Heterogeneity and Stability of Helical Conformations in Peptides: Crystallographic and NMR Studies of a Model Heptapeptide

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

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

Received February 29, 1996[⊗]

Abstract: The helix is a common secondary structural element in proteins. Several subtypes of helical structures are possible, depending on the nature of the backbone hydrogen-bonding pattern. For short peptides in solution, more than one helical conformation may be in equilibrium. Most of the time it is extremely difficult to identify experimentally the interconversion between different helix forms in solution, although one of the helical structures may be trapped in single crystals. This report analyzes crystallographic and NMR studies of a model heptapeptide, Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-OMe, designed to probe conformational heterogeneity and relative stability of different helical structures. In crystals the peptide adopts a right-handed 3_{10}-helical structure, terminated by a *left-handed* helical conformation (α_L) at Aib(6) resulting in a 6→1 hydrogen bond at the C-terminus. The crystals are in space group $P2_1$, $a = 9.843(4)$ Å, $b = 19.944(4)$ Å, $c = 12.456(2)$ Å, $\beta = 102.51(2)°$, $Z = 2$, $R = 10.1\%$, and $R_w = 9.21\%$ for 2301 reflections; $|F| > 3\sigma(F)$. In $CDCl_3$ NMR studies establish that the NH groups of residues 3−6 are hydrogen bonded, consistent with a continuous 3_{10}-helix or the conformation observed in crystals. In $(CD_3)_2SO$, NMR studies reveal that the NH groups of residues 1−3 are solvent exposed, supporting an α-helical structure. However, NMR data suggest appreciable conformational heterogeneity in solution, with population of extended states and conformations resembling the structure in crystals. Helical conformations terminated by α_L residues may be an important subset of helical structure in peptides, particularly when the penultimate residue is achiral or has a propensity to adopt α_L conformations.

Introduction

The stereochemical constraints that accompany the introduction of α-aminoisobutyryl (Aib)[1] residues into acyclic peptide sequences facilitate stabilization of helical folding patterns in short peptides.[2] A large number of crystal structure determinations of Aib-containing peptides have convincingly demonstrated the helix-stabilizing properties of these residues.[3] X-ray crystallographic studies have, however, shown that subtle variation in helix type (3_{10}, α), hydrogen-bonding patterns (4→1, 5→1, 6→1), and backbone solvation are frequently observable features.[3,4] Recent structure determinations reveal an unusual stereochemical feature when an achiral Aib residue occupies a penultimate position at the C-terminal end of a helix. In these cases, the Aib residue adopts a left-handed helical conformation ($\phi \approx +60°$, $\psi \approx +30°$), resulting in the formation of a 6→1 hydrogen bond involving the NH group of the C-terminal residue.[5] Such helix-terminating structural features are indeed observed in proteins with the achiral Gly residue often adopting left-handed helical conformations.[6] The precise definition of variants of peptide helices in crystals provides an opportunity to probe the factors influencing the detailed stereochemistry of helical backbone folding patterns. Relatively little experimental information is available about dynamic interconversions between helix subtypes, although several theoretical analyses have been reported.[7] The unambiguous characterization of the nature of helical conformations in solution has also remained a difficult problem.[8] We describe in this report crystallographic and NMR studies of a helical heptapeptide, Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-OMe (**1**). This sequence was designed by placing an Aib residue at the penultimate position at the C-terminus. The Leu-Aib segment has a propensity to adopt both type II and

* 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.

† Molecular Biophysics Unit.

‡ Department of Physics.

⊗ Abstract published in *Advance ACS Abstracts,* August 1, 1996.

(1) Abbreviations used: Aib, α-aminoisobutyric acid; Boc, *tert*-butyloxycarbonyl. All chiral amino acids are of the L-configuration.

(2) (a) Nagaraj, R.; Shamala, N.; Balaram, P. *J. Am. Chem. Soc*. **1979**, *101*, 16−20. (b) Prasad, B. V. V.; Balaram, P. *CRC Crit. Revs. Biochem*. **1984**, *16*, 307−347. (c) Balaram, P. *Curr. Opin. Struct. Biol*. **1992**, *2*, 845−851. (d) Toniolo, C.; Benedetti, E. *Trends. Biochem. Sci*. **1991**, *16*, 350−353. (e) Marshall, G. R.;Hodgkin, E. E.; Langs, D. A.; Smith, G. D.; Zabrocki, J; Leplawy, M. T. *Proc. Natl. Acad. Sci. (USA)* **1990**, *87*, 487−491. (f) Bosch, R.; Jung, G.; Schmitt, H.; Winter, W *Biopolymers* **1985**, *24*, 961−978. (g) Augspurger, J. D.; Bindra, V. A.; Scheraga, H. A.; Kuki, A. *Biochemistry* **1995**, *34*, 2566−2576.

(3) Karle, I. L.; Balaram, P. *Biochemistry* **1990**, *29*, 6747−6756.

(4) (a) Karle, I. L. *Acta Crystallogr. B* **1992**, *48*, 341−356. (b) Karle, I. L.; Flippen-Anderson, J. L.; Gurunath, R.; Balaram, P. *Protein Sci*. **1994**, *4*, 1547−1555. (c) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Biopolymers* **1993**, *33*, 827−837.

(5) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Int. J. Pept. Protein Res*. **1993**, *42*, 401−410.

(6) (a) Schellman, C. In *Protein Folding*, Janenicke, R., Ed. Elsevier/North Holland Biochemical Press: Amsterdam, 1980; pp 53−61. (b) Nagarajaram, H. A.; Sowdhamini, R.; Ramakrishnan, C.; Balaram, P. *FEBS Lett*. **1993**, *321*, 79−83. (c) Milner-White, E. J. *J. Mol. Biol*. **1988**, *199*, 503−511. (d) Rajashankar, R. K.; Ramakumar, S.; Jain, R. M.; Chauhan, V. S. *J. Am. Chem. Soc.* **1995**, *117*, 11773−11779.

