Jixun Dai
Xu Wang
Yingang Feng
Guibao Fan
Jinfeng Wang
National Laboratory of Biomacromolecules, Center for Molecular Biology, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, Peoples Republic of China

Searching for Folding Initiation Sites of Staphylococcal Nuclease: A Study of N-Terminal Short Fragments

Received 28 March 2004; accepted 22 June 2004
Published online 9 August 2004 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.20121

Abstract: *The N-terminal short fragments of staphylococcal nuclease (SNase), SNase20, SNase28, and SNase36, corresponding to the sequence regions, Ala1–Gly20, Ala1–Lys28, and Ala1–Leu36, respectively, as well as an 8-residue peptide (Ala17–Ile18–Asp19–Gly20–Asp21–Thr22–Val23–Lys24) have been synthesized. The conformational states of these fragments were investigated using CD and NMR spectroscopy in aqueous solution and in trifluoroethanol (TFE)–H_2O mixture. SNase20 containing a sequence corresponding to a bent peptide in native SNase shows a transient population of bend-like conformation around Ala12–Thr13–Leu14 in TFE–H_2O mixture. The sequence region of Ala17–Thr22 of SNase28 displays a localized propensity for turn-like conformation in both aqueous solution and TFE—H_2O mixture. The conformational ensemble of SNase36 in aqueous solution includes populated turn-like conformations localized in sequence regions Ala17–Thr22 and Tyr27–Gln30. The analysis suggests that these sequence regions, which form the regular secondary structures in native protein, may serve as the folding nucleation sites of SNase fragments of different chain lengths starting from the N-terminal end. Thus, the formation of bend- and turn-like conformations of these sequence regions may be involved in the early folding events of the SNase polypeptide chain in vitro.* © 2004 Wiley Periodicals, Inc. Biopolymers 75: 229–241, 2004

Keywords: *staphylococcal nuclease; propensity; fragments; initiation sites; folding*

INTRODUCTION

One of the fundamental questions in protein folding concerns the early events in the folding process of the polypeptide chain. Many studies are devoted to identifying protein folding initiation sites.[1–3] By studying peptide fragments of proteins it was suggested that reverse turns are likely sites for initiation of protein

Correspondence to: Jinfeng Wang; email: jfw@sun5.ibp.ac.cn
Contract grant sponsor: The Pandeng Project of the National Commission for Science and Technology and the National Natural Science Foundation of China (NNSFC). Contract grant number: 39823001 (NNSFC)

folding,[4,5] and they have also been implicated in folding processes involving helices.[6]

The folding initiation sites of staphylococcal nuclease (SNase) have been studied by several groups. The NMR study of a large SNase fragment, $\Delta 131\Delta$, has reported that three-stranded antiparallel β-sheet (β_1–β_2–β_3) persisted at high urea concentration, and segment Leu36–Val39 may stabilize the β-sheet by local non-native interaction.[7,8] The pulsed hydrogen–deuterium exchange method was used to probe the protection of ^1H_N from solvent exchange of SNase variants, H124LSNase and P117GSNase. The results have indicated that the β-hairpin composed by strands β_2 and β_3 (27YKGQP31) is among the earliest structural events in the folding of SNase.[9,10] The NMR studies of two peptides, including the sequence for β-hairpin formed by β_2 and β_3 (27YKGQP31), have suggested that formation of this β-hairpin could be one early event in the folding of SNase.[11,12] The native-like β-turn conformations of the segments Ile18–Asp21 and Tyr27–Gln30 were observed for the partially folded SNase110, SNase121, and SNase135 fragments containing 1–110, 1–121, and 1–135 residues of SNase, respectively, based on the native-like dispersion of the cross peaks in corresponding two-dimensional (2D) heteronuclear singe quantum correlation (HSQC) spectra.[13] Residues Ile18–Asp21, which are in a β-hairpin formed by β_1 and β_2, and Tyr27–Gln30 are within the first 30 residues in SNase sequence. Thus, it is interesting to study the conformational states of N-terminal short fragments of SNase for analyzing the early events in the folding of SNase polypeptide chain of different chain lengths starting from the N-terminal end of the protein in vitro.

Three N-terminal short fragments of SNase have been synthesized. They are SNase20, SNase28, and SNase36 containing residues Ala1–Gly20, Ala1–Lys28, and Ala1–Leu36, respectively, in sequence. A short peptide SNaseτ_1 containing residues Ala17–Lys24 was also synthesized for further investigation. The conformational states of these fragments have been studied by CD and by 2D homonuclear NMR spectroscopy in aqueous solution or in a TFE–H_2O mixture (TFE: trifluoroethanol). The localized structural propensities of bend- and turn-like conformations are found for segments Ala17–Thr22 and Tyr27–Gln30, as well as Ala12–Thr13–Leu14 of SNase20, SNase28, and SNase36, which can be considered as the folding initiation sites of SNase polypeptide chains in vitro.

MATERIALS AND METHODS

Peptide Synthesis and Purification

The peptides corresponding to the N-terminal short fragments of SNase, SNase20, and SNase28 were synthesized on an Applied Biosystems 431A peptide synthesizer using the 9-fluorenylmethyoxycarbonyl (Fmoc)-protection solid-phase strategy.[14,15] The crude peptide products were further purified by high performance liquid chromatography (HPLC) on a PepRPC10 column and the elution buffer was 10 m*M* HCl/acetonitrile. The pure peptide products were sequenced on a Beckman 6300 amino acid analyzer. SNase36 and SNaseτ_1 were synthesized by Genemed Biotechnologies, Inc. Both N- and C-terminals of all peptides were linked with no protective groups to ensure the solubility of the peptides. The purities of the peptides were identified by mass spectrometry, and the molecular weight (MW) of SNase20, SNase28, SNase36, and SNaseτ_1 are 2122.9, 3102.1, 4033.0, and 817.4 Dalton, respectively.

CD Spectroscopy

Far-UV CD spectra of peptides were measured on a Jasco-720 spectropolarimeter at 293 K. Data were recorded over the ranges of 190–260 nm with a 0.1 nm path length. Four scans were averaged for each measurement. The peptides were dissolved in 50 m*M* NaAc/HAc buffer, or in a TFE–H_2O mixture (pH 5.0) with a concentration of 1.0 and 0.75 mg/mL, respectively, except the samples of SNase36 and SNaseτ_1, which were not prepared for the TFE study. The secondary structure contents of the fragments in the TFE–H_2O mixture were estimated by analyzing the CD spectra using software package CONTIN.[16]

NMR Spectroscopy

Two sets of NMR samples were prepared: in the first, the peptides were dissolved in 90% H_2O/10% D_2O solution containing 50 m*M* d4-acetate buffer (pH 5.0) to a final concentration of about 2 m*M*; in the second, the peptides, SNase20 and SNase28, were dissolved in the TFE–H_2O mixture (the sample of SNase36 in the TFE–H_2O mixture was not suitable for NMR study), and the pH was adjusted to a value of 5.0. Added to all NMR samples was 0.003% NaN_3.

The 2D ^{1}H-NMR experiments, correlation spectroscopy (COSY), double quantum filter (DQF)-COSY, total correlation spectroscopy (TOCSY), nuclear Overhauser enhancement spectroscopy (NOESY), and rotating frame NOESY (ROESY) were performed on a Bruker DMX 600 MHz spectrometer at 278 K for the peptides in H_2O and at 300 K for the peptides in the TFE–H_2O mixture. TOCSY data were recorded at the mixing times of 50, 80, and 100 ms. NOESY data were collected at the mixing times of 100, 200, 300, and 400 ms, while the mixing time for ROESY experiments was 200 ms.

