Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide

VICTOR A. JARAVINE,¹ ANDREI T. ALEXANDRESCU,² AND STEPHAN GRZESIEK¹

¹Department of Structural Biology, Biozentrum, University of Basel, Basel CH-4056, Switzerland

²Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269-3125, USA

(RECEIVED November 20, 2000; FINAL REVISION February 6, 2001; ACCEPTED February 7, 2001)

Abstract

Helix formation of an S-peptide analog, comprising the first 20 residues of Ribonuclease A and two additional N-terminal residues, was studied by measuring hydrogen bond (H-bond) $^3J_{\text{NC}'}$ scalar couplings as a function of 2,2,2-trifluoroethanol (TFE) concentration. The $^3J_{\text{NC}'}$ couplings give direct evidence for the closing of individual backbone N-H...O=C H-bonds during the TFE-induced formation of secondary structure. Whereas no $^3J_{\text{NC}'}$ correlations could be detected without TFE, α -helical (i,i +4) H-bond correlations were observed for the amides of residues A5 to M15 in the presence of TFE. The analysis of individual coupling constants indicates that α -helix formation starts at the center of the S-peptide around residue E11 and proceeds gradually from there to both peptide ends as the TFE concentration is increased. At 60% to 90% TFE, well-formed α -helical H-bonds were observed for the amides hydrogens of residues K9 to Q13, whereas H-bonds of residues T5 to A8, H14, and M15 are affected by fraying. No intramolecular backbone H-bonds are present at and beyond the putative helix stop signal D16. As the $^3J_{\text{NC}'}$ constants represent ensemble averages and the dependence of $^3J_{\text{NC}'}$ on H-bond lengths is very steep, the size of the individual $^3J_{\text{NC}'}$ coupling constants can be used as a measure for the population of a closed H-bond. These individual populations are in agreement with results derived from the Lifson-Roig theory for coil-to-helix transitions. The present work shows that the closing of individual H-bonds during TFE-induced helix formation can be monitored by changes in the size of H-bond scalar couplings.

Keywords: S-peptide; RnaseA; J-coupling; NMR; scalar coupling; protein folding

How a specific polypeptide chain folds into a well-defined three-dimensional protein structure remains one of the fundamental questions in structural biology. NMR spectroscopy has long been recognized as an invaluable tool to resolve individual steps of macromolecular folding under natively like conditions in aqueous solution (McDonald et al. 1971; Baldwin 1975). Folding reactions that are slow or

comparable to the timescale of individual NMR experiments can be analyzed by studying chemical shift changes of individual atomic nuclei in entire proteins (Epstein et al. 1971; Westmoreland and Matthews 1973; Kiefhaber et al. 1995; Balbach et al. 1996; Hoeltzli and Frieden 1998; Forge et al. 1999). Detailed structural characterizations by NOESY experiments have been performed for long-lived folding intermediates and molten globule states (Neri et al. 1992; Balbach et al. 1997; Mok et al. 1999). Because of the crucial role in the stabilization of protein secondary and tertiary structures, hydrogen bond (H-bond) formation during protein folding is of particular interest. Hydrogen/deuterium exchange experiments have revealed the stability of individual amide protons against solvent exchange during protein folding and have been used to detect early folding intermediates by fast mixing techniques (Roder and Wuthrich 1986; Roder et al. 1988; Udgaonkar and Baldwin

Reprint requests to: Stephan Grzesiek, Department of Structural Biology, Biozentrum, University of Basel, Klingelbergstr. 70, Basel CH-4056, Switzerland; e-mail: Stephan.Grzesiek@unibas.ch; fax: 41-61-2672109 or Andrei Alexandrescu, Department of Molecular and Cell Biology, University of Connecticut, 75 North Eagleville Road, U-3125, Storrs, CT 06269-3125, USA; e-mail: andrei@uconnvm.uconn.edu.

Abbreviations: TFE, 2,2,2-trifluoroethanol; RNaseA, Ribonuclease A; S-peptide, N-terminal fragment of Ribonuclease A; H-bond, hydrogen bond.

Article and publication are at www.proteinscience.org/cgi/doi/10.1110/ps.48501.

1988). Protection against hydrogen exchange is usually taken as an indication for the presence of H-bonds, although other factors like solvent accessibility are known to have a strong influence on the exchange rates (Wagner and Wuthrich 1982).

Clearly, a more quantitative characterization of individual H-bonds during the transition from the unfolded to the folded state has the potential to reveal important details about the cooperativity and energetics of secondary and tertiary structure formation. Recently, NMR experiments that detect individual $\text{NH}\cdots\text{O}=\text{C}$ H-bonds in proteins through H-bond $^3\text{J}_{\text{NC}'}$ scalar couplings have been described (Cordier and Grzesiek 1999; Cornilescu et al. 1999a). The experiments identify both the donor and the acceptor group of an individual H-bonding pair. In addition, the size of the $^3\text{J}_{\text{NC}'}$ scalar coupling constant provides quantitative information on the ensemble average of the H-bond length (Cordier and Grzesiek 1999; Cornilescu et al. 1999b). We show here that, being successfully applied to the study of folded protein structures, the method is also applicable to marginally stable peptides and small proteins with a large percentage of unfolded structure.

