Cell, Vol. 93, 289–299, April 17, 1998, Copyright ©1998 by Cell Press

NMR Structure of the Bacteriophage λ N Peptide/*boxB* RNA Complex: Recognition of a GNRA Fold by an Arginine-Rich Motif

Pascale Legault,*†‡§ Joyce Li,* Jeremy Mogridge,*† Lewis E. Kay,†‡ and Jack Greenblatt*†
*Banting and Best Department of Medical Research
†Department of Molecular and Medical Genetics
University of Toronto
Toronto, Ontario
Canada, M5S 1A8
‡Protein Engineering Network of Centers of Excellence and Departments of Biochemistry and Chemistry
University of Toronto
Toronto, Ontario
Canada, M5S 1A8

Summary

The structure of the complex formed by the arginine-rich motif of the transcriptional antitermination protein N of phage λ and *boxB* RNA was determined by heteronuclear magnetic resonance spectroscopy. A bent α helix in N recognizes primarily the shape and negatively charged surface of the *boxB* hairpin through multiple hydrophobic and ionic interactions. The GAAGA *boxB* loop forms a GNRA fold, previously described for tetraloops, which is essential for N binding. The fourth nucleotide of the loop extrudes from the GNRA fold to enable the *E. coli* elongation factor NusA to recognize the N protein/RNA complex. This structure reveals a new mode of RNA–protein recognition and shows how a small RNA element can facilitate a protein–protein interaction and thereby nucleate formation of a large ribonucleoprotein complex.

Introduction

The phage λ N protein prevents transcription termination in the two phage early operons (reviewed by Greenblatt et al., 1993; Friedman and Court, 1995). This role of N resembles those of the immunodeficiency virus Tat proteins and depends on an N utilization site (*nut* site) in the transcribed region. The *nut* site consists of RNA and assembles into a ribonucleoprotein complex containing N and four *E. coli* proteins, NusA, NusB, NusG, and ribosomal protein S10. This complex associates stably with RNA polymerase during elongation and inhibits transcription termination.

A *nut* site consists of two genetically defined elements, *boxA* (Olson et al., 1984) and *boxB* (Salstrom and Szybalski, 1978). *boxB* forms an RNA hairpin (Figure 1A) and its 5 bp stem and 5 nt loop are recognized by N (Franklin, 1984; Lazinski et al., 1989). Although all 5 nucleotides of the *boxB* loop are essential for N antitermination function (Doelling and Franklin, 1989; Chattopadhyay et al., 1995a), most studies have shown that only loop nucleotides 1, 3, and 5 are important for N binding in gel mobility shift assays (Chattopadhyay et al., 1995a; Mogridge et al., 1995; Tan and Frankel, 1995; Cilley and Williamson, 1997). Loop nucleotides 2 and 4 are needed for binding of the N protein/*nut* complex to the *E. coli* elongation factor NusA (Mogridge et al., 1995).

The first 22 amino acids of N ($N^{1–22}$) constitute an arginine-rich motif (ARM), which has the same affinity (K_D = 5 nM) and specificity for *boxB* RNA as full-length N (Chattopadhyay et al., 1995a; Tan and Frankel, 1995; Cilley and Williamson, 1997; Van Gilst et al., 1997). N is a disordered protein whose ARM adopts a specific fold upon binding to *boxB* while the other domains of N (NusA- and polymerase-binding domains) remain disordered (Mogridge et al., 1998). Circular dichroism and NMR experiments suggest that the portion of N that binds *boxB* RNA is α helical (Tan and Frankel, 1995; Su et al., 1997a; Van Gilst et al., 1997). Moreover, mutational studies on the ARM of N indicate that amino acids important for N binding and antitermination (Figure 1B) are on one face of an α helix (Franklin, 1993; Su et al., 1997a).

Here, we have used NMR to determine the structure of the $N^{1–22}$/*boxB* RNA complex. The structures of the complexes formed by the ARMs of the BIV Tat (Puglisi et al., 1995; Ye et al., 1995) and HIV-1 Rev (Battiste et al., 1996; Ye et al., 1996) proteins with their RNA ligands have also been determined by NMR. The ARM of N binds RNA differently from either BIV Tat or HIV-1 Rev. $N^{1–22}$ adopts a bent α helix that does not penetrate into the major groove but instead binds exclusively to the 5′ strand of the *boxB* stem and the first three residues of the loop, recognizing primarily the shape and negatively charged surface of the *boxB* hairpin. Remarkably, the *boxB* loop structure is virtually identical to the structure of the GAAA tetraloop (Heus and Pardi, 1991). The fourth nucleotide of the *boxB* GAAGA loop extrudes from the GAAA-like structure, allowing specific binding to NusA in *E. coli* elongation complexes. The structure presented here reveals new principles of RNA folding and RNA–protein recognition and contributes to our growing understanding of transcriptional regulation in bacteriophage λ.

Results and Discussion

A New Mode of Recognition for an Arginine-Rich Motif

The structure of the λ $N^{1–22}$/*nutL boxB* RNA complex was determined from heteronuclear NMR experiments. Preliminary experiments indicated that $N^{1–22}$ and *boxB* RNA form a well-ordered 1:1 complex in the conditions we used for NMR (Mogridge et al., 1998). Formation of a well-ordered complex under these conditions is prevented by a *boxB* loop mutation that abolishes N binding and antitermination (Mogridge et al., 1998). To allow selective detection of NMR signals from each component of the complex, samples were prepared by mixing an isotopically labeled (^{15}N or $^{15}N/^{13}C$) molecule (peptide or RNA) with the corresponding unlabeled ligand. Multiple intermolecular distance restraints were derived from improved isotopically filtered nuclear Overhauser effect spectroscopy (NOESY) experiments (Zwahlen et

§To whom correspondence should be addressed.

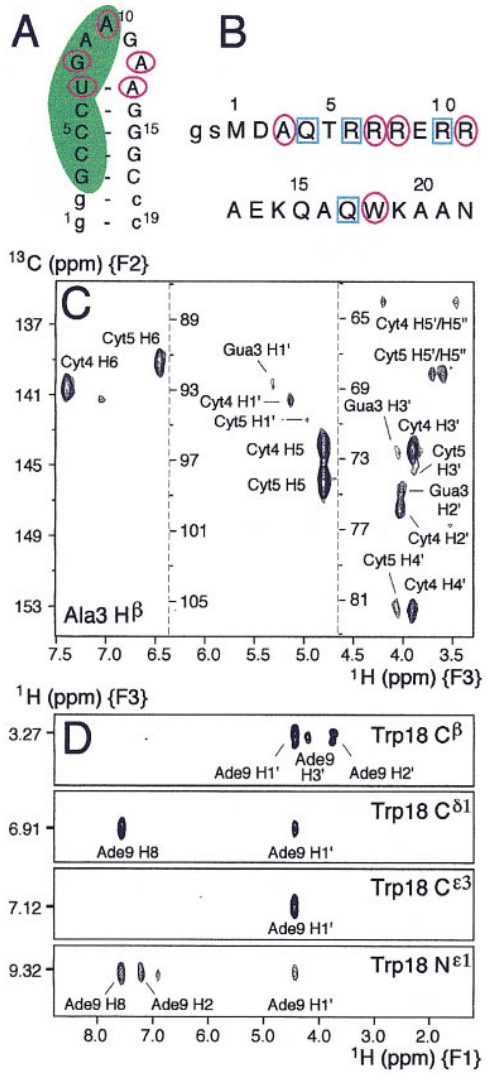

Figure 1. Sequences of *boxB* RNA and N^{1-22} Peptide Used for the NMR Studies and Selected Filtered NOESY Data

(A) Sequence and proposed secondary structure of *boxB* RNA. Two additional G-C base pairs (g1-c19 and g2-c18) were added to the λ *nutL boxB* RNA hairpin stem. Nucleotides important for N binding (see text) are circled and those protected by N from RNase cleavage are covered by a green shadow (Chattopadhyay et al., 1995a).

(B) Sequence of the N^{1-22} peptide used for the NMR studies. The numbering is as for natural N. Lowercase letters indicate GST-fusion-derived amino acids (Mogridge et al., 1998). Framed amino acids are critical (pink circle) or simply preferred (blue squares) for antitermination (Franklin, 1993; Su et al., 1997a).