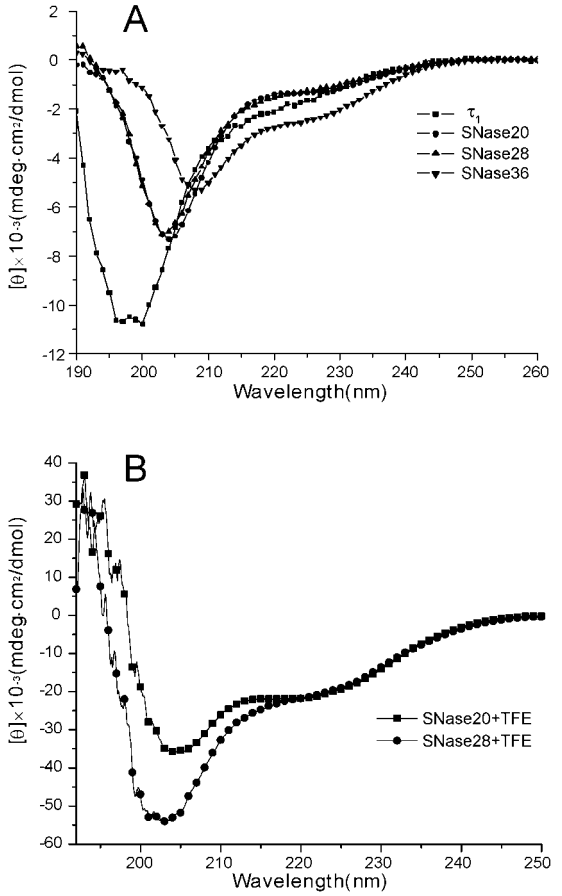


FIGURE 1 Far-UV CD spectra. (A) Peptides, SNase20, SNase28, SNase36, and SNaseτ_1, in 50 m*M* sodium acetate, pH 5.0, at 298 K. (B) Peptides, SNase20 and SNase28 in the TFE–H_2O mixture.

All NMR data were processed and analyzed using FELIX 98 (MSI/Accelrys, Inc.) on SGI workstation ONYX. ^{1}H chemical shifts were referenced to 0 ppm with internal 2, 2-dimethyl-2-silapentane-5-sulfonic acid (DSS).

^{1}H-NMR experiments were also performed for the samples containing 0.1 m*M* peptide fragments to ascertain the monomeric state of all peptides. No significant changes in ^{1}H spectra were found compared with the spectra of the 2 m*M* protein samples, which exclude the possibility of self-association effect of the polypeptides in the experiments above mentioned.

RESULTS

Conformation of SNase20 in H_2O and TFE–H_2O Mixtures

The CD spectra of SNase20 in aqueous buffer (Figure 1A) and in TFE–H_2O mixture (Figure 1B) are not consistent with pure random coil, α-helix, or β-sheet spectra, but rather with a mixture of structured and unstructured conformations. The appearance of weak negative band at 220 nm for SNase20 in a 40% (v/v) TFE–H_2O mixture suggests that TFE may induce a transient population of a structured conformation in SNase20. The CD spectra such as this may depict transient helix-, turn-, or bend-like conformations populated in peptides, since a type I β-turn generally exhibits a CD spectrum similar to that of an α-helix.[17] Analysis of the CD data gave 5% helix, 23% turn, and 37% random coil for SNase20 in the TFE–H_2O mixture.

Almost all ^{1}H resonances were assigned for SNase20 using standard 2D ^{1}H-NMR spectroscopy. The ^{1}H chemical shifts are shown in Table I. There is little chemical shift dispersion (7.9–8.9 ppm) in the amide proton resonance region of the NMR spectra of this peptide. The identified $d_{\alpha N}(i, i+1)$ and $d_{NN}(i, i+1)$ NOE connectivities in the fingerprint region and in the amide proton resonance region of NOESY spectrum, respectively, are shown in Figure 2. The strong $d_{\alpha N}$ $(i, i+1)$ NOEs are observed for segments Thr2–Thr4, Leu7–Glu10, and Pro11–Gly20 (Figure 2A), and the weak $d_{NN}(i, i+1)$ NOEs are obtained for two pairs of residues Leu7–His8 and Lys9–Glu10, and for segments Ala12–Lys16 and Ile18–Gly20 (Figure 2B). The weak $d_{NN}(i, i+1)$ NOEs and strong $d_{\alpha N}$ $(i, i+1)$ NOEs, appearing in the similar sequence regions, indicate that a populated helical-like conformation is in dynamic equilibrium with random coil for segment of Ala12–Gly20 of SNase20 in aqueous solution. Moreover, a number of minor cross peaks were observed for His8, Glu10, Thr13, and Asp19 in the TOCSY spectrum (data not shown), probably due to *cis–trans* isomerization of X-prolyl bond of Glu10–Pro11 and coil–helix transition of the peptide.

The 2D ^{1}H-NMR spectra were recorded also for SNase20 in 40% (v/v) TFE–H_2O mixture. The chemical shifts of ^1H_N and $^1H_\alpha$ resonances are given in Table II. In the fingerprint region of the NOESY spectrum (Figure 3A), the strong $d_{\alpha N}$ $(i, i+1)$ NOE connectivities were identified for segments, Thr2–Leu7, Lys9–Glu10, and Pro11–Gly20. In the ^1H_N–^1H_N resonance region of the NOESY spectrum (Figure 3B), $d_{NN}(i, i+1)$ NOE connectivities were determined for segments Lys6–Glu10, Ala12–Lys16, and Ala17–Gly20, which are similar to those of SNase20 in aqueous solution. However, the intensities of these NOEs are relatively stronger for peptide in the TFE–H_2O mixture than that in the H_2O. Moreover, the weak $d_{NN}(i, i+2)$ NOE between residues Ala12 and Leu14 (Figure 3B), and weak $d_{\beta N}$ $(i, i+2)$ NOEs between Ala12 and Leu14 and between Pro11 and

Table I ^{1}H Chemical Shifts of SNase20 in Aqueous Solution (pH 5.0) at 278 K

Residue	NH	αH	βH	Others
T2	8.74	4.39	4.17	γCH_3 = 1.24
S3	8.67	4.56	3.92, 3.85	
T4	8.46	4.32	4.24	γCH_3 = 1.21
L7	8.42	4.29	1.59	γH = 1.48 δCH_3 = 0.83, 0.91
H8	8.63	4.67	3.18, 3.22	4H = 7.28 2H = 8.60
K9	8.51	4.30	1.76	γCH_2 = 1.39 δCH_2 = 1.68 εCH_2 = 2.98 εNH_3 = 7.60
E10	8.71	4.58	1.87, 2.08	γCH_2 = 2.36
P11		4.40	2.31, 2.04	γCH_2 = 1.93 δCH_2 = 3.83, 3.83
A12	8.58	4.32	1.40	
T13	8.22	4.27	4.15	γCH_3 = 1.20
L14	8.38	4.36	1.60	γH = 1.60 δCH_3 = 0.84, 0.92
I15	8.28	4.12	1.82	γCH_2 = 1.18, 1.48 δCH_2 = 0.88
K16	8.48	4.28	1.78	γCH_2 = 1.43 δCH_2 = 1.71 εCH_2 = 3.00 εNH_3 = 7.60
A17	8.53	4.31	1.39	
I18	8.32	4.13	1.87	γCH_2 = 1.20, 1.47 δCH_3 = 0.92
D19	8.48	4.67	2.67, 2.74	
G20	7.94	3.76		

Thr13 (Figure 3C), were induced by 40% (v/v) TFE. These NOEs illustrate a bent backbone of the segment Ala12–Thr13–Leu14, indicating that TFE induces a bent peptide around residue Thr13.

Secondary chemical shifts can give an indication of α-helix and β-strand structures in protein.[18–21] Figure 4 shows secondary chemical shifts of $^1H_\alpha$ resonances (ΔH_α) of SNase fragments in H_2O and TFE–H_2O mixtures. The changes in the $^1H_\alpha$ chemical shift upon addition of TFE [δH_α= $^1H_\alpha$ (TFE–H_2O)–$^1H\alpha$ (H_2O)] are presented in Figure 4 as well. As is well known, $^1H_\alpha$ experiences upfield shift in the α-helix, but downfield shift in the β-strand, and a piece of helical segment is difficult to differentiate from β-turn or bent peptide. Therefore, a group of four relatively large negative $\Delta H\alpha$ and δH_α values (>0.1 ppm), induced by TFE, indicates a tendency to form local structured conformations in segment Ala12–Ile15 of SNase20, which includes a bent peptide Ala12–Thr13 in native SNase.

