Article:
Davies, RPW, Liu, B, Maude, S et al. (4 more authors) (2018) Peptide strand length controls the energetics of self-assembly and morphology of β-sheet fibrils. Peptide Science, 110 (1). e23073. ISSN 0006-3525

https://doi.org/10.1002/bip.23073

Peptide Strand Length Controls the Energetics of Self-Assembly and Morphology of β-sheet Fibrils.

Robert P.W. Davies[†], *Binbin Liu*[ϕ], *Steven Maude*[‡], *Lisa M. Carrick*[‡] *Irina Nyrkova*[ʃ], *Tom C. McLeish*[§], *Sarah. A Harris**[∂].

[†] Department of Oral Biology, University of Leeds, St James's University Hospital Leeds LS9 7TF. [ϕ] Department of Pediatrics, University of Oxford. Headington, Oxford, OX3 9DU. [‡] School of Chemistry, University of Leeds, Leeds, LS2 9JT. [ʃ] Institute Charles Sandron 67034 France.[§] Department of Physics, Durham University, Durham, DH1 3LE. [∂] School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT

KEYWORDS . Peptide Self-Assembly, Statistical Mechanics, Atomistic Computer Simulation.

ABSTRACT. Self-assembling peptides can be used as versatile, natural and multifunctional building blocks to produce a variety of well-defined nanostructures, materials and devices for applications in medicine and nanotechnology. Here we concentrate on the 1D self-assembly of *de novo* designed Px-2 peptide β-strands into anti-parallel β-sheet tapes and higher order aggregates. We study six members of the Px-2 family, ranging from 3 amino acids (aa) to 13 aa in length, using a range of complementary experimental techniques, computer simulation and theoretical statistical mechanics. The critical concentration for self-assembly (c*) and is found to

increase systematically with decreasing peptide length. The shortest peptide found to self-assemble into soluble β-tapes in water is a 5 amino acid residue peptide. These investigations help decipher the role of the peptide length in controlling self-assembly, aggregate morphology and material properties. By extracting free energies from these data using a statistical mechanical analysis and combining the results with computer simulations at the atomistic level, we can extract the entropy of association for individual β-strands.

1 INTRODUCTION

Peptide self-assembly can give rise to an extensive range of functional or pathological nanostructures [1–5]. In recent years, self-assembling peptides have been used as versatile, natural and multifunctional building blocks to produce a variety of well-defined nanostructures, materials and devices for applications in medicine and nanotechnology [2,6–11]. To develop novel materials from self-assembling peptides, it is imperative that we acquire a thorough understanding of the principal and forces that govern peptide self-assembly as well as a detailed knowledge of the relationship between peptide primary structure, self-assembly, nanostructure formation and material properties.

One class of self-assembling peptides shown in Figure 1, has been found to undergo one-dimensional self-assembly in solution to give rise to β-sheet nanotapes, as well as higher order aggregates: ribbons, fibrils and fibres when equilibrium is reached (as per the previous publications a tape is defined as a aggregated number of β-strands; a ribbon is a double tape; a number of ribbons can stack together to form a fibril; fibers are formed when fibrils interact edge on edge and twist together) [12–16]. Solutions and gels made of these peptides are currently

being explored in a broad range of applications [11,17–22]. One of the best characterised members of this class of tape-forming peptides is the *de novo* designed 11mer peptide P_{11}-2 [13,14,23–25]. Once aggregated, P_{11}-2 gives rise to amphiphilic β-sheet tapes with a hydrophilic face decorated by polar (glutamine, arginine and glutamic acid) side chains and a hydrophobic face decorated mainly by aromatic (phenylalanine and tryptophan) side chains. Electrostatic, hydrogen bonding and aromatic interactions have all been designed into the primary structure of P_{11}-2 in order to promote its self-assembly in 1D into long fibrils which are soluble in water. An odd number of amino acid residues were chosen in order to maximise the hydrogen bonding and electrostatic interactions within the anti-parallel β-sheet. The length of this 11mer peptide provides an experimentally convenient balance between the equilibrium time required for self-assembly (which becomes prohibitively long for longer peptides) and the emergence of a rich variety of higher order fibrillar structures [14] as a function of peptide concentration.

Varying the length of the peptide β-strand can have profound effects on self-assembly. Firstly, as the width of the resulting tape is equal to the length of the peptide strand, aggregated structures can be designed with bespoke dimensions. Secondly, the peptide length determines the number of intermolecular interactions between adjacent β-strands, which controls the energetic driving force for self-assembly and the critical concentration (c*) of monomer strands required. Consequently changing the peptide length can directly affect the morphology of the aggregates produced (e.g. their width, length, persistence length, and scission energy) and therefore alters the material properties of the peptide containing solutions. Properties such as the ability to undergo an isotropic to nematic transition (I/N), and the concentration at the transition $c_{I/N}$, can all be affected, as can the gelation properties and the mechanical robustness of the gels formed. Finally, the peptide length has important commercial implications: the shorter the peptide, the

easier, faster and cheaper it can be synthesized, thus making it more attractive for industrial applications.

Here we study six members of the Px-2 family, ranging from 3 amino acids (aa) to 13 aa in length (Table 1 and Figure. 1). All peptides share the same principle architecture which determines the properties of β-sheet structure: antiparallel β-sheet (due to presence of oppositely charged Arg (R+) and Glu (E-) residues close to the opposite ends of the sequence) with distinct hydrophilic and hydrophobic sides (due to the presence of highly hydrophobic aromatic Phe (F) and Trp (W) residues alternating with hydrophilic Gln (Q) residues). At pH 2, the charge on the Glu (E) is neutralized in the isolated state of the peptides, however it may have a propensity to remain charged when embedded in an anti-parallel β-sheet where it is located near the Arg residue of the neigbouring peptide. The choice of pH 2 as an experimental condition was to ensure that all of the 7 variants of peptides were completely soluble, which was not the case at neutral pH. The self-assembly of all of these peptides in water in the dilute regime at pH 2 and in the model polar organic solvent methanol (presented in supplementary Figures 1 and 2) are studied by circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy, transmission electron (TEM), and computer modelling. These investigations characterise the role of peptide length in controlling self-assembly, aggregate morphology and material properties.

2 EXPERIMENTAL

2.1 Materials. Peptides P_{13}-2 and P_{11}-2 were synthesized in-house as described previously [12], P_9-2, P_7-2 and P_5-2 were custom synthesised by NeoMPS SA, France, and P_3-2 was custom synthesized by GenScript. Peptide quality control was undertaken using mass spectroscopy,

analytical HPLC, amino acid and elemental analysis (carbon, hydrogen, nitrogen and fluorine) and UV spectroscopy, and the result are provided as Supplementary Information in Table S1. Non-peptide molecules present in the dry peptide mass were mainly residual amounts of water and intrinsic counterions; charged ammonium counterions bound on negatively charged peptide groups and trifluoroacetic acid (TFA) counterions bound on positively charged groups. All peptides were stored at -20°C as lyophilised white fluffy powders. D_2O (99.9% D, Fluorochem Ltd/Sigma-Aldrich) DCl (35% weight DCl,99%D,sigma-Aldrich), NaOD (40% weight Sigma-Aldrich), (2,2,3,3-d4)-trimethylsilyl-3-propionic acid (TMSP) (Cambridge Isotope Laboratories, Inc).

2.2 Sample preparation. Samples were prepared by dissolving dry lyophilised powder in the relevant solvent producing a stock solution. This stock solution was subsequently diluted to the correct concentration accordingly. The morphology of the resulting fibrils can be sensitive to the precise method of preparation [14], such as whether the aggregation occurs directly from the hydrated powder or from a stock solution. In this study, it was necessary to prepare the samples using an intermediate stock solution in order to accurately produce samples with low concentrations.

2.3 UV spectroscopy. Solutions were measured in Hellma 1 mm quartz cuvettes and spectra were collected with a Perkin-Elmer Lambda 2 UV-Vis spectrophotometer. Peptide concentration was derived by using the Beer-Lambert law with a molar absorption coefficient of 5500 mol^{-1} dm^3 cm^{-1} at 279 nm [26].

2.4 Secondary structure determination with circular dichroism UV (CD-UV) and high resolution ^{1}H- NMR. To determine the critical concentration (c*) of monomers required for

self-assembly, CD and UV were used to obtain secondary-structure information during aggregation for P_{13}-2, P_{11}-2 and P_9-2. All CD spectra were acquired using a JASCO 715 spectrometer using 1 mm or 10 mm quartz cuvettes at 20^{o}C. Spectra were recorded with a step resolution of 1 nm, a scan speed of 50 nm/minute, a sensitivity of 50 mdegs and a response time of 1 second. Ten scans per sample were typically acquired over a range of 190 to 300 nm. The resulting data was then solvent subtracted, baseline corrected, smoothened and normalised. Regions of the spectra where the high tension voltage exceeded a value of 750 mV were removed before analysis was carried out.

For the shorter peptides, however, the amount of secondary structure present in the self-assembling peptide solutions was measured with ^{1}H NMR. Solutions of peptides were prepared using D_2O and pH adjusted using DCl and NaOD with the addition of (2,2,3,3-d4)-trimethylsilyl-3-propionic acid (TMSP) reference. Peptide solutions (800 μl) were then dispensed into 500 MHz borosilicate glass tubes. Spectra were recorded at room temperature using a DPX300 spectrometer operating at 300MHz (5 mm probe, spectral width 5995 Hz), controlled by XwinNMR software. A presaturation program was used to minimize the water peak, and 1024 scans were measured per spectrum. In order to ascertain the concentration of peptide monomer present in solution, the aromatic peaks of the peptide at chemical shifts 6.9-7.8 ppm were integrated and normalised against a reference peak of 0.125 mM TMSP at 0 ppm. The normalised integrals of the aromatic peaks were then compared with the corresponding values from solutions containing known concentration of monomeric peptides.

2.5 Transmission electron microscopy (TEM) Electron microscope (EM) grids (hexagonal copper grids of mesh size 400 mesh) were carbon coated prior to use, by the flotation of a carbon

film from a mica sheet onto the grids. The EM grids were then glow discharged prior to sample application to ensure adhesion of the sample to the EM grid. Peptide solutions were immediately diluted. The EM grids were touched onto the peptide solutions for one minute. The grid was air-dried and then negatively stained with uranyl acetate solution (4% w/v) for 20 seconds and again air-dried. TEM images were obtained using a Philips CM10 electron microscope operating at 80 kV accelerating voltage. Images were obtained quickly to avoid artefacts and destruction of the sample. The morphological studies of the fibrillar aggregates were performed primarily with TEM.

2.6 Statistical mechanical model fitting to self-assembly data Self-assembly data (i.e. the concentration dependence of the fraction of aggregated peptide measured by NMR or CD) for each peptide were fitted to a statistical mechanical model previously reported by Nyrkova *et al* [14,27]. This model describes peptide aggregation in terms of individual monomers, single and double β-sheet tapes, and fibrils (which are stacks containing *p* tapes, where the number of tapes in these fibrils is estimated from TEM measurements), and calculates the corresponding concentration dependences of their populations (see eqs.(10,50) in ref. [27] for further details). The equilibrium self-assembly of the different possible structures is parameterised by a small set of coarse-grained peptide-peptide interaction energies:

ε_{tr} is the transformation energy, that is the free energy difference between a peptide in a random coil conformation in solution and a peptide in rod-like conformation which is required in order to be incorporated into the β-sheet. We note that these peptides are expected to have significant helical order when free in solution, so the associated entropy change will be far smaller than that obtained for the transition between a truly random coil[28] and an extended β-strand, and has been found to make a small or negligible contribution, depending on the aggregating system

studied [29,30].

