The NMR structure of Escherichia coli ribosomal protein L25 shows homology to general stress proteins and glutaminyl-tRNA synthetases

Matthias Stoldt, Jens Wöhnert, Matthias Görlich and Larry R. Brown

The structure of the Escherichia coli ribosomal protein L25 has been determined to an r.m.s. displacement of backbone heavy atoms of 0.62 ± 0.14 Å by multidimensional heteronuclear NMR spectroscopy on protein samples uniformly labeled with 15N or 15N/13C. L25 shows a new topology for RNA-binding proteins consisting of a six-stranded β-barrel and two α-helices. A putative RNA-binding surface for L25 has been obtained by comparison of backbone 15N chemical shifts for L25 with and without a bound cognate RNA containing the eubacterial E-loop that is the site for binding of L25 to 5S ribosomal RNA. Sequence comparisons with related proteins, including the general stress protein, CTC, show that the residues involved in RNA binding are highly conserved, thereby providing further confirmation of the binding surface. Tertiary structure comparisons indicate that the six-stranded β-barrels of L25 and of the tRNA anticodon-binding domain of glutaminyl-tRNA synthetase are similar.

Keywords: NMR spectroscopy/protein structure/ribosomal protein/ribosome/RNA binding

Introduction

Recently, there has been renewed interest in the structure of ribosomal proteins for several reasons. First, protein biosynthesis is a central step in the realization of genetic information. Given the complexity of the ~2.3 MDa ribosome, recent progress in understanding the structural basis of protein biosynthesis has largely been based on a combination of low resolution methods, e.g. cryo-electron microscopy with three-dimensional image reconstruction together with footprinting, crosslinking experiments and three-dimensional modeling (e.g. Brimacombe, 1995; Frank, 1997; Mueller and Brimacombe, 1997a,h), with high resolution structures obtained by nuclear magnetic resonance spectroscopy (NMR) or X-ray crystallography for ribosomal proteins (for recent reviews see Liljas and Al-Karadaghi, 1997; Ramakrishnan and White, 1998) or fragments of ribosomal RNAs (White et al., 1992; Szewczak et al., 1993; Wimberly et al., 1993; Betzel et al., 1994; Szewczak and Moore, 1995; Fourmy et al., 1996; Correll et al., 1997; Dallas and Moore, 1997; Puglisi et al., 1997; Perbandt et al., 1998). The 5S rRNA constitutes one of the most highly conserved RNA sequences in nature (Specht et al., 1990). It is an essential component of the large ribosomal subunit, and in Escherichia coli it is complexed with the ribosomal proteins L5, L18 and L25 (Chen-Schmeisser and Garrett, 1977). Reconstituted 50S ribosomal subunits lacking the 5S RNA are inactive in protein biosynthesis (Erdmann et al., 1971; Hartmann et al., 1988). The distal end of the E-domain (also referred to as domain IV) of the 5S RNA, which is directly adjacent to the E-loop binding site for the L25 protein (Douthwaite et al., 1979; Garrett and Noller, 1979; Huber and Wool, 1984; Ciesiolka et al., 1992; Shpanchenko et al., 1996), has been shown to reside in proximity to the peptidyl transferase ring and the GTPase-associated area of the 23S rRNA leading to the suggestion that this domain might play an as yet unidentified role in the process of peptide bond formation (Dontsova et al., 1994; Sergiev et al., 1998).

Secondly, it has become increasingly apparent in recent years that in addition to simple A-form helices, RNA molecules often contain other characteristic structural elements, e.g. stable tetraloops (Shen et al., 1995) or A platforms (Cate et al., 1996), that may play important structural and functional roles. The D and E helices of the D/E domain of 5S RNA bound a large internal bulge or E-loop which has a tightly packed helical structure of unusual geometry due to the formation of an array of non-canonical base pairs. The (asymmetric) eukaryote E-loop bulge (Wimberly et al., 1993) has a looped out G-residue which contacts the helix in the major groove in a similar fashion to the α-sarcin/ricin loop of the 23S RNA (Szewczak et al., 1993; Szewczak and Moore, 1995) and constitutes a structure which probably is also similar to domains of potato spindle tuber virus (PSTV) RNA and the hairpin ribozyme (Shen et al., 1995). In contrast, the (symmetric) eubacterial E-loop bulge, which is the binding site for the L25 protein, lacks the looped out G-residue (Correll et al., 1997; Dallas and Moore, 1997). Given the large number of homologues of both the L25 protein and the 5S RNA in various organisms, this system potentially provides a rich source of examples of participation of an apparently common RNA structural element in protein–RNA interactions. At present there are very few protein–RNA complexes for which high resolution structures are available and such data for the L25–E-loop complex is therefore of substantial interest.