(C) Selected 2D (^{13}C, ^{1}H) plane at the F_1 frequency of Ala-3 H^{β} from the 3D ^{13}C F_1-filtered, F_3-edited NOESY-HSQC spectrum (τ_m = 150 ms) of unlabeled N^{1-22}/^{15}N,^{13}C-labeled *boxB* recorded in D_2O (Zwahlen et al., 1997) showing correlations between Ala-3 H^{β} and proximal protons from *boxB* RNA. This spectrum was recorded with extensive aliasing in F_2; the proper ^{13}C reference scale is the one closest to the left of each set of peaks.

(D) Selected 2D (^{1}H, ^{1}H) slices at ^{13}C or ^{15}N frequencies of Trp-18 extracted from the 3D ^{13}C/^{15}N F_1-filtered, F_3-edited NOESY-HSQC spectrum (τ_m = 150 ms) of ^{15}N,^{13}C-labeled N^{1-22}/unlabeled *boxB* recorded in H_2O (Zwahlen et al., 1997). Correlations between protons of *boxB* RNA (F_1) and proximal N^{1-22} protons (F_3) are illustrated.

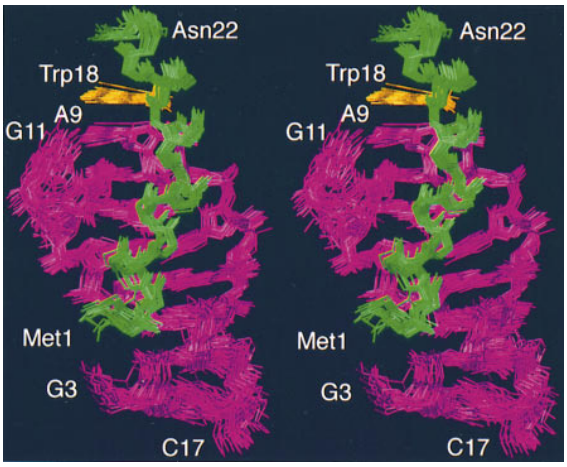

Figure 2. Stereo View Showing the Best-Fit Superposition of 24 Simulated Annealing Structures of N^{1-22}/*boxB* Complex to the Minimized Average Structure

Only the heavy atoms of native residues (capital letters in Figure 1) were used in the fit and shown here. The RNA is pink, the peptide backbone (N, C^{α}, C′, O) green, and the side chain of Trp-18 yellow.

al., 1997). An example of the quality of the NMR data obtained in these experiments is given in Figure 1. Many NOEs define the intermolecular interface, including important interactions between the Ala-3 methyl group and the bases and riboses of C4 and C5 in the *boxB* stem (Figure 1C) and between Trp-18 near the carboxyl terminus of N^{1-22} and A9 of the *boxB* loop (Figure 1D).

Three-dimensional structures were calculated using restrained molecular dynamics and simulated annealing starting from oligomers with randomized torsion angles (Allain et al., 1996). The 24 structures agreeing best with the restraints are superimposed in Figure 2. The level of precision obtained from this ensemble (Table 1) allows good structural definition of N^{1-22}, *boxB*, and their molecular interface.

The overall topology of the N^{1-22}/*boxB* complex, illustrated in Figure 3A, describes a novel mode of protein–nucleic acid recognition. N^{1-22} forms a bent α helix capped by a turn at its amino terminus. The *boxB* RNA forms a hairpin with a well-defined loop and an essentially regular A-form stem. Although the α-helical ARM of N interacts on the major groove face of *boxB*, the intermolecular interface involves only 5′ residues 4–10 and the 5′-phosphate of residue 11 of the RNA hairpin. The amino terminus of N^{1-22} interacts with the 5′ strand of the RNA stem through a number of important contacts, including a well-defined interaction between the Ala-3 methyl group and the hydrophobic pocket created by the bases and riboses of C4 and C5. The carboxy terminus of N^{1-22} interacts with the top of the *boxB* loop, with a well-defined stacking interaction formed between Trp-18 and A9, the second nucleotide of the GAAGA loop. These and many other interactions allow unambiguous orientation of the α helix and define the topology of this complex. In agreement with the NMR data, Chattopadhyay et al. (1995a) concluded that only 5′-phosphates of residues 4–11 are protected from ribonuclease cleavage by N.

Table 1. NMR Structural Statistics and Atomic Root-Mean-Square (Rms) Deviations[a]

Structural Statistics		
	<SA>	$\overline{(SA)}_r$
Experimental restraint violations		
NOE: number > 0.3 Å	0	1 (0.34 Å)
NOE: number > 0.2 Å	1.7 ± 0.7	5
Dihedral angle: number > 1°	0	0
$J_{HN\alpha}$: number > 1.2 Hz	0	0
$J_{HN\alpha}$: number > 1.0 Hz	0.7 ± 0.7	0
Rms deviation from NOE interproton distance restraints (Å)		
All (total of 1984)[b]	0.016 ± 0.001	0.022
Rms deviation from experimental dihedral angle restraints (°)		
All (total of 51)[c]	0.09 ± 0.02	0.11
Rms deviation from coupling constant restraints (Hz)		
$J_{HN\alpha}$ (total of 23)	0.42 ± 0.04	0.36
Rms deviations from ideal covalent geometry		
Bond lengths (Å)	0.0040 ± 0.0001	0.004
Bond angles (°)	0.859 ± 0.004	0.885
Impropers (°)	0.390 ± 0.005	0.416
Atomic Rms Deviations (Å)[d]		
	<SA> versus $\overline{SA}$	<SA> versus $\overline{(SA)}_r$
Peptide backbone	0.54 ± 0.17	0.61 ± 0.25
Peptide	0.97 ± 0.18	1.12 ± 0.30
RNA	0.75 ± 0.12	0.82 ± 0.14
GNRA[e]	0.50 ± 0.12	0.59 ± 0.15
Peptide–RNA complex	0.96 ± 0.12	1.09 ± 0.20

[a] The notation is described as follows: <SA> refers to the final 24 simulated annealing structures; $\overline{SA}$ refers to the average structure obtained by averaging heavy atom coordinates of residues 1–22 of the N^{1-22} peptide and residues 3–17 of *boxB* RNA of each structure from the set of 24 in <SA>; $\overline{(SA)}_r$ was obtained after restrained minimization of $\overline{SA}$.
[b] These include protein intraresidue (365), protein interresidue (545), RNA intraresidue (415), RNA interresidue (484), RNA base-pairing (40), and intermolecular protein–RNA (135) distance restraints.
[c] These include protein χ_1 (15) and RNA δ and γ (36) dihedral angle restraints.
[d] Only the heavy atoms of residues 1–22 of the N^{1-22} peptide and of residues 3–17 of *boxB* RNA were considered.
[e] GNRA refers to the GNRA-like structure formed by residues 1, 2, 3, and 5 of the *boxB* hairpin loop.

ARMs are characterized by a high density of arginine residues in a short sequence of 10–20 amino acids (Lazinski et al., 1989). Arginine side chains play an essential role in RNA recognition, forming intermolecular hydrophobic, hydrogen-bonding, and electrostatic interactions (Puglisi et al., 1995; Ye et al., 1995, 1996; Battiste et al., 1996). However, comparison of the N^{1-22}/*boxB* RNA complex (Figure 3) with complexes of ARMs from HIV-1 Rev (Figure 3B; Battiste et al., 1996) and BIV Tat (Figure 3C; Puglisi et al., 1995) illustrates the rich diversity in the modes of RNA–protein recognition by ARMs. Unlike the N^{1-22}/*boxB* interaction confined to one strand of the hairpin stem and loop, the regular α helix of HIV-1 Rev and the short β hairpin of BIV Tat contact both strands (one in yellow and one in blue in Figure 3) of a widened RNA major groove through interactions with the bases, sugars, and phosphates.

A Bent α Helix for the Arginine-Rich Motif of N

An atomic root-mean-square (rms) deviation from the average structure of 0.61 ± 0.25 Å (Figure 2 and Table 1) was obtained from the ensemble of NMR-derived structures for the backbone heavy atoms of N^{1-22}. Residues 4–21 form an α helix as established by their backbone ϕ and ψ angles, which fall in the α-helical range of the Ramachandran plot (Morris et al., 1992). This agrees with CD spectra indicating formation of an α helix involving 16–18 amino acids (Su et al., 1997a; Van Gilst et al., 1997). This α helix is capped at its amino terminus by intramolecular hydrophobic interactions between the Met-1 methyl group and the methyl and β- and γ-methylene moieties of Thr-5 and Arg-6, respectively, as described for the hydrophobic-staple motif (Muñoz et al., 1995).

