Volume 210, number 1, 85–90 FEB 04316 January 1987

The complete amino acid sequences of the 5 S rRNA binding proteins L5 and L18 from the moderate thermophile *Bacillus stearothermophilus* ribosome

Junko Kimura and Makoto Kimura

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, D-1000 Berlin 33 (Dahlem), Germany

Received 14 October 1986

The complete amino acid sequences of the 5 S rRNA binding proteins L5 and L18 isolated from ribosomes of the moderate thermophile *Bacillus stearothermophilus* are presented. This has been achieved by the sequence analysis of peptides derived by enzymatic digestions with trypsin, chymotrypsin, pepsin, and *Staphylococcus aureus* protease, as well as by chemical cleavage with cyanogen bromide. The proteins L5 and L18 consist of 179 and 120 amino acid residues, and have M_r values of 20 163 and 13 473, respectively. A comparison of the sequences with their counterparts from the *Escherichia coli* ribosome reveals 59% identical residues for L5, and 53% for L18. For both proteins, the distribution of conserved regions is not random along the protein chains: some regions are highly conserved while others are not. The regions which are conserved during evolution may be important for the interaction with the 5 S rRNA molecule.

5 S rRNA binding protein; Amino acid sequence; Primary structure; (*Bacillus stearothermophilus*)

1. INTRODUCTION

The ribosomal proteins from thermophilic organisms, such as *Bacillus stearothermophilus* have proven to be more readily crystallized than those from *E. coli*, and the tertiary structures of several of them are being determined in our institute [1–3] and elsewhere [4].

In previous papers, we have presented the primary structures of 15 ribosomal proteins from *B. stearothermophilus* [5–9]. This knowledge has been of considerable help in the correlation with the corresponding proteins of *E. coli*, and in the refinement of their tertiary structures. Furthermore, a comparison of the homologous proteins from the two organisms can yield useful indirect information about functionally and/or structurally important regions within the proteins. This conclusion is derived from the fact that these important regions should be more conserved during evolution than other regions. Subsequently, we have extended this comparative study to the 5 S rRNA binding proteins L5 and L18 from *B. stearothermophilus*.

The 5 S rRNA-protein complex can be released from the large ribosomal subunit and this complex has been shown to contain one to three proteins, depending on the source of the ribosomes. The structure of the 5 S rRNA protein-complex from the *E. coli* ribosome has been extensively studied. For instance, the complete amino acid sequences of the three proteins (EL5, EL18, and EL25), which specifically bind to the 5 S rRNA, have been determined [10–12]. Furthermore, a fragment of protein EL18 that retains the capability to bind to the 5 S rRNA was prepared by limited tryptic digestion and sequenced [13]. In addition to the 5 S rRNA binding proteins from *E. coli*, several

Correspondence address: J. Kimura, Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, D-1000 Berlin 33 (Dahlem), Germany

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

5 S rRNA binding proteins from *B. stearothermophilus* [14], *Halobacterium cutirubrum* [15], *Saccharomyces cerevisiae* [16], and rat liver [17] have been isolated, characterized and compared with those of the *E. coli* ribosome. Two proteins (BL5 and BL18) from the *B. stearothermophilus* ribosome that can bind to 5 S rRNA have been identified and correlated with EL5 and EL18 from *E. coli*. By reconstitution experiments with 5 S rRNA from *B. stearothermophilus* and proteins from *E. coli* or vice versa, it has been shown that the *B. stearothermophilus* 5 S rRNA binding proteins BL5 and BL18 are interchangeable with EL5 and EL18 [14].

In order to gain a further insight into the structure-function relationship of the 5 S rRNA-protein complex, and to collect more sequence data for a comparative study on ribosomal proteins, we have analyzed the primary structures of L5 and L18 from the *B. stearothermophilus* ribosome and compared their sequences with those of the homologous proteins from *E. coli*.