To directly influence the folding of such molecules, structure-forming solvents such as 2,2,2-trifluoroethanol (TFE) and other halogenated alcohols are frequently used (Nelson and Kallenbach 1986, 1989; Buck 1998; Chiti et al. 1999). TFE preferentially stabilizes α -helical structures (Timasheff 1970; Hamada et al. 1995; Buck 1998). The mechanism of this stabilization remains unclear and both direct binding of TFE to the peptide and weakening of peptide H-bonds to water are being discussed (Luo and Baldwin 1997; Buck 1998).

The S-peptide, corresponding to the first 20 N-terminal residues of Ribonuclease A (RNaseA) (Kim and Baldwin 1984), provides an ideal model system for the study of peptide folding. The 124 amino acid RNaseA is cleaved by subtilisin between residues 20 and 21, generating the S-peptide (1-20 aa) and the S-protein (21-124 aa). The two parts reassociate in stoichiometric amounts to form an S-complex ($K_d < 10^{-7}$ M) that has the full enzymatic activity of RNaseA and a structure nearly identical to that of the intact protein. Whereas at 25°C (pH 3.8, 0.1 M NaCl) the peptide is largely unfolded in aqueous solution, CD spectra indicate that the peptide adopts a stable α -helical structure in the presence of more than 60% (v/v) TFE (Nelson and Kallenbach 1986, 1989). The changes in NMR chemical shifts of the peptide as a function of increasing TFE concentration indicate that this α -helix consists of residues T3 to H12, which also form an α -helix in native RNaseA (Nelson and Kallenbach 1989).

In this paper we follow the transition of an S-peptide analog from the unfolded to the folded state by changing the TFE content of the solution from 0% to 90% (v/v). The H-bond HNCO experiment (Cordier and Grzesiek

Fig. 1. Sections of the long-range HNCO spectra recorded on the S-peptide. Trans-H-bond correlations ($^3\text{J}_{\text{NC}'}$) are marked by the residue number of the HN group followed by the residue number of the acceptor carbonyl. Two-bond $^2\text{J}_{\text{NC}'}$ intrasidue connectivities (Cordier and Grzesiek 1999) are marked by the superscript *i*. Not completely suppressed sequential, one-bond ($^1\text{J}_{\text{NC}'}$) connectivities are marked by the superscript *s*. Two small side bands of the very intense $\text{A}^{22\text{s}}$ resonance are marked by *a*. (A) Small sections of the spectra at 10%, 20%, 30%, 45%, 60%, and 90% (v/v) TFE. (B) Region of $^3\text{J}_{\text{NC}'}$ correlation spectrum at 60% TFE.

1999; Cornilescu et al. 1999a) is used to precisely define the limits and types of H-bonded structure in this peptide and to establish if this structure has alternative H-bonded conformations. The gradual formation of the peptide's α -helical structure is characterized quantitatively by the changes in the individual $^3J_{\text{NC}'}$ coupling constants. The results can be interpreted in terms of the Lifson-Roig coil-to-helix transition theory. This approach shows the power of modern heteronuclear NMR methodology to follow the formation of secondary structure at the level of individual residues.

Results

Observation of the formation of individual amide H-bonds

Figure 1A shows small regions of the long-range HNCO spectra measured at different TFE concentrations. This region comprises α -helical $\text{NH}(\text{i})\cdots\text{O}=\text{C}(\text{i}-4)$ correlations for residues $\text{F10}^{\text{HN}}\text{-A6}^{\text{C}'}$, $\text{E11}^{\text{HN}}\text{-A7}^{\text{C}'}$, and $\text{R12}^{\text{HN}}\text{-A8}^{\text{C}'}$. Clearly, the individual cross H-bond correlations become stronger with increasing TFE concentration. For example, the cross peak corresponding to the $\text{R12}^{\text{HN}}\text{-A8}^{\text{C}'}$ H-bond is invisible in the spectrum at 10% TFE (marked by a dotted circle) and gradually increases in intensity for TFE concentrations of 20% to 90%. The variation in intensity is the result of a change in the ensemble average over the states of the individual H-bonds. Apparently, at low TFE concentrations, the individual backbone H-bonds are only weakly populated such that the $^3J_{\text{NC}'}$ couplings are weak and the cross peak intensities are small. At higher concentrations of TFE, cross peaks corresponding to α -helical H-bonds can be observed for a larger number of residues. Thus at 60% TFE,

sequential α -helical connectivities can be established for all amide hydrogens of residues A6 to M15. With the exception of the overlapping cross peaks for the $\text{K9}^{\text{HN}}\text{-T5}^{\text{C}'}$ and $\text{M15}^{\text{HN}}\text{-E11}^{\text{C}'}$ H-bonds, all other hydrogen bridges are detectable by distinct resonances in the long-range HNCO (Fig. 1B) at this TFE concentration. Therefore, an α -helical conformation with closed H-bonds is clearly established for residues 2 to 15 at 60% TFE. At 90% TFE, the α -helix is extended by another H-bond ($\text{T5}^{\text{HN}}\text{-G1}^{\text{C}'}$) toward the N terminus. Apparently, this α -helix also comprises the two nonnative residues of the artificially prolonged N terminus in this S-peptide analog.

A quantitative analysis of the individual $^3J_{\text{NC}'}$ coupling constants for the different TFE concentrations is presented in Table 1 and Figure 2A. Whereas no H-bond correlations with absolute values of $^3J_{\text{NC}'}$ constants larger than ~ 0.1 Hz were detected for the peptide in water, the number and strength of observed correlations increases rapidly in the range of 10% to 30% TFE. Consistently, the observed $^3J_{\text{NC}'}$ coupling constants for the central residues K9-R12 are larger than those for the terminal residues, indicating that the helix formation is initiated at the center of the peptide and extends from there to both peptide ends.

