Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of ribosomal protein TL5 complexed with RNA provides new insights into the CTC family of stress proteins

Roman Fedorov,[a] Vladimir Meshcheryakov,[a] George Gongadze,[a] Natalia Fomenkova,[a] Natalia Nevskaya,[a] Maria Selmer,[b] Martin Laurberg,[b] Ole Kristensen,[b] Salam Al-Karadaghi,[b] Anders Liljas,[b] Maria Garber[a] and Stanislav Nikonov[a]*

[a]Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia, and [b]Molecular Biophysics, Center for Chemistry and Chemical Engineering, Lund University, Box 124, SE-22 100 Lund, Sweden

Correspondence e-mail: nikonov@vega.protres.ru

Received 15 February 2001
Accepted 12 April 2001

PDB Reference: TL5–RNA complex, 1feu.

The crystal structure of *Thermus thermophilus* ribosomal protein TL5 in complex with a fragment of *Escherichia coli* 5S rRNA has been determined at 2.3 Å resolution. The protein consists of two domains. The structure of the N-terminal domain is close to the structure of *E. coli* ribosomal protein L25, but the C-terminal domain represents a new fold composed of seven β-strands connected by long loops. TL5 binds to the RNA through its N-terminal domain, whereas the C-terminal domain is not included in this interaction. Cd^{2+} ions, the presence of which improved the crystal quality significantly, bind only to the protein component of the complex and stabilize the protein molecule itself and the interactions between the two molecules in the asymmetric unit of the crystal. The TL5 sequence reveals homology to the so-called general stress protein CTC. The hydrophobic cores which stabilize both TL5 domains are highly conserved in CTC proteins. Thus, all CTC proteins may fold with a topology close to that of TL5.

1. Introduction

Ribosomal protein TL5 is a 5S rRNA-binding protein from the large ribosomal subunit of the extreme thermophilic bacterium *T. thermophilus* (Sedelnikova *et al.*, 1987; Gongadze *et al.*, 1993). Its sequence reveals homology (Gryaznova *et al.*, 1996) to the general stress protein CTC found in *Bacillus subtilis* (Völker *et al.*, 1994). Genes of the CTC family are found in genomes of most bacteria. As a rule, they encode proteins of approximately 200 amino-acid residues, except in *E. coli* and *Haemophilus influenzae*, where a shorter sequence homologous to the N-terminal part of full-size CTC proteins is present. In *E. coli*, this short sequence corresponds to a single-domain ribosomal protein L25 binding 5S rRNA (Chen-Schmeisser & Garrett, 1977). TL5 is the counterpart of L25 in *T. thermophilus* ribosomes; its N-terminal domain binds to the same site on 5S rRNA as L25 (Gongadze *et al.*, 1999). Moreover, TL5 can replace L25 functionally in reconstructed *E. coli* ribosomes (Zvereva *et al.*, 2000). Because only one gene for proteins homologous to CTC was found in bacterial species for which the whole genome is known, proteins of this family have two functions, acting as both ribosomal and stress proteins. L25 can be considered as a truncated form and TL5 as a full-size CTC protein.

The structure of the ribosome is conserved in evolution, but there are certain differences between bacterial, archaeal and eukaryal ribosomes. In particular, there is no protein homologous to TL5 or CTC in the Archaea and Eukarya. Even though the structure of the large ribosomal subunit from the halophilic archaeon *Haloarcula marismortui* has recently been

determined at high resolution (Ban *et al.*, 2000), determination of the structures of bacterial ribosome components is still of interest. Three-dimensional structures of *E. coli* L25 free and in complex with a fragment of 5S rRNA have been determined (Stoldt *et al.*, 1998, 1999; Lu & Steitz, 2000). The structure of TL5 described in this paper is the first structure of a full-size protein of the CTC family. The fold of the N-terminal domain is similar to that of the ribosomal protein L25, but the C-terminal domain of TL5 has a unique fold. We determined the structure of TL5 in complex with a fragment of 5S rRNA which is similar to that in the L25–RNA complex. Comparison of the two structures allows an understanding of the differing stabilities of the two complexes.

Ribosomes from *T. thermophilus* are now undergoing structural analysis and the electron-density map of the 70S ribosome at 7.8 Å resolution has recently been published (Cate *et al.*, 1999). At present, this map has been improved to about 5 Å resolution (M. Yusupov, personal communication). The structure of *T. thermophilus* ribosomal protein TL5 described here offers the possibility of fitting this protein into the electron-density map of the *T. thermophilus* ribosome.

2. Experimental

2.1. Crystallization of TL5–RNA complex

TL5 from *T. thermophilus* was overproduced in *E. coli* (Gongadze *et al.*, 1999). Several fragments of 5S rRNA prepared by RNase A digestion were used for crystallization of the TL5–RNA complex. The best crystals were obtained with a 40 nt fragment of *E. coli* 5S rRNA protected by TL5 from RNase A (Gongadze *et al.*, 1999). The purified components, TL5 and the fragment of 5S rRNA, were mixed in equimolar amounts in 10 m*M* sodium cacodylate pH 7.0, 4 m*M* $MgCl_2$, 100 m*M* KCl. The concentration of the complex in the crystallization solution was 2–4 mg ml^{-1}. The solution was mixed with an equal volume of the crystallization solution: 10 m*M* sodium cacodylate pH 7.0, 100 m*M* $MgCl_2$, 4 m*M* $CdCl_2$, 7% 2-methylpentane-2,4-diol (MPD). $CdCl_2$ was used as an additive in the crystallization solution because of its very positive effect on the crystal quality. For crystallization, hanging drops of 10–20 µl of the mixture were placed on siliconized glasses over 0.5 ml reservoir solution containing 200 m*M* KCl and 12% MPD. Long hexagonal lense-shaped crystals appeared after several days at room temperature. For diffraction data collection, a cryoprotective solution was used: 10 m*M* sodium cacodylate pH 7.0, 100 m*M* $MgCl_2$, 100 m*M* KCl, 4 m*M* $CdCl_2$, 15% MPD, 10% PEG 400. The crystals were soaked for 1 h and flash-frozen in liquid nitrogen. The crystals belong to space group $P3_112$, with unit-cell parameters $a = b = 109.9$, $c = 137.5$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$, and diffract to 2.3 Å resolution. There are two complexes per asymmetric unit of the crystal.

