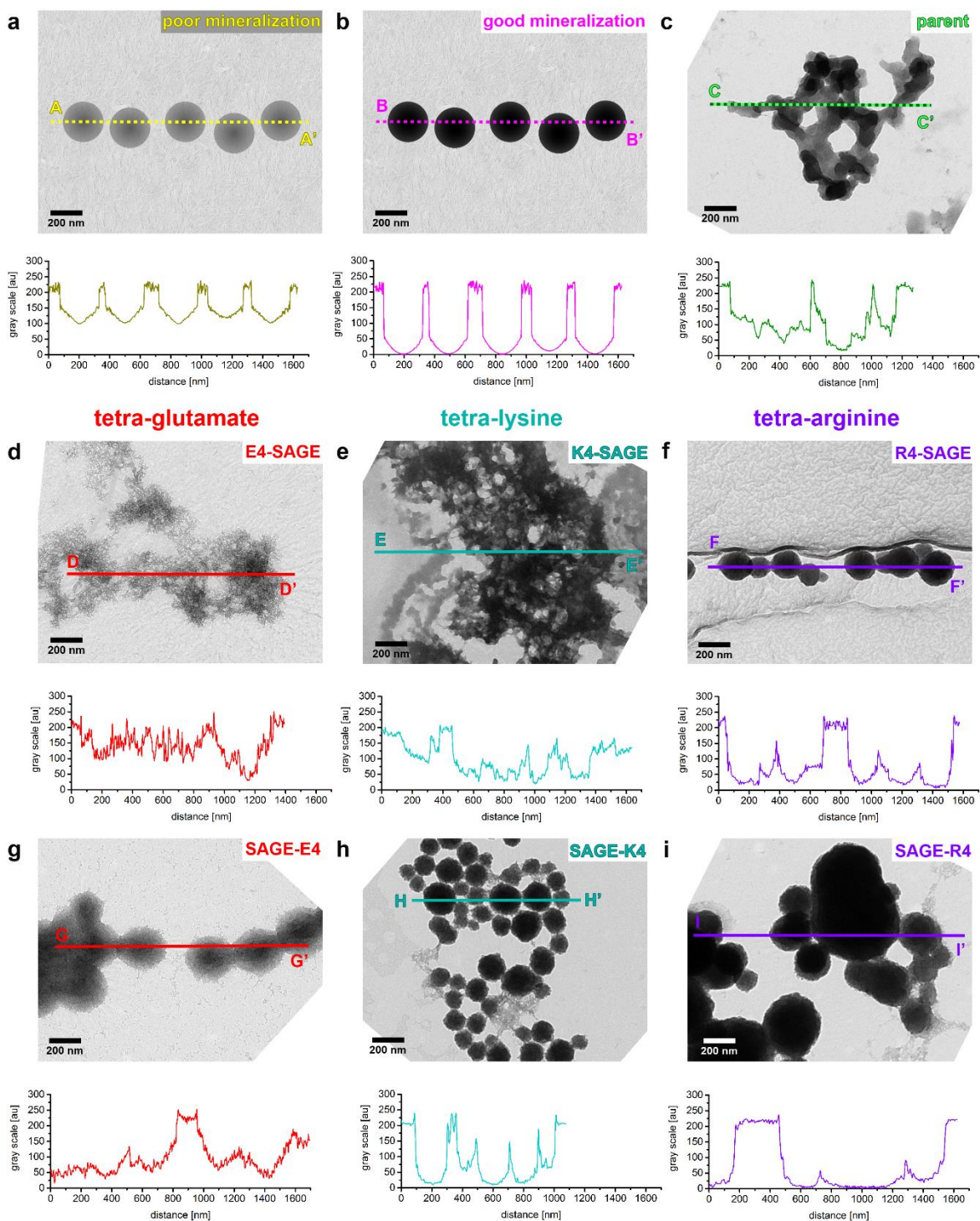
Figure S6. Transmission electron microscope (TEM) images and corresponding line profiles. TEM images (top) are the same as those shown in Figure 2, but they were rotated to align with the line profile (below), which shows the gray scale variation between X and X' from each image. Diagram to show expected image and line profile for (a) a poorly mineralized sample (particles with low electron density) and (b) a well mineralized sample (particles with high electron density). (c) Parent SAGE; and N-terminal decorated (d) E4-SAGE, (e) K4-SAGE, (f) R4-SAGE; and C-terminal decorated (g) SAGE-E4, (h) SAGE-K4 and (i) SAGE-R4. Mineralization conditions: 2 μM peptide, 24 mM silicic acid, 24 hours, PI buffer, 20 °C.
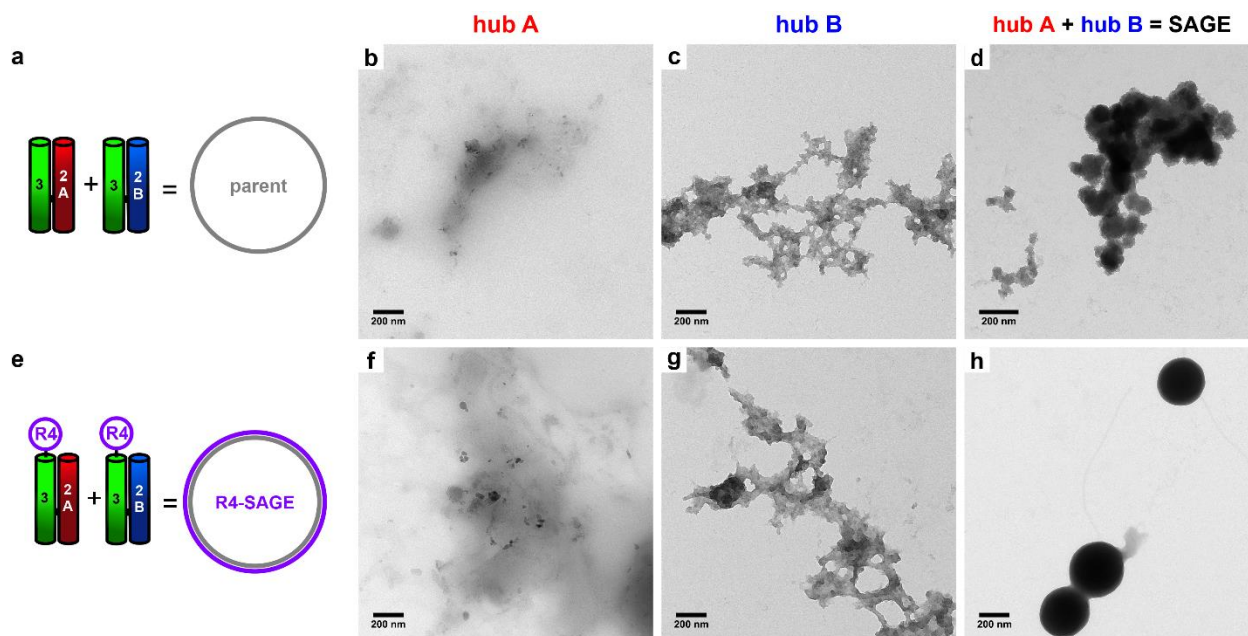
Figure S7. Representative TEM images of SAGE modules and assembled SAGEs after silicification (2 µM peptide, 24 mM silicic acid, 24 hours, PI buffer, 20 °C). (a) Diagram representing the component (left) and assembled (right) parent SAGE. The parent SAGE peptides after silicification of (b) parent hub A (CC-Tri3-CC-Di-A), (c) parent hub B (CC-Tri3-CC-Di-B) and (d) the assembled parent SAGE structures. (e) Diagram representing the component (left) and assembled (right) R4-SAGE. The N-terminal R4 decorated SAGE peptides after silicification of (f) ^{R4}3-A (R4-CC-Tri3-CC-Di-A), (d) ^{R4}3-B (R4-CC-Tri3-CC-Di-B) and (f) the assembled R4-SAGE structures.
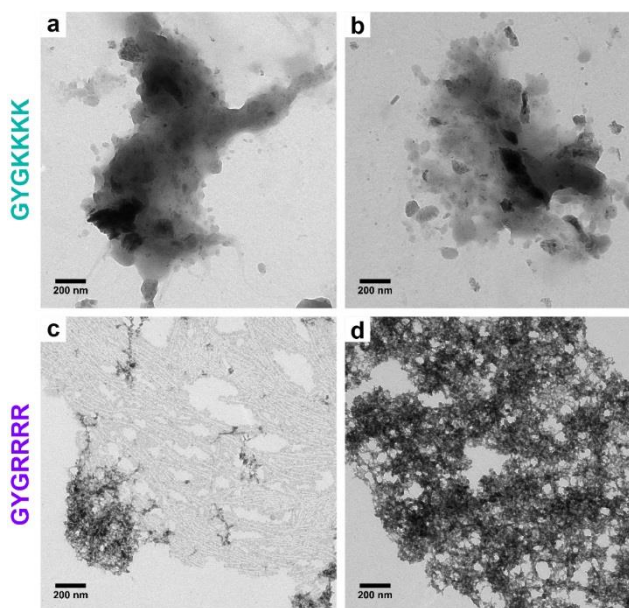

Figure S8. Representative TEM images of silica structures precipitated in the presence of positively charged control peptides (*i.e.* not appended to SAGE coiled-coil peptides, 2 µM peptide, 24 mM silicic acid, PI buffer, 24 hours, 20 °C). (a & b) GYGKKKK (GYG^{K4}) and (c & d) GYGRRRR (GYG^{R4}). No spherical structures are visible in these samples, indicating that the charged peptides need to be attached to the SAGE hubs in order to form self-assembled biomineralized spherical structures.
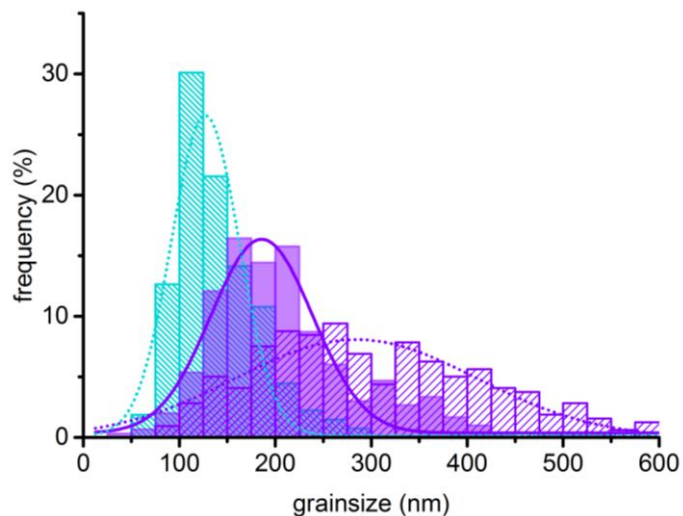
Figure S9. Grainsize plots and fits of positively charged SAGEs (2 µM peptide, 24 mM silicic acid in PI buffer, 20 °C). Plot shows N-terminal R4-SAGE (**violet** fill, **violet** solid line), C-terminal SAGE-R4 (**violet** diagonal fill, **violet** dotted line), and C-terminal SAGE-K4 (**cyan** diagonal fill, **cyan** dotted line). Summary of fit data shown in Table S7.
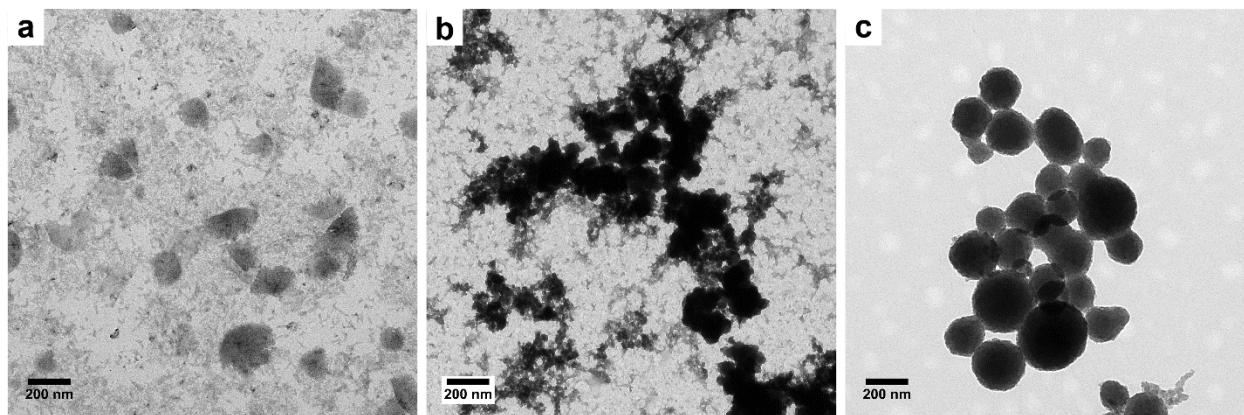

