J. Mol. Biol. (1997) **274**, 101–113

JMB

The RNA Binding Domain of Ribosomal Protein L11: Three-dimensional Structure of the RNA-bound Form of the Protein and its Interaction with 23 S rRNA

Andrew P. Hinck[1], Michelle A. Markus[1], Shengrong Huang[1] Stephan Grzesiek[2], Irina Kustonovich[3], David E. Draper[4] and Dennis A. Torchia[1]*

[1]*Molecular Structural Biology Core, National Institute of Dental Research, National Institutes of Health, Bethesda MD, 20892-4326, USA*

[2]*Forschungszentrum Juelich Juelich, Germany*

[3]*Hebrew University, Jerusalem Israel*

[4]*Department of Chemistry The Johns Hopkins University Baltimore, MD, USA*

**Corresponding author*

The three-dimensional solution structure has been determined by NMR spectroscopy of the 75 residue C-terminal domain of ribosomal protein L11 (L11-C76) in its RNA-bound state. L11-C76 recognizes and binds tightly to a highly conserved 58 nucleotide domain of 23 S ribosomal RNA, whose secondary structure consists of three helical stems and a central junction loop. The NMR data reveal that the conserved structural core of the protein, which consists of a bundle of three α-helices and a two-stranded parallel β-sheet four residues in length, is nearly the same as the solution structure determined for the non-liganded form of the protein. There are however, substantial chemical shift perturbations which accompany RNA binding, the largest of which map onto an extended loop which bridges the C-terminal end of α-helix 1 and the first strand of parallel β-sheet. Substantial shift perturbations are also observed in the N-terminal end of α-helix 1, the intervening loop that bridges helices 2 and 3, and α-helix 3. The four contact regions identified by the shift perturbation data also displayed protein-RNA NOEs, as identified by isotope-filtered three-dimensional NOE spectroscopy. The shift perturbation and NOE data not only implicate helix 3 as playing an important role in RNA binding, but also indicate that regions flanking helix 3 are involved as well. Loop 1 is of particular interest as it was found to be flexible and disordered for L11-C76 free in solution, but not in the RNA-bound form of the protein, where it appears rigid and adopts a specific conformation as a result of its direct contact to RNA.

Keywords: L11, heteronuclear NMR; protein-RNA; 23 S ribosomal RNA; ribosome

Introduction

The C-terminal domain of ribosomal protein L11 interacts specifically with a 58 nucleotide domain of large subunit 23 S ribosomal RNA (Xing & Draper, 1996). Both the protein and its cognate RNA are highly conserved: each is present among archaebacterial, prokaryotic, and eukaryotic organisms and both are known to be components of the large subunit GTPase center. Reconstituted ribosomes lacking native L11 synthesize protein twofold more slowly than normal ribosomes, and are defective in elongation factor-G (EF-G) dependent GTP hydrolysis (Stark & Cundliffe, 1979) and release factor-1 dependent termination (Tate *et al.*, 1984). Native L11 also forms part of the binding

Abbreviations used: L11-C76, recombinant protein produced in *E. coli*, whose amino acid sequence corresponds to the C-terminal 75 residues of *Bacillus stearothermophilus* ribosomal protein L11, with an additional N-terminal methionine residue; *f*L11-C76, designates the form of L11-C76 free in solution; *b*L11-C76, designates RNA-bound form L11-C76 (1:1 complex with its target RNA of 58 nucleotides); nt, nucleotide; RMSD, root mean square deviation; ppm, parts per million; NOE, nuclear Overhauser enhancement; EF, elongation factor; 2D, 3D, two and three-dimensional, respectively; HSQC, Heteronuclear single quantum correlation; NOESY, NOE spectroscopy.

0022–2836/97/460101–13 $25.00/0/mb971379

site for the thiazole family of antibiotics. Such antibiotics bind cooperatively with L11 and inhibit ribosome function by interfering with the interaction of EF-G·GTP and EF-Tu·aminoacyl tRNA·GTP complexes with the large subunit (Thompson *et al.*, 1979).

The specificity of L11 binding for a limited 58 nucleotide domain of 23 S rRNA (nt 1051 to 1108, *Escherichia coli* numbering) was first identified by ribonuclease T_1 digests of L11 bound to naked 23 S rRNA (Schmidt *et al.*, 1981). Subsequent chemical protection studies of the protein-RNA complex (Egebjerg *et al.*, 1990), site-specific mutagenesis studies of the RNA (Ryan & Draper, 1991; Ryan *et al.*, 1991), and analyses of the thermodynamics of unfolding of both the RNA by itself (Laing & Draper, 1994) and the L11-RNA complex (Xing & Draper, 1996) strongly support the idea that the RNA adopts a specific tertiary structure. The tertiary structure of the RNA has been shown to be stabilized by mono- (Draper *et al.*, 1995; Wang *et al.*, 1993) and divalent cations (Laing *et al.*, 1994; Lu & Draper, 1994), ribosomal protein L11 (Xing & Draper, 1996), and the antibiotic thiostrepton (Draper *et al.*, 1995; Ryan *et al.*, 1991; Xing & Draper, 1996).

Limited proteolysis experiments (Xing & Draper, 1996) have shown that native L11 consists of two functional domains: the N-terminal domain is responsible for the cooperative binding of L11 and thiostrepton to RNA, whereas the C-terminal domain is the RNA binding domain. The disassociation constant measured for the C-terminal domain of the protein complexed to nt 1029 to 1126 of 23 S rRNA, as determined by quantitative filter binding assays, is ~ 0.1 μM (Xing & Draper, 1996). Two recent independent NMR studies (Markus *et al.*, 1997; Xing *et al.*, 1997) have shown that the C-terminal 75 residues of recombinant *Bacillus stearothermophilus* L11 (which, along with an N-terminal methionine, is designated L11-C76) folds into a compact structure consisting of a bundle of three α-helices and a short two-stranded parallel β-sheet. Although L11 lacks any detectable sequence homology with other known nucleic acid binding proteins, it was noted that the arrangement of the three helices of L11-C76 free in solution (*f*L11-C76) is strikingly similar to that found in the homeodomain family of eukaryotic transcription factors (Xing *et al.*, 1997). Moreover, several conserved residues in helix 3 which appear to be required for rRNA recognition, align with con-

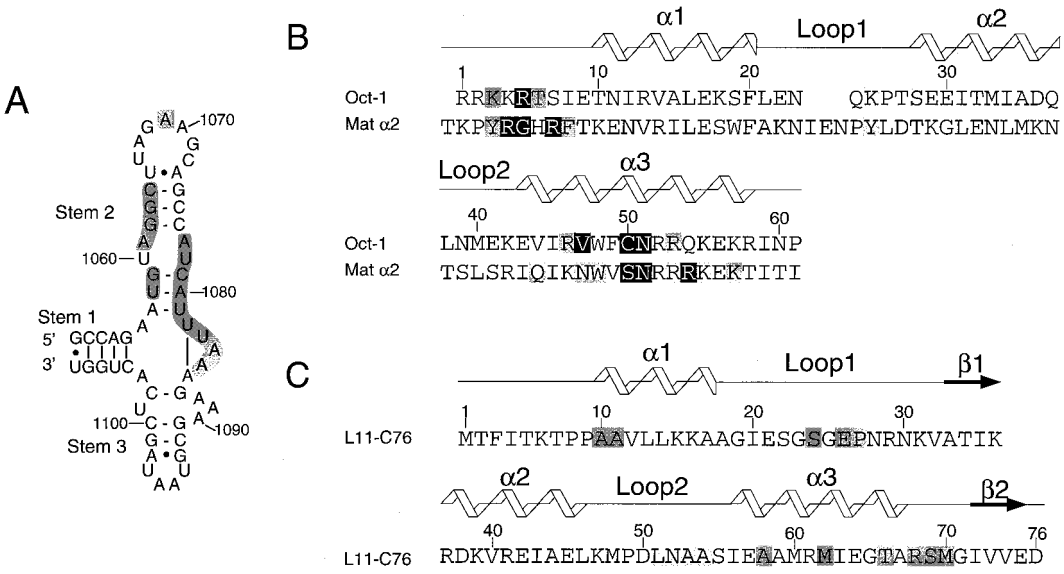


Figure 1. Ribosomal RNA and L11 sequences used for NMR studies and a comparison of L11-RNA and homeodomain-DNA contact sites. (a) A 58 nucleotide fragment of *E. coli* 23 S rRNA, modified at position 1061 (*E. coli* numbering) by a U to A substitution. Bases which are protected by native L11 in hydroxyl radical footprinting experiments are indicated by gray shading (Rosendahl & Douthwaite, 1993). (b) A primary sequence alignment of the Oct-1 (Klemm *et al.*, 1994) and MAT-α2 (Li *et al.*, 1995) homeodomains. Homeodomain residues are numbered according the convention previously established (Li *et al.*, 1995). The helical boundaries and amino acid residues which contact the DNA, are those reported for the Oct-1 (Klemm *et al.*, 1994) and MAT-α2 (Li *et al.*, 1995) homeodomain-DNA complexes, respectively. The three helical regions are indicated symbolically above the amino acid sequences, whereas the protein-DNA contact sites are identified by residue shading. Residues shaded black correspond to those which engage in base-specific contacts, whereas those shaded gray correspond to those which exhibit either phosphate or ribose contacts. (c) Primary sequence, deduced secondary structure,and sites of protein-RNA contacts for the C-terminal fragment (75 residues plus N-terminal initiator methionine) of *Bacillus stearothermophilus* L11. The secondary structure is indicated schematically above the amino acid sequence, whereas the protein-RNA contact sites are indicated by residue shading. The latter were identified on the basis of filtered NOE experiments, as described in Materials and Methods.

served residues in homeodomain helix 3 which engage in base-specific DNA contacts (Xing *et al.*, 1997). In order to provide more detailed structural information as to the extent of protein-RNA interactions, and to investigate the effects of RNA binding on protein structure and dynamics, we have used NMR spectroscopy to study the 27 kDa complex of L11-C76 with its 58 nucleotide binding site in 23 S rRNA (Figure 1). Herein, we report nearly complete backbone and side-chain sequential assignments, the three-dimensional structure for the protein component of the L11-RNA complex, and the sites on L11-C76 that interact with the 58 nt RNA target.

