Intramolecular C——H · · · O Hydrogen-Bond Mediated Stabilization of a *Cis*-^{D}Pro Imide-Bond in a Stereocontrolled Heterochiral Model Peptide

Mohan M. Bhadbhade,[1] Raghuvansh Kishore[2]

[1] *Solid-State and Elemental Analysis Unit, Mark Wainwright Analytical Centre, University of New South Wales, Sydney, NSW 2052, Australia*

[2] *CSIR-Institute of Microbial Technology, Protein Science and Engineering Divison, Sector 39-A, Chandigarh-160 036, India*

Received 8 June 2011; accepted 21 July 2011
Published online 19 August 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bip.21705

ABSTRACT:

*The X-ray diffraction analysis of a stereocontrolled heterochiral designed model peptide Boc-^{D}Pro-Thr-OMe (**1**) revealed the existence of an unusual folded molecular structure, stabilized via an effective unconventional C—H. . .O type intramolecular hydrogen-bond, encompassing a noncovalent 12-membered ring-motif. Together with an uncommon type* a *disposition of the urethane moiety, the tightly folded topology is compounded with a cis-^{D}Pro imide-bond. The overall conformation is suggested to be the reminiscent of specific type VI* β*-turn structures, hitherto, characterized across the Aaa-cis-Pro peptide-bonds in globular proteins and polypeptides. The ^{13}C NMR spectrum of **1** in an apolar $CDCl_3$ environment revealed the presence of approximately an equal population of cis and trans isomers unexpectedly, analogous to Pro side-chain, the ^{13}C NMR chemical-shifts of Thr C^{β}-resonance is observed to be sensitive toward cis-trans isomerization. In conjunction with solid-state FT-IR spectral data, we established that a network of complex intermolecular hydrogen-bonds stabilize a self-complementary noncovalent helical hexagonal self-assembly and crystallographic supramolecular aggregate. The results incline us to highlight that the stabilization of cis-^{D}Pro peptide-bond in crystalline state may be driven by the favorable energy of formation of an unconventional weak C—H. . .O intramolecular hydrogen-bond. © 2011 Wiley Periodicals, Inc. Biopolymers 97: 73–82, 2012.*

Keywords: unusual peptide conformation; cis-trans conformers; X-ray diffraction analysis; C—H. . .O interaction; FT-IR spectroscopy

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

INTRODUCTION

There have been considerable interests in conceiving *de novo* design strategies that restrict overall conformational freedom of short linear peptides comprised of 20 encoded amino acids.[1–7] Amongst the 20 encoded amino acids, the Pro side-chain and C^{β}-branched Val, Ile, and Thr residues are considered to be conformationally unique.[1–15] In case of Pro side-chain, the five-membered pyrrolidine ring specifically restricts the main chain ϕ torsion angle $\sim -65 \pm 15^{\circ}$ and has a relatively high intrinsic propensity for accommodating the *cis* imide-bond preceding the Pro residue.[1] However, the residues bearing C^{β}-branched chemical entities usually impose rather specific constraints

Correspondence to: R. Kishore; e-mail: kishore@imtech.res.in
Contract grant sponsor: Department of Science & Technology (DST)
Contract grant sponsor: Council of Scientific & Industrial Research (CSIR), Government of India

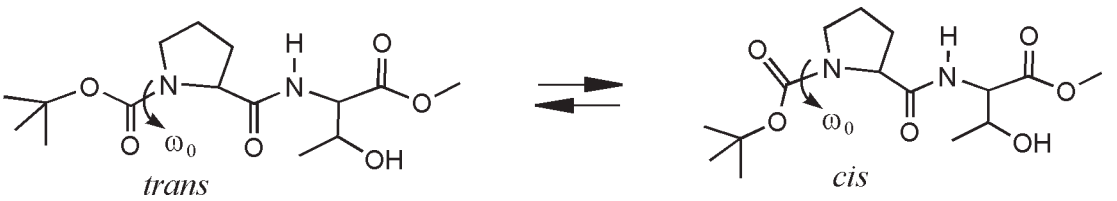

FIGURE 1 Schematic representation of the chemical structure of Boc-^{D}Pro-Thr-OMe signifying the *cis-trans* isomerization across the Boc-^{D}Pro imide-bond involving the ω_0 torsion angle: *trans*-isomer: $\omega_0 \sim 180^\circ$ (left) and *cis*-isomer: $\omega_0 \sim 0^\circ$ (right).

on the backbone conformations.[8–15] The results of protein engineering experiments with C^{β}-branched amino acids indicated that an amphiphilic C^{β}-stereogenic center of the Thr residue exhibits significantly high intrinsic propensity for β-sheet conformations, albeit it showed relatively low P_{β} value as compared with aliphatic C^{β}-branched Val and Ile residues.[1–15] Recently, an account of conformational flexibility scale of encoded amino acids, using fluorescence quenching rate constants in synthetic peptides, revealed that the amphiphilic C^{β}-substituent of the Thr residue somewhat limits the conformational flexibility, but much less severely than the hydrophobic C^{β}-substituent in Val and Ile, and exhibits a much higher tendency to occur in folded and turn structures.[15]

Hydrogen-bondings are ubiquitous and most relevant interactions in many chemical, biochemical, and molecular recognition processes. Together, conventional as well as unconventional hydrogen-bonding interactions[16–26] are known to play a crucial not only in shaping the complex folding-unfolding of compact three-dimensional constructions of proteins and peptides but also their overall stability, including supramolecular constructs and self organization. The existence of unconventional weak C—H. . .O hydrogen-bonds, first invoked in the 1930s, have been the subject of intensive investigation for the last 2–3 decades.[16–26] Although, structural contributions of well-characterized C—H. . .O hydrogen-bonds have been documented quite often in organic, bioorganic, and biomacromolecules[27–33] however, a growing number of reports on C—H. . .O interactions tend to suggest that in the presence of conventional hydrogen-bonds, the potentials of weak C—H. . .O interactions often remain neglected.[20,25] Therefore, we conjecture that solid-state structural analysis of a biologically relevant peptide structure, not influenced and stabilized by conventional main-chain to main-chain or main-chain to side-chain intramolecular hydrogen-bond, may be of considerable general interest. To emphasize, crystal structure analysis of synthetic peptides have unambiguously demonstrated the critical importance of intramolecular C—H. . .O hydrogen-bonds for stabilizing diverse folded structural features, for example, folded β-turn like structural motifs,[27–29] the registry of antiparallel arrangement of α-helices and β-strands[30,31] and peptide chain reversal at the C-terminus of helical segments terminating in 'Schellma-motif.[32,33]