2.2. Data collection, processing and phasing

The native and SeMet derivative data sets were collected on beamline BL711 at the MAX-II synchrotron, Lund, Sweden using a MAR imaging plate. Three-wavelength MAD data were collected at the EMBL beamline BW7A, DESY, Hamburg using a single SeMet crystal and a MAR CCD detector. All data were processed and merged with the *DENZO* and *SCALEPACK* programs (Otwinowski & Minor, 1997). Data statistics are summarized in Table 1. Initial phases were obtained by molecular replacement (MR) using a part of the *E. coli* L25–RNA complex structure containing the 22 nt fragment of RNA and helix α_1 and strands β_2 and β_6 of L25 as a model. The positions of the two molecules in the asymmetric unit were found with a six-dimensional search program (Kissinger *et al.*, 1999). The initial phases were used to calculate difference Fourier maps for a selenomethionine derivative. All six selenium sites for the two molecules in the asymmetric unit were easily found by manual inspection of these maps. The MR and SeMet-derivative phases were subsequently combined and a new electron-density map was calculated. Phase-angle determination, refinement and map calculation were performed using the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994). Although the map was noisy and difficult to interpret, the structures of the RNA and the N-terminal domain of TL5 were recognizable. Subsequently, three-wavelength multiple anomalous dispersion (MAD) data were collected from a SeMet crystal. The selenium sites obtained using MAD data were identical. The MAD phases were combined with phases obtained earlier (Table 1), improved and extended by density-modification procedures (Collaborative Computational Project,

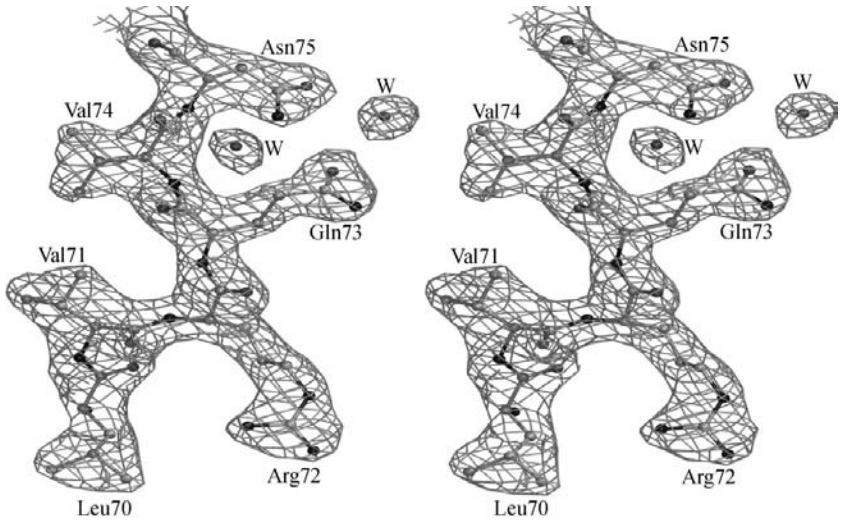

Figure 1
A stereoview showing the final $2F_o - F_c$ map at 2.3 Å resolution contoured at 1.5σ. Figures were produced with *MolScript* (Kraulis, 1991), *BobScript* (Esnouf, 1999), *WebLab ViewerPro* (WebLab ViewerLite 3.20; Molecular Simulations Inc.) and *POV-Ray* (Persistence of Vision Ray Tracer v3.02; http://www.povray.org).

Table 1
Data-collection, phasing and refinement statistics.

Values in parentheses are statistics for the highest resolution shell.

Data-collection and phasing statistics.

			SeMet (2)		
Data set	Native	SeMet (1)	λ_1	λ_2	λ_3
Data-collection statistics					
Space group	$P3_112$				
Unit-cell parameters (Å,°)	$a = b = 109.9$, $c = 137.5$, $\alpha = \beta = 90$, $\gamma = 120$				
λ (Å)	1.108	0.968	0.9794	0.9797	0.9184
Resolution (Å)	20.0–2.3	25–2.9	25.0–3.0	25.0–3.0	25.0–3.0
Total reflections	188440	111437	137093	150010	149411
Unique reflections	41886	20768	36713	36749	36747
Completeness (%)	98.8 (97.8)	98.3 (99.2)	99.9 (99.8)	100.0 (100.0)	100.0 (100.0)
$I/\sigma_i(I)$	17.0 (4.5)	9.3 (3.7)	7.5 (3.1)	11.0 (4.2)	17.3 (4.6)
$R_{sym}(I)$† (%)	5.2 (36.5)	8.7 (37.6)	12.6 (39.9)	8.0 (25.8)	7.6 (24.1)
Phasing statistics					
Number of sites		6	6	—	6
Phasing power‡					
Centric		0.62	0.61	—	1.02
Acentric		0.81	0.91	—	1.51
R_{Cullis}§					
Centric		0.82	0.82	—	0.66
Acentric		0.9	0.86	—	0.73
Anomalous		0.97	0.81	—	0.8
Overall figure of merit at 4 Å resolution	0.5080				

Refinement statistics

Resolution range (Å)	8.0–2.3
Reflections	40543
R factor (%)	20.8
R_{free} (%)¶	24.5
R.m.s. deviation	
Bonds (Å)	0.006
Angles (°)	1.2
Impropers (°)	1.07
Average B factors	
5S rRNA	30.1
TL5	32.3
Solvent	32.8

† $R_{sym} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ Phasing power $= \{\sum_{hkl} |F_H^{calc}(hkl)|^2 / \sum[|F_{PH}^{obs}(hkl)| - |F_{PH}^{calc}(hkl)|]\}^{1/2}$. § $R_{Cullis} = \sum_{hkl} ||F_{PH}^{obs}(hkl) \pm F_P^{obs}(hkl)| - F_H^{calc}(hkl)| / \sum_{hkl} |F_{PH}^{obs}(hkl) \pm F_P^{obs}(hkl)|$. ¶ Calculated for 10% of data in test set.

