Folding of the C-Terminal Fragment V111-D143 of Staphylococcal Nuclease in Aqueous Solution

Yong Geng, Min Wang, Tao Xie, Yingang Feng and Jinfeng Wang*

National Laboratory of Biomacromolecules, Center for Structural and Molecular Biology, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, People's Republic of China

Abstract: Studies of conformational features of fragments SNase(111-143) and SNase(118-143) and segment E122-K136 in 1-139 fragment (SNase139) suggest that the high intrinsic helical propensity can drive segment E122-K136 fold into a stable helix only when the segments V111-H121 and L137-D143 flanked on segment E122-K136 in staphylococcal nuclease (SNase) have stable folding.

Keywords: SNase(111-143), fragment, C-terminal sub-domain, helix-forming tendency, ensemble of interconverting conformations.

INTRODUCTION

Proteins are composed of secondary structural elements which can be referred to intrinsically cooperative units in the protein folding. A new explanation for variation in protein folding behavior proposed that protein folding proceeds by the stepwise assembly of the folden units in a sequential stabilization process by the factors that determine the native state [1]. The foldens refer to cooperative native-like structural units which unfold and refold as concerted folding units. The folding pathway steps are determined by the formation of such cooperative native-like structural units [2]. The number of accessible pathways of a protein is linked to the number of folden contained within the native topology. Thus, proteins can be seen as modular assemblies of competing foldons [3]. For completely folded proteins, foldens tend to be coincident with secondary structural elements, but in partially folded forms folden units need not to be accurately identical with its native secondary elements [2]. Cytochrome c was reported to be composed of five foldens including helices formed in different folding step. Helices appear to be more or less autonomous elements for forming the folding nucleus in the protein, when the corresponding peptide segment has certain degree of helical propensity.

Staphylococcal nuclease (SNase) that behaves as a cooperative unit folds efficiently and reversibly has been studied extensively as a model molecule for protein folding. The tertiary structure of SNase contains a main hydrophobic core which consists of two antiparallel β-pleated sheets forming a "β-barrel" structural region and three α-helices. The N- and C-terminal segments consisting of residues A1-L7 and N144-Q149, respectively are unstructured in 3D structure of native SNase [4, 5]. N-terminal fragments of SNase with different chain lengths have long been used as a model system for studying protein folding and unfolding [6-12]. The 1-110 residues SNase fragment (SNase110) consists of the "β-barrel" region and helices α1 and α2. In aqueous solution, SNase110 exists as an ensemble of coexistent native-like partially folded and unfolded states. The G88W- and V66W-mutaion of 1-110 residues SNase fragment can drive the mutant fragments, G88W110 and V66W110, to form native-like conformations [12]. A detailed investigation including the structural determination and ^{15}N backbone dynamic studies of mutant SNase110 reveals the local and non-local nucleation sites that initiate the cooperative tertiary folding of G88W110 and V66W110. G88W110 and V66W110 can adopt native-like tertiary structures in aqueous solution [13]. An interesting question is about the folding behavior of the C-terminal fragment V111-D143 of SNase in the folding of intact protein. A synthesized peptide corresponding to the C-terminal peptide E129-E142 of SNase was studied by circular dichroism (CD) and 2D ^{1}H nuclear magnetic resonance (NMR) experiments [14]. It was found that a helix-turn conformation of the peptide E129-E142 similar to that in the intact SNase can be induced by 30%-40% 2,2,2-trifluoroethanol. However, the question remains unclear so far whether the C-terminal fragment V111-D143 of SNase will be able to fold into the native-like conformation in H_2O when removed from the rest of the protein which is the 1-110 residues polypeptide chain of SNase.

The present study is to elucidate the folding behaviors of C-terminal fragment V111-D143 of SNase in H_2O and to analyze the correlations of the fragment folding with the folding of SNase. The C-terminal fragment of SNase spanning over residues V111-D143 of native SNase was expressed and purified. The conformational features of the fragment V111-D143 in aqueous solution were studied by CD and heteronuclear multi-dimensional NMR spectroscopic methods. The relationship between the folding of the C-terminal region V111-D143 and the folding of native SNase is described.

*Address correspondence to this author at the National Laboratory of Biomacromolecules, Center for Structural and Molecular Biology, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, People's Republic of China; Tel: +86-10-6488-8490(O); Fax: +86-10-6487-2026; E-mail: jfw@sun5.ibp.ac.cn

0929-8665/07 $50.00+.00

MATERIALS AND METHODS

Construction of Expression Vector

For expression of N-terminal short fragment corresponding to residues V111-D143 of SNase, the fusion expression system, pET-3D-HR52, was adopted. In this plasmid, a His_6-tagged N-terminal fragment containing A1-E52 residues of SNase (HR52) was located between *Nco*I and *Kpn*I endonuclease sites, and followed by a thrombin digestion site [15]. The DNA fragment of SNase(111-143) was generated by polymerase chain reaction (PCR) of the gene of full-length SNase. Two primers, primer 1 (5′-ggggtaccctggttccgcgtggtcc gttgcttatgtttacaaacctaac-3′) and primer 2 (5′-cgggatccattattga cctgaatcagcattgtcttcgctccaaatatttaatttctc-3′) were used as the forward and reverse primers, respectively. The expression plasmid pET-3D-HR52 and the PCR products were cut out by *Kpn*I and *BamH*I endonuclease. Afterwards, the DNA fragment of SNase(111-143) was cloned on the expression vector pET-3D-HR52. The resulting plasmid pET-3D-HR52-SNase(111-143) encodes the full length 33-residue SNase(111-143) as a fusion protein with a N-terminal His_6-tag affinity peptide.

Protein Expression and Purification

The expression plasmid pET-3D-HR52-SNase(111-143) was transformed into E.*coli* BL21(DE3) cells which were grown in LB medium at 37 ℃. Protein expression was induced with 0.4 mM IPTG (isopropyl-D-thiogalactoside) when the bacterial cells were grown to an $O.D_{600}$ value of 0.8. After incubation for 4 hours, the bacterial cells were collected by centrifugation and suspended in buffer A (50 mM Tris-HCl, pH 8.0, 250 mM NaCl). After ultrasonication treatment, the lysate was centrifuged at 12,000 rpm for 20 minutes, and the supernatant was loaded onto an 8 ml Ni^{2+}-chelating column (Pharmacia Biotech) pre-equilibrated with buffer A. The column was extensively washed with buffer A, followed by buffer A with 50 mM imidazole in addition. After washing the column with thrombin buffer (50 mM Tris-HCl, pH 8.0, 140 mM NaCl, 10 mM $CaCl_2$), the thrombin (8 mg) was loaded on the column. The solution in the column was circulated by a pump for 12 h at room temperature. The column was then washed with 50 ml buffer A again. The collected solution was concentrated by lyophilization and loaded to a Superdex 75 column pre-equilibrated with 0.1 mM NH_4HCO_3 buffer. Then, the column was washed with 0.1 mM NH_4HCO_3 buffer. The collected fraction was lyophilized. The purity of eluted fraction of SNase(111-143) was > 99% as verified by 16.5% tricine SDS-PAGE. Uniformly ^{15}N-labeled and $^{15}N/^{13}C$-double labeled SNase(111-143) for NMR studies were obtained by overexpressing the protein in *E. coli* through bacterial growth in minimal media with $^{15}NH_4Cl$ and [$^{13}C_6$]-glucose as the sole nitrogen and carbon sources.

