Volume 210, number 1, 91–96 FEB 04317 January 1987

# The complete amino acid sequence of ribosomal protein S12 from *Bacillus stearothermophilus*

Makoto Kimura and Junko Kimura

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

Received 14 October 1986

The amino acid sequence of ribosomal protein S12 from *Bacillus stearothermophilus* has been completely determined. The sequence data were mainly obtained by manual sequencing of peptides derived from digestion with trypsin, *Staphylococcus aureas* protease and pepsin. A few overlaps of tryptic peptides were established by DNA sequence analysis of a chromosomal fragment containing the *rpsL* gene coding for ribosomal protein S12. The protein contains 138 amino acid residues and has an $M_r$ of 15208. Comparison of this sequence with the sequences of the ribosomal S12 proteins from *E. coli* as well as from *Euglena*, tobacco and liverwort chloroplasts shows that 75% of the amino acid residues are identical within the S12 proteins of all four species. Therefore, S12 is the most strongly conserved ribosomal protein known so far.

Ribosomal evolution; Sequence homology; Protein S12; Amino acid sequence; (*Bacillus stearothermophilus*)

## 1. INTRODUCTION

The primary structures of 15 proteins (L1, L2, L5, L6, L9, L14, L15, L17, L18, L23, L24, L27, L29, L30, and L32) from the 50 S ribosomal subunit of *B. stearothermophilus* have been determined to date, and have been compared with their counterparts from the *E. coli* ribosomes ([1–5], preceding paper). These results allow us to correlate the *B. stearothermophilus* ribosomal proteins to those of *E. coli* and to reveal which regions within the proteins are highly conserved during evolution and which are more variable.

We have extended this comparative study to include the 30 S ribosomal proteins from *B. stearothermophilus*, and at present nearly complete amino acid sequences of nine proteins have been established. In the course of these studies, we found that the amino acid sequence of protein S12 was the most strongly conserved amongst all ribosomal proteins whose primary structures have so far been determined.

It was revealed by reconstitution experiments using the *E. coli* system that S12 is a functionally essential protein which is involved in the initiation step of polypeptide synthesis. 30 S subunits deficient in S12 are only 20% as active as control subunits, although the sedimentation coefficient of the reconstituted particle deficient in S12 is indistinguishable from that of control 30 S subunits [6]. Furthermore, reconstitution experiments using streptomycin sensitive and resistant ribosomes demonstrated that mutated protein S12 confers streptomycin resistance [6]. Subsequent protein-chemical analysis using mutants with altered streptomycin phenotypes showed that Lys-42 and Lys-87 are exchanged by other amino acid in the protein S12 of these mutants [7].

In this paper, we describe the complete amino acid sequence of protein S12 from the *B. stearothermophilus* ribosome and compare it with the sequences of protein S12 from *E. coli* [6], as

Correspondence address: M. 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 

well as from the chloroplast ribosomes in *Euglena* [9], tobacco [10,11], and liverwort [12].

## 2. MATERIALS AND METHODS

### 2.1. *Materials*

The materials and methods used for protein sequencing have been described in the preceding paper, and additional information may be found in [13].

Phage M13 mp19 replicative forms and plasmid pNO1523 [14] containing the *E. coli rpsL* gene and a part of the *rpsG* gene were purchased from Pharmacia. Restriction endonuclease *Sau*3A, $T_4$ ligase and *E. coli* DNA polymerase I large fragment were obtained from Bethesda Research Laboratory. [$\alpha$-$^{32}$P]dCTP (400 Ci/mmol) was from Amersham.

### 2.2. *Protein isolation*

Ribosomal protein S12 from *B. stearothermophilus* (strain NCA 1503) extracted from the 30 S subunit with 66% acetic acid was purified by CM-cellulose column chromatography as described in [15].

### 2.3. *Sequencing of a chromosomal DNA fragment corresponding to a region of S12 from* B. stearothermophilus

Chromosomal DNA was extracted from *B. stearothermophilus* (strain NCA 1503) cells grown in LT medium at 65°C according to Chow et al. [16], and digested completely with *Sau*3A. Restriction fragments were electrophoresed on a 0.8% horizontal agarose gel, transferred to nitrocellulose membranes, and probed with a nick-translated *Hpa*I-*Bam*HI restriction fragment derived from plasmid pNO1523. Hybridization was carried out at 37°C in 5 × SSC, and the hybridized *Sau*3A restriction fragments were purified by electrophoresis in 0.8% low melting agarose gel. The isolated fragments were ligated into the *Bam*HI site of M13 mp19. The ligated products were used to transfect *E. coli* JM105. DNA from the resultant recombinant clones was sequenced by the dideoxynucleotide chain termination method [17].

