Peptide design. Helix–helix motifs in synthetic sequences

Arindam Banerjee,[a] S. Raghothama[b] and P. Balaram*,[a]

[a] *Molecular Biophysics Unit and* [b] *Sophisticated Instruments Facility, Indian Institute of Science, Bangalore-560012, India*

Two centrally positioned α-aminoisobutyryl (Aib) residues have been used to stabilize distinct heptapeptide helical segments in the 15-residue synthetic sequence Boc-Met-Ala-Leu-Aib-Val-Ala-Leu-*Acp*-Val-Ala-Leu-Aib-Val-Ala-Phe-OMe. The helices are connected by the flexible linker ε-aminocaproic acid (*Acp*). NMR studies in $CDCl_3$ establish helical conformations for both independent segments as evidenced by NH–NH nuclear Overhauser effects (NOEs). The peptide strongly aggregates in $CDCl_3$ with the NH groups of Met(1) and Ala(2) participating in intermolecular hydrogen bonds. In $(CD_3)_2SO$ two solvated helical segments are supported by NMR results. Solvent dependent breakdown of aggregates on addition of $(CD_3)_2SO$ to $CDCl_3$ solutions is suggested by analysis of chemical shifts and temperature coefficients of NH protons. The observation of several interhelical NOEs in $CDCl_3$, relatively few NOEs in 10% $(CD_3)_2SO$–$CDCl_3$ and their absence in $(CD_3)_2SO$ provides a means of inferring helix orientations. While an antiparallel arrangement resulting in closed aggregate formation is suggested in $CDCl_3$, a parallel solvated arrangement is favoured in $(CD_3)_2SO$.

Introduction

The success of *de novo* protein design rests on the ability to construct peptide sequences with predictable folding patterns which then assemble to form compact tertiary structures.[1,2] In aqueous solutions, the driving forces for maximizing tertiary interactions are hydrophobic in nature. The construction of amphipathic secondary structures then provides a convenient means of achieving compact polypeptide conformations with a largely hydrophobic core.[3,4] The secondary structural elements are themselves constructed using well-established principles of helix stabilization in short peptides[5–7] and the less widely investigated strategies for β-hairpins.[8,9] The difficulties of precisely designing well packed cores has led to many designed proteins which have molten globule like characteristics.[10–13] The use of metal ion and organic templates to organize structured peptide segments has proved promising.[14–20]

An alternative approach to synthetic polypeptide design which is being investigated in this laboratory uses well-defined secondary structural modules which are soluble in organic solvents. The use of α-aminoisobutyric acid (Aib) and related α,α-dialkylated amino acids to stabilize helices up to 16 residues has been demonstrated by crystallographic and spectroscopic studies.[21–24] The subsequent assembly of apolar peptide helices into helix–linker–helix motifs has also been reported.[25] In this approach peptide chain folding is directed by the use of a limited number of conformationally constrained amino acids. Subsequent interactions of secondary structural elements are determined by the stereochemical properties of the linking segments, hydrogen bonding and side chain packing interactions. The solubility of these systems in relatively inert solvents like chloroform and methanol permits folding processes to be studied in the absence of a strong hydrophobic driving force.

The crystal structure determination of the helix–linker–helix peptide (Boc-Val-Ala-Leu-Aib-Val-Ala-Leu-Acp-Val-Ala-Leu-Aib-Val-Ala-Leu-OMe) reveals that the two helices make an angle of approximately 180°, with the two cylindrical modules being staggered about the connecting ε-aminocaproic (Acp) segment (Fig. 1).[25] This arrangement in crystals is stabilized by a continuous chain of intermolecular 'helix-like' hydrogen bonds involving the NH groups of one molecule and the central CO groups of a neighboring molecule, which constitute the end of the N-terminal segment and the beginning of the C-terminal segment, respectively. Intramolecular folding would have been expected to yield a U-shaped structure in which the two helices are roughly antiparallel, with an interhelical angle close to 0°. In order to address the question of conformational variability about the linking Acp residue, we have carried out detailed NMR investigations of the 15 residue peptide sequence Boc-**Met**-Ala-Leu-Aib-Val-Ala-Leu-*Acp*-Val-Ala-Leu-Aib-Val-Ala-**Phe**-OMe (**1**) in organic solvents [$CDCl_3$ and $(CD_3)_2SO$], which establish solvent dependent changes in structures and aggregation states. The peptide sequence was chosen such that both helical modules (1–7 and 9–15) have centrally positioned Aib residues. The use of Met and Phe as the N- and C-terminal residues respectively, facilitates NMR assignments without changing the apolar character of the peptide and affords the possibility of observing assignable NOEs, diagnostic of tertiary interactions.

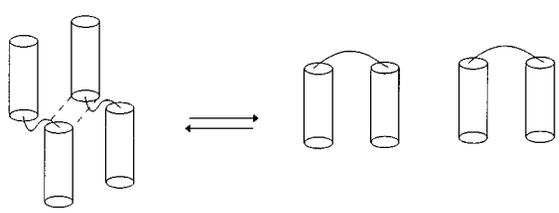

Fig. 1 Interconversions between antiparallel, close packed helical cylinders and open extended forms are illustrated. Two molecules are schematically shown to indicate the nature of intramolecular 'helix like' hydrogen bonds observed in crystals of the peptide Boc-Val-Ala-Leu-Aib-Val-Ala-Leu-*Acp*-Val-Ala-Leu-Aib-Val-Ala-Leu-OMe.[25]

Experimental

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

* E-Mail: pb@mbu.iisc.ernet.in

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-Met-Ala-Leu-Aib-OMe 2. To 1.65 g (4.11 mmol) of Boc-Ala-Leu-Aib-OMe,[26] 10 ml of 98% formic acid was added and the removal of Boc group 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 hydrogen carbonate 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-Met-OH (0.93 g, 3.64 mmol) in 10 ml of dimethylformamide (DMF), followed by 0.74 g (3.7 mmol) of DCC and 0.49 g (3.65 mmol) of HOBT. The reaction mixture was stirred for 3 d under nitrogen atmosphere. The residue was taken in ethyl acetate (60 ml) and dicyclohexylurea (DCU) was filtered off. The organic layer was washed with 2 M HCl (3 × 50 ml), 1M sodium carbonate (3 × 50 ml) and brine (2 × 50 ml), dried over sodium sulfate and evaporated *in vacuo* to yield 1.36 g (70%) of a gummy material. δ_H(400 MHz, $CDCl_3$) 0.90–0.93 (Leu $C^\delta H$), 1.43 (Ala $C^\beta H$), 1.46 (Boc CH_3), 1.50 and 1.53 (Aib $C^\beta H_3$), 1.59 (Leu $C^\beta H$ and $C^\gamma H$), 1.93 (Met $C^\beta H$), 2.12 (Met $C^\varepsilon H$), 2.60 (Met $C^\gamma H$), 3.71 (OCH_3), 4.19 (Ala $C^\alpha H$), 4.38–4.42 (Met $C^\alpha H$ and Leu $C^\alpha H$), 5.27 (Met NH), 6.67 (Ala NH and Leu NH), 6.77 (Aib NH).

