Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of ribosomal protein L30 from *Thermus thermophilus* at 1.9 Å resolution: conformational flexibility of the molecule

Roman Fedorov,[a] Natalia Nevskaya,[a] Alphia Khairullina,[a] Svetlana Tishchenko,[a] Albert Mikhailov,[b] Maria Garber[a] and Stanislav Nikonov[a]*

[a]Institute of Protein Research, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia, and [b]Institute of Crystallography, Russian Academy of Sciences, Leninskii Prospekt, 59, 117333 Moscow, Russia

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

Received 30 March 1999
Accepted 26 July 1999

PDB Reference: ribosomal protein L30, 1bxy.

The crystal structure of ribosomal protein L30 from the extreme thermophilic bacterium *Thermus thermophilus* has been determined at 1.9 Å resolution. The crystals are trigonal and belong to space group $P3_221$, with unit-cell parameters $a = b = 63.5$, $c = 77.8$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$ and two molecules per asymmetric unit. The structure was solved by the molecular-replacement method with *AMoRe* and refined with *X-PLOR* to an R value of 20.3% and an $R_{\rm free}$ of 25.3% in the resolution range 8–1.9 Å. Detailed analyses of the structures of the two molecules in the asymmetric unit and comparison of *T. thermophilus* L30 structure with the structure of homologous L30 from *Bacillus stearothermophilus* reveal two flexible regions at opposite ends of the rather elongated molecule. Such flexibility could be important for the protein fitting in the ribosome. A comparison with *B. stearothermophilus* L30 shows a higher number of salt bridges and unbound positively charged residues and an increased accessible hydrophobic area on the surface of *T. thermophilus* L30. This could contribute to the stability of both the extreme thermophile protein and the ribosome as a whole.

1. Introduction

Ribosomal protein L30 is one of the smallest ribosomal proteins, with a molecular mass of about 7 kDa. Its binding site on the ribosomal RNA is unknown, but cross-linking experiments indicate that elongation factor T_u can be cross-linked to an extended ribosome neighbourhood containing L30 (San Jose *et al.*, 1976). The crystal structure of L30 from a moderate thermophile *B. stearothermophilus* was determined in 1986 by the multiple isomorphous replacement method at 2.5 Å resolution (Wilson *et al.*, 1986).

T. thermophilus is a Gram-negative aerobic eubacterium which can grow at higher temperatures than *B. stearothermophilus*. It is known that the proteins purified from such extreme thermophiles are more stable than their analogues from moderate thermophiles (Merkler *et al.*, 1981).

Protein structures determined by X-ray crystallography are usually described in terms of their atomic coordinates. This gives the impression that proteins are rigid molecules. However, many experiments indicate that proteins are quite flexible. In general, there is an equilibrium between a large number of possible conformations (Lillemoen *et al.*, 1997). Homologous proteins from different sources often crystallize in different space groups, which results in the molecules being in different environments (Hosaka *et al.*, 1997; Wimberly *et al.*, 1997). Intermolecular interactions may cause local conformational changes in flexible parts of the structure. Identifica-

Table 1
Data-collection statistics.

Values for the highest resolution shell are shown in parentheses.

Space group	$P3_221$
Unit-cell parameters (Å, °)	$a = b = 63.5$, $c = 77.8$; $\alpha = \beta = 90$, $\gamma = 120$
Data set 1 (Syntex $P2_1$)	
Resolution range (Å)	80–3.0
Number of unique reflections	4064
Completeness (%)	100
R_{sym}	4.7
Data set 2 (MAX-II synchrotron)	
Resolution range (Å)	10–1.9 (1.94–1.90)
Number of unique reflections	13239 (810)
Redundancy	2.0 (2.1)
Completeness	93.3 (92.7)
R_{merge}	5.1 (22)
Combined data set	
Resolution range (Å)	80–1.9
Number of unique reflections	13880 (810)
Completeness (%)	96.9 (92.7)

tion of such flexible regions may be of great importance. In the case of ribosomal proteins, this may be useful for understanding rRNA–protein and protein–protein interactions in the ribosome (Unge *et al.*, 1997).

In this paper, we report the refined crystal structure of L30 from *T. thermophilus* at 1.9 Å resolution and compare the two crystallographically independent molecules with each other as well as with *B. stearothermophilus* L30.

2. Experimental

2.1. Crystallization and data collection

The crystallization conditions and preliminary crystallographic studies of L30 from the extreme thermophile *T. thermophilus* have been published previously (Shikaeva *et al.*, 1993). The crystals diffracted to 2.3 Å resolution, but the protein was difficult to isolate from ribosomes on a preparative scale. Cloning and sequencing the *spc* operon from the *T. thermophilus* chromosome (Vysotskaya *et al.*, 1997) enabled us to clone the gene encoding L30 and obtain a strain overproducer for this protein (Khairullina *et al.*, 1997).

