The EMBO Journal Vol.18 No.22 pp.6508–6521, 1999

The NMR structure of the 5S rRNA E-domain–protein L25 complex shows preformed and induced recognition

Matthias Stoldt, Jens Wöhnert, Oliver Ohlenschläger, Matthias Görlach[1] and Larry R.Brown

Abteilung Molekulare Biophysik/NMR-Spektroskopie, Institut für Molekulare Biotechnologie, Postfach 100813, 07708 Jena, Germany

[1]Corresponding author
e-mail: mago@imb-jena.de

M.Stoldt and J.Wöhnert contributed equally to this work

The structure of the complex between ribosomal protein L25 and a 37 nucleotide RNA molecule, which contains the E-loop and helix IV regions of the E-domain of *Escherichia coli* 5S rRNA, has been determined to an overall r.m.s. displacement of 1.08 Å (backbone heavy atoms) by heteronuclear NMR spectroscopy (Protein Databank code 1d6k). The interacting molecular surfaces are bipartite for both the RNA and the protein. One side of the six-stranded β-barrel of L25 recognizes the minor groove of the E-loop with very little change in the conformations of either the protein or the RNA and with the RNA–protein interactions occurring mainly along one strand of the E-loop duplex. This minor groove recognition module includes two parallel β-strands of L25, a hitherto unknown RNA binding topology. Binding of the RNA also induces conversion of a flexible loop to an α-helix in L25, the N-terminal tip of which interacts with the widened major groove at the E-loop/helix IV junction of the RNA. The structure of the complex reveals that the E-domain RNA serves as a preformed docking partner, while the L25 protein has one preformed and one induced recognition module.

Keywords: NMR structure/protein–nucleic acid complex/protein–RNA interaction/ribosomal protein/ribosomal RNA

Introduction

Double helical RNA molecules with a canonical A-form conformation pose a dilemma for specific recognition by RNA-binding proteins: sequence-specific hydrogen bond donor and acceptor groups that could govern specific recognition are occluded in the deep and narrow major groove. Conversely, the wider and more accessible minor groove presents an array of functional groups that is too uniform to promote specific recognition by proteins. Functional groups of the RNA can be rendered accessible to recognition by the incorporation of single stranded regions or hairpin loops as, for example, in the U1A and U2B″–U2A′ complexes (Oubridge *et al.*, 1994; Allain *et al.*, 1996; Price *et al.*, 1998) or in the recognition of the anticodon sequence by most aminoacyl-tRNA synthetases (Cusak, 1997; De Guzman *et al.*, 1998a). For an RNA duplex, the major groove can be modified by unpaired bases, unusual base pairs or base triples (e.g. Puglisi *et al.*, 1992, 1995; Battiste *et al.*, 1996; Kalurachchi *et al.*, 1997) and thereby becomes wide enough to accommodate interactions with helical, β-sheet or loop elements of the binding polypeptide chain (Patel, 1999). At the most complex level, tertiary interactions between different parts of an RNA molecule are required to create a surface suited for specific protein binding (Cusak, 1997; De Guzman *et al.*, 1998a; Conn *et al.*, 1999; Wimberly *et al.*, 1999). At present the universe of RNA–protein recognition motifs is far from fully explored. Given growing evidence for an important role of RNA–protein recognition in important biological processes, including signal transduction and control of gene expression (Siomi and Dreyfuss, 1997; Weiss, 1998), a better understanding of the structural basis of these interactions is highly desirable.

A large number of specific RNA–protein and RNA–RNA interactions govern the architecture of the ribosome, but due to the structural complexity of this ~2.3 MDa ribonucleoprotein particle, progress towards a complete high resolution picture of its structure has only been made recently (Ban *et al.*, 1998; Yonath and Franceschi, 1998). The 5S rRNA is the smallest of the ribosomal RNA molecules and one of the most conserved RNA sequences in nature. It is an essential component of the ribosome (Erdmann *et al.*, 1971; Hartmann *et al.*, 1988) and crosslinking to 23S rRNA (Sergeiev *et al.*, 1998; Osswald and Brimacombe, 1999) implicates it in functional domains of the 50S subunit. In *Escherichia coli* the 5S rRNA forms an RNA–protein complex with three ribosomal proteins: L5, L18 and L25 (Chen-Schmeisser and Garrett, 1977). The binding site for protein L25 is contained within the E-domain, i.e. nucleotides 70–106 of 5S rRNA (Douthwaite *et al.*, 1979, 1982; Huber and Wool, 1984). Recently, high resolution structures were reported for this RNA domain (Correll *et al.*, 1997; Dallas and Moore, 1997). The E-domain is characterized by an unusual duplex conformation, the so-called E-loop, with a consecutive stretch of seven non-Watson–Crick base pairs at its center. This prokaryote 5S rRNA E-loop is structurally distinct from the eukaryotic E-loop and related structural motifs (Wimberly *et al.*, 1993; Leontis and Westhof, 1998). The recently determined solution structure of L25 of *E.coli* (Stoldt *et al.*, 1998) showed a new topology for a ribosomal protein composed of a six-stranded, closed β-barrel with two pairs of parallel β-strands and two peripheral α-helices, and indicated that the RNA binding site includes a contiguous surface of four strands of β-sheet as well as a large, unstructured loop of the protein. Since both the protein and the RNA contained structural elements that have not been observed in other protein–RNA complexes, we have determined the NMR solution

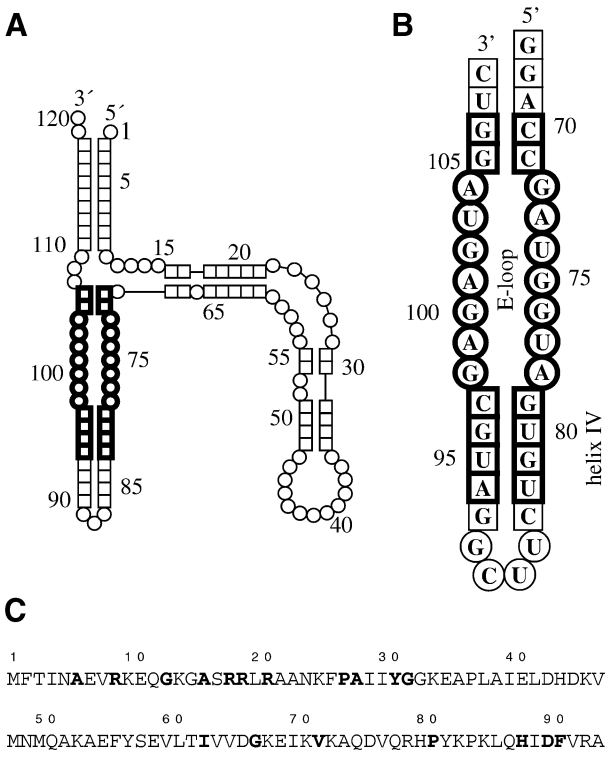

Fig. 1. (**A**) Schematic representation of the secondary structure of *E.coli* 5S rRNA and (**B**) of the 37 nt RNA molecule (5SE). The part of the 5SE molecule that corresponds to the native *E.coli* sequence is emphasized. (**C**) Amino acid sequence of *E.coli* L25. Conserved residues (Stoldt *et al*., 1998) are indicated in bold.

structure of the 22.7 kDa complex of protein L25 and an RNA molecule containing nucleotides 70–82 and 94–106 of the E-domain of *E.coli* 5S rRNA.

Results

Binding of protein L25 to the 5S rRNA E-domain

The 5SE RNA molecule (Figure 1B) has been used to determine the structure of the complex of ribosomal protein L25 (Figure 1C) with its cognate RNA, the E-domain of *E.coli* 5S rRNA. The 5SE RNA molecule includes the full E-loop plus neighboring helical stems, including the two GU base pairs in helix IV. The remaining residues of 5SE are a stable tetraloop sequence and nucleotides necessary for the *in vitro* production of isotope-labeled RNA (see Materials and methods). These latter residues are not in contact with L25 (see below) and are not discussed further.

The dissociation constant for the complex of L25 and 5S rRNA was reported to be ~67 nM (Spierer *et al.*, 1978) and the 5SE RNA molecule also forms a stable complex with L25, which is in slow exchange on the NMR time scale (Stoldt *et al.*, 1998; Wöhnert *et al.*, 1999a). We have compared the ^{1}H-^{15}N HSQC spectrum of L25 bound to 5SE with the ^{1}H-^{15}N TROSY-HSQC (Pervushin *et al.*, 1997) spectrum of L25 bound to the full 5S rRNA (Figure 2B). The virtual identity of backbone amide ^{1}H-^{15}N chemical shifts in these two spectra is very strong evidence that the conformation of bound L25 is not significantly affected by the use of the smaller 5SE molecule. On the other hand, the backbone amide ^{1}H-^{15}N chemical shifts for free L25 and for L25 bound to the 5SE molecule (Figure 2A) indicate that while many backbone amide groups have very similar chemical shifts and therefore presumably show largely unchanged conformations upon RNA binding, other backbone amide groups exhibit substantial changes. Those residues showing strong chemical shift differences between free and bound L25 probably include protein residues involved in RNA binding (Görlach *et al*., 1992; Stoldt *et al.*, 1998). Comparison of the imino NH resonances of 5SE with and without bound L25 (Figure 2C) shows for the most part only modest changes in the observed chemical shifts, suggesting that the conformation of 5SE is not substantially changed by the binding of L25. The most conspicuous difference in the two spectra is the observation for 5SE bound to L25 of two additional signals corresponding to the imino groups of U74 and G75. A ^{1}H-^{15}N TROSY-HSQC spectrum of the imino region of the 50 kDa L25–5S rRNA complex has also been recorded (Figure 2D). Although resonances from the many other residues of the complete 5S rRNA obscure many of the chemical shifts observed for the 5SE–L25 complex, resonances at positions close to the corresponding chemical shifts of U74, G75, U77, U103, G76, G102, U80, U95, G81 and G96 in the 5SE–L25 complex can be observed in the 5S rRNA–L25 complex. These results (Figure 2) suggest that the conformations of both L25 and the E-loop are similar in the 5SE–L25 and 5S rRNA–L25 complexes. Finally, no significant changes in the ^{1}H-^{15}N chemical shifts or linewidths of the imino (5SE) or amide (L25) groups were observed when the 5SE–L25 complex was titrated with $MgCl_2$ up to a concentration of 4 mM (data not shown). Thus, Mg^{2+} appears to be dispensable for stabilization of the bound RNA structure and for stable binding of the protein to its cognate RNA ligand.

Structure determination for the bound L25 protein

The ribosomal protein L25 was overexpressed in *E.coli* BL21(DE3) and purified in unlabeled, uniformly ^{15}N- and uniformly $^{15}N/^{13}C$-labeled forms (see Materials and methods). A nearly complete set of ^{1}H, ^{15}N and ^{13}C NMR assignments (98% backbone, 99% side chain CH) was obtained using multi-dimensional heteronuclear NMR spectroscopy for a complex consisting of $^{15}N/^{13}C$-labeled L25 and unlabeled 5SE. All of the observed chemical shifts for the protein fell within the usual ranges delineated in the DYANA library (Güntert *et al.*, 1997). A series of 3D heteronuclear NOE spectroscopy experiments in H_2O and D_2O, including experiments to distinguish inter- from intramolecular NOEs, were used to extract intramolecular distance constraints for the bound protein. The program DYANA (Güntert *et al.*, 1997) was used to calculate 100 structures for L25 in the bound state. The 20 structures with the lowest DYANA target functions for L25 bound to 5SE were used to characterize the protein structure. A superposition of the best 20 calculated structures is shown in Figure 3A and a summary of the experimental constraints is given in Table I.