(7) (a) Otoda, K.; Kitagawa, Y.; Kimura, S.; Imanishi, Y. *Biopolymers*, **1993**, *33*, 1337−1345. (b) Basu, G.; Kitao, A.; Hirata, F.; Go, N. *J. Am. Chem. Soc*. **1994**, *116*, 6307−6316. (c) Daggett, V.; Kollman, P. A.; Kuntz, I. D. *Biopolymers* **1991**, *31*, 1115−1134. (d) Tirado-Rives, J.; Maxwell, D. L.; Jorgenson, W. L. *J. Am. Chem. Soc*. **1993**, *115*, 11590−11593. (e) Smythe, M. L.; Huston, S. E.; Marshall, G. R. *J. Am. Chem. Soc.* **1993**, *115*, 11594−11595.

(8) (a) Sudha, T. S.; Vijayakumar, E. K. S.; Balaram, P. *Int. J. Pept. 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) Millhauser, G. L. *Biochemistry*, **1995**, *34*, 3874−3877. (e) Toniolo, C.; Polese, A.; Formaggio, F.; Crisma, M.; Kamphuis, J. *J. Am. Chem. Soc.* **1996**, *118*, 2744−2745.

S0002-7863(96)00665-8 CCC: $12.00

type III β-turn conformations, permitting solvent dependent structural changes.[9] The crystal structure determination reveals termination of a right-handed 3_{10}-helix by Aib(6) adopting a left-handed helical conformation, with concomitant formation of a 6→1 hydrogen bond. Maintenance of the crystal state conformation in $CDCl_3$ is broadly consistent with NMR data. In $(CD_3)_2SO$, there is a transition to an α-helical structure together with the population of extended states.

Experimental Section

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 ^{1}H 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 ^{1}H NMR.

Boc-Ala-Leu-Aib-Val-OMe (2). To 1.84 g (4.29 mmol) of Boc-Leu-Aib-Val-OMe,[9] was added 10 mL of 98% formic acid, and the removal of the Boc group was monitored by TLC. After 8 h, the formic acid was removed in vacuo. The residue was taken up in water (20 mL) 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 a highly viscous liquid that gave a positive ninhydrin test. The tripeptide free base was added to an ice-cooled solution of Boc-Ala-OH (0.76 g, 4.02 mmol) in 10 mL of DMF, followed by 0.86 g (4.3 mmol) of DCC, and 0.56 g (4.2 mmol) of HOBT. The reaction mixture was stirred for 3 days. The residue was taken up in ethyl acetate (60 mL), and DCU was filtered off. The organic layer was washed with 2N HCl (3 × 50 mL), 1 M sodium carbonate (3 × 50 mL), and brine (2 × 50 mL), dried over sodium sulfate, and evaporated in vacuo to yield 1.65 g (82%) of a white solid. 400 MHz ^{1}H NMR ($CDCl_3$, δ ppm): 0.90−0.94 (12H, m, Leu C$^\delta$H$_3$ and Val C$^\gamma$H$_3$), 1.35 (3H, d, Ala C$^\beta$H$_3$), 1.44 (9H, s, Boc CH$_3$), 1.52 (3H, s, Aib C$^\beta$H$_3$), 1.56 (3H, s, Aib C$^\beta$H$_3$), 1.69−1.71 (3H, m, Leu C$^\gamma$H and Leu C$^\beta$H), 2.16 (1H, m, Val C$^\beta$H), 3.71 (3H, s, OCH$_3$), 4.13 (1H, br, Ala C$^\alpha$H), 4.31 (1H, br, Leu C$^\alpha$H), 4.47 (1H, br, Val C$^\alpha$H), 4.97 (1H, d, Ala NH), 6.60 (1H, d, Leu NH), 6.80 (1H, s, Aib NH), 7.05 (1H, d, ValNH).

Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-OMe(1): A 1.4 g (2.8 mmol) portion of **2** was deprotected with 98% formic acid and worked up as reported in the preparation of **2**. This was coupled to 1.04 g (2.5 mmol) of Boc-Leu-Aib-Val-OH[9] in 15 mL of DMF using 0.6 g of DCC and 0.34 g of HOBT. After 3 days the reaction was worked up as usual to yield 1.73 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 10 μm, flow rate 1.5 mL/min) and eluted on a linear gradient of methanol−water (70%−90%) with a retention time of 17 min. The peptide was homogeneous on a reverse phase C-18 (5 μm) column and fully characterized by NMR (see results).

X-ray Studies. Crystals of peptide **1** were grown from a methanol−water solution by slow evaporation. X-ray diffraction data were collected from a dry crystal on an automated four circle diffractometer with Cu Kα radiation. Refined unit cell parameters were determined by a least-squares fit of the angular setting of 25 accurately determined reflections in the range $0° < \theta < 25°$. Three dimensional intensity data were collected upto $2\theta = 120°$ with a ω-2θ scan and variable scan speed. Two reflections used as standards, monitored after every 100 measurements, remained constant within 3%. Pertinent parameters concerning data collection and the crystal parameters are listed in Table 1.

Table 1. Diffraction Data for Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-OMe

empirical formula	$C_{39}O_{10}N_7H_{71}$
crystal habit	clear rod shaped
crystal size (mm)	0.2 × 1.0 × 0.2
crystalizing solvent	CH_3OH/H_2O
space group	$P2_1$
cell parameters	
a (Å)	9.843(4)
b (Å)	19.944(4)
c (Å)	12.456(2)
β (deg)	102.51(2)
volume (Å^3)	2387.3
Z	2
molecules/asym unit	1
cocrystallized solvent	none
molecular weight	798.4
density (g/cm) (calcd)	1.167
F(000)	868
radiation (Å)	Cu Kα ($\lambda = 1.5418$)
temperature (°C)	21
2θ range (deg)	120
scan type	$\omega-2\theta$
scan speed	variable
independent reflections	3673
observed reflections. [$\lvert F\rvert > 3\sigma(F)$]	2301
goodness-of-fit (S)	1.08
$\Delta\rho_{max}$ (e Å^{-3})	0.35
$\Delta\rho_{min}$ (e Å^{-3})	−0.40
final *R*	10.1%
final R_w	9.21%
data-to-parameter ratio	4:1

