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

Supplementary information (SI)

Thiol-Michael *click* reaction considerations

The step-wise hybrid physical-covalent assembly has been optimized for the **Peptide 1** and **Peptide 2** system in our previous publication.[1] A typical Thiol-Michael *click* reaction involves the formation of thiolate anion from a free thiol which then reacts with the electron deficient maleimide group to form a Thiol-Michael adduct.[2] In the presence of a nucleophile like TCEP (Tricarboxyethylphosphine), that also prevents disulfide formation of thiols, the thiolate anion is created via the nucleophile-catalyzed route.

Since the Thiol-Michael *click* reaction follows step-growth polymerization kinetics, the number-average degree-of-polymerization (X_N) at full conversion was derived by Flory:[3]

$$X_N = \frac{1+r}{1-r} \qquad \text{(s1)}$$

Where r is the stoichiometric ratio of the thiol to maleimide functional groups, which is used to obtain rods of varying lengths. It is important to note that step-growth polymerization results in a molecular weight distribution which is characterized by Dispersity ($Đ$). After the reaction is complete, Dispersity,

which is the ratio of the weight-average degree of polymerization to the number average degree of polymerization, for an ideal step-growth polymerization is given by:[3]

$$Đ = 1 + \frac{4r}{(1+r)^2} \qquad \text{(s2)}$$

This eq. s2 indicates that a theoretical maximum of $Đ \rightarrow 2$ is expected as the stoichiometric ratio $r \rightarrow 1$. Even in the case of short rods where $r = 0.5$, this equation yields a length dispersity of 1.889. Therefore, the rod systems are expected to be inherently polydisperse in length due to the chemistry employed to connect the coiled coils.

Electron-Spray Ionization Mass Spectroscopy (ESI-MS)

Ultra-performance liquid chromatography (UPLC) and corresponding mass ionization state (*m/z*) of eluting species was measured on Xevo G2-XS QTof Quadrupole Time-of-Flight Mass Spectrometry instrument. Sequence based molecular weight estimates of **Peptide 1** and **Peptide 2** are 3682 Da and 3663 Da respectively. Reaction of bundlemers is expected to result in the formation of a thiol-maleimide adduct i.e. **Peptide 1-2** with an expected molecular weight of 7345 Da. The elution profiles and ionization states of species in the major elution peak for each bundlemer stock solution and the short rigid rods is presented in **Figure S1-S3**. Specifically, in the rigid rod solutions, an elution peak corresponding to a molecular weight of 7346.14 Da is observed which confirms the presence of the **Peptide 1-2** adduct after reaction (see **Figure S3**). Also, since the thiol group containing **Peptide 2** is the limiting reactant, the elution peak corresponding to **Peptide 2** disappears in the rod solution (see **Figure S3 (A)**), while the elution peak indicating presence of excess unreacted **Peptide 1** remains. These results confirm the success of the Thiol-Michael *click* reaction between bundlemers.