Conformation of SNase28 in H_2O and in a TFE–H_2O Mixture

SNase28 provided a similar CD spectrum to those of SNase20 in aqueous solution (Figure 1A). Addition of TFE induces a negative band and a shoulder at 204 and 220 nm, respectively, in the CD spectrum of SNase28 (Figure 1B). Analysis of the CD spectra of SNase28 in the TFE–H_2O mixture gave 4% helix, 23% turn, and 36% random coil. This implies that SNase28 in the TFE–H_2O mixture may have a population of structured conformations.

Nearly complete ^1H resonances of SNase28 in aqueous solution were assigned (Table III). The $d_{\alpha N}$ (i, i+1) NOEs were observed for segments, Thr2–Thr4, Leu7–Glu10, and Pro11–Lys28 (Figure 5A), and $d_{NN}(i, i+1)$ NOEs were determined for residues Ser3–Thr4, Leu7–His8, Lys9–Glu10, Ala12–Thr13–Leu14, Ile15–Lys16, Ala17–Ile18, Asp21–Thr22, Lys24–Leu25, and Tyr27–Lys28 (data not shown). This indicates that the segment prior to Lys16 of SNase28 has very similar backbone conformations to those of SNase20 in aqueous solution. However, the $d_{\beta N}$ (i, i+1) NOEs of Asp19–Gly20 and Asp21–Thr22 (Figure 5B), and a $d_{\alpha N}$ (i, i+2) NOE of Gly20–Thr22 (Figure 5A) in the sequence region after Ala17 were obtained for SNase28 in aqueous solution. These NOEs reveal the conformational differences between SNase28 and SNase20, and indicate a populated turn-like conformation for the segment of Ala17–Thr22 in SNase28.

The standard 2D ^1H-NMR experiments were run with SNase28 in a 20% (v/v) TFE–H_2O mixture.^1H_N and $^1H_\alpha$ chemical shifts are listed in Table II. The $d_{\alpha N}$ (i, i+1) NOE connectivities were determined for segments Thr4–Leu7, Pro11–Asp19, Gly20–Asp21, and Thr22–Tyr27 (data not shown), that span nearly the whole sequence of SNase28. However, $d_{NN}(i, i+1)$ NOEs were determined for residues Ala12–Leu14 in segment Ala1–Ile18, and for Ile18–Val23 and Met26–Lys28 in segment Ile18–Lys28. Two long-range $d_{NN}(i, j)$ NOEs were assigned to pairs of resi-

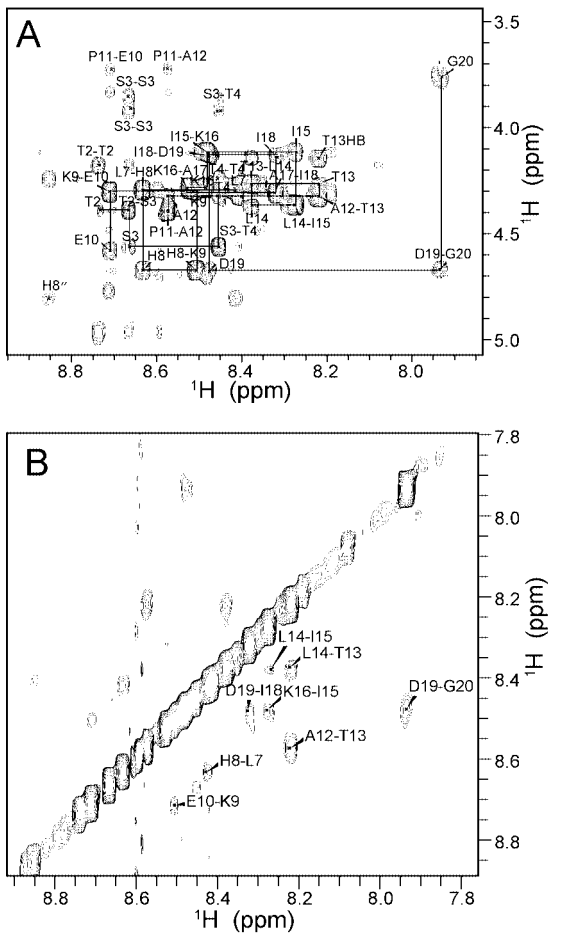


FIGURE 2 The 2D ^{1}H-NOESY spectrum of SNase20 in aqueous solution (pH 5.0) at 278 K. The spectrum was recorded with the mixing time of 300 ms. (A) NOEs in the fingerprint resonance region. (B) NOEs in the amide proton resonance region. The $d_{\alpha N}$ NOE connectivities are presented by continuous lines and arrows. The d_{NN} NOEs are indicated by arrows.

dues Leu14–Leu25 and Ala17–Val23 (Figure 6A). In addition to these NOEs, two $d_{\beta N}$ (i, $i+2$) NOEs were observed between Ala12–Leu14 and between Ala17–Asp19 (Figure 6B). These NOE features suggest that the backbone of segment Leu14–Leu25 of SNase28 in the presence of 20% TFE tends to close up at both N- and C-termini of the segment, and there is a localized structural propensity of reverse turn-like conformation around the segment Ala17–Thr22. The data also indicate a TFE-induced bend at Ala12–Thr13–Leu14 of SNase28.

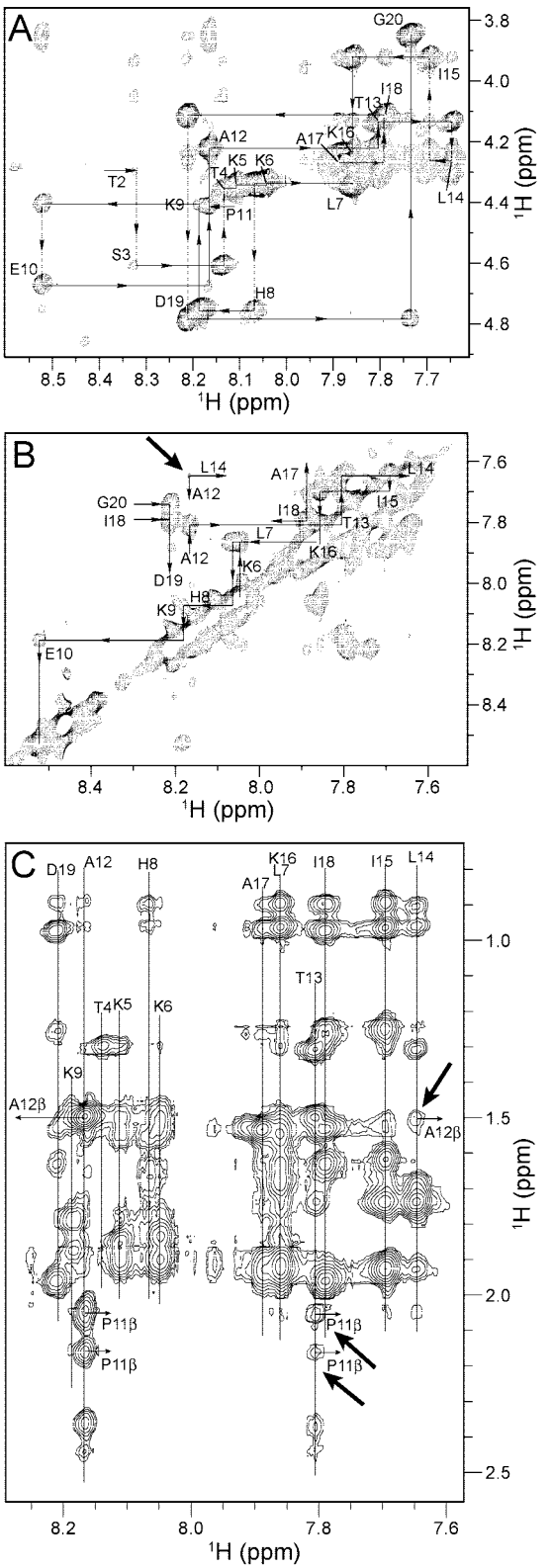
The $\Delta H\alpha$ and δH_α values for residues Ala12–Ile15 of SNase28 in the presence of 20% TFE are very similar to those of SNase20 (Figure 4), indicating a structured polypeptide bend at this segment. The ΔH_α values of residues of SNase28 in H_2O give an indication of a similar conformation around segment Asp19–Val23. However, addition of TFE increases the upfield shift of $^1H_\alpha$ resonances of Thr22–Val23 and Ile15–Lys16, indicated by the values of δH_α for SNase28 in Figure 4. This implies that TFE induces structured conformations in regions adjacent to Ala17–Thr22, which may cause the segment Ala17–Thr22 to have a reverse turn-like conformation.