ε is the β-scission energy, i.e. the total free energy gain when two tapes join into a common β-sheet. This contains the interaction energy between two neighbouring rods in a single β-sheet (ε_β) and the translational entropy associated with peptide localization inside the β-sheet. Defining the effective bond volume (see Supplementary Information for further details) within a β-sheet v_β, and an arbitrary reference volume (chosen to be 10^{-3} nm^3, which corresponds to an effective peptide concentration of 1660M) gives:

$$\varepsilon = \varepsilon_\beta + k_B T \times \ln (v_\beta / v_0) \qquad \text{Eq. 1}$$

ε_p is the fibrillization energy for a fibril made of p single tapes. This is the mean free energy gain per peptide in a stack containing p tapes compared with a peptide in a single tape. For the peptides from Px-2 family in question, $\varepsilon_p \approx \varepsilon_2$ (where ε_2 is the energy of a double tape), as the side-liaison energy is dominated by the energy responsible for the double tape formation (related to strong hydrophobic interactions due to the presence of W- and F- side-groups).

The total free energy gain for peptide association into a fibril is given by:

$$\varepsilon_{tot} = \varepsilon_\beta + \varepsilon_p - \varepsilon_{tr} + k_B T \times \ln (v_\beta / v_0) \cong - k_B T \times \ln (v_0 c^*) \qquad \text{Eq. 2}$$

To physically interpret the terms in Eq. 1, note that ε_{tot} is the chemical potential of a peptide within a fibril, whereas $k_B T \times \ln(v_0 c^*)$ is the one for a free peptide at critical concentration for fibril assembly. The total energy gain for the peptide in a fibril comes from the energies of hydrogen bonding between the tapes within the fibril, from conformation transformation from random coil to rod, and from the entropy of peptide translation within the fibril, respectively. The

right-hand approximate equation in Eq.2 reflects the fact that there should be a chemical equilibrium between free peptides and the peptides associated into fibrils; note that ε_{tot} is the chemical potential of a peptide within a fibril, whereas $k_BT \bullet \ln (v_0 c^*)$ is the one for a free peptide in equilibrium with the fibrils (the critical concentration for fibril assembly c* coincides with the saturation concentration of free peptides in high concentration solutions) (see ref [26] for details) .

We previously showed that the experimental data for self-assembling of P_{11}-2 peptide in pure water were best-fitted with [14,27]:

$$\varepsilon_{tr} \approx 3.0 k_B T$$

$$\varepsilon_t = \varepsilon_\beta + k_B T \times \ln (v_\beta / v_0) \approx 19.3\ k_B T \qquad \text{Eq. 3}$$

$$\varepsilon_t \approx 0.6\ k_B T$$

$$\varepsilon_p - \varepsilon_2 \approx 2 \times 10^{-4}\ k_B T$$

$$p \approx 8$$

The bond volume for P_{11}-2 was estimated, from simulations, as $v_\beta \sim (2 - 15) \times 10^{-6}$ nm^3 with the logarithmic average $v_\beta \sim 5 \times 10^{-6}$ nm^3, this corresponds to the β-strand best-fit interaction energy of ~ 24.6 k_BT (at neutral pH). Note that, from the experimental data alone, it is not possible to extract values for v_β and ε_β separately, however they represent distinct physical contributions and a convincing model must permit reasonable fitting values for both.

2.7 Construction of atomistic models. Peptides P_5-2, P_7-2 and P_{11}-2 were constructed in a β-strand configuration (with torsion angles = -139.0, = 135.0 and = -178.0), and assembled into the anti-parallel β-sheet tape-like structures (see Figure 1) using the NAB molecular building tool v1.3 [31]. The peptide end groups were acetylated and amidated to neutralise the termini. To construct the tapes, we duplicated the β-strand peptides and used a matrix transformation to place them so that the strand-strand backbone hydrogen bond distances were 2.75Å (the distance between neighbouring β-sheets is around twice this value as it includes the contribution from the covalent bond lengths of atoms within the peptides). The side-chain conformations were selected and optimized based on the rotamer libraries in SCWRL and Swiss-PdbViewer [32–34]. The anti-parallel ribbon was constructed by assembling two tapes via the hydrophobic interface

2.8 Molecular dynamics (MD) simulation protocol.

All MD simulations were carried out using the Amber suite of programs [35] and the Amber03 force field [36]. The fast particle-mesh Ewald algorithm was used to treat long-range electrostatic interactions and bonds to hydrogen were constrained with SHAKE [35], allowing an integration time step of 2 fs. The atomistic models of the tape and ribbon (constructed as explained above) were energy minimized and then solvated in a periodic box of TIP3P water molecules with the solute a minimum distance of 10 Å from the edge of the box. To allow direct comparison with the experimental data, Glu residues were protonated to mimic the experimental condition of pH2 and Cl^- counterions were added to neutralize the system. To equilibrate the system, the solvent and solute were firstly energy minimized. This was followed by a restrained MD simulation in which the system was heated up to 300K followed by a multistate relaxation [37] (distance restraints were applied to the backbone hydrogen-bonds for 1 ns). Unrestrained

MD was then performed for over 40ns at constant temperature (300K) and constant pressure (1 atm). The MD simulations were run with 32 processors on the Opteron-Myrinet supercomputer at the University of Leeds.

2.9 MD data analysis

The last 20 ns trajectory from each MD simulation run was used for data analysis. The terminal strands in the aggregates excluded from the analysis in order to avoid edge effects. The hydrogen-bond occupancy and backbone Root Mean Squared Deviations (RMSDs) were calculated with the PTRAJ program in the AMBER suite of programs 31. Energetic analysis was performed on the atomistic structures generated by MD by removing the explicit solvent molecules and using the generalised Born/Surface Area (GB/SA) methodology to compute the energies of the solute [38,39]. The GB/SA model has been successfully applied to study the energy changes in biomolecular systems including amyloid β structures [40–42]. In addition to the configurational energies provided by the MD forcefield, and an estimate of the electrostatic interactions, the GB/SA model includes an approximation for the solvent entropy changes using an empirical term which is proportional to the total solvent accessible surface area. It does not include any estimate of the configurational entropy associated with the different atomic structures adopted by the solute. We introduce the notation $U_{\beta\text{-MD}}$ and $U_{p\text{-MD}}$ for the energetic components calculated from the simulations in order to distinguish them from the free energy extracted from experimental data using the statistical mechanical model.

The association energy of strand-strand in the single β-sheet of tape, $U_{\beta\text{-MD}}$, was calculated using:

$$U_{\beta-MD} = \frac{\langle E_{sheet} \rangle - \left(\langle E_{strand1} \rangle + \langle E_{strand2} \rangle + ... \langle E_{strandN} \rangle \right)}{N-1} \quad \text{Eq. 4}$$

where *N* is the number of strands in the peptide assembly, *N-1* is the number of interfaces, E_x denotes the internal energy of the peptide at position x within the stack and the angular brackets denote the ensemble average. The association energy of a ribbon from two tapes was calculated using:

$$U_{p-MD} = \langle E_{ribbon} \rangle - \left(\langle E_{sheet1} \rangle + \langle E_{sheet2} \rangle \right) \quad \text{Eq. 5}$$

Molecular structures of the tape and ribbon were visualized using VMD v1.8.7 [43] Convergence of these energies with the length of the β-sheet assemblies was checked by calculating the energy as a function of the number of β-strands considered, as shown in Supplementary Figure 3. To demonstrate that the energies of the peptide aggregates are statistically valid and have the expected Gaussian distribution, these were plotted as histograms, as shown in Supplementary Figure 4.

3 RESULTS AND DISCUSSION

To study the aggregation properties of this series of peptides, solutions were prepared as per the method outlined. The self-assembly as a function of peptide concentration was studied in the presence of 10 mM H_3PO_4 (Sigma-Aldrich) (at pH =2). In acidic conditions the peptides are more soluble than at the neutral pH we studied previously [14,25,27]) because they carry +1 net charge per monomer (Table 1). Depending on the length of the peptides, self-assembly was monitored by CD (for P_7-2, P_9-2, P_{11}-2 and P_{13}-2) or by ^{1}H-NMR (for more concentrated

solutions of P_5-2 and P_3-2) to acertain the aggregation status of the solution (until equilibrium was achieved). By performing these experiments for a range of concentrations, the critical concentration c* for self-assembly was determined. The same samples were also used to probe the morphological properties with TEM. The pH was held constant throughout the experiment and the assumed charge state is shown in Table 1.

3.1 Self-assembling properties in acidic aqueous solutions

For the longest oligopeptide P_{13}-2 the self-assembly of monomer peptides into β-sheets occurs at such low concentrations that these were below the threshold of sensitivity of the CD spectroscopy (c*<3μM). Moreover, β-sheet spectra were obtained at all measured concentrations, indicating that no random coil monomeric structures are present (Figure. 2(a)). TEM further revealed that these dilute solutions of P_{13}-2 had formed rigid fibrils with widths ~15 nm and lengths longer than 10μm (Figure. 3(a)).

.

For P_{11}-2 at pH2 we found c* to be 10 μM. The morphologies observed are uniform fibrillar structures, Figure 3(b), with lengths > 10 μm, widths of 9-10 nm and a full twist pitch around 200-240 nm. Self-assembling behavior of the peptide P_{11}-2 has been published previously in water 14. At neutral pH, the observed fibrils were similar to the present ones in width (made of ~8 individual β-sheet tapes), however, the critical concentration c* for aggregation into double tapes was 87 μM and for aggregation into fibrils was 500 μM which reflects the weakness of

electrostatic screening in pure peptide water solutions (hence, the tapes repel each other by a greater amount in pure water).

Typical CD UV spectra of P_9-2 at pH 2 are shown in Figure 2(b). At concentrations $c < 16$ μM, spectra show that the peptide is in a monomeric, random coil conformation illustrated by a negative band at $\lambda_{min} \approx 200$ nm. At $c = 16$ μM, a mixture of random coil and β-sheet bands are obtained. When $c \geq 40$ μM the spectra indicates that the peptide is in a β-sheet conformation. The positive absorption band obtained in the near-UV region at λ_{244} also increases in magnitude with peptide concentration. The solutions were subsequently measured after incubation at 20 °C for 7 days, and no change in the CD spectra was observed, indicating that the solutions were at an apparent equilibrium after 2 days. TEM studies of P_9-2 showed a small number of short curly aggregates are present as well, which may come from intermediate stages of the aggregation process (e.g. they may be super twisted ring-like double tapes); these aggregates represent less than 20% of the peptide and are ignored in the statistical mechanics analysis below.

Peptide P_5-2u ($CH_3COFQWQFNH_2$) was originally designed to complement the studies of its 13, 11, 9 and 7mer counterparts, as an exact 'truncated variant' of P_7-2. However, the reduction in length removed the only two charged residues (arginine and glutamic acid). Synthesis of crude P_5-2u was successful as indicated by mass spectrometry. The absence of charged groups in the primary structure of P_5-2u resulted in it being highly insoluble in solvents suitable for HPLC purification including but not limited to: H_2O, acidified H_2O (TFA), methanol, tetrahydrofuran and toluene. Therefore P_5-2u could not be purified. Ascertaining that charged residues are an

essential component for soluble peptides, P_5-2 (Table 1) was subsequently designed and synthesised to replace P_5-2u. P_5-2 includes the charged residues arginine and glutamic acid on its polar side, but also still retains key design elements such as the lateral hydrogen bond potential along the peptide backbone and distinct faces of opposing hydrophobicity which should drive self-assembly into tapes. Inclusion of these charged peptides successfully increased the solubility of P_5-2 and purification of it was successful. Initial experiments of P_5-2 using CD UV spectroscopy again showed that it was not possible to clearly observe the random coil to β-sheet transition clearly. This is likely due to the high aromatic content in this peptide and subsequent packing of the aromatic residues with each other, which causes significant distortion of the CD spectra (originating from the tryptophan). To overcome this problem, the self-assembly behaviour was monitored using high resolution ^{1}H NMR. The spectra acquired from low concentration solutions consist of sharp peaks indicative of monomeric peptides undergoing random re-orientational motion, whilst the normalised integrals of the aromatic side chains increase linearly with concentration up to ~300 μM. Above this concentration the normalised integrals no longer increase with increasing concentration, denoting the start of self-assembly. Morphologically there were a number of structures present rare flat structures with widths of 17-172 nm and lengths > 0.2 μm (image not shown). Fibrils that were observed more frequently had widths ranging from 15-30 nm and a pitch twist of common structures 75-80 nm (Figure 3(e)).