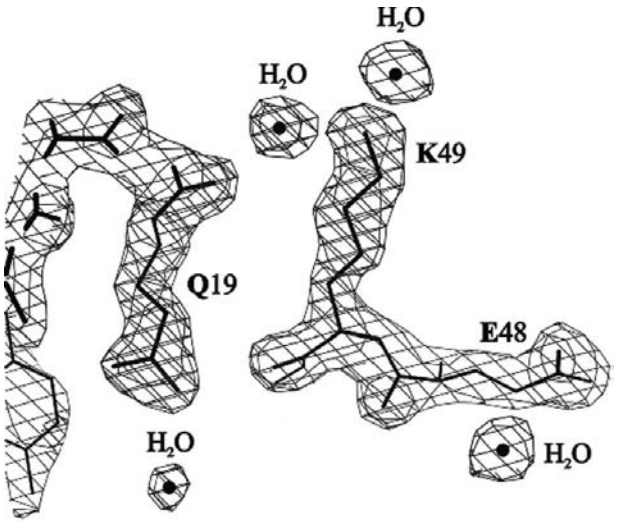

Figure 1
A portion of the $3F_o - 2F_c$ electron density in the region including residues Gln19, Glu48 and Lys49. The map is contoured at the 1.6σ level.

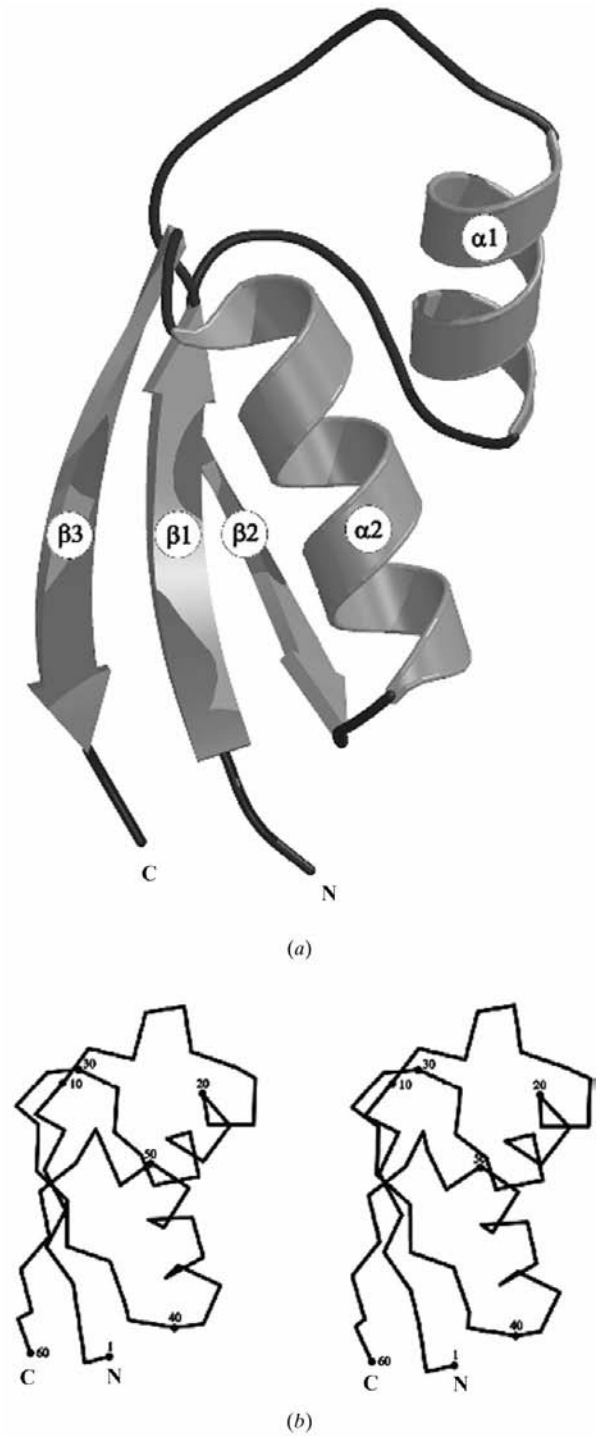

Figure 2
(*a*) Schematic representation of the structure of the ribosomal protein L30 from *T. thermophilus*. Figs. 2(*a*), 4 and 5 were prepared with *MOLSCRIPT* (Kraulis, 1991) and rendered with *Raster3D* (Merritt & Bacon, 1997). (*b*) Stereoview of the C^{α} trace of L30 from *T. thermophilus*.

Table 2
Refinement statistics.

Resolution range (Å)	8–1.9
Number of residues included†	120
Number of non-H atoms†	1030
Number of water molecules†	78
R factor (%)	20.3
R_{free}‡ (%)	25.3
Mean *B* factor (Å^2)	39.6
R.m.s. deviation	
Bond lengths (Å)	0.019
Bond angles (°)	3.079

† Two crystallographically independent molecules in an asymmetric unit. ‡ Calculated for 10% of data in test set (1380 reflections).