The structure was solved by a vector search procedure using the computer program PATSEE.[10] The model used for the search was based on the fragments of the backbone in the peptide Boc-Val-Aib-Phe-Aib-Ala-Aib-Leu-OMe[5] structure. After proper orientation and translation of the model fragment used, the rest of the atoms were found with the partial structure expansion method using the tangent formula.[11]

All the non-hydrogen atoms obtained successively from vector search and the partial structure procedure were initially refined isotropically and then anisotropically using the available refinement package SHELX-76.[12] Full matrix anisotropic least-squares refinement was done on all the non-hydrogen atoms before fixing hydrogen. All the hydrogen atoms were added in idealized positions with C−H = 1.08 Å and allowed to ride with the C or N atom to which each of them was bonded, during the final cycle of refinement. The final *R* factor was 10.1% ($R_w = 9.2\%$). The function minimized during the refinement was $\sum w(\lvert F_o\rvert - \lvert F_c\rvert)^2$ where $w = 1/[\sigma^2(F) - 0.001(F^2)]$. The relatively high *R* value can be attributed to the poor reflection quality of the crystal, low data to parameter ratio as given in Table 1, and probably the large rotational motions in some of the methyl groups of the main chain and side chains of Val(3) and Val(6).

Fractional coordinates for the C, N, and O atoms are listed in Table 2. Bond lengths and bond angles do not show significant or systematic differences from expected values. Torsional angles and hydrogen bonds are listed in Tables 3 and 4, respectively.

Spectroscopic Studies. All NMR studies were carried out on a Bruker AMX-400 spectrometer at the Sophisticated Instruments Facility. Peptide concentrations were in the range of 9−10 mM. The probe temperature was maintained at 298 K for NMR experiments in $CDCl_3$ and 307 K in $(CD_3)_2SO$. Resonance assignments were done using two-dimensional double quantum filtered COSY spectra (1K data points, 512 experiments, 48 transients, spectral width of 4500 Hz). Two-

(9) Karle, I. L.; Banerjee, A.; Bhattacharjya, S.; Balaram, P. *Biopolymers* **1996**, *38*, 515−526.

(10) Egert, E.; Sheldrick, G. M. *Acta Crystallogr. A* **1988**, *41*, 262−268.

(11) Karle, J. *Acta Crystallogr. B* **1968**, *24*, 182 −186.

(12) Sheldrick, G. M. SHELX76, Program for crytal structure determination; University of Cambridge, England, 1976.

(13) Wüthrich, K. In *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986; pp 130−199.

Table 2. Atomic Coordinates ($\times 10^4$) and Equivalent Isotropic Displacement Coefficients ($Å^2 \times 10^3$) for Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-OMe

atom	*x*	*y*	*z*	*U*(eq)[a]	atom	*x*	*y*	*z*	*U*(eq)[a]
C(1)	−4092(17)	2583(10)	−3809(14)	106(08)	N(4)	2754(08)	3784(05)	−1816(07)	73(04)
C(2)	−3617(19)	2308(11)	−2626(13)	137(10)	$C^{\alpha}(4)$	3883(12)	4150(08)	−2188(09)	78(05)
C(3)	−3366(25)	2181(12)	−4569(16)	163(12)	$C^{\beta}(4)$	3283(14)	4498(10)	−3269(10)	109(06)
C(4)	−5707(20)	2463(15)	−4202(18)	183(13)	C′(4)	4571(12)	4642(08)	−1362(09)	85(06)
O	−3979(10)	3278(07)	−3883(08)	106(04)	O′(4)	5772(07)	4865(06)	−1375(06)	100(04)
C′(0)	−2738(19)	3582(12)	−3590(11)	88(08)	N(5)	3847(08)	4897(06)	−639(06)	71(04)
O′(0)	−1585(10)	3306(06)	−3147(08)	108(04)	$C^{\alpha}(5)$	4437(12)	5428(08)	139(10)	78(05)
N(1)	−2892(11)	4245(08)	−3740(08)	79(05)	$C^{\beta}(5)$	3386(20)	6027(09)	49(14)	116(08)
$C^{\alpha}(1)$	−1724(13)	4722(10)	−3474(09)	99(06)	$C^{\gamma}(5)$	2967(23)	6350(10)	−1039(15)	144(10)
$C^{\beta}(1)$	−2512(21)	5458(12)	−3577(19)	186(12)	$C^{\delta 1}(5)$	1999(27)	6955(12)	−936(19)	198(13)
$C^{\gamma}(1)$	−1687(26)	5975(17)	−3113(21)	185(14)	$C^{\delta 2}(5)$	4267(28)	6530(14)	−1532(23)	267(20)
$C^{\delta 1}(1)$	−2403(25)	6609(13)	−2974(22)	201(15)	C′(5)	4826(13)	5156(10)	1311(12)	77(06)
$C^{\delta 2}(1)$	−521(28)	6154(17)	−3887(22)	247(18)	O′(5)	5061(10)	5551(06)	2096(09)	125(05)
C′(1)	−843(11)	4612(08)	−2366(09)	88(05)	N(6)	4850(09)	4484(09)	1423(07)	87(05)
O′(1)	429(07)	4723(04)	−2173(06)	78(03)	$C^{\alpha}(6)$	5389(15)	4179(11)	2494(09)	115(08)
N(2)	−1485(07)	4440(06)	−1528(07)	79(04)	$C^{\beta 1}(6)$	6877(13)	4367(16)	3001(12)	191(15)
$C^{\alpha}(2)$	−754(10)	4478(00)	−383(08)	83(05)	$C^{\beta 2}(6)$	5258(18)	3403(12)	2299(14)	138(09)
$C^{\beta 1}(2)$	−1708(11)	4020(11)	229(01)	112(07)	C′(6)	4454(13)	4323((09)	3342(11)	100(06)
$C^{\beta 2}(2)$	473(14)	5162(08)	36(12)	108(07)	O′(6)	4927(08)	4450(06)	4278(06)	119(04)
C′(2)	622(12)	4025(09)	−275(10)	79(06)	N(7)	3083(10)	4329(06)	2861(07)	96(04)
O′(2)	1654(07)	4195(05)	427(06)	106(04)	$C^{\alpha}(7)$	2069(14)	4417(09)	3601(11)	95(06)
N(3)	553(09)	3452(07)	−838(09)	77(05)	$C^{\beta}(7)$	1163(18)	5057(10)	3156(14)	118(07)
$C^{\alpha}(3)$	1771(14)	2996(10)	−635(12)	107(07)	$C^{\gamma 1}(7)$	2129(29)	5641(11)	3224(30)	274(23)
$C^{\beta}(3)$	1352(23)	2306(10)	−1175(15)	137(10)	$C^{\gamma 2}(7)$	246(20)	5153(13)	4065(16)	186(12)
$C^{\gamma 1}(3)$	1133(29)	2343(17)	−2267(24)	231(11)	C′(7)	1259(16)	3799(11)	3521(14)	112(08)
$C^{\gamma 2}(3)$	2623(25)	1829(18)	−917(28)	308(22)	O′(7)	804(13)	3468(08)	2738(10)	165(06)
C′(3)	3079(12)	3352(08)	−1016(09)	89(05)	O(8)	967(11)	3657(07)	4516(09)	132(05)
O′(3)	4242(08)	3252(06)	−467(07)	115(04)	C(8)	105(20)	3056(12)	4535(18)	171(12)

