

# RNA

## Detailed analysis of RNA-protein interactions within the bacterial ribosomal protein L5/5S rRNA complex

A. Perederina, N. Nevskaya, O. Nikonov, A. Nikulin, P. Dumas, M. Yao, I. Tanaka, M. Garber, G. Gongadze and S. Nikonov

*RNA* 2002 8: 1548-1557

---

### References

Article cited in:

<http://www.rnajournal.org/cgi/content/abstract/8/12/1548#otherarticles>

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

---

### Notes

---

To subscribe to *RNA* go to:  
<http://www.rnajournal.org/subscriptions/>

---

# Detailed analysis of RNA–protein interactions within the bacterial ribosomal protein L5/5S rRNA complex

ANNA PEREDERINA,<sup>1</sup> NATALIA NEVSKAYA,<sup>1</sup> OLEG NIKONOV,<sup>1</sup> ALEXEI NIKULIN,<sup>1</sup>  
 PHILIPPE DUMAS,<sup>2</sup> MIN YAO,<sup>3</sup> ISAO TANAKA,<sup>3</sup> MARIA GARBER,<sup>1</sup> GEORGE GONGADZE,<sup>1</sup>  
 and STANISLAV NIKONOV<sup>1</sup>

<sup>1</sup>Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Moscow region, Russia

<sup>2</sup>Unité Propre de Recherche 902 du Centre National de la Recherche Scientifique,

Institut de Biologie Moléculaire et Cellulaire, 67084 Strasbourg-cedex, France

<sup>3</sup>Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan

## ABSTRACT

The crystal structure of ribosomal protein L5 from *Thermus thermophilus* complexed with a 34-nt fragment comprising helix III and loop C of *Escherichia coli* 5S rRNA has been determined at 2.5 Å resolution. The protein specifically interacts with the bulged nucleotides at the top of loop C of 5S rRNA. The rRNA and protein contact surfaces are strongly stabilized by intramolecular interactions. Charged and polar atoms forming the network of conserved intermolecular hydrogen bonds are located in two narrow planar parallel layers belonging to the protein and rRNA, respectively. The regions, including these atoms conserved in Bacteria and Archaea, can be considered an RNA–protein recognition module. Comparison of the *T. thermophilus* L5 structure in the RNA-bound form with the isolated *Bacillus stearothermophilus* L5 structure shows that the RNA-recognition module on the protein surface does not undergo significant changes upon RNA binding. In the crystal of the complex, the protein interacts with another RNA molecule in the asymmetric unit through the  $\beta$ -sheet concave surface. This protein/RNA interface simulates the interaction of L5 with 23S rRNA observed in the *Haloarcula marismortui* 50S ribosomal subunit.

**Keywords:** 5S rRNA; ribosomal protein L5; RNA–protein interactions; *Thermus thermophilus*; X-ray structure

## INTRODUCTION

5S rRNA complexed with several proteins is an integral part of the large ribosomal subunit. Although 5S rRNA has been the subject of numerous biochemical, chemical, and physical studies, its exact function remains unknown (Noller, 1998). Phylogenetic analysis, chemical and enzymatic probing, and computer modeling have yielded a highly conserved 5S rRNA secondary structure (Westhof et al., 1989; Brunel et al., 1991; Leontis & Westhof, 1998; Szymanski et al., 2002). At present, its tertiary structure is known from the models of archaeal (*Haloarcula marismortui*) and bacterial (*Deinococcus radiodurans*) 50S ribosomal subunits, determined at 2.4 Å (Ban et al., 2000) and 3.1 Å (Harms et al., 2001) resolution, respectively. In contrast to the significant similarity between the RNA folds in the *D. radiodurans* and *H. marismortui* 50S ribosomal sub-

units, the protein folds show remarkable differences (Harms et al., 2001). As regards the 5S rRNA/protein complex, there is a significant difference between ribosomal proteins binding 5S rRNA in Bacteria and Archaea. Only two of the three bacterial ribosomal proteins specifically binding 5S rRNA, L18 and L5, have homologs in Archaea. In the 50S ribosomal subunit from *H. marismortui*, protein L5 (HmaL5) was found near tRNA in the P-site (Nissen et al., 2000). About 80% of the HmaL5 atomic coordinates are now available. The C $\alpha$  coordinates of bacterial L5 within the 50S ribosomal subunit from *D. radiodurans* (DraL5) at 3.1 Å resolution (Harms et al., 2001) and the complete crystal structure of the isolated bacterial L5 from *Bacillus stearothermophilus* (BstL5) at 1.8 Å resolution (Nakashima et al., 2001) have recently been determined. In both Archaea and Bacteria, L5 binds to the same site on 5S rRNA, located at the so-called loop C (Ban et al., 2000; Gongadze et al., 2001).

In this work, we have studied a complex of protein L5 from *Thermus thermophilus* (TthL5) with a specific 34-nt fragment containing helix III and loop C of 5S rRNA

Reprint requests to: Stanislav Nikonov, Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Moscow region, Russia; e-mail: nikonov@vega.protres.ru.

from *Escherichia coli*. TthL5 has 39.6% similarity with its archaeal homolog from *H. marismortui*, and 75.3% and 82.4% similarity with the bacterial L5 proteins from *B. stearothermophilus* and *D. radiodurans*, correspondingly. The sequence of the 5S rRNA fragment used in this work is highly conserved among Bacteria and Archaea. Determination of the TthL5/RNA complex structure allowed comparison between the bacterial and archaeal L5/5S rRNA interfaces and detailed analysis of the peculiarities of RNA–protein interactions crucial for specific RNA–protein recognition and for the complex stability.

## RESULTS

### Structure determination

TthL5 binds efficiently with *T. thermophilus* and *E. coli* 5S rRNAs. Suitable crystals were grown for TthL5 complexed with a 29-nt fragment of *E. coli* 5S rRNA with several extra nucleotides (C, C, C and G, G) added at the 3' and 5' ends of the RNA, respectively, to stabilize

its double helix and to improve crystal quality. Two data sets collected at 2.8 Å and 2.5 Å resolution from crystals containing selenomethionines (SeMet) were used to solve the structure. Both crystals belonged to space group P2<sub>1</sub>, but their cell parameters  $a = 50.93$  Å,  $b = 95.20$  Å,  $c = 62.80$  Å,  $\beta = 102.81$  deg and  $a = 51.89$  Å,  $b = 99.60$  Å,  $c = 65.73$  Å,  $\beta = 100.05$  deg were significantly different. Data statistics are summarized in Table 1. The first data set had a good anomalous contribution at the peak of the white line and was used to find positions of Se atoms and calculate initial phases by the single anomalous dispersion method (SAD). Four Se atoms were attributed to two L5/5S rRNA complexes in the asymmetric unit of the crystal. The initial phases ( $m = 0.265$ ) were improved by density modification and phase extension technique ( $m = 0.900$ ). The resulting electron density map was of a workable quality and allowed us to build about 85% of the two complexes in the asymmetric unit. The structure of this dimer was refined and used as a model to calculate phases for the second data set by the molecular replacement method (MR), which yielded an unambiguous solution with a correlation coefficient of 62.9% and