Results and Discussion

Assignments and secondary structure

Triple-resonance methodology was used to obtain nearly complete sequential backbone and side-chain resonance assignments for *b*L11-C76 using samples in which the protein component of the complex was labeled with either ^{15}N or $^{15}N/^{13}C$, and the RNA component was unlabeled. Upon completion of the assignments, all amide ^{1}H-^{15}N correlations, with the exception of the N-terminal residues M1, T2, and F3, were identified in the 2D $^{1}H/^{15}N$ HSQC spectrum (Figure 2). The N-terminal methionine residue is missing due to its post-translational cleavage from the recombinant protein product, whereas the amides of T2 and F3 are missing owing to rapid exchange of their amide protons with solvent. The former conclusion was based on the observation that (a) only four of the five methionine $H^{\varepsilon}/C^{\varepsilon}$ resonances were identified in a 2D constant time HSQC spectrum (Vuister & Bax, 1992), and (b) that each of these four methyl peaks was sequentially assigned to one of the four internal methionine residues. The latter conclusion follows from the observation that we were able to assign a number of backbone and side-chain ^{1}H and ^{13}C side-chain resonances for residues T2 and F3. These reside in the highly flexible N-terminal sequence, as assessed by backbone ^{15}N relaxation measurements (see Figure 5(c), below), and this accounts for their rapid amide exchange. Side-chain ^{1}H, ^{13}C, and ^{15}N assignments were generally complete, although there were a few exceptions involving Lys $C^{\varepsilon}/H^{\varepsilon}$, Ile $C^{\gamma 1}/H^{\gamma 1}$, and Leu C^{γ}/H^{γ} resonances.

The secondary structure of the protein was deduced from the pattern of short and medium range NOE data, $^{3}J_{HNH\alpha}$ coupling constants, and hydrogen-exchange rate constants, as summarized in Figure 3(a). A consensus of all indices indicates the presence of three α-helical regions and two β-strands. The former was supported by small (<6 Hz) $^{3}J_{HNH\alpha}$ couplings, strong $d_{N\text{-}N}(i,i+1)$, weak $d_{\alpha\text{-}N}(i,i+1)$, observable $d_{\alpha\text{-}N}(i,i+3)$ NOE connectivities, and contiguous stretches of slowly exchanging amide protons. The latter was supported by large (>8 Hz) $^{3}J_{HNH\alpha}$ couplings, and weak or medium $d_{N\text{-}N}(i,i+1)$ and strong $d_{\alpha\text{-}N}(i,i+1)$ NOE connectivities. Evidence supporting pairing of the two strands in a parallel fashion, as shown in Figure 3(b), includes several H^{N}-H^{α} and H^{N}-H^{N} long range NOEs (A33 H^{N}–I72 H^{α}, V73 H^{N}–T34 H^{α}, E75 H^{N}–K36 H^{α}, and V73 H^{N}–I35 H^{N}) and slowly exchanging amide protons (I35, V73, and E75). The regions of secondary structure, as identified by the chemical shift index (Wishart & Sykes, 1994), were in complete agreement with the secondary structure identified using the criteria described above. There were, however, three additional short β-strands predicted by the index, I4-T7, I20-G23, and A53-S55, which are not generally supported by the *J*-coupling, hydrogen

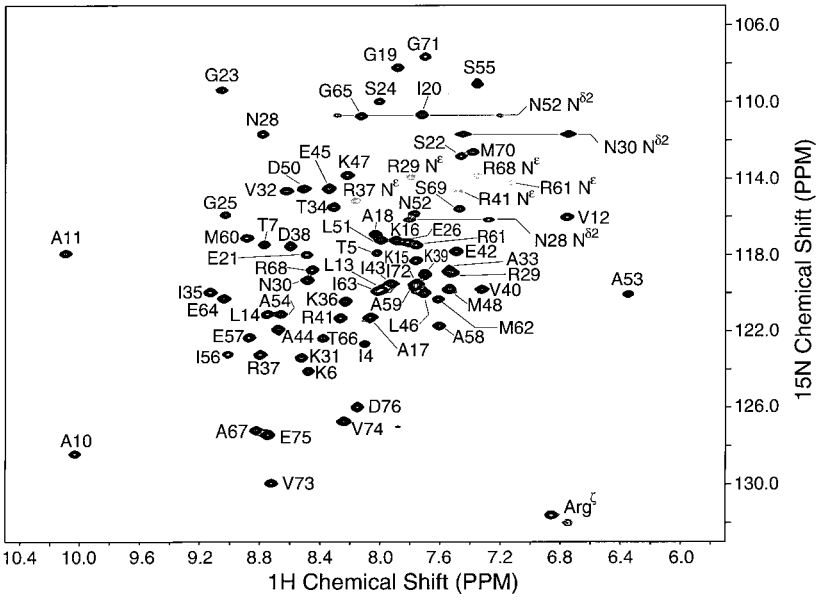
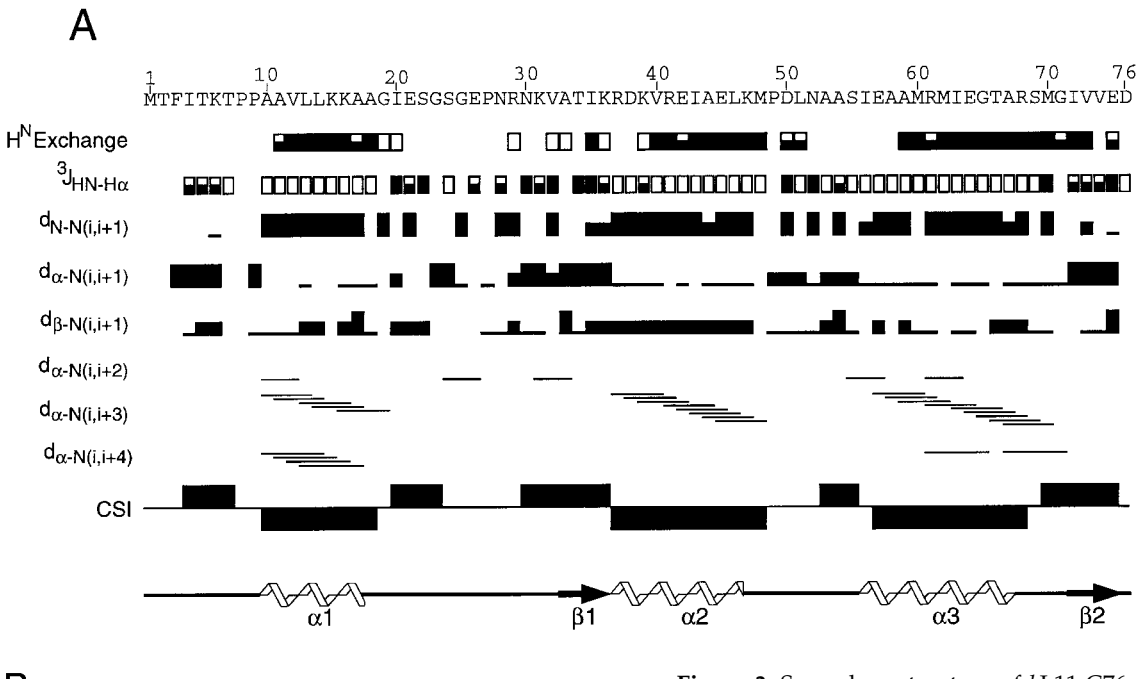


Figure 2. Two-dimensional $^{1}H/^{15}N$ HSQC spectrum of *b*L11-C76. The peaks are labeled by their assignments, with the one-letter amino acid code and their position in the amino acid sequence. The arginine $H^{\varepsilon}/N^{\varepsilon}$ and H^{ζ}/N^{ζ} resonances, are singly and doubly folded into the spectrum, respectively. Their true chemical shifts are their apparent chemical shifts minus 30.2 ppm and 60.4 ppm, respectively.



