SUPPORTING INFORMATION

Self-Assembly Pathway of Peptide Nanotubes Formed from a Glutamatic Acid-Based Bolaamphiphile

Emerson R. Silva,*[a,b] Wendel A. Alves, Valeria Castelletto,[b] Mehedi Reza,[c] Janne Ruokolainen,[c] Rohanah Hussain[d] and [a] Ian W. Hamley *[b]

[a]*Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André 09210-580, Brazil*
[b]*Department of Chemistry, University of Reading, Whiteknights, Berkshire RG6 6AD, UK*
[c]*Dept of Applied Physics, Aalto University School of Science, P. O. Box 15100, FI- 00076, Finland*
[d]*Diamond Light Source Ltd, Harwell Science and Innovation campus, Didcot, Oxfordshire OX11 0DE, UK*

ACKNOWLEDGEMENTS: E.R.S. and W.A.A. thank FAPESP for grants (postdoctoral fellowship abroad BEPE 2014/03514-8 and regular support 2013/12997-0). Work at Reading was supported by EPSRC Platform Grant "Nanostructured Polymeric Materials for Healthcare" EP/L020599/1. I.W.H. is the recipient of a Royal Society-Wolfson Research Merit Award. Mr Nick Spencer and Dr. Askhan Deshsorkhi, at University of Reading, are acknowledged for help with XRD experiments. This work involved use of synchrotron facilities at ESRF, proposal MX1666, and Diamond Light Source, proposal SM10083-1. The authors thank Dr. Adam Round (ESRF-BM29) and Dr. Tamas Javorfi (DLS-B23) for help during SAXS and CD experiments. SAXS data were also collected on I911-4 beamline at MaxLab II (Lund, Sweeden) under proposal number 20140431. Dr. Ana Labrador and Dr. Christopher Söderberg are kindly acknowledged for help during data collection in Lund.

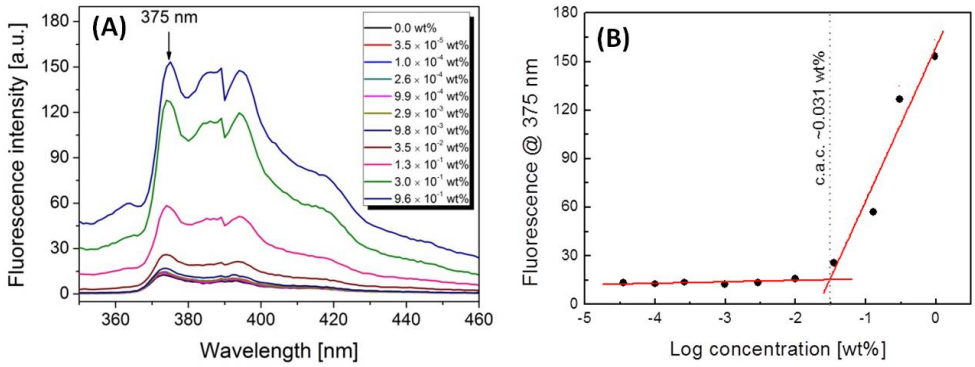
MATERIALS AND METHODS

Peptides and samples preparation: the EFL_4FE bolaamphiphilic peptide was custom-synthesized by CS-Bio (Merlo Park, California) using TFA couterion (batch no. CS12076). Purity was assessed by an Agilent 1200 HPLC system using 0.1 % TFA in H2O (buffer A) and 0.1 TFA in acetonitrile (buffer B), at a flow rate of 1ml/min. through a Phenomenex Luna column C18. Purity was confirmed at 96.70%, with gradient of 25 to 55 % B in 20 min. The mass was found to be Mw = 1022.33 Da (expected 1022.20 Da) by electrospray-mass spectrometry. The dry powder was kept under refrigeration at -20°C, being weighted and dissolved into 20 mM NaOH (Sigma-Aldrich, MO) solutions. The diluent was ultra-pure water from a Barnsted Nanopure system or D_2O (Sigma-Aldrich). Ultrasonication during ca. 20 minutes at 50 °C was performed to ensure solubilization and solutions were kept at room temperature prior to further assays. Control samples were prepared by dissolving peptides into 1:1 NaOH:HCl and H2O:ETOH solutions. The sample was first dissolved into NaOH as above and the equivalent amount of HCl was added immediately after solubilization. **Determination of critical aggregation concentration:** Peptide powder was dissolved into 20 mM NaOH solutions containing pyrene

