J Biochem 130, 695–701 (2001)

The Structure of the Archaebacterial Ribosomal Protein S7 and Its Possible Interaction with 16S rRNA[1]

Harumi Hosaka,* Min Yao,* Makoto Kimura,† and Isao Tanaka*,[2]

*Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810; and †Laboratory of Biochemistry, Faculty of Agriculture, Kyushu University, Fukuoka 812-8581

Received August 22, 2001; accepted September 6, 2001

Ribosomal protein S7 is one of the ubiquitous components of the small subunit of the ribosome. It is a 16S rRNA-binding protein positioned close to the exit of the tRNA, and it plays a role in initiating assembly of the head of the 30S subunit. Previous structural analyses of eubacterial S7 have shown that it has a stable α-helix core and a flexible β-arm. Unlike these eubacterial proteins, archaebacterial or eukaryotic S7 has an N-terminal extension of approximately 60 residues. The crystal structure of S7 from archaebacterium *Pyrococcus horikoshii* (*Pho*S7) has been determined at 2.1 Å resolution. The final model of *Pho*S7 consists of six major α-helices, a short 3_{10}-helix and two β-stands. The major part (residues 18–45) of the N-terminal extension of *Pho*S7 reinforces the α-helical core by well-extended hydrophobic interactions, while the other part (residues 46–63) is not visible in the crystal and is possibly fixed only by interacting with 16S rRNA. These differences in the N-terminal extension as well as in the insertion (between α1 and α2) of the archaebacterial S7 structure from eubacterial S7 are such that they do not necessitate a major change in the structure of the currently available eubacterial 16S rRNA. Some of the inserted chains might pass through gaps formed by helices of the 16S rRNA.

Key words: decoding center, protein–RNA interactions, ribosome, RNA-binding protein, X-ray structure.

Ribosomes are cellular organelles responsible for protein synthesis in all living cells. They synthesize proteins according to the genetic code on the mRNA using aminoacylated tRNAs as substrates. The tRNA first enters the A-site as an aminoacyl tRNA (A-tRNA), then moves to the P-site as a peptidyl tRNA (P-tRNA) after the peptide-bond formation. The deacylated tRNA (E-tRNA) leaves the ribosome through the exit (E-site). Bacterial ribosomes (70S) consist of small (30S) subunits containing 16S rRNA and approximately 20 proteins, and large (50S) subunits containing 23S rRNA, 5S rRNA, and over 30 proteins. The 30S subunit is responsible for decoding mRNA by monitoring base-pairing between the codon on the mRNA and the anticodon on the tRNA, while the 50S subunit catalyzes peptide-bond formation.

Ribosomal protein S7 is one of the primary 16S rRNA-binding proteins in the small (30S) subunit. It is located at the head of the 30S subunit and is known to initiate head assembly (*1, 2*) S7 is also known to be located in the vicinity of the ribosome-decoding site. Experiments using photo-labeled tRNAs at anticodon loop have shown that the ribosomal protein S7 is the protein component that is almost exclusively marked by the tRNAs bound at the A-tRNA, P-tRNA, and E-tRNA sites of the ribosome (*3–6*). A similar experiment using photoreactive mRNA has also shown that S7 is in close contact with the upstream region of the mRNA (*7*) This localization of S7 suggests that it not only plays an important role in assembly of the 30S subunit but may somehow contribute to the decoding process of the 30S subunit.

We and another group have previously solved the crystal structures of S7 from eubacterial *Bacillus stearothermophilus* (*Bst*S7) (*8*) and *Thermus thermophilus* (*Tth*S7a) (*9*), showing that S7 has a stable hydrophobic α-helical core and a β-arm. A model-building experiment has indicated that the most distal part of the β-arm is exposed to the tRNA and the upstream region of the codon-anticodon interaction site of the mRNA, suggesting that S7 might work as a gate keeper for the tRNA (*8, 10*).

Recently, crystal structures of the 30S ribosome subunit from *T. thermophilus* have been reported by two groups at 3.3 and 3.0 Å resolution, respectively (*11, 12*). S7 in the *T. thermophilus* 30S subunit (*Tth*S7b) strongly binds to the junction of helices 29, 30, 41, and 42 in the 3′ major domain of 16S rRNA with a loop between α1 and α2, and a stable part of the β-arm. The N-terminal region that was disordered without rRNA in *Bst*S7 and *Tth*S7a was positioned to bind helices 28 and 43 of 16S rRNA. The crystal structure of the 30S subunit further revealed that ribosomal protein S7 contacts ribosomal protein S11 at the C-terminus. More recently, the crystal structure of the 70S ribosome in complex with bound tRNA has also been reported at 5.5 Å resolution (*13*) In the structure of the 70S complex with

[1] This work was supported in part by a Grant-in-Aid for Scientific Research from The Ministry of Education, Science, Sports, and Culture of Japan and "Ground Research for Space Utilization" promoted by NASDA and the Japan Space Forum The atomic coordinates have been deposited in the Protein Data Bank (PDB ID, 1IQV)

[2] To whom correspondence should be addressed. Tel: +81-11-706-3221, Fax. +81-11-706-4905, E-mail tanaka@castor.sci.hokudai.ac.jp

tRNAs, the C-terminal 3_{10}-helix of S7 packs against the backbone of the anticodon stem, whereas the S7 β-arm is positioned at the Watson-Crick face of the E-tRNA anticodon. Furthermore, the structure also revealed that the exit path for the E-tRNA is blockaded by protein S7 (β-arm and C-terminal 3_{10}-helix), L1, and its rRNA-binding site. In order to release the deacylated tRNA, the movement of one or both of these structures is necessary (*13*).

