The EMBO Journal vol.13 no.1 pp.205-212, 1994

Crystal structure of prokaryotic ribosomal protein L9: a bi-lobed RNA-binding protein

David W.Hoffman, Christopher Davies, Sue Ellen Gerchman[1], J.H.Kycia[1], Stephanie J.Porter, Stephen W.White[2] and V.Ramakrishnan[1,2]

Department of Microbiology, Duke University Medical Center, Durham, NC 27710 and [1]Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA

[2]Corresponding authors

Communicated by K.H.Nierhaus

The crystal structure of protein L9 from the *Bacillus stearothermophilus* ribosome has been determined at 2.8 Å resolution using X-ray diffraction methods. This primary RNA-binding protein has a highly elongated and unusual structure consisting of two separated domains joined by a long exposed α-helix. Conserved, positively charged and aromatic amino acids on the surfaces of both domains probably represent the sites of specific interactions with 23S rRNA. Comparisons with other prokaryotic L9 sequences show that while the length of the connecting α-helix is invariant, the sequence within the exposed central region is not conserved. This suggests that the α-helix has an architectural role and serves to fix the relative separation and orientation of the N- and C-terminal domains within the ribosome. The N-terminal domain has structural homology to the smaller ribosomal proteins L7/L12 and L30, and the eukaryotic RNA recognition motif (RRM).

Key words: RNA-binding motif/ribosome/translation/X-ray crystallography

Introduction

Protein synthesis on the ribosome is a fundamental process in all living organisms. A detailed knowledge of the ribosome structure will be required before its function and mechanisms can be understood at the molecular level. The structure and organization of the ribosome and its subunits have previously been studied using a variety of low resolution techniques including electron microscopy (Frank *et al.*, 1991), immune electron microscopy (Stöffler-Meilicke and Stöffler, 1990), neutron diffraction (Capel *et al.*, 1987; May *et al.*, 1992), protein–protein (Walleczek *et al.*, 1989) and protein–RNA crosslinking (Brimacombe *et al.*, 1990; Osswald *et al.*, 1990) and chemical probing (Stern *et al.*, 1989). The results of such experiments have been used to construct rather detailed models of the small (30S) subunit (Brimacombe *et al.*, 1988; Stern *et al.*, 1988), and preliminary models of the large (50S) subunit (Walleczek *et al.*, 1988; Brimacombe *et al.*, 1990). Concurrently, atomic resolution structure determination methods have been used on isolated ribosomal components. The structures of ribosomal proteins L7/L12 (Leijonmarck *et al.*, 1980), L30 (Wilson *et al.*, 1986), S5 (Ramakrishnan and White, 1992) and L6 (Golden *et al.*, 1993a) have been determined by X-ray crystallography. Also, NMR analyses have been used to resolve the solution structures of ribosomal RNA tetraloops (Cheong *et al.*, 1990; Heus and Pardi, 1991; Varani *et al.*, 1991) and ribosomal protein S17 (Golden *et al.*, 1993b). We report here the crystal structure of ribosomal protein L9 from *Bacillus stearothermophilus*, the first such structure of a primary RNA-binding ribosomal protein.

Prokaryotic ribosomal protein L9 has a molecular weight of 15 700 Daltons (Wittmann, 1982) and its amino acid sequence is highly conserved in the five organisms (three bacterial and two chloroplastid) in which it has been studied (Figure 1). The analogue of the molecule has yet to be identified in eukaryotic organisms. The location and environment of L9 within the *Escherichia coli* 50S subunit has been verified by several independent methods. Immune electron microscopy has localized L9 at, or close to, the lateral protuberance adjacent to L1 and the L2, L5, L16 cluster (Stöffler-Meilicke *et al.*, 1983; Nag *et al.*, 1991). Protein–protein crosslinking studies confirm these interactions and further identify L19 and L28 as neighboring proteins (Traut *et al.*, 1986; Walleczek *et al.*, 1989). L9 does not require the presence of other proteins to bind ribosomal RNA. Early studies showed that L9 binds stoichiometrically to a 3′ 12S fragment of 23S rRNA (Branlant *et al.*, 1976; Marquart *et al.*, 1979; Roth and Nierhaus, 1980). Subsequent protein–RNA crosslinking experiments have indicated that the binding site is located in 23S rRNA helix 58 (Brimacombe *et al.*, 1990) close to nucleotides 1484–1491 (Osswald *et al.*, 1990).

Defining the role of any particular ribosomal protein in protein synthesis is extremely difficult. However, the location of a particular protein within the ribonucleoprotein complex combined with biochemical and genetic data can indicate its function in general terms. L9 has an important role in ribosome function since it is present in all known viable strains of *E.coli* (Dabbs *et al.*, 1986), and anti-L9 antibodies reduce protein synthesis by some 75% (Nag *et al.*, 1991). It is located close to the accepted components of the peptidyltransferase activity (Hackl *et al.*, 1988; Walleczek *et al.*, 1988; Brimacombe *et al.*, 1990; Romero *et al.*, 1990) but is clearly not part of it (Hampl *et al.*, 1981; Nag *et al.*, 1991). Its general role may best be demonstrated by the 50S assembly map (Herold and Nierhaus, 1987) which shows L9 and L1 to be the only proteins that interact directly with the 3′ end of 23S rRNA without cooperating with other L-proteins. L9 may therefore be an essential scaffold molecule that maintains the correct folding of this region of 23S rRNA.

The *B.stearothermophilus* L9 protein was first crystallized nearly 15 years ago in space group $P2_12_12$ with unit cell dimensions a = 135.3 Å, b = 37.5 Å, c = 48.9 Å (Appelt *et al.*, 1979). However, the structure was not solved due to difficulty in finding suitable heavy atom derivatives. Recently, however, the gene for the *B.stearothermophilus*

