# Tryptophan Rich Peptides: Influence of Indole Rings on Backbone Conformation

Radhakrishnan Mahalakshmi,[1] Anindita Sengupta,[2] Srinivasarao Raghothama,[3] Narayanaswamy Shamala,[2] Padmanabhan Balaram[1]

[1] Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

[2] Department of Physics, Indian Institute of Science, Bangalore 560012, India

[3] NMR Research Center, Indian Institute of Science, Bangalore 560012, India

**ABSTRACT:**

Synthetic peptides with defined secondary structure scaffolds, namely hairpins and helices, containing tryptophan residues, have been investigated in this study to probe the influence of a large number of aromatic amino acids on backbone conformations. Solution NMR investigations of **Boc-W-L-W-$^D$P-G-W-L-W-OMe** (peptide **1**), designed to form a well-folded hairpin, clearly indicates the influence of flanking aromatic residues at the $^D$Pro–Gly region on both turn nucleation and strand propagation. Indole–pyrrolidine interactions in this peptide lead to the formation of the less-frequent type I′ turn at the $^D$Pro–Gly segment and frayed strand regions, with the strand residues adopting local helical conformations. An analog of peptide **1** with an Aib–Gly turn-nucleated hairpin (**Boc-W-L-W-U-G-W-L-W-OMe** (peptide **2**)) shows a preference for helical structures in solution, in both chloroform and methanol. Peptides with either one (**Boc-W-L-W-U-W-L-W-OMe** (peptide **3**)) or two (**Boc-U-W-L-W-U-W-L-W-OMe** (peptide **4**)) helix-nucleating Aib residues give rise to the well-folded helical conformations in the chloroform solution. The results are indicative of a preference for helical folding in peptides containing a large number of Trp residues. Investigation of a tetrapeptide analog of peptide **2**, **Boc-W-U-G-W-OMe** (peptide **5**), in solution and in the crystal state (by X-ray diffraction), also indicates a preference for a helical fold. Additionally, peptide **5** is stabilized in crystals by both aromatic interactions and an array of weak interactions. Examination of Trp-rich sequences in protein structures, however, reveals no secondary structure preference, suggesting that other stabilizing interactions in a well-folded protein may offset the influence of indole rings on backbone conformations. © 2006 Wiley Periodicals, Inc. Biopolymers (Pept Sci) 88:36–54, 2007.

*Keywords:* tryptophan peptides; aromatic pairs; conformational interconversion; indole–pyrrolidine interactions; weak interactions; NMR

## INTRODUCTION

Tryptophan has attracted the attention of chemists and biologists alike for several decades, owing to several of its unique properties. Trp residues are most widely present at the membrane interface of transmembrane proteins[1,2] and in hydrophobic clusters.[3] Trp-rich

**Table I  Sequences of Peptides Investigated**

| No. | Peptide Sequence[a] | Expected Structure | Observed Structure[b] |
|---|---|---|---|
| 1 | Boc-W-L-W-$^D$P-G-W-L-W-OMe | Hairpin | Frayed hairpin with a type I′ turn (NMR) |
| 2 | Boc-W-L-W-U-G-W-L-W-OMe | Helix (or) Hairpin | Helix (NMR) |
| 3 | Boc-W-L-W-U-W-L-W-OMe | Helix | Helix (NMR) |
| 4 | Boc-U-W-L-W-U-W-L-W-OMe | Helix | Helix (NMR) |
| 5 | Boc-W-U-G-W-OMe | Incipient helix (or) Incipient hairpin | Incipient hairpin with type I′ turn (NMR); Incipient helix with consecutive type II–I′ turns (X-ray) |

[a] The one letter code is used of amino acids: $^D$P = $^D$Pro; U = Aib.

[b] Spectroscopic technique employed to derive secondary structure information.

sequences are prevalent in antimicrobial peptides, for example, Indolicidin and Gramicidin A.[4,5] From the structural point of view, several well-folded water soluble peptides have been examined, wherein the incorporation of one or more Trp residues imparts protein-like properties to short sequences, for example the TrpZip sequences.[6] Examples of stereospecific aromatic interactions in organic solvents also exist, wherein such interactions stabilize strand segments of peptide hairpins.[7,8] The presence of a large number of Trp residues might be expected to not only contribute to stability but also influence secondary structure, because of the ability of the indole ring to participate in both electrostatic (hydrogen bonding, amide-$\pi$, cation-$\pi$, NH...$\pi$, CH...$\pi$) and $\pi$-stacking interactions.[9–12]

The presence of Trp residues in peptide sequences may influence backbone conformations by means of two distinctly different modes of interaction. These are: (i) weakly polar interactions involving the indole side chain and side chains of aromatic residues that are noncontiguous.[13,14] Such interactions may formally be considered as tertiary or long range interactions. The indole side chain may also preferentially interact with other side chains such as the guanidinium group of arginine.[9] (ii) The positioning of the indole ring in orientations favorable for interaction with the preceding and succeeding peptide units may influence local conformational choice,[15–18] and such interactions may be classified as short range interactions. In peptide hairpins, cross-strand aromatic interactions are favored when the residue pairs occupy the nonhydrogen bonding position. The TrpZip peptides, for example, are highly stabilized by the presence of two interacting Trp–Trp pairs at the nonhydrogen bonding position.[6] At the hydrogen bonding position, the cross-strand interaction geometries are expected to be much less favorable, thereby tilting the balance in favor of interactions with adjacent peptide units.

In this study, the influence of Trp residues on local backbone conformation has been probed using a set of Trp-rich peptides (Table I), which contain centrally positioned, structure-promot-

ing residues. Designed sequences with a tendency to form hairpins and helices have been investigated. The results indicate that in the case of peptides containing a large number of Trp residues, interactions of the indole ring with the peptide backbone as well as the turn region tend to destabilize hairpin formation, but Trp residues seem to be preferentially accommodated in helical scaffolds. Local helical conformations are also observed in the case of small peptide segments, both in the crystal and in solution.

## EXPERIMENTAL SECTION

### Peptide Synthesis

All peptides were synthesized by conventional solution phase chemistry using a fragment condensation strategy.[19] The *tert*-butyloxycarbonyl (Boc-) and methoxy (-OMe) groups were used for N- and C-termini protections, respectively. Deprotections were carried out using 98–100% formic acid and saponification using methanolic NaOH for the N- and C- termini, respectively. Couplings were mediated using dicyclohexylcarbodiimide/hydroxybenzotriazole. The final peptides were purified using medium pressure liquid chromatography (reverse phase, $C_{18}$, pore size 40–60 $\mu$m) and the purity of the peptides assessed using high performance liquid chromatography (HPLC) (reverse phase, $C_{18}$, 10 $\mu$m) using methanol–water gradients. Retention times obtained for a 35-min methanol/water gradient from 75 to 95% methanol on the reverse-phase HPLC: 15.7 min (**1**), 14.7 min (**2**), 13.8 min (**3**), 15.5 min (**4**), 16.2 min (**5**; 65–95% methanol in 40 min). Mass spectra were recorded on a Kompaq SEQ MALDI-TOF mass spectrometer (Kratos Analytical, Manchester, UK). Peptide **1**: $M_{Na}^+ = 1278.0$ Da, $M_K^+ = 1294.2$ Da, $M_{calc} = 1256.0$ Da; Peptide **2**: $M_{Na}^+ = 1267.5$ Da, $M_K^+ = 1283.5$ Da, $M_{calc} = 1244.0$ Da; Peptide **3**: $M_{Na}^+ = 1209.2$ Da, $M_K^+ = 1225.1$ Da, $M_{calc} = 1187.0$ Da; Peptide **4**: $M_{Na}^+ = 1295.8$ Da, $M_K^+ = 1312.1$ Da, $M_{calc} = 1273.0$ Da; Peptide **5**: $M_{Na}^+ = 669.0$ Da, $M_K^+ = 684.9$ Da, $M_{calc} = 646.0$ Da. All the target peptides were fully characterized by 400/500-MHz NMR spectroscopy.

### Nuclear Magnetic Resonance

All NMR experiments were carried out on a Bruker DRX-500 spectrometer using peptide concentrations of ∼10 m$M$ in CDCl$_3$/

CDCl$_3$ + 10% DMSO-d$_6$. Peptide **5** was studied on an AMX-400 spectrometer. Complete resonance assignment was achieved using a combination of TOCSY[20] and ROESY[21,22] experiments. 2D spectra were recorded in the phase-sensitive mode using TPPI methods. 1024 and 512 data points were collected in the $f_2$ and $f_1$ dimensions, respectively. NMR data were processed using Bruker XWINNMR software on a Silicon Graphics Indy workstation. The data were zero-filled to 2 K points in the $f_1$ dimension and a shifted ($\pi/2$) sine-squared window function was applied to both the dimensions prior to Fourier transformation. Hydrogen bonding information was obtained from CDCl$_3$–DMSO titration experiments.[23–25]

### X-ray Diffraction

Single crystals of peptide **5** suitable for X-ray diffraction were grown from chloroform by slow evaporation. The X-ray intensity data were collected at room temperature from a dry crystal (0.25 × 0.13 × 0.01 mm$^3$) mounted on a Bruker AXS SMART APEX CCD diffractometer using MoK$_\alpha$ radiation ($\lambda = 0.71073$ Å). $\omega$ scan type with $2\theta = 40.14°$ was used to obtain a total of 3980 independent reflections. The space group is P2$_1$ with $a = 12.951$ (5), $b = 11.368$ (4), $c = 14.800$ (5) Å, $\beta = 101.411$ (7)°, $V = 2135.8$ (13) Å$^3$, $Z = 2$ for chemical formula C$_{34}$H$_{42}$N$_6$O$_7$.CHCl$_3$, with one molecule per asymmetric unit; $\rho_{calc} = 1.191$ g cm$^{-3}$, $\mu = 0.263$ mm$^{-1}$, $F$ (000) = 804, and $R_{int} = 0.0954$. The structure was solved by direct phase determination using SHELXD.[26] All 47 atoms could be located from the electron density map. Refinement was carried out against F$^2$ with full-matrix least squares methods by using SHELXL-97.[27] Nonhydrogen atoms were all initially refined isotropically followed by full matrix anisotropic least-squares refinement. At the end of isotropic refinement, the $R$-factor was 19.72% and dropped to 14.90% after the anisotropic refinement. Two chloroform sites (I and II with occupancy of 0.50/0.50) were located from the difference Fourier maps. The atom CA1 in site I of chloroform molecule showed nonpositive definite temperature factor during the course of refinement. Hence the atom CA1 was included in the subsequent refinement with isotropic temperature factor. All the hydrogen atoms were fixed geometrically in the idealized positions and refined in the final cycle of refinement as riding over the atoms to which they are bonded. The final $R$-factor was 10.95% (w$R_2$ = 27.06%) for 1777 observed reflections with $|F_o| \geq 4\sigma|F_o|$ and 480 variables, where the data-to-parameter ratio is 3.7:1 and G$_o$F = 1.052. The largest difference peak and hole were 0.501 and $-0.318$ eÅ$^3$, respectively. CCDC-618920 contains the supplementary crystallographic data for peptide **5**. This data can be obtained free of charge from the Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif or by e-mail to deposit@ccdc.cam.ac.uk.

### Database Analysis

A dataset of high-resolution proteins was obtained from the Protein Data Bank (www.rcsb.org/pdb) using the protein sequence culling server list dated July 30, 2006.[28] Parameters used for the choice of the dataset included: resolution cutoff of 2 Å, identity cutoff of 20%, $R$-factor of <0.25. The number of chains retrieved totaled 2346. Search for the Trp-Xxx-Trp (where Xxx = any of the 20 amino acids) sequences were carried out using simple PERL-scripts and the coordinates thus retrieved for each of the 20 amino acids were used for calculation of backbone torsion angles. They were then sorted into helices, sheets, or coils, based on the backbone torsion angles of Trp-Xxx-Trp residues.

## RESULTS AND DISCUSSION

Reports from Cochran's group have clearly defined the role of Trp-pairs, at the nonhydrogen bonding position, in peptide stability.[6] At the hydrogen bonding position, the C$\beta$ atoms of facing residues point outwards from the edges of the hairpin. Appropriate choices of the side chain torsion angles can place aromatic side chains in proximity at this position. Studies of model systems to investigate potential cross-strand aromatic interactions at the hydrogen bonding position have not been reported. To address this issue, peptide **1** was designed and synthesized by conventional solution phase methods. Turn nucleation was achieved using a $^D$Pro–Gly segment.[29] Trp residues were positioned at the hydrogen bonding sites of the strand regions such that two pairs of aromatic interactions are possible in the octapeptide. The indole side chains could then interact with each other either across the strands (1↔8 and 3↔6) or along the strands (1↔3 and 6↔8) (Figure 1).

