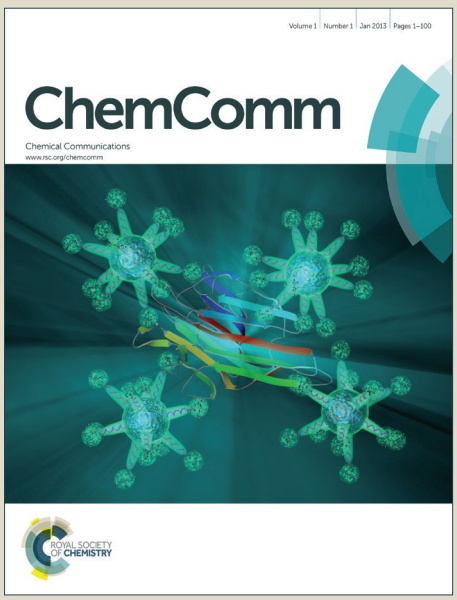
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

COMMUNICATION

Nanofiber Assembly Directed by Non-classical Antiparallel β-Structure from 4S-(OH) Proline Polypeptide

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Nitin D. Bansode,[a] Mahesh V. Sonar[a] and Krishna N. Ganesh[a]*

The antiparallel arrangement of two strands of non-classical β-structure formed exclusively by *cis*-4*S*-(OH) prolyl polypeptide as established by FRET, propagates into self-assembled nanofibers upon conjugation with C12/C14/C16 hydrocarbon chains.

α-Helices, β-sheets and related secondary structural elements in proteins arise from intra and interchain H-bonding of the backbone amide groups.[1] Polyproline peptides being tertiary amides, lack NH groups and hence cannot form standard, well-defined helices or β-sheets mediated through H-bonds. However they adopt helical forms assigned as PPII in water and PPI in organic solvents (eg. n-propanol, trifluoroethanol).[2] These structures arise essentially from steric negotiations of the cyclic prolyl puckers, influencing the ϕ and Ψ backbone dihedral angles. The PPII form prevalent in the collagen triple helix,[3] has growing importance as a preferred transient conformation in protein unfolding.[4] The effects of C4-substituents (F/OH/N_3) on proline pucker in perturbing the PPII/PPI polypeptide conformations has recently attracted considerable attention.[5] We have previously reported that 4-*R*/*S*-(NH_2) proline derived collagen analogs [(X-Y-Gly)$_n$; X/Y=4-NH_2/OH proline] exhibited better triplex stability, depending on the pattern of X/Y substitution, pH and the ionizable NH_2 group.[6] We have unified the different structural parameters such as proline pucker (C4-*endo*/*exo*), H-bonding, n→π* and gauche effects, into a comprehensive model, to explain the triplex forming abilities of such chimeric, cationic collagen analogs.[7]

Continuing our exploration of 4(*R*/*S*)-(NH_2) proline derived polypeptides (**1/2,** X=NH), we recently demonstrated stereo selective, solvent specific formation of unusual β-structure from 4*S*-(NH_2)-prolyl polypeptide (**2**, 4*S*-*amp*$_9$) in trifluoroethanol (TFE), in contrast to PPII conformation in water.[8] In comparison, the 4*R*-(NH_2) prolyl polypeptide (**1**, 4*R*-*Amp*$_9$) exhibited PPII form in both H_2O and TFE. The β-structure of peptide **2** should arise from interchain H-bonding between the side-chain amino group and the backbone carbonyl oxygen (Figure 1, 2B).[8] Its absence in the stereoanalog 4*R*-(NH_2) prolyl peptide **1** (Figure 1, 1A), prompted us to invoke the prime necessity of *cis*-disposition of C4-(NH_2) and the C2-carbonyl substituents in 4*S*-*amp*$_9$ peptide **2** (X=NH) for the formation of the β-structure.[8] To validate this prerequisite, the H-bonding C4-substituent NH_2 was replaced by the alike OH group to synthesise 4*S*-*hyp*$_9$ **2**(X=O) (SI, Schemes 1 and 2). The CD spectra of the resultant 4S-*hyp*$_9$ peptide **2** (X=O) also exhibited β-structure in TFE and PPII form in aqueous medium, similar to 4*S*-*amp*$_9$ **2** (X=NH) (ESI, page S12). This confirmed the obligatory need of H-bonding substituent at C4 in 4*S*-configuration and *cis* to C2-carbonyl amide group for β-structure formation. It should be pointed out that the derived β-structure arises from side chain–backbone H-bonding (Figure 1, 2B) and differs from the classical β-structure formed by H-bonding between the backbone amides of two peptide chains.[1b]