The bend in the α helix of N^{1-22} results from a single localized perturbation at Arg-11 rather than additive effects distributed along the α helix. The average Arg-11 backbone angles (ϕ: −86° and ψ: −5°), which deviate slightly from the narrow distribution of angles for the other α-helical amino acids (ϕ: −48° to −72° and ψ: −19° to −53° for residues 4–10 and 12–20; ϕ: −83° and ψ: −21° for residue 21), create a bend of ~120° between two α-helical segments (residues 4–10 and 12–21) of N^{1-22}. Many residues (amino acids 5, 7–10, 14, and 18) cannot be replaced by the α helix–breaker proline without severely decreasing N binding (Tan and Frankel, 1995) or antitermination (Franklin, 1993). However, proline substitutions at Ala-12, Glu-13, and Gln-15 do not cause substantial functional defects (Franklin, 1993; Su et al., 1997a), indicating that α-helical disruptions in a narrow region of the ARM can still maintain the *boxB* interaction. Although proline replacement at Arg-11 is only partially functional (Franklin, 1993), and the Pro-11

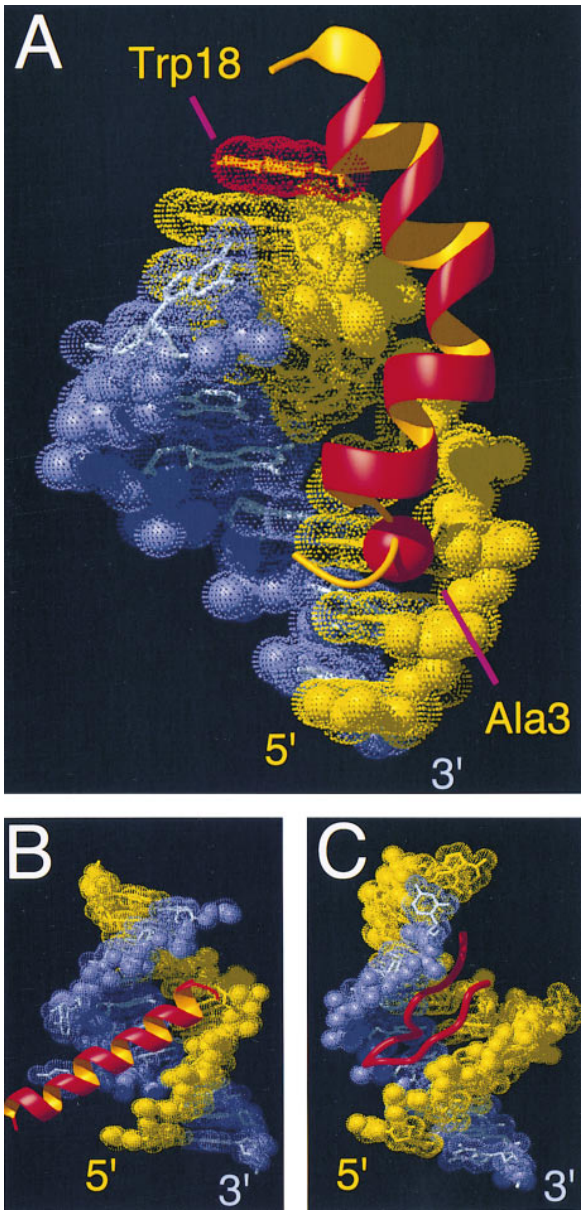

Figure 3. Unique Mode of RNA Binding for the Arginine-Rich Motif of N

(A) Minimized average structure of the N^{1-22}/*boxB* complex. *boxB* residues 3–10 are yellow and 11–17 blue. (B) NMR structure of the HIV-1 Rev peptide/RRE complex (pdb file 1ETF; Battiste et al., 1996). (C) NMR structure of the BIV Tat/TAR interaction (pdb file 1MNB; Puglisi et al., 1995). In (B) and (C), one RNA strand is yellow and the other blue. In (A–C) the RNA is depicted as a dotted surface; bonds connecting heavy atoms of RNA bases and riboses are drawn and the phosphate backbone is represented by spheres for C3′, C5′, O3′, O5′, O1P, O2P, and P atoms. The peptide backbone is represented by a smoothed C^{α} trace and by ribbons for α helices.

substitution might induce an α-helical disruption compatible with *boxB* binding, this substitution would also remove an important side chain interaction with the RNA, as described below.

A GNRA Fold in the *boxB* Pentaloop

From the ensemble of structures, a heavy-atom rms deviation from the average structure of 0.82 ± 0.14 Å (Figure 2 and Table 1) was calculated for the heavy atoms of *boxB* RNA. The hairpin stem consists of approximately a half turn of regular A-form helix. In fact, a heavy-atom rms deviation of only 1.3 Å was obtained after a best-fit superposition of the *boxB* stem (residues 3–7 and 13–17) with an ideal A-form helix.

Remarkably, the 5 nt *boxB* loop adopts a fold identical to that of the GAAA tetraloop (Figure 4; Heus and Pardi, 1991; Su et al., 1997b). The GAAA tetraloop is a member of the family of GNRA tetraloops (N = A, G, C, U; R = A, G) found in many important RNAs such as *E. coli* transcription terminators (d'Aubenton Carafa et al., 1990), ribosomal RNAs (Gutell et al., 1994), RNase P RNAs (James et al., 1988), group I and II introns (Michel and Westhof, 1990), and picornavirus ribosomal entry sites (López de Quinto and Martínez-Salas, 1997). Many GNRA loop structures have been determined by NMR spectroscopy (Heus and Pardi, 1991; Orita et al., 1993; Jucker and Pardi, 1995) and X-ray crystallography (Pley et al., 1994b; Scott et al., 1995; Cate et al., 1996). The GNRA tetraloop fold is characterized by formation of a sheared G-A base pair (type XI; Saenger, 1984) between the first and last loop nucleotides and by a large change in direction of the phosphate backbone between the first and second nucleotides. This leads, in the case of GAAA tetraloops, to sequential stacking of the second, third, and fourth purines on the 3′ stem. In accordance with these characteristics, nucleotides 1, 2, 3, and 5 of the GAAGA *boxB* hairpin loop form a GAAA fold (pink residues in Figure 4A), which excludes nucleotide 4 (white residue in Figure 4A). When the *boxB* GAAA fold was superimposed on the NMR structure of the GAAA tetraloop (Figure 4B; Heus and Pardi, 1991; Jucker et al., 1996), a heavy-atom rms deviation of 1.4 Å was obtained. Similar superpositions (not shown) with GAAA tetraloops from crystal structures of the hammerhead ribozyme (Pley et al., 1994b) and the P4-P6 domain of the group I intron (Cate et al., 1996) produced similar rms deviations (1.4–1.6 Å). In the *boxB* loop, the ribose and base of the fourth loop nucleotide G11 are extruded from the GAAA fold. G11 adopts a syn glycosidic angle that orients its guanine base to permit stacking with the ribose of the third nucleotide in the GAAA fold.

This is the first structural example of a 4 nt GNRA fold in a 5 nt loop. GNRA folds are likely to be found in other loops with more than four nucleotides. Indeed, this one nucleotide extrusion at position 4 of a 5 nt loop suggests that any loop of the form GNR(N_x)A (x = 1, 2, 3, . . .) has the potential to form a GNRA fold. Since a five-membered GAAAA loop can replace the essential GAAA tetraloop of the D5 domain in the group II intron without affecting the enzymatic activity of this ribozyme (Abramovitz and Pyle, 1997), this GAAAA loop mutant presumably also forms a GAAA tetraloop-like structure. Comparison of 16S and 23S ribosomal RNA sequences also indicates conserved hairpin loops, which fit a GNR(N)A or a GNR(NN)A consensus (Gutell et al., 1994).

