J. Biochem. 128, 337–344 (2000)

Structural Study of the N-Terminal Domain of the Alpha Subunit of *Escherichia coli* RNA Polymerase Solubilized with Non-Denaturing Detergents[1]

Takanori Otomo,* Toshio Yamazaki,* Katsuhiko Murakami,† Akira Ishihama,† and Yoshimasa Kyogoku*,[2]

*Institute for Protein Research, Osaka University, Suita, Osaka 565-0871; and †National Institute of Genetics, Mishima, Shizuoka 411-0801

Received May 23, 2000; accepted June 8, 2000

The amino-terminal domain of the α subunit (αNTD) of *Escherichia coli* RNA polymerase consisting of 235 amino acid residues functions in the assembly of the α, β, and β′ subunits into the core-enzyme. It has a tendency to form aggregates by itself at higher concentrations. For NMR structural analysis of αNTD, the solution conditions, including the use of non-denaturing detergents, were optimized by monitoring the translational diffusion coefficients using the field gradient NMR technique. Under the optimal conditions with taurodeoxycholate and with the aid of deuteration of the sample, αNTD gave triple-resonance spectra of good quality, which allowed the assignment of a large part of the backbone resonances. Analysis of the pattern of NOEs observed between the backbone amide and α-protons demonstrated that αNTD has three α-helices and two β-sheets. Although the secondary structure elements essentially coincide with those in the crystal structure, the larger of the two β-sheets has two additional β-strands. The irregular NOE patterns observed for the three positions in the β-sheets suggest the presence of β-bulge structures. The positions of the three helices coincide with the conserved sequence regions that are responsible for the subunit assembly.

Key words: detergent, deuteration, diffusion coefficient, NMR, RNA polymerase.

The *Escherichia coli* RNA polymerase core enzyme has the subunit composition of $\alpha_2\beta\beta'$. It is assembled in the sequence: $\alpha \rightarrow \alpha_2 \rightarrow \alpha_2\beta \rightarrow \alpha_2\beta\beta'$ (*1*). This process starts with the dimerization of the α subunit, and then the resulting dimer binds to the β and β′ subunits. The α subunit, which is the smallest subunit of 329 amino acids, can be divided into two independently folded domains (*2*). The amino-terminal two-third domain (αNTD) is responsible for this enzyme assembly function *in vitro* (*3*) and *in vivo* (*4*, *5*). The carboxyl-terminal one-third domain (αCTD) plays a key role in transcriptional regulation with a group of protein transcription factors and DNA enhancer elements (*6*, *7*). In 1995 our group determined the three-dimensional solution structure of αCTD by NMR spectroscopy and correlated the structure with the activation mechanism (*8*). The relaxation data also showed that αCTD is connected to αNTD by a flexible linker consisting of 13 amino acids (*9*). It was interesting to determine the structure of αNTD and to correlate it with its function. We have tried to determine the solution structure of αNTD by NMR, but it was challenging.

For NMR analysis of αNTD, there were two problems, its high molecular mass and its tendency to aggregate. αNTD comprises 235 amino acid residues, and forms the dimer even at a sub-millimolar concentration. The molecular mass of the dimer, 54 kDa, is close to the upper limit even if deuterated samples and optimized NMR techniques are used (*10–12*). Initially we tried to use mutants which were expected to destabilize the formation of the dimer (*13*), but we could not find any mutant which simply gives a well resolved spectrum. This failure may due to that such a single-site mutation cannot eliminate all non-specific interactions causing aggregation. Since this domain functions in the assembly of other subunits, many of the surface residues are hydrophobic and cause non-specific aggregation at an NMR concentration. Previously we showed that αNTD was essentially invisible in a ^{15}N-^{1}H HSQC spectrum even though the solution was transparent (*9*). αNTD seemed to form aggregates of tens of molecules.

In this paper, we describe some details of optimization of the solution conditions for αNTD, because we believe that the accumulation of knowledge of how to obtain optimal solution conditions must be important for extending the range of NMR applications. The target protein must be soluble without aggregation at a millimolar concentration. Even if it is soluble, non-specific aggregation of the molecules often makes the effective molecular mass much

[1] This work was supported by a Grant-in-Aid for a Specially Promoted Area from the Ministry of Education, Science, Sports and Culture of Japan. T.O. was a Research Fellow of the Japan Society for the Promotion of Science.

[2] To whom correspondence should be addressed at the present address: Fukui Institute of Technology, 3-6-1 Gakuen, Fukui 910-8505. Tel: +81-6-6879-8598, Fax: +81-6-6879-8600, E-mail: kyogoku@protein.osaka-u.ac.jp

Abbreviations: αNTD, the amino-terminal domain of the α subunit of *Eschelichia coli* RNA polymerase; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; DOC, deoxycholate; TDOC, taurodeoxycholate.

higher, leading to the rapid decay of NMR signals.

In general, the pH and salt concentration are the factors initially examined to obtain an optimal solution giving a good spectrum. However, when the hydrophobic interaction is strong, the aid of detergents is necessary. It has been shown that a detergent, CHAPS, was effective for avoiding the aggregation of proteins and for improving NMR spectra (*14*). Moreover, it was recently reported that the solution structure of a protein solubilized with CHAPS was successfully determined (*15*).

Pulsed field gradient (PFG) NMR techniques have been shown to be suitable for determining tranlational diffusion coefficients in solution as estimates of the effective molecular masses of biological macromolecules (*16–18*). Recently, this technique was applied to monitoring of the self-association of protein molecules (*19, 20*). In the present paper, we report effects on the effective molecular masses of detergents using this technique. Finally, we succeeded in observing most of the signals in ^{15}N-^{1}H HSQC spectra under the optimized solution conditions with taurodeoxycholate.

It has been shown that the deuteration of protein samples makes the relaxation time of ^{13}C magnetization six or seven times longer (*10*). This leads to dramatic improvements in the sensitivity of triple resonance experiments such as HN(CA)CB and HN(COCA)CB (*11*). In NOESY experiments, because the spin diffusion effect is greatly reduced by the low proton density, weak H_N-H_N connectivities can be enhanced by employing a longer mixing time (*21*). We obtained a 70% deuterated ^{13}C- and ^{15}N-labeled sample, and performed deuteration-optimized triple-resonance experiments. These experiments allowed the assignment of 84% of the backbone amide resonances. We report the secondary structure of αNTD and its global chain topology, and compare them with the crystal structure, which was published during this study (*22*).