Comparison to secondary structure information from $^{13}\text{C}_\alpha$ secondary shifts

Deviations of $^{13}\text{C}_\alpha$ chemical shifts from random coil values are commonly used as an indicator of backbone ϕ - and ψ -angle conformation (Spera and Bax 1991). These secondary shifts adopt values of around 3 ppm for fully formed α -helical conformations (Spera and Bax 1991). Figure 2C shows $^{13}\text{C}_\alpha$ secondary shifts for the S-peptide in water with-

Table 1. $^3J_{\text{NC}'}$ couplings of S-peptide at different TFE concentrations^a

H-donor ^b	TFE (v/v)						
	0%	10%	20%	30%	45%	60%	90%
5	<0.12	<0.09	<0.12	<0.13	<0.14	<0.15	0.18 $\pm$ 0.06
6	<0.11	ov ^c	ov	0.14 $\pm$ 0.06	0.19 $\pm$ 0.06	0.20 $\pm$ 0.06	0.26 $\pm$ 0.05
7	<0.11	ov	<0.13	0.16 $\pm$ 0.05	0.17 $\pm$ 0.06	0.19 $\pm$ 0.07	0.23 $\pm$ 0.05
8	ov	<0.10	<0.13	0.17 $\pm$ 0.05	0.22 $\pm$ 0.05	0.21 $\pm$ 0.07	0.25 $\pm$ 0.05
9	<0.12	<0.11	0.17 $\pm$ 0.05	0.22 $\pm$ 0.05	0.24 $\pm$ 0.05	ov	0.31 $\pm$ 0.04
10	<0.12	0.13 $\pm$ 0.05	0.18 $\pm$ 0.05	0.23 $\pm$ 0.05	0.29 $\pm$ 0.05	0.28 $\pm$ 0.06	0.30 $\pm$ 0.04
11	<0.13	0.14 $\pm$ 0.05	0.20 $\pm$ 0.04	0.27 $\pm$ 0.04	0.34 $\pm$ 0.04	0.36 $\pm$ 0.05	ov
12	<0.13	<0.12	0.15 $\pm$ 0.06	0.20 $\pm$ 0.05	0.26 $\pm$ 0.05	0.29 $\pm$ 0.05	0.33 $\pm$ 0.04
13	<0.13	<0.12	ov	0.19 $\pm$ 0.06	0.26 $\pm$ 0.05	0.27 $\pm$ 0.06	ov
14	<0.14	ov	<0.15	0.17 $\pm$ 0.06	0.19 $\pm$ 0.07	0.18 $\pm$ 0.09	ov
15	<0.13	ov	ov	0.17 $\pm$ 0.05	ov	ov	ov

^a In Hz. $^3J_{\text{NC}'}$ couplings are given as absolute values with errors. The errors are derived by error propagation from 3 standard deviations of the noise (I_{noise}) and the measured amplitude I_{cross} of the cross peak in the cross experiment (error $\approx 0.5J^*I_{\text{noise}}/I_{\text{cross}}$). Statistical errors in the reference experiment are negligible.

^b Residue number of H-bond donor. All H-bonds correspond to α -helical, $\text{HN}(\text{i})\cdots\text{C}'(\text{i}-4)$ connections.

^c Exact coupling constants could not be determined due to resonance overlap.

Fig. 2. Values of individual $^3J_{NC'}$ coupling constants and $^{13}C_{\alpha}$ secondary shifts of the S-peptide for different TFE concentrations. (A) α -helical $HN(i) \cdots C'(i-4)$ $^3J_{NC'}$ coupling constants as a function of residue number and TFE concentration. Overlapping $^3J_{NC'}$ correlations (Table 1) are omitted. Bars indicate upper limits for the coupling constants that were unobservable at 0% TFE. (B) C_{α} secondary shifts for the S-peptide amides at different TFE concentrations relative to 0% TFE. (C) C_{α} secondary shifts at 0% TFE relative to their random coil values (Spera and Bax 1991).

out TFE. Clearly, many of the residues have almost vanishing secondary shifts that are indicative of a random coil conformation. Exceptions are found at the N and C termini as well as for the AAAK motif in the center of the peptide. This central motif shows secondary shifts of about 0.8 ppm, indicating a slight preference for an α -helical structure. With increasing TFE concentration, residues G1 to M15 experience continuously increasing downfield $^{13}C_{\alpha}$ shifts (Fig. 2B). Similarly, these residues show a gradual decrease of $^3J_{HNHA}$ scalar coupling constants for TFE concentrations from 10% to 45% (data not shown), which is consistent with a trend toward helical backbone ϕ angles. At 90% TFE, most of the residues between K3 and M15 have $^{13}C_{\alpha}$ secondary shift values between 3 and 4 ppm, indicative of a fully formed α -helix. In contrast, residues beyond M15 show a much weaker increase in $^{13}C_{\alpha}$ secondary shifts. The secondary shift of residues D16 to A21 increases by less

than 1 ppm over the range from 10% and 60% TFE. This indicates a trend toward, but not the full formation of, a helical structure. Increasing the TFE content from 60% to 90% continues this trend for residues D16 to S18 and leads to an additional secondary shift of about 1 ppm.