Crystals of the recombinant protein were obtained under conditions described for the protein isolated from ribosomes (Shikaeva *et al.*, 1993). The crystal space group was $P3_221$, with unit-cell parameters $a = b = 63.5$ Å, $c = 77.8$ Å, $\alpha = \beta = 90$, $\gamma = 120°$ and two molecules in the asymmetric unit.

Two native data sets were used. The first was collected at 80–3.0 Å resolution with the four-circle diffractometer Syntex $P2_1$ at the Institute of Crystallography, Moscow, Russia. The second native data set was collected at 10–1.9 Å resolution at beamline BL711 at the MAX-II synchrotron, Lund, Sweden. The latter diffraction data were processed with the *DENZO* and *SCALEPACK* programs (Otwinowski & Minor, 1996). Both data sets were scaled and combined with the *CTAB* program (Nikonov, unpublished work). The characteristics of all data sets are summarized in Table 1.

2.2. Molecular replacement and refinement

Comparison of the *T. thermophilus* and the *B. stearothermophilus* L30 amino-acid sequences revealed 45% identity. This allowed the use of the molecular-replacement method for solving the phase problem. The structure of L30 from *T. thermophilus* was solved with the *AMoRe* package (Navaza, 1994) using the polyalanine model of L30 from *B. stearothermophilus* as a search model. From crystal density considerations (Matthews, 1968), there could be two or three molecules in the asymmetric unit. Peaks in the cross-rotation function used for the two-body translation function (TF) search with a subsequent rigid-body refinement gave a high-contrasted peak with a correlation coefficient of 71.8 and an *R* factor of 40%, while a second peak had corresponding values of 53.9 and 48.5%. A three-body TF search did not reveal any clear peak; two molecules were therefore assumed in the asymmetric unit.

The model of L30 from *T. thermophilus* was subjected to several rounds of model rebuilding with *O* (Jones *et al.*, 1991) and crystallographic refinement with *X-PLOR* (Brünger, 1992). The two molecules in the asymmetric unit were refined separately, although NCS restraints were used during the early stages of refinement. R_{free} was calculated throughout the refinement process, using the same set of test reflections each time. 78 water molecules were localized in the later stages of refinement (Fig. 1). Statistics of the final model assessed with *PROCHECK* (Laskowski *et al.*, 1993) are given in Table 2. Amino-acid sequences were obtained from the SWISS-PROT data bank (Bairoch & Boeckmann, 1993) and were aligned using *CLUSTAL V* (Higgins, 1994).