MATERIALS AND METHODS

Sample Preparation—The amino-terminal domain of the α subunit of 239 amino acid residues (residues 1 to 235 plus a peptide, VKLT, at its C terminus) was expressed in *E. coli* BL21(λDE3) cells transformed with the expression plasmid pGEMAD235 (*6*). To achieve uniform ^{15}N labeling, and ^{15}N and ^{13}C labeling of the protein, the bacteria were grown in M9 minimal medium containing $^{15}NH_4Cl$ (0.5 g/liter) as the sole nitrogen source and ^{13}C-glucose (2 g/liter) as the sole carbon source. The protein was purified as described previously (*6*). The samples for PFG NMR measurements comprised 0.5 mM protein in 100% D_2O, containing 20 mM Tris buffer (pH 8.5, meter direct reading) and 0.1 mM EDTA with or without various detergents.

Triply labeled (70% ^{2}H, uniform ^{13}C and ^{15}N) samples were prepared as follows. Cells transformed with the plasmid were amplified in 5 ml of a H_2O M9 culture up to OD (620 nm) = 0.1 and then transferred to 100 ml of a 70% D_2O/30% H_2O M9 culture containing 2 g/liter of ^{13}C-glucose, 0.5 g/liter of $^{15}NH_4Cl$, and vitamins (thiamin, folic acid, pantothenate, biotin, and pyridoxal; 1 mg/liter of each). The culture was transferred to 1 liter of the same culture when OD reached 0.1. IPTG was added when OD reached 0.5 for induction of the protein, followed by incubation for 5 h. It was important that cells were transferred to a larger culture at a lower cell concentration than 0.1, otherwise, we often saw dead cells. Culturing at a lower temperature like 25°C also seemed a good way of avoiding the death of cells. Through the same purification procedure, 20 mg of ^{2}H-, ^{13}C, and ^{15}N-labeled protein was obtained from 1 liter culture.

For Lys, Arg, and Met specific ^{15}N labeling, the respective ^{15}N-labeled amino acid was added to the M9 culture at the time of induction. For Val, Leu, Ala, and Ile specific labeling, an amino acid culture (*23*) was used. The amino acids in the culture medium were all unlabeled except for that to be labeled. The ^{15}N-labeled amino acid was added to the final concentration of 100 mg/liter at the time of induction with β-chloroalanine (2 mM), which is known as an inhibitor of branched-chain aminotransferase (*24*). Cells were harvested after an additional 2–3 h incubation and then subjected to the standard purification procedure (*6*).

PFG NMR Experiments—All NMR spectra were acquired at 30°C with a 500 MHz spectrometer, Bruker DMX500. The magnetic field gradients were generated with an actively shielded coil assembled around the radiofrequency coils of a 5 mm triple resonance probe. The magnitude of the magnetic field gradient (G) was 0.69 T/m at maximum, as measured by the manufacturer's standard procedure (Bruker). The PFG longitudinal eddy-current delay pulse sequence was used to measure the translational diffusion coefficients of detergents and αNTD. The pulse sequence was: 90°– τ (PFG)–90°– T–90°–τ (PFG)–90°–T_e–90°–Acq (*14*). Here, T is the time for the diffusion of protein molecules. At the beginning of the τ period a sine-bell-shaped field gradient is applied along the z-axis with a top amplitude of G and a width of δ. During the T and T_e periods weak gradient pulses were applied along the x- and y-axes, respectively. For each solution condition a series of 9 spectra was obtained in which the length of the field-gradient pulse (δ) was increased in 0.5 ms steps from 0.5 ms to 4.5 ms. Other fixed parameters were τ = 5 ms, T = 100 ms, and T_e = 2 ms. All NMR spectra consisted of 4,096 points with a spectral width of 8 kHz. For each 1D spectrum 1,024 transients were accumulated for the samples with detergents, and 4,096 transients were accumulated for the sample without a detergent. Spectra were processed and analyzed with the XWIN-NMR software (Bruker).

Analysis of PFG NMR Experimental Results—The NMR signal intensity of a molecule undergoing unrestricted translational diffusion with the diffusion coefficient of D in liquid is described by:

$$I = I_0 \exp[-k^2 D(\Delta + \alpha)],$$

where k ($=2\gamma G\delta/\pi$, for a sine-bell) is the wave number generated by the gradient pulse, Δ ($= T + \tau - \delta$) the time between the gradient pulses, and α ($= 3\delta/4$ for a sine-bell) the correction time including the effect during the gradient pulses. The relative intensities of protein signals were fitted to this equation using a general nonlinear least-squares method to obtain D. The magnitude of the signal from aromatic protons at ~7 ppm was used to determine the signal intensities of the protein. Methyl protons of the protein, which gave more intense signals, could not be used because they overlapped the detergent signals.

Triple Resonance Experiments—Each experiment took 4 days. The ^{2}H-, ^{13}C-, and ^{15}N-labeled protein was concentrated to 0.6 mM in 50 mM Tris-HCl buffer, adjusted at pH 8.5, including 7 mM taurodeoxycholate. All the triple resonance experiments were carried out at 30°C with a 600

MHz spectrometer, Bruker DRX600. HNCA and HN(CO)CA spectra, which gave sequential connections through the C_α chemical shifts, were recorded with and without the constant-time scheme for the C_α chemical shifts (*10*, *25*). HN(CA)CB and HN(COCA)CB spectra, which gave sequential connections through the C_β chemical shifts, were recorded without the constant-time scheme for the C_β chemical shifts (*11*). Typical conditions for the above experiments were a spectral width of 14 ppm centered at 4.7 ppm for the ^1H_N axis, a spectral width of 28 ppm centered at 117 ppm for the ^{15}N axis, a spectral width of 30 ppm centered at 54 ppm for the C_α axis, and a spectral width of 70 ppm centered at 40 ppm for the C_β axis. The repetition time was 1–2 seconds. ^{15}N-edited 3D NOESY spectra were obtained for the deuterated $^{13}C^{15}N$-labeled and protonated ^{15}N-labeled sample. A ^{15}N-edited 3D TOCSY spectrum was obtained for the protonated ^{15}N-labeled sample. Typical parameters for these experiments were a spectral width of 14 ppm centered at 4.7 ppm for ^1H_N (acquisition), a spectral width of 12 ppm centered at 4.7 ppm for the indirect ^1H axis, and a spectral width of 28 ppm centered at 117 ppm for the ^{15}N axis. The mixing time for NOESY of the deuterated sample was 200 ms and that for the protonated sample was 150 ms, and the mixing time for TOCSY was 40 ms. The NOESY spectrum was obtained with a 800 MHz spectrometer, DRX800, and the others with a 600 MHz spectrometer, DRX600.