Finally, there is evidence for the importance of protein–RNA interactions in an increasing variety of biological processes, including signal transduction and regulation of protein synthesis (for a recent review see Siomi and Dreyfuss, 1997). As ancient proteins, ribosomal proteins might be expected to define archetypal protein folds used in RNA binding. For example, structural data on the cold shock domain (CSD) (Newkirk et al., 1994; Schindelin et al., 1994), the ribosomal protein S17 (Jaishree et al., 1994), and the tRNA anticodon-binding domain of glutaminyl-tRNA synthetase (Brown et al., 1995), show that the residues involved in RNA binding are highly conserved, thereby providing further confirmation of the binding surface. Tertiary structure comparisons indicate that the six-stranded β-barrels of L25 and of the tRNA anticodon-binding domain of glutaminyl-tRNA synthetase are similar.

Keywords: NMR spectroscopy/protein structure/ribosomal protein/ribosome/RNA binding

Introduction

Recently, there has been renewed interest in the structure of ribosomal proteins for several reasons. First, protein biosynthesis is a central step in the realization of genetic information. Given the complexity of the ~2.3 MDa ribosome, recent progress in understanding the structural basis of protein biosynthesis has largely been based on a combination of low resolution methods, e.g. cryo-electron microscopy with three-dimensional image reconstruction together with footprinting, crosslinking experiments and three-dimensional modeling (e.g. Brimacombe, 1995; Frank, 1997; Mueller and Brimacombe, 1997a,h), with high resolution structures obtained by nuclear magnetic resonance spectroscopy (NMR) or X-ray crystallography for ribosomal proteins (for recent reviews see Liljas and Al-Karadaghi, 1997; Ramakrishnan and White, 1998) or fragments of ribosomal RNAs (White et al., 1992; Szewczak et al., 1993; Wimberly et al., 1993; Betzel et al., 1994; Szewczak and Moore, 1995; Fourmy et al., 1996; Correll et al., 1997; Dallas and Moore, 1997; Puglisi et al., 1997; Perbandt et al., 1998). The 5S rRNA constitutes one of the most highly conserved RNA sequences in nature (Specht et al., 1990). It is an essential component of the large ribosomal subunit, and in Escherichia coli it is complexed with the ribosomal proteins L5, L18 and L25 (Chen-Schmeisser and Garrett, 1977). Reconstituted 50S ribosomal subunits lacking the 5S RNA are inactive in protein biosynthesis (Erdmann et al., 1971; Hartmann et al., 1988). The distal end of the E-domain (also referred to as domain IV) of the 5S RNA, which is directly adjacent to the E-loop binding site for the L25 protein (Douthwaite et al., 1979; Garrett and Noller, 1979; Huber and Wool, 1984; Ciesiolka et al., 1992; Shpanchenko et al., 1996), has been shown to reside in proximity to the peptidyl transferase ring and the GTPase-associated area of the 23S rRNA leading to the suggestion that this domain might play an as yet unidentified role in the process of peptide bond formation (Dontsova et al., 1994; Sergiev et al., 1998).

Secondly, it has become increasingly apparent in recent years that in addition to simple A-form helices, RNA molecules often contain other characteristic structural elements, e.g. stable tetraloops (Shen et al., 1995) or A platforms (Cate et al., 1996), that may play important structural and functional roles. The D and E helices of the D/E domain of 5S RNA bound a large internal bulge or E-loop which has a tightly packed helical structure of unusual geometry due to the formation of an array of non-canonical base pairs. The (asymmetric) eukaryote E-loop bulge (Wimberly et al., 1993) has a looped out G-residue which contacts the helix in the major groove in a similar fashion to the α-sarcin/ricin loop of the 23S RNA (Szewczak et al., 1993; Szewczak and Moore, 1995) and constitutes a structure which probably is also similar to domains of potato spindle tuber virus (PSTV) RNA and the hairpin ribozyme (Shen et al., 1995). In contrast, the (symmetric) eubacterial E-loop bulge, which is the binding site for the L25 protein, lacks the looped out G-residue (Correll et al., 1997; Dallas and Moore, 1997). Given the large number of homologues of both the L25 protein and the 5S RNA in various organisms, this system potentially provides a rich source of examples of participation of an apparently common RNA structural element in protein–RNA interactions. At present there are very few protein–RNA complexes for which high resolution structures are available and such data for the L25–E-loop complex is therefore of substantial interest.

