Characterization of Helix Terminating Schellman Motifs in Peptides. Crystal Structure and Nuclear Overhauser Effect Analysis of a Synthetic Heptapeptide Helix

Saumen Datta,[†] N. Shamala,[†] Arindam Banerjee,[‡] Animesh Pramanik,[‡] Surajit Bhattacharjya,[‡] and P. Balaram*,[‡]

Contribution from the Department of Physics and Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560012, India

Received February 20, 1997. Revised Manuscript Received June 20, 1997[⊗]

Abstract: The Schellman motif is a widely observed, helix terminating structural motif in proteins, which is achieved by the adoption of a left-handed helical (α_L) conformation by a C-terminus residue. The resulting hydrogen bonding pattern involves an intramolecular $6 \rightarrow 1$ interaction. This helix terminating motif is readily mimicked in synthetic helical peptides by placing an achiral residue at the penultimate position of the helix. The crystal structure of the heptapeptide Boc-Leu-Aib-Val-Gly-Leu-Aib-Val-OMe (**1**) reveals a 3_{10}-helix terminated by a Schellman motif with Aib(6) adopting an α_L - conformation. The crystals are in the space group $P2_1$ with $a = 9.958(3)$Å, $b = 20.447$-(3)Å, $c = 11.723$ (2)Å, $\beta = 99.74(2)°$, and $Z = 2$. The final R_1 and wR_2 values are 7.2% and 18.9%, respectively, for 1731 observed reflections [$I \geq 2\sigma(I)$]. A $6 \rightarrow 1$ hydrogen bond between Aib(2)CO- - - -Val(7)NH and a $5 \rightarrow 2$ hydrogen bond between Val(3)CO- - - -Aib(6)NH are observed. An analysis of several α_L terminated helical peptides in crystals suggests that the medium range C_i^αH- - - -N_{i+3}H [$d_{\alpha N}(i,i+3)$] and C_i^αH- - - -N_{i+4}H [$d_{\alpha N}(i,i+4)$] interproton distances are indeed characteristic of the Schellman motif. NMR studies in CDCl$_3$ establish retention of the 3_{10}-helical conformation with key NOEs demonstrating the persistence of the conformation determined in crystals. The present study demonstrates the identification of the Schellman motif in solution in a synthetic helical peptide.

Introduction

Backbone hydrogen bonding patterns in proteins[1] and peptides[2] have been extensively analyzed. Hydrogen bonds of the type $5 \rightarrow 1$ (C_i=O- - -N_{i+4}H,C_{13}) and $4 \rightarrow 1$ (C_i=O- - -N_{i+3}H,C_{10}) are characteristic of α-helical[3] and 3_{10}-helical[4] structures, respectively. The $6 \rightarrow 1$ hydrogen bonds (C_i=O- - -N_{i+5}H, C_{16}) were originally postulated in the π-helical structures for polypeptides.[5] Such "π-type" ($6 \rightarrow 1$) hydrogen bonds are much less commonly observed in protein structures. Schellman[6] pointed out several years ago that α-helical structures in proteins are frequently terminated by a residue (most often the achiral amino acid Gly) adopting the α_L ($\phi = 50°$, $\psi = 60°$) conformation.[7] This reversal of helix sense at the C-terminus acts as a "helix stop signal" resulting in a structural feature involving concomitant formation of $6 \rightarrow 1$ (C_{16}) and $5 \rightarrow 2$ (C_{10}) hydrogen bonds, with residue "5" occurring in a left-handed helical conformation. This stereochemical motif has

also been termed as a "paper clip".[8] Several more recent analyses have focussed on the occurrence of helix terminating Schellman motifs in a growing data set of protein crystal structures.[9]

Recent crystal structure analyses of peptides containing the achiral residue α-aminoisobutyric acid (Aib or $C^{\alpha,\alpha}$-dimethylglycine) have revealed several examples of α_L-terminated helical conformations, when Aib is positioned at the penultimate position from the C-terminus.[10] Similar features have also been characterized when other achiral residues like α,β-dehydrophenylalanine (ΔPhe) have been used in place of Aib.[11] Thus far, unambiguous characterization of the Schellman motif in peptides *in solution* has not been possible. Parenthetically, it must also

* Address for correspondence: Prof. P. Balaram, Molecular Biophysics Unit, Indian Institute of Science, Bangalore - 560012, India. E-mail: pb@mbu.iisc.ernet.in. Fax: 91-80-3341683.

† Department of Physics.

‡ Molecular Biophysics Unit.

⊗ Abstract published in *Advance ACS Abstracts,* August 1, 1997.

(1) (a) Baker, E. N.; Hubbard, R. E. *Progr. Biophys. Mol. Biol.* **1984**, *44*, 97−179. (b) Stickle, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D. *J. Mol. Biol.* **1992**, *226*, 1143−1159.

(2) Toniolo, C. *CRC Crit. Rev. Biochem.* **1980**, *9*, 1−44.

(3) (a) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 235−240. (b) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601−619.

(4) (a) Taylor, H. S. *Proc. Am. Phil. Soc.* **1941**, *85*, 1−7. (b) Huggins, M. L. *Chem. Rev.* **1943**, *32*, 195−218. (c) Donohue, J. *Proc. Natl. Acad. Sci. U.S.A.* **1953**, *39*, 470−478. (d) Toniolo, C.; Benedetti, E. *Trends Biochem. Sci.* **1991**, *16*, 350 −353.

(5) (a) Low, B. W.; Baybutt, R. B. *J. Am. Chem. Soc.* **1952**, *74*, 5806−5807. (b) Low, B. W.; Grenville-Wells, H. J. *Proc. Natl. Acad. Sci. U.S.A.* **1953**, *39*, 785−801.

(6) Schellman, C. In *Protein Folding;* Jaenicke, R., Ed.; Elsevier/North Holland Biochemical Press: Amsterdam, 1980; pp 53−61.

