Electronic Supplementary Material (ESI) for Soft Matter.

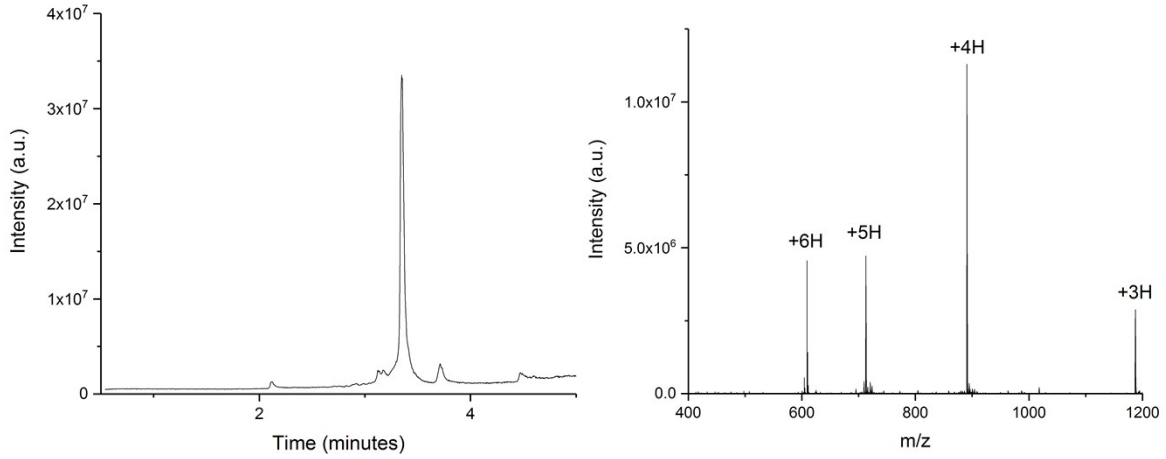
Self-assembly and soluble aggregate behavior of computationally designed coiled-coil peptide bundles

Michael J. Haider,[a] Huixi Violet Zhang,[b] Nairiti Sinha,[a] Jeffrey A. Fagan,[c] Kristi L. Kiick,*[a] Jeffery G. Saven*[b] and Darrin J. Pochan*[a]

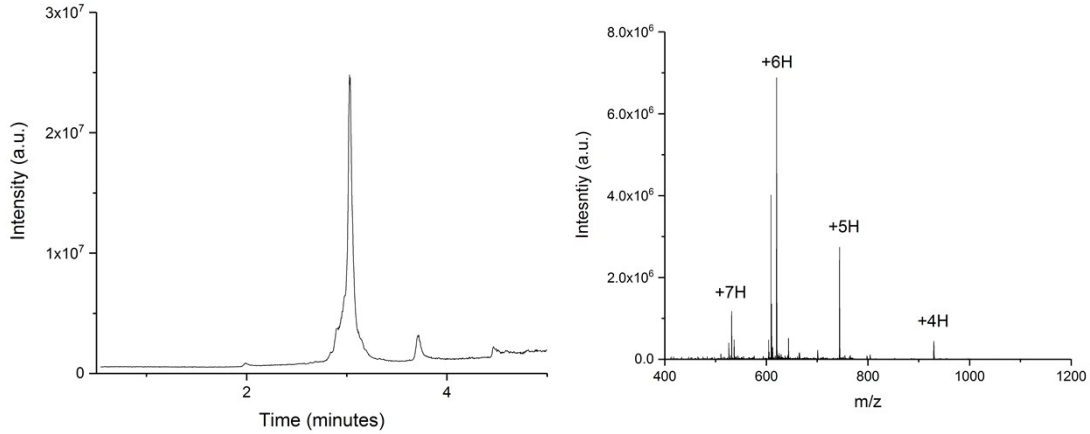
Department of Materials Science and Engineering, University of Delaware, Newark, Delaware 19716, USA. E-mail: kiick@udel.edu, pochan@udel.edu

Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. E-mail: saven@sas.upenn.edu

Materials Science and Engineering Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA. E-mail: jeffrey.fagan@nist.gov



S1: TOP: UPLC-MS data for purified BNDL1. Bottom: BNDL1 designed mass is 3560 g/mol, purified mass from ESI is 3559.2 g/mol.