**TABLE 1.** Data-collection phasing and refinement statistics.

	Data set (1) SeMet	Data set (2) SeMet
Space group	P2 <sub>1</sub>	P2 <sub>1</sub>
Unit-cell parameters (Å, °)	$a = 50.93$ , $b = 95.20$ , $c = 62.80$ , $\alpha = \gamma = 90$ , $\beta = 102.81$	$a = 51.89$ , $b = 99.60$ , $c = 65.73$ , $\alpha = \gamma = 90$ , $\beta = 100.05$
Wavelength (Å)	0.9792	0.9680
Resolution (Å)	20–2.8 (2.87–2.80)	15–2.5 (2.59–2.50)
Unique reflections	14,431 (1,248)	22,852 (1,908)
Completeness (%) <sup>a</sup>	99.9 (99.9)	99.9 (84.1)
Averaged redundancy	6.2 (4.9)	3.02 (3.0)
$I/\sigma_i(I)$ <sup>a</sup>	8.3 (2.9)	13.2 (3.3)
$R_{sym}(I)$ (%) <sup>a,b</sup>	7.1 (24.2)	6.7 (18.6)
Phasing statistics		
Number of Se atoms per complex	2	
Phasing power (centric/acentric) <sup>c</sup>	1.45/1.40	
$R_{Cullis}$ (centric/acentric) <sup>d</sup>	0.60/0.60	
Figure of merit	0.265	
Overall figure of merit at 20–2.8 Å resolution (after density modification):	0.900	
Refinement statistics		
Resolution range (Å)	15–2.5	
Reflections	21,043	
$R$ -factor (%)	21.0	
Free $R$ -factor (%) <sup>e</sup>	28.8	
R.m.s. deviation		
Bond lengths (Å)	0.007	
Bond angles (deg)	1.31	
Improper angles (deg)	1.39	
Average $B$ -factor (Å <sup>2</sup> )	39.4	

<sup>a</sup>Values in parenthesis are statistics for the highest resolution shell.

<sup>b</sup> $R_{sym} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ .

*R*-factor of 46.8%. The reason for the nonisomorphism was found to lie in a 10 deg rotation of the dimer. NCS restraints were used during the early stages of refinement, but the two molecules in the asymmetric unit were finally refined separately to an *R*-factor of 21.0%, and a free *R*-factor of 28.8% for the data between 15 Å and 2.5 Å. Phase and refinement statistics are shown in Table 1. Four Mg<sup>2+</sup>, two K<sup>+</sup> ions, and 187 water molecules were localized in the later stages of refinement. Although the resulting map was of a good quality, some parts of long loops had a weak electron density and no electron density was observed for the N termini of either TthL5 molecule in the asymmetric unit.

A ribbon diagram of the complex is presented in Figure 1. The protein and RNA molecules make a  $\Lambda$ -shaped structure with a small contact area (580 Å<sup>2</sup>) between them. In the crystal, two complexes form a tightly associated dimer with approximate dimensions of 87 × 55 × 60 Å.

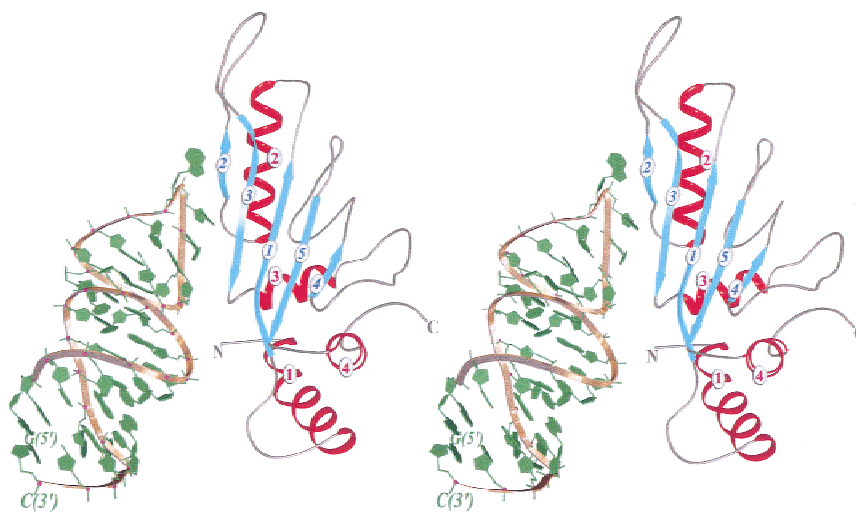
### Structure of TthL5

TthL5 consists of 182 amino acid residues and folds into one  $\alpha/\beta$  domain with the connectivity scheme  $\alpha_1\beta_1\alpha_2\beta_2\beta_3\alpha_3\beta_4\beta_5\alpha_4$  (Figs. 1 and 2). The entire structure is well ordered, with the exception of the N terminus (residues 1–4) and loop  $\beta_2\text{--}\beta_3$ , which seem to be significantly flexible and exhibit a weak electron density. The molecule contains four  $\alpha$ -helices packed on one side of a highly twisted five-stranded antiparallel  $\beta$ -sheet, which is exposed on its other side. This  $\beta$ -sheet has a concave surface and is extended by two loops ( $\alpha_1\text{--}\beta_1$  and  $\beta_2\text{--}\beta_3$ ). The concave shape of the  $\beta$ -sheet is strongly stabilized by an extended hydrophobic core, formed by about 30 residues of the  $\beta$ -sheet,  $\alpha$ -helices, and loops. Helix  $\alpha_1$  is quite separated from the globular part of the molecule and is associated with it by hydrophobic interactions as well as by hydrogen bonds. The

relative position of strands  $\beta_4$  and  $\beta_5$  is additionally stabilized by a Mg<sup>2+</sup> ion in the crystal. This ion is coordinated by protein and solvent groups with five ligands in the first coordination sphere. The ligands are: O atoms of three main chain carbonyl groups and O atoms of an asparagine side chain and of one water molecule. About half of the main chain atoms are involved in the eight loops, some of which contain 3<sub>10</sub> helices. Loops  $\alpha_3\text{--}\beta_4$  and  $\beta_4\text{--}\beta_5$  contain more than 20 residues each. The TthL5 molecule contains 25 negatively and 32 positively charged residues, some of which form a network of hydrogen bonds and salt bridges contributing to the stabilization of the structure.