For clarifying the influence of *cis-trans* isomarization of X-prolyl peptide bond K116-P117, C-terminal fragment N118-D143 of SNase (SNase(118-143)) was also purified following the same protocol as for the fragment SNase(111-143). ^{15}N-labeled SNase and ^{15}N-labeled 1-139 residues fragment of SNase (SNase139) were obtained in addition for description of the folding of SNase(111-143).

Circular Dichroism Measurements

Far-UV CD experiments were performed with peptide fragments SNase(111-143) and SNase(118-143) at concentration of 0.4 mg/ml in 20 mM Tris-HCl buffer (pH 7.4) with 0.0, 2.0, and 4.0 M urea. The spectra were recorded on a Jasco J-720 spectro-polarimeter over a range of 190-250 nm at room temperature. A quartz cuvette with 1.0 mm path length was employed in the measurements. Four scans were averaged for each measurement. The secondary structure contents of the SNase fragments were estimated by analyzing the CD spectra using software package CDPro (http://lamar.colostate.edu/~sreeram/CDPro).

NMR Spectroscopy

All NMR experiments were run on a Bruker DMX 600 spectrometer equipped with a triple-resonance cryo-probe at 300 K. The 3D ^1H-^{13}C-^{15}N HNCACB, CBCA(CO)NH, HN(CA)CO, HNCO, HBHACONH [16] experiments were performed on SNase(111-143) samples for backbone and side-chain ^1H, ^{15}N, and ^{13}C resonance assignments. 1.0 mM $^{13}C/^{15}N$-double labeled protein in 90% H_2O/10% D_2O containing 50 mM deuterated acetate buffer (pH 5.0), 100 mM KCl, and 0.01% NaN_3 was used for these 3D heteronuclear experiments. The 2D ^1H-^{15}N HSQC spectra were collected for 1.0 mM ^{15}N-labeled SNase(111-143), SNase139, and full-length SNase in aqueous solution. For NOE determinations, a 3D ^1H-^{15}N NOESY-HSQC [16] experiment was performed with a sample containing 1.0 mM ^{15}N-labelled SNase(111-143). The NOE mixing time was 300 ms. All NMR data were processed and analyzed using FELIX98 (Accelrys Inc.). The data points in each indirect dimension were usually doubled by linear prediction before zero filling to the appropriate size [17http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=16083422 -B9]. A 70°-90° shifted square sine bell apodization was used for all dimensions before Fourier transformation. ^1H chemical shifts were referenced to internal DSS (2,2-dimethyl -2-silaentane-5-sulphonate). ^{15}N and ^{13}C chemical shifts were referenced indirectly [18http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=16083422-B10].

RESULTS

SNase(111-143) for NMR Experiments

For analysis of the solution conformation of the fragment SNase(111-143), the powerful technique is the multidimensional NMR method which can provide a better signal separation in NMR spectra for resonance assignments. Stable-isotope-labeled SNase(111-143) is then required for this study. For fulfilling this requirement, the expression plasmid pET-3d-HR52-SNase(111-143) was obtained (Fig. **1**). The desired recombinant His_6-tagged HR52-SNase(111-143) was highly expressed in soluble form. In the constructed expression plasmid, there is a thrombin cleavage site (LVPR↓GS) between the fusion partner HR52 and target protein SNase(111-143) (Fig. **1**). After thrombin digestion, the fusion partner was removed and two residues (GS) will be appended to the N-terminus of target protein SNase(111-143). Thus, the resulting fragment SNase(111-143) contains 35

amino acid residues. However, only residues V111-D143 were considered in the analysis of NMR data.

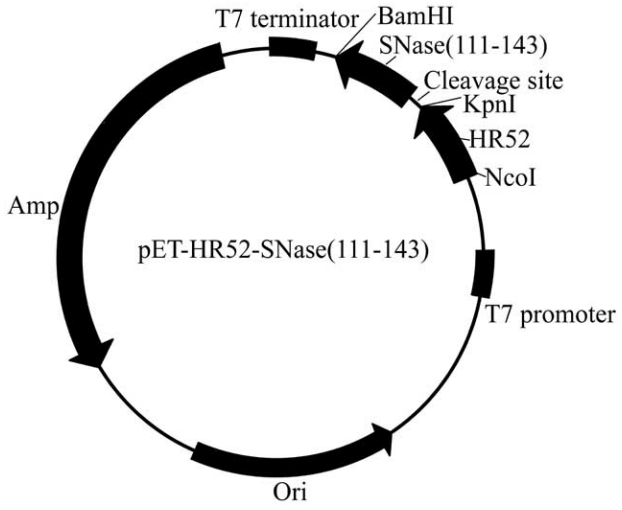

Figure 1. Map of expression plasmid of pET-3d-HR52-SNase(111-143).

For ascertainment of the monomeric state of this 35-residue polypeptide in solution, the 2D ^{1}H-^{15}N HSQC NMR spectra were taken for the samples containing 0.2-1.0 mM purified SNase(111-143). The chemical shifts and linewidths of cross-peaks in the NMR spectra of 1.0 mM sample showed no differences with those of the 0.2 mM sample (spectra not shown). This observation can affirm that SNase(111-143) is in monomeric state under the experimental conditions. Thus, the concentration effect of the sample on the experimental results can be excluded.

Residual Structure of SNase(111-143) Inferred from the CD Spectra

Far-UV CD measurements were used to provide the information of secondary structures of SNase(111-143) in aqueous solution. The far-UV CD spectrum of SNase(111-143) is shown in Fig. (**2A**). The spectrum shows an apparent negative minimum and a shallow shoulder at approximately 200 and 220 nm, respectively. The pronounced minimum at about 200 nm is a characteristic of non-ordered conformation of SNase(111-143) in aqueous solution. The shallow minimum at approximately 220 nm decreased as concentration of urea was increased in sample solution. Fig. **2B** shows the differences in the ellipticities of the wavelength range from 210 nm to 250 nm by subtracting the CD data of SNase(111-143) in the presence of urea from those in aqueous solution. The increased differences in ellipticities on increasing the urea concentrations indicated some residual structures in SNase(111-143). Estimation of the secondary structure contents from CD data gave 7% helix, 36% strand, 19% turn, and 38% random coil for SNase(111-143) in H_2O.

Fig. **2A** shows also the far-UV CD spectrum of SNase(118-143) collected under the same experimental conditions as for SNase(111-143). Comparing ellipticities of the two spectra revealed that the peaks at about 200 and 220 nm in the spectrum of SNase(111-143) was shifted to about 206 and 222 nm, respectively in the spectrum of SNase(118-143). This is consistent with an increased population of helical conformers in SNase(118-143). The negative ellipticities at 206 and 222 nm in the CD spectrum of SNase(118-143) were relatively stronger than those at 200 and 222 nm of SNase(111-143) (Fig. **2A**). The development of relatively strong negative ellipticities in the CD spectrum of SNase(118-143) evidenced also the increased helical conformations. Analysis of the corresponding CD data gave 15% helix, 35% strand, 21% turn, and 29% random coil for SNase(118-143) in H_2O.

