

# Conformational variability of the N-terminal helix in the structure of ribosomal protein S15

William M Clemons Jr<sup>1</sup>, Christopher Davies<sup>2</sup>, Stephen W White<sup>2,3\*</sup> and V Ramakrishnan<sup>1\*</sup>

**Background:** Ribosomal protein S15 is a primary RNA-binding protein that binds to the central domain of 16S rRNA. S15 also regulates its own synthesis by binding to its own mRNA. The binding sites for S15 on both mRNA and rRNA have been narrowed down to less than a hundred nucleotides each, making the protein an attractive candidate for the study of protein–RNA interactions.

**Results:** The crystal structure of S15 from *Bacillus stearothermophilus* has been solved to 2.1 Å resolution. The structure consists of four  $\alpha$  helices. Three of these helices form the core of the protein, while the N-terminal helix protrudes out from the body of the molecule to make contacts with a neighboring molecule in the crystal lattice. S15 contains a large conserved patch of basic residues which could provide a site for binding 16S rRNA.

**Conclusions:** The conformation of the N-terminal  $\alpha$  helix is quite different from that reported in a recent NMR structure of S15 from *Thermus thermophilus*. The intermolecular contacts that this  $\alpha$  helix makes with a neighboring molecule in the crystal, however, closely resemble the intramolecular contacts that occur in the NMR structure. This conformational variability of the N-terminal helix has implications for the range of possible S15–RNA interactions. A large, conserved basic patch at one end of S15 and a cluster of conserved but exposed aromatic residues at the other end provide two possible RNA-binding sites on S15.

## Introduction

Translation of the genetic code occurs on the ribosome, a large composite structure of RNA and protein that provides the framework for recognition of messenger RNA (mRNA) by amino-acylated transfer RNAs (tRNAs), and catalyzes peptidyl-transferase activity. In prokaryotes, the ribosome is a 70S particle that consists of two major parts: the 50S (large) subunit and the 30S (small) subunit. Both of these subunits have activities that are important for the translation of mRNA into protein.

While much is known about protein–DNA interactions, the field of protein–RNA interactions (apart from those involving tRNA) is still in its infancy. As many ribosomal proteins bind specifically to ribosomal RNA (rRNA), they provide excellent opportunities for elucidating the poorly understood principles of protein–RNA recognition. High-resolution crystal and nuclear magnetic resonance (NMR) structures of many individual ribosomal proteins have already provided insights into the nature of their RNA-binding regions [1,2]. Recent work has focussed on the so-called ‘primary’ RNA-binding proteins from the small ribosomal subunit [3–6]. This important subset of ribosomal proteins can bind rRNA

Addresses: <sup>1</sup>Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84132, USA, <sup>2</sup>Department of Structural Biology, St Jude Children’s Research Hospital, 332 North Lauderdale Street, Memphis, TN 38105, USA and <sup>3</sup>Department of Biochemistry, University of Tennessee Memphis, 858 Madison Suite G01, Memphis, TN 38163, USA.

\*Corresponding authors.

E-mail: v.ramakrishnan@m.cc.utah.edu  
stephen.white@stjude.org

**Key words:** multiwavelength anomalous diffraction, protein–RNA interactions, ribosome, RNA binding, X-ray crystallography

Received: 5 January 1998

Revisions requested: 23 January 1998

Revisions received: 29 January 1998

Accepted: 30 January 1998

**Structure** 15 April 1998, 6:429–438

<http://biomednet.com/elecref/0969212600600429>

© Current Biology Ltd ISSN 0969-2126

independently of the other proteins and often have the best characterized binding sites on rRNA. These primary proteins initiate the higher order folding of rRNA which leads to the subsequent binding of the remaining proteins. The wealth of biochemical and electron microscopic structural information on the 30S subunit has resulted in fairly detailed three-dimensional models of the particle. The structures of ribosomal proteins have proved useful in both constraining and clarifying these models [7–9]. The availability of more structures of small subunit proteins, as well as higher resolution image reconstructions by cryo-electron microscopy, will mean that the possible folds of the rRNA molecule of the 30S subunit (16S rRNA) will become increasingly restricted and better defined.

Ribosomal protein S15 is a highly basic protein. The protein from *Bacillus stearothermophilus* has a molecular weight of 10,560 and a predicted *pI* of 10.7. Neutron diffraction and electron microscopic studies have shown that S15 binds in the core of the 30S subunit [10,11]. S15 is one of seven primary RNA-binding proteins in the small subunit that initially bind to 16S rRNA during ribosome assembly [12,13]. Studies have shown that the

important regions on 16S rRNA for S15 binding include helices 21 and 22 [14,15]. The minimal binding site has been defined more precisely using a combination of protection and binding affinity studies. This site consists of 61 nucleotides, and includes conserved unpaired regions located on helix 22 and at a three-way junction between helices 21, 22 and 23 [16,17]. Mutational studies on the rRNA have also shown that S15 is able to recognize specific structural motifs in this region. Important recognition factors include four phylogenetically conserved nucleotides in the three-way junction, a bulged adenosine near the three-way junction, and a G–U base pair just above the internal loop on helix 22 [16].

Several primary rRNA-binding proteins, including S15, are able to regulate the translation of their operon by binding to their own mRNA [18,19]. In many of these cases, the interaction of the protein with its mRNA is similar to its interactions with rRNA [20]. In the case of S15, the similarity is not obvious as the specific binding determinants appear to be different for the mRNA and minimal rRNA-binding sites [21]. It is known, however, that the binding of S15 to mRNA promotes the formation of a pseudo-knot structure that overlaps the ribosome-binding site, thereby preventing translation [22]. Here we present the 2.1 Å crystal structure of ribosomal protein S15 from *B. stearotherophilus*. The structure shows that S15 is a largely helical protein, in which the N-terminal helix is capable of conformational variability. We also present our analysis of the structure and the conformational variability in terms of the RNA-binding function of S15.

## Results

### Crystallization

Crystals of *B. stearotherophilus* S15 were obtained using high concentrations of phosphate as the precipitant. The optimal conditions were between 2.4 M and 3.0 M Na/KPO<sub>4</sub> at 23°C, and within a narrow pH range centered at pH 6.5. Selenomethionyl crystals of S15 were obtained in 3.0 M Na/KPO<sub>4</sub>, 5 mM 2-mercaptoethanol at pH 6.5. All X-ray data were collected in a cold stream at 100K, and the following protocol was necessary for cryoprotection. Fully grown crystals were obtained in about two weeks and transferred into a drop containing mother liquor. The glucose concentration in the drop was raised to 20% in a series of steps, while keeping the phosphate concentration constant and allowing time for equilibration at each step. The crystals were then flash-cooled by rapid immersion in liquid nitrogen. The crystals are in space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>, have unit cell dimensions a = 33.02 Å, b = 33.75 Å, c = 75.67 Å and diffract to 2.1 Å.

