Stereodependent and Solvent Specific Formation of Unusual β-Structure through Side Chain-Backbone H-Bonding in C4(*S*)-(NH_2/OH/NHCHO)-L-Prolyl Polypeptides

Nitin D Bansode, B Madhanagopal, Mahesh V Sonar and Krishna N Ganesh*
Chemical Biology Unit, Discipline of Chemistry,
Indian Institute of ScienceEducation and Research,
Dr. Homi Bhabha Road, Pashan, Pune 411008

Abstract: It is shown that C4(*S*)-NH_2/OH/NHCHO-prolyl polypeptides exhibit PPII conformation in aqueous medium, but in a relatively hydrophobic solvent trifluoroethanol (TFE) transform into an unusual β-structure. The stereospecific directing effect of H-bonding in defining the specific structure is demonstrated by the absence of β-structure in the corresponding C4(*S*)-guanidinyl/(NH/O)-acetyl derivatives and retention of β-structure in C4(*S*)-(NHCHO)-prolyl polypeptides in TFE. The distinct conformations are identified by the characteristic CD patterns and supported by Raman spectroscopic data. The solvent dependent conformational effects are interpreted in terms of intraresidue H-bonding that promotes PPII conformation in water, switching over to interchain H-bonding in TFE. The present observations add a new design principle to the growing repertoire of strategies for engineering peptide secondary structural motifs for innovative nanoassemblies and new biomaterials.

Introduction

The polyproline type II (PPII) helix[1] is a prevalent conformation in both folded and unfolded proteins and an intermediate local conformation during protein unfolding.[2] PPII like structures play an important role in a variety of biological processes, such as signal transduction,[3] transcription,[3] immune response[4] and cell motility.[5] Each strand of collagen triplex consisting of Pro-Hyp-Gly tripeptide repeat motif adopts a left-handed PPII like conformation.[6] Oligoprolines and their derivatives have found utility as cell penetrating agents,[7] specific to target intracellular bacteria[7c] and as molecular spacers in biomimetic systems for energy/electron transport.[8] Recently, phenylproline tetrapeptides, $(PhPro)_4$/(5-Ph)Pro, have been shown to have improved transport across the blood-brain barrier via passive diffusion.[9] The contiguous 4*O*-glycosylation of Hyp residues dramatically increases the thermal stability of the PPII helix.[10] We have recently demonstrated that cationic collagen peptides composed from C4(*R*/*S*)-NH_2/guanidinyl prolyl residues are functional enhancers in transfecting cells with the gene-encoded plasmids.[11] In view of the above objective significance of PPII conformation in oligoprolyl peptides and its conformational importance for collagen structure, we set out to examine the specific role of C4(*R*/*S*)-NH_2/OH/NHCHO/guanidinyl substituents on proline in dictating the conformation of derived homo-polypeptides in different environments.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/bip.22981

Polyprolyl peptides assume two prevalent conformations: (i) the left handed PPII form seen in aqueous medium is a semi-extended structure in which all amide bonds are in the *trans* conformation, nearly perpendicular to the helix axis and (ii) the right-handed PPI form observed in a relatively hydrophobic medium such as isopropanol, with a compact structure composed of all amide bonds in *cis* conformation and oriented parallel to the PPI helix axis.[1] The backbone n(O)→π*(CO) interaction has been suggested to play a key role in stabilizing the PPII helices in C4(*R*)-substituted prolyl peptides by enhancing stereoelectronic effects that favour the *trans* prolyl peptide bond.[12] The recent crystal structure of unsubstituted oligoproline (hexamer) revealed n→π* interaction to be significantly important in addition to hydration in forming PPII helices.[13]

Unsubstituted polyprolines lack H-bond donor sites (NH) in their backbone, leaving unsatisfied H-bond acceptors (carbonyl) free for solvation. They are unable to form either intra or interstrand H-bonds and the observed PPII/PPI helices in prolyl polypeptides arise entirely from sterically favoured dihedral angles. Theoretical studies have even predicted formation of high energy β-sheet like structures in polyproline, for certain combination of dihedral angles, thought such a structure has not been realized experimentally. The presence of H-bonding C4(*S*)-OH[14] or C4(*S*)-NH_2[15] substituents on the proline ring provides opportunity for introducing intra-residue H-bonding that may modulate the PPI and PPII conformations. It is known that substituents at C4 on proline ring affect the *trans/cis* amide ratio towards favoring PPII conformation in water.[15] Electron-withdrawing substituents at C4(*R*) or C4(*S*) site also impact the kinetics of PPII→PPI conversion.[16]

We have earlier reported that C4(*R*)-NH_2-proline at Y-site in collagen peptide stabilizes the triple helix[17] at both acidic and basic pHs. The X/Y-site dependent stability of C4(*R*/*S*)-NH_2/OH substituted cationic collagen peptides arises from an orchestration of several structural factors such as C4(*exo*/*endo*) pucker, stereoelectronic (gauche effect), n→π* interaction and intra-residue H-bonding.[18] In a recent communication,[19] we demonstrated that the C4(*S*)-NH_2-prolyl homo-oligopeptide exists in PPII conformation in water, but shifts to β-structure in a relatively more hydrophobic solvent, 2,2,2-trifluoroethanol (TFE); in contrast, C4(*R*)-aminoprolyl polypeptide retained the PPII form even in TFE. It was postulated that the C4(*S*)-NH_2 group in *cis*-disposition with C2-carbonyl of proline, stabilizes the PPII form *via intra*residue H-bonding in water and switches to *inter*strand H-bonding in TFE inducing β-structure. If such a transition of PPII form to β-structure involves *intra*chain to *inter*chain H-bond rearrangement, it should be as well feasible with other H-bonding C4(*S*)-substituents. To establish the generality of this principle, we have now examined the comparative behavior of the C4(*S*)-OH (*Hyp*), C4(*S*)-NHCHO (*fAmp*) C4(*S*)-$NHCOCH_3$ (*acAmp*) and C4(*S*)-guanidinyl (*Gndp*) substituted prolyl homo-oligopeptides, which all have H-bonding abilities. It is found that only the C4(*S*)-*Amp*/*Hyp*/*fAmp*-proline polypeptides form β-structure in TFE, while polypeptides derived from N/O-acetyl and guanidinyl derivatives retain the PPII conformation in both water and TFE. The corresponding C4(*R*)-substituents in all

cases showed only the PPII form in both solvents. This outcome amplifies the stereospecific role of C4(*S*)-substituent on L-proline directing the secondary structures of the derived polypeptides.

Results

The various C4(*R*/*S*)-substituted prolyloligopeptides **P1-P8** (Figure 1) were synthesized by standard SPPS on rink amide resin using Fmoc chemistry.[17,19] The synthesis of peptides (4*R*)-*Amp*$_9$ (**P1**), (4*S*)-*Amp*$_9$ (**P2**), (4*R*)-*Hyp*$_9$ (**P4**), (4*S*)-*Hyp*$_9$ (**P5**) were achieved using the previously reported N1-(Fmoc)-4(*R*/*S*)NH-(*t*-Boc)-aminoproline,[17,19] N1-(Fmoc)-4(*R*/*S*)-*O*-(*t*-Bu)-hydroxyproline monomers.[18,21] The peptide *Pro*$_9$ (**P3**) was synthesized from the commercially available N1-(Fmoc)-proline and the synthesis of peptide (4*S*)-*Hyp*$_9$ (**P5)** was achieved from (2*S*,4*S*)-N1-(Fmoc)-4-*O*-(*t*-Bu)-hydroxyproline (**4**) (Figure 2). The peptides (4*S*)-*Amp*$_9$ (**P2**) and (4*S*)-*Hyp*$_9$ (**P5**) after purification were treated with acetic anhydride/*N,N*-diisopropylethylamine (DIPEA) in DMF to afford the C4(*S*)-O/N-acetylated peptides (4*S*)-*acHyp*$_9$ (**P6**) and (4*S*)-*acAmp*$_9$ (**P7**). The peptide (4*S*)-*fAmp*$_9$ (**P8**) was synthesized from N1-(Fmoc)-4-formamido proline.[18,21] The peptides, (4*R*)-*Amp*$_9$ (**P9**) and (4*S*)-*Amp*$_9$ (**P10**), synthesized on MBHA resin using *t*Boc strategy with N1-(*t*-Boc)-4(*R*/*S*)NH-(Fmoc)-aminoproline were guanidinylated as reported earlier.[11] Following the deprotection of C4-NHFmoc group on-resin by piperidine, the peptides were treated with N,N'-bis-Boc-1H-pyrazole-1-carboxamidine to obtain peptides **P9** and **P10**.[11] The N-terminus of all peptides carried phenylalanine, whose UV absorption at 257 nm was useful to determine the concentrations of the peptides. All peptides were acetylated at N-terminus after the last coupling and cleaved from the Rink amide resin by 20% TFA in DCM to yield C-amidated peptides. The N/C-end-capping were done to ensure that the observed structural effects arise only from the C4(*R*/*S*)-substituents on proline and not from terminal groups. All peptides were purified by reverse-phase HPLC and their final purity was confirmed by using analytical RP-HPLC on C-18 column. The structural identity of all peptides was confirmed by MALDI-TOF data (Supporting Information, page no. S20).