Boc-Met-Ala-Leu-Aib-OH 3. To 1.26 g (2.37 mmol) of **2** in 15 ml of methanol, 3 ml of 2 M NaOH were added and the reaction was monitored by TLC. On completion of the reaction after 10 h methanol was evaporated, the residue was taken in water and washed with diethyl ether the aqueous layer was acidified with cold 2 M HCl to pH 2 and extracted with ethyl acetate (3 × 30 ml). The organic extracts were pooled, dried over anhydrous sodium sulfate and evaporated *in vacuo* to yield 1.2 g (70%) of a white solid. δ_H(400 MHz, $CDCl_3$) 0.88–0.92 (Leu $C^\delta H$), 1.34 (Ala $C^\beta H$), 1.44 (Boc CH_3), 1.50, 1.54 (Aib $C^\beta H_3$), 1.63 (Leu $C^\beta H$, $C^\gamma H$), 1.93 (Met $C^\beta H$), 2.09 (Met $C^\varepsilon H$), 2.54 (Met $C^\gamma H$), 4.34 (1H, Ala $C^\alpha H$), 4.52–4.56 (2H, Met $C^\alpha H$, Leu $C^\alpha H$), 5.45 (1H, Met NH), 7.23 (1H, Aib NH), 7.52 (1H), 7.93 (2H) (Ala NH, Leu NH).

Boc-Met-Ala-Leu-Aib-Val-Ala-Leu-OMe 4. Boc-Val-Ala-Leu-OMe[26] (1.1 g, 2.65 mmol) was deprotected with 98% formic acid and worked up as reported in the preparation of **2**. This was coupled with 1.2 g (2.25 mmol) of **3** in 15 ml of DMF using 0.48 g (2.40 mmol) of DCC and 0.30 g (2.25 mmol) of HOBT under nitrogen atmosphere. After 3 d the reaction was worked up as described for **3** to yield the crude peptide as a solid (1.17 g, 1.44 mmol). The peptide was purified on a reverse phase C-18 MPLC column using methanol–water mixtures. δ_H(400 MHz, $CDCl_3$) 0.90 (Leu $C^\delta H$), 0.95 (Leu $C^\delta H$), 1.01 (Val $C^\gamma H$), 1.45 (Ala $C^\beta H$), 1.49 (Aib $C^\beta H$), (Ala $C^\beta H$), 1.58 (Boc CH_3), 1.72 (Leu $C^\beta H$, $C^\gamma H$), 1.95 (Val $C^\beta H$), 2.14 (Met $C^\varepsilon H$), 2.50 (Met $C^\beta H$), 2.66 (Met $C^\gamma H$), 3.70 (OCH_3), 4.10 (2H), 4.16 (1H), 4.23 (2H), 4.56 (1H) (Ala, Leu, Met, Val $C^\alpha H$), 5.67 (1H, Met NH), 6.71 (1H), 6.76 (1H), 7.20 (1H), 7.22 (1H), 7.59 (1H) (Ala, Leu, Val NH), 7.33 (1H, Aib NH).

Boc-Met-Ala-Leu-Aib-Val-Ala-Leu-Acp-OMe 5. The peptide **4** (1.16 g) was saponified using 15 ml of methanol and 2 ml of 2 M NaOH under nitrogen atmosphere and worked up as reported in the preparation of **3**. This was coupled to H-Acp-OMe that was isolated from 0.48 g (2.64 mmol) of the corresponding hydrochloride using DCC (0.56 g, 2.80 mmol) and HOBT (0.36 g, 2.67 mmol) under nitrogen atmosphere. The reaction was worked up after 2 d as described for **2** to yield 0.88 g (75%) of a white solid. δ_H(400 MHz $CDCl_3$) 0.89 (Leu $C^\delta H$), 0.95 (Leu $C^\delta H$), 1.04 (Val $C^\gamma H$), 1.47 (Ala $C^\beta H$), 1.49 (Aib $C^\beta H$, Ala $C^\beta H$), 1.58 (Boc CH_3), 1.64–1.70 (Leu $C^\beta H$ and $C^\gamma H$, Acp $C^\delta H$ and $C^\gamma H$), 1.98 (Val $C^\beta H$), 2.12 (Met $C^\varepsilon H$), 2.25 (Met $C^\beta H$), 2.30 (Acp $C^\alpha H$), 2.66 (Met $C^\gamma H$), 3.18 (Acp $C^\varepsilon H$), 3.28 (Acp $C^\varepsilon H$), 3.65 (OCH_3), 3.80 (1H, Val $C^\alpha H$), 4.07–4.11 (3H, Met $C^\alpha H$, Ala $C^\alpha H$, Leu $C^\alpha H$), 4.24 (1H, Ala $C^\alpha H$), 4.40 (1H, Leu $C^\alpha H$), 6.32 (1H, Met NH), 7.13 (1H), 7.15 (1H), 7.31 (1H), 7.34 (1H), 7.58 (1H) (Ala, Leu, Val NH), 7.22 (1H, Acp NH), 7.37 (Aib NH).

Boc-Val-Ala-Phe-OMe 6. Boc-Val-Ala-OH[26] (2.82 g, 10 mmol) was taken in 15 ml of DMF and coupled to H-Phe-OMe isolated from 5.30 g (24.5 mmol) of the corresponding methyl ester hydrochloride using DCC (2.08 g, 10.4 mmol) and HOBT (1.35 g, 10 mmol). The reaction was worked up after 3 d as described for **2** to yield a white solid (2.08 g, 5.7 mmol, 57%). δ_H(400 MHz $CDCl_3$) 0.90 (Val $C^\gamma H$), 1.34 (Ala $C^\beta H$), 1.44 (Boc CH_3), 2.11 (Val $C^\beta H$), 3.10 (Phe $C^\beta H$), 3.70 (OCH_3), 3.91 (Val $C^\alpha H$), 4.46 (Phe $C^\alpha H$), 4.82 (Ala $C^\alpha H$), 5.05 (Val NH), 6.56 (Ala NH), 6.56 (Phe NH), 7.09 (Phe ring H), 7.28 (Phe ring H).

Boc-Val-Ala-Leu-Aib-Val-Ala-Phe-OMe 7. The peptide **6** (1.08 g, 2.34 mmol) was deprotected with 98% formic acid and worked up as reported in the preparation of **2**. This was coupled to 0.97 g (2 mmol) of Boc-Val-Ala-Leu-Aib-OH[26] using DCC (0.42 g, 2.1 mmol) and HOBT (0.27 g, 2 mmol). The reaction was worked up after 3 d as described for **2** to yield the crude peptide as a white solid (0.88 g, 1.06 mmol). The peptide was purified on a reverse phase C-18 MPLC column using methanol–water mixtures. δ_H(400 MHz, $CDCl_3$) 0.89 (Leu $C^\delta H$), 0.95 (Leu $C^\delta H$), 1.04 (Val $C^\gamma H$), 1.47 (Ala $C^\beta H$), 1.49 (Aib $C^\beta H$), (Ala $C^\beta H$), 1.58 (Boc CH_3), 1.64–1.70 (Leu $C^\beta H$ and $C^\gamma H$, Acp $C^\delta H$s and $C^\gamma H$), 1.98 (Val $C^\beta H$), 2.12 (Met $C^\varepsilon H$), 2.25 (Met $C^\beta H$), 2.30 (Acp $C^\alpha H$), 2.66 (Met $C^\gamma H$), 3.18 (Acp $C^\varepsilon H$), 3.28 (Acp $C^\varepsilon H$), 3.65 (OCH_3), 3.80 (1H, Val $C^\alpha H$), 4.07–4.11 (3H, Met $C^\alpha H$, Ala $C^\alpha H$, Leu $C^\alpha H$), 4.24 (1H, Ala $C^\alpha H$), 4.40 (1H, Leu $C^\alpha H$), 6.32 (1H, Met NH), 7.13 (1H), 7.15 (1H), 7.31 (1H), 7.34 (1H), 7.58 (1H) (Ala, Leu, Val NH), 7.22 (1H), Acp NH), 7.37 (Aib NH).