### 2.4. *Nomenclature*

The ribosomal protein S12 of *B. stearothermophilus* was designated according to its amino acid sequence homology to the corresponding *E.*

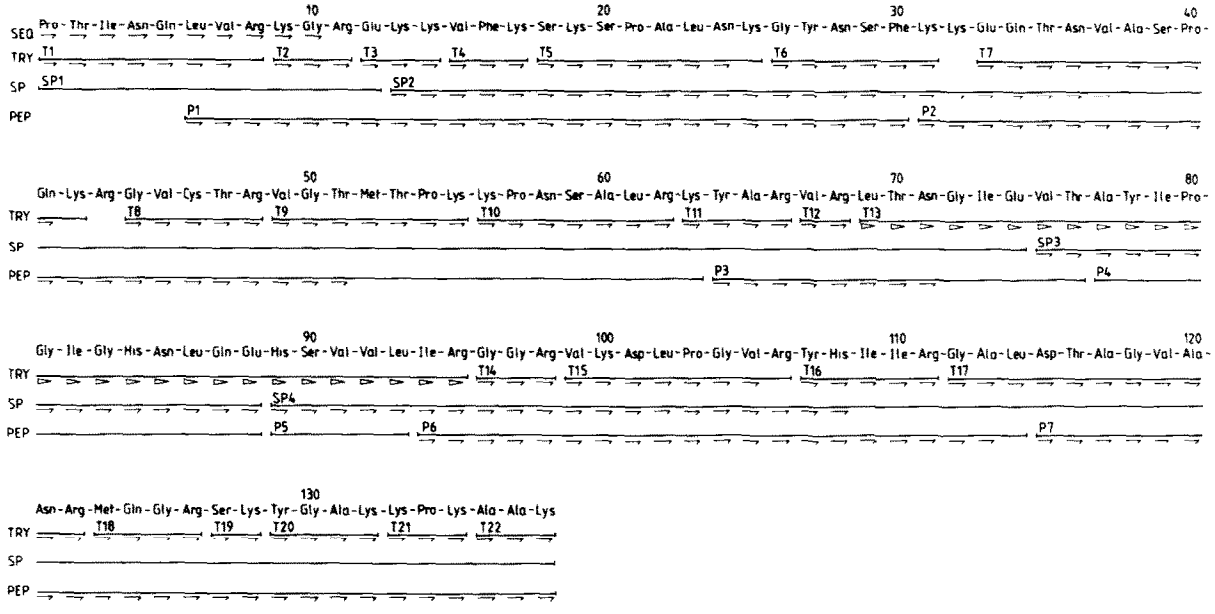


Fig.1. Amino acid sequence of protein S12 from *B. stearothermophilus*. Sequence data on individual peptides are indicated as follows: →, sequenced by the DABITC/PITC double-coupling method; ▷, sequenced by the solid-phase procedure. SEQ indicates a direct degradation of intact protein. TRY, SP and PEP indicate peptides derived from cleavage with trypsin, *Staphylococcus aureus* protease and pepsin, respectively.

*coli* protein S12. When necessary, the prefix B (*B. stearothermophilus*), or E (*E. coli*) is added to differentiate between the homologous proteins.

## 3. RESULTS AND DISCUSSION

### 3.1. *Amino acid sequence of S12*

The complete amino acid sequence of protein S12 from the *B. stearothermophilus* ribosome is presented in fig.1. The sequence of S12 was mainly derived by analysis of peptides obtained by cleavage with trypsin, *Staphylococcus aureus* protease, and pepsin. Tryptic digestion of S12 gave 22 peptides which were isolated by an HPLC system using a Vydac C18 column. All peptides were completely sequenced by the DABITC/PITC double-coupling method or solid-phase method. The alignment of the tryptic peptides was accomplished by sequencing the peptides isolated after digestion with *Staphylococcus* protease or pepsin, with the exception of overlaps between peptides T9 and T11. The order of these peptides was determined by DNA sequence analysis of the *B. stearothermophilus* S12 gene as presented in fig.2.

