

Ribosomal Proteins S5 and L6: High-resolution Crystal Structures and Roles in Protein Synthesis and Antibiotic Resistance

Christopher Davies^{1†}, Dirksen E. Bussiere^{2†}, Barbara L. Golden³
Stephanie J. Porter¹, Venki Ramakrishnan^{4*} and Stephen W. White^{1,5*}

¹Department of Structural Biology, St. Jude Children's Research Hospital, 332 North Lauderdale St., Memphis TN 38105, USA

²Abbott Laboratories Department of Scientific Information, Analysis and Management, 100 Abbott Park Rd., Abbott Park, IL 60064 USA

³Department of Chemistry & Biochemistry, Box 215 University of Colorado Boulder, CO 80309, USA

⁴Department of Biochemistry University of Utah School of Medicine, Salt Lake City UT 84132, USA

⁵Department of Biochemistry University of Tennessee Memphis, 858 Madison Suite G01, Memphis TN 38163, USA

Antibiotic resistance is rapidly becoming a major medical problem. Many antibiotics are directed against bacterial ribosomes, and mutations within both the RNA and protein components can render them ineffective. It is well known that the majority of these antibiotics act by binding to the ribosomal RNA, and it is of interest to understand how mutations in the ribosomal proteins can produce resistance. Translational accuracy is one important target of antibiotics, and a number of ribosomal protein mutations in *Escherichia coli* are known to modulate the proofreading mechanism of the ribosome. Here we describe the high-resolution structures of two such ribosomal proteins and characterize these mutations.

The S5 protein, from the small ribosomal unit, is associated with two types of mutations: those that reduce translational fidelity and others that produce resistance to the antibiotic spectinomycin. The L6 protein, from the large subunit, has mutations that cause resistance to several aminoglycoside antibiotics, notably gentamicin. In both proteins, the mutations occur within their putative RNA-binding sites. The L6 mutations are particularly drastic because they result in large deletions of an RNA-binding region. These results support the hypothesis that the mutations create local distortions of the catalytic RNA component.

When combined with a variety of structural and biochemical data, these mutations also become important probes of the architecture and function of the translational machinery. We propose that the C-terminal half of S5, which contains the accuracy mutations, organizes RNA structures associated with the decoding region, and the N-terminal half, which contains the spectinomycin-resistance mutations, directly interacts with an RNA helix that binds this antibiotic. As regards L6, we suggest that the mutations indirectly affect proofreading by locally distorting the EF-Tu·GTP·aminoacyl tRNA binding site on the large subunit.

© 1998 Academic Press Limited

Keywords: Ribosome architecture; X-ray crystallography; protein-RNA interactions; protein evolution; translocation

*Corresponding authors

Introduction

Efficient protein synthesis is essential to cell growth, and many antibiotics have evolved to incapacitate the ribosome, the cell's ubiquitous translational machinery. In response to this selective pressure, mutations have evolved within bacterial

ribosomal components that confer resistance to these compounds. With the increased use of antibiotics to treat bacterial infections, pathogenic strains are now acquiring antibiotic resistance, and many important drugs that target bacterial ribosomes are becoming ineffective. This represents an extremely serious medical problem that has prompted extensive searches for new antibacterial agents.

An important ribosome function targeted by a number of antibiotics is proofreading, which ensures accurate reading and decoding of the

†These authors contributed equally to this work.

Abbreviations used: dsRBD, double-stranded RNA-binding domain; EF, elongation factor; RRM, RNA recognition motif.

mRNA. Early insights into ribosomal fidelity were provided by the antibiotic streptomycin, which had been shown to reduce the accuracy of protein synthesis. Further studies revealed that fidelity can also be altered by mutations in three 30 S subunit proteins, S4, S5 and S12 (Gorini & Kataja, 1964). Mutations in S4 and S5 affect the ribosome in a way similar to streptomycin and reduce the level of accuracy (the so-called *ram* or ribosome ambiguity mutations), whereas mutations in S12 have the opposite effect and increase translational accuracy (reviewed in Kurland *et al.*, 1990). These factors appear to be additive, since S12 mutations confer resistance to streptomycin, causing the ribosome to become dependent on the antibiotic for optimal accuracy. This dependency can be relieved by subsequent mutations in either S4 or S5. More recently, defined regions of 16 S rRNA have also come to be associated with the ribosome's accuracy function (Montandon *et al.*, 1986; Melançon *et al.*, 1988; Powers & Noller, 1991).

Although largely a feature of the 30 S subunit, translational accuracy can also be affected by two components of the 50 S subunit, ribosomal protein L6 and the 2660 loop region of 23 S rRNA. Mutations have been identified in these components that result in a decreased rate of translation, greater accuracy in protein synthesis, and increased resistance to many of the misreading-inducing aminoglycoside antibiotics, in particular gentamicin (Buckel *et al.*, 1977; Kuhberger *et al.*, 1979; Hummel *et al.*, 1980; Tapprich & Dahlberg, 1990; Melançon *et al.*, 1992).

For several years, we have been studying the structures of individual proteins from the prokaryotic ribosome as part of a concerted effort in many laboratories to understand the structure and mechanism of the translational machinery. The role of ribosomal proteins now appears to be largely architectural, helping to direct and maintain the fold of the catalytic RNA component, and it is of interest to understand how mutations in the proteins exert their effects on ribosome function. Our studies have shown that antibiotic resistance mutations, and other mutations that affect ribosome function and assembly, tend to cluster in their putative RNA-binding sites (Ramakrishnan & White, 1992; Golden *et al.*, 1993a; Hoffman *et al.*, 1996). This suggests that they act indirectly by locally altering the ribosomal RNA conformation. If correct, these proteins represent powerful probes of the structure and function of the ribosome, since their mutant forms and associated phenotypes can be directly correlated with their RNA environments. These results also provide important insights into the mechanism of antibiotic action and the acquisition of antibiotic resistance. We have been investigating the crystal structures of two proteins that exemplify this approach to studying the mechanism of the ribosome, namely

S5 from the 30 S subunit, and L6 from the 50 S subunit. As noted above, mutations within both proteins can influence ribosome accuracy, and they are clearly adjacent to important functional centers of their respective subunits.

S5 is particularly interesting, since three distinct types of mutations have been identified, the *ram* mutations described above (Piepersberg *et al.*, 1975a,b), cold sensitivity (Guthrie & Nomura, 1969) and ribosomal resistance to the antibiotic spectinomycin (Bollen *et al.*, 1969; Funatsu *et al.*, 1972; Piepersberg *et al.*, 1975b). Spectinomycin is usually associated with the translocation mechanism of the ribosome (Burns & Cundliffe, 1973; Wallace *et al.*, 1974), and the S5 locus therefore appears to be close to not one, but two important functions in the 30 S subunit. This central location is supported by RNA footprinting (Stern *et al.*, 1989; Powers & Noller, 1995) and crosslinking experiments (Osswald *et al.*, 1987) that show S5 to be in the vicinity of the three-way junction that connects the major subdomains of the 16 S rRNA molecule.

L6 has been localized by immunoelectron microscopy (Stöffler-Meilicke *et al.*, 1983; Hackl & Stöffler-Meilicke 1988) and protein-protein crosslinking (Traut *et al.*, 1986; Walleczek *et al.*, 1989a,b) to a region between the base of the L7/L12 stalk and the central protuberance (Walleczek *et al.*, 1988). L6 appears to be at the 50 S–30 S interface because it can be cross-linked to S13 in the 30 S subunit (Lambert & Traut, 1981). As regards its 23 S rRNA environment, L6 has been cross-linked to the end of a stem-loop structure (nucleotides 2455 to 2496 in *Escherichia coli*) that projects from domain V (Wower *et al.*, 1981). Domain V is a highly conserved region of the 23 S molecule that has been implicated in peptide bond formation (Vester & Garrett, 1988) and the binding of tRNA at the A and P sites (Steiner *et al.*, 1988; Mitchell *et al.*, 1990; Brimacombe *et al.*, 1993). L6 also binds to a fragment of 23 S rRNA corresponding to domain VI (Leffers *et al.*, 1988). Domain VI contains the α -sarcin loop and is the binding site for the EF-Tu·GTP·aminoacyl tRNA ternary complex (Leffers *et al.*, 1988; Moazed *et al.*, 1988).

Several years ago we reported the medium-resolution crystal structures of S5 and L6 from *Bacillus stearothermophilus*, and showed that each is a two-domain molecule with clear putative RNA-binding regions (Ramakrishnan & White, 1992; Golden *et al.*, 1993b). Here, we present the refined structures of both proteins, and use them to re-evaluate their roles in the organization of the large and small ribosomal subunits. In the case of L6, we have also characterized the mutations that produce gentamicin resistance. Consistent with contemporary ideas on the role of ribosomal proteins, S5 and L6 appear to organize important local regions of the 16 S and 23 S rRNA molecules, respectively.

Results

Ribosomal protein S5

Refinement

The starting point of the refinement was the 2.7 Å structure previously published, which included residues 4 to 148 of the 166 amino acid residue protein (Ramakrishnan & White, 1992). This structure had an *R*-factor of 22% with deviations from ideality in bond lengths and angles of 0.017 Å and 3.74°, respectively. To obtain an initial 2.2 Å structure, a series of Powell minimizations was performed in which the high-resolution data were added in 0.2 Å increments every 100 cycles. This structure was compared to $F_o - F_c$ and $2F_o - F_c$ electron density maps, manual adjustments were made, and water molecules were added. The criteria for the latter were clear peaks in both maps, and adjacency to suitable hydrogen bond donors and acceptors. Following the rebuilding, the model was subjected to 200 cycles of Powell minimization, simulated annealing at 2000 K and individual *B*-factor refinement. Weak electron density now became visible for the disordered N and C termini, and for residues 28 to 30 within a flexible loop. Following three more macrocycles of model adjustment, Powell minimization, simulated annealing and individual *B*-factor refinement, convergence was reached. Summaries of the refinement statistics and the model geometry are shown in Table 1. The density for the N and C termini showed little further improvement, and the termini appear to be essentially disordered and loosely associated with nearby ordered regions within the unit cell. Within the body of the molecule, only residues 28 to 30 are not clearly defined in the electron density. The final model comprised residues 5 to 27 and 31 to 148. The S5 structure has been briefly described (Ramakrishnan & White, 1992), and a detailed description is given below.

S5 comprises two domains

S5 is an α/β class of protein that is organized into two α/β domains, each containing a single β -sheet with α -helical elements on one side (Figure 1). The two domains are directly connected at valine 74. The side-chain of this residue forms part of the C-terminal domain hydrophobic core, and its amide proton is hydrogen-bonded to the carbonyl oxygen atom of glycine 44 in the N-terminal domain. The two halves of the protein pack closely together such that the N-terminal domain β -sheet associates with the C-terminal domain α -helices. The packing of the domains creates alternate layers of α -helices and β -sheets, which results in an unusual α - β - α - β sandwich architecture. The two domains appear to be structurally very similar, but closer inspection reveals that their topologies are in fact quite different (see later).