[a] Equivalent isotropic *U* defined as one-third of the trace of the orthogonaliized U_{ij} tensor.

Table 3. Torsion Angles[a] (deg)

residue	ϕ	φ	ω	χ^1	χ^2
Leu	−48[b]	−38	−165	164	−167, 71
Aib	−55	−38	−172		
Val	−65	−29	177	71, −177	
Ala	−70	−25	−174		
Leu	−108	13	171	−58	−176, −49
Aib	67	36	176		
Val	−115	144[c]	177[d]	−60, −174	

[a] The torsion angles for rotation about bonds of the peptide backbone (ϕ, φ, and ω) and about bonds of the amino acid side chains (χ^1, χ^2) as suggested by the IUPAC−IUB Commission on Biochemical Nomenclature (1970). Estimated standard deviations ~1.0°. [b] C′(0)−N(1)−C^{α}(1)−C′(1). [c] N(7)−C^{α}(7)−C′(7)−O(OMe). [d] C^{α}(7)−C′(7)−O(OMe)−C(OMe).

dimensional NOESY and ROESY spectra were acquired using 1024 points, 512 increments, 64 transients, and a mixing time of 300 ms. All two-dimensional data sets were zero filled to 1024 points with a 90° phase shifted squared sinebell filter in the both dimensions. Coupling constants were calculated from one-dimensional experiments.

Results and Discussions

Crystal Structure. A stereoview of the molecular structure and the residue numbering scheme are illustrated in Figure 1. The molecule forms a 3_{10}-helix encompassing residues 1−5. There is a reversal of the sign of ϕ, ψ values at Aib(6), which adopts an α_L conformation (Table 3). This results in helix termination by the achiral Aib residue and the formation of a 6 → 1 hydrogen bond between Aib(2) CO and Val(7) NH. A feature of the 16-member hydrogen-bonded ring at the C-terminus is that a 4 → 1 hydrogen bond between Aib(6) NH and Val(3) CO group is also accommodated, although the N−O distance is slightly long, 3.36 Å (Table 4). The torsion angles at Leu(5) are significantly distorted from those expected for $3_{10}/\alpha$-helices and lie in the bridge region of the Ramchandran map. These stereochemical features have indeed been observed earlier in helical peptides containing Aib residues at the penultimate position from the C-terminus.[5,9]

Table 4. Hydrogen Bonds

			length (Å)		angle (deg)	
type	donor	acceptor	N−O	H−O	C═O−N	O−HN
intermolecular						
	N(1)[a]	O(6)	2.927	2.123	149	129
	N(2)[b]	O(4)	2.874	2.161	142	121
intramolecular						
4 → 1	N(3)	O(0)	3.183	2.201	134	150
4 → 1	N(4)	O(1)	2.915	2.105	129	129
4 → 1	N(5)	O(2)	3.101	2.181	111	141
4 → 1	N(6)	O(3)	3.365	2.328	106	160
6 → 1	N(7)	O(2)	3.064	2.001	148	167

[a] Symmetry equivalent $-1 + x, y, -1 + z$ to coordinates listed in Table 2. [b] Symmetry equivalent $-1 + x, y, z$ to coordinates listed in Table 2.

A view of the crystal packing is shown in Figure 2. The cylindrical peptide helices are packed in parallel fashion with intermolecular hydrogen bonds formed between Leu(1) and Aib-(2) NH groups and Ala(4) and Aib(6) CO groups of symmetry-related molecules. Interestingly, C-terminal helix distortion leaves the Ala(4) CO group uninvolved in intramolecular hydrogen bonding.

NMR Studies. Conformational characteristics in solution were investigated in two solvents, widely different in their hydrogen-bonding properties, $CDCl_3$ and $(CD_3)_2SO$. In $CDCl_3$, which does not possess any strong hydrogen-bonding groups, folded conformations stabilized by intramolecular hydrogen bonds are favored in peptides. Intrapeptide interactions are expected to prevail over peptide−solvent interactions. In contrast, in $(CD_3)_2SO$, the solvent is capable of acting as a strong hydrogen-bond acceptor and may thus be expected to compete for interaction with peptide NH groups, resulting in destabilization of folded structures. ^1H NMR assignments, summarized in Table 5, were obtained using a combination of 2D COSY and NOESY methods.[13]

Conformations in $CDCl_3$. Delineation of intramolecularly hydrogen-bonded NH groups was carried out by solvent