In the present paper, we report the crystal structure of ribosomal protein S7 from archaebacterial *Pyrococcus horikoshii* (*Pho*S7) at 2.1 Å resolution. Unlike eubacterial S7, the S7 from archaea or eukaryotes has a long N-terminal extension of approximately 60 residues. *Pho*S7 also has a 20-residue insertion at a loop between α1 and α2, at which *Tth*S7b binds most tightly to the 16S rRNA.

EXPERIMENTAL PROCEDURES

Preparation of PhoS7—For the overexpression of *Pho*S7, its gene was produced using a two-step polymerase chain reaction (PCR) strategy with the internal primers shown below: N-terminus, 5′-GAAAAGAGAAGCCAAGCATATGAAGGAGG-3′; C-terminus, 5′-GAAGGTTATGAAAATCAAGC-TTAACGTGAGG-3′; NdeI(CATATG)→(CGTATG), 5′-GAAAGTTAGGGCGTATGAAGTAGTTAAGG-3′,5′-CCTTAACTACTTCATACGCCCTAACTTTC-3′. Purified PCR product was digested with *Nde*I and *Hin*dIII and cloned into the expression plasmid pET-22b (Novagen) at the *NdeI*/*Hin*dIII sites. *Escherichia coli* BL21-Codon-Plus-(DE3)-RIL (Stratagene) was transformed with the pET-22b(+)/*Pho*S7 plasmid. The cells were grown at 37°C in LB medium containing 50 μg/ml ampicillin and 34 μg/ml chloramphenicol. The expression of *Pho*S7 gene in *E coli* BL21 (DE3) was induced after injection of 1 mM IPTG. The cells were harvested by centrifugation at 3,000 ×*g* for 15 min at 4°C and washed in buffer A (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.4 M NaCl, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride). The cells were disrupted using a French press at 8.3 MPa. The homogenate was clarified by centrifugation at 40,000 ×*g* for 30 min at 4°C. The supernatant of the cell extracts was incubated at 65°C for 15 min and centrifuged at 40,000 ×*g* for 30 min. The resultant supernatant was applied to a SP-Sepharose column equilibrated with buffer A. After washing, the bound protein was eluted using a linear gradient of 0.4–1.0 M NaCl in buffer A. In this chromatography, *Pho*S7 was eluted at 0.6 M NaCl and gave a single band on SDS-PAGE (data not shown). The fractions containing *Pho*S7 were pooled, dialyzed against distilled water containing 1 mM DTT, and concentrated by ultrafiltration using CentriPlus-10 and Centricon-10 microconcentrators (Amicon) to a final concentration of 10 mg/ml. The purity of the protein was analyzed by MALDI-TOF mass spectrometry (Voyager DE-PRO, PerSeptive Biosystems), and the molecular mass was estimated to be 24,982 Da, which coincided well with the calculated value (24,976 Da) for *Pho*S7.

Crystallization and X-Ray Data Collection—For the crystallization experiments, the purified sample was dialyzed in distilled water containing 1 mM DTT at a concentration of 10 mg/ml. Crystallization trials were carried out at room temperature using the hanging-drop vapor diffusion technique by mixing 2 μl of protein solution with 2 μl of reservoir. The initial crystallization conditions were determined by the sparse-matrix sampling method using Wizzard I and II of Emerald Bio Structures. Crystals were obtained under condition number 7 of Wizzard II [0.1 M Tris-HCl buffer (pH 7.0), 0.2 M NaCl, and 30% (w/v) PEG3000] and grown up to $0.1 \times 0.1 \times 0.2$ mm^3 in 4 days at room temperature.

The crystals were found to diffract to 2.1 Å and to belong to the hexagonal space group $P3_221$ with unit cell dimensions of $a = b = 58.9$ Å, $c = 118.1$ Å, and $\gamma = 120°$. The asymmetric unit contains one monomer of *Pho*S7, corresponding to the V_M value of 2.1 Å^3 Da^{-1}. X-ray diffraction data of *Pho*S7 were collected from a single crystal using the synchrotron radiation source at the BL44XU station of SPring-8, Japan. The crystal was mounted in nylon loops and flash-frozen under a liquid nitrogen stream (100 K). The conditions of the diffraction experiments were as follows: the wavelength was 0.9 Å, the distance to the crystal was 250 mm, and the exposure time was 10 sec/frame for a 1.0° oscillation angle. The total number of frames was 180. All frames were processed using MOSFLM program (*14*) and scale-merged using SCALA program (*15*) (Table I).

Structure Determination and Refinement—The structure of *Pho*S7 was determined by molecular replacement using program AmoRe (*16*). The search model was constructed from *Bst*S7 (*8*) with deletion of the N-terminus (to residue 18). Based on the sigmaa-weighted (2Fo-Fc) and (Fo-Fc) maps that were calculated after rigid-body refinement, the atomic model was rebuilt using the graphics program O (*17*). The dynamic refinement was carried out using the program CNS with a manual fitting model between rounds

TABLE I **Statistics of data collection and refinement.**

Data collection	
Wavelength (Å)	0.9
Space group	$P3_221$
Resolution[a] (Å)	40–2.1 (2.21–2.1)
Observed reflections	139,697
Independent reflections	14,407
Completeness (%)	99.8 (99.8)
Multiplicity	9.7 (7.1)
$I/\sigma(I)$	7.2 (2.9)
R_{meas}[b] (%)	8.1 (26.2)
Refinement	
Resolution range (Å)	10–2.1
No. of reflections	14,240 ($F \geq 2\sigma$)
Residues included	166
No. of non-hydrogen atoms	1,300
No. of water molecules	140
R-factor[c] (%)	19.6
R_{free}-factor[d] (%)	23.0
Average B-factor (Å^2)	
Main-chain	31.8
Side-chain	36.7
Solvent	50.6
Rms deviations	
Bond lengths (Å)	0.005
Bond angles (°)	1.03
Ramachandran plot[e]	
Residues in most favored regions (%)	92.5
Residues in additional allowed regions (%)	7.5