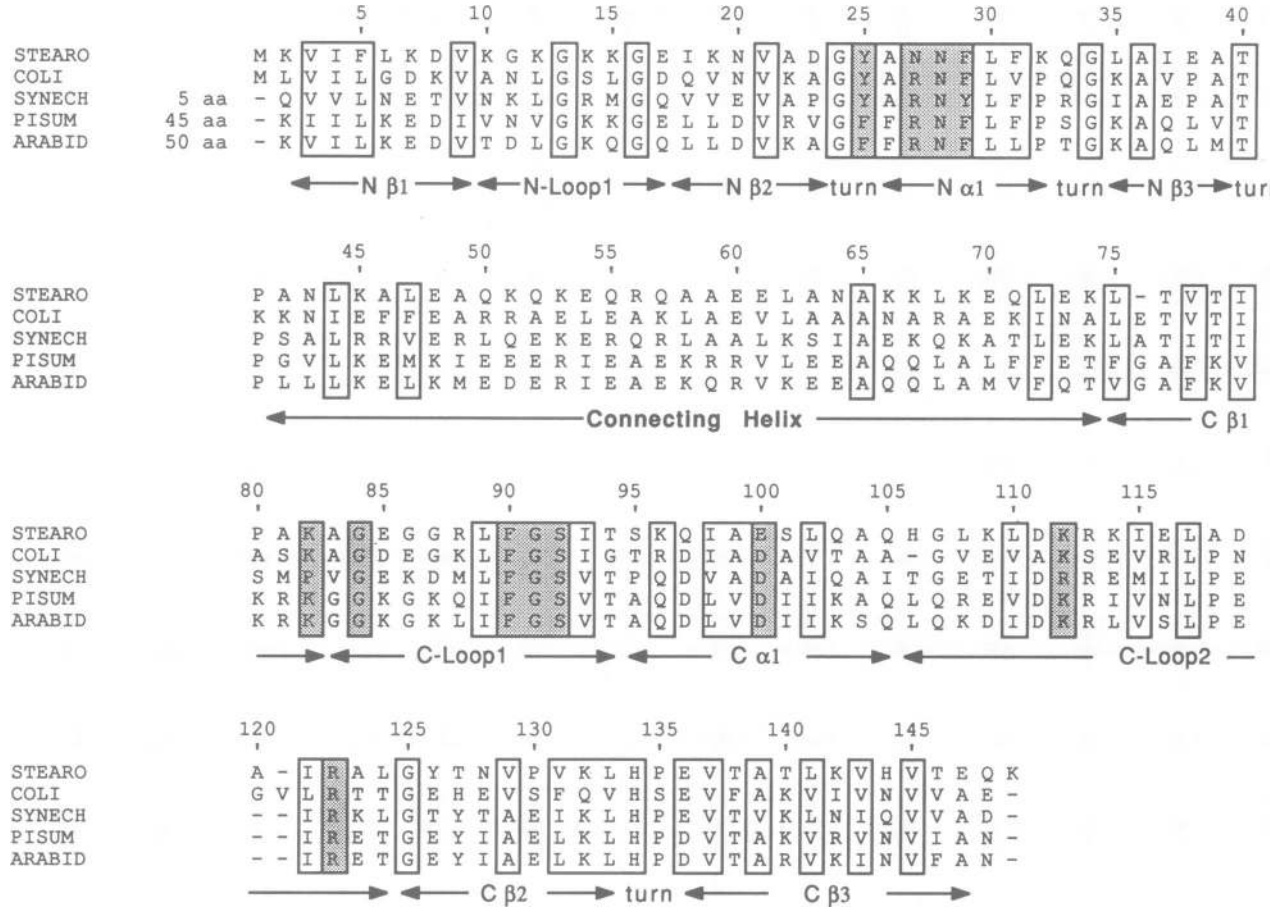

Fig. 1. Amino acid sequence alignments for ribosomal protein L9 using the one-letter code. The sequences are from the bacteria *B.stearothermophilus* (STEARO) (Ramakrishnan and Gerchman, 1991), *E.coli* (COLI) (Schnier *et al.*, 1986), and *Synechococcus elongatus* (SYNECH) (Muehlenhoff *et al.*, 1992), and the chloroplastids of *Pisum sativum* (PISUM) (Gantt, 1988) and *Arabidopsis thaliana* (ARABID) (Thompson, 1992). The secondary structure elements are prefixed with N or C to indicate their locations in the N- or C-terminal domains. Conserved structural residues in the hydrophobic core and at tight turns are boxed. The residues that are proposed to interact with 23S rRNA are boxed and shaded (see text and Figure 4 for details). The simultaneous alignment of the sequences was performed using the program TULLA (Subbiah and Harrison, 1989).

Table I. Data collection, heavy atom refinement and phasing statistics for the structure determination of ribosomal protein L9

Data collection					Phasing at 2.8 Å resolution			
Data set	No. of reflections		Resolution (Å)	*R*sym (%)	Heavy atom sites	*R*c (%)	Centric reflections	Phasing power
	Observed	Unique						
Wild-type	25 621	7789	2.6	7.2				
N27C	22 283	9935	2.9	5.0				
E100C	42 572	10 859	2.2	6.8				
Iodine								
E100C	27 854	8663	2.5	10.3	2	0.62	699	1.89
Thimerosal								
E100C	30 709	9179	2.4	9.4	1	0.73	662	1.65
PCMBS								
E100C	19 263	7645	2.6	10.2	2	0.66	663	1.70
Trimethyl								
lead acetate	17 826	7112	2.5	8.8	2	0.66	575	1.51
E100C								
L35M								
selenomet.	18 970	7474	2.8	4.0				
L124M								
selenomet.	27 370	5636	2.4	4.3				

$R_{sym} = \Sigma|I_i - I_m| / \Sigma I_m$ (I_i is the intensity of the measured reflection and I_m is the mean intensity of the i observations).
$R_c = \Sigma|(F_{PH} \pm F_P) - F_H| / \Sigma|F_{PH} - F_P|$ (F_P, F_{PH} and F_H are, respectively, the protein, derivative and heavy atom structure factors).
Phasing power = $F_H/E_{r.m.s.}$ (F_H is the heavy atom structure factor and $E_{r.m.s.}$ is the residual lack of closure).
PCMBS = *p*-chloromercuribenzenesulfonic acid.

