The Conformation and Aggregation of Proline-Rich Surfactant-Like Peptides

Ian W. Hamley,[a,*] *Valeria Castelletto,*[b] *Ashkan Dehsorkhi,*

Dept of Chemistry, University of Reading, Whiteknights Reading RG6 6AD, UK

Juan Torras,[c] *Carlos Aleman,*[d,*]

Departament d'Enginyeria Química (EEBE) and Barcelona Research Center for Multiscale Science and Engineering, Universitat Politècnica de Catalunya, C/ Eduard Maristany, 10-14, Ed. I2, 08019, Barcelona, Spain

Irina Portnaya and Dganit Danino

Department of Biotechnology and Food Engineering and the
Russell Berrie Nanotechnology Institute, Technion, Haifa, Israel 32000

Revision. For J. Phys. Chem. B

* Authors for correspondence. E-mail: I.W.Hamley@reading.ac.uk; carlos.aleman@upc.edu

[a] ORCID id 0000-00024549-0926
[b] ORCID id 0000-0002-3705-0162
[c] ORCID id 0000-0001-8737-7609
[d] ORCID id 0000-0003-4462-6075

Abstract

The secondary structure of proline rich surfactant-like peptides is examined for the first time and is found to be influenced by charged end groups in peptides P_6K, P_6E, KP_6E and an equimolar mixture of P_6K and P_6E. The peptides exhibit a conformational transition from unordered to polyproline II (PPII) above a critical concentration, detected from circular dichroism (CD) measurements and unexpectedly from fluorescence dye probe measurements. Isothermal titration calorimetry (ITC) measurements provided the Gibbs energies of hydration of P_6K and P_6E, which correspond essentially to the hydration energies of the terminal charged residues. A detailed analysis of peptide conformation for these peptides was performed using density functional theory calculations, and this was used as a basis for hybrid quantum mechanics/molecular mechanics molecular dynamics (QM/MM MD) simulations. Quantum mechanics simulations in implicit water show both peptides (and their 1:1 mixture) exhibit PPII conformations. However, hybrid QM/MM MD simulations suggest that some deviations from this conformation are present for P_6K and P_6E in peptide bonds close to the charged residue, whereas in the 1:1 mixture a PPII structure is observed. Finally, aggregation of the peptides was investigated using Replica Exchange Molecular Dynamics simulations. These reveal a tendency for the average aggregate size (as measured by the radius of gyration) to increase with increasing temperature, which is especially marked for P_6K, although the fraction of the most populated clusters is larger for P_6E.

Introduction

Surfactant-like peptides (SLPs) are remarkable self-assembling molecules in which the aggregation properties (micelle, nanotube, vesicle formation, etc) arising from amphiphilicity are combined with the bioactivity of peptide units. Zhang and his group performed pioneering work on model SLPs such as A_6D, V_6D, V_6D_2 and L_6D_2.[1-2] To our knowledge, similar SLPs with oligo-proline chains have not, as yet, been examined. We expect that such peptides will have interesting self-assembly properties due to the constrained conformation of the oliogo-proline chain, when combined with an ionic amino acid "headgroup". Having recently elucidated several remarkable aspects of the self-assembly of alanine-rich SLPs with cationic headgroups such as A_6K[3-4] and A_6R,[5] we set out to examine the solution self-assembly of peptides P_6E (Pro_6Glu) and P_6K (Pro_6Lys) with proline "tails" rather than alanine ones. Proline is the only amino acid with a secondary amine, leading to a highly constrained conformational space. This is expected to lead to a distinct backbone conformation compared to oligo-alanine repeats. In fact "polyproline" backbone structures are observed for proline-rich peptides, including the polyproline II (PPII) conformation. Actually, PPII is observed not only for peptides containing proline, for instance βAβAKLVF(F) peptides exhibit this conformation (βA denotes β_2-alanine).[6]

The polyproline II (PPII) structure is a left-handed helix with 3 residues per turn and a pitch of 9.3-9.4 Å.[7-8] The peptide bonds are all *trans*. The PPII structure is a significant backbone conformation for unfolded proteins.[9-10] It is involved in important biological processes[11] such as the molecular recognition of peptides,[12] and is observed for natural proteins such as elastin[13]. The structure of collagen (which is proline-rich) comprises a triple helical structure, each of which has a conformation close to that of PPII.[8] In addition to being observed for

proline-rich peptides, it also commonly occurs for short alanine-rich peptides.[6, 9, 14-15] Despite this, we did not observe it for the short alanine-rich peptides A_6K[3] or A_6R[16]. It has also been suggested that PPII may be a significant component in the formation of toxic amyloid pre-fibrillar structures, based on a study of pre-fibrillar structures formed by human lysozyme.[17]

In the present paper we examine the influence of charge on the secondary structure of model proline-rich peptides comprising a hydrophobic hexa-proline sequence and terminal charged residues, either cationic (lysine, K) or anionic (glutamic acid, E). Model peptides with single charged termini P_6K and P_6E are studied as is a di-telechelic peptide KP_6E along with equimolar mixtures of P_6K and P_6E. These peptides are analogues of the alanine or valine-rich SLPs mentioned above, but with proline residues substituted as hydrophobic residues. Very recently the crystal structure has been obtained for a P_6 derivative p-Br-C_6H_4-P_6-OH, providing a detailed molecular conformation for the PPII helix.[18] This work highlights the significant importance of understanding the conformation of polyproline peptides and proteins with important roles in biological processes. Here, the secondary structure of P_6 peptides with charged terminal residues is eludicated for the first time using circular dichroism (CD) and FTIR spectroscopies along with fluorescence probe measurements which reveal the existence of critical aggregation concentrations. Electrostatic interactions between terminal residues are found to influence the content of PPII. This is also modulated by temperature – an increase in temperature favouring reduction in PPII content at the expense of random coil and β-sheet structure. While the secondary structure of short poly(L-proline) oligomers (extending down to PPPPP and PPPPPP etc) has been extensively studied,[19-26] by CD and other techniques, we are not aware of prior studies on conformational properties of proline-rich surfactant-like peptides. A conformational transition from PPI to PPII upon heating (also thermoreversible) has been observed for a P_{12} peptide in *n*-propanol by CD and

it was associated to the cis-trans isomerization of amide bonds.[27] The transition kinetics and thermodynamics were also investigated. The conformation of proline oligopeptides capped at the termini with different aromatic residues (tryptophan, W, or tyrosine, Y) WP_nY (n = 0-5) has been investigated via NMR and computer modelling, with a particular focus on cis-trans isomerization.[28]

The observation of large vesicle-like structures at high temperature (80 ^oC) was recently reported for PBS buffer solutions of P_6E and P_6E.[29] This reversible transition was accompanied by a conformational transition from PPII to unordered. Here, we investigate the conformation of these peptides in water not buffer, which is expected to reduce the screening of the charged residues and hence may significantly influence conformation and aggregation. The conformation is probed via spectroscopic and x-ray scattering methods including measurements as a function of both concentration and temperature. We also report measurements of the thermodynamics of hydration of P_6E and P_6K via isothermal titration calorimetry (ITC) measurements. Finally, a comprehensive suite of computer simulations are performed to analyse the peptide conformation (QM/MM methods) for comparison to the experimental data and also replica exchange molecular dynamics (REMD) simulations of aggregation, with a particular focus on temperature-driven aggregation, again for comparison with the experimental results.

Results

Spectroscopic and X-ray Diffraction Analysis of Peptide Conformation

CD was used to probe the secondary structure of peptides P_6K, P_6E, KP_6E and 1:1 mixtures of P_6K and P_6K. Figure 1 shows CD spectra obtained for P_6E measured for a range of concentrations. Spectra for P_6K, KP_6E and 1:1 P_6K: P_6E are presented in Fig.S2 (ESI). The presence of a deep minimum in a CD spectrum at around 200-205 nm, along with a broad positive maximum at around 225 nm is a signature of PPII secondary structure.[10-11, 21, 30-35] Spectra with a negative band centered at about 195-200 nm with no positive band at about 215-225 nm secondary maximum are characteristic of those of unordered peptide,[21, 32] lacking secondary structure. The spectra in Fig.1 and Fig.S2 (ESI) show that all peptides adopt unordered conformation at low concentration but at high concentration they form PPII structures. In fact, by plotting the intensity of the positive maximum at 227 nm versus concentration we find that, unexpectedly, there is a discontinuity indicating a critical concentration for the ordering transition. As evident from Fig.2, this occurs at (0.35 ± 0.03) wt% for all four of our peptides.