As one of our series of investigations, we have been involved in designing biologically relevant unusual folded-unfolded structural features in Thr containing homochiral or heterochiral model peptides of the type Boc-Aaa-Thr-OMe, where Aaa may be a proteinogenic or nonproteinogenic α-amino acid.[28,32,34–37] We previously reported, using X-ray diffraction analysis and ^{1}H NMR spectroscopy, the characterization of a novel Asx-turns like motif in terminally blocked homochiral peptides Boc-Thr-Thr-OMe and Boc-Thr-Thr-NH_2, stabilized by an unconventional main-chain to side-chain C^{γ}—H. . .O=C intramolecular hydrogen-bond.[28,32] In the past, Benedetti and coworkers[38] attempted to feature the importance of U-shaped molecular structures to produce rather predictable supramolecular self-assembly from enantiomeric and diastereomeric correlated peptides Boc-^{D}Aaa-Aaa-OMe, where Aaa may be a Leu, *a*lle, or Met residue.[38,39] Taken together, these studies tend to point out the propensity(ies) of terminally protected homochiral L,L and heterochiral D,L dipeptides for accommodating diverse biologically relevant folded-unfolded structures[40,41] and probably, their controlled supramolecular self-assembly.[38–41]

In view of prospective implications of C^{β}-branched residues in peptide design, we extend our exploration and describe herein the crystal molecular structure and shape-specific supramolecular self-assembly of a stereocontrolled heterochiral model peptide Boc-^{D}Pro-Thr-OMe (**1**) and compare its ^{13}C NMR spectral characteristics with correlated model peptide Boc-^{D}Thr-Thr-OMe (**2**) and the derivative Boc-Thr-OMe (**3**). The latter peptides serve as control for analysing ^{13}C NMR spectral characteristics associated with the Thr residue.[42–44] The ^{D}Pro residue, besides restricting the ϕ torsion angle in the D-region of the Ramachandran map,[45] compensates for reversed side-chain orientation.[41,45–47] As depicted in Figure 1, the design strategy also involves the possibility of *cis-trans* isomerization ($\omega_0 \approx 0^\circ$ in *cis* and $\omega_0 \approx 180^\circ$ in *trans*) in solution, across the Boc-^{D}Pro imide-

bond. Usually, the *trans* isomer is favored over the *cis* by ~ 0.5 kcal/mol.[45–47] In short peptides, the *cis* and *trans* peptide-bonds can exhibit marked influence on the overall preferred molecular structure and consequently, their shape-specific supramolecular self-assemblies.[33] Unexpectedly, the crystal structure of **1** optimized an unusual folded molecular conformer, accommodating an energetically less favorable *cis*-^{D}Pro imide-bond stabilized via an effective unconventional C—H...O type intramolecular hydrogen-bond. Moreover, analogous to Pro side-chain C^{β}, C^{γ}, and C^{δ} resonances, the observed ^{13}C NMR spectral charatecristics of the Thr C^{β}-resonance demonstrated to be sensitive towards *cis-trans* isomerization.[42–44] To our knowledge, this study represents the first example of a conformational switch, that is, from *cis-trans* isomerization in solution to the *cis* isomer in crystalline-state, where the stabilization of relatively less favorable *cis*-^{D}Pro conformer is driven by the favorable energy of formation of a weak intramolecular C—H...O hydrogen-bond seemingly, in cooperation with an amphiphilic C^{β}-streogenic center.

METHODS

The peptides Boc-^{D}Pro-Thr-OMe (**1**), Boc-^{D}Thr-Thr-OMe (**2**) and the derivative Boc-Thr-OMe (**3**) were synthesized using standard solution-phase methodology[48,49] and purified by silica-gel (60–120 mesh) column chromatography. For the detection of peptides, iodine vapour staining method and/or ninhydrin reaction were employed. Aluminium sheet coated with silica-gel (60 F_{254}: MERCK) was used for determining R_f values. Specific rotations $[\alpha]_D^{25}$ were determined on an Autopol IV Automatic Polarimeter. The melting point of **1** is reported as uncorrected. The identity as well as purity of the covalent structures were ascertained by 1D and 2D ^{1}H NMR. Complete ^{1}H chemical-shift assignments were made using correlation spectroscopy (COSY) spectroscopy on a JEOL 300 MHz spectrometer. The ^{13}C NMR spectra were recoded at 75.43 MHz in $CDCl_3$. The ^{13}C resonance assignment was made from the reported standard chemical-shifts values of the residues.[42–44] All the chemical-shift values for **1–3** are reported downfield relative to tetramethylsilane (TMS) at 0.0 ppm.

Boc-^{D}Pro-Thr-OMe

To a solution of HCl · Thr-OMe (0.85 g; 5.0 mmol) in DMF:DCM (~15 ml; 1:1 v/v), triethylamine (TEA: 0.70 ml; 5.0 mmol), *N*-hydroxybenzotriazole (HOBt: 0.76 g; 5.0 mmol), and Boc-^{D}Pro-OH (1.07 g; 5.0 mmol) were added and cooled the reaction in an ice-salt mixture. After ~5 min dicyclohexylcarbodiimide (DCCI: 1.10 g; ~5.0 mmol) was added and stirred the reaction mixture continuously for 6 hr under cold condition and overnight at room temperature. Dicyclohexylurea (DCU) precipitated was filtered off and after removing DMF, solvent EtOAc (~100 ml) was added and organic layer was washed subsequently with 2*N* HCl (3 × 30 ml), 0.5*M* $NaHCO_3$ (3 × 30 ml), and saturated NaCl solution (~20 ml). The organic layer, dried over anhydrous Na_2SO_4 was evaporated under *vacuuo*. The crude peptide was purified to homogeneity by silica-gel column chromatography using 1–2% MeOH—$CHCl_3$ mixtures as eluents and isolated as a white solid. Yield 0.93 g (~58%), melting point 113°C, R_f 0.48 (5% MeOH—$CHCl_3$ mixture) and $[\alpha]_D^{25}$ = 51.6 (c = 0.5, MeOH). The ^{1}H NMR (conc. = 2.0 mg/0.4 ml) is showed two sets of the Thr amide-NH resonances for *cis-trans* conformers. ^{1}H NMR, δ_{ppm}: 1.25 (3H, *d*, Thr $C^{\gamma}H_3$), 1.48 [9H, *s*, $(CH_3)_3$], 1.92 (2H, *m*, ^{D}Pro $C^{\gamma}H_2$), 2.20 (2H, *m*, ^{D}Pro $C^{\beta}H_2$), 3.40 (1H, broad *s*, Thr $O^{\gamma}H$), 3.52 (2H, *m*, ^{D}Pro $C^{\delta}H_2$), 3.76 (3H, *s*, OMe), 4.36 (2H, *m*, Thr $C^{\beta}H$, ^{D}Pro $C^{\alpha}H$), 4.60 (1H, *dd*, Thr $C^{\alpha}H$), 6.79, 7.15 (1H, broad *d*, Thr NH, *cis-trans* conformers). Noteworthy, the observed ^{13}C NMR spectrum (conc. = 40 mg/0.5 ml) showed two sets of resonances, see Results and Discussion.