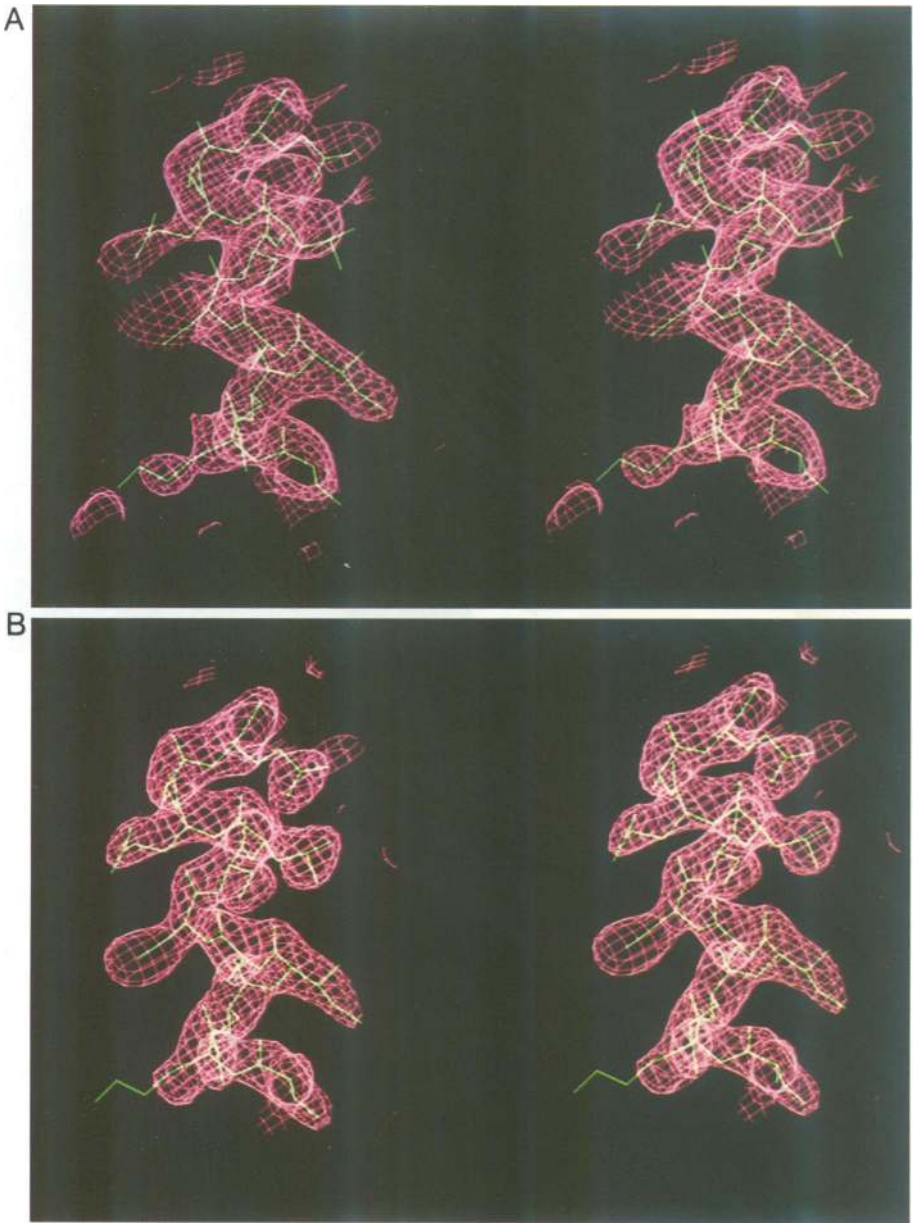

Fig. 2. Stereoview of electron density maps of ribosomal protein L9. Both maps are from the mutant E100C which produced the best diffracting crystals. (**A**) Solvent flattened MIR map at 2.8 Å resolution, showing the electron density in the region of α-helix Cα1 containing residues 95–105. (**B**) $2F_o-F_c$ map of the same region.

protein was cloned and over-expressed in *E.coli* (Ramakrishnan and Gerchman, 1991) which greatly facilitated the crystal structure analysis. The increased availability of the protein helped in the search for heavy atom derivatives, and the cloning permitted the introduction of mutations that expedited this search.

Results

Crystal structure determination

An initial extensive search for heavy atom derivatives using the classic method of soaking soluble heavy atom compounds into crystals was unsuccessful. We therefore prepared two types of site-directed mutants to facilitate the process: cysteine mutants for binding mercury compounds and methionine mutants to act as a vehicle for incorporating selenium atoms in the form of selenomethionine. Eight cysteine mutants were prepared (N20C, N27C, Q33C, T40C, K45C, T94C, E100C and Q105C), but only E100C formed high quality crystals that were isomorphous to the native crystals. In fact, the E100C crystals are superior in quality to the native crystals and diffract to a higher resolution. N27C also formed excellent crystals but in space group $P2_12_12_1$. Two methionine mutants were engineered (L35M and L124M) and both crystallized isomorphously to the wild-type protein.

A mercury derivative of E100C was prepared by soaking the crystals in 2 mM thimerosal for 3 weeks. Crystals soaked in lower concentrations or for shorter times did not derivatize. Concurrently, an additional derivative of E100C was prepared by treating the crystals with I_3^- which results in a substitution of the H3 and H5 protons of a tyrosine

residue with iodine atoms (Covelli and Wolf, 1966; Sigler, 1970). Difference Patterson maps showed a single site for the thimerosal derivative and two sites separated by 6 Å (i.e. one derivatized tyrosine) for the iodine derivative. Phases were calculated from each derivative, and cross-difference Fouriers confirmed the heavy atom locations and the absence of minor sites. Phases based on both derivatives were of sufficient quality to generate a partially interpretable electron density map. These phases were used to calculate difference Fouriers on other derivative data sets that could not be interpreted by difference Pattersons. A two-site trimethyl lead acetate derivative and a two-site *p*-chloromercuribenzenesulfonic acid (PCMBS) derivative were found in this manner. After solvent flattening, the four derivative multiple isomorphous replacement (MIR) map calculated at 2.8 Å was of sufficient quality to build the majority of the structure (Table I and Figure 2A).

The interpretation of the map was straightforward in what turned out to be the C-terminal two-thirds of the molecule. The backbone could be unambiguously traced and most of the side chains were clearly visible in the electron density. A refinement incorporating residues 45–149 using data from 8.0 to 2.8 Å with $F/\text{sig}(F) > 1.0$ yielded an *R*-factor of 25.8%. The quality of the electron density map decreased markedly towards the N-terminal third of the molecule. By inspecting F_o-F_c, $2F_o-F_c$, and phase combined maps, it was clear that the N-terminal domain comprised a three-stranded β-sheet, a short α-helix and the beginning of a long α-helix extending towards the C-terminal domain. However, clear side chain density was often not apparent and the connectivity was unclear.