(7) (a) The definitions of backbone dihedral angles follow the IUPAC-IUB Commission on Biochemical Nomenclature *Biochemistry* **1970**, *9*, 3471−3479. α_L refers to left-handed 3_{10}/α-helical conformations while α_R denotes a right-handed helical conformation. Since both 3_{10} and α-helical conformations fall in a limited region of ϕ, ψ space the terms α_L and α_R are used in a generic sense. (b) Abbreviations used: Aib, U, α-aminoisobutyric acid; Boc, *tert*-butyloxycarbonyl; DCC, dicyclohexylcarbodiimide; HOBt, 1-hydroxybenzotriazole; OMe, methyl ester; pBrBz, *p*-bromobenzoyl; Ac, acetyl; DMF, dimethylformamide; MPLC, medium pressure liquid chromatography; HPLC, high performance liquid chromatography; NOE, nuclear Overhauser effect.

(8) Milner-White, E. J. *J. Mol. Biol.* **1988**, *199*, 503−511.

(9) (a) Nagarajaram, H. A.; Sowdhamini, R.; Ramakrishnan, C.; Balaram, P. **1993** *FEBS Lett. 321*, 79−83. (b) Aurora, R.; Srinivasan, R.; Rose, G. D. *Science* **1994**, *264*, 1126−1130. (c) Rajashankar, K. R.; Ramakumar, S. *Protein Science* **1996**, *5*, 932−946. (d) For an attempt to design a water soluble, synthetic helix terminating in a Schellman motif, based on protein data analyses, see: Viguera, A. R.; Serrano, L. *J. Mol. Biol.* **1995**, *251*, 150−160.

(10) (a) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Int. J. Peptide Protein Res.* **1993**, *42*, 401−410. (b) Karle, I. L.; Banerjee, A.; Bhattacharjya, S.; Balaram, P. *Biopolymers* **1996**, *38*, 515−526. (c) Banerjee, A.; Datta, S.; Pramanik, A.; Shamala, N.; Balaram, P. *J. Am. Chem. Soc.* **1996**, *118*, 9477−9483. (d) Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Santini, A.; Bavoso, A.; Toniolo, C.; Crisma, M.; Sartore, L. *J. Chem. Soc., Perkin Trans. 2* **1990**, 1829−1837. (e) Di Blasio, B.; Pavone,V.; Saviano, R.; Fattorusso, C.; Pedone, E.; Benedetti, E.; Crisma, M.; Toniolo, C. *Peptide Res.* **1994**, *7*, 55−59.
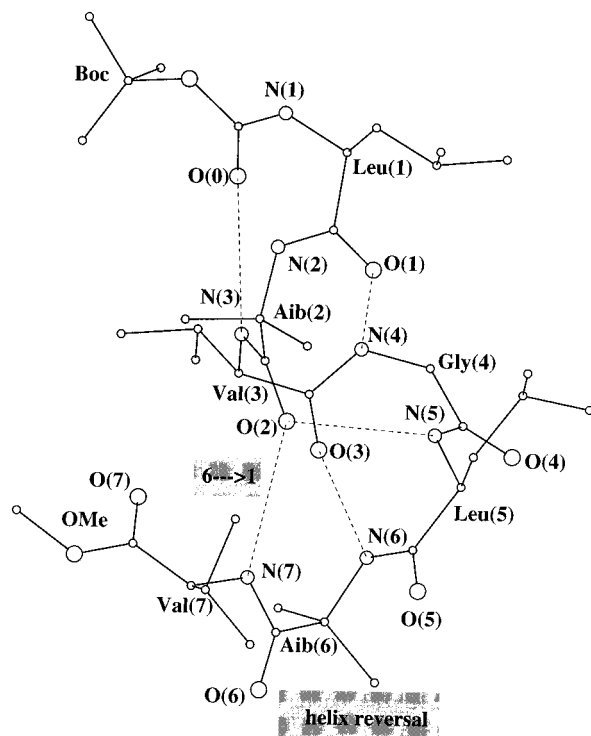
Figure 1. Conformation of the heptapeptide **1** in crystals. Broken lines indicate the 4 → 1 and 6 → 1 hydrogen bonds.

be stressed that even distinction between 3_{10}- and α-helical structures for peptides in solution have proved extremely difficult, with most methodologies being restricted only to very specific examples.[12] We describe in this paper the characterization by X-ray diffraction of an α_L-terminated helix in the model peptide Boc-Leu-Aib-Val-Gly-Leu-Aib-Val-OMe (**1**) in single crystals. Diagnostic NOEs are demonstrated for the Schellman motif in solution. Useful interproton distances have been derived from an analysis of several published peptide crystal structures and idealized polypeptide helices, permitting NMR characterization of this widely observed helix terminating structural feature.

Experimental Procedures

Peptide Synthesis. Peptide **1** was synthesized by conventional solution phase methods by using a racemization free, fragment condensation strategy. The Boc-group was used for N-terminal protection, and the C-terminus was protected as a methyl ester. Deprotections were performed using 98% formic acid or saponification, respectively. Couplings were mediated by dicyclohexylcarbodiimide-1-hydroxybenzotriazole (DCC/HOBt). All the intermediates were characterized by ¹H NMR (400 MHz) and thin layer chromatography (TLC) on silica gel and used without further purification. The final peptide was purified by medium pressure liquid chromatography (MPLC) and high performance liquid chromatography (HPLC) on reverse phase C-18 columns and fully characterized by 400 MHz ¹H-NMR.

Boc-Gly-Leu-Aib-Val-OMe (2). To 1.5 g (3.5 mmol) of Boc-Leu-Aib-Val-OMe[10b] was added 10 mL of 98% formic acid and the removal

Table 1. Torsion Angles[d] (deg) of Boc-Leu-Aib-Val-Gly-Leu-Aib-Val-OMe

residue	ϕ	ψ	ω	χ^1	χ^2
Leu	-50^a	-37	-175	170	$-159, 67$
Aib	-48	-43	-174		
Val	-63	-35	176	$70, -172$	
Gly	-70	-20	-179		
Leu	-106	17	170	-55	$-180, -51$
Aib	65	35	168		
Val	-126	167^b	175^c	$-58, 66$	