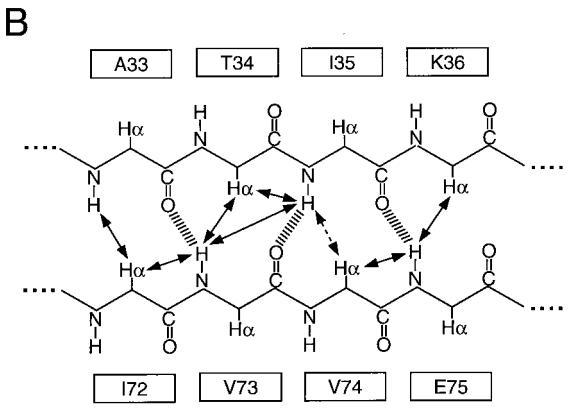


Figure 3. Secondary structure of *b*L11-C76. (a) Secondary structure as deduced from sequential and short range NOE connectivities, hydrogen exchange, $^{3}J_{HNH\alpha}$ couplings, and the chemical shift index (Wishart & Sykes, 1994). Hydrogen exchange: open boxes indicate residues whose backbone amides could be observed at 0.27 hour, half-filled boxes correspond to those which could be detected at 1.53 hours, and filled boxes correspond to those which could be observed longer than 1.53 hours. Three-bond H^{N}-H^{α} couplings: open boxes correspond to $^{3}J_{HNH\alpha} < 6.0$ Hz, half-filled boxes correspond to $6.0\ \text{Hz} < {}^{3}J_{HNH\alpha} < 8.0$ Hz, and filled boxes correspond to $^{3}J_{HNH\alpha} > 8.0$ Hz. Sequential NOE connectivities ($d_{N\text{-}N(i,i+1)}$, $d_{\alpha\text{-}N(i,i+1)}$, $d_{\beta\text{-}N(i,i+1)}$) are classified according to relative intensities, the bar height corresponding to increasing NOE intensity. Short range NOE connectivities are indicated by horizontal bars which link the first residue involved in the crosspeak to the second. Chemical shift index (CSI): C^{α}, C^{β}, C^{O}, and H^{α} consensus values are shown, upward boxes corresponding to β-conformation, downward boxes corresponding to α-conformation, and other regions corresponding to coil. The secondary structure of *b*L11-C76, as deduced by the program PROCHECK (Laskowski *et al.*, 1993), is diagrammed on the lower part of the Figure. (b) A diagramatic illustration of the parallel β-sheet of *b*L11-C76. Sequential and long range NOE connectivities that have been identified are indicated by double-ended arrows; the single broken arrow corresponds to an expected NOE connectivity which could not be verified experimentally owing to resonance overlap. The horizontally broken lines correspond to hydrogen bonds that have been included in the calculation of the *b*L11-C76 structures.

exchange, and NOE data. Although the origin of this discrepancy is not fully understood, one possible source is the fact that these regions adopt generally extended, but non-regular secondary structures.

Solution structure of bL11-C76

The solution structure for *b*L11-C76 was calculated using the distance geometry/simulated annealing protocol as implemented in X-PLOR 3.8. The input data for the calculation, summarized in Table 1, consisted of a total of 684 interresidue NOE-distance and dihedral angle restraints. The ensemble of calculated structures consistent with the experimental restraints (26/50) is shown in Figure 4(a). The set of calculated structures showed good covalent geometry and low overall energies; a summary of the statistics describing these structures is provided in Table 1. The structures have been deposited with the Protein Data Bank (Brookhaven Natiotional Laboratory, Broookhaven, NY, energy minimized average PDB code 1FOY, ensemble of accepted structures PDB code 2FOW). A ribbon diagram depicting the energy minimized average structure is shown in Figure 4(b) and reveals that the overall structure consists of a bundle of three α-helices and a two-stranded parallel β-sheet four residues in length.

Table 1. Distance and dihedral restraint statistics for the structure of *b*L11-C76

A. *Restraint summary*	
Total	684
Total NOE distance restraints	545
Intraresidue ($i-j=0$)	0
Sequential ($\|i-j\|=1$)	241
Short range ($1<\|i-j\|\leqslant 5$)	149
Long range ($\|i-j\|>5$)	155
Total hydrogen bond restraints[a]	46
Total dihedral restraints	60
ϕ	47
$\chi 1$	13
Accepted structures[b]	26/50

B. *RMSD from idealized covalent geometry and experimental restraints*

	$\langle b$L11-C76$\rangle_r^c$	{*b*L11-C76}[c]
Bonds (Å)	0.002	0.002 ± 0.0001
Angles (°)	0.57	0.60 ± 0.01
Impropers (°)	0.44	0.40 ± 0.01
Distance restraints (Å)	0.019	0.025 ± 0.002
Angle restraints (°)	0.000	0.05 ± 0.08
$^3J_{HNH\alpha}$ restraints (Hz)	0.44	0.51 ± 0.03
F_{NOE} (kcal/mol)[d]	11.0	18.2 ± 2.9
$F_{Dihedral}$ (kcal/mol)[e]	0.00	0.02 ± 0.08
$F_{HNH\alpha}$ (kcal/mol)[f]	6.3	8.5 ± 1.1

[a] Includes two restraints per hydrogen bond.
[b] Acceptance criteria: NOE distance violations < 0.3 Å, dihedral angle violations < 3°, and $^3J_{HNHa}$ coupling violations < 1.5 Hz.
[c] {*b*L11-C76} is the ensemble of 26 accepted structures; $\langle b$L11-C76$\rangle_r$ is obtained by restrained minimization of the average of {*b*L11-C76}.
[d] Calculated using a square well potential, center averaging, a scale factor of 50, a square-well constant of 1.0, and ceiling of 1000 kcal/mol (Brünger, 1992).
[e] Calculated using a square well potential and a scale factor of 50 (Brünger, 1992).
[f] Calculated using a harmonic potential and a force constant of 1.0 (Brünger, 1992).

Disorder *versus* flexibility

A superposition of the ensemble of calculated structures relative to the mean reveals that the majority of the protein backbone is well determined. There are, however, two exceptions to this generalization: the six N-terminal residues and residues I20 to N30 in loop 1 both exhibit considerable disorder over the ensemble of calculated structures (Figure 5(a)). The overall RMSD values in the protein backbone coordinates when these and various other amino acid segments are not included in the alignment, are summarized in Table 2. For the ordered part of the protein (P8-G19, K31-D76) the backbone RMSD is 0.64 Å relative to the mean, whereas this value is lowered to 0.50 Å if one considers only the regular elements of secondary structure.

We have observed that the calculated disorder of the N-terminal segment correlates well with internal flexibility in solution, as judged by ^{15}N relaxation data measurements, but the same is not true for residues in loop 1. Measurements of backbone ^{15}N transverse relaxation times (^{15}N T_2) for *b*L11-C76, as shown in Figure 5(c), reveal that with the exception of the N-terminal residues I4, T5, and K6, ^{15}N T_2 values are relatively uniform over the length of the protein backbone. In contrast, the calculated structures reveal elevated RMSD values in the N-terminal segment (M1 to K6) and in loop 1 (I20 to N30) (Figure 5(a)). Residues M1 to K6 of *b*L11-C76 provide a clear example of disorder correlated with flexibility: the ill-defined conformation of this segment is a direct consequence of its internal flexibility and a lack of NOE restraints which define its conformation (Figure 5(b)). Residues I20 to N30 provide an example of disorder uncorrelated with flexibility: the precision of this part of the structure is limited by the fact that only short-range and intraresidue intraprotein NOEs are observed between residues 26 and 30. However, as will be seen, several loop 1 residues have NOEs to the RNA.

Comparison to L11-C76 free in solution

The secondary structure of *b*L11-C76, as deduced by PROCHECK (Laskowski *et al.*, 1993), is in good agreement with that found for L11-C76 free in solution (*f*L11-C76) (Markus *et al.*, 1997), although there are minor differences involving the end points of the helix 1, helix 3, and the two β-strands. Helix 1 extends from A10 to A18 for *f*L11-C76, whereas, it is shortened by one residue in *b*L11-C76; helix 3 extends from I56 to T66 for *f*L11-C76, whereas, it is lengthened by one residue (I56 to A67) in *b*L11-C76. Neither difference appears to be significant as both helix 1 in *b*L11-C76 and helix 3 in *f*L11-C76 have helical ϕ/ψ angles which extend beyond residues A17 and T66, respectively. The parallel β-sheet of *b*L11-C76 differs from that of *f*L11-C76 in that it is extended by one residue at the N-terminal end: A33 to K36 pairs with I72 to E75 for *b*L11-C76 (Figure 3(b)), whereas, T34 to K36 pairs with V73 to E75 for *f*L11-C76. The major observation that supports this difference in structure is an additional NOE restraint in the case of *b*L11-C76 involving the H^N of A33 to the H^α of I72.