2. MATERIALS AND METHODS

2.1. *Protein isolation*

Ribosomal proteins L5 and L18 of *B. stearothermophilus* (strain NCA 1503) were extracted from the 50 S subunit with 66% acetic acid in the presence of 0.1 M $MgCl_2$. The proteins were then separated by CM-cellulose column chromatography as described in [8].

2.2. *Sequence determination*

The proteins were cleaved enzymatically with trypsin, chymotrypsin, pepsin, and *Staphylococcus aureus* protease as well as chemically with cyanogen bromide, as described in [18]. The resulting peptides were separated by gel filtration on Sephadex G-75 (140 × 1 cm) followed by thin-layer chromatography, or directly by reverse-phase HPLC (Vydac C18 column) with an acetonitrile gradient in aqueous trifluoroacetic acid. Amino acid analyses were performed by an HPLC system using *o*-phthaldialdehyde as a derivatized reagent [19]. Sequence determination was performed by the DABITC/PITC double-coupling method [20] or the manual solid-phase technique [21], as described in [22].

2.3. *Nomenclature*

The ribosomal proteins L5 and L18 from *B. stearothermophilus* were designated according to their amino acid sequence homology to the corresponding *E. coli* proteins L5 and L18, respectively. For instance, this new nomenclature replaces BL22 [14] by L18 described in this paper. When necessary, the prefixes B or E are added to differentiate *B. stearothermophilus* and *E. coli*, respectively.

3. RESULTS AND DISCUSSION

The sequences of proteins L5 and L18 from the *B. stearothermophilus* ribosome were determined by a direct protein sequence analysis of the intact proteins, and of peptides derived from them. The strategy by which the complete sequences for L5 and L18 were derived from these data is summarized in figs 1 and 2.

3.1. *The amino acid sequence of protein L5*

The sequence was determined as follows. Protein L5 was first digested with trypsin and the peptides were isolated by HPLC using the Vydac C18 column. Amino acid sequences of the peptides were determined by the DABITC/PITC double-coupling method as shown in fig.1. The alignment of tryptic peptides was obtained from the amino acid sequences of cyanogen bromide peptides, as well as from peptic peptides derived from the intact protein. Cyanogen bromide cleavage of L5 gave eight peptides, CB1~CB8, which were separated by gel filtration on Sephadex G-75, followed by thin-layer chromatography. The peptides thus obtained were sequenced by the DABITC/PITC double-coupling method or solid-phase method, as shown in fig.1. Furthermore, L5 was digested with pepsin and the resulting peptides were isolated and sequenced. From these results, the amino acid sequence of L5 was unambiguously determined. As given in table 1 protein L5 contains 179 amino acid residues, and an M_r of 20163 was calculated from the composition.

3.2. *The amino acid sequence of L18*

The sequence of protein L18 was established by procedures similar to those described above for protein L5. The sequence was determined by the isolation and characterization of the 19 tryptic

