Spectroscopic and Thermodynamic Characterization of the Transcription Antitermination Factor NusE and Its Interaction with NusB from *Mycobacterium tuberculosis*†

B. Gopal,‡,§ K. G. Papavinasasundaram,§ Guy Dodson,‡ M. Jo Colston,§ Sarah A. Major,|| and Andrew N. Lane*,||

Division of Protein Structure, Division of Mycobacterial Research, and Division of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.

Received August 3, 2000; Revised Manuscript Received October 18, 2000

ABSTRACT: N-utilizing proteins (Nus) form a complex involved in the regulation of rRNA biosynthesis in enteric bacteria by modulating the efficiency of transcriptional termination [Nodwell, J. R., and Greenblatt, J. (1993) *Cell 72*, 261−268]. The protein NusE (identical to the protein S10 of the small ribosomal subunit) from the pathogenic mycobacterium *M. tuberculosis* has been cloned and overexpressed in *Escherichia coli*. The pure protein has been characterized by circular dichroism, ultracentrifugation, NMR, and binding to NusB. The near-ultraviolet circular dichroism spectrum of this protein suggests that it has a moderate (ca. 12−16%) α-helical content at 30 °C. The protein undergoes cold denaturation, with a temperature of maximum stability near 40 °C, implying a substantial heat capacity difference between the folded and unfolded states. The sedimentation equilibrium and velocity data indicate that the protein is monomeric and expanded in solution. NMR spectroscopy shows that there is no significant tertiary structure, and confirms the low secondary structure content at low temperatures. Furthermore, there was evidence for more structure at 30 °C than at 10 °C. Well-defined shifts in peaks in the HSQC spectrum of ^{15}N labeled NusE/NusB when the unlabeled counterpart was added at approximately stoichiometric concentrations showed the formation of a NusE−NusB complex in the absence of RNA. The far-UV CD and ultracentrifuge experiments, however, indicated relatively weak binding. Isothermal titration calorimetry showed the binding was weak and endothermic at 15 °C, with a total ΔH of ≥10 kcal/mol. This weak binding is consistent with a small interaction interface and lack of large conformational rearrangements in the predominantly unfolded NusE protein. The conformational flexibility of NusE may be important for its roles in both the ribosome and antitermination complexes.

Transcription of ribosomal RNA in *Escherichia coli* is regulated at a number of levels that tie the production of rRNA and ribosomes to the growth rate of the cells. Rho-dependent termination signals in the leader region of the rRNA transcript normally limit the number of complete transcripts made. However, for rapid growth to occur, specific mechanisms are required to permit read-through; one of these is antitermination, which uses signals embedded in the leader/terminator region of the rRNA to allow RNA polymerase to proceed through termination sites (*1*). This read-through requires a complex of proteins that interact both with RNA polymerase and RNA sequences (termed box A and box B) in the leader region. In phage λ, antitermination is initiated by the λ-encoded N protein, which recruits a number of host proteins called Nus[1] factors (N-utilizing substances) (*2*). Several of the host proteins essential for effective transcription antitermination have been identified, including NusA, NusB, NusE, and NusG (*3*). An analogous ribonucleoprotein complex has been identified in antitermination in the transcription of rRNA in *E. coli*. In particular, NusB interacts with a 12 nucleotide segment called box A (*4*). This association has been reported to be promoted by the formation of a complex of NusB and NusE (*5*). A third protein, NusG, has also been implicated in this association (*6*) but the molecular details of its interactions with the other components in the complex are not clear. Although data are lacking that directly distinguish whether NusE participates in transcription antitermination as a free protein or as a component of the ribosome, the observation that NusB affects the rate of translation and participates in protein secretion is consistent with the latter possibility (*1*). To understand the ternary interactions between NusB, NusE, and box A RNA, it is necessary to consider the possible binary interactions independently of the ternary interaction.

† Supported by the Medical Research Council of the U.K. and in part by the European Commission Science Research and Development Program (Contract ERBIC 18CT970253).

* To whom correspondence should be addressed. E-mail: alane@nimr.mrc.ac.uk. Phone: 44 0208 959 3666. Fax: 44 0208 906 4477.

‡ Division of Protein Structure.

§ Division of Mycobacterial Research.

|| Division of Molecular Structure.

[1] Abbreviations: CD, circular dichroism; DTT, dithiothreitol; TFE, trifluoroethanol; ITC, isothermal titration calorimetry; HSQC, heteronuclear single quantum coherence; NOESY, nuclear Overhauser effect spectroscopy; Nus, N-utilising substance; PCR, polymerase chain reaction; TOCSY, total correlation spectroscopy.

10.1021/bi0018279 CCC: $20.00
Published on Web 12/30/2000

Box A sequences are conserved among enteric bacteria. An analogous system has not been described in detail in other bacteria, but equivalents of the box A and box B sequences have been identified in the leader region of rRNA genes in mycobacteria (*7*), suggesting that a comparable mechanism operates in these species. While most species of eubacteria have multiple rRNA operons, *Mycobacterium tuberculosis*, the causative agent of human tuberculosis, possesses only a single rRNA operon, suggesting that efficient transcription of the operon is likely to be essential for survival (*7*, *8*). *M. tuberculosis* grows slowly, with a doubling time of approximately 15 h under ideal conditions, compared to ca. 20 min for most other eubacteria. The regulation of rRNA transcription and particularly the antitermination mechanism is likely to play an important role in regulating the growth rate of this organism.

The NusE from *M. tuberculosis*, a component of the transcription antitermination complex, is identical in sequence to the ribosomal protein S10. The sequence comparison presented in Figure 1 reveals a high percentage of sequence identity among the various species, as anticipated for a ribosomal protein. Structure prediction algorithms indicate a moderate amount of helix and sheet (Figure 1).

To evaluate the role of NusB·NusE-mediated antitermination in *M. tuberculosis*, we have cloned and expressed both proteins in *E. coli* and purified them to homogeneity. We have characterized their solution properties and ability to form a heteromeric complex in solution. The NusB of *E. coli* has been extensively studied. It is an all-helical protein (*9*, *10*) with a novel fold (*11*, *12*). Essentially the same fold has been found for the NusB from *M. tuberculosis*, except that it forms dimers at millimolar concentrations, unlike the monomeric protein of *E. coli* (*12*, *13*). In contrast, no structural information is available about any NusE protein free in solution. In this manuscript, we describe a spectroscopic characterization of the *M. tuberculosis* NusE and a point variant, NusE C50S, and its interaction with NusB. We show that the NusE protein is only partly folded and is thereby susceptible to proteolysis. NMR spectra recorded on ^{15}N labeled proteins indicate that the binding event involves a small number of specific changes in the structure, as is evident from the shifts in the ^{15}N edited peaks before and after complex formation.