Number 4, 1994). Amino-acid sequences were obtained from the SWISS-PROT data bank (Bairoch & Apweiler, 2000).

2.3. Structure determination and refinement

The resulting solvent-flattened map was of good quality; however, no electron density was observed for the C-termini of both molecules in the asymmetric unit (residues 186–206). The map enabled the construction of a complete model of both RNA fragment and protein with the exception of 30 C-terminal residues (177–206). The initial model was built with the program *O* (Jones *et al.*, 1991) and subjected to several steps of crystallographic refinement using the *CNS* program (Brunger *et al.*, 1998) combined with manual rebuilding. The two molecules in the asymmetric unit were refined separately, although NCS restraints were used during the early stages of refinement. F_o maps were used to incorporate the C-terminal residues 177–185 in both molecules of the asymmetric unit. The progress of all refinement procedures was monitored by using 10% of the reflections to calculate a free R factor (R_{free}). Fig. 1 provides an example of the quality of the final $2F_o - F_c$ electron-density map. The final model, refined to an R factor of 20.8% (R_{free} = 24.5%) at 2.3 Å resolution, includes two complexes each of 185 amino acids and 40 nucleotides. 298 water molecules, 13 Cd^{2+} ions and 14 Mg^{2+} ions were found in the asymmetric unit. The model showed good quality (Table 1) as judged with the program *PROCHECK* (Laskowski *et al.*, 1993) and had no residues in the disallowed regions of the Ramachandran plot. The DALI server (Holm & Sander, 1993) at EMBL Heidelberg, Germany and the program *TOP* (Lu, 2000) were used to search for proteins with similar structural folds to the C-terminal domain of TL5.

3. Results

3.1. Structure of TL5

A ribbon diagram of *T. thermophilus* TL5 in complex with RNA is presented in Fig. 2(*a*). The protein consists of two domains joined at an angle of 90° to form an elongated L-shaped structure with approximate dimensions of 25 × 40 × 65 Å (Figs. 2*b* and 2*c*). The entire structure is well ordered, with the exception of the C-terminus (residues 177–185), which seems to be significantly flexible, resulting in weak electron density. The secondary structure is indicated in Fig. 2(*b*) and the position of secondary-structure elements along the amino-acid sequence is shown in Fig. 3.

The N-terminal domain spans residues 1–91. It has overall dimensions of 25 × 35 × 35 Å and contains a β-barrel formed by two approximately perpendicular β-sheets. One of these β-sheets is composed of two parallel (β_1, β_4) and one antiparallel (β_5) strands. Strand β_5 is bent and shared between the two sheets. The second β-sheet contains four β-strands (β_2, β_3, β_5 and β_6), with two antiparallel β-hairpins formed by strands β_2, β_3 and β_5, β_6. The connectivity scheme of the domain is $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4\beta_5\beta_6$ (Figs. 2*b* and 3). The β-barrel contains an extensive hydrophobic core built up by amino acids Leu5, Ala7, Leu18, Leu24, Pro25, Gly26, Met28, Val37, Val39, Phe44, Val56, Ile57, Leu70, Val71, Val86 and Phe88. Most of

these residues are completely inaccessible and are conserved in all sequences as hydrophobic, implying that the structure of the domain is conserved among different organisms. The domain contains 11 negatively and 18 positively charged residues, which form a network of hydrogen bonds and salt bridges contributing to the stabilization of the structure. Many of these residues make hydrogen bonds with main-chain atoms. The loop between β_5 and β_6 contains five positively charged residues and together with several basic residues of helix α_1 it forms a positively charged pole of the molecule.

The C-terminal domain spans residues 92–176. It is strongly elongated and has a cigar-like shape, with dimensions 22 × 22 × 50 Å. The polypeptide chain of the C-terminal domain is folded five times along the length of the domain. The connectivity scheme of the domain is $\beta_7\alpha_3\beta_8\beta_9\beta_{10}\beta_{11}\beta_{12}\beta_{13}$ (Figs. 2*b* and 3). The seven β-strands of the C-terminal domain form two antiparallel β-sheets (β_7, β_9, β_{12} and β_8, β_{11}, β_{13}) and one two-stranded parallel β-sheet (β_7, β_{10}). Strand β_9 follows β_8 through a bulge formed by His121 and Arg122, while strand β_{13} follows β_{12} after a loop formed by Ala164–Glu169. The bulge and the loop form a wide plateau in the middle of the domain stabilized by interactions between His121 and Glu169 mediated by two Cd^{2+} ions. An extended and tight hydrophobic core runs through the whole domain and includes residues Val100, Leu102, Phe104, Pro108, Leu117, Ile124, Val126, Val128, Pro134, Ile137, Val139, Val141, Leu144, Leu150, Ala152, Leu155, Leu157, Val161, Leu163, Val165, Ile171, Ala172, Val174 and Val175. This hydrophobic core together with an extensive net of hydrogen bonds involving main-chain atoms stabilizes the structure. The amino acids that build up this core are conserved in the CTC family of sequences as hydrophobic residues (boxed in Fig. 3). The turns connecting the β-strands are long and include helical structures of the 3_{10} type. The turn between strands β_7 and β_8 includes helix α_3 and together with the loop between β_{10} and β_{11} forms another pole of the molecule.