The domains interact extensively through a highly conserved network of hydrophobic resi-

Table 1. Data collection and refinement parameters for ribosomal protein S5

<i>A. Data Collection</i>	
Resolution (Å)	50.0–2.2
Total number of observations	66,420
Number of unique reflections	11,138
<i>R</i> -merge (%) ^a	8.15
Completeness of data (%)	93
Completeness of data (%) (2.2 Å–2.3 Å)	66
<i>B. Refinement</i>	
Resolution (Å)	10–2.2
Sigma cut-off	0.0
Residues in model	5–27, 30–148
Number of protein atoms	1,037
Number of water molecules	187
Number of reflections	10,470
<i>R</i> -value (%)	21.66
<i>R</i> -value (%) (2.2 Å–2.3 Å)	26.76
<i>R</i> _{free} ^b value (%)	26.63
Overall <i>G</i> factor ^c	–0.02
r.m.s. deviation on bonds (Å)	0.009
r.m.s. deviation on angles (°)	1.485
r.m.s. deviation on improper angles (°)	2.440
r.m.s. deviation on dihedral angles (°)	27.5
Mean <i>B</i> factor (main-chain) (Å ²)	36.2
r.m.s. deviation in <i>B</i> factor (main-chain) (Å ²)	2.683
Mean <i>B</i> factor (side-chain) (Å ²)	44.3
r.m.s. deviation in <i>B</i> factor (side-chain) (Å ²)	5.556
Residues in most favored regions (%)	87.7
Residues in additional allowed regions (%)	10.5
^a $R\text{-merge} = \sum I_i - I_m / \sum I_m$.	
^b Brünger (1992).	
^c Laskowski <i>et al.</i> (1993).	

dues. Valine 16, 24, 38 and 46, alanine 17, leucine 36, phenylalanine 48 and isoleucine 72 from the N-terminal domain are on the outer surface of the β -sheet, and proline 110 and 133, alanine 113 and 118, valine 114 and 137, leucine 117, isoleucine 134 and phenylalanine 141 from the C-terminal domain are on the outer surface of the α -helices (Figure 2). The two phenylalanine rings, one from each domain, are adjacent and interact in the usual edge-to-face orientation (Burley & Petsko, 1985). A short loop between residues 41 and 44 in the N-terminal domain is adjacent to another loop in the C-terminal domain between residues 118 and 123, and residues 118 and 119 are constrained to be alanine and glycine, respectively, by the close approach. There is a specific hydrogen-bonded interaction between the two loops involving the amide protons of aspartate 41 and lysine 42 and the carbonyl oxygen of atom leucine 117. Finally, there is a conserved interdomain salt-bridge between arginine 14 and glutamate 116.

The N-terminal domain

The N-terminal half of S5 is somewhat elongated and contains five main secondary structural elements, four β -strands and an α -helix, linearly arranged as $\beta\beta\beta\alpha\beta$. The locations of these elements in the sequence are as follows: strand 1, residues 11 to 19; strand 2, residues 33 to 41; strand 3, residues 44 to 53; helix 1, residues 55 to 70; strand 4,

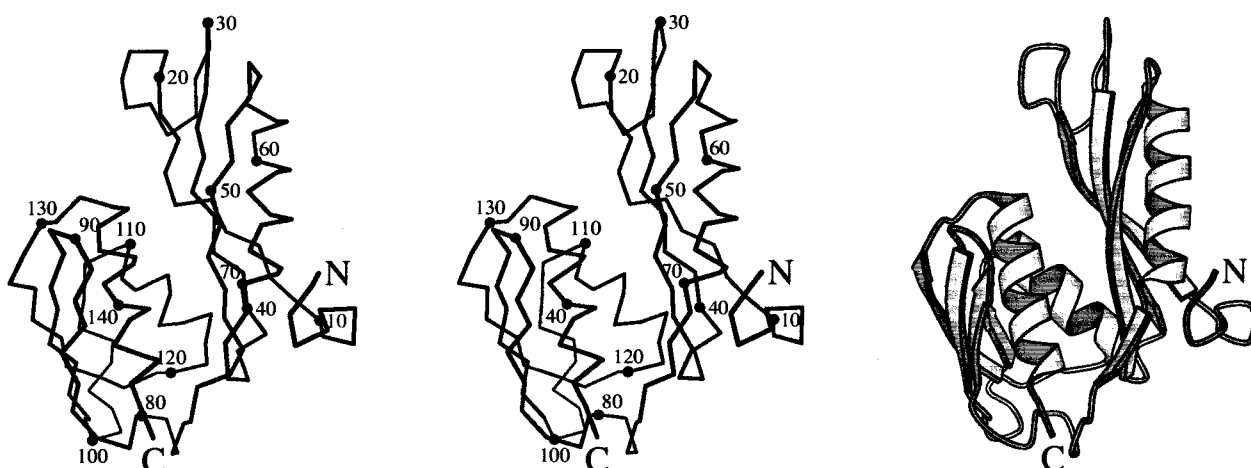


Figure 1. The overall structure of ribosomal protein S5. Left, stereo views of the C α trace of the S5 backbone with every tenth residue labeled and marked with a filled circle. Right, a ribbon representation of S5 showing the elements of secondary structure. The unstructured C-terminal 19 amino acid residues have not been included for clarity. The Figure was produced using MOLSCRIPT (Kraulis, 1991).

residues 71 to 74. These elements are arranged as a four-stranded anti-parallel β -sheet with the single α -helix inserted between strands 3 and 4 (Figure 1). As a result, strand 4 is quite short. The α -helix sits

on one surface of the β -sheet on the outer surface of the S5 molecule. The α -helix appears to be stabilized by a number of salt bridges on its outer surface (arginine 61 and lysine 62 with glutamate 58;

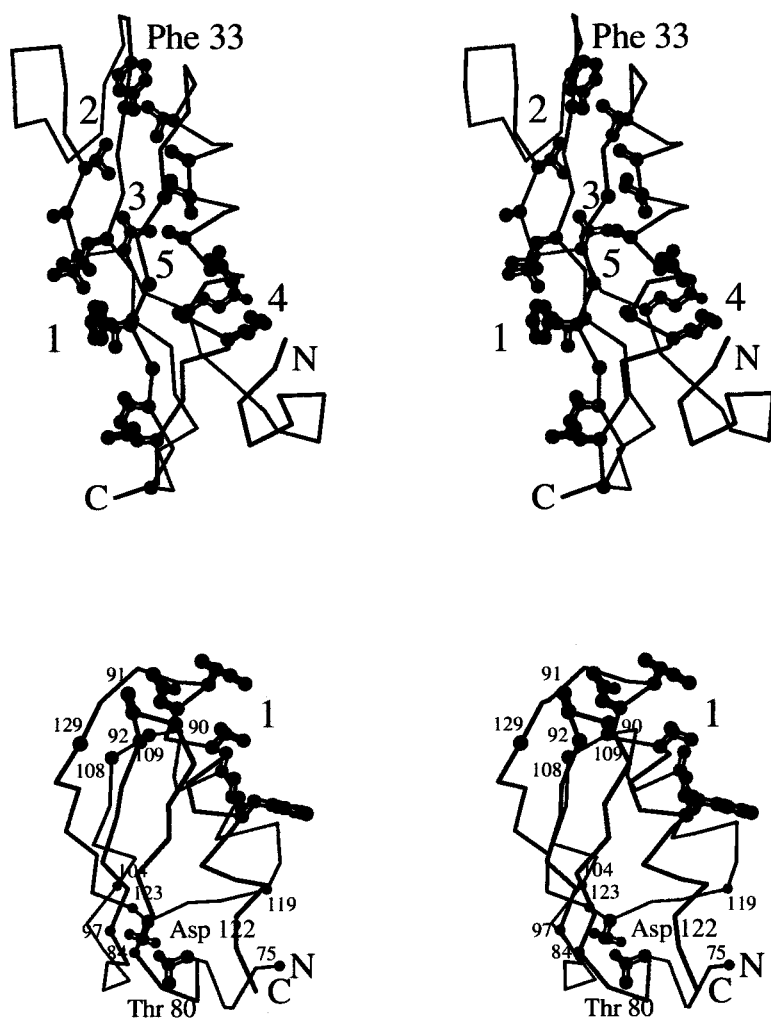


Figure 2. The structurally important features within the two domains of ribosomal protein S5. Top, stereo views of the N-terminal domain. Regions that are discussed in the text are numbered as follows. Region 1 is the interdomain hydrophobic patch. Region 2 is the small hydrophobic core within the putative RNA-binding loop centered on phenylalanine 33. Region 3 is the exposed hydrophobic core. Region 4 is the conserved salt-bridge between glutamate 13 and lysine 68. Region 5 is the close approach between strand 3 and helix 1 that is mediated by conserved glycine and alanine residues. For clarity, individual residues are not labeled. Bottom, stereo views of the C-terminal domain. The residues in the interdomain hydrophobic patch are in the region labeled 1. The conserved glycine and alanine residues in the tight turns at the top of the domain are labeled. The bottom half of the domain is covered by three loops, and the ends of the loops are labeled. Aspartate 122 and threonine 80 form important inter-loop interactions. See the text for details. The unstructured C-terminal 19 amino acid residues have not been included for clarity. The Figure was produced using MOLSCRIPT (Kraulis, 1991).

lysine 68 with glutamate 65; lysine 69 with aspartate 66).

An unusual feature of this domain is that the α -helix is shifted towards one edge of the β -sheet, and makes a very close approach to strand 3. They are so close in fact that the contacting surfaces of strand 3 and helix 1 are constrained to contain exclusively glycine (44, 47, 49 and 51) and alanine residues (59, 63 and 67), respectively (Figure 2). Conversely, the gap between helix 1 and strand 1 on the opposite edge of the β -sheet exposes the hydrophobic core residues in this region (Figure 2). This gap is minimized and held roughly constant by a kink in strand 1 that results from a β -bulge at residues 16 and 17. Specifically, the amide protons of these residues form a bifurcated hydrogen bond to the carbonyl oxygen atom of residue 36 on the adjacent strand 2. At the lower end of the domain, the gap is spanned by a highly conserved salt-bridge between glutamate 13 and lysine 68 (Figure 2).

Loop 2 (residues 19 to 33) that connects strands 1 and 2 is convoluted, flexible, and highly positively charged. Apart from the N and C termini, this region is the most poorly defined in the electron density map, and density for residues 28 to 30 is barely visible. The loop is folded around a mini hydrophobic core that contains a central and highly conserved phenylalanine (residue 33; Figure 2). We have previously identified this loop as a probable RNA-binding region, and subsequent results suggest that it may be specific for double-stranded RNA. This will be discussed later.