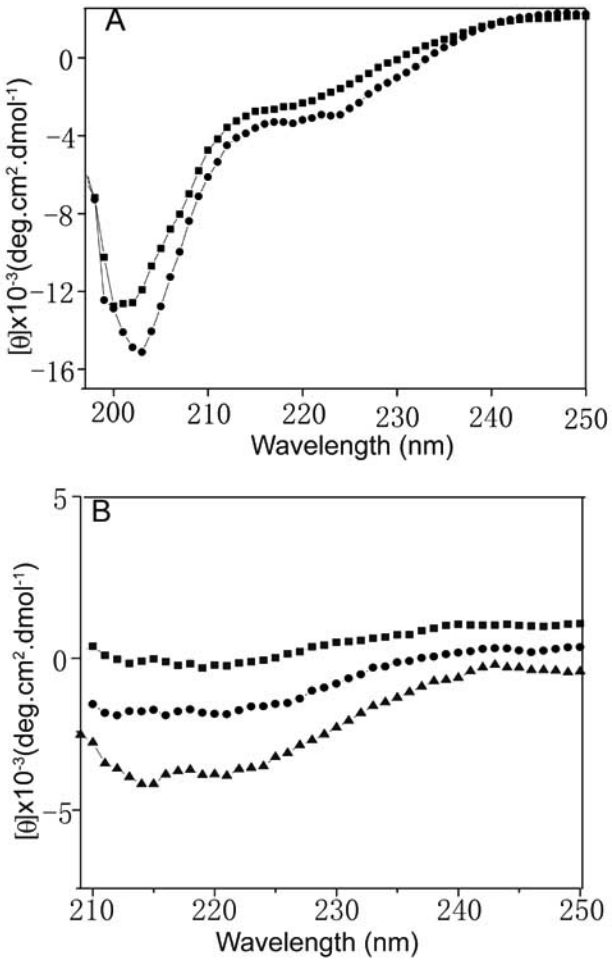

Figure 2. Far-UV CD spectra of SNase fragments recorded at room temperature in H_2O. (A) CD Spectra (197-250 nm) of SNase(111-143) (■) and SNase(118-143) (●) in aqueous solution. (B) Differences in the ellipticities at 210-250 nm between the CD data of SNase(111-143) in aqueous solution and those in the presence of 2.0 (■) and 4.0 (●) M urea. The Differences between the CD data of SNase(118-143) in aqueous solution and those in the presence of 4.0 M urea were also presented (▲) for comparison.

It can be seen that SNase(111-143) and SNase(118-143) have about the same secondary β-strand and turn contents. Truncation of the N-terminal segment V111-P117 of SNase(111-143) containing a K116-P117 peptide bond changed the relative contents of helix and random coil conformations. SNase(118-143) has relatively lower contents of

random coil and higher contents of helical conformations than SNase(111-143). Therefore, the segment V111-P117 enhances a population of random coil conformation of SNase(111-143), and a small amount of nascent helix is populated in the segment N118-D143 of SNase(111-143).

Secondary Chemical Shifts

Sequential backbone ^{15}N, $^{1}H_N$, $^{13}C_\alpha$, $^{1}H_\alpha$, and $^{13}C'$ resonance and side-chain $^{13}C_\beta$ and $^{1}H_\beta$ resonance assignments were determined for SNase(111-143) in H_2O using a series of 3D ^{1}H-^{13}C-^{15}N triple-resonance NMR experiments (Fig. **3**). 100% backbone ^{15}N, $^{1}H_N$, $^{13}C_\alpha$, $^{1}H_\alpha$, and $^{13}C'$resonances were assigned. For the side-chain resonances, 100% $^{13}C_\beta$ and 94% $^{1}H_\beta$ resonances were assigned. The $^{1}H_\beta$ chemical shifts of residues K116 and D143 were unable to be identified. The assigned ^{15}N and $^{1}H_N$ resonances are indicated in Fig. (**4**). Residue 117 is proline and unable to give a $^{1}H_N$-^{15}N cross-peak in the 2D $^{1}H_N$-^{15}N HSQC spectrum. There is little chemical shift dispersion in the amide proton resonance region of the 2D $^{1}H_N$-^{15}N HSQC spectrum of this peptide fragment. Majority of the cross-peaks in the 2D NMR spectrum appeared in a narrow spectral region of 7.8-8.7 ppm for $^{1}H_N$ resonances and 113.5-128.5 ppm for ^{15}N resonances.

Since all the $^{1}H_\alpha$, $^{13}C_\alpha$, $^{13}C_\beta$, and $^{13}C'$ resonance assignments for SNase(111-143) were obtained, combining the ^{1}H and ^{13}C chemical shift indices (CSI) can produce a consensus CSI to identify and locate secondary structures in protein [19]. The secondary chemical shifts, $\Delta^{1}H_\alpha$, $\Delta^{13}C_\alpha$, $\Delta^{13}C_\beta$, and $\Delta^{13}C'$ of SNase(111-143) are presented in Fig. (**5**) [20-23]. It is well known that secondary chemical shifts $\Delta^{1}H_\alpha$ can give an indication of a helix-forming tendency in a peptide if more than four consecutive residues of this peptide have negative $\Delta^{1}H_\alpha$ values of ≥ 0.1 ppm [20]. It is also well known that both $^{13}C_\alpha$ and $^{13}C'$ have positive $\Delta^{13}C_\alpha$ and $\Delta^{13}C'$ values but negative $\Delta^{13}C_\beta$ values for $^{13}C_\beta$ when they are located in helices [22, 23]. Fig. (**5**) shows negative $\Delta^{1}H_\alpha$ values for most residues in sequence region R126-K136. The positive $\Delta^{13}C_\alpha$ and $\Delta^{13}C'$ values for sequence regions S128-L137 and E122-E135, respectively were shown also in Fig. (**5**). It seems sequence R126-L137 may have tendency to helical conformation. As was indicated, $^{13}C_\beta$ resonances can only be used to identify stretches of strand [19], and the consensus CSI for locating helix was produced using $\Delta^{1}H_\alpha$, $\Delta^{13}C_\alpha$, and $\Delta^{13}C'$ of SNase(111-143). The produced consensus CSI indicates that segment E129-K136 can form a helical conformation in SNase(111-143).

For residues V111-S128 and L137-D143, the $^{1}H_\alpha$ resonances show randomly upfield and downfield shifts. The $^{13}C_\alpha$, $^{13}C_\beta$, and $^{13}C'$ resonances have the up- or down-shift values, but are not well consistent with each other. The $^{1}H_\alpha$ CSI are not in conjunction with the $^{13}C_\alpha$, $^{13}C_\beta$, and $^{13}C'$ CSI for identifying the residual secondary structures corresponding to structures in native SNase. Thus, the chemical shift data yield no indication of the secondary structure for both terminal regions of SNase(111-143).