The temperature dependence of the CD spectra at fixed concentration (1 wt% of each peptide) was examined in the range 5 – 80 ^{o}C. The corresponding spectra are shown in Fig.S3 (ESI). Increasing temperature led to a reduction in molar ellipticity at the negative minimum and the positive band at about 202 nm, indicating increased unordered structure at the expense of PPII content with increasing temperature. This was further highlighted by plotting difference spectra with respect to the lowest temperature (5 ^{o}C) spectra, as shown in Fig.S4 (ESI). This highlights the reduction in the minimum at 200-205 nm and the development of a minimum near 220 nm. An isodichroic (isosbestic) point is observed for all peptides at 214-216 nm. A high degree of reversibility in the CD spectra was observed on cooling back to 20 ^{o}C, with recovery of the PPII secondary structure component. The temperature-dependent CD

spectra thus point to a decrease in PPII conformation on heating, with the development of random coil formation. This is a reversible process with no evidence for a discontinuous thermal transition in contrast to the critical point observed as a function of concentration. The thermoreversible transition from PPII to disorder is in agreement with previous results for these peptides in PBS buffer[29] and suggests that charge screening in this buffer does not influence this thermal conformational transition.

FTIR spectra were obtained for samples at concentrations above the critical concentration (Fig. S5, ESI shows the amide I´ region). A broad peak at around 1635 cm^{-1} is observed for all peptides examined. This peak is in the range observed for PPII, structures.[22-23, 26, 36] The additional peak observed for all concentrations at 1672 cm^{-1} is dominated by a contribution from TFA (trifluoro acetate) counterions.[37-39] For this reason, and also because solution FTIR cannot be performed for low concentration samples, the FTIR spectra are much less informative than CD spectra in regard to probing PPII conformation.

Fibre XRD was used to provide more detailed information on peptide conformation. Fig.S6 (ESI) shows representative one-dimensional intensity profiles obtained by integration of the 2D patterns measured for P_6K. In both the XRD profiles (measured for stalks prepared from two different concentrations of peptide), strong broad peaks are observed at (4.6 ± 0.05) Å. This corresponds to half the pitch period of a PPII structure.[7] The broad shoulder at lower wavenumber *q* seems to be associated with the full pitch (9.2 Å). The sharp peaks with *d* = 3.14 Å observed for both samples correspond closely to the expected translation per residue in a PPII structure (side peaks with *d* = 3.08 Å are also observed for the 0.5 wt% P_6K sample). The peak associated with the pitch is broad, as expected for such as short peptide

since for the heptapeptide only just over two turns of a PPII structure will be present. The fibre XRD thus directly confirms the presence of PPII structure.

Peptide Aggregation Studies

We hypothesized that this type of peptide may have surfacant-like properties. We therefore used the fluorescent probe ANS (8-anilonapthalane-1-sulfonic acid) which is sensitive to the the formation of hydrophobic environments[40-44] and performed fluorescence assays. The resulting plots of fluorescence intensity versus peptide concentration are shown in Fig.3. (the original fluorescence spectra are plotted in Figure S7, ESI). The data shows that all samples exhibit critical concentrations similar to those corresponding to the conformational transition located by CD (Figure 2). The values for peptides containing single charged residues (P_6K and P_6E) seem slightly lower, reflecting the electrostatic repulsion between the charged termini. This is the first time that a conformational transiton in a peptide has been correlated to a discontinuity in fluorescence probe assay. It indicates that the transition to PPII is accompanied by a change in the fluorescence dye ANS, indicating a change in local hydrophobicity.

We then examined whether the transition corresponds to the onset of aggregation at high concentration. Small-angle X-ray scattering (SAXS) and dynamic light scattering (DLS) were used to examine the presence of possible aggregates (at room temperature). In fact, SAXS showed the presence of unaggregated peptide even at 1 wt% (i.e. above the critical concentration) as exemplified by the data for P_6E in Fig.S8 (ESI). SAXS intensity profiles were found to be dominated by contributions from coil-like structures as shown. The SAXS profiles measured for all peptides and P_6E:P_6K mixtures studied in fact have a very similar

shape (Fig.S9, ESI). The SAXS data for P_6E could be fitted to the form factor of a Gaussian coil, or alternatively using the software MultiFoXS[45] to a compact conformation based on a pdb file generated for a helical conformation (Fig.S8b, ESI). As expected, the Gaussian coil form factor fit leads to a radius of gyration corresponding to single molecule coils rather than extended nanostructures. Table S1 (ESI) contains parameters obtains from fits to all of the the SAXS data shown in Fig.S9. DLS also showed the presence of mainly monomeric peptide, irrespective of whether the peptide concentration was above or below the critical concentration detected by CD and ANS fluorescence probe assays. Representative DLS size distributions (number average) for P_6E and P_6K are shown in Fig.S10 (ESI).

Thermodynamic Measurements

The enthalpy of hydration, a well-defined quantity that gives properly defined values of ΔG of hydration of amino acids[46] can be determined from isothermal titration calorimetry (ITC). In a classical ITC experiment[47-48] changes in the heat flow are measured during peptide titration. These results for P_6K, as well as the reaction enthalpy and its derivative as a function of concentration are presented in Figure 4, and for the other peptides P_6E, KP_6E and their mixtures (P_6K:P_6E 1:1 and 10:1) in Figs. S11 and S12. The thermodynamic parameters calculated from these measurements are summarized in Tables S2 and S3. All ITC measurements were performed on samples with a concentration 0.2 wt%, which is below the critical aggregation concentration values.

The Gibbs free energy, ΔG, of hydration calculated for P_6K is -6.16 kcal/mol, compared to that for lysine (-4.38 kcal mol^{-1}).[46] Further, ΔG = -6.98 kcal mol^{-1} found here for P_6E is in

excellent agreement with the value for E (glutamic acid) -6.47 kcal mol^{-1}.[46] This indicates that E dominates the hydration free energy of P_6E. Although the experimental value of the Gibbs free energy of hydration cannot be directly compared with the interaction energy, ΔE_{int}, between the peptide and the solvent derived from theoretical calculations because of the omission of different contributions in the latter (e.g. entropic contribution and peptide concentration contribution), both thermodynamic parameters follow the same tendency. More specifically, ΔE_{int} was 183.3 kcal mol^{-1} more favorable for P_6E than for P_6K (details are given in the Supporting Information).

Computer Simulations

Quantum Mechanics Simulations

Backbone conformations of the minimum energy structures were analysed. The results are summarized in Table 1. Conformations are denoted using the nomenclature introduced by Perczel *et al*.,[49] that categorizes the potential energy surface E = E(φ,ψ) of α-amino acids in nine different regions: γ_D, δ_D, α_D, ε_D, ε_{DL}, ε_L, α_L, δ_L, and γ_L. In the case of Pro or proline-like residues, only the γ_L (γ-turn), α_L (α-helix), and ε_L (polyproline II) conformations are accessible because φ is confined to values around -60°. The results show that there is no difference between secondary structures of charged and neutralized systems (P_6E and P_6K versus 1:1 P_6E: P_6K systems). All of them present a PPII secondary structure (ε_L, using Perczel nomenclature). No differences were noted on the backbone dihedral angles beyond to those corresponding to an expected polyproline peptide. Classical MD also shows similar behaviour (not shown). Classical molecular dynamics (MD) simulations also show similar behaviour (not shown).

Hybrid QM/MM - MD simulations

A conformational study of peptide bond dihedral angles of P6E and P6K was performed during the last 5 ps of QM/MM-MD simulations. SI Fig.S13 and Fig.S14 show snapshots illustrating typical conformations of P_6E and P_6K respectively, whilst Fig.S15 shows a corresponding image for 1: 1 P_6E: P_6K. Average conformational parameters are listed in Table 2. This shows that the central proline residues have a PPII conformation (as in the QM calculations). Accordingly, results from hybrid QM/MM MD simulations, where the solvent in included explicitly and the peptide is polarized at the electronic density level, show secondary structures that are in good agreement with experimental data. However, for charged P_6E and P_6K there are some backbone gamma-turns conformations (γ_L) at the position next to charged residues (Table 2). This feature suggests that the interactions between the charged residue and solvent molecules, as well as the repulsive carboxylate⋯carboxylate interactions in the case of P_6E, lead to some secondary structure differences. It should be mentioned that, due to their enormous computational cost, the length of QM/MM MD trajectories is very limited and, therefore, the γ_L state of the charged residues should not be considered as the most stable conformation but as a metastable state, which probably would evolve towards another conformational state in time. In any case, QM/MM MD simulations, which include a better description of intra- and inter-molecular interactions than both classical MD polarization and QM simulations with implicit solvent, clearly indicate that the PPII conformation is lost at the position of the charged residue. In contrast, for the neutralized system 1: 1 P_6E: P_6K presents a PPII conformation for all peptide bonds (ε_L), proving that the conformational preferences of peptides containing charged residues, i.e. P_6E and P_6K, are affected by the interactions mentioned above.