The *Sau*3A restriction digest of the chromosomal DNA from *B. stearothermophilus* was hybridized with the *Hpa*I-*Bam*HI restriction fragment derived from plasmid pNO1523, which contains the intact S12 gene as well as the N-terminal part of the S7 gene from *E. coli*. Specific hybridization was obtained with a fragment of ~0.3 kb. Accordingly, the DNA fragments with a size of 0.3 kb were directly cloned into the M13 mp19 vector and again screened with the *Hpa*I-*Bam*HI fragment. The recombinant DNA obtained was directly sequenced by the dideoxynucleotide chain termination method. It became apparent that the 3'-end of the *Sau*3A fragment contained a part of the gene for protein S12. This sequence yielded the alignment of T9 and T11 and also confirmed the amino acid sequence of residues 1–73 of protein S12, as given in fig.2. Protein S12 contains 138 amino acid residues and has a calculated $M_r$ of 15208, as shown in table 1.

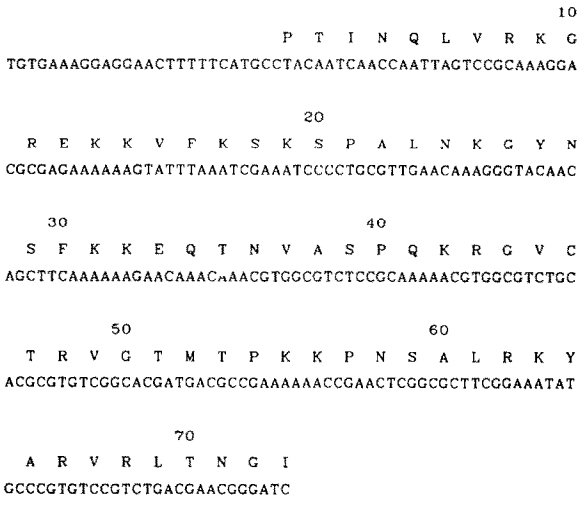


Fig.2. DNA sequence of a chromosomal DNA clone from *B. stearothermophilus* and the deduced protein sequence for part of the protein S12. The protein sequence corresponds to residues 1–73 of S12.

### 3.2. *Comparison of amino acid sequences*

The amino acid sequence of protein S12 from *B. stearothermophilus* is aligned in fig.3 with that of the *E. coli* S12 [8] as well as with the sequences of

Table 1

Amino acid composition and $M_r$ of protein S12 from *B. stearothermophilus*

| Amino acid | Protein S12 |
|---|---|
| Asp | 2 |
| Asn | 8 |
| Thr | 8 |
| Ser | 7 |
| Glu | 4 |
| Gln | 5 |
| Pro | 8 |
| Gly | 14 |
| Ala | 11 |
| Val | 12 |
| Met | 2 |
| Ile | 7 |
| Leu | 8 |
| Tyr | 5 |
| Phe | 2 |
| His | 3 |
| Lys | 18 |
| Arg | 13 |
| Cys | 1 |
| Total | 138 |
| $M_r$ | 15208 |