Boc-Met-Ala-Leu-Aib-Val-Ala-Leu-Acp-Val-Ala-Leu-Aib-Val-Ala-Phe-OMe 1. The peptide **5** (0.87 g, 0.94 mmol) was saponified using 15 ml of methanol and 2 ml of 2 M NaOH under nitrogen atmosphere and worked up as described for **3**. This was coupled to H-Val-Ala-Leu-Aib-Val-Ala-Phe-OMe [isolated from **8** (0.80 g, 0.96 mmol) as described in the preparation of **2**] using 0.11 g of DCC and 0.2 g of HOBT under nitrogen atmosphere. After 3 d the reaction was worked up as usual to yield 0.83 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 Lichrosorb reverse phase C-18 HPLC column (4 × 250 mm, particle size 10 μm, flow rate 1.5 ml min^{-1}) and eluted on a linear gradient of methanol–water (70–95%) with a retention time of 23 min. Mp 109–110 °C. The peptide was homogeneous on a reverse phase C-18 (5 μm) column and fully characterized by NMR (see results).

Spectroscopic studies

All NMR studies were carried out on a Bruker AMX-400 spectrometer at the Sophisticated Instruments Facility. All two-dimensional NMR experiments were carried out using a peptide concentration of 3.7 mM. Concentration dependent 1D NMR experiments were performed over the range 6–0.05 mM and the probe temperature was maintained at 298 K. Resonance assignments were made using two-dimensional double quantum filtered COSY spectra (1K data points, 512 experiments, 48 transients, spectral width 4500 Hz). Two-dimensional NOESY and ROESY spectra were acquired using 1024 points, 512 increments, 64 transients and mixing time 300 ms. All two-dimensional data sets were zero filled to 1024 points with a 90° phase shifted squared sinebell filter in both the dimensions. Coupling constants were calculated from one-dimensional experiments.

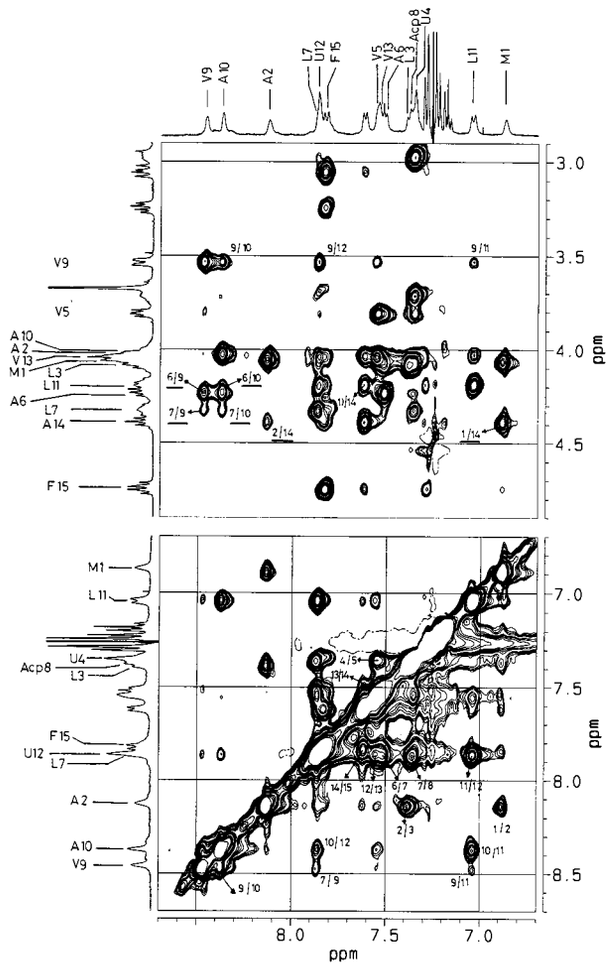

Fig. 2 Partial 400 MHz ^{1}H–^{1}H NOESY spectrum of the peptide **1** in $CDCl_3$. C^αH–NH NOEs (top panel) and amide NH–NH NOEs (bottom) panel are indicated. Specific resonance assignments are indicated on the 1D spectra using one letter code. Several $d_{\alpha N}$ (top) and d_{NN} (botttom) are also labelled. Inter-helix NOEs are underlined (top panel).

CD spectra were recorded on a JASCO J-500 spectropolarimeter equipped with a DP-501N data processor using 1 mm path length cuvettes. Scan speeds were generally kept at 5 nm min^{-1} and the time constant at 8 or 16 s. Spectra were averaged over several scans (typically 4–8). The baseline was obtained under similar conditions as the sample. All spectra are baseline subtracted, converted to a uniform scale of molar ellipticity and replotted. Temperature was kept at 20 °C. The sample concentration used was 181 μM.

Results and discussion

NMR studies of peptide **1** were carried out in $CDCl_3$, $CDCl_3$–$(CD_3)_2SO$ (90:10) and $(CD_3)_2SO$. Complete assignment of the ^{1}H NMR spectrum in $CDCl_3$ and $(CD_3)_2SO$ was performed using a combination of COSY and NOESY/ROESY. The representative NOESY spectra shown in Fig. 2 illustrate observed inter-residue connectivities which permit sequence specific assignments. Chemical shifts for the various resonances are summarized in the Table 1. The temperature coefficients of the chemical shifts of the various NH resonances ($d\delta/dT$) and the $^3J_{HNC^\alpha H}$, vicinal coupling constants in the three solvent systems are listed in Table 2. The observed NOEs in the three solvent systems are summarized in Fig. 3.

Conformational analysis in chloroform

In $CDCl_3$, a succession of strong $d_{NN(i,i+1)}$ ($N_iH \longleftrightarrow N_{i+1}H$) NOE connectivities were observed over the segments 1–7 and 9–15. The corresponding $d_{\alpha N}$ ($C^\alpha_iH \longleftrightarrow N_{i+1}H$) NOEs are weak or not observed (Fig. 2 and 3). These observations support the formation of two helical segments that are separated by the flanking *Acp* residue.

The effect of adding a strong hydrogen bond accepting solvent like $(CD_3)_2SO$ to $CDCl_3$ solutions of the peptide **1** on the amide NH chemical shifts is summarized in Fig. 4. Normally, addition of small amounts of $(CD_3)_2SO$ results in monotonic downfield shifts of exposed NH groups in peptides, leaving solvent shielded NH groups largely unaffected. This permits delineation of intramolecularly hydrogen bonded NH groups in peptides using solvent titration experiments which have found