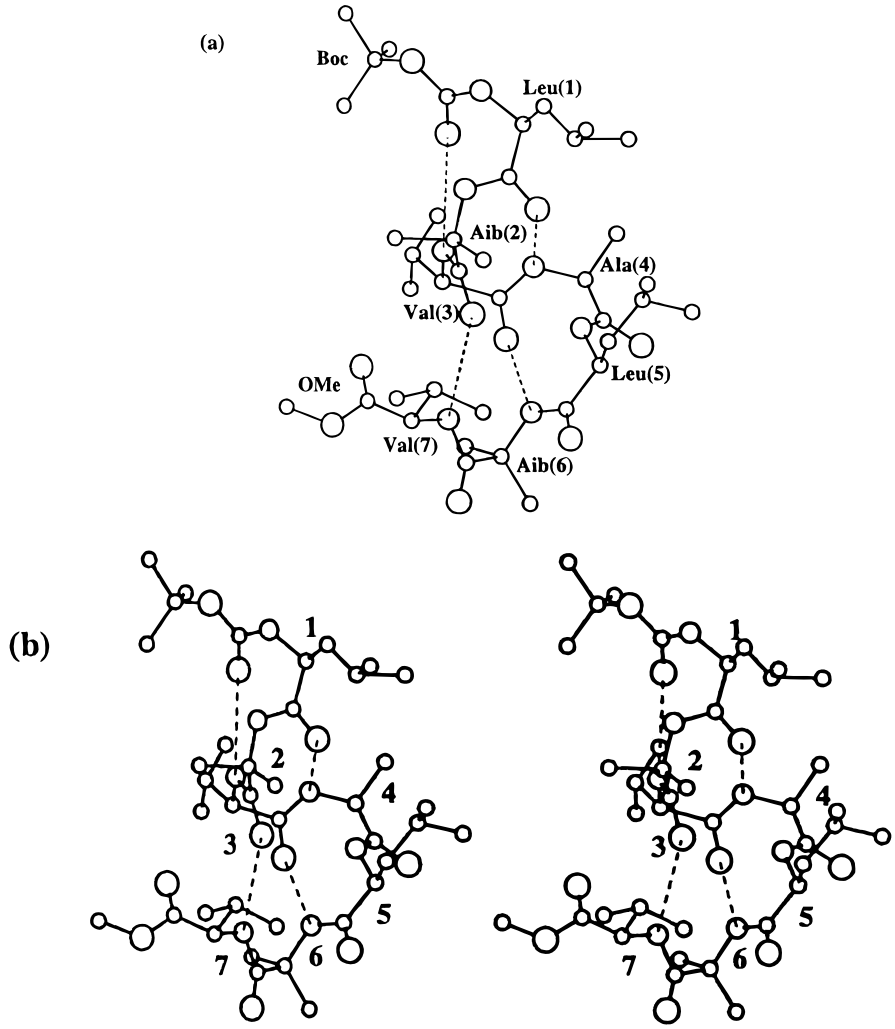

Figure 1. (a) Conformation of the heptapeptide **1** in the crystal. Dashed lines indicate three 4→1 type and one 6→1 type intramolecular hydrogen bonds. (b) Stereoview of the molecule in crystals.

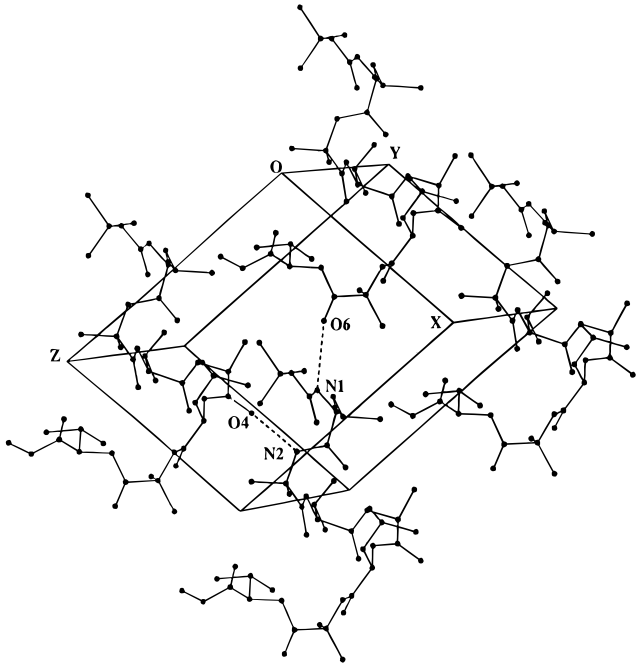

Figure 2. Crystal packing diagram of the heptapeptide **1**. The dashed lines indicate intermolecular hydrogen bonding.

perturbation (Figure 3) in $CDCl_3$, by the addition of small amounts of $(CD_3)_2SO$, which is expected to affect the chemical shift of exposed NH groups. At low concentrations of $(CD_3)_2SO$ structural perturbations are generally not observed, with chemical shift changes arising only as a consequence of solvent hydrogen bonding to exposed NH groups. From the $\Delta\delta$ values in Table 5 it is clear that while Leu(1) and Aib(2) NH groups are exposed, the remaining NH groups of residues 3−7 are solvent shielded, indicative of their involvement in intramolecular hydrogen bonding. These observations are consistent with a 3_{10}-helical conformation in which the NH groups of 3−7 are involved in 4 → 1 hydrogen-bond formation. Indeed, the earlier literature suggests the solvent exposure of only two N-terminal NH groups are indicative of continuous 3_{10}-helical structures.[14] Interestingly, $J_{HNC^{\alpha}H}$ values in Table 5 are very low for residues Leu(1), Val(3), and Ala(4) (<6 Hz), supportive of ϕ values $\approx -60°$, which are necessary for a helical conformation. Surprisingly, the $J_{HNC^{\alpha}H}$ values for Leu(5) is 8.6 Hz, which indicates a large magnitude for ϕ, a feature consistent with crystal structure results which yield a ϕ value of $-108°$. Val(7) which shows a ϕ value of $-115°$ in the crystal structure, yields a $J_{HNC^{\alpha}H}$ value of 8.7 Hz in $CDCl_3$ solution. These results are consistent with the maintenance of the crystal state conformation in solution. Attempts to characterize the conformation further using NOEs were unsuccessful because several critical NOEs in $CDCl_3$ were not observed (Table 5). While

(14) (a) Iqbal, M.; Balaram, P. *J. Am. Chem. Soc.* **1981**, *103*, 5548−5552. (b) Vijaykumar, E. K. S.; Balaram, P. *Biopolymers* **1983**, *22*, 2133−2140. (c) Vijaykumar, E. K. S.; Balaram, P. *Tetrahedron* **1983**, *39*, 2725−2731.