at 1×10^{-5} wt%. After ca. 24 hours of incubation at RT, **fluorescence assays** were performed in a Cary Varian Ellipse spectrometer. Volumes of 0.75 ml were transferred into a 1 cm path length Hellma suprasil cuvette. Excitation was tuned at λ_{exc} = 338 nm and slits were adjusted to 5 × 5 nm. Emission was registered in the range 340-460 nm and intensity at 375 nm was plotted as function of peptide concentration. **Cryo-TEM imaging:** Cryo-TEM imaging was performed as reported elsewhere.[1] In brief, the instrument was a JEOL JEM-3200FSC operating at 300 kV, used in bright-field mode and zero-loss energy filtering with a slit width 20 eV. Images were obtained with a Gatan Ultrascan 4000 CCD camera, from vitrified specimens prepared using a FEI Vitrobot device on Quantifoil 3.5/1 carbon copper grids. Grids were cleaned using a plasma cleaner and then transferred into the environmental chamber of a FEI Vitrobot at RT and 100% humidity. An amount 3 μl of solution was applied on the grid, blotted once for 1 s and then vitrified in a 1:1 mixture of liquid ethane and propane at −180 °C. Grids with vitrified samples were maintained in a liquid nitrogen atmosphere and then cryo-transferred into the microscope. **XRD diffraction:** diffraction data were obtained from oriented stalks prepared by suspending droplets of a 5 wt% peptide solution between the tips of wax-coated capillaries and letting them dry in air at RT. Stalks were vertically positioned onto a RAXIS IV++X-ray diffractometer (Rigaku) with a rotating anode generator. Images were recorded using a Saturn 992 CCD camera, at sample-to-detector distance of 40 mm. Data reduction was performed with Fit2D (ESRF) and unit cell determinations were made with the software CLEARER.[2] **Small-Angle X-ray Scattering:** SAXS experiments were conducted on the bioSAXS beamline BM29 at the ESRF (Grenoble, France).[3] Sample preparation started 10 days prior to the beamtime in order to provide formulations at different aging times (storage at RT). Solutions were dropped into wells of a PCR plate and kept at T = 20 °C using the refrigeration system available on the sample holder. Volumes of 50 μL were injected by an automated device into a 1 mm quartz capillary, horizontally positioned in front of the beam. Cleaning steps, using water and surfactant solutions, were performed between every sample or buffer measurement, followed by drying with nitrogen stream. Sample solutions were flowed during 6 seconds during data acquisition to avoid radiation damage and 30 frames of 0.2 s each were registered per run. Buffer measurements were performed before and after each sample, then averaged prior to subtraction. The X-ray wavelength was λ = 0.91 Å and sample-to-detector distance was 2,864 mm, providing data in the interval 0.04 $nm^{-1} \leq q \leq 4.9\ nm^{-1}$ (where $q = 4\pi/\lambda \times \sin\theta$, with 2θ the scattering angle). The detector was a Pilatus 1M device. Data were radially-averaged, normalized and background subtracted using the beamline software and model fitting was carried out with SASFit program[4] (model details below). Complementary data were recorded on I911-4 beamline at MaxLab II (Lund, Sweeden). The experimental conditions were similar to those described above, but this time samples were oscillated into the capillary. Four frames, 30 s each, were acquired and no radiation damage was detected. **Secondary structure determinations: Fourier transform infra-red (FTIR)** experiments were carried out using a Nicolet Nexus spectrometer. Solutions were prepared in D_2O and droplets were sandwiched in-between CaF_2 windows with mica spacers of 12 μm (for concentrated solutions) or 150 μm (for diluted samples). Data were acquired with a resolution of 4 cm^{-1} and 128 accumulations. Backgrounds with 20 mM NaOH into D_2O were registered prior to measurements and subtracted from the data. **Circular dichroism** was performed using a Chirascan spectropolarimeter (Applied Photophysics, UK) or the module B of the B23 synchrotron radiation CD (SRCD) beamline, at Diamond Light Source (Didcot, UK).[5] Demountable Hellma

cuvettes with path lengths of 0.01 or 0.1 mm were used according to concentration to optimize signal-to-noise ratios and provide absorbance A < 2. Far-UV region was recorded in the interval between 175 and 260 nm. With the Chirascan machine, scans were performed with a 1 nm band width bandwidth, 1 nm per step and 1 second per step. In heating/cooling experiments, a Peltier controller provided temperatures between 20 °C and 90 °C and cuvettes were sealed with parafilm to avoid evaporation. Prior to each temperature-controlled spectrum, a waiting time of 5 minutes was used to allow thermal equilibration. Data were averaged over 4 accumulations. For long-term monitoring, of the order of days, solutions were kept in Eppendorf tubes, incubated at RT and then transferred into cuvettes. UV-degradation experiments were conducted scanning a 1 wt% solution, 10 days after mixing, on SRCD beamline B23 at Diamond Light Source (Didcot, UK) and further details on the set up can be found in the work of Hussain et al.[5] The beamline set up and measurement conditions comprise 1 mm slits, 1 nm per step, 1 s for collection time and 1 nm for bandwidth. A total of 18 scans were recorded to reach full conversion from unordered into β-sheet structures. Experiments were performed at 20 °C and data have been converted into mean residue ellipticity units to allow comparison between different concentrations and path lengths.

CRITICAL AGGREGATION CONCENTRATION



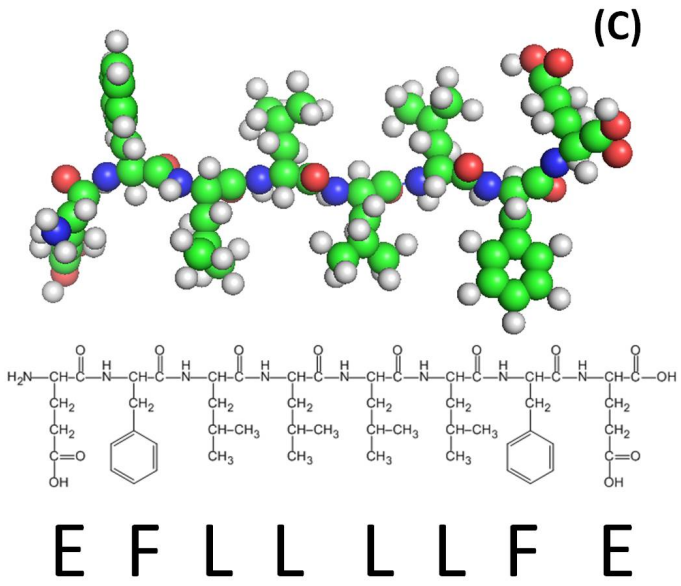


Figure S1: Pyrene assays to estimate the critical aggregation concentration (*cac*) above which self-assemblies are spontaneously formed. (A) Emission profiles from a concentration-series of peptide solutions containing 1×10^{-5} wt% pyrene. Excitation at $\lambda_{exc.}$ = 338 nm. (B) Fluorescence intensity at 375 nm as a function of log[peptide concentration]. Two different regimes are identified with a clear separation around 0.03 wt%, the *cac*. (C) molecular structure of the EFL_4FE bolaamphiphilic peptide.