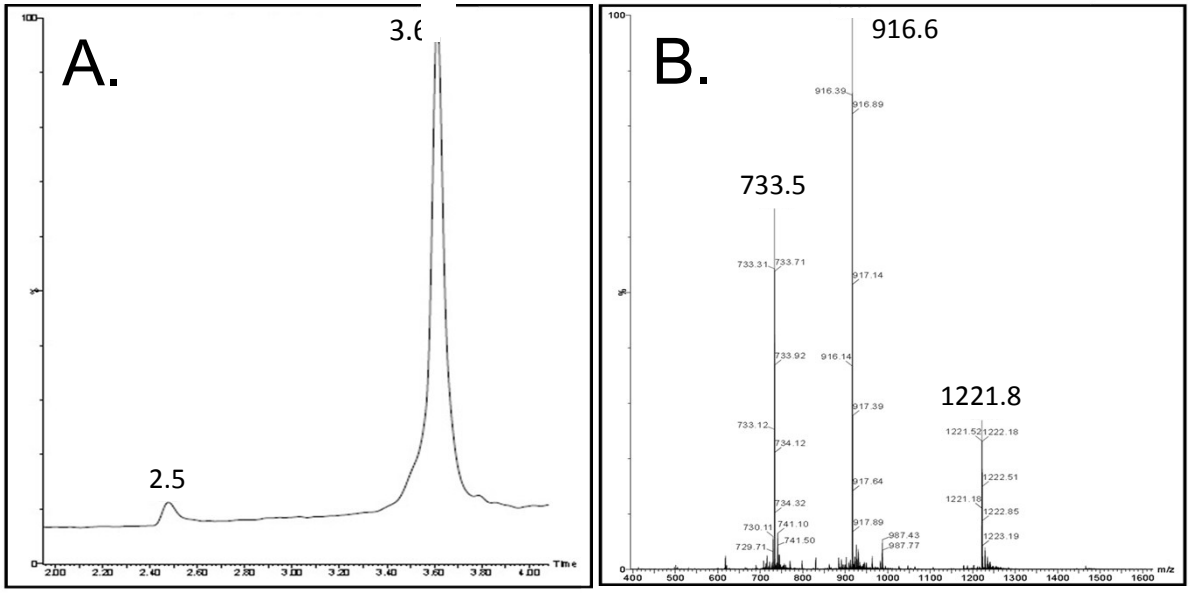

Figure S1: ESI-MS of **Peptide 2**: (A) UPLC elution profile showing a single major peak at 3.6 minutes and (B) ionization peaks for corresponding species yields a peptide molecular weight of 3662.55 Da.

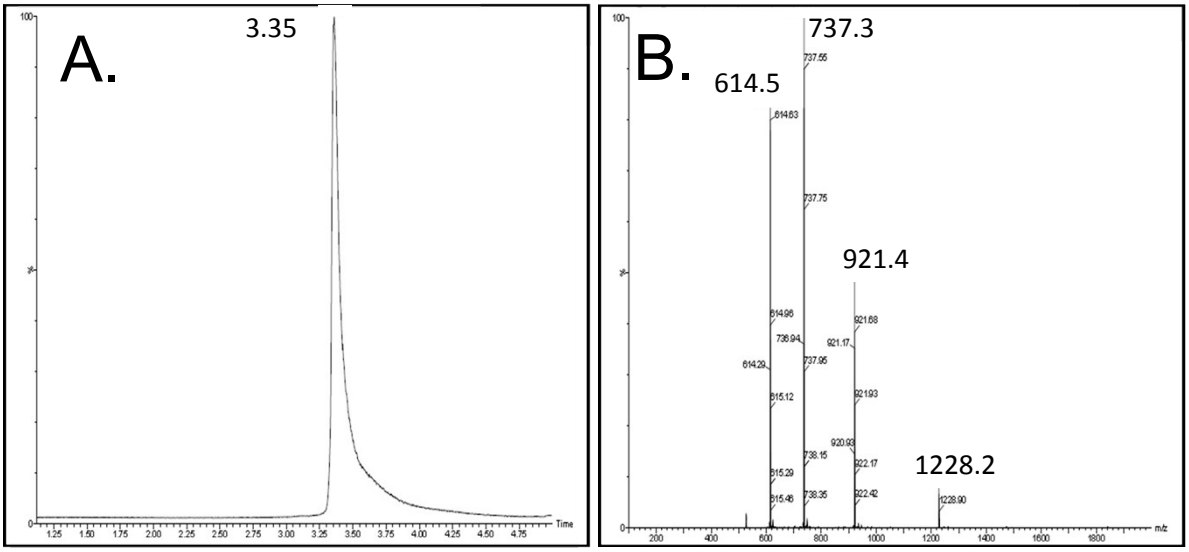

Figure S2: ESI-MS of **Peptide 1**: (A) Elution profile showing a single major peak at 3.35 minutes and (B) ionization peaks for corresponding species yields a peptide molecular weight of 3681.44 Da.