Figure S10. Representative transmission electron microscope (TEM) images of N-terminal tetra-arginine decorated R4-SAGEs. (a) Unstained and unmineralized R4-SAGEs are difficult to see as they are not very electron dense. (b) Negative staining with 1% uranyl acetate stains surrounding material (*e.g.* salt crystals, the film on the grid) as well as the R4-SAGEs, making the SAGEs difficult to discern. (c) SiO₂-R4-SAGEs (mineralization conditions: 2 µM peptide, 24 mM silicic acid, 24 hours, PI buffer, 20 °C) are easy to see as they are electron dense and not surrounded by other contaminating material.
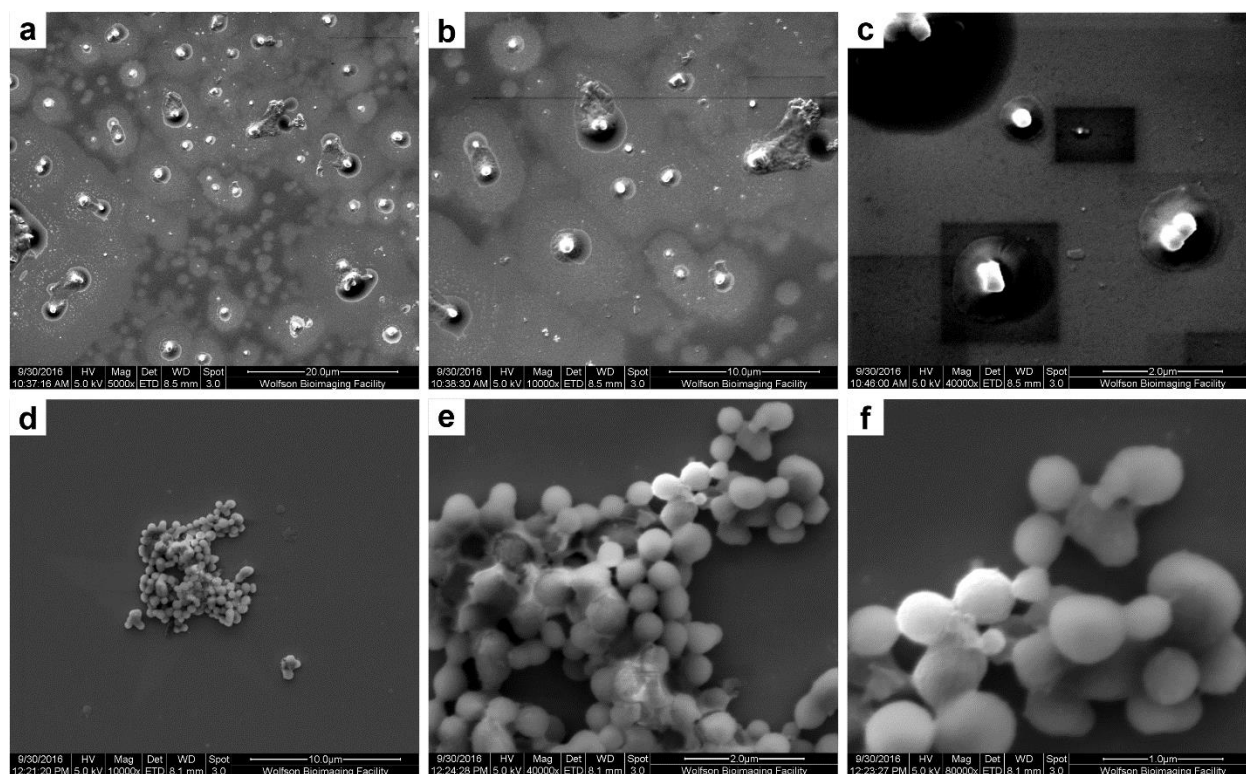
Figure S11. Representative scanning electron microscope (SEM) images of N-terminal R4-SAGEs. Top panels (a – c) show unmineralized sample sputter coated with a thin (≈5 nm) conductive metal layer for imaging at three magnifications. It was difficult to achieve high-magnification images of sputter coated samples due to: electron beam (e-beam) damage (dark rectangles), and charging (white, high contrast particles), see (c). Lower panels (d – f) show mineralized SiO₂-R4-SAGEs (mineralization conditions: 2 µM peptide, 24 mM silicic acid, 24 hours, PI buffer, 20 °C). The silica coated SAGEs produce clearer images at higher magnifications, as e-beam damage and charging minimal in the biomineralized samples.
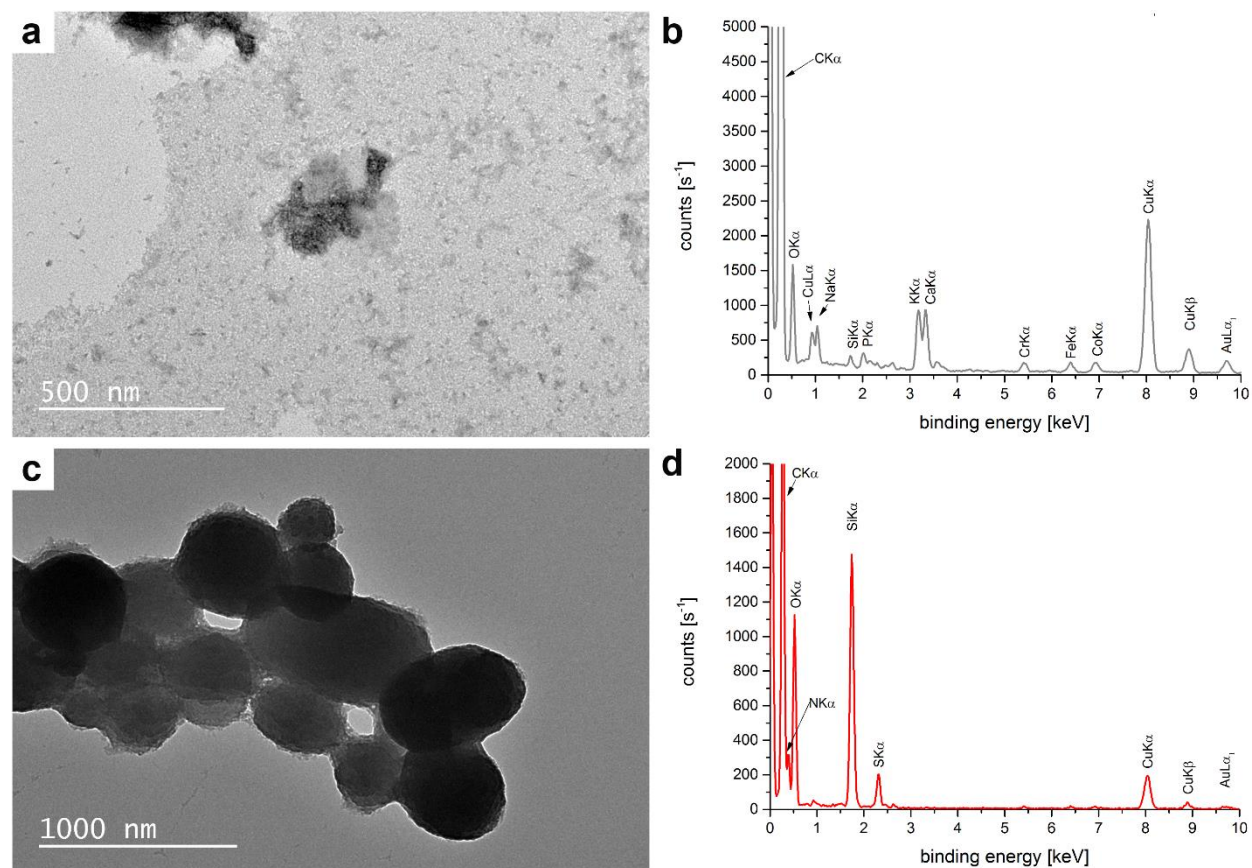
Figure S12. Representative TEM images and energy dispersive X-ray (EDX) spectra of R4-SAGEs. TEM images of (a) unmineralized R4-SAGE and (b) an EDX spectrum from the center of the image. (c) SiO$_2$-R4-SAGEs (mineralization conditions: 2 μM peptide, 24 mM silicic acid, 24 hours, PI buffer, 20 °C) and (d) an EDX spectrum from the center of the image. Both spectra show contributions from the sample holder and grid:

Cu Lα = 0.93 keV, Cu Kα = 8.05 keV and Cu Kβ = 8.90 keV,
Si Kα = 1.74 keV, Cr Kα = 5.41 keV, Fe Kα = 6.40 keV, Co Kα = 6.93 keV, and Au Lα = 9.71 keV;

buffer:
Na Kα = 1.04 keV, P Kα = 2.01 keV, K Kα = 3.31 keV and Ca Kα = 3.69 keV

and peptide:
(C Kα = 0.28 keV, N Kα = 0.39 keV, and O Kα = 0.53 keV).

(d) EDX spectrum from 24 mM silicic acid SiO$_2$-R4-SAGEs shows proportionally a much higher detection of silicon (Si Kα = 1.74 keV) and oxygen (O Kα = 0.53 keV) from the SiO$_2$-R4-SAGEs than observed in the unmineralized R4-SAGE sample (spectrum shown in (b)).
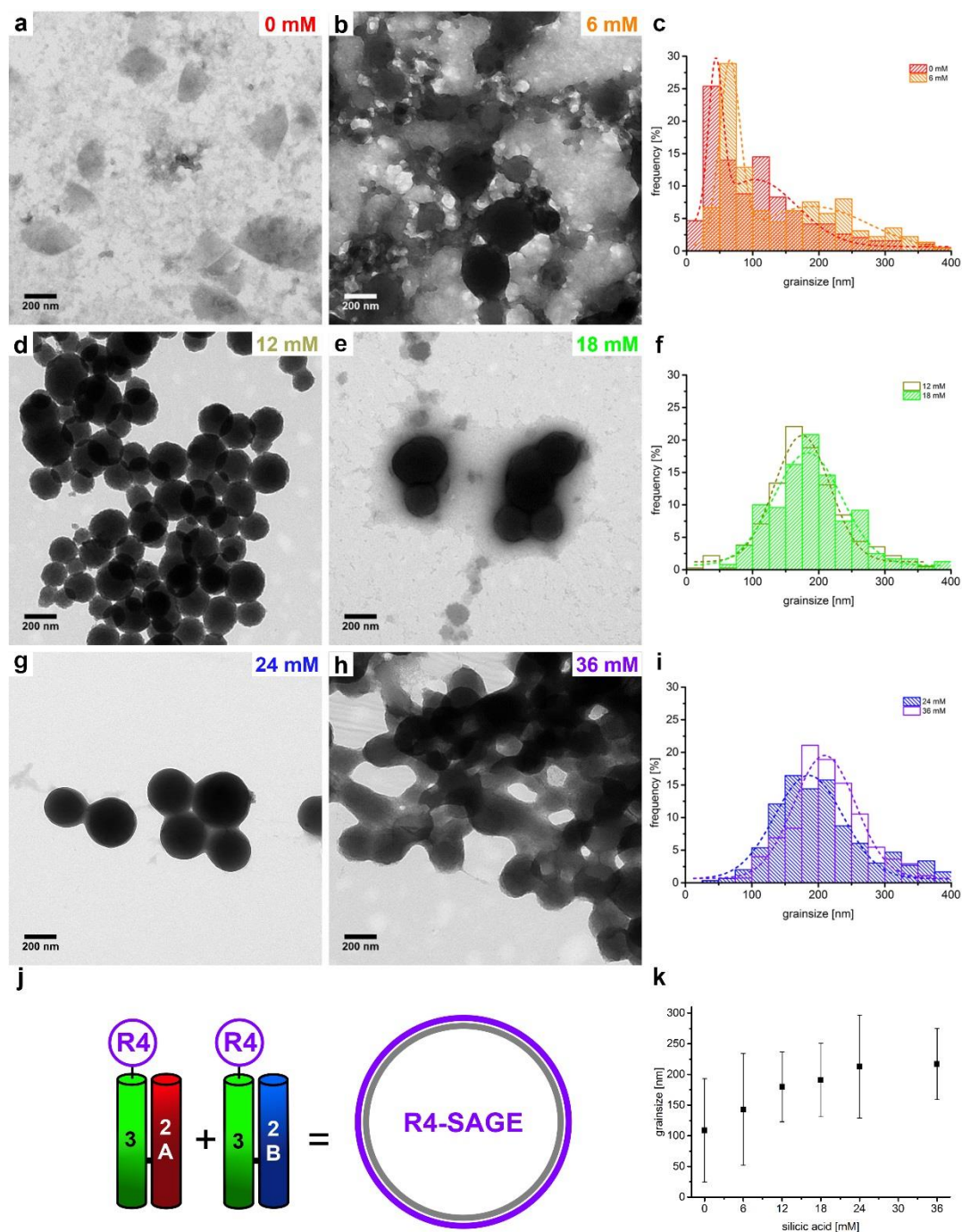
Figure S13. Representative TEM images of R4-SAGEs mineralized with different concentrations (0 mM – 36 mM) of silicic acid (2 µM peptide, 24 hours, PI buffer, 20 °C), and their corresponding grainsize analyses. R4-SAGEs incubated with (a) 0 mM (**red**) and (b) 6 mM (**orange**) silicic acid show (c) a bimodal, polydispersed and possibly asymmetric grainsize profile. (d) 12 mM (**dark yellow**) and (e) 18 mM (**green**) silicic acid show (f) a monodispersed grainsize distribution with smaller sized particles. (g) 24 mM (**blue**) and (h) 36 mM (**violet**) silicic acid show (i) a slightly broader grainsize distribution with larger sized particles, with the highest silicic acid concentration (36 mM) showing clumped together SiO₂-R4-SAGE particles. (j) Diagram of R4-SAGE hub components, and (k) graph summarizing these grainsize data (values shown in Table S7).

Table S7. Summary of grainsize data for the range of silicic acid concentrations used to mineralize R4-SAGEs, graph of these data shown in Figure S13. Also grainsize data for N- or C-terminal SAGE hub functionalization, graph of these data shown in Figure S13.