SUPPLEMENTARY CRYO-TEM IMAGES

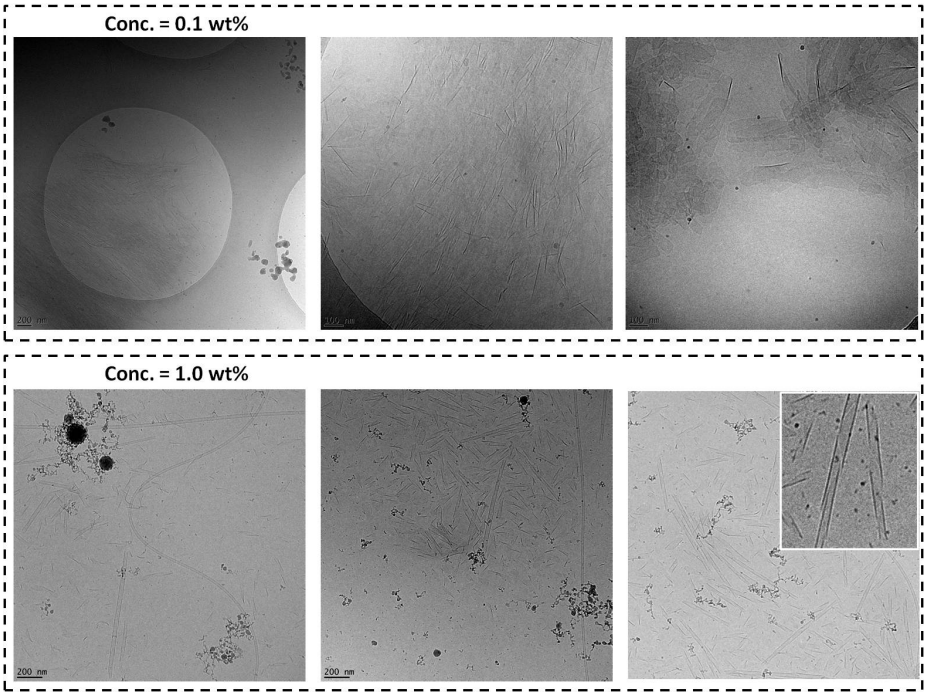


Figure S2: Cryo-TEM images showing the formation of flat sheets at 0.1 wt% EFL_4FE (top row) and the scroll into PNTs at 1 wt% EFL_4FE.

SAXS MODELING

SAXS data have been fitted using least-square routines available in the SASFit program.[4] More details on the form factors and developments of the equations used here can be found in the works of Pabst et al.[6] and Pedersen.[7] In the case of our formulations, total intensities have been modeled using the general expression:

$$I_{tot}(q) = Bkg + I_{bil}(q) + \int_0^\infty D(R) I_{tubes}(q) dR \qquad (1)$$

In Eq. (1), *Bkg* is an additive constant accounting for a flat background, $I_{bil}(q)$ represents intensities scattered from membranes and $I_{tubes}(q)$ is the component associated to PNTs. *D(R)* is a Gaussian distribution accounting for polydispersity in radius in PNTs. Data from diluted or fresh samples have been fitted using only the first two terms in Eq. (1). When oscillations appeared in the low *q* range, indicating the presence of PNTs, the third term has been included.

For the component associated to membranes, we have used a bilayer form factor popular in the description of lipid membranes[8, 9] and which has been successfully used to fit data from amphiphilic peptides in previous work.[6, 10, 11] Since our structures are made up from bolaamphiphilic monolayers, where a hydrophobic core appears surrounded by two polar interfaces, the context is analogous to bilayers made from single-headed lipids. According to this model, the electron density along the cross-sectional profile is represented by a sum of three Gaussian functions, two of them describing polar regions at the outer interfaces and the third accounting for the hydrophobic core in the middle of the structure. The scattering intensities from this component are given by[6]:

$$I_{bil}(q) = \frac{N_{bil}}{q^2} \times \left[\sqrt{2\pi}\, \sigma_{core}\, b_{core} \exp\left(-q \frac{t^2 \sigma_{core}^2}{2}\right) + 2\sqrt{2\pi}\, \sigma_{out}\, b_{out} \exp\left(-\frac{q^{2\sigma_{out}^2}}{2}\right) cos\left(q \frac{t}{2}\right) \right]^2 \qquad (2)$$

N_{bil} is a scaling constant related to volume and number of scattering particles in the sample. The characteristic q^{-2} descent related to flat geometry is also observed in the model. The first term between brackets corresponds to the Gaussian function describing the central core whereas the second term is related to Gaussians describing the polar heads. Parameters b_{out} and b_{core} are, respectively, the electronic contrasts of polar heads and hydrophobic core regarding the solvent; σ_{out} and σ_{core} are standard deviations of the Gaussians related to the size of heads and core. Finally, *t* is the centre-to-centre separation between Gaussians. In Fig. S3, the profile is represented together with the corresponding parameters.