Conformation of SNase36 in H_2O

A negative band and a shoulder at the wavelengths of 208 and 222 nm, respectively, are observed in the CD spectrum of SNase36 in H_2O (Figure 1A). Analysis of the CD data of SNase36 in H_2O gave 3% helix, 19% turn, and 39% random coil, indicating negligible helix content of SNase36. In the native SNase, there is no

Table II Chemical Shifts of ^1H_N and $^1H_\alpha$ Resonances of SNase20 and SNase28 in a TFE–H_2O Mixture

	SNase20 (TFE–H_2O)		SNase28 (TFE–H_2O)	
Residue	NH	αH	NH	αH
A1				
T2				
S3	8.33	4.61	8.40	4.62
T4	8.14	4.35	8.19	4.36
K5	8.11	4.34	8.7	4.33
K6	8.05	4.34	8.14	4.35
L7	7.86	4.34	7.99	4.36
H8	8.07	4.76	8.11	
K9	8.18	4.40		4.38
E10	8.52	4.67	8.52	
P11		4.41		4.40
A12	8.16	4.22	8.27	4.23
T13	7.80	4.13	7.85	4.15
L14	7.65	4.26	7.71	4.29
I15	7.70	3.92	7.80	3.95
K16	7.86	4.24	7.88	4.27
A17	7.89	4.28	7.95	4.31
I18	7.79	4.11	7.93	4.10
D19	8.21	4.78	8.23	
G20	7.74	3.85	8.32	3.97
D21			8.28	
T22			8.08	4.11
V23			7.89	3.94
K24			7.84	4.20
L25			7.79	4.30
M26			7.92	4.37
Y27			7.82	
K28			7.50	4.20



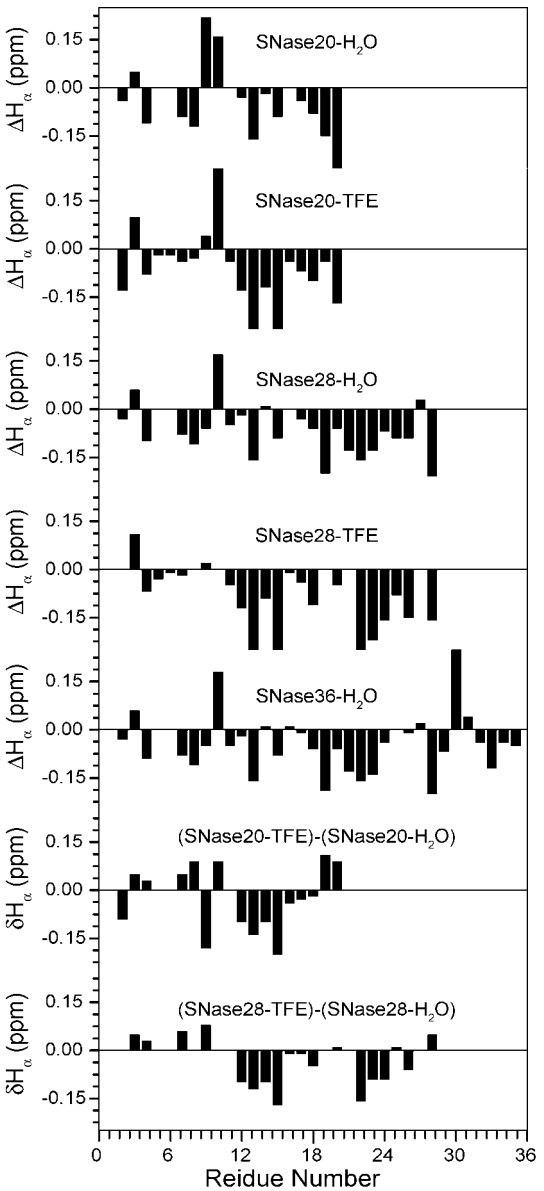


FIGURE 4 Secondary chemical shifts (ΔH_α) for SNase fragments in H_2O and TFE–H_2O mixtures. δH_α represents the difference between $^1H_\alpha$ chemical shift determined in the TFE–H_2O mixture and in H_2O.

FIGURE 3 The 2D ^{1}H-NOESY spectrum of SNase20 in 40% TFE–H_2O mixtures (pH 5.0) at 300 K. The mixing time used for NOESY experiment was 250 ms. (A) NOEs in the fingerprint resonance region. (B) NOEs in the amide proton resonance region. (C) Spectral region for $d_{\beta N}$ NOEs. The $d_{\alpha N}$ NOE connectivities are indicated by lines with arrows, while the d_{NN} and $d_{\beta N}$ NOEs are indicated by arrows at right angles.

Table III ^{1}H Chemical Shifts of SNase28 in Aqueous Solution (pH 5.0) at 278 K

Residue	NH	αH	βH	Others
T2	8.75	4.40	4.18	γCH_3 = 1.24
S3	8.68	4.57	3.93, 3.86	
T4	8.47	4.33	4.25	γCH_3 = 1.22
L7	8.43	4.30	1.59	γH = 1.48 δCH_3 = 0.85, 0.92
H8	8.65	4.68	3.18, 3.23	4H = 7.27 2H = 8.60
K9	8.52	4.30	1.77	γCH_2 = 1.38 δCH_2 = 1.67 εCH_2 = 3.00 εNH_3 = 7.60
E10	8.72	4.59	1.88, 2.09	γCH_2 = 2.37
P11		4.40	2.32, 2.05	γCH_2 = 1.93 δCH_2 = 3.73, 3.84
A12	8.59	4.33	1.40	
T13	8.26	4.27	4.14	γCH_3 = 1.21
L14	8.40	4.39	1.60	γH = 1.60 δCH_3 = 0.85, 0.92
I15	8.32	4.12	1.83	γCH_2 = 1.18, 1.48 δCH_2 = 0.88
K16	8.50	4.28	1.79	γCH_2 = 1.43 δCH_2 = 1.70 εCH_2 = 3.00 εNH_3 = 7.60
A17	8.52	4.32	1.38	
I18	8.38	4.15	1.87	γCH_2 = 1.19, 1.48 δCH_3 = 0.90
D19	8.52	4.62	2.69, 3.73	
G20	8.48	3.96		
D21	8.40	4.69	2.70, 2.76	
T22	8.19	4.27		γCH_3 = 1.22
V23	8.21	4.03	2.06	γCH_3 = 0.94
K24	8.40	4.29	1.79	γCH_2 = 1.46 δCH_2 = 1.74 εCH_2 = 2.98 εNH_3 = 7.60
L25	8.28	4.29	1.60	γH = 1.44 δCH_3 = 0.84, 0.91
M26	8.30	4.43	1.93	γCH_2 = 2.40, 2.49
Y27	8.22	4.61	2.92, 3.11	2,6H = 7.13 3,5 = 6.81
K28	7.93	4.15	1.69	γCH_2 = 1.36 δCH_2 = 1.80 εCH_2 = 3.01 εNH_3 = 7.60

helix formed in the same sequence region as for SNase36.