### Structure determination

The structure of S15 was determined by multiwavelength anomalous dispersion (MAD) on the selenomethionyl protein. Data were collected on beamline X12-C

at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory, and only two wavelengths were used owing to limitations in beam time. The first wavelength was at the peak of  $f''$  (the 'white line') at 0.98 Å, and the second was at 0.93 Å; the difference in  $f'$  between the two wavelengths was expected to give a measurable dispersive difference. These data were collected to a resolution of 2.64 Å. The crystal was stored in liquid nitrogen and used subsequently for data collection to 2.1 Å on an image-plate system mounted on a rotating-anode source (1.54 Å). Despite the differences in detector and source, the rotating-anode data scaled extremely well with the synchrotron data sets, and showed a large peak corresponding to the selenium atom in the difference Patterson map (data not shown). This observation suggests that much of the reduction of systematic error in MAD experiments comes from using the same crystal in the same orientation for the various wavelengths, thus minimizing absorption errors in scaling. Data collection statistics are shown in Table 1. All phasing was done with the rotating-anode data as the reference or 'native' data in treating MAD as a special case of multiple isomorphous replacement (MIR) [23,24]. The phasing calculation was done using the maximum-likelihood program SHARP [25]. Table 2 shows statistics for the phasing calculation. The figure of merit is lower than values usually quoted in structure determinations. One should keep in mind, however, that the figures of merit produced by SHARP are realistic rather than over- or underestimated [24]. Moreover, the relatively unbiased phase probability distributions result in a good starting point for density modification. The original phases to 2.64 Å were solvent flattened and extended to 2.1 Å using the SOLOMON option [26] built into the SHARP interface. This resulted in an interpretable map

**Table 1**

#### Data collection statistics.

	SeMet* ( $\lambda_1$ )	SeMet* ( $\lambda_2$ )	SeMet* ( $\lambda_3$ )
Source	CuK $\alpha$	X12-C (NSLS) <sup>†</sup>	X12-C (NSLS) <sup>†</sup>
Wavelength (Å)	1.54	0.98	0.93
$f'$ of Se (electrons)	-0.88	-7.35	-2.19
$f''$ of Se (electrons)	1.14	5.92	3.46
Resolution (Å)	2.1	2.64	2.64
Total reflections	139,136	24,773	23,475
Unique reflections	6313	2676	2680
Completeness (%)	99.5	93.6	94.2
R <sub>sym</sub> (total/outer shell)	0.075/0.343	0.044/0.101	0.043/0.080

\*SeMet = selenomethionyl protein. <sup>†</sup>NSLS = National Synchrotron Light Source.

**Table 2****Phasing statistics.\***

	$\lambda_3-\lambda_1$ Isomorphous	$\lambda_2-\lambda_1$ Isomorphous	$\lambda_2$ Anomalous	$\lambda_3$ Anomalous
$R_{\text{Cullis}}$ (centrics) <sup>†</sup>	0.56	0.76	–	–
$R_{\text{kraut}}$ (acentrics) <sup>‡</sup>	0.033	0.035	0.038	0.036
Phasing power <sup>§</sup>	0.56	1.1	1.6	1.4

\*Statistics obtained using  $\lambda_1$  as the reference or 'native' data set.

<sup>†</sup> $R_{\text{Cullis}} = \sum(|F_{\text{PH}} - F_{\text{P}}| - F_{\text{H}}) / \sum|F_{\text{PH}} - F_{\text{P}}|$ . <sup>‡</sup> $R_{\text{kraut}} = \sum|F_{\text{PH}}(\text{obs}) - F_{\text{PH}}(\text{calc})| / \sum F_{\text{PH}}(\text{obs})$ , where  $F_{\text{P}}$  and  $F_{\text{PH}}$  are the structure factors for the 'native' and, 'derivative' respectively.

<sup>§</sup>Phasing power =  $F_{\text{H}}/E_{\text{rms}}$ , where  $E_{\text{rms}}$  is the residual lack of closure. The overall figure of merit to 2.64 Å was 0.36 and 0.33 for acentric and centric reflections, respectively.

(Figure 1) in which the helices were clearly visible in the skeletonization. The model was built and refined as described in the Materials and methods section and the statistics for the final model are shown in Table 3. An analysis by PROCHECK [27] showed that of the 85 residues in the model, all but three were in the most favored or additionally allowed regions of the Ramachandran plot, and none were in the disallowed region.

**Structure of S15**

Our model of S15 comprises the entire molecule apart from one residue at the N terminus and two at the C terminus. In addition, experimental density is not present for a partially disordered residue in a loop region. The structure of S15 consists of four  $\alpha$  helices: helix 2 (residues 23–44), helix 3 (residues 51–71) and helix 4 (residues 74–85) which pack together to form the core of the protein; and helix 1 at the N terminus (residues 3–15). Helix 1 extends away from the body of the molecule and associates with another molecule in the crystal through a series of interactions involving conserved hydrophobic residues. The C $\alpha$  positions and a ribbon

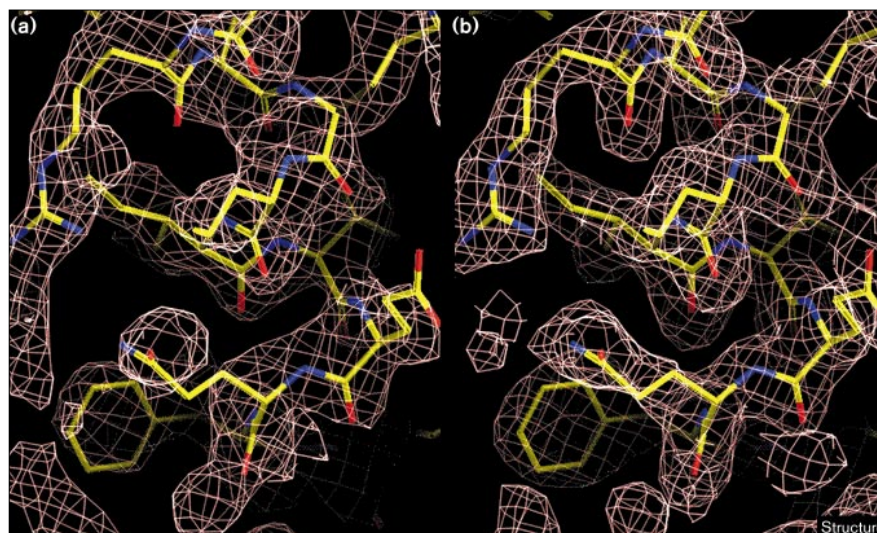
diagram of S15 are shown in Figure 2. The figure also shows a symmetry-related helix 1, which is very likely to represent an alternate conformation of helix 1 in solution (see below). The overall dimensions of S15 are 44 Å  $\times$  22 Å  $\times$  18 Å for the body, with helix 1 protruding out 36 Å from the rest of the molecule.