Results from QM/MM MD simulations have also been used to study peptide⋯solvent interactions. For this purpose, radial distribution function (RDF) pairs involving the oxygen atoms from the peptide and the hydrogen atoms of water ($g_{CO\cdots H}$) and the amide hydrogen atoms and the oxygen atom of water ($g_{NH\cdots O}$) were calculated using the snapshots saved from simulations on P_6E and P_6K and 1:1 P_6E: P_6K systems. The resulting profiles, which are displayed in SI Fig. S15, show a sharp peak centred at 1.7-1.8 Å. This is characteristic of the first hydration shell and contains water molecules bonded through specific H-bonds with the amide and both charged carboxylate and ammonium groups from peptide molecule. Comparison of the RDFs indicates that CO⋯H(water) H-bonds are more abundant than NH⋯O(water) H-bonds for all examined systems. Furthermore, CO⋯H(water) interactions are more abundant for P_6E than for P_6K and 1: 1 P_6E : P_6K, while P_6K is the peptide with the highest amount of NH⋯O(water) interactions.

To provide a deeper characterization of the first solvation shell, both the $g_{CO\cdots H}$ and $g_{NH\cdots O}$ RDFs were calculated for the Pro, Glu and Lys residues of each peptide separately. The resulting profiles (not shown) were used to estimate the hydration number (NW), which refers to the number of water molecules residing in the nearest neighbour shell and bonded to the peptide through CO⋯H(water) or NH⋯O(water) H-bonds, for each residue. More specifically, integration of the first peak with a maximum below a distance of 2.5 Å from the $g_{CO\cdots H}$ and $g_{NH\ldots O}$ RDF gave rise to the NW_{CO} and NW_{NH} hydration humber, respectively. It is worth noting that NW_{CO} refers to number of water molecules H-bonded to both the C=O(amide) and carboxylate groups (if available) of the residue, while NW_{NH} involves water molecules interacting with the NH(amide) and ammonium (if available) groups of the residue. The global hydration number (GNW) for a given residue corresponds to the sum of

NW_{CO} and NW_{NH}. Table 3 lists the position of the first hydration peak in the corresponding RDFs, NW_{CO}, NW_{NH} and GNW for the Pro, Glu and Lys residues of each examined system.

The results indicate that, on average, the amount of specific Pro⋯water interactions is significantly more favourable for P_6E (GNW= 3.3) than for P_6K (GNW= 2.0) and 1: 1 P_6E: P_6K (GNW= 2.3). This should be attributed not only to the fact that Pro residues from the former peptide are involved in more CO⋯H(water) H-bonds than those from the other ones but also to the characteristic N-terminal Pro residue, which is the only one able to form specific NH⋯O(water) interactions. For P_6E this residue interacts on average with 1.2 water molecules, this value decreasing to 0.3 for P_6K and 1: 1 P_6E : P_6K. Considering that each studied peptide molecule contains six Pro residues and only one charged residue, the GNW values listed in Table 3 reflect that the latter is much better hydrated than the former. This is due not only to the charged side group of Glu and Lys but also to the particular chemical structure of Pro. Thus, the distinctive cyclic structure of Pro is incompatible with the formation of N–H⋯O(water) (with the obvious exception of the residue located at the N-terminus position) and also affects the hydration of the carbonyl group. Thus, NW_{CO} is lower for Pro than for Glu and Lys in all cases.

On the other hand, Table 3 indicates the lack of specific (Glu)NH⋯O(water) hydrogen bonds in P_6E. Thus, the first peak in the corresponding $g_{NH\cdots O}$ RDF appears at a distance higher than 3 Å, indicating that water molecules located around the N–H(amide) moiety are not forming specific interactions. This feature is less pronounced in 1: 1 P_6E : P_6K with a NW_{NH} value of 0.6, even though the first peak in the corresponding $g_{NH\ldots O}$ profile is centred at a distance ~0.4 Å less than those typically observed for CO⋯H RDFs (Table 3). Finally, Lys is the most prone to the formation of specific interactions with neighbouring water molecules. In

particular, the NW_{CO} values calculated for Lys, which are the largest ones, indicate that the terminal carboxylate group in P_6K and 1:1 P_6E: P_6K interacts with almost three water molecules simultaneously. This is a remarkable result since the Glu residue in P_6E and 1 : 1 P_6E: P_6K only forms CO⋯H hydrogen bonds with about 2.5 water molecules, even though the concentration of negative charge in P_6E is higher due to the proximity of the carboxylate side group. However, detailed inspection of saved snapshots indicates that the tendency to form peptide⋯peptide interactions is higher for Glu than for Lys, which explains the favoured hydration of the latter with respect to the former. Also, Table 3 indicates that, in average, the ammonium side group of Lys interacts with almost one water molecule, which represents another contribution to the high hydration degree of this residue.

Replica Exchange Molecular Dynamics (REMD) Simulations of Peptide aggregation

The initial stages of the aggregation phenomenon were investigated using REMD simulations, considering systems formed by eight P_6E or eight P_6K molecules (denoted 8P6E and 8PEK, respectively). For each system, 24 replicas were performed in the temperature interval between 274 K up to 453 K that represented 180 ns of trajectory. Structures generated from these simulations were analysed following a density-based algorithm for discovering clusters in large spatial databases with noise (DBSCAN algorithm)[50] to characterize aggregation structure similarities. The results, which are summarized in Table 4, show that the cluster named *0* is the predominant one at all temperatures for the two peptides, even though terminal charges affect significantly its population (i.e. the population of cluster *0* ranges from 39.0% to 45.6% and from 6.7% to 15.2% for 8P6E and 8P6K, respectively). In any case, the charged residue of uncapped peptides is playing an important role favouring the aggregate formation. Figure 5 shows representatives of the most populated cluster of 8P6E

and 8P6K aggregates at 300 K. The figure shows the existence of regions of ionic stabilization among the charged termini and residues. This feature is particularly relevant for 8P6E, in which the all chain form $NH_3^+\cdots^-OOC$ electrostatic interactions, while in the case of 8P6K some molecules interact through conventional hydrogen bond between the amide bonds.

Figure 6 shows the temperature dependence of the averaged R_g plotted against the REMD temperature. A small increase in aggregation for the 8P6E system is observed when the temperature increases, even though there is a small decrease on the radius of gyration, R_g, at intermediate temperatures. This tendency is even more evident for the 8P6K system. However, the fractional occupation of the most populated cluster is much higher for the 8P6E aggregate (39-45 %, Table 4) with a broader R_g than the 8P6K aggregate with a more compact structure but with only about 6-15 % occupation. It is emphasized that due to the constraint on the size of the simulation system, these results only represent tendencies at the initial stages of the aggregation process. It is not possible to simulate 300 – 600 nm peptide vesicles, as observed experimentally for P_6E and P_6K in PBS buffer solution at high temperature,[29] because of the huge number of involved molecules.

Summary and Conclusions

In summary, model proline rich peptides adopt polyproline II structures above a critical concentration which is uniquely identified both from changes in CD spectra and from a fluorescence probe assay. To the best of our knowledge, this type of transition has not been reported for oligo-proline peptides. This suggests that it is due to the presence of the terminal

charged residues, in particular their aggregation propensity. This is supported by our demonstration that this type of surfactant-like peptide exhibits a critical aggregation concentration. We further show that this is not suppressed by electrostatic interactions in either the 1:1 mixture or the peptide KP_6E containing oppositely charged terminal residues. Heating leads to a reversible partial "unfolding" process.

Isothermal titration calorimetry (ITC) provides measures of the Gibbs energies of hydration for P_6E and P_6K, which are close to those measured independently for the charged E and K residues. This shows that solvation of these peptides is dominated by the charged terminal groups.