The overall conformation of the C-terminal domain is unique and has never been observed in other proteins. Nevertheless, a scan of the Protein Data Bank with the program *TOP* (Lu, 2000) revealed that 37 C^{α} positions of the C-terminal domain of TL5 overlap exceedingly well (r.m.s. = 1.47 Å) with a part of domain 2 of the RNA polymerase α-subunit amino-terminal domain from *E. coli* (Zhang & Darst, 1998; Zhang *et al.*, 1999). Overlapping residues of TL5 belong to the C-terminal part of the molecule (residues 125–130, 135–140, 148–164, 166–173) and form a complicated tertiary structure (Figs. 2*b* and 2*c*)

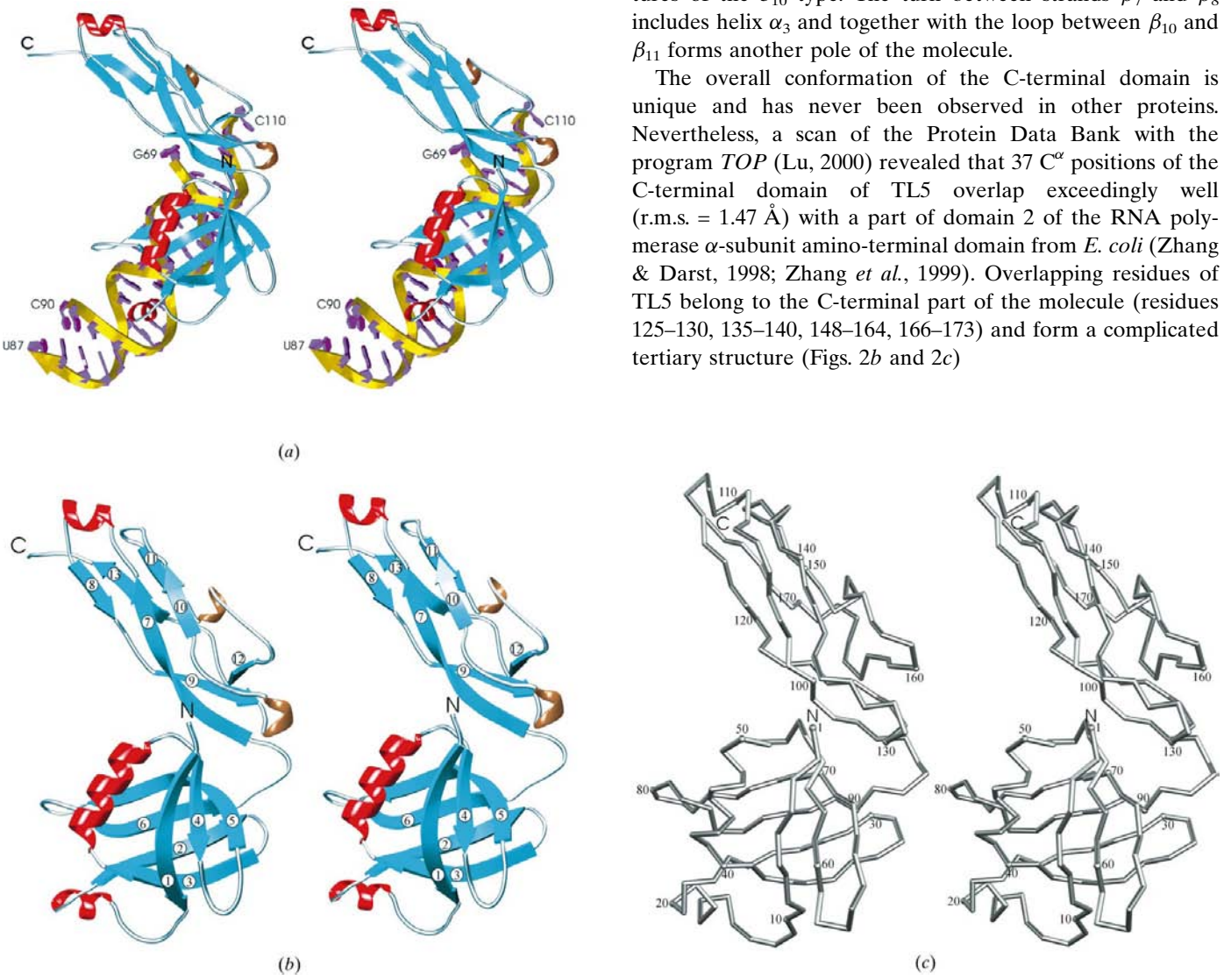

Figure 2
The overall structure of TL5 complexed with a fragment of 5S rRNA. (*a*) Stereo ribbon representation, showing the protein and the RNA. The phosphoribose backbone is in gold, bases are in magenta, β-strands in blue, α-helices in red and 3_{10}-helices in brown. (*b*) Stereoview of the TL5 model. The numbering scheme of the β-strands is shown. (*c*) Stereo C^{α} trace of the TL5 backbone with every tenth atom labelled and marked with a closed circle.

The main body of the domain contains 14 acidic and six basic residues. In addition to this, the C-terminus (residues 177–206) contains 12 glutamic acid residues. This makes the whole domain highly acidic. The C-terminus protrudes out of the body of the domain and is very flexible: the interpretable electron density only extends to Glu185. The sequences in Fig. 3 show that this C-terminal extension is highly variable in length and sequence. In some species it ends immediately after β_{13}.

Several hydrogen bonds and one salt bridge (between Glu94 and Arg31) link the two domains. Their relative position is also stabilized by a hydrophobic core which extends through the entire molecule and involves residues at the domain interface (Phe48, Leu70, Leu91, Val96, Met98, Ile133). The relative position of the domains is in addition stabilized by Cd^{2+} ions in the crystals.

The Cd^{2+} ions are coordinated by solvent and protein groups, usually with six ligands in the first coordination sphere of each ion (Fig. 4). The ligands are N atoms of histidines and O atoms of acidic residues or water molecules. Two Cd^{2+} ions were found to interact with each other through a solvent molecule. Such clusters could contribute to the stabilization of the C-terminal domain and interactions between protein molecules in the crystals and may explain the need for cadmium to be present in the crystallization media.