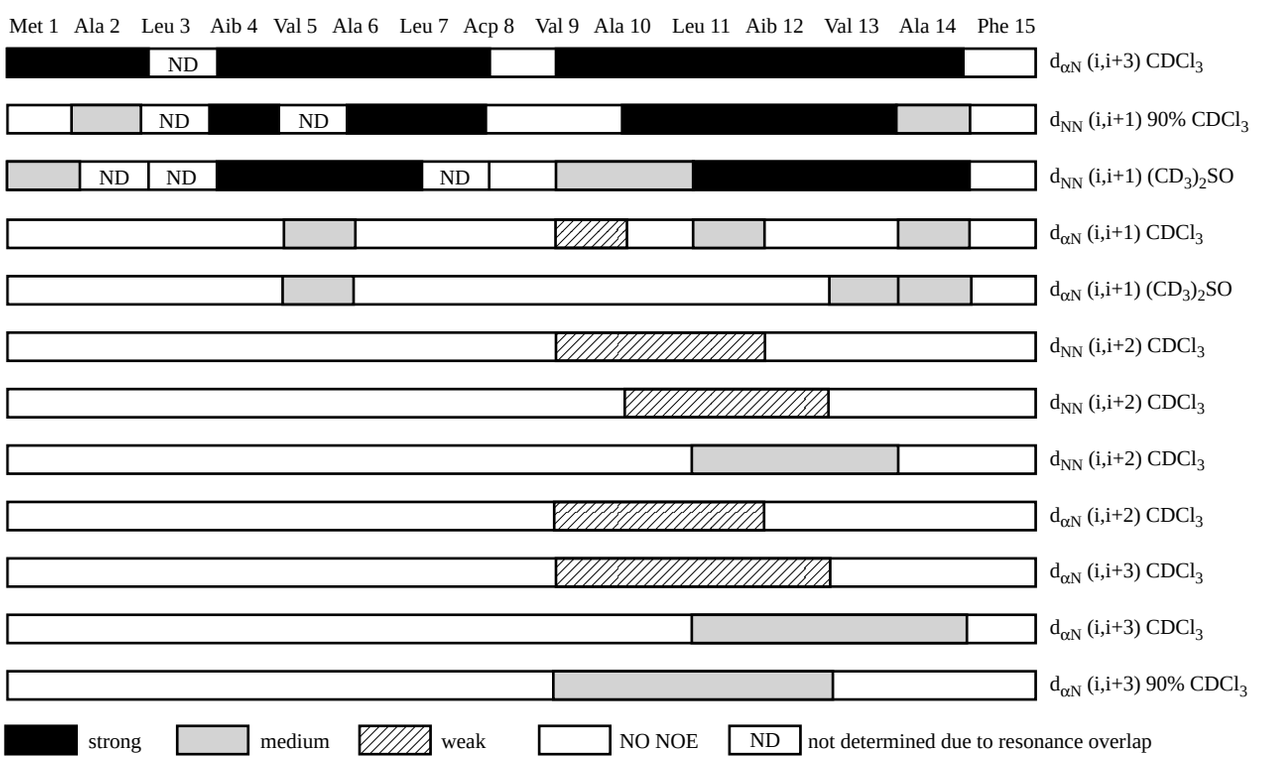

Fig. 3 A summary of inter-residue NOEs for the peptide **1** in three solvent systems

Table 1 Characteristic ^{1}H NMR parameters for the peptide **1**

		δ values[a]				
Residues	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	Others
Met1	6.87	4.07	1.98	2.61		ε-CH_3,
	(7.09)	(3.98)	(2.46)	(1.81)		2.07
						(2.01)
Ala2	8.13	4.04	1.48			
	(8.04)	(4.26)	(1.49)			
Leu3	7.38	4.09	1.72	1.68	0.90, 0.94	
	(7.99)	(4.16)	(1.61)	(1.61)	(0.84)	
Aib4	7.35	—	1.56			
	(8.14)	(—)	(1.49)			
Val5	7.55	3.81	2.18	1.07, 1.08		
	(7.29)	(3.88)	(1.93)	(0.87)		
Ala6	7.51	4.23	1.53			
	(7.80)	(4.20)	(1.46)			
Leu7	7.87	4.33	1.76	1.70	0.84, 0.89	
	(7.47)	(4.17)	(1.58)	(1.58)	(0.86)	
Acp8	7.36	2.24, 2.60	1.62, 1.76	1.62, 1.70	1.62	ε-CH_2
	(7.43)	(2.15)	(1.47)	(1.24)	(1.40)	2.96, 3.71
						(3.02)
Val9	8.46	3.53	2.21	1.09, 0.99		
	(7.87)	(4.07)	(2.15)	(0.83)		
Ala10	8.37	4.03	1.45			
	(8.07)	(4.22)	(1.44)			
Leu11	7.05	4.20	1.79	1.69	0.95, 0.98	
	(7.82)	(4.12)	(1.60)	(1.60)	(0.85)	
Aib12	7.86	—	1.52			
	(8.02)	(—)	(1.48)			
Val13	7.54	4.05	2.27	1.02, 1.03		
	(7.06)	(4.06)	(2.02)	(0.78)		
Ala14	7.62	4.40	1.39			
	(7.77)	(4.24)	(1.46)			
Phe15	7.83	4.75	3.05, 3.25			aromatic H
	(8.13)	(4.44)	(2.91, 3.01)			7.18–7.31
						(7.19–7.30)

[a]Chemical shifts of proton resonances in $CDCl_3$ and, in parentheses, $(CD_3)_2SO$.

Table 2 ^{1}H HNMR parameters for NH groups in peptide **1**

	$d\delta/dT$[a] (ppb K^{-1})			$^3J_{HNC^{\alpha}H}$[b]/Hz		
Residue	$CDCl_3$	$CDCl_3$–$(CD_3)_2SO$ (9:1)	$(CD_3)_2SO$	$CDCl_3$	$CDCl_3$–$(CD_3)_2SO$ (9:1)	$(CD_3)_2SO$
Met1	15.9	9.0	6.7	<2	<2	ND
Ala2	18.4	7.9	5.1	<2	<2	5.0
Leu3	3.7	2.2	4.3	ND	ND	6.2
Aib4	−0.2	−0.6	5.6	—	—	—
Val5	9.2	5.5	3.4	ND	ND	<2
Ala6	1.0	0.8	4.0	5.7	6.2	4.0
Leu7	6.4	5.8	3.6	ND	7.8	9.0
Acp8	2.4	4.9	3.4	4.0	ND	4.5
Val9	8.4	8.1	4.5	<2	3.0	8.1
Ala10	6.7	7.1	5.4	<2	2.9	6.9
Leu11	−2.3	−1.1	4.7	7.6	7.6	6.5
Aib12	3.3	3.9	5.1	—	—	—
Val13	8.1	5.8	1.6	ND	6.3	<2
Ala14	1.2	1.9	4.3	7.1	7.7	ND
Phe15	6.1	3.7	6.3	7.1	8.2	<2

[a] Temperature coefficients in $CDCl_3$, 90% $CDCl_3$–10% $(CD_3)_2SO$ and $(CD_3)_2SO$. [b] Coupling constants in $CDCl_3$, 90% $CDCl_3$–10% $(CD_3)_2SO$ and $(CD_3)_2SO$. ND = not determined due to resonance overlap.

widespread use in the literature.[27,28] The results in Fig. 4 are striking, with several NH groups showing anomalous behaviour. Met1 NH and Ala2 NH which are expected to be solvent exposed in the N-terminal helix move upfield up to a $(CD_3)_2SO$ concentration of *ca.* 5%; at higher $(CD_3)_2SO$ concentration both these NH groups show monotonic downfield shifts with saturation being observed at a $(CD_3)_2SO$ concentration of *ca.* 40%. In a 3_{10}-helical structure the NH groups of the first two residues from the N-terminus are exposed to the solvent. In most of the examples of helical peptides studied so far, addition of $(CD_3)_2SO$ to $CDCl_3$ solutions of the peptide results in a steady downfield shift of the NH resonances for the residues 1 and 2. The strikingly unusual behavior of the Met1 and Ala2 NH groups of peptide **1** is suggestive of peptide aggregation at the concentrations (3.7 mM) used in the NMR experiments. Indeed, helical peptides can associate in apolar solvents by hydrogen bond formation between exposed amino terminal NH groups and carboxy terminal CO groups. This kind of aggregation results in infinite head-to-tail hydrogen bonded arrays in crystals.[29–31]