The C-terminal domain

The C-terminal half of S5 contains six secondary structural elements, four β -strands and two α -helices, linearly arranged as $\beta\beta\beta\alpha\beta\alpha$. The locations of these elements in the sequence is as follows: strand 5, residues 85 to 89; strand 6, residues 92 to 97; strand 7, residues 104 to 107; helix 2, residues 108 to 119; strand 8, residues 123 to 129; helix 3, residues 132 to 146. Like the N-terminal domain, these elements are also arranged as a β -sheet with α -helical elements packed onto one surface, but the topology differs significantly. In this case, the outer short β -strand is parallel with its neighbor and sequentially the third element. One surface of the β -sheet is exposed, and the other is packed against the two α -helices in this domain. Curiously, the domain also resembles the RNA recognition motif (RRM) structure that is found in a number of other ribosomal proteins (Ramakrishnan & White, 1998), but again the topology is quite different.

The connections between the secondary structural elements at the top of the domain are short, and are populated with glycine and alanine residues to accommodate the tight turns (Figure 2). However, those at the bottom are convoluted and serve to cover this end of the protein (Figure 2). Loop 6 between strands 4 and 5 (residues 74 to 85) is

flanked by proline residues and includes a short β -ribbon with a β -turn between residues 77 and 80. Loop 8, between strands 6 and 7 (residues 97 to 104), is a highly conserved region of S5 and contains a β -turn between residues 100 and 103. Proline 98, which initiates the loop, is present in all prokaryotic-like S5 sequences, and threonine 103, whose side-chain oxygen atom makes a hydrogen bond across the β -turn to the amide proton of serine 100, is also highly conserved. Finally, loop 10, between helix 2 and strand 8 (residues 119 to 123), traverses the domain and connects the α -helical side to the β -sheet side. This loop is also conserved and provides the main-chain and side-chain hydrogen bonding elements that knit this region together. One of the side-chain oxygen atoms of aspartate 122 interacts with the amide protons of residues 103 and 104, and the amide protons of residues 121 and 122 bond to the carbonyl oxygen atom of residue 79. Finally, the side-chain oxygen atom of threonine 80 bonds to the amide proton of residue 99 to complete the network.

The N-terminal domain is a double-stranded RNA-binding element

A number of RNA-binding motifs have now been described that can be identified on the basis of their characteristic amino acid sequences (Burd & Dreyfuss, 1994). Structural information is now available for three of these motifs: the RRM (Nagai *et al.*, 1990; Hoffman *et al.*, 1991), the double-stranded RNA-binding domain (dsRBD: Bycroft *et al.*, 1995; Kharrat *et al.*, 1995) and the KH domain (Castiglione *et al.*, 1995). The RRM domain has now been found in several ribosomal proteins (Ramakrishnan & White, 1998), and it was recently shown that the N-terminal half of S5 represents a classic example of a dsRBD (Bycroft *et al.*, 1995). The only major difference is that the dsRBD normally contains an N-terminal α -helix that is positioned next to the C-terminal α -helix on one surface of the β -sheet, but this α -helix is not present in S5.

An analysis of the dsRBD consensus sequence combined with site-directed mutagenesis and RNA-binding experiments have shown that the RNA-binding surface comprises the extended loop between strands 1 and 2, and the N-terminal part of the α -helix (Bycroft *et al.*, 1995). These regions contain a number of highly conserved basic residues that cannot be mutated without seriously impairing the interaction with dsRNA. A model has been proposed in which these two regions of the protein straddle the dsRNA groove and make non-specific ionic interactions with the two sugar-phosphate backbones (Bycroft *et al.*, 1995). Recently, the crystal structure of a dsRBD-dsRNA complex has been determined that confirms the major features of this model (J. Ryter & S. Schultz, personal communication). The N-terminal domain of S5 also contains a number of highly conserved lysine and

arginine residues in these two locations, and is likely to interact with a double-stranded region of 16 S rRNA in the same way. Such an interaction has recently been supported by site-directed hydroxyl radical probing of the S5 RNA environment. A probe placed at residue 22 (21 in *E. coli*) within the extended loop of S5 clearly labeled helix 34 of the 16 S rRNA molecule (Heilek & Noller, 1996). Figure 3 shows two representations of the S5 N-terminal domain that highlight the conserved functional residues and the associated surface electrostatic potential, respectively.

In both of the NMR structures of the dsRBD, the extended loop between strands 1 and 2 is poorly defined and apparently flexible. In the S5 crystal structure, this loop is wrapped around a phenylalanine side-chain in a loosely ordered conformation, and it probably adopts a defined structure when bound to RNA. We have recently determined the crystal structure of ribosomal protein S7, and it contains an extended β -ribbon structure that we propose interacts with a region of double-stranded RNA (Wimberley *et al.*, 1997). The S5 loop could also adopt this type of conformation, and a pair of adjacent conserved glycines residues (27 and 28) could mediate a tight turn at the end of such a β -ribbon. It has now been confirmed from the dsRBD-dsRNA crystal structure that the loop does become extended into a β -ribbon structure,

and that it interacts with the minor groove of the RNA helix (J. Rytter & S. Schultz, personal communication).

Functional regions of the C-terminal domain

S5 is not a primary rRNA-binding protein, but the C-terminal domain does have a surface patch that contains the conserved aromatic and basic residues that are diagnostic of specific interactions with RNA (Oubridge *et al.*, 1994; Valegård *et al.*, 1997). This region, which contains histidine 83 and 88, phenylalanine 89 and arginine 138 is shown in Figure 3. We previously suggested that the C-terminal domain might also interact with 16 S rRNA in the vicinity of glycine 104 and arginine 112 (Ramakrishnan & White, 1992). These amino acid residues are adjacent, highly conserved and the sites of the *ram* mutations. RNA footprinting experiments have shown that the RNA-binding sites of S5 cover several defined regions of the 16 S rRNA molecule (Stern *et al.*, 1989; Powers & Noller, 1995). Recently, more precise site-directed hydroxyl radical cleavage data have been obtained for specific residues within the C-terminal domain (Heilek & Noller, 1996). Position 130 (129 in *E. coli*) shows little interaction with rRNA, whereas position 100 (99 in *E. coli*) shows a more extensive interaction. Neither position is directly within the putative

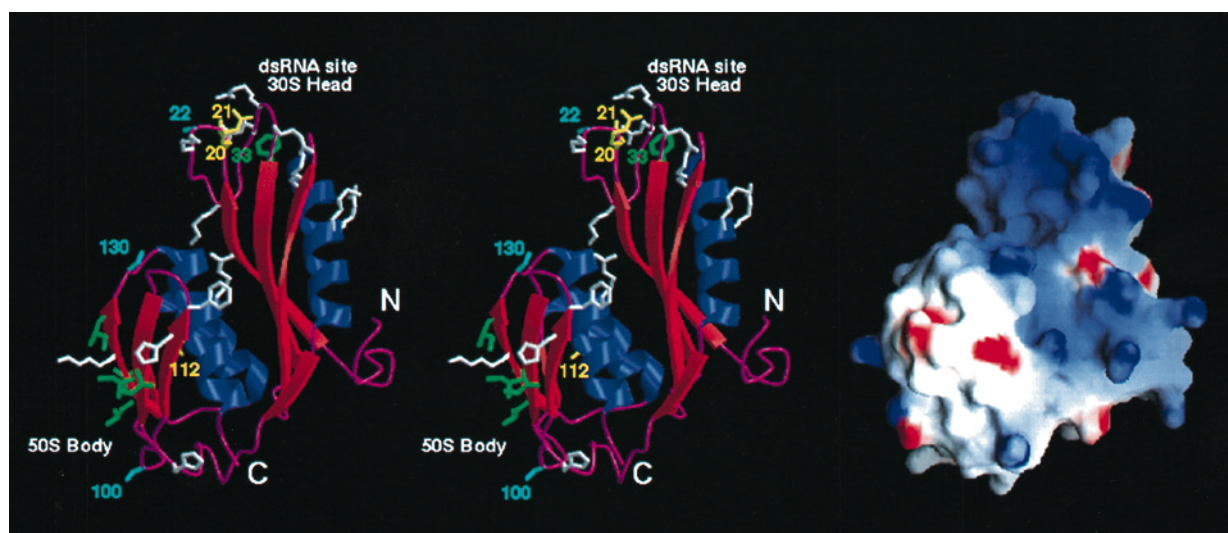


Figure 3. The putative functional sites of ribosomal protein S5. Left, stereo views of the protein with key residues labeled. Individual residues are color-coded as follows: white, basic and aromatic residues; green, exposed hydrophobic residues; yellow, sites of mutations for *ram* (112) and spectinomycin resistance (20, 21 and 22); cyan, residues used for site-specific hydroxyl radical footprinting (22, 100 and 130; Heilek & Noller, 1996). Also shown in green is the conserved phenylalanine 33, which is buried within the putative RNA-binding loop. It is proposed that S5 contains two regions of interaction with ribosomal RNA, one on each domain. The N-terminal domain contains a double-stranded RNA-binding site that interacts with the head of the 30 S subunit. The C-terminal domain interacts with RNA elements in the body of the 30 S subunit. This domain also has an exposed hydrophobic patch on the left that is proposed to interact with another ribosomal protein, possibly S8. The unstructured C-terminal 19 amino acid residues have not been included for clarity. The Figure was produced using MOLSCRIPT (Kraulis, 1991). Right, The surface electrostatic potential of ribosomal protein S5 calculated using GRASP (Nicholls *et al.*, 1991). The orientation of the molecule is identical with that shown on the left. The extreme ranges of red (negative) and blue (positive) represent electrostatic potentials of <-10 to $>+10$ $k_b T$, where k_b is the Boltzmann constant and T is the temperature.

RNA-binding sites, but they are immediately adjacent (Figure 3).

A rather striking surface feature of the C-terminal domain is an extended hydrophobic patch that covers the β -sheet towards the lower end of the protein as viewed in Figure 1. This patch comprises isoleucine 86, 95, 106 and 128, and leucine 124, and is shown in Figure 3. We have reported similar hydrophobic regions on the surfaces of L14 and S8, and suggested that they represent sites of interaction with other ribosomal proteins (Davies *et al.*, 1996a,b). A number of proteins are known to be close to S5 in the ribosome, including S2, S3, S4, S8, S12, S15 and S16 (Lambert *et al.*, 1983; Capel *et al.*, 1987), but the only protein for which there is evidence of a direct interaction is S8. This protein has been shown to form a complex with S5 in solution (Tindall & Aune, 1981), and a protein-protein cross-link between the two molecules has been reported (Allen *et al.*, 1979). We recently determined the structure of S8 and identified a hydrophobic concave surface on the molecule adjacent to the cross-link site that could mediate its interaction with S5 (Davies *et al.*, 1996b).

Murzin (1995) has noted that the C-terminal domain contains a rare left-handed $\beta\alpha\beta$ -unit between the parallel strands 7 and 8 of the mixed β -sheet. He has further noted that the overall topology of the C-terminal domain including this unusual turn is found also in the second domain of the gyrase B molecule and domain IV of elongation factor G. From a detailed comparison of these three protein modules, he has concluded that they probably diverged from a common ancestor, most likely a primitive ribosomal protein.