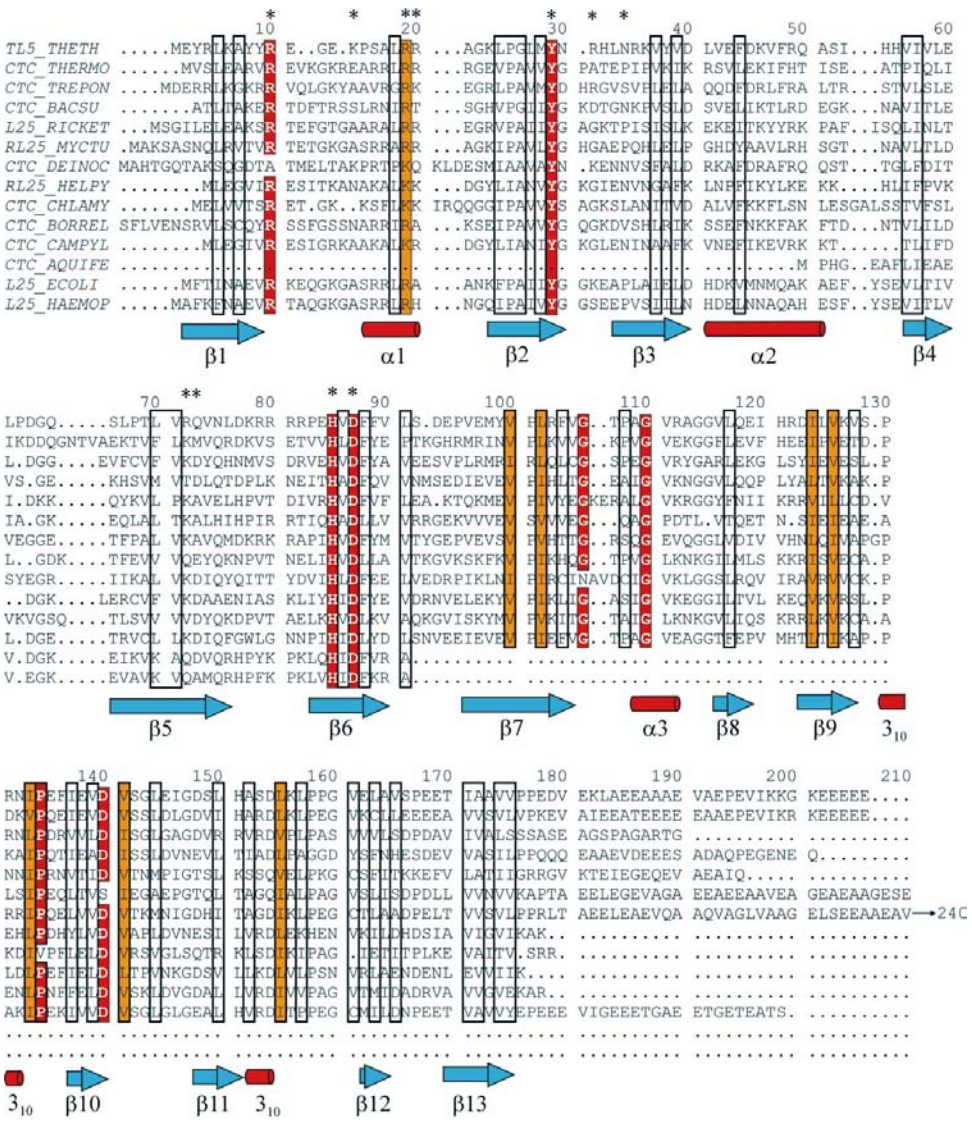

Figure 3
Sequence alignment of the CTC family of proteins from different sources. The numbering corresponds to TL5. Identical residues are drawn with a red background. The background of conserved residues (Ile/Val/Leu, Arg/Lys) is orange. Residues forming the hydrophobic core of TL5 are boxed. Residues of TL5 directly hydrogen bonded with RNA are marked with stars. The sequences are THETH, *Thermus thermophilus* (SWISS-PROT, P56930); THERMO, *Thermotoga maritima* (SWISS-PROT, Q9X1W2); TREPON, *Treponema pallidum* (GenBank/EMBL/DDBJ, AAC65357); BACSU, *Bacillus subtilis* (SWISS-PROT, P14194); RICKET, *Rickettsia prowazekii* (SWISS-PROT, Q9ZCV3); MYCTU, *Mycobacterium tuberculosis* (SWISS-PROT, P96385); DEINOC, *Deinococcus radiodurans* (GenBank/EMBL/DDBJ, AAF10004); HELPY, *Helicobacter pylori* (SWISS-PROT, P56078); CHLAMY, *Chlamydia pneumoniae* (SWISS-PROT, Q9Z6V7); BORREL, *Borrelia burgdorferi* (GenBank/EMBL/DDBJ, AAC67123); CAMPYL, *Campylobacter jejuni* (GenBank/EMBL/DDBJ, CAB72778); AQUIFE, *Aquifex aeolicus* (GenBank/EMBL/DDBJ, AAC06644); ECOLI, *Escherichia coli* (SWISS-PROT, P02426); HAEMOP, *Haemophilus influenzae* (SWISS-PROT, P45281).

The two TL5 molecules in the asymmetric unit have similar conformations, with an r.m.s. deviation between C^{α} atoms of 0.88 Å. Superposition of the N-terminal domains yields an r.m.s. deviation of 0.42 Å and reveals that the C-terminal domains of two molecules show different bends along the line connecting the centre of strand β_7 with the N-termini of strands β_9 and β_{10}. This results in displacements of up to 4.5 Å for C^{α} atoms which belong to the C-terminal pole of the molecule.

The N-terminal domain of TL5 is topologically very close to *E. coli* L25 despite the fact that the comparison of the protein sequences of *T. thermophilus* and *E. coli* reveals only 19% identity. However, there is some variation in the size of secondary-structure elements and loops. Thus, helix α1 contains only six residues in TL5 instead of nine residues in L25, whereas helix α_2 in TL5 is significantly longer and replaces helices α_2 and α_3 in L25. These differences have implications for protein–RNA interactions and are discussed below. The β-barrel structure in the two proteins is very similar. The C^{α} atoms may

be superimposed with an r.m.s. deviation of 0.71 Å. The number of positively and negatively charged side chains is approximately the same in both structures. However, there are 12 arginines and six lysines in TL5 instead of six arginines and 11 lysines in L25.