[a] C'(0)−N(1)−Cᵅ(1)−C'(1). [b] N(7)−Cᵅ(7)−C'(7)−O(OMe). [c] Cᵅ(7)−C'(7)−O(OMe)−C(OMe). [d] The torsion angles for rotation about bonds of the peptide backbone (ϕ, ψ, and ω) and about bonds of the amino acid side chains (χ^1, χ^2) follow the IUPAC-IUB Commission on Biochemical Nomenclature (ref 7). Estimated standard deviations ~1.0°.

of the Boc group was monitored by TLC. After 8 h, the formic acid was removed in vacuo. The residue was taken in water (20 mL) and washed with diethyl ether (2 × 20 mL). The pH of the aqueous solution was then adjusted to 8 with sodium bicarbonate and extracted with ethyl acetate (3 × 30 mL). The extracts were pooled, washed with saturated brine, dried over sodium sulfate, and concentrated to 5 mL of the highly viscous liquid that gives ninhydrin positive test. The tripeptide free base was added to an ice cooled solution of Boc-Gly-OH (0.50 g, 2.86 mmol) in 10 mL of DMF, followed by 0.63 g (2.94 mmol) of DCC and 0.40 g (2.94 mmol) of HOBt. The reaction mixture was stirred for 3 days. The residue was taken in ethyl acetate (60 mL), and dicyclohexylurea (DCU) was filtered off. The organic layer was washed with 2 N HCl (3 × 50 mL), 1 M sodium carbonate (3 × 50 mL), brine (2 × 50 mL), dried over sodium sulfate and evaporated in vacuo to yield 0.87 g (62%) of a gummy material: 400 MHz ¹H NMR (CDCl₃, δ ppm) 0.88−0.96 (12H, m, Leu CᵟH₃ and Val CᵞH₃), 1.45 (9H, s, Boc CH₃), 1.50 (3H, s, Aib CᵝH₃), 1.59 (3H, s, Aib CᵝH₃), 1.68−1.71 (3H, m, Leu CᵞH and Leu CᵝH), 2.16 (1H, m, Val CᵝH), 3.74 (3H, s, OCH₃), 3.90 (2H, m, Gly CᵅH), 4.36 (1H, m, Leu CᵅH), 4.48 (1H, m, Val CᵅH), 5.33 (1H, m, Gly NH), 6.53 (1H, d, Leu NH), 6.86 (1H, s, Aib NH), 7.10 (1H, d, ValNH).

Boc-Leu-Aib-Val-Gly-Leu-Aib-Val-OMe (1). **2** (0.65 g, 1.34 mmol) was deprotected with 98% formic acid and worked up as reported in the preparation of **2**. This was coupled to 0.40 g (0.98 mmol) of Boc-Leu-Aib-Val-OH[9] in 15 mL of DMF using 0.25 g (1.25 mmol) of DCC and 0.15 g (1.11 mmol) of HOBt. After 3 days the reaction was worked up as usual to yielded 0.30 g of the crude peptide. The peptide was purified on a reverse phase C-18 MPLC column using methanol−water mixtures. The peptide was further subjected to HPLC purification on a Lichrosob reverse phase C-18 HPLC column (4 × 250 mm, particle size 10μ, flow rate 1.5 mL/min) and eluted on a linear gradient of methanol−water (70−90%) with a retention time of 16 min. The peptide was homogeneous on a reverse phase C-18 (5μ) column and fully characterized by NMR (see Results).

X-ray Studies. Crystals of Boc-Leu-Aib-Val-Gly-Leu-Aib-Val-OMe were grown from a methanol−water solution by slow evaporation. The crystals were transparent and rectangular in shape. X-ray diffraction data were collected with an automated four circle diffractometer using MoK$_\alpha$ ($\lambda = 0.7107$ Å). ω-2θ scan type was used with a variable scan rate, and $2\theta_{max} = 50°$, for a total 4252 independent reflections and 1731 reflections with intensities $\geq 2\sigma(I)$. Lorentz and polarization correction were applied to the data. The space group is $P2_1$ with $a = 9.958(3)$ Å, $b = 20.447(3)$ Å, $c = 11.723(2)$ Å, $\beta = 99.74(2)°$, $V = 2352.4(9)$, $Z = 2$ for chemical formula $C_{38}H_{69}N_7O_{10}$ with one formula unit per asymmetric unit.

The structure was solved by direct phase determination using the random-tangent formula procedure in the SHELX-86[13] computer program and was refined by full matrix least-squares method using SHELX-93.[14] All the non-hydrogen atoms were anisotropically refined, before fixing hydrogen atoms geometrically in idealized positions. In the final cycle of refinement hydrogen atoms were treated as riding

(11) (a) Rajashankar, K. R.; Ramakumar, S.; Mal, T. K.; Jain, R. M.; Chauhan, V. S. *Angew. Chem., Int. Ed. Engl.* **1996** *35*, 765−767. (b) Rajashankar, K. R.; Ramakumar, S.; Jain, R. M.; Chauhan, V. S. *J. Am. Chem. Soc.* **1995**, *117*, 11773−11779. (c) Rajashankar, K. R.; Ramakumar, S.; Jain, R. M.; Chauhan, V. S. *J. Biomol. Struct. Dyn.* **1996**, *13*, 641−647.
(12) (a) Sudha, T. S.; Vijayakumar, E. K. S.; Balaram, P. *Int. J. Peptide Protein Res.* **1983**, *22*, 464−468. (b) Basu, G.; Kuki, A. *Biopolymers* **1993**, *33*, 995−1000. (c) Miick, S. M.; Martinez, G. V.; Fiori, W. R.; Todd, A. P.; Millhauser, G. L. *Nature* **1992**, *359*, 653−655. (d) Toniolo, C.; Polese, A.; Formaggio, F.; Crisma, M.; Kamphuis, J. *J. Am. Chem. Soc.* **1996**, *118*, 2744−2745. (e) Anderson, N. H.; Liu, Z.; Prickett, K. S. *FEBS Lett.* **1996**, *399*, 47−52.
(13) Sheldrick, G. M. *SHELX 86, Program for Crystal Structure Determination*; University of Gottingen: Germany, 1986.
(14) Sheldrick, G. M. SHELX 93, *Program for Crystal Structure Refinement*; University of Gottingen: Germany, 1993.