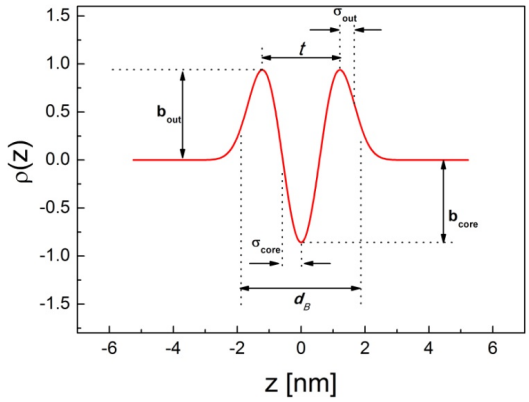


Figure S3: Three-Gaussian model used to describe cross-section electron density in amphiphilic bilayers.[6]

The major structural variable in the context of this work is the bilayer thickness, which is defined as the sum of head-to-head separation plus the full width at half maximum of the outer Gaussians:[6]

$$d_B = t + 2\sqrt{2ln2}\,\sigma_{out} \qquad (3)$$

For hollow tubes, scattering intensities are given by: [7,11]

$$I_{tubes}(q) = \left[\frac{2Si(qL)}{qL} - \left(\frac{\sin\left(\frac{qL}{2}\right)}{\frac{qL}{2}}\right)\right] \times \left(\frac{2J_1(qR)}{qR}(\eta_{core} - \eta_{shell})R^2L\pi + \frac{2J_1(q(R+\Delta R))}{q(R+\Delta R)}(\eta_{shell} - \eta_{solv})(R + \Delta R)^2L\pi\right)^2 \qquad (4)$$

$\boldsymbol{J_1}$ is the regular cylindrical Bessel function of first order and ***Si*** is the sine integral function.[7] Parameters η_{shell}, η_{core} and η_{solv}, respectively, account for electron densities of cylinder shell, cylinder core and solvent. These contrasts behave as multiplying constants and here they were fixed at $\eta_{shell} = 1$, and $\eta_{core} = \eta_{solv} = 0$. ***L*** is the length, also behaving as scaling constant for ***L >> R***. Since cryo-TEM imaging shows lengths reaching the micrometer length scale, we have fixed ***L*** = 1000. ***R*** is the inner radius of the cylinder and **Δ*R*** is the thickness of the shell. These are the most important structural parameters in the context of this work. Since a Gaussian distribution has been included in the fitting to account for polydispersity in radius, ***R*** appears in Table 1 accompanied by the corresponding standard deviation (***std***). Thus, to fit tubular components, we have used only 4 free parameters: ***R***, **Δ*R***, ***std*** and a scaling constant, $\boldsymbol{N_{cyl}}$.

Table S1: Parameters from model fitting of SAXS data. Columns in gray correspond to major parameters related to dimensions of sheets and cylinders.

Sample		**Bilayer form factor**							**Cylinder shell form factor**				
Conc. (wt%)	***time (days)***	N_{bil}	σ_{out} ***(nm)***	b_{out}	σ_{core} ***(nm)***	b_{core}	***t (nm)***	d_B ***(nm)***	N_{cyl}	***R (nm)***	***std (nm)***	ΔR ***(nm)***	***bkg***
0.1	10	7.20E-12	0.40	0.097	1.20	0.024	2.15	3.09	***	***	***	***	0.29
0.5	10	1.51E-11	0.41	0.100	0.80	0.011	2.28	3.25	***	***	***	***	0.07
1	10	1.00E-11	0.38	0.106	1.21	0.008	2.43	3.32	3.17E-07	9.20	1.57	3.87	0.03
1	0	1.85E-10	0.40	0.097	1.22	0.008	2.20	3.14	***	***	***	***	0.05
1	1	1.28E-10	0.41	0.120	1.21	0.008	2.17	3.14	***	***	***	***	0.04
1	3	1.37E-10	0.42	0.110	1.20	0.008	2.16	3.15	9.25E-07	10.47	1.20	3.28	0.05
1	5	1.25E-10	0.44	0.108	1.20	0.008	2.13	3.17	9.80E-07	10.30	0.77	3.67	0.10
1	8	8.20E-11	0.43	0.120	1.21	0.008	2.29	3.30	2.10E-06	10.01	1.20	3.47	0.02
1	10	9.10E-11	0.41	0.110	1.21	0.008	2.32	3.29	2.49E-06	9.65	1.08	3.78	0.05

SAXS FROM SAMPLE INCUBATED FOR LONGER TIME

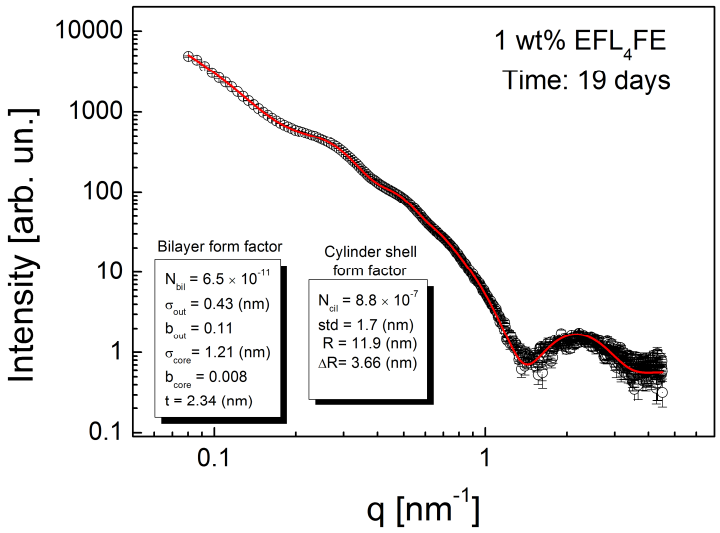


Figure S4: SAXS data after incubation time of 19 days. The signature of PNTs is still observed, with fitting parameters similar to those presented on Table S1.