The structure contains a closed six-stranded β-barrel and three peripheral α-helices with the topology β1(I4-V8)-α1(K14-A23)-β2(K25-Y31)-β3(P37-D43)-α2(H44-Q51)-α3(E55-S58)-β4(L61-V65)-β5(K69-Q78)-β6(H88-R93) (Figure 3B). The β-strands are regular except for a

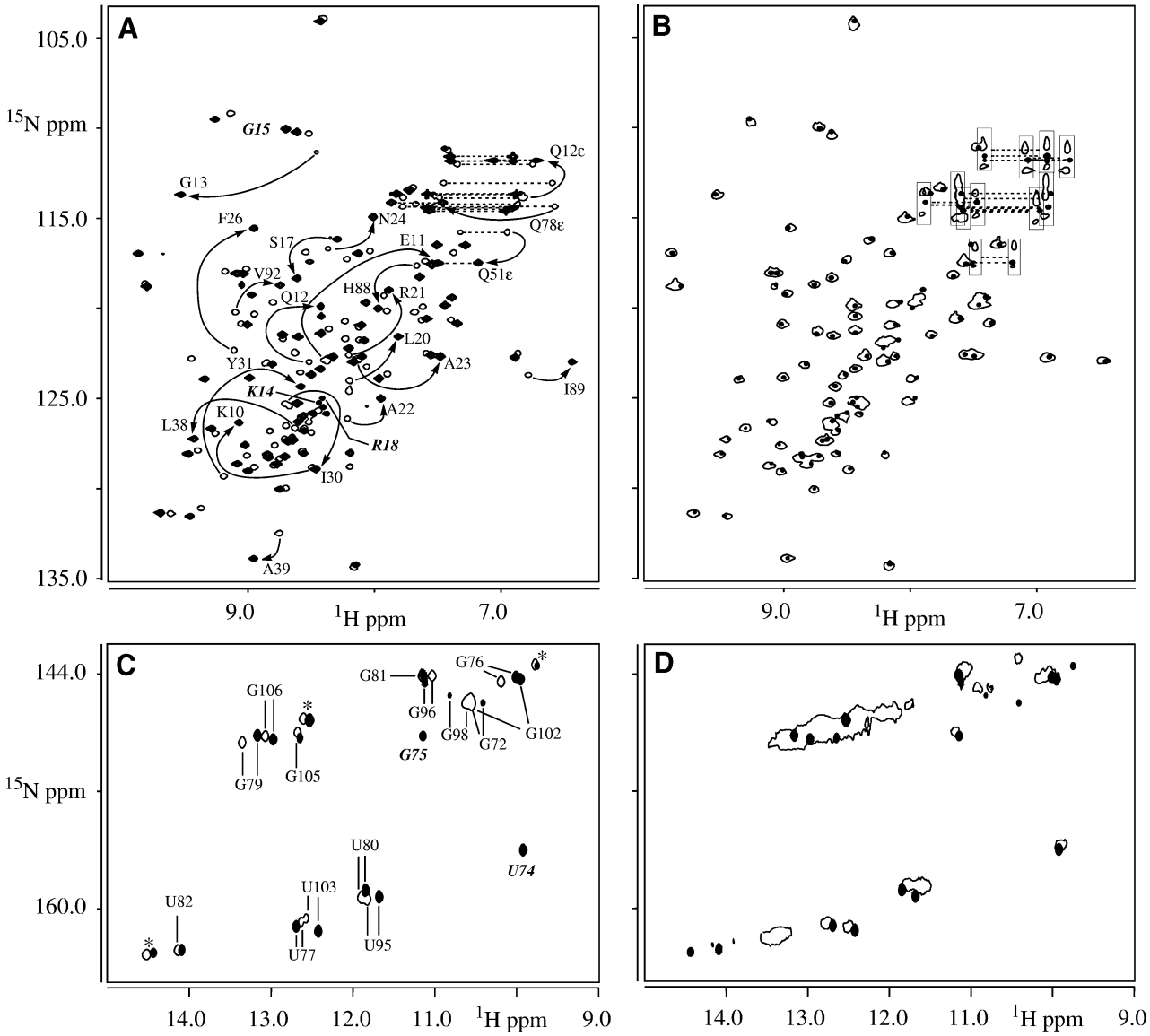

Fig. 2. Comparison of NMR spectra for free and bound states of L25 and RNA. (**A**) Superposition of ^{1}H-^{15}N HSQC spectra of L25 free (open contours) with L25 bound to 5SE (filled contours). Arrows connect the respective resonances of amides exhibiting the most pronounced chemical shift changes upon complex formation. Resonances observable only in the complex are indicated in italic. (**B**) Superposition of the ^{1}H-^{15}N HSQC spectrum of L25 bound to 5SE (filled contours) and of the ^{1}H-^{15}N TROSY-HSQC spectrum of L25 bound to 5S rRNA (open contours). Resonances in the ^{1}H-^{15}N TROSY-HSQC spectrum that are boxed designate the signals of the side chain amide groups. (**C**) Superposition of the nucleotide imino resonances for ^{1}H-^{15}N HSQC spectra of 5SE in the free (open contours) and L25-bound (filled contours) states at 25°C. Resonances visible only in the bound state of 5SE at 25°C are indicated in italic and resonances arising from non-native sequence elements of 5SE are marked by an asterisk. (**D**) Superposition of the nucleotide imino resonances of the ^{1}H-^{15}N HSQC spectrum of 5SE bound to L25 (filled contours) and of the ^{1}H-^{15}N TROSY-HSQC spectrum of 5S rRNA bound to L25 (open contours).

bulge at Gln75, Asp76, which leads the side chains of Gln75, Asp76 and Gln78 to all lie on the same exterior surface of the β-sheet. For the backbone ϕ,ψ angles, 97% lie within favored regions of the Ramachandran plot as determined with the PROCHECK-NMR program (Laskowski *et al.*, 1996).

The structures of bound L25 and free L25 are highly similar (Figure 3B and C). The six β-strands are unaltered in connectivity and form the same closed six-stranded β-barrel previously reported for free L25 (Stoldt *et al.*, 1998). Superposition of the β-barrel structures shows r.m.s. displacements of 1.0 Å (backbone heavy atoms) and 1.44 Å (all heavy atoms). The two short α-helices (α2 and α3 in the bound structure), which close one end of the β-barrel, are also essentially unaltered. Superposition of residues Thr3-Arg9 and Lys25-Ala94, i.e. excluding the disordered loop region (Arg9-Asn24) of free L25, gave r.m.s. displacements of 2.13 Å (backbone heavy atoms) and 2.87 Å (all heavy atoms), which are largely due to changes in the orientation of the loops connecting β2–β3 and β5–β6. The striking difference in the structures is that the large, unstructured loop connecting strands β1 and β2 of free L25 (residues Arg9 to Asn24) forms an α-helix for amino acids Lys14 to Ala23 (Figure 3B) for bound L25. In the 5SE–L25 complex, ^{15}N{^{1}H} NOEs show that all regions of the L25 backbone show similar relaxation behavior (Figure 3D), which is consistent with tumbling of a single globular domain, whereas for free

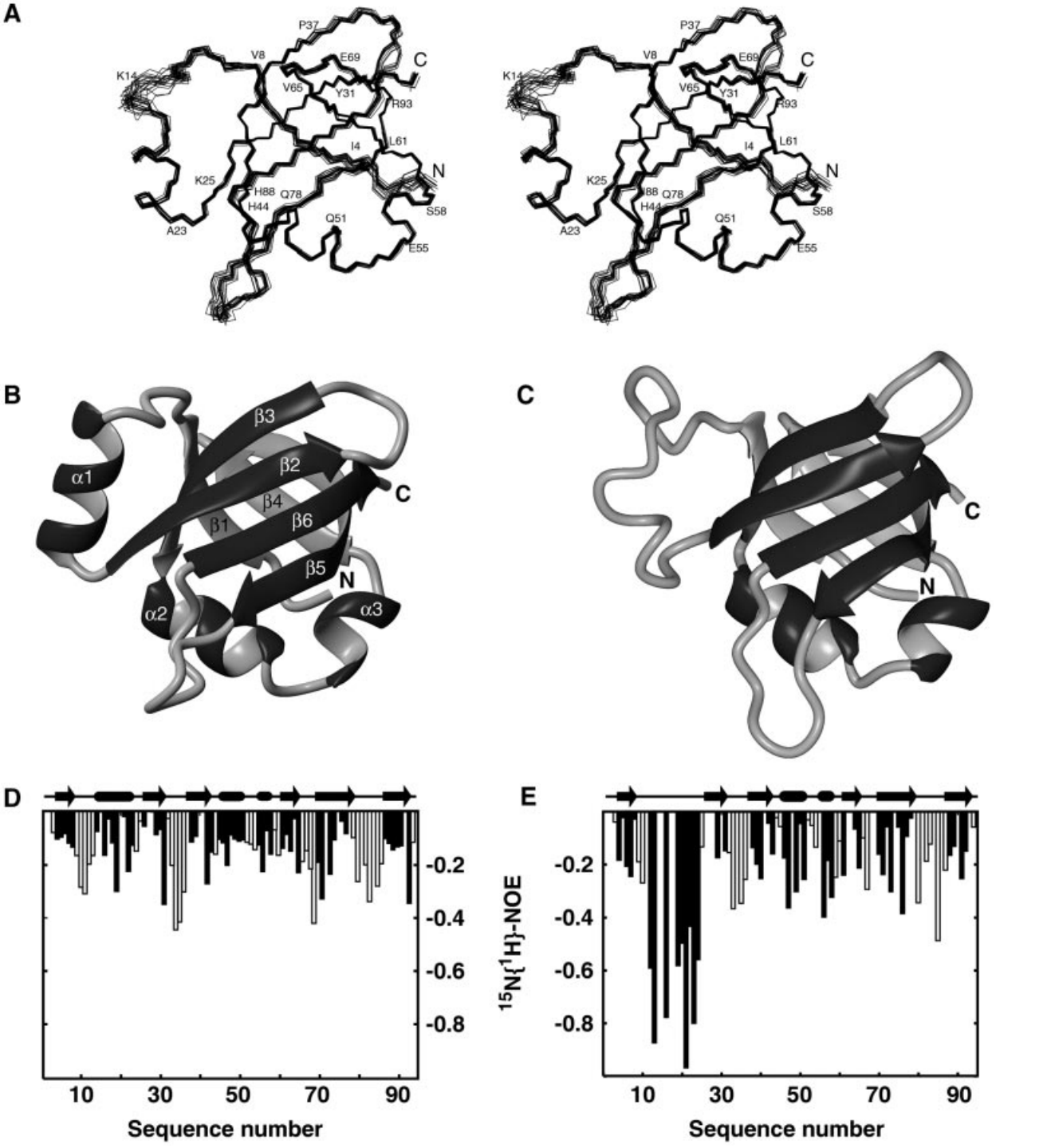

Fig. 3. Characterization of the ribosomal protein L25 bound to 5SE RNA. (**A**) Stereo view of the 20 NMR solution structures with the lowest DYANA target functions for L25 bound to 5SE RNA. The N- and C-termini and the residues at the boundaries of secondary structure elements are labeled. All heavy atoms of the backbone were used for the least square superposition. (**B** and **C**) Ribbon diagrams of the structures of L25 bound to 5SE RNA and of free L25, respectively. (**D** and **E**) $^{15}N\{^{1}H\}$ NOEs for the backbone amide groups of the 5SE–L25 complex and of free L25, respectively.

L25 a high degree of local flexibility is evident for residues Gln12 to Asn24 (Figure 3E). This is clear evidence that binding of 5SE RNA to L25 induces a flexible, unstructured loop to adopt a well defined helical conformation.