Table 2.   Parameters for $4 \rightarrow 1/5 \rightarrow 1/6 \rightarrow 1$ Interactions in Boc-Leu-Aib-Val-Gly-Leu-Aib-Val-OMe

			length (Å)		angle (deg)		
type	donor	acceptor	N- - -O	H- - -O	C=O- - -H	C=O- - -N	O- - - -HN
				Intermolecular			
a	N(1)[b]	O(6)	2.86	2.10	156	146	147
a	N(2)[c]	O(4)	2.93	2.22	141	151	140
				Intramolecular			
$4 \rightarrow 1$[a]	N(3)	O(0)	3.10	2.32	131	134	152
$4 \rightarrow 1$[a]	N(4)	O(1)	2.95	2.30	120	130	132
$4 \rightarrow 1$[a]	N(5)	O(2)	3.08	2.37	102	112	141
$4 \rightarrow 1$[a]	N(6)	O(3)	3.35	2.52	102	105	163
$4 \rightarrow 1$	N(7)	O(4)	6.49	5.92	45	51	129
$5 \rightarrow 1$	N(4)	O(0)	4.27	3.51	135	141	149
$5 \rightarrow 1$	N(5)	O(1)	3.55	2.86	146	155	138
$5 \rightarrow 1$	N(6)	O(2)	3.25	2.93	142	157	104
$5 \rightarrow 1$	N(7)	O(3)	4.90	4.32	85	92	128
$6 \rightarrow 1$[a]	N(7)	O(2)	3.19	2.34	149	146	170

[a] These are interactions satisfying the criteria of good C=O- - - -H-N hydrogen bond geometry (refs 1, 2, and 22). [b] Symmetry equivalent $1+x$, y, $1+z$ to the coordinates of heptapeptide **1**. [c] Symmetry equivalent $1+x$, y, z to coordinates of heptapeptide **1**.

over the non-hydrogen atom to which they are bonded. The final R_1 and wR_2 were 7.2% and 18.9%, respectively, for 1731 observed reflections [$I \geq 2\sigma(I)$]. The maximum and minimum values of the residual electron density map were 0.348 e/Å³ and −0.185 e/Å³, respectively. Goodness-of-fit (S) = 0.937.

NMR Studies. All NMR experiments were carried out on a Bruker AMX-400 spectrometer. Peptide concentrations were in the range of 5−6 mM. Resonance assignments were done using two-dimensional double quantum filtered COSY and rotating frame nuclear Overhauser effect (ROESY) experiments. All 2D data were acquired at 1K data points, 512 experiments with 48−64 transients. A 300 ms mixing time was used for ROESY experiments. The spectral width for all the experiments were set to 4500 Hz. NMR data were processed using UXNMR or FELIX software. All two-dimensional data sets were zero filled to 1024 points with a 90° phase shifted squared sine-bell filter in both dimensions. The probe temperature was maintained at 303 K.

Results and Discussion

Peptide Helix in Crystals. Figure 1 shows a view of the molecular conformation of the heptapeptide **1** in crystals. The backbone dihedral angles and all relevant distances and angular parameters between potential hydrogen bond donor and acceptor groups are summarized in Tables 1 and 2, respectively. Inspection of the ϕ, ψ values in Table 1 reveals that residues 1−4 fall in the $3_{10}/\alpha$-helical regions of conformational space. Aib(6) adopts a left-handed (α_L) conformation with positive ϕ, ψ values. The backbone torsion angles of Leu(5) are appreciably distorted and lie in the bridge region of the Ramachandran map.[15] A comparison of the relevant geometrical parameters for potential C=O- - -HN ($4 \rightarrow 1$ and $5 \rightarrow 1$) (Table 2) establishes that the conformation of the peptide is best described as a 3_{10}-helix spanning residue 1 to 4, stabilized by three intramolecular $4 \rightarrow 1$ hydrogen bonds [O(0)- - -N(3), O(1)- - -N(4), and O(2)- - -N(5)]. At the C-terminus, the helix ends with the formation of a Schellman motif involving a $6 \rightarrow 1$ hydrogen bond between N(7)- - -O(2) and a $4 \rightarrow 1$ hydrogen bond between N(6)- - -O(3). This latter hydrogen bond corresponds to the "$5 \rightarrow 2$" interaction within the 16 atom hydrogen bonded ring, C_{16}, of a $6 \rightarrow 1$ hydrogen bond. This feature has been termed as a "paper clip" structure by Milner−White.[8] The N(6)- - - -O(3) distance of 3.35 Å is rather long suggesting a weak interactions. Two potential multiple interactions may be considered at the C-terminus of the molecule. O(2) is positioned proximate to three hydrogen donors N(5), N(6), and N(7). Of
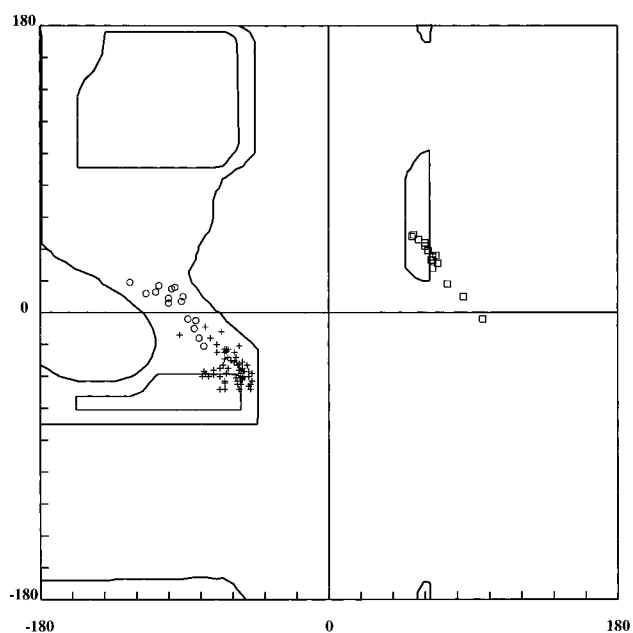
(15) (a) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95−99. (b) Ramakrishnan, C.; Ramachandran, G. N. *Biophys. J.* **1965**, *5*, 909−932.

