Soft Matter

Polyelectrolyte character of rigid rod peptide bundlemer chains constructed via hierarchical self-assembly

Journal:	*Soft Matter*
Manuscript ID	SM-ART-09-2019-001894.R1
Article Type:	Paper
Date Submitted by the Author:	06-Nov-2019
Complete List of Authors:	Sinha, Nairiti; University of Delaware, Materials Science and Engineering; National Institute of Standards and Technology, NIST Center for Neutron Research Wu, Dongdong; University of Delaware, Materials Science and Engineering Kloxin, Christopher; University of Delaware, Materials Science and Engineering; University of Delaware, Chemical and Biomolecular Engineering Saven, Jeffery; University of Pennsylvania, Department of Chemistry Jensen, Grethe ; National Institute of Standards and Technology, NIST Center for Neutron Research Pochan, Darrin; University of Delaware, Materials Science and Engineering

Polyelectrolyte character of rigid rod peptide bundlemer chains constructed via hierarchical self-assembly

Nairiti J. Sinha[1,2], Dongdong Wu[1], Christopher J. Kloxin[1,3], Jeffery G. Saven[4], Grethe V. Jensen*[† 2, 3], Darrin J. Pochan* [1]

[1] Department of Materials Science and Engineering, University of Delaware, Newark, DE, USA
[2] NIST Center for Neutron Research (NCNR), National Institute of Standards & Technology (NIST), Gaithersburg, MD, USA
[3] Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE, USA
[4] Department of Chemistry, University of Pennsylvania, Philadelphia, PA, USA

[†] present address: Danish Technological Institute, Tåstrup, Denmark
*corresponding authors: pochan@udel.edu, gvj@teknologisk.dk

Abstract

Short α-helical peptides were computationally designed to self-assemble into robust coiled coils that are antiparallel, homotetrameric bundles. These peptide bundle units, or 'bundlemers', have been utilized as anisotropic building blocks to construct bundlemer-based polymers via a hierarchical, hybrid physical-covalent assembly pathway. The bundlemer chains were constructed using short linker connections via 'click' chemistry reactions between the N-termini of bundlemer constituent peptides. The resulting bundlemer chains appear as extremely rigid, cylindrical rods in transmission electron microscope (TEM) images. Small angle neutron scattering (SANS) shows that these bundlemer chains exist as individual rods in solution with a cross-section that is equal to that of a single coiled coil bundlemer building block of ≈ 20 Å. SANS further confirms that the interparticle solution structure of the rigid rod bundlemer chains is heterogeneous and responsive to solution conditions, such as ionic-strength and pH. Due to their peptidic constitution, the bundlemer assemblies behave like polyelectrolytes that carry an average charge density of approximately 3 charges per bundlemer as determined from SANS structure factor data fitting, which describes the repulsion between charged rods in solution. This repulsion manifests as a correlation hole in the scattering profile that is suppressed by dilution or addition of salt. Presence of rod cluster aggregates with a mass fractal dimension of ≈ 2.5 is also confirmed across all samples. The formation of such dense, fractal-like cluster aggregates in a solution of net repulsive rods is a unique example of the

subtle balance between short-range attraction and long-rage repulsion interactions in proteins and other biomaterials. With computational control of constituent peptide sequences, it is further possible to deconvolute the underlying sequence driven structure-property relationships in the modular bundlemer chains.

Introduction

Supramolecular self-assembly of biological building blocks into multifunctional materials that display hierarchical structural organization is ubiquitous in nature.[1, 2] Synthetic biomolecule self-assembly has emerged as a bioinspired, powerful tool for the bottom-up construction of materials.[3] While empirical sequence design of non-natural *de novo* biomolecules based on DNA,[4] lipids[5] and proteins[6] to construct non-natural biomaterials has been widely successful, introduction of the computational toolbox for defining biomolecule design has added an exciting new dimension to material discovery while also enabling tunability of structure and function at the nanoscale in a systematic manner. Consequently, *in silico* computational design tools are contributing significantly to the flourishing science of DNA origami[7, 8] and to supramolecular protein assembly,[9] especially via protein-protein interface modification[10] and docking algorithms.[11-13] However, much work remains in terms of sequence-based design and structure predictions for peptide and protein folding and assembly, which is complicated by multiple local interactions, large numbers of potential conformations, and sequence-dependent structural features, the sum of which is commonly referred to as the 'protein folding problem'.[14] As a result, the potential of *in silico* computational design of protein-based biomaterials remains vastly unexploited.

Recently, coiled coils, tertiary or quaternary structural domains frequently found in natural proteins, have garnered attention as candidates for novel biomaterial construction.[15, 16] In nature, coiled coils are commonly left-handed bundles of two to six α-helical peptides, each of which are based on a repeat motif of seven amino acid residues called the heptad (-*abcdefg*-). The coiling of this sequence into a helix produces a 'hydrophobic stripe' along one face of the peptide (amino acids at positions *a* and *d*), which promotes assembly of these peptides into supercoiled bundles, burying their hydrophobic faces while exposing the more hydrophilic residues (amino acids at positions *b,c,e,f* and *g*) to the aqueous

solvent. In case of synthetic coiled coil forming peptides, the solvent-exposed groups can be carefully picked or modified *in silico* such that the resulting inter-bundle interactions can direct bundlemer self-assembly into predetermined nanostructures.[17-19] Important advantages of using coiled coil-based building blocks for nanomaterial construction include assembly that is triggered under mild aqueous conditions, robustness of the bundle assembly, and the use of nonbiological synthesis techniques. Importantly, non-natural amino-acid and arbitrary chemical functional group incorporation is enabled for strategic placement on solvent-exposed bundle sites, facilitating new pathways for hierarchical biomaterial construction. Consequently, advances in the computational design of coiled coils has leveraged their well-understood structural parameter space and cylindrical structure.[20-23]

Zhang et al. demonstrated that computationally designed peptide sequences that form robust, antiparallel homotetrameric coiled coils, i.e. four identical helices packed in an antiparallel fashion providing the bundle with D_2 symmetry, successfully assembled into non-natural 2D lattices with a targeted crystal symmetry.[24] The self-assembled nanostructure of the artificial bundles was responsive to aqueous solution conditions such as pH, salt and temperature.[25, 26] Haider et al. confirmed that these computationally designed bundles have the structure of nano-cylinders that are ≈ 40 Å in length and ≈ 20 Å in diameter in good agreement with their expected computationally designed dimensions.[27] Exploiting the robustness of these bundles, Wu et al. modified the N-termini of the peptides to include either a free thiol (via addition of a cysteine) or a free maleimide (Mal-) group resulting in two tetra-functional bundlemers (see **Figure 1**). When the thiol and maleimide-functional bundlemers were reacted via Thiol-Michael *click* reaction, bundlemer chains that have a 20 Å cross-section were formed having extremely long persistence lengths (i.e., on the order of tens of micrometers), herein referred to as rigid rods.[28] Furthermore, the *click* chemistry-based construction pathway allowed use of a variety of linker moieties and facilitated the control of bundlemer chain length, flexibility, architecture (such as introduction of kinks) and branching; this synthetic approach yields a library of tunable 1D nanomaterials.

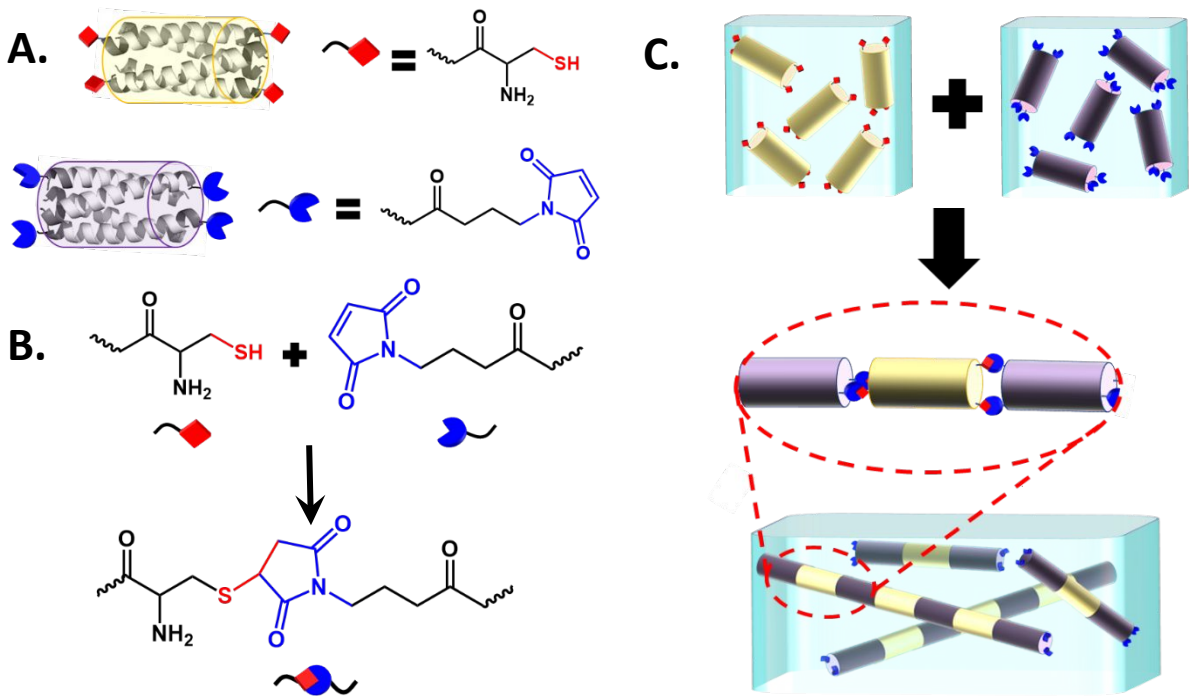


Figure 1: (A) Schematic of homotetramer coiled coils showing the bundlemers **Peptide 1** (purple) decorated with maleimide groups (blue cut circles) and **Peptide 2** (yellow) decorated with thiol groups (red diamonds). **(B)** Reaction scheme of Thiol-Michael *click* reaction between a thiol group (on cysteine) and maleimide group that yields a thiol-maleimide adduct. **(C)** Reaction protocol for hybrid physical-covalent assembly via Thiol-Michael *click* reaction between assembled bundlemers. The thiol-decorated **Peptide 2** bundlemers (yellow with red diamonds) are reacted with Mal-decorated **Peptide 1** bundlemers (purple with blue cut circles) resulting in formation of rigid rod-like chains with alternating **Peptide 1** and **Peptide 2** bundlemers.