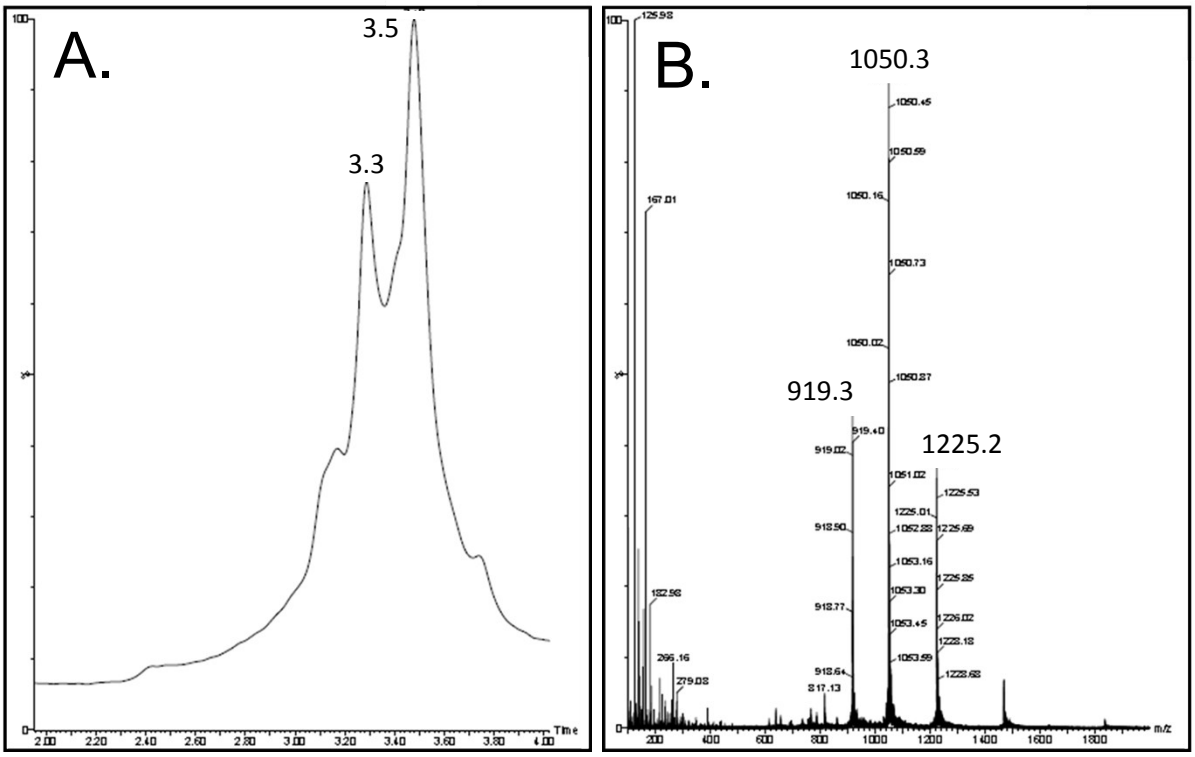

Figure S3: ESI-MS after reaction: (A) Elution profile with peak at 3.3 minutes corresponding to **Peptide 1** (position and molecular weight same as in **Figure S2**) and major peak at 3.5 minutes, (B) ionization peaks in major peak confirming formation of **Peptide 1-2** adduct. Measured molecular weight is 7346.14 Da. Elution peak for **Peptide 2** is not present in (A) indicating that the reaction is complete.

TEM image analyses using ImageJ software

ImageJ analyses of negatively stained TEM micrographs (refer to main text **Figure 5** and **SI Figure S4**) yielded the average length of the rods in each sample.[4] Specifically, the lengths of 50 or more rods were measured using ImageJ's Measure tool, and statistics of average length and standard deviation were calculated from these measurements (see **Table S1**). The average lengths agree well with the target lengths predicted by step-growth length statistics (eq. s1). Overall, we can conclude that a) the average rod length increased as the stoichiometric ratio *r* of the reactants approached unity, and b) the rods formed after reaction is complete are polydisperse, which is characteristic of polymers synthesized via step-growth polymerization as discussed earlier. Using the measured length of rods, the critical concentration

