The structure of the transcriptional antiterminator NusB from *Escherichia coli*

Amanda S. Altieri[1], Marie J. Mazzulla[1,2], David A. Horita[1], R. Heath Coats[1], Paul T. Wingfield[3], Asis Das[4], Donald L. Court[5] and R. Andrew Byrd[1]

[1]*Structural Biophysics Laboratory, National Cancer Institute-FCRDC, P.O. Box B, Building 538, Frederick, Maryland 21702, USA.* [2]*Present address: SmithKline Beecham Pharmaceuticals, 1250 S. Collegeville Road, Collegeville, Pennsylvania 19426, USA.* [3]*Protein Expression Laboratory, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.* [4]*Department of Microbiology, University of Connecticut Health Center, Farmington, Connecticut 06030, USA.* [5]*Gene Regulation and Chromosome Biology Laboratory, National Cancer Institute-FCRDC, P.O. Box B, Building 539, Frederick, Maryland 21702, USA.*

We have determined the solution structure of NusB, a transcription antitermination protein from *Escherichia coli*. The structure reveals a novel, all α-helical protein fold. NusB mutations that cause a loss of function (NusB5) or alter specificity for RNA targets (NusB101) are localized to surface residues and likely affect RNA–protein or protein–protein interactions. Residues that are highly conserved among homologs stabilize the protein core. The solution structure of *E. coli* NusB presented here resembles that of *Mycobacterium tuberculosis* NusB determined by X-ray diffraction, but differs substantially from a solution structure of *E. coli* NusB reported earlier.

Antitermination in bacteriophage λ is a process in which the phage-encoded N protein and host-encoded Nus factors modify RNA polymerase to a termination resistant form at the *nut* site in the transcribed RNA[1–6]. The *nut* site contains a 12-nucleotide strand (boxA), and a stem-loop structure (boxB). Two of the Nus proteins, NusB and NusE, form a heterodimer that specifically binds to *boxA* RNA and enhances antitermination[7,8]. *In vitro*, antitermination is decreased in the absence of NusB, in that it is limited to terminators close to the promoter[9,10]. NusB also competes with a cellular inhibitor that binds to *boxA* and prevents antitermination[11]. *In vivo*, a null mutation in the *nusB* gene produces a cold sensitive phenotype (no cell growth below 32 °C)[12] that is often associated with defects in ribosome assembly. NusB is also required for antitermination of the *E. coli* ribosomal (rrn) RNA operon[13], where it increases the rate of ribosomal rrn boxA-mediated transcription elongation[14].

Description of the structure

The solution structure of NusB determined by NMR is shown as a superposition of 15 of the lowest energy structures in Fig. 1*a*. Structural statistics are given in Table 1. Complete resonance assignments are available at the BMRB (http://www.bmrb.wisc.edu, accession code 4737) and the list of restraints and coordinates are available at the RCSB (http://www.rcsb.org/pdb). NusB can be viewed as two subdomains with helices α1–α3 forming the N-terminal subdomain (Fig. 1*b*,*c*, gold) and α4–α7 forming the C-terminal subdomain (purple). The orientation of the two subdomains relative to each other is 130°, measured as the angle between α1 and α5. Subdomain orientation was determined from observed NOEs between α1 and α5 (Fig. 2) and between α3 and α5 (Table 2). Relative helix orientations and the NOEs that define the fold are listed in Table 2. In addition, several NOEs locate the C-terminal end of α4 near α7: from Ala 130 $CH\beta_3$–Met 66 Hα, Ala 130 $CH\beta_3$–Tyr 69 Hδ, and Ala 130 $CH\beta_3$–Tyr 69 Hε and Ala 130 HN–Tyr 69 Hε. These NOEs are well resolved and their unambiguous assignment was key to defining the global fold of NusB in the early stages of the structure calculation.