Protein Recognition of a GNRA Fold

Formation of the GNRA fold of *boxB* is essential for specific recognition by N. Indeed, all mutations within the *boxB* loop that are not compatible with formation of a GNRA fold also abrogate N binding (Figures 4C and

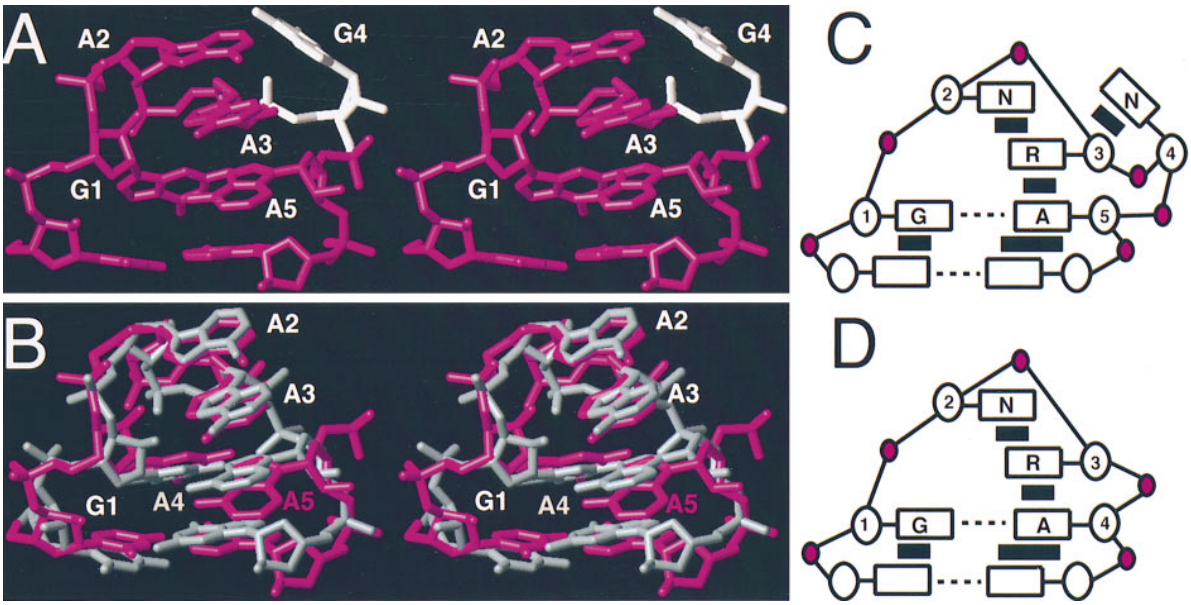

Figure 4. The *boxB* Pentaloop Adopts a GNRA Fold

(A) Stereo view of the *boxB* hairpin loop from the minimized average structure. The *boxB* pentaloop adopts a GNRA fold (pink) from which the fourth loop nucleotide (white) is extruded.

(B) Stereo view showing the best fit superposition between the GNRA folds in *boxB* (in pink; G1, A2, A3, and A5 from Figure 4A) and the GAAA tetraloop (in gray; mean structure generated in MOLMOL from the ten structures of the pdb file 1ZIP; Jucker et al., 1996). In (A) and (B), only bonds connecting heavy atoms are shown.

(C) Diagram showing important structural features of the *boxB* GAAGA loop (from the N^{1-22}/*boxB* complex) and summarizing the GNRNA base requirements for N binding (see text).

(D) Diagram showing important structural features of the GAAA tetraloop (Heus and Pardi, 1991) and summarizing the base requirements of the GNRA fold. In (C) and (D), the open circle, pink circle, and boxes denote ribose groups, phosphate moieties, and bases, respectively. Dotted lines and black boxes represent base-pairing and stacking interactions.

4D). In this regard, most binding studies have shown that the identity of residues 1, 3, and 5 of the *boxB* loop (Figure 4C) is important for N binding, whereas the identity of residues 2 and 4 is important for NusA binding but not for N binding (Chattopadhyay et al., 1995a; Mogridge et al., 1995; Tan and Frankel, 1995; Cilley and Williamson, 1997; Su et al., 1997b; Van Gilst et al., 1997). Mutation of the G at position 1 to A, C, U, or even inosine (I) or the A at position 5 to G, C, or U significantly reduces binding, indicating that the sheared G-A base pair of the GNRA fold is essential for N binding. Replacement of A at position 3 with C or U but not with G decreases the affinity for N, in agreement with the purine requirement at position 3 in the GNRA fold (Heus and Pardi, 1991). Although pyrimidine substitutions at loop position 2 may reduce the affinity for N under certain conditions (Su et al., 1997b), they do not affect binding in other conditions (Chattopadhyay et al., 1995a; Mogridge et al., 1995) or prevent recognition of the *boxB* GNRA fold by N (Su et al., 1997b). Any base substitution at position 4 maintains the high-affinity N-*boxB* RNA interaction. Thus, the *boxB* loop requirements for N binding fit the GNRNA consensus (Figure 4C), indicating that the GNRA fold formed by nucleotides 1, 2, 3, and 5 of the *boxB* loop is necessary for specific binding by N. Interestingly, Franklin (1993) was unable to identify amino acid substitutions in N that suppress the effects of *boxB* loop mutations. These changes would prevent either the formation of the GNRA structure specifically recognized by N or subsequent NusA binding as described below.

The absence of a specific base requirement for extruded nucleotide 4 suggested that N would still bind if this nucleotide were deleted. This was tested by a gel mobility shift assay (Mogridge et al., 1995) under native conditions (Figure 5). Addition of N to *nut* site–containing RNAs with a *boxB* GAAAA loop (the *nut*R site) or a mutant GAAA tetraloop produced mobility shifts indicative of high-affinity interactions between N and the *nut* site. On the basis of the N concentration required for the mobility shift, the affinity of N for the mutant GAAA tetraloop is reduced less than 2-fold relative to the wild-type GAAAA loop. Figure 5 also shows that the N-NusA-*nut* site supershift detected with the wild-type GAAAA loop sequence and not present when the fourth base in the loop is mutated to C (Mogridge et al., 1995) is, as expected, not observed with the *boxB* GAAA tetraloop mutant. These results illustrate the GNRA fold requirement for N binding and the involvement of the extruded purine in NusA binding.

Protein recognition of GNRA motifs likely represents a general class of RNA–protein interactions. Examples include recognition of a GAGA hairpin loop by the cytotoxic protein ricin (Glück et al., 1992), high-affinity binding of HIV-1 integrase to selected RNA ligands containing a GNRA tetraloop (Allen et al., 1995), and involvement of potential GNRA folds in ribosomal protein–RNA interactions (Brimacombe, 1995). The prevalence of GNRA tetraloops in various RNAs might therefore be explained in part by their importance in both RNA–RNA (Michel and Westhof, 1990; Murphy and Cech,

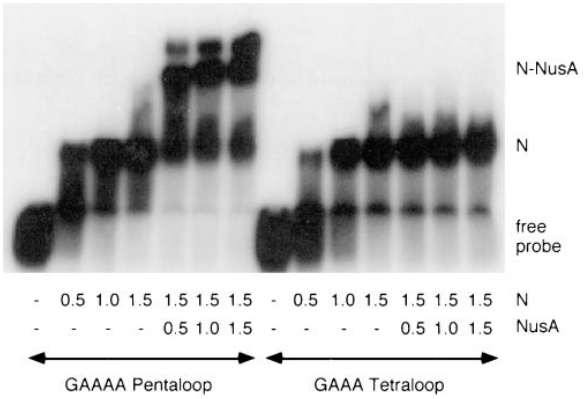

Figure 5. Gel Mobility Shift Assay Showing that Deletion of the Fourth Nucleotide in the *boxB* Loop Affects NusA Binding but Not N Binding

Reactions containing ^{32}P-labeled *nut* site RNAs with the λ *nut*R GAAAA pentaloop or a mutant GAAA tetraloop were submitted to electrophoresis on a 5% nondenaturing gel (Mogridge et al., 1995). The protein components of the gel-shifted RNA–protein complexes are indicated on the right. The concentrations of N and NusA are given in micromolars.

1994) and RNA–protein interactions (Glück et al., 1992; Allen et al., 1995; Brimacombe, 1995).

Description of N^{1-22}/*boxB* RNA Intermolecular Interactions

Figures 6A and 6B show contacts between side chains of N^{1-22} and *boxB* RNA (summarized in Figure 6C). RNA–peptide contacts involve one face of the α helix (residues 3–19) and only the 5′ residues (residues 4–10 and the 5′-phosphate of residue 11) of the RNA hairpin (Figures 3A and 6). The five arginines and two lysines of N^{1-22} create a positively charged surface on one face of the α helix that interacts with the negatively charged phosphodiester backbone of the *boxB* RNA (Figures 6A and 6C). All arginine guanidino groups and the amino group of Lys-14 but not Lys-19 are less than 5 Å from one or more phosphates on the RNA and should make significant electrostatic contributions to *boxB* binding. In agreement with these observations, single alanine replacements at Arg-6, Arg-7, Arg-8, Arg-10, Arg-11, and Lys-14 but not Lys-19 decrease in vitro binding by at least 20-fold (Su et al., 1997a) and also reduce antitermination activity (Franklin, 1993). However, shorter distances (<2.4 Å) between arginine guanidino groups and phosphates in some of the structures suggest stronger ionic interactions and possibly hydrogen bonds between Arg-6 and C4 and C5 and between Arg-8, Arg-10, and Arg-11 and G11, C6, and U7, respectively, as shown in Figure 6C. The structure of the N^{1-22}/*boxB* complex and the mutational data indicate that the five arginines and Lys-14 of N^{1-22} form a cluster of positive charges necessary for *boxB* recognition.

