Structural Characterization of Backbone-Expanded Helices in Hybrid Peptides: \((\alpha\gamma)\) and \((\alpha\beta)\) Sequences with Unconstrained \(\beta\) and \(\gamma\) Homologues of \(L\)-Val*•

Krishnayan Basuroy, Bhimareddy Dinesh, Narayanaswamy Shamala, * and Padmanabhan Balaram*

Helices and hairpins, ubiquitous elements of secondary structures in proteins, have been targets for synthetic mimetic ("foldamer") design. [1] Foldamer research has intensified following the discovery that polypeptides with homologated (i.e. a backbone with one or more additional –CH2– groups in the chain), non-natural backbones can readily adopt diverse helical folds. [2] The Pauling \(\alpha\) helix is an abundantly observed secondary structure in proteins and is characterized by repetitive 5-1, intramolecular, CO(i)···HN(i+4) 13 atom (C13) hydrogen-bonded rings. [3] The 3\(\eta\) helix is a more tightly wound structure, with repetitive CO(i)···HN(i+3), 10 atom (C10) hydrogen-bonded rings. [4] While the 3\(\eta\) helix is less abundant than the \(\alpha\)-helix in proteins, it has been widely characterized in synthetic and natural peptides containing mostly \(\alpha\)-\(\alpha\) dialkylated residues, primarily \(\alpha\)-aminoisobutyric acid (Aib). [5] In polypeptide sequences composed exclusively of \(\alpha\)-amino acids (\(\alpha\alpha\alpha\)), the Pauling \(\alpha\) helix incorporates three residues in the hydrogen-bonded turn, while the 3\(\eta\) helix is composed of two residues in the hydrogen-bonded turn. In both cases, only a single set of backbone torsion angles (\(\varphi, \psi\)) characterize an ideal helical structure (\(\alpha\)-helix, \(\varphi \approx -57.0^{\circ}, \psi \approx -47.0^{\circ}\); 3\(\eta\) helix, \(\varphi \approx -60.0^{\circ}, \psi \approx -30.0^{\circ}\)). [6] Backbone-homologated amino acids, specifically \(\beta\) and \(\gamma\) residues may be incorporated into \(\alpha\) amino acid sequences for the generation of helical structures with hybrid backbones. [2-7]

For example a regular \((\alpha\beta)\), sequence can, in principle, form a regular C11 helix, with the \(\alpha\beta\) segment being the repeating unit resulting in the backbone-expanded analogue of the 3\(\eta\) helix. [8] The three residue (\(\alpha\alpha\beta\alpha\beta\)), hydrogen bond repeat, in an \((\alpha\beta)\)3 sequence, would result in a mixed C10/C11 helix, which is a backbone-expanded analogue of the \(\alpha\) helix. [8a, 8b, 9] Similarly in (\(\alpha\gamma\)), sequences, the analogue of the 3\(\eta\) helix would be the \(\alpha\gamma\) C12 helix. [9d, 10] While the \(\alpha\) helix mimic would be a mixed C10/C12 helix. [11] By analogy with studies on \(\alpha\) peptides, mixed helical structures with variations in the hydrogen-bonding pattern may also be anticipated. Indeed a growing volume of work on peptides with hybrid backbones suggests that a diversity of hydrogen-bonding patterns can be anticipated. [2, 7, 8a, 11] Thus far, definitive structural characterization of helical structures in hybrid peptides with repeating (\(\alpha\beta\)), or (\(\alpha\gamma\)), sequences by X-ray diffraction has been achieved only in the case of stereochemically constrained \(\beta\) and \(\gamma\) residues. [8c, 10] Constraining backbone atoms by cyclization is a device effectively employed by the Selman group. [12] The use of gemically disubstituted \(\gamma\) residues, specifically gabapentin, also facilitates folded hydrogen-bonded conformations, permitting crystallographic characterization of hybrid helical structures. [2b, 13]

We describe herein the characterization of the \(\alpha\gamma\) C12 helix in oligopeptides containing the unconstrained \(\gamma\) residue, \(\gamma'(R)\)-Val. The structure determination of [Aib-\(\gamma'(R)\)-Val], oligomers ranging in length from four to sixteen residues establishes that C12 helices are readily formed. A structural comparison is presented of regular (Aib-Xxx), sequences (where, Xxx = \(\alpha\) (S)Val, \(\beta\) (R)Val, \(\gamma'(R)\)-Val) providing insights into the diversity of hydrogen-bonded structures in hybrid molecules (Boc-\(\beta\) (R)Val-OMe and Boc-\(\gamma'(R)\)-Val-OMe) are formed by homologation of Boc-\(\gamma\)-Val-OMe (Boc-(S)-Val-OMe). Note the change in absolute configuration. The abbreviations follow Seebach et al.) [12]