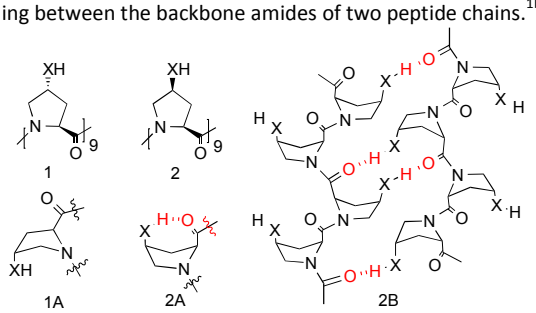


Figure 1. 4(*R*/*S*)-prolyl substituted polypeptides; **1**, 4*R*-*Amp*$_9$ (X=NH); **2**, 4*S*-*amp*$_9$ (X=NH); 4*S*-*hyp*$_9$ (X=O); **1A**, C4(*R*)-*exo*; **2A**, C4(*S*)-*endo* puckers; **2B**, β-structure derived from 4*S*-(X=NH,O)-prolyl peptides

In the β-structure of peptides, the relative orientation of the two chains may be either parallel or antiparallel.[9] In most natural peptides and synthetic models, there is an inherent preference for antiparallel orientation.[10] Although a variety of spectroscopic

[a] *Chemical Biology Unit, Indian Institute of Science Education and Research (IISER) Pune, Dr Homi Bhabha Road, Pune 411008, Maharashtra, India*
† Email: kn.ganesh@iiserpune.ac.in
Electronic Supplementary Information (ESI) available: Structures, synthesis, HPLC and MALDI-TOF of peptides **P1**-**P7**, CD spectra of P1-P7 in buffer and TFE, Fluorescence spectra of peptides **P1**-**P3** in buffer and TFE, SEM of peptides **P4**-**P7** at 300 μM . See DOI: 10.1039/x0xx00000x

techniques (FT-IR,[11] Raman spectroscopy,[12] Circular Dichroism[13]) can fingerprint β-structure, unambiguous assignment of the orientations of the two composing strands is not trivial, except by X-ray diffraction[14] or NMR.[15] We now report a hybrid strategy employing fluorescence spectroscopy (FRET) with imaging (SEM) to clearly distinguish the two orientations and show that the non-classical β-structure in 4*S*-*hyp*$_9$ peptide **2** is antiparallel. Importantly, terminal conjugation with fatty acid chain seeds the formation of nanowires.

Two sets of experiments were designed to delineate the parallel and antiparallel alignment of β-structure (Figure 2): (i) fluorescence resonance energy transfer (FRET) between two peptide strands, each labeled at N-terminus with donor tryptophanyl (Trp) moiety or the acceptor dansyl (Dns) group; FRET signals arise only when the fluorophores are nearby as in parallel β-structure, but not away as in antiparallel orientation (Figure 2A) and (ii) conjugation of the N-terminus with fatty acid chains (C12/C14/C16) to complement the interchain association with additional hydrophobic interactions. A parallel alignment of the peptide chains results in a simple aggregation, while antiparallel association leads to an assembly with alternate stretches of H-bond and hydrophobic interactions (Figure 2B). Such a sequential arrangement reinforcing the binding of two chains should lead to formation of long nanowires.

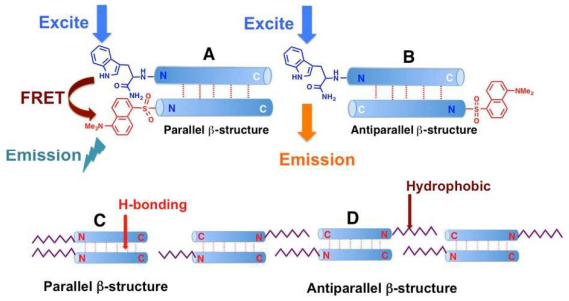
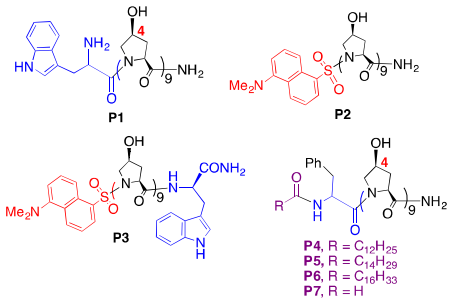