At this point it was decided to solve the structure of the L9 mutant N27C. Unlike wild-type L9 and the other mutants, N27C crystallized in the space group $P2_12_12_1$ with unit cell dimensions 113.7 Å × 37.5 Å × 101.7 Å. However, it was obvious from these cell dimensions and the low resolution features of the Okl zone that the crystal packing is very similar to that of the normal $P2_12_12$ crystals. The only difference is that two tightly packed monomers form a crystallographic dimer around the *c* axis in the $P2_12_12$ space group whereas the same dimer forms two-fold screw related asymmetric units in $P2_12_12_1$, each with local two-fold symmetry. It was therefore rather trivial to solve the N27C structure by molecular replacement using an L9 dimer as the probe at 4 Å resolution. The N27C structure was subsequently refined to 2.9 Å resolution. This afforded two independent views of the N-terminal domain that could be non-symmetrically averaged. To avoid bias, the $2F_o-F_c$ electron density map of N27C used in the averaging only included information from residues 41 to 149. Also, since the two N-terminal domains in the dimer might be oriented differently with respect to the rigid C-terminal domains, only the N-terminal domain density was averaged. This process resulted in a clearer electron density map for the N-terminal domain in which the connectivity and many regions of the sequence could be identified. After exhaustive rounds of refinement and model building only a few ambiguities remained in the N-terminal domain. Finally, the cysteine residues in the models of the mutant proteins were replaced by the wild-type residues, and the resulting model was refined against data from the wild-type crystals at 2.8 Å resolution.

Information from the L9 mutants and their derivatives was invaluable in confirming the correctness of the model. First, the positions of leucines 35 and 124 agree with difference Fourier peaks from the selenomethionine-substituted mutant proteins L35M and L124M. Second, the major difference Fourier peak between wild-type L9 and mutant E100C corresponds to the sulfur atom in the cysteine residue. Third, the position of the ring of tyrosine 126 corresponds to the iodine sites located in a difference Fourier of the iodine derivative. Finally, the four L9 molecules that we had effectively determined at the end of the whole procedure all refined satisfactorily. The statistics and geometries were as follows: wild-type (2.8 Å resolution, $R = 20.9\%$, r.m.s.d. bonds = 0.016 Å, r.m.s.d. angles = 3.3°); E100C (2.8 Å resolution, $R = 21.0\%$, r.m.s.d. bonds = 0.016 Å, r.m.s.d. angles = 3.5°); the non-symmetric dimer of mutant N27C (2.9 Å resolution, $R = 19.2\%$, r.m.s.d. bonds = 0.016 Å, r.m.s.d. angles = 3.4°). A section of the final $2F_o-F_c$ electron density map of E100C is shown in Figure 2b.

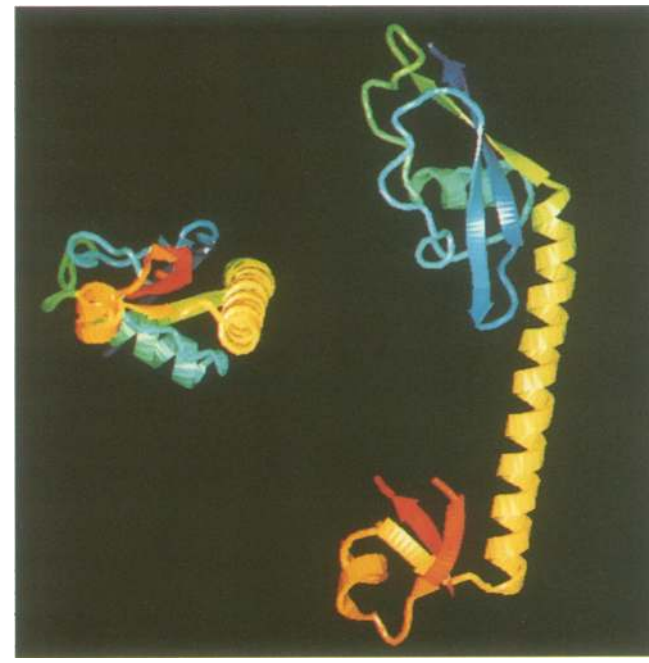

Fig. 3. Ribbon diagrams of ribosomal protein L9 colour-ramped red to blue, N-terminus to C-terminus. The molecule contains two discrete α/β domains joined by a long, nine-turn α-helix. Two orthogonal views are shown to emphasize that the molecule has an extended, somewhat flattened structure. In view (**A**) (right), the smaller N-terminal domain is at the bottom and the larger C-terminal domain is at the top. Residues 11–17 form a disordered loop connecting strands Nβ1 and Nβ2 and are not included in the model. In view (**B**) (left), the molecule is turned by 90° and viewed down the N-terminal domain. Both ribbon diagrams were generated using the program 'O' (Jones *et al.*, 1991).

Molecular structure of L9

The overall fold of the L9 polypeptide chain is shown in Figure 3. The protein has a very unusual structure in which two distinct α/β domains are connected by a nine-turn α-helix. The molecule is highly elongated in one dimension and relatively narrow in a second which results in a flat extended structure with approximate dimensions 25 Å × 30 Å × 82 Å. The nomenclature of the secondary structure elements, and their locations within the primary structure are shown in Figure 1.