NOE Connectivities

As was mentioned above, almost all $^{1}H_N$, $^{1}H_\alpha$, and $^{1}H_\beta$ resonances of SNase(111-143) in aqueous solution were assigned. Thus, sequential and medium or long-range NOE connectivities can be identified for the sequence of SNase(111-143) through d_{NN}, $d_{\alpha N}$ and $d_{\beta N}$ NOEs. Fig. (**6**) shows the selected strips from 3D ^{1}H-^{15}N NOESY-HSQC spectrum which is used to identify NOEs for SNase(111-143). The determined NOE connectivities are summarized in Fig. (**7**). SNase(111-143) did not show any long-range NOEs (more than five residues apart). The strong and medium sequential $d_{NN}(i, i+1)$ NOEs were observed for segments Y113-K116, N119-K127, S128-K136, and L137-S141, and the relatively weak $d_{NN}(i, i+1)$ NOEs were obtained for residues A112-Y113 and S141-E142. The vast majority residues in SNase(111-143) showed strong and medium sequential $d_{\alpha N}(i, i+1)$ NOEs. The medium sequential $d_{\beta N}(i, i+1)$ NOE connectivities were also observed for residues in segments Y113-P117, N119-H121, Q123-L124, and L125-E135, and weak $d_{\beta N}(i, i+1)$ NOEs were identified for residues E135-S141. The sequential $d_{\alpha N}(i, i+1)$, $d_{NN}(i, i+1)$, and $d_{\beta N}(i, i+1)$ NOEs obtained for most of the residues in the peptide usually indicate the presence of conformational averaging in a random coil ensemble [24]. Some medium-range NOEs were identified for sequence region N119-E134 of SNase(111-143). The medium-range $d_{NN}(i, i+2)$ NOE connectivities that is diagnostic of helix formation were identified only for a pair of residues R126-S128. Two medium-range $d_{\beta N}(i, i+2)$ NOEs were observed for pairs of residues S128-A130 and A132-K134, that can be also indicative of helix formation in peptide. Therefore, the helical conformation can be postulated to populate transiently in the conformational ensemble of SNase(111-143). The medium-range $d_{\alpha N}(i, i+2)$ and $d_{\alpha N}(i, i+3)$ NOE connectivities were determined for pairs of residues N119-H121 and N119-E122. The $d_{\alpha N}(i, i+2)$ NOE built between two residues which are three residues apart can appear in turn or 3_{10}-helix. The $d_{\alpha N}(i, i+3)$ NOE connectivity is usually expected for a helix.

In generally, the strong sequential $d_{\alpha N}(i, i+1)$ NOEs are indicative of extended structures in folded proteins. Thus, the strong and medium sequential $d_{\alpha N}(i, i+1)$ NOE connectivities identified for SNase(111-143) indicate the presence of an extended conformation in this peptide fragment. However, the sequential $d_{NN}(i, i+1)$ and $d_{\beta N}(i, i+1)$ NOEs for most of the residues of SNase(111-143) demonstrate that the fragment is not only in an extended conformation, but also can adopt some helix-like folded conformation. Therefore, SNase(111-143) in aqueous solution can be represented as an ensemble of interconverting conformations having transiently populated helix-like structures.

DISCUSSION

Based on the previous studies on SNase fragments with different chain length [6, 12, 13], the C-terminal region V111-D143 is denominated as a α-subdomain of the protein. The polypeptide chain of V111-D143 contains mostly hydrophilic residues and has an ordered tertiary structure under the long-range tertiary interactions in the native structure of SNase [5]. The ordered structure of the α-subdomain contains a 15-residue helix α_3 spanning over residues E122-K136. The polypeptide chain of the N-terminal residues V111-H121 of the α-subdomain has a turn conformation because the *cis*-conformer of peptidyl-proline bond of K116-P117 and a helical-like turn at N118-E122. Segment V111-K116 adopts mainly an extended conformation. The poly-

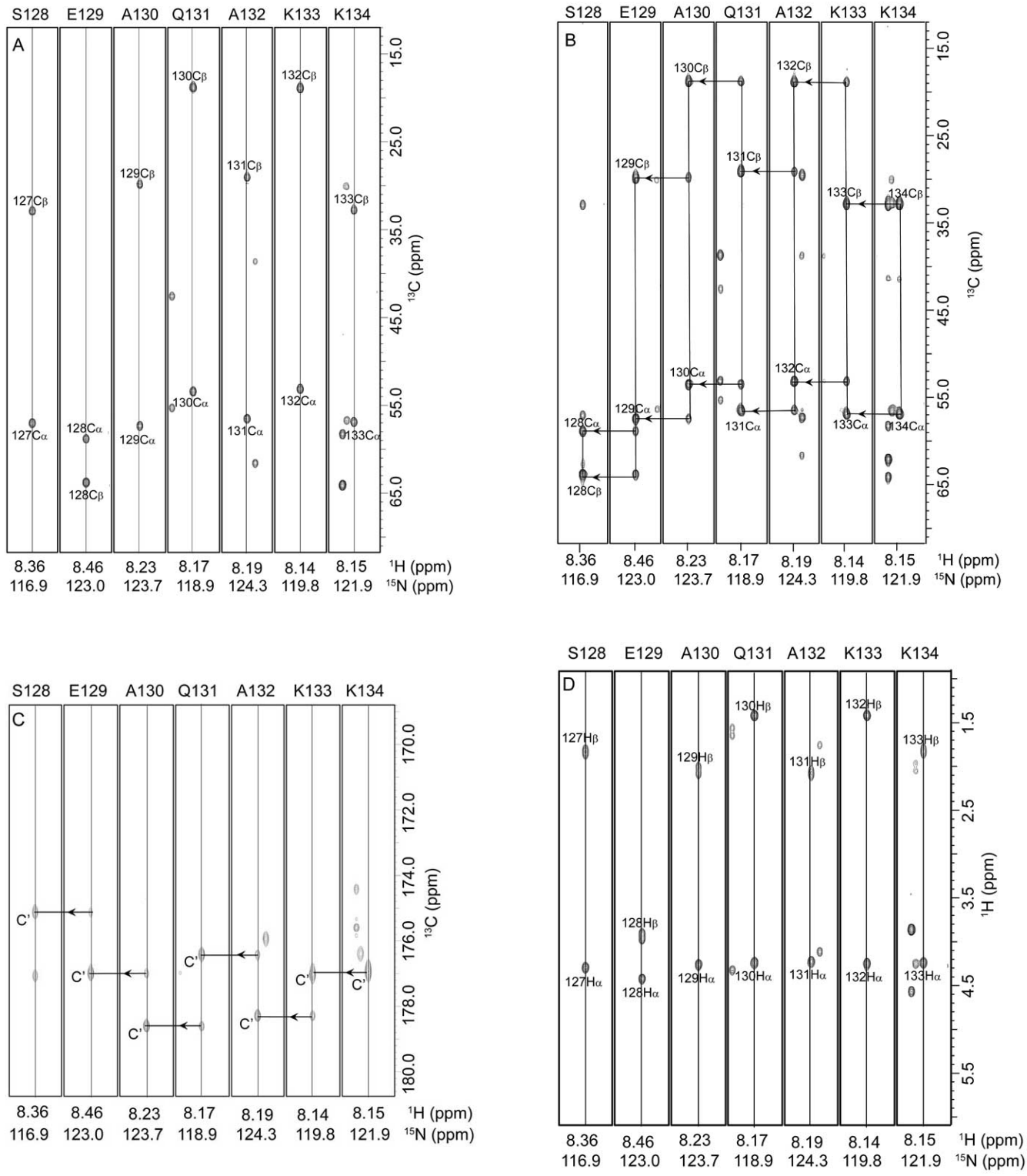

Figure 3. Selected strips from triple resonance correlation spectra of ^{15}N, ^{13}C-double labeled SNase(111-143). The $^{1}H_N$-^{13}C strips are from the 3D CBCACONH (A), HNCACB (B), and HNCACO (C) spectra. The $^{1}H_N$-^{1}H strips are from the 3D HBHACONH spectrum (D). Strips are extracted at ^{15}N chemical shifts of residues S128-K134 from all 3D NMR spectra and have been arranged in sequential order. The $^{1}H_N$ and ^{15}N chemical shifts of selected residues are marked at the bottom of each strip.

peptide chain of the C-terminal residues L137-D143 forms a helical-like loop which makes W140 functionally very important in the folding of full-length SNase, and W140 thus bears information important to both long-range and short-range interactions in the protein [25].