Formation of PPII structure by C4-substituted prolyl polypeptides P1-P10 in water

The conformations assumed by different peptides (4*R*)-*Amp*$_9$ (**P1**), (4*S*)-*Amp*$_9$ (**P2**), *Pro*$_9$ (**P3**), (4*R*)-*Hyp*$_9$ (**P4**), (4*S*)-*Hyp*$_9$ (**P5**), (4*S*)-*acHyp*$_9$ (**P6**), (4*S*)-*acAmp*$_9$ (**P7**), (4*S*)-*fAmp*$_9$ (**P8**), (4*R*)-*Gndp*$_9$ (**P9**) and (4*S*)-*Gndp*$_9$ (**P10**) were examined by their CD spectra recorded at 100 μM peptide concentrations in aq. NaH_2PO_4 buffer (pH 7.2) (Figures 3 and 4). All CD spectra showed a positive band at 220-230 nm and a strong negative band at 200-210 nm that are characteristic of PPII conformation.[22] The PPII helical content[22] as measured by the intensity of the positive band at 224-227 nm was found to decrease in the order (4*S*)-*Gndp*$_9$ (**P10**) > (4*R*)-*Gndp*$_9$ (**P9**) (4*R*)-*Amp*$_9$ (**P1**)> (4*S*)-*acAmp*$_9$ (**P7**) > (4*R*)-*Hyp*$_9$ (**P4**)> (4*S*)-*Amp*$_9$ (**P2**) > (4*S*)-*fAmp*$_9$ (**P8**) > (4*S*)-*Hyp*$_9$

(**P5**) > (4*S*)-*acHyp*$_9$ (**P6**) > *Pro*$_9$ (**P3**). Thus all C4(*R*/*S*)-substituted prolyl polypeptides **P1**-**P10** exhibited PPII conformation in aqueous medium.

Formation of β-structure in C4(*S*)-substituted prolyl peptides in trifluoroethanol

It is well established that non-aqueous solvents such as isopropanol or trifluoroethanol (TFE) that are relatively hydrophobic compared to water encourage PPI conformation.[12] The CD spectra of peptides **P1**-**P10** were therefore recorded in trifluoroethanol (TFE) at 100 μM peptide concentration and the results are shown in Figures 5 and 6.

The polyprolyl peptide *Pro*$_9$ (**P3**) and the C4(*R*)-substituted peptides (4*R*)-*Amp*$_9$ (**P1**), (4*R*)-*Hyp*$_9$ (**P4**) retained the PPII conformation in TFE with characteristic strong negative band at 204-205 nm and a positive band around 223-226 nm in the CD spectral profile. In contrast, the CD spectra of C4(*S*)-peptides (4*S*)-*Amp*$_9$ (**P2**), (4*S*)-*Hyp*$_9$ (**P5**) and (4*S*)-*fAmp*$_9$ (**P8**) in TFE exhibited altered profiles with a positive band at 197-199 nm, a negative maximum around 214-216 nm and a broad shoulder around 228 nm. This CD pattern, clearly distinct from either PPII or PPI helical profiles and seen only in 100% anhydrous TFE, corresponds to the characteristic CD signature of a β-structure.[23] This secondary structure must arise from an interchain association, which requires H-bonding derived from the donor substituent at C4 of proline in one chain with the acceptor carbonyl group of peptide bond in another chain. Since interchain associations should get enhanced at higher concentrations, the CD spectra of C4(*S*)-(NH_2/OH)-substituted peptides were recorded in TFE from 50 μM to 250/300 μM. The strong negative band at 210 nm present in all the peptides at 50 μM, showed noticeable decrease in intensity at 100-250 μM, accompanied by a shift of 5 nm of the negative band to 215 nm in case of C4(*S*)-peptides (4*S*)-*Amp*$_9$ (**P2**) and (4*S*)-*fAmp*$_9$ (**P8**) and only a slight shift with C4(*S*)-peptide (4*S*)-*Hyp*$_9$ (**P5**) at higher concentrations (Supporting information Figure S3). No changes were seen for the corresponding C4(*R*)-substituted peptides (4*R*)-*Amp*$_9$ (**P1**) and (4*R*)-*Hyp*$_9$ (**P4**) with increase in concentration and these remained in PPII form (Supplementary information Figure S2). In TFE, none of the C4(*R/S*)-substituted and *Pro*$_9$ peptides show any conformational transition to PPI form even after 6 days, unlike that seen for polyproline peptides.[15] The specific shift of negative band seen in peptides **P2** and **P8** at higher concentrations suggests possible formation of higher order β-structure in these peptides, while lower intensity of negative band could be a consequence of mere aggregation.[24]

When solutions of C4(*S*)-peptides **P2**, **P5** and **P8** in TFE (β-structure) were titrated with incremental aliquots of water, the CD profile of β-structure instantly switched over to that of PPII helical form (Figure 7).[19] The occurrence of an isodichroic point at 215 nm, hinted at interconversion of the two forms. The interchain H-bonds causing the β-structure of C4(*S*)-substituted peptides **P2**, **P5** and **P8** in TFE are obliterated by water, leading to loss of β-structure in aqueous medium and reformation into PPII form by reinstatement of intraresidue H-bonds. To further substantiate the involvement of H atoms in C4(*S*)-NH_2/OH in interchain hydrogen

bond of β-structure, peptides (4*S*)-*Amp*$_9$ (**P2**) and (4*S*)-*Hyp*$_9$ (**P5**) were N/O-acetylated to obtain (4*S*)-*acAmp*$_9$ (**P7**) and (4*S*)-*acHyp*$_9$ (**P6**) respectively. The peptides **P6** and **P7** retained PPII helical form in both water and TFE, failing to display β-structure in TFE (Figure 6). The inability of O-acetyl peptide (4*S*)-*acHyp*$_9$ (**P6**) to form β-structure is evident due to the absence of H-bonding OH, but the failure of the N-acetyl peptide (4*S*)-*acAmp*$_9$ (**P7**) to form β-structure is surprising since the amide NH (in $NHCOCH_3$) is a better H-bond donor. We supposed that the steric bulk of acetyl group prevents the realization of β-structure in peptide **P7** and hence replaced the acetyl group by the less bulky formyl group in peptide (4*S*)-*fAmp*$_9$ (**P8**). Interestingly, this showed β-structure in TFE, similar to the peptides **P2** and **P5**. The C4(*S*)-guanidine group with stronger cationic substituent should have better H-bonding potential than C4(*S*)-NH_2, but C4(*S*)-guaunidino peptide (4*S*)-*Gdnp*$_9$ **P10** showed PPII form in both water and TFE. The failure to form β-structure in TFE reiterated that C4(*S*) bulky substituents cannot be accommodated in β-structure (Figure 6).

All the peptides (**P1**-**P10**) contained phenylalanine at N-terminus which is known to drive the self-assembly induced by CH-π interactions with adjacent proline.[25] Hence the peptide (4*S*)-*fAmp*$_9$ (**P11)** without phenylalanine at N-terminus was synthesized and its CD spectrum in TFE still exhibited characteristic features of β-structure (Supporting Information, Figure S4) and retained PPII conformation in aqueous medium. These results cumulatively imply that in C4(*S*)-substituted peptides **P2**, **P5** and **P8**, the β-structure originates from an interchain association favoured by the *cis* disposed C4(*S*)-NH_2/OH/NHCHO and the amide carbonyl on same proline residue, not realizable in the corresponding C4(*R*)-NH_2/OH peptides.

Raman spectroscopic studies of peptides in TFE

Additional evidence for exceptional formation of β-structure in polyproline peptides was sought from Raman spectroscopy, which has been amply used for characterization of protein conformations, with well established correspondence between the amide spectroscopic bands and the type of secondary structure.[26] The amide I band in the region 1630 -1640 cm^{-1} arises from α-helix, 1640–1660 cm^{-1} corresponds to random coil and band around 1660-1675 cm^{-1} corresponds to β-structure. The blank Raman spectra of polyproline peptides **P1**-**P5** in TFE recorded using 532 nm frequency-doubled Nd:YAG laser were carefully subtracted from the recorded spectra of C4(*R/S*)-NH_2/OH peptides in TFE (Figure 8). For the C4(*S*)-peptides (4*S*)-*Amp*$_9$ (**P2**) and (4*S*)-*Hyp*$_9$ (**P5**), the observed signals around 1670-1672 cm^{-1} and 1228-1230 cm^{-1} for amide I and amide III bands respectively are signatures of β-structure and these were absent in the spectra of C4(*R)*-peptides (4*R*)-*Amp*$_9$ (**P1**) and (4*R*)-*Hyp*$_9$ (**P4**). These results gave further evidence for β-structure formation exclusively by C4(*S*)-NH_2/OH peptides in TFE and absent in C4(*R*)-NH_2/OH peptides.