The bundlemer chains include physical interactions to stabilize individual bundlemers and covalent connections for hierarchical rigid rod formation. These chains and their hierarchical assembly pathways present unique opportunities to utilize them as model systems to study structure-property relationships of specific amino acid sequences, a knowledge critical in furthering our understanding of 1D assemblies such as collagen, microtubules, Tobacco Mosaic Virus (TMV), fd bacteriophage, actin and amyloid fibers[29-31]. These biological assemblies are usually rod-like polyelectrolytes that are inherently different from their synthetic counterparts, such as colloidal gold nanorods, wormlike micelles and simple polymeric polyelectrolytes such as polystyrene sulfonate, in that they display patchiness in local interactions due to multiple polar, hydrophobic, and oppositely charged side groups.[32-34] The complexity in interactions in biological assemblies is readily mimicked in the computationally-designed bundlemers

due to their sequence tunability, folding, and assembly. We present herein a study of the solution-state structure and interactions of bundlemer rigid rods to shed light on specific sequence driven structure-property relationships relevant to rod-like biological assemblies. Small-angle neutron scattering (SANS) is uniquely suited to studying the solution structure of soft matter systems in a non-destructive manner and can be used alongside transmission electron microscopy (TEM) as together, they unambiguously yield information about the geometric shape of the nanomaterial while also giving insight into interactions between components in solution (obtained via SANS). Three different rigid rod systems with increasing average rod lengths have been investigated for their solution-state structure as a function of peptide concentration. For the long rigid rod system, the impact of monovalent sodium chloride (NaCl) salt and acidic pH on the solution structure has also been investigated. For the interpretation of the SANS data, we model the rigid rods as straight cylinders that are repulsive at low ionic strength and neutral pH conditions using a mean-field expression derived by Schneider et al. [35] Clusters of rods that are evidenced by a characteristic upturn corresponding to large length scales in SANS curves are modeled using Teixeira's fractal scattering expression.[36] Implications of the composite SANS modeling results and important similarities and differences between the rigid rods and other 1D assemblies are discussed.

Experimental section

1. Synthesis and purification of bundlemers:

Two computationally designed peptide sequences that have been previously reported to form bundlemers were employed to construct rigid rods via a hybrid physical-covalent self-assembly pathway.[24, 27] Here, these peptide sequences were modified to include functional groups that enable assembly of polymers of covalently linked bundlemers. The peptide sequences, **Peptide 1** with a maleimide (Mal)-modified N-terminus and amidated C-terminus (Mal DEKIKNM ADQIKHM AWMIDRM AEKIDRE A -NH_2) and **Peptide 2** with a cysteine at the N-terminus and amidated C-terminus (C DEEIRRM AEEIRQM AERIQQM AEQIQQE A -NH_2) were synthesized as reported previously.[24,28] Specifically, peptides were synthesized at a 0.25 x 10^{-3} mol scale via microwave-assisted

121 fmoc-based solid-phase peptide synthesis (fmoc-SPPS) on a CEM Liberty Blue™ Automated Microwave
122 Peptide Synthesizer. Rink-amide resin (ChemPep) that yields an amidated C-terminus after the final
123 cleaving step (see below) was utilized for all syntheses. Amino-acids were purchased from ChemPep and
124 used as is. The solvent for all synthesis steps was analytical-grade di-methyl formamide (DMF) which
125 was purchased from Fischer scientific. Coupling reagents 0.5 mol/L N, N-diisopropylcarbodiimide (DIC,
126 Sigma) and 0.5 mol/L ethyl 2-cyano-2-(hydroxyimino) acetate (Oxyma, CEM) were prepared in DMF
127 while piperidine (Sigma) in DMF having a volume fraction of 20 % was utilized for deprotection cycles.
128 Default microwave and heating cycles were employed for the coupling and deprotection steps which are
129 also recommended by CEM. For **Peptide 1**, the functionalizing agent N-carboxy propyl maleimide
130 (Sigma) was coupled to the N-terminus of the peptidoresin by performing an additional amino-acid
131 coupling cycle on the instrument itself at the end of the synthesis.

132 The final peptidoresin was washed multiples times with analytical-grade dichloromethane (DCM,
133 Fischer) and fresh DMF. The peptide was then cleaved from the resin by suspending the peptidoresin in
134 10 ml cleaving solution containing a volume fraction of 95 % trifluoroacetic acid (TFA, Sigma), 2.5 %
135 Milli-Q water and 2.5 % triisopropylsilane (TIPS, Sigma) and shaking the cocktail for 2 hours. The
136 cleaved peptide solution was collected and precipitated using fresh anhydrous ethyl-ether (Fischer) at
137 least three times. The crude was dried, resuspended in solvent containing a volume fraction of 70 % water
138 and 30 % acetonitrile (HPLC-grade, Fischer) and lyophilized for subsequent purification.

139 The lyophilized crude peptide was dissolved in optimal water-acetonitrile solvent and purified on
140 a high-performance liquid chromatography (HPLC) instrument (Quaternary Gradient Module (Waters
141 2545), Waters Corporation) using a reverse-phase BEH130 Prep C18 10 μm column (XBridge, Waters
142 Corporation, Milford, MA). The elution gradient employed was 85 % solvent A to 15 % solvent B over
143 60 minutes, where solvent A was Milli-Q Water containing 0.1 % TFA and solvent B was acetonitrile
144 with 0.1 % TFA (all percentages are volume fractions). The elutant was continuously monitored for
145 absorbance at 214 nm and 280 nm using a UV-Vis photo-detector (Waters 2489, Waters Corporation).
146 The peptide elution peak was collected and checked for the correct molecular weight using electron-spray

ionization mass spectroscopy (ESI-MS) on Waters Xevo G2-XS QTof Quadrupole Time-of-Flight Mass Spectrometry instrument (see SI for ESI-MS results). The purified peptides were then combined and lyophilized.

2. Hybrid physical-covalent assembly of bundlemers:

Due to the antiparallel packing of four identical α-helical peptides within a coiled coil bundle, N-termini modification yielded two functional groups at each end of the bundlemer. These bundlemers were reacted together via Thiol-Michael *click* reaction resulting in polymers comprised of alternating **Peptide 1** and **Peptide 2** bundlemers linked by two thiol-maleimide adducts between them (**Figure 1**). For this study, the stoichiometric ratio *r* of thiol- groups afforded by cysteines on **Peptide 2** to Mal-groups on **Peptide 1** for constructing polymeric rigid rods having a length distribution centered at a desired length *L*, *i.e.* short (r=0.50, L~105 Å), medium (r~0.88, L~525 Å) and long (r~0.94, L~1085 Å), were calculated using Flory's equation assuming full conversion (See **SI**).[28, 37]

A step-wise hybrid physical-covalent assembly pathway was utilized to construct rigid rods of bundlemers that has been optimized and reported in our previous publication. Stock solutions of lyophilized pure peptides were prepared by suspending them in Milli-Q water at 1 x 10^{-3} mol/L peptide concentration. Tricarboxyethylphosphine hydrochloride (TCEP.HCl, Sigma) was added to **Peptide-2** stock solution at a final concentration of 0.5 x 10^{-3} mol/L to prevent disulfide formation via oxidation of the free thiol groups on the cysteine.

Calculated volumes of **Peptide 1** and **Peptide 2** stock solutions required to assemble rigid rods of target length distributions were mixed together and reacted overnight under constant shaking at 30ºC. The success of the reaction was confirmed via ESI-MS wherein peaks corresponding to the **Peptide 1-2** dimer of the correct molecular weight were observed due to formation of the Thiol-Michael adduct between bundlemers (see **SI** for discussion on ESI-MS). The resulting solution containing the assembled rigid rods was lyophilized and re-suspended in Deuterium Oxide (D_2O, atom fraction of 99.5 %, Aldrich) at desired concentration for subsequent TEM and SANS characterization.

173

174 **3. Characterization of rigid rod-like bundlemers**

175 **a. Transmission Electron Microscopy (TEM):**

176 Before TEM analyses, samples were diluted to a mass by volume fraction of 0.1 %. Sample grids
177 for analyses were prepared on 200-mesh carbon-coated copper grids purchased from Electron Microscopy
178 Solutions Inc. All grids were plasma treated under vacuum for 30 seconds and used within 2 minutes of
179 treatment. A 3-5 μL drop of the respective rod sample was added to the carbon-coated side of a plasma-
180 treated grid. After 60 seconds, excess sample was blotted and 3-5 μL of freshly prepared 1 % (mass by
181 volume fraction) sodium phosphotungstate staining solution ($NaPWO_4$, pH 7.0) immediately was added
182 to the same side of the grid. The excess stain solution was blotted after 60 seconds and the sample grid
183 was dried under air for at least 15 minutes before TEM analyses. TEM was performed on FEI Talos
184 F200X Transmission Electron Microscope using a 200 x 10^3 V accelerating voltage. A Ceta 16 M camera
185 from Gatan was used to record 4 k x 4 k images. ImageJ analysis software was utilized to measure rod
186 lengths on the micrograph and calculate length statistics for each rigid rod sample.[38]

187

188 **b. Small-Angle Neutron Scattering (SANS):**

189 Samples for SANS measurements were prepared by suspending lyophilized rigid rod samples of a
190 target length (short, medium or long) in D_2O and subsequent dialysis against 50 ml D_2O using Thermo
191 Scientific Dialysis Devices (3500 Da Molecular weight cutoff) to remove any small molecule,
192 unassembled peptide, and unreacted bundlemer impurities. The sample with the highest concentration of
193 rigid rods was measured first and subsequently diluted to a lower concentration using fresh D_2O for the
194 next measurement. At each dilution step, the samples were degassed (5 minutes) and sonicated (10
195 minutes) in a sonication bath at 30ºC and rested for at least 30 minutes prior to SANS measurement. All
196 samples and pure D_2O had a pH in the 6 to 7 range as measured using Hydrion pH paper. For studies of
197 the effect of ionic strength on the long rods, a concentrated aqueous solution of sodium chloride (NaCl,
198 salt) was added resulted in a salt concentration of 200 x 10^{-3} mol/L. Similarly, for the acidic pH study,

deuterium chloride (DCl, atom fraction of 99 % D, Sigma-Aldrich) in D_2O was added to the rigid rod solution to prepare a 200 x 10^{-3} mol/L DCl sample with an acidic pH of ca. 1.0 as measured by Hydrion pH paper.