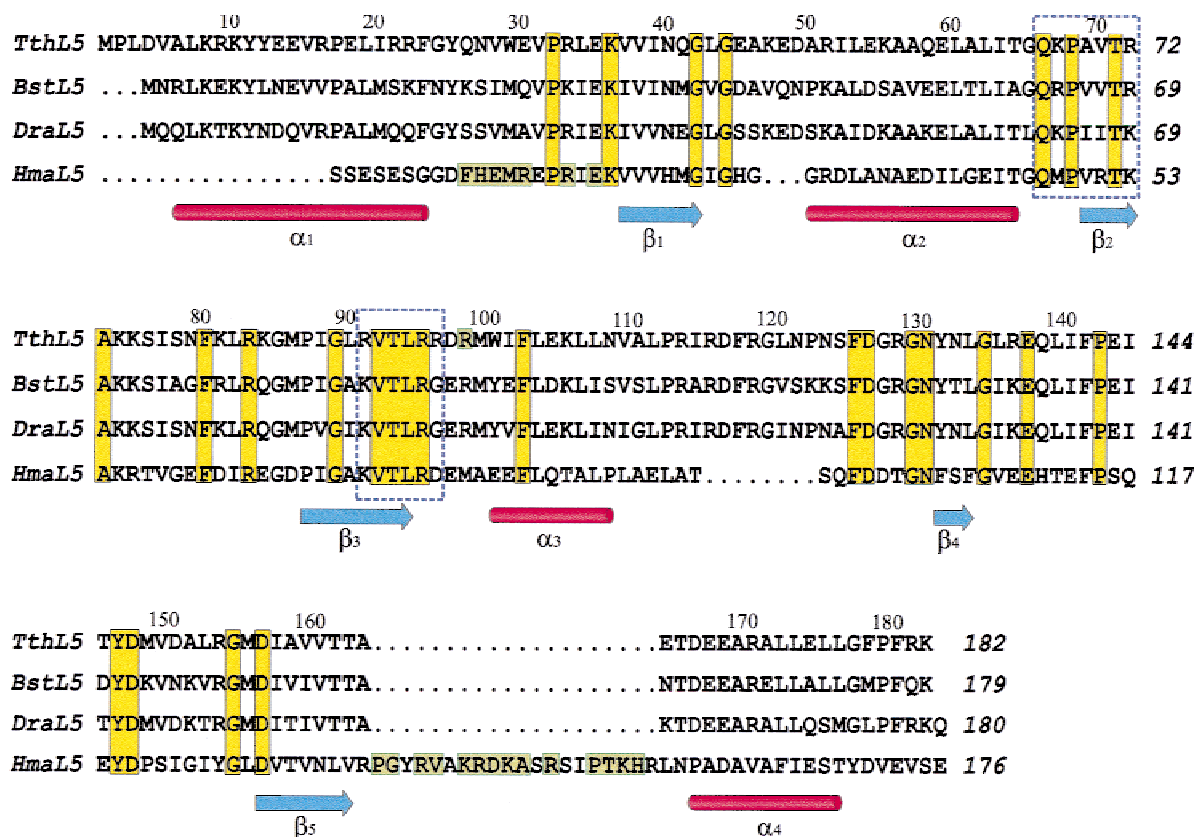
The two molecules in the asymmetric unit of the crystal have slightly different conformations. The main difference between them is in the regions of loops  $\beta_1\text{--}\alpha_2$ ,  $\beta_2\text{--}\beta_3$ , and  $\beta_4\text{--}\beta_5$ . The latter is disordered in one molecule whereas it forms a  $\beta$ -hairpin in another. Superimposition of the two molecules yields an r.m.s. deviation of 2.3 Å for all C <sub>$\alpha$</sub>  atoms and only 0.9 Å without three loops.

The topology of TthL5 is similar to that of the isolated BstL5 (Nakashima et al., 2001) as well as to that of DraL5 (Harms et al., 2001) and HmaL5 (Ban et al., 2000) from the models of the 50S ribosomal subunits. All these proteins have several long loops with very flexible structures. The main differences among bacterial L5 proteins are found in the N terminus, and in loops  $\beta_1\text{--}\alpha_2$ ,  $\beta_2\text{--}\beta_3$ , and  $\beta_4\text{--}\beta_5$ . When optimally superimposed, the overall r.m.s. deviation in C <sub>$\alpha$</sub>  positions is about 3.5 Å, when all residues are included, but only 2.0 Å, when the N terminus and these three loops are excluded. No electron density was observed for protein HmaL5 regions corresponding to helix  $\alpha_1$  and loops  $\beta_1\text{--}\alpha_2$  and  $\beta_4\text{--}\beta_5$  (Ban et al., 2000). Comparison of HmaL5 with the known bacterial L5 structures reveals that archaeal L5 has shortened loop  $\alpha_3\text{--}\beta_4$  and long extra loop  $\beta_5\text{--}\alpha_4$ .



**FIGURE 1.** Stereo ribbon representation of the TthL5/5S rRNA complex. The phosphoribose backbone is in gold, bases are in green,  $\beta$ -strands in blue, and  $\alpha$ -helices in red.

## Structure of TthL5/5S rRNA complex



**FIGURE 2.** Sequence alignment of ribosomal proteins L5 based on their tertiary structures. The sequences are: TthL5: *T. thermophilus*, BstL5: *B. stearothermophilus*, DraL5: *D. radiodurans*, HmaL5: *H. marismortui*. The numbering corresponds to TthL5. Positions of  $\alpha$ -helices and  $\beta$ -strands in TthL5 are shown. The background is yellow for residues strictly conserved (>95%) in all known sequences of bacterial and archaeal L5. Residues involved in L5/5S rRNA interfaces in both TthL5/5S rRNA complex and 50S ribosomal subunit of *H. marismortui* are boxed with dotted lines. The background is green for residues involved in RNA-protein interaction in only one of the two structures.