The variation of NH chemical shifts as a function of peptide concentration over the range 6–0.05 mM in $CDCl_3$ is summar-

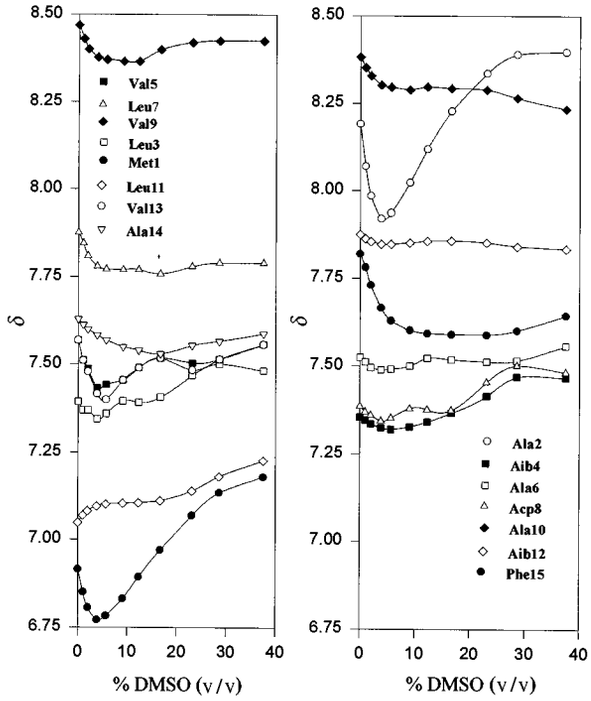

Fig. 4 Solvent dependence of chemical shifts of amide proton resonances of peptide **1** at varying concentrations of $(CD_3)_2SO$ in $CDCl_3$

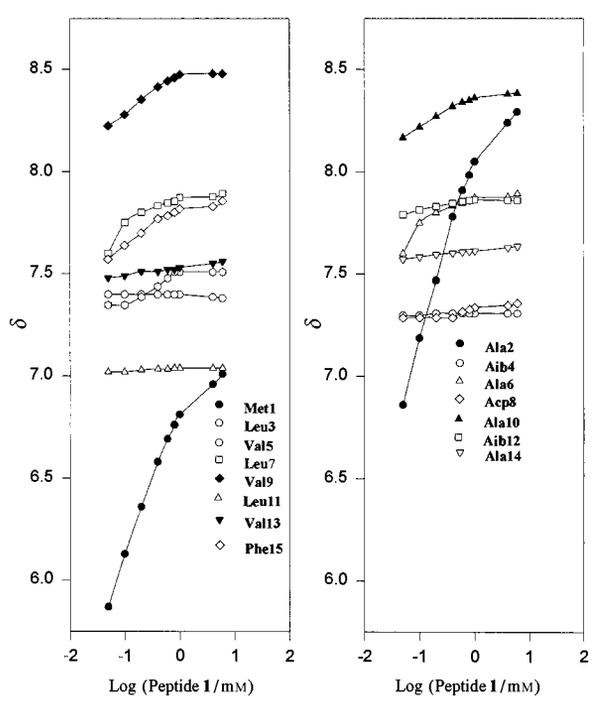

Fig. 5 Concentration dependence of NH chemical shifts of peptide **1** in $CDCl_3$

ized in Fig. 5. The most dramatic concentration dependencies are observed for the Met1 and Ala2 NH groups. This clearly supports their involvement in intermolecular hydrogen bonding. The temperature coefficients ($d\delta/dT$) listed in Table 2 reveal that the largest temperature dependencies are observed for Met1 and Ala2 NH groups. In an apolar, nonhydrogen bonding solvent like $CDCl_3$ large $d\delta/dT$ values have generally been attributed to NH groups involved in intermolecular hydrogen bonds which are readily broken up upon heating. Low $d\delta/dT$ values may arise from strongly intramolecularly hydrogen bonded NH groups or from solvent exposed NH groups. (Note that while distinction between intramolecular hydrogen bonded and free NH groups are possible in strongly hydrogen bonding

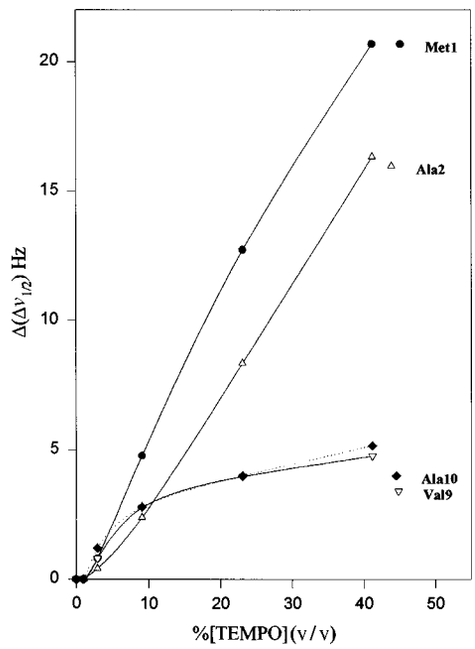

Fig. 6 Free radical (TEMPO) induced broadening of amide protons of peptide **1** in $CDCl_3$. Relative broadening of only four amide protons are indicated

Table 3 Observed long range NOEs for the peptide **1** in two solvent systems[a]

$CDCl_3$	
Phe15 $C^{\alpha}H \longleftrightarrow$ Met1 $C^{\beta}H$ (m)	Ala6 $C^{\alpha}H \longleftrightarrow$ Val9 NH (m)
Ala14 $C^{\alpha}H \longleftrightarrow$ Met1 NH (s)	Leu7 $C^{\alpha}H \longleftrightarrow$ Val9 NH (w)
Met1 $C^{\beta}H \longleftrightarrow$ Phe15 ringH (m)	Leu7 $C^{\alpha}H \longleftrightarrow$ Ala10 NH (w)
Ala14 $C^{\beta}H \longleftrightarrow$ Met1 NH (m)	Leu7 $C^{\alpha}H \longleftrightarrow$ Ala10 $C^{\beta}H$ (w)
Ala14 $C^{\beta}H \longleftrightarrow$ Ala2 $C^{\beta}H$ (m)	Val9 $C^{\alpha}H \longleftrightarrow$ Ala6 $C^{\beta}H$ (s)
Ala14 $C^{\alpha}H \longleftrightarrow$ Ala2 NH (m)	Ala6 $C^{\alpha}H \longleftrightarrow$ Val9 $C^{\beta}H$ (s)
Leu7 $C^{\alpha}H \longleftrightarrow$ Val9 $C^{\beta}H$ (m)	Leu7 $C^{\alpha}H \longleftrightarrow$ Val9 $C^{\gamma}H$ (s)
Ala6 $C^{\alpha}H \longleftrightarrow$ Ala10 NH (m)	
90% $CDCl_3$–10% $(CD_3)_2SO$	
Ala6 $C^{\alpha}H \longleftrightarrow$ Ala10 NH (w)	Ala6 $C^{\alpha}H \longleftrightarrow$ Val9 NH (w)
Leu7 $C^{\alpha}H \longleftrightarrow$ Val9 NH (w)	

[a] s = strong, m = medium, w = weak.

solvents, no such distinction is readily made in $CDCl_3$). NH groups of intermediate $d\delta/dT$ values may arise due to the formation of weak inter- or intra-molecular hydrogen bonds.[32–34]

The NMR results suggest that the peptide is composed of two independent helical segments as evident from NOE data. The molecule aggregates strongly in $CDCl_3$ *via* intermolecular hydrogen bonds involving Met1 and Ala2 NH groups. The size of the aggregate is likely to be small (possibly dimers), because the NMR spectra reveal relatively sharp resonances for various side chain and backbone (particularly $C^{\alpha}H$) resonances. The nature of the aggregate will be considered subsequently.