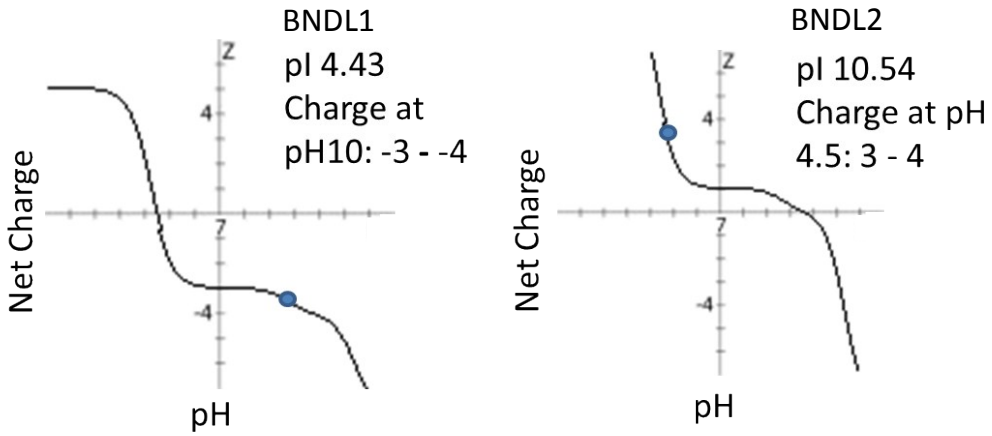


S2: TOP: UPLC-MS data for purified BNDL2. Bottom: BNDL2 designed mass is 3715 g/mol, purified mass from ESI is 3714.8 g/mol.

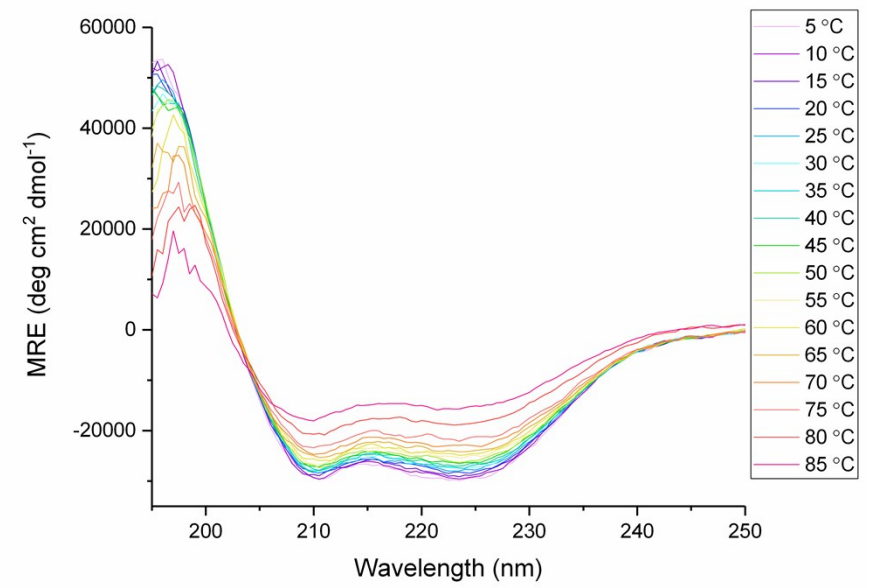
The charge vs pH curves below were calculated at www.pepcalc.com using the equation:

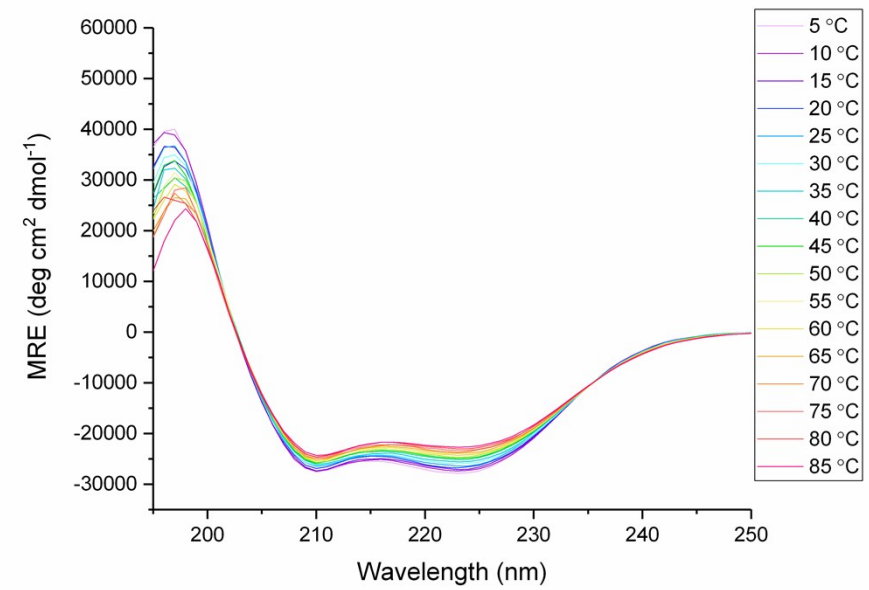
$$Z = \sum_i N_i \frac{10^{pKa_i}}{10^{pH} + 10^{pKa_i}} - \sum_j N_j \frac{10^{pH}}{10^{pH} + 10^{pKa_j}}$$

Where i components are from amine functionalities and j components are from acid functionalities.

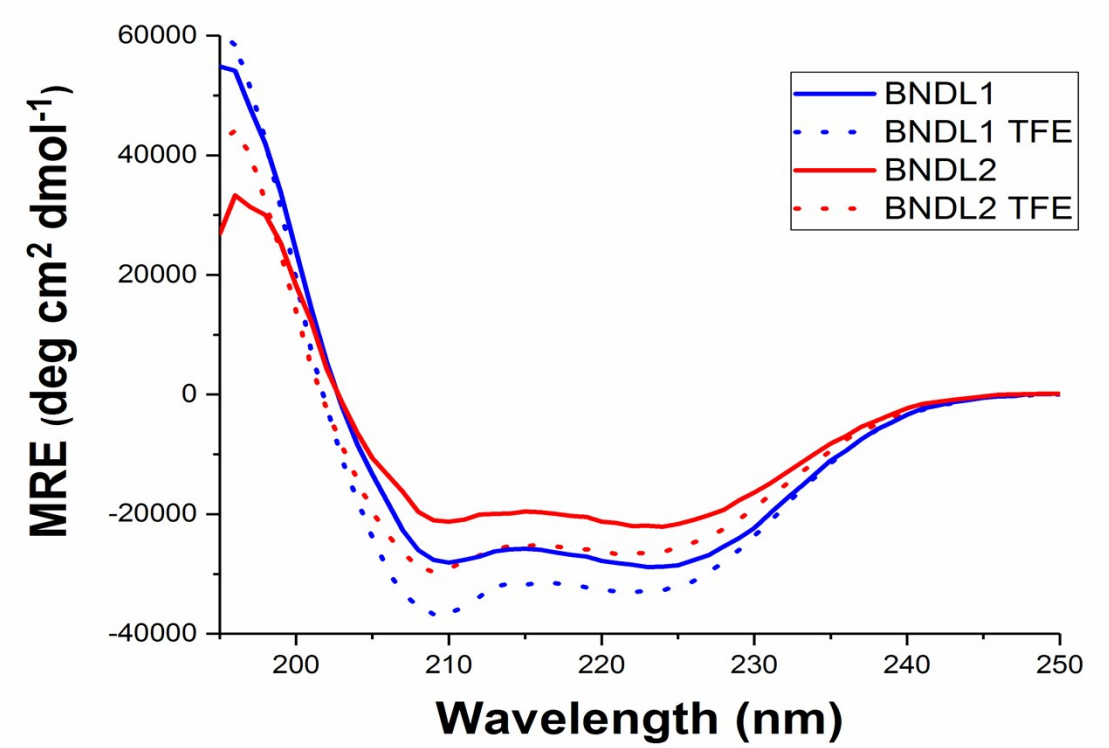


S3: Isoelectric point curves of BNDL1 and BNDL2.