Finally, there is evidence for the importance of protein–RNA interactions in an increasing variety of biological processes, including signal transduction and regulation of protein synthesis (for a recent review see Siomi and Dreyfuss, 1997). As ancient proteins, ribosomal proteins might be expected to define archetypal protein folds used in RNA binding. For example, structural data on the cold shock domain (CSD) (Newkirk et al., 1994; Schindelin et al., 1994), the ribosomal protein S17 (Jaishree et al., 1994), and the tRNA anticodon-binding domain of glutaminyl-tRNA synthetase (Brown et al., 1995), show that the residues involved in RNA binding are highly conserved, thereby providing further confirmation of the binding surface. Tertiary structure comparisons indicate that the six-stranded β-barrels of L25 and of the tRNA anticodon-binding domain of glutaminyl-tRNA synthetase are similar.

Keywords: NMR spectroscopy/protein structure/ribosomal protein/ribosome/RNA binding

© Oxford University Press 6377
1996); the RNA-binding domain of ribosomal protein S1 (Bycroft et al., 1997) and the RNA-binding domains of aspartyl-/lysyl-/phenylalanyl-tRNA synthetases (Cavarelli et al., 1993; Onesti et al., 1995; Mosyak et al., 1995) suggest a potential evolutionary relationship between ribosomal proteins, stress proteins and tRNA synthetases. Because sequence homology of families of RNA-binding proteins may be low, the availability of three-dimensional structures can prove crucial to the recognition of such families.

In the present paper we report the NMR solution structure of the L25 protein from *E.coli* and a determination of the protein surface which is involved in binding the E-loop of 5S rRNA. As with other ribosomal proteins, the availability of a high resolution NMR structure for L25 may aid in future efforts to obtain a precise placement of this protein within low resolution models of the ribosome and hence in assessing possible roles of 5S rRNA in protein biosynthesis. L25 shows a new topology for binding of RNA. Interestingly, many of the residues of L25 involved in RNA binding are widely distributed through the amino acid sequence, but are located on a contiguous surface of the protein structure and are among the most highly conserved positions when compared either with the sequences of the related (Gryaznova et al., 1996) general stress protein CTC or with equivalent ribosomal proteins from other eubacteria. We suggest that the structure of L25 may constitute a prototype for a new family of RNA-binding proteins which recognize irregular dsRNA bulges related to the ribosomal E-loop. Furthermore, structural comparisons indicate that L25 is related to the tRNA anticodon-binding domain of glutaminyl-tRNA synthetase (GlnRS; Rould et al., 1991). This raises interesting questions about evolutionary relationships, even though the present data suggest that L25 and GlnRS probably bind their cognate RNAs in a different fashion.

### Results

**Structure determination**

The ribosomal protein L25 was purified in uniformly $^{15}$N and $^{15}$N/$^{13}$C-labeled form from *E.coli* ribosomes (see Materials and methods). For this 11 kDa protein, a nearly complete set of $^1$H, $^{15}$N and $^{13}$C NMR assignments (94.5% backbone, 99% side chain CH) was obtained using multidimensional heteronuclear NMR spectroscopy. A series of three- and four-dimensional heteronuclear NOE spectroscopy experiments and a HNCA[$^1$H] NMR experiment were used to extract distance and $\phi$ dihedral angle constraints, respectively. This information was used to calculate and refine the structure of L25 with the FOUND (Güntert et al., 1998), DYANA (Güntert et al., 1997) and AMBER (Weiner et al., 1984) programs using recently described protocols (Ohlenschläger et al., 1998). A summary of the experimental constraints and of the structural statistics for the 20 structures with the lowest DYANA target function is given in Table I. A superposition of these 20 structures, which have been used to characterize the solution structure of L25, is shown in Figure 1.