MATERIALS AND METHODS

Cloning and Expression of NusB and NusE from M. tuberculosis. The gene corresponding to the NusE from *M. tuberculosis* was amplified by PCR with pfu polymerase using the following primers: *Nde*1-NusE, TAGAAGA-*CATATG*AGCGTGGCGGGACAGAAGAT; *Bam*H1−NusE: CTTA*GGATCC*TCGTCGAATACCTGCGTCATACCC.

The two primers allowed for the insert containing the gene for NusE to have sites for the restriction enzyme *Nde*1 and *Bam*H1 (in italics). The PCR product was ligated into the PET-15b expression vector between the *Nde*1 and the *Bam*H1 sites. After a double digestion with the two enzymes to verify the size of the insert released, the plasmid was transformed into the *E. coli* strain BL21(DE3)PlysS (Novagen Inc.). The cells were induced with 1 mM isopropylthiogalactoside when the culture density reached $A_{600} = 0.8$ and were grown for a further 6 h before they were harvested. The protein forms inclusion bodies which were solubilized with urea. The His-tagged protein was purified by affinity chromatography on a cobalt resin column (Talon, Clontech Inc.). The protein was bound at pH 8.0 and eluted at pH 6.0 using buffer compositions as suggested by the manufacturer. The refolding/solubilization was achieved by slow dialysis in which the urea was removed in steps of 0.5 M, from 8.0 to 0 M urea in 50 mM Tris buffer, pH 7.5, 200 mM NaCl. The slow rate of dialysis improved the yields of soluble protein obtained. Attempts to improve the yields or the nature of the soluble product by changes in buffer composition such as the Hampton FoldIt screen (Hampton Research Corp., CA) and pH were not successful.

The protein forms a homodimer involving the unique cysteine residue located at position 50. The numbering is based on the actual NusE sequence, ignoring the histidine tag which includes a 20 residue linker containing a unique thrombin cleavage site. DTT at a concentration of 3 mM and 2 mM EDTA were added in the final stages of refolding and in the buffer used for storing the protein. The protein was further purified using size-exclusion chromatography on a Superdex S-75 column (Pharmacia Inc). The purified protein migrated as bands corresponding to a monomeric and dimeric species on a SDS−PAGE (Figure 2). To circumvent the problem of dimerization and possible inhibition of heterodimer formation with NusB, the cysteine at position 50 was mutated to a serine using the protocol for site-directed mutagenesis specified by Strategene Inc. with the following primers (mutations in bold): S10-mut-upper, GAGAA-GAACGTGTAT**A**GCGTCATCCGCTCACC; S10-mut-lower, GGTGAGCGGATGACGC**T**ATACACGTTCTTCTC.

The sequences were verified using an ABI Prism 377 DNA sequencer with the ABI Rhodamine dye terminator cycle kit (PE Applied Biosystems). The resulting C50S variant protein was purified using the same protocol as described for the native protein.

Uniformly ^{15}N-labeled samples for the NMR experiments were prepared by growing the cells in M9 minimal media with $(^{15}NH_4)_2SO_4$ as the nitrogen source. The cloning, expression, and purification of *M. tuberculosis* NusB procedures were as previously described (*13*). ^{15}N-labeled NusB was prepared as described above for NusE. Owing to the low structure content (see below), all experiments were carried out with freshly purified protein.

Circular Dichroism Measurements. CD spectra were recorded on a Jasco J715 spectropolarimeter, with a scan speed of 100 nm/min, using a cell of path-length 0.1 mm. Far-UV CD spectra of NusE and NusB were recorded at 12 μM in 5 mM sodium phosphate, 10 mM sodium chloride buffer, pH 8.0. Spectra were also recorded at different pH values, using an acetate buffer at pH 5.5. There were no significant effects of pH on the spectra in the range $5.5 < pH < 8.0$. Each spectrum was an average of 10 scans. Spectra collected on NusE with various mole fractions of TFE were collected after incubating the protein at the required concentration of TFE for 15 min. The protein concentration was determined using a calculated absorbance coefficient of 3840 M^{-1} cm^{-1} at 280 nm.

The fraction of α-helix was obtained from (*14*, *15*):

$$\%\alpha = 100(\Delta\epsilon_{222} - \Delta\epsilon_P)/(\Delta\epsilon_H - \Delta\epsilon_P) \qquad (1)$$

where $\Delta\epsilon_{222}$ is the observed ellipticity at 222 nm, $\Delta\epsilon_H$ and