Figure 1 shows a view of the molecular conformation determined in crystals for the peptides with the sequence Boc-[Aib-\(\gamma'(R)\)-Val]-O-Me (n = 2, 4, 5, 8). The backbone torsion angles and hydrogen-bond parameters are provided as tables in the Supporting Information. In all cases, successive \(\alpha\gamma\) C12 turns are observed, with the number of C12 hydrogen bonds as follows: \(n = 2, 2; n = 4, 6; n = 5, 8; n = 8, 14\). Notably, the Aib-\(\gamma'(R)\)-Val segment adopted very similar conformations in the various peptides permitting determination of the parameters that describe the C12 helix. The averaged parameters for the C12 helix hydrogen bonds are: N···O (\(\Delta\)) = 2.94 ± 0.07, H···O (\(\Delta\)) = 2.10 ± 0.08, N-H···O (\(\psi\)) = 164.9 ± 7.5. The averaged torsion angles for the C12 helical turns are: Aib: \(\varphi\) (\(\psi\)) = -59.4 ± 3.7, \(\psi\) (\(\varphi\)) = -40.9 ± 4.1; \(\gamma'(R)\)-Val: \(\psi\) (\(\varphi\)) = -125.4 ± 4.9, \(\theta1\) (\(\varphi\)) = 52.6 ± 2.2, \(\theta1\) (\(\varphi\)) = 61.2 ± 2.4, \(\psi\) (\(\varphi\)) = -118.7 ± 6.0. Figure 2 shows a view of an \(\alpha\gamma\) C12 turn, which may be viewed as a backbone-expanded analogue of the type-III \(\beta\) turn in an (\(\alpha\alpha\)) segment. Repetition of the \(\alpha\gamma\) C12 turn leads to the C12 helix, while repeating the type-III \(\beta\) turn generates the polypeptide...
The configuration of the C12 helix in the αγ sequence described above prompted us to compare the analogous (αα)n and (αβ)n sequences. Figure 3 and Figure 4 show the molecular conformations determined in crystals for the tetrapeptides Boc-[Aib-α(S)Val]2-OMe, Boc-[Aib-β(R)Val]2-OMe, and Boc-[Aib-γ(R)Val]2-OMe and the octapeptides Boc-[Aib-α(S)Val]4-OMe, Boc-[Aib-β(R)Val]4-OMe, and Boc-[Aib-γ(R)Val]4-OMe (where α(S)Val = 1-Val).

The structures of the three tetrapeptides reveal two C11 hydrogen bonds corresponding to formation of one 310 helical turn in the case of Boc-[Aib-α(S)Val]2-OMe, as anticipated. The corresponding β and γ analogues reveal the formation of two consecutive C11 and C12 turns, respectively. Upon lengthening the oligopeptide sequences to octapeptides, clear differences in the intramolecular hydrogen-bonding patterns emerge. The (αα)n sequence in Boc-[Aib-α(S)Val]4-OMe shows six successive C10 hydrogen bonds corresponding to an almost ideal 310 helical turn in Boc-[Aib-γ(R)Val]4-OMe and the octapeptides Boc-[Aib-α(S)Val]4-OMe, Boc-[Aib-β(R)Val]4-OMe, and Boc-[Aib-γ(R)Val]4-OMe.
3α helical structure.[15] Backbone torsion angles and hydrogen-bond parameters are provided as tables in the Supporting Information. As already noted, the (αγ)₆ sequence yields a perfect C₁₂ helix. In sharp contrast, the (αβ)₆ sequence yields a mixed C₁₀/C₁₂ helix, with the formation of three C₁₀ (ααα) and two C₁₂ (ββ) hydrogen bonds. In this case, the repetitive units are three-residue hydrogen-bonded turns formed by CO(i)–HN(i+4) interactions, analogous to that observed in the classical α-helix.