The ^{1}H resonance assignments identified for SNase36 are close to that of SNase20 and SNase28 in the corresponding sequence regions, except Lys5, Lys6, and Leu25, which we were unable to assign (Table IV). The $d_{\alpha N}(i, i+1)$ and $d_{NN}(i, i+1)$ NOE connectivities for SNase36 are presented in Figure 7. The $d_{\alpha N}$ $(i, i+1)$ NOEs were assigned for pairs of residues, Met26–Tyr27, Met32–Thr33, and Phe34–Arg35, and for segments Thr2–Thr4, Leu7–Glu10, Pro11–Lys24, and Lys28–Gln30 (Figure 7A). A few $d_{NN}(i, i+1)$ NOEs between residues Leu7 and His8, Lys9 and Glu10, Ala12 and Thr13, Asp21 and Thr22, and Met32 and Thr33, and between residues in the segment Tyr27–Gln30 (Figure 7B) were obtained for SNase36. Unlike the continuous $d_{\alpha N}$ $(i, i+1)$ NOEs throughout the segment Leu14–Lys24, only a $d_{NN}(i, i+1)$ NOE and a strong $d_{\beta N}(i, i+1)$ NOE (data not shown) between residues Asp21 and Thr22 were unambiguously determined for this segment. The relatively large negative values of ΔH_α are obtained for residues Asp19–Val23 (Figure 4), indicating a structured conformation populated in this region similar to that observed in SNase28. Therefore, the conformational ensemble of SNase36 may include a populated turn-like conformation localized in the sequence region around Asp21 and Thr22.

The observed NOEs, including $d_{\alpha N}$ $(i, i+1)$ NOEs for Lys28–Gly29–Gln30, $d_{NN}(i, i+1)$ NOEs for Tyr27–Lys28–Gly29–Gln30, and a weak $d_{\alpha N}$ $(i, i+2)$ NOE between Tyr27 and Gly29 (Figure 7A) represent the NOE features of a turn-like conformation for segment Tyr27–Gln30 of SNase36. Such a turn-like conformation may be difficult to differentiate from a piece of helical segment. The C-terminal residues of SNase36 provide the NOEs $d_{\alpha N}$ $(i, i+1)$, $d_{NN}(i, i+1)$, and $d_{\beta N}(i, i+1)$ between Met32 and Thr33 and between Phe34 and Arg35, characterize an additional structured region.

Conformation of SNaseτ_1 in H_2O

The short peptide SNaseτ_1 consists of eight residues (Ala17–Ile18–Asp19–Gly20–Asp21–Thr22–Val23–Lys24), which are common to three other SNase fragments, SNase20, SNase28, and SNase36. The CD-

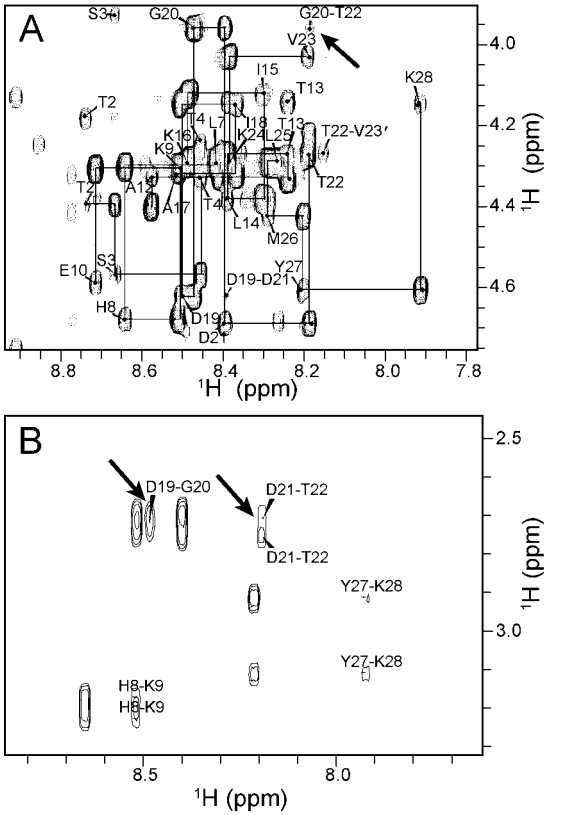


FIGURE 5 The 2D ^{1}H-NOESY spectrum of SNase28 in aqueous solution (pH 5.0) at 278 K collected with a mixing time of 300 ms. (A) NOEs in the fingerprint resonance region. (B) Portion of $^1H_\beta$–^1H_N resonance region. The $d_{\alpha N}$ NOE connectivities are indicated by solid lines and arrows, while the $d_{\beta N}$ NOEs are indicated by arrows.

spectrum of SNaseτ_1 shows a negative ellipticity at about 198 nm (Figure 1A), which characterizes a random coil conformation of the peptide. The assigned ^{1}H resonances for SNaseτ_1 (Table V) give a very narrow spectral region of 8.1–8.8 ppm for dispersed ^1H_N resonances of the peptide. A number of minor cross peaks were observed for residues Asp19, Gly20, Asp21, Thr22, and Val23 in the ^{1}H TOCSY spectrum (data not shown), revealing the multiconformations of this peptide in aqueous solution. A continuous strong $d_{\alpha N}(i, i+1)$ NOE connectivity throughout the sequence of the peptide was identified in the fingerprint region of the ^{1}H NOESY spectrum (Figure 8A). In addition, the weak medium-range $d_{\alpha N}(i, i+2)$ NOEs were obtained for pairs of residues Asp19–Asp21 and Gly20–Thr22 (Figure 8A). Although the ^1H_N–^1H_N resonance region does not provide any d_{NN} NOE connectivity for the peptide, the strong NOEs between side-chain β-protons of residues Asp19–Asp21 and amide protons of residues Gly20–Thr22, respectively, were observed for SNaseτ_1 (Figure 8B). The observed NOE features characterize a fraction of β-turn conformation for this 8-residue peptide.

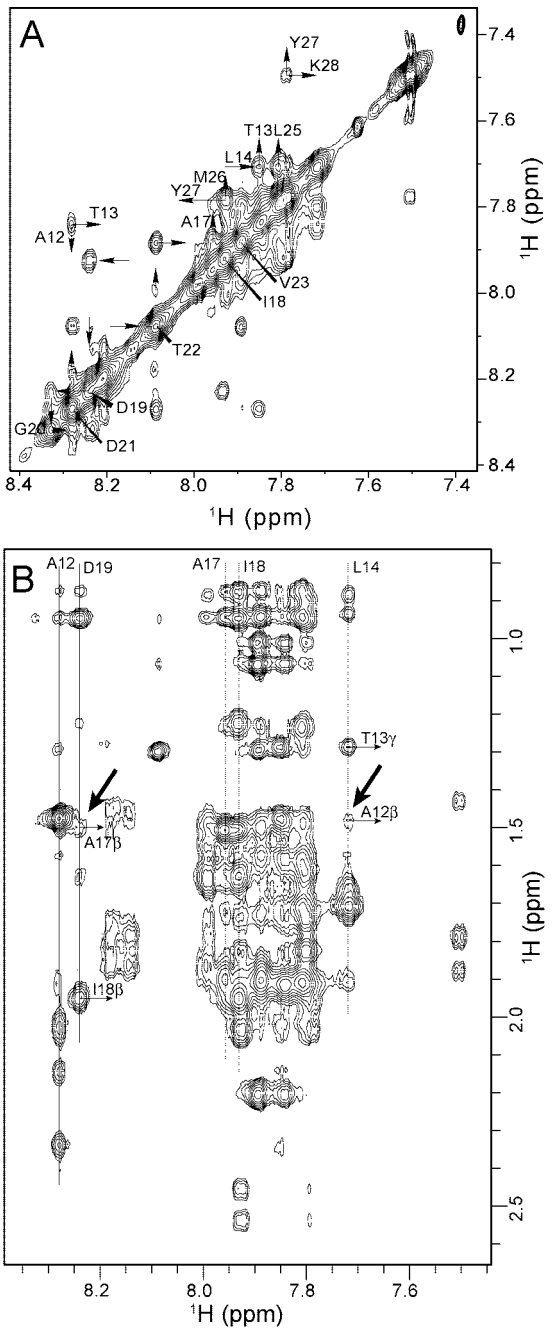


FIGURE 6 The 2D ^{1}H-NOESY spectrum of SNase28 in the 20% TFE–H_2O mixture (pH 5.0) at 300 K. NOE data were collected at mixing time of 400 ms. (A) NOEs in the amide proton resonance region. (B) Portion of the NOESY spectrum showing $d_{\beta N}$ NOEs. The d_{NN} and $d_{\beta N}$ NOEs are indicated by arrows at right angles.