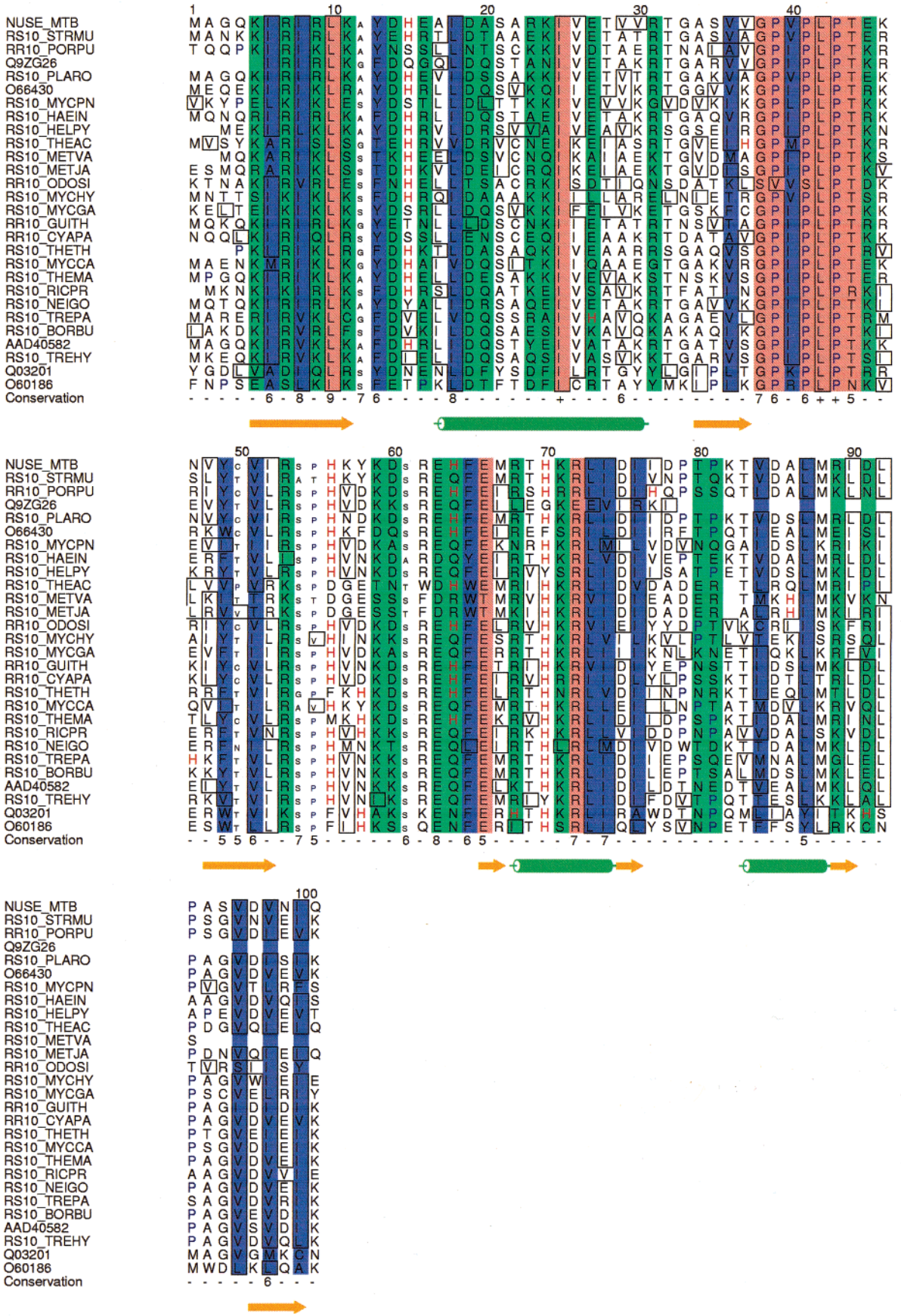

FIGURE 1: Sequence comparison of *M. tuberculosis* NusE with the homologous 30 S Ribosomal protein S10 from other bacteria. The cysteine residue mutated in *M. tuberculosis* NusE is a threonine in the *E. coli* homologue, emphasising that it is not a conserved residue. The overlap shows 61.9% identity in 98 residues. The S10 protein has 101 residues.

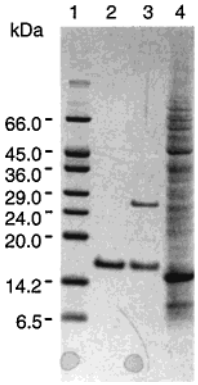

FIGURE 2: SDS−PAGE analysis of *M. tuberculosis* NusE. Proteins were run on a 15% SDS−tricine polyacrylamide minigel lane 1, low-range molecular weight markers (6.5−66 kDa; Sigma); lane 2, C50S variant NusE; lane 3, native NusE with the heterogeneity involving the intermolecular disulfide bridged protein; lane 4, crude cell lysate of induced NusE.

$\Delta\epsilon_R$ are values of $\Delta\epsilon_{222}$ for the fully helical and fully unfolded forms of each sample, and are given by

$$\Delta\epsilon_H = [-44\,000(1 - k/n) + 250T]/3300 \quad (2)$$

$$\Delta\epsilon_R = (2200 - 52.5T)/3300 \quad (3)$$

T is in the temperature (°C), n is the number of residues in the chain, and $k = 2.5$ is a constant that corrects for non-hydrogen-bonded carbonyls that do not contribute to $\Delta\epsilon_H$. The thermal unfolding was monitored at 222 nm and at 230 nm where the coil contribution to the CD is essentially zero, and the helical signal is 0.62 times that at 222 nm (*16*).

For a two state unfolding transition, the equilibrium constant for unfolding, K_u, is

$$K_u = (1 - X)/X \quad (4)$$

and

$$\ln[K_u(T)] = \ln[K_u^0] + (\Delta Cp/R)\ln T/(T_0 + (\Delta H_0 - \Delta Cp T_0)/R)(1/T_0 - 1/T) \quad (5)$$

where K_u^0 is the unfolding equilibrium constant at the reference temperature T_0, X is the mole fraction of the folded state, ΔCp is the difference in heat capacity, and ΔH_0 is the enthalpy change at the reference temperature. For a two-state transition, the mean residue ellipticity is given by

$$\Delta\epsilon = (\Delta\epsilon_f - \Delta\epsilon_c)/[1 + K_u(T)] + \Delta\epsilon_c \quad (6)$$

where $\Delta\epsilon_f$ is the mean residue absorption coefficient for the folded form and $\Delta\epsilon_c$ is that of the coil state. Equations 5 and 6 were used to fit the observed ellipticity as a function of temperature, using nonlinear regression on the parameters $\Delta\epsilon_f$ $\Delta\epsilon_c$, Ku(T_0), $\Delta H(T_0)$, and ΔCp. The data were also analyzed allowing a linear variation of the CD as a function of temperature, requiring two additional slope parameters.

Analytical Ultracentrifugation. The molecular weight and sedimentation coefficient of purified NusB and NusE were determined by centrifugation using a Beckman XLA ultracentrifuge equipped with absorbance optics. Protein solutions at $A(280) = 0.5-1.0$ in Tris buffer, pH 8, were run at 60 000 rpm (sedimentation) and 20−30 krpm (equilibrium) and analyzed using the programs Svedberg (*17*) and Origin (Microcal). The partial specific volumes were calculated from the amino acid compositions (*18*), and were 0.744 mL g^{-1} for NusB and 0.736 g mL^{-1} for NusE. Sedimentation velocity runs were carried out at 5 and 25 °C. The sedimentation coefficients were converted to frictional coefficients making used of the known density of water at different temperatures (*18*). The derived frictional coefficients, f, were then converted to values at 20 °C using the known temperature dependence of the viscosity of water (*19*).

Formation of the NusB:NusE complex was assessed by equilibrium centrifugation of pure NusB, pure NusE and a 1:1 mixture in the same run. The absorption coefficient of NusB at 293 nm is 10 times more than that of NusE, so the absorbance profile largely reflects that of NusB.

NMR Spectroscopic Measurements. NMR spectra were recorded at 11.75 and 14.1 T at 283, 298, or 303 K on Varian UnityPlus or Unity spectrometers, respectively. Heteronuclear ^{15}N-^{1}H NMR spectra were recorded using standard methods (*20, 21*). NOESY and clean-TOCSY (*22*) spectra were recorded in H_2O using Watergate (*23*) to suppress the water signal. Isotropic mixing was achieved using MLEV-17 (*24*) with delays approximately equal to the 90° pulse width to suppress ROESY artifacts.