Ribosomal protein L6

Refinement

The starting point of the refinement was the 2.6 Å structure previously published, which had an *R*-factor of 18.3%, and deviations from ideality in bond lengths and angles of 0.016 Å and 3.37° respectively (Golden *et al.*, 1993b). In this model, which contained 161 residues, seven and nine residues were omitted from the N and C termini, respectively, due to weak or absent electron density. During initial rounds, the model was refined against 2.2 Å resolution data, and the resolution was extended to the 2.0 Å limit at later stages. Between each round of refinement, the model was inspected and manually adjusted as necessary. Water molecules were introduced into the model based on peaks in both the $2F_o - F_c$ and $F_o - F_c$ electron density maps, and suitably adjacent hydrogen-bonding partners. Particular effort was made to extend the N and C termini of the molecule using the gradually improving electron density in these regions. It was eventually possible to include one additional residue at the N terminus and two residues at the C terminus. The final model was inspected using a series of

simulated annealing omit maps in which residues were deleted from the model in blocks of ten along the polypeptide chain. A summary of the refinement statistics and model geometries is shown in Table 2.

Structure description

The structure of the L6 molecule has been described in some detail (Golden *et al.*, 1993b), and only a brief description will be given here, emphasizing features that have been revealed by the refinement process. L6 has an elongated L-shaped structure, and is divided into two structurally homologous domains that appear to have evolved by gene duplication (Figure 4), possibly from an RNA-binding RRM-like motif. The domains are approximately equal in size; the N-terminal domain consists of amino acids 1 through 79, and the C-terminal domain 80 through 177. Each domain comprises a pair of three-stranded antiparallel β -pleated sheets and an α -helix, and these are arranged like a triangular prism in which two β -sheets form two of the faces and the α -helix the third. The two β -sheets contain strands β 1- β 5- β 6 (sheet 1) and β 2- β 3- β 4 (sheet 2), and the α -helix is at the C-terminus. To reflect the homology, corresponding elements within each domain are labeled N β 1, C β 1 etc.

There are two major differences between the two domains. The first is at the N terminus, where N β 1 is very short and contains only four residues,

Table 2. Data collection and refinement parameters for ribosomal protein L6

A. Data Collection	
Resolution (Å)	20.0–2.0
Total number of observations	39,577
Number of unique reflections	13,166
<i>R</i> -merge (%) ^a	8.00
Completeness of data (%)	97.5
Completeness of data (%) (2.00 Å–2.09 Å)	97.4
B. Refinement	
Resolution (Å)	8.0–2.0
Sigma cut-off	0.0
Residues in model	7–170
Number of protein atoms	1251
Number of water molecules	140
Number of reflections	12,972
<i>R</i> -value (%)	23.0
<i>R</i> -value (%) (2.00 Å–2.09 Å)	32.8
<i>R</i> _{free} ^b value (%)	30.1
Overall <i>G</i> factor ^c	0.24
r.m.s. deviation on bonds (Å)	0.014
r.m.s. deviation on angles (°)	1.89
r.m.s. deviation on improper angles (°)	1.45
r.m.s. deviation on dihedral angles (°)	24.9
Mean <i>B</i> factor (main-chain) (Å ²)	29.3
r.m.s. deviation in <i>B</i> factor (main-chain) (Å ²)	1.590
Mean <i>B</i> factor (side-chain) (Å ²)	40.0
r.m.s. deviation in <i>B</i> factor (side-chain) (Å ²)	3.182
Residues in most favored regions (%)	88.8
Residues in additional allowed regions (%)	11.2

^a $R\text{-merge} = \sum |I_i - I_m| / \sum I_m$.

^b Brünger (1992).

^c Laskowski *et al.* (1993).

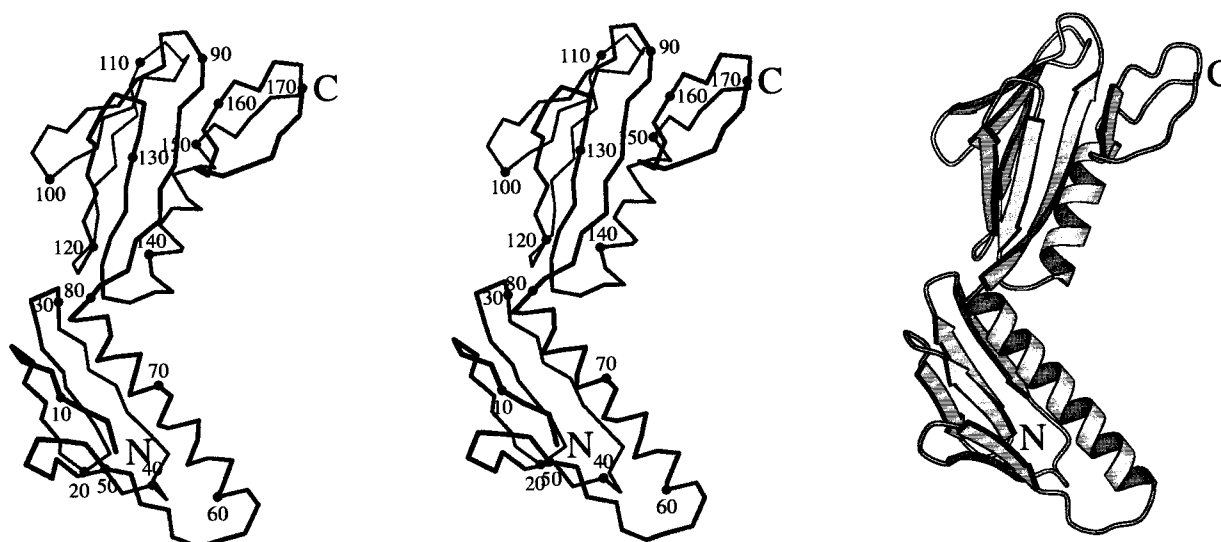


Figure 4. The overall structure of ribosomal protein L6. Left, stereo views of the $C\alpha$ trace of the L6 backbone with every tenth residue labeled and marked with a filled circle. Right, a ribbon representation of L6 showing the elements of secondary structure. The Figure was produced using MOLSCRIPT (Kraulis, 1991).

whereas C β 1 is longer with eight residues. However, the refinement has revealed weak but clear electron density at the N terminus, which indicates that some of the disordered residues at the extreme N terminus extend the strand N β 1. It appears that the end of this outer strand is frayed out from the β -sheet, and the N-terminal domain does contain

an appropriately located hydrophobic depression that could accommodate this short element (Figure 5). As described below, these N-terminal residues probably become more ordered upon binding RNA. The second difference is at the C terminus. N α 1 contains six helical turns and leads directly into C β 1, whereas C α 1 is shorter with four

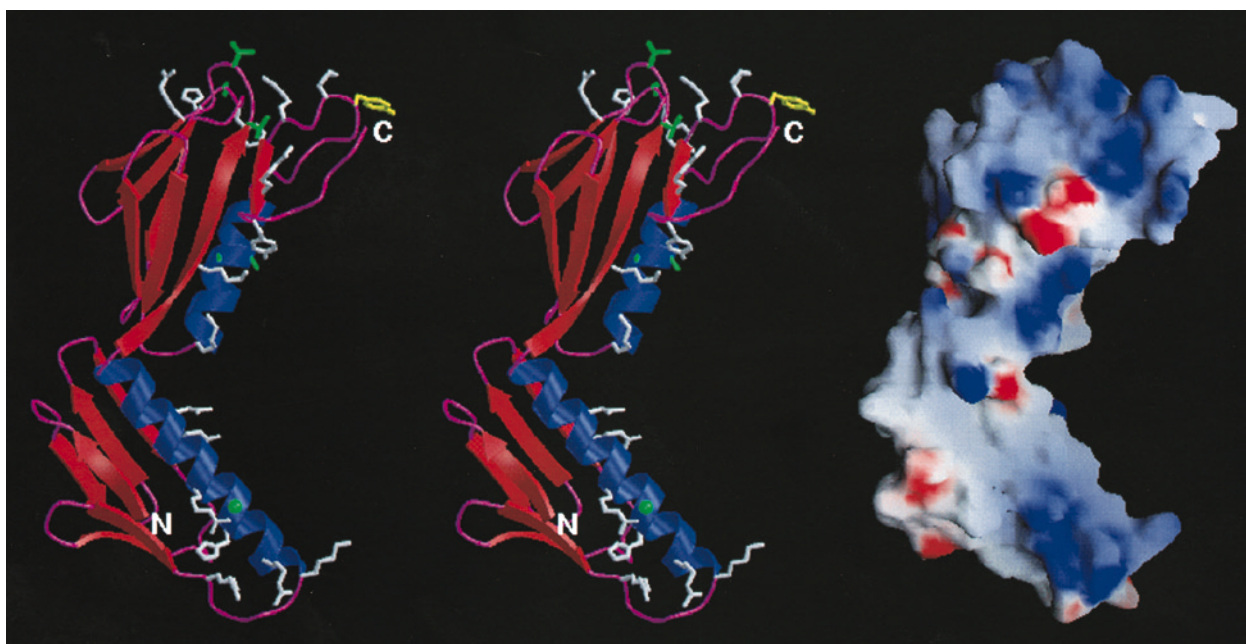


Figure 5. The putative functional sites of ribosomal protein L6. Left, stereo views of the protein with key residues labeled. Basic and aromatic residues are white, exposed hydrophobic residues are colored green, and the yellow side-chain is tyrosine 156 that has been cross-linked to ribosomal RNA (Urlaub *et al.*, 1995). L6 has two putative sites of interaction with ribosomal RNA at opposite ends of the molecules (top and bottom). The Figure was produced using MOLSCRIPT (Kraulis, 1991). Right, the surface electrostatic potential of ribosomal protein L6 calculated using GRASP (Nicholls *et al.*, 1991). The orientation of the molecule is identical with that shown on the left. The extreme ranges of red (negative) and blue (positive) represent electrostatic potentials of <-10 to $>+10$ k_bT , where k_b is the Boltzmann constant and T is the temperature.

helical turns and is followed by a structured loop-strand-loop. The additional β -strand attaches to the outside of sheet 1, which significantly increases the homology of the C-terminal domain to the RRM motif.

RNA-binding regions

We previously identified a putative RNA-binding site on the C-terminal domain of L6 (Golden *et al.*, 1993b). It is centered on the loop-strand-loop at the C terminus, and comprises arginine 34, 94, 162 and 169, lysine 84, 137, 157 and 159, tyrosine 108, 156 and 163, histidine 110, asparagine 73 and serine 109. The disordered C terminus, which contains three conserved lysine residues, 171, 174 and 177, probably contributes to this region. Recently, a cross-link has been obtained between 23 S rRNA and tyrosine 156, which confirms that this region of L6 is immediately adjacent to RNA in the 50 S particle (Urlaub *et al.*, 1995). It should also be noted that this region is located within the RRM-like part of the C-terminal domain, close to the known RRM RNA-binding site on the surface of the β -sheet (Oubridge *et al.*, 1994). This site probably interacts with a highly ordered region of the 23 S rRNA molecule, because the central loop-strand-loop has four conserved features that combine to produce a rigid conformation (Figure 6). First, as described above, the β -strand is secured to the outside of sheet 1. Second, the loop preceding the β -strand contains three proline residues (152, 153 and 155) to provide minimal conformational flexibility. Third, the side-chain of glutamate 154, which is between the proline residues, reaches across and makes hydrogen bonds to the main-chain amide groups further down the loop, just prior to the β -strand. Finally, the second loop is firmly anchored to the end of C α 1 by a conserved

salt-bridge between arginine 148 and glutamate 166.