**Body**

The principal stabilizing feature of S15 is a coiled-coil interaction between the longer helices 2 and 3. This interaction is mediated by a stripe of largely conserved hydrophobic residues that span residues 24–44 in helix 2 and residues 51–69 in helix 3. As viewed in Figure 2, the shorter C-terminal helix 4 packs against the coiled-coil at the top of the molecule to form a three-helix bundle. The packing between the three  $\alpha$  helices is maintained by hydrophobic interactions involving highly conserved residues in this region. These residues include Ile28, Leu31, Ile35, Leu38, Leu55, Met58, Val59 and Leu66 (Figure 3). All four helices are stabilized by capping residues at the N terminus [28]. These residues include Thr3 for helix 1, Ser23 for helix 2, and Asp48 and Asp73,

**Figure 1**

Electron-density maps of S15. (a) Experimental and (b)  $2F_o - F_c$  maps. The region around Phe14 is shown. The maps are contoured at  $1\sigma$  using the program O [43]. The  $2F_o - F_c$  map was generated from a model that had been refined to 2.1 Å. The final model is superimposed in stick representation with atoms shown in standard colors.



**Table 3**

Refinement statistics.	
R factor	0.212
R free	0.318
Resolution range (Å)	6.0–2.1
Number of reflections	6149
$\sigma$ cut-off	0.0
Number of nonhydrogen protein atoms	717
Number of water molecules	59
Rms deviation from ideal geometry	
bond lengths (Å)	0.007
bond angles (°)	1.14

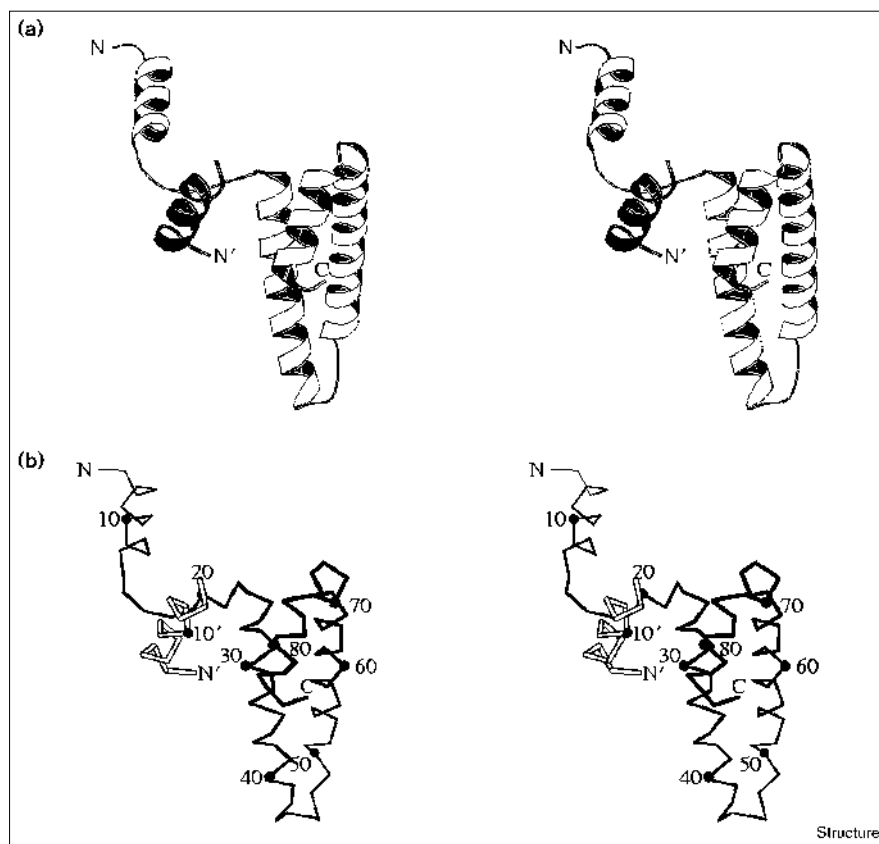
which make hydrogen bonds with mainchain amides of Ser51 of helix 3 and Arg76 of helix 4, respectively. The first of these capping residues is conserved as a serine or threonine in nearly all species (Figure 3) while the others are conserved in eubacteria. Finally, a number of conserved salt-bridge interactions, such as Arg76 to Glu25, add to the overall stability of the molecule.

### Loops

The only residue in the molecule that is not visible in the original experimental electron-density map is Glu18. This

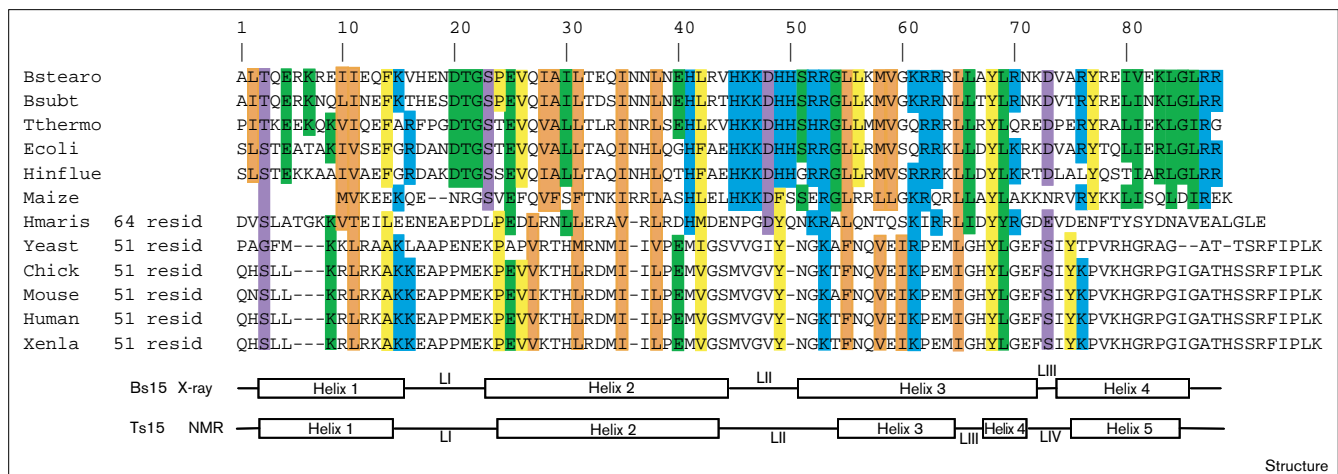
residue is located within loop LI (residues 15–23) which connects helices 1 and 2. The electron density within this solvent-exposed loop is generally weak and it appears to be an inherently flexible region of the molecule. This observation is supported by the amino acid sequence which contains a predominance of polar residues throughout the region and a highly conserved glycine (Gly22) at the C-terminal end prior to helix 2. It appears that the flexibility of loop LI accounts for the extended orientation of helix 1 away from the body of the molecule. As this flexibility is mediated by the conserved residue Gly22, it is probably important for the function of S15.