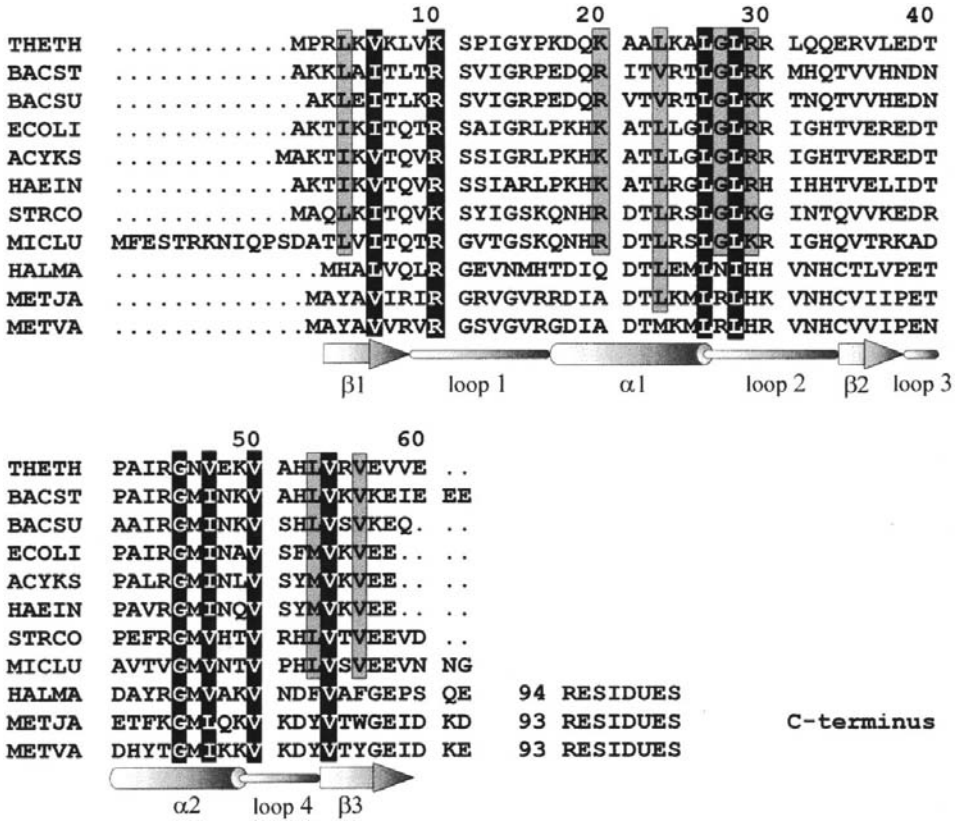

Figure 3
Alignment of L30 sequences. The upper eight sequences are from bacteria; the others are from archaea. The numbering and the location of secondary-structure elements correspond to *T. thermophilus* structure (THETH). Strictly conserved residues or very conservative substitutions (Val/Ile/Leu, Lys/Arg) are indicated with black bars when this applies to all sequences and with grey bars when this applies to bacteria. The abbreviations are as follows: BACST, *B. stearothermophilus*; BACSU, *B. subtilis*; ECOLI, *Escherichia coli*; ACYKS, *Acyrthosiphon kondoi* symbiotic bacterium; HAEIN, *Haemophilus influenzae*; STRCO, *Streptomyces coelicolor*; MICLU, *Micrococcus luteus (M. lysodeikticus)*; HALMA, *Haloarcula marismortui (Halobacterium marismortui)*; METJA, *Methanococcus jannaschii*; METVA, *Meth. vannielii*.

3. Results and discussion

3.1. Structure of the protein

The model of L30 from *T. thermophilus* contains 60 amino acids, giving a total of 476 non-H atoms. The overall view of the structure of *T. thermophilus* L30 and the C^{α}-atom trace are presented in Fig. 2. The location of secondary-structure elements within the sequence is shown in Fig. 3. The protein contains a three-stranded antiparallel β-sheet flanked by two α-helices on one side (referred to as the 'α-side' of the molecule) and exposed on the other (referred to as the 'β-sheet side' of the molecule). The protein includes the so-called split β–α–β motif (Orengo & Thornton, 1993) or abc/αd unit (Efimov, 1994) formed by the β-sheet and helix α2. Helix α1 with two adjacent loops 1 and 2, referred to as the 'loop' end (Wilson *et al.*, 1986), is separated from the split β–α–β motif and is located at one end of the rather elongated L30 molecule. The N and C termini are close to each other on the opposite end of the molecule. Following Wilson *et al.* (1986), this region is referred to as the 'termini' end.

The hydrophobic core of the molecule contains the following residues: Val6, Leu8, Ala22, Leu23, Leu26, Leu28, Val50, Leu53 and Val54. Three of them (Leu26, Val50 and Val54) are invariable in all the known L30 sequences and stabilize the mutual arrangement of the C-termini of both helices and the N-terminus of strand β3. The C^{α} atoms of these residues form an approximately isosceles triangle (with their side chains in its plane) lying between the split β–α–β motif and the loop end of the molecule. A number of hydrophobic residues are partially exposed and most of them form five clearly visible patches on the molecular surface. These patches are located in the loop end (Val9, Pro12, Ile13 and Leu31), in the termini end (Met1 and Pro2), on the β-sheet side (Val36 and Val59) and on the α-side of the molecule (Leu4, Val47, Ala51, Val56, and Val58; Ala21; Ala25, Leu37, Pro41, Ala42 and Ile43). The last two sets of residues are separated by helix α2. All these patches are involved in intermolecular crystal contacts.

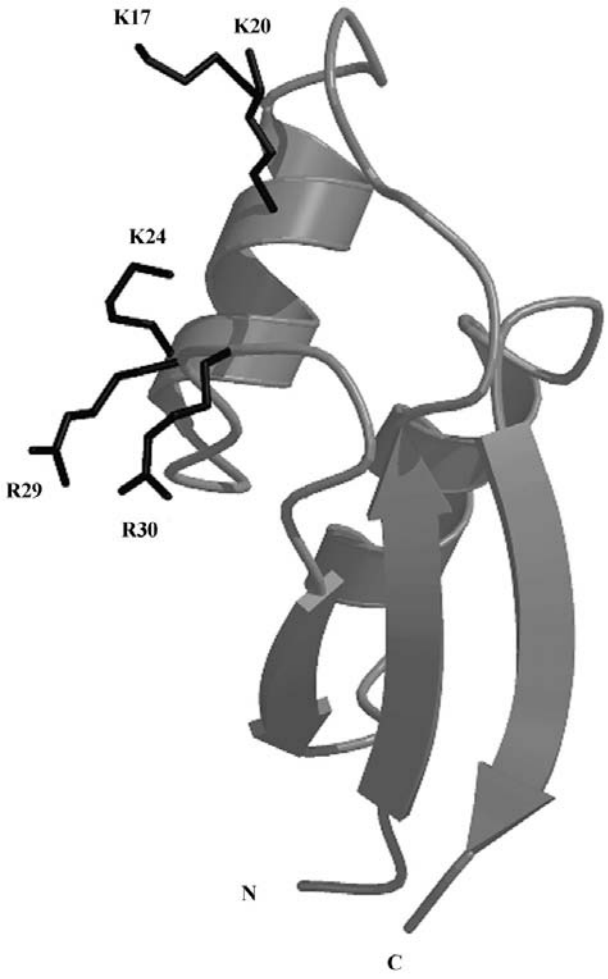

Figure 4
Conserved cluster of unbound positively charged residues which could be involved in interaction with rRNA.