SAGE type	silicic acid concentration [mM]	average [nm]	1 standard deviation [nm]	peak center particle size [nm]	error, FWHM of fit [nm]	number of particles sized
R4-SAGE	0	109	84	44 / 103	21 / 126	193
R4-SAGE	6	143	91	65 / 190	31 / 170	225
R4-SAGE	12	180	57	175	84	367
R4-SAGE	18	191	60	183	103	240
R4-SAGE	24	213	84	186	103	298
SAGE-R4	24	305	120	284	251	319
SAGE-K4	24	139	42	127	71	269
R4-SAGE	36*	217	58	210	93	275

---

[*] The interconnected nature of the 36 mM particles meant that only those particles at the edges of the imaged assemblies, *i.e.* where the edges of the particles could be seen, were measured.
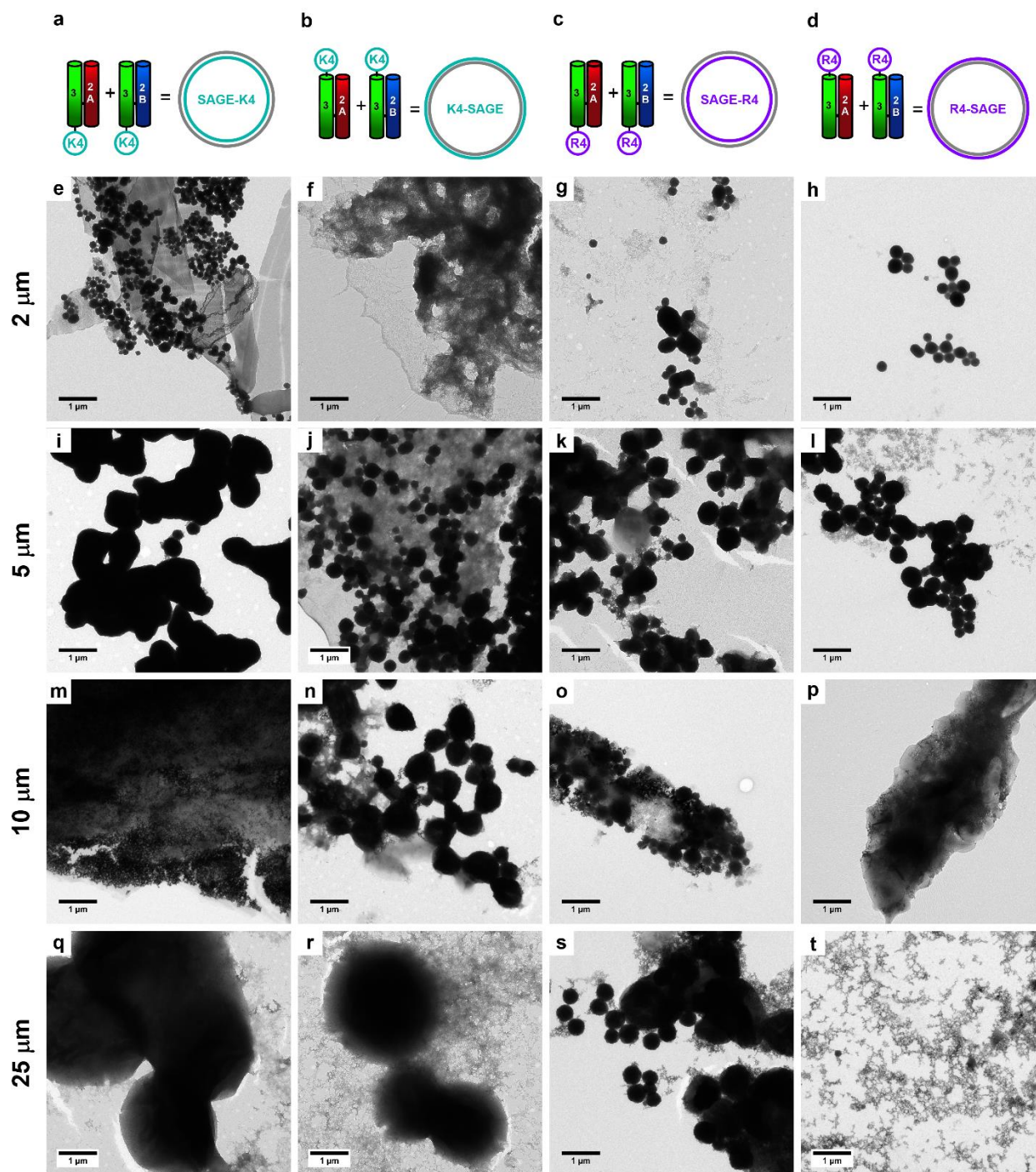
Figure S14. Representative TEM images of positively charged SAGEs assembled at a range of peptide concentrations (2 µM – 25 µM, mineralized using: 24 mM silicic acid, 24 hours, PI buffer, 20 °C). Diagrams to show SAGE module components for SAGEs formed from positive charge decorated homotrimers. (a) C-terminal K4 (SAGE-K4), (b) N-terminal K4 (K4-SAGE), (c) C-terminal R4 (SAGE-R4), and (d) N-terminal R4 (R4-SAGE). Peptide modules A & B were mixed and assembled for 1 hour at 20 °C in PI buffer at concentrations of (e – h) 2 µM, (i – l) 5 µM, (m – p) 10 µM, and (q – t) 25 µM.
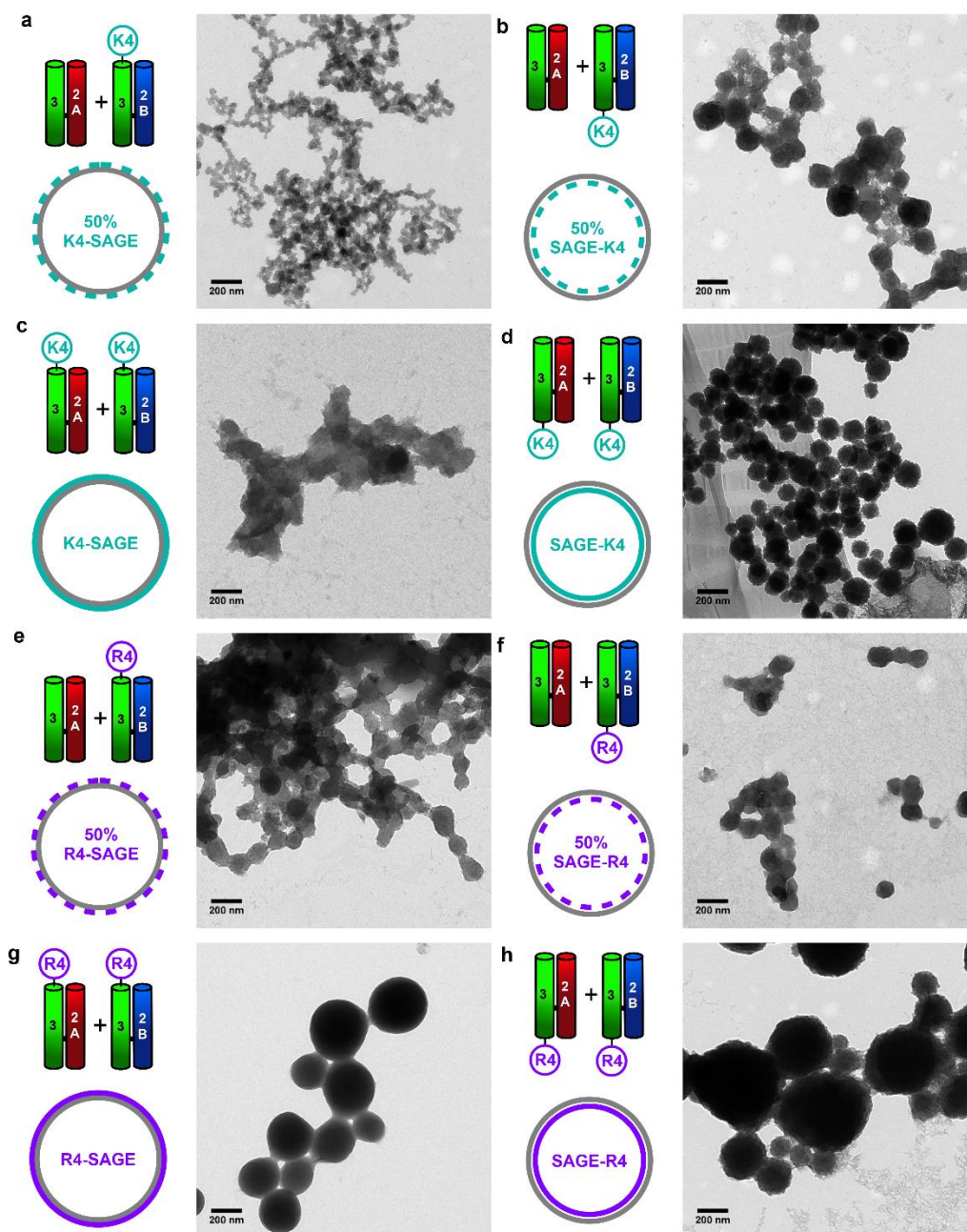
Figure S15. Representative TEM images of the silicified structures formed when the ratio of positively charged hub to parent hub is altered, and diagrams to represent the SAGE assemblies. SAGEs were assembled (2 µM hub peptide, 1 hour assembly in PI buffer, 20 °C) and silicified (24 mM silicic acid, 24 hours, PI buffer, 20 °C). (a) SiO₂-SAGEs formed with 50% of the hubs decorated with N-terminal K4 (50% K4-SAGE) and (b) 50% C-terminal K4 (50% SAGE-K4). (c) 100% N-terminal K4 (K4-SAGE). (d) 100% C-terminal K4 (SAGE-K4). (e) SiO₂-SAGEs formed with 50% of the hubs decorated with N-terminal R4 (50% R4-SAGE) and (f) 50% C-terminal R4 (50% SAGE-R4). (g) 100% N-terminal R4 (R4-SAGE). (h) 100% C-terminal R4 (SAGE-R4).
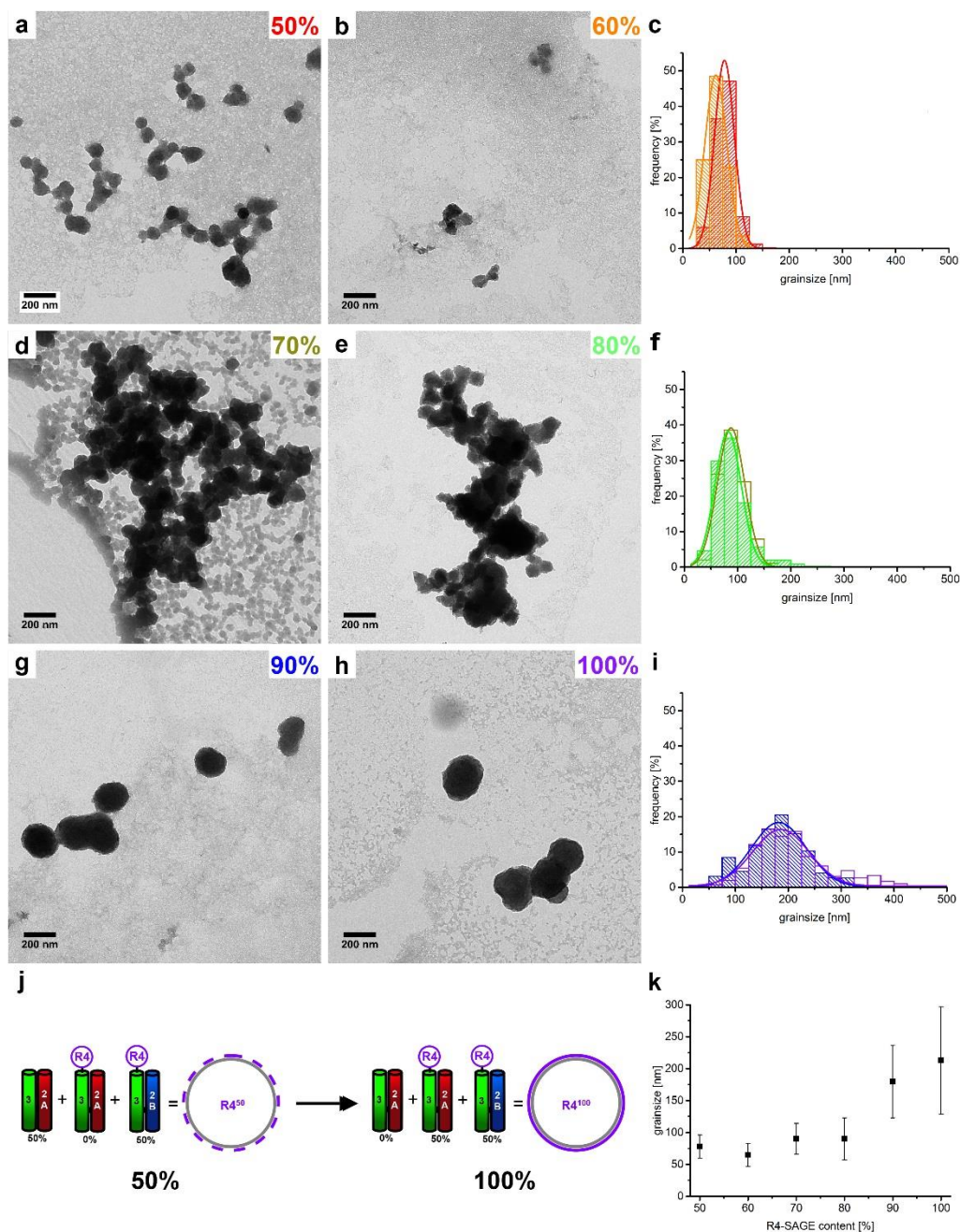
Figure S16. Representative TEM images of SiO$_2$-R4-SAGEs (2 µM peptide, 24 mM silicic acid, PI buffer, 24 hours, 20 °C) doped with different levels (range 0% to 50%) of the undecorated parent SAGE hub A. At higher proportions of the parent hub A, (a) 50% R4-SAGE (0% R4-CC-Di-A + 50% CC-Di-A + 50% R4-CC-Di-B), (b) 60% R4-SAGE (10% R4-CC-Di-A + 40% CC-Di-A + 50% R4-CC-Di-B) and (d) 70% R4-SAGE (20% R4-CC-Di-A + 30% CC-Di-A + 50% R4-CC-Di-B), small, interconnected structures were formed. At (e) 80% R4-SAGE (30% R4-CC-Di-A + 20% CC-Di-A + 50% R4-CC-Di-B), a mixture of smaller and larger particles were observed. Larger individual discrete spheres were biomineralized at (g) 90% R4-SAGE (40% R4-CC-Di-A + 10% CC-Di-A + 50% R4-CC-Di-B) and (h) 100% R4-SAGE (0% R4-CC-Di-A + 0% CC-Di-A + 50% R4-CC-Di-B). Quantification of the grainsize (c, f, i, & k), is also summarized in Table S8; and (j) diagrams representing the proportion of parent and R4 module used to assemble 50% R4 and 100% R4-SAGEs.