$C^*=1/L^3$ for semi-dilute regime for each sample has been calculated.[5] It is important to note here that these values are only a rough estimate for the upper bound of the critical concentration. The true C* may be much lower for each sample due to the polydisperse and polyelectrolyte nature of the assembled rigid rods.[6]

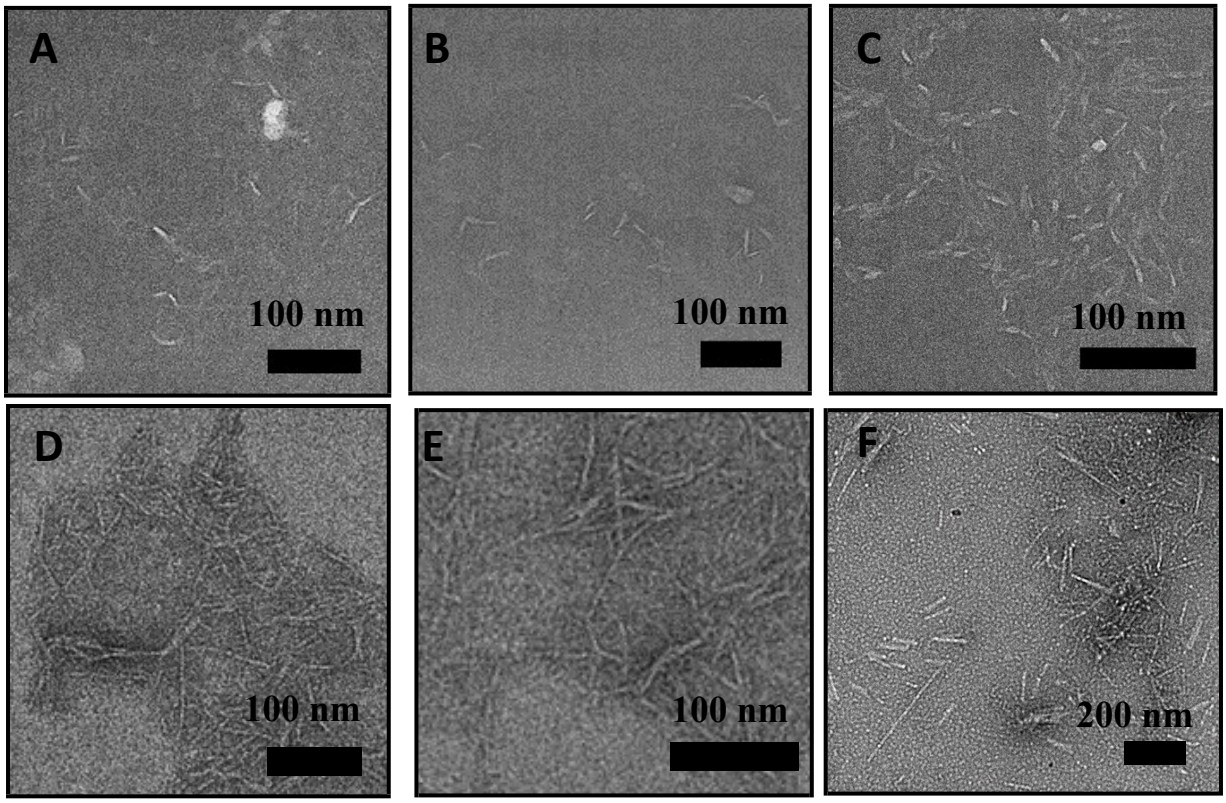

Figure S4: Negatively stained cast-film TEM images of rigid rod samples having an average length distribution of short (**A**, **B**, **C**), medium (**D**, **E**) and long (**F**) rigid rods.

Table S1: Summary of ImageJ length analysis of TEM images for rod samples of different target lengths constructed by varying the stoichiometry of the reacting bundles. The number in parentheses indicates the number of rigid rods used to calculate the length statistics.