RESULTS AND DISCUSSION

Solution Conditions for NMR Samples—At the beginning αNTD seemed to be suitable for an NMR study because its solubility was high, but the spectrum was very poor (*9*). The sample was dissolved in a typical solution, *i.e.* 1 mM protein in 20 mM phosphate buffer, adjusted to pH 6.8, including 1 mM DTT, 0.1 mM EDTA, and 100 mM NaCl. It was also shown by means of dynamic light scattering measurements that αNTD forms aggregates in the range of 150 to 300 kDa (*9*). No cross peaks were detected in the ^{15}N-^1H HSQC spectrum except for several signals arising from the unfolded part because slow overall tumbling caused rapid transverse relaxation of the signals.

To avoid aggregation we roughly examined conditions like the pH, salt concentration, divalent salts, organic solvents and acquired ^{15}N-^1H HSQC spectra. As a result, no significant change was observed except that αNTD precipitated around the isoelectric point (estimated to be 4.9).

Therefore, the conditions for NMR measurements were examined using detergents. Cationic and nonionic detergents having a long aliphatic chain (*e.g.*, 1-laurylpyridinium chloride, *n*-octyl-β-D-glucoside, *etc.*) showed a tendency of precipitation. Detergents which had a steroid-ring structure with an anion, zwitterion, or nonion did not precipitate αNTD and improved the ^{15}N-^1H HSQC spectra. We examined five detergents which seemed to be good for avoiding aggregation in PFG NMR experiments. These detergents have short aliphatic chains connected to the steroid-ring. DOC and cholate have a negative charge on the carboxyl group and TDOC also has a negative charge on the sulfonyl group. CHAPS has both positive and negative charges. The optimal concentrations of detergents have to be determined carefully because too high concentrations of detergents cause the rapid disappearance of signals in hours or days due to denaturation of the protein.

Figure 1 comprises plots of the diffusion coefficient of αNTD *versus* the concentrations of detergents. With no detergent, the diffusion coefficient of αNTD was 5.74×10^{-11} m^2/s. It was found that the titration curves of the detergents had similar patterns. The dissociation of aggregates of αNTD started at the first point of the plots, where the concentration of the detergent was lower than the critical micelle concentration (CMC). A report has indicated that dissociation of aggregates occurs at around CMC (*14*), but different behavior was observed in our case. There was a maximum point corresponding to the smallest molecular mass in each plot. *D* decreases as the concentration of a detergent increases over this point. This is partly explainable by the increase in the effective mass caused by the bound detergent. The self-obstacle effect that reduces

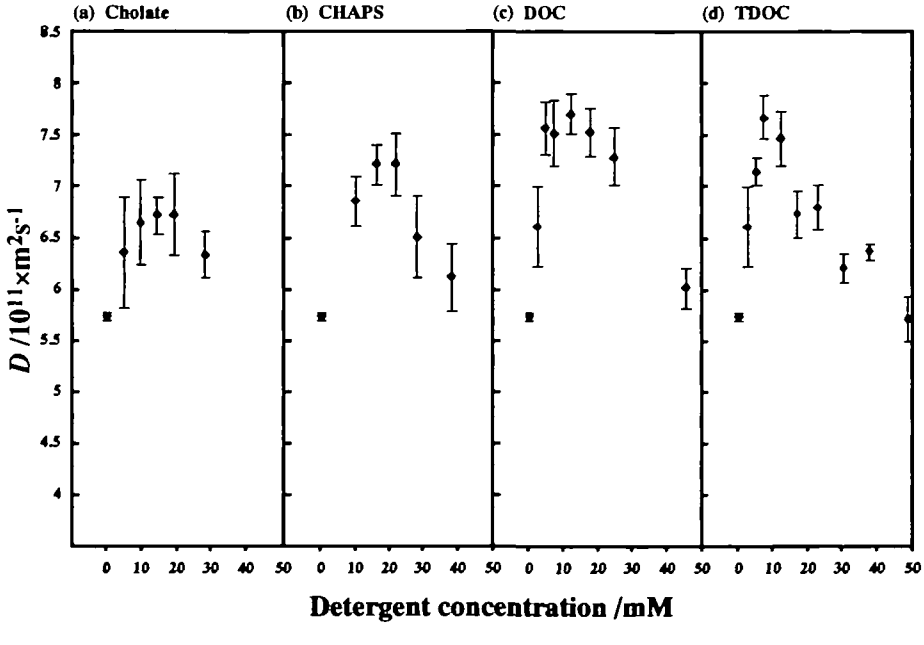

Fig. 1. **Plots of the translational diffusion coefficients of αNTD titrated with detergents, *i.e.*, (a) cholate, (b) CHAPS, (c) DOC, and (d) TDOC.** Error bars denote ±1 standard deviation.

translational diffusion is also enhanced as the concentration of the detergent increases. The detergents used here have very similar chemical structures including a steroid-ring, but the *D* values and concentrations of detergents at the maximum points, and the slopes of the plots were different. Figure 2 shows the ^{15}N-^{1}H HSQC spectrum with each detergent at the concentration for the maximum *D* value. The best spectrum was obtained with either TDOC or DOC, which gave the maximum *D* value among the five examined detergents. As judged from the ratio of the *D*s with and without a detergent, a 2.5 times reduction in molecular mass was achieved. Cholate, which exhibited a lower maximum *D* value, gave a ^{15}N-^{1}H HSQC spectrum of lower quality. There was only a slight difference in the maximum *D* value between CHAPS ($D = 7.22 \times 10^{-11}$ m^2/s at 16 or 21.7 mM) and DOC (7.71×10^{-11} m^2/s at 12.2 mM) or TDOC (7.68×10^{-11} m^2/s at 7.4 mM), but there were remarkable differences in the quality of the ^{15}N-^{1}H HSQC spectra because the effective masses are proportional to D^{-3}. That is, the ^{15}N-^{1}H HSQC spectrum with CHAPS was of a little lower quality than those with DOC and TDOC. DOC and TDOC gave about the same *D* values and ^{15}N-^{1}H HSQC spectra of good quality. We can observe virtually all the signals in the ^{15}N-^{1}H HSQC spectra.