Table S8. Summary of grainsize data for the SiO$_2$-SAGEs assembled with a range of N-terminal R4 decorated hubs between 50-100% R4-SAGE shown in Figure S16. The peak center is the center of the Gaussian fit to the measured grainsize distribution. The error on the grainsize quoted is the full width half maximum (FWHM) of the fitted data.

R4-SAGE hub decoration [%]	average [nm]	1 standard deviation [nm]	peak center [nm]	error, FWHM of fit [nm]	number of particles sized
50*	78	18	78	37	323
60*	65	18	62	41	336
70*	90	24	88	51	379
80*	90	33	82	49	348
90	180	57	182	103	224
100	213	84	186	106	298

---

* The interconnected nature of the 50 – 80% R4-SAGE particles means that only those particles at the edges of the imaged assemblies, *i.e.* where the edges of the particles could be seen, could be measured.
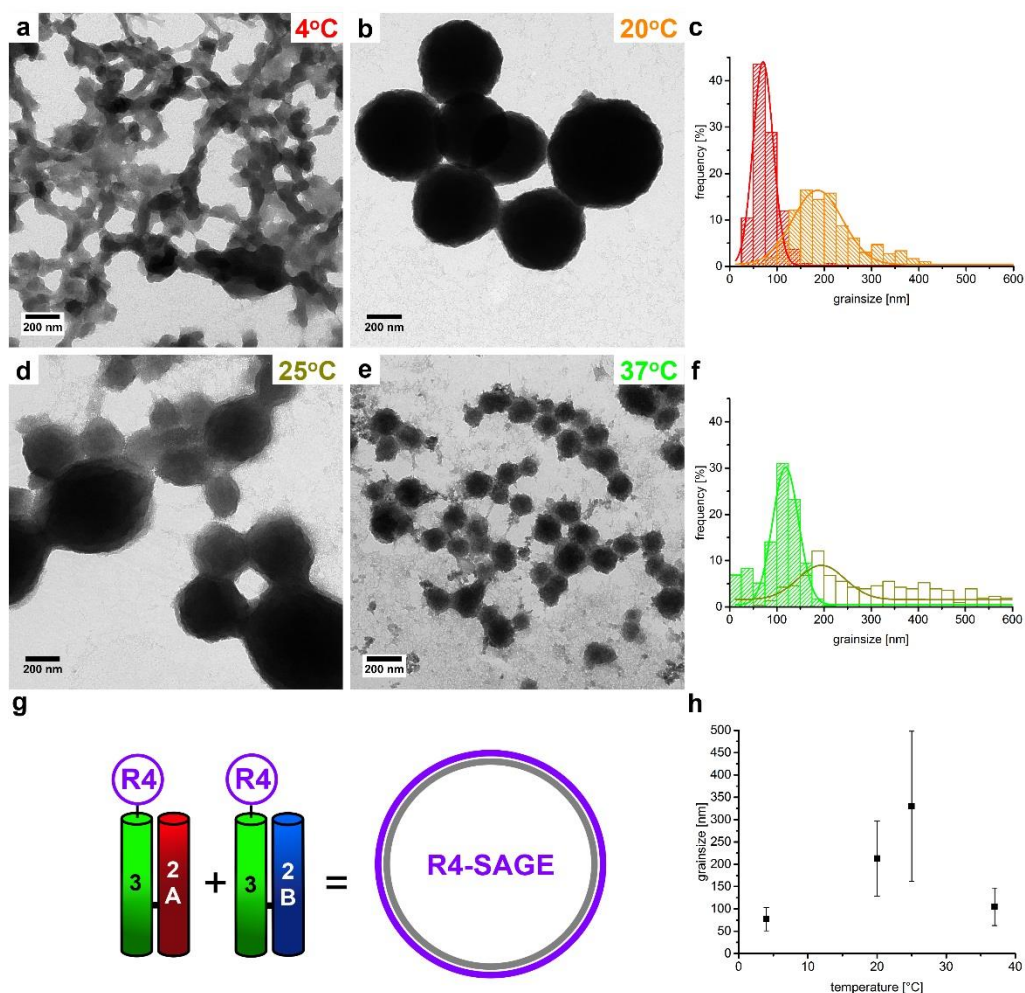
Figure S17. Representative TEM images to show the structures formed by silicification of R4-SAGEs assembled at a range of temperatures. SiO$_2$-R4-SAGEs were assembled and mineralized (2 µM peptide 1 hour, 24 mM silicic acid, 24 hours, PI buffer) at between 4 °C and 37 °C. (a) When cooled (4 °C), small interconnected structures were formed, whereas (b) at room temperature (20 °C), discrete spheres were formed, and (c) grainsize and fits. (d) At 25 °C, larger irregularly sized interconnected spheres were formed, then at (e) 37 °C smaller spheres connected by a fine network were formed, (f) grainsize and fits. (g) Diagram to represent R4-SAGE assembly, and (h) grainsize and distribution of fits (Table S9).

Table S9. Summary of grainsize data for SiO$_2$-R4-SAGEs assembled and mineralized at a range of temperatures (4 - 37 °C), Figure S17. The peak center is the center of the Gaussian fit to the measured grainsize distribution. The error quoted is the full width half maximum (FWHM) of the fitted data.

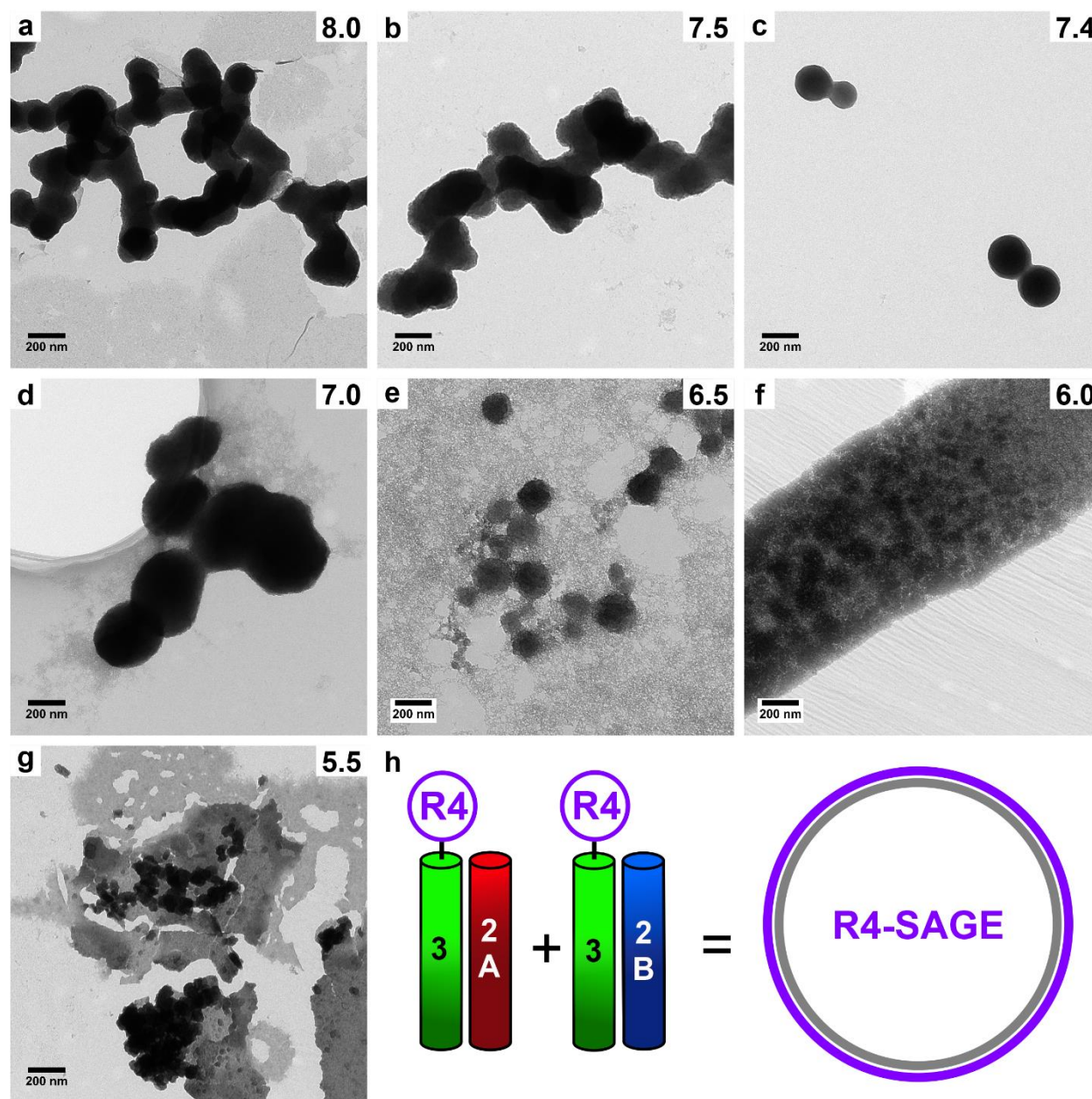| SiO$_2$-R4-SAGE temperature [°C]	average [nm]	1 standard deviation [nm]	peak center [nm]	error, FWHM of fit [nm]	number of particles counted
4	77	26	71	42	326
20	213	84	186	106	298
25	330	168	195	100	307
37	105	42	119	56	349

Figure S18. Representative TEM images of SiO$_2$-R4-SAGEs assembled and silicified (assembled 2 μM peptide, 1 hour, 20 °C) and mineralized (24 mM silicic acid, 24 hours, 20 °C) at different pH in PI buffer. At (a) pH 8.0 and (b) pH 7.5 show that interconnected networks of SiO$_2$-R4-SAGEs were formed. At (c) pH 7.4, individual silicified spheres were produced. The SiO$_2$-R4-SAGEs formed at (d) pH 7.0 were slightly more interconnected than at pH 7.4. At lower pH, the spheres were embedded in a network at (e) pH 6.5, there was more network and fewer spheres at (f) pH 6.0. (g) At pH 5.5, only sheets of silica were precipitated. The specificity of the homotrimer to itself and the heterodimer to its partner relies on interactions between charged residues at the e and g positions of the coiled-coil sequences. Thus, altering pH would be expected to alter these electrostatic interactions, and thus disrupt the self-assembly of the SAGE peptides.
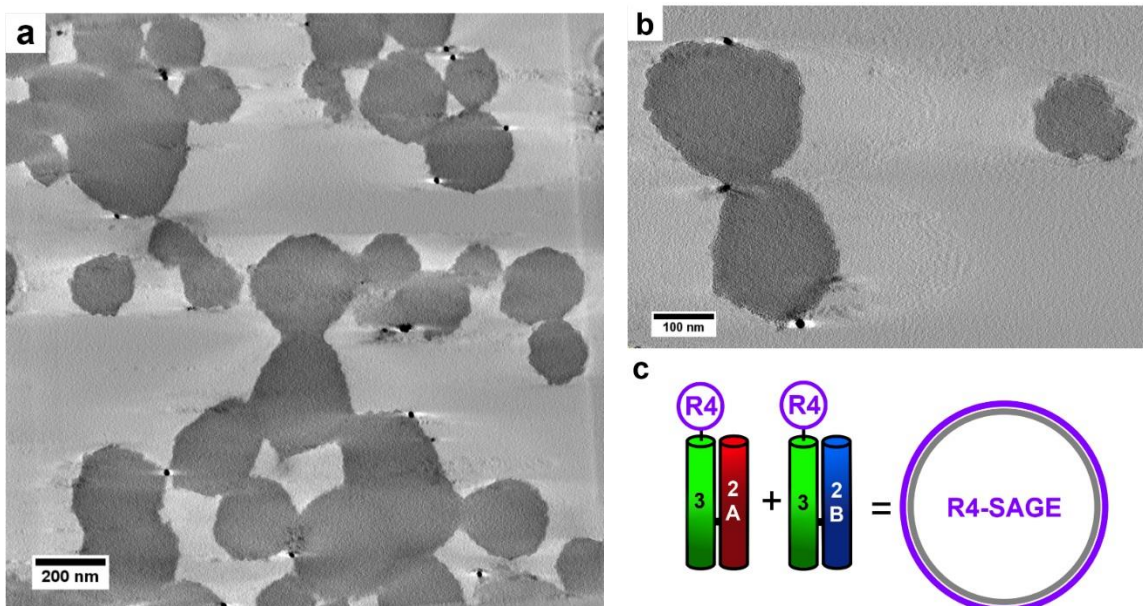
Figure S19. Slices from tomograms reconstructed in Movie S1 – S4. (a) Slice from center of tomogram recorded at 29,000x magnification (Movie S1) of SiO₂-R4-SAGEs mineralized (2 µM peptide, 12 mM silicic acid, 24 hours, PI buffer, 20 °C) with 15 nm diameter gold fiducials added to the grid for alignment, and (b) 50,000x magnification (Movie S2, 3D reconstructions in Movie S3 - S4). These reconstructions show that the dried SiO₂-R4-SAGEs do not collapse, but maintain their 3D shape. (c) Diagram to show R4-SAGE hub components.