SANS experiments were performed at the National Institute of Standards and Technology (NIST) Center for Neutron Research, Gaithersburg, Maryland, USA. Beamlines NG7 SANS and CHRNS NGB30m were utilized for conducting the measurements. A wavelength of 6 Å of cold neutrons with a wavelength spread of ca. 12 % was employed. Sample-to-detector distance of 1 m and 4 m were utilized to collect scattering data in the high to mid-Q region ($0.01 < Q < 0.3$ Å^{-1}). Here, Q is the momentum transfer of the scattered neutron, defined as $Q = \frac{2\pi\, sin\theta}{\lambda}$, where λ is the wavelength of the incoming beam of neutrons and 2θ is the scattering angle. For low Q data collection, samples having peptide concentration C greater than a mass by volume fraction of 0.5 %, were measured with either 13 m sample-to-detector distance configuration and neutron lenses on NGB30 beamline or with 15 m sample-to-detector-distance and neutron lenses on the NG7 beamline, both using a neutron wavelength of 8 Å, resulting in a low-Q cutoff of 0.002 Å^{-1}. Due to low counting statistics for rod samples with concentration C less than mass by volume fraction of 0.5 %, low-Q scattering data was collected at 13 m sample-to-detector distance without the use of neutron lenses, resulting in a larger low-Q cutoff of 0.005 Å^{-1}.

All scattering data were reduced using NIST NCNR's IgorPro data reduction software[39]. For initial analyses, SASView scattering data analysis software was utilized.[40] For in-depth data modeling, an in-house python code was written that uses a non-linear least-squares package to fit the normalized SANS data to model *I(Q)* expressions described in the next section. Best fit results with the lowest possible reduced chi-squared (χ_R^2) values were obtained after multiple fitting routines. The mean and standard deviation (σ) for each fit parameter was calculated and is summarized along with the reduced chi-squared values in the **SI**. Error bars in the SANS data and fit parameters represent ±1σ throughout this article.

Theory:

SANS data modeling:

In a SANS experiment, particles of interest are dispersed in a suitable solvent and bombarded by a coherent beam of neutrons. The neutron scattering pattern is given as the scattering intensity as a function of the scattering vector, Q. For a monodisperse solution of isotropically scattering particles, the resulting scattering intensity profile $I(Q)$ is given by:[41]

$$\mathrm{I(Q) = nV_p^2(\rho_p - \rho_s)^2\ P(Q).S(Q)} \qquad (1)$$

where n is the number density of scatterers having a volume V_p and scattering length density ρ_p dispersed in a solvent with scattering length density ρ_s. The shape and size of the scatterers is captured in the form factor $P(Q) = \langle A^2(Q)\rangle$, where $A(Q)$ is the scattering amplitude, and the brackets denote an average over particle orientations, a requirement for isotropic scattering. The effect of the interparticle interactions is encoded in the structure factor $S(Q)$. We model the form factor of the rigid rods as straight cylinders.[28] For particles with such anisotropic shapes, $S(Q)$ in eq. 1 is approximated by an effective structure factor $\hat{S}(Q)$ to keep eq. (1) consistent with isotropic scattering.[41] Kotlarchyk and Chen have derived the $\hat{S}(Q)$ for cylinders, given by:[42]

$$\hat{S}(Q) = 1 + \beta(Q)(S(Q) - 1) \qquad (2)$$

Where the factor $\beta(Q) = \langle A(Q)\rangle^2 / \langle (A(Q))^2\rangle$ (see eq. 3 and 4) is the decoupling approximation expression which accounts for the assumption that the position and orientation of the anisotropic rigid rods is decoupled. These considerations for model fitting are important also because the 2D scattering intensity obtained via SANS for all samples were radially symmetric, which is indicative of an isotropic solution of rod-like particles at corresponding length scales.

The bundlemer rigid rods were modeled as straight cylinders having a radius R and length L. The form factor of an isotropic dispersion of cylinders is given by:[41]

$$P_{cyl}(Q) = \langle A^2(Q)\rangle = \int_0^{\pi/2} A^2(Q,\alpha)\sin\alpha\, d\alpha \qquad (3)$$

Here, $A(Q)$ is the amplitude function of a cylinder, given by:

$$A(Q,\alpha) = \frac{\sin\left(\frac{QL\cos(\alpha)}{2}\right)}{\frac{QL\cos(\alpha)}{2}} \cdot \frac{2J_1(QR\sin(\alpha))}{QR\sin(\alpha)} \qquad (4)$$

Equations 3 and 4 together indicate that the scattering intensity I follows the power law $I \sim Q^{-1}$ for the Q-range $2\pi L^{-1} < Q < R^{-1}$. If the rods are non-interacting, a Guinier region with $I(Q) = I(0)exp(-Q^2R_g^2/3)$ is observed for $QR_g < 1$, where R_g is the radius of gyration of the monodisperse rod particles.[41] For the rigid rods however, the individual rods interact and form cluster aggregates indicated by the low-Q upturn, and their R_g can therefore not be determined from the data. Also, the length L cannot be determined, but, rather, an apparent length, L_a, that gives the characteristic rod length between points of contact in the cluster. The upturn follows a power law, *i.e.* $I \sim Q^{-\mathcal{D}}$ where $2 < \mathcal{D} < 3$. This can be attributed to the presence of mass fractal-like networked cluster of rods with the fractal dimension $\mathcal{D}$. The structure factor expression used to model this low-Q feature is based on Teixeira's model which was used to describe aggregation of spherical particles.[36] In Teixeira's model, the sphere diameter was applied as a model parameter giving the length scale for the building blocks of the fractals. Here we have replaced it by the apparent length L_a of the rods as shown in **Figure 2**. It was not possible to determine the over-all size of the clusters as it falls outside the length scales probed in the scattering experiment. The model parameter giving the fractal size, ξ, is therefore fixed at an arbitrary, large value ($\xi = 10000$ Å). The expression for the fractal structure factor is thus given by:

$$S_f(Q) = 1 + \frac{\mathcal{D}\Gamma(\mathcal{D}-1)\sin[(\mathcal{D}-1)\arctan(Q\xi)]}{(QL_a)^{\mathcal{D}}(1+(Q\xi)^2)^{\frac{\mathcal{D}-1}{2}}} \qquad (5)$$

As the peptide concentration increases, a depression or 'correlation hole' in the scattering curve is observed in the mid-Q regime, corresponding to a depleted concentration of scatterers around a rod caused by repulsive interrod interactions. We model this feature using a mean-field isotropic structure factor expression derived by Schneider et al.[35] for weakly repulsive, infinitely long thin rods, based on a modified Yukawa-segment model.

$$S_{CH}(Q) = \frac{1}{1 + n \cdot v(Q) P_{thin\ cyl}(Q)} \qquad (6)$$

Where, the $v(Q)$ is the expression that captures the strength of the interaction, given by:

$$v(Q) = \frac{(\mu_L L_a e)^2 / k_B T \epsilon \epsilon_0}{Q^2 + \lambda_D{}^{-2}} \qquad (7)$$

The constant parameters k_B, ϵ, ϵ_0, N_A and e are the Boltzmann constant, permittivity of free space, dielectric constant of water at temperature, T, Avogadro's number, and the charge on an electron, respectively. Here, μ_L is the linear charge density, expressed as the number of charged groups per unit length of rod, and λ_D is the Debye Length in low ionic strength aqueous solution:

$$\lambda_D = \left[\frac{8\pi N_A e^2}{k_B T \epsilon \epsilon_0}\right]^{-0.5} = 27.2 \text{ Å} \qquad (8)$$

The thin-cylinder form factor used in eq. 6 is given by:[41]

$$P_{thincyl}(Q) = \frac{2}{QL_a}\int_0^{QL_a} t^{-1} \sin t \, dt - \left[\frac{\sin(QL_a/2)}{(QL_a/2)}\right]^2 \qquad (9)$$

The form factor expression, $P_{thin\ cyl}(Q)$, is introduced in the denominator of the structure factor $S_{CH}(Q)$, which is justified under the assumptions that rods repel only at contact points along their length and in the absence of any end-effects. A similar analytical expression has been employed for describing sterically interacting polymers via the Polymer-Reference Interaction Site Model, i.e., PRISM theory and random-phase approximation (RPA) theory [41, 43-45]. We chose Schneider et al.'s model over these models due to its simple physical explanation of the system without the introduction of complicated or indirect parameters. Furthermore, using the extracted linear charge density μ_L, the effective diameter (D_{eff}), which is the sum of the cross-section of the rod and the counter-ion double layer around it, can be calculated via Onsager's theory[46] modified for charged rods[47]:

$$D_{eff} = 2R + \lambda_D\{\ln(A) + 0.577 + \ln(2) - 0.5\} \quad (10)$$

Where, A is given by Stroobants et al. [47]

$$A = 2\pi(\mu_L)^2 \lambda_D \lambda_B \exp\left(-2R.\lambda_D^{-1}\right) \qquad (11)$$

And, λ_B is the Bjerrum length in pure water given by:

$$\lambda_B = \frac{e^2}{4\pi\epsilon\epsilon_0 k_B T} \sim 7.1 \text{ Å} \qquad (12)$$

The total scattering expression is therefore:

$$I(\mathrm{Q}) = \mathrm{nV_p^2}(\rho_\mathrm{p} - \rho_\mathrm{s})^2\, \mathrm{P_{cyl}}(\mathrm{Q}).\hat{\mathrm{S}}(\mathrm{Q}) \qquad (13)$$

where, the effective structure factor in **Model 1 i.e.** $\hat{\mathrm{S}}_1(Q)$ (see eq. 13) is a simple combination of the structure factors for fractal structure and repulsive rods:

$$\hat{\mathrm{S}}_1(\mathrm{Q}) = \mathrm{S_{CH}}(\mathrm{Q}).[1 + \beta(\mathrm{Q})(\mathrm{S_f}(\mathrm{Q}) - 1)] \quad (14a)$$

Here, $\hat{\mathrm{S}}_1(Q)$ captures the hierarchy of structures in the semi-dilute and isotropic rod solution, i.e. $S_{CH}(Q)$ is a mean-field isotropic structure factor that describes repulsive inter-rod interactions while $S_f(Q)$ of rods describes a rod cluster that is isotropic by virtue of being corrected by the beta approximation factor $\beta(Q)$. It is important to note that weak anisotropic inter-rod repulsions in semi-dilute solution result in an isotropic distribution of rod particles, which is the simplest model that can be used to satisfactorily describe the SANS data. **Figure 2** illustrates the structure of the rigid rods as described by this combined model. In cases where no correlation hole is observed, such as for neutral rods, low rod concentration, or high salt concentration, we can substitute $S_{CH}(Q) = 1$ i.e. **Model 2**:

$$\hat{\mathrm{S}}_2(\mathrm{Q}) = 1 + \beta(\mathrm{Q})(\mathrm{S_f}(\mathrm{Q}) - 1) \qquad (14b)$$

From **Model 1** (eq. 14 a), we extracted four fit parameters i.e. rod radius R, apparent rod length within cluster L_a, fractal dimension $\mathcal{D}$, and linear charge density μ_L. **Model 2** (eq. 14 b) was used in cases where no correlation hole was apparent, giving three fit parameters, i.e., R, L_a, and $\mathcal{D}$. For both models, the cluster size, ξ, falls outside the length scales that are probed by the experiment and was set to an arbitrary large value of 10000 Å.