Loop LII (residues 44–51), that connects helices 2 and 3, is significant because it is highly ordered and contains a large number of conserved basic and aromatic residues including several histidines. The loop structure is maintained by a network of hydrogen bonds. The fact that S15 crystals require a pH very near 6.5 for growth could be explained by a network of hydrogen bonds involving histidines, which is important for the correct packing interaction of helices 2 and 3. The large number of conserved basic residues within the loop and at the ends of the flanking  $\alpha$  helices make this part of the molecule highly positively charged and probably a major site of interaction with RNA.

**Figure 2**

Overall fold of S15. (a) Stereo view ribbon diagram of S15 with the alternate conformation of the N-terminal helix shown in dark gray. (b) Stereo view C $\alpha$  trace in the same view as (a) with every tenth residue shown as a small black sphere and labeled; the alternate conformation of the helix is shown in white. (The figures were made using the program MOLSCRIPT [46].)

Figure 3



Alignment of S15 amino acid sequences from widely divergent species: Bstearo, *Bacillus stearothermophilus*; Bsubt, *Bacillus subtilis*; Tthermo, *Thermus thermophilus*; Ecoli, *Escherichia coli*; Hinflue, *Haemophilus influenzae*; Maize, *Zea mays* chloroplast; Hmaris, *Haloarcula marismortui*; Yeast, *Saccharomyces cerevisiae*; Chick, *Gallus gallus*; Mouse, *Mus musculus*; Human, *Homo sapiens*; and Xenla, *Xenopus laevis*. The sequences are color-coded: blue, basic putative RNA-binding residues; yellow, hydrophobic putative

RNA-binding residues; orange, conserved hydrophobic core; purple, conserved N cap residues; green, other conserved residues. The sequence numbers given on the top line relate to the *B. stearothermophilus* sequence. The positions of secondary structure elements, according to the X-ray crystal and NMR structures, are indicated below. (The sequences were aligned using the MULTALIN program [47].)

The sequence of S15 in the Genbank database (accession code 133791) appears to be in error. Sequencing of four independent clones showed that residue 45 within LII is a histidine and not an arginine. Mass spectrometry on purified protein (data not shown) confirmed our sequence, and our finding is consistent with the fact that this histidine is completely conserved in all known S15 sequences from prokaryotes (Figure 3).

The residues connecting helices 3 and 4 form a helical hairpin with a short well-structured loop (LIII), in which only a single residue, Asp73, is not part of either helix.

#### N-terminal helix

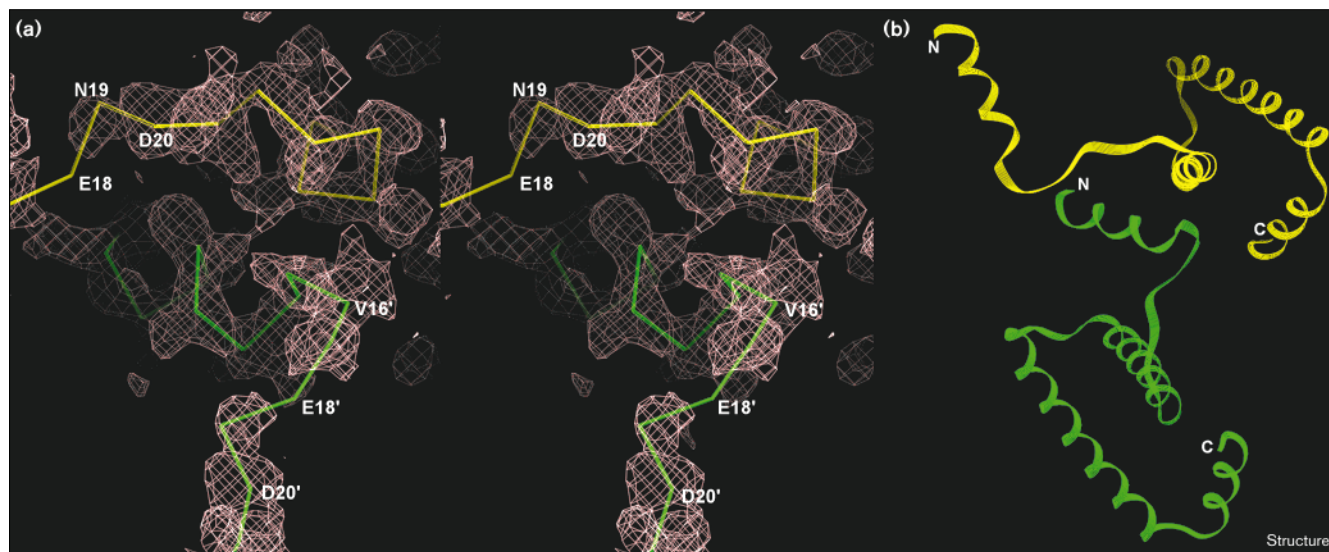
The most intriguing part of the S15 structure is the orientation of helix 1, which protrudes out from the body of the molecule and makes extensive contacts with a symmetry-related molecule. These contacts involve the hydrophobic residues Leu2, Ile10, Ile11 and Phe14 from helix 1 and Ala29, Val26 and Ile30 from the body of the molecule. All of these residues are conserved as hydrophobic residues in S15 from other prokaryotes. It is evident from Figure 2 that helix 1 could occupy one of two possible positions in the crystal. The two possibilities, namely packed against a neighboring molecule or packed against its own molecule, are difficult to distinguish. The first interpretation, however, is likely to be correct because it is consistent with the location and direction of the electron density for loop LI which, although weak, is clearly

visible even in the original experimental map. The chosen assignment only results in missing density for Glu18, whereas the alternative produces both missing density over several residues and unassigned density (Figure 4a). The packing of two symmetry-related molecules is shown in Figure 4b.

#### RNA-binding interface

Our earlier analyses of ribosomal protein structures have suggested that their RNA-binding sites are heavily populated by exposed, conserved, basic and hydrophobic (mainly aromatic) residues [1]. These binding sites are also frequently located in loop regions. Residues in S15 that fit these criteria are displayed in Figure 5a. The residues cluster into two groups at either end of the molecule: the 'top' and 'bottom' as viewed in Figure 5a. The top group includes the highly conserved Phe14, Tyr68 and Tyr77. Of these, Phe14 is located further away from the other two residues in the extended form of the N-terminal helix found in the crystal structure. If this helix were to reassociate with its own molecule, however, it would bring these residues closer together. In any case, the pattern of exposed aromatic residues with nearby basic residues is characteristic of an RNA-binding surface. The bottom group, centered on loop LII, produces a large patch of positive potential on one face of the molecule which is clearly seen in Figure 5b. This patch almost certainly reflects an extensive interaction with RNA.