[a]Values in parentheses are for the outermost resolution shell. [b]$R_{meas} = \Sigma_h[m/(m-1)]^{1/2}\Sigma_j|\langle I\rangle_h - I_{hj}|/\Sigma_h\Sigma_j I_{hj}$, where $\langle I\rangle_h$ is the mean intensity of symmetry-equivalent reflections and m is redundancy. [c]R-factor$=\Sigma||F_{obs}|-|F_{cal}||/\Sigma|F_{obs}|$, where F_{obs} and F_{cal} are observed and calculated structure factor amplitudes. [d]R_{free}-factor value was calculated for R-factor, using only an unrefined subset of reflections data (10%). [e]Ramachandran plot was calculated by PROCHECK program (*26*).

using the program O. Simulated-annealing refinement was carried out in the first round with 2,500 K, and conjugate gradient minimization and restrained individual isotropic temperature-factor refinement were performed in each round with overall anisotropic temperature-factor and bulk-solvent scaling (*18*). At the current stage of refinement, the model has an *R*-factor of 19.6% and a free R-factor of 23.0% for the data in the 20 Å to 2.1 Å resolution region, including 166 residues for the crystallographically independent molecule and 131 water molecules. The refinement statistics are summarized in Table I. The coordinates will be deposited in the Protein Data Bank (*19*).

RESULTS AND DISCUSSION

Structure Description—The stereo view of the overall structure of the *Pho*S7 is given in Fig. 1. The final model of *Pho*S7 consists of six major α-helices (α1-α6), a short 3_{10}-helix (α1′), and two β-strands (β1-β2), with a connectivity of α1′-α1-α2-α3-β1-β2-α4-α5-α6 [for comparison with other S7s, we use the numbering system for secondary structures in reference to *Bst*S7 and *Tth*S7a (*8, 9*)]. Five of six α-helices (α1, α2, α3, α4, and α5) form a hydrophobic core. α1 and α2 make up the helix-turn-helix (HTH) motif. α4 and α5 have a chain trace similar to HTH, but its chain direction is reversed. There is a pseudo-twofold symmetry axis between these two HTH motifs, which are tightly entangled together by the well-conserved hydrophobic residues. At the bottom of these four helices, α3 runs perpendicularly to the pseudo-twofold axis. Two β-strands form a natural right-handed twisted antiparallel β-arm between α3 and α4. The C-terminal helix α6 (residues 203–216) extends halfway along the β-arm. The N-terminal residues 1–17 and a loop between helices α1 and α2 (residues 81–97) were not constructed, as the electron density is poor in these regions. Another long loop of residues 46–63 between α1′ and α1 is also disordered and not included in the present structure.

Overall Comparison of the PhoS7 and Eubacterial S7 Structures—Recent progress in the genome-sequencing project has allowed sequential comparison of ribosomal proteins from wide variety of organisms. Such comparison has shown that although ribosomal proteins are well-conserved among different phylogenies, they often have large insertions (extensions) or deletions in their peptide chains. Since the atomic coordinates of the 30S subunit from *T. thermophilus* (*12*) and the 50S subunit from *Haloarcula marismortui* (*20*) are now available, the significance of such an insertion (extension) or deletion can be evaluated in part based on the structure of the independent molecules.

Comparative analysis of amino acid sequences using the program CLUSTALW (*21*) for all ribosomal proteins in the 30S subunit within archaebacteria (*P. horikoshii*), eukaryote (*Human*), and eubacteria (*T. thermophilus*) has shown that the sequences of archaebacterial ribosomal proteins are generally longer than those of eubacteria, with 15–25% identical residues. They are more similar to those of eukaryotes, with 35–45% identical residues. Most of the large differences (extension/deletion) between archaebacterial *P. horikoshii* and eubacterial *T. thermophilus* are located at the terminal regions. Figure 2 shows a comparison of the sequences of the ribosomal protein S7 from several phylogenies. Among the ribosomal proteins common to eubacteria and archaebacteria, S7 has one of the longest extensions (deletion). The peptide chain of *Pho*S7 is longer than that of *Tth*S7 by 65 residues, of which 45 residues are at the N-terminus and others are in the rRNA-binding loop between α1 and α2

Figure 3a is a superposition of *Pho*S7 and eubacterial S7 from the 30S subunit of *T. thermophilus* (*Tth*S7b) (*12*). Although the N-terminal regions and the two loops (between α1′ and α1, and between α1 and α2) of *Pho*S7 are not defined in the crystal, it is still evident that the two molecules have the same folding topology and a similar structure. The structure of S7 may be divided into three parts: an α-helical domain, a β-arm, and N- and C-termini. The α-helical domain is composed of five helices (α1-α5). This is the core of the S7 structure and acts as an anchor from which rRNA-binding sites extend. This domain has a very fixed structure, as shown by the small r.m.s. deviations of 1.6 Å between *Pho*S7 and *Bst*S7, 1.7 Å between *Pho*S7 and