### Structure of the 5S rRNA fragment

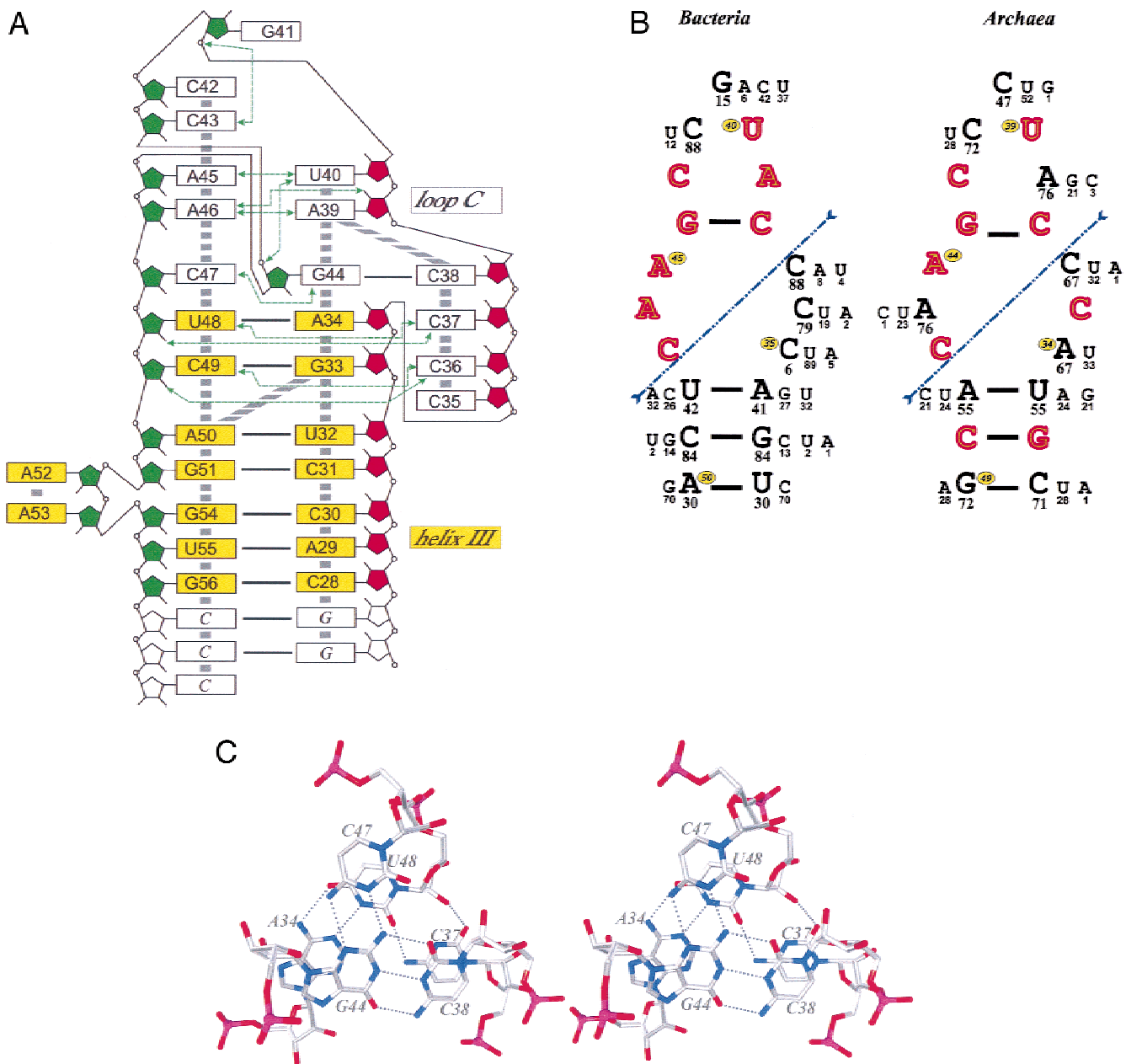
The fragment of *E. coli* 5S rRNA (nt 28–56), used to obtain the protein/RNA complex, includes loop C and helix III (Fig. 3A). The sequence of this fragment is rather conserved in 5S rRNAs among Bacteria and Archaea (Fig. 3B). With the exception of 2 nt (A52 and A53) which are bulged out, helix III displays a regular double-helical structure terminated by the triple C37•(U48-A34) with *cis* Watson–Crick/sugar edge base pair C37•U48 (according to the nomenclature of Leontis & Westhof, 2001). Another triple (C38-G44)•C47 contains *trans* Watson–Crick/sugar edge base pair G44•C47 (Fig. 3C). Watson–Crick base pair C38-G44 rotated approximately by 90 deg with respect to U48-A34 forms the bottom of the “pyramid” involving *cis* Watson–Crick/sugar edge A39•A46 and *trans* Hoogsteen/sugar edge (sheared) U40•A45 base pairs and three nonpaired nucleotides (C43, C42, and G41). Thus, the pyramidal structure is built up by nucleotides of loop C with the exception of C35, C36, and C37, which are bulged out to form two triples C36•(C49-G33) and C37•(U48-A34) with *cis* Watson–Crick/sugar edge base pairs. Almost all nucleotides of the pyramid

are highly conserved in Bacteria and Archaea (Fig. 3B). The positions of bases are stabilized by the intra-molecular interactions, the number of which decreases from the bottom to the top of this structure. It is noticeable that the deep groove of A-helix degenerates on the pyramid surface whereas the shallow groove is retained and terminated in a flat exposed platform containing the phosphor-ribose backbone of G41, C42, C43, and A45 and their base edges. The two RNA molecules in the asymmetric unit reveal a high similarity of the pyramidal structures with an r.m.s. deviation of 0.2 Å between P atoms, whereas that for the whole molecules yields 0.6 Å.

### RNA-protein interactions

5S rRNA contacts TthL5 by nucleotides G41, C42, C43, G44, and A45 of the pyramid, the structure of which is strongly stabilized by a network of conserved hydrogen bonds and stacking interactions (Fig. 3A). The interface region of the protein is formed by the main chain and side chain atoms of the residues belonging to strands  $\beta_2$ ,  $\beta_3$ , and loops  $\alpha_2$ - $\beta_2$ ,  $\beta_3$ - $\alpha_3$ . This region is strongly stabilized





**FIGURE 3.** **A:** Schematic illustration of the tertiary structure of the 34-nt 5S rRNA fragment. Additional nucleotides at 5' and 3' ends are in italics. Noncanonical H-bonds are shown with green arrows. Stacking interactions are shown in gray. **B:** The numbering (yellow background) and the sequences correspond to 5S rRNA from *E. coli* in Bacteria and to 5S rRNA from *H. marismortui* in Archaea. The frequency of occurrence of nucleotides in the loop C and adjacent part of helix III in bacterial and archaeal species (Szymanski et al., 2002) are shown. Nucleotides with frequency more than 95% are red. The "pyramid" structure is underlined. **C:** Detailed stereo view of triples at the top of helix III and at the bottom of the pyramid.

by intramolecular interactions involving strictly conserved residues Gln66, Pro68, Val92, Thr93, Leu94, and Arg95. The side chain location of Gln66 and Thr93 is stabilized by hydrogen bonds, whereas the side chain of Arg95 forms a salt bridge with Glu35.