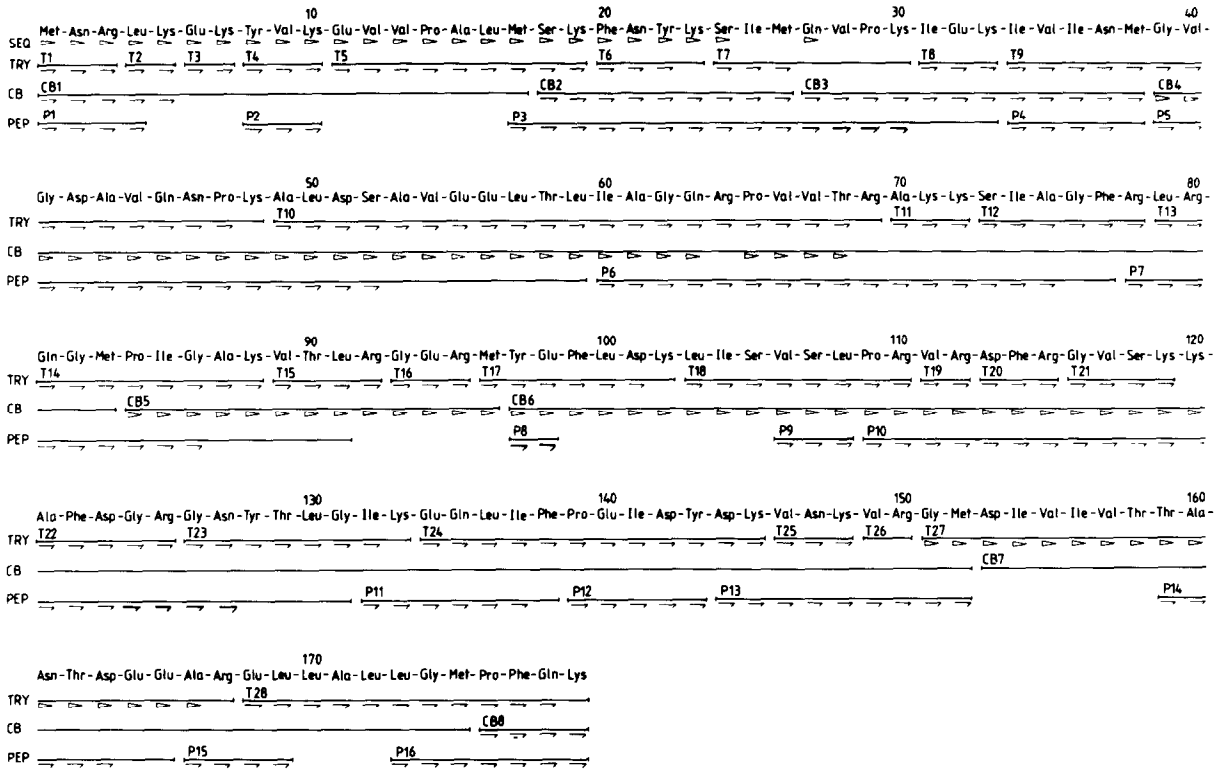

Fig.1. Amino acid sequence of protein L5 from *B. stearothermophilus*. Sequence data of peptides are indicated as follows: →, sequenced by the DABITC/PITC double-coupling method; ▷, sequenced by the solid-phase procedure. SEQ indicates a direct degradation of intact protein. TRY, CB, and PEP indicate peptides derived from cleavage with trypsin, cyanogen bromide, and pepsin, respectively.

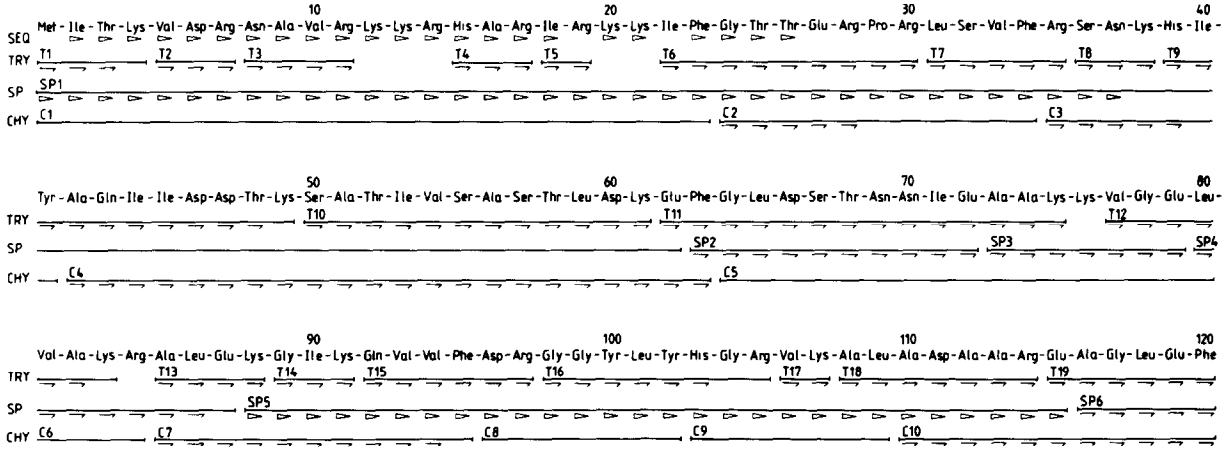

Fig.2. Amino acid sequence of protein L18 from *B. stearothermophilus*. Sequence data of peptides are indicated with symbols and nomenclature as described in fig.1. SP and CHY indicate peptides derived from digestion with *Staphylococcus aureus* protease and chymotrypsin, respectively.