The N-terminal domain, corresponding to residues 1–40,

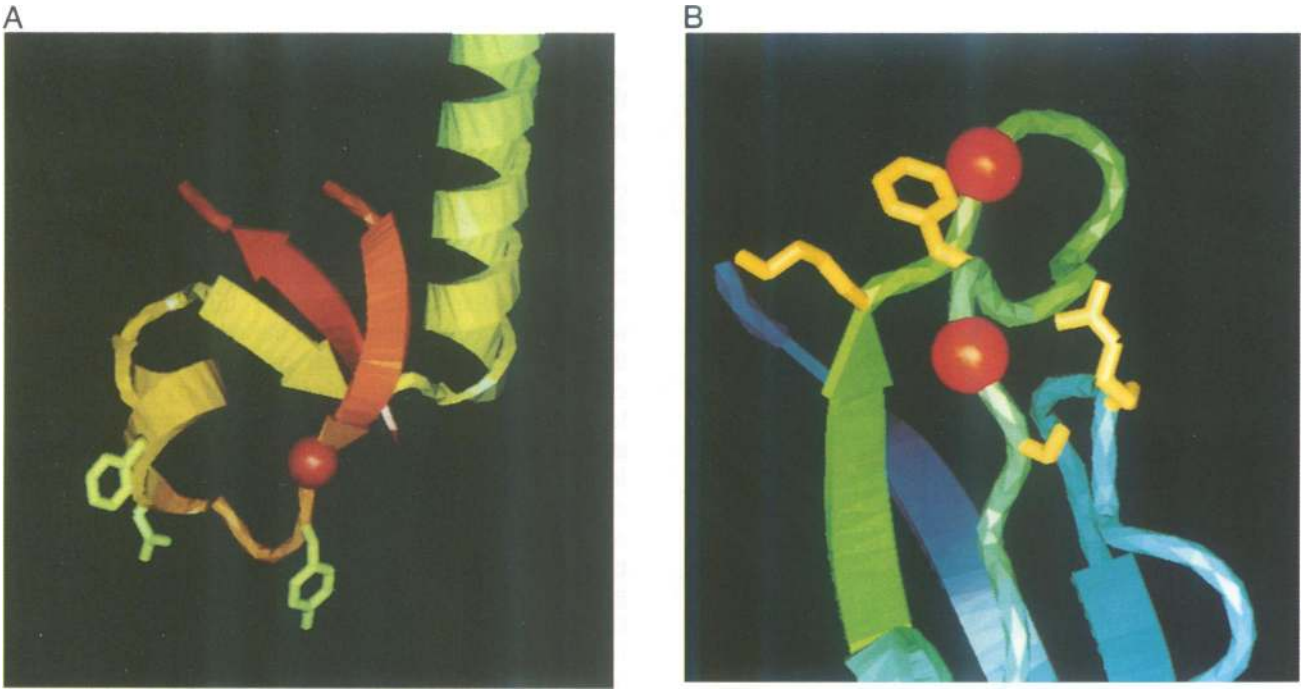

Fig. 4. Expanded views of the putative RNA-binding sites in ribosomal protein L9. In both views the colouring of the main chain is the same as that in Figure 3. (**A**) The site located in the N-terminal domain. This is centered on α-helix Nα1 and contains (right to left in yellow) tyrosine 25, asparagine 28 and phenylalanine 29 (see Figure 1 and text for further details). The red sphere is the conserved glycine 24 in the turn between Nβ2 and Nα1. The domain is in the same orientation as that shown in Figure 3. Residues 10–17 (N-Loop1) which connect the central β-strand Nβ1 to the outside β-strand Nβ2 are missing and apparently disordered. It should be emphasized that (A) is somewhat tentative since the electron densities for these side chains are poor. Nevertheless, it does serve to demonstrate the approximate location and orientation of these conserved amino acids relative to the rest of the molecule. (**B**) The site located in the C-terminal domain centered on C-Loop1. Conserved glycines 84 and 91 are red spheres, and conserved residues (left to right) lysine 82, phenylalanine 90, serine 92 and arginine 122 are shown in yellow (see Figure 1 and text for details). The domain is oriented so that the view is into the left face of the elongated molecule shown in Figure 3. Unlike the N-terminal site, the side chain positions are well-defined by the electron density maps. Both diagrams were generated using the program 'O' (Jones *et al.*, 1991).

contains a three-stranded antiparallel β-sheet (Nβ1-Nβ2-Nβ3) which is sandwiched between a short two-turn α-helix (Nα1) and the N-terminal end of the long Connecting Helix. This small domain contains a well defined hydrophobic core (Figure 1) and three tight turns. The latter are all mediated by prolines and/or glycines (Figure 1) and occur at the junction of Nβ3 and the Connecting Helix, and at both termini of Nα1. The connection between Nβ1 and Nβ2 is via an extended loop (N-loop1) comprising residues 10–17. These residues are exclusively either glycines or charged amino acids (Figure 1) and, not surprisingly, N-loop1 is not visible in the electron density map and is apparently disordered. Residues 24–31, which include the small α-helix Nα1, are well conserved in all of the known L9 sequences (Figure 1). Unfortunately, this is one region of the domain where the side chain electron density is not clear. However, tyrosine 25, asparagines 27 and 28, and phenylalanine 29 appear to form a conserved surface feature.

The C-terminal domain, corresponding to residues 75–149, contains three β-strands (Cβ1-Cβ2-Cβ3), a three-turn α-helix (Cα1) and two extended loops (C-loop1 and C-loop2). Somewhat atypically, the three β-strands create a mixed β-sheet in which Cβ1 and Cβ3 are parallel and Cβ2 is antiparallel. A conserved proline (135) mediates the tight turn between Cβ2 and Cβ3. Apart from the N- and C-termini, there are only two insertions/deletions within the five known L9 sequences, and both occur in the extended C-loop2 of the C-terminal domain (Figure 1). One is located at the beginning of the loop, and the other is within a small surface projection. The domain has a well-defined hydrophobic core that contains ~20 highly conserved residues (Figure 1). In addition, there are other conserved residues that are located on the domain surface. Most notable are phenylalanine 90, glycine 91 and serine 92 which form an exposed loop adjacent to lysine 82, glycine 84 and arginine 122 (Figure 4). These residues are conserved in the five currently known L9 sequences (Figure 1).

The long α-helix that connects the two domains (Connecting Helix) begins after a sharp turn at proline 41 and continues to residue 73. The ends of the α-helix interact with the two domains via several conserved hydrophobic residues and contribute to their hydrophobic cores. Leucines 44 and 47 are integral to the core of the N-terminal domain, and leucines 62, 68, 72 and 75, and alanine 65 augment that of the C-terminal domain. Thus, the α-helix ends can be considered as integral components of the two domains. In contrast, residues 48–60, which correspond to the central exposed region of the α-helix, are not conserved although they are generally hydrophilic. Finally, there appear to be many potential salt bridges between acidic and basic residues along the entire length of the exposed regions of the α-helix.

Discussion

Previous structural investigations by us and others on small proteins that interact with RNA (Nagai *et al.*, 1990; Hoffman *et al.*, 1991; Ramakrishnan and White, 1992; Golden *et al.*, 1993a,b; A.Liljas, personal communication) have indicated that RNA-binding sites tend to be populated with aromatic residues, positively charged residues and glycines. It is

Fig. 5. A superposition of four independently determined crystal structures of ribosomal protein L9. The structures diverge near lysine 53 in the Connecting Helix. The wild-type and mutant E100C structures are nearly identical, and the outlying structures correspond to the two independent molecules in the N27C dimer. The superposition was performed by XPLOR (Brünger, 1988) using only the C-terminal domains in order to show the flexibility of the Connecting Helix. The orientations of the N-terminal domains are fixed relative to the Connecting Helix and have not been included for reasons of clarity.

proposed that the aromatic residues interact with the nucleotide bases, the positively charged residues with the sugar–phosphate backbone, and that the glycines permit the close approach of the nucleic acid. L9 contains two clusters of highly conserved surface residues, one in each domain, that incorporate many of these features. The putative N-terminal RNA-binding site is centered on the small α-helix Nα1 and is shown in Figure 4A. Although somewhat speculative due to poor electron density, the figure demonstrates the exposed location of these conserved and obviously important residues. The proposed C-terminal site is shown in Figure 4B. This is clearly defined and centered on C-loop1. It is striking that both sites are located on the same side of the molecule, opposite the Connecting Helix. One can therefore visualize L9 as a two-pronged RNA-binding molecule that may fix the locations and orientations of two specific regions of 23S rRNA in the ribosome.