Spectra were recorded either at pH 5.5 in 0.05 M sodium acetate buffer or pH 8 in 0.15 M NaCl, 0.05 M sodium phosphate. The low pH was chosen to reduce amide proton exchange, and the higher pH to increase the tendency of the NusE and NusB to associate, as their pI values are 10.06 and 7.27, respectively. For analyzing the interaction of ^{15}N NusB with ^{14}N NusE or ^{14}N NusB with ^{15}N NusE, the two proteins were dialyzed to equilibrium against the same buffer. All spectra were recorded on two or three independently purified samples.

RESULTS

Characterization of NusE. NusE from *M. tuberculosis* was cloned, overexpressed, and purified to homogeneity (Figure 2). To overcome the problem of intermolecular dimerization owing to disulfide formation between the cysteine residue at position 50 (Figure 2, lane 3), a monomeric C50S variant (Figure 1) was engineered and was used in the present studies. A similar protocol was also tried with the *E. coli* protein, with less success; the *E. coli* homologue seems to be rather insoluble in water. The slow resolubilization from urea appears to be essential to obtain a preparation at sufficient concentrations for NMR analysis.

Circular Dichroism. Figure 3a shows far UV CD spectra of NusE at 5 and 30 °C. The positive intensity at wavelengths shorter than 200 nm is consistent with a folded protein (*25*). The negative bands at 208 and 222 nm indicate the presence of α helix and/or type I β-turn structures. According to the intensity (eq 1) this amounts to about 14% or 17 residues at 30 °C. Both the wild-type and the C50S variant exhibited the same features in the far UV CD region. In addition, the same CD spectra and their temperature responses were obtained with thrombin-cleaved protein, which removes the N-terminal hexahis-linker sequence, and implies that these residues do not materially contribute to the observed secondary structure.

To estimate the thermodynamic stability of the structured form of the protein we have monitored the far UV CD as a

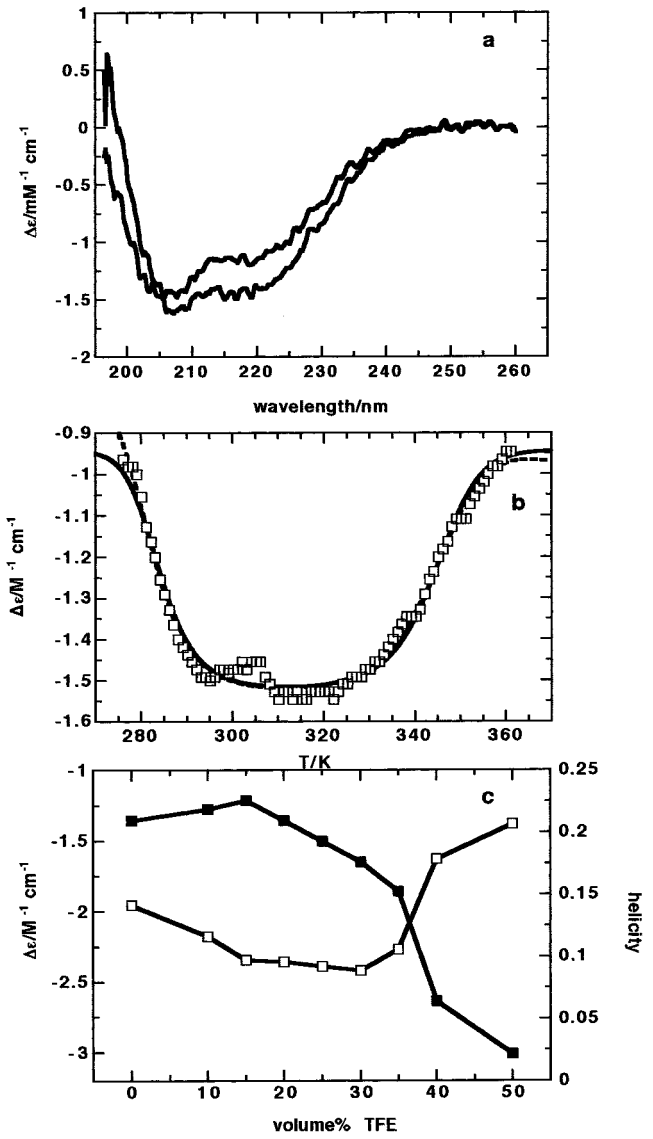

FIGURE 3: Far-UV CD spectra of NusE from *M. tuberculosis*, Spectra were recorded as described in the Materials and Methods. (a) CD spectra at 5 (●) and 30 °C (○). (b) variation of the CD at 222 nm with temperature, (□) $\Delta\epsilon(222)$, (−) fit to eq 5 and (6) assuming no temperature dependence of $\Delta\epsilon_f$ and $\Delta\epsilon_c$. (- - -) Fit including linear dependence of $\Delta\epsilon_f$ and $\Delta\epsilon_c$ on temperature: $\Delta\epsilon(T) = \Delta\epsilon(313) + \partial\Delta\epsilon/\partial T$ (T-313), (c) dependence of the CD at 222 nm at 30 °C on volume fraction of TFE. (■) $\Delta\epsilon$, (□) helicity calculated according to eqs (1−3) using the value for the coil state (*15*). A major transition is observed at ∼35% TFE.

function of temperature and urea concentration as described in the Materials and Methods. The CD of the protein appeared to reach the unfolded baseline (as judged by the ellipticity at 222 nm) by ∼2.0 M urea. This implies that the limited structure present at 30 °C is marginally stable.