Sample	**Ratio r (Thiol:Mal)**	**Target Length L_T (Å)**	**Measured Length L (Å)**	**Critical concentration C* (w/v%)**
Short	0.50	105	145 ± 34 (52)	2.3
Medium	0.88	525	528 ± 149 *(55)*	0.2
Long	0.94	1085	1099 ± 706 *(176)*	0.05

SANS modeling and fit results

Table S2: SASview fit results for cylinder form factor model to scaled SANS data shown in **Figure 4 (A)** of main text.[7] Both samples i.e. **Peptide 2** and ultra-long rigid rods contain peptide mass by volume fraction concentration of 0.1 % in D_2O. The fit range was fixed to *0.01 Å^{-1} <Q<0.3 Å^{-1}*.

Sample	Radius R (Å)	Length L_{cyl} (Å)	Goodness of fit (χ^2/N)
Bundlemer	8.0 ± 0.5	33 ± 2	0.8
Rod	9.2 ± 0.3	3100 ± 776	1.1

In the **Table S3**, fit values and standard deviation (±1σ) for model parameters are summarized. SANS data were fit using either **Model 1** or **Model 2** as indicated. Here *C* (w%) is the mass by volume fraction or the weight percent concentration of peptide rods in solution. In case of SANS data fits for *C* < 0.5 %, as also for added-salt and acidic-pH cases, the fractal dimension *D* values may not be a true estimate since there are not enough points in the low-*Q* upturn regime for this dataset (larger low-*Q* cutoff of 0.004 Å^{-1} since measurements were performed at 13 m sample-to-detector distance without the use of neutron lenses due to poor counting statistics). Also, **Model 1** was not used to fit SANS data for the acidic-pH case since the linear charge density at this pH was not extractable from the data. The mean of the standard deviation ($<\mu>_\sigma$) of each fit parameter is calculated here along with the sample mean ($<\mu>$) and standard error (σ_e) of sample means. These are reported in the last two rows of each table for relevant fit parameters. The goodness of fit was evaluated via the reduced chi-squared value (χ^2/N) where N is the degrees of freedom for the dataset. A reduced chi-squared value of 1 indicates a perfect fit whereas a value greater than 1.2 indicates an increasingly poorer fit (significance level of 0.05).

Table S3: **(A)** Fit values and corresponding standard deviations for each concentration within a dilution series in pure D_2O. The SANS data was fit using **Model 1,** shown in **Figure 5** of the main article. The values highlighted in **bold** were not used in calculating the corresponding sample mean ($<\mu>$) and standard error (σ_e). Specifically, for calculating sample mean of linear charge density μ_L, the lowest concentration sample was omitted since a correlation hole was not present. For fractal dimension *D*, measurements with the larger low-*Q* cutoff were omitted.

(A) Model 1

Name	C	R	σ_R	L_a	σ_L	μ_L	$\sigma\mu$	D	σ_D	χ^2/N

	w %	Å	Å	Å	Å	e/nm	e/nm			
Short (S)										
1	0.5	9.0	0.5	69.8	2.7	**1.18**	**0.36**	2.58	0.07	1.4
2	0.8	9.0	0.2	78.0	2.1	0.68	0.14	2.59	0.05	1.4
3	2.0	9.8	0.1	71.0	0.9	0.76	0.10	2.60	0.02	2.8
Medium (M)										
1	0.1	10.5	0.6	139.6	17.4	**0.75**	**0.59**	**2.17**	**0.25**	2.3
2	0.3	9.5	0.2	170.5	8.4	0.85	0.12	**2.30**	**0.19**	2.3
3	0.7	10.6	0.1	137.9	2.6	0.85	0.03	2.29	0.06	4.1
4	1.5	11.0	0.1	117.7	1.6	0.85	0.03	2.39	0.05	2.5
5	3.0	10.6	0.1	106.4	1.5	0.69	0.02	2.24	0.04	4.3
6	6.0	10.5	0.1	93.0	1.3	0.65	0.01	2.23	0.03	12.6
Long (L)										
1	0.1	12.0	0.3	112.6	9.9	**0.60**	**0.41**	**2.08**	**0.15**	1.7
2	0.2	12.5	0.3	115.1	9.9	**0.50**	**0.52**	**2.08**	**0.16**	1.7
3	0.7	12.1	0.1	108.8	2.8	0.72	0.06	2.47	0.06	2.6
4	1.4	10.8	0.1	103.8	1.4	0.66	0.02	2.49	0.04	2.0
5	2.5	11.9	0.1	88.1	0.9	0.86	0.02	2.45	0.03	2.8
6	4.5	11.9	0.1	76.2	0.8	0.80	0.02	2.51	0.02	4.9
7	9.0	11.7	0.1	59.9	0.6	0.74	0.02	2.56	0.04	15.4
<μ>		**10.8**	**0.2**			**0.76**	**0.05**			
σ_e		**1.1**				**0.08**				