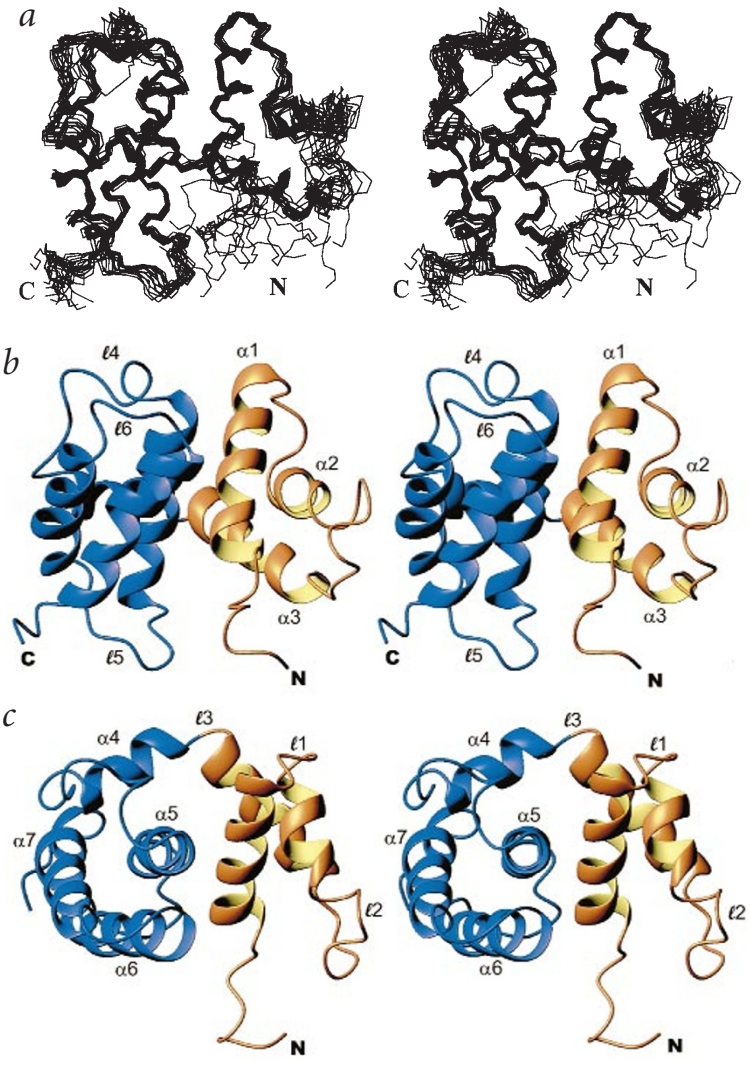

Fig. 1 Solution structure of *E.coli* NusB. ***a***, Stereo view showing the superposition of 15 of the lowest energy structures of NusB. ***b***, Stereo view of a ribbon trace of NusB. The N-terminal subdomain is colored gold, and the C-terminal subdomain is colored purple. Helices and loops are labeled as discussed in the text. ***c***, Stereo view showing a 90_y° rotation of (*b*).

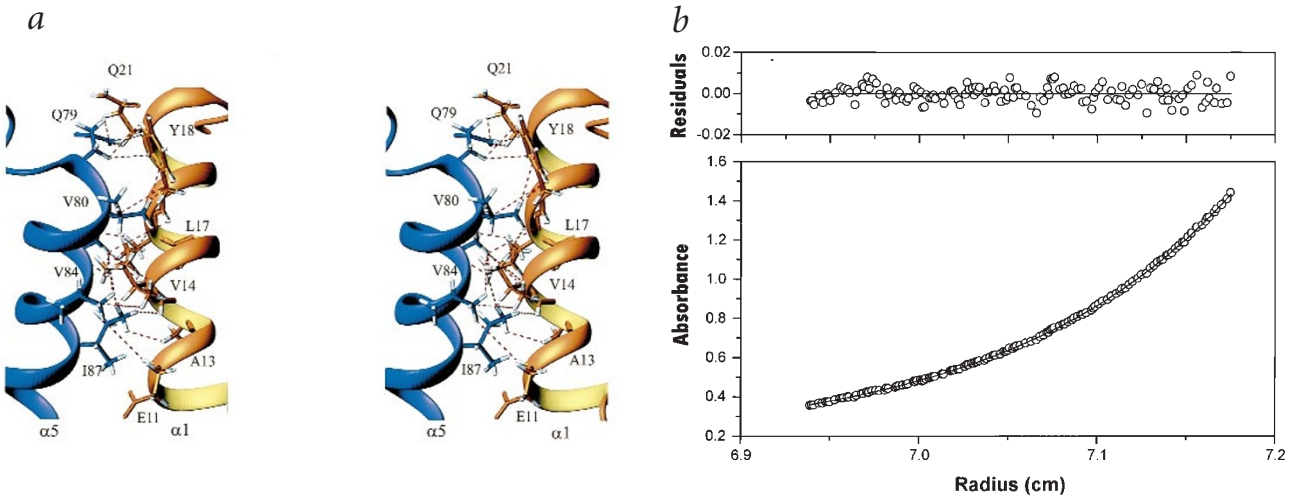

Fig. 2 Orientation of subdomains and momeric state of NusB. ***a***, Long range NOE contacts between α1 (gold) and α5 (blue). These NOEs define the orientation of the N-terminal and C-terminal subdomains relative to each other. ***b***, The determination of the molecular mass of *E. coli* NusB in 50 mM potassium phosphate, pH 6.8, 50 mM NaCl and 1 mM DTT at 20 °C. The absorbance gradient (at 280 nm) in the centrifuge cell after attaining sedimentation equilibrium at 25,000 r.p.m. is shown in the bottom panel. The solid line is the result of fitting to a single ideal species and the open circles are the experimental values. The corresponding top panel shows the difference between the fit and the experimental values as a function of radial position (residuals).