Can the present finding be applied to other proteins? (i) When the charge of a detergent is the same as the net charge of a protein, binding of the detergent increases the net charge of the protein and solubilizes it. On the contrary, detergents with the opposite charge have a tendency to precipitate the protein. (ii) Detergents having a rigid hydrophobic part (*i.e.* steroid rings in this case) have a lower ability to denature a protein than ones having a long and flexible chain. (iii) The fact that deoxycholate was more effective than cholate suggests that detergents of low solubility can bind to proteins at lower concentrations. As an excess amount of a detergent increases the effective molecular mass, it is better to use a minimal amount of a detergent which is effective at lower concentrations to obtain a smaller effective mass.

Other factors, such as the pH, salt concentration and protein concentration, were also important for the protein stability. The final solution conditions that we chose were 0.6 mM protein in 20 mM Tris-HCl buffer, adjusted to pH 8.5, containing no salt and 7 mM TDOC. The stability of αNTD with DOC was lower than that with TDOC. Cross peaks in the ^{15}N-^{1}H HSQC spectrum of αNTD with DOC disappeared within a week but the spectrum with TDOC did not change for a month at 30°C. Therefore, we chose TDOC for αNTD. At a lower pH, αNTD was denatured so quickly even with TDOC. The concentration of the protein could not be very high because the rate of denaturation of the protein became faster at a higher concentration. The concentration of the protein in this experiment with detergents is lower than that preferred for structure determination, but the high sensitivity pulse sequence with a deuterated protein (*10*, *11*, *26*) made it possible to assign the NMR signals of αNTD, as described below.

In order to determine the rotational correlation time (τ_C) of αNTD, T_1, T_2, and NOE of the amide nitrogen were measured using convenient 1D spectra (data not shown). To eliminate the effect of protein internal mobility, we employed the signals at 9 ppm for the analyses in which the amide proton signals mostly arise from the rigid parts of the protein like a β-sheet. The estimated average backbone amide nitrogen T_2 value was about 45 ms, giving an estimated τ_C of 17 ns. This value is comparable to those of the largest proteins for which NMR studies were performed. We also performed a filtered ^{15}N edited 3D NOESY experiment for a 1:1 mixture of the ^{2}H-, ^{13}C-, and ^{15}N-labeled sample and the unlabeled sample. If the dimer state is stable in this solution containing a detergent, it should give intersubunit NOEs, but none were observed. We conclude that αNTD is a monomer in a solution with a detergent. The diffusion coefficients described above were smaller than those for the monomer protein. As the detergent decreases the diffusion coefficient through an increase in the effective mass on binding and also through a self-obstacle effect, the value derived on extrapolation to zero detergent concentration should be a better estimate of the molecular mass.

Backbone Resonance Assignment—In order to obtain the

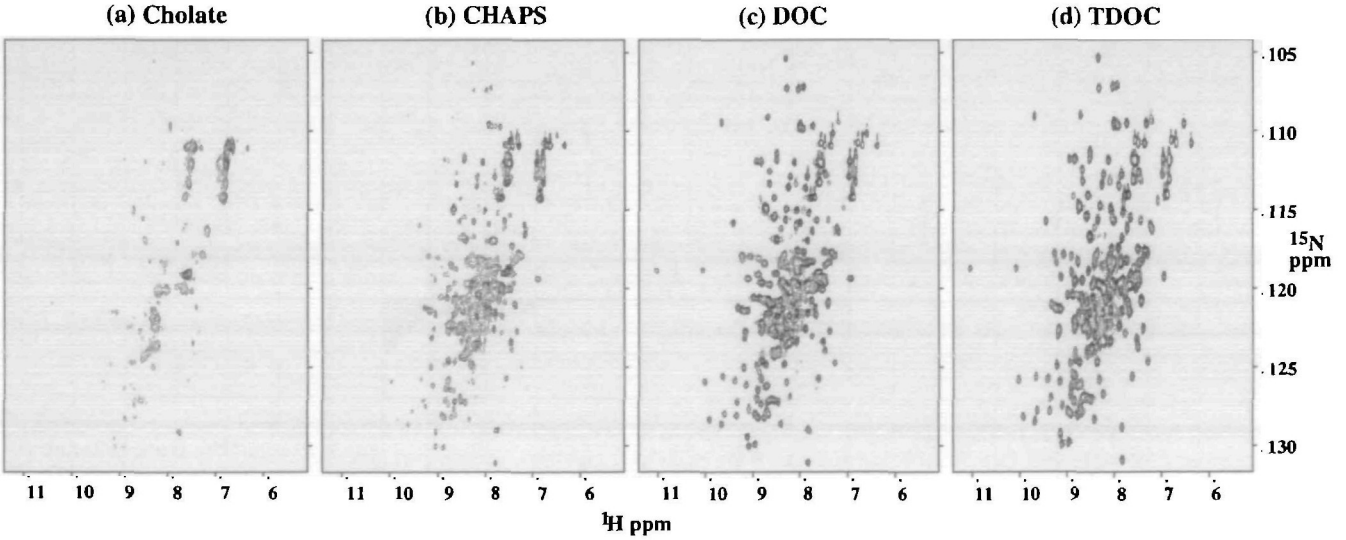

Fig. 2. ^{15}N-^{1}H HSQC spectra of 0.6 mM αNTD acquired in the presence of (a) 16 mM cholate, (b) 20 mM CHAPS, (c) 12 mM DOC, and (d) 7 mM TDOC in 20 mM TrisHCl (pH 8.5) buffer.

highest sensitivity for the large protein under the poor conditions, we used a deuterated sample and conducted amide-based triple resonance experiments optimized for a deuterated sample. The transverse relaxation of the C_α spins can be made as long as 100 ms by replacing the directly attached ^1H with ^2H. H_Ns in the deuterated protein give much sharper resonances because the transverse relaxation of the H_N spins also decreased and the H_N-H_α scalar couplings were eliminated. Taking such effects together, the spectra of the deuterated sample showed essential improvements. The difference was easily seen even in the ^{15}N-^1H HSQC spectrum.