Figure 2. Ramachandran map showing the ϕ, ψ distribution of residues in α_L terminated helices in peptide crystal structures. The sequences of the 14 peptide examples are detailed in Table 3, footnote a (peptide I to XIII): +, residues in the body of the helix; □, α_L residues which are the penultimate residue in all cases; and ○, residues preceding α_L terminator.

these the O- - - -H distance and O- - -HN angle clearly argue against any favorable interaction between O(2)- - -N(6). All other $5 \rightarrow 1$ interactions in this structures are well outside conventional hydrogen bonding limits.[1,2]

Peptide helices in the crystal are linked by intermolecular hydrogen bonds, O(6)- - -N(1) and O(4)- - -N(2). Two C=O group of Leu(5) and Val(7) do not participate in any stabilizing hydrogen bond in the crystal. The extended conformation at Val(7) places its C=O group in the interior of the $6 \rightarrow 1$ hydrogen bonded π-turn.

Comparison of α_L-Terminated Helices in Peptides. Figure 2 shows the Ramachandran plot which represents backbone torsion angles observed in 14 helical peptide crystal structures which terminate in a Schellman motif. Residues in the body of the helix are closely clustered in the $3_{10}/\alpha$-helical region of ϕ, ψ space. A fairly tight cluster is also observed for the α_L residue. In most cases, a significant distortion of backbone ϕ, ψ values from those observed in helices is seen for the residue

Table 3. A Comparison of $6 \rightarrow 1$ and $4 \rightarrow 1$ Hydrogen Bonds in Schellman Motifs Characterized in Peptide Crystal Structures

sequence[a]	type[b]	N- - - -O (Å)	H- - - -O (Å)	C=O- - - -H (deg)	C=O- - - -N (deg)	O- - - -HN (deg)
peptide **I**	$4 \rightarrow 1$	3.29	2.45	97	100	163
	$6 \rightarrow 1$	2.91	2.20	174	166	141
peptide **II**	$4 \rightarrow 1$	3.10	2.25	109	112	166
	$6 \rightarrow 1$	2.92	2.10	175	170	159
peptide **IIIa**	$4 \rightarrow 1$	3.01	2.17	112	115	166
	$6 \rightarrow 1$	2.94	2.21	172	176	143
peptide **IIIb**	$4 \rightarrow 1$	3.09	2.24	104	106	167
	$6 \rightarrow 1$	3.09	2.24	143	141	168
peptide **IV**	$4 \rightarrow 1$	3.37	2.54	102	106	162
	$6 \rightarrow 1$	3.07	2.22	152	149	169
peptide **V**	$4 \rightarrow 1$	3.35	2.52	102	105	163
	$6 \rightarrow 1$	3.19	2.34	149	146	170
peptide **VI**	$4 \rightarrow 1$	3.41	2.59	103	105	161
	$6 \rightarrow 1$	3.02	2.18	147	144	167
peptide **VII**	$4 \rightarrow 1$	4.74[c]	3.74			
	$6 \rightarrow 1$	2.95	2.00	154	150	155
peptide **VIII**	$4 \rightarrow 1$	2.89	1.94	103	108	155
	$6 \rightarrow 1$	3.12	2.17	131	128	157
peptide **IX**	$4 \rightarrow 1$	3.14	2.30	110	112	167
	$6 \rightarrow 1$	3.06	2.22	152	148	163
peptide **X**	$4 \rightarrow 1$	2.91	2.08	112	115	162
	$6 \rightarrow 1$	2.97	2.21	164	158	147
peptide **XI**	$4 \rightarrow 1$	3.25	2.49	105	112	148
	$6 \rightarrow 1$	5.12[d]	4.28			
peptide **XII**	$4 \rightarrow 1$	3.17	2.35	106	111	162
	$6 \rightarrow 1$	3.06	2.34	177	170	143
peptide **XIII**	$4 \rightarrow 1$	2.98	2.14	115	116	163
	$6 \rightarrow 1$	3.08	2.30	147	147	150
peptide **XIV**	$4 \rightarrow 1$	2.99	2.32	112	122	132
	$6 \rightarrow 1$	4.84[d]				

[a] Sequences of peptides and references in parentheses. Peptide I: Boc-Val-Aib-Phe-Aib-Ala-Aib-Leu-OMe (ref 10a). Peptide II: Boc-Val-Aib-Leu-Aib-Ala-Aib-Phe-OMe (ref 10a). Peptide III: Boc-Val-Aib-Leu-Aib-Ala-Aib-Leu-OMe (ref 10a). Peptide IV: Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-OMe (ref 10c). Peptide V: Boc- Leu-Aib-Val-Gly-Leu-Aib-Val-OMe (this study). Peptide VI: Boc-Pro-Aib-Gly-Leu-Aib-Leu-OMe (unpublished). Peptide VII: Ac-ΔPhe-Val-ΔPhe-Phe-Ala-Val-ΔPhe-Gly-OMe (ref 11b). Peptide VIII: Boc-Val-ΔPhe-Phe-Ala-Leu-Ala-ΔPhe-Leu-OH (ref 11c). Peptide IX: pBrBz-(Aib-Ala)$_5$-OMe·2H$_2$O (from aqueous methanol solution) (ref 10d). Peptide X: pBrBz-(Aib-Ala)$_6$-OMe·2H$_2$O (ref 10d). Peptide XI: Boc-Leu-Aib-Val-Gly-Gly-Leu-Aib-Val-OMe (ref 10b). Peptide XII: Boc-Val-ΔPhe-Leu-Phe-Ala-ΔPhe-Leu-OMe (ref 11a). Peptide XIII: pBrBz-(Aib-Ala)$_5$-OMe (from a DMSO−isopropyl alcohol solvent mixture) (ref10e). Peptide XIV: Boc-D(Val-Ala-Leu-Aib-Val-Ala-Leu)-L(Val-Ala-Leu-Aib-Val-Ala-Leu) (ref 19). [b] The $4 \rightarrow 1$ interaction correspond to the "$5 \rightarrow 2$" hydrogen bond within the Schellman motif. [c] The encased $4 \rightarrow 1$ hydrogen bond associated with the π-turn in this case has been formed mediated by a water molecule (O(4)- - - -O(W) = 2.70 Å and O(W)- - - -N(7) = 3.02 Å). [d] In peptides XI and XIV a solvent methanol is inserted into the $6 \rightarrow 1$ hydrogen bond (O(3)- - - -O(M) = 2.76 Å, O(M)- - - -N(8) = 2.93 Å for peptide XI and O(4)- - - -O(M) = 2.72 Å, O(M)- - - -N(9) = 2.86 Å for peptide XIV).