With the aid of the large number of L6 sequences that have been characterized since we first published the structure, it is now apparent that a second region of L6 probably interacts with rRNA. This new region is located at the tip of the N-terminal domain, at the opposite end of the molecule from the first region described above, and includes the disordered N terminus (Figure 5). The residues involved are arginine 2, 53, 61 and 68, lysine 5, 6 and 58, and histidine 64. Glycine 65, which is 100% conserved in all bacterial L6 sequences, is also in this region and appears to create a surface gap, perhaps to allow the intimate contact with rRNA. Residues 1 to 6 are disordered, but as noted above, at least three of these appear to constitute a frayed out β -strand that may become more ordered upon binding rRNA.

Mutations that encode gentamicin resistance

The L6 gene is encoded by nucleotides 2251 to 2784 in the *E. coli* *spc* operon (Cerretti *et al.*, 1983), and positions in the gene described below are with respect to these numbers. Initial attempts to clone the intact L6 gene from the strains GE 20-8 and GS 50-15 failed because the *E. coli* cells that incorporated the recombinant DNA were not viable. Assuming that this was due to a severe toxicity of the mutant L6 protein, the gene was cloned in two segments that corresponded to the N and C-terminal domains of the protein (nucleotides 2251 to 2550 and 2517 to 2850, respectively). Clones for the N-terminal domain were obtained, but cells containing the C-terminal domain construct still failed to grow. Since the C-terminal domain can probably fold in the absence of the N-terminal domain, we surmised that the former was the toxic half of the

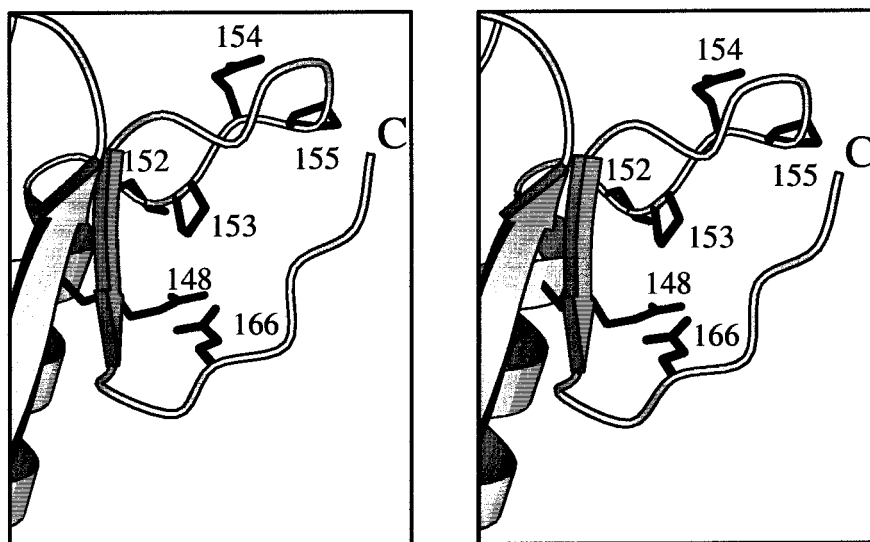


Figure 6. A stereo view of the putative RNA-binding loop-strand-loop region of ribosomal protein L6. Each of the residues shown is highly conserved and provides conformational rigidity to the C terminus of the protein. The majority of this region is deleted in the gentamicin-resistant mutants of L6. See the text for details.

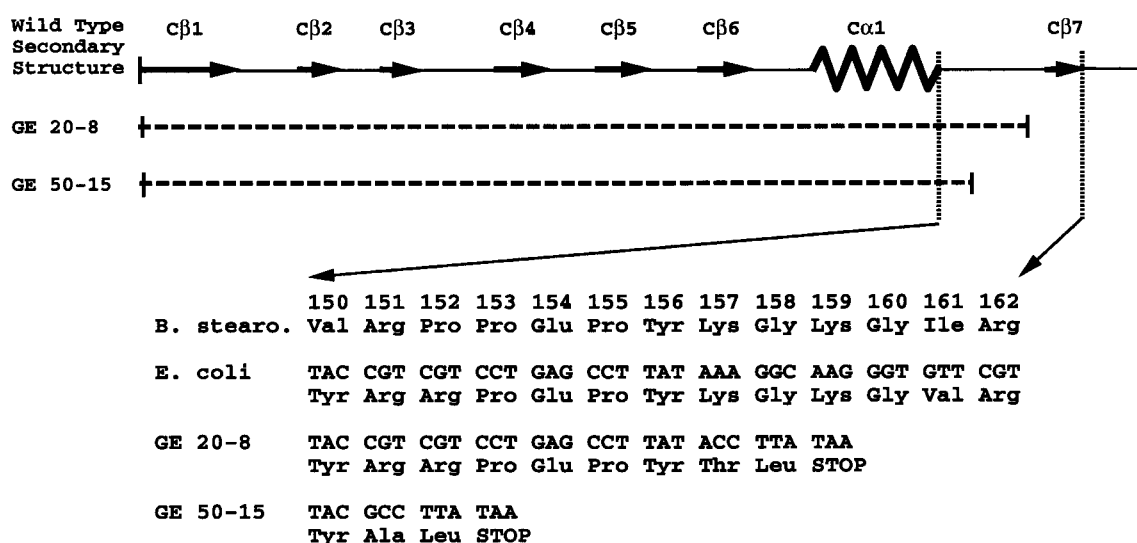


Figure 7. Details of the L6 gentamicin resistance mutants from *E. coli*. Two mutants have been characterized, GE 20-8 and GE 50-15, and both have deletions at the C terminus resulting from the introduction of STOP codons at the 3' end of the gene. The nucleotide and amino acid sequences prior to the mutation are shown for the wild-type and mutant forms, and the corresponding peptide sequence from *B. stearothermophilus* is shown for comparison. The linear arrangement of the L6 secondary structure elements is shown at the top.

protein and further subdivided its gene into two overlapping portions (nucleotides 2516 to 2698 and 2607 to 2854) that would code for two non-folding polypeptides. This strategy was successful, and the entire sequence was obtained.

The two gentamicin-resistant strains contain different mutations, but both result in similar and drastic changes at the extreme C terminus of the molecule. In strain GE 20-8, there was a 7 nt insertion at nucleotide 2722. These 7 nt were a repeat or

a stutter of the previous 7 nt, and this introduced a frame shift and premature stop codon into the gene. When this gene is translated, the wild-type sequence of the L6 protein would end prematurely at residue 158. In strain GS 50-15, an 11 residue deletion occurred after nucleotide 2703, which would result in a frame shift, alterations in residues 151 and 152, and a premature end at residue 152. The results are summarized in Figures 7 and 8. Both of these deletions completely remove

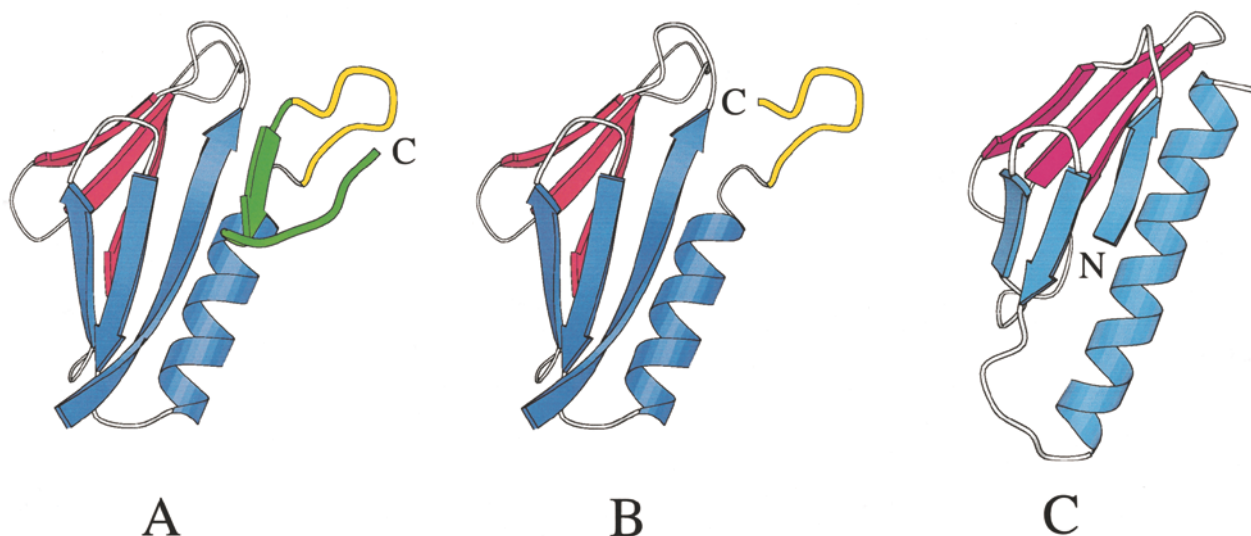


Figure 8. Structural mutagenesis of the two domains of ribosomal protein L6. Shown are the wild-type C-terminal domain (A), the mutated gentamicin-resistant C-terminal domain (B) and the N-terminal domain (C). Secondary structure elements are color-coded to emphasize the homology to the RRM motif (blue/yellow/green), and the sheet 2 insertion (red). The yellow and green regions are deleted in the two gentamicin-resistant mutants of L6. Note that the resulting domain is homologous to the N-terminal domain, suggesting that the mutants would have structurally viable C-terminal domains. The image of the deletion mutant does not take into account any structural rearrangements caused by the deletion and is merely a useful model.

strand C β 7, part of the loop connecting C β 7 to C α 1 and the disordered C terminus. These elements contain the majority of the residues within C-terminal domain RNA-binding site, and they also constitute the major structural differences between the two homologous halves of the L6 molecule.

Discussion

The RNA-binding sites of S5 and L6

It is now generally accepted that the role of ribosomal proteins is to assist in the folding of the ribosomal RNA and to help maintain its functional structure. Consistent with this role, many of the proteins, in particular the primary rRNA-binding proteins, appear to have multiple RNA-binding sites that presumably bring together two or more regions of the RNA at the protein locus (Ramakrishnan & White, 1998). In the "larger" proteins, these binding sites are often on two separate domains that probably originated from the fusion of two primitive RNA-binding proteins. The S5 and L6 molecules both contain these features, and the general arrangement of their RNA-binding sites is in fact quite similar. Each has a well-ordered site on its C-terminal domain, and a more disordered site on the N-terminal domain, largely populated with basic residues. Recent hydroxyl radical footprinting experiments on S5 demonstrate how this type of architecture can mediate the folding of ribosomal RNA (Heilek & Noller, 1996). The protein initially binds rRNA using the site in the C-terminal domain. Although S5 is not a primary RNA-binding protein, this is apparently a highly specific interaction involving a 16 S rRNA locus created early in the assembly process. The second RNA-binding event occurs late in assembly and appears to be a relatively non-specific tethering of an adjacent RNA moiety, in this case a region of dsRNA. L6 is also incorporated into the ribosome at a later stage of assembly (Herold & Nierhaus, 1987), and it probably has a very similar role in folding the 23 S rRNA molecule.