The presence of $^D$Pro in the center of the sequence precludes the formation of alternate secondary structures, like a helix. In peptide **2**, the $^D$Pro residue was replaced by Aib, since the Aib–Gly sequence is capable of forming both type I and I′ $\beta$-turns. The former is compatible with a helical fold, while the latter promotes hairpin formation. An earlier analysis on the hairpin-nucleating ability of Aib–Gly suggested that the peptide adopts a solvent-dependent secondary structure, forming a hairpin in hydrogen-bonding solvents such as methanol and DMSO, and adopting a helical conformation in nonhydrogen bonding solvents such as chloroform and acetonitrile.[30] It was therefore of interest to probe whether an analog of peptide **1** would indeed show such solvent dependence of secondary structure or would adopt one preferred conformation in solution. Additionally, the influence of flanking aromatic pairs on the turn nucleating ability of the Aib–Gly segment could also be examined.

### Conformational Analysis of Peptides 1 and 2

Both peptides examined were highly soluble in nonpolar, organic solvents. However, broad resonances were obtained for backbone NH and C$^\alpha$H protons at 500 MHz in chloroform solutions, at the concentrations used (10 m$M$) for NMR studies. A similar line broadening was also observed at lower concentrations (up to 2 m$M$). This line broadening is possibly due to the peptide association mediated by intermolecular hydrogen bonds.[31] The addition of a strongly hydrogen-bonding solvent like DMSO, leads to the disruption of these
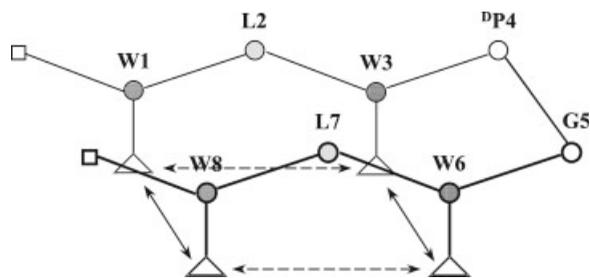
**FIGURE 1** Possible modes of aromatic interactions in peptide **1**. Cross-strand interactions are indicated by bold arrows while interactions within strands are marked by dashed arrows. Residues forming the turn region (^DP and G) are indicated by empty circles, those in the nonhydrogen bonding position (L2 and L7) are indicated by dotted circles, and those in the hydrogen bonding position (W1, W3, W6, and W8) are indicated by checked circles. Squares represent the N- and C-terminal protecting groups.

aggregates, resulting in the observation of sharp backbone NH and $C^{\alpha}H$ resonances. The peptides were consequently studied in $CDCl_3$ + 10% DMSO-$d_6$. Peptide **2** was additionally examined in methanol, to probe solvent dependence of secondary structure, as previous reports of peptides containing an Aib–Gly turn have been shown to form helices in nonhydrogen bonding solvents like chloroform and hairpins in the hydrogen bonding solvents like methanol.[30]

***Backbone Conformations of Peptide 1, Boc-W-L-W-^DP-G-W-L-W-OMe.*** The 500-MHz $^1H$ 1D spectrum of peptide **1** in $CDCl_3$ + 10% DMSO-$d_6$ was characterized by the well-dispersed NH resonances appearing between 5.5 and 8.5 ppm. Secondary structure information was derived from diagnostic NOEs obtained in the ROESY spectrum of the peptide (Figure 2). The $3\alpha$–$4\delta$ NOE was indicative of a *trans* conformation of the Trp3–^DPro4 tertiary amide unit and the turn segment was identified by a Gly5–Trp6 $N_iH$–$N_{i+1}H$ NOE. A striking feature of the ROESY spectrum was the presence of a strong $d_{\delta N}$ $(i - i + 1)$ NOE between ^DPro4 $C^{\delta}H$ and Gly5 NH, supporting $\psi^{DPro}$ of ~ +30°, consistent with a type I' $\beta$-turn conformation at ^DPro–Gly. NOEs indicative of strand registry, namely the $d_{NN}$ Trp3 NH–Trp6 NH and Trp1 NH–Trp8 NH NOEs were absent and the Leu2 $C^{\alpha}H$–Leu7 $C^{\alpha}H$ $(d_{\alpha\alpha})$ NOE was weak. The $^3J_{HN-C^{\alpha}H}$ coupling constants obtained for peptide **1** were in the range 5.1–8.5 Hz, which is vastly different from those expected for well-folded hairpins[32] (8.0–9.5 Hz). The temperature coefficients for strand residues anticipated to be involved in internal hydrogen bonding, in peptide **1**, were Trp1 = −3.9, Trp3 = −7.9, Trp6 = −3.3, Trp8 = −7.2 (all values in ppb/K), while in well-folded hairpins, they are generally <2.0 ppb/K. These observations together suggested that further progression of the

hairpin in peptide **1** beyond the type I' turn was not achieved, and the strand regions were frayed. Investigation of the NH region of the ROESY spectrum showed the presence of sequential $(i - i + 1)$ $d_{NN}$ NOEs $(N_iH \leftrightarrow N_{i+1}H)$ between residues constituting the strands (Figure 2). Strong $d_{NN}$ NOEs characteristic of a local helical conformation ($\phi$–$\psi$ values of −60° ± 20° and −30° ± 20°, respectively) observed at the strand segments additionally supported the conclusion that the backbone torsion angles deviated substantially from those anticipated for an extended conformation.

A strong NOE between ^DPro4 $C^{\alpha}H$–Trp6 $C^{\delta 1}H$ was observed in the ROESY spectrum of peptide **1** (Figure 2), indicative of interactions between the pyrrolidine ring of ^DPro4 and the indole of Trp6. As noted earlier, most ^DPro–Gly $\beta$-turns characterized thus far are of the type II' class.[29,33] It can be anticipated that indole–pyrrolidine interactions in peptide **1** not only lead the formation of a type I' turn at the ^DPro–Gly segment, but also contribute to strand fraying beyond the turn region, resulting in a poorly folded hairpin. Positioning of Trp residues immediately after turn regions probably leads to turn destabilization. Conversely, Trp residues located away from the turn lend stability to the peptide by the formation of strong aromatic interactions, in both aqueous media[6] and organic solvents.[7] Examples of cross-strand (diagonal) aromatic interactions also exist when Trp residues are positioned at the hydrogen-bonding position, but away from the turn segment.[34]

***Backbone Conformations of Peptide 2, Boc-W-L-W-U-G-W-L-W-OMe.*** The ROESY spectrum of peptide **2** in $CDCl_3$ (+10% DMSO-$d_6$) showed the presence of only two sequential $d_{NN}$ NOEs $(N_iH \leftrightarrow N_{i+1}H)$,(1–2, 2–3) intraresidue $(i - i)$ $d_{N\alpha}$ $(N_iH \leftrightarrow C_i^{\alpha}H)$ NOEs of medium intensity and weak sequential $(i - i + 1)$ $d_{\alpha N}$ $(C_i^{\alpha}H \leftrightarrow N_{i+1}H)$ NOEs, NOE evidence being limited to a small segment of the peptide in solution. The scarcity of NOEs obtained suggested that there was no well-formed structure for peptide **2** in $CDCl_3$.

Examination of the ROESY spectrum of peptide **2** in methanol (Figure 3) revealed the presence of several sequential $d_{NN}$ NOEs, strong intraresidue $(i - i)$ $d_{N\alpha}$ $(N_iH \leftrightarrow C_i^{\alpha}H)$ NOEs and weak sequential $(i - i + 1)$ $d_{\alpha N}$ $(C_i^{\alpha}H \leftrightarrow N_{i+1}H)$ NOEs. Additionally, a few $(i - i + 2)$ $d_{\alpha N}$ NOEs $(C_i^{\alpha}H \leftrightarrow N_{i+2}H)$ were also observed (Figure 3), suggesting that the peptide adopted a helical conformation in methanol. In contrast, a previous report on the peptide Boc-Leu-Val-Val-Aib-Gly-Leu-Val-Val-OMe,[30] which is an aliphatic analog of peptide **2**, established a solvent-dependent conformation at equilibrium. In this case the peptide adopted a $\beta$-hairpin fold in methanol, while a helix was seen in $CDCl_3$. The result obtained with peptide **2** suggests that the presence of a large
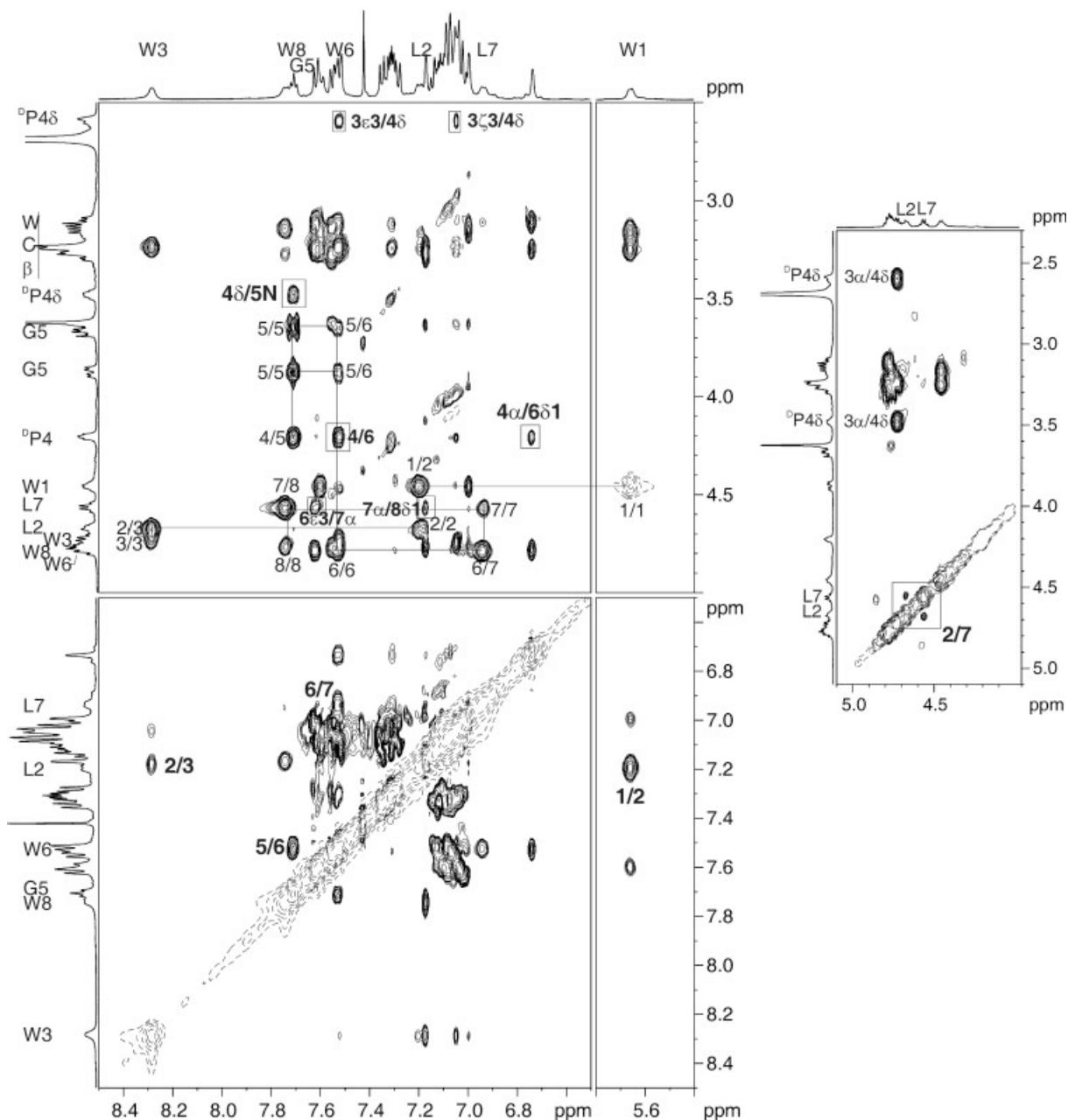
**FIGURE 2**   Partial expansions of the ROESY spectrum of peptide **1**, Boc-W-L-W-$^D$P-G-W-L-W-OMe, recorded on a Bruker DRX 500 instrument in CDCl$_3$ + 10% DMSO-d$_6$ at 303 K. Expansions of the $d_{NN}$ and $d_{N\alpha}$ region are shown to the left and expansion of the $d_{\alpha/\beta}$ region to the right.