Overall, both the H-bond contacts in the HNCO experiment and the $^{13}C_{\alpha}$ secondary shifts establish the gradual formation of the α -helical structure with increasing TFE concentration. In the range 10% to 60% TFE, the formation of all α -helical H-bonds, present in the crystal structure of RNaseA, can be observed unambiguously by $^3J_{NC'}$ correlations. At 90% TFE, the H-bonds of the helix extend from residue G1 to M15. N-terminal fraying of this helix is indicated by the reduced H-bond coupling constants and secondary shifts. C-terminal fraying extends up to residue S18 as indicated by the $^{13}C_{\alpha}$ secondary shifts at 90% TFE. However, no H-bond correlations are found beyond residue M15 in the HNCO experiment, such that the helical conformation is not yet stabilized by a significant presence of H-bonds for residues D16 to S18. This is consistent with the prediction of a helix stop signal at residue D16 derived from the analysis of 1H NMR data (Kim and Baldwin 1984; Nelson and Kallenbach 1989). Very likely, this reduced propensity for helix formation beyond residue M15 is caused by the presence of a contiguous array of hydrophilic, oxygen-containing sidechains in residues D16, S17, S18, T19, and S20. Beyond residue S18, neither H-bond correlations nor secondary shifts indicate any significant amount of α -helical conformation even at high concentrations of TFE.

Comparison to the Lifson-Roig theory of the coil-to-helix transition

Apparently, the state of the S-peptide for the different TFE concentrations is a mixture of different structural conformations. However, only single resonances are observed for all the nuclei of the S-peptide in this mixture. Therefore, the exchange between different conformations, for example, open and closed states of the H-bonds, must be fast on the timescale of chemical shifts, which is on the order of milliseconds for proton resonances. In contrast, absolute sizes of $^3J_{NC'}$ couplings are less than 1 Hz, such that the transfer by these couplings is much slower than the interchange between the different conformations of the peptide. Therefore, $^3J_{NC'}$ values derived from the long-range HNCO experiment correspond to a true-time (and ensemble) average. An exponential correlation between the $^3J_{NC'}$ coupling constant and the $N \cdots O$ distance was described for protein G by the Bax group (Cornilescu et al. 1999b), in which $^3J_{NC'}$ is given as $-59 \cdot 10^3 \text{ Hz} \cdot \exp[-4 \cdot R_{NO}/\text{\AA}]$. Because this dependence on the distance is very steep, it is safe to assume that, in a mixture of random coil and α -helical conformations, only fully formed H-bonds contribute significantly to the H-bond cross peak intensities of the HNCO experiment.

An average value of 0.38 Hz was found for $^{\text{h}^3}\text{J}_{\text{NC}}$ constants in α -helical H-bonds of (fully folded) ubiquitin (Cordier and Grzesiek 1999). Thus, to a first approximation, one could assume that the population p_i of an individual closed H-bond in the S-peptide is given as $^{\text{h}^3}\text{J}_{\text{NC}}/0.38$ Hz.

Figure 3A shows these individual H-bond populations as a function of residue number. The form of this population profile is very close to the bell-shaped profile of the order parameter S^2 of NH bond vector orientations described in a recent NMR relaxation study of backbone HN vector dynamics for the S-peptide in the absence of TFE (Alexandrescu et al. 1998). The motion of HN groups on the pico- to nanosecond timescale was found to be the most restricted ($S^2 > 0.5$) in the center of the peptide between residues 5 and 16, whereas it was least restricted at the peptide ends. These results were interpreted in terms of a model in which compact conformations become increasingly favored moving from the termini to the center of a polypeptide chain. Such a model was rationalized in terms of the relatively low entropic cost of forming contacts in the middle of the chain as compared to contacts between the ends, or between the ends and middle of the chain (Alexandrescu et al. 1998).

A very similar description of the random coil to α -helix transition is contained in the Zimm-Bragg (1959) and Lifson-Roig (1961) matrix theories, which allow the calculation of individual H-bond populations. Although these theories differ to some extent in their assumptions, the underlying mathematics and their results are rather similar (Qian and Schellman 1992). The more convenient formulation for calculating individual H-bond populations is given by the Lifson-Roig theory (Lifson and Roig 1961; Qian and Schellman 1992). This theory assumes three different possible states for a certain residue i : (1) a coil conformation of the backbone with statistical weight u_i ; (2) an α -helical conformation without formation of an H-bond with statistical weight v_i ; and (3) an H-bonded α -helical conforma-

tion with statistical weight w_i . The partition function for the statistical ensemble of peptide conformations is given as:

$$Z = (0 \quad 0 \quad 1) \prod_{i=1}^N M_i \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad (1)$$

in which the product extends over all N residues of the peptide chain and M_i represents a 3×3 Lifson-Roig matrix containing the weights $v_i = v_i'/u_i'$ and $w_i = w_i'/u_i'$, which have been normalized relative to the weight of the coil state. Populations of individual closed H-bonds, p_i , are derived from the partition function by the derivative (Qian and Schellman 1992):

$$p_i = \frac{1}{Z} \frac{\partial Z}{\partial \ln w_i} \quad (2)$$

The results of this theory were compared to the experimentally determined populations by a nonlinear least squares fit. For this procedure, uniform weights v and w were assumed for all residues of the peptide chain. Commonly used values for v range between 0.04 and 0.1 (Qian and Schellman 1992; Rohl et al. 1996), with the results of the theory not being strongly dependent on the precise number (Qian and Schellman 1992). For the fit, v was assumed to be independent of the TFE concentration and arbitrarily fixed to value of 0.05. In contrast, a uniform weight w for the H-bond formation in all residues was optimized for the different TFE concentrations. A peptide chain length N of 14 gave best overall results. Figure 3A shows these fits to the experimental H-bond populations, corresponding to uniform weights w of 1.37, 1.45, 1.54, 1.63, 1.67, and 1.77 for TFE concentrations of 10%, 20%, 30%, 45%, 60%, and