Boc-^{D}Thr-Thr-OMe

This peptide was synthesized by following the procedure essentially described above. The crude yellowish gum was purified to homogeneity by silica-gel column chromatography using 2–3% MeOH-$CHCl_3$ mixtures as eluents. The purified peptide was isolated as colourless gum in ~45% yield (0.60 g). R_f 0.44 (5% MeOH-$CHCl_3$ mixture and $[\alpha]^{25}_D$ = 10.6 (c = 0.5, MeOH). ^{1}H NMR (conc. = 2.0 mg/0.5 ml) δ_{ppm}: 1.23, 1.25 (3H, 3H, *d*,*d*, ^{D}Thr, Thr $C^{\gamma}H_3$), 1.47 [9H, *s*, $(CH_3)_3$], 3.77 (3H, *s*, OMe), 4.12 (1H, *d*, ^{D}Thr $C^{\alpha}H$), 4.40 (2H, *m*, ^{D}Thr, Thr $C^{\beta}H$), 4.55 (1H, *dd*, Thr $C^{\alpha}H$), 5.51 (1H, *d*, ^{D}Thr NH, ^3J = 6.9), and 7.22 (1H, broad *d*, Thr NH). The ^{13}C NMR spectrum (conc. = 35 mg/0.5 ml) showed a single set of resonances (data not shown).

Boc-Thr-OMe

To a solution of HCl · Thr-OMe (0.85 g; 5.0 mmol) in dioxane:water (~15 ml; 1:1 v/v) mixture and TEA (2.1 ml; 15.0 mmol), Boc-azide (0.80 ml; 5.0 mmol) was added and the reaction mixture was stirred for ~48 hr at room temperature, while maintaining the pH alkaline. After evaporating dioxane, the solvent EtOAc (~80 ml) was added and the organic layer was washed subsequently with 2*N* HCl (3 × 20 ml), 0.5*M* $NaHCO_3$ (3 × 20 ml) and finally with brine (~15 ml). The organic layer, dried over anhydrous Na_2SO_4 was evaporated under *vacuuo* and the crude compound isolated was purified to homogeneity by silica-gel column chromatography, as described above. The purified **3** was obtained as a colourless gum in ~58% yield (0.93 g). R_f 0.48 (5% MeOH-$CHCl_3$ mixture) and $[\alpha]_D^{25}$ = 51.4 (c = 0.5; MeOH). The ^{1}H NMR (conc. ~2.5.0 mg/0.5 ml) δ_{ppm} 1.25 (3H, *d*, Thr $C^{\gamma}H_3$), 1.46 [9H, *s*, $(CH_3)_3$], 2.85, (1H, broad *s*, Thr $O^{\gamma}H$), 3.79 (3H, *s*, OMe), 4.12 (2H, *m*, Thr $C^{\alpha}H$, $C^{\beta}H$), and 5.42 (1H, *d*, Thr NH, ^3J = 9.0). The ^{13}C NMR spectrum (conc. ~45 mg/0.5 ml) showed a single set of resonances (data not shown). Our attempt to solidify **2** and **3** however, remained futile.

X-Ray Diffraction Study

Colourless single crystals suitable for X-ray diffraction analysis were obtained from a mixture of MeOH:$CHCl_3$ solution (~1:8 v/v) by slow evaporation at room temperature. A good quality single crystal of dimension 0.95 × 0.25 × 0.20 mm^3 was selected for the determination of cell parameters and data collection. The X-ray diffraction intensities were collected on a Brüker SMART APEX CCD diffractometer equipped with MoKα radiation with highly oriented graphite monochromator (λ = 0.71073 Å) at 293 K. The data were

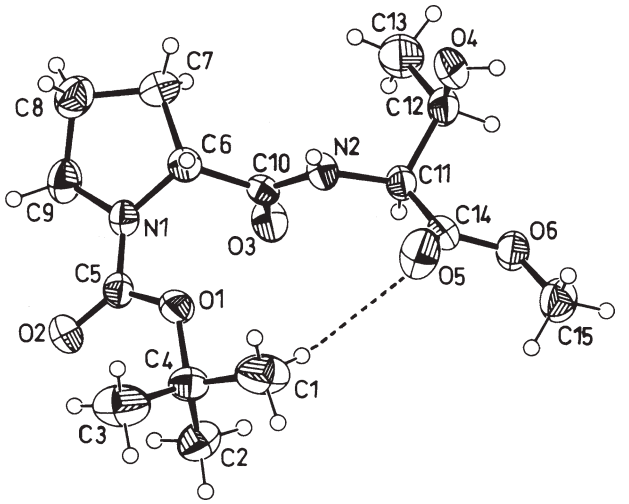

FIGURE 2 An ORTEP presentation of the crystal molecular structure of the model peptide Boc-^{D}Pro-Thr-OMe (**1**). The thermal ellipsoids are shown to the 50% probability level. Dotted line indicates an unconventional C—H. . .O type intramolecular hydrogen-bond encompassing a noncovalent 12-membered ring motif.

corrected for Lorentzian and polarization effects. A semi-empirical absorption corrections based on psi scan was also applied (max t = 0.9971 and min t = 0.9361) using Brüker's SAINT and SADABS programmes. Crystal data **1**: molecular formula: $C_{15}H_{26}N_2O_6$ and molecular weight: M_w = 330.38. The peptide crystallized in hexagonal space group $P6_5$ with unit cell dimensions: a = 10.5780(9), b = 10.5780(9), c = 27.956(5); $\alpha = \beta = 90^\circ$ and $\gamma = 120^\circ$; $F(000)$ = 1068; V = 2709.0(6) Å^3; D_c = 1.215 mg m^{-3} and Z = 6. The structure was solved using direct method program[50] SHELXS-97 and refined with the same program suite SHELXL-97. A total of 3063 unique reflections were measured in a $\omega - 2\theta$ scan mode in the range $2.22^\circ \leq \theta \leq 24.98^\circ$. All the 3063 independent reflections with $I > 2.0\ \sigma(I)$ were considered observed and used in the calculations. The E map computed with the phase set having the best figures of merit revealed all nonhydrogen atoms in the molecule. The nonhydrogen atoms were refined anisotropically, subsequent to the convergence of isotropic refinement. The full matrix least-square refinement of 3063 independent reflections, using SHELX-97 gave the final R factor of 0.049, wR to 0.120 and goodness-of-fit (GOF) on F^2 to 1.11. The absolute structure (Flack) parameter was −0.6(14) and maximum and minimum residual electron densities ($\Delta\rho_{max}$ and $\Delta\rho_{min}$) were 0.41 and −0.13 (e Å^{-3}), respectively. Unit weights were used for all reflections and no constraints of restraints were applied. All hydrogen atoms were placed in idealized positions with assigned isotropic parameters. A total of 237 parameters were used for refinements. The final difference Fourier map was featureless. The crystallographic data of **1** have been deposited at the Cambridge Crystallographic Data Center (CCDC 793870) UK. The crystal data can be obtained free of charges via e-mail: deposit@ccdc.cam.ac.uk or from the Cambridge Crystallographic Data Centre, Cambridge, UK.

Fourier-Transform Infrared

The FT-IR absorption spectrum of the crystalline samples **1** was recorded using KBr pellet (∼1.0 mg per 120 mg KBr) on a Perkin Elmer Spectrum BX FTIR spectrometer equipped with DELL Optiplex G × 1 computer. The IR spectrum was recorded at 2 cm^{-1} resolution and represents an average of 64 scans at an ambient temperature.

