Eur. J. Biochem. **264**, 468–478 (1999) © FEBS 1999

CD and NMR investigations on trifluoroethanol-induced step-wise folding of helical segment from scorpion neurotoxin

Purnima Khandelwal, Subhendu Seth and Ramakrishna V. Hosur

Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai, India

A 14 amino acid residue peptide from the helical region of Scorpion neurotoxin has been structurally characterized using CD and NMR spectroscopy in different solvent conditions. 2,2,2-Trifluoroethanol (TFE) titration has been carried out in 11 steps from 0 to 90% TFE and the gradual stabilization of the conformation to form predominantly α-helix covering all of the 14 residues has been studied by ^{1}H and ^{13}C NMR spectroscopy. Detailed information such as coupling constants, chemical shift indices, NOESY peak intensities and amide proton temperature coefficients at each TFE concentration has been extracted and analysed to derive the step-wise preferential stabilization of the helical segments along the length of the peptide. It was found that there is a finite amount of the helical conformation in the middle residues 5–11 even at low TFE concentrations. It was also observed that > 75% TFE (v/v) is required for the propagation of the helix to the N and C termini and for correct packing of the side chains of all of the residues. These observations are significant to understanding the folding of this segment in the protein and may throw light on the inherent preferences and side chain interactions in the formation of the helix in the peptide.

Keywords: circular dichroism; NMR spectroscopy; protein folding; scorpion neurotoxin; trifluoroethanol.

Protein folding has been the subject of vigorous research activity during the last decade with the availability of sophisticated multidimensional NMR techniques that have enabled determination of high resolution structures of peptides and proteins under a variety of experimental conditions. These in turn have provided the basic knowledge with which to understand the relationships between the primary amino acid sequences, secondary structures, tertiary interactions, and effects of local hydrophobic or hydrophilic environments in proteins [1,2]. Several attempts have been made in the literature to identify intermediates at different stages of folding in the refolding process of denatured proteins [3]. Though these have provided valuable information on the folding pathways, one of the difficulties encountered is that it is never possible to be sure whether the protein is completely unfolded to start with. There have been several reports of residual structures even in extreme denaturing conditions such as 8 M urea or guanidinium hydrochloride [4,5]. In this context working with chosen peptides has an advantage in that no great conformational preferences may be expected in aqueous solutions and therefore propensity vs. induced secondary structure can be studied with greater reliability. One of the approaches used here has been to see whether or not a particular peptide segment of a large protein adopts the same native-like conformation in isolation, in aqueous solutions. Three different kinds of situations have been encountered. (a) The peptide exhibits a potential to form a native-like helical conformation in aqueous solution, and this potential is enhanced by solvents such as trifluoroalcohol (TFE) which provide a hydrophobic surface in the vicinity of the peptide [6–9]. These peptide sequences presumably have high propensity to form helical structures and the fluoroalcohols provide additional stability by removing the water molecules from the surroundings of the peptide; intra-molecular H-bond formation in the helix and H-bond formation between water and the peptide backbone are competing factors. (b) Peptides do not adopt native-like conformations, no matter what the solvent conditions are [10,11]. In these cases, the tertiary interactions in the protein play important roles and the chosen solvent conditions do not simulate those conditions *in vitro*. β-Sheet peptides belong largely to this category. (c) Some peptides adopt non-native-like conformations in the presence of fluoroalcohols [12–14]. For example, a peptide segment that exists as a β-turn in a protein may exhibit a helical conformation in the presence of TFE. In these peptides the amino acid sequence has a high propensity for helix formation, but the tertiary environment in the protein has a major influence in dictating the secondary structure of the peptide. Studies of this type will help in the identification of the initiation sites in the folding process of a given protein and also throw light on the roles of the side chains in dictating secondary and tertiary interactions.

With this view, we describe here a systematic investigation of TFE-induced step-wise helix formation in a short 14 residue peptide, belonging to a scorpion neurotoxin protein. The protein consists of a single polypeptide chain of 65 amino acid residues cross-linked by four disulfide bridges. The secondary structure is composed of 2.5 turns of α-helix and a three-strand stretch of antiparallel β-sheet, several loops projecting from this core, and a surface hydrophobic patch [15]. The peptide with the sequence Ac-ENEGADTEA-KAKNQ-NH_2 derives from the helical region of the scorpion neurotoxin (SNT) protein in which the two cysteine residues

Correspondence to R. V. Hosur, Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India. Fax: + 91 22 215 2110, Tel.: + 91 22 215 2971, E-mail: hosur@tifr.res.in.

Abbreviations: TFE, 2,2,2-trifluoroethanol; ROESY, rotating frame Overhauser effect spectroscopy; DQF-COSY, double quantum filtered correlation spectroscopy; HSQC heteronuclear single quantum correlation; SNT, scorpion neurotoxin.

Note: web site available at URL: http://http://www.tifr.res.in:80/~chp

(Received 18 January 1999, revised 24 May 1999, accepted 11 June 1999)

have been replaced with alanines (Ala5 and Ala9) to eliminate any potential complications due to disulfide formation. The helix formation at different percentage composition of TFE/water mixtures has been monitored by CD and NMR methods and the latter have also been used to determine the details of the helix structure in 90% TFE, when helix formation is complete.

MATERIALS AND METHODS

Peptide synthesis

The peptide was synthesized by Fmoc-solid phase method on an Applied Biosystem 431A peptide synthesizer. The crude peptide was purified by preparative reversed phase HPLC on a Vydac C18 column using 0.1% trifluoroacetic acid and acetonitrile. The resulting product appeared as a single peak on analytical RP-HPLC. Amino acid analysis and MS (FAB) were used to check the composition of the peptide.

CD measurements

CD spectra were recorded on a Jasco J-600 spectropolarimeter. The instrument was calibrated using (+)-10-camphorsulfonic acid. All spectra were recorded at room temperature using 0.1 cm path-length cells. Equal amounts of peptide solutions from a stock were dispensed into TFE (Sigma)/water mixtures (0–90% TFE, v/v). Following peptide addition, the mixture was stirred and incubated at room temperature for 10 min before the spectrum was recorded. The final concentration of the peptide in each of the mixtures was between 50 and 60 μM. The spectra were an average of 32 scans recorded at a speed of 50 nm·min^{-1}, with a band width of 2.0 nm at 0.1 nm step size and a 2-s time constant. The spectra were smoothed and the data converted into ASCII format using the JASCO software. Baseline correction, conversion to mean residue ellipticity and subsequent plotting were performed using the SIGMAPLOT software (Jandel Scientific v3.1). Estimation of percentage helicity was made using the formula given by Farood *et al*. [16].

NMR spectroscopy

For TFE titration, two samples were prepared with peptide concentrations of 5.0–5.5 mM. In one, ≈ 4 mg peptide was dissolved in 0.5 mL of 9 : 1 v/v water/D_2O mixture; in the other, ≈ 4 mg peptide was dissolved in 0.5 mL of 9 : 1 v/v deuterated TFE (ISOTEC Inc.)/water mixture. These two served as the stocks for the peptide in 0% and 90% TFE solutions, respectively. Peptide dissolved in intermediate concentrations of TFE/water mixture was obtained by intermixing these two solutions in different proportions; for example, a volume of 0.05 mL was removed first from both 0% and 90% solutions and then mixed again in an interchanged fashion. Thus peptide solutions of 9% and 81% were obtained. This was done in order to minimize the amount of chemicals used and at the same time the same concentration of peptide is maintained. Using the above procedure, samples with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81 and 90% TFE concentration were prepared. The pH of all the above samples was 2.8–3.5. Sodium 3-(trimethylsilyl)propionate-2,2,3,3-d_4 was added for internal referencing of ^{1}H chemical shifts.