Table 5. Characteristic ^{1}H NMR Parameters for the Heptapeptide **1**

	δ values (ppm)[a]								NOEs		
residues	NH	$C^\alpha H$	$C^\beta H$	$C^\gamma H$	$C^\delta H$	$^3J_{NHC^\alpha H}$ (Hz)[b]	$\Delta\delta$[c]	dδ/dT[d] (ppb/K)	$d_{NN}(i, i+1)$	$d_{\alpha N}(i, i+1)$	$d_{\beta N}(i, i+1)$
Leu(1)	6.91	3.92	1.65	1.58	0.89	5.8	1.30	−6.4	strong	strong	strong
	(5.10)	(3.85)	(1.64)	(1.64)	(0.98)	(2.3)			(*f*)	(*f*)	(*f*)
Aib(2)	8.24	−	1.37			−	1.02	−7.1	medium	−	strong
	(6.75)		(1.40, 1.49)						(*f*)	(−)	(medium)
Val(3)	7.45	3.94	2.08	0.90		5.6	0.01	−5.3	strong	strong	strong
	(7.22)	(3.98)	(2.09)	(0.92)		(4.5)			(strong)	(weak)	(strong)
Ala(4)	7.82	4.13	1.31			6.5	0.12	−3.0	strong	strong	*f*
	(7.69)	(4.23)	(1.5)			(5.6)			(strong)	(*f*)	(strong)
Leu(5)	7.48	4.19	1.56	1.54	0.84	7.4	0.13	−2.4	medium	strong	strong
	(7.18)	(4.48)	(1.63)	(1.62)	(0.88)	(8.6)			(*f*)	(strong)	(*g*)
Aib(6)	7.58	−	1.38			−	0.07	−2.3	strong	−	strong
	(7.16)		(1.53, 1.59)						(*f*)	(−)	(strong)
Val(7)	7.06	4.14	2.00	0.85		8.6	0.06	0.00	−	−	−
	(7.02)	(4.46)	(2.24)	(1.05)		(8.7)			(−)	(−)	(−)

[a] Chemical shifts of proton resonances in $CDCl_3$ and $(CD_3)_2SO$. Values in parentheses correspond to $CDCl_3$. [b] Coupling constants in $CDCl_3$ and $(CD_3)_2SO$. Values in parentheses correspond to $CDCl_3$. [c] $\Delta\delta$ is the chemical shift difference for NH protons in $CDCl_3$ and 18% $(CD_3)_2SO$/$CDCl_3$. [d] dδ/dT is the temperature coefficient of NH chemical shifts in $(CD_3)_2SO$. [e] Some important interresidue NOEs. Values in parentheses correspond to $CDCl_3$. [f] The NOEs which are not observed. [g] The NOE(s) that are not determined due to the overlap of more than one resonance.

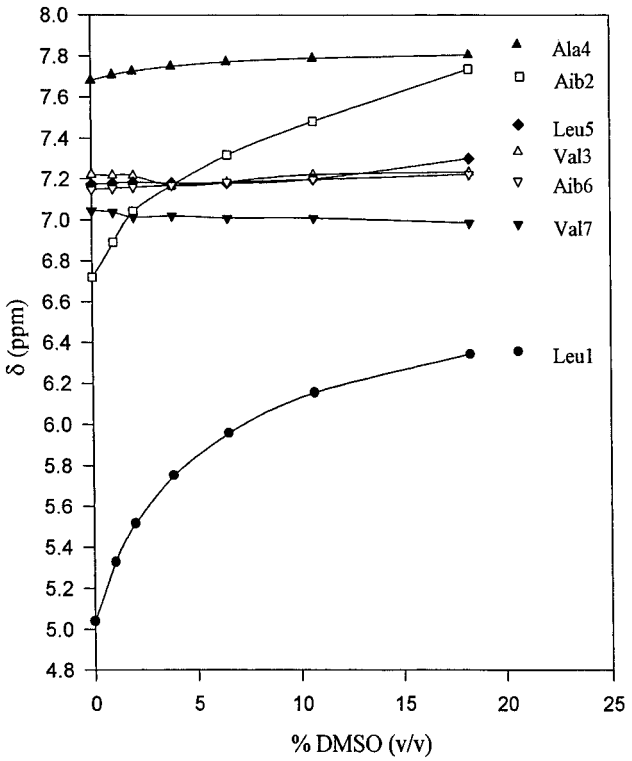

Figure 3. Solvent dependence of the chemical shifts of amide proton resonances of the heptapeptide **1** at varying concentrations of $(CD_3)_2SO$ in $CDCl_3$.

failure to detect NOEs may be rationalized by assuming a population of conformation in which the relevant interproton pairs are >3.5 Å in small peptides, in the present case this explanation seems unlikely. This is because both delineation of hydrogen-bonding NH groups and $J_{HNC^\alpha H}$ values are completely consistent with the helical conformation observed in crystals. The alternative possibility that merits consideration is that in cylindrical peptides anisotropic motion with a preferred axis of rotation coinciding with the helix axis may result in ineffective cross relaxation between pairs of backbone protons. The effects of anisotropic motion in influencing NOE magnitudes have indeed been considered.[15]

(15) (a) Withka, J. M.; Swaminathan, S.; Bolton, P. H. *J. Magn. Reson.* **1990**, *89*, 386−390. (b) Krishnan, V. V. In *Theoretical and Experimental Investigations of Nuclear Overhauser Effect in Nuclear Magnetic Resonance*, Ph D. Thesis, Indian Institute of Science, Bangalore, India, 1991.