Five of the 13 positively charged residues are involved in intramolecular ion-pairing and stabilize the split β–α–β motif. Two of these residues participate in the isolated ion pairs (Arg3$\cdots$Glu38 and Asp39$\cdots$Arg44), while the rest contribute to a network of ion pairs. This network is formed by the extensive alternating charge cluster of the side chains of residues Lys5, Lys7, Glu34, Arg55 and Glu57. This cluster is flat and parallel to the central part of the exposed side of the β-sheet.

The other eight positively charged residues are not involved in intramolecular interactions. Most of them (Lys17, Lys20, Lys24, Arg29 and Arg30) are located in the loop end and contribute to an extensive positively charged cluster on the protein surface (Fig. 4). Moreover, all these residues except Lys17 are conserved in most of the known Leu30 sequences.

The two molecules in the asymmetric unit of *T. thermophilus* L30 crystals are rotated by 175° with respect to each other and have different crystal contacts, which induces some conformational changes. The greatest differences in main-chain conformation occur at the termini ends of the molecules. The root-mean-square (r.m.s.) deviation between the C^{α} atoms excluding residues 1–3, 36–43 and 58–60 is 0.43 Å, in contrast to 0.97 Å for all C^{α} atoms. The β-sheets of the two molecules show different bends along the line connecting residues Arg35, Val6 and Val56, resulting in a maximal displacement of the C^{α} atoms in the termini end of up to 2.3 Å (Fig. 5*a*). The different crystal contacts of the molecules in the asymmetric unit cause changes in the side-chain orientation of the residues involved. Thus, the His52 ring of one of the molecules is at 90° to that of other molecule, owing to the hydrogen bond with Asp18 of the symmetry-related molecule.

3.2. Comparison of L30 from *T. thermophilus* and *B. stearothermophilus*

The overall three-dimensional structure of L30 from *T. thermophilus* is closely related to that of *B. stearothermophilus*. Both structures were superimposed using the least-squares option of the program *O*. This revealed that the main difference is in the mutual arrangement of the split β–α–β motif and the loop end (Fig. 5*b*). Helix α1 is rotated

around its axis by approximately 20° and is shifted perpendicularly to this axis and parallel to the β-sheet surface by 1.5 Å.

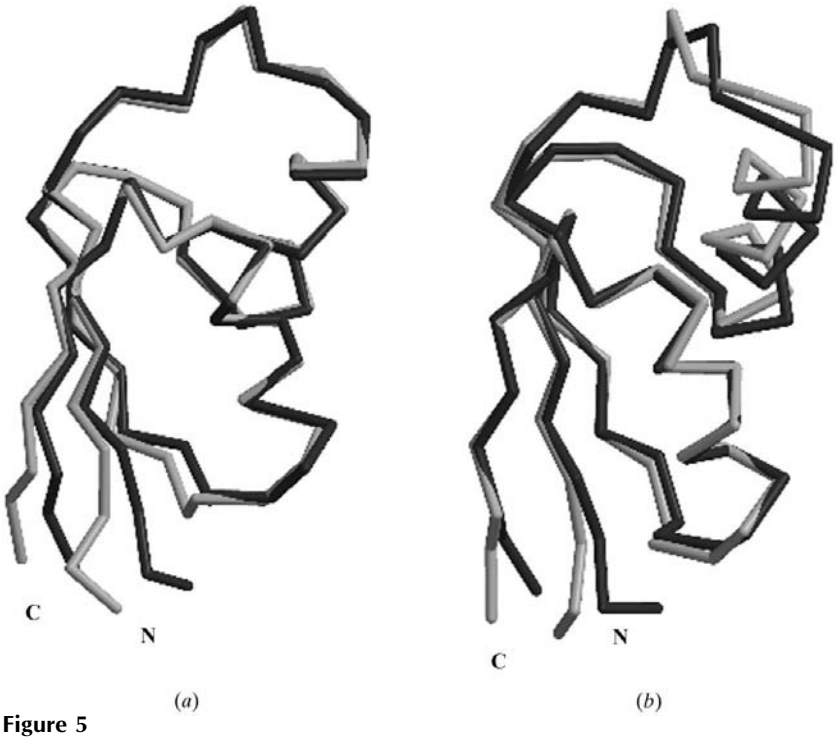

Figure 5
A ribbon representation of the *T. thermophilus* L30 molecule (black) superimposed (*a*) onto the second molecule in the asymmetric unit and (*b*) onto the *B. stearothermophilus* molecule. Two flexible sites in the protein are located at the loop end and the termini end of the molecule.

However, the spatial structures of the split β–α–β motif and part of the loop end (residues 14–27) are strictly conserved, with C^{α} r.m.s. deviations of 0.38 and 0.35 Å, respectively, while superposition of the C^{α} atoms of the whole molecule yields an r.m.s. deviation of 1.08 Å.