Nonpolar ribose and base atoms of G41, C42, C43, and A45 as well as nonpolar protein atoms of residues Gln66, Lys67, Ala69, Val70, Thr71, Arg72, Arg91, Val92, Thr93, Arg95, Arg96, and Arg98 form an intermolecular hydrophobic core. The TthL5/5S rRNA interface is also

stabilized by a network of hydrogen bonds. Ten amino acid residues are hydrogen bonded with 6 nt of 5S rRNA mainly by the side chain atoms (Table 2). Bases of 3 nt, G41, C42, and C43, are involved in these interactions. Additionally the base of C42 stacks on the side chain of Arg91. It is important to note that C42 is involved as an essential part in all interactions with the protein and base-specific interactions between C42 and Thr93 mimic well interactions between two bases in an RNA double helix structure.

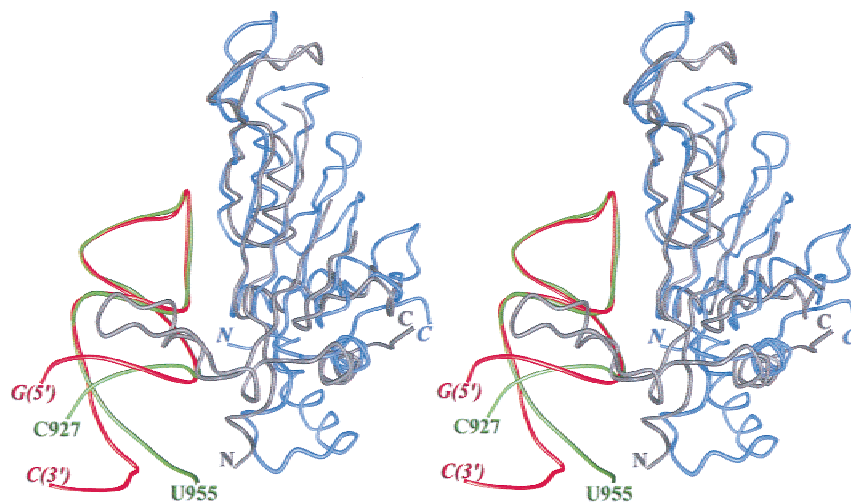
**TABLE 2.** Comparison of protein–RNA hydrogen bonds in the TthL5/5S rRNA complex and HmaL5/5S rRNA within the 50S ribosomal subunit from *H. marismortui*. Conserved hydrogen bonds are highlighted in bold.

TthL5/5SrRNA			HmaL5/5SrRNA*			TthL5/5SrRNA			HmaL5/5SrRNA*		
			Glu12	OE1–O2'	G53	Arg98	NH2–O1P	G44	Pro137	O–O2'	C29
				OE2–O2'					Gly138	CA–O2'	C29
				OE1–C5'	A54				Arg140	CD–O1P	C29
			Met13	CE–O2	C29					NE–O1P	
			Arg14	CD–O2'	A56					CG–O1P	A45
			Arg17	NH1–O1P	C30					NH1–O1P	
<b>Gln66</b>	<b>NE2–O2'</b>	<b>C42</b>	<b>Gln47</b>	<b>NE2–O2'</b>	<b>C41</b>					NH1–O2P	C46
	<b>NE2–O4'</b>	<b>C43</b>		<b>NE2–O4'</b>	<b>C42</b>					NH2–O2P	
<b>Lys67</b>	<b>O–O2'</b>	<b>C42</b>	<b>Met48</b>	<b>O–O2'</b>	<b>C41</b>				Lys143	O–C4'	A45
	<b>O–C1'</b>			<b>O–C1'</b>					Arg144	CA–O1P	C46
	NZ–O1P									NH2–O2'	G58
Ala69	CB–N7	G41	Val50	CG2–O4'	C41				Asp145	N–O1P	C46
	CB–O4'	C42								OD1–C5'	
Val70	O–N1	G41							Arg149	CD–O3'	G58
Thr71	CA–O6	G41								CD–O1P	C59
Arg91	NH2–O6	G41								NH1–O1P	
	NH2–N7								Pro152	CD–O2'	A57
<b>Thr93</b>	<b>N–O2</b>	<b>C42</b>	<b>Thr74</b>	<b>N–O2</b>	<b>C41</b>					CG–O2'	
	<b>OG1–N3</b>			<b>OG1–N3</b>					Thr153	OG1–O1P	G17
	<b>OG1–O2</b>			<b>OG1–O2</b>					Lys154	CE–O2P	U18
<b>Arg95</b>	<b>NE–O4'</b>	<b>A45</b>	<b>Arg76</b>	<b>NE–O4'</b>	<b>A44</b>					NZ–O2P	G17
	<b>NH2–O4'</b>			<b>NH2–O4'</b>					His155	ND1–O2'	A56
	<b>NE–O2</b>	<b>C43</b>		<b>NE–O2</b>	<b>C42</b>					CE1–O2'	
	CD–O2'			NH1–O2						ND1–O4'	A57
Arg96	CD–O1P	A45								NE2–O4'	
	NE–O1P										
	NH2–O2P										
	NH1–O2P	G33									

\*The numbering corresponds to 50S ribosomal subunit from *H. marismortui*.

Protein/RNA interfaces of two complexes in the asymmetric unit of the crystal are the same whereas the comparison of the TthL5/5S rRNA structure with the corresponding part of the *H. marismortui* 50S ribosomal subunit reveals differences in the RNA–protein binding. Thus, the extra loop of the archaeal L5 protein is additionally involved in interaction with 5S rRNA

(Table 2; Fig. 4), forming half of all RNA–protein H-bonds and interacting mainly with the double-helical stem of the 5S rRNA fragment. As a result, the solvent inaccessible region formed upon complex formation is more than twice less in the TthL5/5S rRNA complex (1,160 Å<sup>2</sup>) than in the HmaL5/5S rRNA one (2,690 Å<sup>2</sup>).

**FIGURE 4.** Stereo view of the comparison between the structures of the TthL5/RNA complex and its counterpart in the 50S ribosomal subunit from *H. marismortui*. The protein and RNA backbones from the present structure are in blue and red, respectively. The protein and RNA backbones of *H. marismortui* are in gray and green, respectively.

In the crystal, TthL5 interacts also with another RNA molecule in the asymmetric unit. This protein/RNA interface is formed by the  $\beta$ -sheet concave surface of the protein and the shallow groove of the double-helical RNA. An extended hydrophobic core stabilizes this interface. Hydrogen bonds between Lys36, Asn40, Asp126, Asn132, Arg153, Asp156, and G27, C30, C28, U55, G56, C58 contribute to this stabilization. Both surfaces of the protein/RNA interface show shape complementarity.