Figure 3b shows the ellipticity at 222 nm as a function of temperature. The ellipticity decreases substantially on lowering the temperature from about 30 to 2 °C (ca. 45%), indicating the loss of helical secondary structure; the spectra indicate ∼14% helix at 30 °C and ∼8% at 2 °C. The ellipticity remains more or less constant between 35 and 50 °C, and then decreases at higher temperatures. These temperature-dependent changes were fully reversible. The clear maximum in helicity at around 40 °C (minimum for $\Delta\epsilon$) indicates a large ΔCp of unfolding (*26*). Estimates of ΔCp and the enthalpy change were obtained from analysis of the temperature dependence of K_u (cf. eq 6). The data were fitted in three ways as follows. First, the mean residue CD absorption coefficient for the coil ($\Delta\epsilon_c$) and folded ($\Delta\epsilon_f$) states were treated as temperature-independent constants, which gave values of $\Delta H(313) = 1.1$ kcal mol^{-1}, $K_u(313) = 0.04$, and $\Delta Cp = 1.4$ kcal mol^{-1} K^{-1}. A slightly better fit was obtained assuming a linear temperature dependence of either or both $\Delta\epsilon_c$ and $\Delta\epsilon_f$ (*14*, *15*) (Figure 3b), which requires one or two additional parameters. The value of $\Delta H(313)$ is not well determined, though it is small (in the range 0−2 kcal mol^{-1}), whereas ΔCp was 1.2 ± 0.2 kcal mol^{-1} K^{-1}. Furthermore, the values of $\Delta\epsilon_f$ and $\Delta\epsilon_c$ at 313 K were very similar regardless of whether the temperature dependence was included or not.

The apparent helicity of NusE was calculated as a function of temperature both using eq 3 and assuming a temperature-independent CD spectrum for both coil and helical forms. The main difference between these calculations is in the absolute magnitude of the helicity, in part because of the difference in the CD used for the coil state. However, the shapes of the curves are similar, and both show a temperature of maximum stability near 310 K, with considerable loss of structure at 5 or 80 °C. From the fitted values of $\Delta\epsilon_c$ and $\Delta\epsilon_f$, the maximum helicity was calculated using eq 1, which we estimate as 15%, or about 19 residues. Hence, we conclude that the structured state present near 40 °C contains approximately 19 residues that are mainly helical, and which is a marginally stable state. This is consistent with the low concentration of urea needed to unfold the protein completely.

The CD spectrum at 30 °C suggests an equilibrium between unstructured and helical conformations. The helical signal at 222 nm increased in intensity on adding TFE. Figure 3C shows a 2.5-fold increase of the CD intensity between 0 and 50% TFE, and half of the increase occurred at about 35% TFE. The helical content increased from 12% (15 residues) to 22% (28 residues). This indicates a significant propensity of this protein to form helix, and under native conditions, the protein is only partly folded.

Ultracentrifugation of NusE. The C50S NusE protein is a monomer in solution as is the wild-type protein in highly reducing conditions which prevents the formation of intermolecular disulfide bridges. The molecular weight of NusE C50S was found to be 16 ± 1 kDa by equilibrium ultracentrifugation, at a loading concentration of 90 μM. The actual molecular weight of the construct, including the linker and hexahis tag, is 14.2 kDa. We conclude that under these conditions the NusE protein is monomeric. The translational frictional coefficient, *f*, was determined from the sedimentation velocity. At 25 °C, $f = 6.4 \times 10^{-8}$ g s^{-1}. The frictional ratio $f/f_0 = 2.0$ at 25 °C, where f_0 was calculated for a sphere of molecular weight 14.2 kDa with a monolayer of water. As the value of f/f_0 is much larger than unity, either the protein is very much more strongly hydrated than is typical and/or it is rather elongated or expanded.

NMR Spectroscopy of NusE and NusB. To characterize the conformational state of NusE further, we have recorded NMR spectra of the both NusE in the absence and presence of NusB. Figure 4a shows 1D NMR spectra of the variant NusE at 10 and 30 °C. Clearly, the absence of high-field shifted methyl groups, the random-coil shifts of the bulk

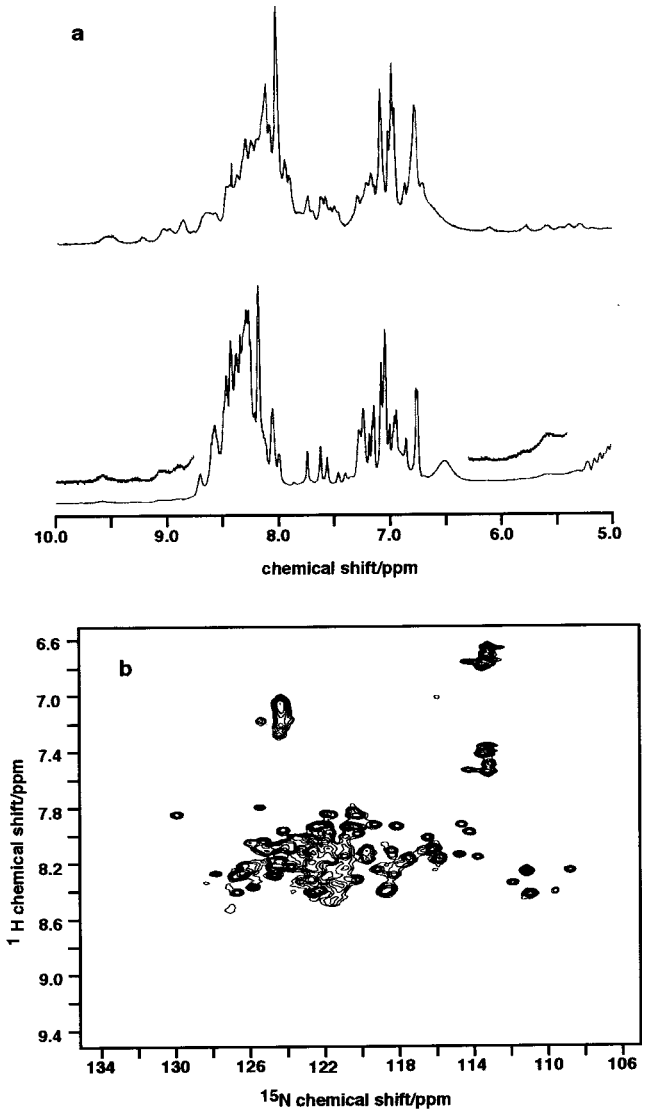

FIGURE 4: NMR spectroscopy of NusE from *M. tuberculosis*, Spectra were recorded as described in the Materials and Methods. The 2D spectra were processed with zero-filling once in t_2 and twice in t_1, and an unshifted Gaussian apodization function in both dimensions. (a) 1D NMR spectra at 10 and 30 °C recorded at 600 MHz at a concentration of 0.6 mM with acquisition times of 1.5 s and recycle times of 3 s. The free induction decays were processed with 1 Hz line broadening. Upper spectrum, 30 °C; lower spectrum, 10 °C (expansion is 4×), (b) ^{15}N HSQC spectrum at 10 °C recorded at 11.75 T with acquisition times of 0.147 s in t_2 and 0.06 s in t_1.

methyl groups at 1 ppm, the low dispersion of the backbone amide NH signals and absence of downfield shifted αCH resonances points to a largely unstructured protein at 10 °C (*27*). There is no evidence of any tertiary structure or the presence of significant amounts of antiparallel β-sheets. However, on increasing the temperature to 30 °C, a number of new resonances appeared in the low field region between 8.7 and 9.6 ppm, which were not present in the ^{15}N-^{1}H HSQC spectrum at 10 °C (Figure 4b). In addition, about seven resonances between 5.2 and 6.1 ppm appear at 30 °C, pointing to the formation of some structure, such as α-helix (see above) and antiparallel β-sheet. The spectra also appear somewhat different in the high field end; at 30 °C the spectra are broader, and contain relatively more intensity between 0.6 and 0.8 ppm than at 10 °C (not shown), indicating the presence of at least some folded structures.