Table IV ^{1}H Chemical Shifts of SNase36 in Aqueous Solution (pH 5.0) at 278 K

Residue	NH	αH	βH	Others
T2	8.74	4.40	4.18	γCH_3 = 1.25
S3	8.68	4.57	3.93, 3.87	
T4	8.47	4.34	4.25	γCH_3 = 1.23
L7	8.42	4.30	1.60	γH = 1.48 δCH_3 = 0.86, 0.91
H8	8.65	4.68	3.18, 3.20	
K9	8.51	4.31	1.78	γCH_2 = 1.39 δCH_2 = 1.68 εCH_2 = 3.00 εNH_3 = 7.61
E10	8.72	4.60	1.88, 2.09	γCH_2 = 2.37
P11		4.41	2.32, 2.02	γCH_2 = 1.94 δCH_2 = 3.73, 3.85
A12	8.58	4.33	1.41	
T13	8.25	4.27	4.14	γCH_3 = 1.21
L14	8.40	4.39	1.61	γH = 1.60 δCH_3 = 0.88, 0.93
I15	8.32	4.13	1.83	γCH_2 = 1.20, 1.48 δCH_2 = 0.89
K16	8.50	4.29		γCH_2 = 1.45 δCH_2 = 1.70, 1.80 εCH_2 = 3.00 εNH_3 = 7.61
A17	8.51	4.34	1.39	
I18	8.37	4.15	1.87	γCH_2 = 1.20, 1.48 δCH_3 = 0.91
D19	8.51	4.63	2.71, 2.75	
G20	8.49	3.96		
D21	8.41	4.69	2.71, 2.76	
T22	8.18	4.27		γCH_3 = 1.22
V23	8.19	4.02	2.08	γCH_3 = 0.95
K24	8.35	4.32	1.82	γCH_2 = 1.47 δCH_2 = 1.66 εCH_2 = 3.01 εNH_3 = 7.61
M26	8.35	4.51	1.96	γCH_2 = 2.43, 2.52
Y27	8.37	4.60	3.01	2,6H = 7.14 3,5 = 6.81
K28	8.51	4.16	1.78	γCH_2 = 1.45 δCH_2 = 1.72 εCH_2 = 3.01 εNH_3 = 7.61
G29	8.02	3.95, 3.82		
Q30	8.18	4.66	1.90	γCH_2 = 2.37
P31		4.49	2.29, 2.03	γCH_2 = 1.90 δCH_2 = 3.68, 3.84
M32	8.65	4.48	1.98	γCH_2 = 2.53
T33	8.13	4.31	4.12	γCH_3 = 1.14
F34	8.35	4.61	3.00, 3.11	2,6H = 7.21
R35	8.27	4.33	1.82	γCH_2 = 1.58, 1.69 δCH_2 = 3.17 NH = 7.22
L36	8.06	4.15	1.61	γH = 1.60, 1.48 δCH_3 = 0.86, 0.93

DISCUSSION

Many peptide fragments of proteins have been found to populate secondary structural conformations, even lacking the stable long-range interactions formed in the intact proteins. The propensities of formation of secondary structural conformations, especially the turn-like conformation, are proposed to play an important role in limiting the conformational space available to the polypeptide chain, and directing the subsequent folding events.[1,2,5,22,23] SNase is composed of β- and α-subdomains.[24,25] The N-terminal 110 residues have a nearly complete sequence of the β-subdomain, which contains a main hydrophobic core named "β-barrel" in the protein. The β-barrel consists of two antiparallel β-pleated sheets, β_I and β_{II}. SNase36 spans β_I, which contains three β-strands: β_1, Thr13–Ala17; β_2, Thr22–Met26; β_3, Pro31–Arg35, and two reverse turns: τ_1, Ile18–Asp21; τ_2, Tyr27–Gln30. The residues Ala12–Thr13–Leu14 of SNase36 form a bent peptide linking strand β_1 and a strand in β_{II} (His8–Pro11). The N-terminal sequence, Ala1–Leu7, is unstructured in the native protein. SNase28 and SNase20 are the portions of the fragment SNase36, and contain the corresponding structural elements in the intact SNase. Peptide SNaseτ_1 (Ala17–Lys24) spans the reverse turn, τ_1, which is the first β-turn from the N-terminus of SNase. As is suggested, nucleation mechanisms that refer the nucleus to the native-like secondary structures formed by some neighboring residues of sequence in protein folding are particularly efficient means for initiating folding.[26] The investigation of localized structural propensity in these short fragments, which match native-like elements, especially the turn-like conformations, may

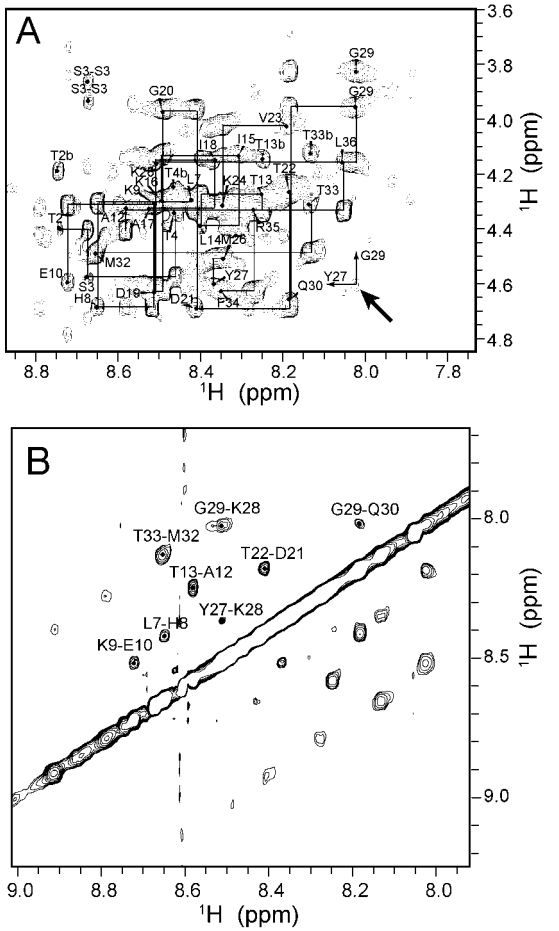


FIGURE 7 The 2D ^{1}H-NOESY spectrum of SNase36 in aqueous solution (pH 5.0) at 278 K. The mixing time of NOESY experiment was 300 ms. (A) NOEs in the fingerprint resonance region. (B) NOEs in the amide proton resonance region. The $d_{\alpha N}$ NOE connectivities are given by solid lines and arrows, while the d_{NN} NOEs are labeled by the pairs of residues.

serve as nucleation sites in the folding process of SNase fragments starting from the N-terminal end and of different chain length.