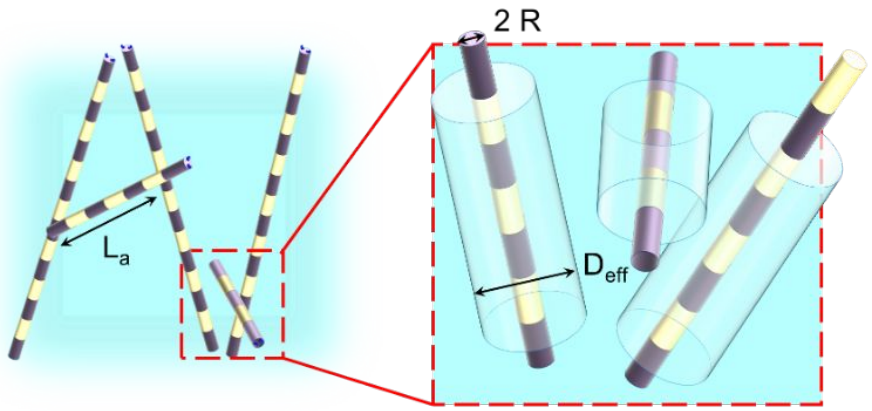


Figure 2: Illustration of the nanostructure of the rigid rod bundlemer chains in low ionic strength solution described by **Model 1**. The rigid rods form fractal networks with an apparent length L_a between contacts. Locally the rods are repulsive with an effective diameter D_{eff} greater than their bare cross section $2R$.

Results and discussion

Representative SANS curves for the homotetrameric building block **Peptide 2** versus ultra-long rigid rods that were constructed using equimolar ratio (r=1) of thiol (**Peptide 2**) to maleimide (**Peptide 1**) groups are shown in **Figure 3 (A)**. TEM analysis of the ultra-long bundlemer chains constructed by the hybrid physical-chemical assembly pathway resulted in rigid rod-like assemblies that were over one micron in length (**Figure 3 (B) & 5**), consistent with previous literature.[28] Therefore, all SANS data has been fitted using a cylinder form factor $P_{cyl}(Q)$ (eq. 3). The fits to individual bundlemers (**Peptide 2**) versus ultra-long rigid rods in **Figure 3 (A)** indicates that both bundlemers and rods have comparable radii R of 8.0 Å ± 0.5 Å and 9.2 Å ± 0.3 Å respectively**,** which is also the computationally designed radius of a single coiled coil bundlemer measured previously.[27] Also, the scattering curve for the long rods follows a power law $I(Q) \sim Q^{-1}$ for $Q < R^{-1}$, indicating a long, cylindrical nanostructure. In contrast, the scattering curve for the individual bundlemer solution plateaus at low Q due to their short length of (33 ± 2 Å). In light of these SANS measurements on the ultra-long rigid rods and since the overall lengths are beyond the measurable Q-regime using SANS, we focused this solution behavior study on three shorter length distributions of rigid rods.

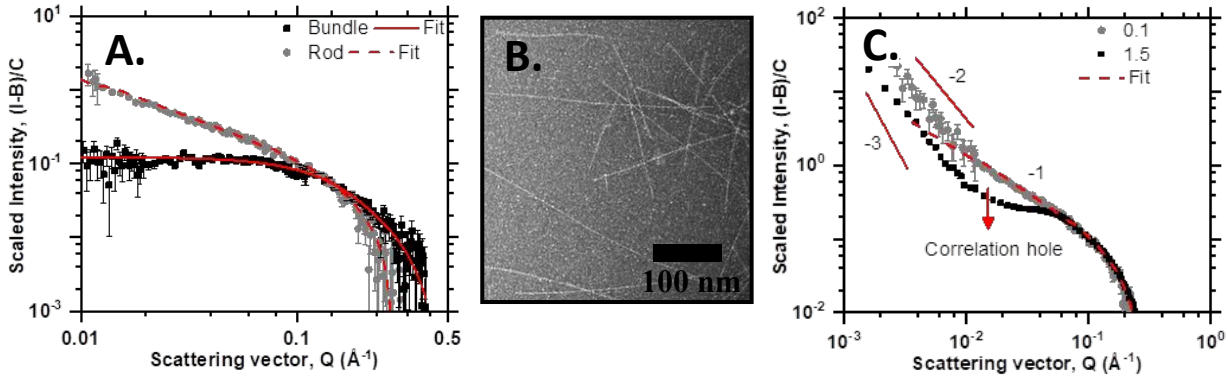


Figure 3: **(A)** SANS curves for **Peptide 2** versus ultra-long rigid rods of alternating **Peptide 1** and **Peptide 2** bundlemers. The data was modeled using a cylinder form factor plotted as red fit lines. **(B)** Negatively-stained TEM micrograph of ultra-long rigid rods that confirms their cylindrical shape. **(C)** Scaled SANS data of rigid rods at low and high concentrations C emphasizing the scattering features, i.e. the mid and high-Q correspond to the shape of the rod, the correlation hole at high concentrations indicate repulsive interactions and the low-Q upturn indicates mass-fractal structures. The legend symbols represents mass by volume fraction (in percent) of peptides in solution. The red solid lines represent slopes on the log-log plot and the red dashed line is a cylinder fit to the 0.1 % data. All scattering intensities I are background subtracted (-B**)** and normalized by sample concentration C.

As discussed earlier, step-growth polymerization kinetics of the Thiol-Michael *click* reaction enables the control of the average rod length, L, of bundlemers by varying the stoichiometric ratio, r, of thiol-to-maleimide functional groups.[37] Three rigid rod samples were constructed by increasing r from a mole ratio of 0.5 to just under 1, resulting in short, medium and long rigid rods. TEM analyses of the samples confirm that average length of the resulting rod bundlemers increased as r approached unity (see **Figure 4**) with polydisperse rod samples comprised mainly of short (145 ± 34 Å), medium (528 ± 149 Å) and long rods (1099 ± 706 Å) (length statistics are summarized in **SI Table S1**). Thus, light scattering techniques that are sensitive to dispersity in lengths and aggregation were not suitable for extracting useful information about the sizes and interactions in the rigid rod systems. Also, in order to avoid artefacts due to beam damage and presence of hydration layer or counterions in solution, we chose neutrons (SANS) over X-ray (SAXS) as our choice of radiation source to further investigate the solution-structure and interactions in the rigid rod system.

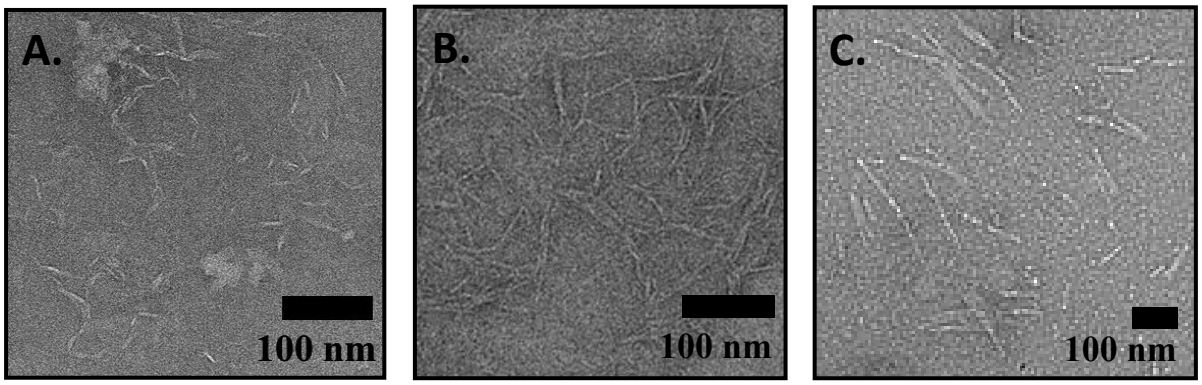


Figure 4: Negatively stained cast film TEM micrographs of 0.1 % (mass by volume) samples of **(A)** short, **(B)** medium and **(C)** long rods constructed by changing the stoichiometric ratio *r* of thiol containing **Peptide 2** to maleimide decorated **Peptide 1** bundlemers.

SANS curves and **Model 1** fits for the dilution series of the three samples under salt-free conditions are presented in **Figure 5**. The SANS curves have four distinct features from which corresponding model parameters can be independently extracted: the curvature at $Q \approx 0.15$ Å^{-1} corresponding to the radius of the rigid rods, a Q-regime with an $I(Q)$ slope of -1 indicative of its 1D shape, an upturn at low-Q having an $I(Q)$ slope between -3 and -2 indicating presence of a mass fractals, and a depression in scattering signal near the transition between these two slopes that becomes significant at higher concentrations alluding to presence of a correlation hole. The sample averages for these parameters are reported in **Table 1** (See **SI Table S3** for all fit results) and each will be discussed in detail in the rest of the article. It is important to emphasize here that the concentrations probed in this study were close to or above the critical overlap concentration for dilute to semi-dilute regime for all rigid rod samples i.e. $C > C^* \approx 1$ rod / L^3.[35] Scattering data at concentrations in the true dilute regime ($C < C^*$) were not accessible due to poor counting statistics at lower concentrations. (see **SI Table S1** for C^* values).