An overlay of the energy-minimized *b*L11-C76 and *f*L11-C76 structures (Markus *et al.*, 1997), as shown in Figure 6, reveals that overall the two structures are quite similar. The largest differences are found in the N-terminal segment from residue M1 to P8 and in loop 1 from residue A18 to V32. The pairwise RMSD for superposition of the backbone atoms excluding these two regions is 1.74 Å, whereas this value is lowered to 1.26 Å if one considers only the regular elements of secondary structure (Table 2). The precision of the ensemble of calculated *b*L11-C76 structures generally parallels that of *f*L11-C76 (Markus *et al.*, 1997): both the N-terminal segment M1–T7 and the loop 1 segment I20–V32 display substantially elevated RMSD values. One notable difference between the two sets of calculated structures is the fact that the RMSD of loop 1 conformations for *b*L11-C76 is roughly one-half that of *f*L11-C76. The primary determinant of this difference appears to be several additional long range NOE

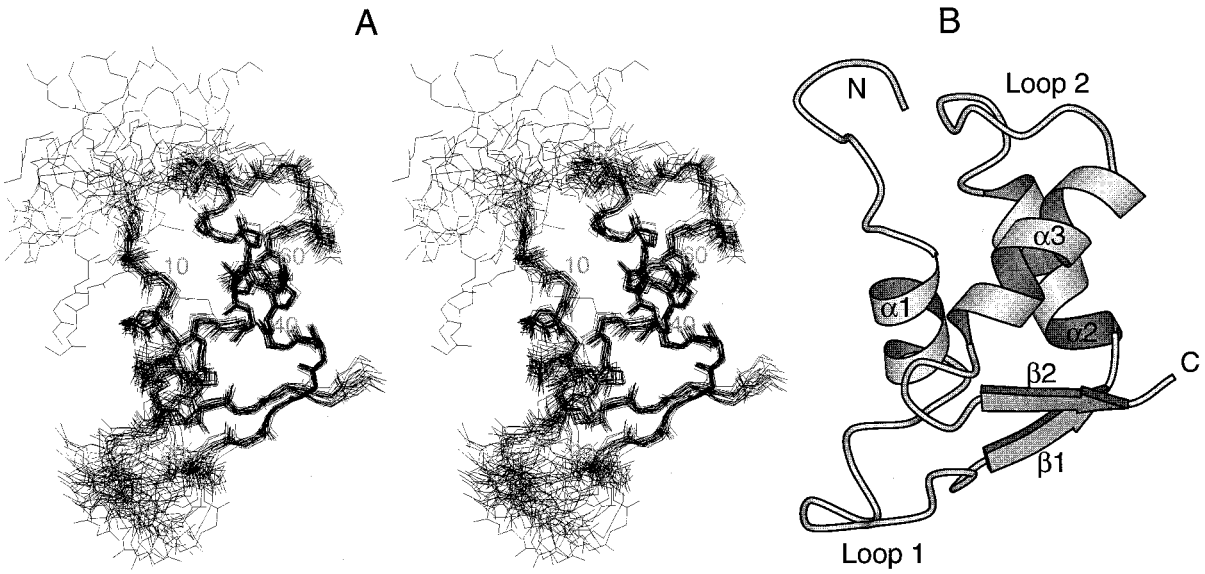


Figure 4. NMR solution structures of *b*L11-C76. (a) A stereoview of the ensemble of 26 accepted structures, superimposed on the basis of the deduced secondary structure (residues 10 to 17, 33 to 46, 56 to 67, 72 to 75). Only backbone heavy atoms are shown. (b) A MOLSCRIPT (Kraulis, 1991) ribbon diagram of the energy-minimized average *b*L11-C76 structure in which the regular elements of secondary structure have been identified.

distance restraints involving residues 24, 25, 31, and 32, which are present in the *b*L11-C76 restraint data set.

Interestingly, *f*L11-C76 differs from *b*L11-C76 in that the disorder among the calculated structures clearly correlates with internal flexibility for both the N-terminal segment from M1 to T7, as well as the loop 1 segment from G19 to V32 (Markus *et al.*, 1997). Loop 1 was shown to be flexible for *f*L11-C76 with internal motions occurring on two timescales: slow (milliseconds to seconds) motions corresponding to *cis:trans* isomerization of the E26–P27 peptide bond, superimposed upon much faster segmental motions (ns to ps) of the loop backbone (Markus *et al.*, 1997). Another major difference between the free and RNA-bound forms of L11-C76 is in the apparent lack of conformational heterogeneity involving *cis:trans* isomerization of the E26–P27 prolyl peptide bond. Evidence supporting this contention follows from the observation that there were no unassigned peaks of significant intensity corresponding to the second isomer present in the 2D $^{1}H/^{15}N$ HSQC spectrum of *b*L11-C76. Attempts to identify such peaks by collecting very high signal-to-noise 2D $^{1}H/^{15}N$ HSQC of *b*L11-C76 revealed several weak peaks in the random coil region of the spectrum, the strongest of which was less than 6% of the average H^{N}/N^{H} peak intensity of *b*L11-C76. Additionally, P8, P9, and P27 have been identified as having a *trans* peptide bond on the basis of strong Xaa H^{α}–Pro H^{δ} NOEs and the lack of Xaa H^{α}–Pro H^{α} NOEs. The former NOE is characteristic of the a *trans* prolyl peptide bond, whereas, the latter is characteristic of a *cis* prolyl peptide bond (Hinck *et al.*, 1993). (For P49, it was not possible to definitively identify either Xaa H^{α}–Pro H^{δ} or Xaa H^{α}–Pro H^{α} NOEs). Taken together, these data indicate that loop 1 becomes ordered upon binding to its cognate RNA. There are two limiting models that could explain the ordering: indirect effects induced by RNA binding, or direct effects that are the result of loop 1 contacts to the RNA.

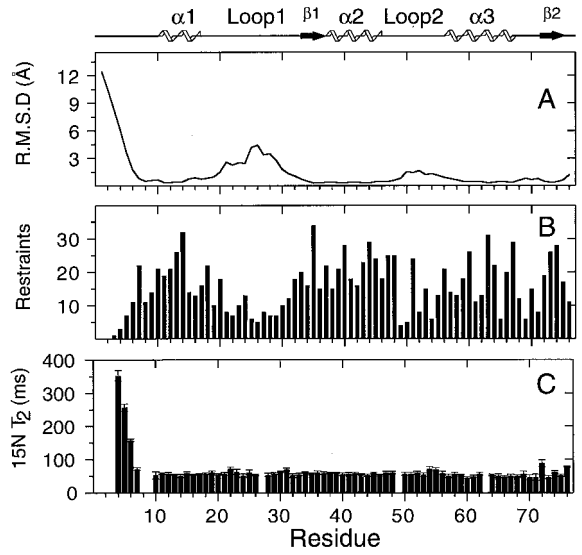


Figure 5. Disorder and flexibility of *b*L11-C76. (a) Residue RMSD values for the ensemble of 26 accepted structures *versus* the mean structure. (b) The number of NOE distance restraints on a per-residue basis. (c) Backbone ^{15}N transverse relaxation times for *b*L11-C76 at 42°C in buffer consisting of 95% H_2O/5% $^{2}H_2O$ with 10 mM Na_2HPO_4, 25 mM KCl (pH 6.1). The data were recorded at a field strength of 11.7 T.

Table 2. Coordinate RMSD of *b*L11-C76

Structures[a]	All		Ordered		Secondary		Helices	
	bb[b]	Heavy	bb	Heavy	bb	Heavy	bb	Heavy
{*b*L11-C76} *versus* ⟨*b*L11-C76⟩	2.26	2.84	0.64	1.24	0.50	1.15	0.48	1.20
⟨*b*L11-C76⟩ *versus* ⟨*b*L11-C76⟩$_r$	1.14	1.49	0.56	0.79	0.46	0.70	0.44	0.72
⟨*b*L11-C76⟩$_r$ *versus* ⟨*f*L11-C76⟩$_r$	3.06	4.00	1.74	2.58	1.26	1.97	1.13	1.88

In units of Å.

[a] {*b*L11-C76} is the ensemble of 26 accepted structures; ⟨*b*L11-C76⟩ is the average of {*b*L11-C76} ⟨*b*L11-C76⟩$_r$ is obtained by restrained minimization of the average of {*b*L11-C76}; ⟨*f*L11-C76⟩$_r$ corresponds to the energy minimized average structure for *f*L11-C76 (Markus, *et al*, 1997; PDB 1FOW).

[b] bb includes the backbone atoms, N^H C^α, and C^O; heavy includes all non-hydrogen atoms; All includes residues 1 to 76, ordered includes 10 to 20, 31 to 50, 57 to 76, secondary structure includes 10 to 18, 33 to 36, 37 to 48, 57 to 68, 72 to 75, and helices includes 10 to 18, 37 to 48, and 57 to 68.

NMR evidence supporting the latter model is presented below.

Chemical shift perturbations and protein:RNA NOEs

The chemical shift is sensitive to small perturbations in local geometry and conformation, the anisotropy of nearby magnetic fields, and local electrostatics (Oldfield, 1995), and has therefore proven to be a useful tool for monitoring the effects of ligand binding and conformational changes that occur within biological macromolecules. A comparison of the 2D $^1H/^{15}N$ HSQC spectrum of *b*L11-C76 with that of *f*L11-C76 showed that the binding to RNA has a major influence on the chemical shifts of the protein (Markus *et al*., 1997). Herein, we have extended such a comparison to include H^α, H^N, C^α, C^β, C^o, and N^H shift perturbations, several examples of which are presented in Figure 7, and used these to calculate a composite shift perturbation index (bottom panel). The largest shift perturbations fall within four regions of the structure: the N-terminal portion of α-helix 1 (T7 to A11), loop 1 (I20 to A33), the C-terminal end of α-helix 2 and loop 2 (L46 to A54), and throughout α-helix 3 and into the turn linking α-helix 3 and the second β-strand (A58 to G71).