Structure of 5SE RNA in the bound state

5SE RNA was prepared in unlabeled, uniformly ^{15}N- and uniformly $^{15}N/^{13}C$-labeled forms by *in vitro* transcription and RNA–protein complexes were obtained as described (Stoldt *et al.*, 1998; Materials and methods). Using 3D heteronuclear through-bond correlation experiments and 3D heteronuclear edited NOE experiments, 97% of the ribose H1′, H2′, H3′, 99% of the base aromatic hydrogens and all observable imino/amino hydrogens (23% were not observable due to exchange with solvent) were assigned. The assignments for ribose H4′, H5′ and H5″ are incom-

Table I. Characterization of the 40 NMR structures of the 5SE–L25 complex with the lowest DYANA target functions after energy minimization

Distance and dihedral angle constraints	
NOE-derived upper distance limits	
RNA–RNA	436
protein–protein	2623
RNA–protein	121
Hydrogen bonds	
RNA	43
protein	28
Dihedral angle constraints	
RNA	98
Total number of constraints	3349
Structure statistics	
DYANA target function (Å^2)	3.4 ± 0.35
Residual NOE violations	
number ⩾ 0.1 Å	3 ± 1
max. (Å)	0.12 ± 0.01
Residual dihedral angle violations	
number ⩾ 2.5°	0
max. (deg)	1.43 ± 0.46
r.m.s. deviations from ideal geometry	
bond length (Å)	0.005 ± 0.0005
bond angles (deg)	1.58 ± 0.18
AMBER94 energy (kcal/mol)	
total	–734 ± 38
van der Waals	–863 ± 12
electrostatic	–1560 ± 28
Pairwise r.m.s. deviation between calculated and energy minimized structures for all heavy atoms (Å)	0.79 ± 0.07
r.m.s. deviations from mean structure (Å)	
Protein	
all backbone heavy atoms (residues 3–94)	0.45 ± 0.20
all heavy atoms (residues 3–94)	0.88 ± 0.24
RNA (all heavy atoms)	
residues 70–82 and 94–106	1.12 ± 0.30
E-loop (residues 72–78 and 98–104)	0.66 ± 0.19
Protein–RNA complex	
all backbone heavy atoms (protein residues 3–94; RNA residues 70–82 and 94–106)	1.08 ± 0.34
all heavy atoms (residues 3–94, 70–82 and 94–106)	1.20 ± 0.32

plete due to the poor dispersion of their resonances even in 3D NMR experiments. The HNN-COSY experiment (Dingley and Grzesiek, 1998) was used to identify the donor and acceptor groups for NH...N type hydrogen bonds for Watson–Crick base pairs in the helical stems and for two reverse Hoogsteen base pairs in the E-loop (Wöhnert *et al.*, 1999a). 2D and 3D heteronuclear NOE experiments in H_2O and D_2O were used to extract intra- and intermolecular distance constraints. Torsion angle constraints for the backbone torsion angles α and ζ were derived from chemical shifts in a ^{31}P NMR spectrum of the bound 5SE RNA (Gorenstein, 1984; Varani *et al.*, 1991). Local conformational analysis with the FOUND module (Güntert *et al.*, 1998) of DYANA indicated that for nucleotides A69–C83 and G93–U107, the glycosidic torsion angles χ are in the 'anti' range. Only C3′-*endo* sugar conformations are observed. For those portions of the 5SE molecule involved in the interaction with L25, i.e. residues C71–U82 and A94–G105 (see below), a superposition of the 20 calculated structures with the lowest DYANA target functions is shown in Figure 4A and a summary of the experimental constraints is given in Table I.

With the exception of the tetraloop sequence and the unpaired G67 at the 5′-end, all other nucleotides are involved in base pairing interactions and the bound form of 5SE displays a double helical structure throughout. Canonical Watson–Crick base pairs are formed by nucleotides G68–C71/G105–C108, G79/C97 and U82–C83/G93–A94 in agreement with the observation of G N1H1/C N3 and U N3H3/A N1 correlations for the GC and AU base pairs, respectively, in the HNN-COSY experiment (Wöhnert *et al.*, 1999a). The most striking features of the conformation of 5SE when bound to L25 are the base pairing and stacking geometries for the seven central, non-canonical base pairs together with the two flanking Watson–Crick GC base pairs. The central region of the E-loop shows three unusual base pairs (U74/G102, G75/A101, G76/G100; Figure 4B). In the U74/G102 base pair a bifurcated hydrogen bond is observed between U74 O4 and the imino and amino groups of G102. G75 and A101 form a base pair with only a single hydrogen bond between the G75 O6 carbonyl and the A101 amino group. A bifurcated hydrogen bond is present in the G76/G100 base pair between the G76 imino and amino groups and G100 O6. The two ends of the E-loop have symmetrical sequences, i.e. C71–A73 paired with U103–G105 and C97–A99 paired with U77–G79 (Figure 1) and form largely identical structures (Figure 4C). The flanking base pair G79/C97 (G105/C71) shows normal Watson–Crick base pairing. The base pair A78/G98 (A104/G72) shows (anti–anti) sheared GA base pairing with hydrogen bonds between the A amino groups and the G N3 nitrogens as well as between the G amino groups and the A N7 nitrogens. The U77/A99 (U103/A73) base pair shows a reversed Hoogsteen geometry and displays hydrogen bonds between the A amino group and the U O2 as well as between the U imino group and the A N7 nitrogen. For each of the two sheared GA/reversed Hoogsteen AU motifs, an unusual cross-strand stacking interaction is observed for the two adenines in the consecutive base pairs (Figure 4C). In helix IV, the two GU base pairs adopt the usual wobble pair geometry. For the two guanines in the GU tandem mismatch this leads to purine/purine cross-strand stacking (Figure 4D). In comparison with regular A-form RNA the minor groove of the E-loop is widened by ~2 Å and the major groove is narrowed by ~2 Å around the central three base pairs, while at the E-loop/helix IV junction the major groove is widened between the base pairs A78/G98 to G81/U95.

All of the unusual structural features noted above for the E-loop and tandem GU base pairs in helix IV of 5SE RNA bound to L25 have previously been observed in crystallographic and NMR studies of the E-domain of *E.coli* 5S rRNA without bound protein (Correll *et al.*, 1997; Dallas and Moore, 1997). The r.m.s. deviation for all heavy atoms between the bound RNA (C71–G79, C97–G105) and the crystal structure of the unbound RNA (Protein Databank entry 354d; Correll *et al.*, 1997) amounts to 1.45 ± 0.16 Å. This indicates that the E-loop does not undergo any major changes in conformation upon binding of L25.

Structure of the protein–RNA complex

An example of the quality of the NOE spectra used to extract intermolecular distance constraints is shown in

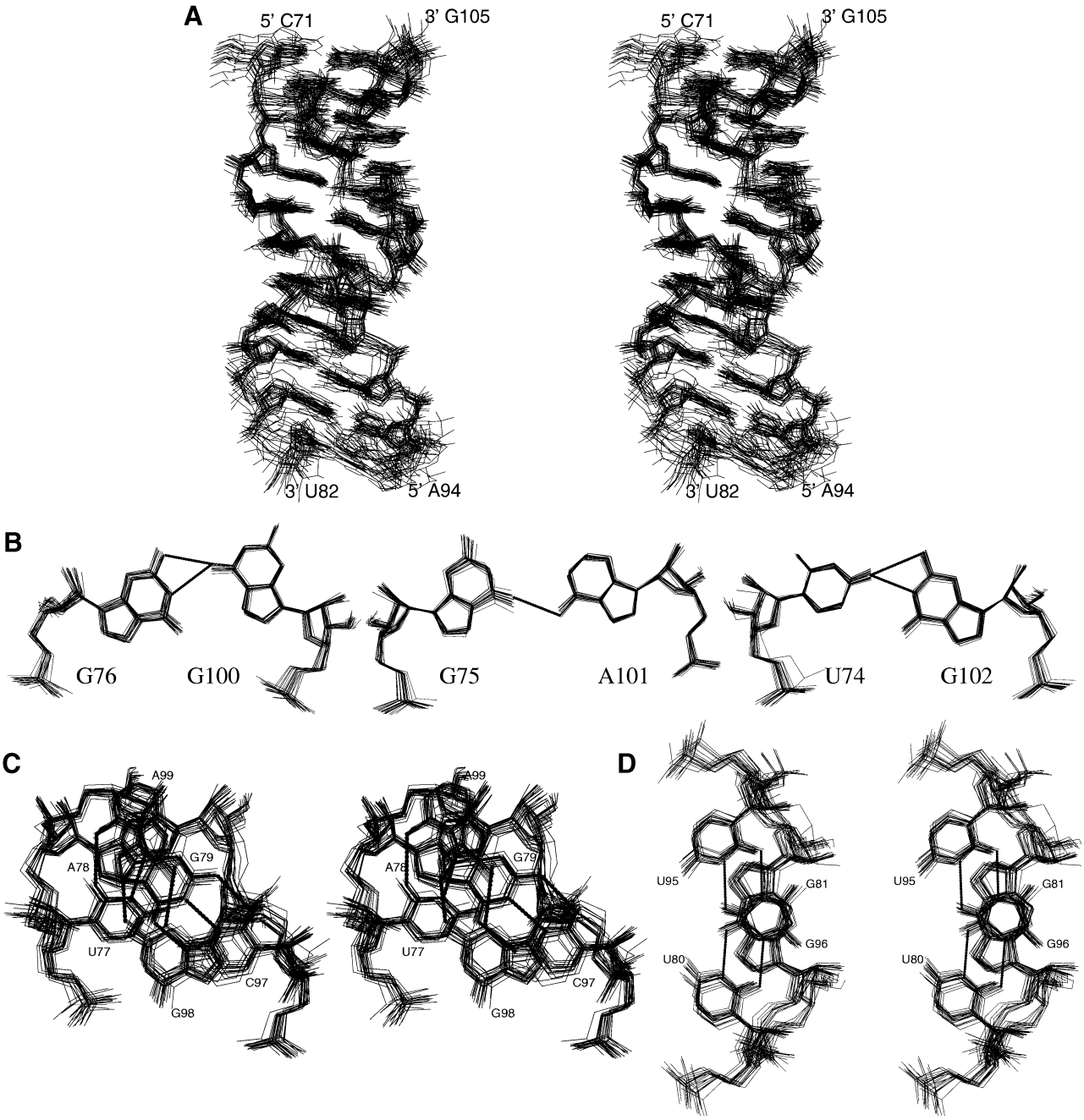

Fig. 4. Structure of 5SE bound to ribosomal protein L25. (**A**) Stereo view of the ensemble of the 20 NMR solution structures of 5SE bound to L25 with the lowest DYANA target functions. The heavy atoms of nucleotides 71–82 and 94–105 are shown and were used for least square superposition of the structures. (**B**) Base pairing in the central part of the E-loop of 5SE. Local superposition of the calculated structures for the base pairs G76/G100, G75/A101 and U74/G102 are shown. Hydrogen bonds are indicated by dotted lines. (**C**) Stereo view showing the U77/A99, A78/G98, G79/C97 structural unit with the reverse Hoogsteen (U77/A99) and sheared (A78/G98) base pairs and the A78/A99 cross strand stacking. (**D**) Stereo view showing the tandem U80/G96 and G81/U95 wobble base pairs and the G81/G96 cross strand stacking.