number of Trp residues tilts the equilibrium in favor of helical conformations, in both chloroform and methanol. The absence of any anomalous chemical shifts of the indole resonances indicates that the Trp rings are not spatially proximal. Aromatic interactions are widely observed when the residues occupy facing regions of strand segments in hairpins.[6] Such interactions are geometrically much less favorable when the residues occupy the $i - i + 3/4$ positions in helices, despite the fact that the residues project out onto the same face of

the helix. Also, very few examples of such interactions have been reported, to date, in proteins.[35]

Conformational analysis of peptide **1** clearly indicates that the interactions of indole rings with the turn segments influence turn type and contribute to the formation of a local helical conformation at the Trp-Leu-Trp segments. Results obtained from the NMR investigations of peptide **2**, in both chloroform and methanol, suggest a predisposition for the helical conformation. To further validate the preference of
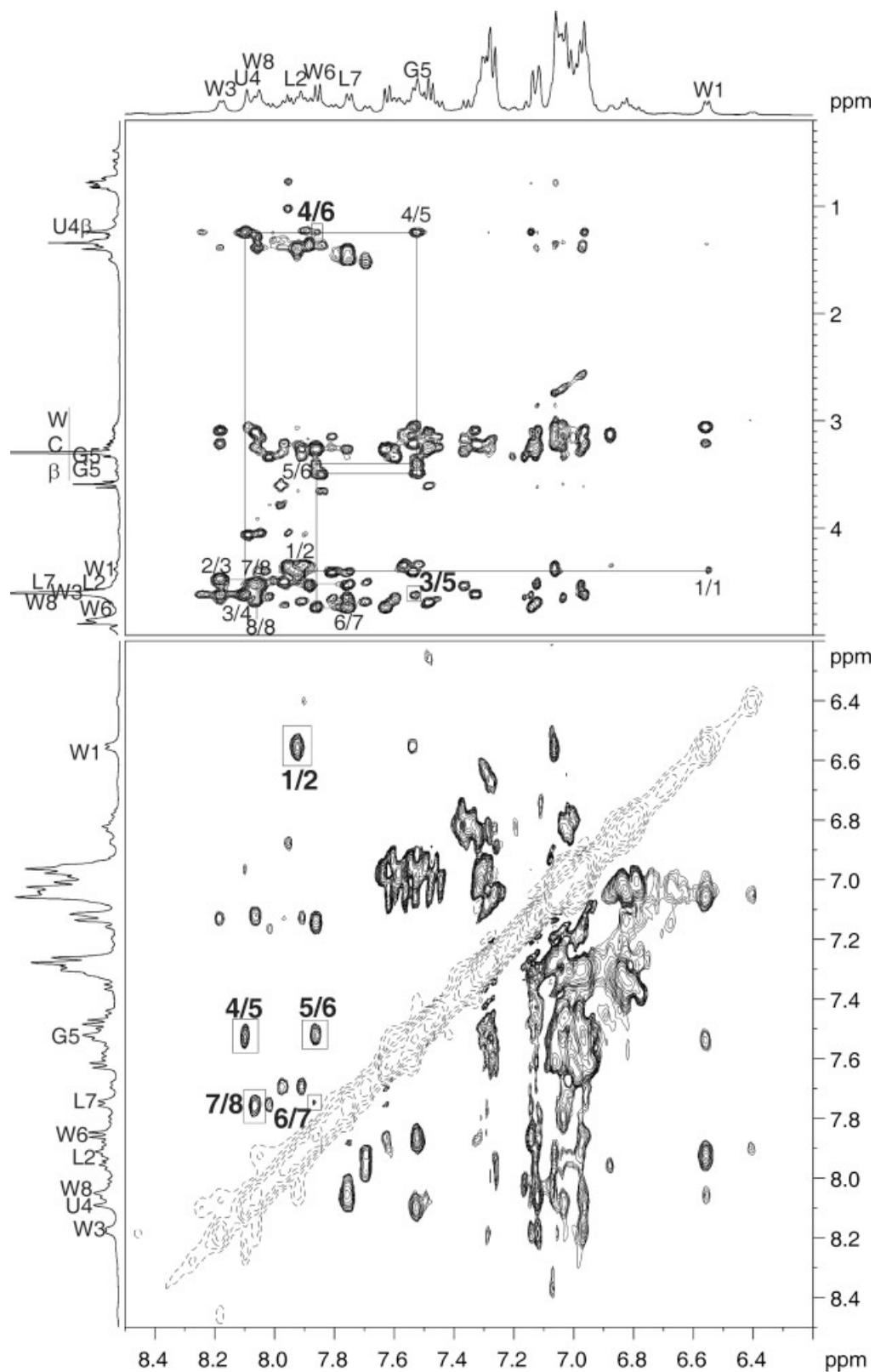
**FIGURE 3** Partial expansions of the ROESY spectrum of peptide **2**, Boc-W-L-W-U-G-W-L-W-OMe in methanol. Signature long-range NOEs characteristic of a helix are boxed.

helical conformations in Trp-rich sequences, helical scaffolds nucleated by Aib residues were synthesized and examined.

## Conformational Analysis of Peptides 3 and 4

Both peptides **3** and **4** were highly soluble in chloroform, giving sharp resonances in the NMR spectrum. Structural studies were therefore carried out in the same solvent. Complete resonance assignment of the backbone protons was achieved using a combination of TOCSY and ROESY experiments.

*Backbone Conformations of Peptide 3, Boc-W-L-W-U-W-L-W-OMe.* The ROESY spectrum of peptide **3** in chloroform (Figure 4) clearly indicated strong sequential $(i - i + 1)$ $d_{NN}$ NOEs ($N_iH \leftrightarrow N_{i+1}H$) and small $^3J_{HN-C^{\alpha}H}$ coupling constants (ranging between 1.0–4.0 Hz, except Leu6 (7.9 Hz) at the C-terminus) for all residues, indicative of a helical scaffold for the peptide. Most of the intraresidue NOEs $(i - i)$ $d_{N\alpha}$ ($N_iH \leftrightarrow C_i^{\alpha}H$) were strong and the sequential $(i - i + 1)$ $d_{\alpha N}$ ($C_i^{\alpha}H \leftrightarrow N_{i+1}H$) NOEs were weak. In addition all $(i - i + 2)$ $d_{\alpha N}$ NOEs ($C_i^{\alpha}H \leftrightarrow N_{i+2}H$) were observed, clearly demonstrating that the peptide adopted a helical conformation in solution. An $(i - i + 3)$ $d_{\alpha N}$ NOE ($C_i^{\alpha}H \leftrightarrow N_{i+3}H$) between residues Trp3 and Leu6 was also observed. It was however surprising to note that the $d\delta/dT$ values obtained from the temperature dependence of backbone NH chemical shifts in CDCl$_3$ + 10% DMSO-d$_6$ gave large $d\delta/dT$ values, which ranged between −4.0 to −5.0 ppb/K (except for residues Leu2 (−7.6 ppb/K) and Trp3 (−8.9 ppb/K)), whereas temperature coefficients of backbone amides in helical peptides are generally small, of the range −1.0 to −3.0 ppb/K. The addition of a strong hydrogen bonding solvent like DMSO may have contributed to partial unfolding of the helix, giving rise to the observed temperature coefficients.

Interestingly, the only difference in the sequences of peptides **2** and **3** is the Gly residue, following Aib, in peptide **2**. Removal of a single residue has not only contributed to increased solubility of peptide **3** (as against peptide **2**) but has also led to the formation of a stable helical scaffold in chloroform. The ability of Gly residues to undergo large conformational variations (as is evident from the allowed regions for this residue in the $\phi$–$\psi$ plot) might account for the poor solubility (and structure) of peptide **2** in nonhydrogen bonding solvents.

*Backbone Conformations of Peptide 4, Boc-U-W-L-W-U-W-L-W-OMe.* The presence of a single Aib residue was found to nucleate a helical scaffold in peptide **3**. To further stabilize the helix, a second Aib residue was incorporated in the sequence. The N-terminus was chosen, as helix nucleation by Aib residues at the N-terminal region has been well established in several examples of synthetic peptide heli-

ces.[36,37] As anticipated, peptide **4**, with the sequence Boc-(Aib-Trp-Leu-Trp)$_2$-OMe showed a well-resolved $^1$H 1D spectrum in CDCl$_3$. The ROESY spectrum of the peptide (Figure 5) clearly showed evidence for the presence of strong sequential $(i - i + 1)$ $d_{NN}$ NOEs ($N_iH \leftrightarrow N_{i+1}H$), strong intraresidue NOEs $(i - i)$ $d_{N\alpha}$ ($N_iH \leftrightarrow C_i^{\alpha}H$) and weak sequential $(i - i + 1)$ $d_{\alpha N}$ ($C_i^{\alpha}H \leftrightarrow N_{i+1}H$) NOEs. In addition, most of the $(i - i + 2)$ $d_{\alpha N}$ NOEs ($C_i^{\alpha}H \leftrightarrow N_{i+2}H$) (others undetectable because of the resonance overlap) and $(i - i + 3)$ $d_{\alpha N}$ NOEs ($C_i^{\alpha}H \leftrightarrow N_{i+3}H$) were observed (Figure 5). These NOEs established beyond doubt that the peptide adopted a stable helical conformation in solution. The temperature coefficients obtained in CDCl$_3$ + 10% DMSO-d$_6$ (all values ranging between −2.0 to −3.5 ppb/K, except for Leu1 = −10.6 ppb/K) also established the formation of strong internal hydrogen bonds along the peptide backbone. It was interesting to note that the backbone amide of Trp2 and Leu3 also showed low $d\delta/dT$ values. The NOEs observed between -CH$_3$ of U1 and the ring protons of W2 in the ROESY spectrum suggested that the ring was possibly oriented over the N-terminus, such that it sterically hindered solvent accessibility to Trp2 NH and Leu3 NH, thereby giving rise to low temperature coefficients for these resonances.

*Chemical Shift Analysis of Peptides 1–4.* A comparison of the C$^{\alpha}$H chemical shifts of the four peptides in CDCl$_3$ + DMSO-d$_6$/CDCl$_3$ was carried out. The C$^{\alpha}$H shifts of peptide **1** were found to be visibly distinct from that of the other three peptides. This observation was supported by NOE data, which suggested that peptide **1** adopted a frayed hairpin conformation, whereas peptides **2–4** adopted helical conformations in solution. It must also be noted that very few NOEs were observed in peptide **2** in CDCl$_3$ + DMSO-d$_6$; comparison of backbone chemical shifts of peptide **2** with those of peptide **1** and peptides **3/4** suggested that the peptide adopted a local helical conformation in solution. Chemical shift indexing of the different peptides was also carried out with random coil values obtained from BMRB.[38] The CSI values supported the structures obtained from NOE data.

The solution structures of peptides **1–4** have also been independently established using vibrational circular dichroism (VCD) experiments in methanol, and have been reported elsewhere.[39] The results of VCD experiments correlate with the observed structures derived from solution NMR methods.