Such perturbations can be broadly classified as being caused by a direct interaction with the RNA, by structural changes that occur within the protein upon RNA binding, or both. Because of the latter effect, chemical shift perturbations must be inter-

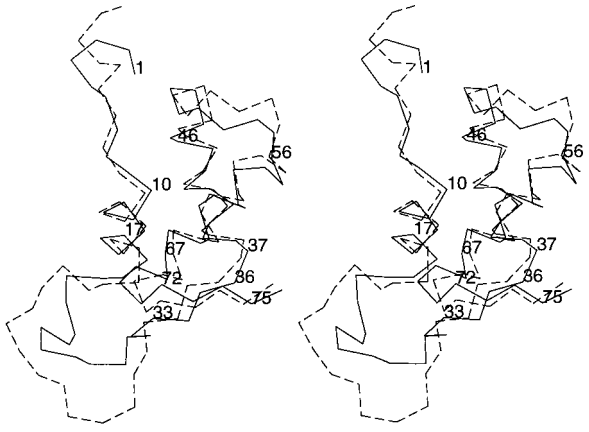


Figure 6. A comparison of the backbone C^α traces for *f*L11-C76 (Markus *et al*., 1997) to *b*L11-C76. The structures have been superimposed on the basis of the deduced secondary structure of *b*L11-C76 (residues 10 to 17, 33 to 46, 56 to 67, and 72 to 75). *b*L11-C76 is indicated by a continuous line, whereas, *f*L11-C76 is indicated by a broken line. Residues are numbered according to the starting and ending points of the regular secondary structure of *b*L11-C76.

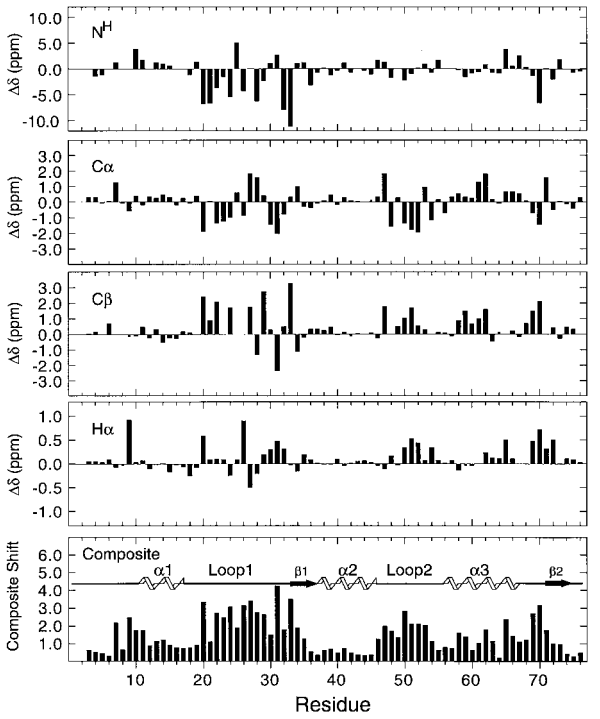


Figure 7. Chemical shift differences between the RNA-bound form of the protein, described here, with those previously reported for L11-C76 free in solution (Markus *et al*., 1997). Panels, from top to bottom, include backbone ^{15}N, C^α, C^β, and H^α shift perturbations, respectively, where the chemical shift difference (Δδ) is calculated by subtracting shift values for the free form of the protein from those of the RNA-bound form of the protein. The bottom panel includes the composite backbone ^{15}N, C^O, C^α, C^β, H^N, and H^α shift perturbations, as described in Materials and Methods.

Table 3. A summary of *b*L11-C76–RNA NOEs identified in ^{13}C-filtered NOE spectra

Protein group	Chemical shift RNA protons (ppm)
A10 H^{β}	4.56, 5.88
A11 H^{β}	3.84, 4.47
S24 H^{N}	6.16
E26 H^{α}	6.30
P27 H^{δ},$H^{\delta\prime}$	6.30
L51 H^{δ}	4.32, 4.42, 6.17
N52 $H^{\delta 2}$,$H^{\delta 2\prime}$	5.50, 5.76
A53 H^{α}	4.09, 4.75
A53 H^{β}	4.32
A54 H^{N}	4.08, 4.74
A54 H^{β}	4.09, 4.71
A58 H^{β}	4.15, 4.33, 4.65
M62 H^{α}	5.90
M62 H^{ε}	4.34, 5.72, 6.17, 7.30
T66 H^{N}	5.91
T66 $H^{\gamma 2}$	5.88
R68 H^{δ},$H^{\delta\prime}$	4.52, 5.84
S69 H^{N}	6.37
M70 H^{N}	6.35
M70 H^{α}	4.38, 4.58, 6.03
M70 H^{β},$H^{\beta\prime}$	4.37
M70 H^{ε}	3.84, 4.37

The H^{δ} protons of P27 are non-degenerate and are designated as P27 H^{δ} and P27 $H^{\delta\prime}$; each displays an NOE to RNA at 6.3 ppm.

Two 3D ^{13}C-filtered NOE experiments were acquired: ^{13}C-selected in F1 with ^{13}C filtering in F3 and ^{15}N-selected in F1 with ^{13}C-filtering in F3.

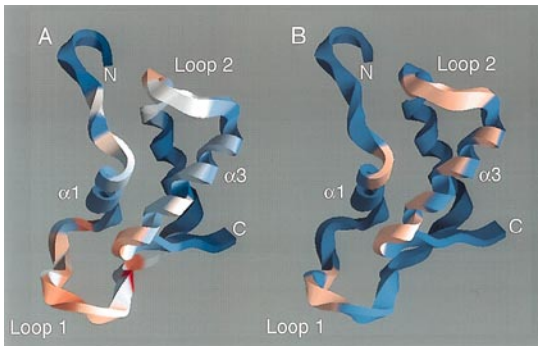


Figure 8. Correspondence of chemical shift perturbations that occur upon RNA binding with protein-RNA intermolecular NOEs that have been identified in ^{13}C-filtered NOE experiments. (a) Composite chemical shift perturbations, calculated as described in Materials and Methods, mapped onto a ribbon diagram of the *b*L11-C76 structure. The magnitude of the shift perturbation is indicated by the differential shading of the protein backbone; the smallest shift perturbations are in blue, intermediate perturbations are in white, and the largest perturbations are in red. (b) A schematic depiction of those parts of the *b*L11-C76 structure which were shown to display NOE contacts to the its cognate RNA. Those residues which displayed protein-RNA NOEs are indicated in a pink color, whereas those which did not are indicated in blue. (a) and (b) were both generated with the program, GRASP (Nicholls *et al.*, 1991).

preted carefully because RNA-induced structural changes are not necessarily relegated to sites of direct protein-RNA contact. The structural comparison of *b*L11-C76 and *f*L11-C76 presented herein shows that there are minimal structural differences in the ordered regions of the sequence, P8 to G19 and K31 to D76. Hence, the shift perturbations observed for T7 to A11, L46 to A54, and A58 to G71 likely arise from direct contacts with the RNA. This conclusion was confirmed by ^{13}C-filtered/^{13}C- or ^{15}N-selected NOESY experiments which strongly select for protein-RNA NOEs on samples containing uniformly ^{15}N/^{13}C labeled protein and unlabeled RNA (Folkers *et al.*, 1993). A total of 40 unique protein-RNA NOEs have been assigned using this approach; a listing is provided in Table 3. Protein residues having NOEs to RNA include A10 and A11 in the N-terminal end of α-helix 1, L51, N52, A53, and A54 in loop 2, A58, M62, and T66 in α-helix 3, and R68, S69, and M70 in loop 3. The data presented in Figure 8 show that there is good agreement between the location of protein-RNA NOEs and the three regions of the protein whose chemical shifts are most strongly affected by RNA binding (T7 to A11, L46 to A54 and A58 to G71). It is interesting to note that the shift perturbations and protein-RNA NOEs all map onto the face of the molecule which includes α-helix 3 (Figure 8). This observation is intriguing in light of the fact that it has been suggested that helix 3 plays an important role in mediating protein-RNA contacts (Markus *et al.*, 1997; Xing *et al.*, 1997).

The other protein residues that have NOEs to the RNA are S24, E26, and P27 in loop 1. Thus, residues located near the center of loop 1 associate intimately with target RNA. Moreover, the dramatic loop 1 structural and dynamic changes detailed earlier, appear to be the direct result of the interaction of L11-C76 with its target RNA, not merely an indirect structural/dynamic change that accompanies RNA binding. In regards to the large shift perturbations which occur within loop 1, two effects appear to predominate: one related to structural perturbations which occur within the loop, and another related to electrostatic and ring-current effects of RNA binding.