A valid question regarding isolated ribosomal proteins is whether they adopt the same structure *in situ* as they do in solution and, indeed, in a crystal environment. We believe that the crystal structure of ribosomal protein L9, like proteins L7/L12 (Leijonmarck *et al.*, 1980), L30 (Wilson *et al.*, 1986), S5 (Ramakrishnan and White, 1992) and L6 (Golden *et al.*, 1993a), is the correct *in situ* structure. All of these proteins incorporate classic structural elements with, for example, conserved hydrophobic cores, glycines and prolines at tight turns and potential conserved functional sites on the molecular surfaces. However, differences may occur in exposed loop regions and in the relative orientations of the subunits, both of which might be affected by the crystalline environment. This is particularly relevant to the long Connecting Helix of L9 which is reminiscent of a similar α-helix in calmodulin that is known to undergo a large conformational change upon binding substrate (Ikura *et al.*, 1992). For a variety of reasons, we believe that the α-helix does exist *in situ*, and that its function is that of a rigid strut that separates and correctly orients the two globular domains. First, in the five known sequences of L9 (Figure 1), the number of residues in the α-helical region is absolutely conserved. This not only preserves the α-helix length but also the relative azimuthal orientations of the distal globular domains. Second, although the five sequences of the α-helical region are not particularly well conserved, they all strongly favour an α-helical secondary structure (data not shown). Third, all five sequences have the potential for many salt bridges along the exposed surfaces of the α-helix, and these are known to stabilize an α-helical conformation (Marqusee and Baldwin, 1987). Finally, two *in situ* observations on L9 suggest that it maintains its highly elongated shape in the ribosome. Immune electron microscopy showed that two exposed epitopes of L9 are separated by ~35 Å (Stöffler-Meilicke *et al.*, 1983), and neutron scattering measurements estimate its axial ratio to be 4:1 (K.H.Nierhaus, personal communication).

The peculiar shape of the protein results in an unusual packing arrangement within the L9 crystals, both in the wild-type $P2_12_12$ type and in the $P2_12_12_1$ form found in N27C. The larger C-terminal domains are tightly packed into two-dimensional sheets and these are rather loosely connected on either side by 'crossbridges' made up of the long Connecting Helix and the distal N-terminal domain. By superimposing the four refined structures of L9 that we have determined (wild-type, E100C and the N27C dimer), it is clear that the Connecting Helix has inherent flexibility centered around lysine 53 (Figure 5). This is the only significant difference between the four structures. We believe that the poor electron density towards the N-terminal third of all the crystal structures results from inherent disorder within the crystal which is caused by a combination of the α-helix flexibility and the crystal packing. This is supported by native data collected at reduced temperature (4°C) which, when used in refinement, produced no improvement in the electron density. It is therefore unlikely that the structure can be significantly improved with the present crystals. Given the apparent stability of the N-terminal domain, we plan to clone and express the corresponding part of the gene and to investigate the possibility of studying its structure in isolation.

Of the six ribosomal protein structures that have now been determined, four have homologous structures centered on

a basic fold consisting of a three-stranded anti-parallel β-sheet and one α-helix. The connection between β-strands 1 and 2 of the β-sheet contains various inserts, but all have identical topologies. L7/L12 (Leijonmarck *et al.*, 1980) has two α-helices, L30 (Wilson *et al.*, 1986) has one α-helix, L6 (Golden *et al.*, 1993a) has four β-strands and S6 (A.Liljas, personal communication) has one β-strand and one α-helix. Interestingly, the eukaryotic RNA recognition motif (RRM) also conforms to this pattern and has a structure very similar to that of S6 (Nagai *et al.*, 1990; Hoffman *et al.*, 1991; A.Liljas, personal communication). The N-terminal domain of L9 also has this motif and is the smallest thus far with a nine residue disordered insert. When the homology between L7/L12 and L30 first appeared, it was suggested that the motif might eventually be found in other ribosomal proteins (Leijonmarck *et al.*, 1988). This now appears to be true, and their common ancestor might have been a primitive RNA-binding protein.

Materials and methods

Cells and strains

Wild-type and mutant *B.stearothermophilus* L9 proteins were produced in the *E.coli* strain BL21(DE3) containing the L9-pET13 plasmid (V.Ramakrishnan and S.E.Gerchman, unpublished results) which provides kanamycin resistance as a selectable marker. Cells were grown in 2 l of Luria broth medium containing 30 mg/l kanamycin, and after reaching an optical density of 0.8 at 550 nm, they were induced with 100 mg/l of isopropyl-β-D-thiogalactopyranoside (IPTG). Three hours after induction, the cells were harvested by centrifugation for 10 min at 6000 *g*.

For labeling with selenomethionine, the plasmid was transformed into the strain B834(DE3) (Wood, 1966; Studier and Moffatt, 1986) which is auxotrophic for methionine. These cells were grown in defined medium containing the following components (per liter): 12 g Na_2HPO_4, 6 g KH_2PO_4, 2 g NH_4Cl, 1 g NaCl, 4 g dextrose, 40 mg kanamycin, 1 mg biotin, 1 mg choline chloride, 1 mg folic acid, 1 mg niacinamide, 1 mg D-pantothenate, 1 mg pyridoxal, 0.1 mg riboflavin, 5 mg thiamine, 2 mM $MgSO_4$, 1 μM $FeCl_3$, 0.2 mM $CaCl_3$, 40 mg/l of each amino acid apart from methionine and 40 mg/l selenomethionine. Since selenomethionine is deleterious to cell growth, it was necessary to adapt the cells to this medium. The cells were first inoculated into 2 ml of medium containing 100% methionine, and this culture was then sequentially diluted into 2 ml cultures containing 50, 80 and 100% selenomethionine. Finally, the last culture was diluted into a second 2 ml culture containing 100% selenomethionine and this was used to inoculate a 2 l culture which was grown and induced as described above.