RESULTS AND DISCUSSION

A perspective view of the crystal molecular structure of **1**, along with atomic numbering scheme, is shown in Figure 2 and the selected torsion angles are listed in Table I. One of the attractive features of the crystal molecular structure is the characterisation of a planar *cis* disposition of the urethane moiety, that is, $\omega_0 = -1.5^\circ$, categorized as an uncommon type *a*.[51] This unusual orientation of urethane moiety exerted a dramatic influence on the overall preferred molecular conformation and possibly, its supramolecular self-assemblage.

While the ^{D}Pro residue in **1** prefers a semi-extended conformation ($\phi = 67.3$; $\Psi = -148.6^\circ$) in the left-handed polyproline II (PPII) region of the Ramachandran map,[45] the C-terminus Thr moiety adopted a significantly extended structure ($\phi = -122.8^\circ$; $\Psi_T = 173.0^\circ$). Interestingly, in combination with Boc-*cis*-^{D}Pro imide-bond orientation, the peptide main-chain conformation adopts an unusual tightly folded U-shaped topology, stabilized by an unconventional C—H. . .O intramolecular hydrogen-bond, encompassing a non-covalent 12-membered ring-motif. The determined hydrogen-bond geometric parameters: [$d_{(C1\ldots O5)}$ = 3.47 Å; $d_{(C1-H\ldots O5)}$ = 2.70 Å and hydrogen-bond angle $\angle_{C1-H\ldots O5}$ = 137.6°] of the C1—H. . .O5 interaction are well accommodated within the accepted definition of an effective C—H. . .O hydrogen-bond.[16–20] In view of the fact that "in some C—H. . .O contacts formed by particular unactivated C—H groups, say a methyl group at the end of an alkyl

Table I Selected Torsion Angles (°) for Boc-^{D}Pro-Thr-OMe

Atoms in the sequence	Designation	Values
C4—O1—C5—N1	θ_1	−175.0
O1—C5—N1—C6	ω_0	−1.5
C5—N1—C6—C10	φ_1	67.3
N1—C6—C10—N2	ψ_1	−148.6
C9—N1—C6—C7	θ	−0.8
N1—C6—C7—C8	χ_1	22.3
C6—C7—C8—C9	χ_2	−34.8
C7—C8—C9—N1	χ_3	33.8
C8—C9—N1—C6	χ_4	−20.6
C6—C10—N2—C11	ω_1	−172.8
C10—N2—C11—C14	φ_2	−122.8
N2—C11—C14—O6	χ_T	173.0
N2—C11—C12—O4	χ_1	66.4
N2—C11—C12—C13	χ_2	−55.6
C11—C14—O6—C15	ω_T	−177.6

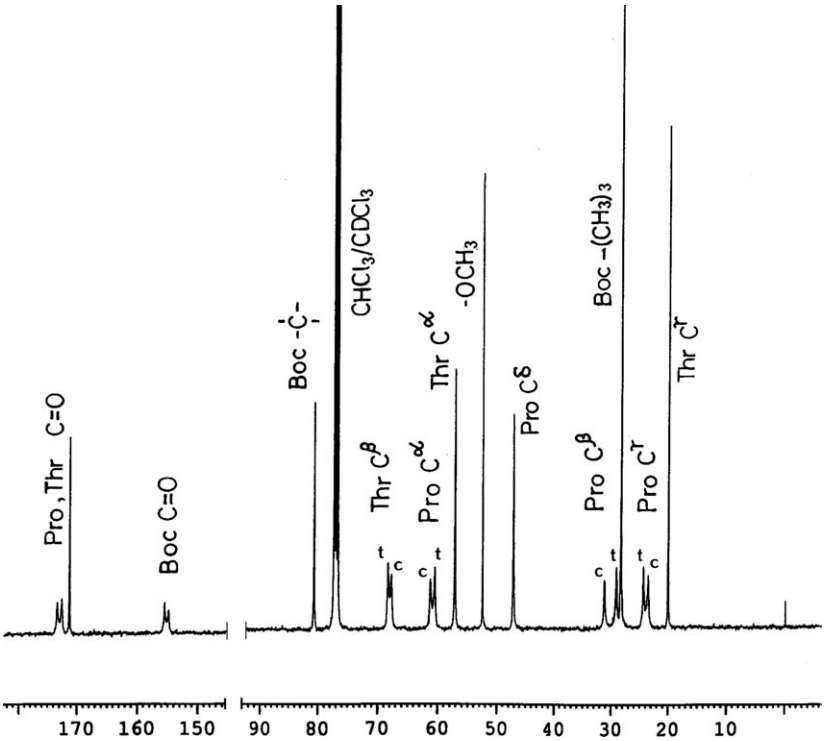

FIGURE 3 ^{13}C NMR spectrum of the peptide Boc-^{D}Pro-Thr-OMe in $CDCl_3$. The *cis* and *trans* conformers are indicated as *c* and *t*, respectively.

chain,"[25] the participation of an unactivated C—H donor of one of the Boc methyl groups in the fabricating an effective C—H. . .O hydrogen-bond in **1** is rather interesting.[32,33] The observed topological features emerged to be the reminiscent of well-characterized type VIb-3 and VId β-turn structures in globular proteins and polypeptides, associated with a *cis*-^{L}-Pro imide-bond.[52–54] Ever since Taylor and Kennard[16] reported the existence and characterization of different unconventional hydrogen-bonds, the weak C—H. . .O interaction has gained wide acceptance as an energetically stable genuine hydrogen-bond. The existing literature provide compelling evidence that ascertains structural as well as functional importance of C—H. . .O interaction in a wide range of organic, bioorganic, and biomacromolecules.[16–33] The ab initio quantum mechanical calculations performed on biologically relevant model systems suggest that energetic contributions of such cohesive interactions may range between ~2.5 and 3.0 kcal/mol in *vacuuo*, that is, approximately one-half the energy of a conventional hydrogen-bond.[20–23] To our knowledge, we are not aware of any example where the stabilization of an isolated *cis* imide-bond, solely by an unconventional C—H. . .O intramolecular hydrogen-bond, has been reported.

The observed ^{13}C NMR spectrum of **1** in an apolar $CDCl_3$ solution showed the presence of approximately an equal population of *cis* (~45%) and *trans* (~55%) isomers.[42–44] It is evident from Figure 3, the incidence of two urethane-carbonyl resonances at ~154.8 and 155.6 ppm suggests that this may be the site of *cis-trans* isomerization. The appearance of Pro side-chain C$^{\beta}$ and C$^{\gamma}$ resonances at ~32.6 and 22.6 ppm, respectively, corresponds to the *cis* conformer.[42] In addition, the peptide also showed distinct chemical-shift differences for the main-chain Pro C$^{\alpha}$ (at ~60.5 and 61.2 ppm) and Pro C=O (at ~172.7 and 173.5 ppm) C-atoms. Unexpectedly, the spectrum also displayed two separate signals for the Thr C$^{\beta}$-atom at ~67.7 and 68.3 ppm. To ascertain that these Thr C$^{\beta}$ resonances are indeed sensitive to *cis-trans* isomerization, the ^{13}C NMR spectra of a correlated model peptide **2** and the derivative **3** were observed under similar experimental condition. The chemical-shift corresponding to Thr C$^{\beta}$ resonance, in fact, appeared as a single conformer at ~67.5 and 68.2 ppm for **2** and **3**, respectively. Here, it may be stated that the urethane moiety in **2** and **3** is not expected to experience the *cis-trans* isomerization. Taken together, the ^{13}C NMR spectral analysis of **1** facilitated the assignment of Thr C$^{\beta}$ signals to *cis-trans* isomerization and the relative downfield chemical-shift could be assigned to the *trans* conformer.