A comprehensive set of simulations provided insight into the conformation of the peptides (hybrid QM/MM MD simulations) which reveals mainly polyproline II structure, although there are backbone γ-turns of the bonds at the charged residues. The simulated PPII structures are in agreement with the CD experiments. Replica exchange molecular dynamics simulations provide information on temperature-dependent aggregation in the initial stage. The radius of gyration increases significantly with temperature for P_6K, but less so for P_6E. On the other hand, the fraction of clusters formed by P_6E at a given temperature is substantially larger.

The surfactant-like proline-rich peptides serve as valuable models for investigation of the tuning of PPII structures via electrostatic interactions. It is hoped that our work stimulates further investigation of these effects via theoretical methods.[20] Further applications of proline-rich surfactant-like peptides in elastin or collagen- mimetic biomaterials are

envisaged and our work provides insight into material design parameters through secondary structure tuning.

ACKNOWLEDGEMENTS

This work was supported by EPSRC grants EP/G067538/1 and EP/L020599/1 to IWH. We are grateful to Diamond for the award of beamtime (references SM8909 and SM10077-1) and to Giuliano Siligardi and Katsuaki Inoue for assistance. We also thank the ESRF for the award of beamtime (references MX-1606 and MX-1869) and Adam Round and Martha Brennich for assistance. The ITC studies were supported by Israel Science Foundation grants 741/11 and 1117/2016 to DD. We thank Ellina Kesselman (Technion) for support. CA and JT thanks the support from MINECO/FEDER (grant MAT2015-69367-R) and from CSUC. Support for the research of C.A. was received through the prize "ICREA Academia" for excellence in research funded by the Generalitat de Catalunya.

Electronic Supplementary Information (ESI) available: [Experimental Methods and characterization data. Additional CD spectra, FTIR and fluorescence spectra, XRD data, SAXS data and SAXS fit parameters additional ITC data and tables of thermodynamic parameters]. See pubs.acs.org.

REFERENCES

(1) Vauthey, S.; Santoso, S.; Gong, H.; Watson, N.; Zhang, S., Molecular Self-Assembly of Surfactant-Like Peptides to Form Nanotubes and Nanovesicles. *Proc. Nat. Acad. Sci USA* **2002,** *99*, 5355-5360.

(2) Zhao, X.; Zhang, S., Fabrication of Molecular Materials Using Peptide Construction Motifs. *Trends in Biotech.* **2004,** *22*, 470-476.

(3) Castelletto, V.; Nutt, D. R.; Hamley , I. W.; Bucak, S.; Cenker, C.; Olsson, U., Structure of Single-Wall Peptide Nanotubes: In Situ Flow Aligning X-Ray Diffraction. *Chem. Comm.* **2010,** *46*, 6270-6272.

(4) Middleton, D. A.; Madine, J.; Castelletto, V.; Hamley , I. W., Insights into the Molecular Architecture of a Peptide Nanotube Using Ftir and Solid-State NMR Spectroscopy Combined with Sample Alignment. *Angew. Chem., Int. Ed. Engl.* **2013,** *52*, 10537-10540.

(5) Hamley , I. W.; Dehsorkhi, A.; Castelletto, V., Self-Assembled Arginine-Coated Peptide Nanosheets in Water. *Chem. Comm.* **2013,** *49*, 1850-1852.

(6) Castelletto, V.; Hamley , I. W.; Cenker, C.; Olsson, U.; Adamcik, J.; Mezzenga, R.; Miravet, J. F.; Escuder, B.; Rodriguez-Llansola, F., Influence of End-Capping on the Self-Assembly of Model Amyloid Peptide Fragments. *J. Phys. Chem. B* **2011,** *115*, 2107-2116.

(7) Creighton, T. E., *Proteins. Structures and Molecular Properties*. W.H.Freeman: New York, 1993.

(8) Voet, D.; Voet, J. G., *Biochemistry*. John Wiley: New York, 1995.

(9) Shi, Z. S.; Chen, K.; Liu, Z. G.; Kallenbach, N. R., Conformation of the Backbone in Unfolded Proteins. *Chem. Rev.* **2006,** *106*, 1877-1897.

(10) Schweitzer-Stenner, R.; Measey, T.; Hagarman, A.; Eker, F.; Griebenow, K., Salmon Calcitonin and Amyloid B: Two Peptides with Amyloidogenic Capacity Adopt Different Conformational Manifolds in Their Unfolded States. *Biochemistry* **2006,** *45*, 2810-2819.

(11) Bochicchio, B.; Tamburro, A. M., Polyproline II Structure in Proteins: Identification by Chiroptical Spectroscopies, Stability, and Functions. *Chirality* **2002,** *14*, 782-792.

(12) Siligardi, G.; Drake, A. F., The Importance of Extended Conformations and, in Particular, the PII Conformation for the Molecular Recognition of Peptides. *Biopolymers* **1995,** *37*, 281-292.

(13) Martino, M.; Bavoso, A.; Guantieri, V.; Coviello, A.; Tamburro, A. M., On the Occurrence of Polyproline Ii Structure in Elastin. *J. Molec. Struct.* **2000,** *519*, 173-189.

(14) Kim, Y. S.; Wang, J. P.; Hochstrasser, R. M., Two-Dimensional Infrared Spectroscopy of the Alanine Dipeptide in Aqueous Solution. *J. Phys. Chem. B* **2005,** *109*, 7511-7521.

(15) Hagarman, A.; Measey, T.; Doddasomayajula, R. S.; Dragomir, I.; Eker, F.; Griebenow, K.; Schweitzer-Stenner, R., Conformational Analysis of XA and AX Dipeptides in Water by Electronic Circular Dichroism and ^{1}H NMR Spectroscopy. *J. Phys. Chem. B* **2006,** *110*, 6979-6986.

(16) Dehsorkhi, A.; Castelletto, V.; Hamley, I. W.; Seitsonen, J.; Ruokolainen, J., Interaction between a Cationic Surfactant-Like Peptide and Lipid Vesicles and its Relationship to Antimicrobial Activity. *Langmuir* **2013,** *29*, 14246-14253.

(17) Blanch, E. W.; Morozova-Roche, L. A.; Cochran , D. A. E.; Doig, A. J.; Hecht, L.; Barron, L. D., Is Polyproline II Helix the Killer Conformation? A Raman Optical Activity Study of the Amyloidogenic Prefibrillar Intermediate of Human Lysozyme. *J. Molec. Biol.* **2000,** *301*, 553-563.

(18) Wilhelm, P.; Lewandowski, B.; Trapp, N.; Wennemers, H., A Crystal Structure of an Oligoproline PPII-Helix, at Last. *J. Am. Chem. Soc.* **2014,** *136*, 15829-15832.

(19) Rothe, M.; Rott, H.; Mazanek, J. In *Solid Phase Synthesis and Conformation of Monodisperse High Molecular Weight Oligo-L-Prolines*, Peptides 1976, Bruxelles, Loffet, A., Ed. Université de Bruxelles: Bruxelles, 1976; pp 309-318.

(20) Manning, M. C.; Woody, R. W., Theoretical CD Studies of Polypeptide Helices: Examination of Important Electronic and Geometric Factors. *Biopolymers* **1991,** *31*, 569-586.

(21) Toniolo, C.; Formaggio, F.; Woody, R. W., Electronic Circular Dichroism of Peptides. In *Comprehensive Chirooptical Spectroscopy*, Berova, N.; Polavarapu, P. L.; Nakanishi, K.; Woody, R. W., Eds. Wiley, New York, 2012; Vol. 2.

(22) Dukor, R. K.; Keiderling, T. A., Reassessment of the Random Coil Conformation: Vibrational CD Study of Proline Oligopeptides and Related Polypeptides. *Biopolymers* **1991,** *31*, 1747-1761.

(23) Dukor, R. K.; Keiderling, T. A., Mutarotation Studies of Poly-L-Proline Using FTIR, Electronic and Vibrational Circular Dichroism. *Chem. Comm.* **1996,** *2*, 83-100.

(24) Kakinoki, S.; Hirano, Y.; Oka, M., Conformational Analysis of Proline Oligopeptides. In *Peptide Science 2004*, Shimohigashi, Y., Ed. The Japanese Peptide Society: Minoh, Japan, 2005; pp 453-456.