Most of the hydrophobic residues in NusB are internal and involved in helix packing. A few are partially exposed, as are all four Phe side chains. Most of the hydrophilic residues are on the surface. The packing quality of the structure is acceptable as evaluated by WhatIf[15] (QUACHK = -1.80 ± 0.07) and by analysis of the distribution of hydrophobic and hydrophilic residues (INOCHK = 1.0 ± 0.02). An interesting feature of the structure is a small, positively charged cavity formed by the termini of Lys 82 and Arg 86, on the backside of the protein from the view in Fig. 1*b*. Since these residues are highly conserved (Fig. 3*a*), this charged cavity is likely a conserved feature in NusB proteins. It is present in the *Mycobacterium tuberculosis* NusB structure[16].

Comparison to another *E. coli* NusB solution structure

Another solution structure of *E. coli* NusB has been reported (Protein Data Bank accession code 1BAQ)[17]. Although some of the helix-to-helix contacts are similar between this structure and ours, the three-dimensional fold is remarkably different. The root mean square (r.m.s.) deviation between the two structures over the Cα atom trace of the helices is 9.9 Å. The Cα trace r.m.s. deviation between the helices in the N-terminal subdomain is 6.0 Å, while it is 8.3 Å for the C-terminal subdomain. In our structure, the subdomain orientation is 130° (roughly antiparallel; Fig 2*a*), while it is 10° in 1BAQ (parallel). Additionally, the nearly perpendicular orientation of α3-α4 (80°) and α4-α5 (101°) in our structure is markedly different from the angles between these helices in 1BAQ (31° and 54°, respectively). Furthermore, the core helices in our structure are α1 and α5 while in 1BAQ they are α3 and α6. The long range contacts reported as α1–α6 NOEs in 1BAQ were not observed in our NOESY data. Key long range NOEs that define our structure are between α1 and α5 (Fig. 2 and Table 2). The solution conditions used in the two studies are reported to be similar, except for the absence of salt for 1BAQ, and the presence of 100 mM NaCl in our buffer. The ^{15}N HSQC spectrum reported by each group appears to be very similar, although there are some differences in sequential assignments[17,18]. Since the full assignments and list of restraints for 1BAQ were not reported, further evaluation and comparisons could not be made. In our structure determination of NusB, 4D NOESY spectra were used to assign the contributing heteronucleus to each H-X pair (where X is either N or C). Complete side chain assignment for observed residues was critical, particularly for the 14 aromatic side chains (Table 2). Additionally, the use of ARIA[19,20] contributed to a resulting completeness of assignment of 93% (2,045 of 2,189 crosspeaks) of the observed peaks in the 4D NOESY spectra. Our structure of *E. coli* NusB has been independently confirmed by the X-ray crystal structure of *M. tuberculosis* NusB[16].

Comparison to *M. tuberculosis* NusB

The structure of the *M. tuberculosis* NusB protein has been solved by X-ray crystallography[16]. The amino acid sequences of *E. coli* and *M. tuberculosis* NusB are 57% homologous and 34% identical (Fig. 3*a*). The two NusB structures were solved independently and compared after the refinement for each protein was completed. The *M. tuberculosis* NusB crystal structure reveals a dimer in the asymmetric unit. Prior characterization of the *E. coli* NusB protein showed it to be a monomer in solution at concentrations up to 1 mM[18,21]. Our data confirm these results; the *E. coli* NusB protein behaves as an ideal monomer during sedimentation equilibrium centrifugation (Fig. 2*b*). There was no tendency for aggregation or self-association over the concentration range studied, and the determined molecular mass of 15,499 Da was within a few percent of that expected from the sequence (15,679 Da).