Figure 5. At the present level of refinement, 121 unambiguously assigned intermolecular NOEs have been identified. The input data for structural computations for the 5SE–L25 complex used distance constraints for these intermolecular NOEs combined with the intramolecular distance and torsion angle constraints obtained individually for the bound protein and the bound RNA. The protein and RNA molecules were joined by a linker sequence consisting of 160 linkage units and all DYANA structure computations were started with completely random torsion angles. Two hundred structures were calculated and the 40 structures with the lowest DYANA target functions were further refined by restrained energy minimization using the AMBER94 force field (Cornell *et al.*, 1995) as implemented in OPAL (Luginbühl *et al.*, 1996). A superposition of the resulting best 20 structures is shown in Figure 6A and a summary of the experimental constraints and structural statistics for the best 40 structures is given in Table I.

The structure of the complex shows that the 5SE RNA and the L25 protein interact through a combination of two different structural motifs (Figure 6B). First, the surface of one side of the β-barrel of protein L25 contacts the

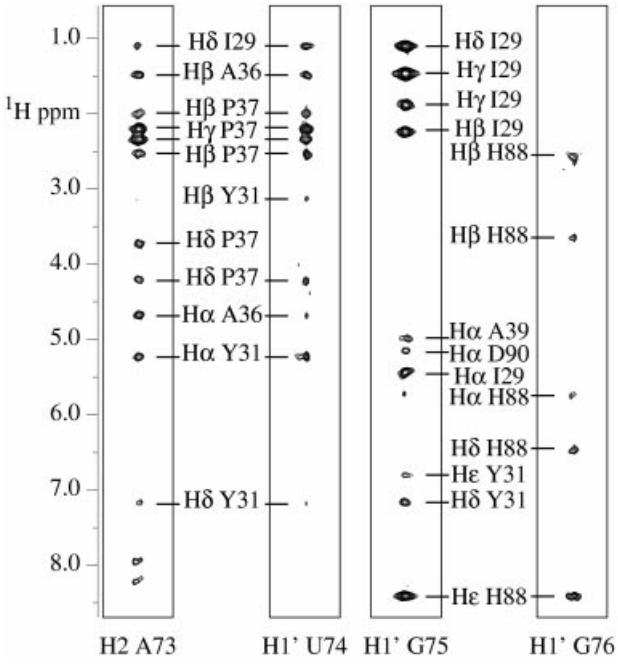

Fig. 5. Intermolecular NOEs between L25 and 5SE. F1–F3 strips of a 3D ^{13}C-F1-filtered,^{13}C-F3-edited NOESY-HSQC (80 ms mixing time) spectrum in D_2O of a sample containing ^{13}C,^{15}N-labeled RNA and unlabeled protein showing NOEs from the protein to the H2 of A73, H1′ of U74, H1′ of G75 and H1′ of G76.

widened minor groove of the E-loop of the RNA at base pairs A73/U103 to U77/A99. Secondly, the N-terminal tip of the inducibly formed α1 α-helix of L25 contacts the widened major groove of the RNA formed by the A99/U77, G98/A78, C97/G79 structural unit and by the adjacent tandem GU base pairs (G96/U80 and U95/G81) in helix IV. Upon complex formation, a solvent-accessible surface of ~1720 Å^2 is buried, which is similar in size to buried surfaces in comparable RNA–protein complexes (Allain *et al*., 1996; DeGuzman *et al*., 1998b; Ryter and Schultz, 1998). The interacting surfaces include areas of matching positive and negative electrostatic potential as well as extensive non-charged areas (Figure 6C).

For the six-stranded β-barrel of L25, amino acid residues in four strands (β2, β3, β5 and β6) show contacts to the RNA (Figures 7 and 8A). The central two (β2 and β6) of these four strands run parallel to each other (Figures 6B and 8A). Most contacts made by the four-stranded β-sheet surface are directed towards one strand of the E-loop, involving the residues A73, U74, G75, G76 and U77 (Figures 5, 7 and 8A). Strands β2 and β3 contribute only side chains of hydrophobic character to the binding surface: Pro27, Ile29, Tyr31 on strand β2, and Pro37 and Ala39 on strand β3. These residues contact a largely hydrophobic part of the minor groove of the E-loop (Figures 6C and 7). Strand β6 adds the hydrophobic side chain of Val92 and the side chains of Asp90 and of His88 to the binding surface, while strand β5 contributes only hydrophilic or charged side chains (Gln75, Asp76 and Gln78) to the contact surface. The hydrophobic face of strands β2 and β3 contributes a number of van der Waals contacts to the interaction surface (Figure 7). The methylene groups of Pro37 direct a hydrophobic surface towards H2 of A73 and H1′ of U74. The side chains of Ile29 and Ala39 are next to the ribose of G75, and Pro27 is in close proximity to the ribose of G76. The aromatic ring of Tyr31 is packed perpendicularly against the edge of the A73/U103 base pair and the side chain of Val92 of β6 points towards the ribose methylene group of A104. Thus, all hydrophobic amino acid side chains that lie in the interaction surface of the β-barrel are buried upon RNA binding (Figure 8A). Strands β5 and β6 show a largely hydrophilic interaction surface for which, at the present level of refinement, three base specific interactions with E-loop nucleotides are seen (Figures 7 and 8B). The side chain of Gln78 of strand β5 is potentially hydrogen-bonded to the base N3 of G76. The side chain carboxyl group of Asp90 of strand β6 is hydrogen-bonded to the base amino group of G75 in almost all of the calculated structures. In addition, the ribose 2′ hydroxyl group of G75 is hydrogen-bonded to the side chain of His88, which is stacked upon the ribose moiety of G76 rather than upon a nucleotide base. Finally, the side chain amide of Gln75 is close to the phosphate group of A104 and the side chain carboxyl of Asp76 approaches the minor groove edge of the G102 base, but specific hydrogen bonds could not be determined with certainty at the present level of refinement.

The newly formed α1 α-helix, together with its flanking N-terminal residues, contacts the RNA only through hydrophilic/charged side chains, including those of Arg9, Lys14, Ser17 and Arg18, at the junction of the U77/A99, A78/G98, G79/C97 structural unit and the tandem U80/G96, G81/U95 base pairs in helix IV (Figures 7 and 8C). Helix α1 inserts the side chain of its N-terminal Lys14 into the major groove of helix IV at the position of the G79/C97 base pair where the groove is widened due to the cross strand purine stacks A99/A78 of the E-loop and G96/G81 of helix IV on either side of the base pair G79/C97. The side chain of Lys14 showed broadened lines in the NMR spectra, which has limited our ability to determine its exact location. The terminal $N^{\zeta}H_3^+$ group fits into a pocket where it is near to the O6 of the base of G79, which is also involved in normal Watson–Crick hydrogen bonding to C97. Interactions of the Lys14 $N^{\zeta}H_3^+$ group with the A78 phosphate group or the base carbonyl groups of the wobble base pair U80/G96 can not be excluded at the present level of refinement (Figure 8C). The side chain of Arg9 is hydrogen-bonded to the G76 phosphate group (Figure 8B). The side chains of Ser17 and Arg18 are directed towards the phosphate backbone of U77 and U95, respectively, and the side chain amide hydrogens of Gln12 show NOEs to the ribose hydrogens of G75, placing them close to the sugar–phosphate backbone (Figure 8C).

The structure of the 5SE–L25 complex is consistent with the observation of a number of slowly exchanging hydrogens. In the protein, the rates of exchange of backbone amide hydrogens with solvent correspond well with the hydrogen bonding patterns in the NMR structure of L25, i.e. the backbone amide hydrogens of the β-sheets all exchange slowly. Overall backbone amide exchange rates are reduced strongly (~40-fold) relative to the free L25 molecule, indicating that binding to 5SE stabilizes the protein structure. The backbone amide groups of Lys14, Gly15 and Arg18 become detectable due to the formation of the α1 α-helix upon RNA binding. The structure of the complex also provides plausible explan-

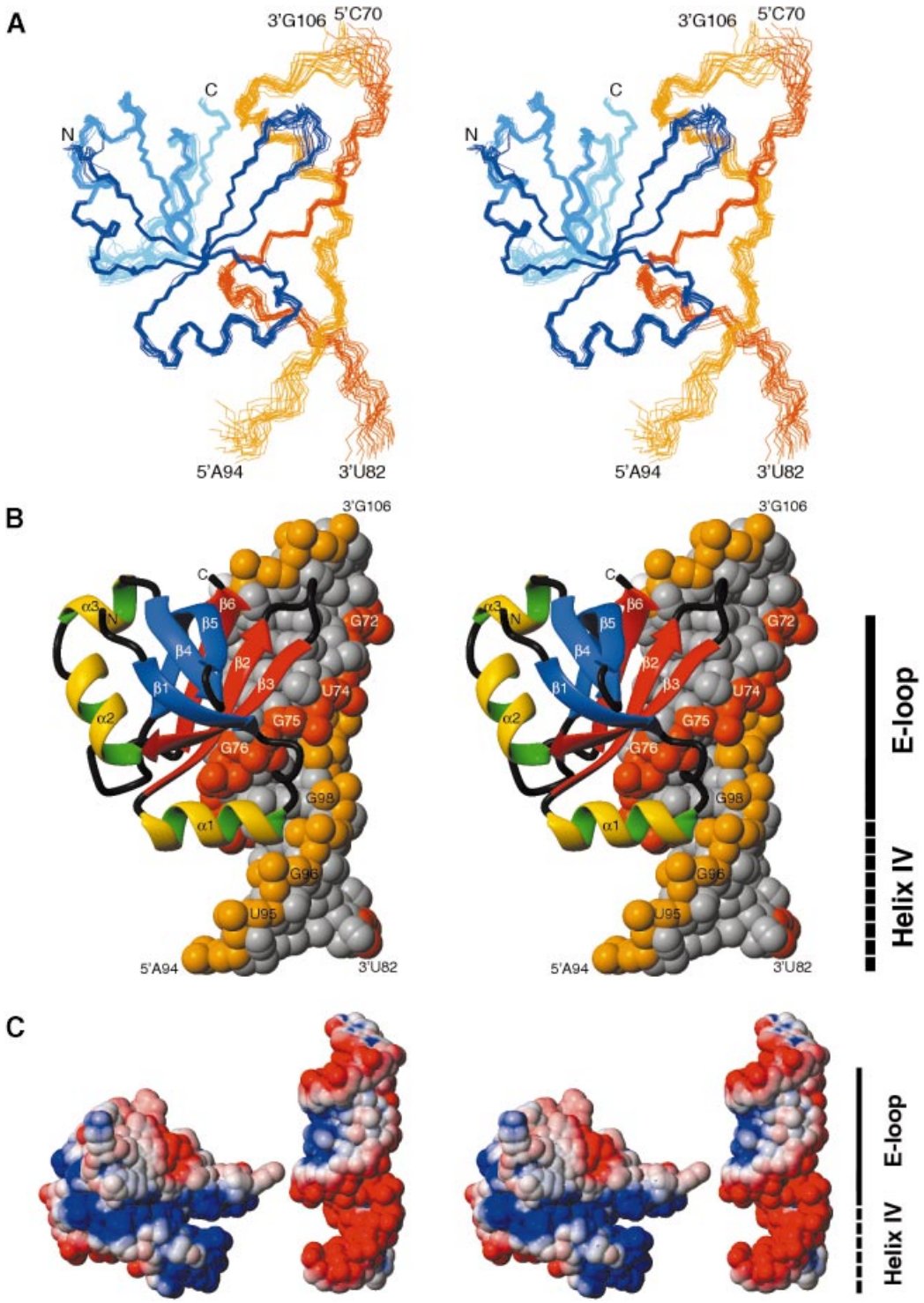

Fig. 6. NMR solution structure of the 5SE–L25 complex. (**A**) Stereo view of the ensemble of the 20 NMR solution structures of the 5SE–L25 complex with the lowest DYANA target functions after energy minimization. The backbone atoms of L25 and of the nucleotides 70–82 and 94–106 of 5SE are shown. Heavy atoms of residues 3–94 of L25 and of nucleotides 70–82 and 94–106 were used for least square superposition of the structures. (**B**) Stereo view of a schematic of the 5SE–L25 complex structure. L25 is shown as a ribbon diagram and 5SE as a CPK representation. (**C**) Surface representation of the electrostatic potential of the binding surfaces for L25 and 5SE. The two molecules of the complex were moved apart horizontally. L25 was rotated by –60° and 5SE was rotated by +60° about a vertical axis in the plane of the page in order to provide a complete view of the two surfaces. Negative charges are shown in red, positive in blue and uncharged regions in white.