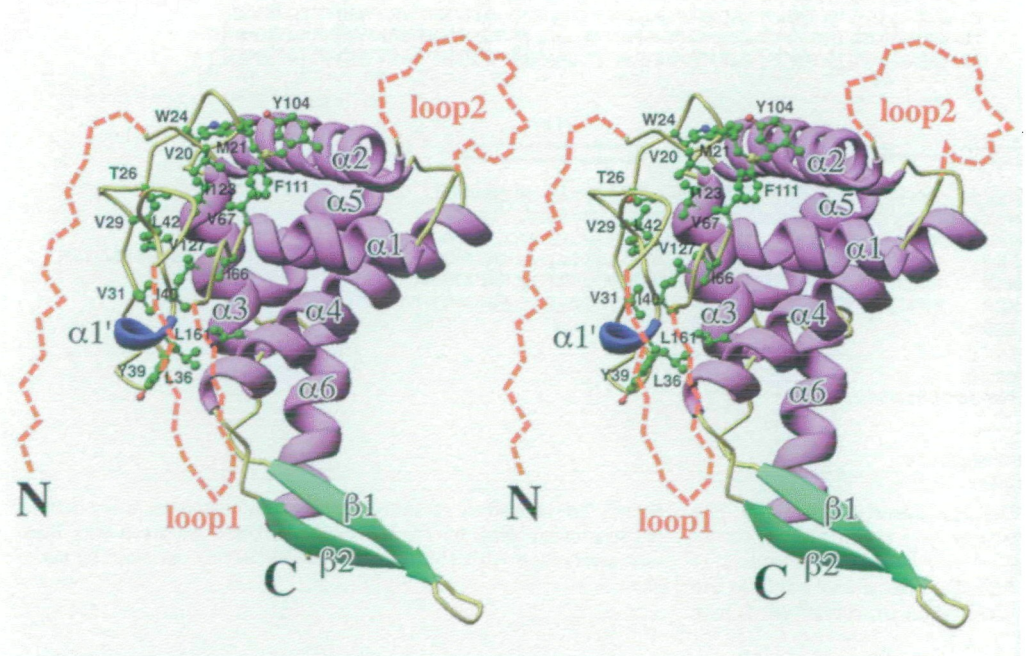

Fig 1 A stereoscopic drawing of the ribosomal protein S7 from *P. horikoshii*. The balls and sticks colored green represent hydrophobic amino acids which make van der Waals contacts between the hydrophobic α-helical core and the N-terminal extension specific to archaebacteria and eukaryotes. The dotted lines are disordered residues of 1–17, 46–63, and 81–97.

*Tth*S7b, and 0.66 Å between *Bst*S7 and *Tth*S7b for 113 Cα atoms in α1, α2, α3, α4, and α5. The residues creating the hydrophobic contacts in the core are well-conserved: especially important are six alanyl residues, Ala110, Ala129, Ala172, Ala188 Ala194, and Ala195. These residues are located at the center of the hydrophobic cluster, at which no larger hydrophobic residues are allowed. In contrast with the α-helical domain, the structure of the β-arm is somewhat flexible, as indicated by the larger r.m.s. deviation values (2.8 Å between *Pho*S7 and *Bst*S7, 2.2 Å between *Pho*S7 and *Tth*S7b, and 1.3 Å between *Bst*S7 and *Tth*S7b for 129 Cα atoms, including β1 and β2). The amino acid residues as well as the secondary structure in the β-arm are moderately conserved. The N- and C-terminal portions are more diverse both in their amino acid sequences and their three-dimensional structure. Moreover, the structures of these portions are often disordered in crystals of isolated molecules from 16S rRNA.

N-terminal Extra Region—As shown in Fig. 2 *Pho*S7 shares only approximately 20% identical residues with eubacterial S7s (*Bst*S7 and *Tth*S7). The largest difference is at the N-terminal region, in which 21 residues of eubacterial S7 are replaced by the unrelated 65 residues of *Pho*S7, corresponding to approximately 30% of the 218 residues of *Pho*S7.

Crystal structure analysis of the isolated S7 molecules from eubacteria (*Bst*S7 and *Tth*S7) has shown that the 10 N-terminal residues are disordered (*8, 9*). This region is structured only when it interacts with 16S rRNA, as shown by the subsequently analyzed 30S subunit (*12*); the N-terminal residues (1–10) stretch into a cleft formed by helices 28, 40, and 43 of 16S rRNA and come into contact with helices 28 and 43 with Ala2, Arg3, Arg4, Arg6, and Arg10 (here the residue number is that of eubacterial S7). Although the amino acid sequence of this region has no similarity to that of archaea or eukaryotes, some of these residues (Ala2, Arg3, Arg4, and Arg10) are well-conserved in the eubacterial S7s. The truncation of this region of *Bst*S7 abolishes the 16S rRNA binding (*22*). Furthermore, a recent analysis using reconstituted ribosomes with the protein S7 tagged with the protein kinase A recognition site showed that the S7 derivative lacking the 17 N-terminal residues causes ribosomes to accumulate on mRNA to abnormally high levels, suggesting that ribosomes containing this S7 derivative are defective in translation elongation or termination (*23*)

Although, as in the case of the isolated S7 molecules of eubacteria, no structure was found around the corresponding region after the superposition of eubacterial S7 and *Pho*S7, the chain trace of the visible part of *Pho*S7 and the