If peptide **1** contains two independent helical segments, NH groups of residues 1,2 and 9,10 should not be involved in intramolecular hydrogen bonds making them available for intermolecular interactions as exemplified in the crystal structures of a related 15 residue peptide.[25] The solvent and concentration dependence of NH chemical shifts shown in Figs. 4 and 5 suggest that the behavior of Met1, Ala2 and Val9, Ala10 NH resonances are distinctly different. The chemical shift changes observed for Val9 and Ala10 NH groups are much smaller than those observed for Metl and Ala2 NH protons. The relative solvent accessibilities of these four NH groups were further probed using the paramagnetic nitroxide 2,2,6,6-tetramethylpiperidin-1-oxyl (TEMPO) as a perturbant. Fig. 6 shows addition of free radical results in much larger broadening for Met1

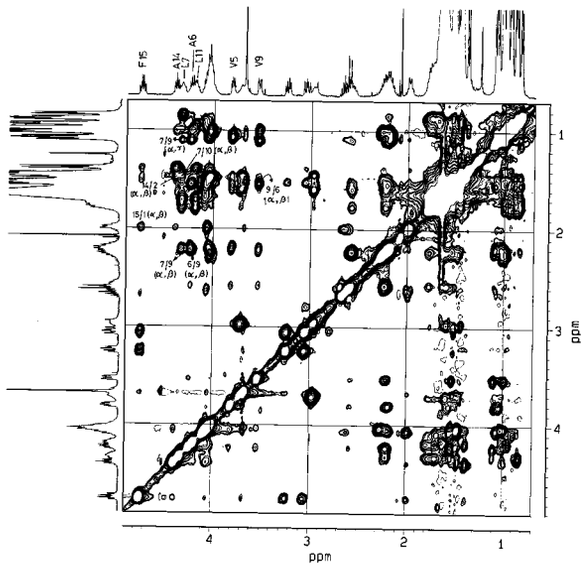

Fig. 7 Partial 400 MHz ^{1}H–^{1}H NOESY spectrum of the peptide **1** in $CDCl_3$. Some observed long range $C^{\alpha}H \longleftrightarrow C^{\beta}H/C^{\alpha}H \longleftrightarrow C^{\gamma}H$ NOE connectivities are indicated.

and Ala2 NH resonances as compared to Val9 and Ala10 NH resonances. The remaining NH resonances in peptide **1** are largely unaffected over the range of radical concentrations studied, confirming their shielded nature. This result strongly suggests that the degree of solvent exposure for Met1 and Ala2 is much greater than the Val9 and Ala10 NH groups.

Inspection of the NOESY spectra of peptide **1** in $CDCl_3$ revealed several long range NOEs, summarized in Table 3. The observed NOEs between residues 7 and residues 9,10 and residue 6 and residues 9,10 may be interpreted in terms of an approximate antiparallel orientation of the two helical segments about the central *Acp* hinge. Interestingly, NOEs are observed between protons of residue 1 and 14 and 15 and residues 2 and 14 (see also Fig. 7). While it is tempting to assign the long range NOEs between the C- and N-terminal end of the molecule as arising through folded structures which brings the two extremities of the molecule in close proximity, the possibility of these NOEs arising due to intermolecular interactions cannot be ruled out. In order to probe solvent effect on peptide aggregation and structure NMR studies were carried out in $(CD_3)_2SO$ and in a $(CD_3)_2SO$–$CDCl_3$ (10:90) mixture.

Conformation in $(CD_3)_2SO$

Several strong d_{NN} connectivities could be assigned in NOESY spectra of the peptide in $(CD_3)_2SO$ (Fig. 3). Some sequential connectivities could not be detected due to resonance overlap. The observed intraresidue $d_{\alpha N}$ ($N_iH \longleftrightarrow C^{\alpha}_iH$) NOEs are significantly more intense relative to d_{NN} NOEs as compared to $CDCl_3$. Further all the inter-residue $d_{\alpha N}$ NOEs are very weak or unobserved. These results favor helical conformations for the 1–7 and 9–15 segments. Interestingly, no long range NOEs could be detected in $(CD_3)_2SO$ suggesting the absence of conformation which involves antiparallel orientation of the two helices. The temperature coefficients ($d\delta/dT$) observed in $(CD_3)_2SO$ are all relatively high with the exception of Val13 NH. This suggests that most of the peptide NH groups of the peptide **1** are reasonably accessible in $(CD_3)_2SO$, although the overall helical folding pattern is supported by the NOE data. Fraying of peptide helices and consequent solvent invasion of the backbone are likely to be favored in a strong hydrogen bond donating solvent like $(CD_3)_2SO$. Indeed, crystallographic and NMR evidence for fragility of short peptide helices in $(CD_3)_2SO$ has been presented.[35] A comparison of the $^3J_{HNC^{\alpha}H}$ values in $CDCl_3$ and $(CD_3)_2SO$ (Table 2) establishes significantly larger values in the latter, indicative of substantial deviation of the dihedral angle φ from an ideal helical φ angle.

Conformations in $CDCl_3$–$(CD_3)_2SO$ (90:10)

While the NMR data in $CDCl_3$ favors aggregated structures involving antiparallel intramolecular helix orientation, the data in $(CD_3)_2SO$ favor a more extended arrangement of the two highly solvated helical segments. An examination of the chemical shift dependence of the NH resonance in $CDCl_3$–$(CD_3)_2SO$ mixtures suggests that aggregates formed by intermolecular hydrogen bonding involving Met1 and Ala2 NH groups are broken at $(CD_3)_2SO$ concentration > 5%. This is undoubtedly due to the ability of $(CD_3)_2SO$ to complete for hydrogen bonding to NH groups. The anomalous behavior of Met1 and Ala2 NH groups in solvent titration experiments provides a convenient means of studying this disaggregation upon addition of $(CD_3)_sSO$. The NH chemical shifts over the range 5–10% $(CD_3)_2SO$ lead to the expectation that while the hydrogen bonded aggregate may be disrupted, intramolecular organization might still be maintained. ROESY spectra in $(CD_3)_2SO$–$CDCl_3$ (10:90) show several strong d_{NN} connectivities in both the helical segments with relatively weak inter-residue $d_{\alpha N}$ NOEs (Fig. 3). The observed $^3J_{HNC^{\alpha}H}$ values are also very close to those determined in $CDCl_3$ suggesting that helical conformations are indeed maintained for residues 1–7 and 9–15. The temperature coefficients ($d\delta/dT$), Table 2, in this solvent mixture are particularly noteworthy. Four NH resonances, Met1, Ala2, Val9 and Ala10 have $d\delta/dT > 7$ ppb K^{-1} which is significantly greater than the values observed for the remaining NH groups. This observation argues for the high degree of exposure to the solvent of the NH groups of the residues 1,2,9 and 10. Of the remaining residues, Val5 NH, Leu7 NH and Val13 NH have $d\delta/dT$ values of 5.5–5.8 ppb K^{-1} suggestive of a moderate degree of solvent exposure. This may be indicative of preferential interaction of $(CD_3)_2SO$ at helix sites susceptible to solvation. In the mixed solvent system only three long range NOEs (Table 3) could be detected. These are Ala6 $C^{\alpha}H \longleftrightarrow$ Ala10 NH, Ala6 $C^{\alpha}H \longleftrightarrow$ Val9 NH and Leu7 $C^{\alpha}H \longleftrightarrow$ Val9 NH, suggesting that the two helical segments are orientated in approximately antiparallel fashion, while the limited number of long range NOEs precludes the very small values of interhelical angle. The data are consistent with a folded geometry about the *Acp* hinge. The disappearance of several long range NOEs in $CDCl_3$ upon the addition of 10% $(CD_3)_2SO$ established that interaction between protons 1/2 and 14/15 are probably a consequence of aggregation.