Figure 4



The extended nature of loop LI of S15 and the packing of a symmetry-related N-terminal helix. **(a)** Stereo view of the packing of symmetry-related N-terminal helix (green) relative to the S15 structure (yellow) showing

experimental density contoured to  $1\sigma$ . **(b)** Ribbon diagram of two symmetry-related molecules in the crystal structure; the N and C termini are marked. (The figures were made using the program O [43].)

## Discussion

### Comparison with the NMR structure

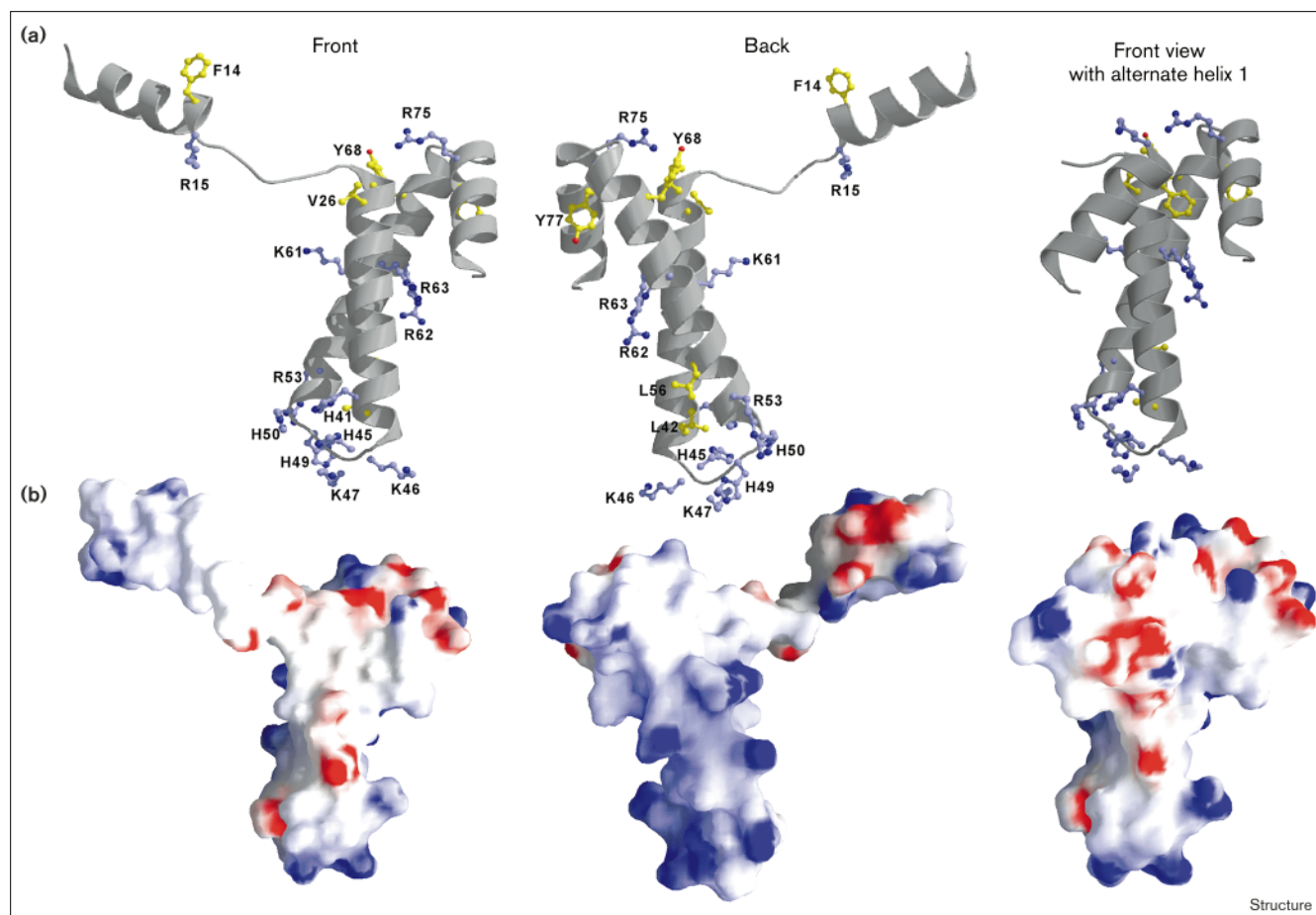
The NMR structure of S15 from *Thermus thermophilus* has been determined [29] and Figure 6 shows a comparison of the X-ray and NMR structures. The  $C\alpha$  root mean square (rms) between the X-ray structure and the entire family of NMR structures was 3.3 Å for the body of the molecule, in which the N-terminal helix and loop LI were excluded from the comparison, and 2.5 Å for just the helical portions of this region. A comparison of the secondary structure assignments is shown in Figure 3. While the overall topology of the two structures is similar, there are significant differences in two regions of the molecule: the position and orientation of helix 1, and the region encompassing loop LII and helix 3.

In the NMR model helix 1 is positioned against the body of S15, whereas in the X-ray structure it is extended away from the molecule to make contacts with a symmetry-related molecule in the crystal. The contacts that helix 1 makes with the neighboring molecule in the crystal structure, however, closely parallel the intramolecular contacts made by this helix in the solution NMR structure. The implication is that helix 1 is capable of peeling off from the molecule and replacing the corresponding helix in a neighboring molecule. This suggests that helix 1 has conformational variability, and also that the connecting loop LI has inherent and possibly functionally important flexibility.

Another difference between the X-ray and NMR models is the conformation of loop LII and helix 3. In the crystal

structure both of these regions are highly ordered; helix 3 is continuous showing the most well-defined region of electron density in the molecule. Moreover, helix 3 is involved in extensive coiled-coil interactions with helix 2. In the NMR structure, however, loop LII is more extensive and disordered and helix 3 is broken up into two helices (NMR helices 3 and 4) which are relatively poorly defined by the family of 26 structures. While it is not surprising to see a different structure for the loop regions, which are susceptible to artificial ordering due to crystal-packing influences, the difference in the helical regions is unusual. Nevertheless,  $C\alpha$  chemical shift data in the NMR study of *T. thermophilus* S15 are entirely consistent with this region being one long continuous helix as in our X-ray structure (T Hård, personal communication), so the difference is probably a result of indeterminacy in the NMR data rather than true disorder in the *T. thermophilus* S15 solution structure. Loop LII is also smaller and more ordered in the crystal structure than it is in the NMR structure. It is difficult to say whether the less ordered loop in the NMR structure is due to indeterminacy in the data or whether it reflects true disorder. The 3 M phosphate in the crystal may well mimic the phosphate environment in the ribosome provided by the RNA and result in some stabilization of a portion of loop LII. An oft-quoted criticism of structural work on individual ribosomal proteins is that the conformations may differ within the ribosome [30], and this is likely to be especially true for loop regions that are involved in RNA-binding. Indeed, an unstructured loop in the C-terminal domain of L11 becomes structured on binding rRNA [31,32]. This may

Figure 5



Potential RNA-binding residues and sites on S15. **(a)** Ribbon diagram of S15 showing the putative RNA-binding residues; basic residues are shown in blue and hydrophobic in yellow. (This figure was made using MOLSCRIPT [46].) **(b)** The electrostatic surface potential of S15 shown in the same orientation as (a). The potential displayed represents a range of  $-12$  to  $+12$   $k_B T$ , shown with red as

negative and blue as positive. The surface potential calculation and display were carried out using the program GRASP [48]. The back view is a  $180^\circ$  rotation of the front view, while the alternate view is in the same orientation as the front view, but with the alternate conformation of helix 1.

well be the case for other small ribosomal proteins like S15, and it should be noted that the NMR structure of S17 also has two large disordered loops that probably become structured upon binding RNA [3,5].