### Structural description

L25 exhibits a compact fold consisting of a six-stranded $\beta$-barrel and two $\alpha$-helices (Figures 1 and 2A). The overall

---

Table I. Constraints and structural statistics for the 20 best NMR structures of protein L25

<table>
<thead>
<tr>
<th>Experimental constraints</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of assigned NOEs</td>
<td>2094</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of distance constraints</td>
<td>1540</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intrareside</td>
<td>209</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sequential</td>
<td>409</td>
<td></td>
<td></td>
</tr>
<tr>
<td>medium range (residue i to i + j = 2...4)</td>
<td>217</td>
<td></td>
<td></td>
</tr>
<tr>
<td>long range (residue i to i + j = 4)</td>
<td>705</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$ dihedral angle constraints</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hydrogen bond constraints</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computational statistics$^a$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYANA target function (Å$^2$)</td>
<td>0.52 ± 0.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMBER energy (kcal/mol)</td>
<td>$-1025 ± 45$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual violations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOE distance constraints, sum (Å)</td>
<td>2.96 ± 0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOE distance constraints, max (Å)</td>
<td>0.17 ± 0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dihedral angle constraints, sum (degrees)</td>
<td>5.22 ± 2.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dihedral angle constraints, max (degrees)</td>
<td>2.17 ± 1.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural statistics$^a$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSD$^b$ (Å) to the mean structure (except residues 10-24)</td>
<td>0.62 ± 0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>backbone heavy atoms</td>
<td>0.62 ± 0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>all heavy atoms</td>
<td>1.14 ± 0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSD$^b$ (Å) to the mean structure (core structure)$^c$</td>
<td>0.33 ± 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>backbone heavy atoms</td>
<td>0.33 ± 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>all heavy atoms</td>
<td>0.81 ± 0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$, $\psi$ angles consistent with Ramachandran plots (%)</td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>most favored regions</td>
<td>60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Average value ± standard deviation.

$^b$Superposition made for all heavy atoms of residues 3–9 and 25–94.

$^c$$\beta$-sheets and $\alpha$-helices.

---

Fig. 1. A stereo view of the ensemble of the 20 NMR solution structures of the L25 protein from *E.coli* with the lowest DYANA target function. The backbone atoms of residues 3–10, 24–31 and 37–94 were used for a least squares superposition of the structures. Only backbone carbon and nitrogen atoms are shown. A ribbon diagram of L25 illustrating the secondary structure elements is given in Figure 2A.
Solution structure of ribosomal protein L25

Fig. 2. (A) Ribbon diagram of the average L25 structure showing the global fold. The ends of the polypeptide chain (N and C), the β-strands (β1–β6) and the α-helices (α1 and α2) are labeled. (B) Topological arrangement of the secondary structural elements of the L25 protein. (C) Stereoview of amino acid residues with large 15N chemical shift changes upon binding to 5SE-rRNA represented on a ribbon diagram of the L25 structure. For these residues the heavy-side chain atoms are shown in stick representation. The color code is: yellow for chemical shift changes from 1 to 1.5 p.p.m., orange for changes from 1.5 to 2 p.p.m., and red for changes beyond this limit. Structural elements are labeled as in (A). Relative to (A), the view of L25 is rotated by 180° about a vertical axis in the plane of the page. (D) Stereoview of the mean structure of the L25 protein (blue and green ribbon diagrams) superimposed on a tube diagram of the two β-barrels of the anticodon-binding domain of E.coli GlnRS (red). (E) Topology of the two β-barrels of E.coli GlnRS forming the anticodon-binding domain.

shape of the protein is flattened with dimensions of \(37\times36\times22 \text{ Å}\). The connectivity of the secondary structure elements is β1-β2-β3-α1-α2-β4-β5-β6 (Figure 2B). Topo- logically, the β-barrel shows two repetitions of a motif consisting of two parallel and one antiparallel β-strand (Figure 2B). Either side of the barrel is formed by two
parallel strands, B1 (residues 3–8)/β4 (residues 61–65) and B2 (residues 26–32)/β6 (residues 88–93) and one antiparallel strand, β3 (residues 37–43) and β5 (residues 69–78), respectively (Figures 1 and 2A). The symmetric six-stranded β-barrel is closed by antiparallel β-sheets between strand B1 and the C-terminal residues of strand B3 as well as between strand B6 and the C-terminal residues of strand B5 (Figures 1 and 2A). Strands B3 and B4 are connected by a two-turn α-helix (α1, residues 45–51) and a second one-turn α-helix (α2, residues 55–58) linked to α1 by a short loop of three residues. The axes of the two helices are oriented perpendicularly to each other. Helix α1 closes the bottom of the β-barrel. The short loop between β4 and β5 is well defined (Figure 1), whereas the loops connecting β2/β3 and β5/β6 are somewhat less well defined by the present data. The 17-residue loop between B1 and B2 showed very few NOE contacts to the core of the protein. Nearly half of the backbone amide hydrogens of this loop exchanged very rapidly with the solvent, and 15N{1H}-NOE data for the backbone amides (data not shown) indicated greater motional freedom for this backbone region. This loop, which contains a cluster of positively charged side chains, i.e. two lysines and four of the six arginine residues, appeared to have limited structural stability in the absence of bound RNA.