## Conformational Analysis of Peptide 5

The results of structural analyses of peptides **1–4** indicated that in the presence of several indole side chains, the favored backbone conformation was a helix. In the case of peptide **2**,
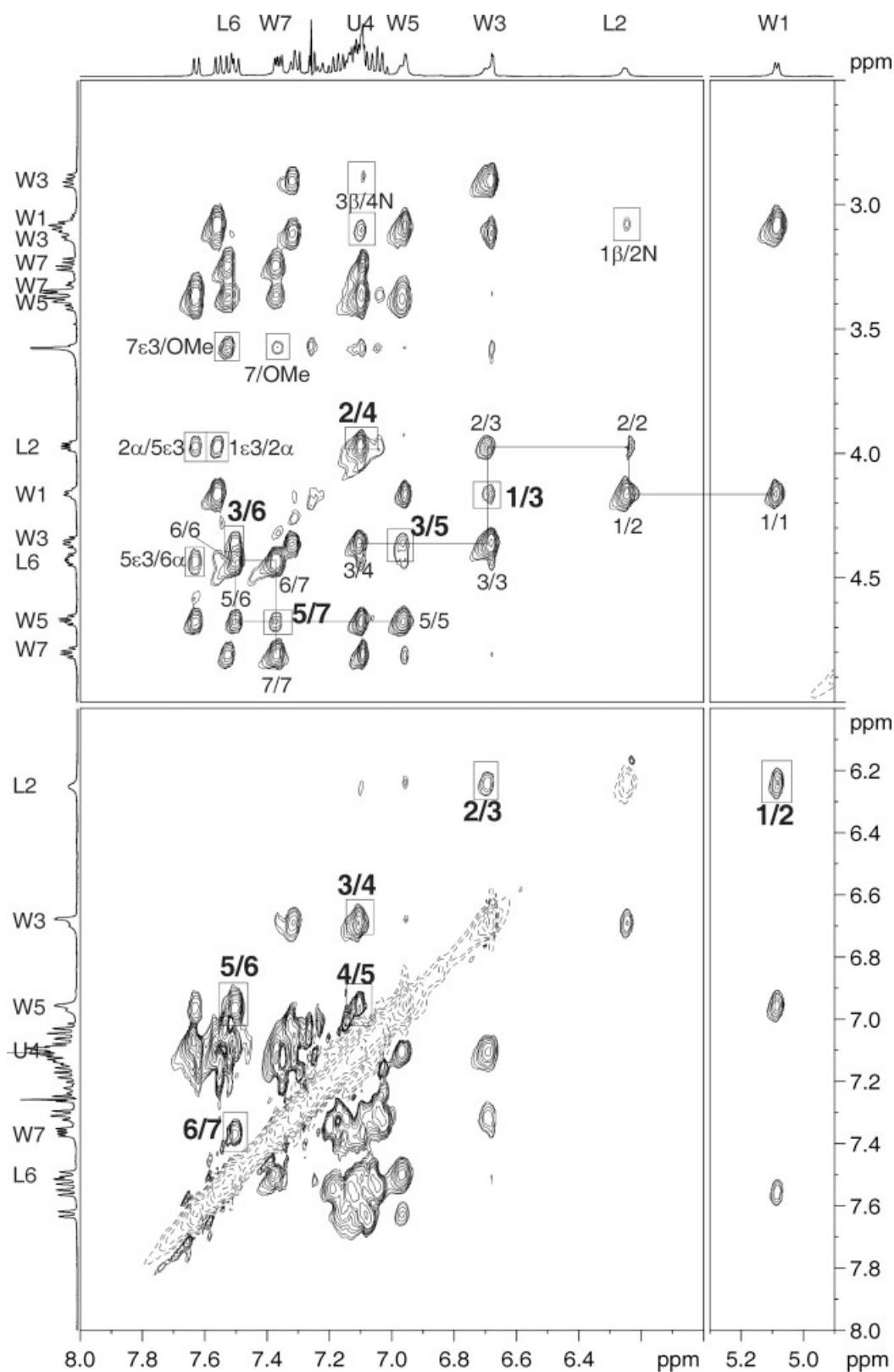
**FIGURE 4** Partial expansions of the ROESY spectrum of peptide **3**, Boc-W-L-W-U-W-L-W-OMe, in CDCl$_3$. Signature long range NOEs characteristic of a helix are boxed.
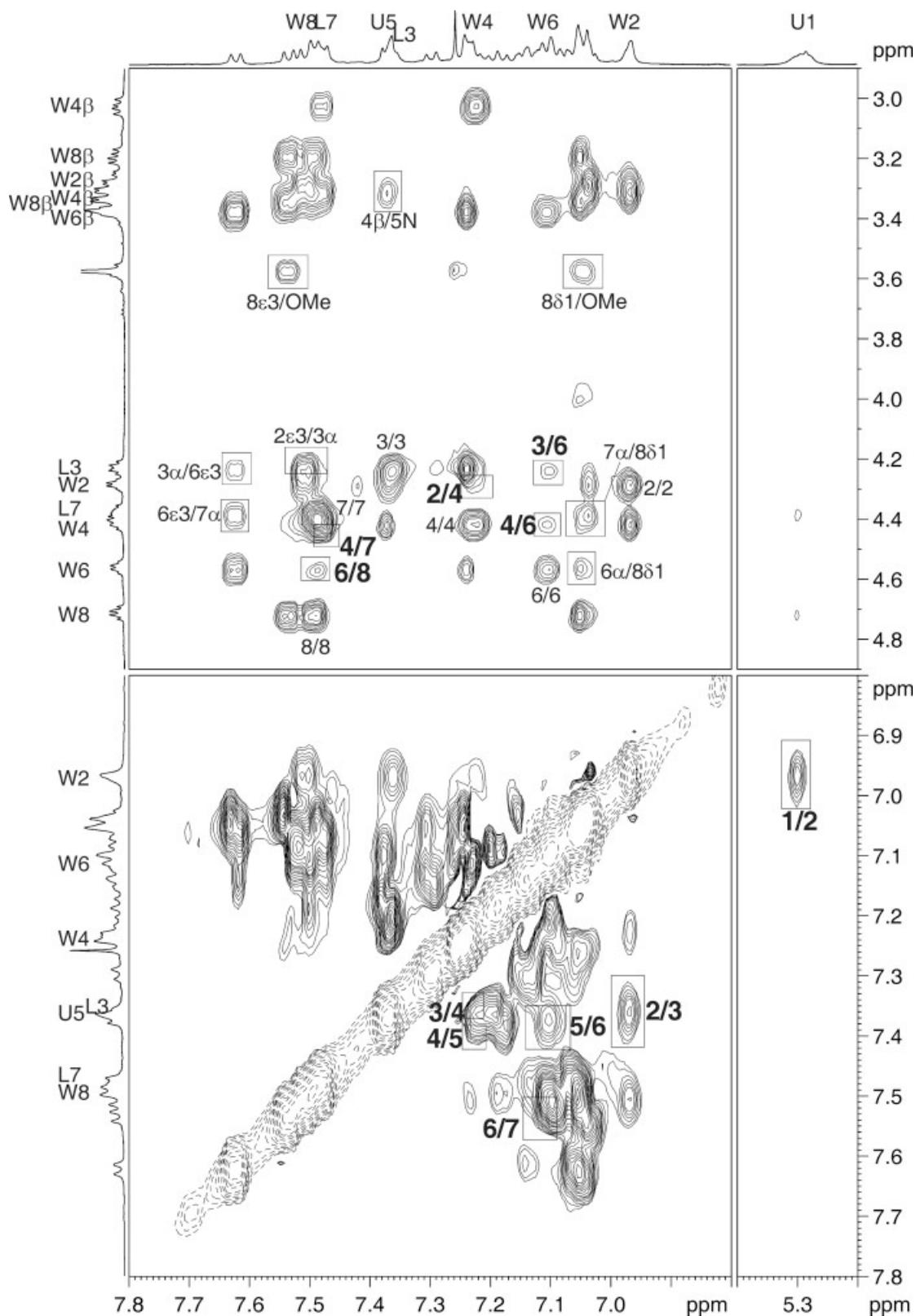
**FIGURE 5** Partial expansions of the ROESY spectrum of peptide **4**, Boc-U-W-L-W-U-W-L-W-OMe, in CDCl₃. Signature long range NOEs characteristic of a helix are boxed.
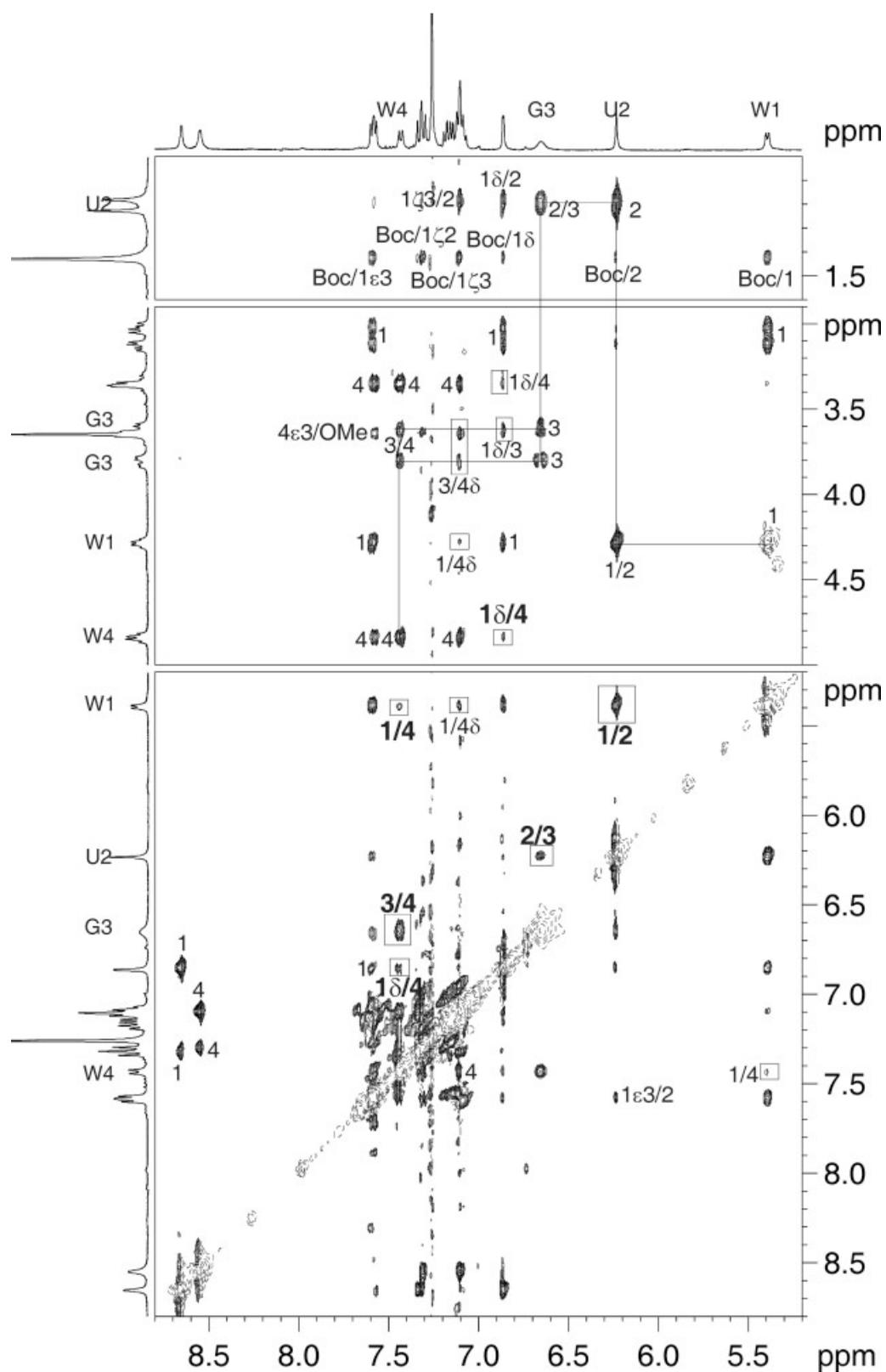
**FIGURE 6** Partial expansions of the ROESY spectrum of peptide **5**, Boc-W-U-G-W-OMe, in CDCl$_3$. Note the presence of several $i - i + 2/3$ NOEs even in a four-residue sequence.
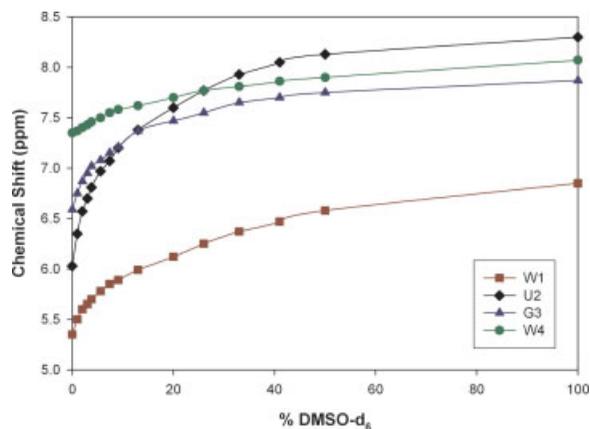
**FIGURE 7** Plot of NH proton chemical shifts of peptide **5** as a function of increasing DMSO concentration.

a helical conformation was favored in both methanol and chloroform, contrary to previous reports on solvent-dependent secondary structure formation in peptides with an Aib–Gly turn.[30] To further understand the influence of flanking aromatic residues on turn nucleation or helix formation in peptide **2**, a model tetrapeptide Boc-Trp-Aib-Gly-Trp-OMe was designed, with the Aib–Gly segment constituting the turn region.