Supplementary Movie Captions

Movie S1. Tomographic reconstruction of SiO₂-R4-SAGEs mineralized with 12 mM silicic acid at 29,000x magnification (15 nm diameter gold fiducials, scaled slice from tomogram shown in Figure S19).

Movie S2. Tomographic reconstruction of SiO₂-R4-SAGEs mineralized with 12 mM silicic acid at 50,000x magnification (15 nm diameter gold fiducials, scaled slice from tomogram shown in Figure S19).

Movie S3. 3D reconstruction of SiO₂-R4-SAGEs biomineralized with 12 mM silicic acid at 50,000x magnification (15 nm diameter gold fiducials, equatorial rotation, tomogram shown in Movie S2).

Movie S4. 3D reconstruction of SiO₂-R4-SAGEs biomineralized with 12 mM silicic acid at 50,000x magnification (15 nm diameter gold fiducials, polar rotation, tomogram shown in Movie S2).

Note on Tomographic Reconstruction

When a tilt series is recorded, there is a missing "cheese wedge" of information (*i.e.* from above ≈70° and below about -70°). This is because the sample holder blocks the beam at extreme tilt angles, preventing images from being recorded. This can lead to some distortion in the *z* direction of the reconstructed tomogram. As this is the direction which we need to compare the width to the height of the SiO₂-SAGEs, AFM rather than TEM has been used to assess the aspect ratio of the coated NPs.
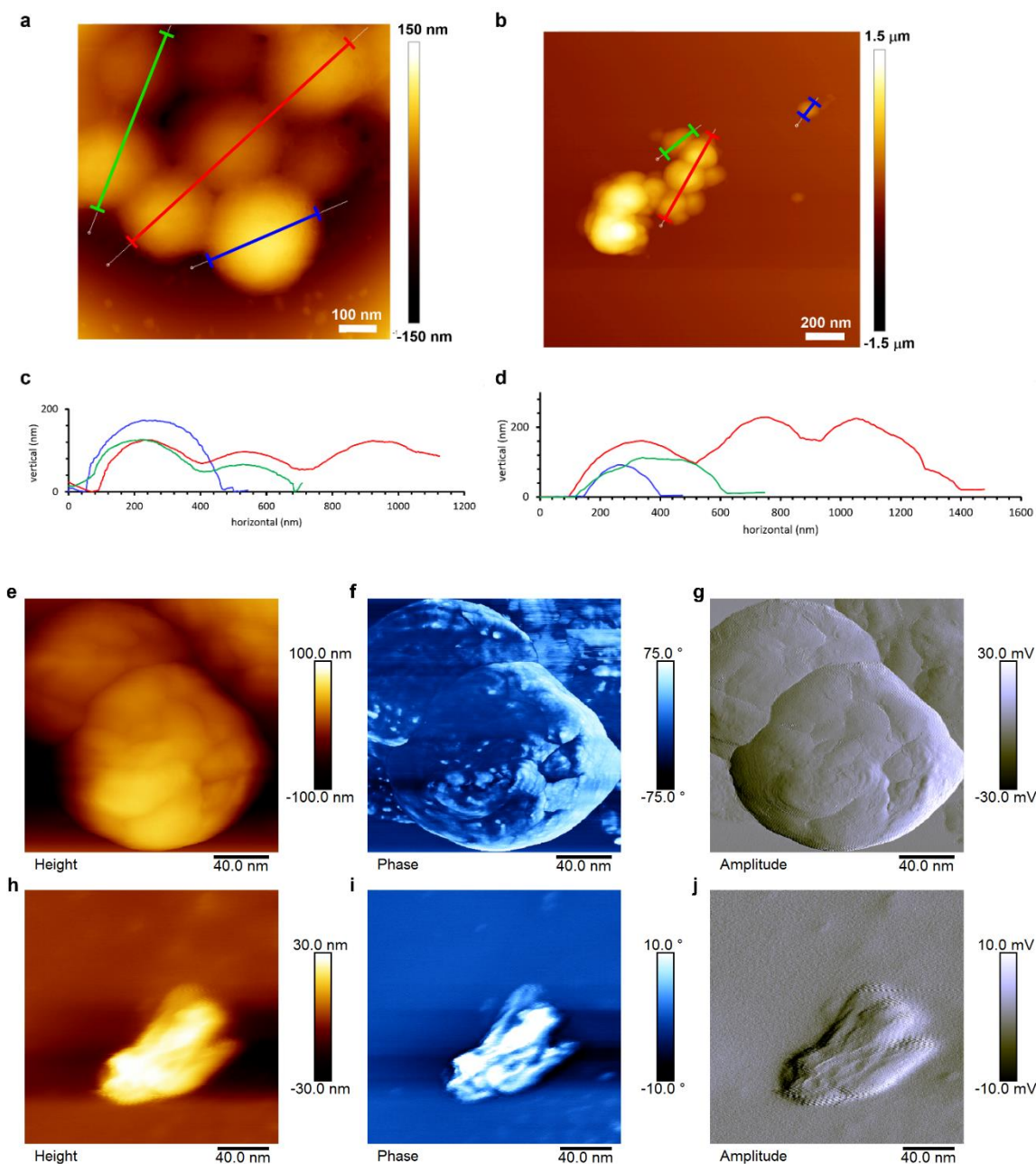
Figure S20. Comparison of height measurements of SiO$_2$-R4-SAGEs recorded using PeakForce (PF-AFM) and Tapping Mode Atomic Force Microscopy (TM-AFM). (a) A 2D height plot recorded using PF-AFM and (b) TM-AFM, which were annotated to show the cross-sectional height measurements shown in (c & d). The height of the SiO$_2$-R4-SAGEs is more than half the width of the particles (aspect ratio ≈0.7), which indicates that there may be some particle collapse, but far less than observe in unmineralized SAGEs (AR ≈0.1, see Fletcher *et al*. 2013[1]). TM-AFM of SiO$_2$-R4-SAGEs and unmineralized R4-SAGEs. SiO$_2$-R4-SAGE maintain their 3D structure shown in the (e) height, and the (f) phase and (g) amplitude plots. Unmineralized R4-SAGEs appear nonspherical, *i.e.* collapsed, the (h) height (z <20 nm) is much less that the diameter in *x* or *y* (≈120 nm), and it is not possible to see a network pattern in the (h) height, (i) phase or (j) amplitude plots from the unmineralized R4-SAGE.
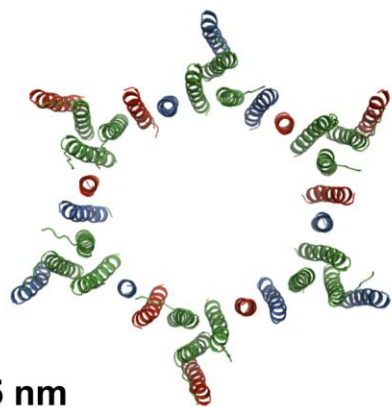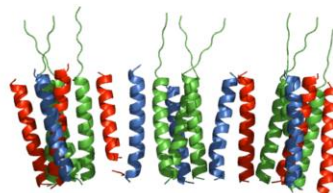
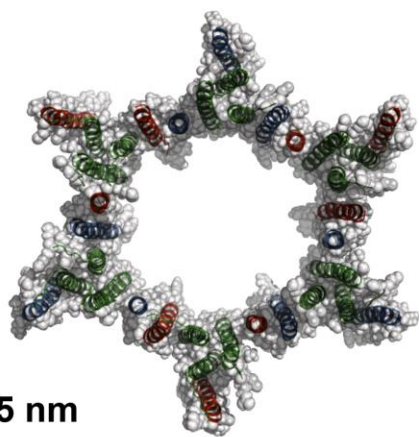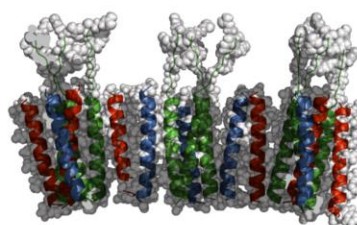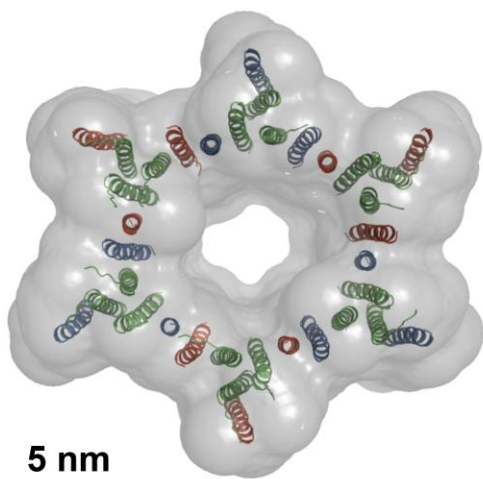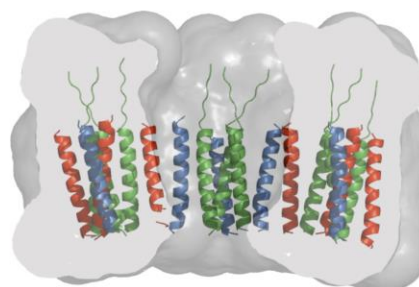**Figure S21**. Illustration of models of the underlying structure and projected surfaces of a patch of a R4-SAGE assembly. The 3-fold symmetric homotrimers (green) plus the 2-fold symmetric heterodimers (blue & red) lead to a hexagonal peptide assembly.[1] (a) Top view (viewed from the N terminus) and (b) cut-through side view of the patch with the coiled-coil helices shown as cartoons. (c & d) The same views as in panels a & b with the van der Waals (VDW) surface of the molecules depicted as gray spheres. (e & f) The same views as in panels a & b with a hypothetical surface calculated using atoms of radius = VDW radius + 10 Å. This is simply to illustrate a 1 nm thick coating on the SAGE surface. In this case, it leaves pores of ≈2 nm diameter in the assembly. If this thickness is increased to 20 Å, the central pore is occluded. These images were created in PyMOL[20] as follows: a & b represented as a cartoon; c & d cartoon plus spheres with standard VDW radii; e & f cartoon plus a surface generated at VDW radii +10 Å.
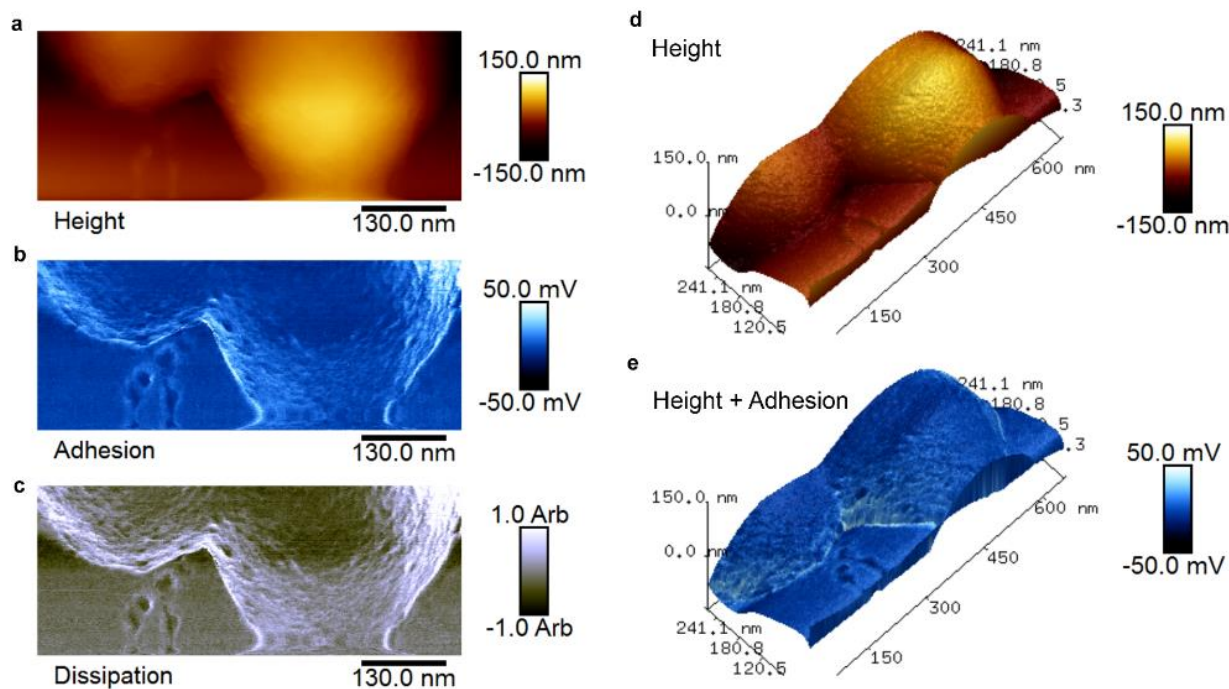