Fig. 3. Fitting of the Lifson-Roig model to experimental H-bond populations $p_i = ^{\text{h}^3}\text{J}_{\text{NC}}/0.38$ for different TFE concentrations. (A) Measured data points are connected by dashes. Data corresponding to the fit of the Lifson-Roig with different average values of the H-bond weight w are represented by continuous lines (see text). (B) The stabilization of the individual H-bond as presented by $RT \ln w$ (see text) as a function of TFE concentration.

90%, respectively. Apparently, the overall trends in the data agree to a reasonable extent with the predictions of this theory, that is, the strongest H-bonds occur in the center of the helix whereas fraying is observed toward the helical ends.

Clearly visible in Figure 3A are also deviations between experimental populations and the theoretical predictions that correspond to a root mean square difference of 0.08. However, such deviations are smaller than the errors in experimentally derived H-bond populations, which range from 0.13 to 0.18 as derived from the noise in the coupling measurement (Table 1). Therefore, at present, the size of the experimental error precludes the use of more detailed models with a larger number of adjustable parameters, such as nonuniform Lifson-Roig statistical weights v and w for individual amino acids or the use of a nonuniform $^{\text{h}3}\text{J}_{\text{NC}'}$ value for a fully closed α -helical H-bond.

A particularly strong deviation between the observed and modeled H-bond population is found for the amide hydrogen of the central residue E11 with consistently larger than expected experimental H-bond populations. One reason for this observation could be an overestimation of the experimental H-bond population because of a $^{\text{h}3}\text{J}_{\text{NC}'}$ constant for the fully closed E11^{HN}-A7^{C'} H-bond that is stronger (more negative) than the α -helical average of -0.38 Hz. Another possible explanation is a stabilization of this H-bond by residue specific interactions (Rico et al. 1983, 1984). In this case, the E11^{HN}-A7^{C'} H-bond would additionally increase the tendency for the helix nucleation in the center of the peptide. In contrast, the population of the H14^{HN}-F10^{C'} H-bond depends only very weakly, and much less than expected, on the TFE concentration (Fig. 3A). Under solution conditions for the S-peptide (Rico et al. 1983) and also in the full structure of RNaseA (e.g., PDB codes 2AAS [Santoro et al. 1993] and 7RSA [Wlodawer et al. 1988]), the side chains of residues H14 and F10 show aromatic stacking interactions. Conceivably, the presence of these interactions could lead to a preferred relative orientation of H14 and F10 thereby reducing the influence of TFE on the stabilization of the H14^{HN}-F10^{C'} H-bond.

Discussion

Standard coil-helix transition theories such as Zimm-Bragg and Lifson-Roig predict that the formation of helical segments is strongly cooperative. Once helix nucleation is achieved, the propagation of a helix is facile and progresses from the nucleation center toward the peptide ends. The present NMR data provide residue-level information about this process and corroborate these theoretical predictions. From the $^{\text{h}3}\text{J}_{\text{NC}'}$ data (Figs. 2, 3), it is clear that the nucleation of the H-bonded helix occurs in the central part of the peptide around residue E11. With increasing TFE concentration, H-bond formation progresses from there to-

ward the N-terminal end and toward the helix stop signal at residue D16. However, even at the highest TFE concentration, the H-bonds are not fully formed at both ends of the α -helix.

The cooperative nature of the coil-to-helix transition is particularly evident from the small changes in the H-bond weight w needed for an individual amino acid to bring the peptide from the coil to the helix state: For the 14 amino acid segment, the transition from almost random coil at 10% TFE to the almost completely formed helix at 90% TFE is caused by a change in the H-bond weight w of only 29%, that is, from 1.37 to 1.77. In the Lifson-Roig theory, the cooperativity of the transition results from the interaction of all the amino acids in the helical state. This interaction is implicitly incorporated into the summation of the partition function. Longer helical segments have a higher number of interacting amino acids and therefore show a sharper transition as a function of the H-bond weight (Qian and Schellman 1992).

The results of the preceding analysis are in good agreement with a study by Luo and Baldwin (1997), who were able to fit the Lifson-Roig theory to circular dichroism data of the TFE-induced coil-to-helix transition for a number of small model peptides over a wide range of experimental parameters. In particular, it was found that within a range of 0 to 4 M ($\sim 25\%$) TFE, the average of the H-bond weight $\langle w \rangle$ obeys a relation of the type:

$$RT \cdot \ln \langle w \rangle = RT \cdot \ln \langle w_0 \rangle - mC \quad (3)$$