(25) Kakinoki, S.; Hirano, Y.; Oka, M., On the Stability of Polyproline-I and II Structures of Proline Oligopeptides. *Polym. Bull.* **2005,** *53*, 109-115.

(26) Chi, H.; Lakhani, A.; Roy, A.; Nakaema, M.; Keiderling, T. A., Inter-Residue Coupling and Equilibrium Unfolding of PPII Helical Peptides. Vibrational Spectra Enhanced with ^{13}C Isotopic Labeling. *J. Phys. Chem. B* **2010,** *114*, 12744-12753.

(27) Kuemin, M.; Engel, J.; Wennemers, H., Temperature-Induced Transition between Polyproline I and II Helices: Quantitative Fitting of Hysteresis Effects. *J. Peptide Sci.* **2010,** *16*, 596-600.

(28) Poznański, J.; Ejchart, A.; Wierzchowski, K. L.; Ciurak, M., ^{1}H- and ^{13}C- NMR Investigations on *cis-trans* Isomerization of Proline Peptide-Bonds and Conformation of Aromatic Side-Chains in H-Trp-(Pro)(N)-Tyr-Oh Peptides. *Biopolymers* **1993,** *33*, 781-795.

(29) Felip-Leon, C.; Galindo, F.; Miravet, J. F.; Castelletto, V.; Hamley, I. W., Thermally Regulated Reversible Formation of Vesicle-Like Assemblies by Hexaproline Amphiphiles. *J. Phys. Chem. B* **2017,** *121*, 7443-7446.

(30) Woody, R. W., Circular Dichroism. *Meth. Enzymol.* **1995,** *246*, 34-71.

(31) Shi, Z. S.; Woody, R. W.; Kallenbach, N. R., Is Polyproline II a Major Backbone Conformation in Unfolded Proteins? In *Unfolded Proteins*, Rose, G. D., Ed. Academic Press Inc: San Diego, 2002; Vol. 62, pp 163-240.

(32) Paramonov, S. E.; Jun, H.-W.; Hartgerink, J. D., Self-Assembly of Peptide-Amphiphile Nanofibers: The Roles of Hydrogen Bonding and Amphiphilic Packing. *J. Am. Chem. Soc.* **2006,** *128*, 7291-7298.

(33) Woody, R. W., Circular Dichroism Spectrum of Peptides in the Poly(Pro)II Conformation. *J. Am. Chem. Soc.* **2009,** *131*, 8234-8245.

(34) Chen, K.; Liu, Z. G.; Kallenbach, N. R., The Polyproline II Conformation in Short Alanine Peptides Is Noncooperative. *Proc. Nat. Acad. Sci USA* **2004,** *101*, 15352-15357.

(35) Drake, A. F.; Siligardi, G.; Gibbons, W. A., Reassessment of the Electronic Circular-Dichroism Criteria for Random Coil Conformations of Poly(L-Lysine) and the Implications for Protein Folding and Denaturation Studies. *Biophys. Chem.* **1988,** *31*, 143-146.

(36) Keiderling, T. A.; Xu, Q., *Unfolded Peptides and Proteins Studied with Infrared Absorption and Vibrational Circular Dichroism Spectra*. Academic Press: San Diego, 2002; Vol. 62, p 111-161.

(37) Pelton, J. T.; McLean, L. R., Spectroscopic Methods for Analysis of Protein Secondary Structure. *Analytical Biochemistry* **2000,** *277*, 167-176.

(38) Gaussier, H.; Morency, H.; Lavoie, M. C.; Subirade, M., Replacement of Trifluoroacetic Acid with Hcl in the Hydrophobic Purification Steps of Pediocin Pa-1: A Structural Effect. *Appl. Environ. Microbiol.* **2002,** *68*, 4803-4808.

(39) Eker, F.; Griebenow, K.; Schweitzer-Stenner, R., $A\beta_{1-28}$ Fragment of the Amyloid Peptide Predominantly Adopts a Polyproline II Conformation in an Acidic Solution. *Biochemistry* **2004,** *43*, 6893-6898.

(40) Lindgren, M.; Sorgjerd, K.; Hammarstrom, P., Detection and Characterization of Aggregates, Prefibrillar Amyloidogenic Oligomers, and Protofibrils Using Fluorescence Spectroscopy. *Biophys. J.* **2005,** *88*, 4200-4212.

(41) Gasymov, O. K.; Glasgow, B. J., ANS Fluorescence: Potential to Augment the Identification of the External Binding Sites of Proteins. *Biochim. Biophys. Acta-Proteins and Proteomics* **2007,** *1774*, 403-411.

(42) Hawe, A.; Sutter, M.; Jiskoot, W., Extrinsic Fluorescent Dyes as Tools for Protein Characterization. *Pharm. Res.* **2008,** *25*, 1487-1499.

(43) Castelletto, V.; Kirkham, S.; Hamley, I. W.; Kowalczyk, R.; Rabe, M.; Reza, M.; Ruokolainen, J., Self-Assembly of the Toll-Like Receptor Agonist Macrophage-Activating Lipopeptide MALP-2 and of Its Constituent Peptide. *Biomacromolecules* **2016,** *17*, 631-640.

(44) Castelletto, V.; Ryumin, P.; Cramer, R.; Hamley, I. W.; Taylor, M.; Allsop, D.; Reza, M.; Ruokolainen, J.; Arnold, T.; Hermida-Merino, D., et al., Self-Assembly and Anti-

Amyloid Cytotoxicity Activity of Amyloid Beta Peptide Derivatives. *Sci. Rep.* **2017,** *7*, 43637.

(45) Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A., Foxs, Foxsdock and Multifoxs: Single-State and Multi-State Structural Modeling of Proteins and Their Complexes Based on SAza Profiles. *Nucleic Acids Research* **2016,** *44*, W424-W429.

(46) Sitkoff, D.; Sharp, K. A.; Honig, B., Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models. *J. Phys. Chem.* **1994,** *98*, 1978-1988.

(47) Moitzi, C.; Portnaya, I.; Glatter, O.; Ramon, O.; Danino, D., Effect of Temperature on Self-Assembly of Bovine Beta-Casein above and Below Isoelectric pH. Structural Analysis by Cryogenic-Transmission Electron Microscopy and Small-Angle X-Ray Scattering. *Langmuir* **2008,** *24*, 3020-3029.

(48) Danino, D.; Abezgauz, L.; Portnaya, I.; Dan, N., From Discs to Ribbons Networks: The Second Critical Micelle Concentration in Nonionic Sterol Solutions. *J. Phys. Chem. Lett.* **2016,** *7*, 1434-1439.

(49) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J. L.; Marcoccia, J. F.; Csizmadia, I. G., Peptide Models .1. Topology of Selected Peptide Conformational Potential-Energy Surfaces (Glycine and Alanine Derivatives). *J. Am. Chem. Soc.* **1991,** *113*, 6256-6265.

(50) Ester, M.; Kriegel, H.; Sander, J. In *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*, Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Simoudis, E.; Han, J.; Fayyad, U., Eds. American Association for Artificial Intelligence: 1996; pp 226-231.

Tables

Table 1. Conformational parameters (torsion angles **Φ** and **Ψ** in degrees, and backbone conformation, **Cf**) from P_6E, P_6K and 1:1 P_6E: P_6K of the minimum energy structure derived at the M06L/6-31+G(d) level.

Peptide	**$(Pro)_6Glu$**			**$(Pro)_6Lys$**		
Bond	**Φ**	**Ψ**	**Cf**	**Φ**	**Ψ**	**Cf**
			P6E			
0	-	177.8				
1	-59.5	152.3	ε_L			
2	-55.1	149.5	ε_L			
3	-54.0	142.9	ε_L			
4	-57.6	146.6	ε_L			
5	-55.6	148.0	ε_L			
6	-59.4	-				
			P6K			
0				-	176.2	
1				-62.9	163.5	ε_L
2				-60.1	161.4	ε_L
3				-53.8	147.4	ε_L
4				-55.4	147.8	ε_L
5				-53.5	145.9	ε_L
6				-62.2	-	
			P6EP6K			
0	-	177.8		-	175.6	
1	-59.5	152.3	ε_L	-59.7	146.9	ε_L
2	-55.0	145.0	ε_L	-55.1	143.0	ε_L
3	-56.6	147.4	ε_L	-58.4	147.1	ε_L
4	-54.8	145.5	ε_L	-54.5	145.3	ε_L
5	-54.2	145.0	ε_L	-56.1	145.2	ε_L
6	-59.7	-		-62.9	-	

Table 2. Averaged conformational parameters (torsion angles **Φ** and **Ψ** in degrees, and backbone conformation, **Cf**) from P_6E, P_6K and 1:1 P_6E: P_6K simulations using the QM/MM MD approach. Standard deviations are also shown.