The NMR results establish that broadly helical conformations are maintained for the segments 1–7 and 9–15 in $CDCl_3$, $(CD_3)_2SO$–$CDCl_3$ (10:90) and $(CD_3)_2SO$. The orientation of the helices about the central flexible *Acp* hinge is however sensitive to solvent composition as established by the observation of NOEs between residues 6/7 and 9/10 in $CDCl_3$ and $(CD_3)_2SO$–$CDCl_3$, (10:90) whereas no interhelical interactions are detected in $(CD_3)_2SO$. The peptide is strongly aggregated in $CDCl_3$ with the addition of a small amount of $(CD_3)_2SO$ resulting in dissociation, a feature inferred from the solvent dependence of NH chemical shifts. This type of self-assembly formation in relatively apolar non-hydrogen-bonding solvents has been reported in the literature and the nature of the solvent plays a crucial role for noncovalent association,[36] which is entropically disfavored but enthalpically favored due to the formation of intermolecular hydrogen bonds. The fact that sharp well-resolved resonances are observed at high concentrations of $CDCl_3$ favors the formation of very small aggregates with dimers being an attractive possibility. A schematic model to rationalize the observed NMR data is presented in Fig. 8.

The formation of closed dimeric aggregates formed by hydrogen bonding between exposed NH groups of the N-terminal helical segments and the exposed CO groups of the C-terminal helical segments is postulated to account for the strong

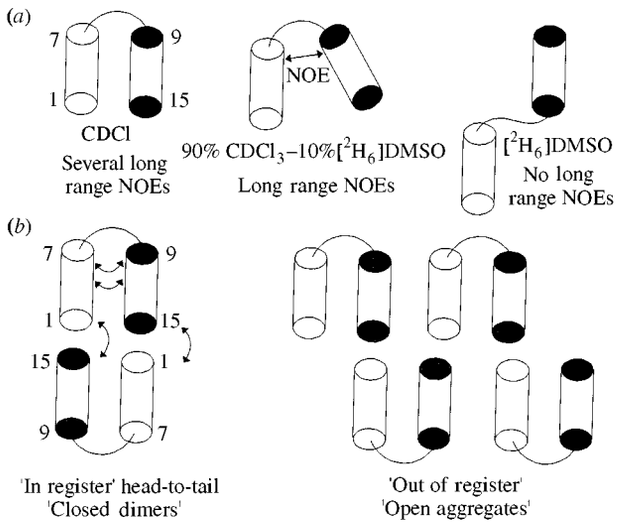

Fig. 8 Schematic representations of (*a*) solvent dependent structural changes of the peptide **1**; (*b*) possible aggregation patterns for peptide **1** in $CDCl_3$

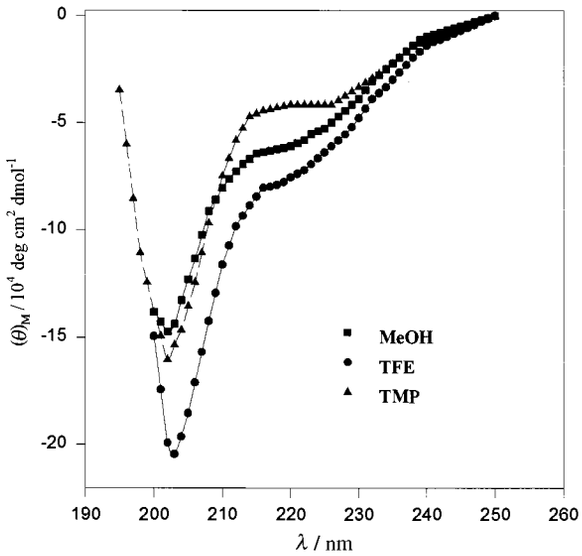

Fig. 9 A comparison of CD spectra of the peptide **1** in methanol, TMP (trimethyl phosphate) and TFE (2,2,2-trifluoroethanol)

concentration dependence of NH chemical shifts of Met1 NH and Ala2 NH resonances. In this model Val9 and Ala 10 NH groups are solvent exposed, but do not participate appreciably in intermolecular interactions. Indeed, concentration dependence of NH chemical shifts for these residues are substantially lower than that observed for Met1 and Ala2 NH protons. An alternative aggregated species which may be considered is the open structure, illustrated in Fig. 8. This may be ruled out since such species would be expected to propagate to large sizes, a requirement that is not supported by the observed narrow NMR resonances. The effects of addition of a strong hydrogen bonding solvent like $(CD_3)_2SO$ are also rationalized in Fig. 8. While small amounts of $(CD_3)_2SO$ successfully disrupt intermolecular hydrogen bonding, the peptide retains an overall folded super-secondary structure, consistent with the observation of a limited number of interhelical NOEs in $(CD_3)_2SO$–$CDCl_3$ (10:90). In pure $(CD_3)_2SO$ solutions, there is no evidence for compact arrangements of helices and it is likely that conformations similar to that observed in the related peptide Boc-Val-Ala-Leu-Aib-Val-Ala-Leu-*Acp*-Val-Ala-Leu-Aib-Val-Ala-Leu-OMe (**UV15**) are obtained. In that case, crystal structure determination reveals an approximate 180° interhelical angle.[25]

Circular dichroism

The CD spectra of the peptide **1** in methanol, TMP and TFE are shown in Fig. 9. The peptide **1** displays CD spectra which are identical in nature but of varying intensities in the three solvents. The long wavelength component of the π–π* band appears at 203 nm in TFE and at 202 nm in both methanol and TMP. The n–π* transition appears as a shoulder at 220 nm in all the three different solvents. The spectra resemble those observed for well-characterized helical peptides of similar length.[25,37] In TFE, both π–π* and n–π* bands are more intense indicating greater helicity in that solvent. The lack of any dramatic solvent dependence of the CD spectra suggests that the overall backbone conformation of the two independent helical modules is retained in solvent media of widely different polarity and hydrogen bonding ability. Detailed conclusions regarding the nature of the peptide helix (3_{10} or α) and the extent of the helicity are fraught with uncertainty, because of the problem of subtle conformational heterogeneity in solution[38–40] and the difficulty of providing definitive conformational interpretation of relative intensities of the far UV CD bands in short peptide helices.[41–43]