in which C is the TFE concentration, m and $\langle w_0 \rangle$ are constants, and R and T have their usual meaning. At higher TFE concentrations, $\langle w \rangle$ quickly reaches a saturation level in a number of homogeneous model peptides (Luo and Baldwin 1997). Figure 3B shows this dependence on TFE concentration for the values $RT \ln w$ derived from the fit of the S-peptide J-coupling data in Figure 3A. Within the experimental error, the dependence is close to linear between 10% and 30% TFE and begins to saturate at higher concentrations. The slope within this linear regime is similar to the data reported by Luo and Baldwin (1997), such that the increase from 10% to 30% TFE corresponds to a free-energy stabilization per single H-bond of approximately 0.07 kcal/mole. Beyond this linear regime, the transition to the saturation is not as sharp as previously reported (Luo and Baldwin 1997), and an additional stabilization of 0.08 kcal/mole is still observed between 45% and 90% TFE (Fig. 3B). This different saturation behavior cannot be ascribed to the inaccuracies of the fit in Figure 3A, because the experimentally observed increase in the S-peptide H-bond populations is statistically significant between 45% and 90% TFE. Similar significant changes are reported for circular dichroism data of the S-peptide at these TFE concentrations (Nelson

and Kallenbach 1986). Therefore, these observations reflect genuine differences between the S-peptide and the peptides investigated by Luo and Baldwin. In the latter case, very homogeneous amino acid compositions were used, and the helix formation occurred over the full peptide lengths. In our case, the amino acid sequence is heterogeneous and contains the helical stop signal at D16. Such heterogeneities are expected to “smear out” the transition between the coil and the helix state. Overall, the increase from 10% to 90% TFE stabilizes a single H-bond by 0.15 kcal/mole (Fig. 3B). For the assumed chain of 14 residues containing 12 H-bonds, this corresponds to a total stabilization of the α -helical structure by 1.8 kcal/mole or $3.1 \cdot RT$. The modest size of this stabilization clearly illustrates the delicate balance between folded and unfolded states of polypeptide chains.

It is relevant to note that all the observed $^3J_{\text{NC}'}$ correlations correspond to genuine (i,i + 4) α -helical H-bonds and not to other types of H-bond connections. The detection limit for the $^3J_{\text{NC}'}$ correlations in the present experiments is approximately 0.1 Hz. This implies that populations of less than about 20% to 30% for H-bonds in typical secondary structure elements with a relatively straight $\text{NH}\cdots\text{O}=\text{C}$ geometry, and standard donor-acceptor distances would not lead to detectable correlations. On the other hand, it can be concluded that the populations of other types of regular H-bonds, for example, of H-bonds corresponding to 3_{10} -helices, cannot be considerably larger than 30% in the S-peptide at all TFE concentrations. Such 3_{10} -helical H-bonds have been proposed for short alanine peptides based on CD, ESR, and ^1H NMR experiments (Miick et al. 1992; Millhauser et al. 1997). Similarly, a set of turnlike structures was proposed by Dyson et al. (1988) as an early structure for small peptides, and MD simulations (Tobias and Brooks 1991; Smith et al. 1998) indicate reverse turns as important intermediates along the helix folding/unfolding pathway. Our experiments cannot exclude such types of structure as long as their populations for regular, straight H-bonds are smaller than about 30%. However, this statement does not preclude the existence of larger populations of H-bonds with a non-linear $\text{NH}\cdots\text{O}=\text{C}$ geometry, such as bifurcated H-bonds, because recent evidence indicates that the H-bond scalar couplings are strongly reduced when the $\text{H}\cdots\text{O}=\text{C}$ angle deviates substantially from 180° (Dingley et al. 2000).

In summary, we have shown that the long-range HNCO experiment cannot only be used for the detection of H-bonds in folded proteins, but also in flexible peptides with a large proportion of unfolded structure. The value of $^3J_{\text{NC}'}$ constants provides a measure for the population of a closed H-bond that can be quantitatively compared to the predictions of coil-to-helix transition theories. Future applications will include a more thorough testing of this experimental and theoretical approach for a variety of different solvent conditions and designed polypeptide sequences.

Materials and methods

Materials

Expression and purification of S-peptide was performed as described previously (Alexandrescu et al. 1998). Compared with the original S-peptide of RNaseA, this peptide contains two additional residues at its N terminus such that the entire sequence is given as GSKETAAAKF ERQHMSSTS AA. Residue numbering in the present paper starts with the N-terminal glycine and is thus shifted by two residues relative to the RNaseA numbering scheme (e.g., residue H14 in the recombinant peptide corresponds to residue H12 in the wild-type sequence). 2,2,2-trifluoroethanol- d_2 -OH (TFE) was purchased from Euriso-top. NMR samples were prepared in Shigemi microtubes with volumes of 300 μL containing 3.4 mM uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched S-peptide (~ 2 mg) in 0% to 90% vol TFE/water, 50 mM NaCl, 95% $\text{H}_2\text{O}/5\%$ D_2O . For every concentration of TFE, the pH was adjusted individually to a value of 3.7 (uncorrected meter reading). No additional buffer was used.