Preparation of L9 mutant proteins

The L9 cysteine mutants were constructed by a site-directed mutagenesis procedure that involves three steps of polymerase chain reaction (PCR) (Nelson and Long, 1989). In the first step, two oligonucleotides were used to make a 'superoligo' that contains the required mutation within the gene and a unique DNA sequence (duso) in the 3′ flanking region. The mixture contained 20 ng of template DNA, 50 pmol of each oligonucleotide, 0.2 mM dNTPs and 2.5 units of *Taq* DNA polymerase. The template DNA was the pET-13 plasmid that contains the wild-type L9 gene. Twenty-five cycles of PCR were performed (94°C for 1 min, 45°C for 1 min and 72°C for 2 min) and this was followed by a single round at 45°C for 1 min, and 72°C for 7 min. The second step, which fills in the remaining gene sequence, involved a single cycle of PCR (95°C for 5 min, 45°C for 2 min, and 72°C for 15 min) using a mixture containing 100 ng of superoligo, 20 ng of template DNA, 0.2 mM dNTPs and 2.5 U of *Taq* DNA polymerase. The conditions for the final PCR step, which specifically amplifies the mutant gene via the duso sequence, were identical to the first.

Protein purification

Cells were resuspended in 30 ml of 20 mM Tris−Cl, pH 8.0, 100 mM NaCl, and lysed by the addition of 5 mg hen egg white lysozyme followed by one cycle of freeze and thaw. The viscosity of the mixture was reduced by adding 1 mg DNase and 10 mM $MgCl_2$ to cut the cell DNA, and the cell debris was removed by centrifugation for 30 min at 12 000 *g*. The resulting supernatant was loaded directly onto a 50 ml S-Sepharose column (Pharmacia) equilibrated in 20 mM Tris−Cl, pH 7.5, washed extensively with this buffer and eluted with a 0.0−1.0 M NaCl gradient. The fractions containing L9 were identified by SDS−PAGE and pooled. A typical yield was 30 mg/l of cell culture. The purified L9 was prepared for crystallization by dialysis against 1 M ammonium sulfate, 20 mM Tris−Cl, pH 7.8 and concentrating to 10 mg/ml (A_{280} = 3.0) using Centricon microconcentrators (Amicon).

Crystallization

Crystals were grown by the hanging drop vapor diffusion method (reviewed in Ollis and White, 1990) using 10 μl drops and 1 ml well solutions. Wild-type crystals with typical dimensions of 0.3 mm × 0.5 mm × 1.0 mm were grown in 2.4 M ammonium sulfate, 20 mM Tris−Cl, pH 7.8 at 22°C. Comparable crystals were grown under identical conditions for the L9 mutants L35M and L124M, and their selenomethionine-substituted counterparts. Crystals of L9 mutant E100C were grown at the same pH and ammonium sulfate concentration, but with the addition of 1 mM β-mercaptoethanol.

Crystallography

Wild-type, mutant and derivative crystals were initially characterized on an Enraf−Nonius precession camera mounted on a Rigaku RU300 rotating anode X-ray generator operating at 40 kV and 80 mA. The precession camera had a crystal-to-film distance of 100 mm and a 0.3 mm collimator, and the CuK_α radiation (1.54 Å) was nickel filtered.

All data were collected and processed on a Rigaku RAXIS-II image plate system (Molecular Structure Corporation) mounted on the X-ray generator described above. The crystal-to-detector distance was set to 130 mm and the collimator size was 0.3 mm. Data were collected by the standard oscillation method using 2° ranges around the long *a* crystal axis to avoid problems of overlap. Typically, each 2° range was collected in 30 min, and 90° of data were collected in 24 h. The L9 crystals were quite robust, and the quality of the diffraction did not decay significantly during data collection. When collecting derivative data, the crystal a^* axis was accurately aligned parallel to the rotation axis so as to record anomalous pairs simultaneously.

Structure solution and refinement

The protein structure was solved by the multiple isomorphous replacement (MIR) method. Iodine and mercury (thimerosal) positions were independently located from difference Patterson maps and used for initial phasing. These phases were subsequently used to locate the sites of a second mercury derivative (PCMBS) and a lead derivative on difference Fourier maps. The MIR phases based on all four derivatives yielded an electron density map with many interpretable features, and the quality of the phases was further improved by solvent flattening. The map allowed two-thirds of the molecule to be built unambiguously and this was refined by simulated annealing. The final third of the molecule was built in several iterative steps involving a combination of the original solvent flattened map, F_o-F_c maps, $2F_o-F_c$ maps and phase combined maps, and simulated annealing. In addition, the mutant N27C that crystallized in space group $P2_12_12_1$, was essential in resolving this part of the molecule. The dimeric asymmetric unit in this crystal form was solved by molecular replacement, and the poor electron density was improved and eventually interpreted by several rounds of non-crystallographic symmetry averaging, model building and simulated annealing. Difference Fourier maps of the two selenomethionine-substituted mutant proteins L35M and L124M clearly showed the selenium positions and were invaluable in the chain tracing process. Crystallographic calculations were performed using the PHASES package (Furey and Swaminathan, 1990) operating on Silicon Graphics 4DTG25 and 4DTG30 Personal Irises. Model building was carried out on Evans and Sutherland ESV workstations using the programs O (Jones *et al.*, 1991) and FRODO (Jones, 1985). All refinements and molecular replacements were performed by the program XPLOR (Brünger, 1988) running on a Silicon Graphics 4D440S machine.

Acknowledgements

We thank Barbara Golden for many useful discussions and for help in using the 'O' program. We also thank Anders Liljas for providing unpublished data, Bill Furey for patiently explaining his non-crystallographic symmetry averaging programs and F.William Studier for providing the B834(DE3) cells for incorporating selenomethionine. This work was supported by NIH grant GM 44973 (to S.W.W. and V.R.) and by a grant from the Office of Health and Environmental Research of the US Department of Energy (to V.R.).

References

Appelt,K., Dijk,J. and Epp,O. (1979) *FEBS Lett.*, **103**, 66–70.

Branlant,C., Krol,A., Sriwidada,J. and Brimacombe,R. (1976) *Eur. J. Biochem.*, **70**, 483–492.

Brimacombe,R., Atmadja,J., Stiege,W. and Schüler,D. (1988) *J. Mol. Biol.*, **199**, 115–136.

Brimacombe,R., Gornicki,P., Greuer,B., Mitchell,P., Osswald,M., Rinke-Appel,J., Schüler,D. and Stade,K. (1990) *Biochim. Biophys. Acta.*, **1050**, 8–13.

Brünger,A.T. (1988) *J. Mol. Biol.*, **203**, 803–816.