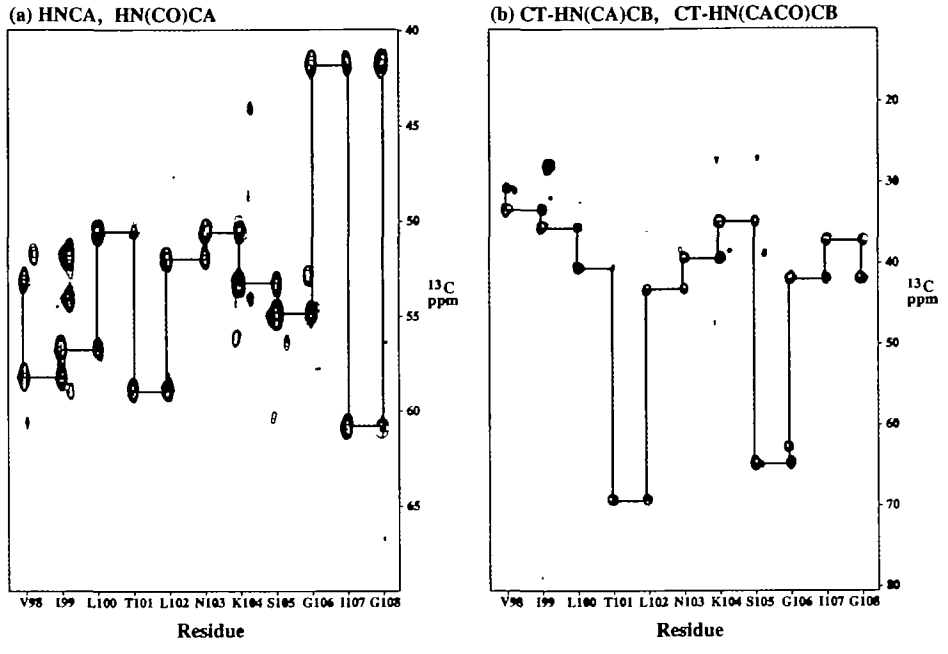

Fig. 3. **Strip plots displaying the sequential connectivities for the β-5 strand (residues 98–108) in (a) HNCA and HN(CO)CA (red), and (b) CT-HN(CA)CB and CT-HN(CO-CA)CB (red) spectra.**

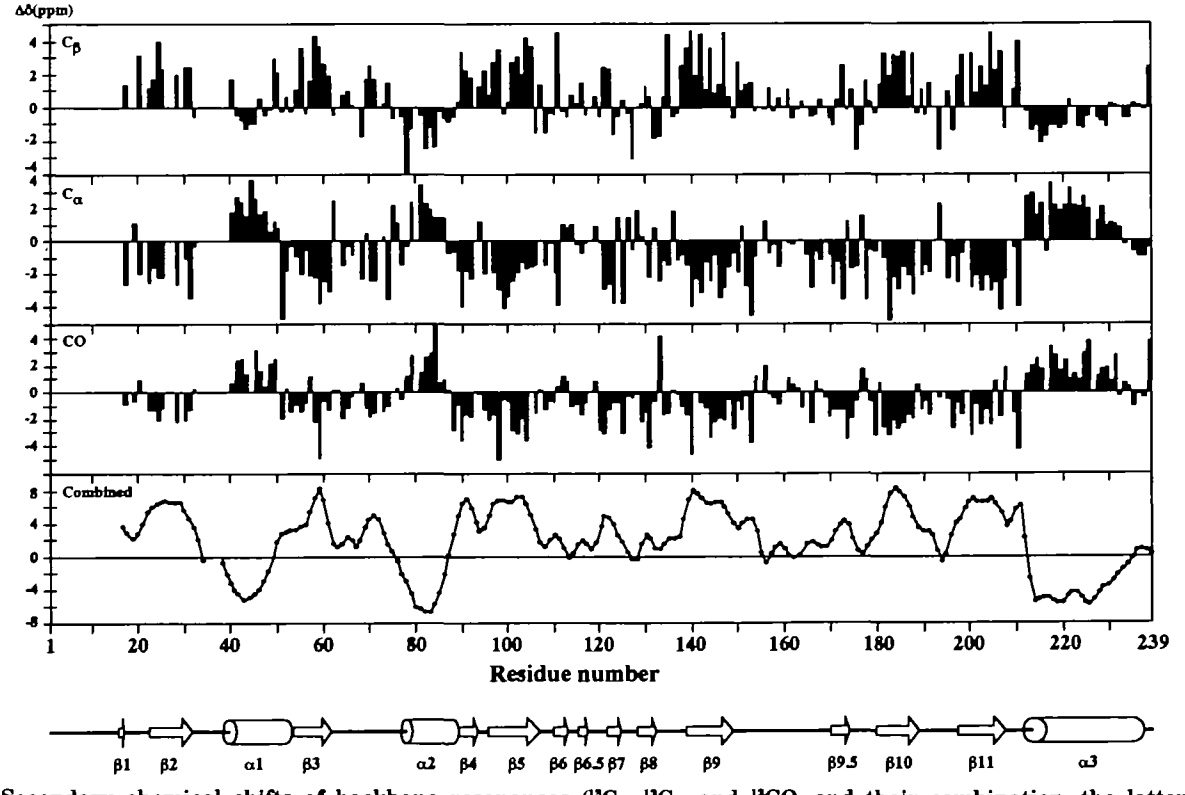

Fig. 4. **Secondary chemical shifts of backbone resonances ($^{13}C_\alpha$, $^{13}C_\beta$, and ^{13}CO, and their combination, the latter being smoothed by means of a box-car filter with a band width of 5 residues) *versus* the residue number of αNTD, and the determined secondary structure of αNTD.**

The number of peaks observed in the triple-resonance spectra was 200, which amounts to 87% of the 229 observable amide ^{1}H-^{15}N pairs. The backbone assignments were mainly made with the HNCA, HN(CO)CA, HN(CA)CB, and HN(COCA)CB spectra. In the HNCA and HN(CA)CB spectra, sequential peaks, which were weaker than intra-residue ones, were also observed for roughly 140 residues. Although the constant-time version of the HNCA spectrum exhibited much higher resolution in the $^{13}C_{\alpha}$ dimension, the peak intensities were about half those in the variable-time HNCA spectrum. Since many of the peaks were not intense enough, probably because of exchange with the solvent, we mainly used the variable-time versions of HNCA and HN(CO)CA spectra. The constant-time version spectra were used for resolving closely located cross peaks, and also for confirming the matches of the C_{α} chemical shifts of intra-residue and sequential peaks. 180 quartets [$C_{\alpha}(i)$, $C_{\alpha}(i-1)$, $C_{\beta}(i)$, $C_{\beta}(i-1)$] were observed. We also obtained ^{15}N-^{1}H HSQC spectra for amino acid-specifically ^{15}N-labeled samples (Lys, Arg, Met, Ala, Ile, Val, and Leu). These spectra, except the Met- and Leu-labeling ones, showed peaks exclusively for the labeled amino acids. Met-labeling suffered from cross-labeling slightly, but the most intense peaks