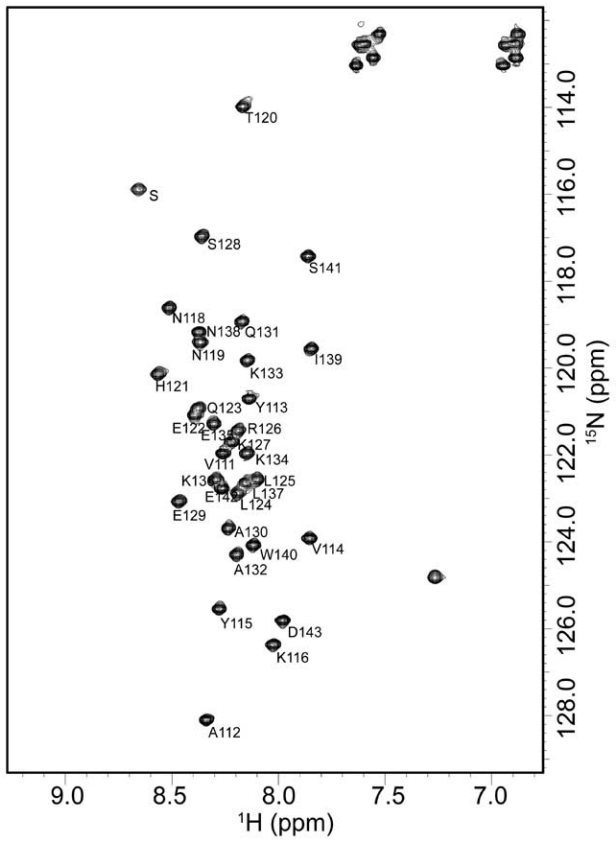

Figure 4. 2D ^{1}H-^{15}N HSQC spectrum of SNase(111-143) in aqueous solution. NMR resonance assignments were given by one-letter amino acid code and residue numbers.

Conformation of SNase(111-143) in Aqueous Solution

Many peptide fragments of proteins have been found to populate secondary structural conformations of the corresponding sequences in the parent protein, even lacking the long-range interactions for intact proteins. This is depicted by conformational preferences of peptide fragments in aqueous solution. CD and NMR experiments are usually done in investigation on the conformational preferences of peptides. The deviations from random coil values of CD data and NMR parameters can be indicative of the conformational preferences of peptides. The far-UV ellipticities, secondary chemical shifts and NOE connectivities of SNase(111-143) show deviations from the random coil values, that indicates the presence of some helical conformations in the fragment.

The determined NOE connectivities for SNase(111-143) indicate that the conformational averaging exists in SNase(111-143) in H_2O. This is expected for SNase(111-143) because of the flexibility and conformational interconversion as peptide usually behaviors in solution. Thus, most of assigned NOEs for SNase(111-143) are sequential and only a few identified medium-range NOEs can be used to describe the population of residual secondary structures in SNase(111-143). In tertiary structure of native SNase, segment E122-K136 is a helix α3. In aqueous solution, the fragment SNase(111-143) produces one $d_{NN}(i, i+2)$ and two $d_{\beta N}(i, i+2)$ medium-range NOEs for pairs of residues R126-S128, S128-A130, and A132-K134 in sequence region

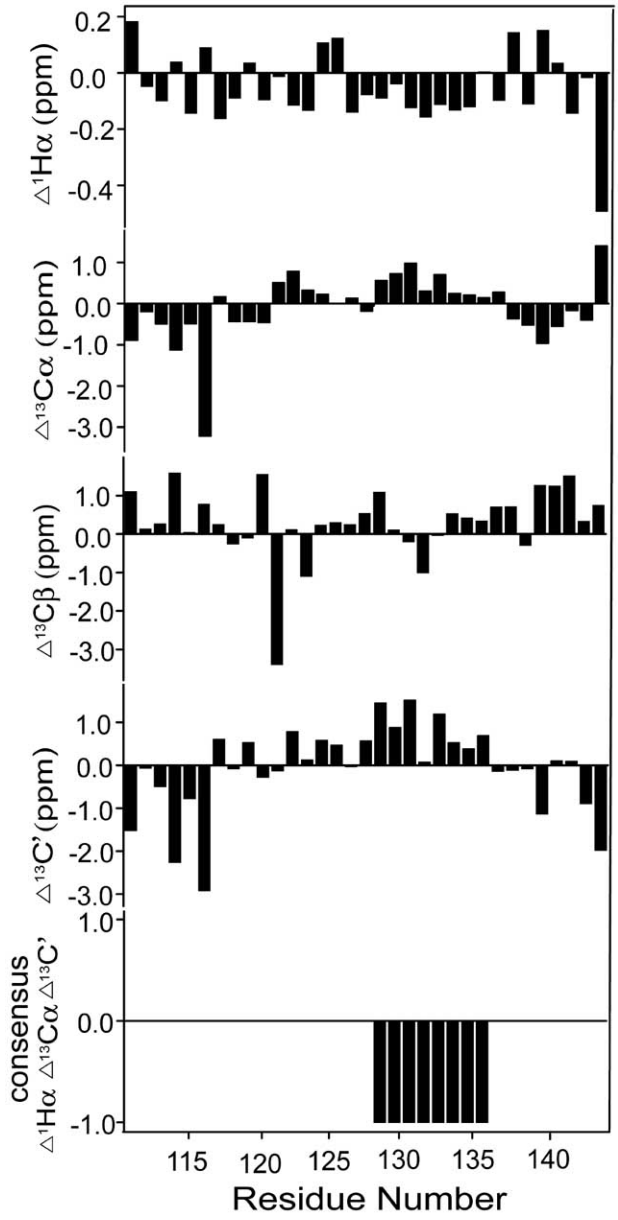

Figure 5. Histograms of the secondary chemical shifts for $^1H_\alpha$, $^{13}C_\alpha$, $^{13}C_\beta$, and $^{13}C'$ resonances of SNase(111-143) in H_2O. The Δ^1H_α, $\Delta^{13}C_\alpha$, $\Delta^{13}C_\beta$, and $\Delta^{13}C'$ values represent the chemical shift differences between those of the particular residues and the random coil values corrected by sequence. The consensus CSI derived by CSI [19, 20] is given for residues of SNase(111-143).