ations for the observation of a number of more rapidly exchanging hydrogens that could only be detected in the 5SE–L25 complex. These include the imino hydrogens of U74 and G75 of the RNA, the $N^{\delta}H$ of His88, the $O^{\eta}H$ of Tyr31 and the $N^{\zeta}H_3^+$ group of Lys14, all of which are buried at the binding interface. Finally, some imino hydrogens continued to show broad linewidths in the complex, e.g. G72 and G98 in the E-loop, which is

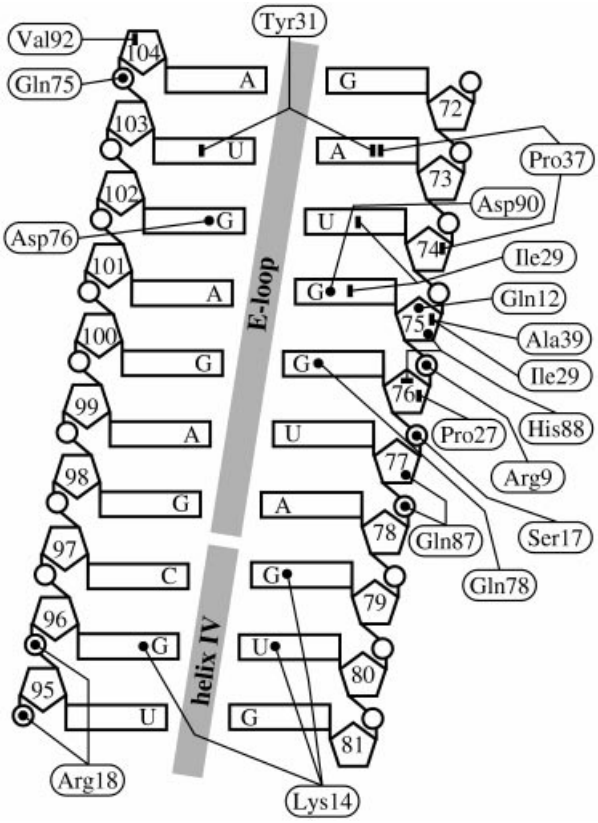

Fig. 7. Schematic depiction of the interaction between L25 and 5SE RNA. Dots indicate hydrophilic/charged interactions and filled bars designate hydrophobic/van der Waals contacts.

consistent with the continued accessibility of these hydrogens to water in the complex.

Discussion

The overall structure (Figure 6B) is in agreement with a substantial body of biochemical and biophysical data obtained previously. The binding site of the ribosomal protein L25 on the RNA (Figure 6B) is at the center of the binding region proposed on the basis of RNase protection studies (Douthwaite *et al*., 1979, 1982; Huber and Wool, 1984). Furthermore, the conversion of a large flexible loop from Arg9 to Asn24 to a stable α-helical secondary structure (Lys14 to Ala23) upon complex formation (Figure 3) and the binding of this α1 α-helix of L25 to the E-loop/helix IV junction of the E-domain (Figure 6) is consistent with the reduced trypsin susceptibility of the N-terminal part of L25 upon E-loop binding (Newberry and Garrett, 1980). Our present experimental observations (Figure 2C and D) as well as the similarity between the conformation observed for 5SE with bound L25 (Figure 4) and the conformation observed for a free E-domain RNA molecule (Correll *et al.*, 1997; Dallas and Moore, 1997) are also consistent with chemical probing experiments (Garrett and Noller, 1979; Ciesiolka *et al.*, 1992). Most importantly, the chemical shifts of the backbone amide groups of protein L25 in complex with either the 5SE RNA or the entire 5S rRNA (Figure 2) are virtually identical. Taken together, these observations provide strong evidence that the 5SE RNA construct contains the full cognate binding site for L25 and that the present structure is relevant to the overall architecture of the complex formed between 5S rRNA and ribosomal proteins L5, L18 and L25. Additional features might emerge from future studies involving the other two proteins, i.e. L18 and/or L5, although biochemical data suggest that there is at most a minor synergistic effect on the RNA structure upon binding of L25 together with L18 (Baer *et al.*, 1977; Ciesiolka *et al.*, 1992).

Many of the amino acid residues of L25 showing large ^{1}H-^{15}N chemical shift changes upon complex formation (Stoldt *et al.*, 1998) are found to be directly involved in the RNA–protein contacts, i.e. Gln12, Ser17, Ile29, Tyr31, Ala39, Gln75, Asp76, His88, Asp90 and Val92 (Figure 7). Other residues exhibiting significant chemical shift perturbations (Lys10, Glu11, Gly13, Arg19–Ala22, Asn24, Phe26, Glu41, His44, Lys85) do not directly contact the 5SE RNA, but are located at or in proximity to locations in the three-dimensional structure where structural changes in the protein are necessary for complex formation (Figure 3).

The overall arrangement of the 5SE–L25 complex is clearly determined by the topological locations of two bipartite structural elements in both L25 (the surface of strands β2, β3, β5, β6 and the α1 α-helix) and the 5SE RNA (the E-loop and the distorted helix IV). The RNA structure, with its unique ensemble of unusual base pairs in the E-loop and the tandem GU base pairs in helix IV, shows very little change in conformation upon complex formation and therefore constitutes a preformed recognition element for the protein. Likewise, the β-sheet surface of L25 shows little change upon RNA binding and provides a preformed recognition element for the RNA ligand. In contrast, the formation of the α1 α-helix of L25 upon binding at the E-loop/helix IV junction is clearly an induced-fit interaction. These considerations suggest that while the interaction of the α1 α-helix of L25 with the E-loop/helix IV junction of the 5SE RNA may serve to anchor the position of L25 along the length of 5SE, e.g. relative to the symmetrical G105/C71, A104/G72, U103/A73 structural unit at the other end of the E-loop, the detailed orientation of L25 relative to the 5SE RNA is probably determined by the interaction of the two preformed recognition elements, i.e. the β-sheet surface of L25 and the widened minor groove of the E-loop of 5SE.

On the basis of the structure of the free E-domain RNA, it was suggested that the intricate array of hydrogen bond donor and acceptor groups displayed in the minor groove of the E-loop by the central three unusual base pairs (U74/G102, G75/A101, G76/G100) could potentially provide a distinctive binding surface for L25 (Corell *et al.*, 1997; Dallas and Moore, 1997). Plausible interactions between the protein and the three nucleotide bases of G75, G76 and G102 are apparent in the complex structure. These involve the side chain of Asp90 and the exocyclic amino group of G75, the side chain of Gln78 and N3 of G76, and possibly the side chain of Asp76 and the exocyclic amino group of G102 (Figure 8B). The latter interaction might involve a water-mediated hydrogen bond since a water molecule was found to mediate the hydrogen bonding between the U74 NH and the exocyclic amino group of G102 in the free E-domain RNA molecule (Correll *et al.*, 1997). The network involving these base specific interactions together with the hydrogen bonds between

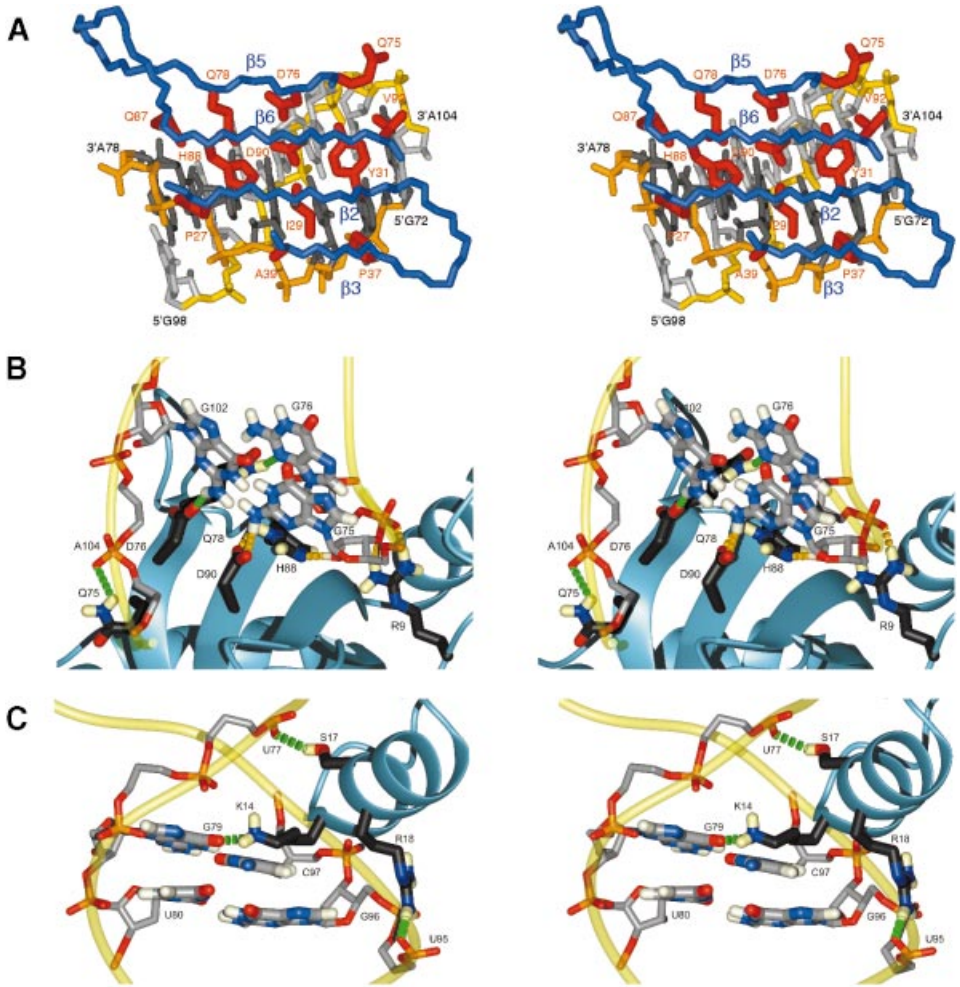

Fig. 8. Stereo views of the intermolecular interactions in the mean structure of the 5SE–L25 complex. (**A**) The interaction of strands β3, β2, β6 and β5 of the β-barrel of L25 with the minor groove of the E-loop of 5SE. Amino acid side chains of L25 that contact the RNA are shown in red. The protein backbone is shown in blue, the RNA bases in gray and the RNA backbone in yellow. (**B**) The network of interactions involving the amino acid side chains of Arg9, Gln75, Asp76, Gln78, His88 and Asp90 and the nucleotides G75, G76, G102 and A104 (phosphate only) is shown (yellow, definite hydrogen bonds; green, potential hydrogen bonds for which exact donor/acceptor groups were not fixed at the present level of refinement). The protein backbone is shown as a light blue ribbon diagram and the RNA backbone as a light yellow tube representation. (**C**) The network of interactions involving the amino acid side chains of Lys14, Ser17 and Arg18 and the nucleotides U77 (phosphate only), G79, U80, U95 (phosphate only), G96 and C97 is shown. The color scheme is the same as in (B).