3.2. Structure of the 5S rRNA fragment

The fragment of *E. coli* 5S rRNA (nucleotides 69–87/90–110) used to obtain the RNA–protein complex includes helix IV, loop *E*, helix V and part of loop *A* (Fig. 5*a*). With the exception of U87 and G107, A108, A109, C110 at the 3′-ends and G69 at the 5′-end of the RNA chains, all other nucleotides are involved in base-pairing interactions, displaying a double-helical structure. The structure is very close to that of RNA in the complex with *E. coli* L25 (Stoldt *et al.*, 1999; Lu & Steitz, 2000). The most striking features, also observed in the unbound RNA structure (Correll *et al.*, 1997), are the base pairing and stacking geometry for the base pairs of loop *E*, which forms a bulge on the double-helical structure. The distances between C1′ atoms of base pairs at the central part of this loop and at its ends change from 14.7 to 9.4 Å. As a result, the minor groove of loop *E* is widened and the major groove is narrowed around the central three base pairs. The comparison of the two RNA molecules in the asymmetric unit reveals a small flexibility in the region 79–81/95–97 of helix IV, which causes a difference in the width of the major groove. The superposition of these two molecules yields an r.m.s. deviation of 1.21 Å for P atoms of the whole molecule, and superposition of loop *E* and helix V yields an r.m.s. deviation of 0.60 Å, while the superposition of Watson–Crick pairs of helix IV gives a value of 0.37 Å.

The nucleotide content and the three-dimensional structure of the RNA fragment used in this work are very close to that of the RNA fragment in the *E. coli* L25–RNA complex (Lu & Steitz, 2000). The only difference between the nucleotide sequences is substitutions at the 5′- and 3′-ends of both chains. Superposition of the atomic coordinates yields an r.m.s. deviation of 1.03 Å for all P atoms of the double-helical part, approximately the same as for the two molecules in the asymmetric unit of the TL5–RNA crystal.

3.3. Protein–RNA interactions

The structure of the TL5 complex shows that only the N-terminal domain of the protein interacts with the RNA fragment. Protein–RNA interactions involve ten of the 17 base pairs. Upon complex formation, 1680 Å^2 of protein and RNA surface is buried. A contiguous area formed by the side chains of nine residues belonging to the four-stranded β-sheet and four residues belonging to the α_1 region makes direct or water-mediated contacts with the minor groove of loop *E*. A schematic view of the RNA fragment and the interactions it makes with protein side chains is shown in Fig. 5(*a*). A majority of the interactions are made with the phosphoribose backbone. Five invariant and highly conserved residues (Arg10, Arg19, Tyr29, His85 and Asp87) interact with the phosphoribose backbone and/or with bases directly or *via* solvent. The conformation of the side chains of these residues is stabilized by interactions within the protein. Thus, Arg10 and Arg19 are hydrogen bonded to carbonyl groups of the main chain, whereas Asp87 is hydrogen bonded to Tyr29 and His85. The main characteristic feature of this conformation is that all protein and RNA atoms involved in interactions lie approximately in two planes, which are approximately parallel to the plane of the four-stranded β-sheet (Fig. 5*b*). The RNA-interacting groups of Arg31 and Arg72 furthermore increase the total area of the contact plane, which is one of the most stable parts of the protein structure. This is supported by molecular-dynamics simulations of the free N-terminal domain of TL5 in water medium at different temperatures (R. Fedorov, unpublished results).

The comparison of the TL5 and L25 complexes with fragments of *E. coli* 5S rRNA reveals some interesting differences (Fig. 6). Thus, although the number of RNA–protein contacts in both complexes is approximately the same, in the L25–RNA complex most of them are clustered in the helix α1 region (Fig. 3), which is inserted in the major groove of RNA. In TL5 this helix is shorter and is located outside the major groove. Most interactions in this case are made by residues from the four-stranded β-sheet and involve all β-strands (Fig. 3). In the *E. coli* L25–RNA complex the atoms of the five most conserved and two non-conserved residues which form the interactions also lie in a plane, although the area of the plane is about half that observed in the TL5 complex.

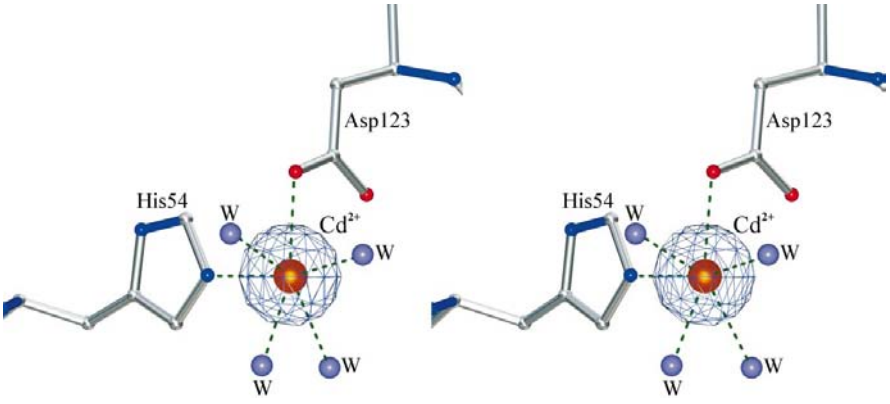

Figure 4
Stereoview of the first coordination sphere of a Cd^{2+} ion. The electron density of the final $2F_o - F_c$ map is contoured at 11σ. Water molecules are shown as violet spheres.

3.4. Crystal packing

The two complexes in the asymmetric unit interact entirely through RNA. The fragments are approximately perpendicular to each other and make contacts mainly through the surfaces of the minor grooves formed by loops *E* and helices IV.