The structures of the *E. coli* and *M. tuberculosis* NusB monomers are very similar. Helices α1 and α3–α7 superimpose well with an r.m.s. deviation between Cα atoms of 2.0 Å. Helix α2 is closer to α1 in the *E. coli* structure than it is in the *M. tuberculosis* one. This difference may be due to the conformational heterogeneity of α2 and loop ℓ2 in solution (residues Glu 31–Asp 44 are more disordered; Fig. 1*a*). The difference in

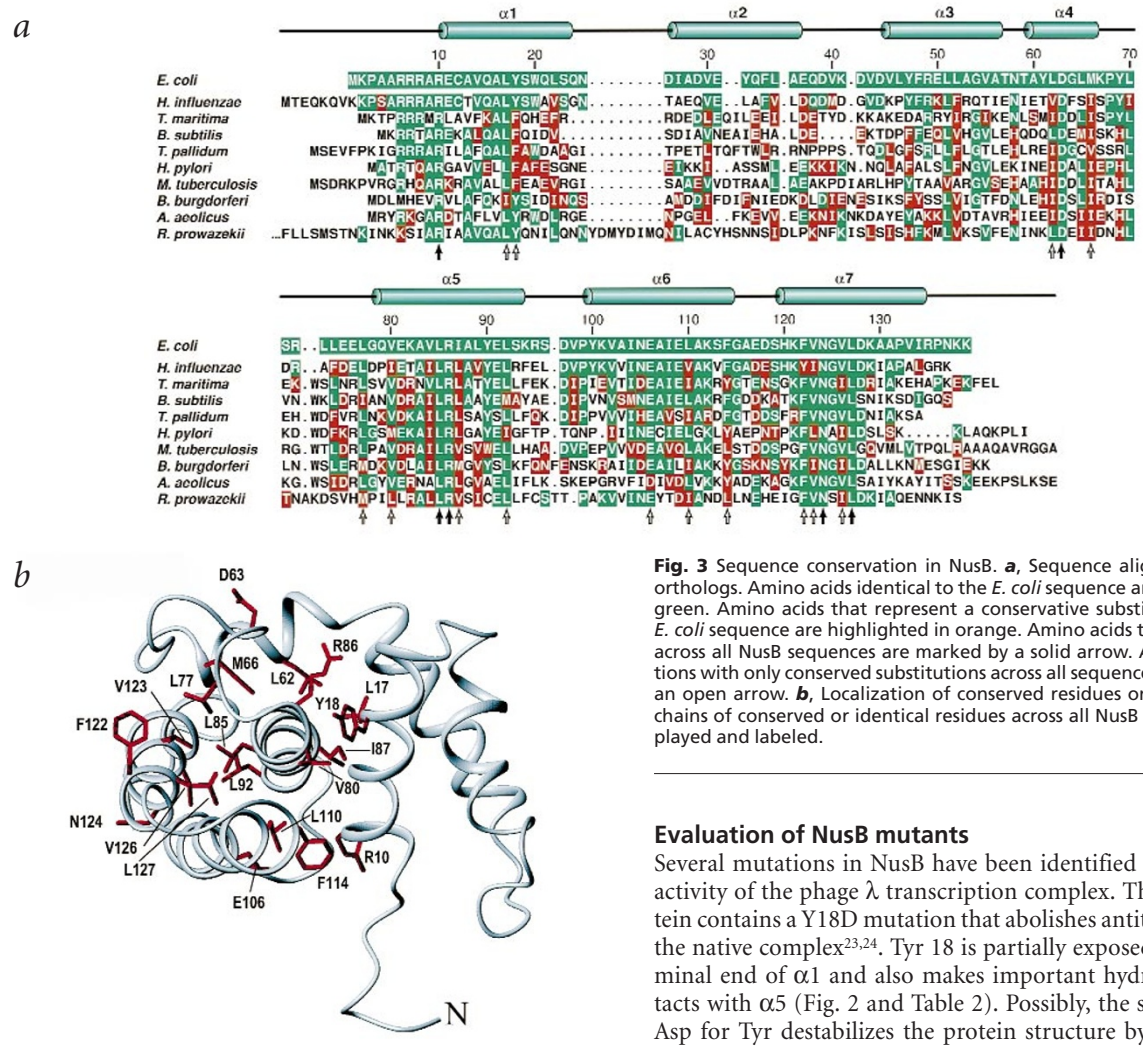

Fig. 3 Sequence conservation in NusB. ***a***, Sequence alignment of NusB orthologs. Amino acids identical to the *E. coli* sequence are highlighted in green. Amino acids that represent a conservative substitution from the *E. coli* sequence are highlighted in orange. Amino acids that are identical across all NusB sequences are marked by a solid arrow. Amino acid positions with only conserved substitutions across all sequences are marked by an open arrow. ***b***, Localization of conserved residues on NusB. The side chains of conserved or identical residues across all NusB proteins are displayed and labeled.

the positions of α2 may also be due to dimer contacts in the *M. tuberculosis* protein, since α2 makes contacts with α2′. The similarity of the *E. coli* and *M. tuberculosis* NusB structures supports the correctness of the protein fold presented here.