Hydrophobic interactions are also crucial for *boxB* recognition (Figure 6B). Indeed, the role of arginine and lysine residues is not restricted to ionic interactions since, in certain cases, the aliphatic portion of these side chains also contacts the RNA (Figures 6B and 6C). For example, the H^{β} and H^{γ} protons of Arg-7 are close to the C6 ribose and the base H5 and H6 of C6 and U7, while the H^{γ} and H^{δ} protons of Arg-8 are near the A10 ribose protons. Intermolecular hydrophobic contacts between the A9 ribose and Lys-14 and Lys-19 are also observed (Zwahlen et al., 1997). In addition, Ala-3, Gln-15, and Trp-18 are involved in hydrophobic interactions. The aliphatic side chain of Gln-15 lies between the A9 and A10 riboses (Zwahlen et al., 1997) and one of its side chain amino protons is close to the main-chain carbonyl of Arg-11, possibly forming a hydrogen bond that would stabilize the α-helical bend. These interactions may explain why glutamine is preferred at position 15, although many mutations at this site do not affect antitermination (Franklin, 1993; Su et al., 1997a).

The Ala-3 methyl group lies between the C4 and C5 riboses, closer to H5 and H6 of C5 than H5 and H6 of C4. Replacement of C5 by G reduces antitermination activity by less than 2-fold, and a full inversion of the four G-C base pairs in the stem ($^{5'\text{-GCCC}}_{3'\text{-CGGG}}$ to $^{5'\text{-CGGG}}_{3'\text{-GCCC}}$) causes only moderate reduction in activity (~6-fold; Chattopadhyay et al., 1995a), consistent with a nonspecific interaction between Ala-3 and the RNA stem. The contribution of Ala-3 to N binding is nevertheless significant, since any amino acid replacement except serine reduces *boxB* binding (Su et al., 1997a). Interestingly, only alanines and serines are found at corresponding positions in *boxB*-binding proteins from other lambdoid phages (N of P22 and ϕ21 and Nun of HK022; Lazinski et al., 1989). With the exception of Nun (Chattopadhyay et al., 1995b), which binds specifically to λ *boxB*, these N proteins interact with different *boxB* RNAs, suggesting that the Ala-3 nonspecific hydrophobic interaction of λ N is conserved among these lambdoid phages.

The Trp-18 indole ring stacks on the A9 base and the H^{β} protons of Trp-18 contact the A9 ribose protons. The Trp-18 stacking interaction is important for *boxB* loop recognition since all 19 possible mutations at position 18 reduce the affinity for *boxB* (Su et al., 1997a). Substitutions to aromatic residues (Tyr and Phe) are the least damaging, decreasing the affinity by less than 2-fold. Intermolecular stacking interactions between aromatic side chains and RNA bases are very common in RNA–protein complexes and may provide substantial enthalpic contributions to the overall binding energy (LeCuyer et al., 1996; Varani, 1997). Trp-18 stacking is specific to recognition of λ *boxB* since other lambdoid phage N proteins do not have an aromatic residue at the analogous position, whereas HKO22 Nun protein (which binds λ *boxB*) has tyrosine at this position (Lazinski et al., 1989).

The RNA–protein interface of the N^{1-22}/*boxB* complex is dominated by ionic interactions between positively charged side chains and the phosphate backbone (Figure 6A) and hydrophobic contacts between peptide side chains and RNA bases and riboses (Figure 6B). In contrast, only a few intermolecular hydrogen bonds can be proposed (Figures 6B and 6C). Side chain amino protons of Gln-4 and Arg-7 could interact with hydrogen-bond acceptors of U7 or G8 in the major groove. Intermolecular NOEs between the G8 imino proton and the Gln-4 amino protons indicate that these amino protons are near hydrogen-bond acceptors on the bases of U7 or G8. Only for Arg-7 in N^{1-22} is rotation hindered about

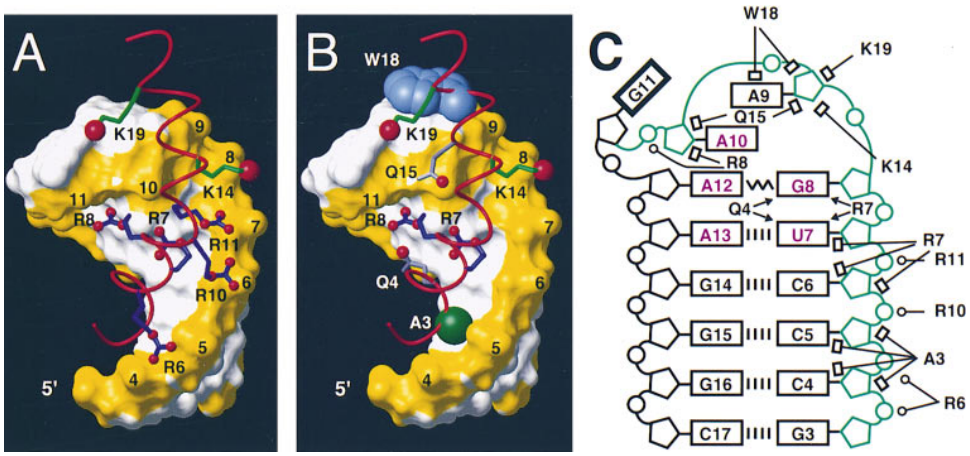

Figure 6. Summary of the N^{1-22}/*boxB* interactions

(A) Ionic interactions between positively charged side chains and the phosphate backbone.

(B) Hydrophobic and hydrogen-bonding interactions with *boxB*. In (A) and (B), the RNA is depicted as a white surface with the phosphate backbone yellow and the N^{1-22} C^{α} trace shown in red. Bonds connecting heavy atoms of selected side chains are shown in blue for ARG, green for LYS, and gray for GLN, and side chain nitrogens of these residues are shown as red spheres. The Ala-3 side chain is represented by a green sphere. Side chain heavy atoms of Trp-18 are blue spheres.

(C) Schematic representation of the *boxB* structure and RNA–protein interactions observed in the N^{1-22}/*boxB* complex. Open circles, pentagons, and rectangles represent phosphates, riboses, and bases, respectively. Riboses and phosphates drawn in green are protected by N from RNase cleavage, and bases labeled in pink are important for N binding (see text). Dashed lines represent Watson-Crick base pairing, and the zig-zag line denotes a sheared G-A base pair (type XI; Saenger, 1984). All ribose puckers are 3′-endo, and all glycosidic angles are anti except for G11, which is syn (bold rectangle). Open circles, open rectangles, and arrowheads depict ionic, hydrophobic, and hydrogen-bonding interactions, respectively.

both the N^{ϵ}-C^{ζ} and C^{ζ}-N^{η} bonds (P. L., et al., unpublished data), suggesting that the Arg-7 guanidino protons form strong and stable hydrogen bonds. A more precise description of the hydrogen-bonding interactions involving Gln-4 and Arg-7 awaits structural refinement, which may explain why glutamine is favored at position 4 and all tested mutations at position 7 abolish antitermination (Franklin, 1993).

Principal Determinants of Binding Specificity

In many DNA–protein interactions, sequence-specific binding results from hydrogen bonds formed between protein side chains and the exposed edges of the bases in the major groove of a regular DNA duplex (Steitz, 1990). Recognition of conformational features is also important in certain DNA–protein complexes (Steitz, 1990) but is absolutely central in RNA–protein recognition because of the unique shapes and charge distributions created by the diversity of RNA structures (Varani, 1997), as illustrated here for the N^{1-22}/*boxB* complex.