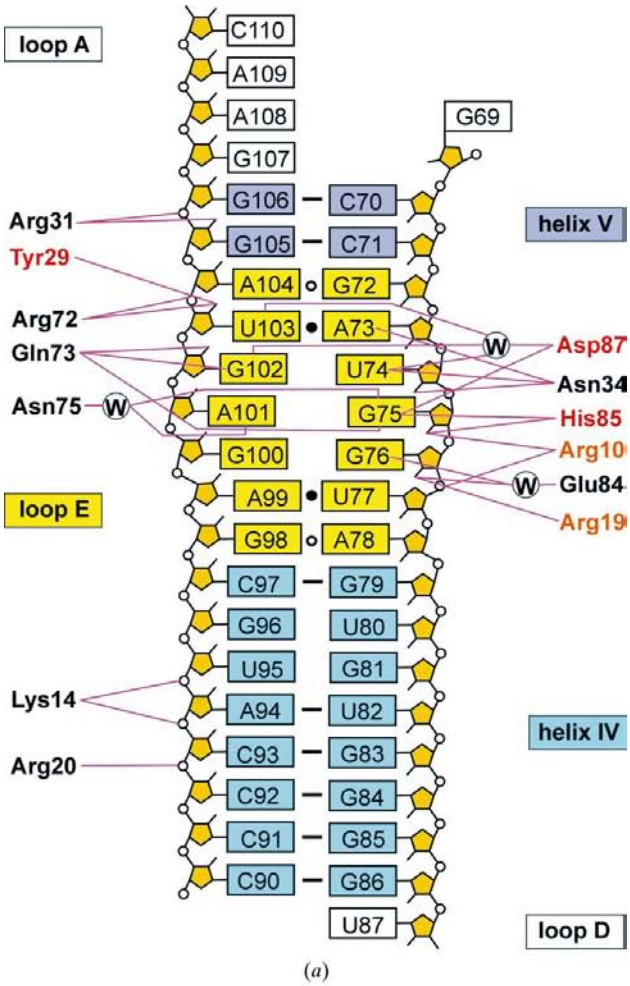

Figure 5
(*a*) Schematic illustration of the secondary structure of the 40 nt fragment of 5S rRNA and its contacts with TL5. The contacts are only with the side chains of the amino acids. The conserved amino-acid residues are coloured as in Fig. 3. (*b*) Detailed view of the protein–RNA interactions involving the most conserved amino-acid residues. The side-chain atoms which contact the RNA lie in a plane parallel to the flat surface formed by the contact atoms of the RNA.

Nucleotides A78, G79, U80 and G96, A99, G100 are involved in RNA–RNA interactions. There are 20 hydrogen bonds between the molecules, 13 of which include atoms of nucleotide bases. These nucleotides are not involved in RNA–protein interactions and are located on the side of the RNA opposite to the protein.

Symmetry-related complexes contact each other through the protein and RNA molecules. RNA–protein contacts are more stable. Thus, the base of G69 forms two hydrogen bonds with the main chain of Asn34, whereas the minor groove of the helix IV interacts with the loop between β_1 and α_1. The N- and C-terminal domains of symmetry-related protein molecules contact through the two Cd^{2+} ions, which are hydrogen bonded to Asp93 of one molecule and Asp154 and His151 of the other. Such interactions improve the crystal quality.

4. Discussion

The portion of 5S rRNA from *T. thermophilus* containing loop *E* is closely related (Walters & Erdmann, 1988) to that from *E. coli*. The nucleotides, which are different in the two species, are not involved in interactions with the protein. Thus, it is very likely that TL5 will bind to *T. thermophilus* and *E. coli* 5S rRNA in a similar way.

TL5 and L25 are functional analogues in ribosomes. Reconstruction experiments show that *E. coli* 50S ribosomal subunits with L25 replaced by TL5 are functionally active (Zvereva *et al.*, 2000). Earlier, it had also been shown that TL5 could displace *E. coli* L25 from its complex with 5S rRNA (Gongadze *et al.*, 1993). This may be caused by the larger flat RNA-interacting surface of TL5 compared with *E. coli* L25. A more stable complex could be a specific feature of the protein from *T. thermophilus*, which has to function at elevated temperatures. Upon formation of the complex, the RNA is hydrogen bonded directly to 11 residues of TL5. A substantial proportion of the protein and RNA contact surfaces is hydrophobic. Flat hydrophobic surfaces are efficient for complex formation, whereas specific arrangement of interacting atoms on such surfaces would be important for specific recognition.

The structure of TL5 described in this paper is the first structure of a full-size protein from the CTC family of stress proteins, which have not yet been characterized with regard to their structure and cell destination. Most proteins of the CTC family are composed of N- and C-terminal domains (Fig. 3). Only in *E. coli* and in *H. influenzae* are the L25 type of proteins with short sequences corresponding to the N-terminal domain of TL5 found. A contrary case is the CTC protein from *Aquifex aeolicus*, which contains a shortened N-terminal sequence (less than half of L25) linked to a full C-terminal domain. Sequence alignment and crystal structure analysis reveal that the hydrophobic cores which mainly stabilize the structures of both TL5 domains are highly conserved in the CTC family of proteins. Thus, the N- and C-terminal domains of all CTC proteins may fold with a topology close to that of TL5. As for the N-terminal domain, this is confirmed by the known L25 protein structure.

Since the most conserved residues of the N-terminal domain interact with 5S rRNA in both TL5 and L25, we suggest that all proteins of the CTC family can bind to the ribosome through their N-terminal domains. The function of the C-terminal domain is not known. However, it is known that it does not bind 5S rRNA (Gongadze *et al.*, 1996). The C-terminal domain is probably needed for interactions with some cellular components under stress conditions. Thus, the two domains of the CTC family of proteins may have widely different functions and may function independently in cells of different organisms. The isolated N-terminal fragment of ribosomal protein TL5, containing 91 residues, forms a stable globular protein which interacts specifically with 5S rRNA (Gongadze *et al.*, 1999).

The two domains of TL5 are joined at a relative angle of 90° to form an elongated L-shaped molecule, which reveals a marked similarity with the overall shape of tRNA. How would such a mimicry work in the cell? Several proteins are known to mimic tRNA and bind to a tRNA-binding site on the ribosome. We can propose that induced to high concentrations under stress conditions CTC may have the ribosome as its receptor as do several other stress proteins. Alternatively, the similarity of a part of the C-terminal domain of TL5 to part of the α-subunit of RNA polymerase may suggest that CTC interacts with some partner of RNA polymerase under stress conditions.

The structure of the large ribosomal subunit from the halophilic archaeon *Haloarcula marismortui* has been determined at high resolution (Ban *et al.*, 2000). The structure of the ribosome is conserved in evolution, but there are certain differences between bacterial and archaeal ribosomes. In particular, there is no protein homologous to TL5 or CTC in the Archaea. Bacterial ribosomes (from *T. thermophilus*) are now undergoing structural analysis and the low-resolution structure of the 70S ribosome was recently published (Cate *et al.*, 1999). The structure described here offers a possibility of fitting TL5 into the electron-density map of the *T. thermophilus* ribosome.