L30 from *T. thermophilus* contains 20 charged residues instead of the 17 found in the *B. stearothermophilus* protein. A cluster of unbound positive charges in both proteins is located at the loop end, but this cluster is more extensive in *T. thermophilus* L30 (Fig. 6). Moreover, L30 from *T. thermophilus* possesses an extra extensive alternating charge cluster of residues bound by a network of ion pairs. These factors could contribute to the stabilization of the structure of the protein from the extreme thermophile.

The locations of the four hydrophobic patches on the *T. thermophilus* L30 surface are very similar to those on the *B. stearothermophilus* protein. The extra patch at the termini end of L30 from *T. thermophilus* essentially increases the hydrophobic accessible surface area of this molecule. This may allow additional interactions between L30 and neighbouring protein(s) within the ribosome of the extreme thermophile to occur. Analysis of the crystal packing of L30 molecules from *T. thermophilus* and *B. stearothermophilus*

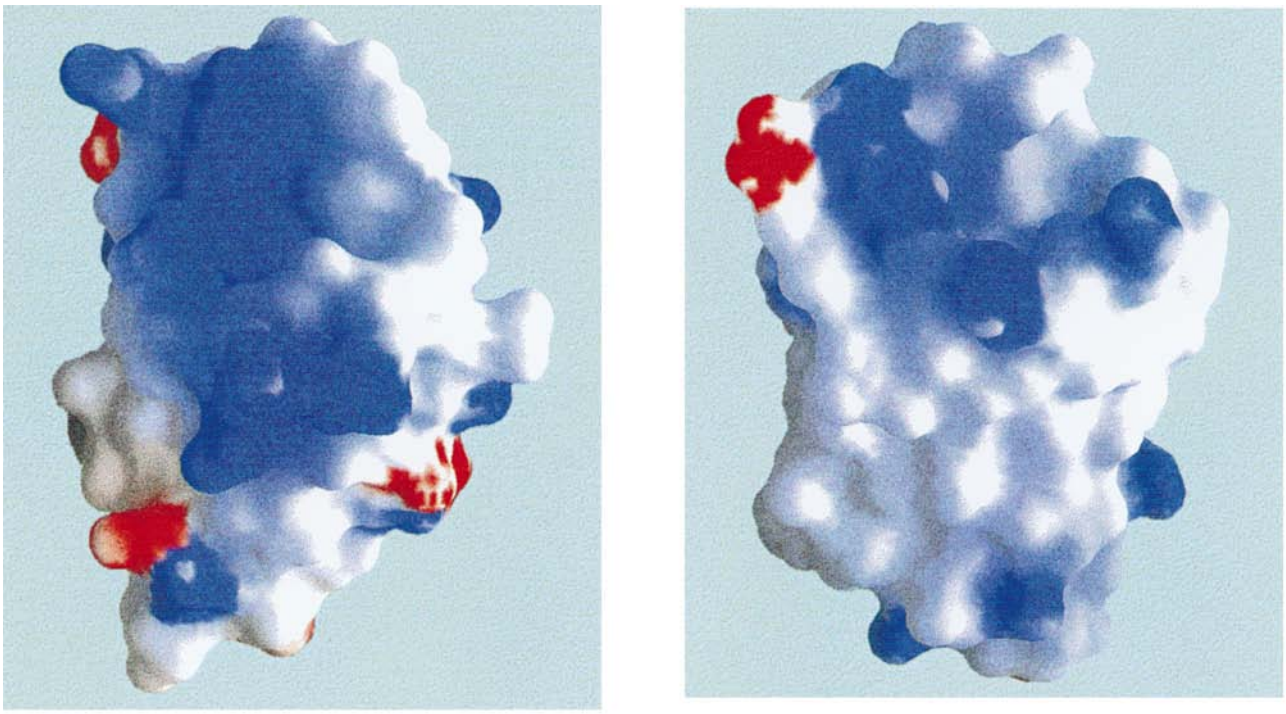

Figure 6
The surface electrostatic potential (blue for positive and red for negative) calculated with *GRASP* (Nicholls *et al.*, 1991) for (*a*) *T. thermophilus* L30 and (*b*) *B. stearothermophilus* L30. Enlargement of the positive cluster is clearly seen in *T. thermophilus* L30.

showed that all the hydrophobic patches on the molecular surfaces are involved in intermolecular interactions and that the contacts through the hydrophobic patches appear to be the most extensive.

The two molecules in the asymmetric unit of *T. thermophilus* L30 crystals interact through the identical hydrophobic patches on the α-side of the molecule and form an extensive intermolecular hydrophobic core. This core, together with the intramolecular cores of both molecules, creates a joint extended hydrophobic core of the dimer, which is packed so tightly that it appears like a single molecule (Fig. 7). The dimer is additionally stabilized by electrostatic interactions of charged residues in the contact region.

L30 from *B. stearothermophilus* crystallized in space group $P4_32_12$ with one molecule in asymmetric unit. Crystal packing of this protein reveals another type of intermolecular contact. Two symmetry-related molecules interact through a joint extended β-sheet containing six antiparallel β-strands, which looks like a half a cylinder with four α-helices packed on its outside.