Table 1

Amino acid compositions and M_r of protein L5 and L18 from *B. stearothermophilus*

Amino acid	Protein	
	L5	L18
Asp	9	7
Asn	7	4
Thr	7	7
Ser	7	6
Glu	12	7
Gln	6	2
Pro	8	1
Gly	13	8
Ala	12	14
Val	18	9
Met	8	1
Ile	13	9
Leu	16	8
Tyr	5	3
Phe	7	5
His	0	3
Lys	18	14
Arg	13	12
Total	179	120
M_r	20163	13473

peptides which were aligned by overlapping peptides isolated after *Staphylococcus aureus* protease and chymotryptic digestion. The combination of the sequence information obtained from these peptides allowed us to establish the complete amino acid sequence of L18, as shown in fig.2. Protein L18 contains 120 amino acid residues, and an M_r of 13473 was calculated as given in table 1.

3.3. *Comparison of amino acid sequences*

The amino acid sequences of L5 and L18 from *B. stearothermophilus* are compared with those of L5 and L18 from *E. coli*, respectively (fig.3). BL5 has one extra residue (a methionine at the N-terminus), when the residues are aligned to optimize the homology. There are identical residues at 105 of the 179 positions, i.e. 59% of the amino acid residues are identical in the two L5 proteins. The amino acid sequence homology extends throughout most of two protein chains but is stronger within the C-terminal half than within the N-terminal half. This homology suggests that the C-terminal half, particularly positions 80–100 and 109–154, might have an important role in the interaction of L5 with the 5 S rRNA molecule.

The BL18 has four extra residues at the N-terminus and requires the introduction of an insertion of two residues after Gly-64 and a deletion of Phe-23, when compared with its *E. coli* homologue. There are identical residues at 61 of the 120 positions, i.e. the identity is 53%. There exist two long runs of highly conserved residues, namely Val-81–Lys-91 and Tyr-102–Phe-120. Both of these regions are present within the C-terminal halves of the two molecules. In contrast, a comparison of the N-terminal halves shows a high level of variation in the amino acid sequences.

It was reported previously that a fragment extending from pos. 16/17 to 117 of EL18 retains the capability to bind to the 5 S rRNA, and it was therefore concluded that the N-terminal region, pos. 1–16, is not essential for the L18-5 S rRNA interaction [13]. Our result on the comparison of the amino acid sequences of EL18 and BL18 is quite consistent with this conclusion, showing that the C-terminal two-third is well conserved in contrast to the N-terminal region.

A relatively strong conservation in the positions of basic amino acid residues along the entire length of the protein chains can be observed. This may be a reflection of the functional or structural significance of these residues. Newberry et al. [13] have proposed that the N-terminal region of L18 might be involved in either the interaction with the other 5 S rRNA binding proteins L5 and L25, or the attachment of the 5 S rRNA-protein complex to the 23 S rRNA.