It is well known that Thioflavin-T (ThT) specifically stains aggregates formed by peptide β-sheets.[27] and the C4(*S*)-peptides were therefore assayed with Thioflavin-T in TFE. The fluorescence emission spectra of the peptides **P2** and **P5** were recorded after incubating with the dye ThT for 1 hour in TFE or aqueous phosphate by exciting

at 420 nm or 440nm respectively (Figure 9). It was seen that the fluorescence spectra of C4(*S*)-peptides **P2** and **P5** in phosphate buffer (pH 7.2) had low fluorescence, while that in TFE showed 6-8-fold enhancement of fluorescence intensity. This provides additional support for β-structure formation by C4(*S*)-peptides in TFE.

Discussion

Polyproline peptides adopt PPII conformation in aqueous solutions and prefer PPI conformation in a more hydrophobic solvent. Being tertiary amides, these peptides are devoid of H-bond donor sites and have only acceptor carbonyl groups in the amide backbone. Thus unlike other helical structures found in peptides (α, 3_{10}, π etc), PPII and PPI helices are formed without any intrachain H-bonds and dictated solely by steric compulsions. Introduction of H-bond donors such as NH_2/OH in proline side-chain should induce intra/interchain H-bonding abilities, thereby modulating the secondary structure in such substituted polyproline peptides. The present study involves comparative conformational features of homo-oligopeptides derived from C4(*R*/*S*)-NH_2/OH/NHCHO substituted prolines in aqueous and non-aqueous (TFE) solvents.

In C4(*S*)-peptides **P2**, **P5**, **P7** and **P8**, the C4(*S*)-NH_2/OH/*OAc*/*NHCHO* substituents and the C2-amide carbonyl are in *cis*-disposition, promoting *intraresidue* H-bonding that is not feasible in C4(*R*)-peptides (**P1** and **P4**) where the C2 and C4-substituents are *trans* to each other (Figure 10A, and 10D). The C4(*R*)-peptides exhibit PPII structure in water due to favoured stereoelectronic effects, while PPII form in C4(*S*)-peptides is favoured by intraresidue H-bonding (Figure 10C, and 10E). The relative propensities of PPII helices formed by these peptides in aqueous solvent, decreased in the order **P10>P9>P1**>**P7**>**P4**>**P2**>**P8>P5**>**P6**>**P3**.

In a fluorinated hydrophobic solvent TFE that favours *interchain* H-bonding, arising from the C4(*S*)-(NH_2/OH/NHCHO) substituent on prolines on one chain and the C2-amide carbonyl of another chain, leads to formation of β-structure in C4(*S*)-peptides **P2**, **P5** and **P8**. Since this is seen only in case of C4(*S*)-peptides, the *trans* disposition of C4(*R*)-substituents and C2-amide carbonyl group does not seem to be conducive for interchain H-bonding. The prerequisite of H on C4(*S*)-substituent for β-structure is supported by the failure C4(*S*)-O-acetyl peptide (4*S*)-a*cHyp_9* (**P6**) to form β-structure. The absence of β-structure in C4(*S*)-N-acetyl peptide (4*S*)-a*cAmp_9* (**P7**) and its presence in N-formyl peptide **P8** suggests that bulky substituents at C4 inhibit β-structure, although the amide nature of N-acetyl substituent is expected to promote stronger H-bonding. However this feature enhances the propensity of **P7** and **P8** to form PPII among all C4(*S*)-peptides. Raman spectroscopic data, showing characteristic bands in TFE at 1228-1230 cm^{-1} (amide I) and 1670-1672 cm^{-1} (amide II), seen specifically for peptides (4*S*)-*Amp_9* (**P2**) and (4*S*)-*Hyp_9* (**P5**) and absent in C4(*R*)-peptides along with the observed 6-8 fold enhancement in fluorescence intensity for C4(*S*)-peptides (**P2** and **P5**) in TFE with ThT provided additional support for β-structure.

The growth of β-structure content (and eventual saturation) with increased concentrations of C4(*S*)-peptides supports an associated intermolecular process. The switching of β-structure back to PPII by addition of tiny amounts of water implies disruption of interchain H-bonds assembling the β-structure. In aqueous solution, no β-structure was seen even at high peptide concentrations. Thus β-structure involving *interchain* H-bonding is promoted under absolute anhydrous environment, with TFE providing this by local dehydration, which is mitigated in aqueous environment. Molecular dynamics simulation studies on the effects of TFE in stabilizing secondary structures of peptides point to preferential accumulation of TFE molecules around the peptides.[28] Such a coating of TFE displaces water as alternative hydrogen-bonding partners providing a low dielectric environment to strengthen H-bonds. In case of C4(*S*)-substituted prolines, H-bonding substituents reinforce the NH_2/OH---OC *intraresidue* H-bonds and initiate new *interchain* H-bonds. This leads initially to two-stranded β-structure at lower peptide concentration that propagates to sheet like aggregates at higher concentrations as indicated by broadening of CD.

The rapid formation of β-structure by C4(*S*)-peptides in TFE with increased concentration and its instant conversion to PPII form on addition of water may be attributed to two types of structural changes: change in amide bond geometry (*trans*↔*cis*) and/or change in proline pucker (C4-*endo* ↔ C4-*exo*). Conversion of PPII to PPI form in TFE/isopropanol requires change of peptide bond from *trans* to *cis* geometry, which is kinetically slow and takes days.[16] Upon storage of C4(*R/S)*-peptides in water or TFE even for six days, no PPI form was observed. The instantaneous conversion between β-structure and PPII form in presence of even tiny amounts of water implies retention of the *trans* amide bond of PPII form in β-structure. The CD spectrum of peptide (4*S*)-*fAmp*$_9$ (**P11**) devoid of phenylalanine hinted that N-terminal phenylalanine has no role in nucleating the β-structure.

Figure 11 illustrates the possible structures of C4(*S*)-substituted polypeptides (X=O, NH, NCHO) in water and TFE. The PPII structure (Figure 11A) in water maximizes all favorable interactions C4-*endo* pucker, gauche effect, n→π* interaction[18] and intraresidue (C4)-NH_2---OC-(C2) H-bond. The classical β-structure in peptides arises from reciprocal H-bonding among the backbone amides of two polypeptide chains. The proposed β-structure here (Figure 11B) is unusual in the sense that it originates from H-bonding of acceptor carbonyl amide on backbone with the side chain donor amino group on the opposite chain. By employing FRET and FESEM, we have recently shown that the orientation of two strands of β-structure in (4*S*)-*Hyp*$_9$ is antiparallel.[29] In this orientation, by retaining the C4-*endo* pucker, C4(*S*)-NH_2/OH/NHCHO peptides can in principle form both interchain and intraresidue H-bonds, (Figure11B and 11D). Alternate possibility of change in proline pucker to C4-*exo* places the axial C4(*S*)-H-bonding substituent in equatorial position, wherein it can access the C2-amide carbonyl of another chain, but with loss of intraresidue H-bonding (Figure11C). The constituent chains of β-structure can further interact with more chains (Figure 11D) to grow into sheets or fibers at higher concentrations. The shift in negative band at 210 nm and broadened shoulder in the CD profiles of

peptides (4*S*)-*Amp*$_9$ (**P2**) and (4*S*)-*fAmp*$_9$ (**P8**) at higher concentrations are perhaps indicative of such possibilities.

An ideal PPII helix has dihedral angles φ and ψ of -75° and +145°, respectively, and all amide bonds are in *trans* conformations (ω= 180°).[12] Here we conjecture that the two chains in β-structure may have similar dihedral angles as in PPII conformation, and β-structure may arise from hydrogen bonding between side chain (OH/NH) groups of one chain and the backbone carbonyl oxygen of another chain. We emphasize that the present results do not provide any direct experimental evidences for the microstructural details of proline pucker, amide bond geometry or the exact hydrogen bonding patterns, but the antiparallel orientation of the two chains of β-structure is evident from our recent work.[29]

Conclusions

Prolyl polypeptides are well known to exhibit PPII conformation in water and PPI form in a more hydrophobic solvent TFE or isopropanol, the helical structures arising entirely from steric constraints imposed by proline pucker and amide bond geometry. In C4(*S*)-(NH_2/OH/NHAc/NHCHO/OAc/NH-guanidino) substituted prolyl peptides **P1-P10,** the H-bonding substituents being *cis* to C2-amide carbonyl form intraresidue H-bonding in a C4-*endo* ring pucker and adopt PPII structure in aqueous medium. In a relatively more hydrophobic medium (TFE), the peptides C4(*S*)-(NH_2/OH/NHCHO) exhibit β-structure, arising from interchain H-bonding, unprecedented for any prolyl polypeptides. This unusual β-structure in TFE is a consequence of *inter*chain hydrogen bonds from the side-chain C4(*S*)-(NH_2/OH/NHCHO) groups and the backbone C2-amide carbonyl. The solvent specific structural change illustrates a fine balance between stereoelectronic and H-bonding effects in tuning the secondary structure of C4(*R*/*S*)-NH_2/OH/NHCHO proline polypeptides. Such unusual β-structure as a new motif in peptide secondary structures is amenable for aggregation to higher order structures akin to classical β-sheet. The current results of specific β-structure in a hydrophobic medium, may have potential applications of these peptides in cell penetration (in hydrophobic membrane environment) and in developing C4-substituted proline polypeptides as stimulus responsive materials. The present outcomes also add a new design principle to a growing repertoire of strategies for engineering peptide secondary structural motifs for new biomaterials and nanoassemblies.[31]

Acknowledgements

NDB acknowledges UGC, New Delhi for award of a fellowship. BM thanks IISER Pune for research fellowship. MVS acknowledges CSIR, New Delhi and LTMT, Mumbai and KNG is a recipient of JC Bose fellowship of DST, New Delhi and Honorary Fellow, JNCASR, Bengaluru.