Peptide	**$(Pro)_6Glu$**			**$(Pro)_6Lys$**		
Bond	**Φ**	**Ψ**	**Cf**	**Φ**	**Ψ**	**Cf**
			P6E			
0	-	150.2±20.8				
1	-71.8±24.4	112.3±10.7	γ_L			
2	-86.7±17.3	167.6±9.6	ε_L			
3	-63.2±8.0	169.0±8.0	ε_L			
4	-57.7±7.0	154.2±6.2	ε_L			
5	-84.8±7.8	59.3±14.8	γ_L			
6	-172.1±16.2	-				
			P6K			
0					87.4±12.6	
1				-61.0±10.5	158.9±8.0	ε_L
2				-63.3±7.4	170.2±7.7	ε_L
3				-61.2±8.8	141.2±10.4	ε_L
4				-58.2±7.4	161.6±7.4	ε_L
5				-81.0±7.7	67.6±7.3	γ_L
6				-85.2±19.4		
			P6EP6K			
0		146.1±10.0 [a]			158.5±11.7	
1	-69.8±8.5[a]	155.3±7.2 [a]	ε_L	-75.4±11.3	167.7±9.7	ε_L
2	-62.9±6.8 [a]	138.9±12.5 [a]	ε_L	-74.4±11.3	172.1±9.8	ε_L
3	-67.5±7.8 [a]	161.0±8.2 [a]	ε_L	-57.4±10.7	140.2±15.3	ε_L
4	-56.3±9.8 [a]	164.1±6.7 [a]	ε_L	-70.0±11.8	166.6±8.1	ε_L
5	-58.8±8.0 [a]	148.0±8.3 [a]	ε_L	-58.0±10.3	154.0±8.6	ε_L
6	-65.8±10.7 [a]			-177.5±10.3		

[a] Dihedral angles derived from MM atoms of QM/MM MD simulation

Table 3. Data from the $g_{CO\cdots H}$ and $g_{NH\cdots O}$ RDFs calculated for the Pro, Glu and Lys residues of each studied system: position of the peak associated to the first hydration shell ($d_{O\cdots H}$ and $d_{H\cdots O}$; in Å), hydration number associated to the CO and NH moieties of the peptide (NW_{CO} and NW_{NH}, respectively), global hydration number (GNW= NW_{CO} + NW_{NH}). Results derived from quantum regions of P_6E, P_6K and P_6E: P_6K simulations using the QM/MM MD approach.

System	**Pro**			Glu			**Lys**		
	$d_{O\cdots H}$ / $d_{H\cdots O}$	**NW_{CO} / NW_{NH}**	**GNW**	**$d_{O\cdots H}$ / $d_{H\cdots O}$**	**NW_{CO} / NW_{NH}**	**GNW**	**$d_{O\cdots H}$ / $d_{H\cdots O}$**	**NW_{CO} / NW_{NH}**	**GNW**
P6E	1.81 / 1.78	2.1 / 1.2	3.3	1.73 / - [a]	2.3 / - [a]	2.3	-	-	-
P6K	1.75 / 1.85	1.7 / 0.3	2.0	-	-	-	1.77 / 1.83	2.9 / 0.9	3.8
P6EP6K	1.77 / 1.88	1.9 / 0.3	2.2	1.71 / 2.13	2.6 / 0.6	3.2	1.72 / 1.83	2.7 / 0.7	3.4

[a] The first peak in the $g_{NH\ldots O}$ RDF is centered at a distance higher than 3 Å.

Table 4. Fractional occupation, number of hydrogen bonds and radius of gyration of the most populated clusters of 8P6E and 8P6K aggregates at different temperatures.

Temperature	#[a]	Fractional Occupation[b]	Averaged distance[c]	HB[d]	R_g(Å)[e]	R_g^{max}(Å)[f]
			8P6E			
280.42	0	0.390	3.874 ± 1.185	14	23.4 ± 4.3	40.9 ± 8.7
300.00	0	0.456	4.240 ± 1.352	17	23.3 ± 4.3	40.6 ± 8.2
321.25	0	0.435	3.964 ± 1.229	17	23.0 ± 4.4	40.2 ± 8.2
350.86	0	0.420	3.959 ± 1.240	11	22.4 ± 4.0	39.2 ± 7.3
			8P6K			
280.42	0	0.152	3.584 ± 1.032	3	18.5 ± 3.3	35.5 ± 6.6
	1	0.098	2.220 ± 0.599	3	21.0 ± 0.8	40.6 ± 2.4
300.00	0	0.099	3.573 ± 1.069	3	17.0 ± 2.2	32.3 ± 4.4
321.25	0	0.066	3.220 ± 0.897	2	16.4 ± 1.5	31.3 ± 2.7
350.86	0	0.067	3.109 ± 0.731	2	16.3 ± 1.5	31.6 ± 2.9

[a] Cluster number with an occupational fraction of trajectory higher than 5%
[b] Size of cluster as fraction of total trajectory
[c] Average distance between points in the cluster. Standard deviation of points in the cluster is also shown
[d] Number of intermolecular hydrogen bonds with and occupation higher than 2%
[e] Radius of Gyration derived from the representative frame of the cluster
[f] Maximum Radius of Gyration derived from the representative frame of the cluster

Figures

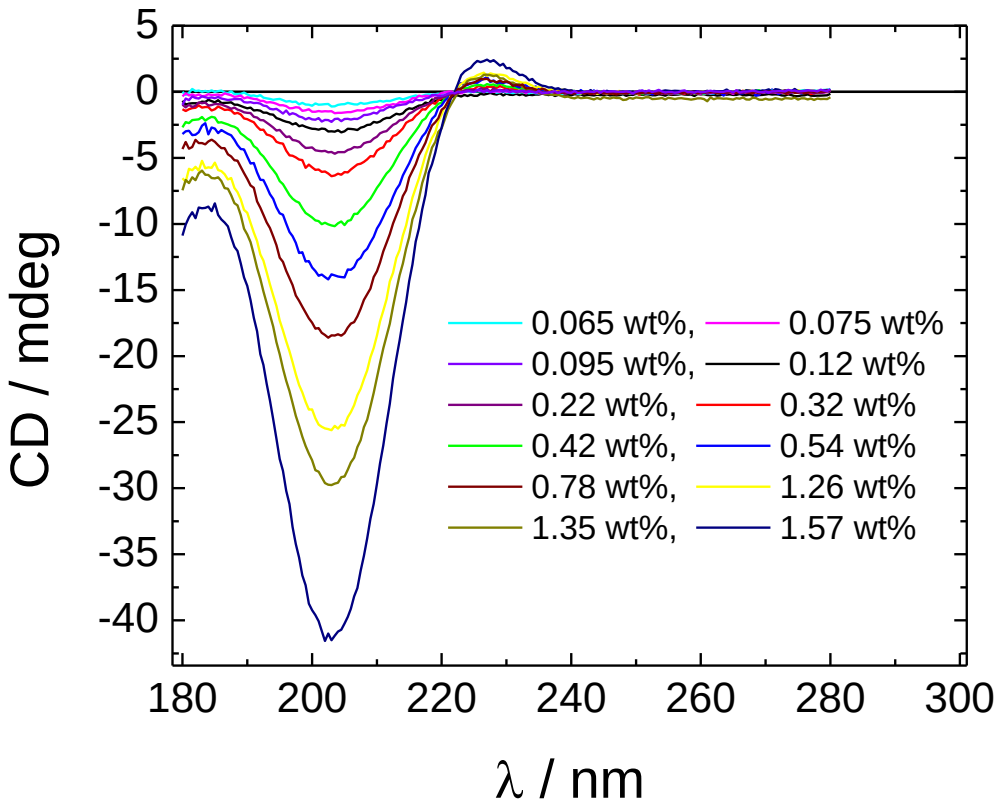


Figure 1. CD spectra from solutions of P_6E, at the concentrations indicated.