Two factors undoubtedly contribute to the stability of oligopeptide helices. First, the number of intramolecular hydrogen bonds formed over the length of the peptide must favor the tighter helices formed by repetitive turns with fewer atoms in the hydrogen-bonded ring. For example, the 3α helix might be expected to be favored over the α helices in short peptides. Indeed a study of mixed α/β (1:1) sequences has revealed a pronounced chain-length dependence of helix type.[88] With increasing chain length, the differences in the hydrogen-bonding contribution diminish. Second, α helix stability must be strongly influenced by non-bonded interactions, with large diameter helices being disfavored by poor hydrogen-bonding contribution diminish. Second, a type.[8b] With increasing chain length, the differences in the revealed a pronounced chain-length dependence of helix peptides. Indeed a study of mixed α/β sequences. These studies with the unconstrained might be expected to be favored over the α peptides and proteins.[2g,16] In hybrid sequences containing already noted, the (αβ)₆ (in co-crystallized solvent) was treated with PART commands. In peptide 5, data collection was carried out in phi and omega scan-type mode using dry crystals at 296 K for peptides 1, 3, 4, 5, 6, 7, and 8. Crystals of peptide 2 were fragile and data was collected using a glass capillary with mother liquor (dichloromethane) at 296 K. For peptide 3, data collection was carried out at low temperature (240 K) to resolve the positional disorder (i.e. whether it is static or dynamic) of the γ(Val)₈ side chain, which turned out to be positional static disorder. All peptide structures were solved using iterative dual-space direct methods in SHELXD.[49] After the initial solution methods, all the structures were refined against F² isotropically followed by full-matrix anisotropic least-squares refinement using SHELXL-97.[49] The high quality of the diffraction data enabled location of many H atoms in these peptides directly from the difference Fourier map. All H atoms attached to backbone N atoms could be located in peptides 1, 2, 3, 4, and 5, and also with proper restraints and constraints to obtain chemically meaningful geometry of the disordered groups. Apart from tetrapeptide 5, the remaining hydrogen atoms, other than those which were attached to backbone N atoms could be located in tetrapeptide 7, five only free atoms attached to the atoms N₁, N₂, C₄A, C₂A, and C₄B were located from the difference Fourier map. Positional disorder in peptides 3, 4 (in γ(Val)₈ side chain), and 2 (in co-crystallized solvent) was treated with PART commands and also with proper restraints and constraints to obtain chemically meaningful geometry of the disordered groups. Apart from tetrapeptide 5, the remaining hydroxyl groups, other than those which were located from difference Fourier map, were fixed geometrically in the idealized position and allowed to ride on the C or N atoms to which they were bonded, in the final cycles of refinement. All H atoms attached to backbone N atoms could be located in peptides 1, 2, 3, 4, and 5. In peptide 6, all the hydrogen atoms were fixed geometrically in the idealized positions and allowed to ride on the C or N atoms to which they were bonded, in the final cycles of refinement. The details of the peptide sequence and relevant physical parameters.

### Table 1: Peptide sequences and relevant physical parameters.

<table>
<thead>
<tr>
<th>Peptides (No.)</th>
<th>Melting point [°C]</th>
<th>ESI-MS [Da]</th>
<th>Space group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boc-[U]-γ(Val)₄[Ome]₄ (1)</td>
<td>162–164</td>
<td>557.2, 579.2, 595.1</td>
<td>P₂₁₁₁</td>
</tr>
<tr>
<td>Boc-[U]-γ(Val)₄[Ome]₂ (2)</td>
<td>232–234</td>
<td>981.1, 1003.1, 1019.0</td>
<td>P₂₁₂₁</td>
</tr>
<tr>
<td>Boc-[U]-γ(Val)₄[Ome]₃ (3)</td>
<td>251–252</td>
<td>1193.0, 1215.0, 1230.9</td>
<td>P₂₁₂₁</td>
</tr>
<tr>
<td>Boc-[U]-γ(Val)₄[Ome]₄ (4)</td>
<td>195–197</td>
<td>592.9, 551.1, 567.1</td>
<td>P₂₁</td>
</tr>
<tr>
<td>Boc-[U]-γ(Val)₄[Ome]₅ (5)</td>
<td>242–243</td>
<td>925.0, 947.0, 962.9</td>
<td>P₂₂</td>
</tr>
<tr>
<td>Boc-[U]-γ(Val)₄[Ome]₆ (6)</td>
<td>151–152</td>
<td>501.2, 523.2, 539.1</td>
<td>P₂₂</td>
</tr>
<tr>
<td>Boc-[U]-γ(Val)₄[Ome]₇ (7)</td>
<td>240–241</td>
<td>861.0, 891.0, 907.0</td>
<td>P₂₁</td>
</tr>
<tr>
<td>Boc-[U]-γ(Val)₄[Ome]₈ (8)</td>
<td>232–234</td>
<td>981.1, 1003.1, 1019.0</td>
<td>P₂₁₂₁</td>
</tr>
</tbody>
</table>

[a] U = Aib. [b] V = Val. [c] Did not melt up to 300°C.