Notes and references

1 (a) Cowan, P. M.; McGavin,S. Nature 1955, 176, 501-503; (b) Traub,W.; Shmueli, U. Nature 1963, 198, 1165-1166. (c) Zhong, H.; Carlson, H. A., J. Chem. Theory Comput. 2006, 2, 342-353.

2 (a) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. Chem. Rev. 2006, 106, 1877-1897; (b) Dukor, R. K.; Keiderling, T. A. Biopolymers 1991, 31, 1747-1761; (c) Krimm, S.; Tiffany, M. L. Isr J. Chem. 1974, 12, 189–200; (d) Rucker, A. L.; Creamer, T. P. Protein Sci. 2002, 11, 980-985.

3 Kay, B. K.; Williamson, M. P.; Sudol, M. FASEB J. 2000, 14, 231–241.

4 (a) Srinivasan, M.; Eri, R.; Zunt, S. L.; Summerlin, D.; Brand, D. D.; Blum, J. S. Arthritis Rheum 2007, 56, 498-508; (b) Jardetzky, T. S.; Brown, J. H.; Gorga, J. C.; Stern, L. J.; Urban, R. G.; Strominger, J. L.; Wiley D. C. Proc Natl. Acad. Sci. USA. 1996, 93, 734-738.

5 Holt, M. R.; Koffer, A. Trends Cell Biol. 2001, 11, 38–46.

6 (a) Brodsky, B.; Thiagarajan, G.; Madhan, B.; Kar, K. Biopolymers 2008, 89, 345–353; (b) Shoulders, M. D.; Raines, R. T. Ann. Rev. Biochem. 2009, 78, 929–958.

7 (a) Farrera-Sinfreu, J.; Giralt, E.; Castel, S.; Albericio, F.; Royo, M. J. Am. Chem. Soc. 2005, 127, 9459–9468; (b) Fillon, Y. A.; Anderson, J. P.; Chmielewski, J. J. Am. Chem. Soc. 2005, 127, 11798–11803; (c) Nepal, M.; Thangamani, S.; Seleem, M. N.; Chmielewski, J. Org. Biomol. Chem. 2015, 13**,** 5930-5936.

8 (a) Doose, S.; Neuweiler, H.; Barsch,H.; Sauer, M. Proc. Natl. Acad. Sci. USA. 2007, 104, 17400–17405; (b) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. Proc. Natl. Acad. Sci. USA 2005, 102, 2754–2759.

9 Gibert, P. A.; Guixer, B.; Malakoutikhah, M.; Muttenthaler, M.; Guzman, F.; Teixidó, M.; Giralt, E. J. Am. Chem. Soc. 2015, 137, 7357−7364.

10 Owens, N. W.; Stetefeld, J.; Lattová, E.; Schweizer, F. J. Am. Chem. Soc. 2010, 132, 5036–5042.

11 (a) Nanda, M. N.; Ganesh, K. N. J. Org. Chem. 2012, 77, 4131-4135; (b) Umashankara, M.; Nanda, M. N.; Sonar, M. V.; Ganesh, K. N. Chimia, 2012, 66, 936-940.

12 Horng, J.-C.; Raines, R. T. Protein Sci. 2006, 15, 74–83.

13 Wilhelm, P.; Lewandowski, B.; Trapp, N.; Wennemers, H. J. Am. Chem. Soc. 2014, 136, 15829-15832.

14 Shoulders, M. D.; Kotch, F. K.; Choudhary, A.; Guzei, I. A.; Raines, R. T. J. Am. Chem. Soc. 2010, 132, 10857–10865.

15 Kuemin, M.; Nagel, Y. A.; Schweizer, S.; Monnard, F. W.; Ochsenfeld, C.; Wennemers, H. Angew. Chem. Int. Ed 2010, 49, 6324–6327.

16 Chiang, Y.C.; Lin, Y.-J.; Horng, J.-C. Protein Sci2009, 18, 1967–1977.

17 (a) Babu, I. R.; Ganesh, K. N. J. Am. Chem. Soc. 2001, 123, 2079-2080; (b) Umashankara, M.; Babu, I. R.; Ganesh, K. N. Chem. Comm. 2003, 20, 2606–2607.

18 Umashankara, M.; Sonar, M. V.; Bansode, N. D.; Ganesh, K. N. J. Org. Chem. 2015, 80, 8552−8560.

19 Sonar, M. V.; Ganesh, K. N. Org. Lett. 2010, 12, 5390-5393.

20 Seki, M.; Matsumoto, K. Biosci Biotech. Biochem 1995, 59, 1161-1162.

21 Erdmann, R. S.; Wennemers, H. Angew. Chem. Int. Ed 2011, 50, 6835−6838.

22 Woody, R. W.; Adv. Biophys. Chem. 1992, 2, 37–79.

23 (a) Seebach, D.; Overhand, M.;Kühnle, F. N. M.; Martinoni,B.Oberer, L.; Hommel, U.; Widmer, H. Helv. Chim. Acta, 1996, 79, 913–941; (b) Woody, R. W. Method Enzymol. 1995, 246, 34-71.

24 (a) Casallanovo, F.et al. Biopolymers 2006, **84**, 169-180; (b) Davies,R. P. W.; Aggeli, A.J Pept. Sci. 2011, 17, 107-114; (c) Aggeli, A.; Nyrkova, I. A.; Bell, M.; Harding, R.; Carrick, L.; McLeish, T. C. B.; Semenov, A. N. ; Boden,N. Proc. Natl. Acad. Sci. USA, 2001, 98, 11857-11862.

25 Pandey, A. K.; Thomas, K. M.; Forbes, C. R.; Zondlo, N. J. Biochemistry 2014, 53, 5307–5314.

26 Podstawka, E.; Ozaki, Y.; Proniewicz, L. M. Appl. Spect. 2005, 59, 1516-1526; (b) Podstawka, E.; Ozaki, Y.; Proniewicz, L. M.; Appl. Spect. 2004**,** 58, 570-580.

27 (a) Wolfe, L. S.; Calabrese, M. F.; Nath, A.; Blaho, D. V.; Miranker, A. D.; Xiong,Y. Proc. Natl. Acad. Sci. USA, 2010, 107, 16863-16868; (b) Sabate, R.; Rodriguez-Santiago, L.;Sodupe, M.; Saupe, S. J.; Ventura, S. Chem Commun. 2013, 49, 5745-5747; (c) Biancalana, M.; Koide, S.; Biochim. Biophys. Acta 2010, 1804, 1405-1412.

28 Roccatano,D.;Colombo,G.;Fioroni,M.; Mark, A. E. Proc. Natl. Acad. Sci. USA 2002, 99, 12179-12184; (b) Luo, P.; Baldwin, R. L. Biochemistry 1997, 36, 8413-8421.

29 Bansode, N. D.; Sonar, M. V.; Ganesh, K. N. Chem. Commun. 2016, 52, 4884-4887.

30 Kuemin, M.; Schweizer, S.; Ochsenfeld, C.; Wennemers, H., J. Am. Chem. Soc. 2009, 131, 15474-15482.

31(a)Tsai, C.J.; Zheng, J.; Zanuy, D.; Haspel,N.; Wolfson, H.; Alemán, C.; Nussinov, R. Proteins 2007, 68, 1–12; (b) Zhao, X.; Pan, F.; Xu, H.; Yaseen, M.; Shan, H.; Hauser, C. A. E.; Zhang, S.; Lu, J. R. Chem. Soc. Rev. 2010, 39, 3480–3498; (c) Apostolovic, B.; Danial, M.; Klok, H.-A. Chem. Soc. Rev. 2010, 39, 3541–3575; (d) Zotti, M. D.; Formaggio, F.; Crisma, M.; Peggion, C.; Moretto, A.; Toniolo, C., J Pept. Sci. 2014, 20, 307-322.