FTIR SPECTRA

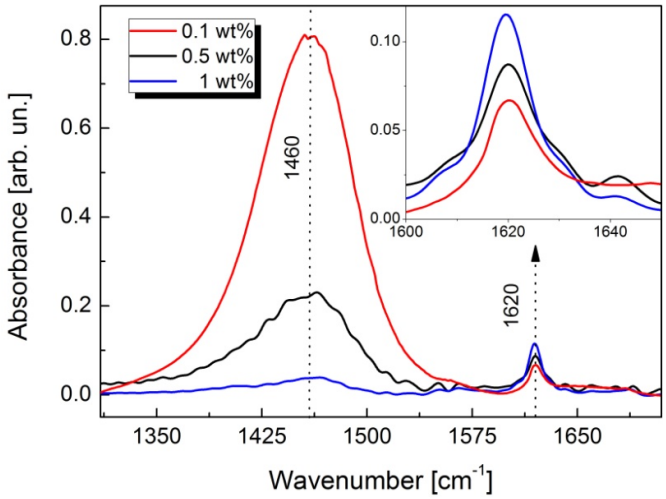


Figure S5: Typical FTIR spectra from EFL_4EF solutions prepared at different concentrations above *cac*. Amide I region is magnified in the inset. Data have been normalized to concentration and path lengths to make them comparable. Dotted lines are guides for the eyes to indicate wave numbers where notable peaks appear.

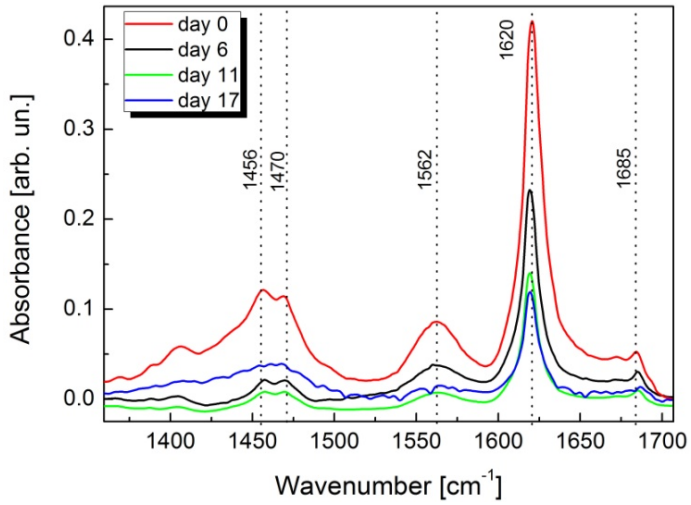


Figure S6: FTIR spectra from a peptide solution prepared at 1 wt% EFL_4EF sampled over a period of days.

FIBRE DIFFRACTION EXPERIMENTS

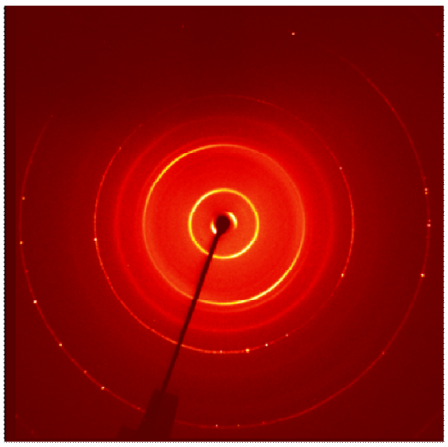


Figure S7: Fibre diffraction XRD pattern from a partially-oriented stalk obtained from EFL_4FE solution at 2 wt%.

Table 2: Observed XRD reflections with the calculated indexation according to an orthorhombic unit cell with lattice parameters *a* = 31.15 Å, *b* = 11.17 Å and *c* = 9.82 Å. Reflections corresponding to (010) and (002) planes are ascribed to separations between β-strands and β-sheets. The peaks in bold are most intense.

Reflection	Calculated (Å)	Experimental (Å)
(100)	**31.15**	**31.73**
(010)	**11.17**	**11.22**
(002)	**4.91**	**4.88**
(012)	4.49	4.48
(030)	3.72	3.74
(203)	3.20	3.21
(123)	2.81	2.81
(024)	2.25	2.25
(244)	1.83	1.83

The lattice cell parameters may be correlated, respectively, to major features of the EFL_4FE super-structure. For instance, the *d* spacing associated with the (100) reflection d_{100} may be associated to the length of the octamer (8×3.4 Å = 27.2 Å)[12] plus monolayers of Na^+ ions (Na^+ diameter ~ 3.7 Å), [13] and d_{010} could be related to the separation between β-sheets and the reflection (001) is twice the β-strand separation, 4.88 Å, and its extinction reflects the anti-parallel nature of the crystal organization.