The ^{1}H NMR spectra were recorded on Bruker AMX 500 MHz and Varian Unity*plus* 600 MHz spectrometers, with time proportional phase incrementation (TPPI) [17] for quadrature detection in the F1 dimension in the former and States–Haberkorn [18] in the latter. Standard presaturation techniques were used for solvent suppression. TOCSY [19], rotating frame Overhauser effect spectroscopy (ROESY) [20] and NOESY [21] experiments were carried out with mixing times of 75 ms, 250 ms and (250 ms and 400 ms), respectively, for spin system and sequential assignments [22]. The receiver gain was kept equal for NOESY experiments at all the TFE concentrations for direct comparison of peak intensities. Double quantum filtered correlation spectroscopy (DQF-COSY) [23] spectra were acquired with 4096 data points in the F2 dimension and 600 increments with 16 scans to obtain enough resolution to measure coupling constants. This experiment was also used to assign the Thr7 α-proton. To calculate temperature coefficients of amide protons, TOCSY spectra were recorded at a series of temperatures from 288 to 313 K in steps of 5 K with 400 t_1 increments and four scans.

The spectra were processed using UXNMR (Bruker), VNMR (Varian) and Felix v.97.0 (Molecular Simulations Inc.) softwares. For two-dimensional experiments, the F2 dimension was zero-filled to 2048 real data points, F1 dimension was zero-filled to 1024 data points and π/2 phase-shifted sine squared window function was applied along both dimensions. Baseline corrections were applied as needed. In some spectra, π/4 phase-shifted sine squared window function was applied to obtain better resolution when the peaks overlapped. The spectra were referenced with respect to 3-(trimethylsilyl)propionate at 0.0 p.p.m.

To estimate distances from the NOESY spectra for calculation of the ordered helical structure in 90% TFE, the peaks in a 250-ms spectrum were broadly classified as strong, medium and weak based on the number of contour levels in the peaks. Well resolved NOESY peaks between geminal protons of side chain amide groups of Asn2, Asn13 and Gln14 were used as reference peaks as the distance (1.8 Å) between geminal protons is independent of conformation. Final NOE distance constraints were created by assigning the strong, medium and weak intensities to the distance ranges of 1.80–2.70, 2.70–3.50 and 3.50–5.00 Å. These distance ranges for the different proton pairs are internally consistent and represent fairly loose bounds for most peaks. Pseudoatom corrections were used for methyl and methylene protons where stereospecific assignments were not available. There were a total of 190 constraints.

$^3J_{NH\alpha}$ coupling constants for all nonglycine residues were estimated from high resolution DQF-COSY spectra. Horizontal one-dimensional slices (at fixed F1) were picked from the centres of the antiphase peaks and these were inverse fast Fourier transformed (FFT) after a Hilbert transformation using FELIX v.97.0 software. Zero filling was done to 65 536 points to obtain a good lineshape and a very strong sine-squared window function shifted by 7° was applied along the F2 dimension to reduce overlap of antiphase components and to narrow the linewidths. This trace was then complex FFT and the separation between the antiphase components was taken to be the coupling constant. For a few peaks, sine-squared window function shifted by 30° was applied when a stronger window function led to distortion of the peaks

^{1}H-^{13}C Gradient-heteronuclear single quantum correlation (HSQC) [24] experiments with sensitivity enhancement were recorded at 0, 27, 36, 45, 54, 63 and 90% TFE to monitor the changes in ^{13}Cα chemical shifts. Spectral width used for the F1 dimension was 12 000 Hz and 16 scans were added for each of the 600 t_1 increments. Indirect referencing to 2,2-dimethyl-2-silapentane-5-sulfonic acid was used along this dimension where 0 p.p.m. ^{13}C = −13933.465 Hz absolute frequency.

Molecular dynamics simulations

The structure calculations were performed using Discover and INSIGHTII v.95 softwares (Molecular Simulations Inc.) on a Silicon Graphics Indigo2 workstation. Using the Biopolymer module, a right handed α-helix with acetylated N terminus and amidated C terminus and ionized side chains of Asp, Glu and Lys to simulate the pH in the NMR studies was constructed. The dielectric constant was set as 1.0*r to simulate aqueous environment implicitly. The energy of the system was calculated with the CFF91 [25] forcefield with no cross-terms included in the energy expression. The energy of the system was minimized first with 500 steps of steepest descents, followed by 1000 steps of conjugate gradients to remove any strain in the starting conformation of the peptide. A forcing potential of 100 kcal·mol^{-1}·rad^{-2} was applied to all omega angles to keep them in *trans* configuration during the simulation. A total of 190 NOEs were used as distance restraints with the force constant set to 60.0 kcal·mol^{-1}·Å^{-2}. The maximum force that was allowed to be applied to satisfy the constraint was 1000 kcal·mol^{-1}. The following protocol was used for restrained molecular dynamics calculations and simulated annealing. The molecule was 'heated' to a temperature of 1000 K and equilibrated at this temperature for 5 ps. Dynamics was continued for 200 ps during which Newton's equation was solved by the VERLET algorithm [26] with the integration time step of 1 fs. The molecular dynamics trajectory was sampled every 1 ps to generate a total of 200 structures. Equilibration was done by velocity scaling and a weak coupling to a temperature bath of time constant 0.1 ps was used during the sampling period. All of these structures were then 'cooled' to 300 K in steps of 50 K; at each step the molecule was equilibrated for 1 ps and then cooling resumed for 5 ps. At the end of the simulated annealing, all the structures were energy minimized for 500 steps of steepest descents followed by 1000 steps of conjugate gradients. Among all the structures, 10 structures with low energy and least restraint violations were collected for further analysis.

RESULTS

CD

The progressive formation of α-helix on increasing the concentration of TFE in the peptide solution was monitored by CD spectroscopy. Fig. 1A shows that the peptide does not have significant absorption at 222 nm at 0% TFE but on increasing TFE concentration minima appear at 222 and 208 nm which continuously deepen with a maximum appearing at 190 nm. These features are characteristics of progressive helix formation in the peptide. However, it was also observed that the CD spectra as a function of TFE concentration exhibit

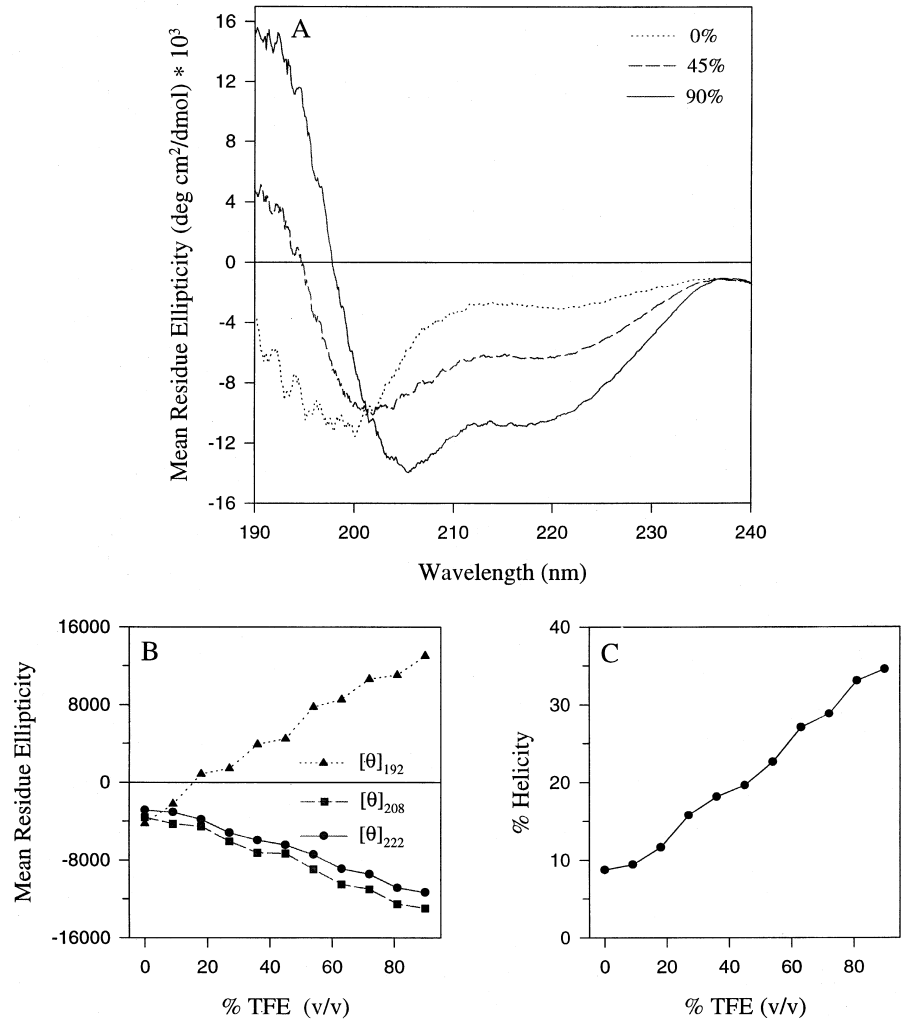

Fig. 1. Characteristics of the SNT-like peptide in increasing concentrations of TFE. (A) CD spectra: mean residue ellipticity (θ) plotted at 0%, 45% and 90% TFE. The presence of isodichroic point at 202 nm signifies a two-state equilibrium between random coil and helix. (B) Plot of $[\theta]_{192}$, $[\theta]_{208}$ and $[\theta]_{222}$ vs. TFE concentration. The ellipticity values change steadily in the direction of increasing helix population. (C) Calculation of percentage helicity from $[\theta]_{222}$. The maximum helical percentage reached at 90% TFE is 35%.