FIGURE 1. Chemical structures of peptides **P1-P10**

FIGURE 2. Synthesis of C4(*S*)-(O*tBu*)-prolyl monomer

FIGURE 3. CD profiles of peptides **P1-P5** and **P9** at a concentration of 100 μM in phosphate buffer (pH 7.2): (a) (4*R*)-*Hyp*$_9$, **P1**; (b) (4*S*)-*Amp*$_9$, **P2**; (c) *Pro*$_9$, **P3**; (d) (4*R*)-*Amp*$_9$, **P4**; (e) (4*S*)-*Hyp*$_9$, **P5**; (f) (4*R*)-*Gndp*$_9$, **P9**

FIGURE 4. CD profiles of peptides **P6-P8** and **P10** at a concentration of 100 μM in phosphate buffer (pH 7.2): (a) (4*S*)-*fAmp*$_9$, **P6**; (b) (4*S*)-*acAmp*$_9$, **P7**; (c) (4*S*)-*acHyp*$_9$, **P8**; (d) (4*S*)-*Gndp*$_9$, **P10**

FIGURE 5. CD profiles of peptides **P1-P5** and **P9** at a concentration of 100 μM in TFE: (a) (4*S*)-*Amp*$_9$, **P1**; (b) (4*S*)-*Hyp*$_9$, **P2**; (c) (4*R*)-*Amp*$_9$, **P3**; (d) (4*R*)-*Hyp*$_9$, **P4**; (e) *Pro*$_9$, **P5**; (f) (4*R*)-*Gndp*$_9$, **P9**.

FIGURE 6. CD profiles of peptides **P6-P8** and **P10** at a concentration of 100 μM in TFE: (a) (4*S*)-*fAmp*$_9$, **P6**; (b) (4*S*)-*acAmp*$_9$, **P7**; (c) (4*S*)-*acHyp*$_9$, **P8**; (d) (4*S*)-*Gndp*$_9$, **P10**.

FIGURE 7.(A) CD spectra of peptide (4*S*)-*fAmp*$_9$ (**P8**) in TFE with incremental addition of phosphate buffer (pH 7.2) from 0.2%-2.5% (a-f) and (g) 3.0%. (B) CD spectra of peptide (4*S*)-*Hyp*$_9$ (**P5**) in TFE with incremental addition of phosphate buffer (pH 7.2) from 0.2%-2.0% (a-e) and (f) 4.0%.

FIGURE 8. Raman spectra of C4(*R/S*)-NH_2/OH prolyl peptides: (A) (4*R*)-*Amp*$_9$ (**P1**) and (4*S*)-*Amp*$_9$ (**P2**); (B) (4*R*)-*Hyp*$_9$ (**P4**) and (4*S*)-*Hyp*$_9$ (**P5**).

FIGURE 9. Fluorescence emission spectra of ThT in the presence of peptides **P2** and **P5** in phosphate buffer (pH 7.2) and TFE: (a) (4*S*)-*Amp*$_9$ (**P2**) in phosphate buffer; (b) (4*S*)-*Hyp*$_9$ (**P5**) in phosphate buffer; (c) (4*S*)-*Amp*$_9$ (**P2**) in TFE; (d) (4*S*)-*Hyp*$_9$ (**P5**) in TFE.

FIGURE 10. Factors dictating conformational features of C4(*R/S*)-substituted proline polypeptides. (A)C4(*R*)-substitution favours C4-*exo*, gauche and $n\rightarrow\pi^*$ interactions in trans amide geometry (B) C4(*S*)-substitution favors C4-*endo* form with loss of gauche, and $n\rightarrow\pi^*$ interactions (C) H-bonding between C4(*S*)-substituent and C1-amide carbonyl restores favorable amide geometry for $n\rightarrow\pi^*$interaction (D) and (E) cumulative effects that favor PPII form in C4(*R*)- and C4(*S*)-substituted prolyl polypeptides respectively.

FIGURE 11. Possible structures of C4(*S*)-*aminoproline* polypeptide in (A) PPII form in water, (B) antiparallel β-structure in TFE, C4-endo pucker, intraresidue and interchain H-bonding (C) parallel β-structure in TFE, C4-exo pucker, interchain H-bonding. In all structures, the *trans* geometry of tertiary amide bond is retained and (D) structure of possible higher aggregates of antiparallel β-structure through extended H-bondings indicated by arrows

Figure 12. For Graphic Table of Contents

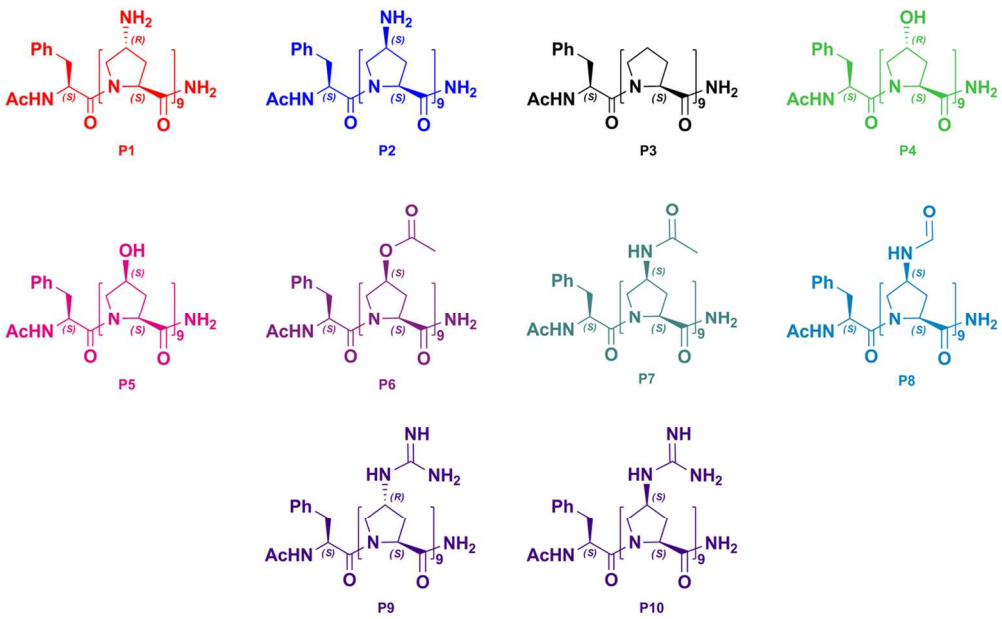


Figure 1

119x73mm (300 x 300 DPI)

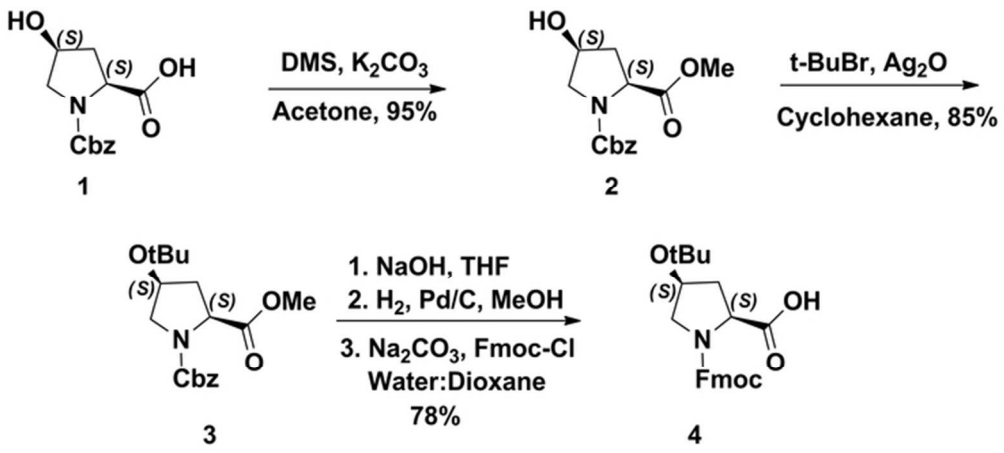


Figure 2

62x28mm (300 x 300 DPI)

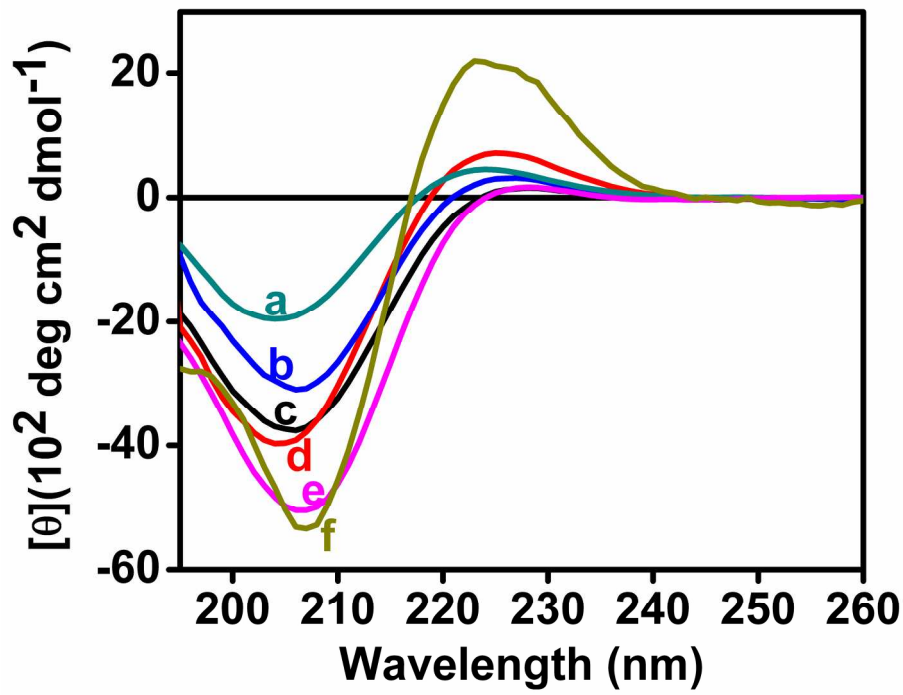


Figure 3

201x156mm (300 x 300 DPI)