the $N^{\varepsilon 2}$ of His88 and the 2′OH of G75 and between the guanidino group of Arg9 and the phosphate of G76 (Figure 8B) is formed at the center of the E-loop. Two potential additional interactions between the β-sheet of L25 and the RNA backbone, which involve the side chains of Gln75 and Gln87 with the phosphate groups of A104 and A78, respectively, are situated at the upper and lower end of the E-loop and are symmetrically placed with respect to the two C71/G105 (C97/G79), G72/A104 (A78/G98), A73/U103 (U77/A99) structural units. These 'guiding' interactions are complemented by a large number of hydrophobic and van der Waals contacts observed between hydrophobic amino acid side chains (Pro27, Ile29, Pro37, Tyr31, Ala39, His88, Val92) of the β-sheet and a subset (A73, U74, G75, G76, A104) of the nucleotides forming the unusual base pairs of the E-loop and which are centered along one strand (A73–G76) of the E-loop (Figure 7). Much of this interaction involves contacts with backbone ribose groups. The array of hydrophobic side chains involved in the RNA–protein interaction closely matches the E-loop (Figure 6C), the shape of which is determined by its array of seven unusual base pairs. Furthermore, the hydrophobic surface of 5SE RNA, which interacts with strands β2, β3 of L25, is flanked by regions of positive potential in the minor groove at the level of the unusual base pairs U74/G102, G75/A101, G76/G100 and negative electrostatic potential along the phosphate backbone, which interact with complementary regions of the protein surface (Figure 6C). This suggests that shape recognition between preformed structures is a significant contributing element to the specific molecular recognition between protein L25 and the E-domain of 5S rRNA.

Shape-complementary recognition (Wimberly *et al.*, 1999) of a flattened minor groove by hydrogen bonding and hydrophobic interactions has been described very recently for the ribosomal protein L11–RNA complex. However, in L11 α-helical and loop elements of the C-terminal domain mediate these interactions, which include base pair and RNA backbone recognition by the protein backbone (Conn *et al.*, 1999; Wimberly *et al.*,

1999). The recognition of the E-loop by the β-sheet of L25 is radically different to other known RNA–protein structures involving β-sheet interactions. For the four-stranded antiparallel β-sheet of the canonical RNP motif (Nagai *et al.*, 1990; Hoffman *et al.*, 1991; Wittekind *et al.*, 1992) examples include the snRNP U1A in complex with the U1 snRNA hairpin II or with the 3′-UTR of its own mRNA, and the ternary complex consisting of proteins snRNP U2B″–U2A′ and the stem–loop IV of the U2 snRNA. Specific recognition in these complexes involves loop elements (Oubridge *et al.*, 1994; Allain *et al.*, 1996) or an auxiliary protein (Price *et al.*, 1998). The hallmarks are extensive hydrogen bonding and hydrophobic interactions with single stranded RNA, including the stacking of highly conserved aromatic side chains with splayed out bases. In contrast, in the double-stranded 5SE–L25 complex, no stacking of an aromatic side chain with a base is observed and the hydrophobic surface of L25 includes many interactions with the ribose moieties of the RNA backbone.

The interaction of the N-terminal tip of the newly formed α1 α-helix of L25 with the modified major groove at the E-loop/helix IV junction mainly involves hydrophilic and charged side chain to RNA interactions. The arrangement of a central Watson–Crick base pair flanked by a sheared A/G-reverse Hoogsteen AU mismatch on one side and by tandem wobble GU pairs on the other side, each with a cross strand purine–purine stack, opens the major groove (Correll *et al.*, 1997; Dallas and Moore, 1997). Contacts to the E-loop/helix IV junction by the α1 α-helix are made through location of the side chain $N^{\zeta}H_3^+$ group of Lys14 near the O6 of G79 of the G79/C97 Watson–Crick base pair, location of the guanidino group of Arg18 near the phosphate group of U95 and interaction of the side chain of Ser17 with the phosphate of U77. Interestingly, binding of a lysine residue located at the N-terminal tip of an α-helix has been observed in the structurally distinct dsRBD–dsRNA complex (Ryter and Schultz, 1998) where a widened major groove is induced by cross strand stacking of two purines at the interface of two RNA molecules in the crystal but there the lysine contacts the phosphate backbone in a sequence-unspecific manner. The binding-induced formation of the α1 α-helix observed here is reminiscent of the adaptive recognition of RNA by peptides and protein modules, which has become evident recently mainly from structural studies of viral and bacteriophage systems (Frankel and Smith, 1998; Weiss and Narayana, 1998b; Patel, 1999). A common theme is that the peptide chain adopts defined α-helical (Battiste *et al.*, 1996), 3_{10}-helical (De Guzman *et al.*, 1998b) or β-hairpin structures (Puglisi *et al.*, 1995) upon binding to the RNA ligand, which in turn also adapts its conformation. The binding occurs in the major groove, which is made available for binding by looping out of bases, mismatched base pairs and base triples (Puglisi *et al.*, 1995; Battiste *et al.*, 1996; De Guzman *et al.*, 1998b; Patel, 1999). Some of the interactions observed for the peptide–RNA complexes show resemblance to the binding of L25 to the major groove edge of the G79/C97 and A78/G98 base pairs. Moreover, an adaptive transition/reorientation within an RNP motif RNA-binding domain has been observed for the snRNP U1A–3′UTR complex (Allain *et al.*, 1996) but the structurally rigid behavior of the 5SE RNA is a major difference to the peptide–RNA complexes and also to the U1A–RNA complex.

It has been suggested that as ancient proteins, ribosomal proteins might define families of protein folds, which are also used in other protein–RNA interactions (Bycroft *et al.*, 1997). The family represented by Asp-/Phe-/Lys-tRNA synthetases (Cavarelli *et al.*, 1993; Mosyak *et al.*, 1995; Onesti *et al.*, 1995), the cold shock domain proteins (Newkirk *et al.*, 1994; Schindelin *et al.*, 1994) and the ribosomal proteins S1 and S17 (Jaishree *et al.*, 1996; Bycroft *et al.*, 1997) seems to be one good example. Sequence homologues of L25 include CTC stress proteins (Gryaznova *et al.*, 1996) and the topology of the β-barrel of L25 has so far only been observed in glutaminyl-tRNA-synthetases (GlnRS; Rould *et al.*, 1991; Stoldt *et al.*, 1998), albeit in the latter case with very little sequence homology to L25. The sequence homologies observed between L25 (and equivalent ribosomal proteins from other organisms) and the CTC stress proteins include many conserved residues (Stoldt *et al.*, 1998), which the 5SE–L25 structure now confirms to be involved in the binding interaction, thereby suggesting that these two groups of proteins are strongly related. In contrast, although the same β-barrel topology observed for L25 is used by GlnRS for binding of the tRNA anticodon loop (Rould *et al.*, 1991), this binding differs strongly from that observed in the 5SE–L25 complex. The tRNA anticodon loop is bound at the interface between two homologous β-barrels of GlnRS and involves short segments of β-strands and loops of the GlnRS molecule. Furthermore, the RNA-binding surface is on the opposite side of the β-barrels of GlnRS with respect to the L25 β-barrel. The anticodon bases are splayed out allowing for hydrophobic interactions and for a direct read out of the base identities by hydrogen bonding (Rould *et al.*, 1991). Thus, the highly homologous β-barrel topology between L25 and GlnRS is contrasted by a very different mode of RNA recognition and by the complete lack of sequence homology (Stoldt *et al.*, 1998). Therefore, it seems unlikely that GlnRS constitutes a true member of a family of RNA-binding proteins, which includes ribosomal L25 and the CTC-type general stress proteins, although the evolutionary relationship among these proteins (Stoldt *et al.*, 1998) remains an open question.

In conclusion, even though some characteristics of the protein–RNA interaction bear resemblance to other systems, the present complex between L25 and the 5SE RNA represents a unique structure. In particular, the preformed interaction of a β-sheet including parallel β-strands with the minor groove of the RNA, which is certainly facilitated by the stretch of seven non-Watson–Crick base pairs in the E-loop, represents a feature not found in other protein–RNA or peptide–RNA complexes of known structure.

Materials and methods

Preparation of E.coli L25 protein

The *E.coli* L25 gene was amplified by PCR using appropriate primers and *E.coli* genomic DNA as template and cloned into the plasmid vector pET11a for protein overexpression. *Escherichia coli* strain BL21(DE3) was transformed and grown at 37°C in Luria–Bertani medium or M9 minimal medium containing 2 g/l [^{13}C]glucose and 1 g/l [^{15}N]NH_4Cl. Protein expression was induced at OD_{600} = 0.8 with 1 mM isopropyl-

β-D-thiogalactopyranoside (IPTG) and cells were harvested after a further 4 h of growth. The cells were lysed by sonication and treated with 2 vol. of acetic acid. The proteins of the supernatant were precipitated with 5 vol. of acetone, dissolved in 50 mM NaH_2PO_4/Na_2HPO_4 buffer at pH 6.5, containing 6 M urea and 4 mM 2-mercaptoethanol and separated on an SP-Sepharose High Performance column (Amersham Pharmacia Biotech) developed with a linear NaCl gradient. L25 was refolded and further purified by HPLC as described (Stoldt *et al.*, 1998). The identity of the overexpressed protein with wild-type L25 was established by sequencing of the complete gene, by sequencing of the protein by Edman degradation and by comparison of the ^1H-^{15}N HSQC NMR spectra.

Preparation of RNA and the protein–RNA complex

^{15}N- and $^{15}N/^{13}C$-labeled nucleotide triphosphates were prepared from *E.coli* grown on M9 minimal medium supplied with [^{15}N]ammonium chloride or [^{15}N]ammonium chloride/[^{13}C]glucose as the sole nitrogen/carbon source as described (Batey *et al.*, 1992; Grüne *et al.*, 1996). [*U*-^{15}N]- and [*U*-^{13}C,^{15}N]-5SE RNA was prepared by *in vitro* transcription with T7 RNA polymerase (Milligan *et al.*, 1987) from linearized plasmid DNA templates and purified and folded as described previously (Stoldt *et al.*, 1998). Samples of the protein–RNA complex were prepared as described previously (Stoldt *et al.*, 1998) in NMR sample buffer consisting of 20 mM NaH_2PO_4/Na_2HPO_4 buffer pH 7.2, 100 mM KCl and 3 mM NaN_3 and 5% D_2O. For NMR experiments in D_2O the protein and RNA were lyophilized separately and redissolved in 99.996% D_2O. The following samples were prepared: [*NA*]-5SE:[*U*-^{13}C,^{15}N]-L25 in H_2O (2.0 mM) and D_2O (2.2 mM), [*U*-^{15}N]-5SE:[*NA*]-L25 in H_2O (1.8 mM), [*U*-^{15}N]-5SE:[*U*-^{15}N]-L25 in H_2O (1.0 mM) and [*U*-^{13}C,^{15}N]-5SE:[*NA*]-L25 in H_2O (1.8 mM) and D_2O (1.7 mM). [*U*-^{15}N]-G,U 5S rRNA was prepared as described (Grüne *et al.*, 1996) and exchanged into NMR sample buffer. The following sample was prepared: [*U*-^{15}N]-G,U 5S rRNA:[*U*-^{15}N]-L25 (0.9 mM).