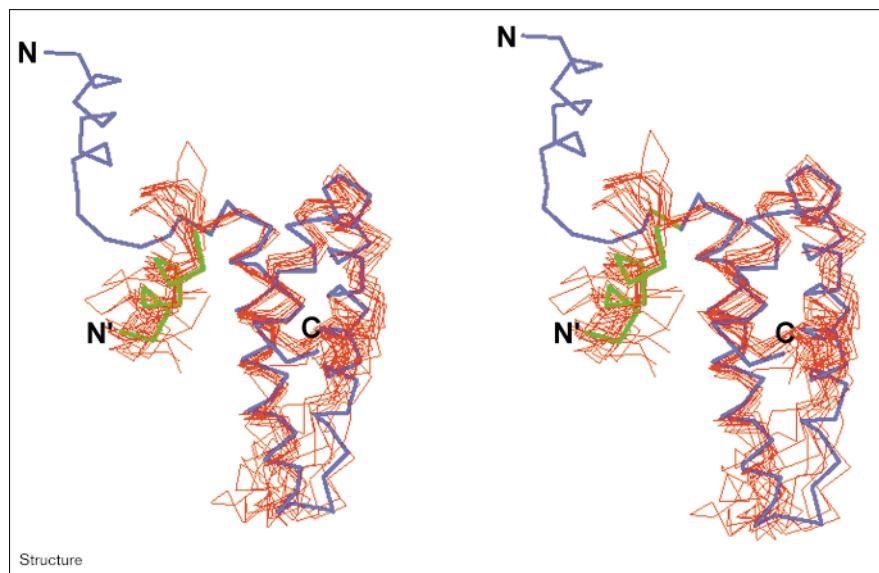
#### RNA binding implications

An extensive analysis of the S15-binding site on rRNA revealed that helix 22 and the three-way junction connecting helices 20, 21 and 22 are both crucial [16,17]. The identification of two putative RNA-binding sites at either end of S15 is consistent with the protein straddling these two areas with its basic face pointing towards the RNA. It is impossible to predict the orientation of the binding sites relative to the RNA but, with the structure in hand, this question can be answered using site-directed hydroxyl radical cleavage techniques to map the

details of protein–RNA contacts [33,34]. Computer modeling studies based on protection assays and gel-shift experiments have shown that the binding of S15 may cause the rRNA three-way helical junction to adopt a Y conformation in which helices 21 and 22 are almost parallel [35]. The basic face of S15 may cause the junction to bend around it, thereby promoting this conformation. Further research is needed to support this hypothesis.

What is the significance of the conformational variability of helix 1 and the connecting loop LI? Clearly, intramolecular contacts between helix 1 and the body of S15 have been replaced by intermolecular contacts in the crystal, in which helix 1 peels away from its own molecule and makes essentially equivalent contacts with a neighboring molecule. This variability could easily be dismissed as a

Figure 6



Stereo view comparison of the X-ray structure (this work) and the NMR structures [29] of S15. The diagram shows a superposition of the C $\alpha$  trace of the X-ray structure of S15 and the ten NMR structures from the ensemble (thin red lines) that are most similar to the X-ray structure. The X-ray structure is shown as a thick blue line, with the alternate conformation of helix 1 shown in green. (The least squares fits and transformations were calculated in the program O [43] and the figure was made using MOLSCRIPT [46].)

crystallization artifact, but one should also keep in mind more interesting possibilities. Firstly, the ability of the helix to peel off under conditions that may mimic an RNA environment suggests that, at least in principle, S15 could have the N-terminal helix dissociated from the body of the molecule when making contacts with RNA. Such a conformational change on binding RNA would be much more extensive than the simple ordering of exposed loops discussed earlier. However, just such a scenario occurs in the case of the binding of the spliceosomal protein U1A to its cognate RNA [36]. In the U1A protein alone, a C-terminal helix makes hydrophobic contacts with a  $\beta$  sheet that is known to be involved in RNA-binding [37]. In the presence of RNA, however, this helix swings out 135°, not only exposing the  $\beta$  sheet for interaction with RNA but also directly interacting with RNA itself [36]. In view of our finding of the conformational variability of the N-terminal helix of S15, an analogous mode of RNA binding must be kept open as a possibility. Secondly, it is known that S15 binds both to 16S rRNA during ribosome assembly and to its own mRNA as part of a translational feedback mechanism. It has been established that the mRNA-binding site adopts a pseudo-knot structure with little obvious structural similarity to the rRNA site [38]. It is therefore possible that S15 adopts two different conformations in binding to these two types of RNA, and that these different conformations are mediated by the flexibility of loop LI and consequently the variable orientation of helix 1. Finally, it is possible that helix 1 in its extended form makes a hydrophobic interaction with another ribosomal protein, possibly with its immediate neighbor S8 which contains a suitable hydrophobic patch [4].

### Biological implications

Proteins are essential components of all ribosomes and play a crucial role in the function of the modern ribosome. Primary RNA-binding proteins, such as S15, can bind directly to ribosomal RNA independently of other proteins and direct the subsequent folding of the complex. S15 is located in the core of the 30S (small) ribosomal subunit where it plays an important structural role. S15 is also able to regulate its own synthesis by binding to its own messenger RNA. This dual RNA-binding property of S15 makes it a good system to understand how the same protein can specifically recognize two different RNA molecules.

Here we describe the crystal structure of ribosomal protein S15 at 2.1 Å resolution. S15 is composed of three compact  $\alpha$  helices and an N-terminal helix that is variable in its orientation. S15 contains a large conserved basic patch at one end, and a cluster of conserved but exposed aromatic residues at the other end. These two sites are likely to be involved in RNA binding. The conformational variability of the N-terminal helix could play a role in RNA-binding. It is also possible that different conformations of S15 may be involved in the recognition of the two different types of RNA (ribosomal and messenger).