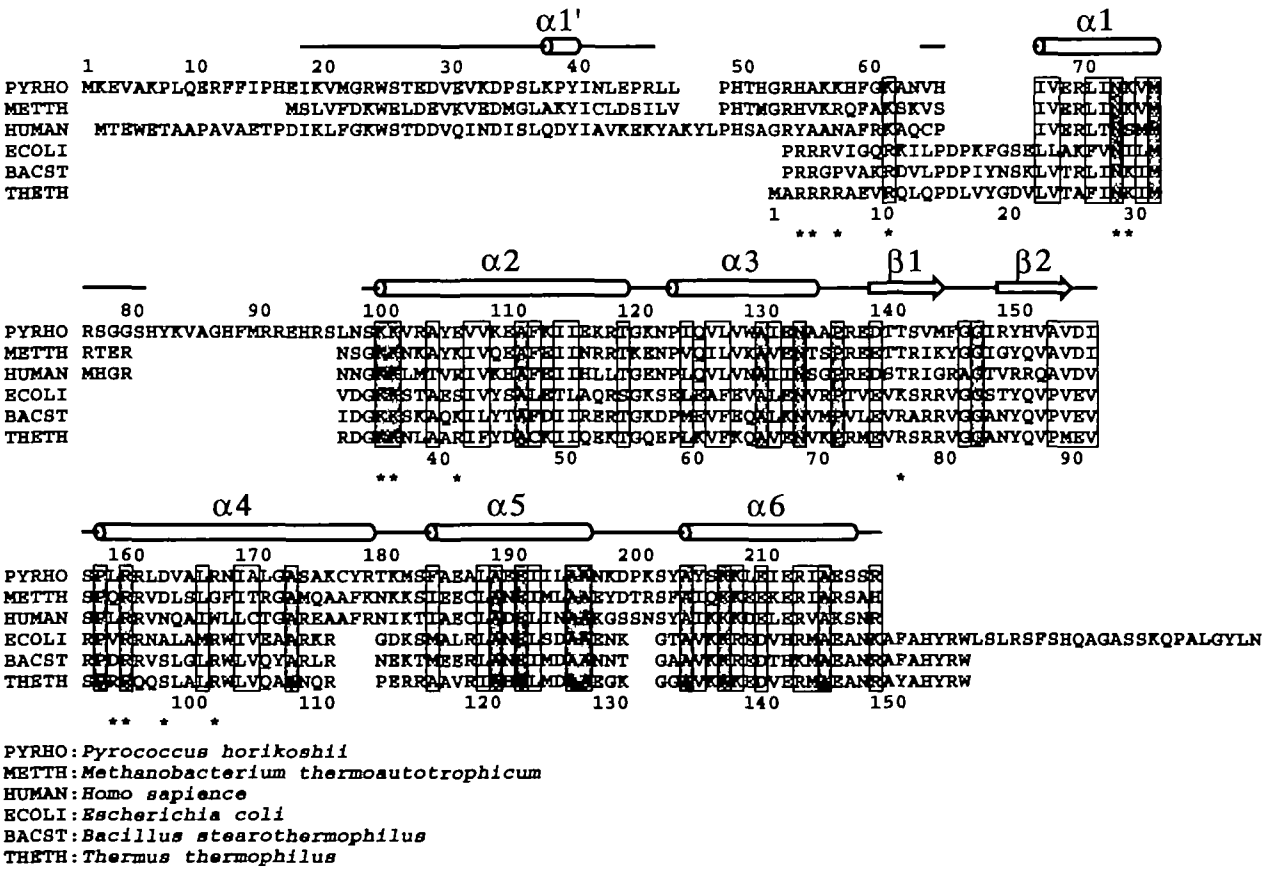

Fig. 2. **Sequence comparison of the ribosomal protein S7s.** The alignment of amino acid sequences was first calculated by use of CLUSTALW (*21*) and modified based on secondary structures of *Pho*S7 (present work) and *Tth*S7b (*12*) The amino acid residues are shown as follows completely identical (shaded), conserved change (boxed) The secondary structure elements indicated are those defined by the present work for *Pho*S7 using the DSSP program (*25*). Residues interacting with 16S rRNA in the *Tth*S7b are marked by asterisks.