NMR spectroscopy and spectral assignments

NMR experiments were carried out on a Varian ^{UNITY}*INOVA* 750 MHz spectrometer. Triple resonance NMR spectra used for assignments were processed with the VNMR software (Varian Associates, Inc.). PROSA (Güntert *et al.*, 1992) and/or NMRPipe (Delaglio *et al.*, 1995) were used for processing of the 2D NOESY, 2D CPMG-NOESY and 3D NOESY spectra. XEASY (Bartels *et al.*, 1995) was employed for visualization and analysis of the spectra. In general, experiments were carried out at 25°C. Additional experiments at 15°C were recorded for the observation of NOEs involving exchangeable RNA hydrogens.

^1H, ^{13}C and ^{15}N resonance assignments for the bound protein were obtained at 25°C using the 3D HNCA, HNCO, C(CO)NH, H(CCO)NH, HCCH-TOCSY and HCCH-COSY experiments. Methionine and arginine ε hydrogen resonances and the Lys14-$N^{\zeta}H_3^+$ resonance were assigned through 3D ^{13}C- or ^{15}N-edited NOESY experiments. Glutaminyl and asparaginyl side chain amide resonances were assigned in the C(CO)NH spectrum. Aromatic resonances were assigned through a selective 3D ^1H-^{13}C NOESY-HSQC spectrum (80 ms mixing time). Amide hydrogens that exchanged slowly with solvent were identified through ^1H-^{15}N HSQC spectra recorded 12, 24 and 48 h after dissolving a lyophilized H_2O sample of the complex in D_2O. NOE constraints for structural calculations were obtained from the following experiments: 3D ^1H-^{15}N NOESY-HSQC (80 ms mixing time) using the [*U*-^{15}N]-5SE:[*U*-^{15}N]-L25 sample and 3D ^1H-^{13}C NOESY-HSQC (80 ms mixing time) selective for aliphatic carbons using the [*NA*]-5SE:[*U*-^{13}C,^{15}N]-L25 sample in H_2O. Initially, NOEs were unambiguously assigned from the combination of the ^{13}C- and ^{15}N-edited NOE spectra both recorded on a H_2O sample. Additional NOEs were then automatically assigned using the NOAH algorithm (Mumenthaler *et al.*, 1997) as implemented in the DYANA program after the r.m.s. displacement for backbone heavy atoms dropped to 1 Å for a calculated structure ensemble.

Sequential assignments for the H6/H8/C6/C8 and C1′/H1′ resonances of the bound RNA were obtained by establishing a sequential walk using H6/H8$_{(n)}$-H1′$_{(n)}$-H6/H8$_{(n+1)}$ and H6/H8$_{(n)}$-H6/H8$_{(n+1)}$ NOE connectivities (Wüthrich, 1986) observed in 3D ^1H-^{13}C NOESY-HSQC spectra (150 ms mixing time) separately optimized for aromatic and sugar ^{13}C chemical shifts in H_2O and D_2O. H5/C5 resonances were assigned by their strong NOEs to H6 resonances of the same base. Sequential assignments for the exchangeable resonances of the pyrimidines were obtained using the H5(C5C4N)H experiment (Wöhnert *et al.*, 1999b). The remaining guanine imino resonances were assigned using sequential and intra-base pair imino/imino and imino/aromatic NOE connectivities in 3D ^1H-^{15}N NOESY-HSQC and 2D ^1H-^1H NOESY experiments. Adenine H2 resonances were distinguished from other aromatic resonances by the unique chemical shift of their attached C2 carbon and sequentially assigned using NOEs to imino/amino, aromatic and H1′ resonances. H2′ and H3′ ribose resonances were assigned using 3D HCCH-COSY and HCCH-RELAY-COSY experiments assisted by 3D ^1H-^{13}C NOESY-HSQC experiments and are 97% complete. Due to the severe spectral overlap encountered even in 3D experiments only a few H4′ and H5′/H5″ resonances could be assigned. NH...N type hydrogen bonds were identified using the HNN-COSY experiment (Dingley and Grzesiek, 1998). NOE constraints for structural calculations were taken from 2D ^1H-^1H-NOESY and 3D ^1H-^{15}N-NOESY (80 ms mixing time) experiments recorded in H_2O using the [*U*-^{15}N]-5SE:[*NA*]-L25 sample and 3D ^1H-^{13}C-NOESY (150 ms mixing time) experiments selective for aromatic or aliphatic carbons recorded in H_2O and D_2O using the [*U*-^{13}C,^{15}N]-5SE:[*NA*]-L25 sample at 25°C. Additional NOEs involving imino/amino resonances were derived from 2D ^1H-^1H-NOESY, 2D ^1H-^{15}N-CPMG-NOESY (Mueller *et al.*, 1995) and 3D ^1H-^{15}N-NOESY (80 ms mixing time) spectra recorded at 15°C. Intermolecular NOEs were identified in a 3D $^{13}C/^{15}N$-F1-filtered, $^{13}C/^{15}N$-F3-edited NOESY-HSQC (80 ms mixing time) spectrum in H_2O (Zwahlen *et al.*, 1997), and in a 3D ^{13}C-F1-filtered,^{13}C-F3-edited NOESY-HSQC (80 ms mixing time) spectrum in D_2O using the [*NA*]-5SE:[*U*-^{13}C,^{15}N]-L25 sample and a 3D ^{13}C-F1-filtered,^{13}C-F3-edited NOESY-HSQC (80 ms mixing time) spectrum in D_2O (Zwahlen *et al.*, 1997) using the [*U*-^{13}C,^{15}N]-5SE:[*NA*]-L25 sample. Heteronuclear $^{15}N\{^1H\}$-NOE measurements were made according to the pulse sequence of Kay *et al.* (1989) using the uniformly ^{15}N-labeled protein either free or bound to 5SE RNA.

Structure calculations

The structure calculations were performed with the program DYANA (Güntert *et al.*, 1997) starting from random conformations. The 20% of structures with the lowest DYANA target functions were subjected to energy minimization using the AMBER94 force field (Cornell *et al.*,1995), with scaling of the electrostatic interaction by 0.5 to avoid excessive charge on the RNA backbone for *in vacuo* calculations, and used to characterize the solution structure of the bound L25 protein, the bound 5SE–RNA and the protein–RNA complex, respectively. Illustrations of structures were generated with MOLMOL (Koradi *et al.*, 1996).

Protein. NOE crosspeak intensities were classified as strong, medium and weak, corresponding to upper distance limit constraints of 2.6, 3.6 and 5.2 Å, respectively. The cross peak intensities of NOEs between protons with known distances were used for calibration. For NOEs involving methyl groups upper limit distance constraints of 2.8, 3.9 and 5.6 Å for strong, medium or weak interactions were used. In addition, the pseudo-atom correction of DYANA was applied. At the last stage, hydrogen bonds consistently formed in previous calculations and supported by the detection of backbone amides with slow exchange with solvent were introduced as additional upper distance limits of 2.6 Å.

RNA. Upper limit distance constraints for the non-exchangeable hydrogens were classified according to their intensity in the NOESY spectra to strong, medium and weak, corresponding to distance limits of 3.0, 3.8 and 5.4 Å, respectively. For the exchangeable hydrogens only two distance classes corresponding to 4.0 and 5.6 Å, respectively, were introduced. NOE intensities corresponding to fixed H5–H6 distances and intra-base pair HN to H2 distances were used for calibration, respectively. Initially, hydrogen bond constraints (four upper and four lower limit constraints/base pair) were introduced for Watson–Crick base pairs, when the presence of a uridine N3H3/adenine N1 or guanine N1H1/cytosine N3 hydrogen bond was detected in the HNN-COSY experiment (Dingley and Grzesiek, 1998), and for the uridine N3H3/adenine N7 hydrogen bonds of the reversed Hoogsteen base pairs A73/U103 and U77/A99 detected in the same experiment (Wöhnert *et al.*, 1999a). Additional hydrogen bond constraints were introduced for hydrogen bonds consistently present in the calculated structures (one upper limit constraint/hydrogen bond). Backbone torsion angles α and ζ of residues 68–83 and 93–108 were restricted to be in a range of ± 60° of A-form helical values since no downfield-shifted ^{31}P resonances were detected in the ^{31}P NMR spectrum of the bound RNA (Gorenstein, 1984; Varani *et al.*, 1991). Angle constraints for the glycosidic torsion angle χ of residues 69–83 and 93–107 were obtained from a local conformational analysis using the FOUND module (Güntert *et al.*, 1998) in DYANA (Güntert *et al.*, 1997).

Protein–RNA complex. The structure of the complex was calculated from starting structures with completely random torsion angles using

the constraint sets described for the protein and the RNA combined with 121 intermolecular constraints classified as for the protein. In order to ensure a flexible adjustment of protein versus RNA, the sequences of the molecules were linked by 160 linker elements consisting exclusively of pseudoatoms as described for the DYANA program (Güntert *et al*., 1997).

Database depositions. The atomic coordinates and the experimental constraints of the complex have been submitted to the Protein Data Base (entry code:1d6k).

Acknowledgements

The authors thank A.Figuth and R.Lamprecht for technical assistance in the preparation of isotope labeled materials, E.Nyakatura for Edman sequencing of the recombinant L25, and J.Leppert for help with NMRPipe. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (Br 1487/2-3). J.W. is grateful to the Fond der Chemischen Industrie for financial support.

References

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.

Baer,D.G., Schleich,T., Noller,H.F. and Garrett,R.A. (1977) Alternations of 5S RNA conformation by ribosomal proteins L18 and L25. *Nucleic Acids Res.*, **4**, 2511–2526.

Ban,N., Freeborn,B., Nissen,P., Penczek,P., Grassucci,R.A., Sweet,R., Frank,J., Moore,P.B. and Steitz,T.A. (1998) A 9 Å resolution X-ray crystallographic map of the large ribosomal subunit. *Cell*, **93**, 1105–1115.

Bartels,C., Xia,T., Billeter,M., Güntert,P. and Wüthrich,K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR*, **6**, 1–10.

Batey,R.T., Inada,M., Kujawinski,E., Puglisi,J.D. and Williamson,J.R. (1992) Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA. *Nucleic Acids Res.*, **20**, 4515–4523.

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

Bycroft,M., Hubbard,T.J.P., Proctor,M., Freund,S.M.V. and Murzin,A.G. (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell*, **88**, 235–242.

Cavarelli,J., Rees,B., Ruff,M., Thierry,J.C. and Moras,D. (1993) Yeast tRNA (Asp) recognition by its cognate class II aminoacyl-tRNA synthetase. *Nature*, **362**, 181–184.

Chen-Schmeisser,U. and Garrett,R.A. (1977) A new method for the isolation of a 5S RNA complex with proteins L5, L18 and L25 from *Escherichia coli* ribosomes. *FEBS Lett.*, **74**, 287–291.

Ciesiolka,J., Lorenz,S. and Erdmann,V.A. (1992) Structural analysis of three prokaryotic 5S rRNA species and selected 5S rRNA-ribosomal–protein complexes by means of Pb(II)-induced hydrolysis. *Eur. J. Biochem.*, **204**, 575–581.

Conn,G.L., Draper,D.E., Lattman,E.E. and Gittis,A.G. (1999) Crystal structure of a conserved ribosomal protein–RNA complex. *Science*, **284**, 1171–1174.

Cornell,W.D. *et al.* (1995) A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.

Correll,C.C., Freeborn,B., Moore,P.B. and Steitz,T.A. (1997) Metals, motifs and recognition in the crystal structure of a 5S rRNA domain. *Cell*, **91**, 705–712.

Cusack,S. (1997) Aminoacyl-tRNA synthetases. *Curr. Opin. Struct. Biol.*, **7**, 881–889.

Dallas,A. and Moore,P.B. (1997) The loop E–loop D region of *Escherichia coli* 5S rRNA: the solution structure reveals an unusual loop that may be important for binding ribosomal proteins. *Structure*, **5**, 1639–1653.