The self-assembly of P_3-2 was also monitored by high resolution ^{1}H NMR. Figure 2(c) illustrates that with increasing peptide concentration the relative size of the reference TMSP peak decreased upon comparison with the peptide peaks. Indeed the normalised integrals of the aromatic side chains increased linearly with increasing concentration for all concentrations

studied up to 3,400 μM with a positive gradient of 0.0073 ± 0.0002. This gradient is one third the value of the gradient of P_5-2 below c* which is due to P_3-2 having only one third of the aromatic residues of P_5-2. These results are consistent with the peptide not self-assembling at any concentration studied; indeed, for c< 3,143 μM, no aggregates were observed by negative staining TEM (not even amorphous aggregates).

By performing a systematic analysis of the self-assembly of a series of peptides of varying lengths but well defined physical chemical properties, we have demonstrated that to self-assemble into fibrils at some critical concentration c*, peptides should contain at least 5 amino acids, and should possess three or more aromatic residues on one side of the beta-sheet. The typical lengths of such resulting fibrils are at least μm in length (the longest length that can be observed is restricted by the size of the TEM grids). The critical concentration c* decreases monotonically and regularly with the peptide length *x,* see Figure. 4(a) . The width of the fibrils also decreases regularly with the peptide length *x* (Figure. 4(b)). Unambiguously determining the pitch-length, *h*, of the fibril twist proved to be difficult, as a mixture of flat quasi-crystalline aggregates and shorter, twisted morphologies are seen by TEM.

3.2 Self-assembling data analysis based on the statistical mechanical coarse-grained fibrillization model

The self-assembly data on the fraction of beta-sheet aggregated peptides f_β as a function of the total peptide concentration c can be fitted using the statistical mechanical model proposed in refs [14,27]. An appropriate asymptotic form is given by Eqn. (S1), presented in Supplementary

Information. Our experiments provide the critical concentration of peptides required for self-assembly (c*) as a function of their length *x*. By fitting the data to f_β (see Supplementary Figure 5 (a to d) for the fits and or the individual energetic contributions Table 2), it is possible to derive the total free energy gain for a peptide association into a fibril, ε_{tot} , see Table 2:

The results for ε_{tot} have a simple linear dependence on *x* that can be fitted by:

$$\varepsilon_{tot} \cong 0.6\ k_BT \times x + 13\ k_BT \qquad \text{Eq. 6}$$

To interpret this result, consider the total energy ε_{tot} as a sum of individual contributions to fibril formation. The transformation energy ε_{tr} is the energy of transformation from a random coil to a rod-like conformation, it is mainly of entropic origin and for short chains (or in so-called theta-solvents where the chain statistics is as in an ideal coil) it is proportional to the peptide length *x*. We previously estimated ε_{tr} for x=11 (P_{11}-2) to be $3.0k_BT$ (Eq. 3, [39].),giving a per-peptide dependance of:

$$\varepsilon_{tr} \approx x \times 0.27 k_BT \qquad \text{For x= 5,7,11} \qquad \text{Eq. 7}$$

Because all these peptides share a common hydrophobic core (the FWF residues, see Table 1), which dominates the tape stacking energy, we assume the interaction energies which are responsible for the tape association into double tapes are practically the same for all these peptides. Note also that $\varepsilon_p - \varepsilon_2 << \varepsilon_2$ due to the hydrophilic nature of the outer sides of the double tapes (the attraction between them is weak in water. Hence (taking into account the value of ε_2 for P_{11}-2, Eqn 3):

$\varepsilon_p \approx \varepsilon_2 \approx 0.6\ k_BT$ For x=5,7,11 Eq. 8

For x=13 however, which possess an FFWF aromatic core (which is around 33% larger than the FWF motif), we estimate this energy to be $\varepsilon_p \approx 0.8\ k_BT$.

As the length x of the peptide increases, so does the number of available hydrogen bonding interactions between adjacent rods within the β-sheets. In addition to H-bonding, the β-sheet is also stabilized by the specific bonding related to hydrophobic side-chain stacking (primarily within the FWF block). Hence, it is reasonable to assume that the corresponding β-strand interaction free energy ε is a linear function of the peptide length x:

$$\varepsilon_\beta \approx x \times \varepsilon_H + \varepsilon_{FWF} \quad \text{Eq. 9}$$

By combining Eq 2 (for ε_{tot}) with Eqs 5-8, we get the following estimates for the main inputs into the beta-sheet stabilizing energy ε_β:

$$\varepsilon_H \approx 1.15\ k_BT,\ \ \varepsilon_{FWF} \approx 14.7\ k_BT \quad \text{Eq. 10}$$

where ε_H is the typical energy gain from one hydrogen bond in a β-sheet and ε_{FWF} is the free energy associated with the hydrophobic interactions between the two blocks containing FWF (the three aromatic side groups) in adjacent peptides. A more detailed physical discussion of the contributory terms in Eq 2 is provided in Supplementary Material.

3.3 Molecular modelling of tapes and ribbons in acidic aqueous solution

Fully atomistic molecular models of the β-sheet tapes for P_5-2, P_7-2 and P_{11}-2, and ribbons made of two tapes joined by their hydrophobic lateral sides were constructed using the AMBER

modelling software. The Glu (E) residues on the peptides were fully protonated to mimic the acidic aqueous solution at pH=2. The aggregates were solvated, equilibrated and then an ensemble of representative conformers was generated by MD simulation (see Methods). The interaction energies maintaining the β-sheet assemblies were calculated to compare the energetic stability of tapes and ribbons (see Figure 5). Our results are consistent with the results of quantum mechanics calculations of Perczel *et al* who computed the thermodynamic stability of beta-peptides at the B3LYP/6-31G(d) level of theory [44]. Indeed, within their polarised-continuum model, the interaction energy difference between the two separate extended beta-stranded peptides Ala_5 and the same peptides placed into a backbone $(Ala_5)_2$ in the standard anti-parallel beta-sheet conformation was found to be -32.52 kcal/mol. Our classical 20-ns Amber MD simulations of a β-sheet tape peptides gives a consistent result for the interaction energy of this interface for peptide of a similar size, P_5-2: $E_{P5\text{-}2TP}$ = -33.87 kcal/mol (see Figure. 5a).

Figure 5 (a) and (c) show that the interaction energy holding the β-strands together within the tapes (blue bars) and the ribbons (yellow bars) increases with the length of the individual peptide strands, as it is assumed in the statistical mechanical model (Eq. 8). The interaction energy per residue is approximately constant for tapes (blue bars in Figure 5(c), but shows a slight increase with increasing peptide length for ribbons (yellow bars in Figure 5(c)). The relative contributions of the intra- and inter-sheet energetic terms in the MD simulations are compared for the three peptides in Figures. 5(b) and 5(d). Although the self-assembly from tapes to ribbons is enthalpically favorable, the energetic contribution from the hydrophobic core is considerably smaller to that from β-sheet hydrogen bonds.

To interpret these data, a detailed analysis of hydrogen bonding in the tapes and ribbons was performed for the MD ensemble of aggregates constructed from P_5-2, P_7-2 and P_{11}-2. Figure 4(c) shows the mean number of hydrogen bonds formed via the backbone and sidechain interactions. While the number of hydrogen bonds per amino acid formed by the backbone atoms in tapes and ribbons is similar in all three peptides, and so increases in proportion to the number of residues, the hydrogen bonds formed by side-chain interactions increases in ribbons with increasing peptide length. This side-chain hydrogen bonding is mainly contributed by Gln-Gln interactions, and arise so frequently between these sidechains as they are sufficiently long and flexible that they can interact across adjacent tapes within a ribbon, and with residues above and below within a β-sheet stack. The tapes composed of peptides containing only 5 and 7 amino acids show reduced levels of anti-parallel β-sheet content relative to P_{11}-2, particularly in the tapes (see Figure. 6). The disordered nature of the 5mers and 7mer tapes compared to the 11mers demonstrates that there is a level of co-operativity of the hydrogen bonding interactions as the peptide length increases, in agreement with the experimental observation that if there are too few potential interactions between the β-sheets, as is the case for the 3mer peptide, then self-assembly is suppressed.

4 CONCLUSION

This systematic study of the effect of peptide length on self-assembly of the Px-2 motif indicates that peptide sequences containing 5 or more amino acid residues undergo aggregation into fibrils under the conditions studied. On reducing peptide length, c* increases due to the decrease in the magnitude of the favourable interaction enthalpy between the individual β-strands. We also observe differences in morphology between the aggregates of peptides containing different

numbers of amino acids. In particular, the proportion of thicker fibrillar aggregates increases, as the peptides get shorter.

By using a statistical mechanical model for the thermodynamics of self-assembly, we have extracted key thermodynamic parameters by fitting to the experimental data. In accordance with Eq. 8, this procedure implies that each extra hydrogen bonding interaction (associated with extending the sequence by one amino acid) supplies an extra 1.15 kT to the free energy stabilising the interactions between the individual peptide strands in a β-sheet (as in Eq. 9). In contrast, the bare interaction energies between the β-strands extracted from simulation ($U_{\beta\text{-MD}}$) are significantly more stabilising (see Supplementary Figure 6). Fitting the extracted values to a linear dependence on peptide length x, we find for the MD calculations:

$$U_{\beta\text{-MD}} = (-12.6 \times x + 8.8)\, kT \qquad \text{Eq. 11}$$

The difference between the energies $U_{\beta\text{-MD}}$ and the free energies ε_β (see Figure S6(a)) emerges from the two different physical length-scales at which they model the interaction. The energetic parameter $U_{\beta\text{-MD}}$ at the atomistic level is a combination of all of the pair-wise van der Waals and local electrostatic energies. However, at the coarse-grained level of a rod, ε_β is a free energy that contains additional contributions not explicitly represented by the atomistic model, such as the entropy of localisation of the stacked side-chains when the β-strands self-assemble into sheets and reorganisation of the solvent environment, which includes a contribution from the counter-ion localisation entropy (long fibrils will be neutralised by Manning condensation of anions [45,46]). These entropic terms are local to the peptides at this coarse-grained level, and

contribute to the free energy of association, which is proportional to peptide length. There are also entropic penalties associated with global modes of the peptides independent of their length, such as rotation and bending.

By comparing the free energy calculated by model fitting with the enthalpies obtained from the simulation data we can obtain an estimate for the entropy penalty associated with localising a β-strand into a fibril aggregate as a function of peptide length (see Figure S6(a) in Supplementary Information). This procedure shows that there is a larger reduction in entropy for longer peptides, but that the entropy per peptide (Supplementary Figure 6(b)) converges towards a fixed value of ~11k as the length of the peptides increases. The implied number of entropic degrees of freedom per amino acid associated with assembly, (of order 10) is consistent with the restriction of bond angles and positions observed in MD simulations, in which the monomers are constrained into aggregates. In contrast to the length-independent terms in the entropy penalty, favourable enthalpic attractions increase with peptide length. So the compensating enthalpy becomes increasingly dominant in the overall free energy change during aggregation for longer peptide strands, than for shorter. This also explains why the molecular structures observed in the MD contain more disorder for the shorter strands – the reduced available compensating enthalpy restricts the degree of ordering (and its entropic cost) possible.