an isodichroic point at 202 nm indicating the presence of a two-state equilibrium between random coil and α-helix conformations. Fig. 1B records the experimentally observed values of $[\theta]_{192}$, $[\theta]_{208}$ and $[\theta]_{222}$ with the gradually varying concentration of TFE and it is seen that all of the three mean residue ellipticity values change in the direction pointing towards an increasing percentage of α-helix in the solution The helical content in the peptide calculated using the formula given by Farood et al [16] is shown in Fig. 1C. The maximum helicity observed at 90% TFE is 35%. However, the CD spectra do not provide information about the location and extent of helical structure in the peptide.

NMR spectroscopy

In order to derive the sequence-wise details of helix formation, as a function of TFE percentage, we have carried out NMR analysis on the peptide, at each of the TFE percentages mentioned above.

^1H Resonance assignment. TOCSY and NOESY/ROESY spectra were recorded at each of the 11 concentrations of TFE and sequence specific assignments were obtained following standard procedures [22]. Fig. 2 presents the fingerprint regions of NOESY spectra at 36% and 90% TFE and NOESY and ROESY spectra at 0% TFE with the sequential connectivities drawn in as illustrations. Likewise, Fig. 3 shows the NH–NH region of NOESY spectra for 36% and 90% TFE and NOESY and ROESY spectra for 0% TFE. The NOESY spectra displayed many noticeable features including a great increase in the dispersion of α proton chemical shifts, increase in the intensity of peaks and appearance of many medium range αH-NH(i, i + 3) NOE peaks as the TFE percentage increased from 0% to 90% TFE indicating progressive structure formation.

^1H-^1H Coupling constants. The values of coupling constants $^3J_{NH\alpha}$ measured at various concentrations of TFE further reinforce the structural transition observed by other methods. The coupling constants are generally averaged to 7–8 Hz in short peptides because of flexibility of the structure that results in the presence of an equilibrium between populations of different structures. Hence, if the values of $^3J_{NH\alpha} < 6$ Hz, it is taken to indicate a predominantly α-helical structure. We observe that at 0% TFE, the value of $^3J_{NH\alpha}$ for all of the residues is greater than 6 Hz, mostly 7–8 Hz, signifying that the peptide is present mainly as a random coil. At 90% TFE the coupling constants are reduced to 5–6 Hz with the exception of the C-terminal residue Gln14, which may experience fraying at the terminus. Among these, the central residues have lower values, suggesting higher helical propensity.

Stepwise helix formation

To probe further residue-wise structure formation, as the TFE concentration was increased, the following analyses have been carried out.

NOE connectivity patterns. The NOESY spectra were scanned to pick out peaks characteristic of helix and their intensities were estimated by counting the number of contours. This is indicated by the height of the bars in Fig. 4 that shows the

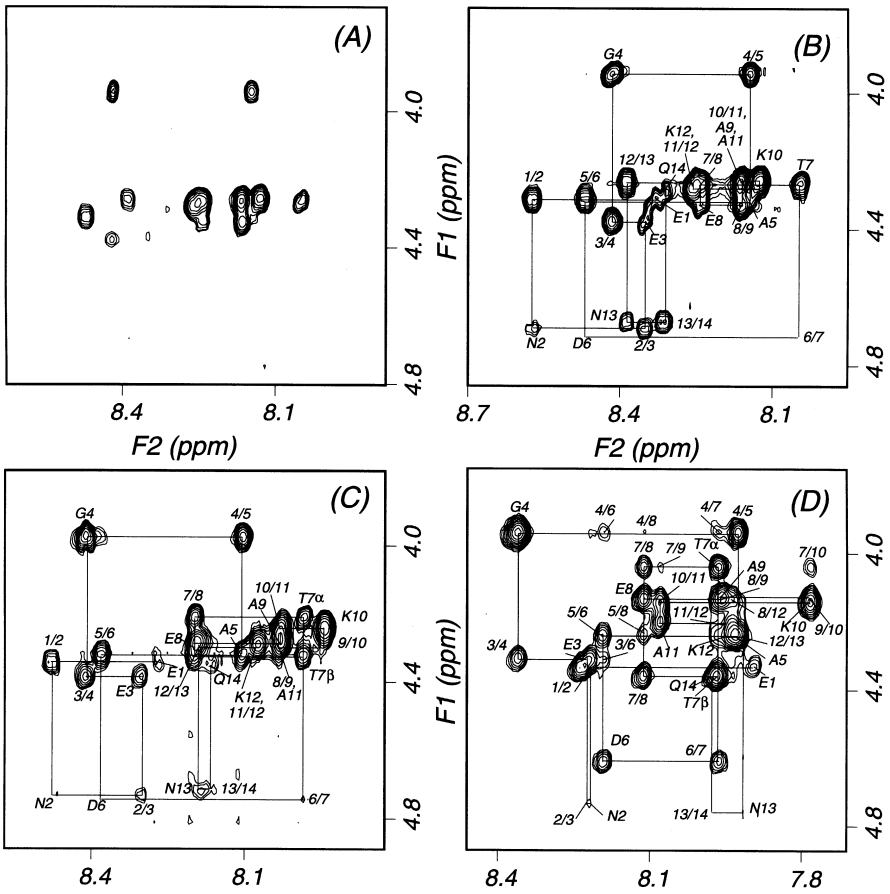

Fig. 2. 600 MHz fingerprint region of NOESY (400 ms mixing time) and ROESY spectra (250 ms mixing time) for the SNT-like peptide at different concentrations of TFE. The spectra are assigned and labelled except for the NOESY spectrum of 0% TFE (A) for which the ROESY spectrum (B) was used for assignments. The inter-residue cross-peaks are labelled as α(i)/NH(i + *n*) where *n* = 1, 2,3 and 4. (C) and (D) are NOESY spectra at 36% and 90% TFE respectively.