We thank Dr T. A. Steitz for the coordinates of the L25–RNA complex prior to publication, the staff of the EMBL Outstation in Hamburg for assistance during data collection and Dr Lucy Malinina for very helpful consultation on the molecular-replacement method. We are grateful to Elena Mudrik for excellent technical assistance. This work was supported by the Russian Academy of Sciences, the Russian Foundation for Basic Research and the Swedish National Science Research Council. The research of MBG and AL was supported in part by the International Research Scholar's award from the Howard Hughes Medical Institute.

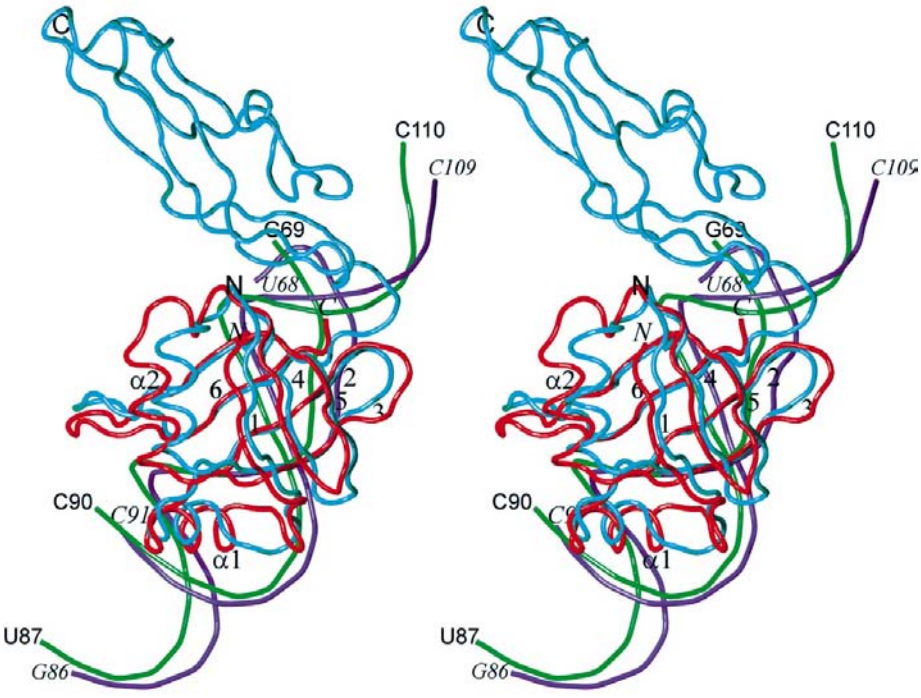

Figure 6
Stereoview of the superposition of the present structure and the structure of the L25–RNA complex. The comparison was made by superimposing only the C^{α} atoms of the four-stranded β-sheet within two structures, which gave an r.m.s. deviation of 1.0 Å. TL5 is in blue, L25 in red, the RNA backbone from the present structure is in green and that from the L25–RNA complex in magenta. The α-helices and β-strands of the N-terminal domain of TL5 are labelled.

References

Bairoch, A. & Apweiler, R. (2000). *Nucleic Acids Res.* **28**, 45–48.
Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000). *Science*, **289**, 905–920.
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. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cate, J. H., Yusupov, M. M., Yusupova, G. Zh., Earnest, T. N. & Noller, H. F. (1999). *Science*, **285**, 2095–2104.
Chen-Schmeisser, U. & Garrett, R. A. (1977). *FEBS Lett.* **74**, 287–291.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Correll, C. C., Freeborn, B., Moore, P. B. & Steitz, T. A. (1997). *Cell*, **91**, 705–712.
Esnouf, R. M. (1999). *Acta Cryst.* D**55**, 938–940.
Gongadze, G., Kashparov, I., Lorenz, S., Schroeder, W., Erdmann, V. A., Liljas, A. & Garber, M. (1996). *FEBS Lett.* **386**, 260–262.
Gongadze, G. M., Meshcheryakov, V. A., Serganov, A. A., Fomenkova, N. P., Mudrik, E. S., Jonsson, B.-H., Liljas, A., Nikonov, S. V. & Garber, M. B. (1999). *FEBS Lett.* **451**, 51–55.
Gongadze, G. M., Tishchenko, S. V., Sedelnikova, S. E. & Garber, M. B. (1993). *FEBS Lett.* **330**, 46–48.
Gryaznova, O. I., Davydova, N. L., Gongadze, G. M., Jonsson, B. H., Garber, M. B. & Liljas, A. (1996). *Biochimie*, **78**, 915–919.
Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lu, G. (2000). *J. Appl. Cryst.* **33**, 176–183.
Lu, M. & Steitz, T. A. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 2023–2028.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sedelnikova, S. E., Agalarov, S. C., Garber, M. B. & Yusupov, M. M. (1987). *FEBS Lett.* **220**, 227–230.
Stoldt, M., Wöhnert, J., Görlach, M. & Brown, L. R. (1998). *EMBO J.* **17**, 6377–6384.
Stoldt, M., Wöhnert, J., Ohlenschläger, O., Görlach, M. & Brown, L. R. (1999). *EMBO J.* **18**, 6508–6521.
Völker, U., Engelmann, S., Maul, B., Riethdorf, S., Völker, A., Schmid, R., Mach, H. & Hecker, M. (1994). *Microbiology*, **140**, 741–752.
Walters, J. & Erdmann, V. A. (1988). *Nucleic Acids Res.* **16**(*Suppl.*), 1–70.
Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. & Darst, S. A. (1999). *Cell*, **98**, 811–824.
Zhang, G. & Darst, S. A. (1998). *Science*, **281**, 262–266.
Zvereva, M. I., Shpanchenko, O. V., Nierhaus, K. H. & Dontsova, O. A. (2000). *Dokl. Akad. Nauk RAS*, **374**, 830–832.