Figure S22. PeakForce Atomic Force Microscopy (PF-AFM) of SiO₂-R4-SAGEs (mineralization conditions: 12 mM silicic acid, 24 hours, PI buffer, 20 °C). These scans include the edges of two SiO₂-R4-SAGEs shown at the bottom of the cluster in Figure 4. 2D plots show variation in (a) height, (b) adhesion and (c) dissipation, with 3D plots showing (d) height and (e) height overlain with adhesion. The network pattern on the surface can be seen on all sides of the SiO₂-R4-SAGEs. The pattern is more difficult to image on the particle sides, as the steeper incline leads to changes in how the tip interacts with the surface, distorting the image of the curved surface.
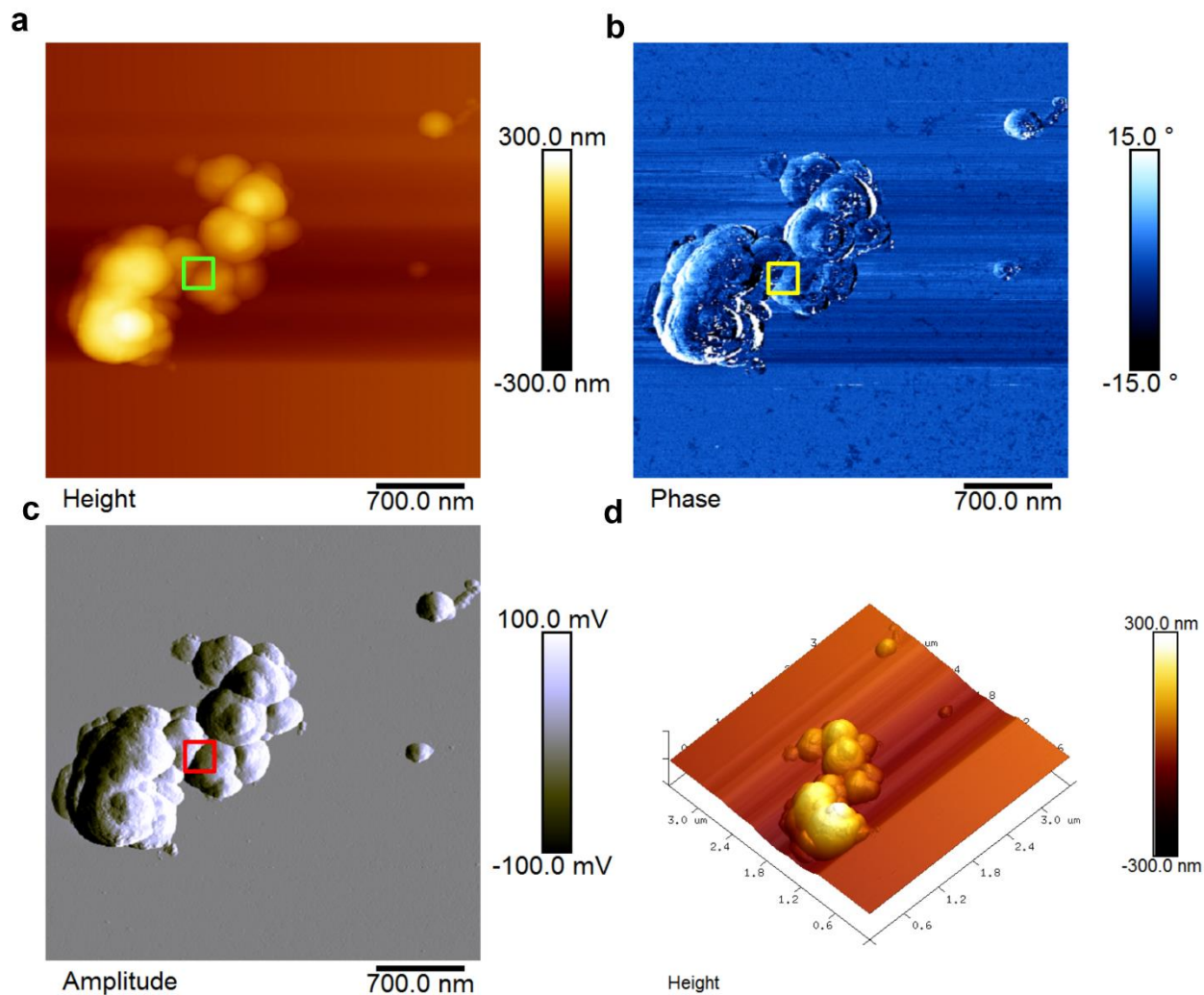
Figure S23. Tapping Mode Atomic Force Microscopy (TM-AFM) of a group of SiO$_2$-R4-SAGEs (mineralization conditions: 24 mM silicic acid, 24 hours, PI buffer, 20 °C). 2D plots and scale bars of the (a) height, (b) phase, and (c) amplitude contrast and (d) a 3D height plot of a group of SiO$_2$-R4-SAGEs on a mica surface.