CD ASSAYS

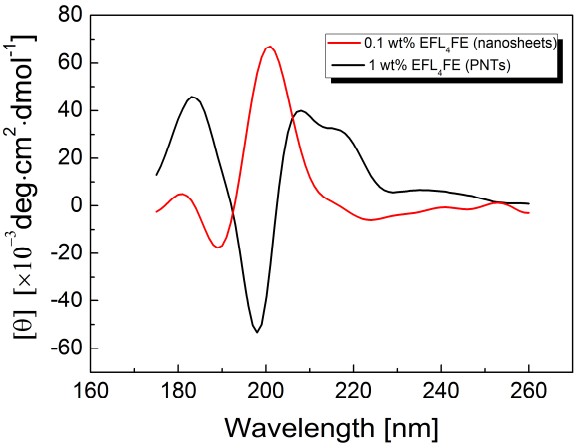


Figure S8: CD spectra from solutions at indicated concentrations. For the sample populated exclusively by nanosheets the CD spectrum corresponds to that usually ascribed to β-sheet secondary structure. In contrast, at higher concentration where PNTs appear, the profile corresponds to "unordered" structures.

CONTROL SAMPLES DISSOLVED INTO 1:1 H2O:ETOH AND 1:1 NaCl:HCl MIXTURES

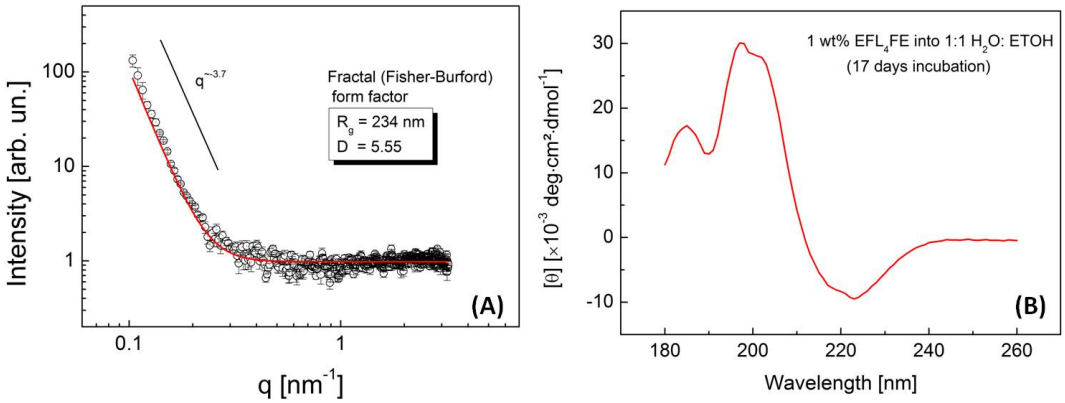


Figure S9: SAXS and CD data from a 1 wt% peptide solution dissolved into 1:1 ETOH:H2O. (A) the intensity decrease with $q^{-3.7}$ is characteristic of surface fractal aggregates.[14] The fit has been performed with a fractal Fisher-Burford form factor.[15]

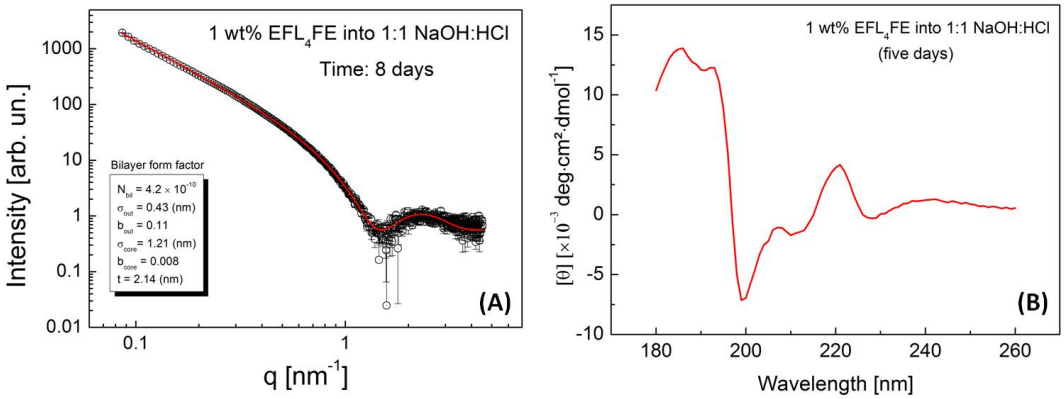


Figure S10: SAXS and CD data from a 1 wt% peptide solution dissolved into 1:1 NaOH:HCl (20 mM). SAAXS data are fitted according to the same bilayer form factor used in the main text.