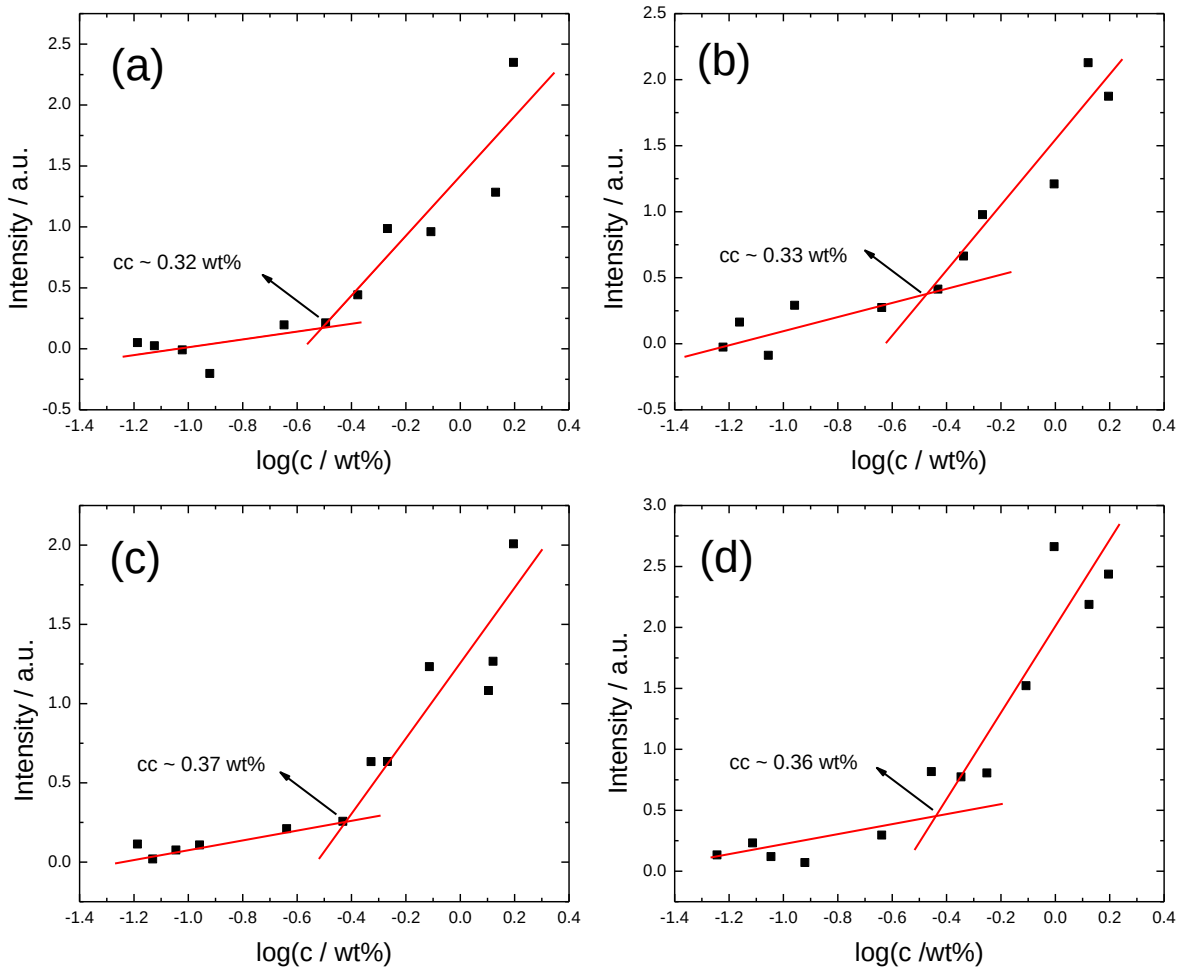


Figure 2. Concentration dependence of the CD signal at 227 nm for (a) P_6E, (b) P_6K, (c) KP_6E, (d) 1:1 P_6E: P_6K.

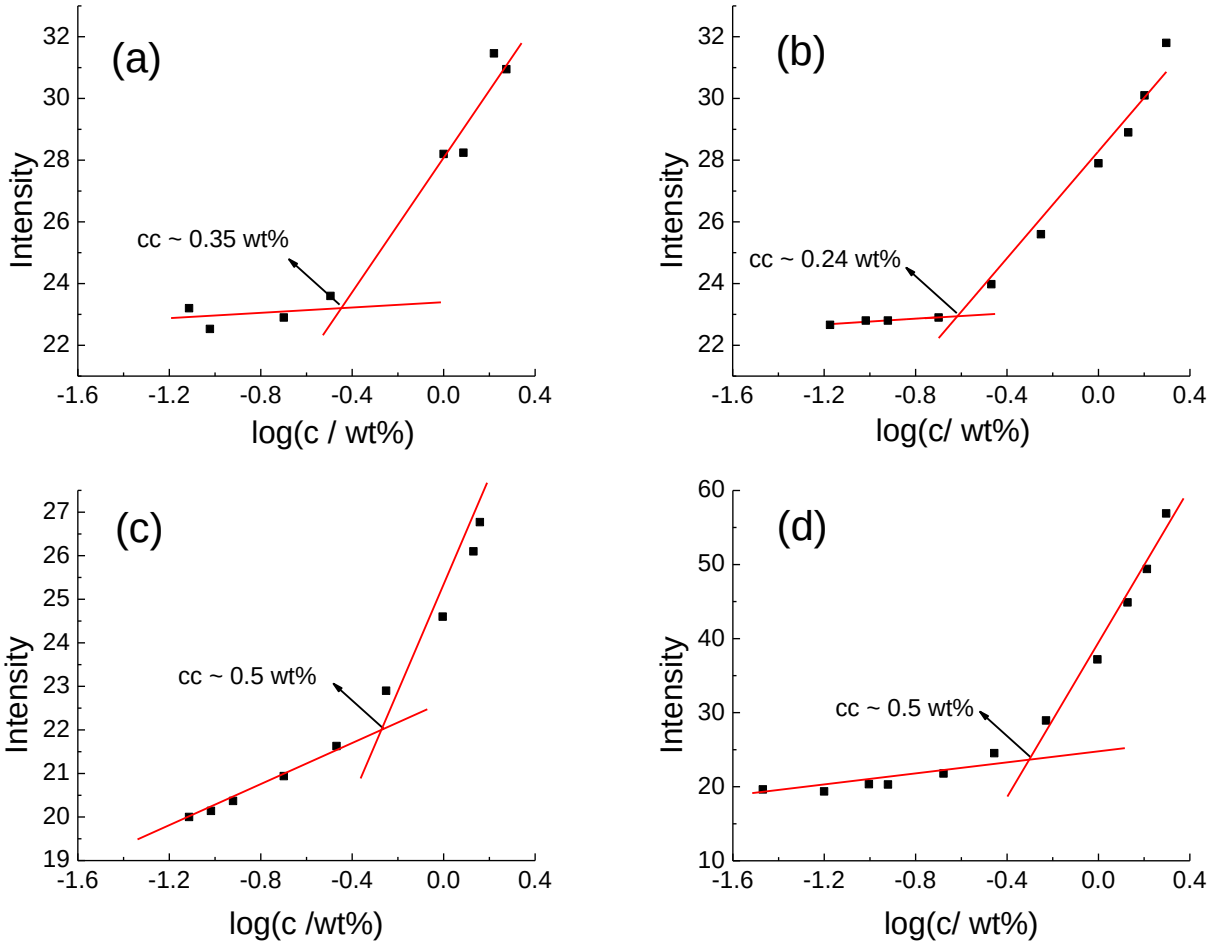


Figure 3. Fluorescence intensity assays using ANS to determine critical concentrations (a) P_6E, (b) P_6K, (c) KP_6E, (d) 1:1 P_6K: P_6E.