## DISCUSSION

At present, three models of the 50S ribosomal subunits from three different sources and at different resolutions are available (Ban et al., 2000; Harms et al., 2001; Yusupov et al., 2001). The conformations of the C loop pyramidal parts of bacterial and archaeal 5S rRNAs are essentially identical, and it is in agreement with conservation of the sequences (Fig. 3B). As for the L5 protein, only the *H. marismortui* 50S subunit model contains the coordinates of side chain atoms of the protein and can thus be used for a detailed comparative analysis of the protein/RNA interfaces.

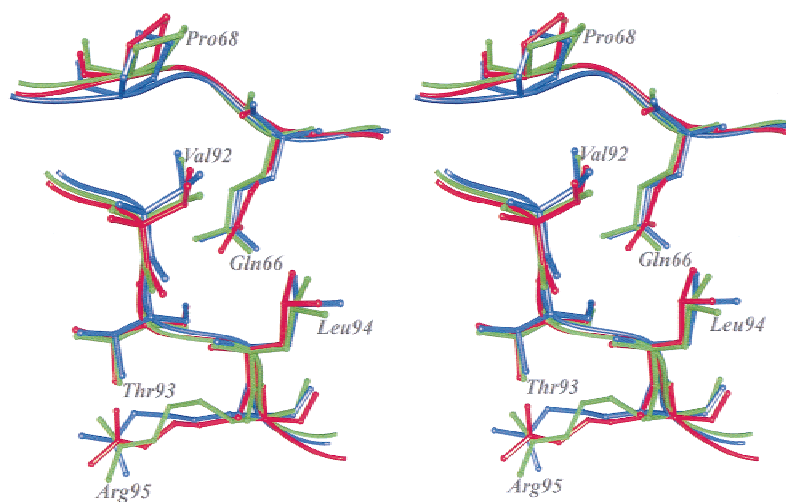
### Recognition modules on the protein and RNA surfaces

It is known that many ribosomal proteins are structurally and functionally interchangeable within the ribosomes from different species (Gourse et al., 1981). This is also true for 5S rRNA binding proteins. Thus, hybrid protein/5S rRNA complexes can be reconstituted from *B. stearotherophilus* and *E. coli* 5S rRNAs and ribosomal proteins (Horne & Erdmann, 1972). Protein TthL5 binds specifically to *T. thermophilus* and *E. coli* 5S rRNAs with similar affinities (Gongadze et al., 1993, 2001). Moreover, the reconstitution experiments (Erdmann et al., 1980) revealed that different bacterial and

archaeal 5S rRNAs could be incorporated into biologically active *B. stearotherophilus* 50S ribosomal subunits. This implies that RNA-recognition modules on the proteins and protein-recognition modules on the corresponding RNA molecules are structurally highly conserved. Upon complex formation, some additional interactions, which are not essential for recognition but are significant for complex stabilization, can be induced. Regions involved in these additional interactions together with structurally conserved recognition modules form specific binding sites on the protein and RNA surfaces. This complicates the structure of the protein/RNA interface and makes isolation of the recognition module more difficult.

We suggest that RNA and protein molecules recognize their specific sites through the atoms forming a conserved network of RNA–protein hydrogen bonds and the recognition modules include atoms, relative positions of which are very close in protein/RNA complexes from different sources. To search for the recognition modules, the structures of the TthL5/5S rRNA complex and the corresponding part of the *H. marismortui* 50S ribosomal subunit were superimposed and the networks of protein–RNA hydrogen bonds were compared (Table 2). Superposition of TthL5, HmaL5, and BstL5 shows that the recognition module atoms have very close positions in the RNA-bound and in the isolated L5 proteins (Fig. 5). Atoms of the protein- and RNA-recognition modules form rigid body structures on two corresponding approximately parallel planar narrow layers. A similar arrangement of interacting atoms was earlier observed in the CTC/5S rRNA and S8/16S rRNA complexes (Fedorov et al., 2001; Tishchenko et al., 2001). Such planar arrangement of contacting atoms probably facilitates specific recognition.

Two of the 3 nt forming the recognition module on the RNA are strictly conserved, whereas C42 is replaced by U in some organisms (Fig. 3B). The relatively fre-



**FIGURE 5.** Superposition of RNA-recognition modules from TthL5 (green), HmaL5 (red), and BstL5 (blue).



quent replacement of C42 by U is accounted for by the fact that OG1 of Thr93 can serve as a donor or as an acceptor in hydrogen bonding with the N3 of C or U, respectively. As a result, this replacement does not affect the conserved network of RNA–protein hydrogen bonds.

Because strictly conserved Thr93 and Arg95 of the RNA-recognition module are not involved in the L5 structure stabilization, mutation of these residues should not change the protein conformation, but should disturb the conserved network of H-bonds with the RNA. This was indeed observed with a Thr93Ala mutation on BstL5 not affecting the stability of the protein but resulting in a 4.2-fold decrease in the binding affinity (Nakashima et al., 2001). A detailed analysis shows that the replacement of Arg95 could result in a lack of recognition and binding. One more conserved residue, Gln66, is involved in hydrogen bonding with C42 in the TthL5/5S rRNA complex. Moreover this residue stabilizes the structure of the protein-recognition module through the H-bond with the main chain of Thr93, and retains its position in the structure of the isolated protein L5 from *B. stearothermophilus* (Fig. 5). The replacement of Gln66 by Ala in BstL5 results in some destabilization of the recognition module on the protein and in the loss of one of the RNA–protein H-bonds. This decreases the binding affinity 12.5-fold compared with the wild-type protein (Nakashima et al., 2001).

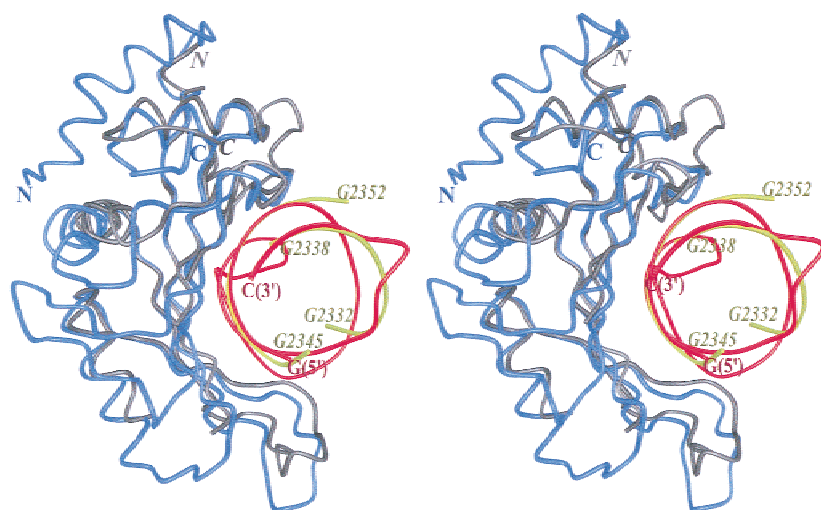
### Structural and evolutionary constraints of 5S rRNA-binding site on L5

The residues comprising the TthL5/RNA and HmaL5/RNA interfaces are shown on Figure 2. Residues Gln66, Pro68, Thr71, Val92, Thr93, Leu94, and Arg95, which are responsible for the stabilization of the recognition module and the recognition of 5S rRNA, are strictly conserved in Bacteria and Archaea. Side chains of res-

idues at positions 67, 70, 91, 96, and 98 form the surface of the interface and can form solvent-accessible hydrogen bonds only. As a result, their replacements do not significantly affect the interface stability. It is interesting to note that only TthL5 has an arginine at position 96, forming four hydrogen bonds with the 5S rRNA. This may be responsible for the increased stability of the hybrid complex between TthL5 and 5S rRNA from *E. coli* (Gongadze et al., 1993) relative to the native *E. coli* L5/5S rRNA complex.