By using a combination of statistical mechanics and atomistic computer simulation, we have been able to provide a structural and thermodynamic rationale for the experimental data obtained on peptide self-assembly as a function of length. The entropy/enthalpy compensation has a

subtle but important dependence on chain length that determines both the ability of a peptide of a given length to form aggregates, and the degree of ordering present in the assemblies that form. This work validates the coarse-grained model designed originally for a single peptide to a much broader range of amino acid sequences, and combined with atomistic simulation can provide a general multi-scale method for rational design of self-assembling nanomaterials.

5 ASSOCIATED CONTENT

Details on the calculation of assembled monomers and bond volume estimates along with Figures S1-S6 and Table S1 are provided as supporting information.

REFERENCES

1 Iconomidou VA, Hamodrakas SJ. Natural Protective Amyloids. *Curr. Protein Pept. Sci.* 2008; **9**: 291–309; DOI:10.2174/138920308784534041.

2 Morgan C, Colombres M, Nunez MT, Inestrosa NC. Structure and Function of Amyloid in Alzheimer's Disease. *Prog. Neurobiol.* 2004; **74**: 323–349; DOI:10.1016/j.pneurobio.2004.10.004.

3 Jiang D, Rauda I, Han S, Chen S, Zhou F. Aggregation Pathways of the Amyloid β (1–42) Peptide Depend on Its Colloidal Stability and Ordered β-Sheet Stacking. *Langmuir* 2012; **28**: 12711–12721.

4 Knowles TPJ, Buehler MJ. Nanomechanics of Functional and Pathological Amyloid Materials. *Nat. Nanotechnol.* 2011; **6**: 469–479.

5 Wong S, Shim MS, Kwon YJ. Synthetically Designed Peptide-Based Biomaterials with Stimuli-Responsive and Membrane-Active Properties for Biomedical Applications. *J. Mater. Chem. B* 2014; **2**: 595–615.

6 Webber MJ, Kessler JA, Stupp SI. Emerging Peptide Nanomedicine to Regenerate Tissues and Organs. *J. Intern. Med.* 2010; **267**: 71–88; DOI:10.1111/j.1365-2796.2009.02184.x.

7 Yemini M, Reches M, Rishpon J, Gazit E. Novel Electrochemical Biosensing Platform Using Self-Assembled Peptide Nanotubes. *Nano Lett.* 2005; **5**: 183–186; DOI:10.1021/nl0484189.

8 Scanlon S, Aggeli A. Self-Assembling Peptide Nanotubes. *Nano Today* 2008; **3**: 22–30;

DOI:10.1016/s1748-0132(08)70041-0.

9 Rajagopal K, Lamm MS, Haines-Butterick LA, Pochan DJ, Schneider JP. Tuning the pH Responsiveness of Beta-Hairpin Peptide Folding, Self-Assembly, and Hydrogel Material Formation. *Biomacromolecules* 2009; **10**: 2619–2625; DOI:10.1021/bm900544e.

10 Zhou M, Smith AM, Das AK, Hodson NW, Collins RF, Ulijn R V, Gough JE. Self-Assembled Peptide-Based Hydrogels as Scaffolds for Anchorage-Dependent Cells. *Biomaterials* 2009; **30**: 2523–2530; DOI:10.1016/j.biomaterials.2009.01.010.

11 Brunton P a, Davies RPW, Burke JL, Smith a, Aggeli a, Brookes SJ, Kirkham J. Treatment of Early Caries Lesions Using Biomimetic Self-Assembling Peptides--a Clinical Safety Trial. *Br. Dent. J.* 2013; **215**: E6; DOI:10.1038/sj.bdj.2013.741.

12 Aggeli A, Bell M, Boden N, Keen JN, Knowles PF, McLeish TCB, Pitkeathly M, Radford SE. Responsive Gels Formed by the Spontaneous Self-Assembly of Peptides into Polymeric Beta-Sheet Tapes. *Nature* 1997; **386**: 259–262; DOI:10.1038/386259a0.

13 Aggeli A, Bell M, Boden N, Keen JN, McLeish TCB, Nyrkova I, Radford SE, Semenov A. Engineering of Peptide Beta-Sheet Nanotapes. *J. Mater. Chem.* 1997; **7**: 1135–1145; DOI:10.1039/a701088e.

14 Aggeli A, Nyrkova IA, Be[1] A. Aggeli, I.A. Nyrkova, M. Bell, R. Harding, L. Carrick, T.C.B. McLeish, A.N. Semenov, N. Boden, Hierarchical self-assembly of chiral rod-like molecules as a model for peptide β-sheet tapes, ribbons, fibrils, and fibers, Proc. Natl. Acad. Sci. 98 (M, Harding R, Carrick L, McLeish TCB, Semenov AN, Boden N. Hierarchical Self-Assembly of Chiral Rod-like Molecules as a Model for Peptide β-Sheet Tapes, Ribbons, Fibrils, and Fibers. *Proc. Natl. Acad. Sci.* 2001; **98**: 11857–11862.

15 Davies RPW, Aggeli a., Beevers a. J, Boden N, Carrick LM, Fishwick CWG, Mcleish TCB, Nyrkova I, Semenov a. N. Self-Assembling β-Sheet Tape Forming Peptides. *Supramol. Chem.* 2006; **18**: 435–443; DOI:10.1080/10610270600665855.

16 Davies RPW, Aggeli A. Self-Assembly of Amphiphilic Beta-Sheet Peptide Tapes Based on Aliphatic Side Chains. *J. Pept. Sci.* 2011; **17**: 107–114; DOI:10.1002/psc.1335.

17 Meegan JE, Aggeli a, Boden N, Brydson R, Brown a P, Carrick L, Brough a R, Hussain a, Ansell RJ. Designed Self-Assembled Beta-Sheet Peptide Fibrils as Templates for Silica Nanotubes. *Adv. Funct. Mater.* 2004; **14**: 31–37; DOI:10.1002/adfm.200304477.

18 Bell CJ, Carrick LM, Katta J, Jin Z, Ingham E, Aggeli A, Boden N, Waigh TA, Fisher J. Self-assembling Peptides as Injectable Lubricants for Osteoarthritis. *J. Biomed. Mater. Res. Part A* 2006; **78**: 236–246.

19 Kirkham J, Firth A, Vernals D, Boden N, Robinson C, Shore RC, Brookes SJ, Aggeli A. Self-Assembling Peptide Scaffolds Promote Enamel Remineralization. *J. Dent. Res.* 2007; **86**: 426–430; DOI:10.1177/154405910708600507.

20 Kyle S, Felton SH, McPherson MJ, Aggeli A, Ingham E. Rational Molecular Design of Complementary Self-Assembling Peptide Hydrogels. *Adv. Healthc. Mater.* 2012; **1**: 640–645.

21 Maude S, Miles DE, Felton SH, Ingram J, Carrick LM, Wilcox RK, Ingham E, Aggeli A. De Novo Designed Positively Charged Tape-Forming Peptides: Self-Assembly and Gelation in Physiological Solutions and Their Evaluation as 3D Matrices for Cell Growth. *Soft Matter* 2011; **7**: 8085–8099; DOI:10.1039/c0sm00974a.

22 Guilliatt R, Davies R, Wilshaw SP, Aggeli A, Ingham E. Biocompatibility and Haemocompatibility of Self-Assembling Peptides. *J. Tissue Eng. Regen. Med.* 2012; **6**: 218–219.

23 Whitehouse C, Fang J, Aggeli A, Bell M, Brydson R, Fishwick CWG, Henderson JR, Knobler CM, Owens RW, Thomson NH. Adsorption and Self-Assembly of Peptides on Mica Substrates. *Angew. Chemie* 2005; **117**: 2001–2004.

24 Carrick L, Tassieri M, Waigh TA, Aggeli A, Boden N, Bell C, Fisher J, Ingham E, Evans RML. The Internal Dynamic Modes of Charged Self-Assembled Peptide Fibrils. *Langmuir* 2005; **21**: 3733–3737.

25 Nyrkova I a, Semenov a N, Aggeli a, Boden N. Fibril Stability in Solutions of Twisted Beta-Sheet Peptides: A New Kind of Micellization in Chiral Systems. *Eur. Phys. J. B* 2000; **17**: 481–497; DOI:10.1007/s100510070127.

26 Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. HOW TO MEASURE AND PREDICT THE MOLAR ABSORPTION-COEFFICIENT OF A PROTEIN. *Protein Sci.* 1995; **4**: 2411–2423.

27 Nyrkova IA, Semenov AN, Aggeli A, Bell M, Boden N, McLeish TCB. Self-Assembly and Structure Transformations in Living Polymers Forming Fibrils. *Eur. Phys. J. B* 2000; **17**: 499–513; DOI:10.1007/s100510070128.

28 Ghosh K, Dill KA. Theory for Protein Folding Cooperativity: Helix Bundles. *J. Am. Chem. Soc.* 2009; **131**: 2306–2312; DOI:10.1021/ja808136x.

29 van Raaij ME, van Gestel J, Segers-Nolten IMJ, de Leeuw SW, Subramaniam V. Concentration Dependence of α-Synuclein Fibril Length Assessed by Quantitative Atomic Force Microscopy and Statistical-Mechanical Theory. *Biophys. J.* 2008; **95**: 4871–4878; DOI:10.1529/biophysj.107.127464.

30 Schreck JS, Yuan J. A Statistical Mechanical Approach to Protein Aggregation. *J. Chem. Phys.* 2011; **135**: 235102; DOI:10.1063/1.3666837.

31 Macke TJ, Case DA. Modeling Unusual Nucleic Acid Structures. *Am. Chem. Soc* 1998; 379–393.

32 Lovell SC, Word JM, Richardson JS, Richardson DC. The Penultimate Rotamer Library. *Proteins-Structure Funct. Genet.* 2000; **40**: 389–408; DOI:10.1002/1097-0134(20000815)40:3<389::aid-prot50>3.0.co;2-2.

33 Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An Environment for Comparative Protein Modeling. *Electrophoresis* 1997; **18**: 2714–2723; DOI:10.1002/elps.1150181505.

34 Krivov GG, Shapovalov M V, Dunbrack RL. Improved Prediction of Protein Side-Chain Conformations with SCWRL4. *Proteins-Structure Funct. Bioinforma.* 2009; **77**: 778–795; DOI:10.1002/prot.22488.

35 Case DA, Darden TA, Cheatham III TE, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM. AMBER 12; University of California: San Francisco, 2012. *There is no Corresp. Rec. this Ref.*

36 Duan Y, Wu C, Chowdhury S, Lee MC, Xiong GM, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang JM, Kollman P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* 2003; **24**: 1999–2012; DOI:10.1002/jcc.10349.

37 Shields GC, Laughton CA, Orozco M. Molecular Dynamics Simulations of the d(T Center Dot A Center Dot T) Triple Helix. *J. Am. Chem. Soc.* 1997; **119**: 7463–7469; DOI:10.1021/ja970601z.

38 Hawkins GD, Cramer CJ, Truhlar DG. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* 1996; **100**: 19824–19839; DOI:10.1021/jp961710n.

39 Tsui V, Case DA. Theory and Applications of the Generalized Born Solvation Model in Macromolecular Simulations. *Biopolymers* 2001; **56**: 275–291.

40 Berryman JT, Radford SE, Harris SA. Thermodynamic Description of Polymorphism in Q- and N-Rich Peptide Aggregates Revealed by Atomistic Simulation. *Biophys. J.* 2009; **97**: 1–11; DOI:10.1016/j.bpj.2009.03.062.

41 Hills RD, Brooks CL. Hydrophobic Cooperativity as a Mechanism for Amyloid Nucleation. *J. Mol. Biol.* 2007; **368**: 894–901; DOI:10.1016/j.jmb.2007.02.043.