According to the NOE connectivity patterns for turns,[27,28] the type I turn shows relatively weak $d_{\alpha N}$ (2, 3) and strong $d_{\alpha N}$ (3, 4), d_{NN}(2, 3), and d_{NN}(3, 4) NOEs; however, in the type II turn, both $d_{\alpha N}$ (2, 3) and $d_{\alpha N}$ (3, 4) NOEs are strong, and both d_{NN}(2, 3) and d_{NN}(3, 4) NOEs are weak. Some $d_{\alpha N}$ (i, i+2) and $d_{\beta N}$(i, i+2) NOEs can be observed for both types of turn conformations. Based on this NOE criteria, a common sequence region Ala17–Thr22 in fragments SNase28, SNase36 and SNaseτ_1 as well, and a sequence region Tyr27–Gln30 of SNase36, can be considered as the nucleation sites, forming the turn-like conformations as shown by d_{NN}(i, i+1) and $d_{\alpha N}$ (i, i+2), and some $d_{\beta N}$ (i, i+2) NOEs in aqueous solution (Figure 9). The localized propensity for forming a turn-like conformation of the sequence region Tyr27–Gln30 in aqueous solution was evidenced by the NMR study of a synthesized short peptide Tyr27–Lys28–Gly29–Gin30–Pro31 in aqueous solution.[11] A populated bend conformation [as shown by d_{NN}(i, i+1) NOEs] in dynamic equilibrium with unfolded conformations was observed for short segment Ala12–Thr13–Leu14 of all fragments, which is crucial for different trends of strands β_1 and β_4 in conformational space of native SNase, and thus for the spatial packing of two antiparallel β-sheets (β_I and β_{II}) in the "β-barrel" of the protein. The N-terminal short fragments of SNase, lacking the long-range tertiary interactions, are in such a denatured state that is a heterogeneous collection of rapidly interconverting conformations having some populated native-like structural elements. The localized bend- and turn-like structures, as the folding nucleus, of the SNase fragments in aqueous solution are, therefore, transiently populated in different polypeptides, at different times in conformational ensemble of the fragments. As was reported, TFE as a "hydrophobic" media can facilitate the rearrangement of non-native tertiary contacts by increasing the solvent hydrophobicity,[29,30] and can induce the acceleration of folding by its ability to

Table V ^{1}H Chemical Shifts of SNaseτ_1 in Aqueous Solution (pH 5.0) at 278 K

Residue	NH	αH	βH	Others
A17		4.11	1.53	
I18	8.66	4.22	1.87	δCH_3 = 0.92 γCH_2 = 1.19, 1.46
D19	8.64	4.61	2.70, 2.75	
G20	8.51	3.97		
D21	8.37	4.71	2.72, 2.77	
T22	8.20	4.33	4.19	γCH_3 = 1.21
V23	8.31	4.09	2.08	γCH_3 = 0.95
K24	8.17	4.19	1.82	δCH_2 = 1.71 γCH_2 = 1.41 εNH_3 = 7.57 εCH_2 = 3.00

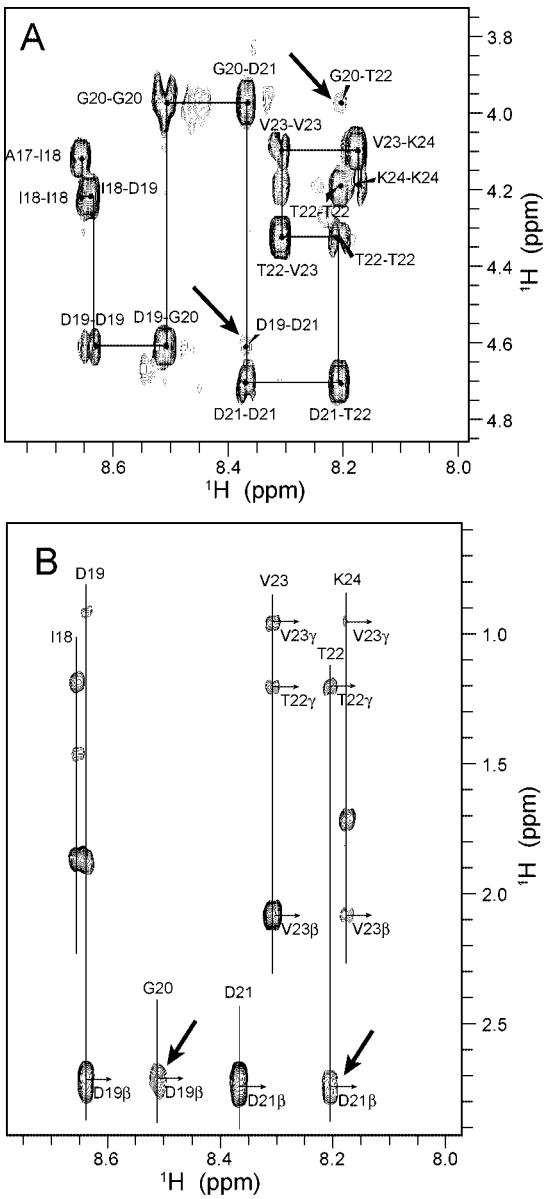


FIGURE 8 The 2D ^{1}H-NOESY spectrum of SNaseτ_1 in aqueous solution (pH 5.0) at 278 K. The spectrum was recorded at mixing time of 300 ms. (A) The sequential $d_{\alpha N}$ NOE connectivities which indicated by solid lines. The medium range $d_{\alpha N}$ NOEs indicated by arrows and labeled by a pair of residues. (B) The interresidual NOEs of aliphatic protons with amide protons.

stabilize native-like local hydrogen bonds.[31] The contributions of TFE to stabilize the folding nucleus of SNase fragments are revealed by the new NOEs (Figure 9) and the changes in ΔH_α (Figure 4) obtained for SNase20 and SNase28 in the TFE–H_2O mixture. The long-range $d_{NN}(i, j)$ NOEs for the sequence Leu14–Leu25 of SNase28 and $d_{NN}(i, i+2)$ and $d_{\beta N}(i, i+2)$ NOEs for the sequence Pro11–Leu14 of SNase20 and SNase28 indicate more developed folding nucleus of the fragments, a reverse turn-like conformation of Ala17–Thr22, and a bent peptide around Ala12–Thr13, respectively, in the presence of TFE.

The above-indicated localized structural propensities of the N-terminal short fragments of SNase reflect closely the secondary structural elements of the same sequence regions in the β-barrel of native SNase. The preferences of formation of bend- and turn-like conformations are observed not only for the N-terminal short fragments, but also for the same sequence regions of N-terminal large fragments of SNase.[13] This suggests that segments Ala17–Thr22 and Tyr27–Gln30, as well as Ala12–Thr13–Leu14, could act as folding nucleation sites of SNase fragments of different chain lengths starting from the N-terminal end. Since the bend- and turn-like conformations of the N-terminal short segment of SNase are important for limiting the conformational space of the corresponding SNase polypeptide chain in the "β-barrel" tertiary fold of SNase, the formation of turn-like and bend-like conformations of segments Ala17–Thr22 and Tyr27–Gln30, as well as Ala12–Thr13–Leu14, may play important roles in the early folding event of SNase polypeptide chains in vitro.

The above analysis indicates that folding of SNase is initiated by forming the turn-like conformations in the β-barrel region. However, the native-like tertiary interactions forming in the β-barrel region of SNase are required for the nucleus sites to be accounted for initiating the full structure formation of SNase. Studies of 110-residue SNase fragments, SNase110, and its mutants, G88W110 and V66W110,[13,32] provided the experimental evidences that the overall structure formation of the SNase fragments is correlated to the stability and conformation of the folded β-barrel. Therefore, the β-barrel acts as a large nonlocal nucleus for the folding of SNase. In the nucleation–condensation mechanism, nucleation and overall structure formation are coupled; the formation of the nucleus and the formation of tertiary structure are concerted.[26] In the light of this folding mechanism, the different folding pathways may exist for the folding of SNase in accordance with the different folded states of the β-barrel. For proline-free SNase, the multiple parallel-pathway folding was observed using stopped-flow techniques.[33,34] However, the relationship of this observation for SNase with the initiation folding of β-barrel needs further investigation.