Table 4. Key Interproton Distances in Peptide Helices

	α_L-terminated helix[a] (Å)	α-helix[b] (Å)	3_{10}-helix[b] (Å)	$\alpha/3_{10}$-helix[b] (Å)	ideal α-helix[c] (Å)	ideal 3_{10}-helix[c] (Å)
C^β_iH↔N$_{i+3}$H (mean)	3.62 ± 0.24	5.44 ± 0.63	5.11 ± 0.92	5.25 ± 0.82	5.37	3.38
(min.)		3.23	3.20	2.88		
(max.)		6.08	7.03	6.17		
C^β_iH↔N$_{i+5}$H (mean)	3.69 ± 0.18	7.76 ± 0.58	9.27 ± 0.91	7.89 ± 0.81	7.88	7.66
(min.)		6.39	7.45	5.48		
(max.)		8.69	10.67	9.71		
C^α_iH↔N$_{i+2}$H (mean)	3.92 ± 0.06	4.47 ± 0.18	3.88 ± 0.34	4.29 ± 0.30	4.53	3.99
C^α_iH↔N$_{i+3}$H (mean)	3.48 ± 0.15	3.66 ± 0.22	3.87 ± 0.25	3.69 ± 0.22	3.43	3.49
C^α_iH↔N$_{i+4}$H (mean)	3.61 ± 0.11	4.26 ± 0.16	6.43 ± 0.32	4.72 ± 0.49	4.20	5.54

[a] These mean distances have been calculated taking the corresponding values (≤4.0 Å) from the crystal structures of peptides given in the Table 3. [b] Distances have been calculated from the crystal structures of the peptides given in the following list. In case of C^β_iH↔N$_{i+3}$H and C^β_iH↔N$_{i+5}$H calculations all the H-atoms of two β-CH$_3$ groups of the Aib residue has been considered. Peptide helices of purely α-helical conformation: Peptide I: Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-Ala-Leu-Aib-OMe (unpublished). Peptide II: Boc-Aib-Ala-Aib-Ala-Leu-Ala-Leu-Aib-Leu-Aib-OMe (ref 17a). Peptide III: Boc-Val-Ala-Leu-Aib-Val-Ala-Leu-(Val-Ala-Leu-Aib)$_2$-OMe (ref 17b). Peptide helices of purely 3_{10}-helical conformation: Peptide I: Boc-(Ala-Aib)$_4$-OMe (ref 18a). Peptide II: pBrBz-(Aib)$_8$-OMe (ref 18b). Peptide III: pBrBz-(Aib)$_3$-Val-Gly-Leu-(Aib)$_2$-OMe (ref 18b). Peptide helices of mixed (3_{10}/α)-helical conformation: Peptide I: Boc-Trp-Ile-Ala-Aib-Ile-Val-Aib-Leu-Aib-Pro-OMe (ref 23). Peptide II: Boc-Aib-Val-Ala-Leu-Aib-Val-Ala-Leu-Aib-OMe (ref 24). Peptide III: Boc-Aib-Val-Aib-Aib-Val-Val-Val-Aib-Val-Aib-OMe (ref 25). [c] Distances have been calculated from the computer-generated model of α- and 3_{10}-helix using standard bond lengths, bond angles, and backbone torsion angles[3b] with INSIGHT-II (Biosym Technologies, San Diego, CA).

immediately preceding the α_L-residue. This observation suggests that the stereochemistry of these helix termination motifs is fairly sharply defined. Table 3 compares the geometrical parameters for the $6 \rightarrow 1$ and $4 \rightarrow 1$ hydrogen bonds within the Schellman motif. In most examples, the parameters for $6 \rightarrow 1$ interaction are indicative of a stronger hydrogen bond as compared to the $4 \rightarrow 1$ interaction. Examples of solvent insertion into both $4 \rightarrow 1$ and $6 \rightarrow 1$ hydrogen bonds of Schellman motifs are found in Table 3.

Interproton Distances Characteristic of the Schellman Motif. Three-dimensional structure determination in peptides and proteins in solution relies on the ability to observe short (3.5 Å → 4.0 Å) interproton distances by detection of nuclear Overhauser effects (NOEs).[16] The availability of several high
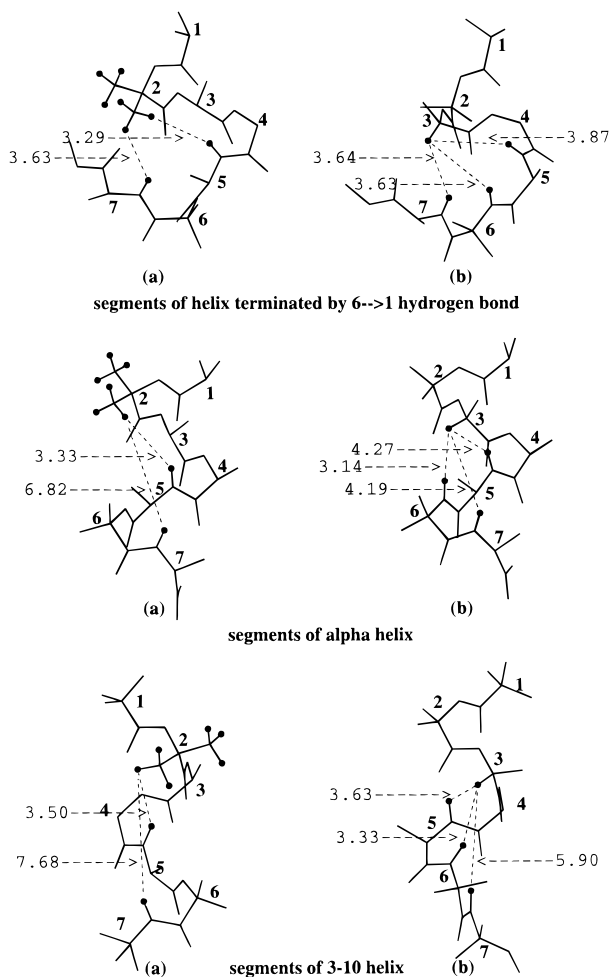
Figure 3. (top) Partial structure of the heptapeptide **1** in crystals. The hydrogen atoms are shown in filled circles. The broken lines indicate the key interproton distances in the Schellman motif. Similar illustration of distances in crystal structure of (center) α-helix segment in Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-Ala-Leu-Aib-OMe (unpublished) and (bottom) 3_{10}-helix segment in pBrBz-(Aib)$_3$-Val-Gly-Leu-(Aib)$_2$-OMe (ref 18b). (a) Distances between the N(5)H/N(7)H proton and one of the pro-**R**-β-CH$_3$ protons of Aib(2) are highlighted. (b) The C_i^αH- - - -N$_{i+2}$H [$d_{\alpha N}(i,i+2)$], C_i^αH- - - -N$_{i+3}$H [$d_{\alpha N}(i,i+3)$] and C_i^αH- - - -N$_{i+4}$H [$d_{\alpha N}(i,i+4)$] connectivities are highlighted.