NMR methods

The NMR experiments were performed at 25°C on a 600 Bruker DMX spectrometer, equipped with a triple resonance, z-axis pulsed field gradient probehead. Measurement of individual $^3J_{\text{NC}'}$ couplings was performed by using the quantitative J-correlation, long-range HNCO experiment as described previously (Cordier and Grzesiek 1999). Both reference and through-H-bond cross correlation spectra were recorded as 50° (t_1 , $^{13}\text{C}'$) $\times$ 1024° (t_2 , $^1\text{H}^{\text{N}}$) data matrices (in which n° refers to complex points) with acquisition times of 36 msec (t_1) and 86 msec (t_2). Total experimental times were 58 hr and 0.5 hr for each cross and reference experiment, respectively. The ^{15}N - $^{13}\text{C}'$ de/rephasing delay (T) was set to a value of 129 msec for optimal suppression of sequential $^1J_{\text{NC}'}$ -interactions. Data were processed using the NMRPipe processing package (Delaglio et al. 1995), and peak intensities and positions were determined with the program PIPP (Garrett et al. 1991). The absolute value of $^3J_{\text{NC}'}$ was obtained from the ratio of the intensities in the cross and reference spectra using the relation $J_{\text{NC}'} = (I_{\text{cross}}/I_{\text{ref}})^{1/2}/(\pi T)$ (Cordier and Grzesiek 1999). In cases in which cross peak amplitudes were too small to be detected, upper limits for the values of the coupling constants were derived from the intensity in the reference spectra as described previously (Dingley et al. 1999). $^1\text{H}^{\text{N}}$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}'$, and ^{15}N backbone assignments were made using 2D and 3D versions of HNCO, HNCA, and HN(CA)CO experiments. ^1H chemical shifts were referenced relative to external TSP. For this purpose a thin, closed glass capillary containing TSP was placed into the center of the NMR tube containing the S-peptide sample. ^{13}C and ^{15}N chemical shifts were referenced indirectly relative to the TSP ^1H frequency (Markley et al. 1998).