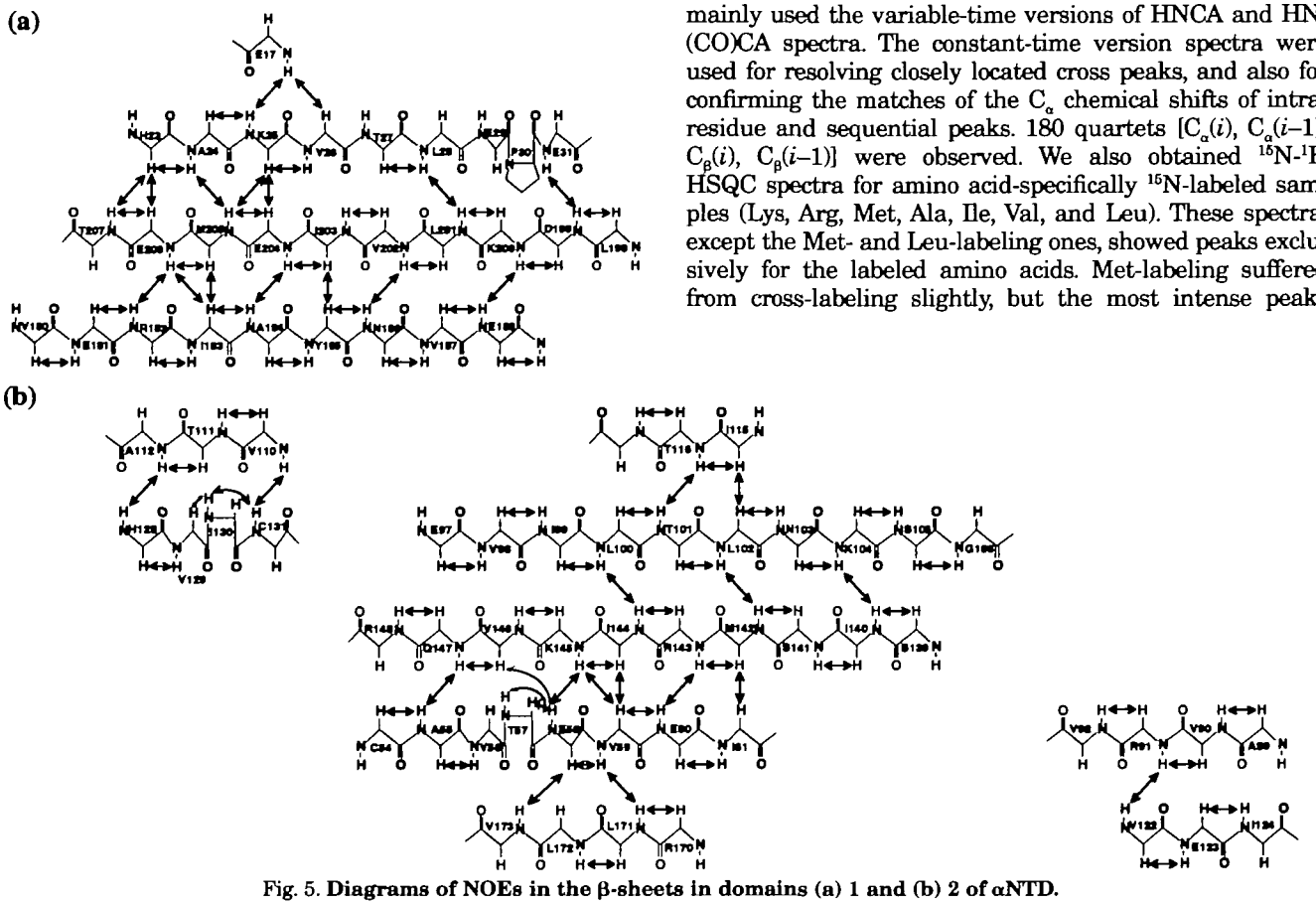

Fig. 5. **Diagrams of NOEs in the β-sheets in domains (a) 1 and (b) 2 of αNTD.**

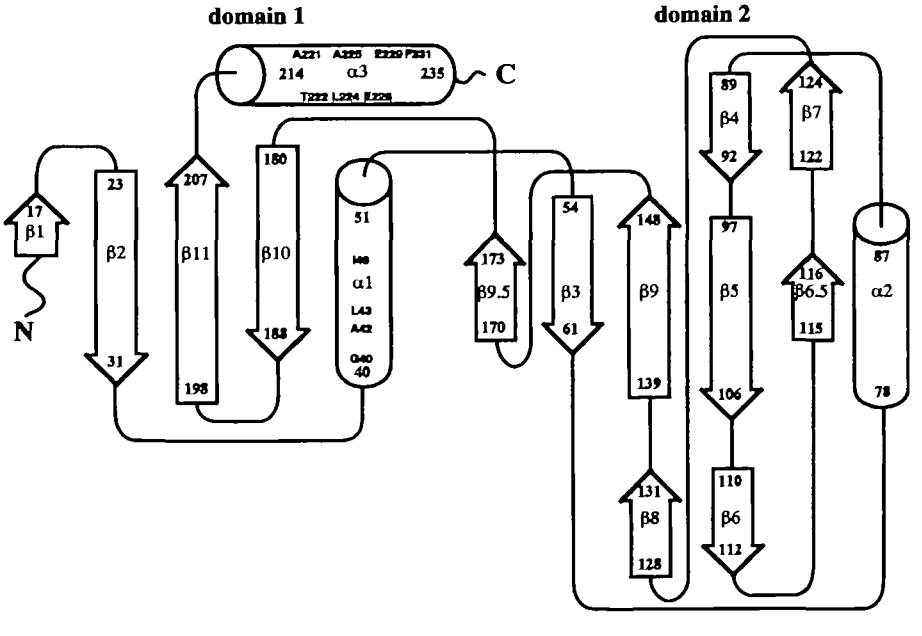

Fig. 6. **Secondary structure and chain topology of αNTD.** The residues which give the inter-molecular NOEs between αNTD and TDOC are shown in blue.

were Met ones. This domain has 25 Leu residues; 19 of the signals observed for the Leu-labeled sample exhibited typical C_β chemical shifts (~40 ppm) and could be assigned to individual residues. The remaining peaks must be due to Leu residues but were too weak to observe signals in the triple resonance spectra.