42 Zheng J, Jang H, Ma B, Tsai CJ, Nussinov R. Modeling the Alzheimer A beta(17-42) Fibril Architecture: Tight Intermolecular Sheet-Sheet Association and Intramolecular Hydrated Cavities. *Biophys. J.* 2007; **93**: 3046–3057; DOI:10.1529/biophysj.107.110700.

43 Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. *J. Mol. Graph. Model.* 1996; **14**: 33–38; DOI:10.1016/0263-7855(96)00018-5.

44 Perczel A, Hudaky P, Palfi VK. Dead-End Street of Protein Folding: Thermodynamic Rationale of Amyloid Fibril Formation. *J. Am. Chem. Soc.* 2007; **129**: 14959–14965; DOI:10.1021/ja0747122.

45 Manning GS. The Molecular Theory of Polyelectrolyte Solutions with Applications to

the Electrostatic Properties of Polynucleotides. *Q. Rev. Biophys.* 1978; **11**: 179–246; DOI:10.1017/S0033583500002031.

46 Oosawa F. Polyelectrolytes. In *Polyelectrolytes*. Marcel Dekker, 1971.

ATHOR INFORMATION

Corresponding Author

*E-mail: S.A.Harris@leeds.ac.uk

Funding

Funding for this work was provided by the Dutch Polymer Institute (DPI/26102006). The work was also partly supported by the Leeds Centre of Excellence in Medical Engineering funded by the Wellcome Trust and EPSRC, WT088908/z/09/z.

ACKNOLEDGMENTS

We would like to gratefully acknowledge Dr A. Aggeli for thoughtful discussion.

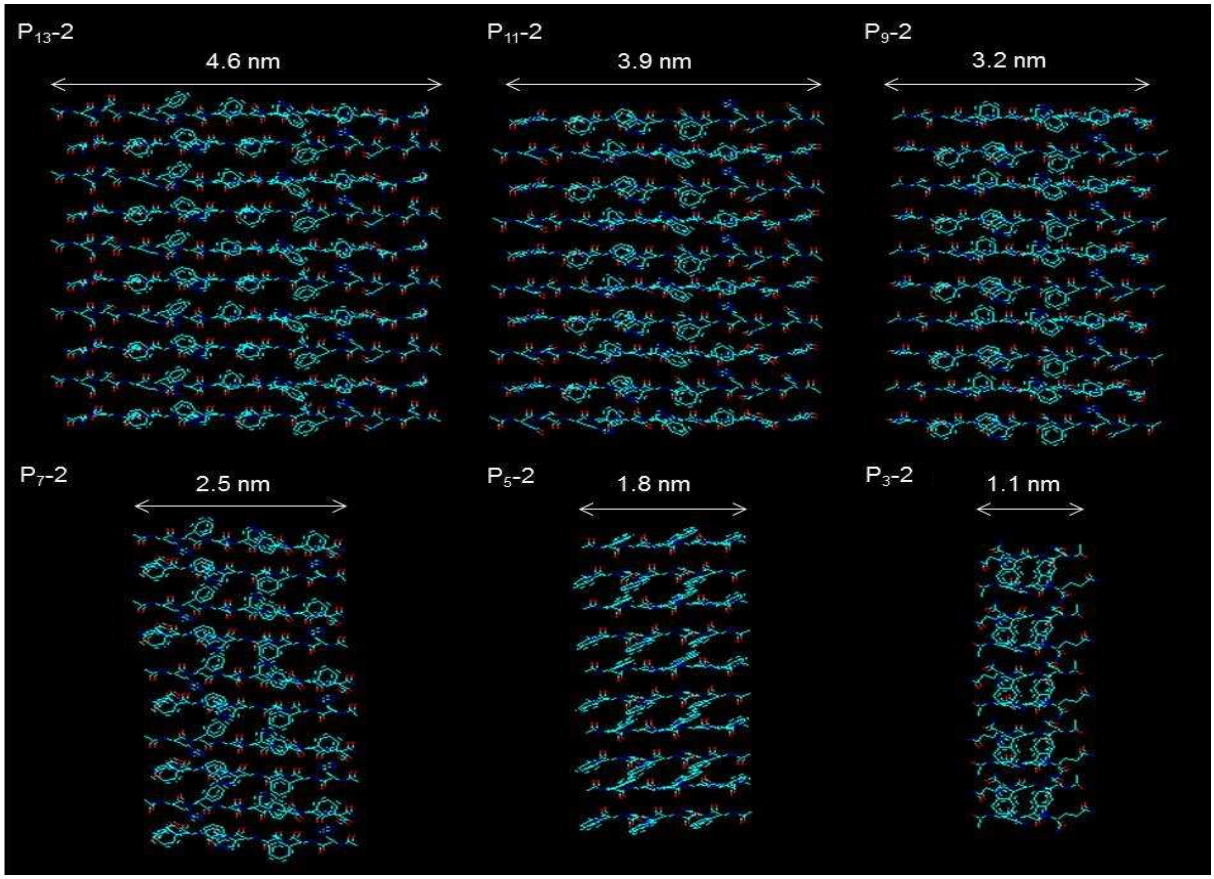


Figure 1. Molecular models of P_{13}-2, P_{11}-2, P_9-2, P_5-2 and P_3-2 β-sheet tapes, indicating the decrease in tape width with decrease of peptide and β-strand length; peptides are represented as individual β-strands arranged in an antiparallel conformation.

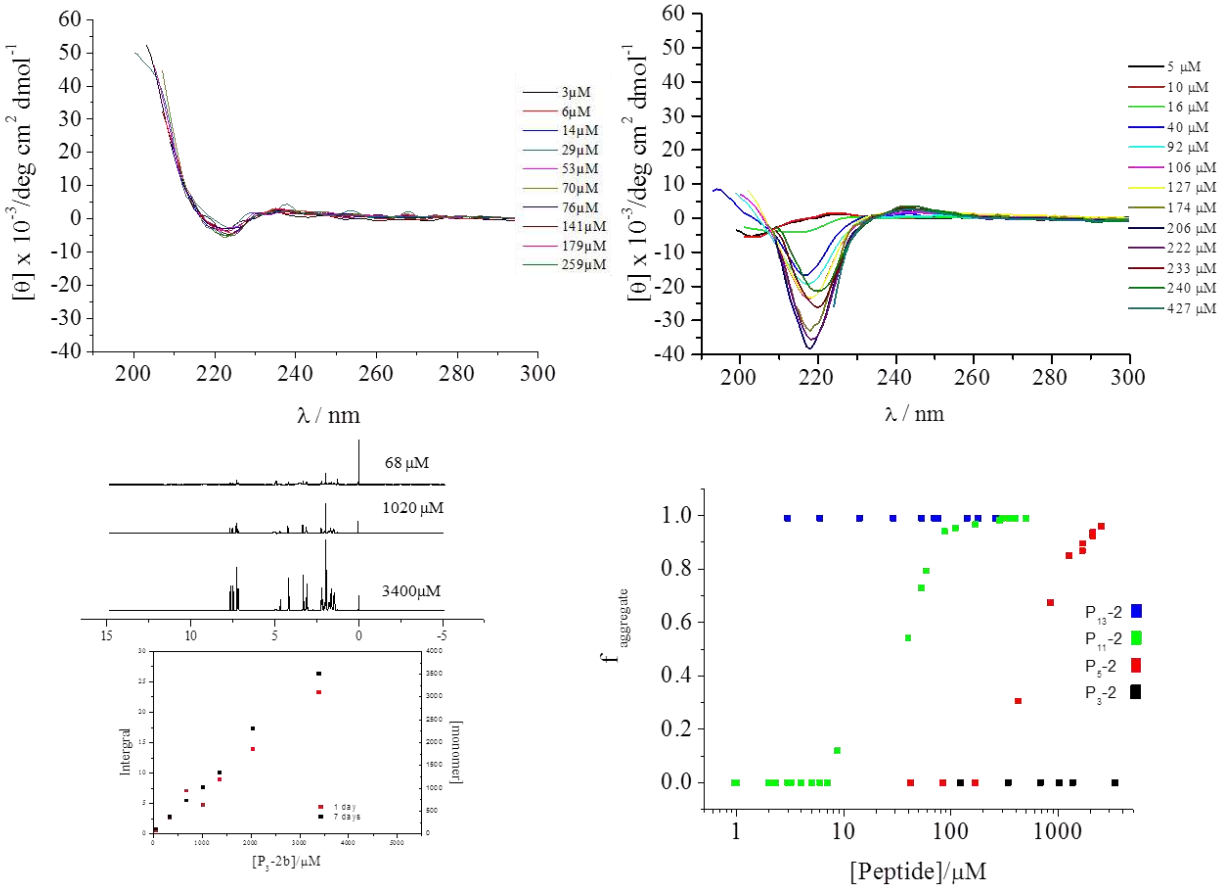


Figure 2. Self-assembly of P_{13}-2, P_9-2 and P_3-2 in water at pH 2: (a) CD spectra for P_{13}-2 obtained 2 days after sample preparation, (b) CD spectra of P_9-2 obtained 2 days after sample preparation (c) High resolution ^{1}H NMR spectra of P_3-2b at increasing peptide concentration and the corresponding graph of NMR integrals of aromatic peaks vs concentration; (d) Self-assembly curves showing fraction of peptide in β-sheet aggregate vs peptide concentration in water at pH 2 .

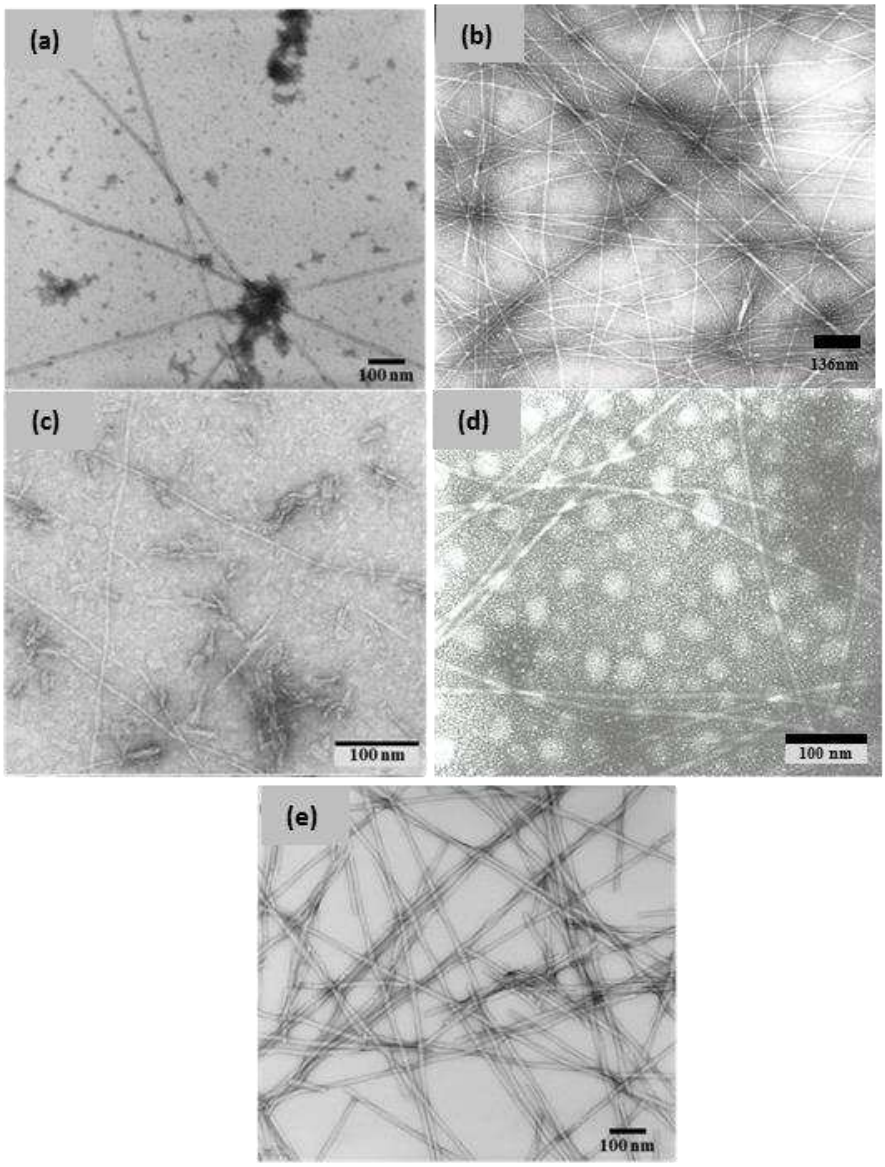


Figure 3. Negatively stained transmission electron micrographs of peptide nanostructures in water: (a) P_{13}-2 at 256 μM pH 2; (b) P_{11}-2 at 286 μM pH 2.(c) P_9-2 at 350 μM pH 2; (d) P_7-1 at c = 4.6 μM in H_3PO_4 pH 2. (e) P_5-2 at 2100 μM pD 2