resolution crystal structures, where peptides adopt α-helical,[17] 3_{10}-helical,[18] and α_L-terminated helical conformations,[10,11,19] permits an analysis of interresidue interproton distances which may be used as a conformational diagnostic in NOE experiments.

Figure 3 shows a view of the relevant segment of the heptapeptide **1** with crucial short interproton distances indicated. Comparable fragments chosen from crystallographically characterized α- and 3_{10}-helices are also shown. In all the three cases an Aib residue is present at position 2 with both C^β-methyl groups being shown. A key feature that is immediately apparent

is the short distances between the N(7)H proton and one of the protons of the pro-**R**-β-CH$_3$ of Aib(2) (3.63 Å). This distance is dramatically larger in α- (6.82 Å) and 3_{10}- (7.68 Å) helical examples. The use of the C^βH(2) → NH(7) distance as a diagnostic is, of course, limited to situations where a pro-**R**-substituent is present at the residue 2 α-carbon atom. Two additional short interproton distances which are potentially useful are between C^αH(3)- - -NH(6) and C^αH(3)- - -NH(7), both of which are ∼3.6 Å. The corresponding distances in the α-helical examples in Figure 3 are 3.14 and 4.19 Å, while in the 3_{10}-helical case these are 3.33 and 5.90 Å. Encouraged by this observation we have computed key interproton distances in several experimentally determined crystal structures of α_L-terminated helices, 3_{10}-helices, and α-helices and also in idealized 3_{10}- and α-helical structures. The results are summarized in Table 4. It is clear that the simultaneous observation of both C_i^αH- - -N$_{i+3}$H [$d_{\alpha N}(i,i+3)$] and C_i^αH- - -N$_{i+4}$H [$d_{\alpha N}(i,i+4)$] will be characteristic of a π-turn, involving a 6 → 1 hydrogen bond between C=O($i-1$) and NH($i+4$). It may be noted that the backbone stereochemistry of the π-turn, involving a C_{16} hydrogen bonded ring, is determined by the backbone conformational angles (ϕ, ψ) at residues i, $i+1$, $i+2$, and $i+3$. The Schellman motif corresponds to a peptide segment with the conformational assignment α_R-α_R-α_R-α_L.[9a] In the case of the heptapeptide **1** this encompasses residues Val(3)-Gly(4)-Leu(5)-Aib(6). When the second substituent at $C^\alpha(i-1)$ is present, as in the case of Aib(2) in peptide **1**, additional short interproton distances $C^\beta(i-1)$H- - -N($i+2$)H and $C^\beta(i-1)$H- - -N($i+4$)H are observable. The data in Table 4 strongly suggest that in the case of relatively small peptides where interproton NOEs are usually observed up to distances of 3.5 → 4.0 Å, a clear identification of the Schellman motif is indeed possible.

NMR Studies of Heptapeptide 1. Solution NMR studies were carried out in the apolar solvent CDCl$_3$ in order to facilitate intramolecular hydrogen bond formation, without perturbations due to competing solvent−peptide interactions. Assignments of all resonances were achieved in a straightforward manner using a combination of 2D COSY and ROESY experiments.[20] The chemical shifts are summarized in Table 5. The solvent accessibility of peptide NH groups was probed using a solvent perturbation experiment carried out by addition of increasing amount of DMSO. The solvent shift ($\Delta\delta$ values) in Table 5 confirm that only Leu(1) and Aib(2) NH groups are solvent exposed ($\Delta\delta > 0.8$ ppm). The remaining five NH groups are relatively unperturbed at low DMSO concentration, suggesting their involvement in intramolecular hydrogen bonds. The simplest interpretation of solvent inaccessibility is to conclude that the peptide folds into a continuous 3_{10}-helical conformation, stabilized by five successive 4 → 1 hydrogen bonds. α-Helical structure involving bifurcated hydrogen bond formation, where Val(3)NH and Gly(4)NH groups interact simultaneously with the Boc C=O group cannot, of course, be excluded.

Figure 4 shows a partial ROESY spectrum displaying successive NH−NH connectivities. $N_iH \leftrightarrow N_{i+1}H$ [$d_{NN}(i,i+1)$] NOEs are observed for residues 1−6 suggesting that each of

(16) Neuhaus, D.; Williamson, M. In *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; VCH Publishers, Inc.: New York, 1989.

(17) (a) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Biopolymers* **1990**, *30*, 719−731. (b) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Sukumar, M.; Balaram, P. *J. Am. Chem. Soc.* **1990**, *112*, 9350−9356.

(18) (a) Otoda, K.; Kitagawa, Y.; Kimura, S.; Imanishi, Y. *Biopolymers* **1993**, *33*, 1337−1345. (b) Bavoso, A.; Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 1988−1992.

(19) Banerjee, A.; Raghothama, S. R.; Karle, I. L.; Balaram, P. *Biopolymers* **1996**, *39*, 279−285.

(20) Wuthrich, K. in *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.