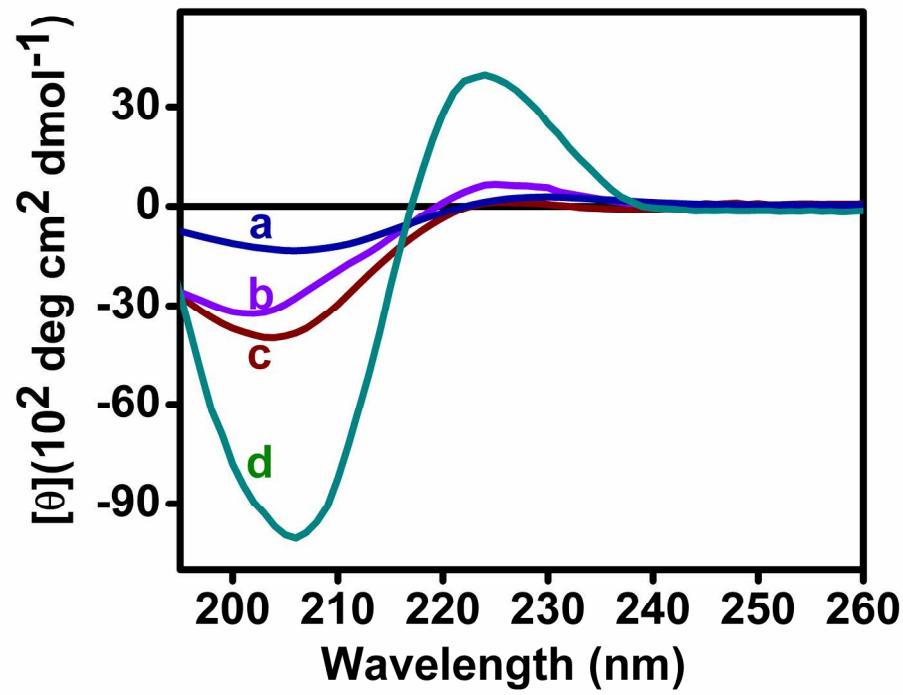


Figure 4

201x156mm (300 x 300 DPI)

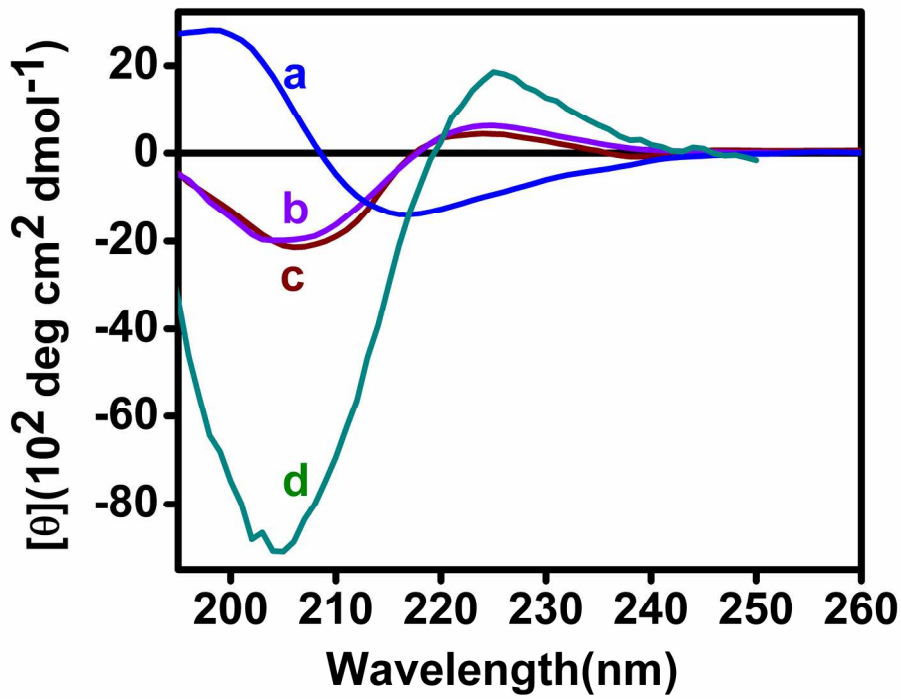


figure 6

201x156mm (300 x 300 DPI)

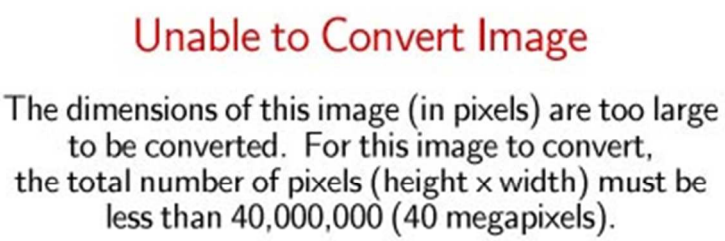
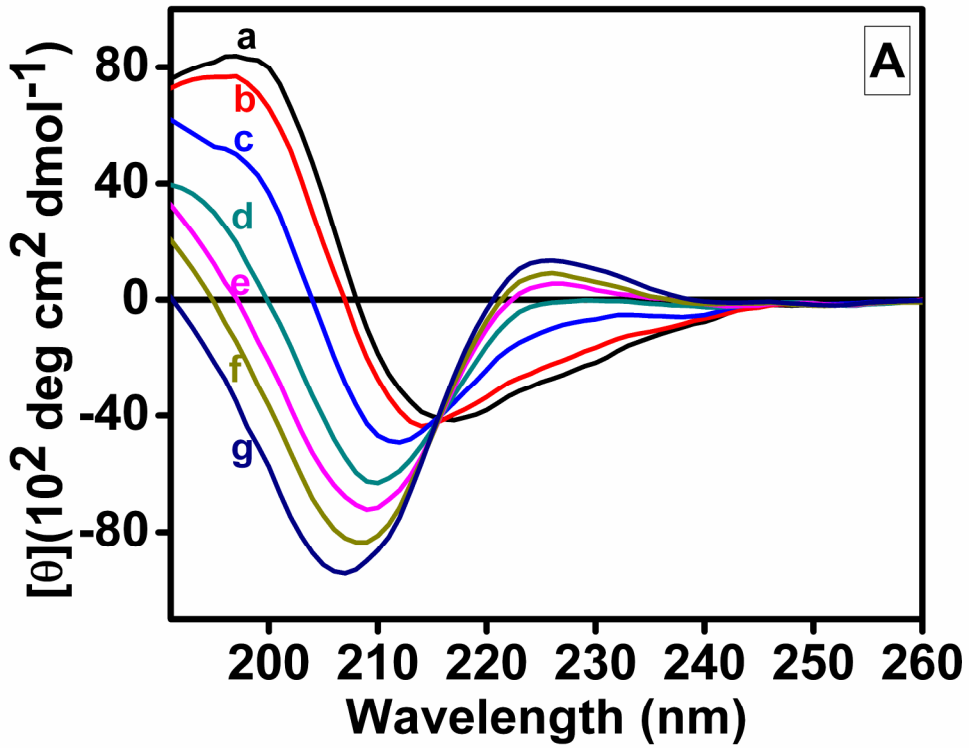
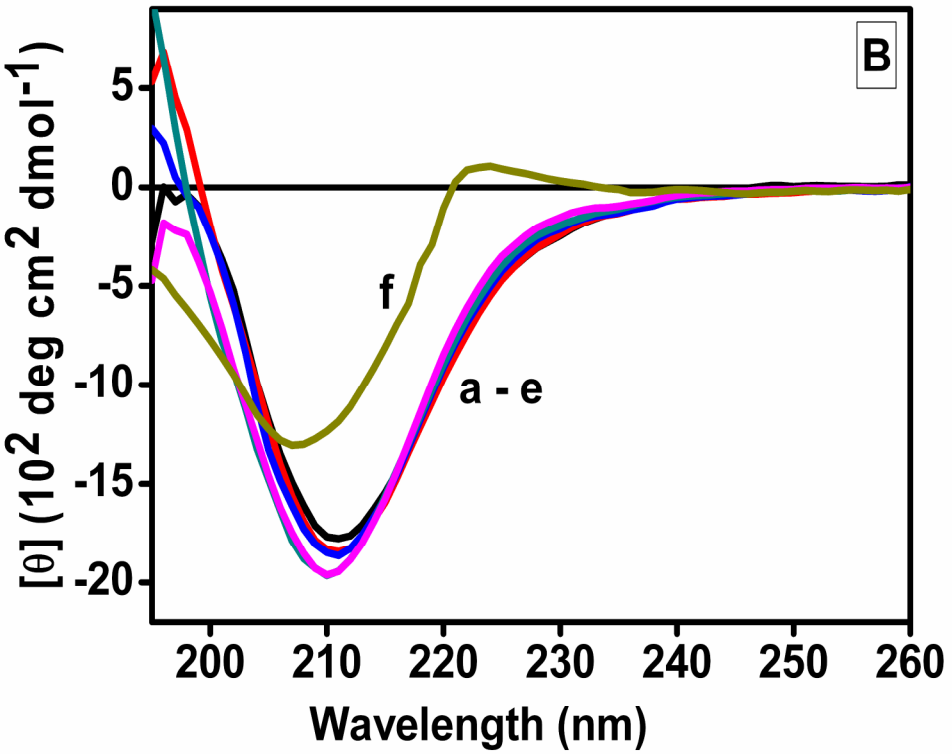