The specific determinants of N^{1-22}-*boxB* recognition include the shape and electrostatic surface of the GNRA fold and the U-A base pair closing the loop (Figure 6). Indeed, although multiple intermolecular contacts involve the *boxB* stem, the identity of the four G-C base pairs in the stem is not critical for antitermination provided that Watson-Crick base pairing is maintained (Chattopadhyay et al., 1995a). That the GNRA fold is necessary for specific N recognition is supported by the many intermolecular contacts with loop nucleotides 1, 2, and 3. Furthermore, hydrogen bonding of U7 with Gln-4 or Arg-7 (see above), as well as hydrophobic contacts between the aliphatic portion of Arg-7 and the H5 and H6 of U7, explain why changing the U7-A13 base pair to A7-U13 (Tan and Frankel, 1995) or G7-C13 (Cilley and Williamson, 1997) reduces N binding more than 20-fold. Finally, it is interesting that arginines that interact with the *boxB* stem (Arg-6 and Arg-10) are involved only in electrostatic interactions and can be replaced by lysines, whereas arginines that interact with the *boxB* loop and its U-A closing base pair (Arg-7, Arg-8, and Arg-11) cannot be replaced by lysines, suggesting that their contributions to *boxB* binding are more than simple charge effects (Franklin, 1993; Su et al., 1997a).

Comparison with the GAAA Tetraloop-Receptor Interaction of the Group I Intron

Previously, there was only detailed structural information describing RNA interactions, and not protein interactions, with GNRA folds (Pley et al., 1994a; Cate et al., 1996). Recognition of the GAAA tetraloop (the L5b loop) of the P4-P6 domain of the *Tetrahymena thermophila* group I intron by its intramolecular RNA receptor (the J6a/6b element; Cate et al., 1996) is fundamentally different from the recognition of the *boxB* loop by N. The ARM of N contacts the major groove face of the *boxB* stem and GAAGA loop mainly through ionic and hydrophobic interactions (Figure 7A), whereas the GAAA-tetraloop receptor of the P4-P6 domain recognizes the minor groove face of the GAAA tetraloop through multiple hydrogen-bonding contacts with bases and riboses of the loop (Figure 7B; Cate et al., 1996). Similarly, the minor groove face of a GAAA tetraloop contacts an RNA hairpin in an intermolecular crystal contact of a hammerhead ribozyme structure (Pley et al., 1994a). Indeed, the many available hydrogen-bond donors and acceptors

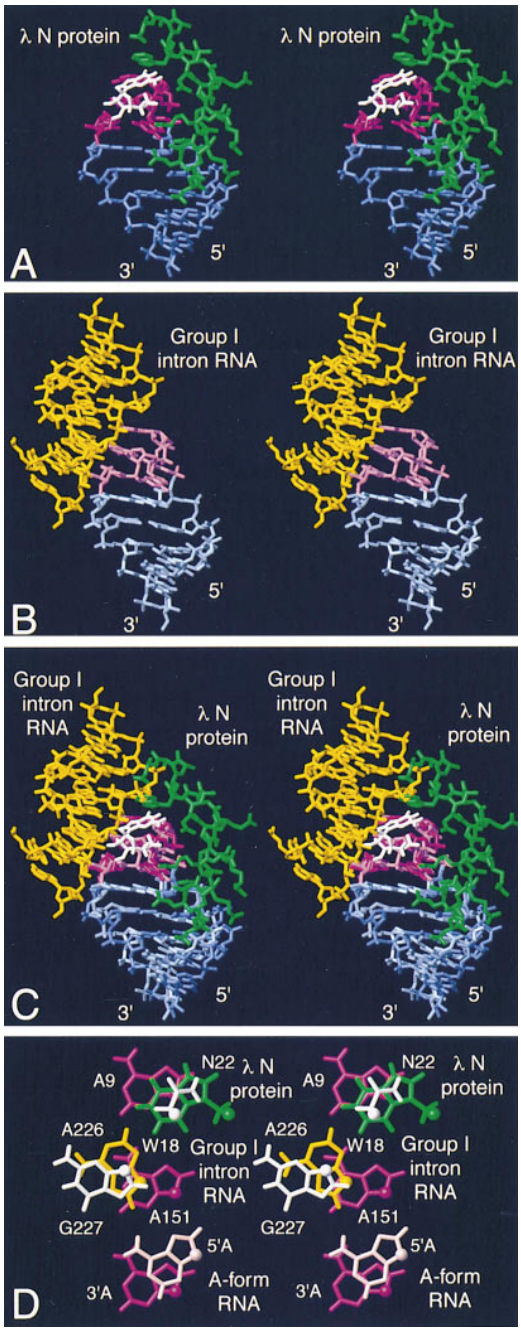

Figure 7. Stereo Views Showing Protein and RNA Binding to a GAAA Fold

(A) The λ N peptide (green) binds on the major groove face of the *boxB* loop (GAAA fold, pink; extruded G, white; stem residues, blue).

(B) The J6a/6b GAAA tetraloop receptor (yellow) binds on the minor groove face of the L5b GAAA tetraloop (GAAA fold, pale pink; stem residues, pale blue) in the *Tetrahymena thermophila* group I intron ribozyme (pdb file 1GID; Cate et al., 1996).

(C) Superposition of the structures shown in (A) and (B) optimizing the fit of their GAAA fold (pink residues).

(D) Intermolecular stacking against the second adenine of the GAAA fold (A9 in *boxB* and A151 in the group I intron ribozyme). Spheres represent N^9 (RNA) or C^α atoms (protein).

Only heavy atoms are shown in (A–C) and hydrogen atoms are added in (D).

on the minor groove face of a GNRA fold provide many potential intermolecular and tertiary contacts (Jucker and Pardi, 1995). However, the structure of the N^{1-22}/*boxB* complex presented here (Figure 7A) demonstrates that the major groove face of a GNRA fold can also be involved in macromolecular recognition.

The GAAA folds of the P4-P6 domain of the group I intron and the *boxB* loop are superimposed in Figure 7C to further contrast the diverse modes of recognition in these systems. This superposition also shows how a GAAAA pentaloop might interact with a GAAA tetraloop receptor, as postulated for a group II intron where the D5 GAAA tetraloop, which forms a tertiary interaction similar to that of the group I intron (Figure 7B), was mutated to a functional GAAAA loop (Abramovitz and Pyle, 1997).

Despite the global difference in recognition of a GAAA fold by N and the GAAA tetraloop receptor (Figure 7), residues from the GAAA tetraloop ligand (protein or RNA) interact in both cases with the base at the top of the GAAA fold and extend the 3 adenine stack, which stabilizes the tetraloop (Figure 7D). In the group I intron, A226 from one helical strand of the GAAA tetraloop receptor stacks on the second base of the GAAA loop (Figure 7D; Cate et al., 1996). In the N^{1-22}/*boxB* complex (Figure 7D), Trp-18 stacks on the second base of the *boxB* loop, acting as a pseudobase at the RNA–peptide interface, and the Asn-22 side chain contacts the other face of the tryptophan ring. Interaction of Asn-22 with the tryptophan ring is well defined by the NMR data and not an artifact of the short N^{1-22} peptide construct. Specific NMR chemical shifts and NOE contacts indicate that this interaction exists also in complexes of *boxB* RNA with N^{1-47} and full-length N (Mogridge et al., 1998). Since mutation of Asn-22 to Ala reduces *boxB* binding only 2-fold (Su et al., 1997a), Asn-22 plays a relatively minor role in stabilizing the complex. The similar stacking interaction between a GAAA fold and the N peptide or GAAA tetraloop receptor is a striking example of how important intermolecular interactions can be preserved by either protein or RNA ligands.

Future Directions: Studies of a *nut* Site/N Protein/NusA Ternary Complex

The NMR structure presented here reveals a new mode of RNA recognition for an ARM. N^{1-22} does not penetrate a distorted RNA major groove like the ARMs of BIV Tat and HIV-1 Rev. Instead, N^{1-22} adopts a bent α helix that binds exclusively to the 5′ strand of the *boxB* stem and the first three residues of the loop, recognizing primarily the shape and negatively charged surface on the major groove side of the *boxB* hairpin. The structure of the N^{1-22}/*boxB* RNA complex also provides new insights into RNA folding. Indeed, nucleotides 1, 2, 3, and 5 of the 5 nt *boxB* GAAGA loop adopt the GNRA fold commonly found in tetraloops, while the base of extruded nucleotide 4 stacks against the ribose at position 3 of the loop. Specific recognition of *boxB* RNA by the ARM of N depends on the formation of this GNRA fold and the Watson-Crick U-A base pair closing the *boxB* loop.