Conformations in $(CD_3)_2SO$. In $(CD_3)_2SO$, the temperature coefficients of NH chemical shifts (Table 5) reveal that the NH groups of residues 1−3 are solvent exposed, while the remaining four NH groups have low dδ/dT, suggestive of their involvement in intramolecular hydrogen bonding. $J_{HNC^\alpha H}$ values are also consistent with a helical conformation for residues 1−5. Interestingly, Leu(5) has an appreciably lower $J_{HNC^\alpha H}$ value in $(CD_3)_2SO$ (7.4 Hz) as compared to $CDCl_3$ (8.6 Hz). Furthermore, the ROESY spectrum of the peptide in $(CD_3)_2SO$ (Figure 4) reveals a succession of $d_{NN}(i, i+1)$ connectivities characteristic of a significant population of helical conformations. Taken together with the observation of three solvent-exposed amino terminal NH groups, the NOE data are supportive of a major population of an α-helical conformation for the heptapeptide in $(CD_3)_2SO$. Earlier studies of Aib rich oligopeptides have suggested that while 3_{10}-helical conformations are favored in an apolar solvent like $CDCl_3$, α-helices may be preferentially populated in $(CD_3)_2SO$.[14b,c] The ROESY spectra in Figure 4 also show the presence of strong $C_i^\alpha H \leftrightarrow N_{i+1}H$ connectivities. Indeed, the interresidue $d_{\alpha,N}$ connectivities are stronger than the corresponding intraresidue $d_{\alpha,N}$ NOEs. Simultaneous observation of $d_{NN}(i, i+1)$ and $d_{\alpha N}(i, i+1)$ NOEs is strongly suggestive of conformational averaging, with both α-helical conformations and unfolded forms being populated in $(CD_3)_2SO$. This is consistent with the known tendency of $(CD_3)_2SO$ to invade helix backbones resulting in unfolding.[4c] Examination of the ROESY spectrum in Figure 4 reveals the presence of two weak NOEs, Val(3)$C^\alpha H \leftrightarrow$ Val(7)NH and Val(3)$C^\alpha H \leftrightarrow$ Ala(6)NH. These medium range NOEs correspond to $d_{\alpha,N}(i, i+3)$ and $d_{\alpha,N}(i, i+4)$ interactions. The corresponding interproton distances in the crystal structure are Val(3)$C^\alpha H$ to Val(7)NH = 3.49 Å and Val(3)$C^\alpha H$ to Ala(6) = 3.40 Å. The corresponding distances in idealized 3_{10}- and α-helices are as follows: 3_{10}-helix (−60°, −30°) $d_{\alpha,N}(i, i+3)$ = 3.42 Å, $d_{\alpha,N}(i, i+4)$ = 5.54 Å; α-helix (−57°, −47°) $d_{\alpha,N}(i, i+3)$ = 3.32 Å, $d_{\alpha,N}(i, i+4)$ = 4.18 Å. Thus it is clear that the NOE data are consistent with at least a small population of conformations resembling the crystal state conformation (i.e. α_L-terminated 3_{10}-helix) in $(CD_3)_2SO$. Hydrogen bond delineation results which establish the solvent exposure of the residues 1−3 and the observation of strong $d_{\alpha,N}(i, i+1)$ NOEs suggest that appreciable conformational heterogeneity exists in $(CD_3)_2SO$ solution. While the former favors an α-helical conformation, the latter suggests a significant population of frayed structures with ψ values in the extended region. Comparison of the

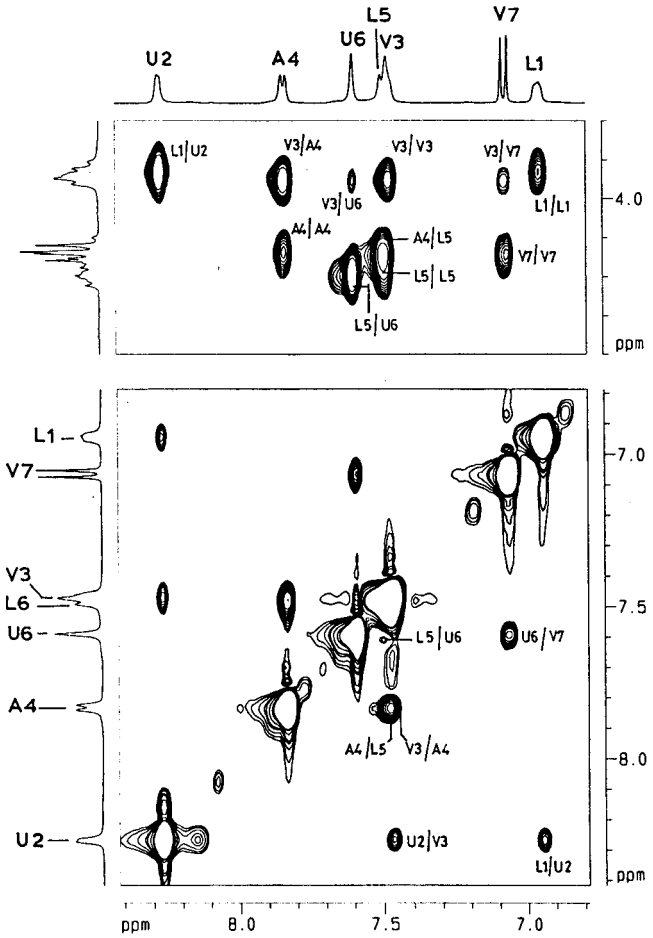

Figure 4. Partial 400 MHz ROESY spectrum of the heptapeptide **1** in $(CD_3)_2SO$ at 307 K. $C^{\alpha}H-NH$ NOEs (Top panel) and amide NH−NH NOEs (bottom panel) are indicated.