*Conformations of Peptide 5, Boc-W-U-G-W-OMe, in Solution.* The 1D spectrum of the peptide in chloroform gave

well-resolved resonances that could be readily assigned using TOCSY and ROESY experiments. It was interesting to note that the methyl resonances of Aib that usually are observed at $\sim 1.25$ and $\sim 1.40$ ppm were upfield shifted to 1.17 and 1.22 ppm in this sequence, indicating that the peptide was well-structured, with at least one of the aromatic rings positioned over the turn segment. Examination of the ROESY spectrum of the peptide indicated the presence of all sequential $d_{NN}$ NOEs, of which Trp1–Aib2 and Gly3–Trp4 were very strong (Figure 6). In addition, a very weak $d_{NN}$ NOE between the amides of Trp1 and Trp4 was also seen. The peptide also showed an unusually large number of sequential $(i - i + 1)$ and long-range $(i - i + 2, i - i + 3$ etc.) NOEs between the indole rings and the peptide backbone. DMSO titration experiments (Figure 7) indicated the presence of two hydrogen bonds in the peptide, with one of the NH donors being the backbone amide of Trp4.

Interestingly, NOEs observed for the peptide in the ROESY spectrum were incompatible with a single structure for the sequence, suggesting that the peptide underwent a conformational interconversion between more than one stable form in solution. The peptide can adopt several conformations in solution, of which the theoretically possible conformers with a minimum of two hydrogen bonds (as inferred from the DMSO titration data) are listed in Table II. Presence of $d_{NN}$ NOEs between Gly3 and Trp4 as well as Trp1 and

**Table II Theoretically Possible Structures for Peptide 5, With two Hydrogen Bonds and Diagnostic NOEs**

| Conformer | Residue | | | | Description | Diagnostic NOE(s) |
|---|---|---|---|---|---|---|
| | Trp | Aib | Gly | Trp | | |
| A | $P_{II}{}^a$ | $\alpha_L$ | $\alpha_L$ | | $\beta_{II}–\beta_{I'}$: Helical turn | W1C$^\alpha$H–U2NH; G3NH–W4NH |
| B | $\alpha_R$ | $\alpha_R$ | $\alpha_R$ | $\alpha_R$ | $\beta_{III}–\beta_{III}$: Incipient $3_{10}$ helix | Sequential $d_{NN}$ NOEs |
| C | $E^b$ | $\alpha_L$ | $\alpha_L$ | $E^a$ | $\beta$-hairpin with one type I′/III′ turn | W1NH–W4NH |

| NOEs[c] | Conformer A[d] | Conformer B[d] | Conformer C[d] |
|---|---|---|---|
| 1NH–2NH | − | ++ | −− |
| 2NH–3NH | + | ++ | + |
| 3NH–4NH | ++ | ++ | + |
| 1NH–4NH | −− | −− | ++ |
| 3αH–4NH | ++ | − | + |
| 1αH–3NH | −− | ++ | −− |
| 1αH–4NH | −− | + | −− |
| 2αH–4NH | −− | ++ | −− |

[a] Polyproline ($P_{II}$) conformation (Idealized parameters[40]: $\phi = -78.0°$; $\psi = +149°$).
[b] E: Extended conformation (antiparallel $\beta$-sheet; Idealized parameters[40]: $\phi = -139°$; $\psi = +135°$).
[c] Distances between protons range between 2.5 Å and 3.5 Å.
[d] Symbols '++' and '−−' indicate that presence of these NOEs are mandatory for the given conformer, while symbols '+' and '−' indicate that the NOEs may be present in the conformer.
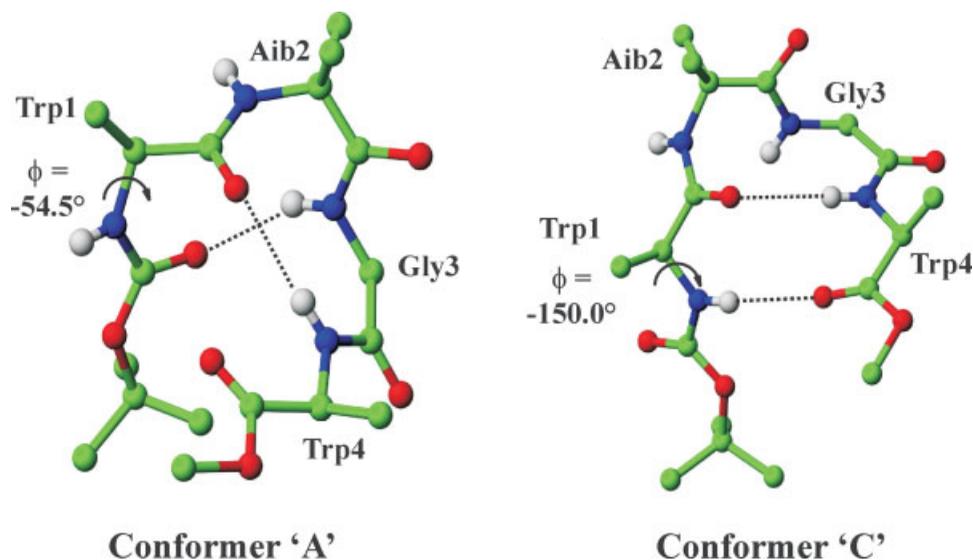
**FIGURE 8** Conversion of conformer 'A' (left) to conformer 'C' (right), obtained by changing Trp1 $\phi$ (indicated) from $-54.5°$ to $-150.0°$. Indole side chains are not shown for simplicity. Hydrogen bonds that will be formed in both conformations are indicated.

Trp4 indicated turn formation along residues 2–3 and formation of the first hydrogen bond between residues 1–4, respectively. A weak $d_{NN}$ NOE between residues 2 and 3 suggested the formation of a type I′/III′ turn along residues Aib2–Gly3, suggesting the presence of conformer 'C' in solution (Table II) in peptide **5**. Strong NOEs between Trp1 $C^{\alpha}H$–Aib2 NH and Gly3 NH–Trp4 NH also suggested the existence of conformer 'A' in solution. Although sequential $d_{NN}$ NOEs were obtained, differences in the intensity of these NOEs and the absence of diagnostic $d_{\alpha N}\ i-i+2/3$ NOEs indicated that an all-helical conformation (conformer 'B'), if present, was

probably adopted by only a minor population of the peptide molecules in chloroform.

The DMSO titration experiment also supported the presence of both conformers 'A' and 'C' in solution. A $\beta$-turn structure (conformer C) for the peptide should give rise to large chemical shift variations for amides of residues Aib2 and Gly3 and small chemical shift variations for residues Trp1 and Trp4 in the DMSO titration experiment (Figure 7). On the other hand, in conformer 'A', amides of residues Trp1 and Aib2 are exposed and should show large chemical shift changes on addition of DMSO. In both structures, Trp4 is hydrogen bonded and Aib2 is solvent exposed; this is confirmed by the chemical shift variation of these residues in Figure 7. Interestingly, the extents of change in chemical shift of the amides of Trp1 and Gly3, upon addition of DMSO, are intermediate to that of Aib2 and Trp4. It is well known that NMR spectra provide aver-
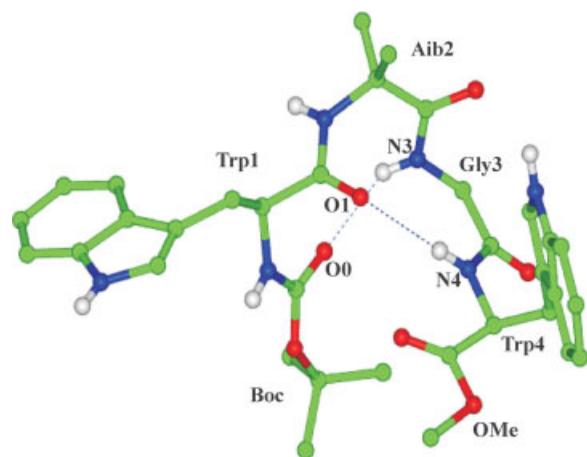


**FIGURE 9** Crystal-state conformation of peptide **5**, Boc-Trp-Aib-Gly-Trp-OMe (type II–I′ consecutive $\beta$-turns). Two intramolecular hydrogen bonds are shown in dotted lines.

**Table III  Torsion Angles (degrees)[43] for Boc-Trp-Aib-Gly-Trp-OMe (5)**

| Residue | $\phi$ | $\psi$ | $\omega$ | $\chi^1$ | $\chi^2$ |
|---|---|---|---|---|---|
| Trp1 | $-52.3^a$ | 129.5 | 175.8 | $-55.0$ | $-54.3$; 130.1 |
| Aib2 | 61.3 | 20.2 | 175.7 | | |
| Gly3 | 93.6 | 4.4 | $-180.0$ | | |
| Trp4 | $-155.7$ | $-177.4^b$ | $177.7^c$ | 68.9 | $-71.9$; 114.3 |

[a] $C0'-N1-C1A-C1'$.
[b] $N4-C4A-C4'-O0M$.
[c] $C4A-C4'-O0M-C0M$.

**Table IV  Hydrogen Bond Parameters for Boc-Trp-Aib-Gly-Trp-OMe (5)**

| Type | Donor | Acceptor | N...O (Å) | H...O (Å) | C—O...H (degree) | C—O...N (degree) | O...HN (degree) |
|---|---|---|---|---|---|---|---|
| Intramolecular | | | | | | | |
| 4→1[a] | N3 | O0 | 2.947 | 2.178 | 143.3 | 141.1 | 148.7 |
| 4→1 | N4 | O1 | 3.402 | 2.545 | 117.3 | 118.8 | 174.1 |
| Intermolecular[a] | | | | | | | |
| | N1ε1 | O1[b] | 2.955 | 2.114 | 133.9 | 133.4 | 165.5 |
| | N2 | O2[c] | 3.125 | 2.276 | 156.8 | 154.1 | 169.0 |

[a] These are acceptable hydrogen bonds.
[b] Symmetry related by $(-x + 1, y - 1/2, -z)$.
[c] Symmetry related by $(-x, y - 1/2, -z)$.