We showed previously that *E. coli* NusA, an essential component of the antitermination complex, interacts

with *E. coli* RNA polymerase and a central domain of N (Greenblatt and Li, 1981a, 1981b; Mogridge et al., 1998). Given that NusA interacts directly with N and RNA polymerase, why is a *nut* site needed for antitermination? Evidently, the *boxB* hairpin supplements the direct N/NusA protein–protein interaction that is important for antitermination. Although NusA does not bind directly to *nut* site RNA, pyrimidine substitutions of nucleotides 2 and 4 in the *boxB* loop abolish binding of NusA to the *nut* site/N protein complex (Mogridge et al., 1995). Moreover, as shown here, deletion of loop nucleotide 4 does not affect N binding but prevents formation of the *nut* site/N/NusA complex. Loop nucleotides 2 and 4 are well defined in the ensemble of NMR structures and partially exposed for NusA recognition. However, NusA's interaction with the *nut* site/N complex might also involve the accessible minor groove face of the GNRA fold, which is rich in hydrogen-bond donors and acceptors. The N peptide/*boxB* RNA structure presented here demonstrates that a GNRA fold with an additional nucleotide can act as a module to link two proteins, N and NusA, and initiate the assembly of a ribonucleoprotein complex. Such GNRA-based scaffolds could also in principle specifically bring together two RNAs or a protein and an RNA. In this regard, the *boxB* hairpin has an analogous function to protein modules like the SH2, SH3, PTB, PDZ, and WW domains, which organize signaling complexes (Pawson and Scott, 1997). Future structural studies on ternary complexes containing the *nut* site, N, and NusA will establish exactly how these components are assembled to form a functional antitermination complex.

Experimental Procedures

NMR Spectroscopy

NMR samples were prepared as described previously (Zwahlen et al., 1997; Mogridge et al., 1998) and sample concentrations were 2–3 mM. All NMR experiments were performed at 25°C on Varian 500 and 600 MHz spectrometers. Complete assignments and structural restraints were obtained from a large number of NMR experiments (data not shown). Spectra were processed with the NMRPipe/NMRDraw programs (Delaglio et al., 1995) supplemented with in-house written routines and analyzed with NMRView (Johnson and Blevins, 1994).

Intramolecular interproton distance restraints were obtained from 2D, 3D, and 4D NOESY spectra recorded using mixing times of 50 and 150 ms, and intermolecular restraints were established from filtered spectra recorded with mixing times of 150 ms (Zwahlen et al., 1997). All NOE intensities were estimated semiquantitatively on the basis of cross-peak intensities from spectra collected with a 50 ms mixing time. Because of strong NMR evidence for the formation of A-U and G-C Watson-Crick base pairs in the *boxB* stem (Wijmenga et al., 1993), canonical distance restraints were employed to define the hydrogen-bonding pattern and the planarity of these base pairs (Saenger, 1984).

J-coupling restraints from the HNH$^{\alpha}$ experiment (Vuister and Bax, 1993) were input directly into X-PLOR, whereas protein χ_1 angles obtained from 3D HNCOC$^{\gamma}$ and 3D HNC$^{\gamma}$ experiments (Konrat et al., 1997) along with some RNA δ and γ angles were input as dihedral angle restraints. RNA δ torsion angles, derived from NOE intensities, were set to define the 3′-endo ($\delta = 86° \pm 10°$) sugar pucker conformation for nucleotides where (H6/H8 - H2′ intraresidue NOE) $<$ (H6/H8 - H3′ intraresidue NOE) and where the χ angle is anti (Wijmenga et al., 1993). Gauche+ ($\gamma = 50° \pm 15°$) or *trans* ($\gamma = 180° \pm 40°$) γ conformations were derived from intraresidue H6/H8 to H5′ and H6/H8 to H5′′ NOE intensities for nucleotides with a 3′-endo sugar pucker and an anti χ angle (Wijmenga et al., 1993).

Structure Calculation

Structure calculations were performed with the program X-PLOR (version 3.851; Brünger, 1992) using the experimental restraints described in Table 1 along with restraints to maintain RNA and protein covalent structure and stereochemistry. A van der Waals repulsive term was included in the potential, but electrostatic contributions were not. Standard X-PLOR topology and parameter files were used with the exception that the ribose bond angles were modified to those given by Saenger (1984) and the proper sugar chirality was enforced (Schultze and Feigon, 1997). Our structure calculation protocol was derived from the one kindly provided by Allain et al. (1996) and uses restrained molecular dynamics and simulated annealing starting from random ϕ and ψ torsion angles for N$^{1\text{-}22}$ and from random backbone (α, β, γ, δ, ϵ, and ζ) and χ angles for *boxB*. Refined structures with no NOE violations greater than 0.3 Å, no torsion angle violations greater than 1°, and no J violations greater than 1.2 Hz were accepted and retained for analysis. Using this protocol 80% (24/30) of all calculated structures satisfied these criteria. The average structure ($\overline{\text{SA}}$) was computed with X-PLOR from the 24 accepted structures ($<\overline{\text{SA}}>$) and minimized ($(\overline{\text{SA}})_r$) using 1000 steps of restrained energy minimization. The program MOLMOL (Koradi et al., 1996) was used for visualization of the structures, and both MOLMOL and X-PLOR were used for analysis.

Acknowledgments

We thank V. Sklenář for providing his improved HCN experiments prior to publication, F. Allain and G. Varani for their X-PLOR protocol, L. Serrano for useful discussions, C. M. Kay and L. Hicks for hydrodynamic studies of the N$^{1\text{-}22}$/*boxB* complex, A. Pardi for critical reading of the manuscript, and R. Muhandiram for assistance in recording NMR experiments. This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada (L. E. K.) and the Medical Research Council of Canada (L. E. K. and J. G.). P. L. is a Terry Fox Research Fellow of the National Cancer Institute of Canada supported with funds provided by the Terry Fox Run. J. G. is a Distinguished Scientist of the Medical Research Council of Canada. J. G. and L. E. K. are International Research Scholars of the Howard Hughes Medical Institute.

Received December 5, 1997; revised March 11, 1998.

References

Abramovitz, D.L., and Pyle, A.M. (1997). Remarkable morphological variability of a common RNA folding motif: the GNRA tetraloop-receptor interaction. J. Mol. Biol. *266*, 493–506.

Allain, F.H.-T., Gubser, C.C., Howe, P.W.A., Nagai, K., Neuhaus, D., and Varani, G. (1996). Specificity of ribonucleoprotein interaction determined by RNA folding during complex formation. Nature *380*, 646–650.

Allen, P., Worland, S., and Gold, L. (1995). Isolation of high-affinity RNA ligands to HIV-1 integrase from a random pool. Virology *209*, 327–336.

Battiste, J.L., Mao, H., Rao, S.N., Tan, R., Muhandiram, D.R., Kay, L.E., Frankel, A.D., and Williamson, J.R. (1996). α helix-RNA major groove recognition in a HIV-1 Rev peptide-RRE RNA complex. Science *273*, 1547–1551.

Brimacombe, R. (1995). The structure of ribosomal RNA: a three-dimensional jigsaw puzzle. Eur. J. Biochem. *230*, 365–383.

Brünger, A.T. (1992). X-PLOR Version 3.1: A System for X-Ray Crystallography and NMR (New Haven, CT: Yale University Press).

Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R., and Doudna, J.A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. Science *273*, 1678–1685.

Chattopadhyay, S., Garcia-Mena, J., DeVito, J., Wolska, K., and Das, A. (1995a). Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage λ. Proc. Natl. Acad. Sci. USA *92*, 4061–4065.

Chattopadhyay, S., Hung, S.C., Stuart, A.C., Palmer, A.G., III, Garcia-Mena, J., Das, A., and Gottesman, M.E. (1995b). Interaction between the phage HK022 Nun protein and the *nut* RNA of phage λ. Proc. Natl. Acad. Sci. USA *92*, 12131–12135.

Cilley, C.D., and Williamson, J.R. (1997). Analysis of bacteriophage N protein and peptide binding to *boxB* RNA using polyacrylamide gel coelectrophoresis (PACE). RNA *3*, 57–67.

d'Aubenton Carafa, Y., Brody, E., and Thermes, C. (1990). Prediction of Rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. J. Mol. Biol. *216*, 835–858.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277–293.

Doelling, J.H., and Franklin, N.C. (1989). Effects of all single base substitutions in the loop of *boxB* on antitermination of transcription by bacteriophage λ's N protein. Nucleic Acids Res. *17*, 5565–5577.