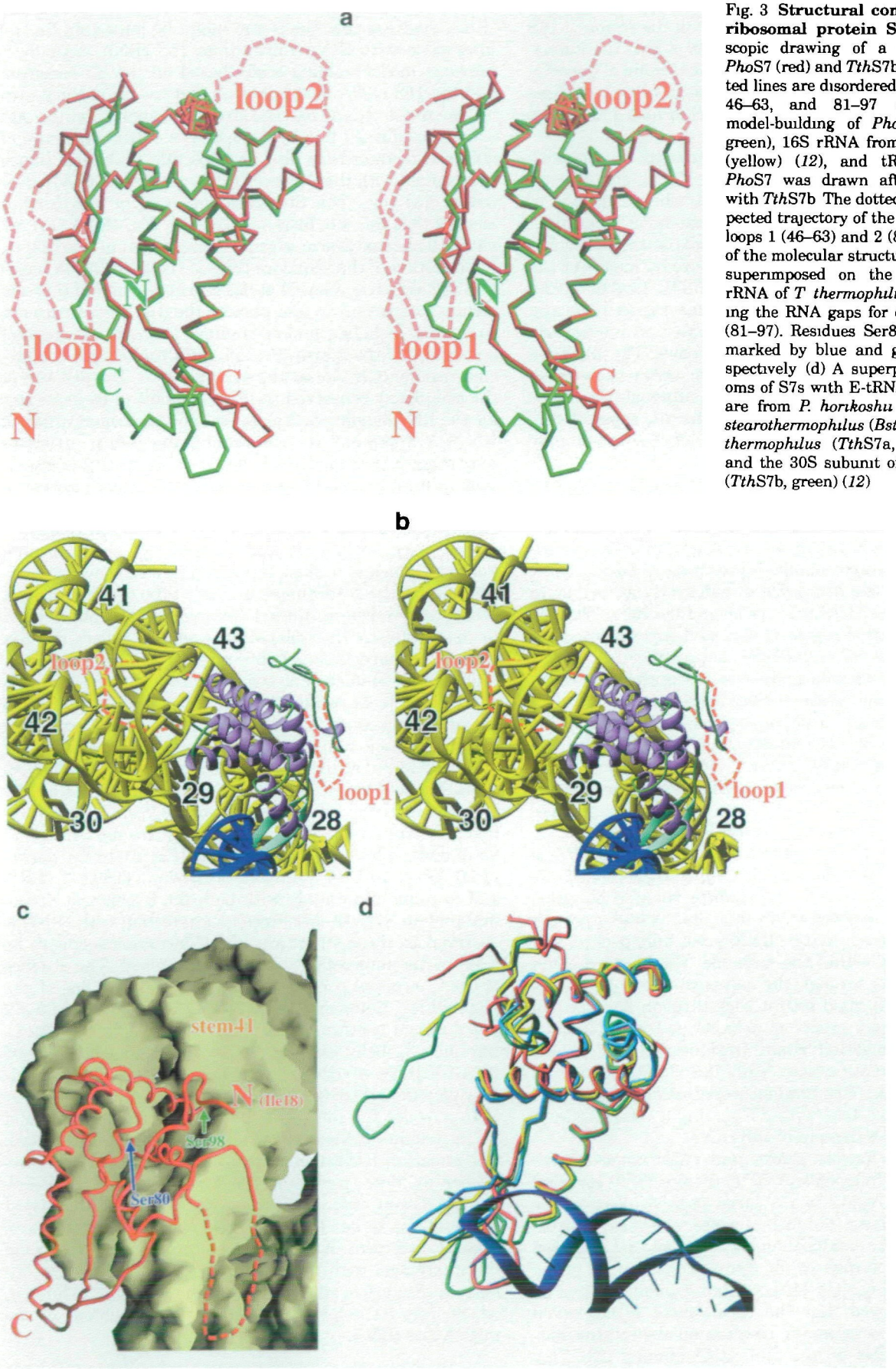

Fig. 3 **Structural comparison of the ribosomal protein S7s.** (a) A stereoscopic drawing of a superposition of *Pho*S7 (red) and *Tth*S7b (green) The dotted lines are disordered residues of 1–17, 46–63, and 81–97 (b) Stereoscopic model-building of *Pho*S7 (purple and green), 16S rRNA from *T thermophilus* (yellow) (*12*), and tRNA (blue) (*13*). *Pho*S7 was drawn after superposition with *Tth*S7b The dotted lines are the expected trajectory of the peptide chains of loops 1 (46–63) and 2 (81–97). (c) A view of the molecular structure of *Pho*S7 (red) superimposed on the surface of 16S rRNA of *T thermophilus* (yellow) showing the RNA gaps for disordered loop 2 (81–97). Residues Ser80 and Ser98 are marked by blue and green arrows, respectively (d) A superposition of C^{α} atoms of S7s with E-tRNA (blue) The S7s are from *P. horikoshii* (*Pho*S7, red), *B stearothermophilus* (*Bst*S7, yellow) (*8*), *T thermophilus* (*Tth*S7a, light blue) (*9*), and the 30S subunit of *T thermophilus* (*Tth*S7b, green) (*12*)

partial trace of the poor electron density suggests that residues 46–63 (loop1) come into contact with the region of 16S rRNA where N-terminal residues (1–10) of *Tth*S7b interact. Thus, instead of the end of the N-terminal chain of eubacterial S7, *Pho*S7 seems to come into contact with the helices 28 and 43 by means of residues 46–63 in loop 1 (Fig. 3, a and b).

Residues 11–21 of eubacterial S7 come into contact with the hydrophobic α-helical core (α1, α2, α3, α4, and α5) and are exposed on the surface of the 30S subunit. Instead of these residues of *Tth*S7b, part of the extended N-terminal region (residues 18–45) of *Pho*S7 attends to the hydrophobic α-helical core (α1, α2, α3, α4, and α5) by means of the Val20, Met21, Trp24, Thr26, Val29, Val31, Leu36, Tyr39, Ile40, and Leu42 residues (Fig. 1). This region is on the opposite side of 16S rRNA and is exposed on the particle surface. As the sequence alignment shows (Fig. 2), these hydrophobic residues are conserved in archaebacteria as well as in eukaryotic ribosomes. Thus, although the N-terminal region seems to have evolved after the separation of archaebacteria and eubacteria, the role of this portion may be similar in the two phylogenies.

16S rRNA-Binding Loop (Loop2)—In the crystal structure of the 30S subunit, S7 tightly binds to 16S rRNA by a helix-turn-helix (HTH) between α1 and α2 with Asn28, Lys29, Asp33, Lys35 Lys36, Asn37, Arg41, and Gly34 residues (here the residue numbers are those of *Tth*S7). Most of these residues are conserved in eubacteria as well as in *Pho*S7 (Asn72, Lys73, Asn97, Lys99, and Lys100 in *Pho*S7). The structure of this region is also well-superimposed between eubacterial S7 and *Pho*S7, suggesting that *Pho*S7 binds to 16S rRNA similarly by these residues in the HTH. However, as mentioned above, *Pho*S7 has an insertion of 20 amino acids (residues 76–95) at the turn region of the HTH between residues Met31 and Arg32 of *Tth*S7 (Fig. 2). The sequence of the 16S rRNA at the 3′ major domain is highly conserved, and no major change in the tertiary structure of RNA is expected. Thus, it is an open question as to where this inserted chain runs with respect to 16S rRNA.