aged information of the entire population. This, in turn, suggested that both conformers 'A' and 'C' existed in solution and at any given time point, Trp1 NH and Gly3 NH were hydrogen bonded in ∼50% of the molecules, in conformers 'C' and 'A', respectively, giving rise to the intermediate variation in amide chemical shifts for these residues,
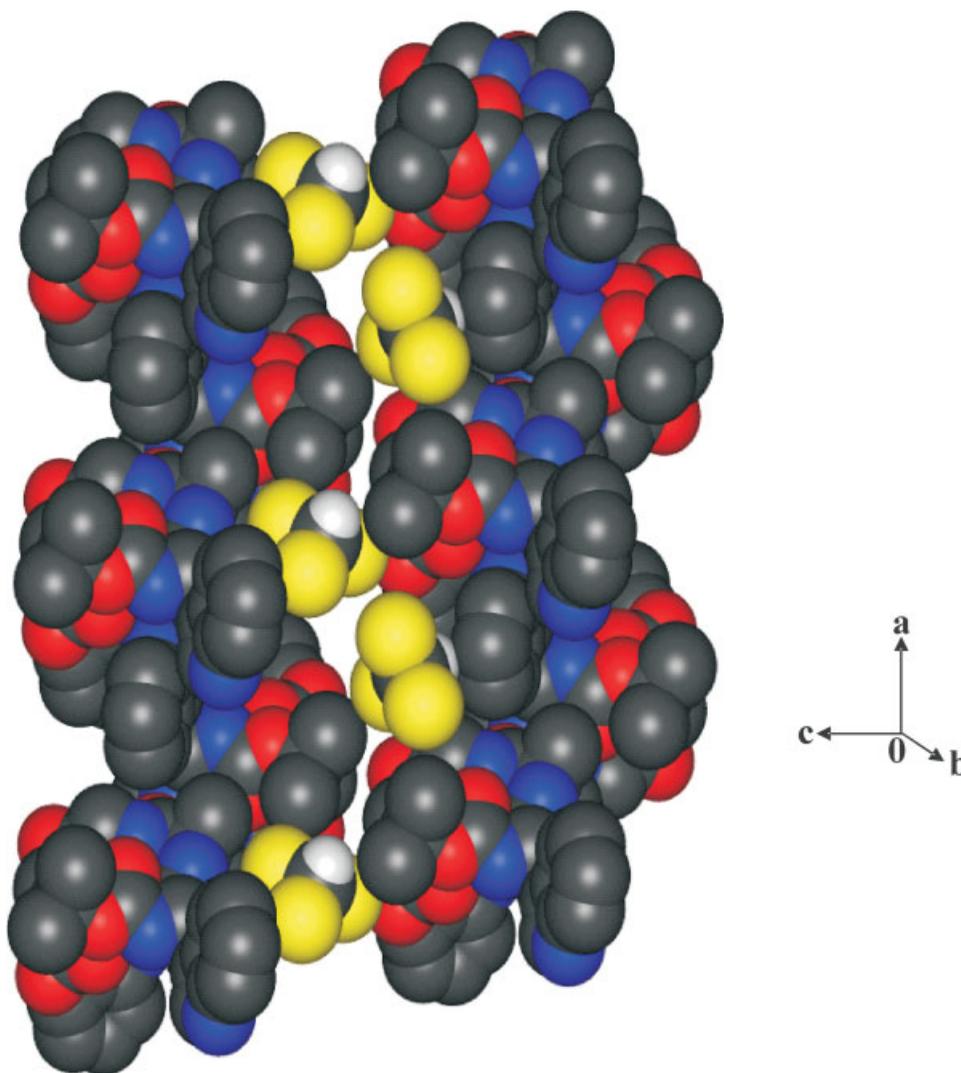


**FIGURE 10**  Space filling representation of molecular packing illustrated of peptide **5**. The chloroform molecules in two distinct sites are occupied the cavities between two adjacent columns.
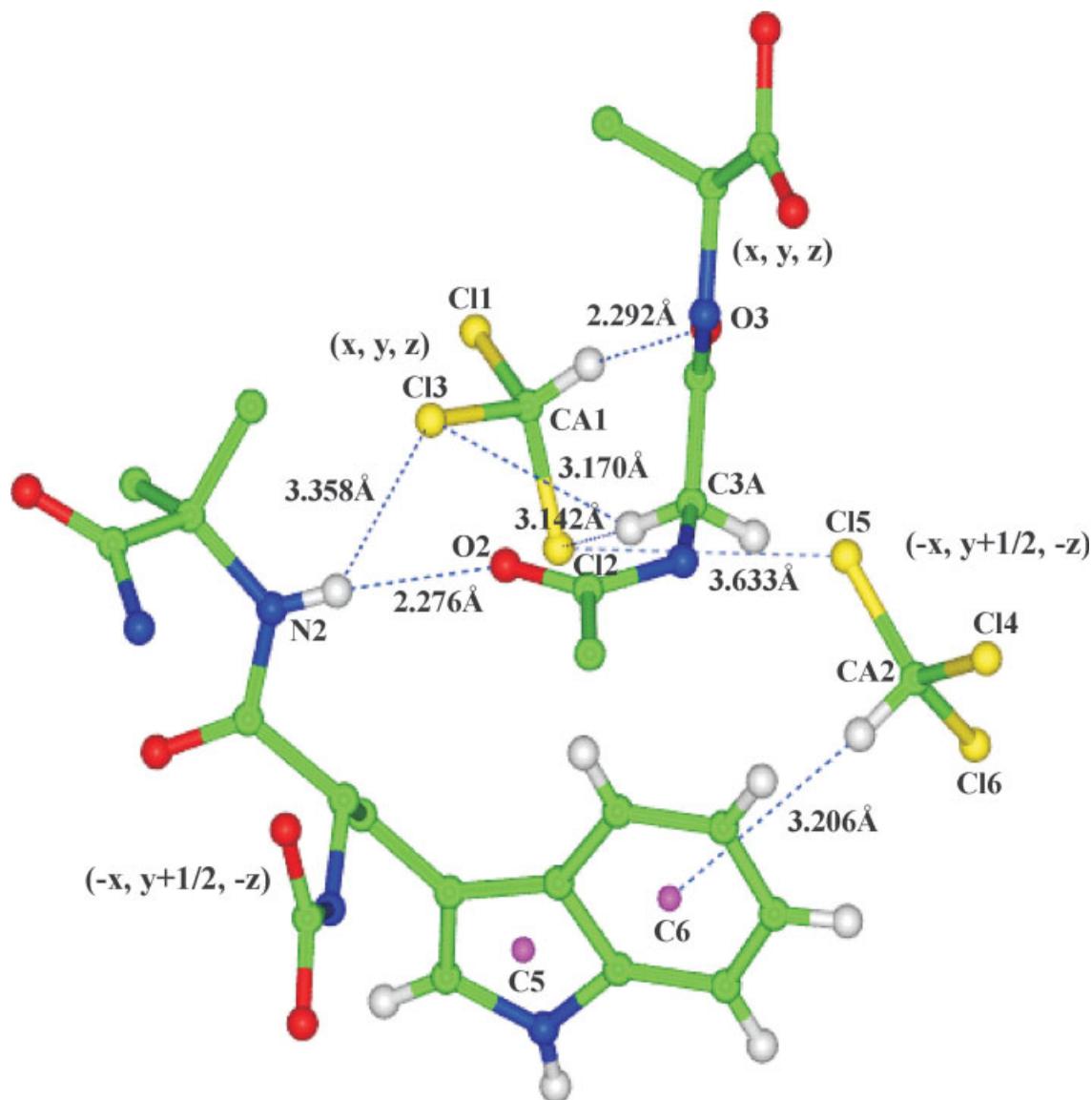
**FIGURE 11** Environment of chloroform molecule in peptide **5**. The intermolecular hydrogen bonds are shown in dotted lines.

when compared with the other two residues. Examination of the structures of both conformers indicates that a switch from 'A' to 'C' can be easily achieved by a change in the torsion angle $\phi$ of residue Trp1 from $-54.5°$ to the extended form, which is illustrated in Figure 8.

Notably, peptide **5** shows structural variation in comparison with its parent peptide **2**. In the case of peptide **2**, only the helical conformation was observable in both chloroform and methanol, which led us to the conclusion that presence of Trp residues flanking turn regions caused disruption of the hairpin-nucleating turn segment and instead gave rise to a helical conformation in solution, wherein the aromatic rings are positioned away from the backbone. However, an

NOE between the amides of Trp1 and Trp4 unambiguously established the presence of a type I′ hairpin nucleating element in the case of peptide **5**. This can be explained when one considers the contributions of hydrogen bonds to secondary structure stability. It should be noted that in a four residue sequence, both $3_{10}$ helical and $\beta$-hairpin structures have two hydrogen bonds, whereas in an 8-residue peptide, the helix will have up to six hydrogen bonds ($3_{10}$ structure), while the hairpin will have only four. It can therefore be argued that destabilization of the turn region because of the indole rings along with the stabilizing forces of up to six hydrogen bonds leads to the formation of a helical conformation in peptide **2**. When the contributions of hydrogen bonds

**Table V    Parameters for Potential Weak Interactions Observed in Crystals of Peptide 5**

| NH...π Interaction[53] | N | X | N...X (Å) | H...X (Å) | N—H...X (degree) | γ (degree) |
|---|---|---|---|---|---|---|
| NH...π Interaction (side chain) | N4ε1 | C6[a] | 3.554 | 2.777 | 151.2 | 82.5 |
| NH...π Interaction (main chain) | N1 | C4ζ2[b] | 3.832 | 3.135 | 139.8 | 46.5 |
| CH...π Interaction[51] | C | X | C...X (Å) | H...X (Å) | C—H...X (degree) | |
| | CA2 | C6 | 4.183 | 3.206 | 174.7 | |
| π...π Interaction[54] | X | X | $R_{5cen}$ (Å) | $R_{6cen}$ (Å) | $R_{clo}$ (Å) | γ (degree) |
| | C6 | C6[b] | | 5.312 | 3.56 | 82.5 |
| | C5 | C5[b] | 5.094 | | | |
| | C6 | C6[c] | | 6.314 | 3.556 | 54.5 |
| | C5 | C5[c] | 4.557 | | | |
| NH...Cl Interaction[49,50,55] | N | Cl | N...Cl (Å) | H...Cl (Å) | N—H...Cl (degree) | |
| | N2[d] | Cl3 | 3.746 | 3.358 | 110.5 | |
| CH...Cl Interaction[56,57] | C | Cl | C...Cl (Å) | H...Cl (Å) | C—H...Cl (degree) | |
| | C3A | Cl2 | 3.934 | 3.142 | 139.9 | |
| | C3A | Cl3 | 3.966 | 3.170 | 140.4 | |
| CH...O Interaction[46–48] | C | O | C...O (Å) | H...O (Å) | C—H...O (degree) | C=O...H (degree) |
| | CA1 | O3 | 3.182 | 2.292 | 150.5 | 121.3 |
| Cl...Cl Interaction[e58–60] | Cl | Cl | Cl...Cl (Å) | C—Cl...Cl (degree) | | |
| | Cl2 | Cl5[f] | 3.633 | CA1 − Cl2...Cl5 = 104.6 | | |
| | | | | CA2 − Cl5...Cl2 = 122.5 | | |

[a] Symmetry related by ($x, y + 1, z$).

[b] Symmetry related by ($x, y − 1, z$).

[c] Symmetry related by ($−x + 1, y − 1/2, −z$).

[d] Symmetry related by ($−x, y + 1/2, −z$).

[e] CA1 and CA2 represent the carbon atoms of chloroform solvent in two sites I and II, respectively. C6 and C5 are the centroids of 6- and 5-membered rings of Trp residue respectively. γ is the interplanar angle.

[f] Symmetry related by ($−x, y + 1/2, −z$).

to secondary structure stability are nullified, as in peptide **5**, both helix-nucleating and hairpin-nucleating turns are detected in solution. Nonetheless, the presence of both conformers in solution for peptide **5** suggests that Trp-mediated turn destabilization probably does occur even in the tetrapeptide.

### *Conformation of Peptide 5, Boc-W-U-G-W-OMe, in Crystals.*

In the preceding section three stereochemically acceptable, two hydrogen bonded conformations have been considered of peptide **5** in solution (Table II). Peptide crystal structures usually provide a detailed view of one of the energetically accessible conformational state of flexible peptides. Examples of crystal structures accommodating distinctly different conformational states are rare.[41,42] Efforts were made to obtain diffraction quality single crystals of peptide **5**. Crystals of moderate quality were obtained from chloroform solution, while efforts from other solvents were unsuccessful. Figure 9 shows the view of molecular conformation in crystals, while Tables III and IV list the backbone and side chain torsion angles and hydrogen bond parameters, respectively. The molecule adopts a folded, compact conformation, stabilized by two intramolecular 4→1 hydrogen bonds: Boc0 C=O...HN Gly3 and Trp1 C=O...HN Trp4.