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., Sillers,I.-Y., Yabuki,S. and Moore,P.B. (1987) *Science*, **238**, 1403–1406.

Cheong,C., Varani,G. and Tinoco,I. (1990) *Nature*, **346**, 680–682.

Covelli,I. and Wolf,J. (1966) *Biochemistry*, **5**, 860–866.

Dabbs,E.R. (1986) In Hardesty,B. and Kramer,G. (eds), *Structure, Function and Genetics of Ribosomes*. Springer-Verlag, New York, pp. 733–748.

Frank,J., Penczek,P., Grassucci,R. and Srivastava,S. (1991) *J. Cell Biol.*, **115**, 597–605.

Furey,W. and Swaminathan,S. (1990) PA33, *Am. Crystallogr. Assoc. Abstr., ser. 2*, **18**, 73.

Gantt,J.S. (1988) *Curr. Genet.*, **14**, 519–528.

Golden,B.L., Ramakrishnan,V. and White,S.W. (1993a) *EMBO J.*, **12**, 4901–4908.

Golden,B.L., Hoffman,D.W., Ramakrishnan,V. and White,S.W. (1993b) *Biochemistry*, in press.

Hackl,W., Stöffler-Meilicke,M. and Stöffler,G. (1988) *FEBS Lett.*, **233**, 119–123.

Hampl,H., Schulze,H. and Nierhaus,K.H. (1981) *J. Biol. Chem.*, **256**, 2284–2288.

Herold,M. and Nierhaus,K.H. (1987) *J. Biol. Chem.*, **262**, 8826–8833.

Heus,H.A. and Pardi,A. (1991) *Science*, **253**, 191–193.

Hoffman,D.W., Query,C.C., Golden,B.L., White,S.W. and Keene,J.K. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 2495–2499.

Ikura,M., Clore,G.M., Gronenborn,A.M., Zhu,G., Klee,C.B. and Bax,A. (1992) *Science*, **256**, 632–638.

Jones,T.A. (1985) *Methods Enzymol.*, **115**, 157–171.

Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta Crystallogr.*, **A47**, 110–119.

Leijonmarck,M., Eriksson,S. and Liljas,A. (1980) *Nature*, **286**, 824.

Leijonmarck,M., Appelt,K., Badger,J., Liljas,A., Wilson,K.S. and White,S.W. (1988) *Proteins*, **3**, 243–251.

Marquardt,O., Roth,H.E. and Nierhaus,K.H. (1979) *Nucleic Acids Res.*, **6**, 3641–3650.

Marqusee,S. and Baldwin,R.L. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 8898–8902.

May,R.P., Nowotny,V., Nowotny,P., Voss,H. and Nierhaus,K.H. (1992) *EMBO J.*, **11**, 373–378.

Muehlenhoff,U., Haehnel,W., Witt,H.T. and Hermann,R.G. (1992) *EMBL/GENBANK/DDBJ data bank.*

Nag,B., Akella,S.S., Cann,P.A., Tewari,D.S., Glitz,D.G. and Traut,R.R. (1991) *J. Biol. Chem.*, **266**, 22129–22135.

Nagai,K., Oubridge,C., Jessen,T.H., Li,J. and Evans,P.R. (1990) *Nature*, **348**, 515–520.

Nelson,R.M. and Long,G.L. (1989) *Anal. Biochem.*, **180**, 147–151.

Ollis,D. and White,S.W. (1990) *Methods Enzymol.*, **182**, 646–659.

Osswald,M., Greuer,B. and Brimacombe,R. (1990) *Nucleic Acids Res.*, **18**, 6755–6760.

Ramakrishnan,V. and Gerchman,S.E. (1991) *J. Biol. Chem.*, **266**, 880–885.

Ramakrishnan,V. and White,S.W. (1992) *Nature*, **358**, 768–771.

Romero,D.P., Arredondo,J.A. and Traut,R.R. (1990) *J. Biol. Chem.*, **265**, 18185–18191.

Roth,H.E. and Nierhaus,K.H. (1980) *Eur. J. Biochem.*, **103**, 95–98.

Schnier,J., Kitakawa,M. and Isono,K. (1986) *Mol. Gen. Genet.*, **204**, 126–132.

Sigler,P.B. (1970) *Biochemistry*, **9**, 3609–3618.

Stern,S., Weiser,B. and Noller,H.F. (1988) *J. Mol. Biol.*, **204**, 447–481.

Stern,S., Powers,T., Changchien,L.-M. and Noller,H.F. (1989) *Science*, **244**, 783–790.

Stöffler-Meilicke,M. and Stöffler,G. (1990) In Hill,W.E., Dahlberg,A.D., Garrett,R.A., Moore,P.B., Schlessinger,D. and Warner,J.R. (eds), *The Ribosome: Structure Function and Evolution*. American Society for Microbiology, Washington, DC, pp. 123–133.

Stöffler-Meilicke,M., Noah,M. and Stöffler,G. (1983) *Proc. Natl Acad. Sci. USA*, **80**, 6780–6784.

Studier,F.W. and Moffatt,B.A. (1986) *J. Mol. Biol.*, **189**, 113–130.

Subbiah,S. and Harrison,S.C. (1989) *J. Mol. Biol.*, **209**, 539–548.

Thompson,M.D., Jacks,C.M., Lenvik,T.R. and Gantt,J.S. (1992) *Plant Mol. Biol.*, **18**, 931–944.

Traut,R.R., Tewari,D.S., Sommer,A., Gavino,G.R., Olson,H.M. and Glitz,D. (1986) In Hardesty,B. and Kramer,G. (eds), *Structure, Function and Genetics of Ribosomes*. Springer Verlag, New York, pp. 286–308.

Varani,G., Cheong,C. and Tinoco,I. (1991) *Biochemistry*, **30**, 3280–3289.

Walleczek,J., Schüler,D., Stöffler-Meilicke,M., Brimacombe,R. and Stöffler,G. (1988) *EMBO J.*, **7**, 3571–3576.

Walleczek,J., Redl,B., Stöffler-Meilicke,M. and Stöffler,G. (1989) *J. Biol. Chem.*, **264**, 4231–4237.

Wilson,K.S., Appelt,K., Badger,J., Tanaka,I. and White,S.W. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 7251–7255.

Wittmann,H.G. (1982) *Annu. Rev. Biochem.*, **51**, 155–182.

Wood,W.B. (1966) *J. Mol. Biol.*, **16**, 118–133.

Received on August 31, 1993