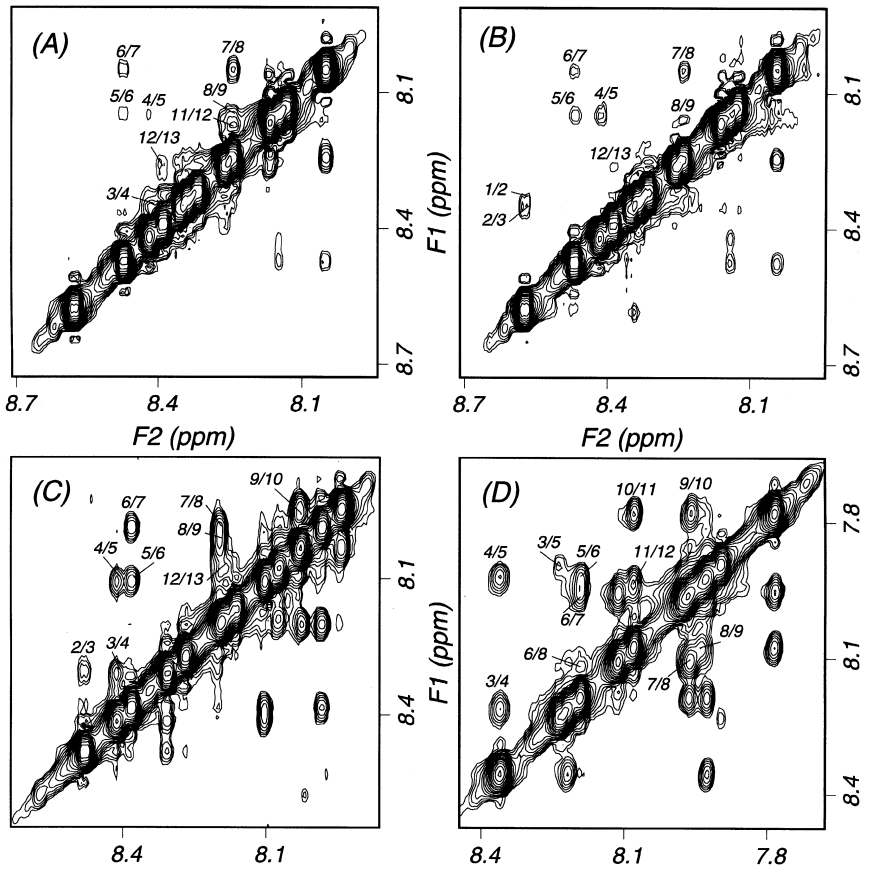

Fig. 3. Backbone amide region of NOESY (400 ms) and ROESY (250 ms) spectra for the SNT-like peptide at different concentrations of TFE. The peaks are labelled as NH(i)/NH(i + *n*) where *n* = 1,2. Cross-peaks in the ROESY spectrum have the opposite phase to the diagonal. (A) NOESY and (B) ROESY spectra at 0% TFE. NOESY spectra at (C) 36% and (D) 90% TFE.

summary of sequential and other connectivities at six different concentrations. We note that helix-specific NOEs, αH-NH (i, i + 3), αH-NH(i, i + 4) and αH-βH (i, i + 3) appear preferentially for different residues as the concentration of TFE is increased, with 18% TFE showing only two αH-NH (i, i + 3) peaks in the middle region (residues 4–10). Besides, the intensities of the peaks also increase in the same preferential manner. The presence of some structurally important peaks such as NH-NH (i, i + 1) and NH-NH (i, i + 2) cross-peaks for the middle residues even at 0% TFE indicates that a part of the backbone does have a finite population of helical conformation and the peptide is not completely random coil at 0% TFE. The CD measurements indicated ≈ 8% helicity at 0% TFE (Fig. 1) but qualitatively the NMR estimated percentage may be slightly higher as otherwise characteristic NOEs may not be seen. Such differences between NMR and CD studies regarding the estimation of helicity present in the system have been reported earlier in the literature [12,27,28] and various explanations have been offered for the same; the primary reason is the different time scales of observations. At 90% TFE, we are able to see complete helix specific connectivities through the entire chain. However, this must be taken to imply that every residue exhibits a finite population of helical conformation, and the NOEs may be taken to characterize these structures.

Chemical shifts analysis. Changes in chemical shift values of peptide resonances on TFE addition are of utmost significance in detecting and characterizing increasing populations of folded structures in solution. Chemical shift analyses for the αH, NH and αC have been carried out and the results for αH are depicted in Fig. 5. It is important to mention here that αH profiles are less distorted (than NH) when temperature or solvent changes are used to promote helix formation as the chemical shift values of such signals do not change very much in random peptides [29,30]. As the concentration of TFE is increased, αH chemical shifts for residues Ala5 to Ala11 move upfield, which is characteristic of progressive helix formation [31]. The chemical shifts of the end residues show a very interesting pattern: they are almost unchanged until 9–18% TFE, then move more downfield until 36–45% TFE when a plateau is reached; they then finally move a little upfield at 72–81% TFE which may signify that these residues move into the helical conformation very late, after 72% TFE only. When all of these αH chemical shifts are compared with the random coil values [31], which have been checked by Wright *et al.* [30] to be unchanged even in the presence of large concentration of TFE, it is found that the residues at the N and C termini show a fluctuating behaviour and have smaller chemical shift differences with respect to the random coil values. On the other hand, for residues Ala5 to Lys12, the chemical shifts are significantly lower than the random coil values (except Asp6) and progressively move more upfield up to 90% TFE.

Similar characteristic trends are seen in αC chemical shifts. As shown in Fig. 6, the dispersion of the backbone resonances substantially increase from 0 to 90% TFE. Notable changes are also observed in the positions of the side chain carbon resonances. The large dispersion of the chemical shifts at 90% TFE is typical of a well-defined folded conformation that exhibits unique local environments for individual residues. When the αC chemical shifts at the various concentrations of TFE are compared with the random coil values, these also show a similar pattern as observed for the αH. The αC resonate more and more downfield as the percentage of TFE is increased. Residues from Ala5 to Ala11 show a higher downfield shift than the terminal residues.

Packing of side chains. Analysis was also performed of NOESY cross-peaks in the side chains region to see the TFE dependence of these NOEs. Table 1 shows how the peaks of the side chain protons belonging to some residues change from negative to zero to positive intensities. Negative intensities mean that the protons belong to a region which is very mobile and therefore has a small rotational correlation time such that the product of spectrometer frequency and the molecular rotational correlation time, $\omega_0\tau_c$, < 1. This situation is present for residues Glu1, Glu2, Glu3, Glu8, Lys10, Asn13 and Gln14 at low percentages of TFE. The peaks slowly become positive at higher concentration of TFE, passing through a stage when these are absent ($\omega_0\tau_c \approx 1$) which happens at the intermediate concentrations of TFE. It is observed that cross-peaks involving backbone amide protons become positive at a lower concentration of TFE than those arising from αH or side chain amide protons. Besides, the cross-peaks involving γH remain negative or unobserved until a higher concentration of TFE as compared with those involving βH. Moreover, cross-peaks for the middle residues, Glu3, Glu8 and Lys10, undergo transition at an earlier concentration of TFE as compared with those ones involving the end residues. All of the cross-peaks become positive by 81% TFE, except the one involving the amide protons in the long side chain of Gln14. The cross-peak Gln14 $\delta NH_2 \rightarrow \gamma H$ protons remains negative till the very end of the titration signifying high mobility of this side chain. This graded change of intensities indicates that the observed change of NOE sign is not a result of viscosity changes in the medium, but must indicate packing of the side chains as the secondary structure is formed. One noteworthy point here is that only the residues with long side chains show this characteristic pattern and this must be because these side chains are highly mobile in solution. This analysis thus tells us how helix stabilization is taking place where not only the backbone but the side chains are also getting packed together with the increasing concentration of TFE. Furthermore, the fact that the percentage of TFE at which the different side chains, depending on their length and position in the sequence, order themselves suggests step-wise helix