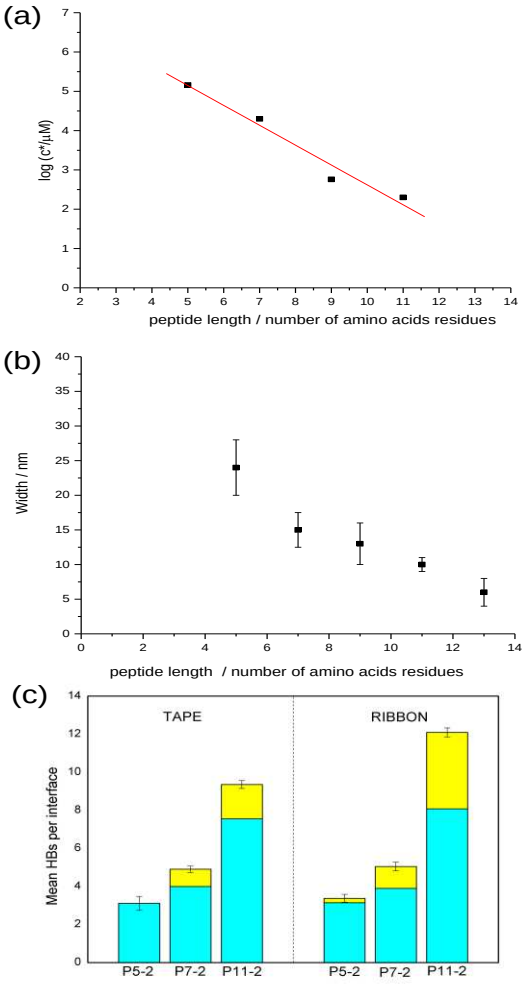


Figure 4. (a) The natural log of the critical concentration for self-assembly vs peptide length in water at pH 2; (b) Width of peptide fibrils vs peptide length; (c) Number of hydrogen bonding interactions formed by backbone (in blue) and sidechain atoms (in yellow) for tapes and ribbons.

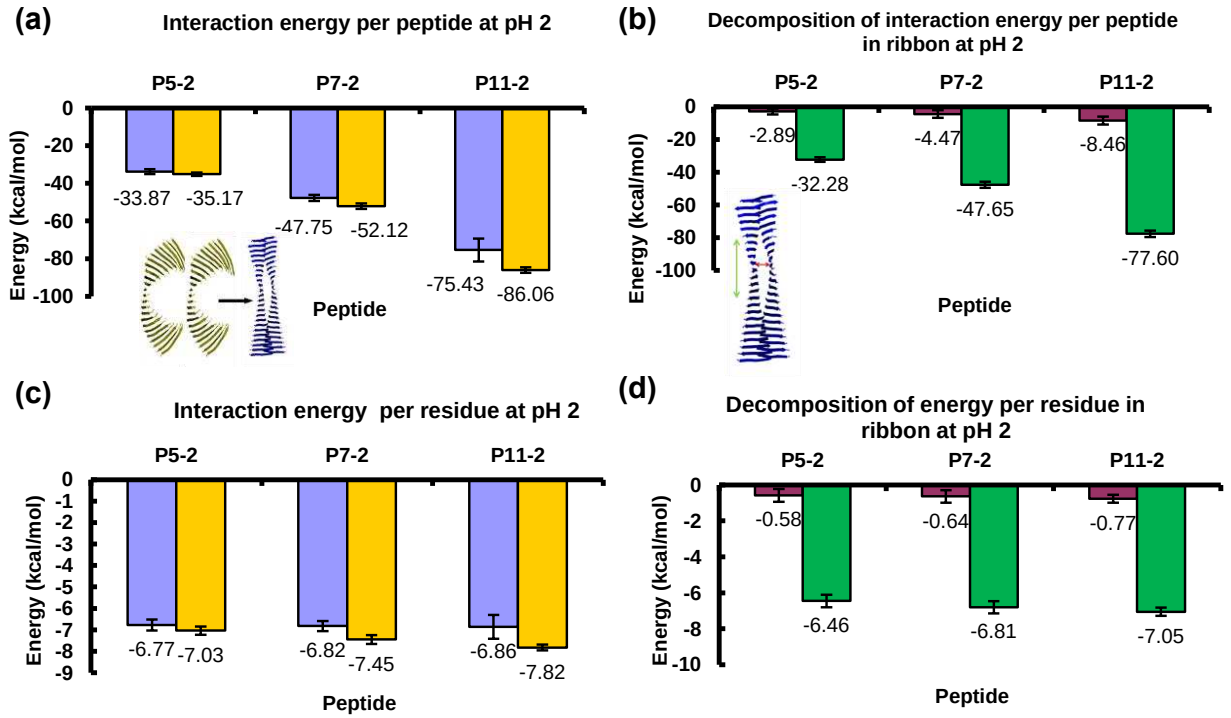


Figure 5. Interaction energies of tapes (indicated by yellow bars) and ribbons (blue bars) *per peptide* (a) and *per residue* (c) at pH 2 and decomposition of enthalpies of ribbons along the fibril axis (green bars) and across the pair of stacked β-sheets (red bars) *per peptide* (b) and *per residue* (d) at pH 2.

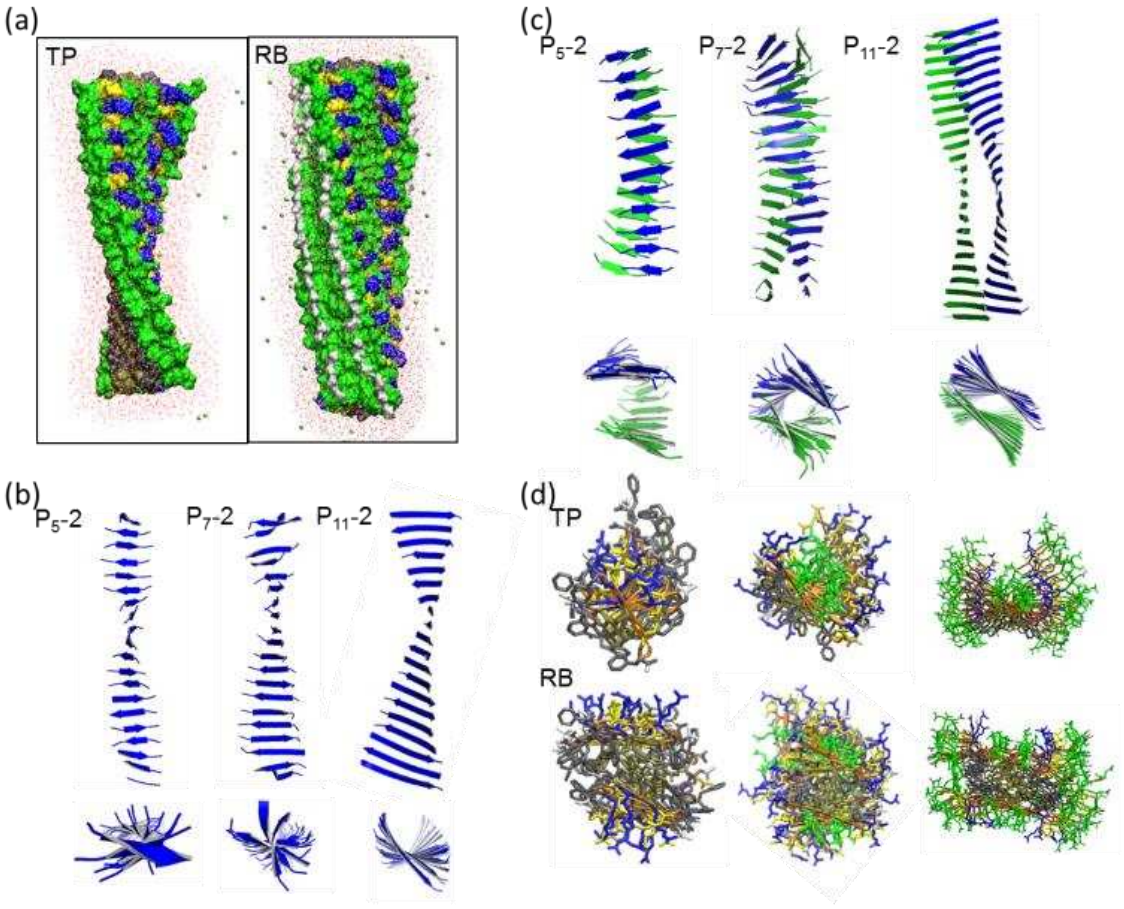
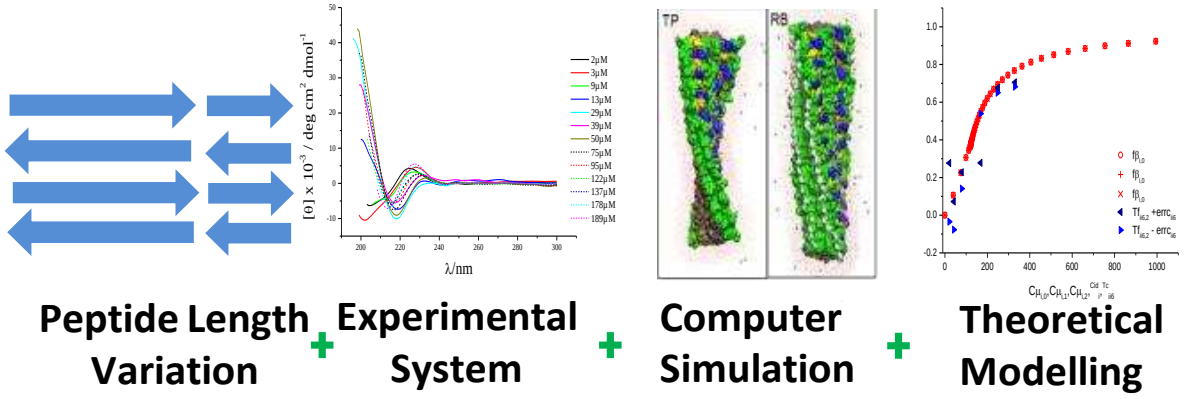


Figure 6. a) Final structure of P_{11}-2 tapes (TP) and ribbons (RB) in molecular dynamic simulations; b) side-view (top) and bottom-view (bottom) of TP conformations from last snapshots of P_5-2, P_7-2 and P_{11}-2; c) side-view (top) and bottom-view (bottom) of RB conformations from last snapshots of P_5-2, P_7-2 and P_{11}-2; d) bottom view of atomic details of P_5-2, P_7-2 and P_{11}-2 in TP and RB.