Structural biology of sequence conservation

NusB homologs are present in a variety of organisms. A subset of NusB orthologs found using the COGnitor program (COG0781, http://www.ncbi.nlm.nih.gov/COG/)[22] is aligned in Fig. 3*a*. There is a high degree of identity to the *E. coli* sequence across helices α1, α5, α6 and α7. The location of conserved residues is shown in Fig. 3*b*. Four conserved residues are on the surface of the protein (Arg 10, Asp 63, Glu 106 and Asn 124) and all three conserved aromatic residues (Tyr 18, Phe 114 and Phe 122) are partially exposed to solvent. Aromatic residues exposed on the surface are often involved in recognition and specificity at intermolecular binding surfaces. The biological significance of Tyr 18 was established by defective antitermination of the *nusB5* mutant[23,24]. Phe 114 and Phe 122 may have similar functional significance. The remaining conserved residues are in the core of the protein and are involved in multiple helix-to-helix contacts. This distribution suggests that the main role of conservation in NusB is to stabilize the protein fold. It also implies that the fold of the orthologous NusB proteins is likely to be very similar.

Evaluation of NusB mutants

Several mutations in NusB have been identified that affect the activity of the phage λ transcription complex. The NusB5 protein contains a Y18D mutation that abolishes antitermination of the native complex[23,24]. Tyr 18 is partially exposed at the C-terminal end of α1 and also makes important hydrophobic contacts with α5 (Fig. 2 and Table 2). Possibly, the substitution of Asp for Tyr destabilizes the protein structure by reducing the number of hydrophobic contacts between α1 and α5; alternatively, the mutation may change the electrostatic nature of the region. Another possibility is that Tyr 18 is involved in a specific interaction with RNA, perhaps stacking with nucleic acid bases. In this case, substitution of any nonaromatic amino acid at position 18 would have the same effect as the NusB5 mutation.

A second NusB mutant, NusB101, is a D118N substitution that has no effect on the native transcription complex, yet rescues defective antitermination caused by the NusA1 mutant[25]. NusB101 similarly rescues an antitermination defect resulting from the NusE71 mutation[24]. The Asp to Asn substitution is normally considered conservative, and the location of residue 118 on a surface loop makes it unlikely to destabilize the structure. More likely, Asp 118 is part of or near contact surfaces for nucleic acid interactions. The observation that the rescue of antitermination defect of NusB101 requires the presence of *boxA*[24] supports this hypothesis as does the report that NusB101 has native binding affinity for NusE. This mutation is specific to phage λ, since the NusB101 mutation does not have the same affect on phage 21 N-mediated antitermination. The overall implication is that the NusB101 mutation might directly enhance binding of NusB (or a NusB–protein complex) to phage λ *boxA*.

Structural homology

No protein with a structure similar to NusB has been found; hence, this represents a new protein fold. However, structural

Table 1 Summary of restraints and structural statistics

Restraints		
NOEs[1]		
Intraresidue	1,474	
Sequential	429	
Medium range (i < 5)	341	
Long range	460	
Unambiguous	2,699	
Ambiguous	5	
Total NOEs	2,704	
Others		
ϕ,ψ	150	
Hydrogen bonds[2]	48 (× 2)	
Total number of restraints	2,948	
Deviations from experimental	<SA>[3]	Lowest Energy
R.m.s. deviation of NOE	0.020 (0.001)	0.018
NOE violations > 0.3 Å	0.4 (0.6)	0
ϕ,ψ violations > 5°	0.3 (0.5)	0
Deviations from ideal geometry		
Bonds (Å)	0.0018 (0.0001)	0.0016
Angles (°)	0.360 (0.006)	0.351
Impropers (°)	0.277 (0.011)	0.270
Precision		
Backbone helices[4]	0.43 (0.07)	
Heavy atoms helices[5]	0.91 (0.10)	
Heavy atoms 10–34, 48–136	1.09 (0.17)	
Structure quality		
Procheck (%; mf / aa / ga / da)[6]	74 / 18 / 6 / 2	77 / 17 / 4 / 2
WhatIf[7]	-1.80 (0.07)	-1.79
X-PLOR energy[8]	163.1 (6.5)	149