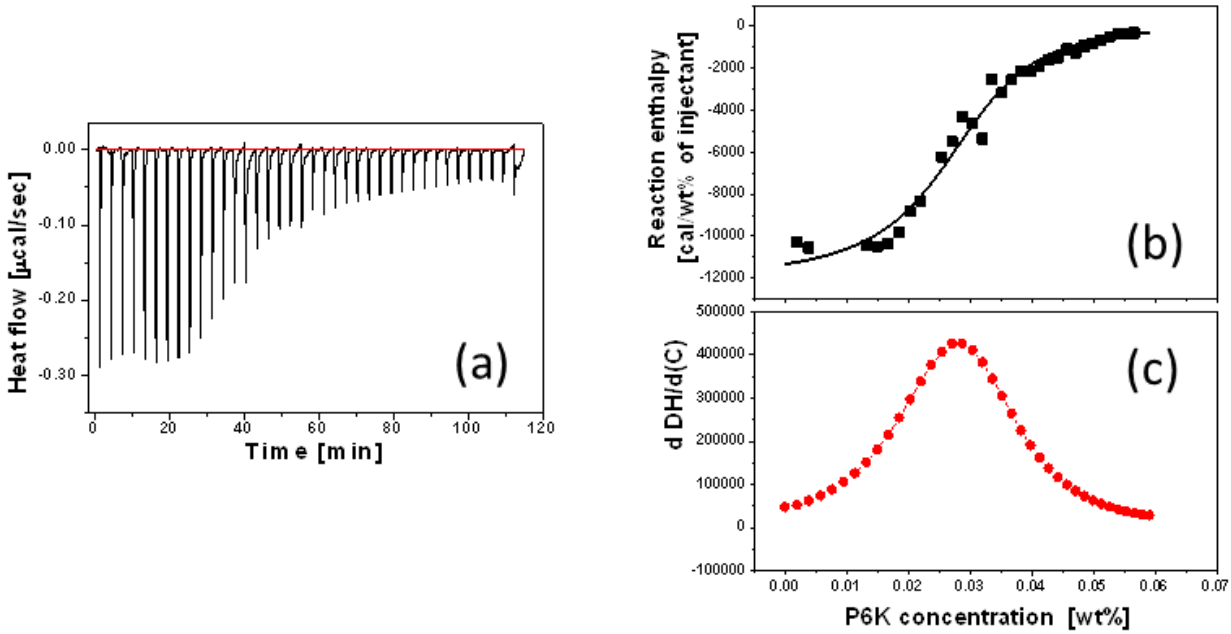


Figure 4. Isothermal Titration calorimetry data for titration of P_6K (0.2 wt%) into water. (a) Calorimetric titration trace, (b) Integrated heat data providing the enthalpy as a function of P_6K concentration (indicated), (c) Derivative of enthalpy with respect to concentration.

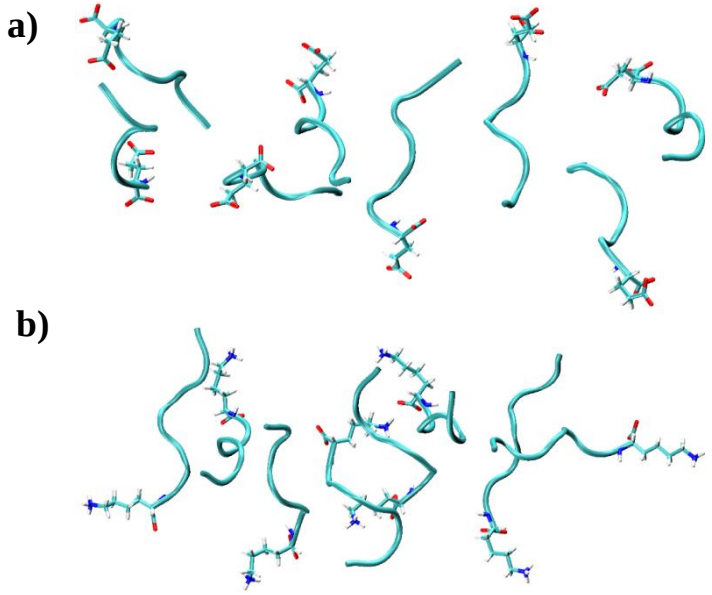


Figure 5. Representative frames of the most populated clusters of (a) 8P6E and (b) 8P6K at 300 K. The last ionic peptide residue is highlighted from the main backbone chain.

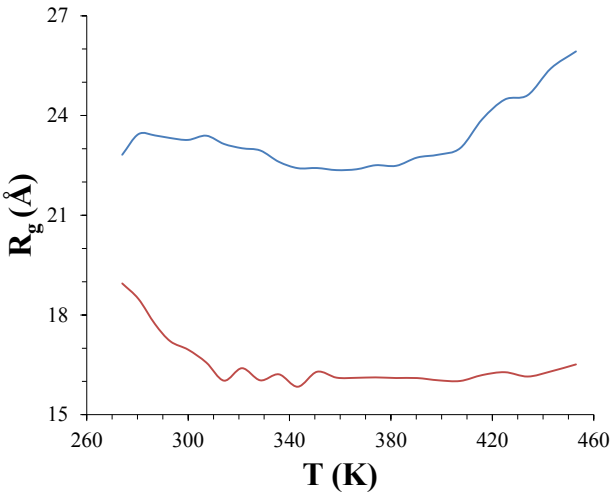
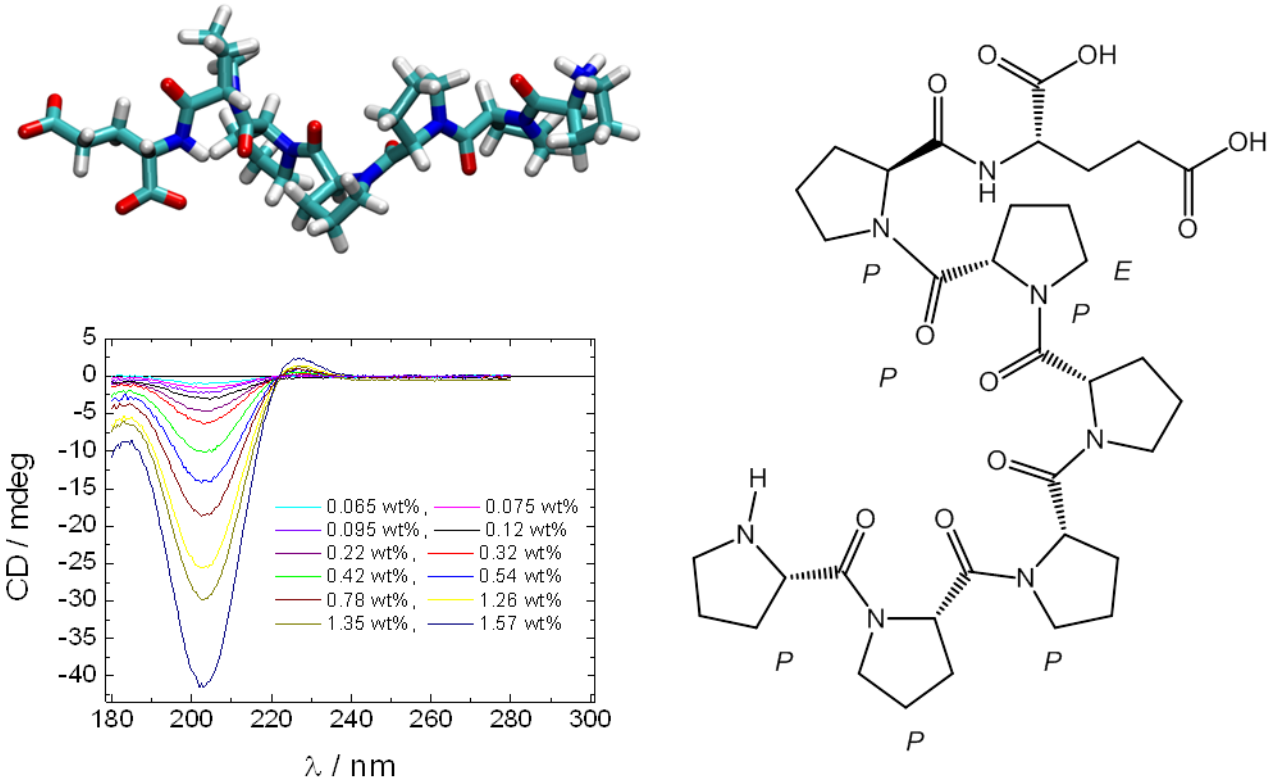


Figure 6. Temperature dependence of averaged radius of gyration of (a) 8P6E (red) and (b) 8P6K systems (blue). Values derived from the most populated cluster (#0) at each REMD temperature.

TOC Graphic


CD / mdeg
λ / nm
0.065 wt%
0.075 wt%
0.095 wt%
0.12 wt%
0.22 wt%
0.32 wt%
0.42 wt%
0.54 wt%
0.78 wt%
1.26 wt%
1.35 wt%
1.57 wt%
P
E
P
P
P
P
P