**Experimental Section**

Table 1 lists the sequences and relevant physical parameters determined for each of the peptides in this study. The amino acids, Boc-γ(Val)₄OH and Boc-β(Val)₄OH were synthesized by previously described procedures.[20,17] All the peptides were prepared by solution-phase synthesis using the tert-butyloxy carbonyl (Boc) group for N terminal protection. The C terminus was protected using a methyl ester. Deprotects were performed using 98% formic acid and 1.1 trifluoroacetic acid/dichloromethane (TFA/DCM) to remove the Boc group, whereas the methyl ester was removed by alkaline hydrolysis. Couplings were mediated by iso-thiazolocyclohexyl (1H)-benzotriazole HOBr; 1.01 equiv) for dipeptides and the successive peptides with 7-(7-azabenzotriazol-1-yl)-N,N,N',N-tetramethyluronium hexafluorophosphate (HATU) and HOBt. All intermediates were characterized by electrospray ionization mass spectrometry (ESI-MS), 500/700 MHz 1H NMR spectroscopy, and thin-layer chromatography (TLC) on silica gel (SiO₂, CHCl₃/MeOH 9:1) and monitored at 226 nm. The peptides were characterized by ESI-MS and 700 MHz 1H NMR spectroscopy.

X-ray diffraction datasets were collected using CuKα (1.54178 Å) radiation for peptides 1, 2, 3, 4, and 8 and MoKα (0.7073 Å) radiation for peptides 5, 6, and 8, using BRUKER AXS SMART APEX ULTRA CCD (rotating anode X-ray generator) and BRUKER AXS KAPPA APEXII CCD diffractometer respectively. Data collection was carried out in phi and omega scan-type mode using dry crystals at 296 K for peptides 1, 3, 4, 5, 6, 7, and 8. Crystals of peptide 2 were fragile and data was collected using a glass capillary with mother liquor (dichloromethane) at 296 K. For peptide 3, data collection was carried out at low temperature (240 K) to resolve the positional disorder (i.e. whether it is static or dynamic) of the γ(Val)₈ side chain, which turned out to be positional static disorder. All peptide structures were solved using iterative dual-space direct methods in SHELXD.[49] After the initial solution methods, all the structures were refined against F² isotropically followed by full-matrix anisotropic least-squares refinement using SHELXL-97.[49] The high quality of the diffraction data enabled location of many H atoms in these peptides directly from the difference Fourier map. All H atoms attached to backbone N atoms could be located in peptides 1, 2, 3, 4, and 6. In peptide 5, all the hydrogen atoms were fixed geometrically in the idealized positions and allowed to ride on the C or N atoms to which they were bonded, in the final cycles of refinement. In tetrapeptide 7, five only free atoms attached to the atoms N₁, N₂, C₄A, C₂A, and C₄B were located from the difference Fourier map. Positional disorder in peptides 3, 4 (in γ(Val)₈ side chain), and 2 (in co-crystallized solvent) was treated with PART commands and also with proper restraints and constraints to obtain chemically meaningful geometry of the disordered groups. Apart from tetrapeptide 5, the remaining hydroxyl groups, other than those which were located from difference Fourier map, were fixed geometrically in the idealized position and allowed to ride on the C or N atoms to which they were bonded, in the final cycles of refinement. The details of the...
crystal data and structure refinement for all the peptides mentioned above are provided as tables in the Supporting Information.

CCDC deposition numbers for the peptides are 881181 (1), 881182 (2), 881177 (3), 881178 (4), 881179 (5), 881180 (6), 881183 (7), and 882909 (8) and contain the supplementary crystallographic data for this paper. These have been deposited free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

Received: June 7, 2012
Published online: July 25, 2012

Keywords: backbone conformations · helices · hybrid peptides · hydrogen bonds · X-ray diffraction