The determination of the primary structures of proteins BL5 and BL18 described in this paper, together with those of proteins sequenced previously, enable us now to correlate 15 proteins from the *B. stearothermophilus* 50 S ribosomal subunit with the homologous proteins from the *E. coli* ribosome. Table 2 summarizes a comparison of sequenced *Bacillus* ribosomal proteins with their homologues from the *E. coli* ribosome. The degree of homology for the corresponding proteins from the two bacteria varies considerably depending upon the protein pairs. Although it is presently difficult to provide a convincing explanation for this observation, the data from two additional studies are suggestive to this problem. Firstly, Auron and

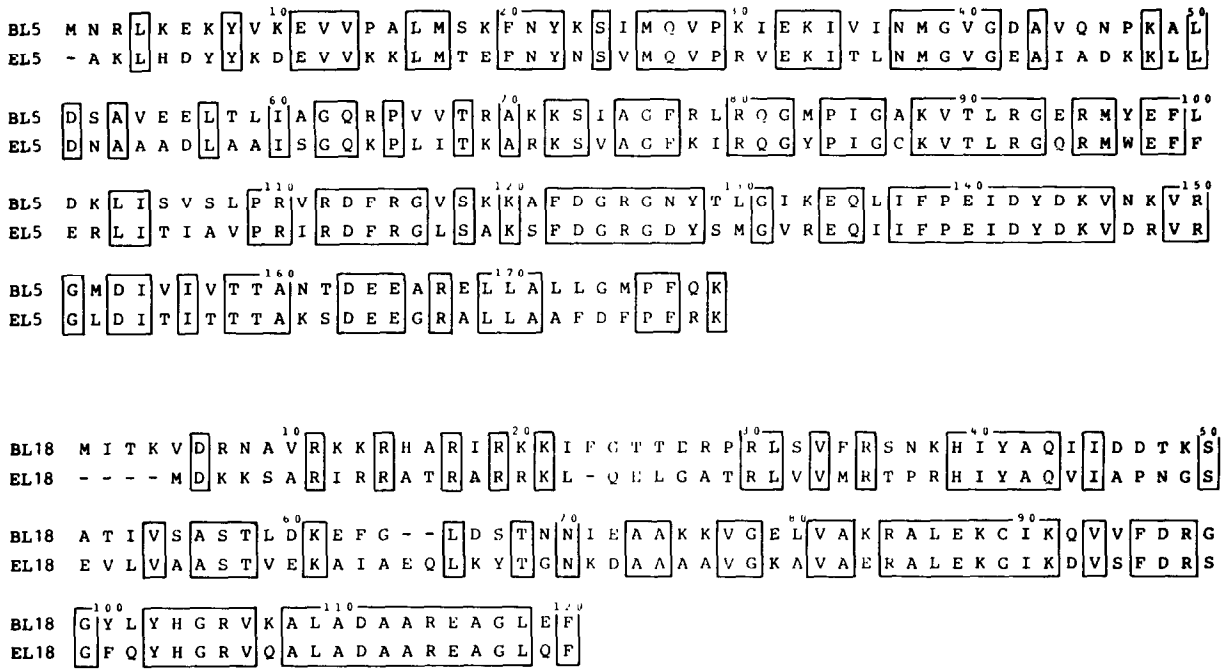

Fig.3. Comparison of the amino acid sequences of L5 and L18 from *B. stearothermophilus* with the homologous sequences from *E. coli*. A maximum homology was obtained by a programme ALIGN [25] with a break penalty of 25 residues. Identical residues are enclosed in boxes.

Table 2

Comparison of 50 S ribosomal proteins from *B. stearothermophilus* and *E. coli*

Protein	Number of residues		Identity (%)	References
	B. stearo-thermo-philus	*E. coli*		
L1	232	233	51	8
L2	274	272	60	9
L5	179	176	59	this paper
L6	177	176	49	6
L9	147	148	33	6
L14	122	123	69	8
L15	146	144	44	8
L17	117	127	49	6
L18	120	117	53	this paper
L23	95	100	28	8
L24	103	103	44	8
L27	87	84	54	6
L29	66	63	46	8
L30	61	58	53	5
L32	56	56	25	7

Fahnestock [24] reconstituted the *B. stearothermophilus* 50 S ribosomal subunit from a mixture of purified proteins and RNA, and analyzed the functional significance for individual proteins by single protein omission tests. Secondly, Dijk and co-workers (unpublished) extracted proteins from *B. stearothermophilus* 50 S ribosomal subunits by consecutive salt treatment and then isolated several proteins in pure form. Combining the results from these two studies, an interesting correlation emerges: proteins which are found to be strongly involved in polyPhe and peptidyltransferase activities, and which are not easily extracted by salt treatment are, highly conserved. On the other hand, proteins which do not affect both functional activities and are readily extracted by salt are rather poorly conserved.