Figure 7





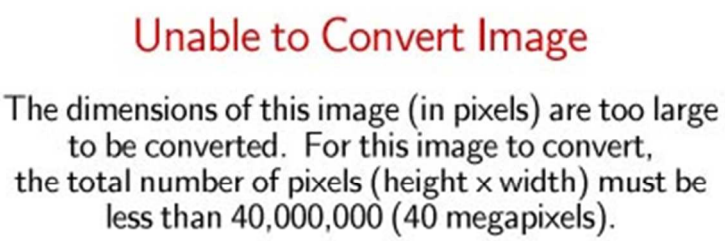
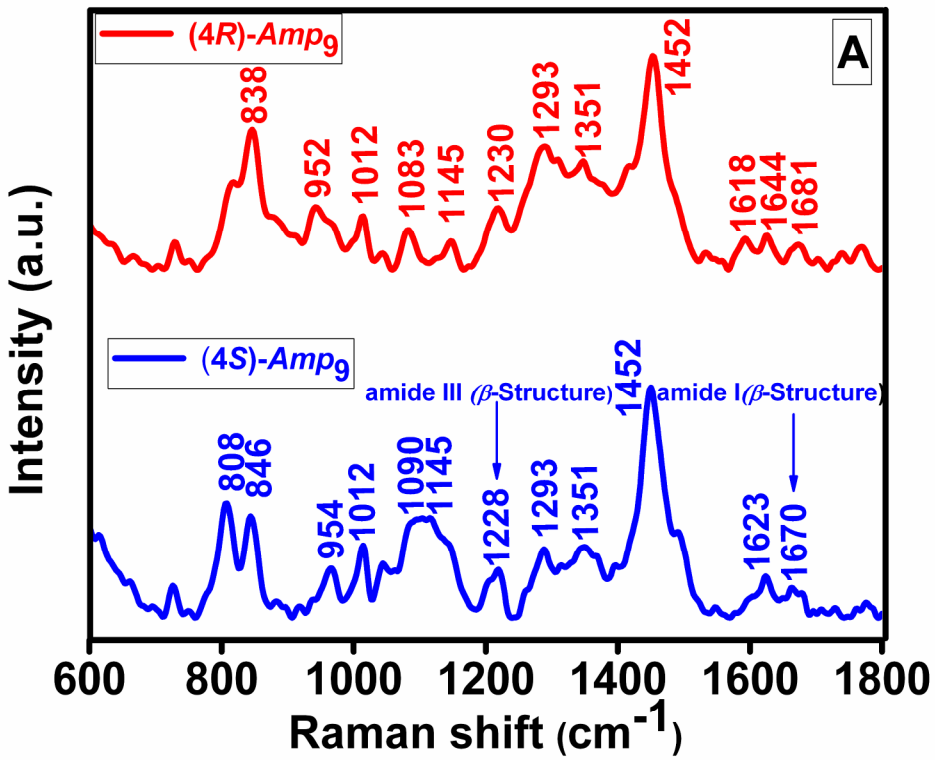
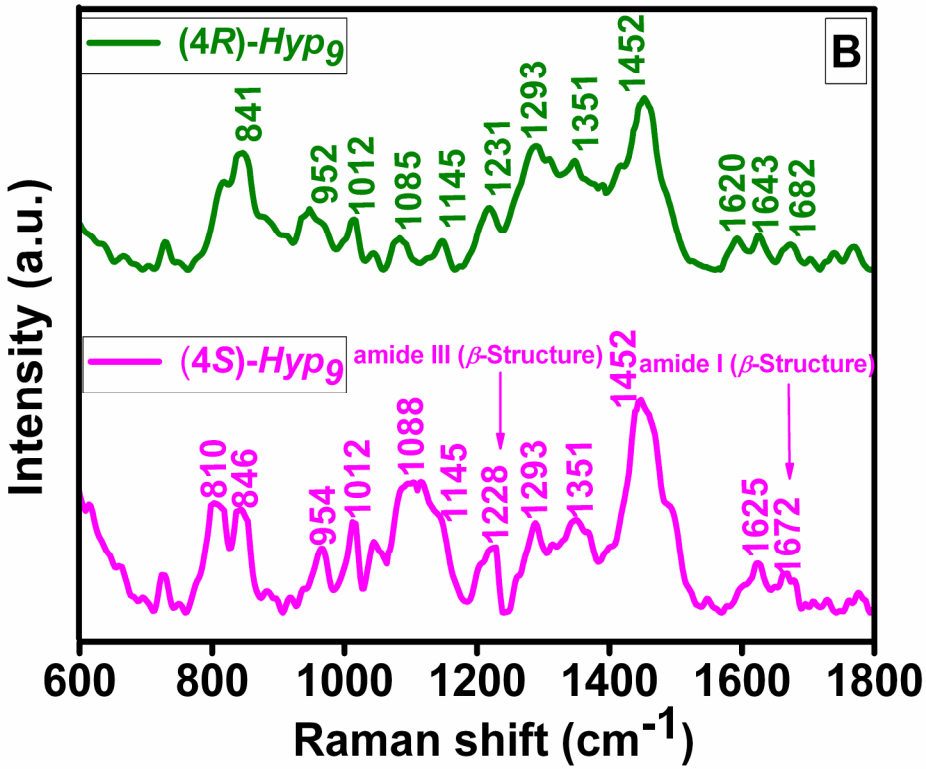


Figure 8





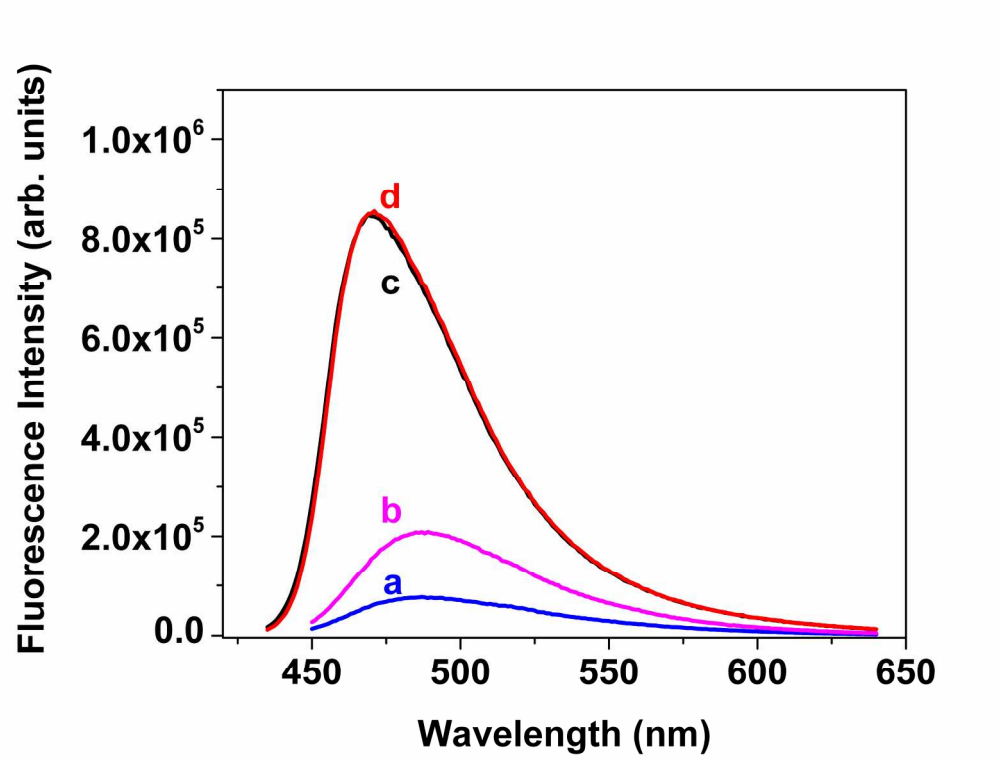


Figure 9

207x158mm (300 x 300 DPI)