In all SANS curves, the scattering feature in the vicinity of $Q \approx 0.15$ Å^{-1} is related to the rod cross-section. **Model 1** fit results for the radius R of rigid rods are in good agreement with the coiled coil radius as discussed before, with a global average value of 11 Å ± 1 Å. This corroborates the design strategy employed for bundlemer assembly, confirming that the rigid rods are indeed formed by an end-

to-end assembly of the bundles as depicted in **Figure 1 (C)**. Also, as discussed earlier, for the lowest concentration samples, the continuation of the scattering feature corresponding to the radius of the rigid rod into a power law $I \sim Q^{-1}$ (0.01 Å^{-1} < Q < 0.1 Å^{-1}) corresponds to the presence of rods in solution.

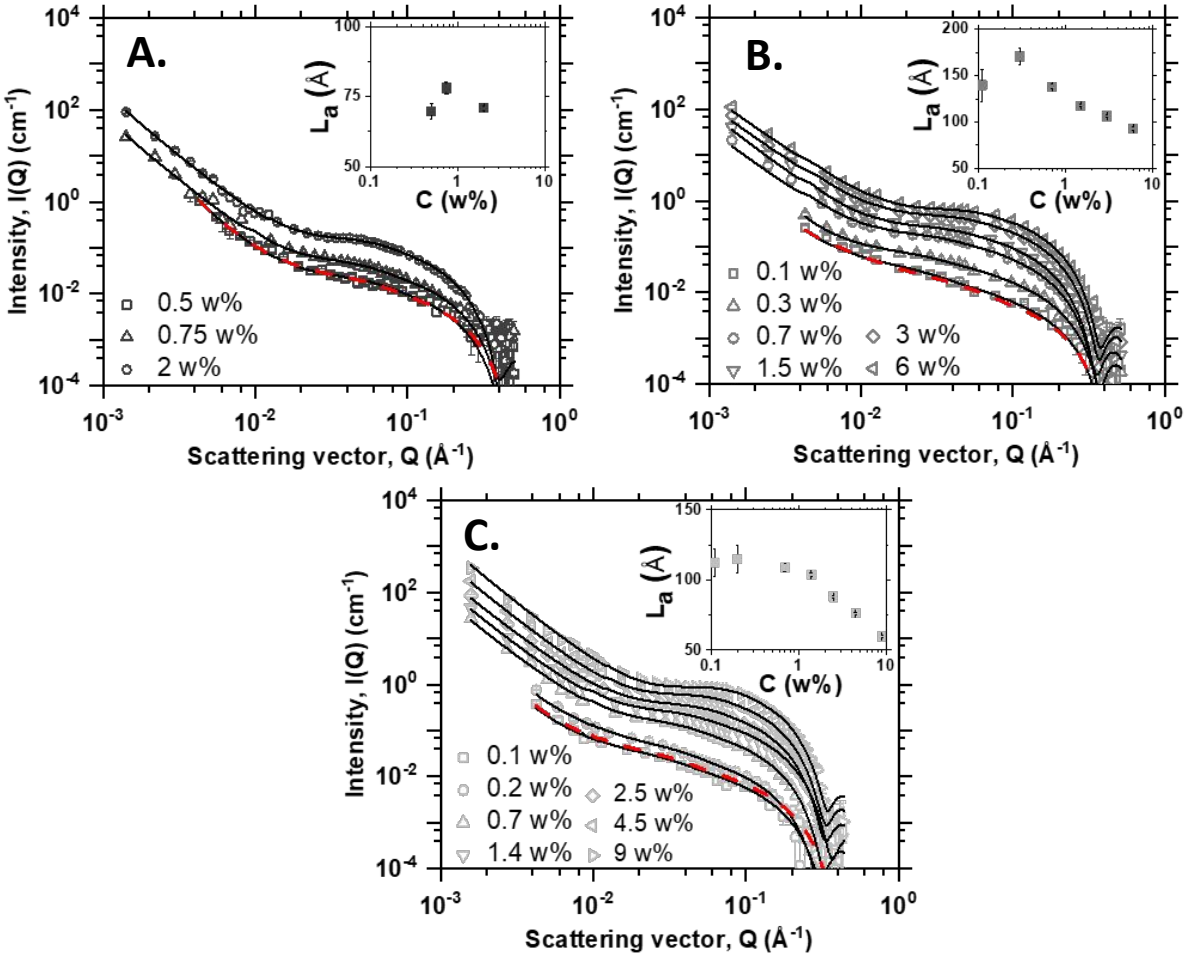


Figure 5: SANS data and corresponding fits for a dilution series of **(A)** short, **(B)** medium and **(C)** long rigid rods in pure D_2O. Here, w% notation in the legends and insets is the mass by volume fraction concentration C of peptides in solution. Black lines are fits using **Model 1**. The lowest concentration fit to **Model 2** is shown as a red dashed line in each case. The inset (top-right) within a graph is a plot of the fitted parameter apparent length L_a versus C that shows the decreasing relation between L_a and C.

Table 1: Model 1 fit results for radius, R, charge density, μ_L, and fractal dimension, D, for data shown in **Figure 5.** Standard error and sample means of R, μ_L, and D are reported here. The number in parentheses indicates number of individual measurements used for the calculation. The effective diameter, D_{eff} is calculated via eq. (10) and error propagation is used to calculate uncertainty (±1σ) in its value (see **SI**).

Sample	Radius R (Å)	Charge density μ_L (e/Å)	Effective diameter D_{eff} (Å)	Fractal dimension D
Short	9.3 ± 0.4 *(3)*	0.07 ± 0.01 *(2)*	71 ± 3	2.6 ± 0.1 *(3)*

Medium	10.5 ± 0.5 *(6)*	0.08 ± 0.01 *(5)*	75 ± 6	2.3 ± 0.1 *(5)*
Long	11.8 ± 0.5 *(7)*	0.07 ± 0.01 *(5)*	71 ± 8	2.5 ± 0.1 *(5)*

We were unable to extract the true length or even the persistence length of the rigid rods via form factor analysis for all samples since the rods are in the semi-dilute concentration regime ($C > C^*$). This results in significant overlap of rod excluded volume, giving rise to an upturn in scattering intensity at the lowest Q values ($Q < 0.01$ Å^{-1}). The upturn follows the power law $I \sim Q^{-D}$ ($D > 1$) which is characteristic of mass fractal-like cluster aggregates with fractal dimension D, where a larger value of D indicates a denser fractal structure. Here, we comment on the origin of the low-Q upturn for low-ionic strength rod solution (see **Figure 5**) measured for samples having a mass by volume fraction concentration C greater than 0.5 % since for these samples, the fractal dimension was recorded over a decade of Q-values (0.001 Å^{-1} < Q < 0.01 Å^{-1}). Phenomenologically, aggregation of rods can take place in semi-dilute concentrations via the previously reported Diffusion Limited Cluster Aggregation (DLCA) pathway that results in aggregates having a fractal dimension $D \approx 1.8$ for spheres having aspect ratio of 1 to $D \approx 2.2$ for rods having an aspect ratio (L/R) of 30.[29, 32] Multiple studies suggest that rod-like colloidal particle suspensions may also form percolated clusters as a direct result of their anisotropic shape.[48, 49] However, clustering due to competing short range attraction versus long range repulsion (SALR) [50] interactions have been reported for semi-dilute solutions of polyelectrolytes such as peptide fibrils,[51] DNA[52] and cellulose nanocrystals (CNCs)[53]. Polyelectrolyte clustering is therefore mechanistically different from the aforementioned entropically-driven DLCA pathways that describe aggregation in colloidal rod-like particles such as single-walled CNTs.[34, 54, 55] Since the peptide-based rigid rods investigated here have multiple solvent-exposed oppositely-charged groups as well as polar and hydrophobic surface patches, they can interact by a combination of interactions, *i.e.* weak short-range van der Waals or London dispersive forces and strong long-range electrostatic and entropic (hydrophobic) forces, that can drive formation of denser cluster aggregates with a larger fractal dimension of $2.3 < D < 2.6$ as observed here.[56, 57] Interestingly, while a variation in D with rigid rod sample type (short *vs* medium *vs* long) may

be attributed to length polydispersity, sample handling, and preparation, the modeling results indicate that D remains unchanged within a dilution series, alluding to a favored packing of sequestered rods within a cluster aggregate.

In the mid- to low-Q regime, the cross-over from $I \sim Q^{-1}$, signifying a straight cylinder, to $I \sim Q^{-D}$, signifying a fractal-like object, yields an apparent length, L_a, of rods in the rigid rod sample. The fitted parameter L_a characterizes the average length of rods between contact points within the cluster and is therefore shorter than the actual length L of the rigid rods measured via TEM.[54, 55, 58] This parameter is also related to the average pore size or mesh size within a network of polyelectrolyte chains.[59] As the concentration of the rods increases, the fitted L_a decreases (see insets in **Figure 5**). This can be attributed to crowding of rods within a cluster that results in an increase in the number of inter-rod contacts and a decrease in the average distance between these points of contact. For the lowest concentration rigid rod samples, the data can be satisfactorily modeled using **Model 2**, describing a fractal-network structure of rods (see **Figure 2**). However, at higher concentrations, repulsive interactions lead to significant correlation in the spacing of the individual rods. This is manifested as a depression in the scattering intensity in the mid-Q region, i.e. a correlation hole, for $Q \approx 0.01 Å^{-1}$. The approximate location of the depression increases to larger Q with increasing weight percent of the peptide, paralleling the trend observed for L_a^{-1} and suggesting that this interaction distance corresponds to the mesh size of the networked cluster. While rigid rods constructed using bundlemers may be described as polyelectrolyte-like chains of charged bundlemer units, the correlation hole in the rigid rod systems is distinct from the correlation hole driven 'polyelectrolyte peak' that has been widely reported in scattering experiments from low-ionic strength suspensions of polyelectrolytes.[34] The correlation peaks at $Q^* \sim C^{1/2}$ in traditional polyelectrolyte systems are indicative of long-range perturbations and preferred distances.[30, 60] In the present study, no such peak in scattering intensity was recorded even at high peptide concentrations. Sequence-based charge prediction[61] of the peptides in **Peptide 1** and **Peptide 2** bundlemers can give us insight into this behavior of the rigid rod assemblies via charge analysis of the individual bundlemers. At neutral pH, assuming all ionizable groups are fully dissociated, bundlemers of **Peptide 1** and **Peptide 2**

can carry a maximum net electronic charge of ca. -4*e* and -12*e*, respectively. Thus, the rigid rods that comprise of alternating **Peptide 1** and **Peptide 2** bundlemers must have a patchy charge distribution of alternating net charges and can display a maximum net linear charge density of ≈ -0.2 *e*/Å. However, complete dissociation of oppositely charged groups is unlikely, resulting in weakly repulsive rigid rods with a net charge density smaller than widely studied rod systems such as DNA (–0.6 *e*/ Å)[62, 63], TMV (–1-2 *e*/ Å)[64] and amyloid fibers (–0.5 *e*/ Å)[65]. Thus, we infer that the rigid rods are weak polyelectrolytes at neutral pH, i.e., the repulsive interactions between them are not strong enough to cause long-range perturbations typically responsible for the correlation peak observed in other polyelectrolyte-like suspensions.[30, 45, 52, 60, 66]