The L6 gentamicin resistance mutations

Two mutations in L6 that give rise to gentamicin resistance result in severe truncations at the C terminus of the C-terminal domain; in GE 20-8 the C-terminal 18 residues have been deleted, and in GS 50-15 the C-terminal 25 residues are missing. These deletions have a dramatic effect on the putative specific RNA-binding site of L6 (Figures 5 and 6), and also result in the loss of β -strand C β 7, which is at the heart of the RRM homology. Although most of the previously characterized protein mutations that produce antibiotic resistance are the result of single amino acid changes (Ramakrishnan & White, 1992; Golden *et al.*, 1993a), the results with L6 support the general notion that the mutations affect the translational

machinery by creating locally distorted regions of RNA. In support of this general idea, it has been shown that mutations in S4 and S12, and translational miscoding agents all elicit structural changes within regions of the 16 S rRNA associated with translational accuracy (Allen & Noller, 1989; Powers & Noller, 1994).

When a mutation results in the deletion of a protein region, the possibility always exists that the phenotype is simply the result of a misfolded or unfolded protein. We were unable to verify this directly for L6, because it was impossible to clone either the full-length mutant proteins or their C-terminal domains to perform structural analysis. We believe that the proteins are folded for three reasons. First, attempts to clone the mutant C-terminal domain failed, probably due to their toxicity to *E. coli*. This suggests that low-level expression of the domain is restrictive to cell growth, implying a folded, functional structure. Secondly, as is evident from Figures 6 and 8, the deleted region is packed on the outside of the C-terminal domain and contributes very little to the domain's integrity. Indeed, the deletion simply generates a structure similar to that of the N-terminal domain, and the overall stability of L6 is unlikely to be severely compromised. Finally, the mutated L6 proteins have actually been shown to be present within the ribosomes of the gentamicin-resistant *E. coli* strains (Buckel *et al.*, 1977). This raises the question of how an L6 molecule with a substantially reduced major RNA-binding site can assemble into the ribosome. L6 is added late in the 50 S assembly process (Herold & Nierhaus, 1987), and protein-protein interactions are likely to be important for its incorporation into the ribonucleoprotein complex. We have previously suggested which surfaces of L6 are involved in protein-protein interactions, and these are in regions of the molecule unaffected by the deletion (Golden *et al.*, 1993b).

The role of S5 in the 30 S subunit

Based on extensive biochemical and functional data, we have previously suggested that S5 interacts with two regions of 16 S rRNA (Ramakrishnan & White, 1992). The C-terminal half, which contains the sites of mutation relating to ribosomal accuracy, is located at the so-called proofreading domain that contains all the 30 S components associated with ribosomal accuracy (Lambert *et al.*, 1983; Stöffler & Stöffler-Meilicke, 1984; Capel *et al.*, 1987; Moazed & Noller, 1987; Stern *et al.*, 1989; Powers & Noller, 1995). The N-terminal half, which contains the sites of mutation that confer resistance to the antibiotic spectinomycin, binds non-specifically to helix 34 (Mueller *et al.*, 1995). This helix encompasses nucleotides 1046 to 1067 and 1211 to 1189, and contains the specific binding site for spectinomycin (Sigmund *et al.*, 1984; Moazed & Noller, 1987). This proposal was controversial because it disagreed with contemporary models of the 30 S subunit (Brimacombe *et al.*,

1988; Stern *et al.*, 1988), but it has since been supported by a number of findings.

Brimacombe and co-workers have shown that mRNA at the A-site and the P-site can be cross-linked to the 530 loop, helix 34, the decoding site of 16 S rRNA, and proteins S5 and S7 (Dontsova *et al.*, 1991, 1992). These data place S5 at a central location in the 30 S subunit, not only adjacent to helix 34 and the 530 loop, but also to the decoding site. The data also suggest that the proofreading and decoding regions are not separated, as they were in previous models, but adjacent. This important conclusion was supported by additional cross-linking data, which revealed that the 530 loop is close to S7 and the 1400 region of 16 S rRNA, which are both accepted components of the decoding site (Muralikrishna & Cooperman, 1994; Alexander *et al.*, 1994). Recently, Noller and co-workers showed directly that the putative RNA-binding loop in the N-terminal domain of S5 is indeed adjacent to helix 34, and further showed that the interaction occurs only in the assembled 30 S particle (Heilek & Noller, 1996). Finally, the finding that the N-terminal domain of S5 is a dsRBD provides a simple model for its interaction with helix 34. This homology is particularly satisfying, since the structure of the N-terminal domain was originally thought to be very unique, but it now has a clear function that is consistent with the S5 environment in the 30 S subunit.

The decoding and proofreading regions are located in the body of the 30 S subunit, and the spectinomycin binding site within helix 34 is located in the head region. Thus, a picture emerges that S5 domains straddles these major subdomains. This location of S5 may have implications for the translocation function of the ribosome, since spectinomycin is thought to inhibit this process (Burns & Cundliffe, 1973; Wallace *et al.*, 1974). For example, could translocation involve a relative movement of the head and body subdomains that is mediated by the binding of S5 to helix 34? It has long been suspected that the translational process is associated with dynamic changes in the ribosome structure, and that these may be mediated by conformational switches in the RNA. Recently, clear evidence for such a conformational switch in 16 S rRNA has been reported, which appears to be mediated by S5 (Lodmell & Dahlberg, 1997). Interestingly, alternative base-pairing schemes are also possible within helix 34 (Glotz & Brimacombe, 1980), and these have been suggested to be the basis of a spectinomycin-sensitive conformational switch (Sigmund *et al.*, 1984).

The role of L6 in the 50 S subunit

It has been shown that the mechanism of L6-mediated antibiotic resistance is distinct from that by which S12 mutations negate the effects of streptomycin and *ram* mutations in the 30 S subunit (Hummel *et al.*, 1980). In general, the 30 S mutations produce resistance to misreading-indu-

cing antibiotics directly, by reducing their binding affinities (Powers & Noller, 1991; Leclerc *et al.*, 1991). However, the 50 S mutations do not affect these affinities (Melançon *et al.*, 1992) but rather slow the translational machinery and increase the opportunity for proofreading at the decoding center. The underlying mechanism appears to involve the process by which the EF-Tu·GTP·aminoacyl tRNA ternary complex binds to the ribosome and is subsequently turned over (Thompson, 1988). Aminoacyl tRNA is brought to the ribosome as a complex with elongation factor Tu (EF-Tu) and GTP. However, the interaction of the 3' acceptor stem with the peptidyl transferase site on the 50 S subunit and subsequent peptide bond formation occurs only after GTP turnover and the dissociation of the EF-Tu·GDP complex (Moazed & Noller, 1989). Recently, this proofreading mechanism was directly observed in an electron microscope image of the ribosome with the bound ternary complex (Stark *et al.*, 1997). Any mutation of a 50 S component that alters the rate of this dissociation either increases or decreases the speed of translation.

L6 is well placed to be such a component. First, it is directly at the subunit interface (Lambert & Traut, 1981) and is positioned close to the tRNA-binding site below the central protuberance (Stöffler-Meilicke *et al.*, 1983; Hackl & Stöffler-Meilicke, 1988). Second, L6 is close to the GTPase center in the region of the L7/L12 stalk (Walleczek *et al.*, 1988; Spirin & Vasiliev, 1989), and can be cross-linked to EF-G (Sköld, 1982), which binds in the same location as EF-Tu. Third, L6 has been associated with domains V and VI of 23S rRNA (Wower *et al.*, 1981; Leffers *et al.*, 1988), both of which have been implicated in the binding of the EF-Tu·GTP·aminoacyl tRNA ternary complex (Leffers *et al.*, 1988; Moazed *et al.*, 1988; Moazed & Noller, 1989; Mitchell *et al.*, 1990).

The only other 50 S component in which mutations affect proofreading is the α -sarcin loop. Mutations at nucleotide 2661 within the loop result in an increase in the binding affinity of the ternary complex, a reduction in the rate of translation, and negation of the misreading effects of streptomycin, gentamicin and neomycin without reducing their binding affinities (Tappich & Dahlberg, 1990; Melançon *et al.*, 1992). These results are consistent with independent data that show that the α -sarcin loop forms part of the ternary complex binding site (Leffers *et al.*, 1988; Moazed *et al.*, 1988). The mutations at nucleotide 2661 cause a phenotype that is very similar to that of the L6 mutants, and it has been proposed that L6 stabilizes or modulates the RNA structure that forms the ternary complex binding site (Melançon *et al.*, 1992). Our analysis of the L6 mutants suggests that this interaction would be mediated by the highly ordered RNA-binding site in the C-terminal domain. If correct, L6 should prove to be a useful vehicle for probing this important local RNA structure in the same

way that S5 has been used to study the 16 S RNA molecule (Heilek & Noller, 1996).

Materials and Methods

High-resolution data collection

New native data were collected for both S5 and L6 to extend the resolution of the two models to the diffraction limits of the crystals. Diffraction data were collected using an RAXIS II image plate system equipped with Yale mirror optics (Molecular Structure Corporation), mounted on an RU300 X-ray generator operating at 40kV and 80mA. Data were collected at room temperature using the standard oscillation method (Wonacott, 1977) and a crystal-to-detector distance of 100 mm. S5 crystallizes in the trigonal space group $P3_221$ with cell dimensions $a = b = 59.3 \text{ \AA}$ and $c = 109.8 \text{ \AA}$. Due to crystal morphology, which results in a preferred alignment of the c axis normal to the X-ray beam, it was necessary to collect two data sets from two separate crystals mounted orthogonally to fill in the blind region. The data comprised $45 \times 2^\circ$ oscillation frames collected around the a^* axis, and $30 \times 2^\circ$ oscillation frames around the c^* axis, using an exposure time of 15 minutes per degree. These data were processed and merged using the RAXIS suite of programs to give a final data set that was 93% complete to 2.2 \AA (Table 1). L6 crystallizes in the hexagonal space group $P6_122$ with cell dimensions $a = b = 71.8 \text{ \AA}$ and $c = 124.5 \text{ \AA}$. A single data set was collected comprising $80 \times 0.35^\circ$ frames, using an exposure time of 45 minutes per degree. The data were integrated, scaled and merged using HKL (Otwinowski, 1993) to give a final data set that was 97.5% complete to 2.0 \AA (Table 2).

Refinement

The refinement was performed using the program X-PLOR (Brünger *et al.*, 1987), and model building was done using the O program (Jones *et al.*, 1991). The stereochemical parameters of the final model were analyzed by the program PROCHECK (Laskowski *et al.*, 1993).