The low-temperature HSQC spectrum (Figure 4b) shows low spectral dispersion in both the ^{1}H and ^{15}N dimensions, typical of largely unfolded proteins (*27*, *28*). Some assignments to residue type were obtained from TOCSY spectra of the C50S variant recorded in H_2O at 10 °C, e.g., Gly, Ile, Ala, and Asx residues. In the fingerprint region, the αCH show essentially random coil CαH shifts (*29*), and the NOESY spectrum shows a mixture of intraresidue and presumably sequential NH-CαH NOEs (data not shown). Similarly, the ^{15}N-^{1}H HSQC spectra (Figure 4b) shows a clustering of cross-peaks typical of dynamically unstructured proteins (*28*, *30*). Only a small number of weak NH−NH cross-peaks were detected in the NOESY spectrum. The chemical shift dispersion of the HN, NH, and CαH, absence of obviously shifted resonances from tertiary structure indicates a largely unfolded structure at 10 °C (*27*, *28*, *30*). Hence at 10 °C, there is only a small fraction of the protein in a helical conformation, consistent with the CD data (see above). The NMR spectra indicate that there is unlikely to be a single continuous α-helix present at low temperature.

At 30 °C, where the CD spectra show significantly more structure (see above), the NMR spectrum is more dispersed than at 10 °C, with the appearance of extra nonrandom coil resonances. Furthermore, the HSQC spectrum at 30 °C shows increased spectral dispersion (see below), consistent with the additional peaks seen in the 1D spectrum (Figure 4a), indicating that more backbone NH are participating in hydrogen-bonded structures. Again, this is in general agreement with the CD data, which indicate the presence of more secondary structure at 30 °C than at 5 °C. However, the NMR and CD spectra at 30 °C show only a relatively small fraction of folded structures. Unfortunately, the increased chemical exchange of NH with solvent at 30 °C also made it impossible to assign particular residues by ^{15}N-edited NOESY and TOCSY spectroscopy. Taken together, the CD data, hydrodynamic results and the NMR spectroscopy indicate a partially structured protein at room temperature, with a maximal stability at ca. 40 °C.

Interaction of NusE with NusB. If NusB and NusE bind as a heterodimer to box A RNA (*4*), then there must be some tendency for the two proteins to form a complex in the absence of RNA. Such a complex has been detected for the *E. coli* enzymes (*5*). We have therefore used a number of methods to detect and characterize the interaction in the absence of RNA.

NMR Spectroscopy. The ^{15}N HSQC spectrum of NusE at low temperature (Figure 4b) indicates a largely unstructured protein. Changes in shifts of NusE on addition of unlabeled NusB should therefore reflect either a direct interaction or an induction of folding. Figure 5a shows the ^{15}N HSQC spectrum of NusE in the absence and presence of an excess of unlabeled NusB. Changes in line widths are expected as NusB is a dimer of 34 kDa, whereas NusE is a largely unfolded monomer of 14.2 kDa. There are clearly differential changes in intensity of a number of peaks; at least 15 cross-peaks present in the free protein either shift or disappear, with the appearance of at least 12 new or shifted peaks in the mixture. This observation indicates that a complex is being formed under these conditions.

Figure 5b shows the inverse experiment. The HSQC spectrum of ^{15}N-labeled NusB in the absence of NusE at pH 8 is comparatively well dispersed, though probably

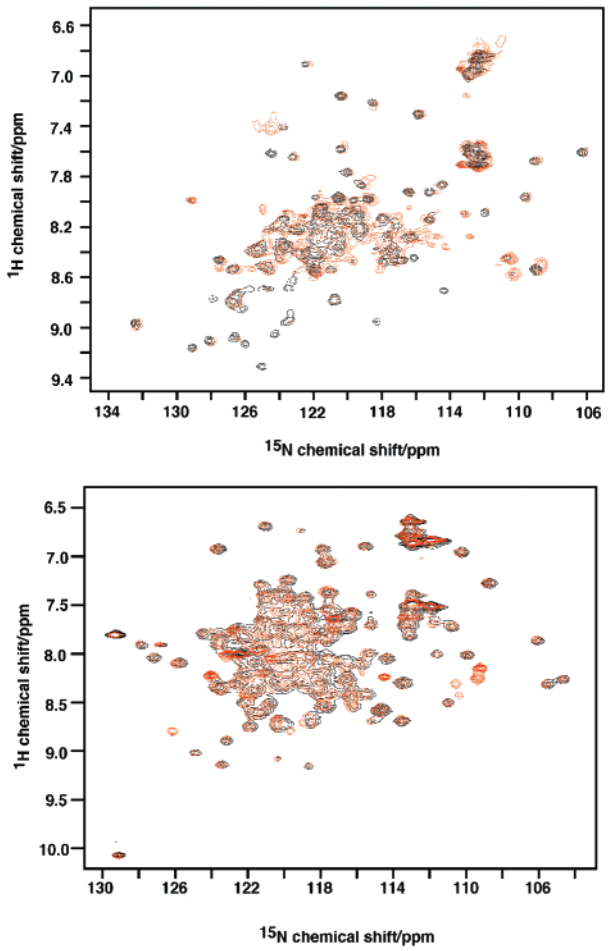

FIGURE 5: Interaction of *M. tuberculosis* NusE with NusB monitored by ^{15}N-HSQC, ^{15}N-^{1}H HSQC spectra were recorded at 11.75 or 14.1 T as described in the methods, with acquisition times of 0.15 s in t_2 and 0.06 s in t_1. The data tables were zero-filled once in t_2 and twice in t_1, and apodized using an unshifted Gaussian function. (a) 500 μM ^{15}N NusE (black); + 800 μM ^{14}N NusB (red), pH 5.5, 30 °C 11. 75 T. (b) 400 μM ^{15}N NusB (black); + 1.5 M excess of ^{14}N NusE (red). pH 8, 25 °C, 14.1 T.