The role of helix 3 in RNA binding

Previous structural work on *f*L11-C76 revealed the interesting structural similarity between the homeodomain family of eukaryotic helix-turn-helix transcription factors and the three-helix bundle of *f*L11-C76 (Markus *et al.*, 1997; Xing *et al.*, 1997). The structure of *b*L11-C76 is similarly related to the homeodomain motif. A structural based search of the Protein Data Bank using the DALI algorithm (Holm & Sander, 1993) reveals that the three helical domains of *b*L11-C76 superimpose with pairwise RMSD values of 1.6 and 1.3 Å with the Oct-1 Pou and MAT-α2 homeodomains, respectively (Table 4). Homeodomains recognize their target DNA sequences specifically by positioning helix 3 into the major groove. This arrangement is stabil-

Table 4. Structural alignments of *b*L11-C76 to homeodomains

	DALI[a]	Three helices[b]	Helix 2/helix 3[c]
⟨*b*L11-C76⟩ *versus* Oct-1	2.68	1.62	0.90
⟨*b*L11-C76⟩ *versus* MAT-α2	2.46	1.34	0.78

[a] Alignment of L11-C76 and homeodomain amino acid sequences on the basis of structural homology of *b*L11-C76 with the Oct-1 and MAT-α2 homeodomains, respectively. For the L11-C76–Oct-1 comparison, residues K6 to P9, V12 to K16, A18 to E21, V32 to M48, and A59 to G71 of *b*L11-C76 align with E9 to I12, V14 to K18, S19 to E22, Q24 to N39, and E41 to R53 of Oct-1, respectively. For the L11-C76-MAT-α2 comparison, residues K6 to P9, V12 to K16, A18 to E21, V32 to M48, and A59 to G71 of *b*L11-C76 align with residues T9 to N12, R14 to S18, W19 to K22, N23 to S39, and S41 to R53 of MAT-α2, respectively. Numbers reported are pairwise RMSD values in Å between backbone atoms of the two structures.

[b] Sequences are aligned as in column 1, but include only residues of the helical domains (residues V12 to K16, R37 to L46, and A59 to A67 of L11-C76).

[c] Sequences are aligned as in column 1, but include only residues of helix 2 and helix 3 (residues R37 to L46 and A59 to A67 of L11-C76).

ized by an array of basic residues that interact with the phosphate backbone, along with base-specific contacts involving a universally conserved Asn at position 51 (Figure 1(b)). Additional major groove base-specific contacts that typify most homeodomain-DNA complexes involve residues found at positions 47, 50, and 54. Additional homeodomain-DNA interactions occur by wrapping the N-terminal tail into the adjacent minor groove and by interacting with the DNA *via* conserved basic residues.

The current NMR data strongly support the hypothesis that helix 3 of L11-C76 plays an important role in RNA binding: G65, R68 and S69, all undergo significant shift perturbations in response to RNA binding, whereas the latter two display protein-RNA NOEs as well (Figure 1(c)). Moreover, the chemical shift perturbations and protein-RNA NOEs generally correlate with mutagenesis studies of the protein (Xing *et al.*, 1997) which have shown that substitution of helix 3 residues G65 and T66 by alanine and valine, respectively, lowers the binding affinity for the target RNA by greater than 35-fold. Alanine substitutions for other residues in helix 3 and loop 3 have smaller effects on binding, R61A weakens binding 5.0-fold, R68A 2.7-fold, and S69A 9.5 fold. These observations highlight the structural and thermodynamic importance of helix 3 residues in RNA binding.

The RNA fragment to which L11-C76 binds is composed of three helical stems, one of which has a number of sites protected in hydroxyl radical footprinting studies of native L11 bound its target RNA (Rosendahl & Douthwaite, 1993; Figure 1(a)). Although an *A*-form RNA does not form a suitable interaction surface for an α-helix, irregular RNA helices that result from internal loops, bulge loops, and mismatches can accommodate an α-helix, as has been recently suggested by the solution NMR structure of the HIV rev peptide bound to its target site, the rev response element (Battiste *et al.*, 1996). The present work does not provide any direct information regarding the L11 contact sites on the RNA, although it is apparent that helix 3 must be in close proximity to the bases of the RNA as helix 3 residue M62 displays an NOE with a chemical shift of 7.30 ppm (Table 3). On the basis of chemical shifts alone, these are likely to be adenine or guanine H8, adenine H2, or uridine or cytosine H6 protons, which would be consistent with the interaction of α-helix 3 with a distorted major or minor groove. It is tempting to speculate that the internal loop of stem 2 leads to distortions in the helical framework, thereby exposing a suitable interaction surface for helix 3 of L11-C76. Several lines of evidence, including universal conservation of A1077 and mutagensis studies of U1060 and the U1060-U1078 mismatch (Lu & Draper, 1994; Ryan *et al.*, 1991) implicate the internal loop as being important for RNA tertiary structure formation. Structural studies of the RNA component of the complex, which are clearly needed, should provide information regarding the extent to which stem 2 forms a helical structure and the details by which it interacts with helix 3.

Interactions outside helix 3 are important in RNA binding

On the basis of the shift perturbation data and the protein-RNA NOE data, three additional regions of L11-C76 outside of helix 3 contact the RNA. These include the N-terminal end of helix 1, loop 1, and the C-terminal end of helix 2 including loop 2. It is of interest that the flexible N-terminal tail of *b*L11-C76 is not among these. The N terminus is a conserved region of homeodomain sequences, and is flexible in the free homeodomain, but upon binding DNA, it inserts into the minor groove in an ordered conformation (Otting *et al.*, 1990).

We note that unlike L11-C76, simple three helix homeodomains generally do not have contact sites that fall outside of helix 3, although protein-DNA backbone contact sites are sometimes found within loop 1 and helix 2 (Figure 1(b)). More recent structural work on the PU.1 ETS (Kodandapani *et al.*, 1996) and HNF-3/forkhead (Clark *et al.*, 1993) DNA binding proteins complexed with their target DNA sequences has shown that sequence-specific DNA recognition can often be achieved by the use of additional loops and β-strands flanking α-helix 3

of a core homeodomain fold. These additional regions are referred to as wings. The fold that these proteins adopt, which is classified as α + β helix-turn-helix, consists of a core homeodomain plus additional β-sheet regions. Loop extensions and β-strand insertions are common among the loops connecting the three homedomain helices. Critical DNA contact sites often include residues located within the loops connecting the homeodomain helices, as in the case of PU.1 ETS, or by the loops which serve to connect the β-strands, as in the case of HNF-3/forkhead.

Although the topology and three-dimensional structure of L11-C76 is not identical to any of the winged homedomains whose structure has been thus determined, it appears that L11-C76 employs a similar strategy to interact with its target RNA. Loop 1 is considerably longer than corresponding loops in minimalist homeodomains such as Oct-1, engrailed, and Mat α2. Additionally, the loop contains a single β-strand insertion, and is flexible in *f*L11-C76 (Markus *et al.*, 1997), but is rigid and residues S24, E26, and P27 display RNA contacts upon complex formation. Although none of the L11-C76 loop 1 residues is fully conserved across the three major phylogenetic domains, several, such as G23, S24, and P27 of eubacteria, are highly conserved within a single domain (Xing *et al.*, 1997). Additionally, mutagenesis studies (Xing *et al.*, 1997) have shown that these three residues in particular contribute significantly to L11-C76–RNA binding affinities: the G23P, S24A, and P27G substitutions diminish protein-RNA binding affinities 13, 2.4, and 13-fold, respectively. Within loop 2 of L11-C76, there are four residues, L51, N52, A53, and A54, which display both substantial RNA shift perturbations and protein-RNA NOEs. Although these residues are not well conserved among the three major phylogenetic domains (Xing *et al.*, 1997), their RNA contacts may be relevant, as loop 2 is considerably longer than that found in minimalist homeodomains (Figure 1(b) and (c)) and several protein-DNA contacts were found among PU.1 ETS loop 2 residues and the DNA oligomer with which it was co-crystallized. The L11-C76–RNA contacts identified in the N-terminal segment of α-helix 1 may also play a role in RNA recognition, in a manner similar to residues located in the N-terminal end of α helix-1 of PU.1 ETS which contact the phosphate backbone.

In summary, the major RNA contact sites in *b*L11-C76 include α-helix 3 and residues located within loop regions flanking its two ends. The extensive interaction surface that is formed from these contacts, as delineated by the shift perturbation and protein-RNA NOE data (Figure 8), suggests a model for RNA recognition in which the primary interaction surface, α-helix 3, is anchored on both ends by additional protein-RNA interactions involving loops 1 and 2. On the basis of the ^{1}H chemical shifts of the protein-RNA NOEs, it appears that helix 3 may contact the RNA in the context of a groove, although at present the exact nature of the RNA interaction surface is not known. The chemical shifts of the RNA protons contacted by loop 1 residues are in the 6.2 to 6.3 ppm range and are probably either anomeric ribose protons or cytosine H6 protons, whereas loop 2 residues have NOE contacts to RNA protons primarily in the 4.0 to 5.0 ppm, and are probably ribose H2′, H3′, H4′, H5′, or H5″ protons. The exception to this are L51 and N52 which display NOEs to protons at 6.2 and 5.5/5.8 ppm, respectively, which again are probably either anomeric ribose protons or cytosine H6 protons. Further NMR studies are clearly needed to assign the RNA resonances in the L11-C76–RNA complex, and to use these along with NOE data to elucidate a three-dimensional structure for the complex.

Conclusion

The C-terminal domain of ribosomal protein L11 is the newest member of the family of ribosomal proteins whose structure has been studied at near-atomic resolution (Ramakrishnan *et al.*, 1995). Significantly, this is the first ribosomal protein whose interaction with ribosomal RNA has been studied using a structure-based approach. These studies have shown that the RNA binding domain of L11 interacts with its target RNA by employing α-helix 3 of a homeodomain DNA-binding motif, and by the use of loop regions that serve to connect the three helices of the homeodomain. The loop regions of L11-C76 are lengthened relative to those of minimalist homeodomains and probably have adapted through evolution to optimize binding to 23 S rRNA. The RNA-binding strategy employed by L11-C76 appears to be similar in principle to that used by winged homeodomain DNA binding proteins, such as HNF-3/forkhead and PU.1 ETS, although the placement of β-strands and use of loop regions differ in detail. In the absence of any direct structural information about the RNA itself, it is premature to speculate further about the interaction. One potential practical outcome of the type of structural information reported here would be to aid in the design of new classes of antibiotics that specifically disrupt the interaction between key ribosomal proteins and their cognate RNA. The design of such drugs are of increasing importance as many of the drugs that selectively target prokaryotic ribosomes are no longer effective.