[1]NOEs were counted with explicit inclusion of all H atoms of methyl and methylene groups (that is, no pseudoatoms). Trivial distances were not included.
[2]Hydrogen bonds were included as a restraint of 1.5 (0.8) Å between HN_i and O_{i-3} atoms and a restraint of 2.5 (0.8) Å between N_i and O_{i-3} for those residues whose amides where determined to be in slow to intermediate exchange within helices. The value in parentheses is the upper bound on the restraint.
[3]Values are reported as the average values over 15 of the lowest energy structures with standard deviations in parentheses.
[4]The average r.m.s. deviation for the coordinate set was calculated by superimposing each of the 15 structures onto the mean coordinate set. This superposition was over backbone N, C, O and Cα atoms of residues 13–20, 27–30, 48–66, 79–92, 100–114, 120–133.
[5]This superposition was over nonhydrogen atoms of residues 13–20, 27–30, 48–66, 79–92, 100–114, 120–133.
[6]Procheck analysis[39]: mf, most favored; aa, additionally allowed; ga, generously allowed; da, disallowed.
[7]WhatIf score (QUACHK)[15].
[8]Energy calculated from X-PLOR 3.851[38] with force constants of 50 kcal mol^{-1} for the NOE restraints and 200 kcal mol^{-1} rad^{-2} for the torsion angle restraints. All other force constants used were the default values.

similarities were found for subsets of the NusB structure and several of these proteins contain homeodomains[26,27]. Both the N-terminal and C-terminal subdomains of NusB contain helix-turn-helix (HTH) features, including a helix rich in basic residues that corresponds to the recognition helix of the homeodomains. Although the HTH motif is prevalent in transcription factors, it is not yet known whether the HTH-like domains in NusB function as nucleic acid binding domains.

Implications for nucleic acid binding

In the context of specific single stranded RNA binding, the all helical structure of NusB is unique. Recent structures of single stranded nucleic acid–protein complexes show that the nucleic acid binding surface of the protein commonly consists of a β-sheet[28–31]. However, there is no obvious groove or cleft in NusB containing positive charges and exposed aromatic residues that would make specific contacts similar to those in the single stranded nucleic acid–protein complexes. Specific *boxA* binding to NusB has only been demonstrated in the presence of NusE[8]. NusE may assist recognition by either contributing part of the RNA binding site, or by changing the structure of NusB to create a specific nucleic acid binding site. Future work on the details of specificity will undoubtedly reveal more novel features in the transcription regulation machinery.

Methods

Sample preparation. NusB was prepared as described[18] to produce $^{2}H/^{15}N$, $^{13}C/^{15}N$ and ^{15}N labeled samples. The sample conditions were ~1 mM protein concentration in 50 mM sodium phosphate buffer, 0.1 M NaCl and 2 mM dithiothreitol at pH 6.8 and 25 °C. The NusB protein was active in an *in vitro* antitermination assay[32].

Analytical ultracentrifugation. Analytical ultracentrifugation was performed at 25,000 rpm, 20 °C, using a Beckman XL-I Optima analytical ultracentrifuge with an An-60 Ti rotor and standard double sector centerpiece cells. Solvent density was calculated according to Laue *et al*.[33]. The partial specific volume of the protein (0.744) was calculated from the predicted amino acid composition[34]. Centrifugation data were analyzed using the Beckman-Origin software.

NMR spectroscopy. NMR spectra were acquired on a Varian Unity Plus 600 MHz spectrometer. Spectra were processed using NMRPipe[35] and assigned using ANSIG 3.3[36,37]. ^{1}H, ^{15}N and ^{13}C assignments have been made for NusB and are available at the BMRB. Sequential assignments are 94% complete. Out of the 139 residue in NusB, 4 residues (Arg 6, Arg 7, Arg 8 and Asp 44) are completely unassigned. In addition, the NHs of 11 residues (Met 1–Arg 10 and Asp 44) have no assignments; only one atom pair in each of the four residues (Met 1, Arg 10, Asp 42 and Val 43) are assigned. For six residues (Ala 4, Ala 5, Ala 9, Phe 34, Phe 122, Lys 138), only one atom in each is unassigned. Aromatic side chains were assigned using 2D CB(CGCD)HD and CB(CGCDCE)HE experiments combined with NOESY and HSQC spectra. The aromatic side chains are completely assigned with the exception of Phe 34 CξHξ and Phe 122 CξHξ. Interproton distances were measured from the following spectra with the given mixing times (t_{mix}): ^{15}N edited NOESY-HSQC (t_{mix} = 120 ms); $^{15}N/^{13}C$ (simultaneous) edited NOESY-HSQC (t_{mix} = 100 ms); $^{15}N/^{15}N$ HSQC-NOESY-HSQC (t_{mix} = 200 ms); $^{13}C/^{15}N$ HMQC-NOESY-HSQC (t_{mix} = 100ms), and $^{13}C/^{13}C$ HMQC-NOESY-HSQC (t_{mix} = 100 ms).