The scattering profile due to the weak polyelectrolyte character of rigid rods is adequately described by the mean-field structure factor developed by Schneider et al. for weakly-repulsive thin rods, which uses a linear solution to the Poison-Boltzmann potential, i.e., the Debye-Hückel potential, to model the interrod repulsion.[35] This expression gives a physically accurate description of the rigid rod system and has been shown to satisfactorily describe solution structure of weakly-interacting rods observed in both experiments and simulations.[67, 68] The net linear charge density, μ_L, for the rods is obtained by fitting the data with this structure factor expression and is summarized in **Table 1**. The resulting linear charge density averaged over all rigid rod samples is -0.08 ± 0.01 *e*/ Å, which is much smaller than the estimate of -0.2 *e*/ Å based on amino acid sequence of the bundlemer. Thus, the model suggests that not all ionizable groups are dissociated under the conditions considered, and the bundlemers carry a weak net charge of approximately -3*e* per bundlemer along the rod. The magnitude of this effective charge per length may also be attenuated by the presence of counter ions. More information can be extracted via Onsager's theory[46] modified for charged rods[47] that gives the effective diameter (D_{eff}) of the rigid rods in solution and is summarized in **Table 1**. The calculated D_{eff} of the rods is ≈ 71 Å, which is larger than their bare rod cross-section $D = 2R$ ($D_{eff} \approx 3.5\ D$); this is expected in a low ionic strength solution due to poor screening of net charge on the rigid rod surface, which is effectively the Stern Layer. However, the Oosawa-Manning criterion for counterion condensation ($\lambda_B.\mu_L > 1$) [66] , albeit most relevant to solutions

with multivalent counterions, is not satisfied for this system ($\lambda_B.\mu_L \sim 0.54$). The Stern layer of charge at the rod surface is likely not formed or sparsely populated with counterions. Therefore, we conclude that the electric double layer, i.e., the Gouy-Chapman Double Layer, consists primarily of loosely bound and diffuse counterions under neutral pH conditions**.**

To confirm that the correlation hole results from electrostatic repulsion between like-charged rods at these low ionic strengths, we performed SANS experiments with addition of 200 x 10^{-3} mol/L NaCl to the long rigid rod sample. The SANS data for these samples are shown in **Figure 6 (B)**. In this case, the correlation hole disappears; the scattering curve shows a direct transition from a $I \sim Q^{-1}$ dependence to a $I \sim Q^{-D}$ dependence without an intermediate depression in scattering. This solution behavior is because the charged side groups are now screened by ions forming a tightly bound electric double layer around the rigid rods. Thus, the effective diameter of the rods decreases ($D_{eff,\ \mathrm{NaCl}} < D_{eff,\ water}$) and repulsion between rods is screened at contact points within the cluster.[30, 66] This difference is depicted in the schematic in **Figure 6 (A)**. Consequently, the scattering curve can be modeled more effectively as neutral rods via **Model 2**. The results for fitted parameters obtained from **Model 2** for this case are given in **Table 2**. We also used **Model 1** to fit the data assuming a fixed linear charge density $\mu_L = 0.08\ e/Å$, corresponding to the value obtained for the samples without added NaCl, and fixed salt concentration of $C_s = 200$ x 10^{-3} mol/L for comparison (see **Figure 6 (B), SI Table S3** for all fit results). The results for the high ionic strength case are further corroborated by studies at acidic pH, which resulted from the addition of 200 x 10^{-3} mol/L deuterium chloride (DCl, pH ≈ 1) to the long-rigid rods solution. Although sequence-based charge analysis suggests that the rigid rods should carry a large positive charge, no depression in the curve is evident in this case (see **Figure 6 (C)**). We infer that the presence of ample chloride ions (Cl^{-1}) in this solution effectively screens the positive charge on the rod in a similar manner to the high ionic strength (200 x 10^{-3} mol/L NaCl) case. We fit this data using **Model 2** and the results are presented in **Table 2**. The similar effect of either salt and pH on the scattering curve confirms the electrostatic origin of the correlation hole, which is sensitive to the presence of electrolytes in solution.

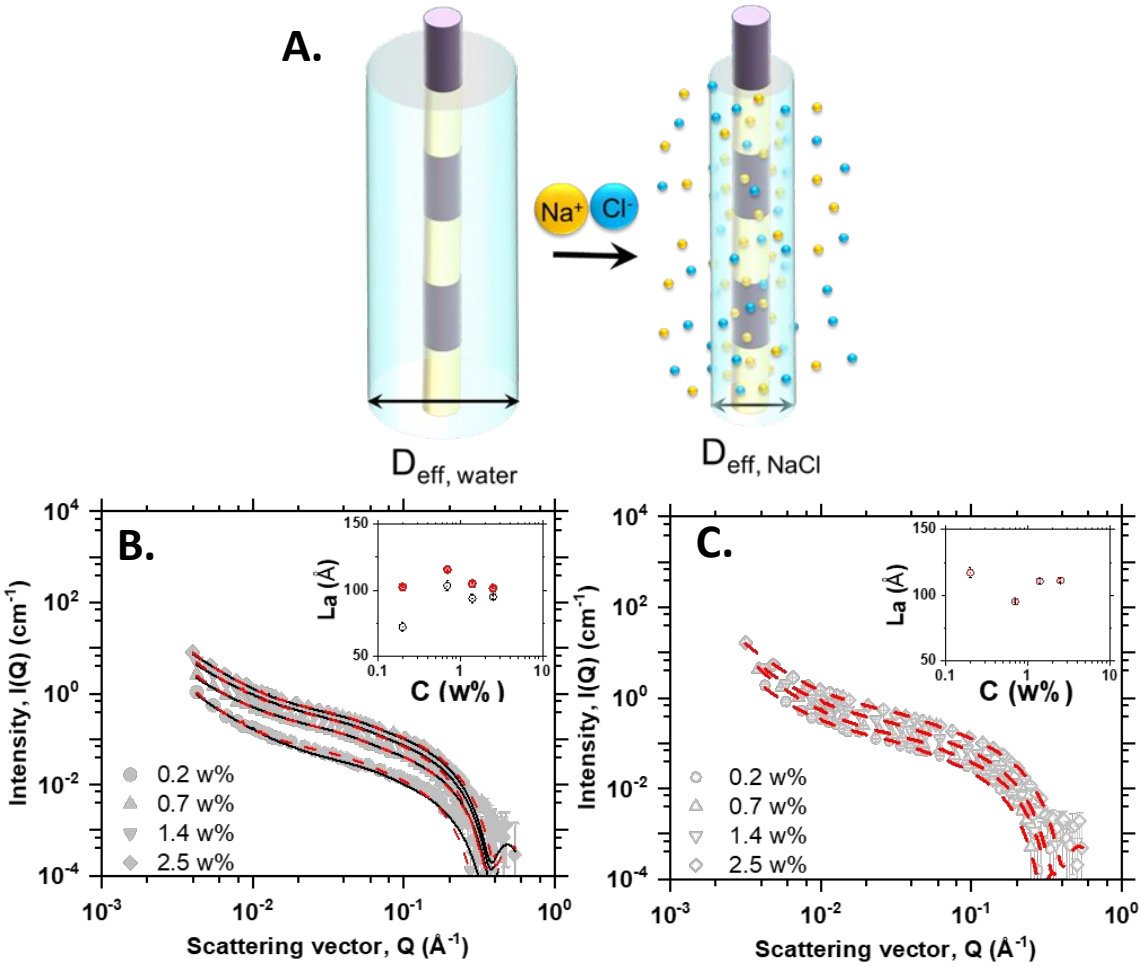


Figure 6: (A) Schematic illustrating the decrease in the effective diameter of the charged rods on addition of monovalent sodium chloride (NaCl) salt. The salt ions screen the net charge on the rods, thus reducing the thickness of the electric double layer. **(B)** SANS data for long rods in 200 x 10^{-3} mol/L NaCl. The data is fit using Models 1 (black line) and 2 (red dashed line). The apparent length L_a vs rod concentration C is plotted as inset. **(C)** SANS data for long rods in 200 x 10^{-3} mol/L deuterium chloride (DCl, pH ≈ 1). These data were fit using Model 2 and fitted L_a versus C in this case is shown (inset). The notification of w% in the legend and inset in each plot represents the mass by volume fraction concentration *C* of the rods.

Table 2: Results for fits to SANS data shown in **Figure 6** using **Model 2** i.e. long rods with added 200 mM sodium chloride (NaCl salt) and in acidic pH with added 200 mM deuterium chloride (DCl). The sample mean and standard error in radius, *R* and fractal dimension, *D* are reported here. The effective diameter, D_{eff} and its uncertainty (±1σ) are calculated via eq.(10) and error propagation respectively (see **SI**).