The significantly extended main-chain torsion angles of the Thr residue in **1** indicate the possibility for an intraresidue main-chain to main-chain N*i*—H. . .O=C*i* type C_5-

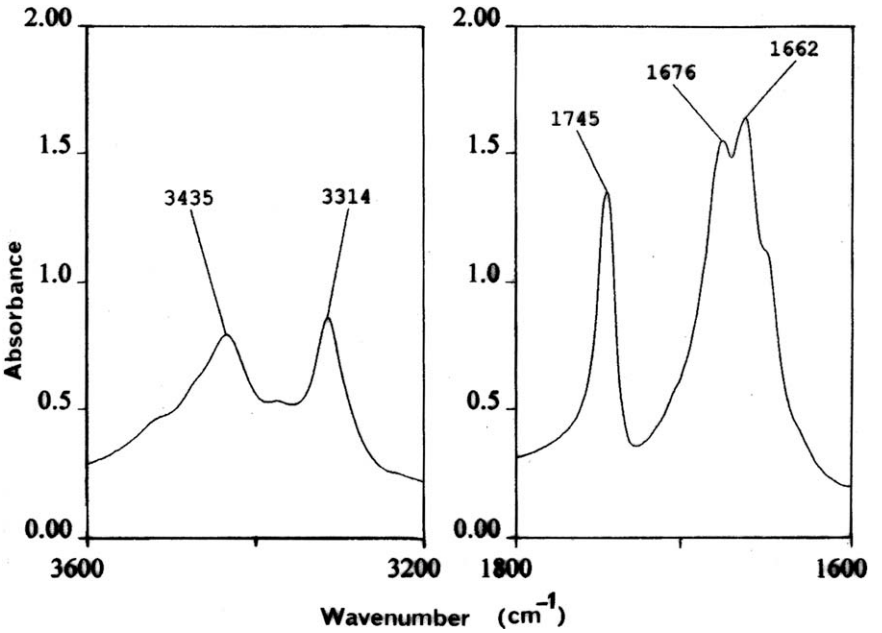

FIGURE 4 Partial FT-IR spectrum of Boc-^{D}Pro-Thr-OMe in KBr disc. The intense absoption bands in the informative amide-A (left) and amide-I (right) regions are indicated.

interaction, that is, C_5-secondary structure, nevertheless, an offensive hydrogen-bond angle, that is, $\angle_{N-H\ldots O} = 89.1^{\circ}$, inclined us to preclude its contribution towards structural stabilization.[34–36] Moreover, energetically preferred *gauche*$^{+}$ orientation of the Thr O$^{\gamma}$H group, that is, N$^{\alpha}$—C$^{\alpha}$—C$^{\beta}$—O$^{\gamma}$ (χ_1) = 66.0°, in view of topological consideration of the semi-extended ϕ_{Thr} torsion angle ~−123°, also precluded the fabrication of an intraresidue main-chain to side-chain N_i—H…O_i^{γ} type interaction, that is, *pseudo* C_5-structure.[1,14,36] Regarding Pro side-chain conformations, of the two energetically equivalent puckered orientations of the five-membered pyrrolidine ring, that is, A or "UP" or C$^{\gamma}$-endo and B or "DOWN" or C$^{\gamma}$-exo orientations, the ^{D}Pro moiety in **1** preferred a C$^{\gamma}$-exo orientation, characterized by positive χ_1, χ_3, and negative χ_2, χ_4 torsion angles (Table I).[45,52–54] Noteworthy, an analysis of 236 *cis*-^{L}Pro counts in high-resolution protein crystal structures also revealed that the DOWN pucker is predominantly favored over the UP pucker, that is, 188 C$^{\gamma}$-exo orientations against 29 C$^{\gamma}$-endo puckers, respectively.[55]

To further sustain that the crystal structure of **1** is primarily stabilized by hydrogen-bonding interactions, solid-state FT-IR absorption spectrum was analyzed.[56–60] As shown in Figure 4, the observed FT-IR spectrum of the crystalline sample **1** in KBr disc displays intense absorption bands at ~3435 and 3314 cm^{-1} in amide-A region and at ~1745, 1676, and 1662 cm^{-1} in amide-I region.[33,58,59] The fact that molecular conformation of **1** is devoid of a conventional intramolecular hydrogen-bond, the appearance of two intense peaks in amide-A region may indicate the involvement of Thr O$^{\gamma}$H and amide-NH donor groups in strong intermolecular interactions.[56–60] The presence of a deep carbonyl absorption at ~1745 cm^{-1} in amide-I region may be ascribed to the methyl ester participating in hydrogen-bonding interactions.[58,59] The remaining two IR bands at ~1676 and 1662 cm^{-1} are expected to arise from the urethane and Pro moieties, which may be attributed to hydrogen-bonded C=O acceptor group, fully consistent with the one established from the crystal structure analysis.

Another major object of present analysis is the illumination of supramolecular self-assembly engendered from an unusual folded building-block **1**. A view of the molecular packing, along the crystallographic *c* axis, is shown in Figure 5. The description of crystal packing is complete with the involvement of three roughly parallel symmetry related molecules. The potential hydrogen-bond donors, that is, Thr O$^{\gamma}$H and amide-NH groups, comprised two conventional intermolecular hydrogen-bonds, that is, N2—H…O3 and O4—H…O2 interactions, fully consistent with FT-IR data. The self-complementary building-block also optimized the creation of two unconventional intermolecular hydrogen-bonds, that is, C1—H…O6 and C6—H…O3 interactions, which may further strengthen the overall supramolecular self-organization. The determined hydrogen-bond geometric parameters of these interactions are summarized in Table II. The molecular self-assembly of **1** produced an attractive helical

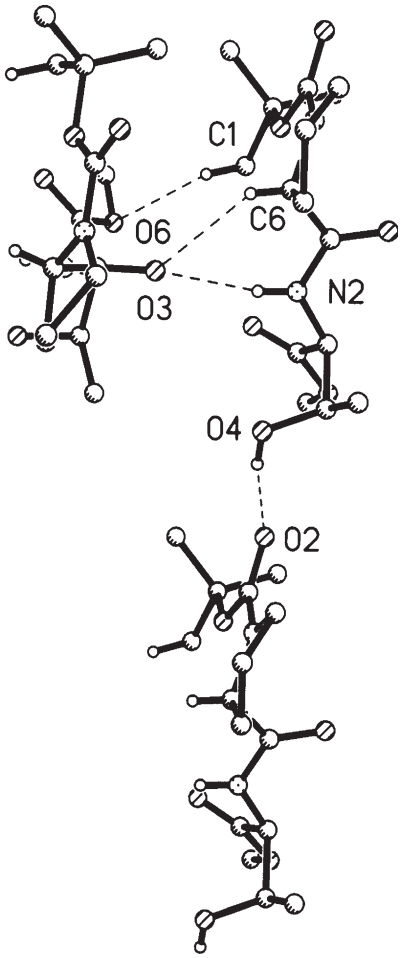

FIGURE 5 An illustration of intermolecular hydrogen-bonds for Boc-^{D}Pro-Thr-OMe as viewed along the *bc* plane. For the sake of clarity, only those hydrogen-atoms involved in intermolecular interactions are shown. Dotted lines indicate intermolecular hydrogen-bonds.