ACKNOWLEDGEMENTS

We would like to thank Dr H.G. Wittmann for his interest and encouragement, Dr J. McDougall for critical reading of the manuscript, Mrs G.

Breithaupt for excellent technical assistance and Mr D. Kamp for performing the amino acid analyses.

REFERENCES

[1] Appelt, K., Dijk, J. and Epp, O. (1979) FEBS Lett. 103, 66–70.
[2] White, S.W., Appelt, K., Dijk, J. and Wilson, K.S. (1983) FEBS Lett. 163, 73–75.
[3] Appelt, K., Tanaka, I., White, S.W. and Wilson, K.S. (1984) FEBS Lett. 165, 43–45.
[4] Liljas, A. and Newcomer, M.E. (1981) J. Mol. Biol. 153, 393–398.
[5] Kimura, M. (1983) J. Biol. Chem. 259, 1051–1055.
[6] Kimura, M. and Chow, C.K. (1984) Eur. J. Biochem. 139, 225–234.
[7] Tanaka, I., Kimura, M., Kimura, J. and Dijk, J. (1984) FEBS Lett. 166, 343–346.
[8] Kimura, M., Kimura, J. and Ashman, K. (1985) Eur. J. Biochem. 150, 491–497.
[9] Kimura, M., Kimura, J. and Watanabe, K. (1985) Eur. J. Biochem. 153, 289–297.
[10] Chen, R. and Ehrke, G. (1976) FEBS Lett. 69, 240–245.
[11] Brosius, J., Schiltz, E. and Chen, R. (1975) FEBS Lett. 56, 359–361.
[12] Bitar, K.G. and Wittmann-Liebold, B. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1343–1352.
[13] Newberry, V., Brosius, J. and Garrett, R. (1976) Nucleic Acids Res. 5, 1753–1766.
[14] Horne, J.R. and Erdmann, V.A. (1972) Mol. Gen. Genet. 119, 337–344.
[15] Smith, N., Matheson, A.T., Yaguchi, M., Willick, G.E. and Nazar, R.N. (1978) Eur. J. Biochem. 89, 501–509.
[16] Nazar, R.N., Yaguchi, M., Willick, G.E., Rollin, C.F. and Roy, C. (1979) Eur. J. Biochem. 102, 573–582.
[17] Terao, K., Takahashi, Y. and Ogata, K. (1975) Biochim. Biophys. Acta 402, 230–237.
[18] Wittmann-Liebold, B. and Lehmann, A. (1980) in: Methods in Peptide and Protein Sequence Analysis (Birr, C. ed.) pp.49–72, Elsevier, Amsterdam, New York.
[19] Ashman, K. and Bosserhof, A. (1985) in: Modern Methods in Analyical Protein Chemistry (Tschesche, H. ed.) pp.155–171, De Gruyter, Berlin.
[20] Chang, J.Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[21] Chang, J.Y. (1979) Biochim. Biophys. Acta 578, 188–195.
[22] Wittmann-Liebold, B. and Kimura, M. (1984) in: Methods in Molecular Biology, vol.1, Proteins (Walker, J.M. ed.) pp.221–242, Humana Press, New York.
[23] Nazar, R.N., Yaguchi, M. and Willick, G.E. (1982) Can. J. Biochem. 60, 490–496.
[24] Auron, P.F. and Fahnestock, S.R. (1981) J. Biol. Chem. 256, 10105–10110.
[25] Dayhoff, M.O. (1978) in: Atlas of Protein Sequence and Structure, vol.5, suppl.3, National Biomedical Research Foundation, Washington, DC.