Of these, the N4...O1 interaction is significantly weaker with somewhat large N...O distance (3.402 Å). O1 is involved in a strong intermolecular hydrogen bond with the Trp1 indole NH group of a symmetry related molecule. Inspection of the Ramachandran angles reveals that the molecule adopts a consecutive type II–I′ β-turn conformation (conformer 'A' described in Table II). Similar consecutive β-turn conformations for short peptides have previously been observed in Aib-containing sequences.[36,44,45]

### *Molecular Packing in Crystals.*

Two intermolecular hydrogen bonds involving the indole NH of Trp1 and Aib2 NH hold the molecules together in crystals. Peptide molecules pack in a monoclinic space group (P2₁) into columns with large cavities formed between adjacent columns. The crystal contains a stoichiometric amount of chloroform. Two distinct sites, I and II, are obtained for chloroform each with an occupancy factor of 0.5. In the view shown in Figure 10 the solvent molecules are distributed such that for each peptide molecule only one of the two possible sites is occupied. The occupancies of site I and II alternate in the interstitial sites between the two columns. Figure 11 shows a view of the environment of two chloroform molecules. At site I, the sol-
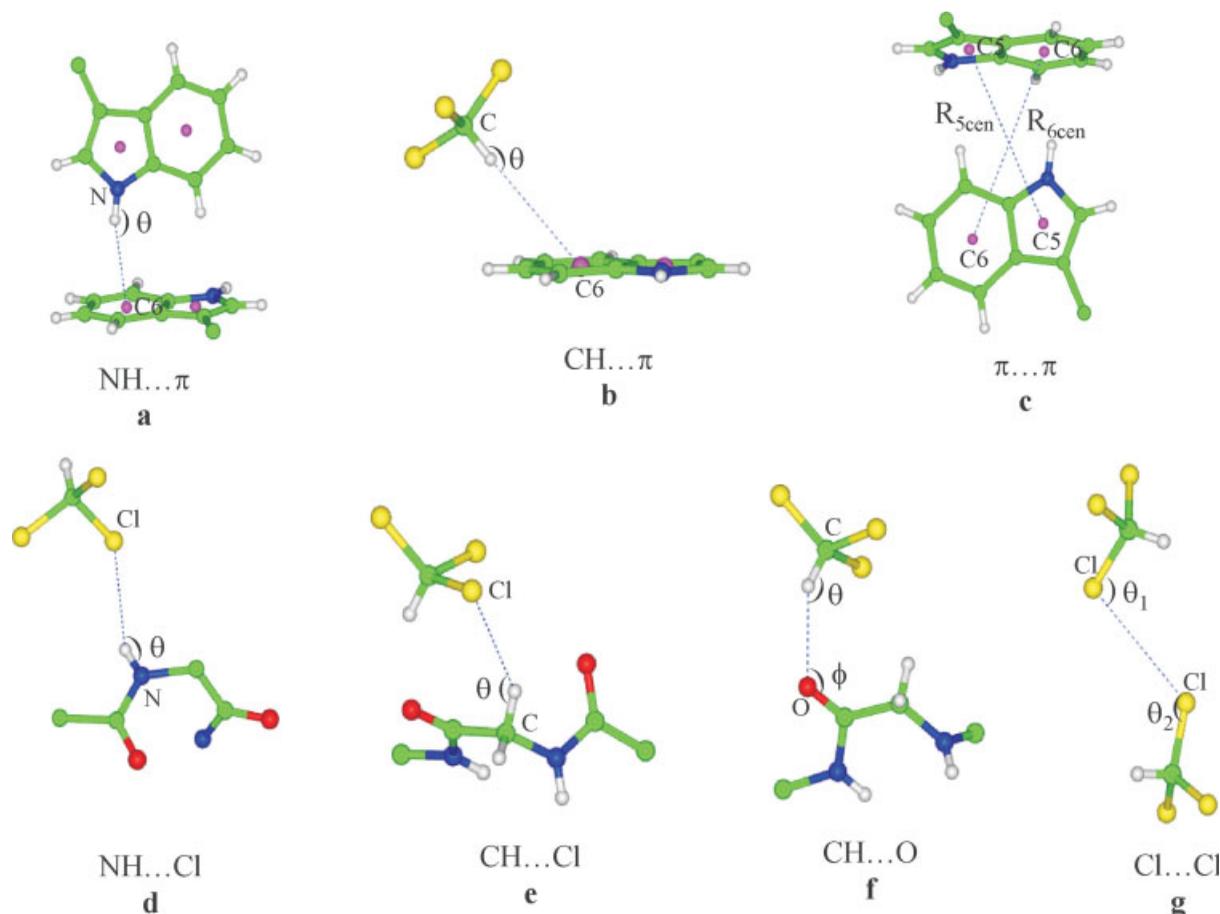
**FIGURE 12** Schematic views of potential weak interactions in crystals. The parameters are indicated: (a) N$\ldots$C6 $\leq$ 3.8 Å, N—H$\ldots$C6 ($\theta$) $\geq$ 120°, $\gamma$ > 30°; (b) C$\ldots$C6 $\leq$ 4.5 Å, C—H$\ldots$C6 ($\theta$) $\geq$ 120°; (c) 4.5 Å $\geq$ $R_{cen}$ $\leq$ 7.0 Å, 0° $\leq$ $\gamma$ $\leq$ 90°; (d) N$\ldots$Cl $\leq$ 3.7 Å, N—H$\ldots$Cl ($\theta$) $\geq$ 100°; (e) C$\ldots$Cl $\leq$ 3.8 Å, C—H$\ldots$Cl ($\theta$) $\geq$ 120°; (f) 3.0 Å $\leq$ C$\ldots$O $\leq$ 3.8 Å, 110° $\leq$ C—H$\ldots$O ($\theta$) $\leq$ 180°; (g) Cl$\ldots$Cl $\leq$ 3.6 Å, $\theta_1 = \theta_2 \approx$ 90°.

vent is held firmly by the CH$\ldots$O hydrogen bond[46–48] between the chloroform CH and the Gly3 C=O group. The second interaction observed is between the Gly2 NH group and one of the Cl atoms of chloroform. The N$\ldots$Cl distance of 3.746 Å is consistent with values reported in crystal structures of peptide–chloroform solvates.[49,50] At site II, the chloroform CH is positioned so as to participate in a weak CH$\ldots$π interaction[51] with the six membered aromatic ring of Trp moiety. The shortest Cl$\ldots$Cl distance (3.633 Å) observed between the chloroform molecule at sites I and II is between Cl2 and Cl5. This is slightly more than the expected van der Waals contact distance between two nonbonded Cl atoms (3.5 Å). These observations suggest that simultaneous occupation of both sites I and II is sterically feasible. The relatively poor packing of peptide molecules in crystals of **5** and the consequent incorporation of solvent molecules at multiple sites may be responsible for the relatively poor quality of crystal, resulting in a rather high *R*-factor of 10.95%. It must

be noted that the conformational conclusions and an analysis of intra and intermolecular interactions are largely unaffected by the resolution of the structure determination ($\sim$1.03 Å).

*Weak Interactions in the Crystal Structure of Peptide 5.* Designed peptides with Trp residues provide an opportunity to examine the nature of weak interactions involving the indole side chains.[18,52] Table V lists the parameters for all the potential weak intermolecular interactions observed in peptide **5**. Considerable recent attention has been focused on weak intramolecular interactions observed in protein structures[10,35,51,61–63] and in the packing of organic molecules in crystals.[64–66] The structure of peptide **5** provides examples of potential NH$\ldots$π, CH$\ldots$π, π$\ldots$π, NH$\ldots$Cl, CH$\ldots$Cl, and CH$\ldots$O interactions. Schematic views of the weak interactions in crystals are summarized in Figure 12. In peptide **5**, the closest distance of approach in Cl atoms at the two proximal sites is somewhat longer than that observed in a recent
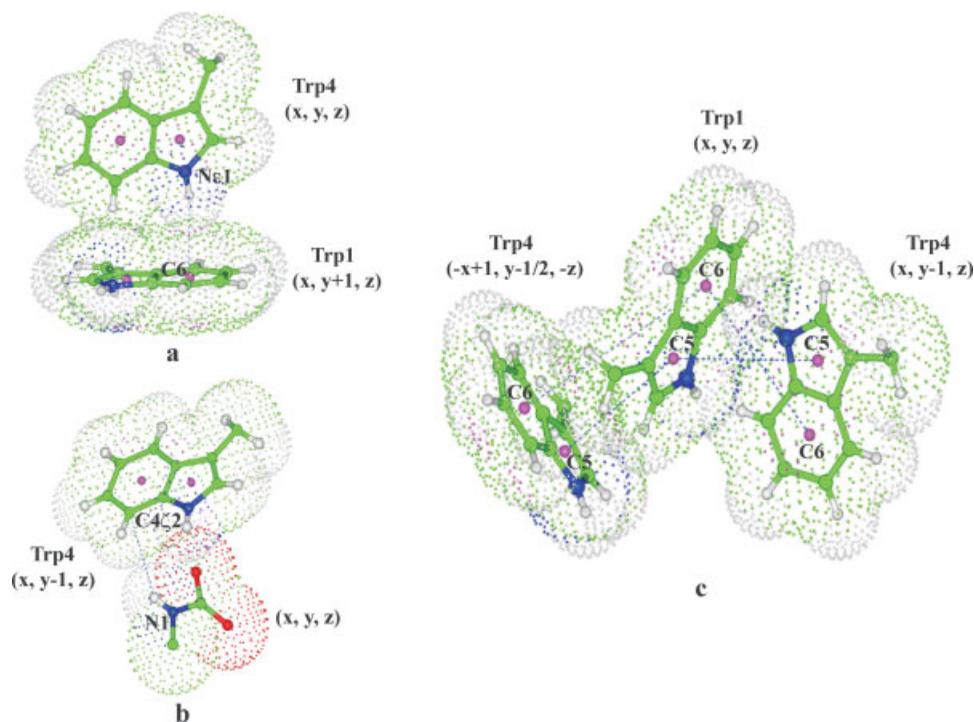
**FIGURE 13** Close packing of proximal Trp rings illustrating weak interactions observed in the crystals of peptide **5**. (a) Indole NH...$\pi$ interaction; (b) backbone NH...$\pi$ interaction; (c) aromatic $\pi$...$\pi$ interaction between the Trp residues. The van der Waals surfaces are shown in all three cases. The parameters defining the above weak interactions are listed in Table V.

analysis of halogen...halogen interactions.[58] Figure 13 shows the observed NH...$\pi$ and $\pi$...$\pi$ interactions in peptide **5**.

### Trp-Xxx-Trp Sequences in Proteins

Trp-rich sequences are not abundant in nature, but play a significant role when present in proteins.[67] The presence of a large number of aromatic residues is a major deciding factor of both local conformation and stability in proteins. In the case of peptides, packing of aromatic residues greatly influences the secondary structure of the peptide, in both solution and the crystal state. Analyses of the Trp-rich peptides **1–4** have revealed that such sequences have a tendency to preferentially adopt helical conformations in nonhydrogen bonding solvents and disfavor an extended sheet-like structure. To investigate the preferred local conformations of Trp-rich segments in proteins, a database analysis of the Trp-Leu-Trp segment was carried out on a collection of high-resolution protein structures. Out of a total database of 2346 structures, only 5 proteins were found to contain the Trp-Leu-Trp sequence. Analysis of the local conformations and preferred orientation of the indole ring in these proteins revealed no preference for specific secondary structures. A more general analysis using Trp-Xxx-Trp sequences was therefore carried out. The results obtained (Table VI) clearly indicated that

Trp-Xxx-Trp sequences were equally abundant in helical and extended structures in proteins, although a marginal preference for extended strand segments was discernable.

The results obtained from the database analysis do not agree with the observed conformational preferences in the case of Trp-rich peptides discussed in this study. This can be explained when one takes into consideration the different interactions that Trp residues in peptides and proteins involve in, when placed in defined secondary structure scaffolds. Preferred torsion angles for aromatic side chains have been observed in proteins, especially for Trp residues, in helices, and strand segments.[68] In the case of strands, the preferred $\chi 1$ values of *trans* ($180°$) or *gauche*− ($−60°$) will orient the indole rings of Trp residues, positioned at hydrogen bonding sites, away from each other and towards the peptide plane, leading to the formation of aromatic–amide interactions. In the case of proteins, the availability of several other stabilizing interactions overrides the influence of aromatic residues on backbone conformation. However, in the absence of other interactions in isolated peptides, the indole ring tends to maximize its interactions with the backbone, thereby influencing local torsion angles. In the case of peptide **1**, such interactions lead to the formation of a type I' turn at the $^D$Pro–Gly segment and strand fraying beyond the turn

**Table VI    Observed Secondary Structures[a] for WXW Sequences Retrieved from a Protein Database[b] Search**

| Sequence | Total | Helix | Sheet | Other |
|---|---|---|---|---|
| WAW | 12 | 6 | 4 | 2 |
| WCW | 1 | – | 1 | – |
| WDW | 10 | 3 | 4 | 3 |
| WEW | 6 | 1 | 4 | 1 |
| WFW | 4 | 1 | 2 | 1 |
| WGW | 10 | 3 | 1 | 6 |
| WHW | 8 | 3 | 4 | 1 |
| WIW | 6 | 2 | 1 | 3 |
| WKW | 5 | 1 | 2 | 2 |
| WLW | 5 | 2 | 3 | – |
| WMW | 1 | – | 1 | – |
| WNW | 5 | – | 3 | 2 |
| WPW | 4 | 1 | – | 3 |
| WQW | 7 | 3 | 2 | 2 |
| WRW | 10 | 4 | 2 | 4 |
| WSW | 13 | 4 | 5 | 4 |
| WTW | 13 | 2 | 5 | 6 |
| WVW | 5 | 1 | 4 | – |
| WWW | 1 | – | – | 1 |
| WYW | 2 | 2 | – | – |

[a] Total WXW sequences retrieved: 128; Total WXW sequences in helical conformation: 39 (30.4%); Total WXW sequences in extended conformation: 48 (37.5%); Total WXW sequences in other (nonhelical, nonextended) conformations: 41 (32.0%).