By combining this information, we safely made backbone assignments of 198 residues (87%). About 20 more residues were assigned tentatively. Figure 3, (a) and (b), shows examples of traces through the C_α and C_β resonances, respectively. The HNCO and HN(CA)CO spectra were used for confirmation of the assignments. Assignments of the H_α chemical shifts were partially (for 128 residues) made from the TOCSY spectrum of the ^{15}N-labeled sample, for which the short mixing period of 40 ms was used to obtain only H_N-H_α connectivities. The coupling constants between H_α and H_N are relatively large (~8 Hz) when the residue takes on a β-strand structure. In such cases cross peaks between H_N and H_α were observed.

Several segments of this domain remained unassigned, although almost all peaks which appeared in all HNCA, HN(CO)CA, HN(CA)CB, and HN(COCA)CB spectra could be assigned. The triple resonance experiments performed here were relatively tolerant as to the fast amide proton exchange because the water flip back scheme was used (*27*). As the number of amide groups which did not give peaks was not so high, even residues on the molecular surface seemed to give peaks in the triple-resonance spectra and thus could be assigned. This implies that the unassigned segments correspond to very flexible parts and that they are completely exposed to the solvent. The greatest deficiency in the assignment arose from the N-terminal 16 residues, which seem not to have a definite structure (E17 could be involved in β1). This is consistent with the results of the deletion experiments, which showed that deletion of the N-terminal 20 residues did not disrupt the function of αNTD (*28*). The G34–L39 region did not give any signals. This segment could be a flexible loop connecting secondary structure elements. There are some other small breaks in the assignment list at Gly and Ser residues. This is explained by their faster exchange rates. The amide exchange rate depends largely on the type of amino acid. On the contrary, Glu and Asp exhibit slower exchange rates and give intense peaks even if they are in loops.

Secondary Structure Elements and Chain Topology—The deviations of the chemical shifts of the C_α, CO, and C_β resonances from those for a random coil are shown in Fig. 4. We can easily pick up possible regions for α-helices and β-strands. Characteristic NOEs for these secondary structure elements were also collected on ^{15}N-edited 3D NOESY of the deuterated sample for the H_N-H_N connectivies and that of the protonated ^{15}N-sample for the H_N-H_α connectivities. The H_N-H_N connectivities observed in the spectrum of the deuterated sample were much more intense than those of the protonated sample. In spite of the large molecular mass, we observed intense sequential connectivities in the α-helix regions and weaker but clear inter-strand connectivities in the β-strand regions (Fig. 5). By combining this information, secondary structure elements were determined definitely (Fig. 4). However, a problem we have here is that the exchange of amide protons can decrease the intensities of the amide signals at the terminal ends of the secondary structure elements and also on the surface of the molecule. Even if the assignments of such residues were made, NOESY gave very weak signals because it required a longer mixing time. Therefore, the secondary structure elements shown in the figure might be shorter than they actually are.

Figure 6 shows the topology of αNTD, which comprises 3 α-helices and 13 β-strands. The strands form two β-sheets and two strand pairs by aligning in an anti-parallel manner for all the strands. The smaller β-sheet has 4 strands (β1, β2, β11, and β10) and the larger one has 5 strands (β6.5, β5, β9, β3, and β9.5). In strand β1, only E17 was assigned. The topology of the β-sheets and the observed NOEs between strands are given in Fig. 5. Inter-strand H_N-H_N connectivities are reliable information for determining the arrangement of the strands. Inter-strand H_N-H_α and H_α-H_α were also observed on ^{15}N-edited 3D NOESY of the ^{15}N-labeled sample and ^{13}C-edited 3D NOESY of the ^{13}C-labeled sample in D_2O, respectively. One particular point of the topology is that it includes many loops. The longest loop (G149–G169) between β9 and β9.5 includes the EEDE sequence, which is proposed to be the contact region with the catabolite activator protein (CAP) transcription activator at Class II CAP-dependent promoters (*29*). This region exhibited relatively strong resonances in the triple-resonance spectra, suggesting that this region is more flexible than the others, but that the rate of exchange of the amide protons is reduced because the negative charges of this region (EEDE) may exclude OH^- ions. This is consistent with that this region is disordered in the crystal structure (*22*).

The crystal structure was reported by Zhang and Darst (*22*) during this study. The secondary structure elements in it are essentially identical to ours except for the following points. We observed two extra β-strands (β6.5 and β9.5), which run anti-parallel on both sides of the large β-sheet in domain 2, making a larger β-sheet. In the crystal structure, the corresponding regions adopt an extended structure and run close to the edge of the β-sheet. From the NMR data revealing inter-strand NOEs for the extra strands, we concluded that they were members of the β-sheet. We also observed three β-bulge structures, which were not mentioned in their report. P31 in β2, T57 in β3, and I130 in β8 seem to be β-bulge structures suggested by the pattern of interstand NOEs and sequential H_N-H_N NOEs (Fig. 5).

The detergent molecules cover the hydrophobic patch and make the protein molecule more soluble. Intermolecular NOE experiments can show where the detergent molecules are bound. These surfaces must be important for subunit-subunit interactions. We performed a ^{15}N-edited 3D NOESY experiment with the deuterated, ^{15}N- and ^{13}C-labeled protein, and the unlabeled detergent. We looked at the NOE from the detergent (in the ^{15}N dimension the protein signal was eliminated by ^{13}C filtering and also by deuteration) to the amide ^{15}N-^1H of αNTD. Probably because a detergent molecule does not stay long enough at a single site, we could not observe any signals with the mixing time of 100 ms. However, we could observe some cross peaks with the longer mixing time of 450 ms. The peaks arose from the amides of helices-1 and 3 (Fig. 6), which are involved in the dimerization of α subunits (*22*). The binding of two α subunits is not very strong because a monomer–dimer equilibrium was observed in the sedimentation equilibrium experiment (data not shown). It is easily under-

stood that the weak interaction between α subunits is disrupted on the covering of the hydrophobic surface with detergent molecules.

In conclusion, we found that taurodeoxychlate was effective for covering the hydrophobic surface of αNTD and for preventing aggregation of the molecules, and optimized the solution conditions. We showed that monitoring of translational diffusion coefficients using the PFG NMR technique was very useful for optimizing the concentration of the detergent. The secondary structure and overall topology were determined by means of triple resonance experiments with the aid of deuteration. The molecule comprises three α-helices and two β-sheets. It seemed that αNTD molecules were dispersed as monomers in the optimal solution. The dimer interface comprising helices-1 and 3 was covered with the detergent.

We wish to thank Dr. Shirakawa for the use of the DRX800 MHz spectrometer.