(21) (a) Srinivasan, N.; Anuradha, V. S.; Ramakrishnan, C.; Sowdhamini, R.; Balaram, P. *Int. J. Peptide Protein Res*. **1994**, *44*, 112−122. (b) Richardson, J. S. *Adv. Protein Chem*. **1981**, *34*, 164−339.

(22) Datta, S.; Shamala, N.; Banerjee, A.; Balaram, P. *Int. J. Peptide Protein Res.* **1997** In press.

(23) Karle, I. L.; Flippen-Anderson, J. L.; Sukumar, M.; Balaram, P. *Int. J. Peptide Protein Res*. **1988**, *31*, 567−576.

(24) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Int. J. Peptide Protein Res*. **1988**, *32*, 536−543.

(25) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, H.; Balaram, P. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 765−769.

Table 5. Chacteristic ^{1}H NMR Parameters of the Peptide **1**

residues	NH	δ values^a (ppm)				$^3J_{NHC\alpha H}$^{b}	$\Delta\delta^c$
		C$^\alpha$H	C$^\beta$H	C$^\gamma$H	C$^\delta$H	(Hz)	
Leu1	5.15	3.87	1.53	1.53	0.90	2.5	1.18
Aib2	6.71		1.41, 1.47				0.88
Val3	7.32	4.04	2.25	1.07	4.4	0.03	
Gly4	7.99	4.15, 3.69				5.9, 6.8	0.15
Leu5	7.24	4.55	1.63	1.63	0.99	8.8	0.07
Aib6	7.12		1.54, 1.58				0.09
Val7	7.02	4.41	2.10	0.93		8.9	0.05

^{a} Chemical shift values of proton resonances in CDCl$_3$. ^{b} Coupling constants in CDCl$_3$. ^{c} $\Delta\delta$ is the chemical shift difference for NH protons in CDCl$_3$ and 9.9%(CD$_3$)$_2$SO/CDCl$_3$.

Figure 5. Partial 400 MHz ^{1}H−^{1}H ROESY spectra showing some important C$^\beta$H−NH connectivities (top panel) and C$^\alpha$H−NH connectivities (bottom panel) (Aib ≡ U).

Figure 4. Partial 400 MHz ^{1}H-^{1}H ROESY spectra for peptide **1** in CDCl$_3$ showing some important NH-NH connectivities. The abbreviation U is used for Aib.

these residues in peptide **1** adopts ϕ, ψ values which lie in the helical region of the Ramachandran map [$\phi \cong \pm 50°$, $\psi \cong \pm 50°$]. The observation of d$_{NN}$(i,i+1) NOEs, of course, establishes <u>local helical</u> conformation at each residue but does not provide an indication of the signs of the ϕ, ψ values or the helix sense.

Since the crystal structure of peptide **1** reveals a helix reversal at Aib(6), it is important to establish whether this feature is retained in solution. An unambiguous answer to this question is not provided by the NH group accessibility data since the NH groups of residue 3−7 remain hydrogen bonded in both the continuous 3$_{10}$-helix and an α_L-terminated 3$_{10}$-helix. Figure 5 shows the ROESY spectrum which depicts NOEs between the C$^\alpha$H-NH and C$^\beta$H-NH protons. It is evident that NOEs are observable between one of the two C$^\beta$-methyl resonances of Aib(2) ($\delta = 1.36$ ppm) and Leu(5) NH and Val(7) NH groups. Furthermore, two almost equally intense NOEs are observed between the Val(3) C$^\alpha$H and Aib(6) and Val(7) NH groups. These are precisely the short interproton distances expected from the conformation characterized in the crystal structure (Figure 1).

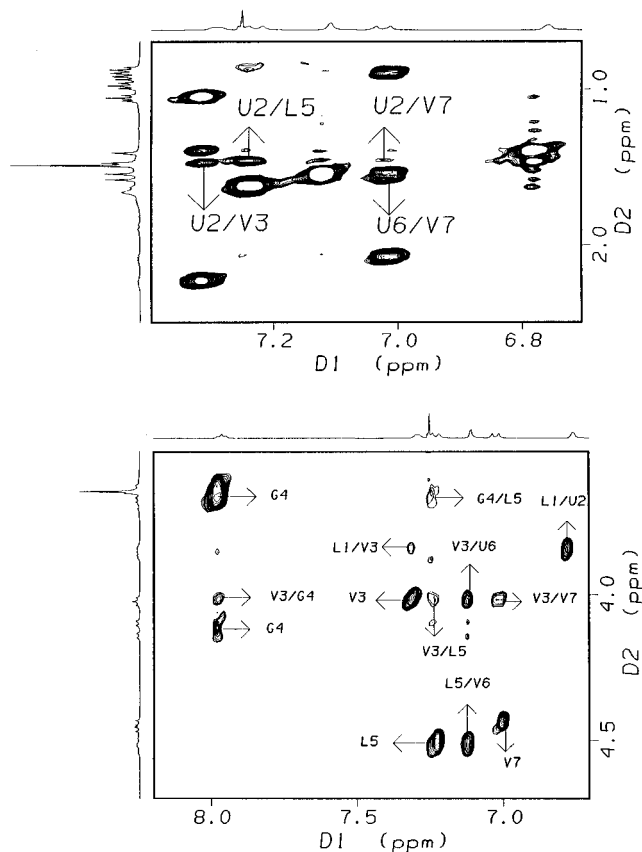The NMR results thus provide unambiguous evidence for the retention in solution of the 6 → 1 hydrogen bond terminated 3$_{10}$-helical conformation observed in crystals. The results presented in this paper suggest that characteristic NOE patterns permit identification of the Schellman motif in solution. In discussing stability of helical conformations in peptides, α_L-terminated structures must clearly be considered as important contributors, when achiral residues (cf. Gly/Aib) or residue with a high α_L propensity, for example, Asn^{21}, are present at the C-terminus of the helix.

Supporting Information Available: Tables of crystal data and structure refinement details, atomic coordinates, bond lengths, bond angles, anisotropic temperature factors and hydrogen atom coordinates for peptide **1** (11 pages). See any current masthead page for ordering and Internet access instructions.

JA970551Y