The S15 crystal structure will improve our understanding of ribosome structure in several ways. The structure can be used to design better biochemical probing experiments to identify protein-RNA contacts in the ribosome and with mRNA. Moreover, a high-resolution structure should also be useful in conjunction with

modest resolution electron microscopic structures of the 30S subunit in constraining and refining current models of the three-dimensional fold of RNA within the subunit.

## Materials and methods

### Protein purification

Degenerate primers based on the protein sequence were used to amplify the gene for *B. stearothermophilus* S15 from genomic DNA, using previously described methods [39]. The gene was introduced into the T7-based expression vector pET-13a, and four independent clones were sequenced. The gene was overexpressed in the *Escherichia coli* strain BL21 (DE3). Cells were grown to an OD<sub>600</sub> of approximately 0.8, induced with 0.4 mM isopropyl-β-D-thiogalactoside (IPTG) and harvested 3 h after induction. All purification buffers contained 0.05 M phenyl-methane sulfonyl fluoride (PMSF), 6 mM 2-mercaptoethanol, and all procedures were carried out at 0–4°C. The cells were resuspended in lysis buffer (50 mM Tris, 1 mM EDTA, pH 8.0) treated with 1 mg/ml lysozyme for 1 h and lysed with 0.1% Na-deoxycholate. The viscosity was reduced by treatment with DNaseI in the presence of 10 mM MgCl<sub>2</sub> and MnCl<sub>2</sub>. S15 was made completely soluble by raising the salt concentration of the lysate to 0.7 M NaCl. After a low speed spin for 30 min at 18K to remove cell debris, the supernatant was diluted to 0.3 M NaCl, filtered through a 0.45 micron filter, loaded onto a cation exchange column (Fractogel SO<sub>3</sub><sup>-</sup>, Merck) and eluted with a NaCl gradient from 0.5–1.5 M. Fractions containing S15 were pooled and diluted fivefold with 20 mM phosphate buffer pH 7.0 and then loaded onto a hydroxyapatite column (Biorad Macro-Prep Ceramic Hydroxyapatite). The protein was eluted using a 0.2–1.0 M NaCl gradient. The fractions were pooled, concentrated to about 20 mg/ml, mixed with an equal volume of glycerol and stored at –70°C. Prior to crystallization, the samples were thawed and passed through a gel filtration column (Superdex 75, Pharmacia). The peak containing S15 was essentially pure as determined by a Coomassie stained SDS PAGE gel. The protein was dialyzed against 5 mM MES and 6 mM 2-mercaptoethanol at pH 6.0 and then concentrated to 12 mg/ml. Crystallization trials were performed at room temperature (23°C) using the hanging-drop method by mixing 5 μl of protein solution with 5 μl of the reservoir.

### Data collection

All data were collected from a single, flash-cooled selenomethionyl crystal (see Results section) in a cryostream at 100K. Data at two of the wavelengths, λ<sub>2</sub> and λ<sub>3</sub>, corresponding to the 'white line' and a remote point of the K-edge of selenium, were collected on beamline X12-C at the NSLS at Brookhaven National Laboratory. These data were collected to 2.64 Å on a CCD detector with 1024 × 1024 pixels [40]. Data at a third wavelength, λ<sub>1</sub>, was subsequently collected on the same crystal using CuKα radiation from a Rigaku RU-200 rotating anode with focusing mirrors and an R-AXIS IV image plate system. All data were collected in images corresponding to 1 degree rotations, and were integrated and scaled using the programs DENZO and SCALEPACK [41]. The integrated, scaled intensities were reduced to structure factors, and scaled relative to each other using the programs TRUNCATE and SCALEIT from the CCP4 package [42].

### Phasing and refinement

The maximum likelihood program SHARP [25] was used for heavy-atom refinement and phasing. The solvent content was estimated at 38% based on the unit cell dimensions and total molecular weight, and this was used in density modification by solvent flattening, which was done using the program SOLOMON [26] by launching a script from within SHARP. The solvent-flattened map was used as the experimental map for model building using the program O [43]. Refinement was carried out using the λ<sub>1</sub> data set from 6.0–2.1 Å in the program X-PLOR [44]. The free R factor [45] was used to guide the refinement and no reflections were omitted based on the standard errors of the intensities.

### Accession numbers

Coordinates and structure factors have been deposited in the Brookhaven Protein Data Bank, with accession code LIA3, and can also be obtained by E-mail from the authors.

## Acknowledgements

We thank SE Gerchman for cloning and sequencing the gene for ribosomal protein S15, and RM Sweet, SJ Sclafani and BT Wimberly for help with data collection at the NSLS. This work was supported by NIH grant GM 44973 (to SWW and VR) and by a grant from the National Science Foundation and the Lucille B Markey Charitable Trust to the University of Utah. The diffraction facility at beamline X12-C at the NSLS at Brookhaven is supported by the United States Department of Energy Offices of Health and Environmental Research and of Basic Energy Sciences, and by the National Science Foundation.