This work was supported in part by the Pandeng Project of the National Commission for Science and Technology and by the National Natural Science Foundation of China (NNSFC) 39823001.

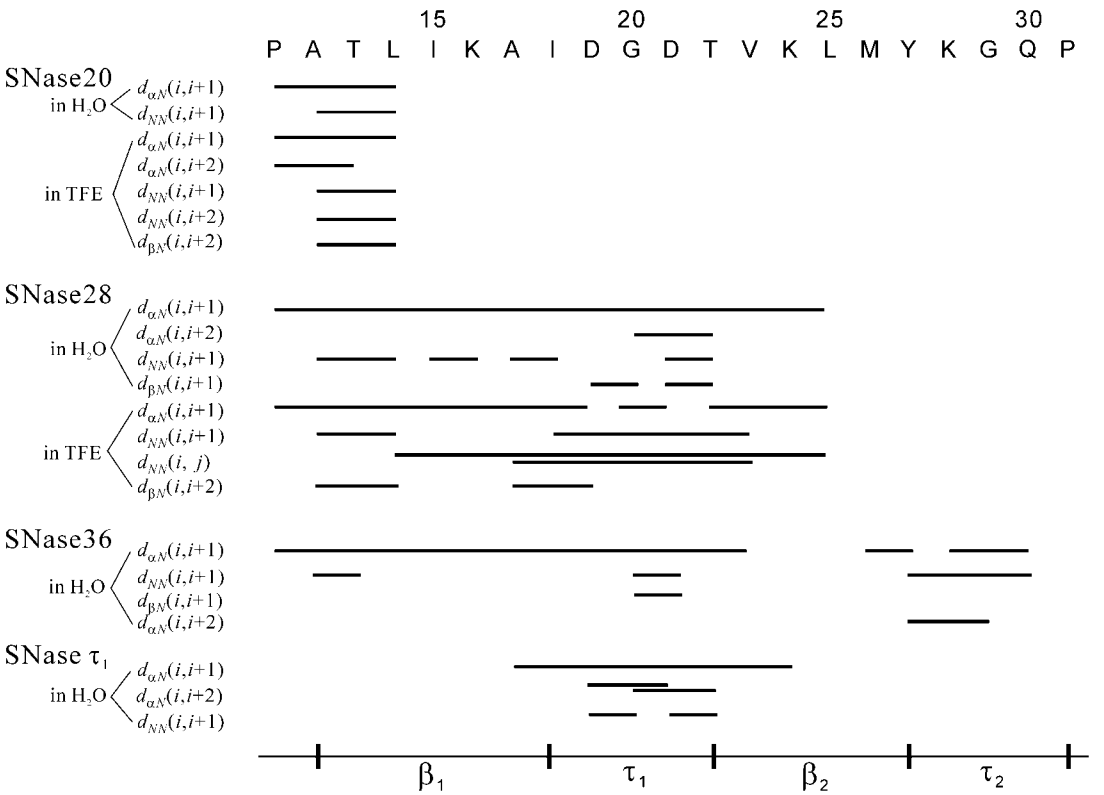


FIGURE 9 Summary of NOE data for sequence regions Pro11–Leu14 of SNase20; Pro11–Leu25 of SNase28; Pro11–Pro31 of SNase36; and Ala17–Lys24 of SNaseτ_1 recorded with samples dissolved in H_2O or a H_2O–TFE mixture. The horizontal bars denote the corresponding NOEs. The secondary structures of native SNase are indicated in the figure.

REFERENCES

1. Waltho, J. P.; Feher, V. A.; Merutka, G.; Dyson, H. J.; Wright, P. E. Biochemistry 1993, 32, 6337–6347.
2. Shin, H. C.; Merutka, G.; Waltho, J. P.; Wright, P. E.; Dyson, H. J. Biochemistry 1993, 32, 6348–6355.
3. Shin, H. C.; Merutka, G.; Waltho, J. P.; Tennant, L. L.; Dyson, H. J.; Wright, P. E. Biochemistry 1993, 32, 6356–6364.
4. Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. J Mol Biol 1988, 201, 161–200.
5. Wright, P. E.; Dyson, H. J.; Lerner, R. A. Biochemistry 1988, 27, 7167–7175.
6. Dyson, H. J.; Rance, M.; Houghten, R. A.; Wright, P. E.; Lerner, R. A. J Mol Biol 1988, 201, 201–217.
7. Wang, Y.; Shortle, D. Biochemistry 1995, 34, 15895–15905.
8. Wang, Y.; Shortle, D. Protein Sci 1996, 5, 1898–1906.
9. Jacobs, M. D.; Fox, R. O. Proc Natl Acad Sci USA 1994, 91, 449–453.
10. Walkenhorst, W. F.; Edwards, J. A.; Markley, J. L.; Roder, H. Protein Sci 2002, 11, 82–91.
11. Ramakrishna, V.; Sasidhar, Y. U. Biopolymers 1997, 41, 181–191.
12. Ramakrishna, V.; Sasidhar, Y. U. Ind J Biochem Biophys 1998, 35, 333–338.
13. Feng, Y.; Liu, D.; Wang, J. J Mol Biol 2003, 330, 821–837.
14. Chang, C. D.; Meienhofer, J. Int J Pept Protein Res 1978, 11, 246–249.
15. Meienhofer, J.; Waki, M.; Heimer, E. P.; Lambros, T. J.; Makofske, R. C.; Chang, C. D. Int J Pept Protein Res 1979, 13, 35–42.
16. Greenfield, N. J. Anal Biochem 1996, 235, 1–10.
17. Gierasch, L. M.; Deber, C. M.; Madison, V.; Niu, C. H.; Blout, E. R. Biochemistry 1981, 20, 4730–4738.
18. Schwarzinger, S.; Kroon, G. J.; Foss, T. R.; Wright, P. E.; Dyson, H. J. J Biomol NMR 2000, 18, 43–48.
19. Wishart, D. S.; Sykes, B. D.; Richards, F. M. J Mol Biol 1991, 222, 311–333.
20. Wishart, D. S.; Sykes, B. D.; Richards, F. M. Biochemistry 1992, 31, 1647–1651.
21. Wishart, D. S.; Sykes, B. D. J Biomol NMR 1994, 4, 171–180.
22. Dyson, H. J.; Wright, P. E. Annu Rev Biophys Biophys Chem 1991, 20, 519–538.
23. Muga, A.; Surewicz, W. K.; Wong, P. T.; Mantsch, H. H. Biochemistry 1990, 29, 2925–2930.

24. Hynes, T. R.; Fox, R. O. Proteins 1991, 10, 92–105.
25. Wang, J.; Truckses, D. M.; Abildgaard, F.; Dzakula, Z.; Zolnai, Z.; Markley, J. L. J Biomol NMR 1997, 10, 143–164.
26. Fersht, A. R. Curr Opin Struct Biol 1997, 7, 3–9.
27. Wilmot, C. M.; Thornton, J. M. J Mol Biol 1988, 203, 221–232.
28. Wüthrich, K. NMR of Proteins and Nucleic Acids; John Wiley & Sons: New York, 1986.
29. Lu, H.; Buck, M.; Radford, S. E.; Dobson, C. M. J Mol Biol 1997, 265, 112–117.
30. Kentsis, A.; Sosnick, T. R. Biochemistry 1998, 37, 14613–14622.
31. Hamada, D.; Chiti, F.; Guijarro, J. I.; Kataoka, M.; Taddei, N.; Dobson, C. M. Nat Struct Biol 2000, 7, 58–61.
32. Ye, K.; Jing, G.; Wang, Biochim Biophys Acta 2000, 1479, 123–134.
33. Kamagata, K.; Sawano, Y.; Tanokura, M.; Kuwajima, K. J Mol Biol 2003, 332, 1143–1153.
34. Walkenhorst, W. F.; Green, S. M.; Roder, H. Biochemistry 1997, 36, 5795–5805.

Reviewing Editor: Dr. David Wemmer