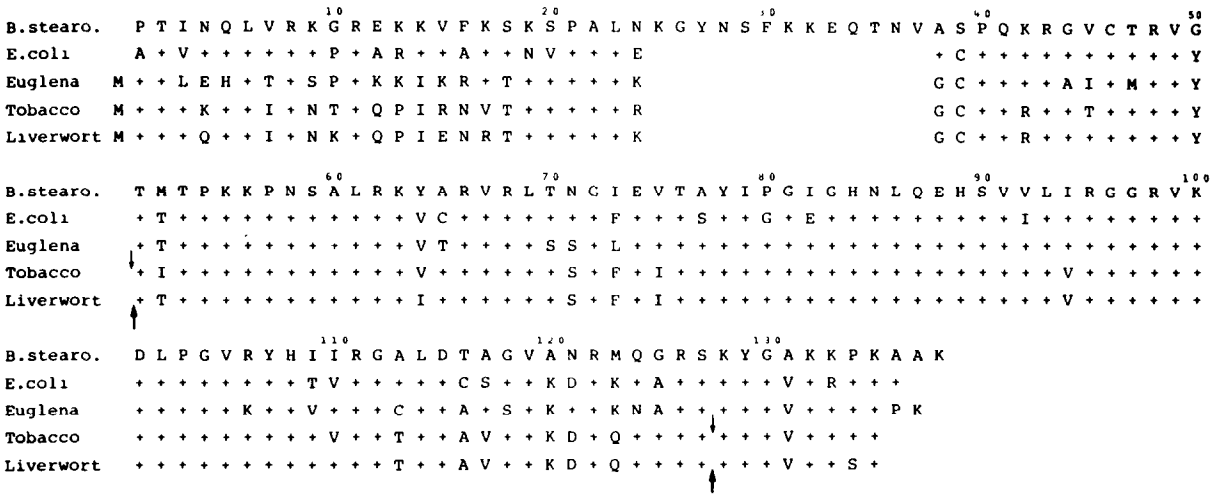


Fig.3. Comparison of the amino acid sequence of the protein S12 from the *B. stearothermophilus* ribosome with the sequences of the corresponding proteins from *E. coli* [8], and from chloroplast of *Euglena gracilis* [9], *Nicotiana tabacum* [10,11] and *Marchantia polymorpha* [12]. +, indicates residue identical with *Bacillus* S12. Numbering is based on the residues of the protein S12 from *B. stearothermophilus*. Vertical arrows indicate sites of splicing junctions in the ribosomal proteins S12 from tobacco and liverwort.

the homologous proteins from *Euglena* [9], tobacco [10,11] and liverwort [12] chloroplasts. BS12 can be easily aligned with ES12 and with the chloroplast proteins, showing identities between 71% and 76%. The degree of homology between the various proteins is given in table 2. BS12 has the highest homology to its homologue from *E. coli* (76% identical residues) and only a slightly lower homology (71%) to the *Euglena* protein. BS12 appears to be somewhat more closely related to the chloroplast proteins than ES12.

Table 2

Degree of homology of individual proteins, as percent identity, calculated for each pair of protein S12

| | *B. stearothermophilus* | *E. coli* | *Euglena* | Tobacco |
|---|---|---|---|---|
| *B. stearothermophilus* | | | | |
| *E. coli* | 76 | | | |
| *Euglena* | 71 | 70 | | |
| Tobacco | 75 | 71 | 74 | |
| Liverwort | 75 | 70 | 75 | 92 |

The sequence homology found for protein S12 is particularly high as compared to the other homologous pairs of *B. stearothermophilus* and *E. coli* ribosomal proteins sequenced so far (table 3). This result indicates a very slow rate of evolutionary change in protein S12 and it is also found by a comparison of the chloroplast S12 proteins with the *E. coli* S12: chloroplast S12 shows about 70% homology to the *E. coli* S12, whereas other chloroplast ribosomal proteins have only ~40% homology with their counterparts of *E. coli* [12]. This high conservation of the amino acid sequence of protein S12 reflects the essential role of protein S12 in bacterial and chloroplast ribosomes.

Several interesting features of the *B. stearothermophilus* sequence are apparent by comparing it with the sequences of ES12 and the chloroplast S12 proteins. Firstly, the cysteine content of BS12 is lower and only one cysteine (Cys-45) is conserved whereas the methionine content of BS12 is higher and this protein contains two methionine residues (Met-52 and Met-123) instead of threonine and lysine, respectively, in protein ES12. In contrast, the three histidine residues in BS12 (His-84, His-89 and His-108) are conserved in all four proteins. As described above, the lysine residues in positions 42

Table 3

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

| Protein | Number of residues | | Identity (%) | References |
|---|---|---|---|---|
| | *B. stearothermophilus* | *E. coli* | | |
| S2[a] | 234 | 240 | 53 | unpublished |
| S4[a] | 197 | 203 | 52 | unpublished |
| S5 | 166 | 166 | 55 | 1 |
| S8[a] | 129 | 129 | 47 | unpublished |
| S9 | 129 | 129 | 54 | 2 |
| S12 | 138 | 123 | 76 | this paper |
| S15 | 88 | 88 | 57 | unpublished |
| S16[a] | 88 | 88 | 45 | unpublished |
| S20[a] | 83 | 86 | 51 | unpublished |
| L1 | 232 | 233 | 51 | 4 |
| L2 | 274 | 272 | 60 | 5 |
| L5 | 179 | 176 | 59 | preceding paper |
| L6 | 177 | 176 | 49 | 2 |
| L9 | 147 | 148 | 33 | 2 |
| L14 | 122 | 123 | 69 | 4 |
| L15 | 146 | 144 | 44 | 4 |
| L17 | 117 | 127 | 49 | 2 |
| L18 | 120 | 117 | 53 | preceding paper |
| L23 | 95 | 100 | 28 | 4 |
| L24 | 103 | 103 | 44 | 4 |
| L27 | 87 | 84 | 54 | 2 |
| L29 | 66 | 63 | 46 | 4 |
| L30 | 61 | 58 | 53 | 1 |
| L32 | 56 | 56 | 25 | 3 |

[a] Sequences have been provisionally established

and 87 of ES12 are exchanged by other amino acids in streptomycin-resistant mutants of *E. coli* [7]. It is interesting that both lysine residues are conserved in BS12 as well as in the chloroplast ribosomal S12 proteins.