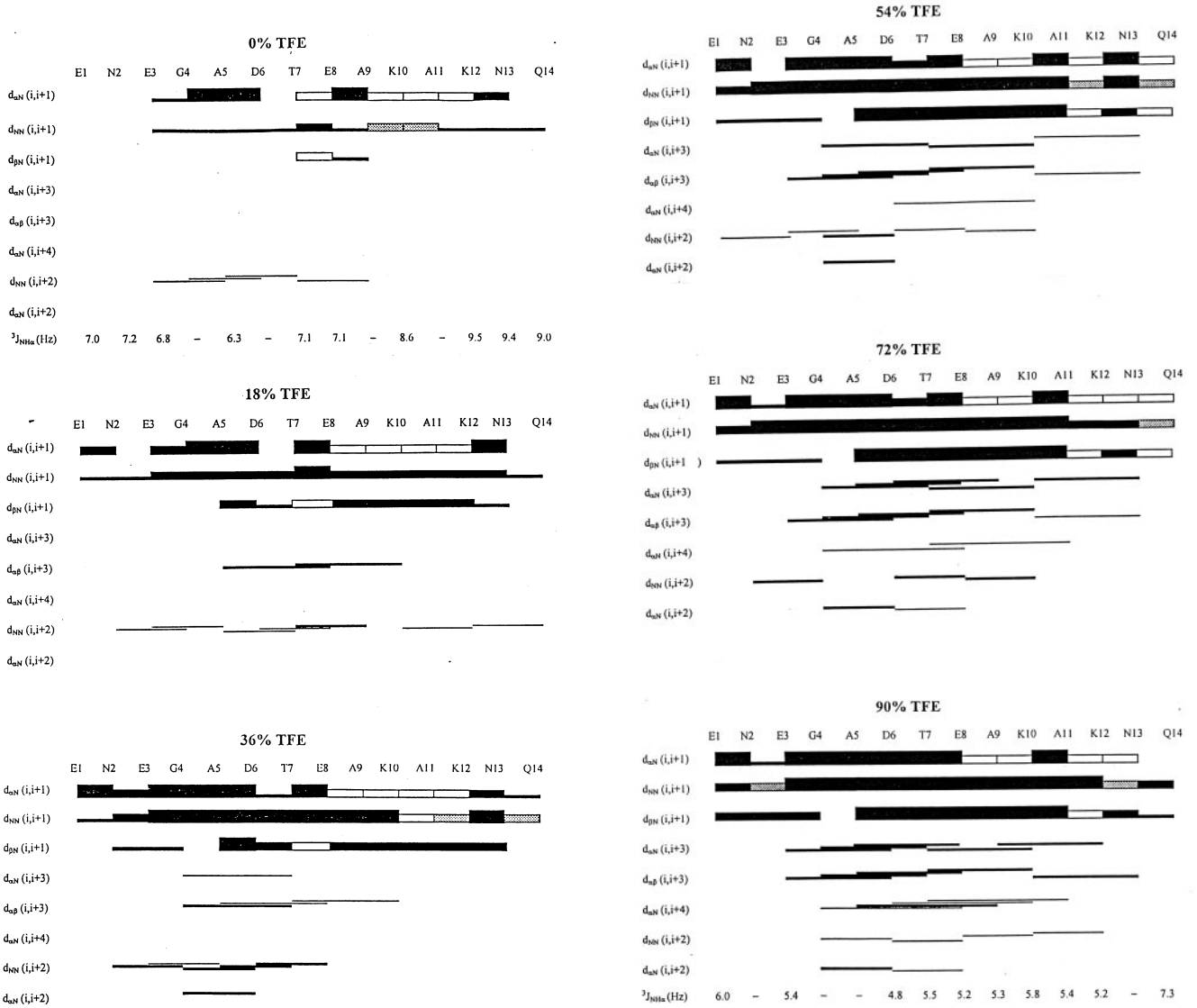

Fig. 4. Summary of the sequential, short- and medium-range NOE connectivities for the SNT-like peptide in 0%, 18%, 36%, 54%, 72% and 90% TFE. The intensities of the observed NOEs are represented by the thickness of the bars. Open bars represent overlapping peaks and grey-shaded bars denote near diagonal peaks. Overlapping and near diagonal peaks are not included for short- and medium-range NOEs $-d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+2)$, $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$ and $d_{\alpha N}(i,i+4)$. $^3J_{NH\alpha}$ coupling constants (except Gly) are also given at 0% and 90% TFE. Some coupling constants could not be measured because of overlap of peaks.

Table 1. Summary of NOEs observed in SNT-like peptide that undergo transition from negative to positive intensity from 0% to 90% TFE. Negative (–) and positive (+) intensites and peaks absent from their expected positions (0) are shown.

	TFE concentration (%)										
	0	9	18	27	36	45	54	63	72	81	90
Glu1 α → β	–	–	–	–	0	0	+	+	+	+	+
Glu1 α → β′	–	–	–	0	0	0	+	+	+	+	+
Glu3 α → β	–	–	0	+	+	+	+	+	+	+	+
Gln14 α → β	–	–	–	0	0	+	+	+	+	+	+
Gln14 α → β′	–	–	–	–	0	+	+	+	+	+	+
Glu1 α → γ	–	–	–	–	–	0	0	0	0	+	+
Glu3 α → γ	–	–	0	+	+	+	+	+	+	+	+
Glu8 α → γ	–	+	+	+	+	+	+	+	+	+	+
Lys10 α → γ	–	0	0	0	0	+	+	+	+	+	+
Lys10 α → δ	–	–	–	0	0	+	+	+	+	+	+
Gln14 α → γ	–	–	–	–	–	–	0	0	0	+	+
Glu1 NH → β	–	–	0	0	+	+	+	+	+	+	+
Glu1 NH → β′	–	–	0	0	+	+	+	+	+	+	+
Asn2 NH → β	–	0	0	+	+	+	+	+	+	+	+
Glu3 NH → β	–	–	+	+	+	+	+	+	+	+	+
Gln14 NH → β	–	0	0	+	+	+	+	+	+	+	+
Gln14 NH → β′	–	0	0	+	+	+	+	+	+	+	+
Glu1 NH → γ	–	–	–	–	+	+	+	+	+	+	+
Glu3 NH → γ	–	–	–	0	+	+	+	+	+	+	+
Gln14 NH → γ	–	–	–	–	0	0	+	+	+	+	+
Asn2 γNH_2 → β/β′	–	–	–	–	–	0	+	+	+	+	+
Asn13 γNH_2 → β/β′	–	–	–	–	–	0	+	+	+	+	+
Gln14 δNH_2 → γ	–	–	–	–	–	–	–	–	–	–	–

Table 2. Summary of experimental restraints used for structure calculation of SNT-like peptide at 90% TFE and statistical analysis of the family of structures obtained by simulated annealing using NOE distance constraints.

Parameter	Value
Distance restraints	
All	190
Intra-residue	96
Inter-residue	94
Sequential	50
Medium range	44
i, i + 2	11
i, i + 3	27
i, i + 4	6
Long range	0
RMS violations/constraint (nm)	0.019 ± 0.001
Average number of violations/structure	14.60 ± 0.52
RMSDs with average structure (nm)	
Backbone atoms	0.023 ± 0.008
All atoms	
maximum	0.095
minimum	0.063
average	0.078 ± 0.010
Average pair-wise RMSD (nm)	0.103 ± 0.070
Total energy (kcal/mol)	–118.16 ± 1.27

formation and this must point to different helical propensities of the residues and their position along the sequence.

TFE dependent amide protons temperature coefficients

Amide proton temperature coefficients serve as good monitors for the formation of intra-molecular H-bonds in a peptide or