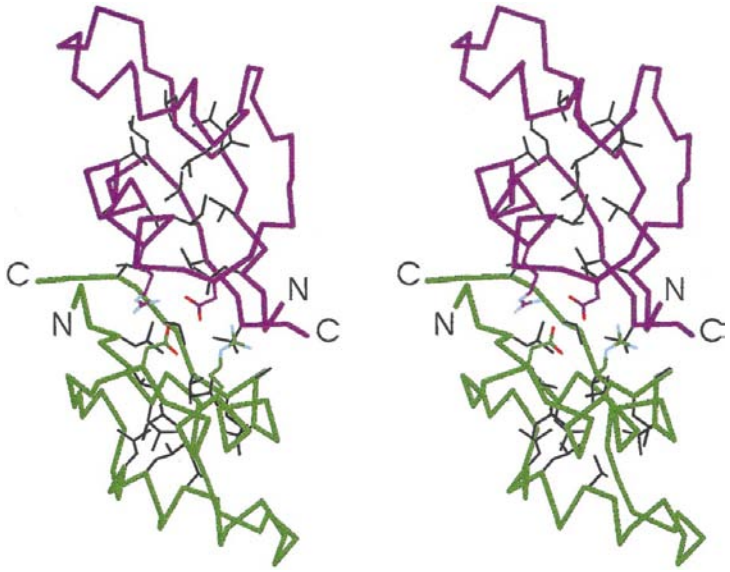

Figure 7
A dimer in the asymmetric unit of *T. thermophilus* L30 crystals possesses an extended hydrophobic core (black) formed by core residues of both molecules as well as by surface hydrophobic residues in the contact region. Charged residues cover hydrophobic residues in this region. Figs. 7 and 8 were prepared with *MOLSCRIPT* (Kraulis, 1991).

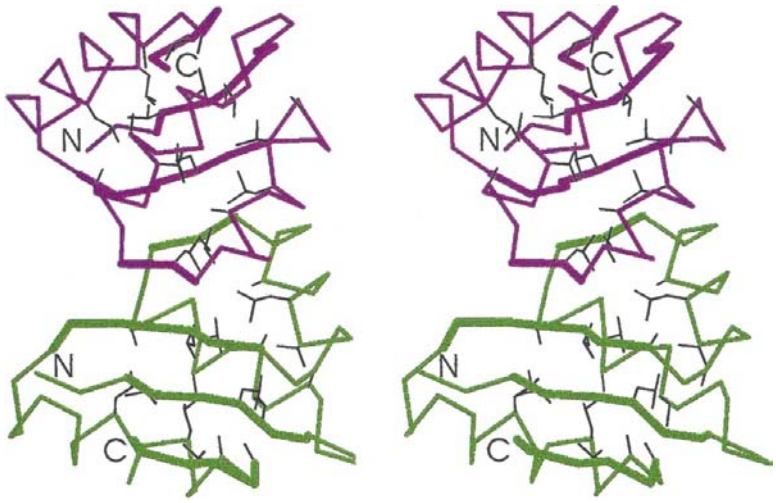

Figure 8
The tightest contact between symmetry-related molecules (magenta and green) in CTF crystals is provided by a joint extended hydrophobic core (black) together with the two extended β-sheets (bold).

3.3. RNA-binding sites in L30

Ribosomal protein L30 is not a core ribosomal protein. Its binding site on the ribosome probably involves other ribosomal proteins and/or rRNA. Superposition of the L30 structures from two bacteria reveal two flexible regions located on the opposite ends of the molecule. The first is the loop-end, including helix α1 with adjacent loops 1 and 2. The second region is the termini end of the molecule, including loop 3 and the part of the β-sheet which is differently bent in the two molecules within the asymmetric unit of L30 from *T. thermophilus*. This conformational flexibility may be important for L30 to fit to other components of its binding site in the ribosome.

It has previously been proposed that potential RNA-binding sites on the proteins involve patches of conserved positively charged residues (Wilson *et al.*, 1986; Davies *et al.*, 1996). Such patches are located in the flexible loop regions in the *T. thermophilus* and *B. stearothermophilus* L30 structures. These regions are likely to be involved in rRNA binding. The cluster of unbound positively charged residues is more extensive in L30 from *T. thermophilus* and could add to the stability of the protein–RNA binding. An increase in the number of salt bridges and unbound positively charged residues and an increase in the accessible hydrophobic area on the surface of the L30 protein from *T. thermophilus* could contribute to the stability of both the extreme thermophile protein and the ribosome as a whole.

3.4. Protein–protein interactions

Protein–protein interactions are often mediated through exposed hydrophobic patches. The crystal packing of two homologous L30 proteins clearly illustrates this. There is a number of exposed hydrophobic patches closely related in both proteins. Despite the different space groups of the crystals of L30 from the two bacteria, all these patches are involved in intermolecular crystal contacts.

Moreover, formation of a tightly packed dimer in the asymmetric unit of *T. thermophilus* L30 crystals

is accompanied by amalgamation of the hydrophobic cores of both molecules through the hydrophobic residues in the contact region. Such a type of protein–protein association could be used by L30 in interactions with neighbouring ribosomal protein(s).