**Binding of the E-loop of 5S rRNA to the L25 protein**

Formation of a complex between L25 and an RNA construct containing the E-loop (Materials and methods) was monitored with a two dimensional 1H-15N HSQC correlation NMR experiment. A 1:1 protein–RNA complex was readily obtained and showed a high quality, very well dispersed 1H-15N HSQC NMR spectrum. Sequential assignments for the backbone amide 1H and 15N resonances were obtained with a three-dimensional HNCA NMR experiment. These experiments also revealed that a number of hydrogens which could not be observed in the free protein due to rapid exchange with solvent, i.e. five arginine NHs (Arg9, 18, 19, 21, 79) and four backbone amide hydrogens (Lys14, Gly15, Arg18, Leu86), became observable in the RNA–protein complex. This observation suggested that binding to the E-loop of the RNA stabilizes the conformation of the L25 loop residues 9–25 that were poorly defined in the solution structure of the free protein (Figure 1). At the present stage, comparison of the 15N chemical shifts of the L25 backbone amide nitrogens in the free and RNA-bound protein reveals that all substantial changes in chemical shifts are restricted to residues (Lys10 to Gly13, Ser17, Arg19–Ala22, Asn24, Phe26, Ile29 to Tyr31, Ala39, Gln41, His44, Gln57, Asp76, Lys85, His88, Asp90 and Val92) located on one side of the L25 structure (Figure 2C). These results suggest that the lysine/arginine rich loop region (residues 9–25) as well as the surface of β-strands β2 and β6 and the N- and C-terminal residues of β3 and β5, respectively, constitute the regions of L25 involved in binding the cognate E-loop RNA.

**Proteins with sequence homology to L25**

Comparison of the sequence of L25 to functionally equivalent ribosomal proteins and to a variety of open reading frames in other bacteria reveals that the L25 residues with the most pronounced chemical shift perturbations upon rRNA binding (Figure 2C) show a strong correlation with the most highly, if not absolutely, conserved amino acid sequence positions (Figure 3). In particular, residues in the loop joining strands B1 and B2 (Arg9, Arg18 and Arg21), in β2 (Ile29 and Tyr31) and in β6 (His88 and Asp90), all of which lie in the region implicated as the binding surface for the E-loop on L25, are conserved. Other residues that are conserved (Ala6, Pro27, Ala28, Gly67, Val72, Pro81 and Phe91) appear to have clear roles in maintaining the structure of L25, e.g. Val72 and Phe91 contribute to the hydrophobic core of the β-barrel and Gly67 is involved in the tight loop connecting β4 and β5. On the basis of moderate, pairwise sequence homologies (18–28%; Gryaznova et al., 1996), it has previously been proposed that L25, the functional analogue of L25 in *Thermus thermophilus*, the protein TL5 (Gongadze et al., 1993), and the general stress protein CTC from *Bacillus subtilis* are structurally homologous (Gryaznova et al., 1996). From the structure of L25 and the additional sequence data that have become available recently (Figure 3), it is now apparent that most of the residues that have been implicated in the RNA-binding surface of L25 are conserved in the general stress protein CTC.

**Proteins with tertiary structure homology to L25**

To search for proteins with similar structural folds to L25, the protein data bank was searched using the DALI server (Holm and Sander, 1993) at the EMBL Heidelberg, Germany. Only one known structure was found to contain the same fold: GlnRS (Rould et al., 1991) from *E.coli* (Figure 2D and E) which is highly related by sequence to GlnRSs from *Haemophilus influenzae* (Fleischmann et al., 1995), yeast (Ludmerer and Schimmel, 1987), the plant lupine [DDBJ/EMBL/GenBank database accession No. X91787, deposited by Siatecka (1995)] and humans (Lamour et al., 1994). In this protein, the anticodon-binding domain of the tRNA synthetase is formed by two β-barrels with an essentially identical topology. These β-barrels are arranged in an antiparallel orientation and each of them shows a high degree of structural similarity to the β-barrel of L25 (Figure 2D and E). For superposition of the backbone heavy atoms of the six β-strands, the r.m.s. displacements between the L25 protein and the two domains of GlnRS are 1.68 and 1.53 Å. However, there is one difference in the overall arrangement of the strands of the β-barrels of L25 and GlnRS, i.e. through topological ‘exchange’ strand B1 replaces strand B1’ in barrel 2 and vice versa (Figure 2E). Major differences between the two proteins exist both in the length and the orientation of the loops connecting the strands, the occurrence of helices, and additional β-ribbons in the case of GlnRS. The long arginine-rich loop of L25 (residues 9–25) is absent in GlnRS and even within the structurally homologous six-stranded β-barrel, there is very limited sequence homology between the positions constituting the β-barrels of L25 and GlnRS, respectively (Figure 3).