Graphical abstract

Supplementary Data

Fraction of assembled monomers: If ε_{tr} is high enough (this condition depends on other parameters of the system (see eqs.(28,36-38) in ref.[26]) but in practice is always valid for transitions from monomers to multi-tapes (fibrils), e.g. for peptides made of x≥5 aminoacids with strong tendency to double tape formation due to hydrophobicity of one side of the beta-sheet), the fraction of aggregated peptides f_β follows an asymptotic pattern everywhere apart from a narrow region near the critical concentration c* (see eqs.(35,42) in ref. [26]):

$$f_\beta \cong 2 \exp[(\varepsilon_\beta - 2\varepsilon_{tr})/k_BT]\, v_\beta c << 1 \quad \text{for } c<c^*-\Delta c, \quad f_\beta \cong 1 - c^*/c \quad \text{for } c>c^*+\Delta c \qquad \text{Eq.(S1)}$$

$$c^* v_\beta \cong \exp[(\varepsilon_{tr} - \varepsilon_\beta - \varepsilon_p)/k_BT] \qquad \text{Eq.(S2)}$$

For $c<c^*-\Delta c$ the peptide solution is dominated by monomers with a small fraction cf_β of dimers. For $c>c^*+\Delta c$ concentration of monomers saturates at $c \cong c^*$ whereas all extra peptide material ($c-c^*$) goes to fibrils.

Details of bond volume estimate: Dependence of the bond volume v_β on the length x is less straight-forward. However, it is possible to follow the line of ref.[26] for the estimation of v_β: we assume that adjacent rods are connected by x identical elastic springs corresponding to the number of H-bonds. The result is

$$v_\beta \sim k_1 (x^{-0.5} + k_2)^5 ((x+x_0)^{-1} + k_3)^2 \qquad \text{Eq.(S3)}$$

where k_i are constants. The factor $(x^{-0.5} + k_2)^3$ comes from the rod's center-of-mass slipping fluctuations (k_2 reflects the fact that the rod itself cannot be considered as completely rigid for $x>k_2^{-2}$). The factor $(x^{-0.5} + k_2)^2 / x^2$ comes from similar angular fluctuations, however in order to exclude the divergence at small x, in eq.(D) the factor $1/x^2$ is replaced by $((x+x_0)^{-1} + k_3)^2$. We

expect that $x_0 \sim 2$, $k_2 \sim 0.3$, $k_3 \sim 0.1$. The factor $k_1 \sim 2 \times 10^{-3}$ nm^3 can estimated from the known value $v_\beta \sim 5 \times 10^{-6}$ nm^3 for P_{11}-2 peptide (cf.Eq.(A)), For x=5-11 the bond fluctuation volume eq.(D) can be approximated linearly as

$$\ln(v_\beta/v_0) \approx - 2.3 - x \times 0.28 \qquad \text{Eq.(S4)}$$

Figure 1S. Self-assembling properties of P_{13}-2, P_7-2 and P_5-2 in methanol: (a) CD spectra for P_{13}-2 and the corresponding graph of molar ellipticity at 220 nm vs concentrations after 2 and 5 days from sample preparation; (b) CD spectra of P_7-2 and corresponding graph of molar ellipticity at 218 nm vs concentration; (c) High resolution ^{1}H NMR spectra of P_5-2 at increasing concentration (42μM, 336μM and 2100μM) and corresponding graph of aromatic peak integrals vs concentration; (d) Self-assembly curves showing fraction of peptide in β-sheet aggregate vs peptide concentration in methanol for P_{13}-2, P_9-2, P_7-2, P_5-2 and P_3-2.

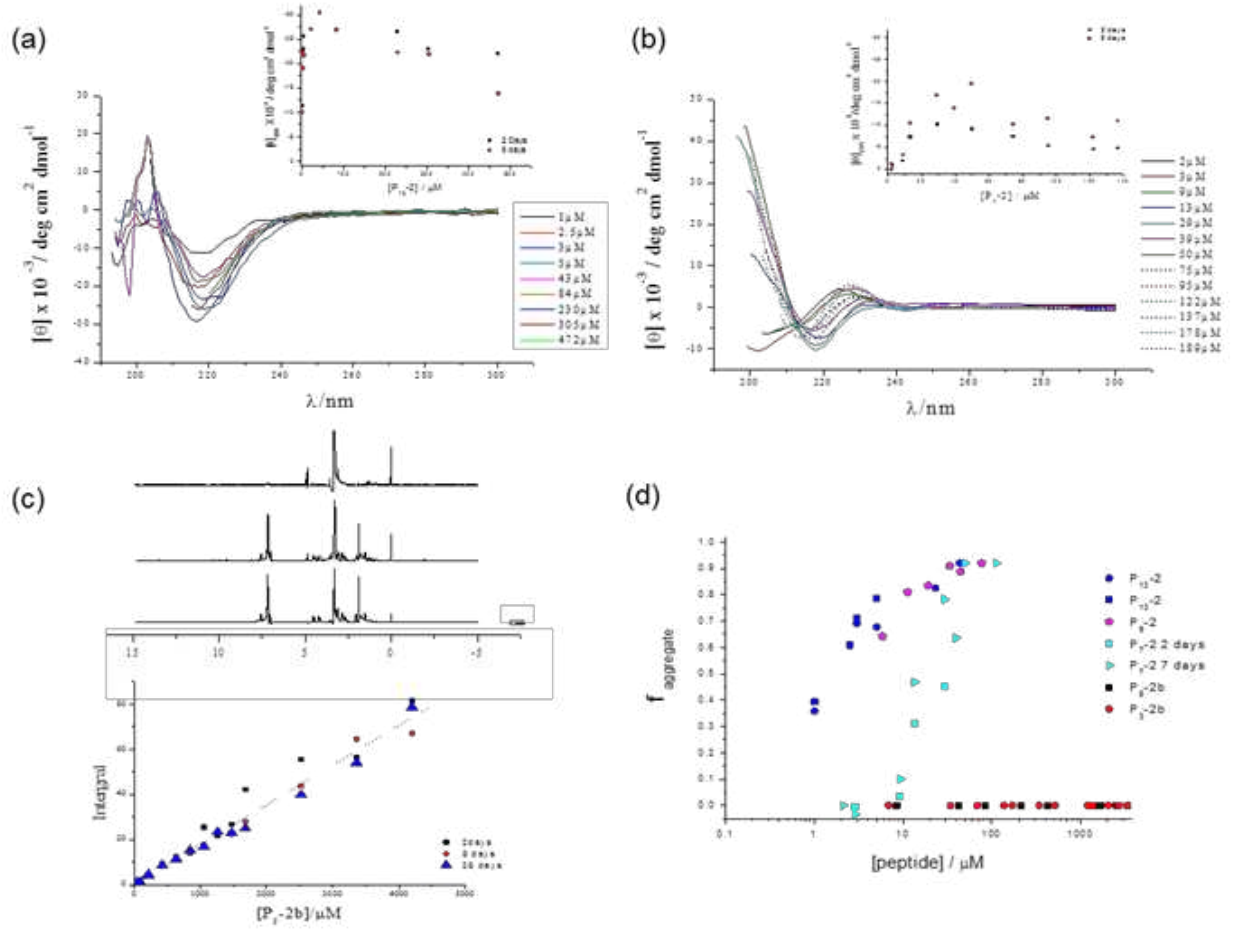


Figure 2S. Negatively stained transmission electron micrographs of peptide nanostructures in methanol: (a) P_{13}-2 at 80μM; (b) P_{11}-2 at 83 μM; (c) P_9-2 at 64μM; (d) P_7-2 at 40 μM.

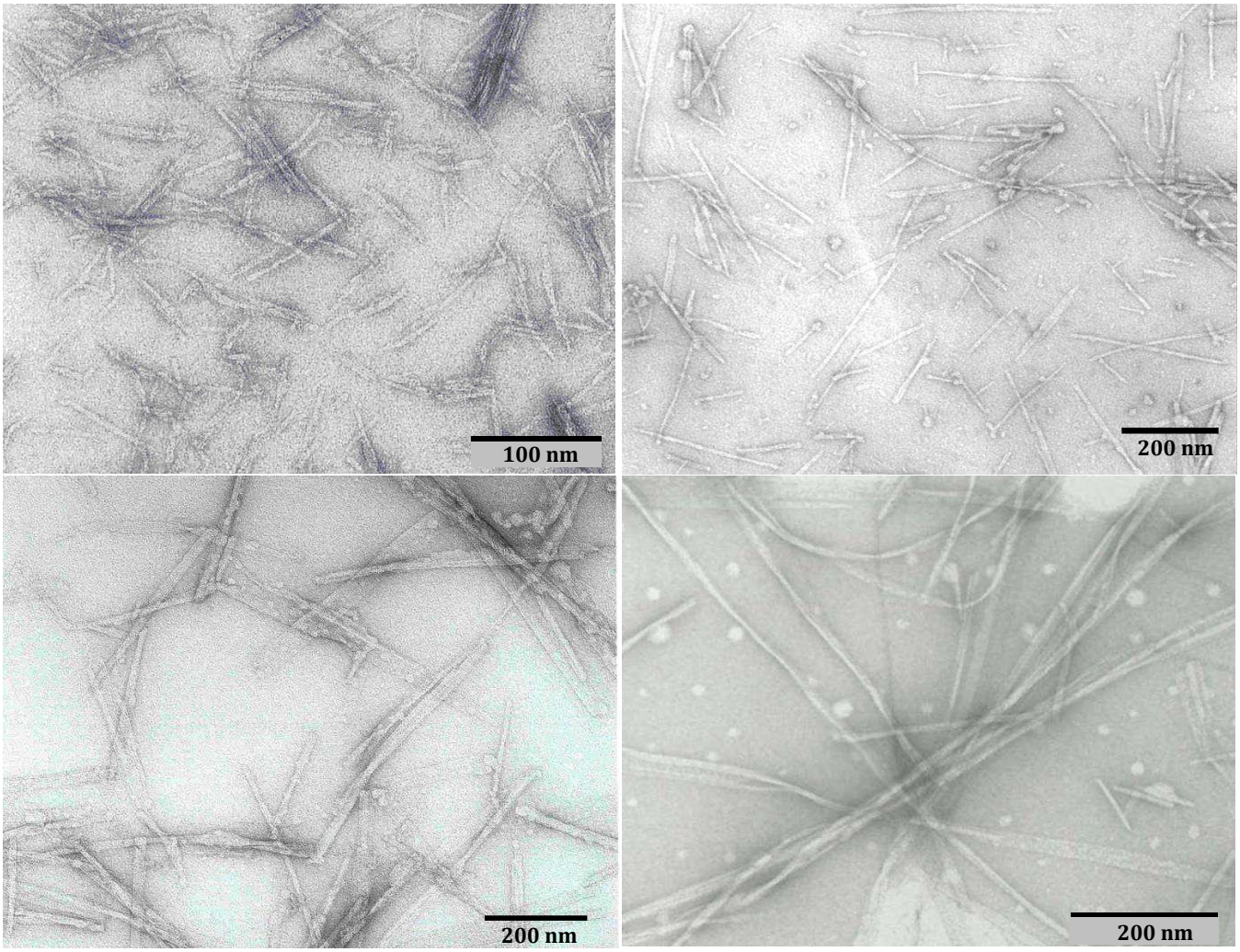


Figure 3S. Cooperativity of tape (a) and ribbon (b) enthalpies at pH 2; (c and d) cooperativity of decomposed ribbon enthalpies.