REFERENCES

1. Ishihama, A. (1981) Subunit assembly of *Escherichia coli* RNA polymerase. *Adv. Biophys.* **14**, 1–35
2. Negishi, T., Fujita, N., and Ishihama, A. (1995) Structural map of the alpha subunit of *Escherichia coli* RNA polymerase: structural domains identified by proteolytic cleavage. *J. Mol. Biol.* **248**, 723–728
3. Igarashi, K., Fujita, N., and Ishihama, A. (1991) Identification of a subunit assembly domain in the alpha subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **218**, 1–6
4. Hayward, R., Igarashi, K., and Ishihama, A. (1991) Functional specialization within the α-subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **221**, 23–29
5. Kimura, M. and Ishihama, A. (1996) Subunit assembly *in vivo* of *Escherichia coli* RNA polymerase: role of the amino-terminal assembly domain of alpha subunit. *Genes Cells* **1**, 517–528
6. Igarashi, K. and Ishihama, A. (1991) Bipartite functional map of the *E. coli* RNA polymerase α subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **32**, 319–325
7. Ishihama, A. (1992) Role of the RNA polymerase α subunit in transcription activation. *Mol. Microbiol.* **6**, 3283–3288
8. Jeon, Y.H., Negishi, T., Shirakawa, M., Yamazaki, T., Fujita, N., Ishihama, A., and Kyogoku, Y. (1995) Solution structure of the activator contact domain of the RNA polymerase α subunit. *Science* **270**, 1495–1497
9. Jeon, Y.H., Yamazaki, T., Otomo, T., Ishihama, A., and Kyogoku, Y. (1997) Flexible linker in the RNA polymerase alpha subunit facilitates the independent motion of the C-terminal activator contact domain. *J. Mol. Biol.* **267**, 953–962
10. Yamazaki, T., Lee, W., Revington, M., Mattiello, D., Dahlquist, F.W., Arrowsmith, C.H., and Kay, L.E. (1994) An HNCA pulse scheme for the backbone assignment of ^{15}N, ^{13}C, ^{2}H-labeled proteins: application to a 37-kDa Trp repressor-DNA complex. *J. Am. Chem. Soc.* **116**, 6464–6465
11. Yamazaki, T., Lee, W., Arrowsmith, C.H., Muhandiram, D.R., and Kay, L.E. (1994) A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins with high sensitivity. *J. Am. Chem. Soc.* **116**, 11655–11666
12. Shan, X., Gardner, K.H., Muhandiram, D.R., Rao, N.S., Arrowsmith, C.H., and Kay, L.E. (1996) Assignment of ^{15}N, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and HN resonances in an ^{15}N, ^{13}C, ^{2}H labeled 64 kDa Trp repressor-operator complex using triple-resonance NMR spectroscopy and 2H-decoupling. *J. Am. Chem. Soc.* **118**, 6570–6579
13. Kimura, M. and Ishihama, A. (1995) Functional map of the alpha subunit of *Escherichia coli* RNA polymerase: amino acid substitution within the amino-terminal assembly domain. *J. Mol. Biol.* **254**, 342–349
14. Anglister, J., Grzesiek, S., Ren, H., Klee, C.B., and Bax, A. (1993) Isotope-edited multidimensional NMR of calcineurin B in the presence of the non-deuterated detergent CHAPS. *J. Biomol. NMR* **3**, 121–126
15. Matsuo, H., Li, H., McGuire, A.M., Fletcher, M., Gingras, A., Sonenberg, N., and Wagner, G. (1997) Structure of translation factor elF4E bound to m^{7}GDP and interaction with 4E-binding protein. *Nat. Struct. Biol.* **4**, 717–724
16. Stejskal, E.O. and Tanner, J.E. (1965) Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys.* **42**, 288–292
17. Tanner, J.E. (1970) Use of the stimulated echo in NMR diffusion studies. *J. Chem. Phys.* **52**, 2523–2526
18. Gibbs, S.J. and Johnson, C.S. (1991) A PFG NMR experiment for accurate diffusion and flow studies in the presence of eddy currents. *J. Magn. Reson.* **93**, 395–402
19. Dingley, A.J., Mackay, J.P., Morris, M.B., Chapman, B.E., Kuchel, P.W., Hambly, B.D., and King, G.F. (1995) Measuring protein self-association using pulsed-field-gradient NMR spectroscopy: application to myosin light chain 2. *J. Biomol. NMR* **6**, 321–328
20. Dingley, A.J., Mackay, J.P., Shaw, G.L., Hambly, B.D., and King, G.F. (1995) Measuring macromolecular diffusion using heteronuclear multiple-quantum pulsed-field-gradient NMR. *J. Biomol. NMR* **10**, 1–8
21. Grzesiek, S., Wingfield, P.T., Stahl, S., Kaufman, J.D., and Bax, A. (1995) Four-dimensional ^{15}N-separated NOESY of slowly tumbling perdeuterated ^{15}N-enriched proteins. Application to HIV-1 Nef. *J. Am. Chem. Soc.* **117**, 9594–9595
22. Zhang, G. and Darst, S.A. (1998) Structure of the *Escherichia coli* RNA polymerase α subunit amino-terminal domain. *Science* **281**, 262–266
23. Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E., and Dahlquist, F.W. (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol.* **177**, 44–73
24. Whalen, W.A., Wang, M., and Berg, C.M. (1985) β-Chloro-L-alanine inhibition of the *Escherichia coli* alanine-valine transaminase. *J. Bacteriol.* **164**, 1350–1352
25. Kay, L.E., Ikura, M. Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Mag. Reson.* **89**, 496–514
26. Yamazaki, T., Tochio, H., Furui, J., Aimoto, S., and Kyogoku, Y. (1997) Assignment of backbone resonance for larger proteins using the ^{13}C-^{1}H coherence of a $^{1}H_{\alpha}$-, ^{2}H-, ^{13}C-, and ^{15}N-labeled sample. *J. Am. Chem. Soc.* **119**, 872–880
27. Grzesiek, S. and Bax, A. (1993) The importance of not saturating H_2O in protein NMR. Application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* **115**, 12593–12594
28. Kimura, M., Fujita, N., and Ishihama, A. (1994) Functional map of the alpha subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **242**, 107–115
29. Niu, W., Kim, Y., Tau, G., Heyduk, T., and Ebright, R.H. (1996) Transcription activation at Class II CAP-dependent promotors: two interactions between CAP and RNA polymerase. *Cell* **87**, 1123–1134