R126-K134 (Fig. **7**) which were observed also in native SNase [4, 5, 26]. Thus, the identified medium-range NOEs are indicative of propensity towards helix formation in the fragment. This is consistent with the estimation of consensus CSI that segment E129-K136 can form a helical conformation in SNase(111-143). The NOE connectivity pattern in sequence region N119-E122 of SNase(111-143) is very similar to that seen in native SNase. Two medium-range $d_{\alpha N}(i, i+2)$ and $d_{\alpha N}(i, i+3)$ NOE connectivities are observed for pairs of residues N119-H121 and N119-E122, respectively in SNase(111-143) (Fig. **7**). The similar $d_{\alpha N}(i, i+2)$ and $d_{\alpha N}(i,$

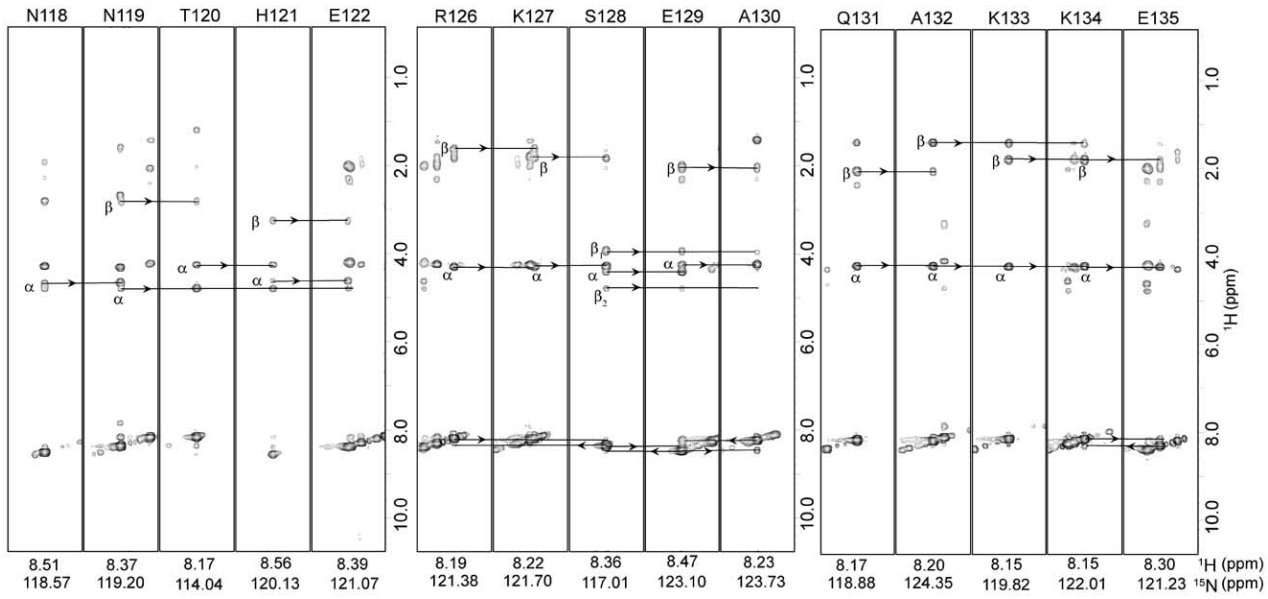

Figure 6. Strip plots of 2D ^{1}H-^{1}H planes from the 3D ^1H_N-^{15}N NOESY-HSQC spectrum of ^{15}N labeled SNase(111-143). The selected strips demonstrate the NOE connectivities in segments N118-E122, R126-A130, and Q131-E135 of SNase(111-143).

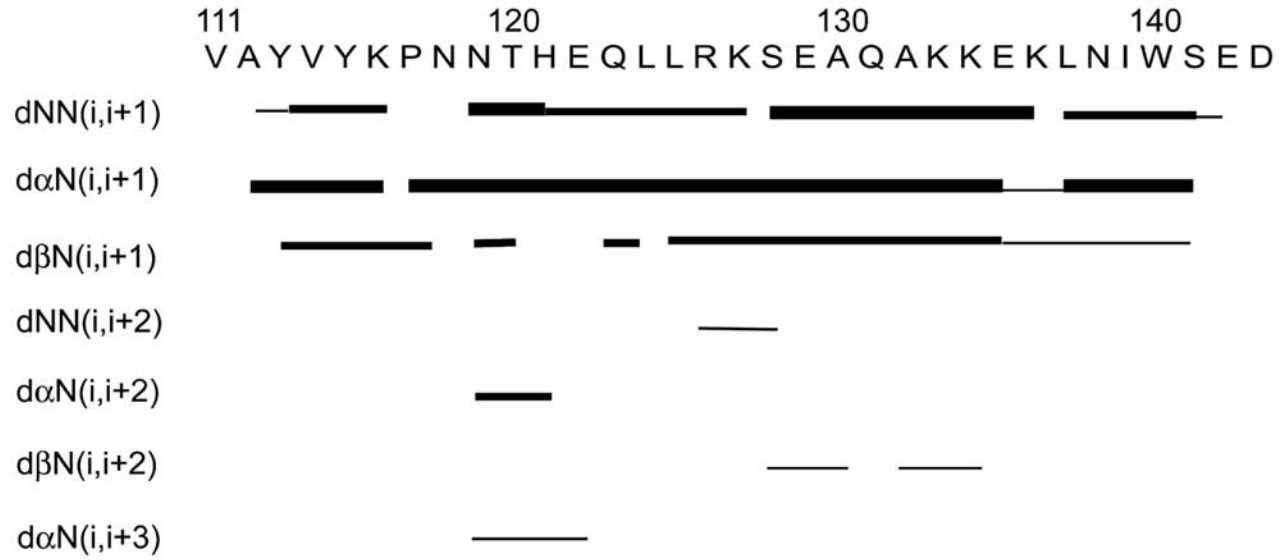

Figure 7. Summary of sequential and medium-range NOE connectivities obtained for SNase(111-143). The d_{NN}, $d_{\alpha N}$, and $d_{\beta N}$ NOE connecticities are indicated by bars. The relative NOE intensities are indicated by the thickness of the bars, and the medium-range NOEs are indicated by lines connecting the two coupled residues.

i+3) NOEs were determined for N118-T120 and N119-E122, respectively in native SNase. Besides, the medium and weak $d_{NN}(i, i+1)$ and $d_{\alpha N}(i, i+1)$ NOEs were also determined for N118-H121 and N118-E122, respectively, in native SNase [26]. These data indicate that the conformation of segment N118-E122 in SNase(111-143) is very similar to the conformation formed by N-terminal capping residues of helix E122-K136 in native SNase.

In consequence, SNase(111-143) shows the localized propensities for formation of helix- and turn-like structures in the region of N118-K136. The partially folded structure with helical segment can be postulated to populate transiently in the ensemble of interconverting conformations of SNase(111-143).