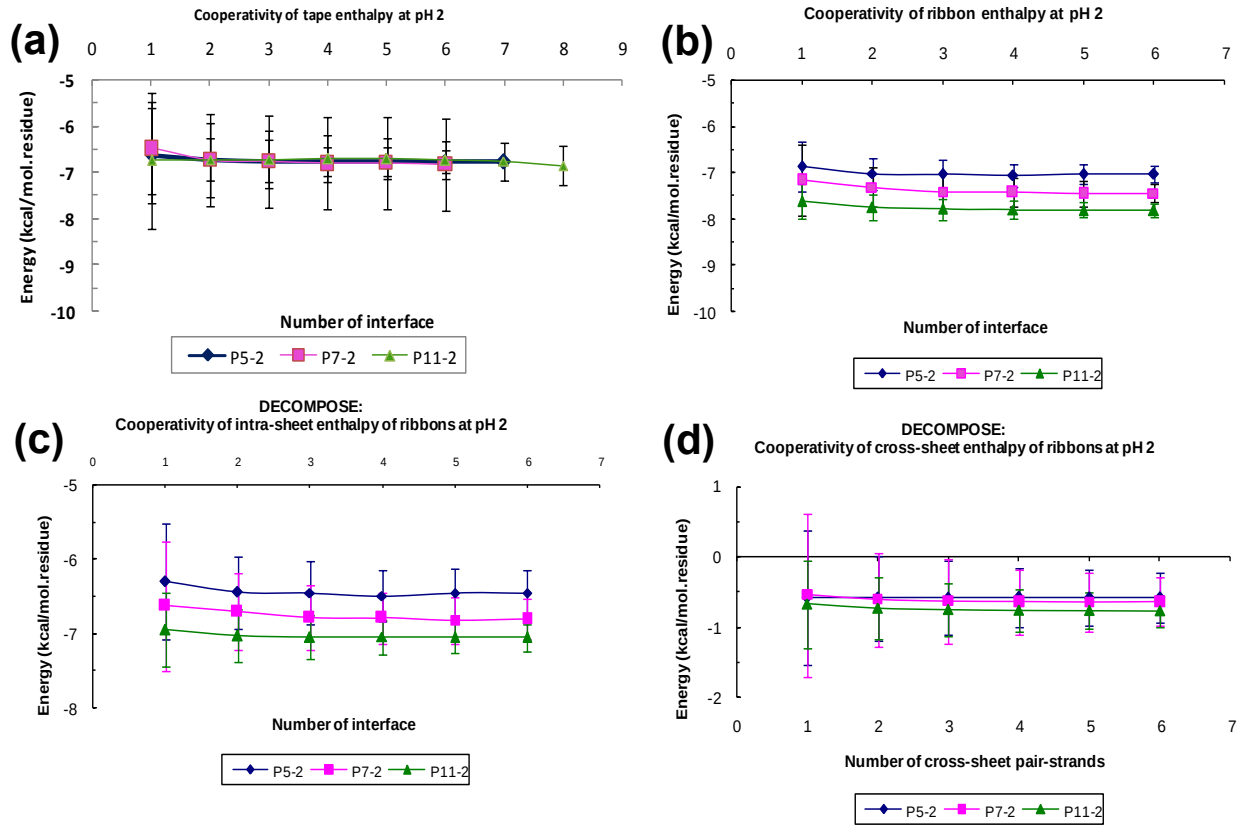


Figure 4S. Histogram of single β-strand enthalpies of (A) P_5-2 in tape; (B) P_5-2 in ribbon; (C) P_7-2 in tape; (D) P_7-2 in ribbon; (E) P_{11}-2 in tape; (F) P_{11}-2 in ribbon.

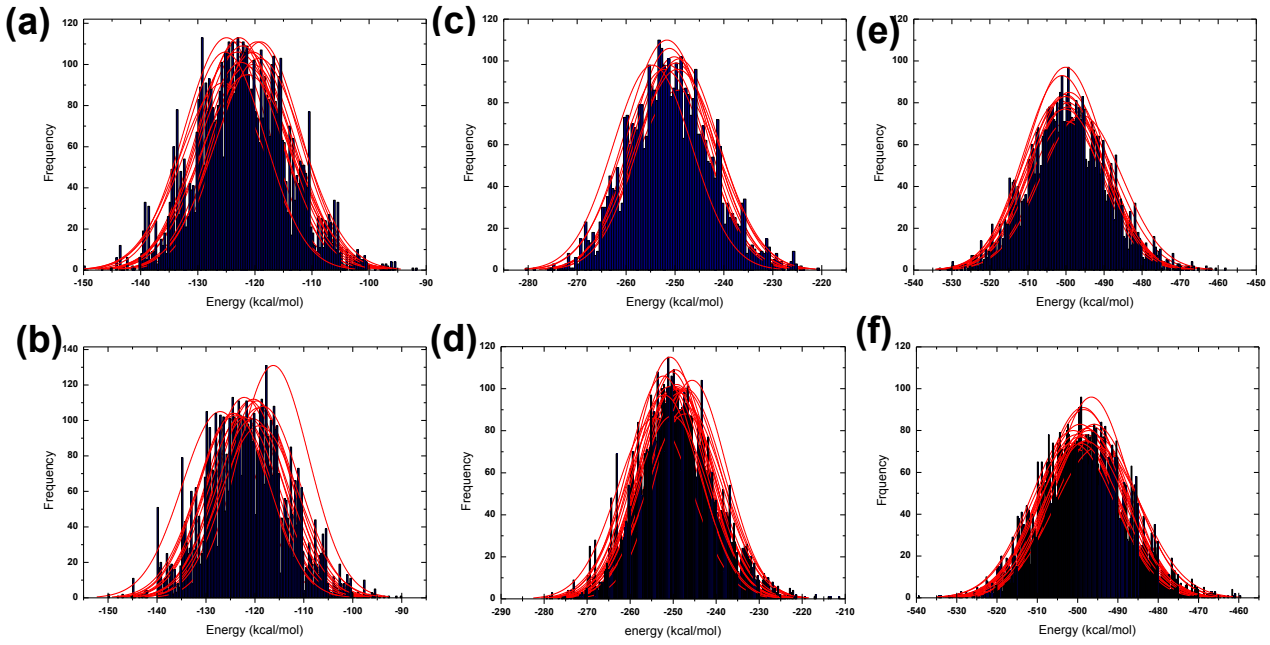


Figure 5S. a) Theoretical fit of P_{11}-5 using a C* of 170 μM, p=19, ε_{tr}= 13.5, ε_{p}=0.6 b) Theoretical fit of P_{11}-7 using a C* of 76 μM, p=12, ε_{tr}= 1.9, ε_{p}=0.6 c) Theoretical fit of P_{11}-9 using a C* of 14 μM, p=10, ε_{tr}= 2.4, ε_{p}=0.6 d) Theoretical fit of P_{11}-11 using a C* of 4.5 μM, p=8, ε_{tr}= 3, ε_{p}=0.6

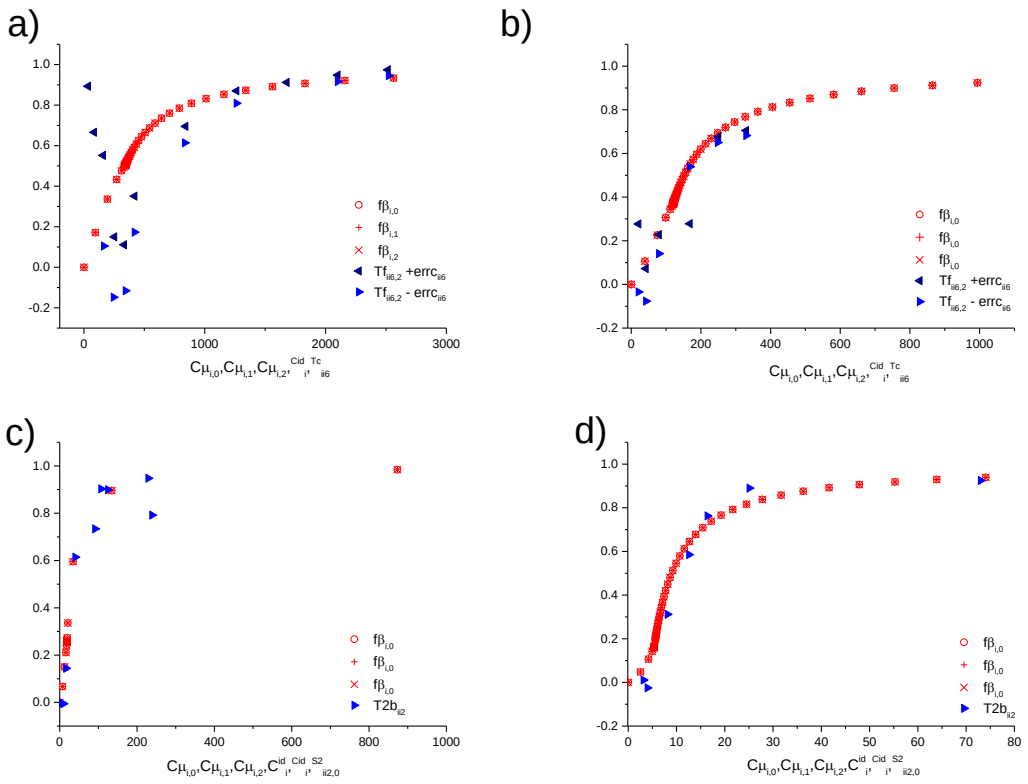
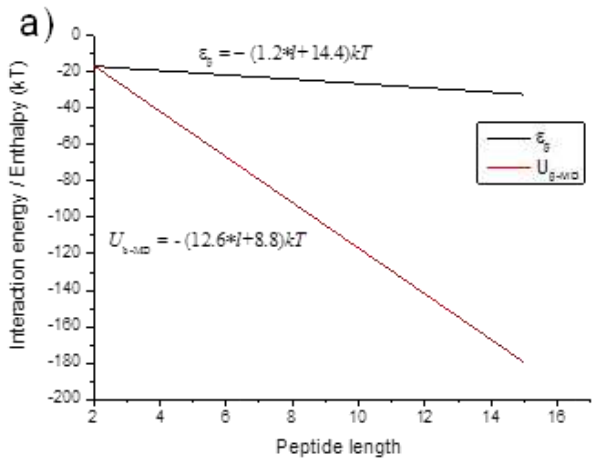


Figure 6S. (a). Free energy (red) and enthalpy (black) obtained from model fitting to the experimental data and MD simulations respectively. (b). Effective entropy per amino acid as a function of peptide length.


a)
Interaction energy / Enthalpy (kT)
Peptide length
$\varepsilon_\beta = -(1.2*l+14.4)kT$
$U_{\beta\text{-sh}} = -(12.6*l+8.8)kT$
ε_β
$U_{\beta\text{-sh}}$

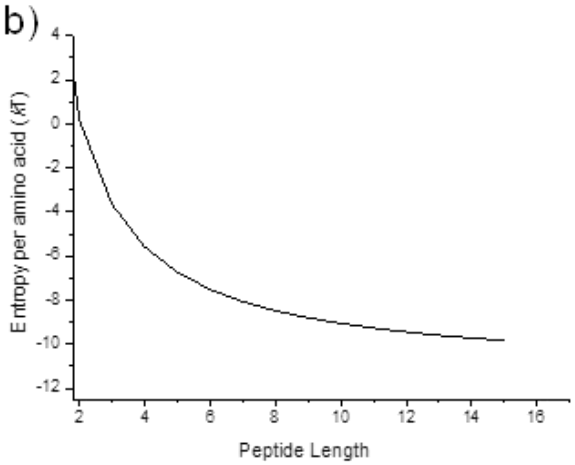

b)
Entropy per amino acid (kT)
Peptide Length

Table 1S. Manufacturing details of peptides, including the peptide purity by HPLC and the net peptide content which is an average of results from amino acid analysis, UV concentration determination and elemental analysis (from the nitrogen content)

Peptide	Manufactured	Theoretical mass / Da	Purity (%)	Net Peptide content (%)
P_{13}-2	In house	1334.4	80.0	85.0
P_{11}-2	In house	1593.7	96.1	84.4
P_9-2	NeoMPS	1334.4	96.1	92.2
P_7-2	NeoMPS	1080.0	97.6	86.5
P_5-2	NeoMPS	823.9	99.2	98.0
P_3-2	GenScript	530.0	97.7	66.7