## References

- Ramakrishnan, V., *et al.*, & White, S.W. (1995). Structures of prokaryotic ribosomal proteins: implications for RNA binding and evolution. *Biochem. Cell Biol.* **73**, 979-986.
- Liljas, A. & Garber, M. (1995). Ribosomal proteins and elongation factors. *Curr. Opin. Struct. Biol.* **5**, 721-727.
- Golden, B.L., Hoffman, D.W., Ramakrishnan, V. & White, S.W. (1993). Ribosomal protein S17: characterization of the three-dimensional structure by <sup>1</sup>H- and <sup>15</sup>N-NMR. *Biochemistry* **32**, 12812-12820.
- Davies, C., Ramakrishnan, V. & White, S.W. (1996). Structural evidence for specific S8–RNA and S8–protein interactions within the 30S ribosomal subunit: ribosomal protein S8 from *Bacillus stearothermophilus* at 1.9 Å resolution. *Structure* **4**, 1093-10104.
- Jaishree, T.N., Ramakrishnan, V. & White, S.W. (1996). Solution structure of prokaryotic ribosomal protein S17 by high-resolution NMR spectroscopy. *Biochemistry* **35**, 2845-2853.
- Wimberly, 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.
- Mueller, F., Stark, H., van Heel, M., Rinke-Appel, J. & Brimacombe, R. (1997). A new model for the three-dimensional folding of *Escherichia coli* 16 S ribosomal RNA. III. The topography of the functional centre. *J. Mol. Biol.* **271**, 566-587.
- Mueller, F. & Brimacombe, R. (1997). A new model for the three-dimensional folding of *Escherichia coli* 16 S ribosomal RNA. II. The RNA–protein interaction data. *J. Mol. Biol.* **271**, 545-565.
- Mueller, F. & Brimacombe, R. (1997). A new model for the three-dimensional folding of *Escherichia coli* 16 S ribosomal RNA. I. Fitting the RNA to a 3D electron microscopic map at 20 Å. *J. Mol. Biol.* **271**, 524-544.
- Capel, M.S., *et al.*, & Moore, P.B. (1987). A complete mapping of the proteins in the small ribosomal subunit of *Escherichia coli*. *Science* **238**, 1403-1406.
- Stöffler-Meilicke, M. & Stöffler, G. (1987). The topography of ribosomal proteins on the surface of the 30S subunit of *Escherichia coli*. *Biochimie* **69**, 1049-1064.
- Held, W.A., Ballou, B., Mizushima, S. & Nomura, M. (1974). Assembly mapping of 30S ribosomal proteins from *Escherichia coli*. Further studies. *J. Biol. Chem.* **249**, 3103-3111.
- Mizushima, S. & Nomura, M. (1970). Assembly mapping of 30S ribosomal proteins from *E. coli*. *Nature* **226**, 1214-1218.
- 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.
- Muller, R., Garrett, R.A. & Noller, H.F. (1979). The structure of the RNA binding site of ribosomal proteins S8 and S15. *J. Biol. Chem.* **254**, 3873-3878.
- Batey, R. & Williamson, J. (1996). Interaction of the *Bacillus stearothermophilus* ribosomal protein S15 with 16 S rRNA: I. Defining the minimal RNA site. *J. Mol. Biol.* **261**, 536-549.
- Serganov, A.A., *et al.*, & Ehresmann, C. (1996). The 16S rRNA binding site of *Thermus thermophilus* ribosomal protein S15: comparison with *Escherichia coli* S15, minimum site and structure. *RNA* **2**, 1124-1138.
- Portier, C., *et al.*, & Ehresmann, C. (1990). Translational control of ribosomal protein S15. *Biochim. Biophys. Acta* **1050**, 328-336.
- Portier, C., Dondon, L. & Grunberg-Manago, M. (1990). Translational autocontrol of the *Escherichia coli* ribosomal protein S15. *J. Mol. Biol.* **211**, 407-414.
- Zengel, J.M. & Lindahl, L. (1994). Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*. *Prog. Nucleic Acid Res. Mol. Biol.* **47**, 331-370.

21. Philippe, C., Benard, L., Portier, C., Westhof, E., Ehresmann, B. & Ehresmann, C. (1995). Molecular dissection of the pseudoknot governing the translational regulation of *Escherichia coli* ribosomal protein S15. *Nucleic Acids Res.* **23**, 18-28.
22. Philippe, C., Eyermann, F., Benard, L., Portier, C., Ehresmann, B. & Ehresmann, C. (1993). Ribosomal protein S15 from *Escherichia coli* modulates its own translation by trapping the ribosome on the mRNA initiation loading site. *Proc. Natl. Acad. Sci. USA* **90**, 4394-4398.
23. Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. & Sweet, R.M. (1993). Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* **362**, 219-223.
24. Ramakrishnan, V. & Biou, V. (1997). Treatment of multiwavelength anomalous diffraction data as a special case of multiple isomorphous replacement. In *Methods in Enzymology Vol. 276*. (Carter, C.W., Jr. & Sweet, R.M., eds), pp. 538-557, Academic Press, New York, NY.
25. de la Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In *Methods in Enzymology Vol. 276*. (Carter, C.W., Jr. & Sweet, R.M., eds), pp. 472-493, Academic Press, New York, NY.
26. Abrahams, J.P. & Leslie, A.G.W. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Cryst. D* **52**, 30-42.
27. 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. Cryst.* **26**, 283-291.
28. Richardson, J.S. & Richardson, D.C. (1988). Amino acid preferences for specific locations at the ends of alpha helices. *Science* **240**, 1648-1652.
29. Berglund, H., Rak, A., Serganov, A., Garber, M. & Härd, T. (1997). Solution structure of the ribosomal RNA binding protein S15 from *Thermus thermophilus*. *Nat. Struct. Biol.* **4**, 20-23.
30. Yonath, A. & Franceschi, F. (1997). New RNA recognition features revealed in ancient ribosomal proteins. *Nat. Struct. Biol.* **4**, 3-5.
31. Markus, M.A., Hinck, A.P., Huang, S., Draper, D.E. & Torchia, D.A. (1997). High resolution solution structure of ribosomal protein L11-C76, a helical protein with a flexible loop that becomes structured upon binding to RNA. *Nat. Struct. Biol.* **4**, 70-77.
32. Hinck, A.P., *et al.*, & Torchia, D.A. (1997). The RNA binding domain of ribosomal protein L11: three-dimensional structure of the RNA-bound form of the protein and its interaction with 23 S rRNA. *J. Mol. Biol.* **274**, 101-113.
33. Heilek, G.M. & Noller, H.F. (1996). Site-directed hydroxyl radical probing of the rRNA neighborhood of ribosomal protein S5. *Science* **272**, 1659-1662.
34. Heilek, G.M. & Noller, H.F. (1996). Directed hydroxyl radical probing of the rRNA neighborhood of ribosomal protein S13 using tethered Fe(II). *RNA* **2**, 597-602.
35. Batey, R.T. & Williamson, J.R. (1996). Interaction of the *Bacillus stearothermophilus* ribosomal protein S15 with 16S rRNA: II. Specificity determinants of RNA-protein recognition. *J. Mol. Biol.* **261**, 550-567.
36. Avis, J.M., Allain, F.H., Howe, P.W., Varani, G., Nagai, K. & Neuhaus, D. (1996). Solution structure of the N-terminal RNP domain of U1A protein: the role of C-terminal residues in structure stability and RNA binding. *J. Mol. Biol.* **257**, 398-411.
37. 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.
38. Ehresmann, C., Philippe, C., Westhof, E., Benard, L., Portier, C. & Ehresmann, B. (1995). A pseudoknot is required for efficient translational initiation and regulation of the *Escherichia coli* rpsO gene coding for ribosomal protein S15. *Biochem. Cell Biol.* **73**, 1131-1140.
39. Ramakrishnan, V. & Gerchman, S.E. (1991). Cloning, sequencing and overexpression of genes for ribosomal proteins from *Bacillus stearothermophilus*. *J. Biol. Chem.* **266**, 880-885.
40. Stanton, M., Phillips, W.C. & O'Mara, D. (1994). CCD-based detector for X-ray crystallography. *Proc. Soc. Photo-Opt. Instr. Eng.* **2278**, 16-20.
41. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology Vol. 276*. (Carter, C.W., Jr. & Sweet, R.M., eds), pp. 307-325, Academic Press, New York, NY.
42. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760-763.
43. 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 Cryst. A* **47**, 110-119.
44. Brünger, A.T. (1988). Crystallographic refinement by simulated annealing. Application to a 2.8 Å structure of aspartate aminotransferase. *J. Mol. Biol.* **203**, 803-816.
45. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.
46. Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950.
47. Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881-10890.
48. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281-296.