Franklin, N.C. (1984). Conservation of genome form but not sequence in the transcription antitermination determinants of bacteriophages λ, φ21 and P22. J. Mol. Biol. *181*, 75–84.

Franklin, N.C. (1993). Clustered arginine residues of bacteriophage λ N protein are essential to antitermination of transcription, but their locale cannot compensate for *boxB* loop defects. J. Mol. Biol. *231*, 343–360.

Friedman, D.I., and Court, D.L. (1995). Transcription antitermination: the λ paradigm updated. Mol. Microbiol. *18*, 191–200.

Glück, A., Endo, Y., and Wool, I.G. (1992). Ribosomal RNA identity elements for ricin A-chain recognition and catalysis. Analysis with tetraloop mutants. J. Mol. Biol. *226*, 411–424.

Greenblatt, J., and Li, J. (1981a). Interaction of the sigma factor and the *nusA* gene protein of *E. coli* with RNA polymerase in the initiation-termination cycle of transcription. Cell *24*, 421–428.

Greenblatt, J., and Li, J. (1981b). The *nusA* gene protein of Escherichia coli: its identification and a demonstration that it interacts with the gene *N* transcription antitermination protein of bacteriophage lambda. J. Mol. Biol. *147*, 11–23.

Greenblatt, J., Nodwell, J.R., and Mason, S.W. (1993). Transcriptional antitermination. Nature *364*, 401–406.

Gutell, R.R., Larsen, N., and Woese, C.R. (1994). Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. Microbiol. Rev. *58*, 10–26.

Heus, H.A., and Pardi, A. (1991). Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. Science *253*, 191–194.

James, B.D., Olsen, G.J., Liu, J., and Pace, N.R. (1988). The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. Cell *52*, 19–26.

Johnson, B.A., and Blevins, R.A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. J. Biomol. NMR *4*, 603–614.

Jucker, F.M., and Pardi, A. (1995). GNRA tetraloops make a U-turn. RNA *1*, 219–222.

Jucker, F.M., Heus, H.A., Yip, P.F., Moors, E.H.M., and Pardi, A. (1996). A network of heterogenous hydrogen bonds in GNRA tetraloops. J. Mol. Biol. *264*, 968–980.

Konrat, R., Muhandiram, D.R., Farrow, N.A., and Kay, L.E. (1997). Pulse schemes for the measurement of $^{3}J_{C'C\gamma}$ and $^{3}J_{NC\gamma}$ scalar couplings in ^{15}N, ^{13}C uniformly labeled proteins. J. Biomol. NMR *9*, 409–422.

Koradi, R., Billeter, M., and Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graphics *14*, 51–55.

Lazinski, D., Grzadzielska, E., and Das, A. (1989). Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. Cell *59*, 207–218.

LeCuyer, K.A., Behlen, L.S., and Uhlenbeck, O.C. (1996). Mutagenesis of a stacking contact in the MS2 coat protein–RNA complex. EMBO J. *15*, 6847–6853.

López de Quinto, S., and Martínez-Salas, E. (1997). Conserved structural motifs located in distal loops of aphtovirus internal ribosome entry site domain 3 are required for internal initiation of translation. J. Virol. *71*, 4171–4175.

Michel, F., and Westhof, E. (1990). Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J. Mol. Biol. *216*, 585–610.

Mogridge, J., Mah, T.-F., and Greenblatt, J. (1995). A protein–RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the λ N protein. Genes Dev. *9*, 2831–2844.

Mogridge, J., Legault, P., Li, J., Van Oene, M.D., Kay, L.E., and Greenblatt, J. (1998). Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage λ N protein. Mol. Cell *1*, 265–275.

Morris, A.L., MacArthur, M.W., Hutchinson, E.G., and Thornton, J.M. (1992). Stereochemical quality of protein structure coordinates. Proteins *12*, 345–364.

Muñoz, V., Blanco, F.J., and Serrano, L. (1995). The hydrophobic-staple motif and a role for loop-residues in α-helix stability and protein folding. Nat. Struct. Biol. *2*, 380–385.

Murphy, F.L., and Cech, T.R. (1994). GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. J. Mol. Biol. *236*, 49–63.

Olson, E.R., Tomich, C.-S., and Friedman, D.I. (1984). The *nusA* recognition site. Alteration in its sequence or position relative to upstream translation interferes with the action of the N antitermination function of phage lambda. J. Mol. Biol. *180*, 1053–1063.

Orita, M., Nishikawa, F., Shimayama, T., Taira, K., Endo, Y., and Nishikawa, S. (1993). High-resolution NMR study of a synthetic oligoribonuclotide with a tetranucleotide GAGA loop that is a substrate for the cytotoxic protein, ricin. Nucleic Acids Res. *21*, 5670–5678.

Pawson, T., and Scott, J.D. (1997). Signaling through scaffold, anchoring, and adaptor proteins. Science *278*, 2075–2080.

Pley, H.W., Flaherty, K.M., and McKay, D.B. (1994a). Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. Nature *372*, 111–113.

Pley, H.W., Flaherty, K.M., and McKay, D.B. (1994b). Three-dimensional structure of a hammerhead ribozyme. Science *372*, 68–74.

Puglisi, J.D., Chen, L., Blanchard, S., and Frankel, A.D. (1995). Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. Science *270*, 1200–1203.

Saenger, W. (1984). Principles of Nucleic Acid Structure (New York: Springer-Verlag).

Salstrom, J.S., and Szybalski, W. (1978). Coliphage λ *nutL*$^{-}$: a unique class of mutants defective in the site of gene *N* product utilization for antitermination of leftward transcription. J. Mol. Biol. *124*, 195–221.

Schultze, P., and Feigon, J. (1997). Chirality errors in nucleic acid structures. Nature *387*, 668.

Scott, W.G., Finch, J.T., and Klug, A. (1995). The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. Cell *81*, 991–1002.

Steitz, T.A. (1990). Structural studies of protein–nucleic acid interaction: the sources of sequence-specific binding. Quart. Rev. Biophys. *23*, 205–280.

Su, L., Radek, J.T., Hallenga, K., Hermanto, P., Chan, G., Labeots, L.A., and Weiss, M.A. (1997a). RNA recognition by a bent α-helix regulates transcriptional antitermination in phage λ. Biochemistry *36*, 12722–12732.

Su, L., Radek, J.T., Labeots, L.A., Hallenga, K., Hermanto, P., Chen, H., Nakagawa, S., Zhao, M., Kates, S., and Weiss, M.A. (1997b). An RNA enhancer in a phage transcriptional antitermination complex functions as a structural switch. Genes Dev. *11*, 2214–2226.

Tan, R., and Frankel, A.D. (1995). Structural variety of arginine-rich RNA-binding peptides. Proc. Natl. Acad. Sci. USA *92*, 5282–5286.

Van Gilst, M.R., Rees, W.A., Das, A., and von Hippel, P.H. (1997). Complexes of N antitermination protein of phage λ with specific and nonspecific RNA target sites on the nascent transcript. Biochemistry *36*, 1514–1524.

Varani, G. (1997). RNA–protein intermolecular recognition. Acc. Chem. Res. *30*, 189–195.

Vuister, G.W., and Bax, A. (1993). Quantitative J correlation: a new approach for measuring homonuclear three-bond J($H^{N}H^{\alpha}$) coupling constants in ^{15}N-enriched proteins. JACS *115*, 7772–7777.

Wijmenga, S.S., Mooren, M.M., and Hilbers, C.W. (1993). NMR of nucleic acids; from spectrum to structure. In NMR of Macromolecules. A Practical Approach, G.C.K. Roberts, ed. (New York: IRL Press), pp. 217–288.

Ye, X., Gorin, A., Ellington, A.D., and Patel, D.J. (1996). Deep penetration of an α helix into a widened RNA major groove in the HIV-1 Rev peptide–RNA aptamer complex. Nat. Struct. Biol. *3*, 1026–1033.

Ye, X., Kumar, R.A., and Patel, D.J. (1995). Molecular recognition in the bovine immunodeficiency virus Tat peptide–TAR RNA complex. Chem. Biol. *270*, 827–840.

Zwahlen, C., Legault, P., Vincent, S.J.F., Greenblatt, J., Konrat, R., and Kay, L.E. (1997). Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: application to a bacteriophage λ N-peptide/*boxB* RNA complex. JACS *119*, 6711–6721.