Table S3: **(B)** Fit values and corresponding standard deviations for dilution series in pure D_2O modeled using **Model 2,** shown in **Figure 5** of the main article. In case of medium and long rods, the fit parameters were evaluated only for the low concentrations. The poor fits indicated by large reduced chi-squared values as concentrations increase are highlighted here in **bold**. Note that both **Model 1** and **Model 2** give comparable values for the radius of rigid rods.

(B) Model 2

Name	**C**	**R**	**σ_R**	**L_a**	**σ_L**	**Ɗ**	**σ_D**	**χ^2/N**
	w %	Å	Å	Å	Å			
Short (S)								
1	0.5	8.3	0.4	75.5	1.6	2.58	0.05	1.1
2	0.8	8.5	0.2	81.6	2.0	2.64	0.03	1.0
3	2.0	9.2	0.1	84.2	1.4	2.61	0.04	15.6
Medium (M)								
1	0.1	9.0	0.4	87.5	16.4	1.69	0.13	1.6
2	0.3	8.2	0.1	103.2	7.8	1.58	0.06	6.5
3	0.7	9.5	0.1	136.5	11.4	2.50	0.20	**46.5**

Long (L)								
1	0.1	10.0	0.4	136.5	8.3	2.15	0.12	1.9
2	0.2	10.0	0.2	134.3	3.3	2.26	0.06	1.6
3	0.7	10.8	0.1	114.5	2.7	2.48	0.06	8.0
4	1.4	10.0	0.1	112.8	4.0	2.48	0.08	**27.6**
<μ>		**10.2**	**0.2**					
σ_e		**0.3**						

Table S3: **(C)** Fit results for dilution series of long rods in the presence of 200 x 10^{-3} M sodium chloride salt using **Model 1**, shown in **Figure 6(B)** of the main article. The linear charge density μ_L was set to 0.08 e/Å.

(C) Model 1 added-salt case

Name	**C**	**R**	**σ_R**	**L_a**	**σ_L**	**Ɗ**	**σ_D**	**χ^2/N**
	w %	Å	Å	Å	Å			
Long (L)								
2	0.2	11.0	0.3	72.3	3.3	2.06	0.05	1.5
3	0.7	10.4	0.1	103.5	3.7	1.90	0.05	3.2
4	1.4	10.5	0.1	93.9	3.4	1.92	0.04	3.8
5	2.5	10.4	0.1	95.1	1.5	2.15	0.03	2.5
<μ>		**10.6**	**0.2**			**2.0**	**0.04**	
σ_e		**0.2**				**0.10**		

Table S3: **(D)** Fit results for dilution series of long rods in the presence of 200 x 10^{-3} M sodium chloride salt using **Model 2** shown in **Figure 6(B)** of the main article.