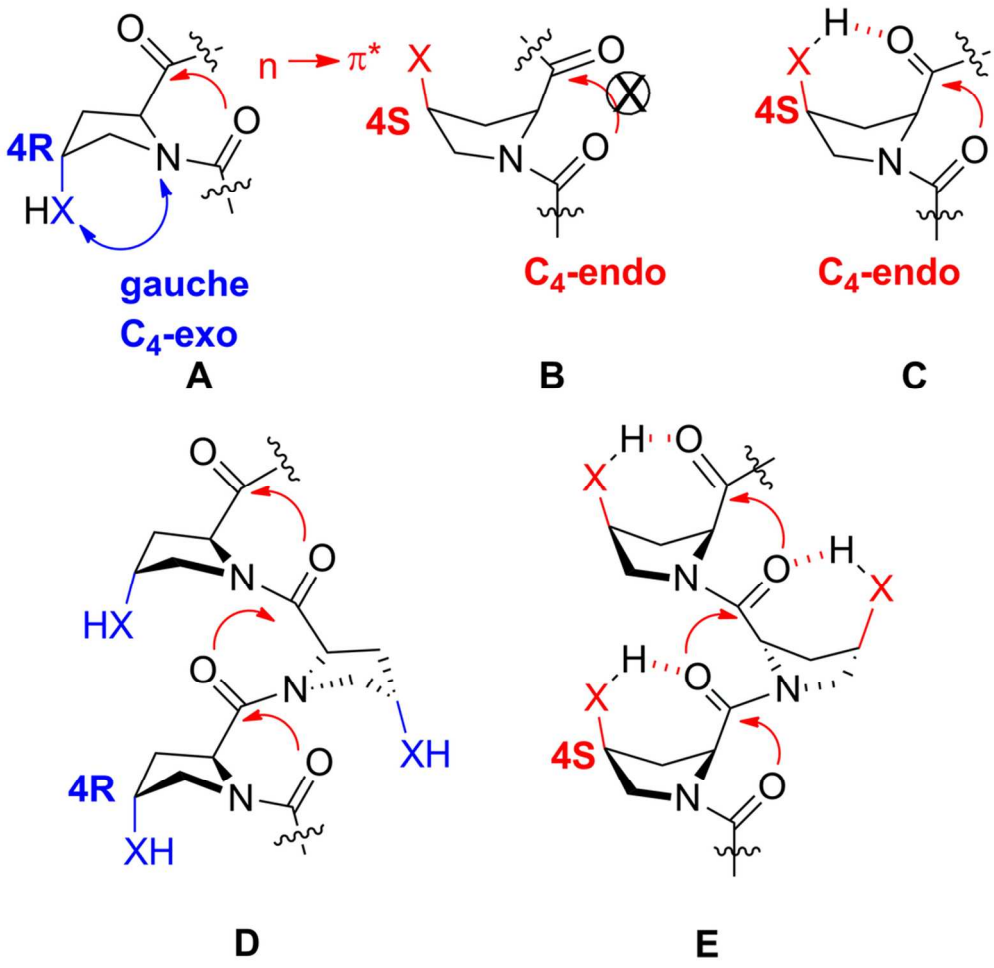


Figure 10

92x90mm (300 x 300 DPI)

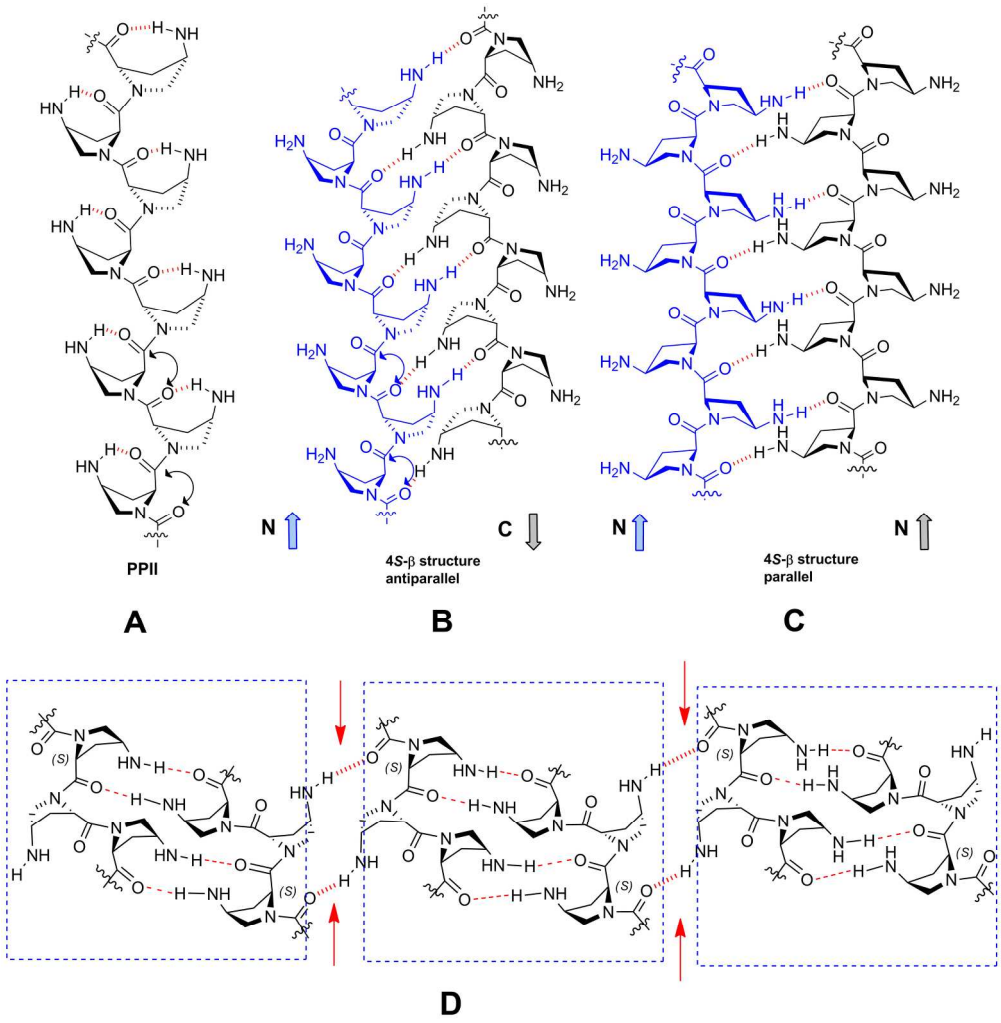


Figure 11

205x209mm (300 x 300 DPI)

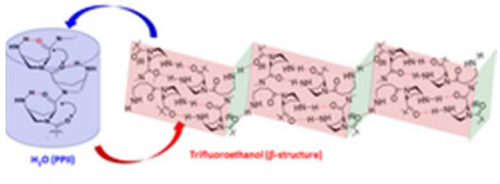


Figure 12 - Table of Contents Graphics

20x8mm (300 x 300 DPI)

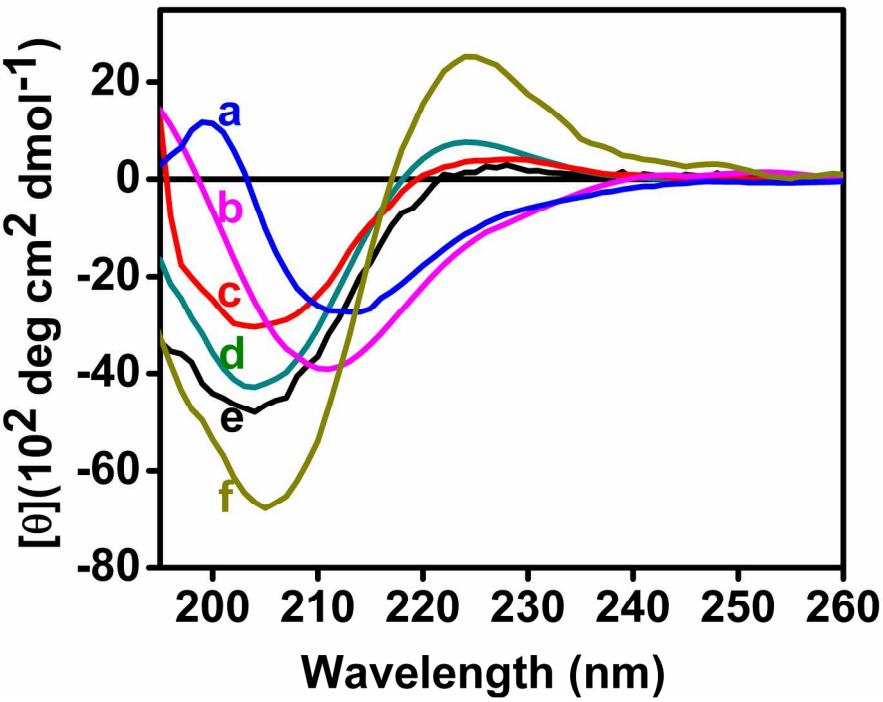


Figure 5

201x156mm (300 x 300 DPI)