Figure 2. Distinguishing parallel and antiparallel modes. (i) FRET experiment: fluorescence emission from dansyl expected only in parallel orientation (A) and not in antiparallel orientation (B). (ii) Conjugation of fatty acids to reinforce H-bonds (peptide chains) with hydrophobic interactions (fatty chains). Antiparallel alignment (D) leads to long chains and parallel arrangement ends in simple aggregation to short rods (C).

Based on the above premise, the mono-labeled peptides Trp-4*S*-*hyp*$_9$ **(P1)** and Dns-4*S*-*hyp*$_9$ (**P2**) containing tryptophan and dansyl at N-terminus respectively and the double-labelled peptide Dns-4*S*-*hyp*$_9$-Trp **(P3)** with dansyl at N-terminus and Trp at C-terminus were synthesised. The peptides **P4-P6**, having fatty acid (C12/C14/C16) conjugated at N-terminus or N-acetate (**P7**) were synthesised to monitor the orientation dependent morphology through microscopic imaging.



All peptides were synthesized on solid phase using rink amide resin by standard procedures, proceeding from C-terminus to N-terminus by sequential addition of appropriate monomer as per previous protocols[8] (SI, Schemes 1 and 2). The synthesis was completed by coupling with either N-protected tryptophanyl (**P1**) / dansyl (**P2**) monomers or the corresponding fatty acid activated esters (**P4**-**P6**) ($C_{12/14/16}$-4*S*-*hyp*$_9$). The peptide **P3** (Dns-4*S*-*hyp*$_9$-Trp) labeled with Dns at N-terminus and Trp at C-terminus was synthesized by first coupling resin with lysine, and then N-protected tryptophanyl monomer, followed by prolyl monomer couplings and terminated by dansyl coupling at N-terminus. All peptides were cleaved from the resin, purified by RP HPLC and identified by mass spectral data (SI, p3-10). The concentrations of the peptides for CD and fluorescence studies were determined based on the spectral properties of Dns (λ_{max} = 330 nm; ε = 4800 $M^{-1}cm^{-1}$) or Trp (λ_{max} = 280 nm; ε = 5500 $M^{-1}cm^{-1}$) (for peptides **P1**-**P3**) and Phe (λ_{max} = 254 nm; ε = 195 $M^{-1}cm^{-1}$) (for peptides **P4**-**P7**).

The CD spectra of peptides **P1-P6** recorded in sodium phosphate buffer (pH 7.1; 50 μM-250 μM concentration) was similar to that of unconjugated peptide Phe-4*S*-*hyp*$_9$ (**P7**) (SI, p11) with positive and negative bands around 222-225 nm and 205 nm respectively (Figure 3A,C) characteristic of PPII form. The CD spectra of peptides **P1**-**P7** in TFE exhibited a positive band at 195 nm and negative band ~210 nm, typical of β-structure (Figure 3B,D).[9,14a] Thus conjugation of Trp, Dns and fatty acid chains retained the secondary structures of peptides **P1**-**P7** both in buffer (PPII) and TFE (β-structure). The lower ellipticities observed with C14-peptide **P5** compared to Trp-peptide **P1**, may arise from (i) a lower β-content and (ii) lack of trp residue that contributes to ellipticity in CD of peptide **P5**.

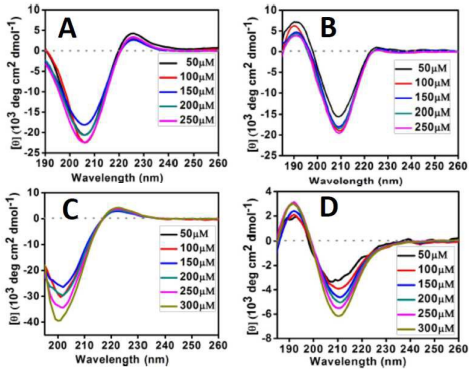


Figure 3. CD spectra of representative peptides **P1** (Trp-4*S*-*hyp*$_9$) and **P5** (C_{14}-4*S*-*hyp*$_9$) in phosphate buffer (pH 7.1) (A and C) and in TFE (B and D) respectively, at different concentrations.