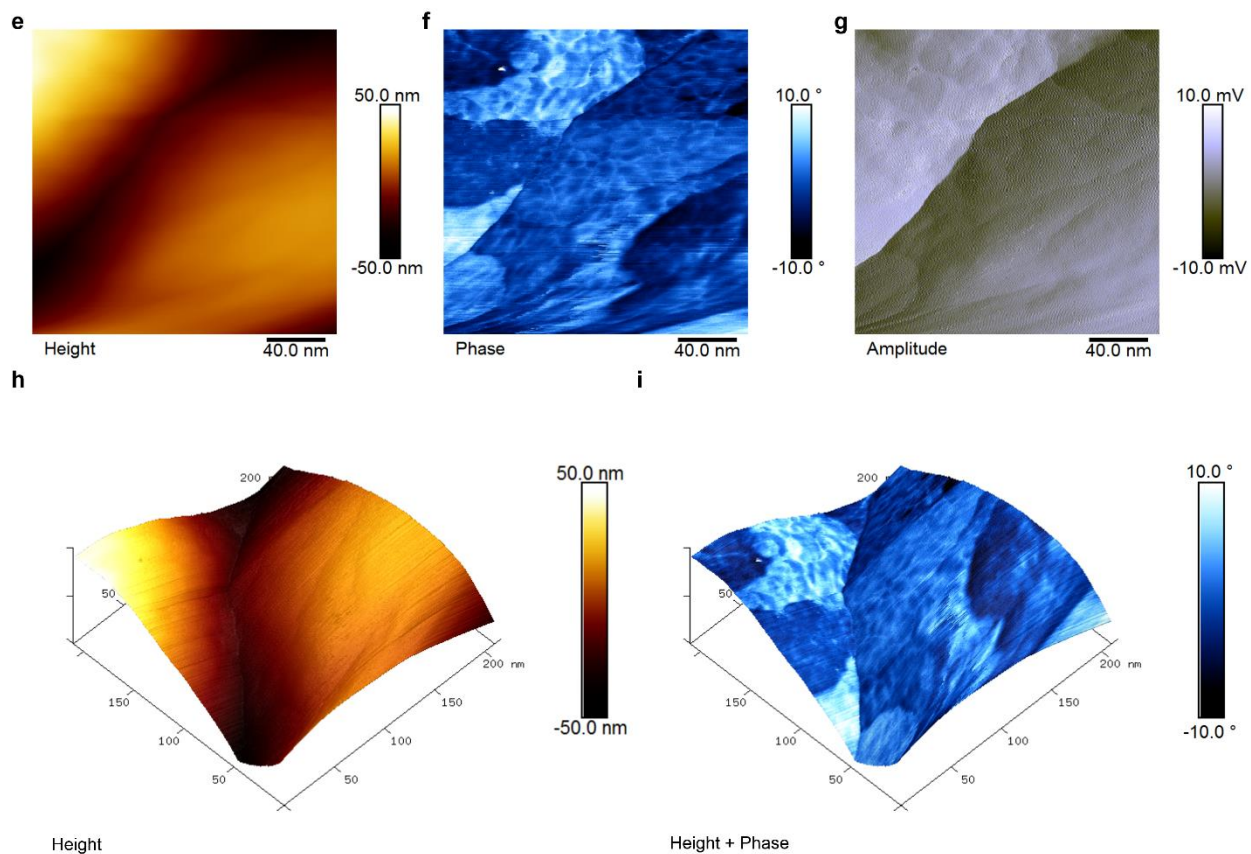**Figure S23 continued.** Tapping Mode Atomic Force Microscopy (TM-AFM) of SiO$_2$-R4-SAGEs (mineralization conditions: 24 mM silicic acid, 24 hours, PI buffer, 20 °C). 2D plots and color scale bars of the (e) height, (f) phase, and (g) amplitude contrast of junction between two SiO$_2$-R4-SAGEs on a mica surface. (h) Shows a 3D plot of the topography, and (i) a 3D height plot overlain with the phase contrast measurements.
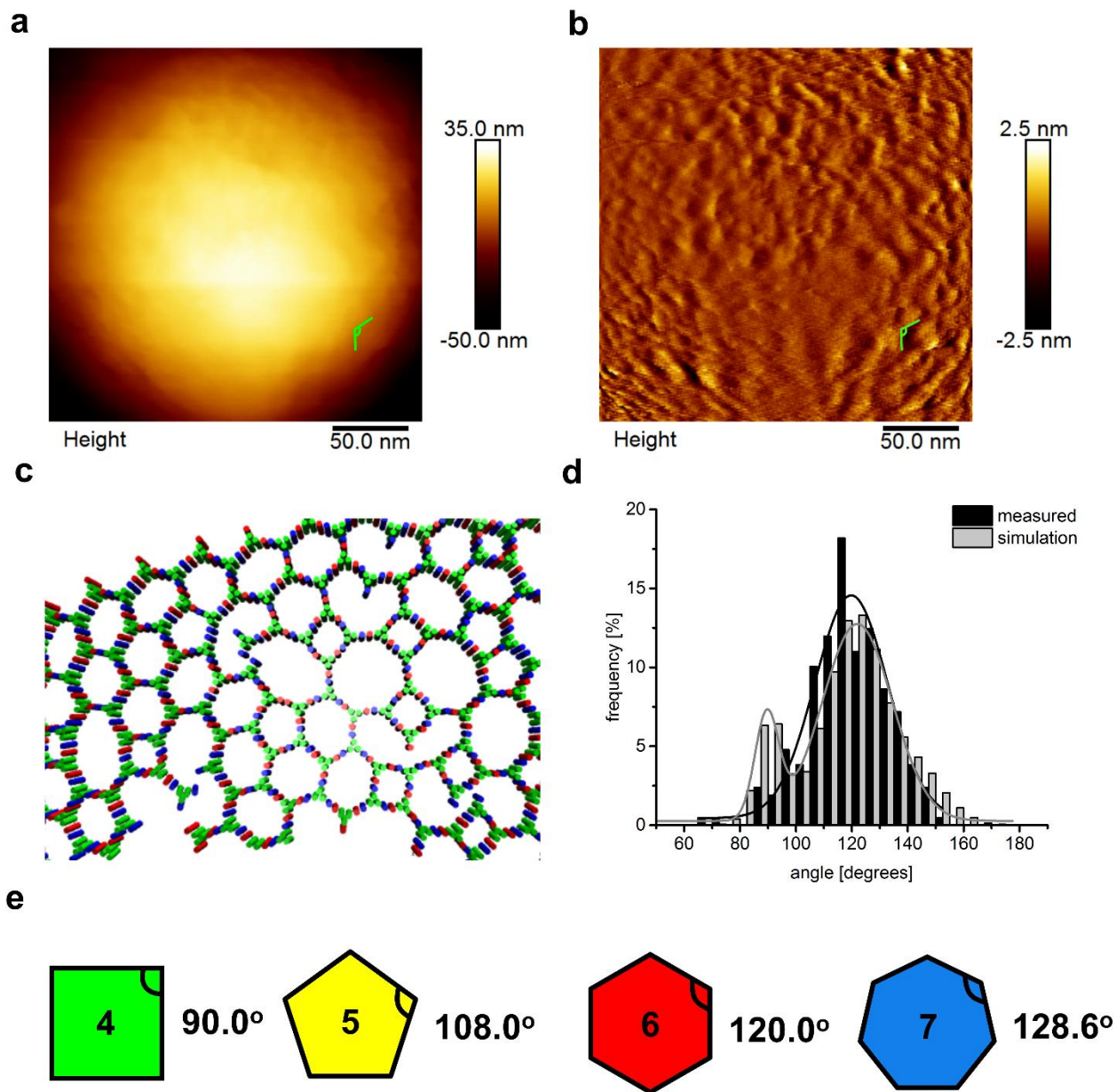
Figure S24. Plots of PF-AFM data and coarse grained computational simulations. (a) As measured topography plot (also shown in Figure 5a) annotated with an example of a measured angle. (b) A plot with the spherical curvature of the SAGE surface removed mathematically (see methods), which allows irregularities in the surface topography (*e.g.* ridges and pores) to be seen more clearly, annotated with the measured angle shown in (a). (c) A snapshot from a coarse-grained computational simulation shows a partially formed SAGE after annealing (CC-Tri3 shown in green, CC-Di-A in red and CC-Di-B in blue, each coiled-coil is ≈3 nm long). There was an abundance of nonhexagonal polygons, which we attribute to the flexibility of the interacting hubs. (d) Distribution of internal angles of polygons measured from AFM plots (n = 209, 119.8° ± 26.1°) and computational simulation (main peak 122.0° ± 25.0°) are centered on hexagonal (*i.e.* 120°). Coarse-grained simulations also show a smaller peak (89.6° ± 8.6°) corresponding to squares (*i.e.* 90°) rather than odd sided polygons. As the model does not allow homomerization between CC-Di-A and CC-Di-B or mixed hub formation by homotrimer exchange, this drives the simulation towards forming even-sided polygons. (e) Illustration to show internal angles of some regular polygons.
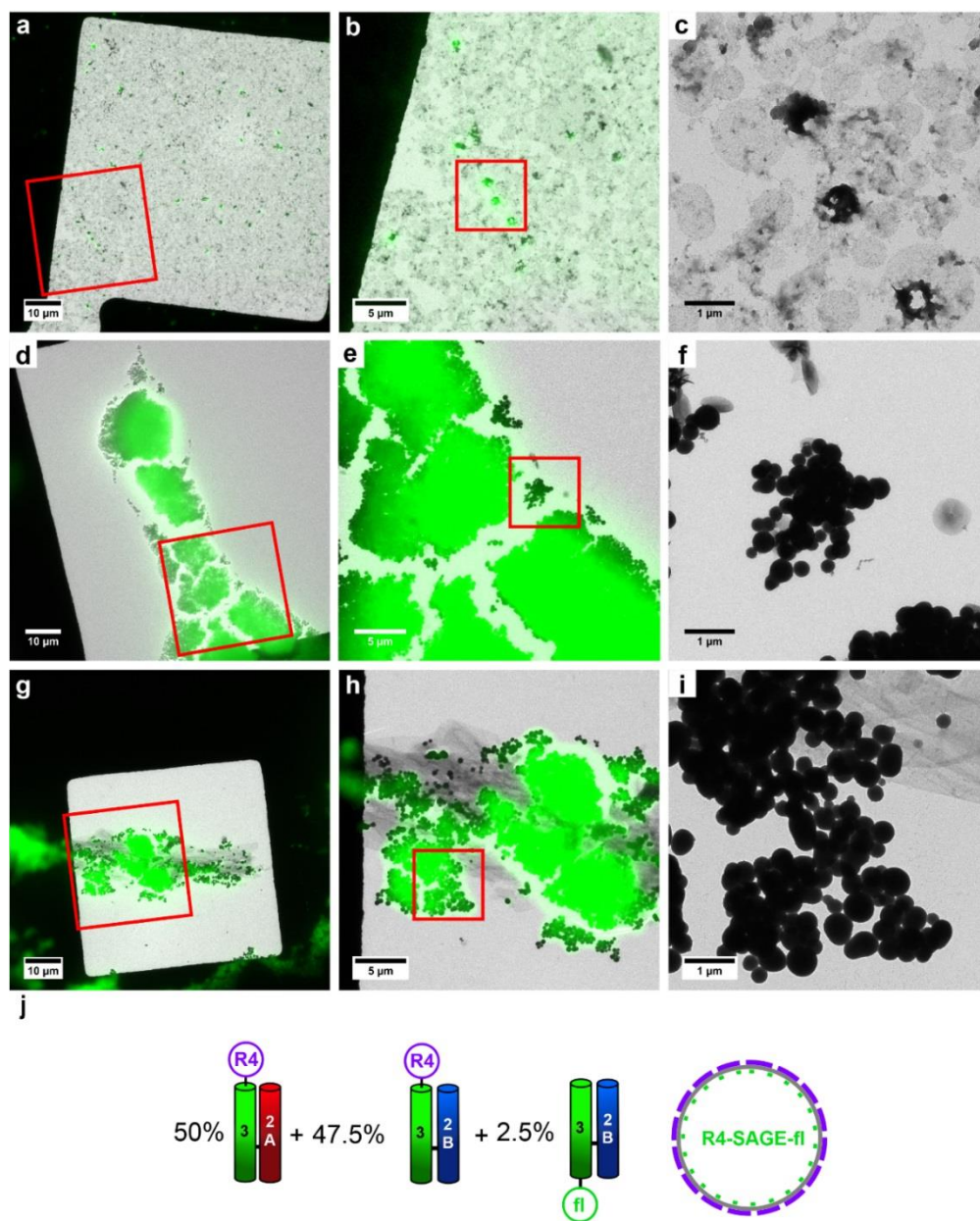
Figure S25. Correlative light and electron microscopy (CLEM) images of tetra-arginine decorated fluorescent SAGEs (R4-SAGE-fls). Fluorescence microscope images were aligned and overlain with TEM images of R4-SAGE-fl to create CLEM images (red squares highlight area of zoom). (a – c) Different magnifications of unmineralized, unstained R4-SAGE-fl showed that fluorescence is co-localized with darker, more electron dense and ruptured structures. Mineralization (24 hours, PI buffer, 20 °C) with (d – f) 12 mM silicic acid had similar fluorescence intensity to uncoated R4-SAGE-fl particles, and those mineralized with 24 mM silicic acid (Figure 5b – d). (g – i) 36 mM silicic acid had weaker fluorescence, (gains on fluorescence microscope were 3x greater to compensate for lower fluorescence detection from the 36 mM silica SAGEs). (j) A diagram to show the peptide module ratios that were used to assemble R4-SAGE-fl.
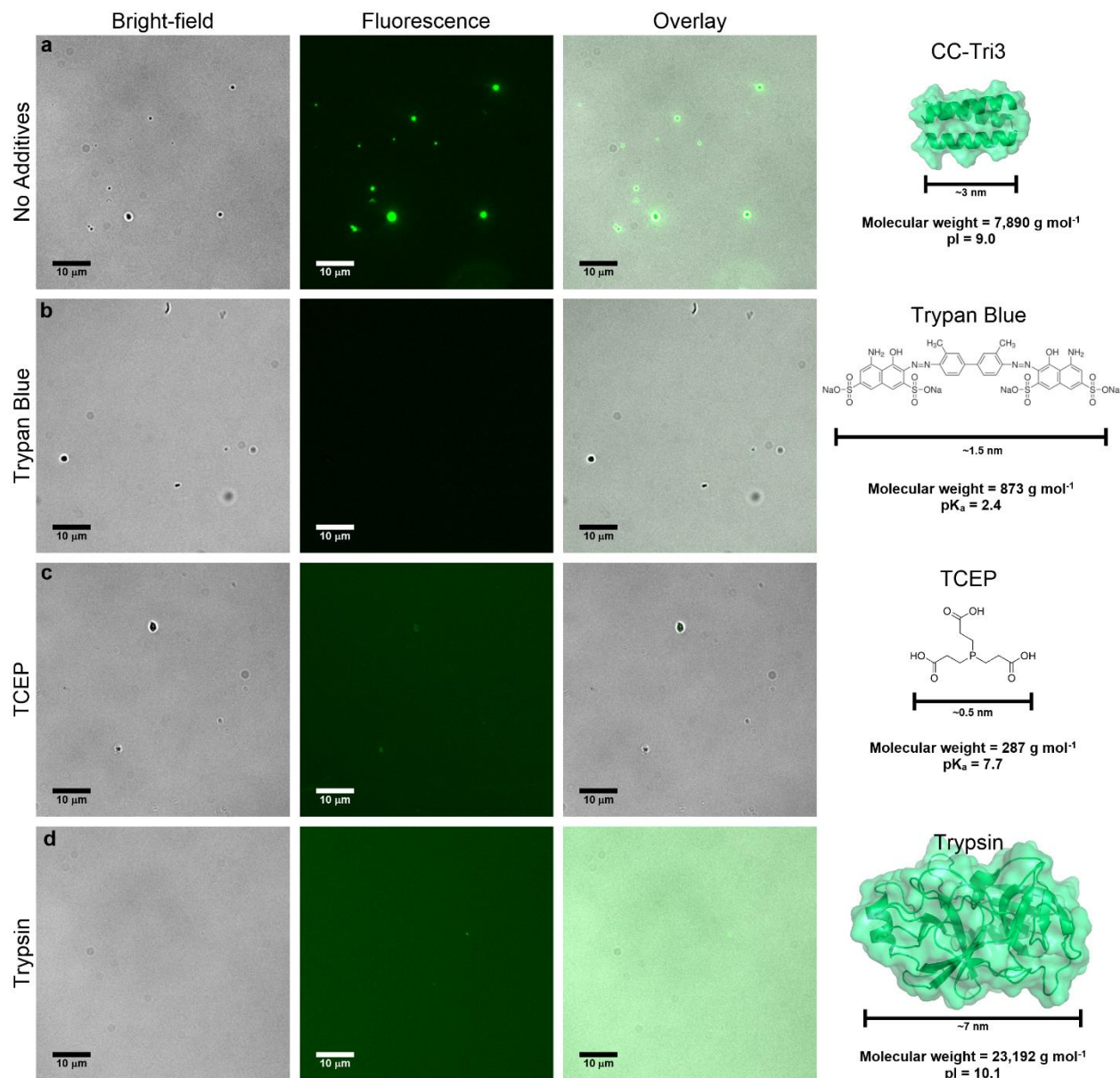
Figure S26. Bright-field and fluorescence microscope images of unmineralized R4-SAGE-fls. Representative images show (left) bright-field and (center) fluorescence channels, and (right) an overlay of these. To test permeability, R4-SAGE-fls were treated with (a) no additives, (b) the fluorescence quencher Trypan blue, (c) the reducing agent tris(2carboxyethyl)phosphine hydrochloride (TCEP), and (d) the protease trypsin. The loss of fluorescent structures indicates that R4-SAGE-fls were subjected to disruption by TCEP and digestion by trypsin, and quenching by Trypan blue indicates that the quenching molecule can penetrate the peptide structures. On the right are diagrams of each molecule (CC-Tri3, Trypan blue, TCEP and trypsin), their approximate size (peptide & protein are on the same scale as each other, quenching and reducing molecules are on same scale as each other), their mass, and pKₐ or pI.
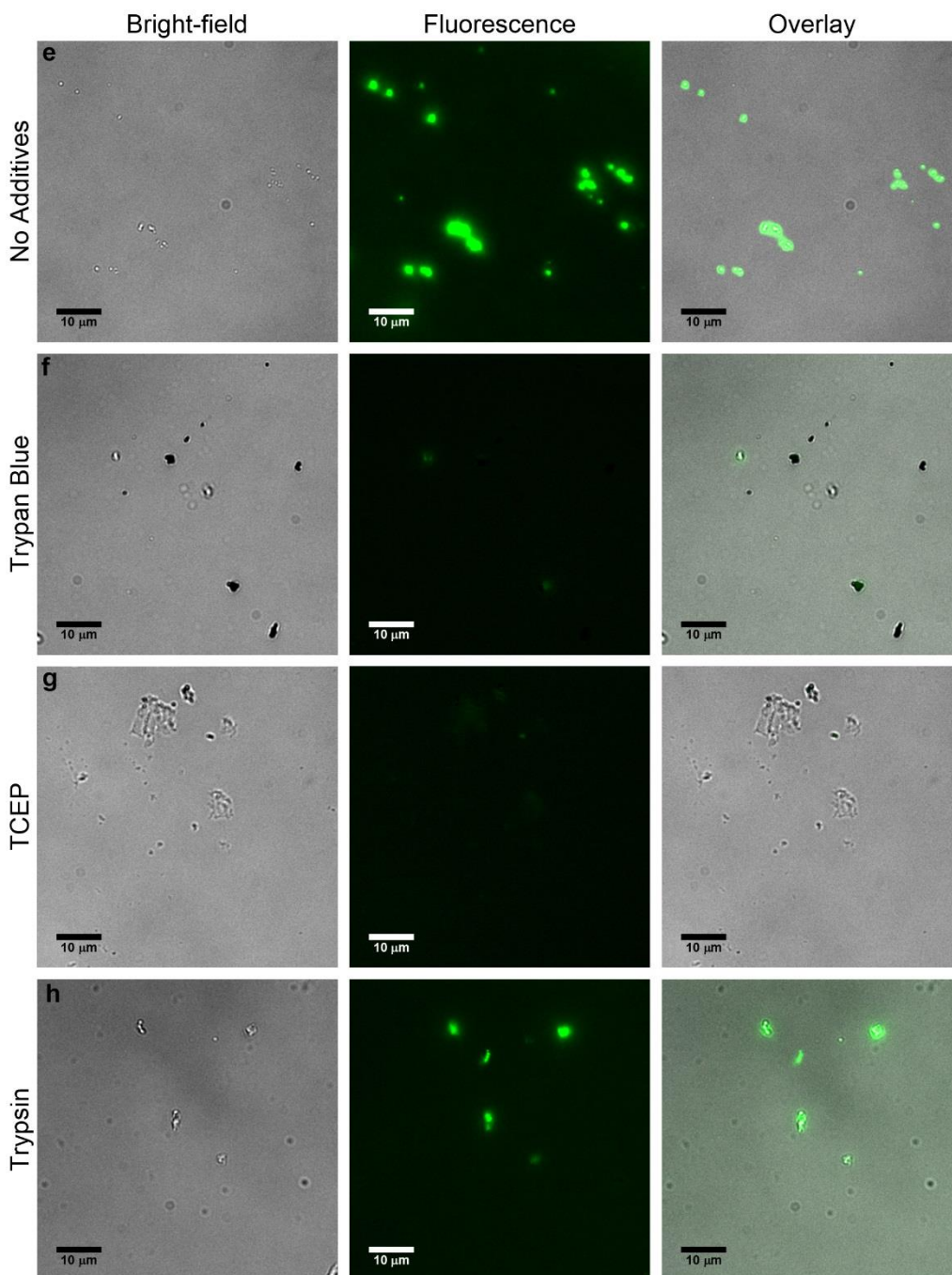
Figure S26 continued. Bright-field and fluorescence microscope images of SiO$_2$-R4-SAGE-fls mineralized with 12 mM silicic acid. Images (left) bright-field and (center) fluorescence, and (right) an overlay. To test permeability, SiO$_2$-R4-SAGE-fls were treated with (e) no additives, (f) Trypan blue, (g) TCEP, and (h) trypsin. Effective quenching by Trypan blue indicates that the quenching molecule can penetrate the silica coated peptide structures. TCEP treatment leaves nonfluorescent structures, which indicates that the hubs within SiO$_2$-R4-SAGE-fls were disrupted by TCEP and the peptides could escape the silica shell. Fluorescent structures remain after treatment with trypsin, which indicates that the protease trypsin cannot penetrate the mineral coating to digest the peptide hubs.
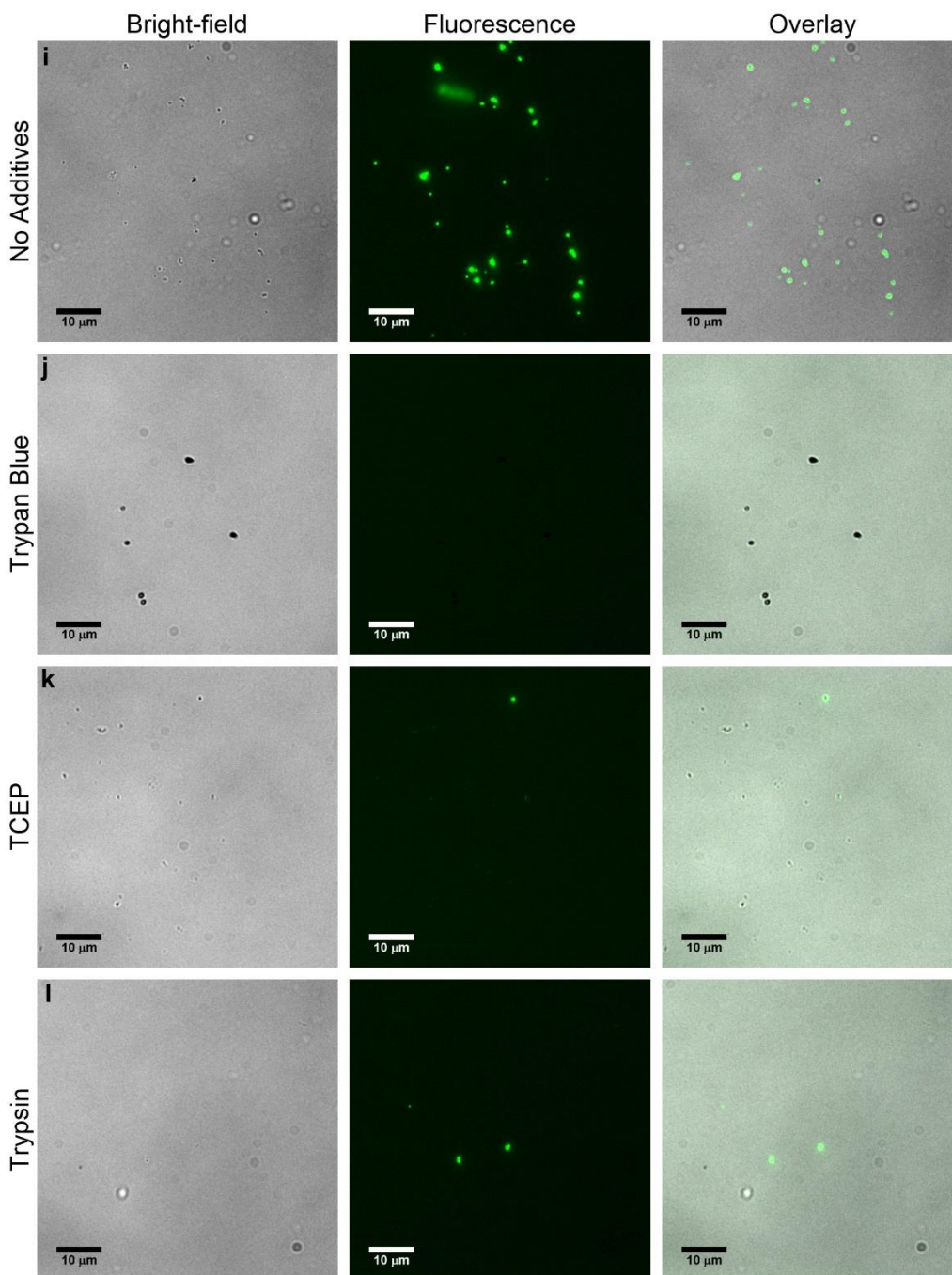
Figure S26 continued. Bright-field and fluorescence microscope images of SiO$_2$-R4-SAGE-fls mineralized with 24 mM silicic acid. Images (left) bright-field and (center) fluorescence, and (right) an overlay. To test permeability, SiO$_2$-R4-SAGE-fls were treated with (i) no additives, (j) Trypan blue, (k) TCEP, and (l) trypsin. Effective quenching by Trypan blue indicates that the quenching molecule can penetrate the silica coated peptide structures. TCEP treatment leaves nonfluorescent and fluorescent structures, which indicates that partial disruption and escape of the hubs within SiO$_2$-R4-SAGE-fls was possible. Fluorescent structures remain after treatment with trypsin, which indicates that trypsin cannot penetrate the mineral coating to digest the peptide hubs.
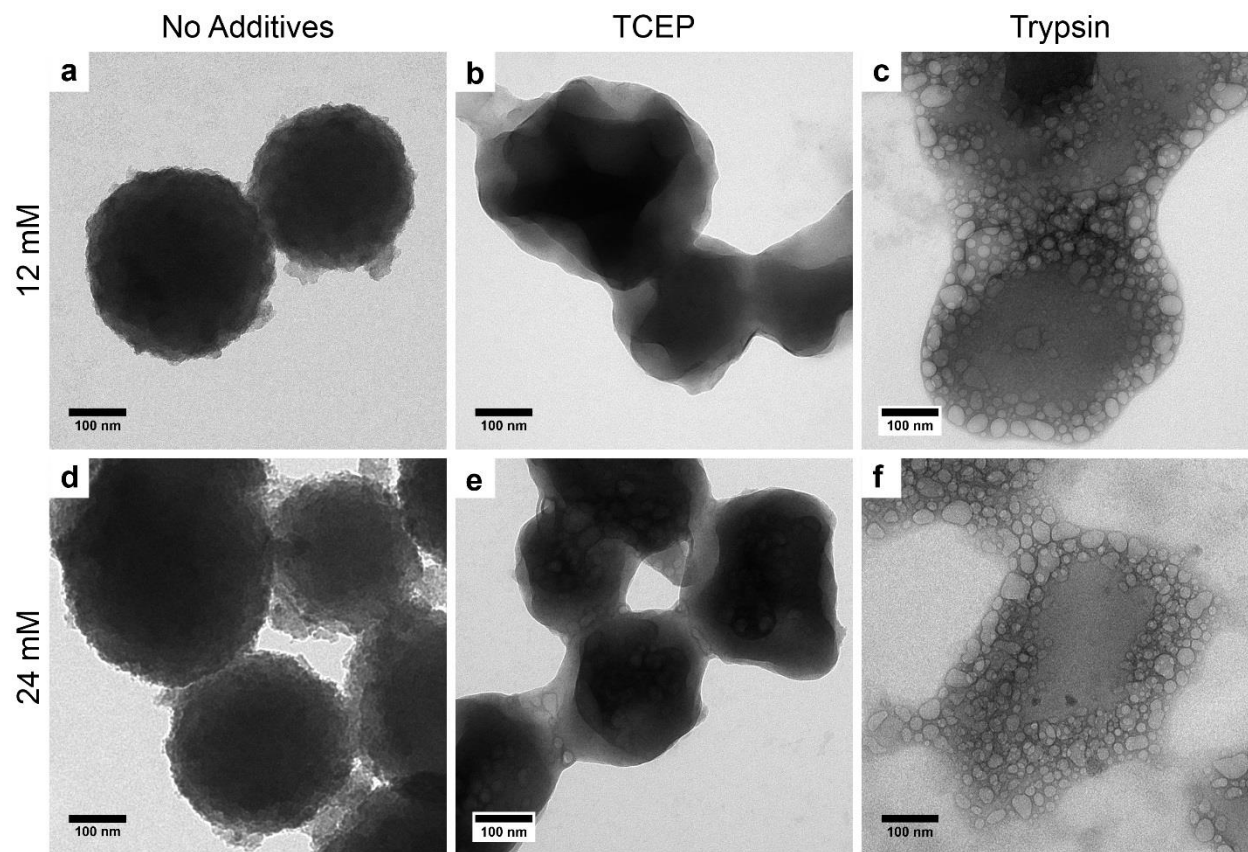
Figure S27. Representative TEM images of SiO$_2$-R4-SAGE-fls before and after treatment with a reducing agent and proteolyzing enzyme. SiO$_2$-R4-SAGE-fls mineralized with (a – c) 12 mM silicic acid and (d – f) 24 mM silicic acid. Silicified particles (a & d) with no additive, after treatment with TCEP (b) 12 mM silicic acid mineralized SiO$_2$-R4-SAGE-fls appear to be more collapsed than (e) 24 mM silicic acid mineralized particles. (c & f) After treatment with trypsin, the particles are coated in a layer that readily undergoes beam damage in TEM, which is probably the protease adhering to the SiO$_2$-SAGE surface.