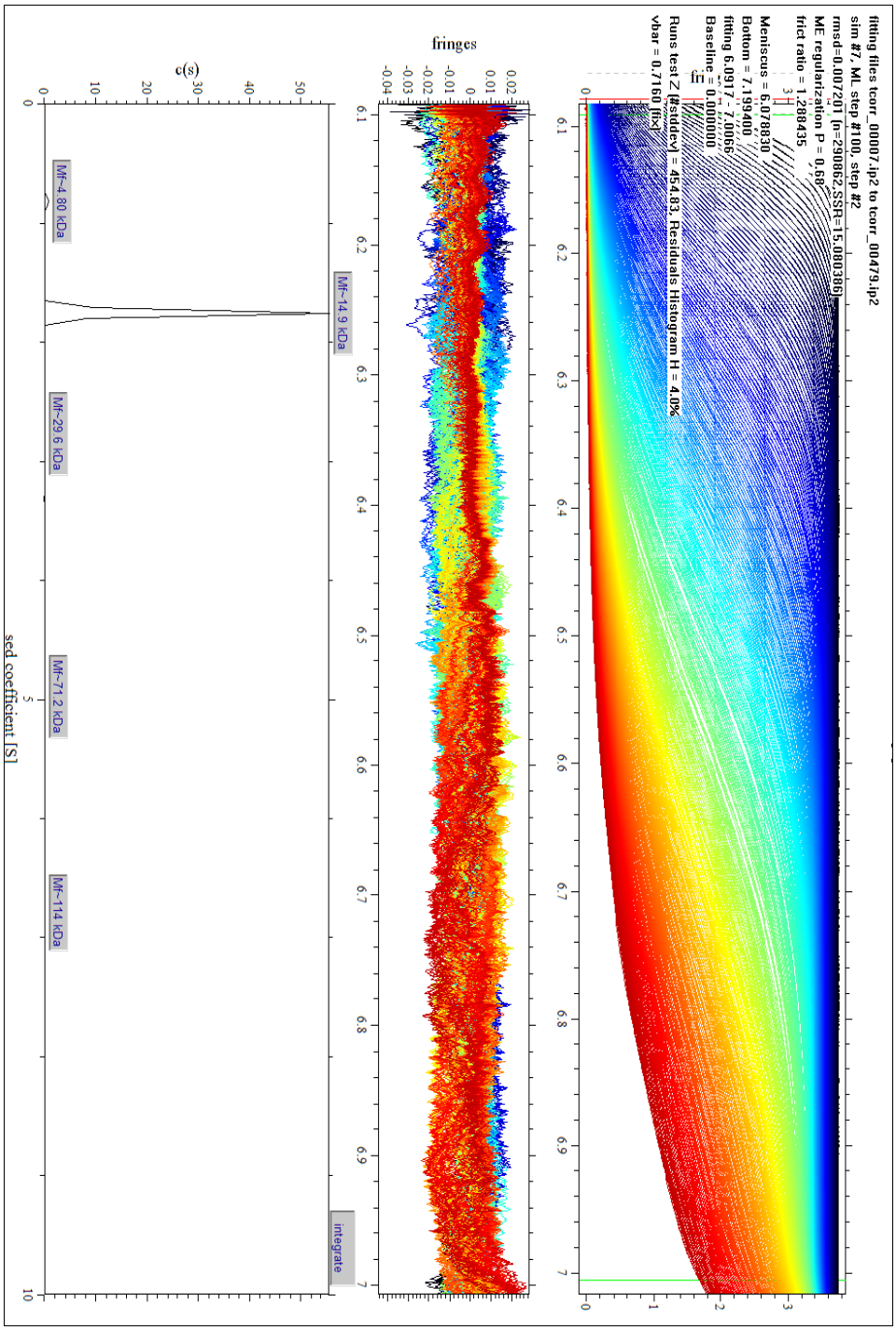




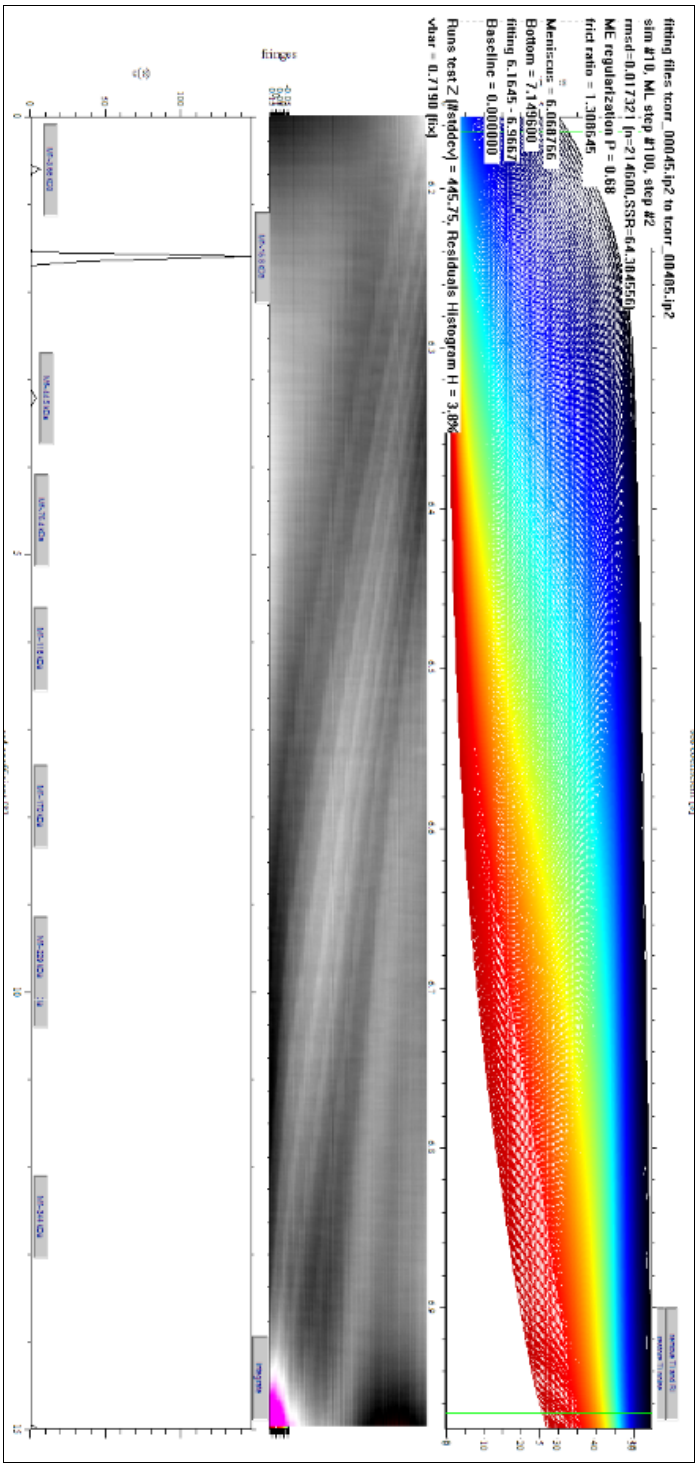
S4: Full CD curves for heating from 5 °C to 85 °C for BNDL1 (Top) and BNDL2 (bottom). Concentrations are 0.1 mM peptide in 50 mM pH 10 borate buffer and 50 mM pH 4.5 sodium acetate buffer for BNDL1 and BNDL2 respectively.