Fluorescence spectra of individual peptides **P1**-**P3** in buffer or TFE were done by excitation of Trp (λ_{ex} 287 nm) or Dns (λ_{ex} 330 nm). The peptides 4*S*-*hyp*$_9$ **P1** and Dns-4*S*-*hyp*$_9$ **P2** in buffer gave the corresponding fluorescence emission bands at 357 nm and 570 nm respectively (ESI, Fig S2). In TFE, the emission bands of **P1** and **P2** were seen at 350 nm and 546 nm respectively with a slight blue shift as expected in a non-polar solvent (ESI, Fig S3). The concentration dependent fluorescence spectra indicated no self-quenching up to 150 mM (ESI, Fig S2 and S3). The conjugated Trp and Dns fluorophores are an efficient FRET pair as seen from an

effective overlap of λ_{em} of donor Trp in **P1** 357 nm with λ_{ex} of acceptor Dns in **P2** at 330 nm) (ESI, Fig S1).

In order to determine the relative orientation of two strands in β-structure, the N-terminus acceptor-tagged **P2** (Dns-4*S*-*hyp*$_9$) was mixed with the donor-tagged **P1** (Trp-4*S*-*hyp*$_9$) in stoichiometric ratios (1:0.5; 1:1) in TFE. The reverse experiment consisted of the donor Trp-tagged **P1** stoichiometrically (1:0.5; 1:1) mixed with the acceptor Dns-tagged peptide **P2**.

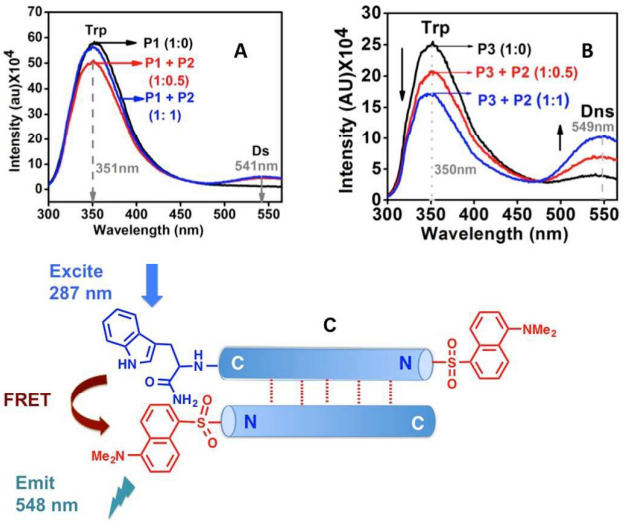


Figure 4. (A) Emission spectra of **P1** (Trp-4S-hyp$_9$) before and after addition of **P2** (Dns-4S-hyp$_9$), excitation λ_{287} nm; emission λ_{350} nm (B) Emission spectra of double labeled **P3** (Trp-4S-hyp$_9$-Ds) before and after addition of **P2** (Dns-4S-hyp$_9$), excitation λ_{287} nm; emission λ_{548} nm (B) antiparallel orientation of peptides in β-structure.

In both experiments (Figure 4A), upon excitation of Trp at λ_{ex} 287 nm, emission was seen only from Trp at λ_{em} 350 nm and not from Dns at λ_{em} at 546 nm (Figure 4A). This ruled out a parallel orientation of the two strands as per Figure 2A, suggesting the alternate antiparallel orientation. A definitive proof for antiparallel alignment was sought from the third experiment involving stoichiometric addition of the double-labelled **P3** (Dns-4*S*-*hyp*$_9$-Trp; N-terminal Dns/C-terminal Trp) to peptide **P2** (Dns-4*S*-*hyp*$_9$; N-terminal Dns). Upon excitation of donor Trp in **P3** at 287 nm, fluorescence emission was seen from the acceptor Dns of **P2** at 548 nm, (Figure 4B). This fluorescence emission from Dns is a result of energy transfer, since the excitation was at Trp, whose emission at 350 nm gets transferred to excite Dns (λ_{ex} 350 nm) and yields its emission at 548 nm. This is also accompanied significant drop in Trp emission at 350 nm, strongly corroborating with the appearance of Dns emission. Because of the specific labelling configurations in **P2** and **P3**, FRET from Trp to Dns is caused via antiparallel arrangement of the two chains (Figure 6C).