numerous amide peaks are absent owing to rapid exchange with solvent at pH 8 and 298 K. Nevertheless, on addition of unlabeled NusE, at least six new peaks appear that are not present in the free NusB, and other changes in intensity of peaks compared with the free protein, indicating that there are significant differences between the free and complexed forms. As NusB is folded in solution, these spectra changes reflect direct interactions between NusB and NusE. Again, the relatively small number of shifted/new peaks indicates a comparatively small binding surface, which is consistent with the weak interaction (see above), but also argues against purely nonspecific interactions with the entire surface of the protein. These experiments have been carried out at both pH 5.5 and pH 8, and also at low and high temperature. Although there are quantitative differences between these experiments, the main results are similar, suggesting that the changes are not due to nonspecific electrostatic interaction at neutral pH. Furthermore, most of the resonances of the NusB remain unshifted, suggesting that the tertiary structure is not greatly perturbed by the interaction with NusE.

Thus, complexes of NusB:NusE do form at high protein concentration (800 μM NusB and 500 μM NusE). We also detected this complex at lower concentrations of NusE and NusB (200 and 500 μM), but the data available do not permit an estimate of the stoichiometry nor the dissociation constant of this binding event. However, the spectra suggest a comparatively small interaction surface, and no significant change in the conformation of the NusB protein.

Analytical Ultracentrifugation. NusB, which contains two tryptophan residues, absorbs relatively strongly at 293 nm, whereas NusE, which contains only tyrosine and phenylalanine, absorbs very weakly. Hence, the absorbance profile in an equilibrium run reflects primarily the molecular weight of the Nus B. We have compared the apparent molecular masses of NusB, NusE, and a mixture of the two in the analytical ultracentrifuge in the same 6-sector cell run at the same time. The NusB molecular weight was 28 300 ± 1800 Da, verifying that the protein is indeed a dimer under the conditions, and that of NusE was 16 000 ± 2000 Da. The apparent mass of the mixture was 33 900 ± 2300 Da. This indicates a small increase in effective mass of NusB under these conditions. The expected mass of a complex of 1 NusB dimer and 2 NusE proteins is 61 kDa. The observed mass then indicates the presence of a relatively small amount of such a complex, and therefore a relatively high dissociation constant ($\gg$5 μM). However, we do not know whether the binding of NusE to the NusB dimer is cooperative or not. For noncooperative binding, we would expect a mixture of free NusB and the 1:1 and 1:2 complexes (M = 31 kDa, 45 and 61 kDa). As the initial protein concentrations were 109 and 120 μM for NusB and NusE, respectively, the observed small increase could indicate a dissociation constant of >20 μM. This would make determination of the stoichiometry and dissociation constants difficult.

Preliminary isothermal titration calorimetry results showed an endothermic reaction at pH 8 and 288 K, with ΔH > 10 kcal mol^{-1}. The binding appeared to be weak and the affinity increased when the proteins were buffered at pH 8.0 compared with pH 5. The pI values for NusB and NusE are 7.27 and 10.06, respectively. Thus at pH 8.0 (and presumably at physiological pH), the NusB and NusE have opposite charge which may stabilize the interaction.

DISCUSSION

We have characterized the physical properties of the mycobacterial NusE protein and shown that it is only partly folded in solution (ca. 10−12% α-helix at 30 °C) as shown by the NMR data, far-UV CD, and sedimentation coefficient. The helical content decreases on lowering the temperature, and is disrupted by low (<2 M) concentrations of urea. However, NusE has a significant propensity to form α-helix, as shown by the large increase in helicity at >30% TFE. This indicates a marginal degree of stability, that is strongly affected by the solution conditions. The presence of up to ca. 20 helical residues at the optimal temperature is consistent with the secondary structure prediction (Figure 1). The observation of cold denaturation is unusual for proteins in the absence of denaturants or low pH values. However, we cannot be certain whether there is one long helix present, or several shorter ones.

The CD and NMR spectra both indicate that at 30 °C the protein is only partially folded. If the formation of a NusB:NusE complex requires folding of NusE, then the simplest mechanism to consider is

$$\begin{array}{ccc} & K_1 & \\ E + B & <=> & EB \\ K_E \; \| & & \\ E^* & & \end{array}$$

where E and B represent NusE and NusB, respectively, and E* is an unstructured state of NusE present at high and low temperature. The observed dissociation constant will be:

$$K_{obs} = K_1(1 + K_E) \quad (7)$$

This reflects the presumed unfavorable population distribution biased in favor of E* at low temperature. If $K_E > 1$, then $K_{obs} = K_1K_E$, and $\Delta H_{obs} = \Delta H_1 + \Delta H_E$. For a weak interaction involving a small number of residues, ΔH_1 is expected to be small and negative (binding direction). In the forward direction (i.e., toward EB), E undergoes a net increase in folding. For stable proteins, ΔH(fold) is negative, but for proteins below the temperature of maximum stability, ΔH(fold) is positive. As the analysis of the temperature dependence of the CD spectra showed (Figure 3) the value of ΔH(fold) at 310 K was near zero, but the large ΔCp implies a large positive ΔH(fold) at 288 K [ca. 1.5 kcal (mol residue)$^{-1}$]. Hence, even for a small number of residues becoming helical, the net observed enthalpy change at 288 K would be positive, and decrease toward zero as the temperature is raised above 303 K.

NusE has been implicated in modulating the efficiency of transcription in the phage λ system (*31*) and in ribosomal RNA transcription (*32*). It seems likely that the concentration of free NusE in the cell is low, as it is complexed either in the ribosome or with NusB in the antitermination complex. Its structure may therefore be determined by the interactions it makes with other molecules. The structure of S10 within the intact 30S subunit of *T. thermophilus* has recently been reported at 3 Å (*33*, *34*). S10 contains a core comprising three short β strands and two α helices, with a long β hairpin extension. The total amount of secondary structure elements is quite low, ca. 16% helix and 20% β. The β hairpin is stabilized by multiple interactions with rRNA, and the helix/sheet globular part is stabilized by hydrophobic interactions with S14 and S3. The *B*-factors of the backbone atoms of the beta extension are large (average ca. 100 Å^2), compared with overall *B*-factor of ca. 70 Å^2, suggesting considerable disorder. As this structure would be completely solvent exposed in solution, it is unlikely to be very stable, and probably be dynamically disordered. Thus, the failure to fold in solution to the same structure as in the 30S ribosomal subunit is understandable. It is also possible that NusE folds to an alternative structure on forming a complex with different proteins such as NusB, which would provide a possible rationale for its low intrinsic structure content in the free state.

