The amino acid sequence of two small ribosomal proteins from *Bacillus stearothermophilus*

Isao Tanaka, Makoto Kimura, Junko Kimura and Jan Dijk

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, Berlin-Dahlem, Germany

Received 30 November 1983

The low-M_r proteins (tentatively called protein I and II) were purified from 2 M NaCl extracts of the *Bacillus stearothermophilus* ribosome. Their amino acid sequences have been determined from the peptides obtained by digestion with trypsin, chymotrypsin, and pepsin, and by cleavage with CNBr, using the micro-DABITC/PITC double-coupling method [FEBS Lett. (1978) 93, 205–214]. Protein I contains 56 residues and has an M_r of 6514. Protein II had 37 residues with an M_r of 4361. The amino acid sequence of protein I shows significant similarity to L32 from *E. coli*, whereas that of protein II is slightly, if at all, related to ribosomal protein L34 from *E. coli*.

B. stearothermophilus *ribosome* *Ribosomal protein* *Protein L32* *Amino acid sequence* *Sequence homology*

1. INTRODUCTION

The elucidation of the tertiary structure of several *Bacillus stearothermophilus* ribosomal proteins by X-ray structure analysis is in progress [1–4]. To avoid denaturation of the proteins during their isolation, a method has been developed which avoids denaturing agents and handling [5]. Using this method we have isolated two proteins (tentatively called I and II) which had not previously been identified. Here we report their complete amino acid sequences and their possible relationship with two ribosomal proteins from *E. coli*.

2. MATERIALS AND METHODS

2.1. *Protein isolation*

Ribosomal proteins of *B. stearothermophilus* were isolated under mild conditions essentially as in [5] but with the following modification. Instead of the 50 S subunits we used 70 S ribosomes isolated by Sephacryl S-300 gel filtration as starting material for the salt extraction to enable large-scale purification. Proteins I and II were among proteins extracted by 2 M NaCl and were purified by CM-Sepharose CL-6B chromatography followed by G-50 gel filtration.

2.2. *Protein digestion*

Tryptic digestions of carboxymethylated proteins were carried out in 0.2 M *N*-methylmorpholine acetate buffer (pH 8.1) for 4 h at 37°C using an enzyme:substrate ratio of 1:50 (w/w). The resulting peptides (T-peptides) were separated by thin-layer fingerprinting.

Chymotrypsin digested peptides (C-peptides) were obtained by digestion in 0.2 M *N*-methylmorpholine acetate buffer (pH 8.1) for 6 h at 37°C. The enzyme:substrate ratio was 1:50 (w/w). The peptides were separated by fingerprinting.

Protein I was also digested with pepsin in 5% formic acid at 37°C for 6 h. The peptides (P-peptides) were separated by fingerprinting.

Cleavage with CNBr was carried out in 70% for-

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

mic acid for 20 h at room temperature. The peptides (CB peptides) were separated after carboxymethylation on a Sephadex G-50 sf column (1 × 140 cm) in the dark.

2.3. *Sequence analysis*

All the peptides obtained by proteolysis, and the intact proteins were sequenced by the DABITC/PITC. (4-*N*,*N*-dimethylaminoazobenzene-4′-isothiocyanate/phenylisothiocyanate) double coupling method [6,7]. The sequencing of the CNBr-treated peptides was performed by the manual solid phase method as in [6].

3. RESULTS AND DISCUSSION

3.1. *Isolation*

Proteins I and II were eluted at 0.48 and 0.44 M NaCl, respectively, from the CM Sepharose CL-6B column. The proteins, purified further by gel filtration on Sephadex G-50, migrated as a single component on SDS–PAGE. From the positions on two-dimensional gels, proteins I and II most likely correspond to BL37 and BL38 as described in [8].

3.2. *Sequence determination*

The complete amino acid sequences of proteins I and II are shown in fig. 1,2. The proposed sequences were deduced as follows. The proteins were first digested with trypsin which produced 12 peptides from protein I and 8 from protein II. Peptides isolated by fingerprinting were completely sequenced by the DABITC/PITC double coupling method, except T7 (positions 21–36) in protein I, and T2 (positions 3–13) and T5 (positions 23–31) in protein II. To arrange the tryptic peptides in order and also to determine the missing sequences, peptic (for protein I only) and chymotryptic peptides were also sequenced. The combined information gave almost the entire sequence, but several overlaps were based on a single residue only. Hence, the proteins were further cleaved with CNBr and the resulting peptides were purified by gel filtration on Sephadex G-50 sf (1 × 140 cm). As expected from the number of methionine residues, two major peptides (CB1 and CB2) were isolated from each protein. Sequencing of these peptides by the manual solid phase technique completed the amino acid sequences of proteins I and II as presented in fig. 1,2.