Conclusions

The present study establishes that antiparallel conformations may be populated for a designed helix–linker–helix motif Boc-**Met**-Ala-Leu-Aib-Val-Ala-Leu-*Acp*-Val-Ala-Leu-Aib-Val-Ala-**Phe**-OMe in an apolar solvent like $CDCl_3$. Peptide association through head-to-tail hydrogen bonding may in fact stabilize an antiparallel orientation. Strongly hydrogen bonding solvents like $(CD_3)_2SO$ not only disrupt peptide aggregation but also favor formation of noncompact structures in which the helical segments extend away from one another, probably in a manner similar to that in crystals of a related peptide **UV15**. In the solid state the formation of intermolecular hydrogen bonds between all CO and NH groups which do not participate in intramolecular hydrogen bond formation is likely to be the driving force for the extended (open) orientation. Such open structures may also be facilitated in $(CD_3)_2SO$, by the ability of the solvent to interact strongly with the exposed NH groups. The present study demonstrates that designed synthetic peptides can mimic helical hairpin structures under defined conditions. The use of a conformationally mobile linking segment results in the complete absence of stereochemical control at the hinge region. The use of conformationally constrained residues in segments linking secondary structural elements is presently being explored in order to permit greater orientational control in the designed super secondary structural motifs. The present study establishes that compact polypeptide structures possessing a limited level of tertiary interactions can indeed be obtained in designed synthetic peptides in nonaqueous media. Further elaboration to larger systems may be of relevance in studies aimed at mimicking helical hairpins in membrane proteins.

References

1 S. F. Betz, D. P. Raleigh and W. F. DeGrado, *Curr. Opin. Struct. Biol.*, 1993, **3**, 601.
2 J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. O'Neil and W. F. DeGrado, *Science*, 1995, **270**, 935.
3 S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik and M. H. Hecht, *Science*, 1993, **262**, 1680.
4 M. Munson, R. O'Brien, J. M. Sturtevant and L. Regan, *Protein Sci.*, 1994, **3**, 2363.
5 K. R. Shoemaker, P. S. Kim, E. J. York, J. M. Stewart and R. L. Baldwin, *Nature*, 1987, **326**, 536.
6 S. Padmanabhan and R. L. Baldwin, *Protein Sci.*, 1994, **3**, 1992.
7 A. J. Doig and R. L. Baldwin, *Protein Sci.*, 1995, **4**, 1325.
8 F. J. Blanco, G. Ravis and L. Serrano, *Nature Struct. Biol.*, 1994, **1**, 584.

9 M. S. Searle, D. H. Williams and L. C. Packman, *Nature Struct. Biol.*, 1995, **2**, 999.
10 M. Beauregard, K. Goraj, V. Goffin, K. Haremans, E. Goormaghtigh, J.-M. Ruysschaert and J. A. Martial, *Protein Eng.*, 1991, **4**, 745.
11 D. P. Raleigh and W. F. DeGrado, *J. Am. Chem. Soc.*, 1992, **114**, 10079.
12 T. Tanaka, H. Kimura, M. Hayashi, Y. Fujiyoshi, K. Fuhuhara and H. Nakamura, *Protein Sci.*, 1994, **3**, 419.
13 T. Tanaka, Y. Kuroda, H. Kimura, S. Kidokoro and H. Nakamura, *Protein Eng.*, 1994, **7**, 969.
14 M. R. Ghadiri, C. Soares and C. Choi, *J. Am. Chem. Soc.*, 1992, **114**, 825.
15 J. M. Berg, *Curr. Opin. Struct. Biol.*, 1993, **3**, 585.
16 J. M. Berg, *Acc. Chem. Res.*, 1995, **28**, 14.
17 M. Klemba, K. H. Gardner, S. Marino, N. D. Clarke and L. Regan, *Nature Struct. Biol.*, 1995, **2**, 368.
18 M. Mutter and S. Vuillemier, *Angew. Chem., Int. Ed. Engl.*, 1989, **28**, 535.
19 M. Mutter, G. G. Tuchscherer, C. Miller K.-H. Altman, R. I. Carey, D. F. Wyss, A. M. Labhardt and J. E. Rivier, *J. Am. Chem. Soc.*, 1992, **114**, 1463.
20 G. Tuchsherer and M. Mutter, *J. Peptide Science*, 1995, **1**, 3.
21 P. Balaram, *Pure Appl. Chem.*, 1992, **64**, 1061.
22 P. Balaram, *Curr. Opin. Struct. Biol.*, 1992, **2**, 845.
23 I. L. Karle, R. B. Rao, S. Prasad, R. Kaul and P. Balaram, *J. Am. Chem. Soc.*, 1994, **116**, 10355.
24 I. L. Karle, R. Gurunath, S. Prasad, R. Kaul, R. B. Rao and P. Balaram, *J. Am. Chem. Soc.*, 1995, **117**, 9632.
25 I. L. Karle, J. L. Flippen-Anderson, M. Sukumar, K. Uma and P. Balaram, *J. Am. Chem. Soc.*, 1991, **113**, 3952.
26 K. Uma, *Modular design of synthetic protein mimics. Construction of helices*. PhD Thesis, Indian Institute of Science, Bangalore, 1990.
27 T. P. Pitner and D. W. Urry, *J. Am. Chem. Soc.*, 1972, **94**, 1399.
28 P. A. Raj and P. Balaram, *Biopolymers*, 1985, **24**, 1131.
29 I. L. Karle, *Acta Crystallogr., Sect. B*, 1992, **48**, 341.
30 I. K. Karle, J. L. Flippen-Anderson, M. Sukumar and P. Balaram, *J. Med. Chem.*, 1992, **35**, 3885.
31 I. L. Karle, J. L. Flippen-Anderson, M. Sukumar and P. Balaram, *Protein. Struct. Funct. Genet.*, 1992, **12**, 324.
32 E. S. Stevens, N. Sugawara, G. M. Bonora and C. Toniolo, *J. Am. Chem. Soc.*, 1980, **102**, 7048.
33 M. Iqbal and P. Balaram, *Biopolymers*, 1982, **21**, 1427.
34 C. Toniolo, G. M. Bonora, V. Bavoso, E. Benedetti, B. DiBlasio, P. Gramaldi, F. Lelj, V. Pavone and C. Pedone, *Macromolecules*, 1985, **18**, 895.
35 I. L. Karle, J. L. Flippen-Anderson, K. Uma and P. Balaram, *Biopolymers*, 1993, **33**, 827.
36 G. M. Whitesides, E. E. Simanek, J. P. Mathias, C. T. Seto, D. N. Chin, M. Mammen and D. M. Gordon, *Acc. Chem. Res.*, 1995, **28**, 37.
37 I. L. Karle, J. L. Flippen-Anderson, K. Uma and P. Balaram, *Current Science*, 1990, **59**, 875.
38 A. Banerjee, S. Dutta, A. Pramanik, N. Shamala and P. Balaram, *J. Am. Chem. Soc.*, 1996, **118**, 9477.
39 S. M. Miick, G. V. Martinez, W. R. Fiori, A. P. Todd and G. L. Millhauser, *Nature*, 1992, **359**, 653.
40 J. Tirado-Rives, D. L. Maxwell and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1993, **115**, 11590.
41 T. S. Sudha, E. K. S. Vijaykumar and P. Balaram, *Int. J. Peptide Protein Res.*, 1983, **22**, 464.
42 C. Toniolo, A. Polese, F. Formaggio, M. Crisma and J. Kamphuis, *J. Am. Chem. Soc.*, 1996, **118**, 2744.
43 N. H. Andersen, Z. Liu and K. S. Prickett, *FEBS Lett.*, 1996, **399**, 47.

Paper 7/00624A
Received 28*th January* 1997
Accepted 13*th May* 1997