hexagonal supramolecular architecture, as viewed along the *ab* plane. A complex network of intermolecular hydrogen-bonds, along with Van der Waals interactions, seems to stabilize the complex three-dimensional supramolecular self-organization, as depicted in Figure 6. Since, the less favorable *cis*-^{D}Pro imide-bond in solid-state is favored to the *trans*-^{D}Pro isomer, we therefore, reiterate that "crystal structures are not necessarily the global free energy minima of solid supramolecular self-assemblies."[61] In sum, the results incline us to preclude that the observed tightly folded molecular structure and the consequential hexagonal supramolecular self-organization of **1**, is merely a consequence of intermolecular interactions.

The nonexistence of conventional intramolecular hydrogen-bond in **1** evidently suggests that the structural role played by the unconventional C—H. . .O hydrogen-bond could be of paramount importance. It should be noted that the observed weak interaction predominantly restrict the main-chain torsion angles suitable for accommodating the *cis* imide-bond orientation. The compelling evidence for this inference is offered from the reported crystal molecular structures of three strongly correlated model peptides, listed in Table III. The first imperative distinguishing feature of these peptides, in marked contrast to **1**, is the absence of an uncommon type *a* disposition of the urethane moiety. Besides, while main-chain torsion angles of the Thr residue in **1** are significantly extended ($\phi \sim -123$; $\Psi_T \sim 173$), the corresponding angles of the C-terminus residue in correlated peptides adopted either semi-folded ($\phi = -145 \pm 5$; $\Psi_T = -45 \pm 15^{\circ}$ in Boc-^{D}Leu-Leu-OMe and Boc-^{D}*a*Ile-*a*Ile-OMe) or folded ($\phi = 65$; $\Psi_T = -45^{\circ}$ in Boc-^{D}Met-Met-OMe) conformation.[38,39] However, despite gross conformational differences at the molecular level, the former correlated peptides show remarkably similar molecular self-assembly and supramolecular architecture.[38] The results clearly suggest that the overall stereochemical restrictions imposed via side-chains in peptide **1** clearly enforced the fabrication of a weak C—H. . .O intramolecular hydrogen-bond. And also, in the realm of crystal design and engineering, we highlight the importance of a heterochiral dipeptide building-block with a *cis* imide-bond, which could be constructive to understand the principles of analogous or inimitable self-assembly and supramolecular aggregates.

In the absence of competing interactions, high proportion of favorable *trans*-Pro conformers is usually expected in unconstrained Pro containing peptide sequences. A survey of the distribution of Ψtorsion angles of the ^{D}Pro residues, extracted from Cambridge Structure Database, revealed that of the 111 ^{D}Pro entries, 71 counts (∼63%) favored the PPII

Table II The Determined Hydrogen-Bond Geometric Parameters for Intramolecular and Intermolecular Interactions for the Model Peptide Boc-^{D}Pro-Thr-OMe

	Distance (Å)		
D–H · · · A	D · · · A	H · · · A	Angle D–H · · · Å
Intramolecular			
C1—H · · · O5	3.47	2.70	137.6
Intermolecular			
N2—H · · · O3[a]	3.03	2.16	169.9
O4—H · · · O2[b]	2.81	2.00	165.9
C6—H · · · O3[a]	3.34	2.59	145.4
C1—H · · · O6[a]	3.33	2.65	128.5

[a] Symmetry equivalent positions: $1 + x - y, x, -0.16666 + z$.
[b] Symmetry equivalent positions: $x, -1 + y, z$.

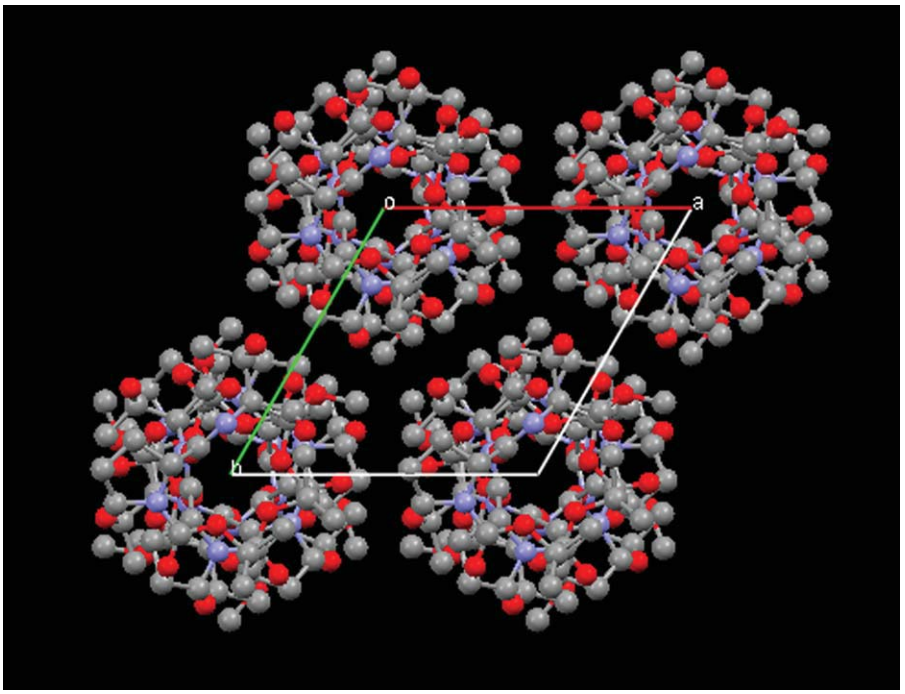

FIGURE 6 Solid-state hexagonal self-assembly of the U-shaped topology of Boc-^{D}Pro-Thr-OMe as viewed along the crystallographic *c* axis. For the sake of clarity hydrogen-atoms and hydrogen bonds are omitted.