**Discussion**

Structures containing β-barrels or β-barrel-like motifs are already known for the ribosomal proteins L14 (Davies
et al., 1996), S1 (Bycroft et al., 1997) and S17 (Jaishree et al., 1996). However, these are five-stranded with the topology $\beta_3\beta_2\beta_1\alpha_1\beta_3$ (S1) and $\beta_3\beta_2\beta_1\alpha_1\beta_3$ (S17). The closed, six-stranded $\beta$-barrel of L25 containing four parallel strands and two short $\alpha$-helices with the topology $\beta_3\beta_2\beta_1\alpha_1\beta_3\alpha_1\beta_3$ represents a new fold for ribosomal proteins. This adds a new member to the group of RNA-binding proteins dominated by $\beta$-strands which are found in all organisms (Nagai, 1996; Siomi and Dreyfuss, 1997; Varani, 1997).

Protease protection studies on $E.coli$ L25 indicated that amino acid residues 1–22 are susceptible to trypsic hydrolysis in the absence of 5S rRNA and that the $\Delta N22$ fragment of L25 lacks RNA-binding activity (Newberry and Garrett, 1980) indicating that the arginine-rich N-terminus is required for binding and/or for the structural integrity of L25. In addition, 1D NMR studies of L25, both free (Kime et al., 1981; van deVen et al., 1983) and bound to 5S rRNA, suggested that arginine side-chains play an important role in RNA binding (Kime and Moore, 1984). In the structure presented here this N-terminal region comprises strand $\beta_1$ and the loop connecting $\beta_1$ with $\beta_2$. The backbone amide $^{15}$N chemical shift perturbations of L25 upon RNA binding strongly suggest that this loop together with $\beta_2$ and $\beta_6$ constitute the major RNA-binding elements of L25 (Figure 2C). However, since small changes of the local geometry and conformation can also lead to chemical shift perturbations upon binding of a ligand (Oldfield, 1995), the residues in the present work which show substantial chemical shift changes may not all be directly interacting with RNA. Conversely, in the case of the RNP motif, mapping of an RNA-binding site on a $\beta$-sheet structure via chemical shift changes has proven to be valid (Görlach et al., 1992; Oubridge et al., 1994; Allain et al., 1996). In addition, for the large loop joining strands $\beta_1$ and $\beta_2$, which are poorly defined in the solution structure of the free protein (Figure 1), a marked increase in the conformational stability in the RNA–protein complex is indicated by the $^{15}$N-{$^1$H}-NOEs (data not shown) and by a reduced exchange with solvent observed for the backbone amide hydrogens. This suggests an ‘induced fit’ type of binding of this loop to RNA. Similarly, in the ribosomal protein L11 an arginine- or lysine-rich loop is flexible in the free protein, but becomes ordered upon binding to its cognate RNA (Hinck et al., 1997; Markus et al., 1997). On the basis of the present data it appears that L25 may combine two features found for other RNA-binding proteins: the $\beta$-sheet region may act as a ‘platform’ (Görlach et al., 1992) which could interact with the flat and broadened minor groove of the central portion of the E-loop while a restructured arginine/lysine-rich loop in turn could fit into the widened major groove (Battiste et al., 1996) of the helix IV adjacent to the E-loop. While a full structure of the L25–E-loop complex is clearly still needed, this interpretation of the present data appears to be consistent with a suggested
model of RNA recognition deduced from the structure of E-loop rRNA (Correll et al., 1997; Dallas and Moore, 1997).