CD TIME SERIES

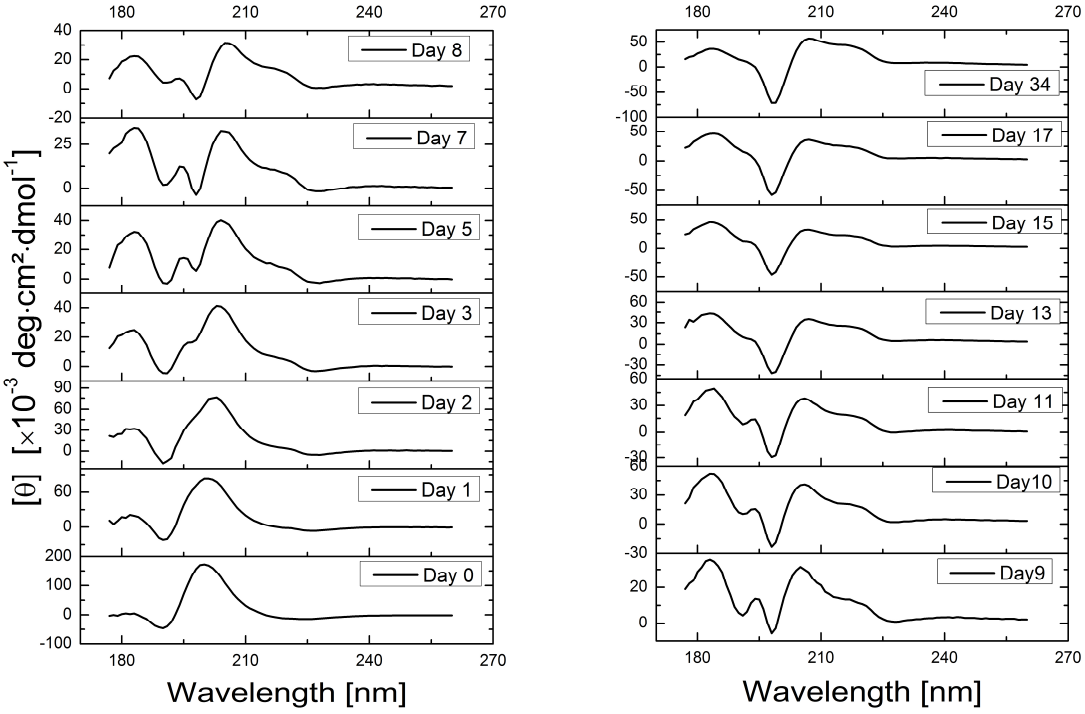


Figure S11: CD spectra from solutions probed at different times.

ESTIMATION OF β-STRUCTURE CONTENT FROM CD DATA

The fraction of β-structures has been estimated using the ellipticity at 199 nm, through the formula:[16]

$$\alpha_\beta = \frac{([\theta]_t - [\theta]_u)}{([\theta]_\beta - [\theta]_u)} \qquad (5)$$

In Eq. (5), α_β is the fraction of β-sheet content, $[\theta]_t$ is the ellipticity at 199 nm at a given time t, $[\theta]_u$ is the ellipticity when solution is fully dominated by unordered-like structures and $[\theta]_\beta$ is the ellipticity when sample is dominated by β-sheet structures. In the case of the kinetic measurements shown in Fig. 3 of the main text, $[\theta]_\beta$ = 170,451 deg·cm²·dmol^{-1} has been obtained from the very first spectrum, obtained from fresh solutions (around 30 minutes after mixing). In this case, we have assumed no significant amount of bent/twisted sheets in the very early stage of self-assembly. The value of $[\theta]_u$ = -71,282 deg·cm²·dmol^{-1} has been obtained from a spectrum registered much later in time, 34 days after sample preparation. In this case, since the general profile is essentially the same for samples after *ca.* 10 days from preparation, we have assumed that the fraction of β structures in this aged sample can be disregarded. Then, the time course has been constructed using only independent spectra, excluding data used to estimate $[\theta]_\beta$ and $[\theta]_u$.

To describe the fraction during UV-degradation scanning shown in Fig. 4 of main text, the procedure also involved Eq. (5). The value for $[\theta]_\beta$ has been assumed to be the ellipticity found in the last frame (18th) of the measurement series ($[\theta]_\beta$ = 72,316 deg·cm²·dmol^{-1}) which appears in a region of saturation (see inset Fig. 4). However, since the sample had age of 10 days from preparation, it is presumed that a fraction of β-like structures is also present and need to be considered in the calculations. From the time-course obtained in Fig. 3, we estimated a content *ca.* 20% of β-sheets contributing to this first spectrum. The ellipticity at 199 nm in the first frame has been measured to be $[\theta]$ = -41,503 deg·cm²·dmol^{-1}. In this case, the reference value for $[\theta]_u$ for a solution containing "only" PNTs could be obtained using Eq. (5):

$$[\theta]_u = \frac{-41{,}503 - 0.2 \times 72{,}316}{0.8} = -69{,}958\ \mathrm{deg \cdot cm^2 \cdot dmol^{-1}}$$

The reference value found above is consistent with ellipticity measured at the solution used as reference to calculate the time-course, which has been measured 34 days after preparation.