To further ratify the antiparallel arrangement, the morphologies of self assembled 4*S*-substituted prolyl polypeptide conjugated with fatty acid chains were examined by FE-SEM imaging. The 4*S*-peptides, **P4** (C_{12}-4*S*-*hyp*$_9$), **P5** (C_{14}-4*S*-*hyp*$_9$), **P6** (C_{16}-4*S*- *hyp*$_9$) and **P7** (4*S*-*hyp*$_9$) that form PPII structure in water (as shown by CD) showed formation of nanorods of 100-200 nm length in water at 50 μM concentration (Figure 5). The C_{16}-peptide **P6** exhibited somewhat longer rods compared to C_{12}-peptide **P4** and C_{14}-peptide **P5**. Thus, the PPII conformation of 4*S*-polyproline peptides seen by CD in water translates into rod type nanostructures, with sizes depending on the length of hydrocarbon chain.

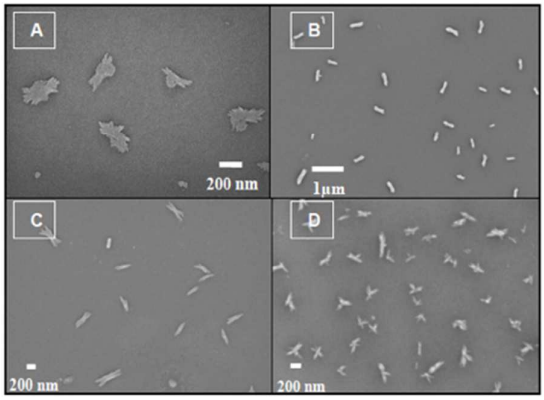


Figure 5. FESEM images for 4S-peptides (A) **P7**, 4S-hyp$_9$, (B) **P4**, (C_{12}-4S-hyp$_9$), (C) **P5**, C_{14}-4S-hyp$_9$, (D) **P6,** C_{16}-4S-hyp$_9$ in water at 50 μM.

In contrast to rod structures seen in water, the fatty acid conjugated polypeptides **P4-P6** exhibited formation of long nanofibers in (TFE) (Figure 6, B-D). The peptide **P7** (4*S*-*hyp*$_9$) devoid of fatty acid chain had no distinct morphological features (Figure 6A) even at higher concentrations. The nanofibers found in TFE should be a consequence of contiguously connected antiparallel β-structures growing from the hydrophobic overlap of N-terminal hydrocarbon chains on adjacent units of β-structure comprised of interchain H-bonds (Figure 6D). FESEM imaging needs samples from drying that promotes highly ordered nanostructures. The results of dynamic light scattering (DLS) studies also supported the formation of assembled structures in solution in TFE (~400 nm), compared to ~50 nm in water (ESI, S18).

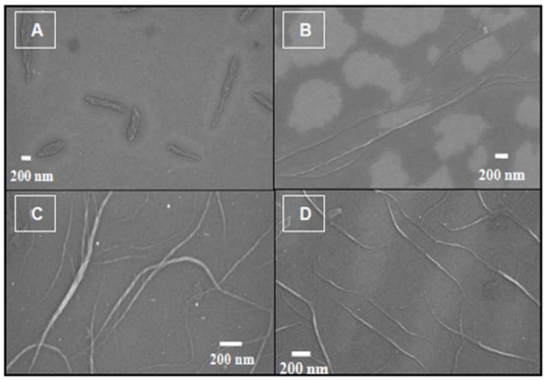


Figure 6. FESEM images for peptides (A) **P7** (4S-hyp$_9$), (B) **P4** (C_{12}-4S-hyp$_9$), (C) **P5** (C_{14}-4S-hyp$_9$) (D), **P6** (C_{16}-4S-hyp$_9$) in TFE at 50 μM**.** For higher concentrations, see ESI, Fig S4)

Peptide **P7** devoid of fatty chain showed only nanorods (Figure 8A) and hence the formation of well-defined and extended nanowires cannot arise simply from interchain hydrogen bonds of β-structure. The covalently linked fatty chain has definite role in seeding the supramolecular creation of nanowires assembled into nanofibers through additional hydrophobic interactions. Nanowire formation is a synergistic effect of hydrogen bonding in the peptide core and hydrophobic interactions at the termini and inducing

aggregation of wires into fibers. Such an ordered assembly into wires/fibers is possible only in antiparallel orientation of two peptide strands, since a parallel orientation would truncate the elongation process, giving short rods (Figure 7).