Implication in the Folding of Native SNase

In generally, the conformational preference of peptide for helical structure arises from the amino acid composition of the peptide. Sequence of SNase(111-143) predicts that the fragment has an inherent helix-forming tendency. The sequence region E122-K136 of the fragment SNase(111-143) (Fig. **7**) contains two alanine residues which is an excellent

amino acid for forming helix and two leucines, two glutamines, three glutamics, four lysines, and one arginine which are good residues in forming helix, and one serine residue as well. Thus, the amino acid residues in the segment E122-K136 have higher order of helix propensity. However, both the N- and C-terminal regions of the fragment SNase(111-143) have only four residues (A112, K116, L137, and E142) good for helix formation and the other 14 residues have low order of helix propensity. In the sequence region E122-K136, E129 and K133 are the two residues located on the same face of the corresponding helix α_3 in native SNase. As was indicated, the side-chain-to-side-chain lactam bridges of (*i*, *i*+4) spaced residues glutamic-lysine can induce or stabilize helical structure in peptides [27, 28]. This implies that the possible interactions between side chains of E129 and K133 in SNase(111-143) could induce the helical structure in the corresponding segment. The similar (*i*, *i*+4) spaced residues are observed for residues E122 and R126. Therefore, the amino acid composition of the fragment SNase(111-143) provides such an information that residues E122-K136 have strong intrinsic tendency towards the helical structure in aqueous solution.

Usually, peptides are shown to be flexible and have no ordered structure in solution owing to the lack of long-range interactions. However, some small peptide fragments of proteins have been found to form secondary structures seen in native protein. One of them is N-terminal 1-26 residues fragment of human protein PDCD5 (PDCD5(1-26)) [29]. Segment A3-A20 of human PDCD5(1-26) contains 16 residues having higher order of helix propensity. The far-UV ellipticities and secondary chemical shifts obtained for PDCD5(1-26) suggest strongly that residues A3-A20 of PDCD5(1-26) has high conformational preference for helix structure. The determined NOEs generated a stable well-formed α-helix of A3-A20. However, SNase(111-143) has low conformational preference for helix structure and can not generate a stable helical conformation in aqueous solution, although the amino acid composition of E122-K136 shows high intrinsic helical propensity. PDCD5(1-26) appear to be an autonomous folding element, having no long-range NOE contacts and tertiary interactions with the residues in main part of the protein. Nevertheless, the segment V111-D143 couples with β-subdomain of SNase (A1-K110 of SNase) by long-range interactions in native SNase. Tertiary structure of native SNase shows that some residues in the segments V111-H121 and L137-D143 have hydrogen-bonding and hydrophobic contacts with residues in the N-terminal subdomain of SNase, stabilizing the folding of segment V111-D143 in native SNase. Lacking of long-range interactions with parent protein causes very flexible N- and C-terminal segments of SNase(111-143) and influences the folding of segment E122-K136 having high intrinsic helical propensity.

Segments E57-A69 and M98-Q106 of SNase forming helices α1 and α2, respectively in native SNase have different features in amino acid sequences [5]. In the helix α1 having 50% hydrophobic residues, most of the hydrophobic residues are located on the same side of the helix. However,

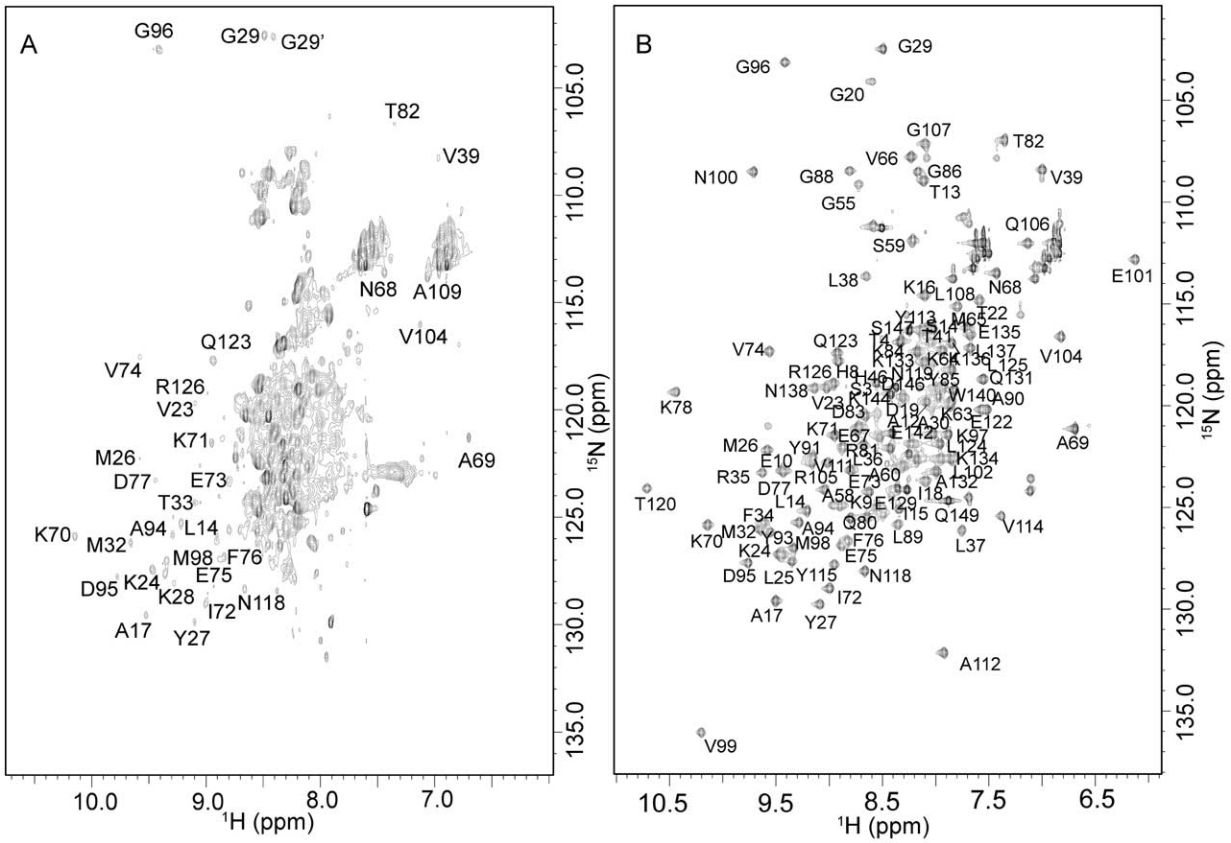

Figure 8. 2D ^{1}H-^{15}N HSQC spectra of full-length SNase and C-terminal truncated SNase. (A) Spectrum of SNase139 which is obtained by truncation of C-terminal residues W140-Q149 of intact SNase. (B) Spectrum of native SNase.

helix α2 contains 60% hydrophobic residues that are distributed along the helix in native SNase. The previous studies have indicated that formation of native-like helices in segments E57-A69 and M98-Q106 require non-local interactions through transient contact with hydrophobic residues in other parts of native SNase [30]. Unlike the segments E57-A69 and M98-Q106 of helices α1 and α2, segment E122-K136 of SNase(111-143) contains only 27% hydrophobic residues (two leucines and two alanines). Only residue L125 in the segment E122-K136 may have hydrophobic interaction with hydrophobic residues from main part of native SNase. Thus, it can be postulated that the hydrophobic environment is not very crucial for the folding of segment E122-K136 in native SNase.