### Crystal packing simulates the interaction of L5 with 23S rRNA

In the 50S ribosomal subunit, protein L5 interacts specifically with 5S rRNA and nonspecifically with 23S rRNA (Ban et al., 2000; Harms et al., 2001). The interface between HmaL5 and 23S rRNA is formed by the concave surface of the  $\beta$ -sheet and the double helical RNA shallow groove. The relative position of HmaL5 and 23S rRNA is very close to that of TthL5 and the NCS-related 5S rRNA molecule in the crystal of the TthL5/5S rRNA complex (Fig. 6). It is thus tempting to suggest that the double-helical part of the NCS-related 5S rRNA molecule in the crystal resembles the 23S rRNA in the ribosome. This allows one to attribute a 23S rRNA binding capacity to the region of TthL5 involved in the above-mentioned intermolecular interactions in the crystal. Among the residues forming this nonspecific contact, Lys36, Asp126, and Asp156 are strictly conserved, and positions of their side chain atoms are stabilized by intramolecular interactions in the structures of the free (BstL5) and bound proteins (TthL5 and HmaL5). Thus, the side chain of the lysine in all three structures is elongated in the middle of the hydrophobic core formed by nonpolar atoms of neighboring residues, whereas conformations of the asparagines are stabilized by hydrogen bonds. The distances between NZ and NH at-



**FIGURE 6.** TthL5 with NCS-related 5S rRNA molecule is superimposed onto HmaL5 with 23S rRNA fragment (stereo view). The protein and RNA backbones from the present structure are in blue and red, respectively.

oms of the conserved lysine and arginine residues are very close in all three proteins. To form L5/23S rRNA hydrogen bonds through these residues, protein L5 should be specifically oriented relative to the 23S rRNA shallow groove. In the crystal of TthL5/5S rRNA and in the 50S ribosomal subunit of *H. marismortui*, the  $\beta$ -strands of L5 are approximately perpendicular to the axis of the RNA A-helix (Fig. 6). As a result, the position of L5 is fixed relative to 5S rRNA as well as to 23S rRNA. Thus, though the two RNAs interact directly (Sergiev et al., 1998; Nissen et al., 2001), the protein should also help to orient these ribosomal RNAs in a proper way within the 50S subunit central protuberance. In the *H. marismortui* 50S ribosomal subunit protein, L5 was found near the tRNA in the P-site (Nissen et al., 2000). The conserved network of hydrogen bonds between L5 and two ribosomal RNAs strongly stabilizes this location of the protein and raises the possibility that L5 might help stabilize the positioning of ribosome-bound tRNA (Nissen et al., 2000).

## MATERIALS AND METHODS

### Sample preparation and crystallization

L5 from *T. thermophilus* was cloned and overproduced in *E. coli* as described by Gongadze et al. (2001). The protein was purified by combination of hydrophobic and cation-exchange chromatography as previously reported for other *T. thermophilus* ribosomal proteins (Vysotskaya et al., 1994; Meshcheryakov et al., 1997). A fragment of RNA corresponding to nt 28–56 of *E. coli* 5S rRNA was synthesized by in vitro transcription using T7 RNA polymerase. The plasmid contained the following elements: a T7 promoter, a template for the target RNA sequence, and an *Sma*I site used for plasmid linearization. RNA was isolated as described by Gongadze et al. (2001). The purified components, L5, and the fragment of 5S rRNA, were mixed in equimolar amounts in 10 mM sodium cacodylate at pH 7.0, 10 mM MgCl<sub>2</sub>, 200 mM KCl. The concentration of the complex in the crystallization solution was 2 mg/mL. The complex solution was mixed with an equal volume of the precipitant solution: 50 mM sodium cacodylate at pH 6.5, 100 mM Mg(CH<sub>3</sub>COO)<sub>2</sub>, 50 mM KF, 10% PEG 8000. KF was used in the crystallization solution because of its very positive effect on the crystal quality. For crystallization, hanging drops of 2–5  $\mu$ L of the mixture were placed on siliconized glasses over 0.25–0.5 mL reservoir solution containing 50 mM sodium cacodylate at pH 6.5, 100 mM Mg(CH<sub>3</sub>COO)<sub>2</sub>, 200 mM KCl, 10% PEG 8000. Crystals were grown at room temperature within a few days to maximum dimensions of 0.05  $\times$  0.05  $\times$  0.1 mm. Subsequently, a selenomethionine derivative for L5 was prepared, its complex with the fragment of 5S rRNA was obtained, and crystals of this complex were produced under conditions identical to those described for the native L5/RNA complex. For diffraction data collection, the following cryoprotective solution was used: 25 mM sodium cacodylate at pH 6.5, 50 mM Mg(CH<sub>3</sub>COO)<sub>2</sub>, 100 mM KCl, 25 mM KF, 5% PEG 8000, 20% 2-methylpentane-2,4-diol (MPD). The crystals were soaked for 2 h and flash-frozen in liquid nitrogen.

### Data collection and structure determination

The data collection experiments were performed at beamline BL44B2 of Spring-8 (Japan) and at beamline ID14.1 of European Synchrotron Radiation Facility (Grenoble, France). The reflections were indexed and integrated using the Denzo and Scalepack programs (Otwinowski & Minor, 1997). Determination of heavy atom positions, initial phasing, and refinement were executed using the CNS program suit (Brünger et al., 1998). Preliminary phases were assigned by the molecular replacement method using the *AmoRe* package (Navaza, 1994) for data, collected at 2.5 Å. The map interpretation and model building were performed with the program O (Jones et al., 1991). The program PROCHECK (Laskowski et al., 1993) was used to check the model quality. To analyze solvent-accessible areas and the area difference of protein L5 and fragment of 5S rRNA in free and bound states, the program AREAIMOL of the CCP4 program suit was used (Bailey, 1994). The probe sphere radius was set to 1.4 Å. The residues and nucleotides were incorporated into the protein/RNA interface if their solvent-accessible surfaces decreased more than 10%. The structural data and the coordinates have been deposited in the Protein Data Bank (accession code 1MJJ).