Preparation of genomic DNA

The *E. coli* strains GE 20-8 and GS 50-15 containing the mutant L6 genes were grown in LB containing 20 mg/l gentamicin. Total genomic DNA was isolated from cells grown in 50 ml overnight cultures. The cells were harvested by centrifugation at 12,000 rpm for 15 minutes, washed with TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA), and frozen at -80°C to promote cell lysis. The cell pellet was thawed, and resuspended in 4 ml of TE buffer and 0.1 M EDTA, and SDS was added to 0.5% (w/v). To complete the cell lysis, 1 mg of lysozyme was added to the solution, which was then placed on ice for five minutes. To remove protein, 2 mg of proteinase K was added, and the solution was further incubated at 65°C overnight. The mixture was then brought to room temperature, and extracted twice with phenol/chloroform/isoamyl alcohol (25:24:1, by vol.). The DNA was precipitated by overlaying the solution with ice-cold 100% ethanol, and it was removed from the interface by spooling onto a Pasteur pipette. The DNA was finally transferred to an Eppendorf tube, and re-dissolved in TE buffer at room temperature.

Cloning and sequencing of the mutant L6 genes

The two mutant genes encoding ribosomal protein L6 were PCR-amplified and cloned in three parts because it was impossible to clone the intact genes. The oligonucleotides used in the PCR amplification were 29 to 40 bases in length and designed to amplify the nucleotides encoding the N-terminal domain, and overlapping halves of the C-terminal domain. The oligonucleotides also introduced unique restriction sites for *Xba*I and *Bam*HI at the 5' and 3' ends for cloning. The oligonucleotide sequences are listed below, and the restriction sites that were incorporated are underlined.

oligo 1a	5'-ATGACTGAT <u>CTAG</u> AAGCGCGCCAA-GCTGGTCTTGGTGGCG-3'
oligo 1b	5'-CCTTTAACGGATCCACGGTAACCT-ACACCAACCAGCTGC-3'
oligo 2a	5'-GGTTGGT <u>CTAG</u> ATTACCGTGCAGC-GGTTAAAGGGAATG-3'
oligo 2b	5'-CGCGCGGATCCCGCTGCAACCTGGC-CGATC-3'
oligo 3a	5'-GGTATCACT <u>CTAG</u> AATGTCCGACT-CAGACTG-3'
oligo 3b	5'-GCTCCTGGATCCTGCGGCGTGCGC-GGGTCGCACG-3'

PCR products were quantified on a 1% (w/v) agarose gel, and digested with *Xba*I and *Bam*HI. The digested DNA fragments were gel-purified, and ligated into the *Xba*I/*Bam*HI sites of pET13a. These were transformed into competent *E. coli* DH5a cells and were then plated onto LB agar with 25 mg/ml kanamycin and grown overnight at 37°C . Plasmid DNA from resulting colonies was extracted and sequenced using oligonucleotides complementary to the plasmid sequences flanking the gene

Acknowledgments

This work was supported by grant GM44973 from the National Institutes of Health (to S.W.W. and V.R.), and by the American Lebanese Syrian Associated Charities (ALSAC).

References

- Alexander, R. W., Muralikrishna, P. & Cooperman, B. S. (1994). Ribosomal components neighboring the conserved 518–533 loop of 16S rRNA in 30S subunits. *Biochemistry*, **33**, 12109–12118.
- Allen, G., Capasso, R. & Gualerzi, C. (1979). Identification of the amino acid residues of proteins S5 and S8 adjacent to each other in the 30S ribosomal subunit of *Escherichia coli*. *J. Biol. Chem.* **254**, 9800–9806.
- Allen, P. N. & Noller, H. F. (1989). Mutations in ribosomal proteins S4 and S12 influence the higher order structure of 16 S ribosomal RNA. *J. Mol. Biol.* **208**, 457–468.
- Bollen, A., Davies, J., Ozaki, M. & Mizushima, S. (1969). Ribosomal protein conferring sensitivity to the antibiotic spectinomycin in *Escherichia coli*. *Science* **165**, 85–86.

- Brimacombe, R., Atmadja, J., Stiege, W. & Schüler, D. (1988). A detailed model of the three-dimensional structure of *E. coli* 16 S ribosomal RNA *in situ* in the 30 S subunit. *J. Mol. Biol.* **199**, 115–136.
- Brimacombe, R., Mitchell, P., Osswald, M., Stade, K. & Bochkariov, D. (1993). Clustering of modified nucleotides at the functional center of bacterial ribosomal RNA. *FASEB J.* **7**, 161–167.
- Brünger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
- Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R-factor refinement by molecular dynamics. *J. Mol. Biol.* **203**, 803–816.
- Buckel, P., Buchberger, A., Böck, A. & Wittmann, H.-G. (1977). Alteration of ribosomal protein L6 in mutants of *Escherichia coli* resistant to gentamicin. *Mol. Gen. Genet.* **158**, 47–54.
- Burd, C. G. & Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.
- Burley, S. K. & Petsko, G. A. (1985). Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*, **229**, 23–28.
- Burns, D. J. & Cundliffe, E. (1973). Bacterial protein synthesis: a novel system for studying antibiotic action *in vivo*. *Eur. J. Biochem.* **37**, 570–574.
- Bycroft, M., Grunert, S., Murzin, A. G., Proctor, M. & St. Johnson, D. (1995). NMR solution structure of a double-stranded RNA-binding domain from *Drosophila* staufer protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J.* **14**, 3563–3571.
- Capel, M. S., Engleman, D. M., Freeborn, B. R., Kjeldgaard, M., Langer, J. A., Ramakrishnan, V., Schindler, D. G., Schneider, D. K., Schoenborn, B. P., Yabuki, S. & Moore, P. B. (1987). A complete mapping of the proteins in the small ribosomal subunit of *Escherichia coli*. *Science*, **238**, 1403–1406.
- Castiglione Morelli, M. A., Stier, G., Gibson, T., Joseph, C., Musco, G., Pastore, A. & Trave, G. (1995). The KH module has an $\alpha\beta$ fold. *FEBS Lett.* **358**, 193–198.
- Cerretti, D. P., Dean, D., Davis, G. R., Bedwell, D. M. & Nomura, M. (1983). The *spc* ribosomal protein operon of *Escherichia coli*: sequence and cotranscription of the ribosomal protein genes and a protein export gene. *Nucl. Acids Res.* **11**, 2599–2616.
- Davies, C., White, S. W. & Ramakrishnan, V. (1996a). The crystal structure of ribosomal protein L14 reveals an important organizational component of the translational apparatus. *Structure*, **4**, 55–66.
- Davies, C., Ramakrishnan, V. & White, S. W. (1996b). Structural evidence for specific S8-RNA and S8-protein interactions within the 30S ribosomal subunit: the structure of ribosomal protein S8 from *Bacillus stearothermophilus* at 1.9 Å. *Structure*, **4**, 1093–1104.
- Dontsova, O., Kopylov, A. & Brimacombe, R. (1991). The location of mRNA in the ribosomal 30S initiation complex; site-directed cross-linking of mRNA analogues carrying several photo-reactive labels simultaneously on either side of the AUG start codon. *EMBO J.* **10**, 2613–2620.
- Dontsova, O., Dokudovskaya, S., Kopylov, A., Bogdanov, A., Rinke-Appel, J., Jünke, N. & Brimacombe, R. (1992). Three widely separated positions in the 16S RNA lie in or close to the ribosomal decoding region; a site-directed cross-linking study with mRNA analogues. *EMBO J.* **11**, 3105–3116.
- Funatsu, G., Nierhaus, K. & Wittmann-Liebold, B. (1972). Ribosomal proteins. XXII. Studies on the altered protein S5 from a spectinomycin-resistant mutant of *Escherichia coli*. *J. Mol. Biol.* **64**, 210–209.
- Glötz, C. & Brimacombe, R. (1980). An experimentally-derived model for the secondary structure of the 16S ribosomal RNA from *Escherichia coli*. *Nucl. Acids Res.* **8**, 2377–2395.
- Golden, B. L., Hoffman, D. W., Ramakrishnan, V. & White, S. W. (1993a). Ribosomal protein S17: characterization of the three-dimensional structure by ^1H and ^{15}N NMR. *Biochemistry*, **32**, 12812–12820.
- Golden, B. L., Ramakrishnan, V. & White, S. W. (1993b). Ribosomal protein L6: structural evidence of gene duplication from a primitive RNA-binding protein. *EMBO J.* **12**, 4901–4908.
- Gorini, L. & Kataja, E. (1964). Phenotypic repair by streptomycin of defective phenotypes in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **51**, 487–493.
- Guthrie, C. & Nomura, M. (1969). Structure and function of *E. coli* ribosomes, VIII, Cold-sensitive mutants defective in ribosome assembly. *Proc. Natl Acad. Sci. USA*, **63**, 384–391.
- Hackl, W. & Stöffler-Meilicke, M. (1988). Immunoelectron microscopic localization of ribosomal proteins from *Bacillus stearothermophilus* that are homologous to *Escherichia coli* L1, L6, L23 and L29. *Eur. J. Biochem.* **174**, 431–435.
- Heilek, G. M. & Noller, H. F. (1996). Site-directed hydroxyl radical probing of the rRNA neighborhood of ribosomal protein S5. *Science*, **272**, 1659–1662.
- Herold, M. & Nierhaus, K. H. (1987). Incorporation of six additional proteins to complete the assembly map of the 50S subunit from *Escherichia coli* ribosomes. *J. Biol. Chem.* **262**, 8826–8833.
- Hoffman, D. W., Query, C. C., Golden, B. L., White, S. W. & Keene, J. D. (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.
- Hoffman, D. W., Cameron, C. S., Davies, C., White, S. W. & Ramakrishnan, V. (1996). Ribosomal protein L9: a structure determination by the combined use of X-ray crystallography and NMR spectroscopy. *J. Mol. Biol.* **264**, 1058–1071.
- Hummel, H., Piepersberg, W. & Böck, A. (1980). 30S subunit mutations relieving restriction of ribosomal misreading caused by L6 mutations. *Mol. Gen. Genet.* **179**, 147–153.
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect A*, **47**, 110–119.
- Kharrat, A., Macias, M. J., Gibson, T. J., Nilges, M. & Pastore, A. (1995). Structure of the dsRNA binding domain of *E. coli* RNase III. *EMBO J.* **14**, 3572–3584.
- Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
- Kuhberger, R., Piepersberg, W., Petzet, A., Buckel, P. & Böck, A. (1979). Alteration of ribosomal protein L6 in gentamicin-resistant strains of *Escherichia coli*. Effects on fidelity of protein synthesis. *Biochemistry*, **18**, 187–193.