(D) Model 2 added-salt case

Name	**C**	**R**	**σ_R**	**L_a**	**σ_L**	**Ɗ**	**σ_D**	**χ^2/N**
	w %	Å	Å	Å	Å			
Long (L)								
2	0.2	12.63	0.6	102.5	2.5	2.29	0.05	3.7
3	0.7	10.6	0.1	115.7	1.8	2.27	0.03	2.2
4	1.4	11.3	0.1	104.9	1.6	2.14	0.03	2.6
5	2.5	9.8	0.1	101.6	1.1	2.26	0.02	2.7
<μ>		**11.1**	**0.2**			**2.24**	**0.03**	
σ_e		**1.0**				**0.06**		

Table S3: **(E)** SANS fit results for dilution series of long rods in the presence of deuterium chloride pH 1.0 using **Model 2** shown in **Figure 6(C)** of the main article.

(E) Model 2 acidic pH case

Name	C	R	σ_R	L_a	σ_L	Ɗ	$\sigma_Ɗ$	χ^2/N
	w %	Å	Å	Å	Å			
Long (L)								
2	0.2	13.5	0.6	117.3	4.1	2.16	0.08	4.5
3	0.7	11.2	0.1	110.9	1.8	2.14	0.03	2.3
4	1.4	9.7	0.1	111.4	1.3	2.27	0.02	2.7
5	2.5	11.8	0.1	95.4	1.0	2.27	0.02	4.4
<μ>		**11.6**	**0.2**			**2.21**	**0.04**	
σ_e		**1.4**				**0.06**		

Calculation of uncertainty in effective diameter D_{eff}

For the salt-free low ionic strength solution, the effective diameter is given by:

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

Where: $A = 2\pi(\mu_L)^2\lambda_D\lambda_B \exp\left(-2R.\lambda_D^{-1}\right)$ and $\lambda_B = \frac{e^2}{4\pi\epsilon\epsilon_0 k_B T} \sim 0.71\ nm$

In low ionic strength water, the Debye length is approximated by $\lambda_D = \left[\frac{8\pi N_A e^2}{k_B T\epsilon\epsilon_0}\right]^{-0.5}$ = 2.72 nm

In 200 mM monovalent electrolyte solution (added salt and DCl case), $\lambda_D = \left[\frac{N_A e^2 \sum_{i,j} c_i z_i^2}{k_B T\epsilon\epsilon_0}\right]^{-0.5}$ = 0.69 nm

The uncertainty in D_{eff}, i.e. ΔD_{eff} was calculated using the following expression:

$$\Delta D_{eff} = \left(\frac{\partial D_{eff}}{\partial R}\right)\Delta R + \left(\frac{\partial D_{eff}}{\partial \mu_L}\right)\Delta\mu_L = (0)\Delta R + \left(\frac{2\lambda_D}{\mu_L}\right)\Delta\mu_L = 2\lambda_D.\left(\frac{\Delta\mu_L}{\mu_L}\right)$$

Here, Δ is the standard error in sample mean of corresponding fit parameters. Therefore, $\mu_L \pm \Delta\mu_L$ obtained via **Model 1** fits were used to calculate the uncertainty value ΔD_{eff}.

While using **Model 2**, the rigid rods are assumed to be neutral due to screened charges in solution. For this case, the effective diameter is simply $D_{eff} = D = 2R$, therefore, the uncertainty in its value:

$$\Delta D_{eff} = \left(\frac{\partial D_{eff}}{\partial R}\right)\Delta R = 2.\Delta R$$

These values are summarized in **Table 1** and **Table 2** of the main text.

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. 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.
2. D. P. Nair, M. Podgórski, S. Chatani, T. Gong, W. Xi, C. R. Fenoli and C. N. Bowman, *Chemistry of Materials*, 2014, **26**, 724-744.
3. P. J. Flory, *Journal of the American Chemical Society*, 1936, **58**, 1877-1885.
4. C. A. Schneider, W. S. Rasband and K. W. Eliceiri, *Nature Methods*, 2012, **9**, 671.
5. J. Schneider, W. Hess and R. Klein, *Journal of Physics A: Mathematical and General*, 1985, **18**, 1221.
6. M. Muthukumar, *Macromolecules*, 2017, **50**, 9528-9560.
7. http://www.sasview.org/.