conformation, that is, $\Psi = -145 \pm 25^{\circ}$ (N. Shamala, personal communication). And, of these 71 counts, only 8 ^{D}Pro residues (~11%) were compounded with *cis*-^{D}Pro imide-bond. None of this geometries was found to be accommodated in the helical region of the Ramachandran map.[45] However, the analyses of globular protein X-ray structures revealed the occurrence of only ~5 − 6% of the Pro residues in *cis* orientation and are suggested to be of structural and/or functional importance.[52–55] Noteworthy, despite its low abundance, the favorable *cis*-Pro orientations are predisposed to facilitate the polypeptide chain reversal, while displaying marked preference for specific type VI β-turn structures.[52–54] A comprehensive analysis of *cis*-Pro imide-bond mediated type VI β-turns in well-characterized globular proteins of known three-dimensional structures revealed the characterization of two separate categories of tightly folded reverse turn structures categorized as type VIb-3 and VId β-turns, each stabilized by an unconventional C^{α}_{i}—H. . .O_{i+1} hydrogen-bond.[52–54] Nevertheless, in proteins this noncovalent interaction encompasses a 7-membered ring-motif. Therefore, the overall folded topology of **1** is suggested to be the reminiscent of type VIb-3 and VId β-turn structures where the favorable weak C—H. . .O hydrogen-bond stabilizes the *cis*-^{D}Pro peptide-unit. Although, we recognized the C—H. . .O hydrogen-bond mediated stabilization of a *cis*-^{D}Pro imide-bond in an isolated model peptide nevertheless, analogous secondary structural features in well-structured synthetic peptides and can cause polypeptide chain reversal and presumed to have functional significance.[52–55] Many different factors are known to contribute to the experimentally observed *cis* conformer however, the analysis incline us to offer that varying intrinsic propensities of C^{β}-branched residue on either side of the Aaa-^{D}Pro or Aaa-^{L}Pro imide-bond may play a beneficial role in orientating the activated or unactivated C—H donor to facilitate C—H. . .O hydrogen-bond formation which in turn may dictate the relative high-

Table III A Comparison of Main-Chain Torsion Angles (°) in Correlated Model Peptides of the Type Boc-^{D}Xaa-Yaa-OMe

	^{D}Xaa		Xaa		
Peptides	ϕ	ψ	ϕ	ψ_T[a]	Ref.
Boc-^{D}Pro-Thr-OMe	67.3	−148.6	−122.8	173.0	This paper
Boc-^{D}Leu-Leu-OMe	88.0	−134.0	−138.0	62.0	38
Boc-^{D}aIle-Ile-OMe	105.0	−126.0	−151.0	−28.0	38
Boc-^{D}Met-Met-OMe	71.8	−73.8	−65.0	38.3	39

Boc- and OMe represent *t*-butyloxycarbonyl and methyl ester groups, respectively.

[a] The torsion angle ψ_T is defined as N—C^{α}—C′—O.

population of *cis* isomeric form in solution and/or in solid state(s). Taking this into account, we believe that more comprehensive structural exploration of strongly correlated homochiral as well as heterochiral sequence motifs may be imperative to appraise the "β-carbon hypothesis or β-carbon rule" advocated by Hodel et al., through protein engineering experiments.[62]

This analysis is clearly indicative of steric requirement, imposed via an amphiphilic C$^{\beta}$-branched Thr residue, for the stabilization of a *cis*-^{D}Pro conformer in crystalline-state. The enhanced stability of the *cis*-isomer intended to be associated with a smaller entropy loss in the *trans* → *cis* conversion in the Aaa-Pro peptide-bond as compared to Aaa-non-Pro amide-bond. A number of unique conformational as well as functional properties have been shown to be associated with *cis*-Pro or *cis*-non-Pro peptide-bonds not only in folded-unfolded globular and transmembrane proteins but also in bioactive peptides with potential therapeutic implications.[63–70] In general, the *cis* peptide-bonds, despite having differences in their locations in protein three-dimensional structures, are shown to be critically important in controlling protein folding-unfolding, biorecognition events, enzyme catalysis, signal transduction, mechanism that opens the receptor pore, post-translational modifications, etc.[63–70] For these reasons, we infer that the manifestation of C—H. . .O hydrogen-bond mediated stabilization of *cis*-^{D}Pro imide-bond in a short synthetic peptide might be crucial as it exhibited marked influence on the preferred molecular structure and supramolecular self-assembly. In an attempt to derive a plausible mechanism of the structural conversion, that is, from *cis-trans* isomerization in solution to *cis*-conformer in solid-state, we incline to propose that the direction as well as strength of the C—H. . .O hydrogen-bond in the *cis*-conformer may dictate the nucleation process during crystallization, which concurrently triggers the stabilization of specific folded molecular structure and supramolecular self-assembly process.

CONCLUSIONS

We have demonstrated, from X-ray diffraction analysis, the existence and characterization of an unusual U-shaped molecular structure in a stereocontrolled heterochiral model peptide Boc-^{D}Pro-Thr-OMe. To the best of our knowledge, the investigation provided the first experimental observation where an isolated *cis*-Pro imide-bond is solely stabilized by an intramolecular C—H. . .O hydrogen-bond and the enforced topology perceived to be the reminiscent of specific type VI β-turn secondary structures, recognized in globular proteins and polypeptides.[52–54] This study unambiguously established that in the absence of conventional intramolecular interaction, a weak C—H. . .O hydrogen-bond could be of paramount importance for stabilizing unusual folded biological structures, which could be of relatively higher energy. Of particular interest, analogous to Pro side-chain, the ^{13}C NMR spectral characteristics of the Thr C$^{\beta}$ resonance is found to be sensitive towards *cis-trans* isomerization and the correlated model systems unambiguously facilitated the resonance assignments to either of the two forms. Such heterochiral sequence motifs may simultaneously permit an extensive verification and validation of "β-carbon hypothesis" nevertheless, by designing new stereocontrolled synthetic peptides while incorporating appropriate nonproteinogenic constrained amino-acids. Finally, the observed hexagonal molecular self-assembly incline us to highlight that a little rationale seem to govern the overall supramolecular self-assembly which could be advantageous for introducing shape-specific subtle variations in physico-chemical properties.

The authors thank C.P. Midha for providing NMR spectra and V. Kumar and S.K. Nath for technical assistance. This is IMTech Communication No. 63/2010.

REFERENCES

1. Richardson, J. S.; Richardson, D. V. In Prediction of Protein Structure and the Principles of Protein Conformation; Plenum Press: New York, 1990; pp 1–98.
2. Kim, C. A.; Berg, J. M. Nature 1993, 362, 267–270.
3. Minor, D. L., Jr.; Kim, P. S. Nature 1994, 367, 660–663.
4. Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado. W. F. Science 1995, 270, 935–941.
5. Dahiyat, B. I.; Mayo, S. L. Science 1997, 278, 82–87.
6. DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. Ann Rev Biochem 1999, 68, 779–819.
7. Serrano, L. Adv Protein Chem 2000, 53, 49–85.
8. Siemion, I. Z.; Sobczyk, K.; Lisowski, M. Int J Pept Protein Res 1986, 27, 127–137.
9. Altmann, K.-H.; Flörsheimer, A.; Mutter, M. Int J Pept Protein Res 1986, 27, 314–319.
10. Griffiths-Jones, S. R.; Sharman, G. J.; Maynard, A. J.; Searle, M. S. J Mol Biol 1998, 284, 1597–1609.
11. Koca, J. Prog Biophys Mol Biol 1998, 70, 137–173.
12. Monne, M.; Hermansson, M.; von Heijne, G. J Mol Biol 1999, 288, 141–145.
13. Liou, Y. C.; Tocilj, A.; Davies, P. L.; Jia, Z. Nature 2000, 406, 322–324.
14. Pal, D.; Chakrabarti, P. Acta Crystallogr Sect D 2000, 56, 589–594.
15. Huang, F.; Nau, W. M. Angew Chem Int Ed Engl 2003, 42, 2269–2272.
16. Taylor, R.; Kennard, O. J Am Chem Soc 1982, 104, 5063–5070.
17. Baker, E. N.; Hubbard, R. E. Prog Biophys Mol Biol 1984, 44, 97–179.
18. Burley, S. K.; Petsko, G. A. Adv Protein Chem 1988, 39, 125–189.