In conclusion, it is shown that peptides composed of *4S*-(OH) proline adopt an unusual, non-classical, antiparallel β-structure in TFE arising from *inter*chain association of two polyproline peptide

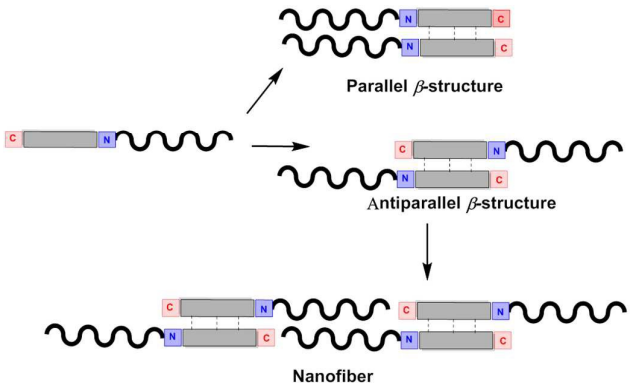


Figure 9. Role of fatty chain in formation of nanowires

strands. The two chains are held by H-bonding between backbone carbonyl and side-chain amino group, possible only in 4*S*-prolyl moieties. Observance of FRET between fluorescent donor (Trp) at C-terminus and acceptor (Dns) at N-terminus provided definitive evidence of antiparallel orientation of the two strands in β-structure. Further, conjugating fatty acid tail at the N-terminus of these peptides led to elongated nanowires and nanofibers in TFE, which can be accounted only by an antiparallel alignment of two strands in β-structure. The PPII form seen in water gave short nanorods.

The current results add a new design principle to a growing repertoire of strategies for engineering peptide structural motifs to create new biomaterials and nanoassemblies.[16] It enlarges the scope of potential applications of polyproline peptides including models for β-structure and collagen.[17] The present findings suggest the development of new tunable nanostructured morphologies based on steric disposition of structural elements.[18] and may help understanding the origin of unusual peptide aggregates. Amyloid fibrils containing misfolded protein with β-sheet structures are the basis of several neurogenerative diseases such as Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD).[19] The present findings of nanofibre formation in polyprolyl peptides may have relevance in understanding the molecular basis for formation of amyloid structures.

NB thanks UGC, New Delhi for a research fellowship. KNG acknowledges DST, New Delhi for award of JC Bose Fellowship and is a Honorary Professor at JNCASR, Bengaluru.