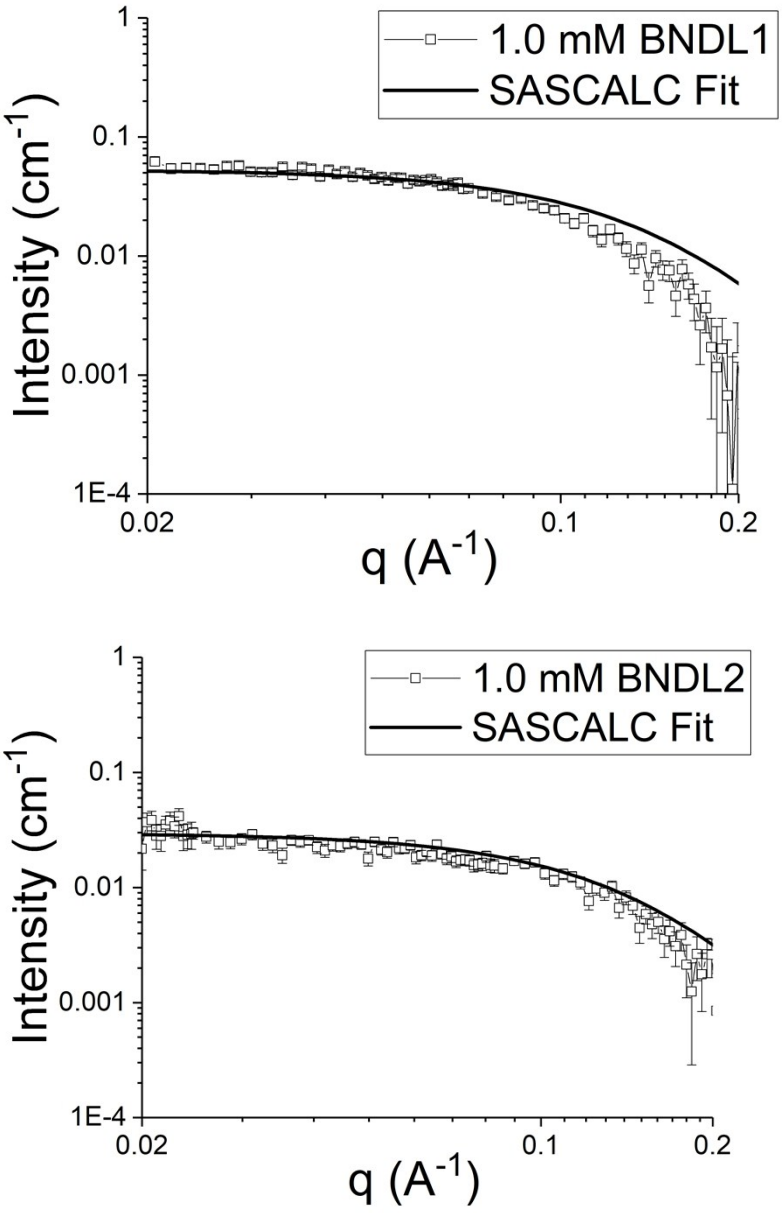
S5: CD of 0.1 mM peptide in solutions prepared from 50% TFE and 50% 50 mM pH 10 borate buffer for BNDL1 and 50% TFE and 50% 50 mM pH 4.5 sodium acetate buffer for BNDL2.



S6: AUC Data for BNDL1 (top) and BNDL2 (bottom) strongly suggest that the tetramer is the dominant species in solution. Low intensity peaks are observed at larger masses indicating the presence of soluble aggregates.



S6: AUC Data for BNDL1 (top) and BNDL2 (bottom) strongly suggest that the tetramer is the dominant species in solution. Low intensity peaks are observed at larger masses indicating the presence of soluble aggregates.