Acknowledgments

We thank Richard Kammerer for the gift of the S-peptide clone and Klara Rathgeb-Szabo for the peptide purification, as well as Thomas Kiefhaber and Florence Cordier for very valuable and stimulating discussions. This work was supported by SNF grant 31-43'091.95 grant to A.T.A and S.G.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Alexandrescu, A.T., Rathgeb-Szabo, K., Rumpel, K., Jahnke, W., Schulthess, T., and Kammerer, R.A. 1998. ¹⁵N backbone dynamics of the S-peptide from ribonuclease A in its free and S-protein bound forms: Toward a site-specific analysis of entropy changes upon folding. *Protein Sci.* **7**: 389–402.
- Balbach, J., Forge, V., Lau, W., Jones, J., van Nuland, N., and Dobson, C. 1997. Detection of residue contacts in a protein folding intermediate. *Proc. Natl. Acad. Sci.* **94**: 7182–7185.
- Balbach, J., Forge, V., Lau, W.S., van Nuland, N.A., Brew, K., and Dobson, C.M. 1996. Protein folding monitored at individual residues during a two-dimensional NMR experiment. *Science* **274**: 1161–1163.
- Baldwin, R.L. 1975. Intermediates in protein folding reactions and the mechanism of protein folding. *Annu. Rev. Biochem.* **44**: 453–475.
- Buck, M. 1998. Trifluoroethanol and colleagues: Cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* **31**: 297–355.
- Chiti, F., Taddei, N., Webster, P., Hamada, D., Fiaschi, T., Ramponi, G., and Dobson, C.M. 1999. Acceleration of the folding of acylphosphatase by stabilization of local secondary structure. *Nat. Struct. Biol.* **6**: 380–387.
- Cordier, F. and Grzesiek, S. 1999. Direct observation of hydrogen bonds in proteins by interresidue ³J_{NC'} scalar couplings. *J. Am. Chem. Soc.* **121**: 1601–1602.
- Cornilescu, G., Hu, J.S., and Bax, A. 1999a. Identification of the hydrogen bonding network in a protein by scalar couplings. *J. Am. Chem. Soc.* **121**: 2949–2950.
- Cornilescu, G., Ramirez, B.E., Frank, M.K., Clore, G.M., Gronenborn, A.M., and Bax, A. 1999b. Correlation between ³J_{NC'} and hydrogen bond length in proteins. *J. Am. Chem. Soc.* **121**: 6275–6279.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**: 277–293.
- Dingley, A.J., Masse, J.E., Feigon, J., and Grzesiek, S. 2000. Characterization of the hydrogen bond network in guanosine quartets by internucleotide ³J_{NC'} and ²J_{NN} scalar couplings. *J. Biomol. NMR* **16**: 279–289.
- Dingley, A.J., Masse, J.E., Peterson, R.D., Barfield, M., Feigon, J., and Grzesiek, S. 1999. Internucleotide scalar couplings across hydrogen bonds in Watson-Crick and Hoogsteen base pairs of a DNA triplex. *J. Am. Chem. Soc.* **121**: 6019–6027.
- Dyson, H.J., Rance, M., Houghten, R.A., Wright, P.E., and Lerner, R.A. 1988. Folding of immunogenic peptide fragments of proteins in water solution. II. The nascent helix. *J. Mol. Biol.* **201**: 201–217.
- Epstein, H.F., Schechter, A.N., Chen, R.F., and Anfinsen, C.B. 1971. Folding of staphylococcal nuclease: Kinetic studies of two processes in acid renaturation. *J. Mol. Biol.* **60**: 499–508.
- Forge, V., Wijesinha, R.T., Balbach, J., Brew, K., Robinson, C.V., Redfield, C., and Dobson, C.M. 1999. Rapid collapse and slow structural reorganization during the refolding of bovine alpha-lactalbumin. *J. Mol. Biol.* **288**: 673–688.
- Garrett, D.S., Powers, S., Gronenborn, A.M., and Clore, G.M. 1991. A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Res.* **95**: 214–220.
- Hamada, D., Kuroda, Y., Tanaka, T., and Goto, Y. 1995. High helical propensity of the peptide fragments derived from beta-lactoglobulin, a predominantly beta-sheet protein. *J. Mol. Biol.* **254**: 737–746.
- Hoeltzli, S. and Frieden, C. 1998. Refolding of [6-19F]tryptophan-labeled Escherichia coli dihydrofolate reductase in the presence of ligand: A stopped-flow NMR spectroscopy study. *Biochemistry* **37**: 387–398.
- Kiefhaber, T., Labhardt, A., and Baldwin, R. 1995. Direct NMR evidence for an intermediate preceding the rate-limiting step in the unfolding of ribonuclease A. *Nature* **375**: 513–515.
- Kim, P.S. and Baldwin, R.L. 1984. A helix stop signal in the isolated S-peptide of ribonuclease A. *Nature* **307**: 329–334.
- Lifson, S. and Roig, A. 1961. On the theory of helix-coil transitions in polypeptides. *J. Chem. Phys.* **34**: 1963–1974.
- Luo, P. and Baldwin, R.L. 1997. Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **36**: 8413–8421.
- Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E., and Wuthrich, K. 1998. Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J. Mol. Biol.* **280**: 933–952.
- McDonald, C.C., Phillips, W.D., and Glickson, J.D. 1971. Nuclear magnetic resonance study of the mechanism of reversible denaturation of lysozyme. *J. Am. Chem. Soc.* **93**: 235–246.
- Miick, S.M., Martinez, G.V., Fiori, W.R., Todd, A.P., and Millhauser, G.L. 1992. Short alanine-based peptides may form 3(10)-helices and not alpha-helices in aqueous solution. *Nature* **359**: 653–655.
- Millhauser, G.L., Stenland, C.J., Hanson, P., Bolin, K.A., and van de Ven, F.J. 1997. Estimating the relative populations of 3(10)-helix and alpha-helix in Ala-rich peptides: a hydrogen exchange and high field NMR study. *J. Mol. Biol.* **267**: 963–974.
- Mok, Y.-K., Kay, C., Kay, L.E., and Forman-Kay, J. 1999. NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol.* **289**: 619–638.
- Nelson, J.W. and Kallenbach, N.R. 1986. Stabilization of the ribonuclease S-peptide alpha-helix by trifluoroethanol. *Proteins* **1**: 211–217.
- . 1989. Persistence of the alpha-helix stop signal in the S-peptide in trifluoroethanol solutions. *Biochemistry* **28**: 5256–5261.
- Neri, D., Billeter, M., Wider, G., and Wuthrich, K. 1992. NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science* **257**: 1559–1563.
- Qian, H. and Schellman, J.A. 1992. Helix-coil theories: A comparative study for finite length polypeptides. *J. Phys. Chem.* **96**: 3987–3994.
- Rico, M., Gallego, E., Santoro, J., Bermejo, F.J., Nieto, J.L., and Herranz, J. 1984. On the fundamental role of the Glu 2–... Arg 10+ salt bridge in the folding of isolated ribonuclease A S-peptide. *Biochem. Biophys. Res. Commun.* **123**: 757–763.
- Rico, M., Nieto, J.L., Santoro, J., Bermejo, F.J., Herranz, J., and Gallego, E. 1983. Low-temperature ¹H-NMR evidence of the folding of isolated ribonuclease S-peptide. *FEBS* **162**: 314–319.
- Roder, H., Elove, G.A., and Englander, S.W. 1988. Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature* **335**: 700–704.
- Roder, H. and Wuthrich, K. 1986. Protein folding kinetics by combined use of rapid mixing techniques and NMR observation of individual amide protons. *Proteins* **1**: 34–42.
- Rohl, C., Chakrabarty, A., and Baldwin, R. 1996. Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci.* **5**: 2623–2637.
- Santoro, J., Gonzalez, C., Bruix, M., Neira, J.L., Nieto, J.L., Herranz, J., and Rico, M. 1993. High-resolution three-dimensional structure of ribonuclease A in solution by nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* **229**: 722–734.
- Smith, L.J., Mark, A.E., Dobson, C.M., and van Gunsteren, W.F. 1998. Molecular dynamics simulations of peptide fragments from hen lysozyme: insight into non-native protein conformations. *J. Mol. Biol.* **280**: 703–719.
- Spera, S. and Bax, A. 1991. Empirical correlation between protein backbone conformation and C.alpha. and C.beta. ¹³C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**: 5490–5492.
- Timasheff, S.N. 1970. Protein-solvent interactions and protein conformation. *Acc. Chem. Res.* **3**: 62–68.
- Tobias, D.J. and Brooks, C.L. 1991. Thermodynamics and mechanism of alpha helix initiation in alanine and valine peptides. *Biochemistry* **30**: 6059–6070.
- Udgaonkar, J.B. and Baldwin, R.L. 1988. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature* **335**: 694–699.
- Wagner, G. and Wuthrich, K. 1982. Amide protein exchange and surface conformation of the basic pancreatic trypsin inhibitor in solution. Studies with two-dimensional nuclear magnetic resonance. *J. Mol. Biol.* **160**: 343–361.
- Westmoreland, D.G. and Matthews, C.R. 1973. Nuclear magnetic resonance study of the thermal denaturation of ribonuclease A: Implications for multistate behavior at low pH. *Proc. Natl. Acad. Sci.* **70**: 914–918.
- Wlodawer, A., Svensson, L.A., Sjolin, L., and Gilliland, G.L. 1988. Structure of phosphate-free ribonuclease A refined at 1.26 Å. *Biochemistry* **27**: 2705–2717.
- Zimm, B.H. and Bragg, J.K. 1959. Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem Phys* **31**: 526–535.