Previous work had suggested that *E. coli* NusE forms a 1:1 complex with NusB (*4*) and only the complex was able to bind significantly to a 12-base RNA motif called box A (*3*, *4*). A dissociation constant has been reported for the NusE·NusB complex of 0.1 or 3 μM, depending on the method used (*5*). The apparent dissociation constant for box A RNA binding was estimated from band shifts to be in the submicromolar range (*4*) which is comparable to the NusB·NusE dissociation constant. Box A RNA does perturb the NMR spectrum of *E. coli* NusB in the absence of NusE (*11*), indicating at least moderate affinity of this protein for this RNA. However, more recent work has shown that NusE and NusB, while necessary for transcription antitermination, are not sufficient (*6*). The complexity involved in these interactions can be gauged from the phage λ model, where all the components seem to be required to achieve successful antitermination (*3*). We have shown that the homologous NusB and NusE proteins of *M. tuberculosis* can form a weak complex in the absence of RNA. In contrast with *E.coli* NusB, *M. tuberculosis* NusB is a stable homodimer in both the crystalline and solution states (*13*) (and this work). However, we have so far been unable to detect a ternary complex with an *E. coli* box A RNA or a putative box A RNA analogous to that found in *E. coli*. This may indicate that other proteins are in fact essential for this complex to form, such as NusG and/or RNA polymerase, and/or the Mycobacterial box A is not yet defined.

ACKNOWLEDGMENT

We thank Dr. Steve Martin for assistance with the CD measurements and fruitful discussions and Dr. J. Feeney for helpful comments on the manuscript. We wish to thank Dr. Stewart Cole of the Institute Pasteur for kindly providing the *M. tuberculosis* cosmids from which the genes encoding NusB and NusE were amplified.

SUPPORTING INFORMATION AVAILABLE

Table of thermodynamic analysis and four figures of CD, ITC, and NMR spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Richardson, J. P., and Greenblatt, J. (1996) in *Cellularand Molecular Biology* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaecher, M., and Umbarger, H. E., Eds.) pp 822−848, American Society for Microbiology, Washington, DC.
2. Greenblatt, J., Nodwell, J. R., and Mason, S. W. (1993) *Nature 364*, 401−406.
3. Mogridge, J., Mah, T.-F., and Greenblatt, J. (1998) *J. Biol. Chem. 273*, 4143−4147.
4. Nodwell, J. R., and Greenblatt, J. (1993) *Cell 72*, 261−268.
5. Mason, S. W., Li, J., and Greenblatt, J. (1992) *J. Mol. Biol. 223*, 55−66.
6. Zellers, M., and Squires, C. L. (1999) *Mol. Microbiol. 32*, 1296−1304.
7. Ji, Y. E., Colston, M. J., and Cox, R. A (1994) *Microbiol. 140*, 123−132.
8. Ji, Y. E, Kempsell, K. E., Colston, M. J., and Cox, R. A (1994) *Microbiol. 140*, 1763−1773.
9. Altieri, A. S., Mazzulla, M. J., Zhou, H., Costantino, N., Court, D. L., and Byrd, R. A. (1997) *FEBS Lett. 415*, 221−226.
10. Berglechner, F., Richter, G., Fischer, M., Bacher, A., Gschwind, R. M., Huenges, M., Gemmecker, G., and Kessler, H. (1997) *Eur. J Biochem. 248*, 338−346
11. Huenges, M., Rolz, C., Gschwind, R., Peterander, R., Berglechner, F., Richter, G., Bacher, A., Kessler, H., and Gemmecker, G. (1998) *EMBO J. 17*, 4092−4100.
12. Altieri, A. S., Mazzulla, M. J., Horita, D. A., Coats, R., H., Wingfield, P. T., Das, A., Court, D. L., and Byrd, R. A. (2000) *Nat. Struct. Biol. 7*, 470−474.
13. Gopal, B., Haire, L. F., Cox, R. A., Colston, M. J., Major, S. A., Brannigan, J. A., Smerdon, S. J., and Dodson, G. G. (2000) *Nat. Struct. Biol. 7*, 475−478.

14. Scholtz, J. M., Hong, Q., York, E. J., Stewart, J. M., and Baldwin, R. L. (1991) *Biopolymers 31*, 1463−1470.
15. Luo, P., and Baldwin, R. L. (1997) *J. Mol. Biol*. *36*, 8413−8421.
16. Woody, R. W. (1995) *Methods Enzymol. 246*, 34−71.
17. Philo, J. S. (1994) in *Modern Analytical Ultracentrifugation* (Schusterm T. M., and Laue, T. M., Eds.) Birhaüser, Boston.
18. McRorie, D. K., and Voelker, P. J. (1993) in *Self-associating systems in the analytical ultracentrifuge*, Beckman Instruments Inc. Fullerton, CA.
19. Hardy, R. C., and Cottington, R. L. (1949) *J. Chem. Phys*. *17*, 509−510.
20. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett*. *69*, 185−189.
21. Fesik, S. W., and Zuiderweg, E. R. P. (1988) *J. Magn. Reson*. *78*, 588−593.
22. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. R. (1988). *J. Am. Chem. Soc*. *110*, 7870−7872.
23. Piotto, M., Saudek. V., and Sklenar, V. (1992) *J. Biomol. Struct. 2*, 661−665.
24. Bax, A, and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
25. Woody, R. W. (1992) *Adv. Biophys. Chem. 2*, 37−79.
26. Robertson, A. D., and Murphy, K. P. (1997) *Chem. Rev. 97*, 1251−1267.
27. Yao, J., Dyson, H. J., and Wright, P. E. (1997) *FEBS Lett*. *419*, 285−289.
28. Schwalbe, RL, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J., and Dobson, C. M. (1997) *Biochemistry 36*, 8977−8991.
29. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR 5*, 67−81.
30. Eleizer, D., Yao, J., Dyson, H. J., and Wright, P. E. (1998) *Nat. Struct. Biol*. *5*, 148−155.
31. Das, A., and Wolska, K. (1984) *Cell 38*, 615−173.
32. Friedman, D. I., Schaur, A. T., Baumann, M. R., Baron, L. S., and Adhya, S. L. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 1115−1118.
33. Wimberly, B. T., Broderson, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) *Nature 407*, 327−339.
34. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., and Yonath, A. (2000) *Cell 102*, 615−623.

BI0018279