[b] Cullpdb list dated July 30, 2006; Resolution cutoff: 2 Å; % identity cutoff: 20%; *R*-factor: 0.25; total number of chains used: 2346.

region. In the case of Trp residues found in helices, the preferred $\chi 1$ of *gauche+* ($+60°$) orients the aromatic side chain away from the backbone; the indole is no longer interacting with the peptide plane, thereby resulting in the formation of stable helical scaffolds.

## CONCLUSIONS

Aromatic interactions have been widely implicated in protein folding and stability. Presence of a large number of Trp residues in a given sequence can be anticipated to lend stability to the peptide by close packing of the indole rings, thereby forming a hydrophobic cluster. Reports on TrpZip peptides and related sequences support the involvement of tryptophan residues in hairpin stabilization by the formation of the strong T-shaped aromatic interactions in the polar solvents.[6] Structural characterization of peptide **1**, however, reveals that the peptide does not adopt the anticipated $\beta$-hairpin conformation with two pairs of aromatic interactions. Instead, the data clearly suggests that interactions between the indole and the pyrrolidine rings of Trp and Pro residues, respectively, not only lead to the formation of a type I′ turn but also results in strand

fraying. Short-range interactions, for example, between the indole ring and amide plane, seem to predominate over long-range aromatic interactions, when Trp residues are positioned at the hydrogen bonding sites of short peptide hairpins.

Results obtained from structural characterization of peptides **2–4** suggest that multiple Trp residues can be accommodated on a helical peptide backbone, as this structure offers greater conformational freedom to the bulky indole side chain. The conformation adopted by peptide **2** in both nonhydrogen bonding and hydrogen bonding solvents supports this argument. Designed peptides containing a large number of Trp residues in peptides **3** and **4**, with helix-nucleating Aib residues also give rise to the formation of stable helical scaffolds in solution. It is worth mentioning that the presence of a single centrally positioned Aib residue is sufficient to nucleate a helical conformation in Trp-rich sequences, as demonstrated in peptide **3**.

Peptide **5** is unique, as NMR studies suggest the presence of a folded conformation even for a tetrapeptide. In chloroform, evidence for conformational interconversion between the type II–I′ to a type I′/III′ turn, accompanied by rearrangement of the hydrogen bonds, is obtained, with both conformers being almost equally populated in solution. NOEs observed between the indole rings of Trp1 and Trp4 to the backbone is suggestive of strong aromatic-backbone interactions stabilizing both conformers. Structure of the peptide in the crystal state, derived from X-ray crystallographic studies, indicates a consecutive type II–I′ $\beta$-turn structure with the two indole rings involved in strong intermolecular T-shaped aromatic interactions. A large number of weak interactions aid in packing of the peptide in the crystal.

The results of a database analysis of Trp-Xxx-Trp sequences in proteins, however, reveal that the specific preferences for secondary structure are not seen. Presence of several other stabilizing interactions may explain the absence of an overwhelming choice for a particular secondary structure for Trp-rich sequences in proteins.

## REFERENCES

1. Schiffer, M.; Chang, C.-H.; Stevens, F. J. Protein Eng 1992, 5, 213–214.
2. Yau, W.-M.; Wimley, W. C.; Gawrisch, K.; White, S. H. Biochemistry 1998, 37, 14713–14718.

3. Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S. B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L. J.; Dobson, C. M.; Schwalbe, H. Science 2002, 295, 1719–1722.

4. Doyle, D.; Wallace, B. J Mol Biol 1997, 266, 963–977.

5. Strøm, M. B.; Rekdal, O.; Svendsen, J. S. J Pept Sci 2002, 8, 431–437.

6. Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. Proc Natl Acad Sci USA 2001, 98, 5578–5583.

7. Mahalakshmi, R.; Raghothama, S.; Balaram, P. J Am Chem Soc 2006, 128, 1125–1138.

8. Rai, R.; Raghothama, S.; Balaram, P. J Am Chem Soc 2006, 128, 2675–2681.

9. Gallivan, J.; Dougherty, D. Proc Natl Acad Sci USA 1999, 96, 9459–9464.

10. Samanta, U.; Pal, D.; Chakrabarti, P. Proteins: Struct Funct Genet 2000, 38, 288–300.

11. Pejov, L. Chem Phys Lett 2001, 339, 269–278.

12. Meyer, E. A.; Castellano, R. K.; Diederich, F. Angew Chem Int Ed Eng 2003, 42, 1210–1250.

13. Burley, S. K.; Petsko, G. A. Adv Protein Chem 1988, 39, 125–189.

14. Singh, J.; Thornton, J. Atlas of Protein-side Chain Interactions; IRL Press: Oxford, 1992.

15. Duan, G.; Smith, V. H. J.; Weaver, D. F. Int J Quantum Chem 2000, 80, 44–60.

16. Duan, G.; Smith, V. H. J.; Weaver, D. F. J Phys Chem A 2000, 104, 4521–4532.

17. Gervasio, F. L.; Chelli, R.; Procacci, P.; Schettino, V. Proteins: Struct Funct Genet 2002, 48, 117–125.

18. Sengupta, A.; Mahalakshmi, R.; Shamala, N.; Balaram, P. J Pept Res 2005, 65, 113–129.

19. Awasthi, S. K.; Raghothama, S.; Balaram, P. J Chem Soc Perkin Trans 2 1996, 2701–2706.

20. Braunschweiler, L. E.; Ernst, R. R. J Magn Reson 1983, 53, 521–528.

21. Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. J Am Chem Soc 1984, 106, 811–812.

22. Bax, A.; Davis, D. G. J Magn Reson 1985, 63, 207–213.

23. Pitner, T. P.; Urry, D. W. J Am Chem Soc 1972, 94, 1399–1400.

24. Iqbal, M.; Balaram, P. J Am Chem Soc 1981, 103, 5548–5552.

25. Raj, P. A.; Balaram, P. Biopolymers 1985, 24, 1131–1146.

26. Schneider, T. R.; Sheldrick, G. M. Acta Crystallogr D 2002, 58, 1772–1779.

27. Sheldrick, G. M. SHELXL-97: A program for crystal structure refinement, University of Göttingen, Göttingen, Germany, 1997.

28. Wang, G.; Dunbrack, R. L. Bioinformatics 2003, 19, 1589–1591.

29. Venkatraman, J.; Shankaramma, S. C.; Balaram, P. Chem Rev 2001, 101, 3131–3152.

30. Awasthi, S. K.; Shankaramma, S. C.; Raghothama, S.; Balaram, P. Biopolymers 2001, 58, 465–476.

31. Aravinda, S.; Harini, V. V.; Shamala, N.; Das, C.; Balaram, P. Biochemistry 2004, 43, 1832–1846.

32. Awasthi, S. K.; Raghothama, S.; Balaram, P. Biochem Biophys Res Commun 1995, 216, 375–381.

33. Mahalakshmi, R.; Balaram, P. Methods Mol Biol 2006, 340, 71–94.

34. Butterfield, S. M.; Waters, M. L. J Am Chem Soc 2003, 125, 9580–9581.

35. Bhattacharyya, R.; Samanta, U.; Chakrabarti, P. Protein Eng 2002, 15, 91–100.

36. Prasad, B. V. V.; Balaram, P. CRC Crit Rev Biochem 1984, 16, 307–384.

37. Toniolo, C.; Crisma, M.; Formaggio, F.; Valle, G.; Cavicchioni, G.; Precigoux, G.; Aubry, A.; Kamphuis, J. Biopolymers 1993, 33, 1061–1072.

38. Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Wright, P. E.; Dyson, H. J. J Biomol NMR 2000, 18, 43–48.

39. Mahalakshmi, R.; Shanmugam, G.; Polavarapu, P. L.; Balaram, P. ChemBioChem 2005, 6, 2152–2157.

40. Creighton, T. E. Protein: Structures and Molecular Properties; Freeman: New York, 1993.

41. Prasad, S.; Mitra, S.; Subramanian, E.; Velmurugan, D.; Balaji Rao, R.; Balaram, P. Biochem Biophys Res Commun 1994, 198, 424–430.

42. Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, H.; Balaram, P. Proc Natl Acad Sci USA 1989, 86, 765–769.

43. IUPAC-IUB Commission on Biochemical Nomenclature.Biochemistry 1970, 9, 3471–3479.

44. Prasad, B. V. V.; Balaram, P. Conformation in Biology; Adenine: Schenectady, NY, 1982; pp 133–140.

45. Crisma, M.; Valle, G.; Toniolo, C.; Prasad, S.; Rao, R. B.; Balaram, P. Biopolymers 1995, 35, 1–9.

46. Desiraju, G. R. Acc Chem Res 1991, 24, 290–296.

47. Desiraju, G. R. Acc Chem Res 1996, 29, 441–449.

48. Steiner, T. Chem Commun 1997, 727–734.

49. Fisk, J. D.; Powell, D. R.; Gellman, S. H. J Am Chem Soc 2000, 122, 5443–5447.

50. Schmitt, M. A.; Choi, S. H.; Guzei, I. A.; Gellman, S. H. J Am Chem Soc 2005, 127, 13130–13131.

51. Brandl, M.; Weiss, M. S.; Jabs, A.; Sühnel, J.; Hilgenfeld, R. J Mol Biol 2001, 307, 357–377.

52. Mahalakshmi, R.; Sengupta, A.; Raghothama, S.; Shamala, N.; Balaram, P. J Pept Res 2005, 66, 277–296.

53. Mitchell, J. B. O.; Nandi, C. L.; McDonald, I. K.; Thornton, J. M. J Mol Biol 1994, 239, 315–331.

54. Burley, S. K.; Petsko, G. A. Science 1985, 229, 23–28.

55. Aullón, G.; Bellamy, D.; Brammer, L.; Bruton, E. A.; Orpen, A. G. Chem Commun 1998, 653–654.

56. Wheeler, G. L.; Colson, S. D. Acta Crystallogr B 1975, 31, 911–913.

57. Taylor, R.; Kennard, O. J Am Chem Soc 1982, 104, 5063–5070.

58. Reddy, C. M.; Kirchner, M. T.; Gundakaram, R. C.; Padmanabhan, K. A.; Desiraju, G. R. Chem—Eur J 2006, 12, 2222–2234.

59. Desiraju, G. R.; Parthasarathy, R. J Am Chem Soc 1989, 111, 8726–8727.

60. Price, S. L.; Stone, A. J.; Lucas, J.; Rowland, R. S.; Thornley, A. E. J Am Chem Soc 1994, 116, 4910–4918.

61. Babu, M. M.; Singh, S. K.; Balaram, P. J Mol Biol 2002, 322, 871–880.

62. Fabiola, G. F.; Krishnaswamy, S.; Nagarajan, V.; Pattabhi, V. Acta Crystallogr D 1997, 53, 316–320.

63. Tóth, G.; Watts, C. R.; Murphy, R. E.; Lovas, S. Proteins: Struct Funct Genet 2001, 43, 373–381.

64. Steiner, T. Angew Chem Int Ed Eng 2002, 41, 48–76.

65. Dunitz, J. D.; Gavezzotti, A. Angew Chem Int Ed Eng 2005, 44, 1766–1787.

66. Desiraju, G. R.; Steiner, T. The Weak Hydrogen Bond in Structural Chemistry and Biology; Oxford University Press: Oxford, UK, 1999.

67. Meyer, E. F.; Tollett, W. J. J. Acta Crystallogr D 2001, 57, 181–186.

68. Chakrabarti, P.; Pal, D. Protein Eng 1998, 11, 631–647.