It has recently been proposed that as ancient proteins, ribosomal proteins may define archetypal folds for groups of now functionally distinct RNA-binding proteins that have evolved from a common ancestor (Bycroft et al., 1997). One such family, based on a five stranded β-barrel or a β-barrel-like motif, appears to exist for RNA-binding domains of the type found in the ribosomal protein S1 (Bycroft et al., 1997), the ribosomal protein S17 (Jaishree et al., 1996), the CSD (Newkirk et al., 1994; Schindelin et al., 1994) and domains of aspartyl-lysoyl-phenylalanyl-tRNA synthetases (Cavarelli et al., 1993; Mosyk et al., 1995; Onesti et al., 1995). On the present evidence, it seems that a further such family, based on a six-stranded β-barrel, exists for ribosomal protein L25, general stress protein CTC and GlnRS. The structural data reported here, as well as the elucidation of a corresponding motif for conserved residues which is distributed over the entire protein sequence, suggest that L25 and the general stress protein CTC may recognize similar double-stranded RNA bulge motifs. This information should help in identifying further members of this putative family of RNA-binding proteins. On the other hand, the recognition of the single-stranded anticodon region of tRNA by GlnRS involves a cleft between two six-stranded β-barrels (Roulde et al., 1991), i.e. an apparently different mode of RNA recognition. The very low level of sequence homology between L25/CTC and GlnRS (Figure 3) might simply reflect the different prerequisites for recognition of irregular dsRNA, i.e. the E-loop rRNA, versus ssRNA, i.e. the tRNA anticodon loop, which may have been acquired through divergent evolution. However, a recent model for the evolution of the glutamyl-/glutaminyl-tRNA synthetase family suggests the acquisition of GlnRS only for a subgroup of the eubacterial kingdom, excluding Bacillus subtilis which lacks GlnRS (Ibba et al., 1997), via horizontal gene transfer from an eukaryotic source (Lamour et al., 1994). Hence, it remains possible that the six-stranded β-barrel, which constitutes a RNA-binding fold for L25 and the anticodon-binding domain of GlnRS, represents a case of convergent evolution.

Materials and methods

Preparation of E. coli L25 samples

Uniformly 15N and 15N/13C-labeled L25 was purified from an acetic acid extract of 70S polysomes of E.coli MRE 600 cells grown in M9 media containing the appropriate isotopes as the sole carbon and/or nitrogen source. Ribosomal proteins were precipitated by acetone fractionation. Proteins from a 65–78% aceton/acetone fraction were dissolved in 50 mM sodium phosphate buffer pH 6.5, containing 6 M urea and 4 mM 2-mercaptoethanol, and separated by ion-exchange chromatography on Sepharose Fast Flow (Amersham Pharmacia Biotech). Fractions containing L25 were refolded by dialysis against 50 mM sodium phosphate buffer pH 7.0, containing 0.1 M potassium chloride. L25 was further purified by HPLC on a Mono S HR 5/5 or SOURCE 15S PE 4.6/100 column (Amersham Pharmacia Biotech). Protein samples were concentrated with Centricron-3 concentrators (Amicon, Inc.), and the buffer was exchanged against NMR buffer consisting of 20 mM sodium phosphate pH 7.0, 0.1 M potassium chloride, 3 mM sodium azide and 5% D2O. Protein concentrations were determined by UV spectrophotometry (ε278 = 3840 l/mol/cm) to be 1 and 1.4 mM for the 15N and 15N/13C-labeled NMR samples, respectively. For NMR experiments in D2O the protein samples were lyophilized and redissolved in 99.996% D2O (Cambridge Isotopes Laboratories, Inc).

Preparation of RNA and the protein–RNA complex

A 37 nt oligonucleotide of the sequence 5′-GGACGAGUGUGUCUCGAGAUAGGUGUC-3′ was bound to an optimized 30 ml transcription reaction was carried out essentially as described elsewhere (Grüne et al., 1996). Self-cleavage during transcription released the 5SE-rRNA with a uniform 3′-end. The 5SE-rRNA was purified on a DEAE-Sepharose FF column (Amersham Pharmacia Biotech) with 50 mM potassium phosphate and 2 mM tetramethylammonium hydrogen sulfate, pH 5.9, in water employing an acetonitrile gradient. The product was lyophilized and desalted with a NAP-25 gel-filtration column (Amersham Pharmacia Biotech), precipitated twice with a 5-fold volume of 2% lithium perchlorate in acetone, folded into a monomeric hairpin form by denaturing at 368 K at a concentration of 0.25 mM and subsequent dilution to 0.05 mM with ice cold water and finally exchanged into NMR buffer using Centricron-3 microconcentrators (Amicon, Inc.). Samples of the protein–RNA complex were prepared in buffer containing 20 mM sodium phosphate, pH 7.2, 0.1 M potassium chloride, 4 mM magnesium chloride, 3 mM sodium azide and 5% D2O. RNA was added and the titration monitored by 2D 1H-15N HSQC spectra until peaks arising from free protein disappeared, i.e. until a 1:1 complex was obtained.