Although most of the inserted chain (residues 81–97) is not visible in the present crystal, a superposition of *Pho*S7 with the 30S subunit of *T. thermophilus* shows a plausible trajectory of this inserted chain (Fig. 3b). Unlike a protein molecule, the interior of 16S rRNA is not fully packed, and holes are present within the molecule. The inserted chain in HTH may pass through the gap surrounded by helices 29, 30, 41, and 43, then return back through another gap formed at the major groove of helix 41 of 16S rRNA (Fig. 3c) With this inserted chain (residues 81–97), *Pho*S7 makes more intimate contact with 16S rRNA, making 16 rRNA more stable. This binding model of *Pho*S7 and 16S rRNA is also consistent with S7 playing a crucial role in folding the 3′ major domain of 16S rRNA.

tRNA-Binding Region: β-Arm and C-Terminus—Structural analysis of independent S7 molecules (*8, 9*) revealed that the β-arm region has a large positive electrostatic potential and a large number of conserved basic and exposed hydrophobic residues. As such, it was first expected that the concave surface of the β-arm would be one of the main rRNA-binding sites. However, subsequent analysis by mutagenesis showed that the replacement of conserved positive residues or aromatic residues on the concave surface had little effect on the 16S rRNA binding (*22*). This result, together with the close location of the β-arm to the tRNA, suggests that the β-arm might be important for the interaction with tRNA rather than 16S rRNA. Actually, a previous model-building study based on the S7 structure and the 16S rRNA model showed that the tip of the β-arm is close to the tRNA (*10*). The structural analysis of the 30S subunit confirmed this hypothesis. The crystal structure of 70S of *T. thermophilus* also revealed that the β-arm comes into contact with the Watson-Crick face of the E-tRNA anticodon (*13*) (Fig. 3d). Structural analysis of 70S further showed that this site blocks the release of E-tRNA (*13*). S7 with a truncated β-arm is poorly represented in 70S (*23*).

Inspection of the sequence (Fig. 2) showed that no insertion or deletion is allowed at the β-arm region and that the residues important to maintaining the β-arm structure are conserved, including proline residues at the beginning and end points of the β-arm (Pro135 and Pro157), and the glycine residue (Gly146) at the turn position. The absence of the completely conserved residues thought to be necessary for the RNA-binding at this region and the conservation of residues important to maintaining the β-arm structure may suggest that the flexible nature of the β-arm is important to its function. Figure 3d shows the superposition of S7s from the crystal structures analyzed thus far, which clearly demonstrates the flexible nature of the S7 β-arm.

The β-arm is halfway accompanied by the C-terminal helix α6, which is well-superposed (Fig. 3d). However, as is the case for the N-terminus, the end of the C-terminus (approximately seven residues) following the α6 is not ordered in the crystals of the isolated molecules of eubacterial S7s (*8, 9*) The importance of this region has been shown by mutagenesis studies of *E. coli* S7, with a mutant having a short deletion (31 residues) at the C-terminus appearing to be more detrimental to cell growth than the S7 wild type (*24*). In addition, a mutant with a longer deletion including α6 (47 residues) at the C-terminus assembled into 30S subunits, but was found to be very poorly represented in 70S (*23*). Interestingly, while the C-terminal residues (and 3_{10}-helix) are conserved in eubacteria, they are not present in S7 of eukaryotes and archaebacteria (Fig. 2). In the eubacterial 30S or 70S, this portion was shown to form a 3_{10}-helix and to come into contact with both the β-arm and ribosomal protein S11 (*12, 13*). Since no interaction with rRNA is observed in these structures, the C-terminus seems to be fixed by the interaction with the S11 protein. The absence of the C-terminal portion of *Pho*S7 may suggest one of two possibilities. Compared with *Tth*S7, the β-arm of *Pho*S7 has different contact and does not need the help of the C-terminal 3_{10}-helix site, or an insertion [residues 86–93 (PGGSKSKT)] of ribosomal protein S11 of *P. horikoshii* may play the role of the C-terminal 3_{10}-helix.

To consider the role of the β-arm and the C-terminus of S7 in protein synthesis, several ideas have been proposed. The importance of this region was first suggested as a gatekeeper (*8*). More specifically, the region may help to dissociate the E-site codon-anticodon (*12*). Although the actual role remains to be elucidated, the flexibility of the β-arm may be important. It may be worth noting that because of the interaction with the C-terminal α6 (including the 3_{10}-helix in the case of eubacteria), β-arm movement is limited, which may be helpful in the directed movement of the mRNA and tRNA.

We would like to thank Drs. E Yamashita, A Nakagawa, and T Tsukihara, of Institute for Protein Research, Osaka University of Japan, for their kind help with data collection using synchrotron radiation at BL44XU, SPring-8

REFERENCES

1 Nowotny, V and Nierhaus, K.H (1988) Assembly of the 30S subunit from *Escherichia coli* ribosomes occurs *via* two assembly domains which are initiated by S4 and S7. *Biochemistry* **27**, 7051–7055

2 Mandiyan, V, Tumminia, S, Wall, J S, Hainfeld, J F, and Boublik, M (1989) Protein-induced conformational changes in 16 S ribosomal RNA during the initial assembly steps of the *Escherichia coli* 30 S ribosomal subunit *J Mol Biol* **210**, 323–336

3 Abdurashidova, G G, Tsvetkova, E.A., and Budowsky, E I (1989) Nucleotide residues of tRNA, directly interacting with proteins within the complex of the 30 S subunit of *E coli* ribosome with poly(U) and NAcPhe-tRNA(Phe) *FEBS Lett* **243**, 299–302