REFERENCES

(1)   Fletcher, J. M.; Harniman, R. L.; Barnes, F. R. H.; Boyle, A. L.; Collins, A.; Mantell, J.; Sharp, T. H.; Antognozzi, M.; Booth, P. J.; Linden, N.; Miles, M. J.; Sessions, R. B.; Verkade, P.; Woolfson, D. N. Self-Assembling Cages from Coiled-Coil Peptide Modules. *Science* **2013**, *340*, 595–599.

(2)   Fields, G. B.; Noble, R. L. Solid Phase Peptide Synthesis Utilizing 9-Fluorenylmethoxycarbonyl Amino Acids. *Int. J. Pept. Protein Res.* **1990**, *35*, 161–214.

(3)   Ruiz-Gayo, M.; Albericio, F.; Pons, M.; Royo, M.; Pedroso, E.; Giralt, E. Uteroglobin-like Peptide Cavities I. Synthesis of Antiparallel and Parallel Dimers of Bis-Cysteine Peptides. *Tetrahedron Lett.* **1988**, *29*, 3845–3848.

(4)   Kelly, S. M.; Jess, T. J.; Price, N. C. How to Study Proteins by Circular Dichroism. *Biochim. Biophys. Acta - Proteins Proteomics* **2005**, *1751*, 119–139.

(5)   Myers, J. K.; Pace, C. N.; Scholtz, J. M. A Direct Comparison of Helix Propensity in Proteins and Peptides. *Proc. Natl. Acad. Sci.* **1997**, *94*, 2833–2837.

(6)   Scholtz, J. M.; Qian, H.; York, E. J.; Stewart, J. M.; Baldwin, R. L. Parameters of Helix-Coil Transition Theory for Alanine-Based Peptides of Varying Chain Lengths in Water. *Biopolymers* **1991**, *31*, 1463–1470.

(7)   Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **2000**, *78*, 1606–1619.

(8)   Hurton, T.; Wright, A.; Deubler, G.; Bashir, B. SEDNTERP Daemon Version: 20120828 BETA http://sednterp.unh.edu/ (accessed Feb 22, 2017).

(9)   UTHSCSA. Ultrascan http://ultrascan.uthscsa.edu/ (accessed Feb 22, 2017).

(10)   Alexander, G. B. The Preparation of Monosilicic Acid. *J. Am. Chem. Soc.* **1953**, *75*, 2887–2888.

(11)   Shimizu, K.; Del Amo, Y.; Brzezinski, M. A.; Stucky, G. D.; Morse, D. E. A Novel Fluorescent Silica Tracer for Biological Silicification Studies. *Chem. Biol.* **2001**, *8*, 1051–1060.

(12)   Abramoff, M. D.; Magalhaes, P. J.; Ram, S. J. Image Processing with Image J. *Biophotonics Int.* **2004**, *11*, 36–42.

(13)   Schindelin, J.; Rueden, C. T.; Hiner, M. C.; Eliceiri, K. W. The ImageJ Ecosystem: An Open Platform for Biomedical Image Analysis. *Mol. Reprod. Dev.* **2015**, *82*, 518–529.

(14)   Kremer, J. R.; Mastronarde, D. N.; McIntosh, J. R. Computer Visualization of Three-Dimensional Image Data Using IMOD. *J. Struct. Biol.* **1996**, *116*, 71–76.

(15)   Mastronarde, D. N. Dual-Axis Tomography: An Approach with Alignment Methods That Preserve Resolution. *J. Struct. Biol.* **1997**, *120*, 343–352.

(16)   de Boer, P.; Hoogenboom, J. P.; Giepmans, B. N. G. Correlated Light and Electron Microscopy: Ultrastructure Lights Up! *Nat. Methods* **2015**, *12*, 503–513.

(17)   Müller-Reichert, T.; Verkade., P. *Correlative Light and Electron Microscopy*; Müller-Reichert, T.; Verkade., P., Eds.; Oxford Academic Press: Oxford, UK, 2012.

(18)   Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; Tinevez, J.-Y.; White, D. J.; Hartenstein, V.; Eliceiri, K.; Tomancak, P.; Cardona, A. Fiji: An Open-Source Platform for Biological-Image Analysis. *Nat. Methods* **2012**, *9*, 676–682.

(19)   Mosayebi, M.; Shoemark, D. K.; Fletcher, J. M.; Sessions, R. B.; Linden, N.; Woolfson, D. N.; Liverpool, T. B. Beyond Icosahedral Symmetry in Packings of Proteins in Spherical Shells. *Proc. Natl. Acad. Sci.* **2017**, *114*, 9014–9019.

(20)   The PyMOL Molecular Graphics System, Version 1.7.4.5, 2012.