### ACKNOWLEDGMENTS

This work was supported by the Russian Academy of Sciences and the Russian Foundation for Basic Research (No. 02-04-48806 and No. 01-04-48396). The research of MG was supported in part by an International Research Scholar's award from the Howard Hughes Medical Institute.

Received May 23, 2002; returned for revision  
July 27, 2002; revised manuscript received  
September 9, 2002

### REFERENCES

- Bailey S. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D50*:760–763.
- Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289:905–920.
- Brunel C, Romby P, Westhof E, Ehresmann C, Ehresmann B. 1991. Three-dimensional model of *Escherichia coli* ribosomal 5S RNA as deduced from structure probing in solution and computer modeling. *J Mol Biol* 221:293–308.
- Brünger AT, Adams PD, Marius Clore G, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu, NS, Read RJ, Rice LM, Simonson T, Warren GL. 1998. Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D54*:905–921.
- Erdmann VA, Appel B, Digweed M, Kluwe D, Lorenz S, Luck A, Schreiber A, Schuster L. 1980. Structure and function of 5S and 5.8S ribosomal RNAs. In: Osawa S, Ozeki H, Uchida H, Yura T, eds., *Genetics and evolution of RNA polymerase, tRNA and ribosomes*, Tokyo: University of Tokyo Press. pp 553–568.
- Fedorov R, Meshcheryakov V, Gongadze G, Fomenkova N, Nevskaya N, Selmer M, Laurberg M, Kristensen O, Al-Karadaghi S, Liljas A, Garber M, Nikonov S. 2001. Structure of ribosomal protein TL5 complexed with RNA provides new insights into the CTC family of stress proteins. *Acta Crystallogr D57*:968–976.
- Gongadze GM, Perederina AA, Meshcheryakov VA, Fedorov RV, Moscalenko SE, Rak AV, Serganov AA, Shcherbakov DV, Nikonov SV, Garber MB. 2001. The *Thermus thermophilus* 5S rRNA-protein complex: Identification of specific binding sites for

- proteins L5 and L18 in the 5S rRNA. *Mol Biol (Moscow)* 35:521–526.
- Gongadze GM, Tishchenko SV, Sedelnikova SE, Garber MB. 1993. Ribosomal proteins, TL4 and TL5, from *Thermus thermophilus* form hybrid complexes with 5S ribosomal RNA from different microorganisms. *FEBS Lett* 330:46–48.
- Gourse RL, Thurlow DL, Gerbi SA, Zimmermann RA. 1981. Specific binding of a prokaryotic ribosomal protein to an eukaryotic ribosomal RNA: Implications for evolution and autoregulation. *Proc Natl Acad Sci USA* 78:2722–2726.
- Harms J, Schluenzen F, Zarivach R, Bashan A, Gat S, Agmon I, Bartels H, Franceschi F, Yonath A. 2001. High resolution structure of the large ribosomal subunit from mesophilic eubacterium. *Cell* 107:679–688.
- Horne JR, Erdmann VA. 1972. Isolation and characterization of 5S RNA–protein complexes from *Bacillus stearothermophilus* and *Escherichia coli* ribosomes. *Molec Gen Genet* 119:337–344.
- Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M. 1991. Improved methods for the building of protein models in electron density maps and the location of errors in these. *Acta Crystallogr A* 47:110–119.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
- Leontis NB, Westhof E. 1998. The 5S rRNA loop E: Chemical probing and phylogenetic data versus crystal structure. *RNA* 4:1134–1153.
- Leontis NB, Westhof E. 2001. Geometric nomenclature and classification of RNA base pairs. *RNA* 7:499–512.
- Meshcheryakov VA, Gryaznova OI, Davydova NL, Mudrik ES, Pederina AA, Vasilenko KS, Gongadze GM, Garber MB. 1997. RNA-binding properties of an unusual ribosomal protein TL5 from *Thermus thermophilus*. *Biochemistry* 62:629–634.
- Nakashima T, Yao M, Kawamura S, Iwasaki K, Kimura M, Tanaka I. 2001. Ribosomal protein L5 has a highly twisted concave surface and flexible arms responsible for rRNA binding. *RNA* 7:692–701.
- Navaza J. 1994. AmoRe: An automated package for molecular replacement. *Acta Crystallogr A* 50:157–163.
- Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. The structural basis of ribosomal activity in peptide bond synthesis. *Science* 289:920–930.
- Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA. 2001. RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. *Proc Natl Acad Sci USA* 98:4899–4903.
- Noller HF. 1998. Ribosomal RNA. In: Simons RW, Grunberg-Manago M, eds. *RNA structure and function*. Plainview, New York: Cold Spring Harbor Laboratory Press. pp 73–92.
- Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
- Sergiev P, Dokudovskaya S, Romanova E, Topin A, Bogdanov A, Brimacombe R, Dontsova O. 1998. The environment of 53 rRNA in the ribosome: Cross-links to the GTPase-associated area of 23S rRNA. *Nucleic Acids Res* 26:2511–2525.
- Szymanski M, Barciszewska MZ, Erdmann VA, Barciszewski J. 2002. 5S ribosomal RNA database. *Nucleic Acids Res* 30:176–178.
- Tishchenko S, Nikulin A, Fomenkova N, Nevskaya N, Nikonov O, Dumas P, Moine H, Ehresmann B, Ehresmann C, Pienl W, Lamzin V, Garber M, Nikonov S. 2001. Detailed analysis of RNA–protein interactions within the ribosomal protein S8–rRNA complex from the archaeon *Methanococcus jannaschii*. *J Mol Biol* 311:331–334.
- Vysotskaya V, Tishchenko S, Garber M, Kern D, Mougél M, Ehresmann C, Ehresmann B. 1994. The ribosomal protein S8 from *Thermus thermophilus* VK1: Sequencing of the gene, expression of the protein in *Escherichia coli* and interaction with rRNA. *Eur J Biochem* 223:437–445.
- Westhof E, Romby P, Romaniuk PJ, Ebel JP, Ehresmann C, Ehresmann B. 1989. Computer modeling from solution data of spinach chloroplast and of *Xenopus laevis* somatic and oocyte 5S rRNAs. *J Mol Biol* 207:417–431.
- Yusupov MM, Yusupova GZh, Baucom A, Lieberman K, Earnest TN, Cate JHD, Noller HF. 2001. Crystal structure of the ribosome at 5.5 Å resolution. *Science* 292:883–896.