TEMPERATURE-CONTROLLED CD EXPERIMENTS

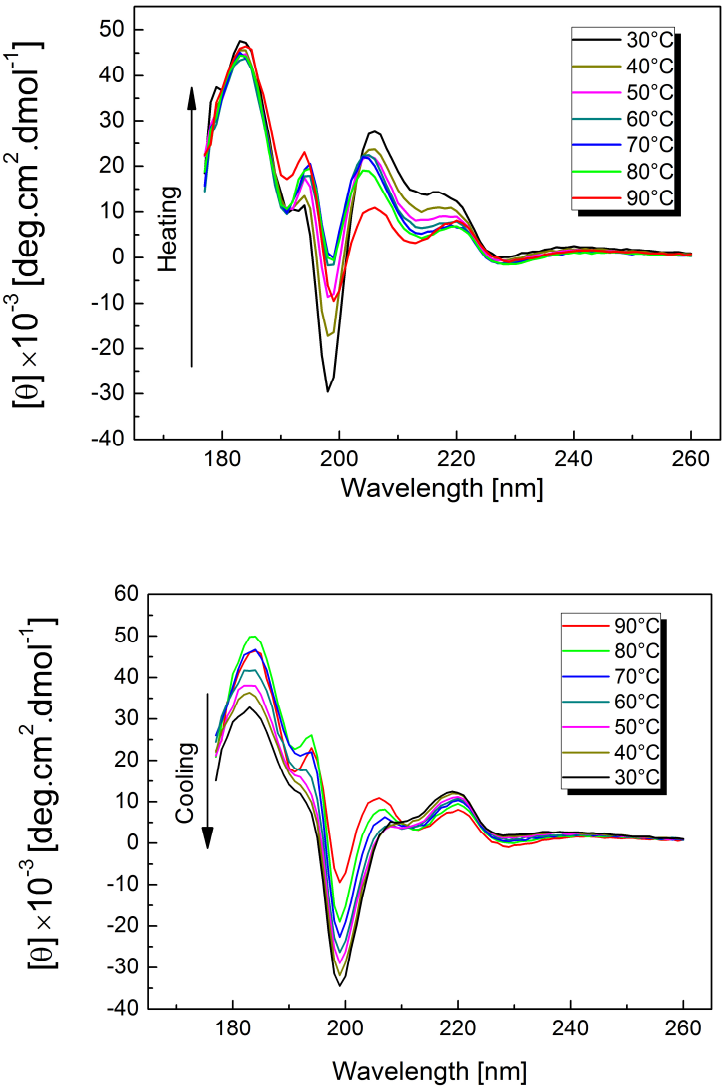


Figure S12: Temperature series of CD spectra from a 1wt% EFL_4FE solution, after 10 days of incubation at RT, showing heating (left) and cooling (right) ramps. Intensity of the minimum at 199 nm decreases upon temperature increasing; however, unordered-like signature is kept up to 90°C. During the cooling process, the intensity at 199 nm increases again, indicating growth of unordered conformations.

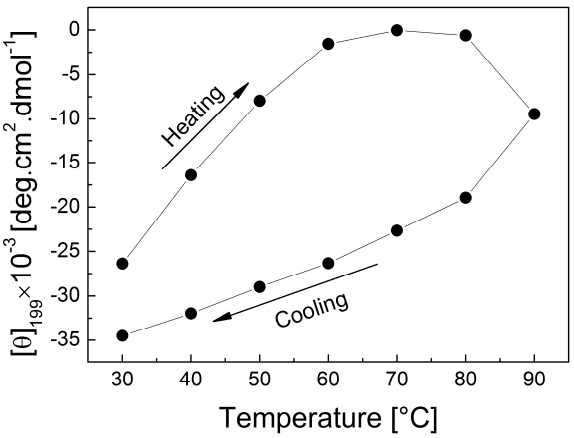


Figure S13: Molar elipticity from CD spectrum at 199 nm as a function of temperature during the heating/cooling cycle. A breakdown in the curve is observed at ~ 60°C, when the peak reaches its maximum value. Above this value, the content of disordered structures increases up to the maximum temperature, 90°C. Unlike observations made upon UV-degradation, the process here is reversible, exhibiting a clear loop-like signature.

1. V. Castelletto, R. M. Gouveia, C. J. Connon, I. W. Hamley, J. Seitsonen, A. Nykanen and J. Ruokolainen, *Biomater Sci-Uk*, 2014, **2**, 362-369.
2. O. S. Makin, P. Sikorski and L. C. Serpell, *Journal of Applied Crystallography*, 2007, **40**, 966-972.
3. P. Pernot, A. Round, R. Barrett, A. D. Antolinos, A. Gobbo, E. Gordon, J. Huet, J. Kieffer, M. Lentini, M. Mattenet, C. Morawe, C. Mueller-Dieckmann, S. Ohlsson, W. Schmid, J. Surr, P. Theveneau, L. Zerrad and S. McSweeney, *J. Synchrot. Radiat.*, 2013, **20**, 660-664.
4. J. Kohlbrecher, Bressler, I., Software package SASfit for fitting small-angle scattering curves

https://kur.web.psi.ch/sans1/SANSSoft/sasfit.html).

5. R. Hussain, T. S. Javorfi and G. Siligardi, *J Synchrotron Radiat*, 2012, **19**, 132-135.
6. G. Pabst, M. Rappolt, H. Amenitsch and P. Laggner, *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, 2000, **62**, 4000-4009.
7. J. S. Pedersen, *Advances in Colloid and Interface Science*, 1997, **70**, 171-210.
8. C. L. P. Oliveira, B. B. Gerbelli, E. R. T. Silva, F. Nallet, L. Navailles, E. A. Oliveira and J. S. Pedersen, *Journal of Applied Crystallography*, 2012, **45**, 1278-1286.
9. N. Kucerka, J. Pencer, J. N. Sachs, J. F. Nagle and J. Katsaras, *Langmuir*, 2007, **23**, 1292-1299.
10. I. W. Hamley, A. Dehsorkhi and V. Castelletto, *Langmuir*, 2013, **29**, 5050-5059.

11. I. W. Hamley, A. Dehsorkhi and V. Castelletto, *Chem Commun (Camb)*, 2013, **49**, 1850-1852.
12. T. E. Creighton, *Proteins : structures and molecular properties*, W.H. Freeman, New York, 2nd edn., 1993.
13. R. Schmid, A. M. Miah and V. N. Sapunov, *Phys Chem Chem Phys*, 2000, **2**, 97-102.
14. T. Zemb and P. Lindner, *Neutrons, X-rays and light : scattering methods applied to soft condensed matter*, Elsevier, Amsterdam ; Boston, 1st edn., 2002.
15. J. K. a. I. Bressler, *Journal*, 2011.
16. N. J. Greenfield, *Nat Protoc*, 2006, **1**, 2891-2899.