Materials and Methods

Sample preparation

L11-C76, whose sequence is shown in Figure 1(c), was produced in *E. coli* using a pET11a expression vector (Novagen, Madison, WI) and was labeled uniformly with either ^{15}N or $^{15}N/^{13}C$ and purified as described (Markus *et al.*, 1997; Xing & Draper, 1996). The RNA, whose nucleotide sequence is shown in Figure 1(a), was produced using T7-based *in vitro* runoff transcription with a plasmid template and was purified by preparative scale polyacrylamide gel electrophoresis (8% (w/v)

gel) and electroelution as described (Laing & Draper, 1994). Samples of the L11-C76–RNA complex for NMR spectroscopy were prepared by first exchanging each of the components into buffer containing 5 % (v/v) ^2H_2O, 10 mM K_2HPO_4, 25 mM KCl, 2 mM $MgCl_2$ at pH 6.1 (Centricon-10; Amicon, Beverly, MA) followed by stepwise addition of small aliquots of the RNA solution. The addition of RNA was continued until peaks arising from the non-liganded form of the protein were no longer present in 2D $^1H/^{15}N$ HSQC spectra. The samples were then lyophilized, redissolved in either 5% ^2H_2O/95% H_2O or 99.99% ^2H_2O in a volume of approximately 220 µl, and transferred to a 5 mm NMR microcell (Shigemi, Allison Park, PA). The final concentration of the L11-C76-RNA complex, as judged on the basis of the extinction coefficient of the RNA (1.00 A_{260} = 37.2 µg/ml, S.H., unpublished observation), was 0.5 mM.

NMR Spectroscopy

NMR spectra were acquired at field strengths of either 11.7 or 14.2 Tesla using either a Bruker AMX or DMX spectrometer equipped with a pulsed-field gradient unit and a triple-resonance 5 mm probe. The majority of data were collected at a temperature of 47°C, although the 3D CBCA(CO)NH data set was collected at 42°C. Spectral quality improved at 47°C owing to the more favorable relaxation properties of the complex. Sequential backbone assignments of the protein were made using the 3D CT-HNCA (Grzesiek & Bax, 1992b) and 3D CBCA(CO)NH (Grzesiek & Bax, 1992a) pair of experiments. Side-chain assignments were made by extending the sequentially assigned H^N resonances using the 3D C(CO)NH (Grzesiek *et al.*, 1993) and 3D HBHA(CO)NH (Grzesiek & Bax, 1993) pair of experiments, and by using the 3D HCCH-TOCSY (Bax *et al.*, 1990) experiment, although the 3D ^{13}C-edited NOESY was used in a few instances. Asparagine H^δ protons were assigned from their NOESY crosspeaks to the assigned H^β resonances, methionine H^ε resonances were assigned from a 2D HMBC experiment (Bax *et al.*, 1994), and arginine H^ε resonances were assigned by their NOESY crosspeaks to the assigned H^δ resonances.

Distance restraints were derived from three 3D NOESY data sets: ^{13}C-edited, ^{15}N-edited, and an ^{15}N-edited experiment with ^{13}C-editing and ^{13}C chemical shift evolution in the F2 dimension. Protein-RNA NOE contacts were delineated by a pair of ^{15}N and ^{13}C-edited NOESY experiments, modified with a filter that selects for ^{12}C-bound protons following the NOE transfer step. The $^3J_{HN\text{-}H\alpha}$ coupling constant was derived by averaging the calculated couplings from two independently acquired 3D HNHA (Vuister & Bax, 1993) data sets. Side-chain torsion angle restraints for Ile, Thr, and Val residues were made by measuring $^3J_{N\text{-}C\gamma}$, $^3J_{CO\text{-}C\gamma}$, and $^3J_{H\alpha\text{-}H\beta}$ using the 2D ^{13}CO spin echo difference constant time HSQC (SED-ctHSQC; Grzesiek *et al.*, 1993), 2D ^{15}N SED-ctHSQC (Vuister *et al.*, 1993), and 3D HAHB experiments (Grzesiek *et al.*, 1995), respectively.Hydrogen exchange rates were measured by dissolving a lyophilized, protiated sample of ^{15}N L11-C76, in standard NMR buffer containing 99.99% ^2H_2O and by following amide proton intensities with a series of 2D $^1H/^{15}N$ HSQC spectra at 37°C. Spectra were acquired at 0.27, 0.60, 0.93, 1.53, 3.18, 6.15, 10.1, 14.1, 18.1, 22.0 and 155.5 hours after ^2H_2O addition. Backbone ^{15}N transverse relaxation times were measured by the ^1H-detected pulse scheme as described (Kay *et al.*, 1992).

Structure calculations

Structure calculations were carried out with the distance geometry/simulated annealing protocol with XPLOR 3.8. The input data are described below. Distance restraints for structure calculations were derived by interactively analyzing the three NOE data sets using the program PIPP (Garrett *et al.*, 1991). The initial set of NOE crosspeaks were limited to those defining the secondary structure of the protein, along with several unambiguous NOEs involving residues that pack in the hydrophobic core. In successive iterations, NOEs were added to the restraint tables provided that either the symmetry-related crosspeak was present in the ^{13}C-edited NOE data set, as in the case of aliphatic-aliphatic NOEs, or the observed NOE crosspeak was present in both the ^{15}N-edited and ^{15}N-edited/^{13}C-edited NOE data sets, as in the case of amide-aliphatic NOEs. NOE crosspeak intensities were converted into distances by (a) first normalizing each to the intensity of the corresponding diagonal peak, and (b) by using a modified form of the isolated two-spin approximation (Suri & Levy, 1993, 1995) to calculate distances. The calculated distances, along with a 1.8 Å lower distance bound and appropriate pseudo-atom corrections for methylene, methyl, and non-stereospecifically assigned methyl groups of leucine and valine residues, were then used as input for structure calculations Distance restraints for hydrogen bonds, N to O and HN to O, were added for those residues which were found in regular elements of secondary structure and whose amide proton was protected from exchange with solvent (Figure 3(a)).

Backbone torsion angle restraints for ϕ, derived from HN-H^α coupling constants, were applied to only those residues found in regular elements of secondary structure, and were set to $-65(\pm25)^\circ$ for $^3J_{HNH\alpha} < 6$ Hz and $-120(\pm40)^\circ$ for $^3J_{HNH\alpha} > 8.0$ Hz. Phi torsion angles among the regular elements of secondary structure, were additionally restrained by refining the measured $^3J_{HNH\alpha}$ couplings against a reparametrized form of the Karplus equation (Wang & Bax, 1995). Side-chain torsion angles for 6/7 Ile, 1/4 Thr, and 5/5 Val residues, derived from 2D N-C^γ SED-ctHSQC, 2D CO-C^γ SED-ctHSQC, and 3D $H^\alpha H^\beta$ data sets, were restrained to one of the staggered rotamers, $+60^\circ$, -60°, or 180°, with an error range of $\pm30^\circ$.

An additional calculation was carried out by placing all of the calculated distances into either strong ($d = 2.5$ to 3.0 Å), medium ($d = 3.0$ to 4.0 Å), weak ($d = 4.0$ to 5.0 Å), or very weak (d> 5.0 Å) categories. The structures calculated from such restraints compared favorably with those calculated using NOE distance restraints derived from the isolated spin pair approximation. Thus, all calculations reported here were carried out using the former approach.

Chemical shift perturbations

Chemical shift perturbations of L11-C76 that occur upon RNA binding were calculated by subtracting the values for *f*L11-C76 previously reported (Markus *et al.*, 1997) from the values reported here for *b*L11-C76. Since the temperature at which the *f*L11-C76 shifts were recorded (25°C) differed from that of *b*L11-C76 (47° C), the *b*L11-C76 shifts were first temperature-corrected to a common reference temperature of 25 °C. A composite absolute shift perturbation index was calculated by first calculating the absolute shift for all residues of a given atom type and by normalizing the maximum to a value

of 1.0. These absolute shifts were then summed over all of the atom types compared (H^N, H^α, C^α, C^β, C^O, and N^H) to give the composite absolute shift perturbation.

Acknowledgments

We gratefully acknowledge Frank Delaglio for developing and supporting the NMR processing package nmrPipe/nmrDraw, and Dan Garrett for the data analysis package, PIPP. We also acknowledge Dr Larry Fischer for providing assistance in purifying the ^{15}N-labeled protein. This work was supported by the AIDS Targeted Antiviral Program of the Office of the Director on the National Institutes of Health.

References

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

Bax, A., Clore, G. M., Driscoll, P. C., Gronenborn, A. M., Ikura, M. & Kay, L. E. (1990). Practical aspects of proton-carbon-carbon-proton three-dimensional correlation spectroscopy of 13C Labeled Proteins. *J. Magn. Reson.* **87**, 620–627.