L30 proteins and the C-terminal fragment of L7/L12 (CTF; accession code 1ctf) from *E. coli* (Liljas, 1982; Leijonmarck & Liljas, 1987) have very close topology. These proteins contain the so-called split β–α–β motif, which was subsequently found in a number of other ribosomal proteins: S6, L1, L6, L9 and L22. This motif is suggested to be responsible for interaction with rRNA (Liljas & Al-Karadaghi, 1997). The striking similarity of L30 from *B. stearothermophilus* to CTF has previously been discussed in detail (Wilson *et al.*, 1986).

Detailed analysis of CTF crystal packing reveals intermolecular association through the extended joint hydrophobic core, similar to that found in *T. thermophilus* L30 crystals. Moreover, this core is enclosed by two extended β-sheets (Fig. 8). Protein–protein contact, where a β-sheet extends from one molecule into another, was found in *B. stearothermophilus* L30 crystals. Thus, both types of protein–protein interactions, observed in different L30 crystals, co-exist within one and the same contact region between symmetry-related molecules in the CTF crystals.

Protein–protein contacts through the joint extended β-sheets in ribosomal proteins were subsequently observed in crystals of S8 from *T. thermophilus* (Nevskaya *et al.*, 1998) and L14 from *B. stearothermophilus* (Davies *et al.*, 1996). Therefore, this type of protein–protein interaction may be common for some ribosomal proteins and may be realised within the ribosome.

We thank Anders Svensson for his kind help in data collection at MAX-II synchrotron, Salam Al-Karadaghi for helpful discussion and Anders Liljas for support. We thank Stephen White for the permission to use atomic coordinates of *B. stearothermophilus* L30 and valuable remarks. The work was supported by the Russian Foundation for Basic Research, Russian Academy of Sciences. The studies of MG were supported in part by an International Research Scholar's award from the Howard Hughes Medical Institute.

References

Bairoch, A. & Boeckmann, B. (1993). *Nucleic Acids Res.* **21**, 3093–3096.
Brünger, A. T. (1992). *X-PLOR Version 3.1. A System for Crystallography and NMR.* Yale University, New Haven, Connecticut, USA.
Davies, C., White, S. W. & Ramakrishnan, V. (1996). *Structure*, **4**, 55–66.
Efimov, A. V. (1994). *FEBS Lett.* **355**, 213–219.
Higgins, D. G. (1994). *Methods Mol. Biol.* **25**, 307–318.
Hosaka, H., Nakagawa, A., Tanaka, I., Harada, N., Sano, K., Kimura, M., Yao, M. & Wakatsuki, S. (1997). *Structure*, **5**, 1199–1208.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Khairullina, A. R., Scherbakov, D. V., Tischenko, S., Nikonov, S. V. & Garber, M. B. (1997). *Biochemistry (Moscow)*, **62**, 221–224.
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.
Leijonmarck, M. & Liljas, A. (1987). *J. Mol. Biol.* **195**, 555–580.
Liljas, A. (1982). *Prog. Biophys. Mol. Biol.* **40**, 161–228.
Liljas, A. & Al-Karadaghi, S. (1997). *Nature Struct. Biol.* **4**, 767–771.
Lillemoen, J., Cameron, C. S. & Hoffman, D. W. (1997). *J. Mol. Biol.* **268**, 482–493.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Merkler, D. J., King Farrington, G. & Wedler, F. C. (1981). *Int. J. Pept. Protein Res.* **18**, 430–442.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **227**, 505–524.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Nevskaya, N., Tishchenko, S., Nikulin, A., Al-Karadaghi, S., Liljas, A., Ehresmann, B., Ehresmann, C., Garber, M. & Nikonov, S. (1998). *J. Mol. Biol.* **279**, 233–244.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Orengo, C. A. & Thornton, J. M. (1993). *Structure*, **1**, 105–120.
Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **276**, 307–326.
San Jose, C., Kurland, C. G. & Stoffler, G. (1976). *FEBS Lett.* **71**, 133–137.
Shikaeva, O. S., Sedelnikova, S. E., Fomenkova, N. P., Nikonov, S. V. & Garber, M. B. (1993). *J. Mol. Biol.* **230**, 1309–1310.
Unge, J., Al-Karadaghi, S., Liljas, A., Jonsson, B.-H., Eliseikina, I., Ossina, N., Nevskaya, N., Fomenkova, N., Garber, M. & Nikonov, S. (1997). *FEBS Lett.* **411**, 53–59.
Vysotskaya, V. S., Shcherbakov, D. V. & Garber, M. B. (1997). *Gene*, **193**, 23–30.
Wilson, K. S., Appelt, K., Badger, J., Tanaka, I. & White, S. W. (1986). *Proc. Natl Acad. Sci. USA*, **83**, 7251–7255.
Wimberly, B. T., White, S. W. & Ramakrishnan, V. (1997). *Structure*, **5**, 1187–1198.