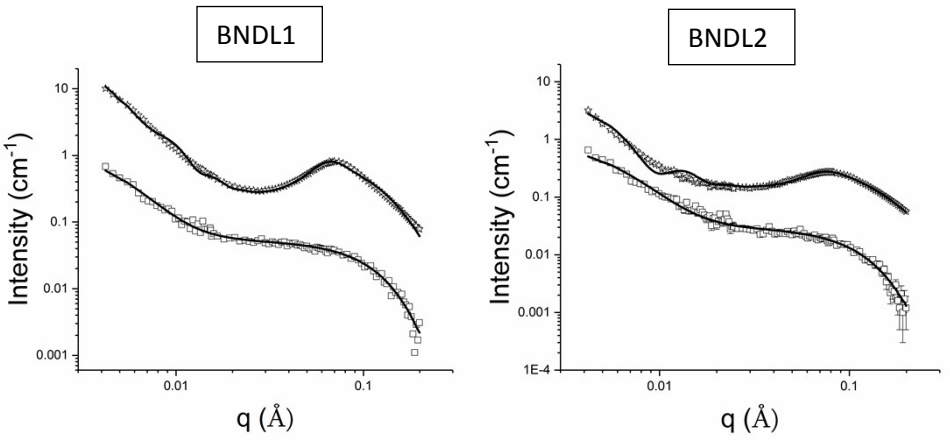


S7: BNDL1 (top) and BNDL2 (bottom) at 1.0 mM (squares) with SASCALC fit (solid line).

BNDL1											
C	R	dR	varR	L	dL	varL	N	D	sigD	R^2	X^2
1	14.1	0.423	0.1789	40.4	1.945	3.783	1	N/A	N/A	0.987	1.282
2.5	12.4	0.142	0.0202	40.9	0.693	0.480	1	N/A	N/A	0.986	3.706
5	12.3	0.060	0.0036	45.9	0.497	0.247	69	47.1	4.32	0.997	3.992
10	12.7	0.032	0.0010	46.1	0.273	0.075	76	35.2	4.73	0.996	14.286
15	12.8	0.024	0.0006	45.4	0.191	0.036	79	30.3	12.46	0.995	44.067
Average	12.9	0.202		43.7	0.961						

BNDL2											
C	R	dR	varR	L	dL	varL	N	D	sigD	R^2	X^2
1	13.8	0.514	0.2642	42.0	2.475	6.126	1	N/A	N/A	0.972	1.55
2.5	10.8	0.169	0.0286	44.1	0.868	0.753	1	N/A	N/A	0.965	2.661
5	10.7	0.087	0.0076	45.1	0.771	0.594	60	43.7	10.50	0.99	1.193
10	10.2	0.052	0.0027	45.3	0.524	0.275	62	34.3	8.91	0.994	1.601
15	9.7	0.044	0.0019	42.5	0.391	0.153	60	30.4	10.52	0.979	9.949
Average	11.0	0.247		43.8	1.257						

ST1: Fit results for SANS concentration series of BNDL1 and BNDL2. Fit parameters include radius of cylinder (R, units Å), length of cylinder(L, units Å), number of bundles per aggregate (N), distance between bundles within aggregates (D, units Å), and the standard deviation of D (sigD, units Å). The values dR and dL are the standard deviation errors calculated by python while curve fitting. The standard deviation in the average values of R and L was calculated by taking the square root of the average of the variances. R^2 is the coefficient of correlation and X^2 is the goodness of fit.



S8: BNDL1 (left) and BNDL2 (right) at 1.0 mM (squares) and 15.0 mM (stars) fit for the whole q range to reveal low-q scattering intensity. The fit of the low-q data approximates the characteristic distance between soluble aggregates/concentration fluctuations at larger lengthscales.

BNDL1						
C	A	dA	E	dE	R^2	X^2
1	1.53	0.199	195	10	0.987	1.28
15	29.88	0.766	227	2	0.995	44.07

BNDL2						
C	A	dA	E	dE	R^2	X^2
1	0.96	0.07	154	5	0.972	1.55
15	1.6	0.256	199	12	0.979	9.95

ST2: Fit results for 1 mM and 15 mM BNDL1 and BNDL2 for the full q range. Fit parameters include A, weighting factor, and E, the distance between aggregates in units of Å.