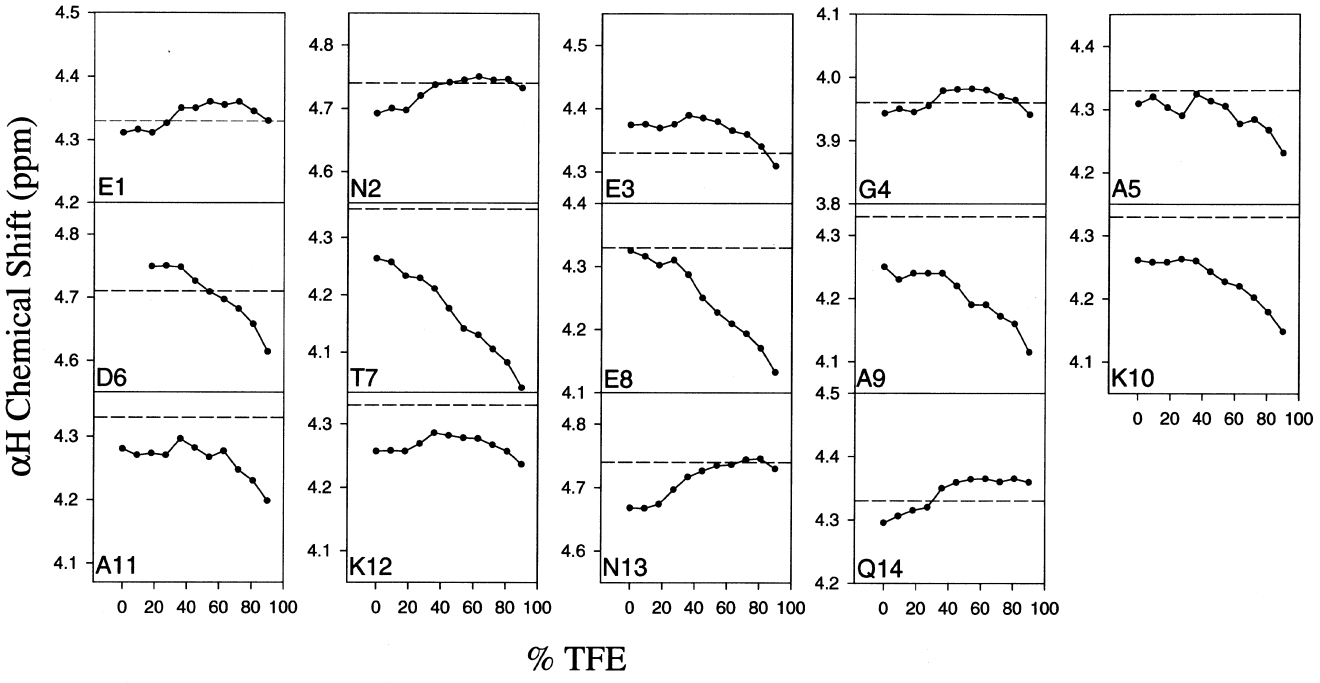

Fig. 5. Change in αH chemical shifts of SNT-like peptide with increasing concentration of TFE. The broken horizontal line denotes the random coil chemical shift for that residue taken from the literature [31]. The first two points could not be ascertained for Asp6 because of close proximity of the resonance to the water line resulting in saturation transfer.

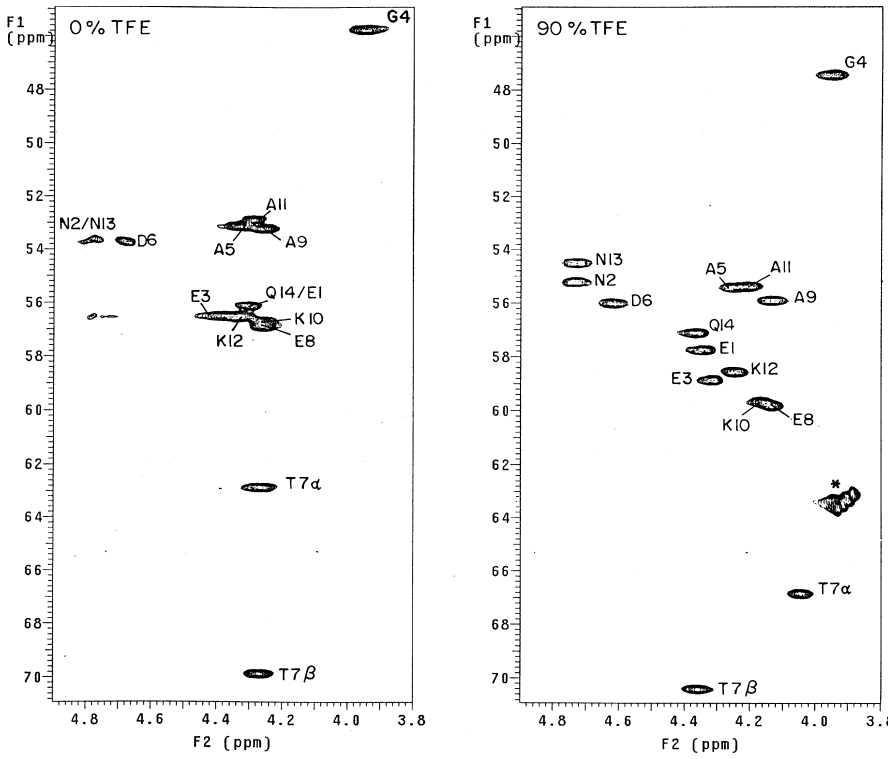

Fig. 6. ^{13}C-HSQC spectra of SNT-like peptide at 0% and 90% TFE. The overlapping peaks were assigned with the help of ^{1}H spectra. The peak marked with an asterisk is an artefact arising from the excess TFE in the solution.

protein molecule. In aqueous solutions, low values below a particular cut-off of 5 p.p.b.·K^{-1} established empirically is taken to indicate the participation of the amide proton in intra-molecular H-bonds [32]; this cut-off value depends, of course, on the nature of the solvent with which the free amide proton is solvated. Thus in a TFE environment, the cut-off value for indication of intra-molecular H-bond formation by the amide proton, will be different from that in aqueous solution. Fig. 7 compares the observed temperature coefficients of the backbone amide protons at 0, 27, 45, 63 and 90% TFE concentrations. The pattern shows an interesting behaviour, with a minimum occurring at about 45% TFE. Now because we know that, at higher TFE concentrations, the helix is very well formed, the data in Fig. 7 suggests that the temperature coefficient cut-off

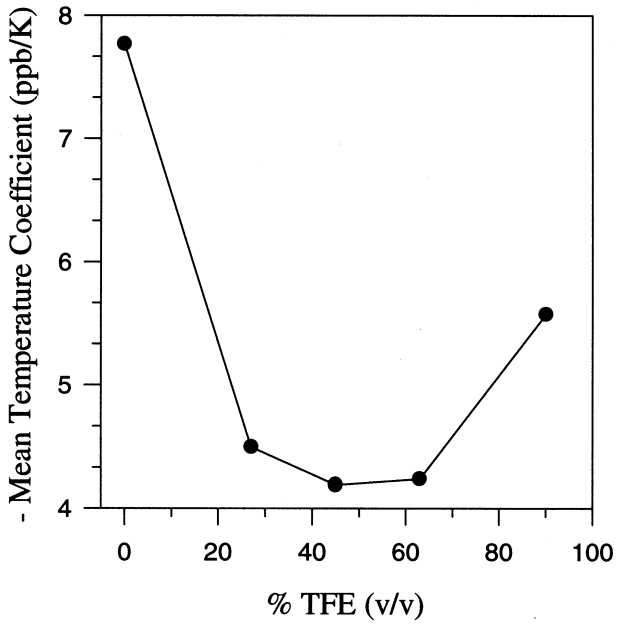

Fig. 7. Plot of mean temperature coefficient vs TFE concentration. The mean was calculated from the observed temperature coefficients of amide protons of individual residues at a particular concentration of TFE.

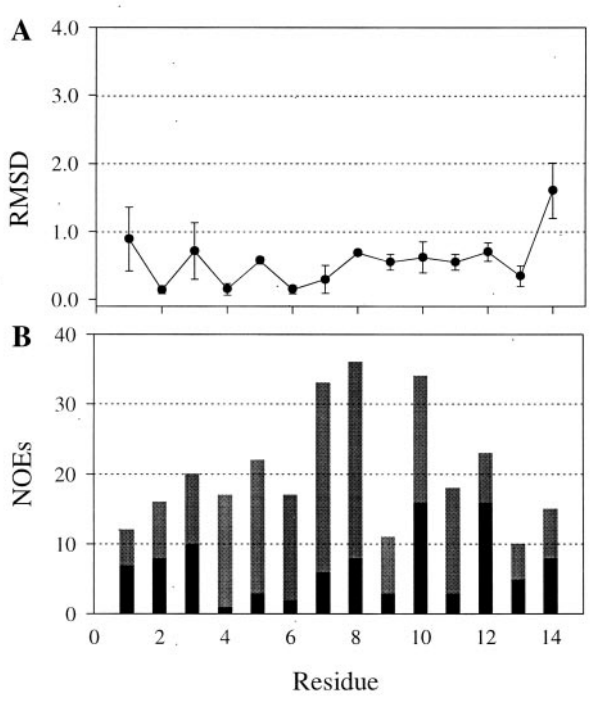

Fig. 8. RMS deviation of the 10 structures of SNT-like peptide about the mean structure for all atoms per residue. (A) The bars denote SDs. (B) Black columns, intra-residue NOEs; grey columns, inter-residue NOEs.