4 Abdurashidova, G G, Tsvetkova, E.A., and Budowsky, E.I (1990) Determination of tRNA nucleotide residues directly interacting with proteins in the post- and pretranslocated ribosomal complexes. *FEBS Lett* **269**, 398–401

5 Sylvers, L A., Kopylov, A.M, Wower, J, Hixson, S S, and Zimmermann, R.A. (1992) Photochemical cross-linking of the anticodon loop of yeast tRNA(Phe) to 30S-subunit protein S7 at the ribosomal A and P sites *Biochimie* **74**, 381–389

6. Döring, T, Mitchell, P, Osswald, M, Bochkariov, D, and Brimacombe, R (1994) The decoding region of 16S RNA, a cross-linking study of the ribosomal A, P and E sites using tRNA derivatized at position 32 in the anticodon loop. *EMBO J* **13**, 2677–2685

7 Stade, K., Rinke-Appel, J, and Brimacombe, R (1989) Site-directed cross-linking of mRNA analogues to the *Escherichia coli* ribosome; identification of 30S ribosomal components that can be cross-linked to the mRNA at various points 5′ with respect to the decoding site. *Nucleic Acids Res.* **17**, 9889–9908

8 Hosaka, H, Nakagawa, A, Tanaka, I, Harada, N, Sano, K., Kimura, M, Yao, M., and Wakatsuki, S (1997) Ribosomal protein S7 a new RNA-binding motif with structural similarities to a DNA architectural factor. *Structure* **5**, 1199–1208

9 Wimberly, B.T, White, S W, and Ramakrishnan, V (1997) The structure of ribosomal protein S7 at 1 9 Å resolution reveals a beta-hairpin motif that binds double-stranded nucleic acids *Structure* **5**, 1187–1198

10. Tanaka, I, Nakagawa, A, Hosaka, H, Wakatsuki, S, Mueller, F, and Brimacombe, R (1998) Matching the crystallographic structure of ribosomal protein S7 to a three-dimensional model of the 16S ribosomal RNA. *RNA* **4**, 542–550

11. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J, Gluehmann, M., Janell, D, Bashan, A., Bartels, H, Agmon, I, Franceschi, F., and Yonath, A. (2000) Structure of functionally activated small ribosomal subunit at 3 3 angstroms resolution *Cell* **102**, 615–623

12. Wimberly, B.T, Brodersen, D E, Clemons, W M, Jr, Morgan-Warren, R.J., Carter, A.P, Vonrhein, C., Hartsch, T, and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit *Nature* **407**, 327–339

13 Yusupov, M M, Yusupova, G Z, Baucom, A., Lieberman, K., Earnest, T N, Cate, J H., and Noller, H F (2001) Crystal structure of the ribosome at 5 5 Å resolution. *Science* **292**, 883–896

14 Leslie, A.G W (1993) Auto-indexing of rotation diffraction images and parameter refinement in *Proceeding of the CCP4 Study Weekend* (Sawyer, L, Isaacs, N, and Bailey, S, eds.) pp 44–51, SERC Daresbury Laboratory, Warrington, UK

15. Evans, P R (1997) Scaling of MAD data in *Proceedings of the CCP4 Study Weekend* (Wilson, K.S, Davies, G, Ashton, A W, and Bailey, S, eds) pp 97–102, CCLRC Daresbury Laboratory, Warrington, UK

16 Navaza, J. (1994) AmoRe An automated package for molecular replacement *Acta Crystallogr* **A50**, 157-163

17 Jones, T.A, Zou, J Y, Cowan, S W, and Kjeldgaard (1991) Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr* **A47**, 110–119

18 Brunger, A.T, Adams, P D, Clore, G M, DeLano, W L, Gros, P, Grosse-Kunstleve, R W, Jiang, J S, Kuszewski, J, Nilges, M, Pannu, N S, Read, R J, Rice, L M, Simonson, T, and Warren, G L (1998) Crystallography & NMR system A new software suite for macromolecular structure determination *Acta Crystallogr* **D54**, 905–921

19. Abola, E E, Sussman, J L, Prilusky, J, and Manning, N O (1997) Protein Data Bank archives of three-dimensional macromolecular structures. *Methods Enzymol* **277**, 556–571

20 Ban, N, Nissen, P, Hansen, J, Moore, P B, and Steitz, T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2 4 Å resolution *Science* **289**, 905–920

21 Thompson, J D, Higgins, D G, and Gibson, T J (1994) CLUSTAL W. improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice *Nucleic Acids Res.* **22**, 4673–4680

22 Miyamoto, A., Usui, M, Yamasaki, N, Yamada, N, Kuwano, E, Tanaka, I, and Kimura, M (1999) Role of the N-terminal region of ribosomal protein S7 in its interaction with 16S rRNA which binds to the concavity formed by the beta-ribbon arm and the alpha-helix *Eur J Biochem* **266**, 591–598

23 Fredrick, K., Dunny, G M., and Noller, H F (2000) Tagging ribosomal protein S7 allows rapid identification of mutants defective in assembly and function of 30 S subunits. *J Mol Biol* **298**, 379–394

24 Robert, F and Brakier-Gingras, L (2001) Ribosomal protein S7 from Escherichia coli uses the same determinants to bind 16S ribosomal RNA and its messenger RNA. *Nucleic Acids Res.* **29**, 677–682

25 Kabsch, W and Sander, C (1983) Dictionary of protein secondary structure pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637

26 Laskowski, R.A., MacArthur, M W, Moss, D S, and Thornton, J M (1993) PROCHECK. a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26**, 283–291