De Guzman,R.N., Turner,R.B. and Summers,M.F. (1998a) Protein–RNA recognition. *Biopolymers*, **48**, 181–195.

De Guzman,R.N., Wu,Z.R., Stalling,C.C., Pappalardo,L., Borer,P.N. and Summers,M.F. (1998b) Structure of the HIV-1 nucleocapsid protein bound to the SL3 ψ-RNA recognition element. *Science*, **279**, 384–388.

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.

Dingley,A.J. and Grzesiek,S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^{2}J_{NN}$ couplings. *J. Am. Chem. Soc.*, **120**, 8293–8297.

Douthwaite,S., Garrett,R.A., Wagner,R. and Feunteun,J. (1979) A ribonuclease-resistant region of 5S rRNA and its relation to the RNA binding sites of proteins L18 and L25. *Nucleic Acids Res.*, **6**, 2453–2470.

Douthwaite,S., Christensen,A. and Garrett,R.A. (1982) Binding sites of ribosomal proteins on prokaryotic 5S ribonucleic acids: a study with ribonucleases. *Biochemistry*, **21**, 2313–2320.

Erdmann,V.A., Fahnestock,K.H. and Nomura,M. (1971) Role of 5S RNA in the functions of 50S ribosomal subunits. *Proc. Natl Acad. Sci. USA*, **68**, 2932–2936.

Frankel,A.D. and Smith,C.A. (1998) Induced folding in RNA–protein recognition: more than a simple molecular handshake. *Cell*, **92**, 149–151.

Garrett,R.A. and Noller,H.F. (1979) Structures of complexes of 5S RNA with ribosomal proteins L5, L18 and L25 from *Escherichia coli*: identification of kethoxal-reactive sites on the 5S RNA. *J. Mol. Biol.*, **132**, 637–648.

Gorenstein,D.G. (1984) *Phosphorus-31 NMR: Principles and Applications*. Academic Press, New York, NY.

Görlach,M., Wittekind,M., Beckman,R.A., Mueller,L. and Dreyfuss,G. (1992) Interaction of the RNA-binding domain of the hnRNP C proteins with RNA. *EMBO J.*, **11**, 3289–3295.

Grüne,M., Görlach,M., Soskic,V., Klussmann,S., Bald,R., Fürste,J.P., Erdmann,V.A. and Brown,L.R. (1996) Initial analysis of 750 MHz NMR spectra of selectively ^{15}N-G,U labelled *Escherichia coli* 5S rRNA. *FEBS Lett.*, **385**, 114–118.

Gryaznova,O.I., Davydova,N.L., Gongadze,G.M., Jonsson,B.H., Garber,M.B. and Liljas,A. (1996) A ribosomal protein from *Thermus thermophilus* is homologous to a general shock protein. *Biochimie*, **78**, 915–919.

Güntert,P., Dötsch,V., Wider,G. and Wüthrich,K. (1992) Processing of multi-dimensional NMR data with the new software PROSA. *J. Biomol. NMR*, **2**, 619–629.

Güntert,P., Mumenthaler,C. and Wüthrich,K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, **273**, 283–298.

Güntert,P., Billeter,M., Ohlenschläger,O., Brown,L.R. and Wüthrich,K. (1998) Conformation analysis of protein and nucleic acid fragments with the new grid search algorithm FOUND. *J. Biomol. NMR*, **12**, 543–548.

Hartmann,R.K., Vogel,D.W., Walker,R.T. and Erdmann,V.A. (1988) *In vitro* incorporation of eubacterial, archaebacterial and eukaryotic 5S rRNAs into large ribosomal subunits of *Bacillus stearothermophilus*. *Nucleic Acids Res.*, **16**, 3511–3524.

Hoffman,D.W., Query,C.Q., Golden,B.L., White,S.W. and Keene,J.W. (1991) RNA-binding domain of the A protein component of the U1 small nuclear ribonucleoprotein analyzed by NMR spectroscopy is structurally similar to ribosomal proteins. *Proc. Natl Acad. Sci. USA*, **88**, 2495–2499.

Huber,P.W. and Wool,I.G. (1984) Nuclease protection analysis of ribonucleoprotein complexes: use of the cytotoxic ribonuclease α-sarcin to determine the binding sites for *Escherichia coli* ribosomal proteins L5, L18 and L25 on 5S rRNA. *Proc. Natl Acad. Sci. USA*, **81**, 322–326.

Jaishree,T.N., Ramakrishnan,V. and White,S.W. (1996) Solution structure of prokaryotic ribosomal protein S17 by high-resolution NMR spectroscopy. *Biochemistry*, **35**, 2845–2853.

Kalurachchi,K., Uma,K., Zimmermann,R.A. and Nikonowicz,E.P. (1997) Structural features of the binding site for ribosomal protein S8 in *Escherichia coli* 16S rRNA defined using NMR spectroscopy. *Proc. Natl Acad. Sci. USA*, **94**, 2139–2144.

Kay,L.E., Torchia,D.A. and Bax,A. (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, **28**, 8972–8979.

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

Laskowski,R.A., Rullmann,J.A., MacArthur,M.W., Kaptein,R. and Thornton,J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.

Leontis,N.B. and Westhof,E. (1998) A common motif organizes the structure of multi-helix loops in 16S and 23S ribosomal RNAs. *J. Mol. Biol.*, **283**, 571–583.

Luginbühl,P., Güntert,P., Billeter,M. and Wüthrich,K. (1996) The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR*, **8**, 136–146.

Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, **15**, 8783–8798.

Mosyak,L., Reshetnikova,L., Goldgur,Y., Delarue,M. and Safro,M.G. (1995) Structure of phenylalanyl-tRNA synthetase from *Thermus thermophilus*. *Nature Struct. Biol.*, **2**, 537–547.

Mueller,L., Legault,P. and Pardi,A. (1995) Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino-protons. *J. Am. Chem. Soc.*, **117**, 11043–11048.

Mumenthaler,C., Güntert,P., Braun,W. and Wüthrich,K. (1997) Automated combined assignment of NOESY spectra and three-dimensional protein structure determination. *J. Biomol. NMR*, **10**, 351–362.

Nagai,K., Oubridge,C., Jessen,T.H., Li,J. and Evans,P. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*, **348**, 515–520.

Newberry,V. and Garrett,R.A. (1980) The role of the basic N-terminal region of protein L18 in 5S RNA–23S RNA complex formation. *Nucleic Acids Res.*, **8**, 4131–4142

Newkirk,K., Feng,W., Jiang,W., Tejero,R., Emerson,S.D., Inouye,M. and Montelione,G.T. (1994) Solution structure of the major cold shock protein (CspA) from *Escherichia coli*: identification of a binding epitope for DNA. *Proc. Natl Acad. Sci. USA*, **91**, 5114–5118.

Onesti,S., Miller,A.D. and Brick,P. (1995) The crystal structure of the lysyl-tRNA synthetase (LysU) from *Escherichia coli*. *Structure*, **3**, 163–176.

Osswald,M. and Brimacombe,R. (1999) The environment of 5S rRNA in the ribosome: cross-links to 23S rRNA from sites within helices II and III of the 5S molecule. *Nucleic Acids Res.*, **27**, 2283–2290.

Oubridge,C., Ito,N., Evans,P.R., Teo,C.-H. and Nagai,K. (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432–438.

Patel,D.J. (1999) Adaptive recognition in RNA complexes with peptides and protein modules. *Curr. Opin. Struct. Biol.*, **9**, 74–87.

Pervushin,K., Riek,R., Wider,G. and Wüthrich,K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl Acad. Sci. USA*, **94**, 12366–12371.

Price,S.R., Evans,P.R. and Nagai,K. (1998) Crystal structure of the spliceosomal U2B″–U2A′ protein complex bound to a fragment of U2 small nuclear RNA. *Nature*, **394**, 645–650.

Puglisi,J.D., Tan,R., Calnan,B.J., Frankel,A.D. and Williamson,J.R. (1992) Conformation of the TAR RNA–arginine complex by NMR spectroscopy. *Science*, **257**, 76–80.

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

Rould,M.A., Perona,J.J. and Steitz,T.A. (1991) Structural basis of anticodon loop recognition by glutaminyl-tRNA synthetase. *Nature*, **352**, 213–218.

Ryter,J.M. and Schultz,S.C. (1998) Molecular basis of double-stranded RNA–protein interaction: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.*, **17**, 7505–7513.

Schindelin,H., Jiang,W., Inouye,M. and Heinemann,U. (1994) Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **91**, 5119–5123.

Sergiev,P., Dokudovskaya,S., Romanova,E., Topin,A., Bogdanov,A., Brimacombe,R. and Dontsova,O. (1998) The environment of 5S rRNA in the ribosome: cross-links to the GTPase-associated area of 23S rRNA. *Nucleic Acids Res.*, **26**, 2519–2525.

Siomi,H. and Dreyfuss,G. (1997) RNA-binding proteins as regulators of gene expression. *Curr. Opin. Genet. Dev.*, **7**, 345–353.

Spierer,P., Bogdanov,A.A. and Zimmermann,R.A. (1978) Parameters for the interaction of ribosomal proteins L5, L18 and L25 with 5S RNA from *Escherichia coli*. *Biochemistry*, **17**, 5394–5398.

Stoldt,M., Wöhnert,J., Görlach,M. and Brown,L.R. (1998) The NMR structure of *Escherichia coli* ribosomal protein L25 shows homology to general stress proteins and glutaminyl-tRNA synthetases. *EMBO J.*, **17**, 6377–6384.

Varani,G., Cheong,C. and Tinoco,I.,Jr (1991) Structure of an unusually stable RNA hairpin. *Biochemistry*, **30**, 3280–3289.

Weiss,M.A. (1998) RNA-mediated signaling in transcription. *Nature Struct. Biol.*, **5**, 329–333.

Weiss,M.A. and Narayana,N. (1998) RNA recognition by arginine-rich peptide motifs. *Biopolymers*, **48**, 167–180.

Wimberly,B., Varani,G. and Tinoco,I.,Jr (1993) The conformation of loop E of eukaryotic 5S ribosomal RNA. *Biochemistry*, **32**, 1078–1087.

Wimberly,B., Guymon,R., McCutcheon,J.P., White,S.W. and Ramakrishnan,V. (1999) A detailed view of a ribosomal active site: the structure of the L11–RNA complex. *Cell*, **97**, 491–502.

Wittekind,M., Görlach,M., Friedrichs,M., Dreyfuss,G. and Mueller,L. (1992) ^{1}H, ^{13}C and ^{15}N NMR assignments and global folding pattern of the RNA-binding domain of the human hnRNP C proteins. *Biochemistry*, **31**, 6254–6265.

Wöhnert,J., Dingley,A.J., Stoldt,M., Görlach,M., Grzesiek,S. and Brown,L.R. (1999a) Direct identification of NH...N hydrogen bonds in non-canonical base pairs of RNA by NMR spectroscopy. *Nucleic Acids Res*, **27**, 3104–3110.

Wöhnert,J., Ramachandran,R., Görlach,M. and Brown,L.R. (1999b) Triple resonance experiments for correlation of H5 and exchangeable pyrimidine base hydrogens in ^{13}C, ^{15}N labelled RNA. *J. Magn. Reson.*, **139**, 430–433.

Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley, New York, NY.

Yonath,A. and Franceschi,F. (1998) Functional universality and evolutionary diversity: insights from the structure of the ribosome. *Structure*, **6**, 679–684.

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 l N-peptide/*boxB* RNA complex. *J. Am. Chem. Soc.*, **119**, 6711–6721.

Received August 17, 1999; revised and accepted September 28, 1999