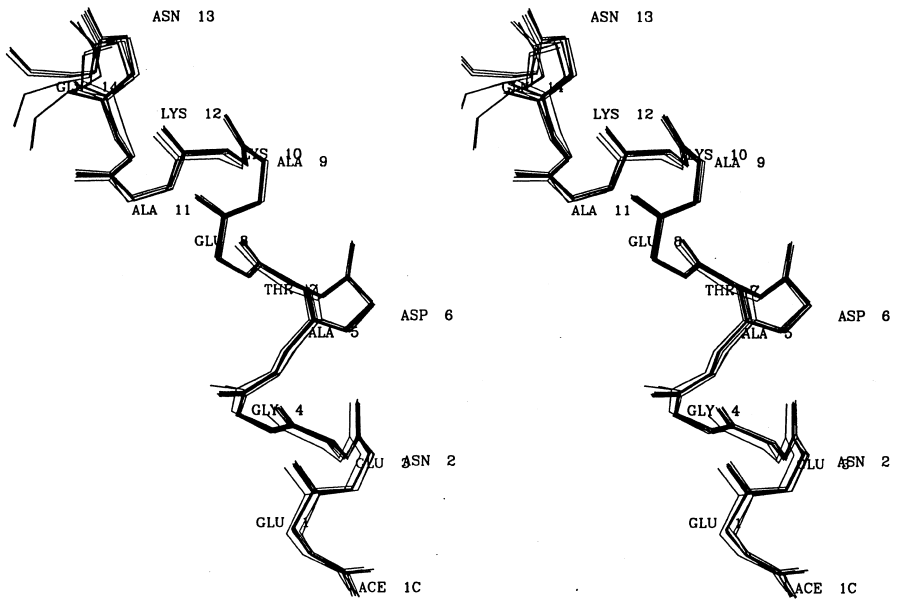

Fig. 9. Stereo view of the backbone of the 10 final simulated annealing structures of the SNT-like peptide. Structures were superimposed minimizing the RMSDs between the backbone atoms – N, αC, C and O. The αC of each residue is labelled.

value to indicate intra-molecular H-bond is much higher in TFE than that in aqueous solutions. Residue-wise analysis of the temperature coefficients indicated that all of the residues exhibit similar patterns with TFE concentration as shown in Fig. 7. But the individual magnitudes were seen to depend on the nature of the residue in addition to position in the sequence.

Helical structure at 90% TFE

With a view to see the side chain interactions in the helix formed at 90% TFE, the ordered structure of the peptide was calculated using the structural constraints derived from NMR data at 90% TFE, as described in Materials and methods. A summary of statistics associated with final simulated annealing structures and parameters used for the restrained molecular dynamics simulation on the 90% TFE sample is given in Table 2. A total of 190 distance constraints were used with 96 intra-residue and 94 inter-residue distances. Dihedral angle restraints were not used as the coupling constants were found to be averaged due to conformational heterogeneity. Fig. 8B shows the distribution of NOEs used for deriving distances across the length of the peptide. Residues Thr7, Glu8 and Lys10 show the maximum number of NOEs and consequently the RMSDs of these residues are lower as compared with others and the error bars are also very low, as depicted in Fig. 8A. Ten best convergent structures were selected for detailed analysis: the structures were well-defined with a maximum pair-wise RMSD of 0.023 nm for backbone atoms and 0.103 nm for all atoms of all residues. An overlay of the 10 final structures in stereo mode is shown in Fig. 9: a good α-helix is formed in the stretch Glu1–Lys12, and the helix is distorted at the C-terminal end. The structure shows a small kink at Gly4. Glycine is known to be a helix-breaker [33] and so it is not surprising to find a twisted α-helix in presence of this high concentration of structure-inducing solvent. The C-terminus is not as clearly defined as the rest of the backbone because of the relatively smaller number of distance constraints obtained for this region because of spectral overlap. In the Ramachandran plot (not

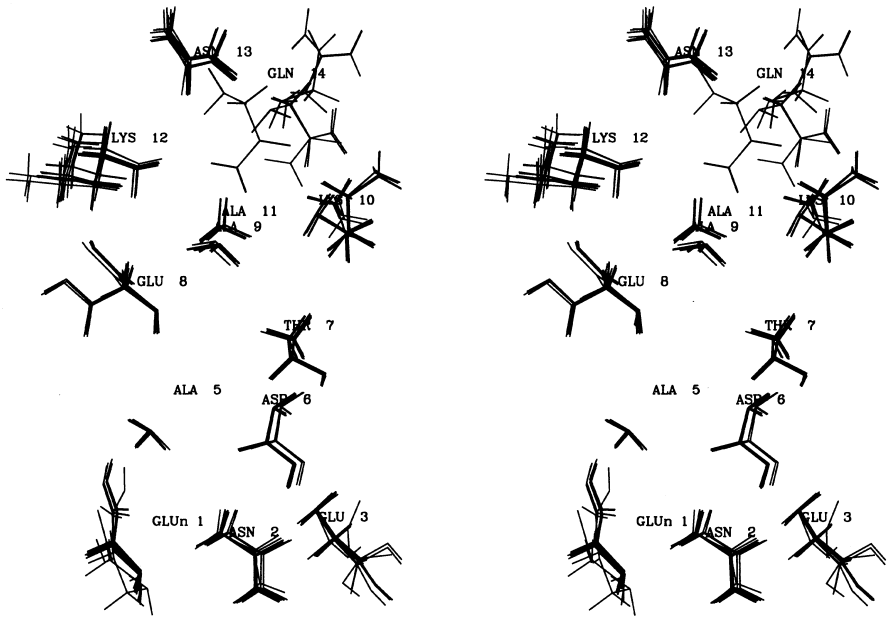

Fig. 10. Stereo-view of the superposition of only the side chains of the 10 final simulated-annealing structures of the SNT-like peptide.

shown) all of the backbone dihedral angles were seen to be distributed properly in the energetically favourable region characteristic of α-helix. Fig. 10 shows the packing of the side chains in the superposition of the 10 final structures of the peptide. It may be seen that the side chains from Asn2 to Asn13 have well-defined positions in the structure indicating a good packing of the side chains. In contrast, the end residues side chains span a large conformational space. Thus, one may conclude that the side chain interactions which are dependent on the sequence play important roles in dictating the secondary structure of the peptide.

DISCUSSION

We have characterized in this paper the conformational preferences for the SNT-like peptide in detail at various concentrations of TFE using CD and NMR spectroscopy. Far UV CD spectroscopy allowed us to monitor the overall structure present in the ensemble of molecules in solution, that is, α-helix or random coil. It was observed that the population of α-helix increased to 35% at 90% TFE. The maximum helical content of only 35% indicated that the inherent helical propensity in the peptide to form α-helix may not be very high. To gain further insight into residue level information important for studying the mechanism of folding of the peptide in detail, NMR spectroscopy was used which enabled us to study folding using changes occurring at each step of TFE titration by means of chemical shifts, coupling constants and NOESY peak patterns.

Chemical shifts of αH and NH (data not shown) protons and αC carbons showed clear cut upfield shifts in the former nuclei and downfield shifts in the latter, with respect to the random coil values, denoting helix formation in the peptide. This was very useful for the present study as it provided information about even small changes occurring in the peptide backbone caused by small variations in the solvent condition. Both the proton and carbon chemical shifts indicated that the stretch from Ala5 to Ala11 has a high population of helix even at low TFE concentration, and as the concentration of TFE is increased, the end residues also exhibit characteristics of helical conformation. This can also be interpreted to mean that residues from Ala5 to Ala11 are stabilized first when the peptide folds providing a nucleation centre and later on, the end residues also slowly move into helical conformation. The same conclusions were derived from the coupling constants which showed values removed from the random coil values indicating large populations of specific structures. The $^{3}J_{NH\alpha}$ coupling constants values changed from 7 to 8 Hz at 0% TFE to 5–6 Hz at 90% TFE, the central portion of the peptide having lower values. This showed that the population of α-helix had increased at 90% TFE, although it was much less than 100% (for which $^{3}J_{NH\alpha}$ should be less than 4 Hz) consistent with the results of CD studies.

Although the above data gave us information about the backbone conformation, the NOE intensities provided information about both backbone and side chains rigidities. NOE transfer efficiency between two protons is directly proportional to the rotational correlation time, and inversely proportional to the sixth power of the distance between the two protons. So, besides providing inter-proton distances in the molecule which have unique patterns for different structures, it gave very important information about the change in flexibility observed by different side chains depending on their lengths and location in the sequence as the concentration of TFE was increased. While the patterns and intensities of NOESY cross-peaks from the fingerprint and backbone amide regions of the spectra clearly showed the presence of medium range NOEs (increase in the intensity of NH–NH sequential cross-peaks) implying formation of a stable helical structure as the concentration of TFE is increased, the side chain regions showed how the movement of long side chains was becoming restricted, signifying a greater packing of the structure.