NMR spectroscopy and spectral assignments

NMR experiments were carried out on a Varian Unity NMR spectrometer operating at 750 MHz proton frequency, at 298 K. Triple resonance NMR spectra used for assignments were processed with the VNMR software (Varian Associates, Inc.). PROSA (Güntert et al., 1992) and NMRPipe (Delaglio et al., 1995) were used for processing of the 3D NOESY and 4D NOESY spectra, respectively. XEASY (Bartels et al., 1995) was employed for visualization and analysis of the spectra. Backbone 1H, 13C and 15N chemical shift assignments were obtained through a 3D HNCA experiment (Wittekind and Mueller, 1993), 3D HNCO (Ikura et al., 1990) and 3D CCH-TOSY-NH (Grzesiek et al., 1993) experiments. Side chain aliphatic 1H and 13C chemical shift assignments were completed by using 3D HCH-TOSY-NH (Grzesiek et al., 1993), 3D doubly se-H(C)CH-TOCSY (Sattler et al., 1995) and 3D 1H-1H-15N TOCSY-HSQC (Marion et al., 1989) experiments. Aromatic resonances were assigned through a selective 3D 1H-1H-13C NOESY-HSQC (Zaiderweg and Fesik, 1989) spectrum (100 ms mixing time). Slowly exchanging amide protons were identified through a 1H-15N HSQC spectrum after dissolving the lyophilized 15N-labeled H2O NMR sample in D2O. NOE constraints for structural calculations were obtained from the following experiments: 3D 1H-1H-15N NOESY-HSQC (100 ms mixing time), 3D 1H-1H-15N NOESY-HSQC (80 ms mixing time) selective for aromatic carbons and 3D 1H-1H-13C HSQC-HSQC (80 ms mixing time) selective for aliphatic carbons. Assignment of the 15N-edited NOESY spectrum was assisted by a 3D 1H-1H-13C/HMQC-NOESY-HSQC (Muhandiram et al., 1993). Virtually all 1H-edited NOEs as well as 13C-edited NOEs originating from aromatic, H2O and resolved methyl groups were assigned manually. Additional NOEs were then automatically assigned using the NOAH algorithm (Mumenthaler et al., 1997) as implemented in the DYANA program.

Structure calculation

NOE crosspeak intensities were classified as strong, medium and weak, corresponding to upper distance limit constraints of 2.7, 3.8 and 5.5 Å, respectively. Inter-amide proton NOE constraints were generally transferred into upper distance limits of 4.4 Å. Dihedral angle constraints were generated by local conformational analysis with the HABAS algorithm (Güntert et al., 1991) as implemented in FOUND (Güntert et al., 1998) based on 13Cα, 13Cβ vicinal coupling constants and intraresidual and sequential NOEs. Typically, 100 structures were generated by making use of the hybrid distance geometry/simulated annealing protocol of DYANA starting...
from random conformations. At the last stage, hydrogen bonds consistently formed in previous calculations were introduced as additional constraints. Out of 100 structures, the 20 with the lowest DYANA target functions were subjected to energy minimization using the AMBER force field (Weiner et al., 1984) and used to characterize the solution structure of the L25 protein. Illustrations of protein structures were generated with MOLMOL (Karoda et al., 1996).

Acknowledgements

The authors thank A. Figraphy for technical assistance in the preparation of isotope-labeled materials, O. Ohlenschläger for help on structural calculations and helpful discussions, and J. Leppert for processing of the four-dimensional NMR spectra. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (Br 1487/2-3). J.W. is grateful to the Fond der Chemischen Industrie for financial support.

References


Bitar, K.G. and Wittmann-Liebold, B. (1975) The primary structure of the L25 protein. Illustrations of protein structures were based on UNIX pipes. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (Br 1487/2-3). J.W. is grateful to the Fond der chemischen Industrie for financial support.

Solution structure of ribosomal protein L25


M.Stoldt et al.


Received July 17, 1998; revised September 4, 1998; accepted September 7, 1998