Structural calculation. Structures were calculated using X-PLOR 3.851 (ref. 38). A fully extended starting conformation was used on which 24,000 steps of simulated annealing at 1,200 K followed by 15,000 cooling steps of 0.005 ps to 100 K were carried out. Initial structures (100) were calculated using 1891 unambiguous NOE restraints. Fifteen of the lowest energy structures (backbone r.m.s.d. 1.1 Å) were used as starting coordinates for ARIA[19,20]. The following ARIA protocol included 343 NOE crosspeaks as ambiguous distance restraints ('P' values are listed first followed by the assignment cutoff distance in parentheses): 0.999 (5), 0.999 (2), 0.99 (1), 0.99 (0.5), 0.98 (0.5), 0.96 (0.25), 0.93 (0.25), 0.90 (0.20), 0.80 (0.2).

Structural homology. The DALI program (http://www2.ebi.ac.uk/dali) was used to search protein databases for structures similar to NusB. Selection criteria were a Z-score ≥ 3.0 or an r.m.s. deviation between Cα atoms ≤ 3.0 Å over subsets comprising ≥ 3 helices.

Coordinates. Coordinates have been deposited in the Protein Data Bank (accession code 1EY1).

Table 2 Interhelix angles[1] and residues with observed NOEs between helices

Helices	Angle (°)	NOEs[2]
α1-α2	123	Ala 9–**Tyr 32**, Gln 15–**Phe 34**, Ala 16–Val 30, Ala 16–**Phe 34**, Leu 17–Ile 27, Ser 19–Val 30, **Trp 20**–Val 30
α2-α3	151	Ile 27–Leu 52, Ile 27–A56, Glu 31–Leu 52, **Tyr 32**–Leu 51, Gln 33–Arg 49, **Phe 34**–Val 40, Leu 35–Val 40, Leu 35–**Phe 48**, Leu 35–Arg 49
α1-α3	51	Ala 9–**Phe 48**, Ala 9–Leu 51, Arg 10–Leu 51, Cys 12–**Phe 48**, Ala 13–**Phe 48**, Ala 13–Leu 51, Ala 13–Leu 52, Ala 16–Leu 52, Leu 17–Leu 52, **Trp 20**–Ala 56, **Trp 20**–Thr 59
α3-α4	80	
α4-α5	101	**Tyr 69**–Lys 82, Met 66–Lys 82, Met 66–Leu 85, Met 66–Arg 86, Met 66–Leu 89, Leu 62–**Tyr 90**, Leu 62–Glu 91
α6-α5	161	Leu 92–Ala 103, Leu 92–Ile 104, Ala 88–Ala 103, Ala 88–Leu 110, Leu 85–Ala 107, Val 84–Ala 107, Val 84–Ala 111, Val 80–**Phe 114**, Val 80–Gly 115, Val 80–Ala 116
α7-α5	32	Leu 85–Val 123, Leu 85–Val 126, Leu 85–Leu 127, Ala 88–Leu 127, Leu 89-Ala 130, **Tyr 90**–Ile 134, Leu 92–Ala 131

[1]The angles were calculated by defining a cylinder based on Cα atom position for each helix using the Molmol graphics program[40].
[2]NOEs involving aromatic residues are in bold.

Acknowledgments
The authors would like to thank S. Sparks, P. Domaille and F. Delaglio for valuable discussions and L. Pannell for mass spectroscopy. We also thank the staff and administration of the Advanced Biomedical Computing Center for their support of this project. Part of this work was performed under the auspices of the ABL-Basic Research Program under contract to the NCI, DHHS.

Correspondence should be addressed to R.A.B. *email: rabyrd@ncifcrf.gov*

Received 1 February, 2000; accepted 26, April, 2000.