The results of the above studies suggest that the peptide has the potential to form a stable helical structure which encompasses the whole of the length of the peptide with the side chains also being packed correctly in the presence of a high concentration of the structure stabilizing solvent TFE. This fully helical structure is formed from an initial structure in which the helix spans only the middle residues and the side chains are free to move. Slowly, the end residues are also forced to arrange themselves in a helical fashion with side chains becoming more and more rigid thus forming a highly stable helical structure in which all of the atoms have well-defined positions. This provides an important insight into the folding mechanism of peptides and proteins: a nucleation step occurs first, involving only a few residues in the beginning, and the helix then propagates to the N- and C- terminal residues together with packing of the side chains.

The mean as well the individual temperature coefficients of the amide protons displayed an interesting dependence on TFE concentration. They both decreased first, reached a plateau at around 40–60% TFE and then increased again. Because at 90% TFE, the peptide is maximally helical spanning the whole length of the peptide, the above observations indicated that for an amide proton which is H-bonded intra-molecularly, the temperature coefficient in TFE is much higher than that in aqueous solutions; caution should be exercised in interpreting temperature coefficient data at any one discrete solvent condition as indicative of intra-molecular H-bond or otherwise, solely on the basis of a reported cut-off value of temperature coefficient.

ACKNOWLEDGEMENTS

We thank the National Facility for High Field NMR at TIFR, supported by Department of Science and Technology, India, for all the facilities. We thank M. M. Dhingra for providing the crude peptide and S. C. Panchal for help in purification of the peptide.

REFERENCES

1. Shortle, D. (1996) Structural analysis of non-native states of proteins by NMR. *Curr. Opin. Struct. Biol.* **6**, 24–30.
2. Wuthrich, K. (1994) NMR assignments as a basis for structural characterization of denatured states of globular proteins. *Curr. Opin. Struct. Biol.* **4**, 93–99.
3. Kim, P.S. & Baldwin, R.L. (1990) Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **59**, 631–660.
4. Dobson, C.M. (1992) Unfolded proteins, compact states and molten globules. *Curr. Opin. Struct. Biol.* **2**, 6–12.
5. Dill, K.A. & Shortle, D. (1991) Denatured states of proteins. *Annu. Rev. Biochem.* **60**, 795–825.
6. Segawa, S.-J., Fukuno, T., Fujiwara, K. & Noda, Y. (1991) Local structures in unfolded lysozyme and correlation with secondary structures in the native conformation: Helix-forming or -breaking propensity of peptide segments. *Biopolymers* **31**, 497–509.
7. Yang, J.J., Pitkeathly, M., Kotik, M., Haynie, D.T., Dobson, C.M. & Radford, S.E. (1995) Conformational properties of four peptides spanning the sequence of Hen lysozyme. *J. Mol. Biol.* **252**, 483–491.
8. Munoz, V., Serrano, L., Jimenez, M.A. & Rico, M. (1995) Structural

analysis of peptides encompassing all α-helices of three α/β parallel proteins. Che-Y, Flavodoxin and P21-Ras: implications for α-helix stability and the folding of α/β parallel proteins. *J. Mol. Biol.* **247**, 648–669.

9. Lu, Z., Fok, K., Erickson, B.W. & Hugli, T.E. (1984) Conformational analysis of COOH-terminal segments of Human C3a. *J. Biol. Chem.* **259**, 7367–7370.
10. Reymond, M.T., Merutka, G., Dyson, H.J. & Wright, P.E. (1997) Folding propensities of peptide fragments of myoglobin. *Protein Sci.* **6**, 706–716.
11. Dyson, H.J., Sayre, J.R., Merutka, G., Shin, H.-C., Lerner, R.A. & Wright, P.E. (1992) Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding II Plastocyanin. *J. Mol. Biol.* **226**, 819–835.
12. Kemmink, J. & Creighton, T.E. (1995) Effects of trifluoroethanol on the conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor. *Biochemistry* **34**, 12630–12635.
13. Blanco, J.F., Jimenez, M.A., Pineda, A., Rico, M., Santoro, J. & Nieto, J.L. (1994) NMR solution structure of the isolated N-terminal fragment of protein-G B_1 domain. Evidence of trifluoroethanol induced native-like β-hairpin formation. *Biochemistry* **33**, 6004–6014.
14. Hamada, D., Kurod, Y., Tanaka, T. & Goto, Y. (1995) High helical propensity of the peptide fragments derived from β-lactoglobulin, a predominantly β-sheet protein. *J. Mol. Biol.* **254**, 737–746.
15. Zhao, B., Carson, M., Ealick, S.E. & Bugg, C.E. (1992) Structure of scorpion toxin variant-3 at 1.2 A° resolution. *J. Mol. Biol.* **227**, 239–252.
16. Farood, B., Filiciano, E.J. & Nambiar, K.P. (1993) Stabilization of α-helical structures in short peptides via end capping. *Proc. Natl Acad. Sci. USA* **90**, 838–842.
17. Marion, D. & Wuthrich, K. (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H-^{1}H spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113**, 967–974.
18. States, D.J., Haberkorn, R.A. & Ruben, D.J. (1982) A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson.* **48**, 286–292.
19. Braunschweiler, L. & Ernst, R.R. (1983) Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. *J. Magn. Reson.* **53**, 521–528.
20. Bothner-By, A.A., Stephens, R.L., Lee, J., Warren, C.D. & Jeanloz, R.W. (1984) Structure determination of a tetrasaccharide: transient nuclear Overhauser effects in the rotating frame. *J. Am. Chem. Soc.* **106**, 811–813.
21. Macura, S., Huang, Y., Suter, D. & Ernst, R.R. (1981) Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *J. Magn. Reson.* **43**, 259–281.
22. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. Wiley, New York.
23. Piantini, U., Sorensen, O.W. & Ernst, R.R. (1982) Multiple quantum filters for elucidating NMR coupling networks. *J. Am. Chem. Soc.* **104**, 6800–6801.
24. Kay, L.E., Keifer, P. & Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665.
25. Maple, J., Dinur, U. & Hagler, A.T. (1988) Derivation of force fields for molecular mechanics and dynamics from *ab initio* energy surfaces. *Proc. Natl Acad. Sci. USA* **85**, 5350–5354.
26. Verlet, L. (1967) Computer 'experiments' on classical fluids. I. Thermodynamical properties of Lennard–Jones molecules. *Phys. Rev.* **159**, 98–103.
27. Bradley, E.K., Thomason, J.F., Cohen, F.E., Kosen, P.A. & Kuntz, I.D. (1990) Studies of synthetic helical peptides using circular dichroism and nuclear magnetic resonance. *J. Mol. Biol.* **215**, 607–622.

28 Sonnichsen, F.D., Van Eyk, J.E., Hodges, R.S. & Sykes, B.D. (1992) Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. *Biochemistry* **31**, 8790–8798.

29. Jimenez, M.A., Nieto, J.L., Herranz, J., Rico, M. & Santoro, J. (1987) ^{1}H NMR and CD evidence of the folding of the isolated ribonuclease 50–61 fragment. *FEBS Lett* **221**, 320–324.
30. Merutka, G., Dyson, H.J. & Wright, P.E. (1995) 'Random coil' ^{1}H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J. Biomol. NMR* **5**, 14–24.
31. Wishart, D.S. & Sykes, B.D. (1994) Chemical shifts as a tool for structure determination. In *Methods in Enzymology*, Vol. 239 (James, T.L. & Oppenheimer, N.J., eds), pp. 363–392. Academic Press Inc., San Diego.
32. Dyson, H.J., Rance, M., Houghten, R.A., Lerner, R.A. & Wright, P.E. (1988) Folding of immunogenic peptide fragments of proteins in water solution. I Sequence requirement for the formation of a reverse turn. *J. Mol. Biol.* **201**, 161–200.
33. Strehlow, K.G. & Baldwin, R.L. (1989) Effect of the substitution Ala→Gly at each of the five residue positions in the C-peptide helix. *Biochemistry* **28**, 2130–2133.