Bax, A., Delaglio, F., Grzesiek, S. & Vuister, G. W. (1994). Resonance assignment of methionine methyl groups and chi 3 angular information from long range proton-carbon and carbon-carbon *J* correlation in a calmodulin-peptide complex. *J. Biomol. NMR,* **4**, 787–797.

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

Clark, K. L., Halay, E. D., Lai, E. & Burley, S. K. (1993). Co-crystal structure of the HNF-3/forkhead DNA-recognition motif resembles histone H5. *Nature,* **364**, 412–420.

Draper, D. E., Xing, Y. & Laing, L. G. (1995). Thermodynamics of RNA unfolding: stabilization of ribosomal RNA tertiary structure by thiostrepton and ammonium ion. *J. Mol. Biol.* **249**, 231–238.

Egebjerg, J., Douthwaite, S. R., Liljas, A. & Garrett, R. A. (1990). Characterization of the binding sites of protein L11 and the L10(L12)$_4$ pentameric complex in the GTPase domain of 23 S ribosomal RNA from *Escherichia coli*. *J. Mol. Biol.* **213**, 275–288.

Folkers, P. J. M., Folmer, R. H. A., Konings, R. N. H. & Hilbers, C. W. (1993). Overcoming the ambiguity problem encountered in the analysis of nuclear Overhauser magnetic resonance spectra of symmetric dimer proteins. *J. Am. Chem. Soc.* **115**, 3798–3799.

Garrett, D. S., Powers, R., Gronenborn, A. M. & Clore, G. M. (1991). A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Reson.* **95**, 214–220.

Grzesiek, S. & Bax, A. (1992a). Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291–6293.

Grzesiek, S. & Bax, A. (1992b). Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.* **96**, 432–440.

Grzesiek, S. & Bax, A. (1993). Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins. *J. Biomol. NMR,* **3**, 185–204.

Grzesiek, S., Anglister, J. & Bax, A. (1993). Correlation of backbone amide and aliphatic side chain resonances in $^{13}C/^{15}N$ enriched proteins by isotropic mixing of ^{13}C magnetization. *J. Magn. Res.* **101**, 114–119.

Grzesiek, S., Kuboniwa, H., Hinck, A. P. & Bax, A. (1995). Multiple-quantum line narrowing for measurement of H^α-H^β *J* couplings in isotopically enriched proteins. *J. Am. Chem. Soc.* **117**, 5312–5315.

Hinck, A. P., Eberhardt, E. S. & Markley, J. L. (1993). NMR strategy for determining Xaa-Pro peptide bond configurations in proteins: mutants of staphylococcal nuclease with altered configuration at proline-117. *Biochemistry,* **32**, 11810–11818.

Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233(1)**, 123–128.

Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A. & Torchia, D. A. (1992). Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear *T*1 and *T*2 values in proteins. *J. Magn. Reson.* **97**, 359–375.

Klemm, J. D., Rould, M. A., Aurora, R., Herr, W. & Pabo, C. O. (1994). Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell,* **77**, 21–32.

Kodandapani, R., Pio, F., Ni, C.-Z., Piccialli, G., Klemsz, M., McKercher, S., Maki, R. A. & Ely, K. R. (1996). A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature,* **380**, 456–460.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.

Laing, L. G. & Draper, D. E. (1994). Thermodynamics of RNA folding in a conserved ribosomal RNA domain. *J. Mol. Biol.* **237**, 560–576.

Laing, L. G., Gluick, T. C. & Draper, D. E. (1994). Stabilization of RNA structure by magnesium ions. *J. Mol. Biol.* **237**, 577–587.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

Li, T., Stark, M. R., Johnson, A. D. & Wolberger, C. (1995). Crystal structure of the MATa1/MATα2 homeodomain bound to DNA. *Science,* **270**, 262–269.

Lu, M. & Draper, D. E. (1994). Bases defining ammonium and magnesium ion-dependent tertiary structure within the large subunit ribosomal RNA. *J. Mol. Biol.* **244**, 572–585.

Markus, M. A., Hinck, A. P., Huang, S., Draper, D. E. & Torchia, D. A. (1997). High resolution solution structure of ribosomal protein L11–C76, a helical protein with a flexible loop that becomes structured upon binding to RNA. *Nature Struct. Biol.* **4(1)**, 70–77.

Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**(4), 282–292.

Oldfield, E. (1995). Chemical shifts and three-dimensional protein structures. *J. Biomol. NMR,* **5**, 217–225.

Otting, G., Qian, Y. Q., Billeter, M., Müller, M., Affolter, M., Gehring, W. J. & Wüthrich, K. (1990). Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.* **9**, 3085–3092.

Ramakrishnan, V., Davies, C., Gerchman, S. E., Golden, B. L., Hoffman, D. W., Jaishree, T. N., Jadwiga, H. K., Porter, S. & White, S. W. (1995). Structures of prokaryotic ribosomal proteins: implications for RNA binding and evolution. *Biochem. Cell Biol.* **73**, 979–986.

Rosendahl, G. & Douthwaite, S. (1993). Ribosomal proteins L11 and L10.(L12)4 and the antibiotic thiostrepton interact with overlapping regions of the 23 S rRNA backbone in the ribosomal GTPase center. *J. Mol. Biol.* **234**, 1013–1020.

Ryan, P. C. & Draper, D. E. (1991). Detection of a key tertiary interaction in the highly conserved GTPase center of large subunit ribosomal RNA. *Proc. Natl Acad. Sci. USA,* **88**, 6308–6312.

Ryan, P. C., Lu, M. & Draper, D. E. (1991). Recognition of the highly conserved GTPase center of 23 S ribosomal RNA by ribosomal protein L11 and the antibiotic thiostrepton. *J. Mol. Biol.* **221**, 1257–1268.

Schmidt, F. J., Thompson, J., Lee, K., Dijk, J. & Cundliffe, E. (1981). The binding site for ribosomal protein L11 within 23 S ribosomal RNA of *Escherichia coli*. *J. Biol. Chem.* **256(23)**, 12301–12305.

Stark, M. & Cundliffe, E. (1979). On the biological role of ribosomal protein BM-L11 of *Bacillus megaterium*, homologous with *Escherichia coli* ribosomal protein L11. *J. Mol. Biol.* **134**, 767–779.

Suri, A. K. & Levy, R. M. (1993). Estimation of interatomic distances in proteins from NOE spectra at longer mixing times using an empirical two-spin equation. *J. Magn. Reson. ser. B,* **101**, 320–324.

Suri, A. K. & Levy, R. M. (1995). A relaxation-matrix analysis of distance-constraint ranges for NOEs in proteins at long mixing times. *J. Magn. Reson. ser. B,* **106**, 24–31.

Tate, W. P., Dognin, M. J., Noah, M., Stöffler-Meilicke, M. & Stöffler, G. (1984). The NH_2-terminal domain of *Escherichia coli* ribosomal protein L11. Its three-dimensional location and its role in the binding of release factors 1 and 2. *J. Biol. Chem.* **259**, 7317–7324.

Thompson, J., Cundliffe, E. & Stark, M. (1979). Binding of thiostrepton to a complex of 23 S rRNA with ribosomal protein L11. *Eur. J. Biochem.* **98(1)**, 261–265.

Vuister, G. W. & Bax, A. (1992). Resolution enhancement and spectral editing of uniformly ^{13}C enriched proteins by homonuclear broadband ^{13}C decoupling. *J. Magn. Reson.* **98**, 428–435.

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

Vuister, G. W., Wang, A. C. & Bax, A. (1993). Measurement of three-bond nitrogen-carbon *J* couplings in proteins uniformly enriched in ^{15}N and ^{13}C. *J. Am. Chem. Soc.* **115**, 5334–5335.

Wang, A. C. & Bax, A. (1995). Reparametrization of the Karplus relation for $^3J_{(H\alpha\text{-}N)}$ and $^3J_{(HN\text{-}C')}$ in peptides from uniformly $^{13}C/^{15}N$-enriched human ubiquitin. *J. Am. Chem. Soc.* **117**, 1810–1813.

Wang, Y.-X., Lu, M. & Draper, D. E. (1993). Specific ammonium ion requirement for functional ribosomal RNA tertiary structure. *Biochemistry,* **32**, 12279–12282.

Wishart, D. S. & Sykes, B. D. (1994). The ^{13}C Chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical shift data. *J. Biomol. NMR,* **4**, 171–180.

Xing, Y. & Draper, D. E. (1996). Cooperative interactions of RNA and thiostrepton antibiotic with two domains of ribosomal protein L11. *Biochemistry,* **35(5)**, 1581–1588.

Xing, Y., GuhaThakurta, D. & Draper, D. E. (1997). The RNA binding domain of ribosomal protein L11 is structurally similar to homeodomains. *Nature Struct. Biol.* **4(1)**, 24–27.

Edited by P. E. Wright

(Received 4 June 1997; received in revised form 8 August 1997; accepted 20 August 1997)

http://www.hbuk.co.uk/jmb

Supplementary material, consisting of five Tables, including a summary of NMR experiments and their parameters, a complete listing of resonance assignments, backbone and side-chain coupling constants, statistics comparing structures calculated with distance restraints derived from either the isolated spin pair approximation or standard (strong, medium, weak, very weak) restraint categories, and NOE and dihedral angle restraints in X-PLOR format, is available from JMB Online.