Secondly, with regard to the distribution of conserved and variable residues in the S12 proteins, there are long runs of highly conserved sequences, namely in the regions Thr-53–Tyr-64 and Leu-93–His-108. Both regions are located in the middle of the protein chain. On the other hand, the N-terminal 50 residues are less conserved, and this region contains an insertion of 13 amino acid residues. The amino acid sequences of the S12 proteins from tobacco and liverwort chloroplasts have recently been deduced from DNA sequence analyses of the corresponding genes [10–12]. The DNA sequence data revealed that the genes for these two S12 proteins are divided into exons and introns. It is of interest that one of these exons (exon 2) corresponds well to the conserved region in S12 proteins.

Finally, as mentioned above, in BS12 there is a long run of 13 amino acid residues inserted within the N-terminal region (fig.3). In spite of this rather long insertion, the *E. coli* S12 can be completely replaced by the *B. stearothermophilus* S12 in poly(U)-dependent polyPhe synthesis [18]. This might suggest that the inserted residues of BS12 are not essential for the ribosomal function. However, when a natural mRNA, e.g. f2 phage RNA, and not an artificial mRNA like poly(U) is used, there is a distinct difference between the ribosomes of *E. coli* and *B. stearothermophilus*: the f2 phage RNA can be translated by the former but not the latter ribosomes [19]. It has been found that the 30 S components responsible for this difference are mainly S12 and the 16 S rRNA (see discussion in [20]). It is possible that this different ability of ribosomes of the two species is correlated with the long insertion in the N-terminal region of BS12 protein, especially since the other parts of the S12 protein are very strongly conserved.

## ACKNOWLEDGEMENTS

We would like to thank Dr H.G. Wittmann for his interest and valuable discussion, 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] Kimura, M. (1983) J. Biol. Chem. 259, 1051–1055.
[2] Kimura, M. and Chow, C.K. (1984) Eur. J. Biochem. 139, 225–234.
[3] Tanaka, I., Kimura, M., Kimura, J. and Dijk, J. (1984) FEBS Lett. 166, 343–346.
[4] Kimura, M., Kimura, J. and Ashman, K. (1985) Eur. J. Biochem. 150, 491–497.
[5] Kimura, M., Kimura, J. and Watanabe, K. (1985) Eur. J. Biochem. 153, 289–297.

[6] Ozaki, M., Mizushima, S. and Nomura, M. (1969) Nature 222, 333–339.
[7] Funatsu, G., Nierhaus, K.H. and Wittmann, H.G. (1972) Biochim. Biophys. Acta 287, 282–291.
[8] Funatsu, G., Yaguchi, M. and Wittmann-Liebold, B. (1977) FEBS Lett. 73, 12–17.
[9] Montandon, P.E. and Stutz, E. (1984) Nucleic Acids Res. 12, 2851–2859.
[10] Fromm, H., Edelman, M., Koller, B., Goloubinoff, P. and Galun, E. (1986) Nucleic Acids REs. 14, 883–898.
[11] Torazawa, K., Hayashida, N., Obokata, J., Shinozaki, K. and Sugiura, M. (1986) Nucleic Acids Res. 14, 3143.
[12] Fukuzawa, H., Kohchi, T., Shirai, H., Ohyama, K., Umesono, K., Inokuchi, H. and Ozeki, H. (1986) FEBS Lett. 198, 11–15.
[13] 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.
[14] Dean, D. (1981) Gene 15, 99–102.
[15] Isono, S. and Isono, K. (1975) Eur. J. Biochem. 50, 483–488.
[16] Chow, L.T., Kahmann, R. and Kamp, D. (1977) J. Mol. Biol. 113, 591–609.
[17] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[18] Higo, K., Held, W., Kahan, L. and Nomura, M. (1973) Proc. Natl. Acad. Sci. USA 70, 944–948.
[19] Lodish, H.F. (1970) Nature 226, 705–707.
[20] Higo, K. and Loertscher, K. (1974) J. Bacteriol. 118, 180–186.