Notes and references

1. (a) L. Pauling, R. B. Corey, H. R. Branson, *Proc. Natl. Acad. Sci. USA*, 1951, **37**, 205-211; (b) J. S. Richardson, Adv. Protein Chem., 1981, **34**, 167-339.
2. (a) S. Knof, J. Engel, *Isr. J. Chem*. 1974, **12**, 165–177; (b) M. Mutter, T. Wohr, S. Gioria, M. Keller, *Biopolymers,* 1999, **51**, 121–128; (c) S. Kakinoki, Y. Hirano, M. Oka, *Polym. Bull*. 2005, **53**, 109–115.
3. (a) B. Brodsky, G. Thiagarajan, B. Madhan, K. Kar, *Biopolymrs,* 2008, **89**, 345–353; (b) M. D. Shoulders, R. T. Raines*, Annu. Rev. Biochem*. 2009, **78**, 929–958.
4. (a) Z. Shi, K. Chen, Z. Liu, N. R Kallenbach, *Chem. Rev*. 2006, **106**, 1877–1897; (b) R. K. Dukor, T. A. Keiderling, *Biopolymers*, 1991, *31*, 1747–1761; (c) S. Krimm, M. L. Tiffany, *Isr. J. Chem*. 1974, **12**, 189–200; (d) A. L. Rucker, T. P. Creamer, *Protein Science*, 2002, **11**, 980-985.
5. (a) H. Jia-Cherng, R. T. Raines*, Protein Sci*. 2006, **15**, 74–83; (b) M. Kuemin, L. S. Sonntag, H. Wennemers, *J. Am. Chem. Soc*. 2007, **129**, 466–467.
6. (a) I. R. Babu, K. N. Ganesh, *J. Am. Chem. Soc.* 2001, **123***,* 2079-2080-81; (b) M. Umashankara, I. R. Babu, K. N. Ganesh, *Chem. Commun***.** 2003, **20**, 2606–2607.
7. M. Umashankara, M. V. Sonar, N. D. Bansode, K. N. Ganesh, *J. Org*. *Chem*. 2015, **80**, 8552−8560.
8. M. V. Sonar, K. N. Ganesh, *Org. Lett*. 2010, **12**, 5390-5393.
9. S. Lifson, C. Sander, *Nature*, 1979, **282**, 109-111.
10. (a) J. S. Nowick, *Acc. Chem. Res*. 2008, **41**, 1319-1330; (d) W. A. Loughlin, J. D. A. Tyndall, M. P. Glenn, D. P. Fairlie, *Chem. Rev*. 2004, **104**, 6085-6117; e) W. A. Loughlin, J. D. A. Tyndall, M. P. Glenn, T. A. Hill, D. P. Fairlie, *Chem. Rev*. 2010, **110**, PR32-PR69.
11. (a) D. Marsh, M. Muller, F. J. Schmitt, *Biophys. J.* 2000, **78**, 2499-2510; (b) L. K. Tamm, S. A. Tatulian, *Q. Rev. Biophys.* 1997, **30**, 365-429.
12. (a) S. A. Overman, M. Tsuboi, G. J. Thomas Jr., *J. Mol. Biol.* 1996, **259**, 331-336; (b) S. H. Lee, S. Krimm, *J. Raman Spectrosc.* 1998, **29**, 73-80; (c) S. A Overman, P. Bondre, N. C. Maiti, G. J. Thomas, *Biochemistry*, 2005, **44**, 3091-3100.
13. (a) A. Micsonai, F. Wien, L. Kernya, Y-H. Lee, T. Goto, M. Refregiers, J. Kardos, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, E3095-E3103; (b) N. J. Greenfield, *Nat. Protoc*. 2006, **1**, 2876-2890.
14. (a) P-N. Cheng, C. Liu, M. Zhao, D. Eisenberg, J. S. Nowick, *Nat. Chem*. 2012, **4**, 927-933; (b) J. S. Richardson, D. C. Richardson, *Biophys. J*. 2014, **106**, 510-525; (c) P-N. Cheng, J. D. Pham, J. S. Nowick, *J. Am. Chem. Soc*. 2013, **135**, 5477-5492.
15. (a) B. Bechinger, C. Aisenbrey, P. Bertani, *Biochim. biophys. Acta,* 2004, **1666**, 190-204; (b) U. H. Durr, K. Yamamoto, S-C. Im, L. Waskell, A. Ramamoorthy, *J. Am. Chem. Soc.* 2007, **129**, 6670-6671.
16. (a) X. Zhao, F. Pan, H. Xu, M. Yaseen, H. Shan, C. A. E. Hauser, S. Zhang, J. R Lu, *Chem. Soc. Rev*. **2010**, *39*, 3480–3498; (b) B. Apostolovic, D. M. Maarten, H.-M. Klok, *Chem. Soc. Rev*. **2010**, *39*, 3541–3550.
17. J. A. Fallas, L. E. R. O'Leary, J. D. Hartgerink, *Chem. Soc. Rev*. 2010, **39**, 3510–3527.
18. (a) H. Tao, S.C. Lee, A. Moeller, R. S. Roy, F. Y. Siu, J. Zimmermann, R. C. Stevens, C. S. Potter, B. Carragher, Q. Zhang, *Nat. Methods*, 2013, **10**, 759-761; (b) Y. Kim, W. Li, S. Shin, M. Lee, *Acc. Chem. Res*. 2013, **46**, 2888-2897.
19. (a) C. A. Ross, M. A. Poirier, *Nat. Med*. 2004, **10**, S10-S17. (b) J. Tyedmers, A. Mogk, B. Bukau, 2010, *Nat. Rev. Mol. Cell. Biol.*, 2010, **11**, 777-788. (c) T. P. J. Knowles, M. Vendruscolo, C. M. Dobson, *Nat. Rev. Mol. Cell. Biol.*, 2014, **15**, 384-396.