SANS Fitting Procedure

In static SANS experiments, if a sample has an isotropic distribution of monodisperse scatterers with no orientation preference, the scattered intensity I is a function of the scattering vector Q, given by the expression:

$$I(Q) = NV^2\Delta\rho^2 P(Q)S(Q)\ldots(1)$$

Thus, the number density of scatterers N, the volume V and difference between the Scattering Length Densities (SLD) of the scatterer and solvent, $\Delta\rho$ contribute to the total scattering at scattering vector Q equal to 0. P(Q), the form factor, depends on the size and shape of individual scatterers, while the interparticle interactions are described by the structure factor S(Q).

Theoretical form factor based scattering profiles for each designed homotetrameric bundle was calculated from the atom coordinates (PDB file) using SasCalc, a free analysis tool available on SASIEweb [2]. SasCalc employs a “golden vector” approximation i.e. orientations of scattering vector Q are generated on a quasi-uniform spherical grid using the golden number, which in-turn enables speedy calculations of a rotationally averaged scattering profile using the exact coordinates of all atoms in the scattering entity [1].

A python script that employs a non-linear least squares algorithm based on Scipy’s Optimize package was used to fit the scattering expression (1) to the experimental SANS data in the Q-range of 0.02 $Å^{-1}$ to 0.2 $Å^{-1}$. The form factor P(Q) was calculated based on a cylindrical model, described by the expression:

$$P(Q) = \int_0^{\pi/2} f^2(Q,\alpha)\sin\alpha d\alpha$$

$$f(Q) = 2 * \frac{sin(Q\ L\ cos\alpha/2)}{Q\ L\ cos\alpha/2} * \frac{J_1(Q\ R\sin\alpha)}{(Q\ R\sin\alpha)}$$

The radius of the cylinder R and the length of the cylinder L were varied within 20% of the average values predicted from the cylindrical form factor fits by SasView. The structure factor was modeled using an expression for interacting cylinders obeying a Gaussian distribution as derived by Kratky and Porod. [3]

$$S(q) = 1 + \frac{2}{N}\left[\sum_{k=1}^{N}(N-k)\cos(kDq)\exp\left(\frac{-kq^2\sigma_D^2}{2}\right)\right]$$

The corresponding fit parameters were: the number of bundles in an aggregate, N; the distance between bundles within the aggregate, D and; the Gaussian error in the distance distribution within the aggregate, σ_D. The reduced chi-squared value χ^2_R was calculated to estimate the goodness of fit, where χ^2_R value of 1 indicates a perfect fit.

$$\chi^2_R = \frac{1}{N-n}\sum_{i=1}^{N}\left(\frac{I_{fit}(Q_i) - I_i}{s_i^2}\right)^2$$

Where $I_{fit}(Q_i)$ is the analytical expression (1) with n fitted parameters, I_i is the ith scattering data point and s_i^2 is the corresponding variance in I_i.[4]

The up-turn at lower Q values has been previously shown to fit an empirical expression [5]

$$I_{total}(Q) = I(Q) + \frac{A}{\left(1 + q^2E^2\right)^2}$$

Here, E corresponds to a long range correlation distance between loose clusters of peptide aggregates and A is a relative weighting factor. $I_{total}(Q)$ was fit with A and E as fit variables using a similar non-linear least-squares algorithm, wherein the expression (1) with previously fitted parameters R, L, N, D and σ_D was used to analytically calculate $I(Q)$. The goodness of fit was estimated by reduced chi-squared values.

Reference(s) and Citations

1. Rapid and accurate calculation of small-angle scattering profiles using the golden ratio M.C. Watson, J.E. Curtis, J. Appl. Crystallogr. 46, 1171-1177 (2013)).
2. SASSIE: A program to study intrinsically disordered biological molecules and macromolecular ensembles using experimental scattering restraints J. E. Curtis, S. Raghunandan, H. Nanda, S. Krueger, Comp. Phys. Comm. 183, 382-389 (2012).

3. Diffuse small-angle scattering of X-rays in colloid systems, O. Kratky and G. Porod, J. Colloid Science 4, 35–70 (1949).
4. *Statistical methods in experimental physics*. James, F. World Scientific Publishing Co. Inc. (2006)
5. SANS from poly(ethylene oxide)/water systems, B. Hammouda, D.L. Ho, and S. Kline, Macromolecules,35, 8578–8585 (2002).