Sample	**Radius R (Å)**	**Effective diameter $D_{eff,NaCl}$ (Å)**	**Fractal dimension D**
Added Salt	11.2 ± 1.0	22 ± 2	2.2 ± 0.1
Acidic pH	11.6 ± 1.4	23 ± 3	2.2 ± 0.1

We have discussed earlier in this article that **Peptide 1** and **Peptide 2** bundlemers carry different theoretical net charges at neutral pH, -4 *e* and -12 *e* respectively, which alludes to a patchy distribution of alternating net charge densities along the lengths of the rigid rods. Therefore, the depiction of a homogenous effective diameter in **Figure 2** and **6 (A)** may be an over-simplification of the true counterion cloud which may display local counterion distribution and hydration differences. However, the fit quality of **Model 1** to SANS data in low-ionic strength solution and **Model 2** in high-ionic strength solution suggests that an average bundlemer charge of 3*e* in neutral pH is sufficient to describe their solution-state structure and interactions under probed concentration and solution conditions.

In our previous publication, we reported that the persistence length l of the rigid rods is unprecedented, with initial analyses estimating l to be upwards of ten micrometers. Using the linear charge density μ_L extracted from SANS structure factor fitting, we can estimate the contribution to this persistence length l by the electrostatic repulsion between bundlemer units, which is expected to result in stiffening of polyelectrolyte chains in the low salt limit. Using the Odijk-Skonick-Fixman (OSF) model,[69, 70] the total persistence length l can be expressed as the sum of the intrinsic persistence length l_o and the electrostatic persistence length l_e (i.e., $l = l_o + l_e$), where the electrostatic persistence length l_e was derived using the Debye-Hückel interaction potential between charged monomers along the chain. For the rigid rods constructed using alternating **Peptide 1** and **Peptide 2** bundlemers, this results in:

$$\mathrm{l_e} = {\lambda_B \cdot \mu_L^2}\Big/{4\lambda_D^{-2}} = 7.6\ \text{Å}$$

Thus, the contribution of the electrostatic persistence length to the total persistence length is small ($l \gg l_e$), implying that the rigid rods are in fact intrinsically stiff polymers of coiled coil bundlemers ($l \sim l_o$).

Since the rigid rods are cylinders with a large persistence length, one might expect the formation of liquid crystalline phases at high rigid rod concentrations. However, we do not observe birefringence in any rigid rod sample at neutral pH. Onsager's theory for rigid rods that interact via hard-core repulsive interactions can be used to calculate an approximate onset concentration for liquid crystal formation. For

monodisperse stiff rod-like particles, the theoretical cross-over concentration from isotropic to nematic ($I \rightarrow N$) liquid crystal phase is given by the expression:[71]

$$C_{I\rightarrow N} = 4\frac{D_{eff}}{L}\vartheta$$

Where, ϑ is the density of the rods ($\vartheta = 1.35\ g/ml$[72, 73]). This equation suggests that longer and thinner rods form liquid crystal (LC) phases at lower concentrations. The theoretical $C_{I\rightarrow N}$ for the long rigid rods (L=1095 Å, D_{eff} = 71 Å) is calculated to be roughly a mass by volume fraction of 35 %. For the medium and short rigid rod samples, the onset concentration will be even higher. Thus, in all cases, the onset concentrations are much higher than the concentrations that were presently studied. It has also been proposed by Odijk that the $C_{I\rightarrow N}$ in suspensions of charged rods increases due to twisting of charged rods with respect to one another.[47] These theoretical considerations support the observation that LC phases do not form at neutral pH conditions for the rigid rod samples at probed concentrations. Polydispersity in rod lengths has been reported to result in a broad biphasic concentration regime in which longer rod populations sequentially separate into LC enriched phases at concentrations much lower than the sharp $C_{I\rightarrow N}$ of an equivalent monodisperse rod solution.[74, 75] The absence of such LC phases in the polydisperse rigid rod solutions studied here may be due to attractive forces that hold together rod clusters and hinder macroscopic rearrangement of the rods into LC phases. This argument is also supported by the constant slope of the upturn in low-Q region of SANS indicating that the internal fractal structure of the cluster aggregate remains unchanged for a given rigid rod sample. Similar arguments can be made for the absence of birefringence in the added salt case wherein the interrod repulsions are further screened by salt ions. Interestingly, lamellar-like liquid crystalline droplets form specifically in acidic pH conditions for the long rigid rod samples, which we have reported elsewhere (data not shown). Overall, a balance of the magnitude of net charge and the balance of opposite charges along the rod length combined with average rod length and distributions inhibits lyotropic liquid crystal formation here.

Conclusion:

Computationally designed coiled coils that self-assemble into target nanostructures under mild aqueous conditions provide a toolbox for material construction and discovery. We have utilized a hybrid physical-covalent assembly pathway to construct polymers of bundlemers via thiol-Michael *click* reactions. The resultant bundlemer chains were observable via negatively stained TEM characterization. SANS confirmed that these bundlemer rigid rod chains have the same diameter as the computationally designed coiled coils.

The solution structure and interactions in the model rigid rod system were studied in detail via a combination of SANS, TEM and bundlemer-design. The scattering data were modeled using a straight cylinder form factor and a composite structure factor comprised of a fractal scattering function modified for rods coupled with a mean-field structure factor that encodes weak, inter-rod repulsion within the fractal network. The rigid rod solution was heterogeneous wherein the rods formed fractal-like networks or cluster aggregates giving rise to a low-Q upturn in the scattering intensity. The apparent length of the rods within the cluster, related to its mesh size, decreased as the rigid rod concentration increased as more rods were incorporated into the clusters. Furthermore, we showed that the rods are electrostatically repulsive due to the presence of a small net charge as evidenced by the presence of a correlation hole at higher rigid rod concentrations. The presence of a weak net charge is also supported by sequence-based charge analysis of the peptides used to build the rods as well as from fits of an inter-rod structure factor to mid-Q SANS data. Consequently, the addition of monovalent ions to the solution screened the net charge, effectively suppressing the correlation hole. This weak polyelectrolyte character results in formation of fractal clusters of repulsive rods at neutral pH in contrast to other widely reported polyelectrolyte rod materials where a scattering correlation peak, indicating longer-range order, is observed due to larger net charge density.

The balance between short-range attraction and long-range repulsion interactions in protein-based biomaterials is a product of their underlying amino-acid sequence that is often over-simplified while describing their structure-property attributes. The model bundlemer rigid rod system studied here is unique in that the rods are intrinsically stiff polymers that are readily viewable and easily modifiable by

virtue of their hierarchical assembly and computational design. The patchy local distribution of interaction sites on naturally occurring assemblies can also be readily mimicked on the surface of the bundlemer-based polymers since these are modularly built using orthogonal chemical reactions between designable bundlemer units. The bundlemer-based assemblies can thus provide insights into important sequence-driven structure-property relationships in complex, protein-based materials via future systematic sequence manipulation enabled by their computational design.

Acknowledgements:

This manuscript was prepared under cooperative agreement #370NANB17H302 from National Institute of Standards and Technology (NIST), U.S. Department of Commerce. We acknowledge the support of the NIST, U.S. Department of Commerce in providing the neutron research facilities used in this work. Access to NGB 30m SANS was provided by the Center for High Resolution Neutron Scattering (CHRNS), a partnership between the National Institute of Standards and Technology and the National Science Foundation under Agreement No. DMR-1508249. This work utilized facilities supported in part by the National Science Foundation (NSF) under Agreement No. DMR-0944772. This work also benefited from the use of the SasView application, originally developed under NSF award DMR-0520547. SasView contains code developed with funding from the European Union's Horizon 2020 research and innovation programme under the SINE2020 project, grant agreement No 654000. Support for peptide design, synthesis, purification, and bundlemer rod formation was provided by the Department of Energy, Office of Basic Energy Sciences, Biomolecular Materials Program under grant No. DE-SC0019355 and DE-SC0019282.

Disclaimer:

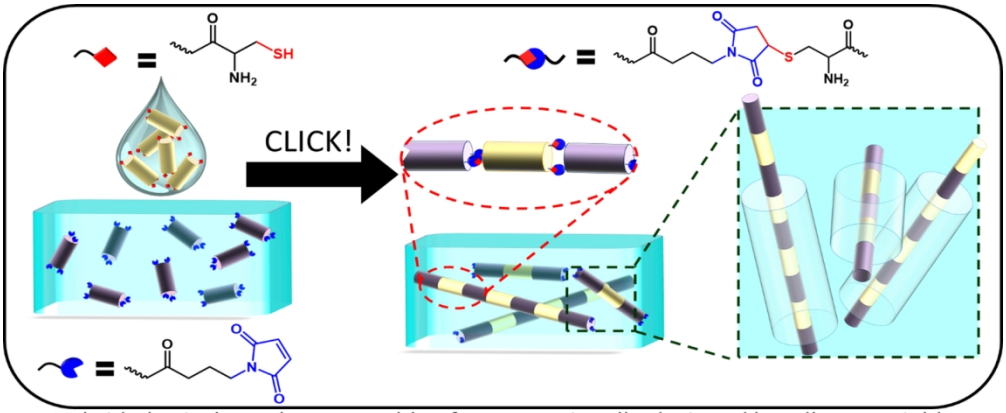
The statements, findings, conclusions and recommendations are those of the authors and do not necessarily reflect the view of NIST or the U.S. Department of Commerce. Certain commercial equipment, instruments, materials, suppliers and software are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the NIST, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

References:

1. D. Philp and J. F. Stoddart, *Angewandte Chemie International Edition in English*, 1996, **35**, 1154-1196.
2. G. M. Whitesides and M. Boncheva, *Proceedings of the National Academy of Sciences of the United States of America*, 2002, **99**, 4769-4774.
3. S. Zhang, *Nature Biotechnology*, 2003, **21**, 1171.
4. E. Winfree, F. Liu, L. A. Wenzler and N. C. Seeman, *Nature*, 1998, **394**, 539.
5. M. Antonietti and S. Förster, *Advanced Materials*, 2003, **15**, 1323-1333.
6. N. P. King, W. Sheffler, M. R. Sawaya, B. S. Vollmar, J. P. Sumida, I. André, T. Gonen, T. O. Yeates and D. Baker, *Science*, 2012, **336**, 1171.
7. T. Tørring, N. V. Voigt, J. Nangreave, H. Yan and K. V. Gothelf, *Chemical Society Reviews*, 2011, **40**, 5636-5646.
8. J. N. Zadeh, C. D. Steenberg, J. S. Bois, B. R. Wolfe, M. B. Pierce, A. R. Khan, R. M. Dirks and N. A. Pierce, *Journal of Computational Chemistry*, 2010, **32**, 170-173.
9. Y.-T. Lai, N. P. King and T. O. Yeates, *Trends in Cell Biology*, 2012, **22**, 653-661.
10. T. Kortemme and D. Baker, *Current Opinion in Chemical Biology*, 2004, **8**, 91-97.
11. B. I. Dahiyat and S. L. Mayo, *Science*, 1997, **278**, 82.
12. G. Colombo, P. Soto and E. Gazit, *Trends in Biotechnology*, 2007, **25**, 211-218.
13. S. Lyskov and J. J. Gray, *Nucleic Acids Research*, 2008, **36**, W233-W238.
14. K. A. Dill and J. L. MacCallum, *Science*, 2012, **338**, 1042.
15. H. Robson Marsden and A. Kros, *Angewandte Chemie International Edition*, 2010, **49**, 2988-3005.
16. P. Burkhard, J. Stetefeld and S. V. Strelkov, *Trends in Cell Biology*, 2001, **11**, 82-88.
17. C. Cohen and D. A. D. Parry, *Proteins: Structure, Function, and Bioinformatics*, 1990, **7**, 1-15.
18. D. N. Woolfson, in *Advances in Protein Chemistry*, Academic Press, 2005, vol. 70, pp. 79-112.
19. J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. Neil and W. F. DeGrado, *Science*, 1995, **270**, 935.
20. C. J. Lanci, C. M. MacDermaid, S.-g. Kang, R. Acharya, B. North, X. Yang, X. J. Qiu, W. F. DeGrado and J. G. Saven, *Proceedings of the National Academy of Sciences*, 2012, **109**, 7304.
21. A. R. Thomson, C. W. Wood, A. J. Burton, G. J. Bartlett, R. B. Sessions, R. L. Brady and D. N. Woolfson, *Science*, 2014, **346**, 485.
22. C. Xu, R. Liu, A. K. Mehta, R. C. Guerrero-Ferreira, E. R. Wright, S. Dunin-Horkawicz, K. Morris, L. C. Serpell, X. Zuo, J. S. Wall and V. P. Conticello, *Journal of the American Chemical Society*, 2013, **135**, 15565-15578.
23. E. N. G. Marsh and W. F. DeGrado, *Proceedings of the National Academy of Sciences*, 2002, **99**, 5150.
24. H. V. Zhang, F. Polzer, M. J. Haider, Y. Tian, J. A. Villegas, K. L. Kiick, D. J. Pochan and J. G. Saven, *Science Advances*, 2016, **2**, e1600307.
25. Y. Tian, H. V. Zhang, K. L. Kiick, J. G. Saven and D. J. Pochan, *Organic & Biomolecular Chemistry*, 2017, **15**, 6109-6118.
26. Y. Tian, F. B. Polzer, H. V. Zhang, K. L. Kiick, J. G. Saven and D. J. Pochan, *Biomacromolecules*, 2018, **19**, 4286-4298.
27. M. J. Haider, H. V. Zhang, N. Sinha, J. A. Fagan, K. L. Kiick, J. G. Saven and D. J. Pochan, *Soft Matter*, 2018, **14**, 5488-5496.
28. D. Wu, N. Sinha, J. Lee, B. P. Sutherland, N. I. Halaszynski, Y. Tian, J. Caplan, H. V. Zhang, J. G. Saven, C. J. Kloxin and D. J. Pochan, *Nature*, 2019, **574**, 658-662.
29. M. J. Solomon and P. T. Spicer, *Soft Matter*, 2010, **6**, 1391-1400.

30. M. Muthukumar, *Macromolecules*, 2017, **50**, 9528-9560.
31. F. Horkay, P. J. Basser, A.-M. Hecht and E. Geissler, *The Journal of Chemical Physics*, 2018, **149**, 163312.
32. A. Mohraz, D. B. Moler, R. M. Ziff and M. J. Solomon, *Physical Review Letters*, 2004, **92**, 155503.
33. B. A. Schubert, E. W. Kaler and N. J. Wagner, *Langmuir*, 2003, **19**, 4079-4089.
34. C. E. Williams, M. Nierlich, J. P. Cotton, G. Jannink, F. Boué, M. Daoud, B. Farnoux, C. Picot, P. G. DeGennes, M. Rinaudo, M. Moan and C. Wolff, *Journal of Polymer Science: Polymer Letters Edition*, 1979, **17**, 379-384.
35. J. Schneider, W. Hess and R. Klein, *Journal of Physics A: Mathematical and General*, 1985, **18**, 1221.
36. J. Teixeira, *Journal of Applied Crystallography*, 1988, **21**, 781-785.
37. P. J. Flory, *Journal of the American Chemical Society*, 1936, **58**, 1877-1885.
38. C. A. Schneider, W. S. Rasband and K. W. Eliceiri, *Nature Methods*, 2012, **9**, 671.
39. S. Kline, *Journal of Applied Crystallography*, 2006, **39**, 895-900.
40. http://www.sasview.org/.
41. T. Zemb and P. Lindner, *Neutrons, X-rays and Light: Scattering Methods Applied to Soft Condensed Matter*, Elsevier, 2002.
42. M. Kotlarchyk and S. H. Chen, *The Journal of Chemical Physics*, 1983, **79**, 2461-2469.
43. K. S. Schweizer and J. G. Curro, *Physical Review Letters*, 1987, **58**, 246-249.
44. L. Arleth, M. Bergström and J. S. Pedersen, *Langmuir*, 2002, **18**, 5343-5353.
45. E. Josef and H. Bianco-Peled, *Soft Matter*, 2012, **8**, 9156-9165.
46. L. Onsager, *Annals of the New York Academy of Sciences*, 1949, **51**, 627-659.
47. A. Stroobants, H. N. W. Lekkerkerker and T. Odijk, *Macromolecules*, 1986, **19**, 2232-2238.
48. L. Wu, C. P. Ortiz and D. J. Jerolmack, *Langmuir : the ACS journal of surfaces and colloids*, 2017, **33**, 622-629.
49. P. van der Schoot and T. Odijk, *The Journal of Chemical Physics*, 1992, **97**, 515-524.
50. Y. Liu and Y. Xi, *Current Opinion in Colloid & Interface Science*, 2019, **39**, 123-136.
51. M. Owczarz, S. Bolisetty, R. Mezzenga and P. Arosio, *Journal of Colloid and Interface Science*, 2015, **437**, 244-251.
52. R. Borsali, H. Nguyen and R. Pecora, *Macromolecules*, 1998, **31**, 1548-1555.
53. H. Oguzlu, C. Danumah and Y. Boluk, *Current Opinion in Colloid & Interface Science*, 2017, **29**, 46-56.
54. R. S. Justice, D. H. Wang, L.-S. Tan and D. W. Schaefer, *Journal of Applied Crystallography*, 2007, **40**, s88-s92.
55. W. Zhou, M. F. Islam, H. Wang, D. L. Ho, A. G. Yodh, K. I. Winey and J. E. Fischer, *Chemical Physics Letters*, 2004, **384**, 185-189.
56. J. Ray and G. S. Manning, *Macromolecules*, 2000, **33**, 2901-2908.
57. A. Chremos and J. F. Douglas, *The Journal of Chemical Physics*, 2017, **147**, 241103.
58. D. W. Schaefer, J. Zhao, J. M. Brown, D. P. Anderson and D. W. Tomlin, *Chemical Physics Letters*, 2003, **375**, 369-375.
59. Y. Zhang, J. F. Douglas, B. D. Ermi and E. J. Amis, *The Journal of Chemical Physics*, 2001, **114**, 3299-3313.
60. S. H. Chen, E. Y. Sheu, J. Kalus and H. Hoffman, *Journal of Applied Crystallography*, 1988, **21**, 751-769.
61. D. S. Moore, *Biochemical Education*, 1985, **13**, 10-11.
62. M. T. Record, S. J. Mazur, P. Melancon, J. H. Roe, S. L. Shaner and L. Unger, *Annual Review of Biochemistry*, 1981, **50**, 997-1024.
63. R. Gilbert, *Crystallography Reviews*, 2009, **15**, 285-288.

64. S. Fraden, in *Observation, Prediction and Simulation of Phase Transitions in Complex Fluids*, eds. M. Baus, L. F. Rull and J.-P. Ryckaert, Springer Netherlands, Dordrecht, 1995, DOI: 10.1007/978-94-011-0065-6_3, pp. 113-164.
65. R. Mezzenga, J.-M. Jung and J. Adamcik, *Langmuir*, 2010, **26**, 10401-10405.
66. A. V. Dobrynin and M. Rubinstein, *Progress in Polymer Science*, 2005, **30**, 1049-1118.
67. H. Löwen, *The Journal of Chemical Physics*, 1994, **100**, 6738-6749.
68. J.-P. Hansen and H. Löwen, *Annual Review of Physical Chemistry*, 2000, **51**, 209-242.
69. J. Skolnick and M. Fixman, *Macromolecules*, 1977, **10**, 944-948.
70. T. Odijk, *Journal of Polymer Science: Polymer Physics Edition*, 1977, **15**, 477-483.
71. G. J. Vroege and H. N. W. Lekkerkerker, *Reports on Progress in Physics*, 1992, **55**, 1241.
72. I. D. Kuntz and W. Kauzmann, in *Advances in Protein Chemistry*, eds. C. B. Anfinsen, J. T. Edsall and F. M. Richards, Academic Press, 1974, vol. 28, pp. 239-345.
73. K. Gekko and H. Noguchi, *The Journal of Physical Chemistry*, 1979, **83**, 2706-2714.
74. T. Odijk and H. N. W. Lekkerkerker, *The Journal of Physical Chemistry*, 1985, **89**, 2090-2096.
75. P. A. Buining and H. N. W. Lekkerkerker, *The Journal of Physical Chemistry*, 1993, **97**, 11510-11516.



Hybrid physical-covalent assembly of computationally-designed bundlemers yields polyelectrolyte-like rigid rods displaying rich solution behavior

78x38mm (600 x 600 DPI)