1. Das, A. *J. Bacteriol.* **174**, 6711–6716 (1992).
2. Greenblatt, J. in *Transcriptional regulation* (eds McKnight, S.L. & Yamamoto, K.R.) 203–226 (Cold Spring Harbor Laboratory Press, New York; 1992).
3. Das, A. *Annu. Rev. Biochem.* **62**, 893–930 (1993).
4. Roberts, J.W. *Cell* **72**, 653–655 (1993).
5. Greenblatt, J., Nodwell, J.R. & Mason, S.W. *Nature* **364**, 401–406 (1993).
6. Friedman, D.I. & Court, D.L. *Mol. Microbiol.* **18**, 191–200 (1995).
7. Mason, S.W., Li, J. & Greenblatt, J. *J. Mol. Biol.* **223**, 55–66 (1992).
8. Nodwell, J.R. & Greenblatt, J. *Cell* **72**, 261–268 (1993).
9. Mason, S.W. & Greenblatt, J. *Genes Dev.* **5**, 1504–1512 (1991).
10. DeVito, J. & Das, A. *Proc. Natl. Acad. Sci. USA* **91**, 8660–8664 (1994).
11. Patterson, T.A. *et al. J. Mol. Biol.* **236**, 217–228 (1994).
12. Taura, T., Ueguchi, C., Shiba, K. & Ito, K. *Mol. Gen. Genet.* **234**, 429–432 (1992).
13. Squires, C.L., Greenblatt, J., Li, J. & Condon, C. *Proc. Natl. Acad. Sci. USA* **90**, 970–974 (1993).
14. Zellars, M. & Squires, C.L. *Mol. Microbiol.* **32**, 1296–1304 (1999).
15. Hooft, R.W., Vriend, G., Sander, C. & Abola, E.E. *Nature* **381**, 272 (1996).
16. Gopal, B. *et al. Nature Struct. Biol.* **7**, 475–478 (2000).
17. Huenges, M. *et al. EMBO J.* **17**, 4092–4100 (1998).
18. Altieri, A.S. *et al. FEBS Lett.* **415**, 221–226 (1997).
19. Nilges, M. *J. Mol. Biol.* **245**, 645–660 (1995).
20. Nilges, M., Macias, M.J., O'Donoghue, S.I. & Oschkinat, H. *J. Mol. Biol.* **269**, 408–422 (1997).
21. Berglechner, F. *et al. Eur. J. Biochem.* **248**, 338–346 (1997).
22. Tatusov, R.L., Galperin, M.Y., Natale, D.A. & Koonin, E.V. *Nucleic Acids Res.* **28**, 33–36 (2000).
23. Friedman, D.I., Baumann, M. & Baron, L.S. *Virology* **73**, 119–127 (1976).
24. Court, D.L. *et al. J. Bacteriol.* **177**, 2589–2591 (1995).
25. Ward, D.F., DeLong, A. & Gottesman, M.E. *J. Mol. Biol.* **168**, 73–85 (1983).
26. Ohlendorf, D.H., Anderson, W.F., Fisher, R.G., Takeda, Y. & Matthews, B.W. *Nature* **298**, 718–723 (1982).
27. Sauer, R.T., Yocum, R.R., Doolittle, R.F., Lewis, M. & Pabo, C.O. *Nature* **298**, 447–451 (1982).
28. Horvath, M.P., Schweiker, V.L., Bevilacqua, J.M., Ruggles, J.A. & Schultz, S.C. *Cell* **95**, 963–974 (1998).
29. Handa, N. *et al. Nature* **398**, 579–585 (1999).
30. Bogden, C.E., Fass, D., Bergman, N., Nichols, M.D. & Berger, J.M. *Mol. Cell* **3**, 487–493 (1999).
31. Ding, J. *et al. Genes Dev.* **13**, 1102–1115 (1999).
32. Das, A. *et al. Methods Enzymol.* **274**, 374–402 (1996).
33. Laue, T.M., Shah, B.D., Ridgeway, T.M. & Pelletier, S.L. in *Analytical ultracentrifugation in biochemistry and polymer science* (eds Harding, S.E., Rowe, A.J. & Horton, J.C.) 90–125 (Royal Society of Chemistry, London; 1992).
34. Cohn, E.J. & Edsall, J.T. *Proteins, amino acids and peptides* (Van Nostrand-Reinhold, Princeton, New Jersey; 1943).
35. Delaglio, F. *et al. J. Biomol. NMR* **6**, 277–293 (1995).
36. Kraulis, P.J. *et al. Biochemistry* **28**, 7241–7257 (1989).
37. Kraulis, P.J., Domaille, P.J., Campbell-Burk, S.L., Van Aken, T. & Laue, E.D. *Biochemistry* **33**, 3515–3531 (1994).
38. Brunger, A.T. *X-PLOR (version 3.1) A system for X-ray cyrstallography and NMR* (Yale University Press, New Haven, 1987).
39. Laskowski, R.A., Rullman, J.A., MacArthur, M.W., Kaptein, R. & Thornton, J.M. *J. Biomol. NMR* **8**, 477–486 (1996).
40. Koradi, R., Billeter, M. & Wüthrich, K., *J. Mol. Graphics* **14**, 51–55 (1996).