The CD data obtained for SNase(111-143) and SNase(118-143) indicate that the flexible segment V111-P117 on the N-terminal end of segment E122-K136 in SNase(111-143) enhances a population of random coil conformation of the fragment. Truncation of C-terminal residues W140-Q149 of SNase generating a 1-139 residues SNase fragment (SNase139) destroys the native long-range interactions of segment L137-D143 with β-subdomain exhibited in full-length SNase. SNase139 lacks the long-range interactions between its C-terminal residues and the parent protein, but may retain the native-like interactions of segment V111-H121 on the N-terminal end of segment E122-K136 with the main portion of SNase139. The 2D ^{1}H-^{15}N HSQC spectrum of SNase139 (Fig. **8**), providing the cross-peaks for residues A109, N118, Q123, and R126 which have the same chemical shifts as native structure, demonstrates a native-like folding for the N-terminal region of segment V111-I139 in SNase139, but not for the C-terminal region. The above results imply that the stable helical folding of segment E122-K136 requires the interactions of segments V111-P117 and L137-D143 on both ends of the segment E122-K136 with parent protein. In SNase(111-143), the segment E122-K136 can adopt very little helical structure and show low conformational preference for helix structure, probably because of that both terminal regions of SNase(111-143) are flexible by lacking the long-range interactions with the parent protein. The above analysis suggests that the stable folding of fragments V111-H121 and L137-D143 flanked on both ends of segment E122-K136 is crucial for formation of α-helical structure of segment E122-K136 in the C-terminal subdomain of SNase.

In conclusion, this study has indicated that both the N- and C-terminal regions of the fragment SNase(111-143) are flexible and the segment E122-K136 can adopt little helical structure centralized in region E129-K136. Lacking the long-range interactions with the parent protein, the partially folded structure with helical segment can be postulated to populate transiently in the ensemble of interconverting conformations of SNase(111-143). A comprehensive analysis of the conformational features of the fragments SNase(111-143) and SNase(118-143) and the segment E122-K136 in SNase139 suggests that the high intrinsic helical propensity can drive segment E122-K136 fold into a stable helix only when the fragments V111-H121 and L137-D143 flanked on both ends of segment E122-K136 in the C-terminal sub-domain of SNase, namely segment V111-D143, have stable folding.

ACKNOWLEDGEMENTS

This research was supported by the National Natural Science Foundation of China (NNSFC 30570375).

REFERENCES

[1] Maity, H. Maity, M. Krishna, M.M.G. Mayne, L. and Englander, S. (**2005**) *Proc. Natl. Acad. Sci. USA*, *102*, 4741-4746.
[2] Krishna, M.M.G. Maity, H. Rumbley, J.N. Lin, Y. and S.W. (**2006**) *J. Mol. Biol.*, *359*, 1410–1419.
[3] Fersht, A.R. and Daggett,V. (**2007**) *Curr. Opin. Struct. Biol.*, *17*, 1–2.
[4] Loll, P.J. and Lattman, E.E. (**1989**) *PROTEINS: Structure, Function, and Genetics*, *5*, 183-201.
[5] Wang, J. Truckses, D.M. Abildgaard, F. Dzakula, Z. Zolnai, Z. and Markley, J.L. (**1997**) *Journal of Biomolecular NMR*, *10*, 143-164.
[6] Ye, K. Jing, G. and Wang, J. (**2000**) *Biochim. Biophys. Acta.*, *1479*, 123-134.
[7] Alexandrescu, A.T. Gittis, A.G. Abeygunawardana, C. and Shortle, D. (**1995**) *J. Mol. Biol.*, *250*, 134-143.
[8] Wang, Y. and Shortle, D. (**1995**) *Biochemistry*, *34*, 15895-15905.
[9] Ermacora, M.R. Ledman, D.W. and Fox, R.O. (**1996**) *Nat. Struct. Biol.*, *3*, 59-66.
[10] Eftink, M.R. Ionescu, R. Ramsay, G.D. Wong, C. Wu, J. and Maki, A.H. (**1996**) *Biochemistry***,** *35*, 8084-8094.
[11] Hirano, S. Mihara, K. Yamazaki, Y. Kamikubo, H. Imamoto, Y. and Kataoka, M. (**2002**) *PROTEINS: Structure, Function, and Genetics*, *49*, 255-265.
[12] Feng,Y. Liu, D. and Wang, J. (**2003**) *J. Mol. Biol.*, *330*, 821-837.
[13] Xie, T. Liu, D. and Wang, J. *Biophys.J.*, In press.
[14] Mark, W. Zehfus, M.H. (**1995**) *Biochemistry*, *34*, 5795-5800.
[15] Cheng, Y. Liu, D. Feng, Y. and Jing, G. (**2003**) *Protein Pept. Lett.*, *10*, 175-81.
[16] Sattler, M. Schleucher, J. and Griesinger, C. (**1999**) *Progress in Nuclear Magnetic Resonance Spectroscopy*, *34*, 93–158.
[17] Zhu, G. and Bax, A. (**1992**) *J. Magn. Reson.*, *100*, 202–207.
[18] Cavanagh, J. Fairbrother, W.J. Palmer, A.G. and Skelton, N.J. (**1996**) *Protein NMR Spectroscopy*, Academic Press, San Diego.
[19] Wishart, D.S. Sykes, B.D. (**1994**) *J. Biomol. NMR*, *4*, 171–180.
[20] Wishart, D.S. Sykes, B.D. and Richards, F.M. (**1992**) *Biochemistry*, *31*, 1647–1651.
[21] Schwarzinger, S. Kroon, G.J. Foss, T.R. Chung, J. Wright, P.E. and Dyson, H.J. (**2001**) *J. Am. Chem. Soc.*, *123*, 2970–2978.
[22] Wishart, D.S. Sykes, B.D. and Richards, F.M. (**1991**) *J. Mol. Biol.*, *222*, 311–333.
[23] Wishart, D.S. and Sykes, B.D. (**1994**) *Methods Enzymol.*, *239*, 363-392.
[24] Fiebig, K.M. Schwalbe, H. Buck, M. Smith, L.J. and Dobson, C.M. (**1996**) *J. Phys. Chem.*, *100*, 2661-2666.
[25] Hirano, S. Kamikubo, H. Yamazaki, Y. and Kataoka, M. (**2005**) *Proteins: Structure, Function, and Bioinformatics*, *58*, 271-277.
[26] Wang, J. LeMaster, D.M. and Markley, J.L. (**1990**) *Biochemistry*, *29*, 88-101.
[27] Mierke, D.F. Maretto, S. Schievano, E. DeLuca, D. Bisello, A. Mammi, S. Rosenblatt,M. Peggion, E. and Chorev, M. (**1997**) *Biochemistry*, *36*, 10372-10383.
[28] Houston, M.E. Campbell, A.P. Lix, B. Kay, C.M. Sykes, B.D. and Hodges, R.S. (**1996**) *Biochemistry*, *35*, 10041-10050.
[29] Liu, D. Yao, H. Chen, Y. Feng, Y. Chen, Y. and Wang, J. (**2005**) *Biochem. J.*, *392*, 47–54.
[30] Wang, M., Shan, L. and Wang, J. (**2006**) *Biopolymers*, *83*, 268-279.

Received: March 27, 2007 Revised: June 10, 2007 Accepted: June 13, 2007