conformational energies of minimized structures obtained by using the CVFF force field[16] available in the DISCOVER molecular modeling package in INSIGHT II (BIOSYM, CA) on a Silicon Graphiscs Iris Workstation for ideal 3_{10}- and α-helices and the experimentally determined structures in crystals suggested that all three conformations have comparable energies. Solvation may thus play a major role in determining conformer populations. In the peptide under study, the presence of the Leu-Aib segment raises the possibility of conformational flipping between type I(III) and type II β-turn conformation which can easily be achieved by an approximately 180° flip of the Leu-Aib peptide bond. Such transitions are estimated to be almost barrierless for a concerted process involving correlated rotation of ϕ_{i+1} and ψ_{i+2} (Gomathy, L.; Chandrasekhar, J.; Balaram, P., unpublished results) and have indeed been observed in Leu-Aib sequences.[9] The possibility of such processes may also add the complexity of the conformational mixture in strongly solvating media like $(CD_3)_2SO$, where disruption of the intrahelical hydrogen bonds are compensated by solvent−peptide interactions. In the present study three distinct conformational states appear broadly compatible with experimental data and with known stereochemical preferences of Aib-containing peptides. These are (1) an ideal 3_{10}-helix containing five successive 4→1 hydrogen bonds of the type C═O(i)−NH-(i + 3), (2) an ideal α-helix stabilized by four successive 5→1 hydrogen bonds of the type C═O(i)−NH(i + 4), and (3) A 3_{10}-helical structure terminated by an α_L conformation at Aib-(6) which is indeed the observed conformation in single crystals. The factors controlling 3_{10}- to α-helical transitions have been investigated systematically by Kuki and co-workers using a series of isomeric octapeptides containing as many as six Aib residues.[17] These authors have demonstrated, using NMR studies, that sequence permutations can indeed induce the transition between 3_{10}- and α-helical structures. The solvent exposure of N-terminal NH groups as monitored by temperature and solvent perturbation of NH chemical shifts has been the major indicator of the nature of the helical conformations. Similar approaches have indeed been used in the earlier reports.[14b,c] Extensive crystallographic studies of Aib peptides have revealed that even largely α-helical structures can sometimes contain a single 3_{10}-helical turn at the N-terminus; the formation of the mixed helices containing both 4→1 and 5→1 hydrogen bonds has been frequently observed in crystals.[4a,b] The maintenance of such conformations in solution may further complicate the interpretation of NMR spectroscopic data. The relative ease of disruption of the folded conformations correlates well with the spectroscopic results in $(CD_3)_2SO$, which indicate that solvent competition for hydrogen-bonding sites on the peptide can result in population of nonhelical conformations.

While NMR studies in $CDCl_3$ have been interpreted in terms of a conformation similar to that observed in crystals, it should be noted that a completely 3_{10}-helical conformation is also consistent with the experimentally determined nature of hydrogen-bonded NH groups. The crystal state conformation and the ideal 3_{10}-helix differ primarily in the helix sense (signs of ϕ, ψ values) at Aib(6). In the present study, the high value of $J_{HNC^{\alpha}H}$ at Leu(5) has been used to favor the distortion of the ϕ value of the Leu(5) residues, similar to that observed in the crystal.

Conclusions

Extensive studies of Aib-containing sequences over the last two decades suggest that both 3_{10}- and α-helical structures are energetically accessible with subtle factors determining the precise conformational preference.[2] Even in crystals variants of helical conformations have sometimes been observed in polymorphic forms or among multiple molecules in an asymmetric unit.[3,4] Conformational heterogeneity in solution must therefore be an important factor complicating the interpretation of NMR spectroscopic data. When the Aib content of the sequence is high, i.e. >50% with residues being uniformly distributed, the conformational flexibility is considerably limited, permitting characterization of well-defined conformational states, even in polar solvents. In peptides with lower Aib content, spectroscopic interpretations may be less definitive. The present attempt to correlate crystallographic and spectroscopic studies on a designed peptide containing Aib at positions 2 and 6 suggests that several closely related helical conformations, which differ in a subtle manner, are indeed accessible in short Aib-containing peptides. While substantial earlier discussion has centered on 3_{10}- and α-helical structures,[2−4,7] the present study together with earlier crystal structure determinations[5,6d] suggests that " α_L-terminated helices" are yet another structural type when the achiral Aib residue is placed at the penultimate position from the C-terminus.

A further point of interest that merits consideration is the relationship between the structure determination in crystals and the conformation in solutions. In single crystals, most commonly, a unique conformational state is characterized. Less frequently, when multiple molecules are present in the asymmetric unit, more than one closely related conformation may

(16) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, V. A.; Wolff, J.; Genest,M.; Hagler, A. T. *Proteins: Sturct. Funct. Genet.* **1988**, *4*, 31−47.

(17) (a) Basu, G.; Bagchi, K; Kuki, A. *Biopolymers* **1991**, *31*, 1763−1744. (b) Basu, G.; Kuki, A. *Biopolymers* **1992**, *32*, 61−71.

be determined. From the very large body of Aib-containing crystal structures it has become apparent that there appears to be no easily interpretable relationship between molecular packing and peptide backbone conformations.[3,4] It appears reasonable to suggest that the process of nucleation and crystallization selects from an ensemble of conformational states that are present in solution. If kinetic factors determine the process of crystal formation, even minor conformations which are not readily detectable in spectroscopic experiments may in fact be trapped in the crystalline state. Small energy differences (on the order of ≈2 kcal/mol) can effectively render the high-energy conformers undetectable in spectroscopic experiments. Rapid equilibration in solution between conformational states can however facilitate the trapping of the minor conformation in crystals. Crystal growth may be determined by the effectiveness with which a specific peptide conformation is accommodated within the requirements of crystal symmetry. Crystallization then selects in a Darwinian manner the conformation which is the best suited for "adaptation" to the crystalline environment. Packing forces may serve as an element of selective pressure rather than as a means of distorting molecular conformation. Indeed, in favorable situations, conformational transitions have been demonstrated upon dissolution of peptide single crystals at low temperature using NMR spectroscopy.[18] The tendency of many oligopeptides to crystallize in several polymorphic forms may permit structural characterization of several conformational variants,[3,4] although control over the formation of polymorphs may remain elusive.[19]

Acknowledgment. The use of the AMX-400 NMR spectrometer at the Sophisticated Instruments Facility, Indian Institute of Science is gratefully acknowledged. A.P. was supported by the award of a Research Associateship from the Council of Scientific and Industrial Research, India.

Supporting Information Available: Tables of atomic coordinates, bond lengths, bond angles, anisotropic temperature factors, and hydrogen atom coordinates for peptide **1** (9 pages). See any current masthead page for ordering and Internet access instructions.

JA960665U

(18) Balaram, H.; Prasad, B. V. V.; Balaram, P. *J. Am. Chem. Soc.* **1983**, *105*, 4065−4071.
(19) Dunitz, J. D.; Bernstein, J. *Acc. Chem. Res.* **1995**, *28*, 193−200.