- Kurland, C. G., Jørgensen, F., Richter, A., Ehrenberg, M., Bilgin, N. & Rojas, A.-M. (1990). Through the accuracy window. In *The Ribosome: Structure, Function and Evolution*, pp. 513–526, American Society for Microbiology, Washington, DC.
- Lambert, J. M. & Traut, R. R. (1981). The subunit interface of the *Escherichia coli* ribosome: identification of proteins at the interface between the 30 S and 50 S subunit by crosslinking with 2-iminothiolane. *J. Mol. Biol.* **149**, 451–476.
- Lambert, J. M., Boileau, G., Cover, J. A. & Traut, R. R. (1983). Cross-links between ribosomal proteins of 30S subunits in 70S tight couples and in 30S subunits. *Biochemistry*, **22**, 3913–3920.
- 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.
- Leclerc, D., Melançon, P. & Brakier-Gingras, L. (1991). Mutations in the 915 region of *Escherichia coli* 16S ribosomal RNA reduce the binding of streptomycin to the ribosome. *Nucl. Acids Res.* **19**, 3973–3977.
- Leffers, H., Egebjerg, J., Anderson, A., Christensen, T. & Garrett, R. A. (1988). Domain VI of *Escherichia coli* 23 S ribosomal RNA. Structure, assembly and function. *J. Mol. Biol.* **204**, 507–522.
- Lodmell, J. S. & Dahlberg, A. E. (1997). A conformational switch in *E. coli* 16S ribosomal RNA during decoding of messenger RNA. *Science*, **277**, 1262–1267.
- Melançon, P., Lemieux, C. & Brakier-Gingras, L. (1988). A mutation in the 530 loop of *E. coli* 16S ribosomal RNA causes resistance to streptomycin. *Nucl. Acids Res.* **16**, 9631–9639.
- Melançon, P., Tappich, W. E. & Brakier-Gingras, L. (1992). Single-base mutations at position 2661 of *Escherichia coli* 23S rRNA increase efficiency of translational proofreading. *J. Bacteriol.* **174**, 7896–7901.
- Mitchell, P., Osswald, M., Schueler, D. & Brimacombe, R. (1990). Selective isolation and detailed analysis of intra-RNA crosslinks induced in the large ribosomal subunit of *E. coli*: a model for the tertiary structure of the tRNA binding domain in 23S RNA. *Nucl. Acids Res.* **18**, 4325–4333.
- Moazed, D. & Noller, H. F. (1987). Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature*, **327**, 389–394.
- Moazed, D. & Noller, H. F. (1989). Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E sites. *Cell*, **57**, 585–597.
- Moazed, D., Robertson, J. M. & Noller, H. F. (1988). Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S RNA. *Nature*, **334**, 362–364.
- Montandon, P. E., Wagner, R. & Stutz, E. (1986). *E. coli* ribosomes with a C912 to U base change in the 16S rRNA are streptomycin resistant. *EMBO J.* **5**, 3705–3708.
- Mueller, F., Döring, T., Erdemir, T., Greuer, B., Jünke, N., Osswald, M., Rinke-Appel, J., Stade, K., Thamm, S. & Brimacombe, R. (1995). Getting closer to an understanding of the three-dimensional structure of ribosomal RNA. *Biochem. Cell Biol.* **73**, 767–773.
- Muralikrishna, P. & Cooperman, B. S. (1994). A photolabile oligodeoxyribonucleotide probe of the decoding site in the small subunit of the *Escherichia coli* ribosome: identification of neighbouring ribosomal components. *Biochemistry*, **33**, 1392–1398.
- Murzin, A. G. (1995). A ribosomal protein module in EF-G and DNA gyrase. *Nature Struct. Biol.* **2**, 25–26.
- Nagai, K., Oubridge, C., Jessen, T. H., Li, J. & Evans, P. R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*, **348**, 515–520.
- 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**, 281–296.
- Osswald, M., Greuer, B., Brimacombe, R., Stöffler, G., Bäumert, H. & Fasold, H. (1987). RNA-protein crosslinking in *Escherichia coli* 30S ribosomal subunits: determination of sites on 16S RNA that are cross-linked to proteins S3, S4, S5, S7, S8, S9, S11, S19 and S21 by treatment with methyl *p*-azidophenyl acetimidate. *Nucl. Acids Res.* **15**, 3221–3240.
- Otwinowski, Z. (1993). Data collection and processing. In *Proceedings of the CCP4 Study Weekend*, pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
- Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H. & 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.
- Piepersberg, W., Böck, A. & Wittmann, H.-G. (1975a). Effect of different mutations in ribosomal protein S5 of *Escherichia coli* on translational fidelity. *Mol. Gen. Genet.* **140**, 91–100.
- Piepersberg, W., Böck, A., Yaguchi, M. & Wittmann, H.-G. (1975b). Genetic position and amino acid replacements of several mutations in ribosomal protein S5 from *Escherichia coli*. *Mol. Gen. Genet.* **143**, 43–52.
- Powers, T. & Noller, H. F. (1991). A functional pseudoknot in 16S ribosomal RNA. *EMBO J.* **10**, 2203–2214.
- Powers, T. & Noller, H. F. (1994). Selective perturbation of G530 of 16 S rRNA by translational miscoding agents and a streptomycin-dependence mutation in protein S12. *J. Mol. Biol.* **235**, 156–172.
- Powers, T. & Noller, H. F. (1995). Hydroxyl radical footprinting of ribosomal proteins on 16S rRNA. *RNA*, **1**, 194–209.
- Ramakrishnan, V. & White, S. W. (1992). The structure of ribosomal protein S5 reveals sites of interaction with 16S rRNA. *Nature*, **358**, 768–771.
- Ramakrishnan, V. & White, S. W. (1998). Ribosomal protein structures: Insights into the architecture, mechanism and evolution of the ribosome. *Trends Biochem. Sci.* In the press.
- Sigmund, C., Ettayebi, M. & Morgan, E. A. (1984). Antibiotic resistance mutations in 16S and 23S ribosomal RNA genes of *Escherichia coli*. *Nucl. Acids Res.* **12**, 4653–4663.
- Sköld, S.-E. (1982). Chemical cross-linking of elongation factor G to both subunits of the 70S ribosomes from *Escherichia coli*. *Eur. J. Biochem.* **127**, 225–229.
- Spirin, A. S. & Vasiliev, V. D. (1989). Localization of functional centers on the prokaryotic ribosome: immuno-electron microscopy approach. *Biol. Cell* **66**, 215–223.
- Stark, H., Rodnina, M. V., Rinke-Appel, J., Brimacombe, R., Wintermeyer, W. & van Heel, M. (1997). Visualization of elongation factor Tu on the *Escherichia coli* ribosome. *Nature*, **389**, 403–406.
- Steiner, G., Kuechler, E. & Barta, A. (1988). Photo-affinity labeling at the peptidyl transferase center

- reveals two different positions for the A- and P-sites in domain V of 23S rRNA. *EMBO J.* **7**, 3949–3955.
- Stern, S., Weiser, B. & Noller, H. F. (1988). Model for the three-dimensional folding of 16S ribosomal RNA. *J. Mol. Biol.* **204**, 447–481.
- Stern, S., Powers, T., Changchien, L.-M. & Noller, H. F. (1989). RNA-protein interactions in 30S ribosomal subunits: folding and function of 16S rRNA. *Science*, **244**, 783–790.
- Stöffler, G. & Stöffler-Meilicke, M. (1984). Immunoelectron microscopy of ribosomes. *Annu. Rev. Biophys. Bioeng.* **13**, 303–330.
- Stöffler-Meilicke, M., Epe, B., Steinhäuser, K. G., Woolley, P. & Stöffler, G. (1983). Immunoelectron microscopy of ribosomes carrying a fluorescence label in a defined position: Location of proteins S17 and L6 in the ribosome of *Escherichia coli*. *FEBS Letters*, **163**, 94–98.
- Tappich, W. E. & Dahlberg, A. E. (1990). A single-base mutation at position 2661 in *Escherichia coli* 23S ribosomal RNA affects the binding of ternary complexes to the ribosome. *EMBO J.* **9**, 2649–2655.
- Thompson, R. C. (1988). EF-Tu provides an internal kinetic standard for translational accuracy. *Trends Biochem. Sci.* **13**, 91–93.
- Tindall, S. H. & Aune, K. C. (1981). Assessment by sedimentation equilibrium analysis of a heterologous macromolecular interaction in the presence of self-association: interaction of S5 with S8. *Biochemistry*, **20**, 4861–4866.
- Traut, R. R., Tewari, D. S., Sommer, A., Gavinno, G. R., Olson, H. M. & Glitz, D. G. (1986). Protein topography of ribosomal functional domains: effects of monoclonal antibodies to different epitopes in *Escherichia coli* protein L7/L12 on ribosome function and structure. In *Structure, Function, and Genetics of Ribosomes*, pp. 286–308, Springer-Verlag, New York, NY.
- Urlaub, H., Kruff, V., Bischof, O., Mueller, E. C. & Wittmann-Liebold, B. (1995). Protein-rRNA binding features and their structural and functional implications in ribosomes as determined by cross-linking studies. *EMBO J.* **14**, 4578–4588.
- Valegård, K., Murray, J. B., Stonehouse, N. J., van den Worm, S., Stockley, P. G. & Liljas, L. (1997). The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein-RNA interactions. *J. Mol. Biol.* **270**, 724–738.
- Vester, B. & Garrett, R. A. (1988). The importance of highly conserved nucleotides in the binding region of chloramphenicol at the peptidyl transfer center of *Escherichia coli* 23S ribosomal RNA. *EMBO J.* **7**, 3577–3587.
- Wallace, B. J., Tai, P.-C. & Davis, B. D. (1974). Selective inhibition of initiating ribosomes by spectinomycin. *Proc. Natl Acad. Sci. USA* **71**, 1634–1638.
- Walleczek, J., Schüler, D., Stöffler-Meilicke, M., Brimacombe, R. & Stöffler, G. (1988). A model for the spatial arrangement of the proteins in the large subunit of the *Escherichia coli* ribosome. *EMBO J.* **7**, 3571–3576.
- Walleczek, J., Martin, T., Redl, B., Stöffler-Meilicke, M. & Stöffler, G. (1989a). Comparative cross-linking study on the 50S ribosomal subunit from *Escherichia coli*. *Biochemistry*, **28**, 4099–4105.
- Walleczek, J., Redl, B., Stöffler-Meilicke, M. & Stöffler, G. (1989b). Protein-protein cross-linking of the 50S ribosomal subunit of *Escherichia coli* using 2-iminothiolane. *J. Biol. Chem.* **264**, 4231–4237.
- Wimberley, B. T., White, S. W. & Ramakrishnan, V. (1997). The structure of ribosomal protein S7 reveals a β -hairpin motif that binds double-stranded nucleic acids. *Structure*, **5**, 1187–1198.
- Wonacott, A. J. (1977). Geometry of the rotation method. In *The Rotation Method in Crystallography*, pp. 75–105, North-Holland Publishing, Oxford.
- Wower, I., Wower, J., Meinke, M. & Brimacombe, R. (1981). The use of 2-iminothiolane as an RNA-protein crosslinking agent in *Escherichia coli* ribosomes, and the localization on 23S RNA of sites crosslinked to proteins L4, L6, L21, L23, L27 and L29. *Nucl. Acids Res.* **9**, 4285–4302.

Edited by D. Rees

(Received 4 November 1997; received in revised form 26 February 1998; accepted 9 March 1998)