19. Jeffery, G. A. An Introduction to Hydrogen Bonding; Oxford University Press: New York, 1997.
20. Desiraju, G. R.; Steiner, T. The Weak Hydrogen Bond in Structural Chemistry and Biology; Oxford University Press: Oxford, 1999.
21. Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. J Am Chem Soc 2000, 122, 4750–4755.
22. Scheiner, S.; Kar, T.; Gu, Y. J Biol Chem 2001, 276, 9832–9837.
23. Senes, A.; Ubarretxena-Belandia, I.; Engelman, D. M. Proc Natl Acad Sci U S A 2001, 98, 9056–9061.
24. Loll, B.; Raszewski, G.; Saenger, W.; Biesiadka, J. J Mol Biol 2003, 328, 737–747.
25. Desiraju, G. R. Chem Commun 2005, 2995–3001.
26. Desiraju, R. Cryst Growth Design 2011, 896–898.
27. Thakur, A. K.; Kishore, R. Biopolymers 2000, 53, 447–454.
28. Thakur, A. K.; Kishore, R. Tetrahedron Lett 2001, 42, 4691–4694.
29. Cheung, E. Y.; McCabe, E. E.; Harris, K. D.; Johnston, R. L.; Tedesco, E.; Raja, K. M.; Balaram, P. Angew Chem Int Ed Engl 2002, 41, 494–496.
30. Babu, M. M.; Kumar Singh, S.; Balaram, P. J Mol Biol 2002, 322, 871–880.
31. Singh, S. K.; Babu, M. M.; Balaram, P. Proteins 2003, 51, 167–171.
32. Thakur, A. K.; Kishore, R. Biopolymers 2006, 81, 440–449.
33. Kumar, N.; Venugopalan, P.; Kishore, R. Biopolymers 2010, 93, 927–931 and references therein.
34. Ashish; Kishore, R. Tetrahedron Lett 1997, 38, 2767–2770.
35. Ashish; Kishore, R. FEBS Lett 1997, 417, 97–100.
36. Thakur, A. K.; Ashish; Kishore, R. Chem Commun 1999, 1643–1644.
37. Banumathi, S.; Velmurugan, D.; Subramanian, E.; Ashish; Kishore, R. Acta Crystallogr Sect C 1999, 55, 78–79.
38. Di Blasio, B.; Saviano, M.; Del Duca, V.; De Simone, G.; Rossi, F.; Pedone, C.; Benedetti, E.; Lorenzi, G. P. Biopolymers 1995, 36, 401–408.
39. Immirzi, A.; Avena, P.; Ciajolo, M. R.; Becker, J. M.; Naider, F. Acta Cryst 1978, B 34, 179–183.
40. Di Blasio, B.; Benedetti, E.; Pavone, V.; Pedone, C.; Spiniello, O.; Lorenzi, G. P. Biopolymers 1989, 28, 193–201.
41. Durani, S. Acc Chem Res 2008, 41, 1301–1308 and references therein.
42. Wüthrich, K. NMR in Biological Research: Peptides and Proteins, of Proteins; North-Holland Publishers Company: Amsterdam, Oxford, 1976.
43. Wüthrich, K. NMR of Proteins and Nucleic Acids; Wiley: New York, 1986.
44. Evans, J. N. S. Biomolecular NMR Spectroscopy; Oxford University Press: Oxford, 1995.
45. Ramachandran, G. N.; Sasisekharan, V. Adv Protein Chem 1968, 23, 283–438.
46. Rose, G. D.; Gierasch, L. M.; Smith, J. A. Adv Protein Chem 1985, 37, 1–109.
47. Venkatraman, J.; Shankaramma, S. C.; Balaram, P. Chem Rev 2001, 101, 3131–3152 and references therein.
48. Bodansky, M.; Bodansky, M. The Practice of Peptide Synthesis,2nd ed.; Springer Verlag: New York, 1994.
49. Okada, Y. Curr Org Chem 2001, 5, 1–43.
50. Sheldrick, G. M. SHELXS-97 and SHELXL-97, Programs for the Solution and Refinement of Crystal Structures; University of Göttingen: Göttingen, Germany, 1997.
51. Benedetti, E.; Pedone, C.; Toniolo, C.; Nemethy, G.; Pottle, M. S.; Scheraga, H. A. Int J Pept Protein Res 1980, 16, 156–172.
52. Pal, D.; Chakrabarti, P. J Mol Biol 1999, 294, 271–288.
53. Chakrabarti, P.; Pal, D. Prog Biophys Mol Biol 2001, 76, 1–102 and references therein.
54. Bhattacharyya, R.; Chakrabarti, P. J Mol Biol 2003, 331, 925–940.
55. Ho, B. K.; Coutsias, E. A.; Seok, C.; Dill, K. A. Protein Sci 2005, 14, 1011–1018.
56. Krimm, S.; Bandekar, J. Adv Protein Chem 1986, 38, 181–364.
57. Jackson, M.; Mantsch, H. H. Crit Rev Biochem Mol Biol 1995, 30, 95–120.
58. Shaw, R. A.; Perczel, A.; Fasman, G. D.; Mantsch, H. H. Int J Pept Protein Res 1996, 48, 71–78.
59. De Wael, K.; Zeegers-Huyskens, T. Biopolymers 1997, 41, 205–212.
60. Vass, E.; Hollosi, M.; Besson, F.; Buchet, R. Chem Rev 2003, 103, 1917–1954.
61. Nangia, A. Acc Chem Res 2008, 41, 595–604.
62. Hodel, A.; Kautz, R. A.; Fox, R. O. Protein Sci 1995, 4, 484–495.
63. Tugarinov, V.; Zvi, A.; Levy, R.; Anglister, J. Nat Struct Biol 1999, 6, 331–335.
64. Hoffmann, T.; Lanig, H.; Waibel, R.; Gmeiner, P. Angew Chem Int Ed Engl 2001, 40, 3361–3364.
65. Lummis, S. C.; Beene, D. L.; Lee, L. W.; Lester, H. A.; Broadhurst, R. W.; Dougherty, D. A. Nature 2005, 438, 248–252.
66. Che, Y.; Marshall, G. R. Biopolymers 2006, 81, 392–406.
67. Dugave, C. Cis-Trans Isomerization in Biochemistry; Weinheim: Wiley-VCH, 2006.
68. Vartak, A. P.; Skoblenick, K.; Thomas, N.; Mishra, R. K.; Johnson, R. L. J Med Chem 2007, 50, 6725–6729.
69. Lee, J.; Dubey, V. K.; Longo, L. M.; Blaber, M. J Mol Biol 2008, 377, 1251–1264.
70. Kuemin, M.; Nagel, Y. A.; Schweizer, S.; Monnard, F. W.; Ochsenfeld, C.; Wennemers, H. Angew Chem Int Ed 2010, 49, 6324–6327 and references therein.

Reviewing Editor: David A. Case