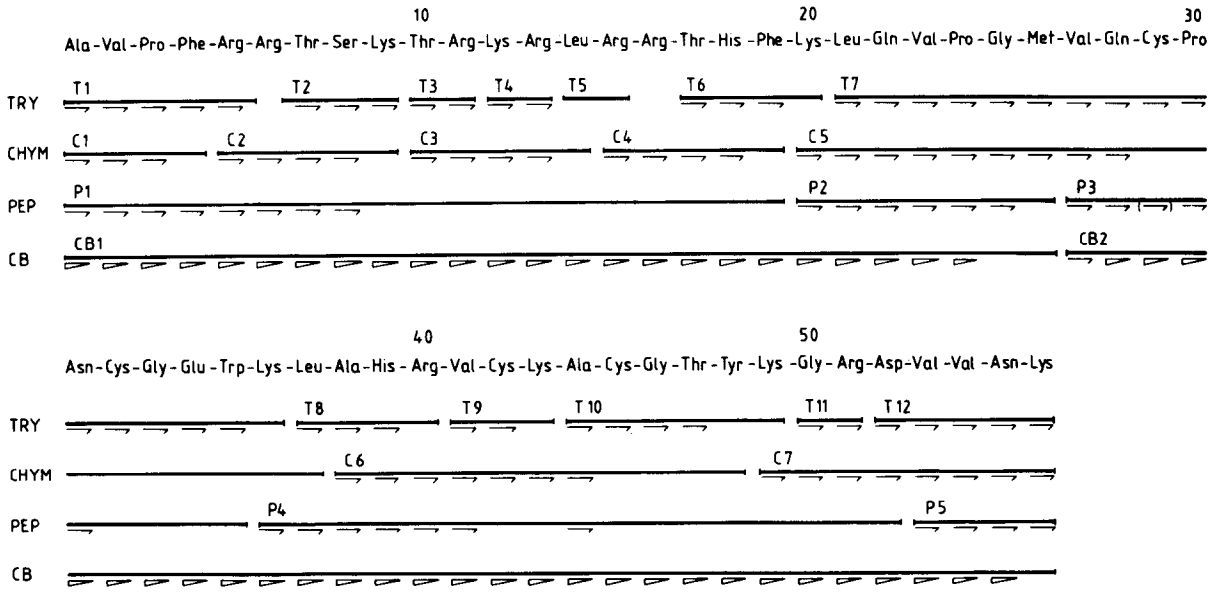

Fig.1. The amino acid sequence of protein I from *B. stearothermophilus*. Sequence data on individual peptides are indicated as follows: ⟶, sequenced by the DABITC/PITC double coupling method; ⟹, sequenced by the manual solid-phase procedure. TRY, CHYM and PEP indicate peptides derived from digestion with trypsin, chymotrypsin and pepsin, respectively. CB, cleavage with CNBR.

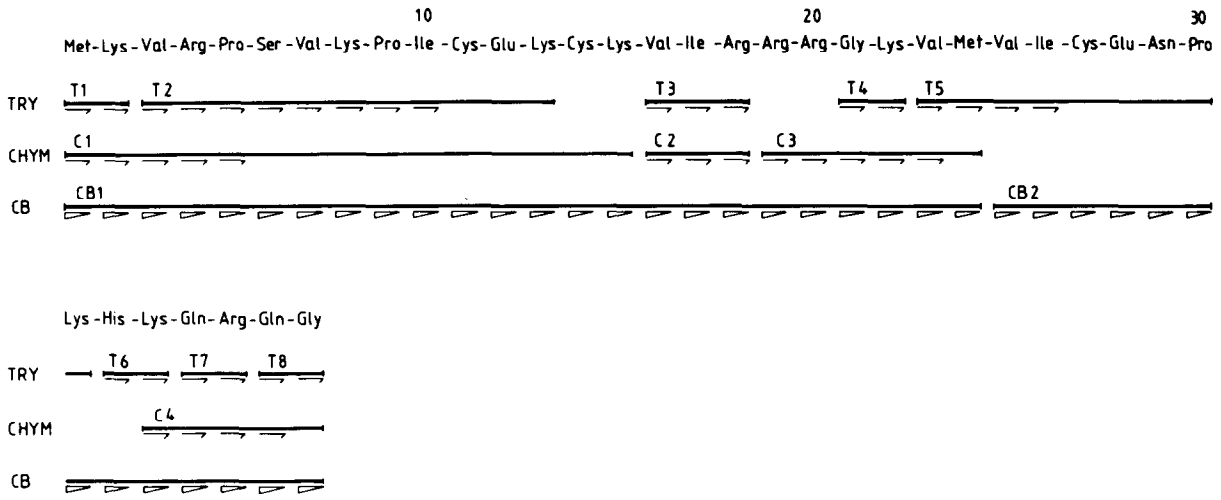

Fig.2. The amino acid sequence of protein II from *B. stearothermophilus*. Sequence data on individual peptides and abbreviations are indicated in the same way as in fig.1.

The amino acid compositions and M_r values calculated from the sequence are given in table 1. Of particular interest is the high number of cysteine residues (4 residues in I and 3 residues in II), a relatively rare amino acid in ribosomal proteins. Among the 53 ribosomal proteins from *E. coli*, all of which are now completely sequenced [9], only S12 and L31 show a similar level of cysteine content.

3.3. *Comparison of the sequences of proteins I and II with those of ribosomal proteins from* E. coli

In order to find the corresponding proteins in *E. coli*, both sequences were compared to all *E. coli* ribosomal proteins by the computer program ALIGN [10]. A mutation matrix with a break penalty of 25 was employed.

To protein I, only ribosomal protein L32 of *E. coli* gave a significant similarity at a level of 9 SD. The two sequences are aligned in fig.3. There are 14 amino acid residues identical out of a total of 56 (25% identity). This is an exceptionally low value among the pairs between *B. stearothermophilus* and *E. coli* ribosomal proteins studied so far [11]. There are no long stretches of the conserved sequence.

Nevertheless, there are 3 reasons to identify protein I as protein L32 from *B. stearothermophilus*. Firstly, the similarity to *E. coli* L32 is significantly

Table 1

Amino acid compositions and M_r values of proteins I and II

Residue	Protein	
	I	II
Asp	1	0
Asn	2	1
Thr	4	0
Ser	1	1
Glu	1	2
Gln	2	2
Pro	3	3
Gly	4	2
Ala	3	0
Val	6	5
Met	1	2
Ile	0	3
Leu	3	0
Tyr	1	0
Phe	2	0
His	2	1
Lys	7	7
Arg	8	5
Cys	4	3
Trp	1	0
Total	56	37
M_r	6514	4361

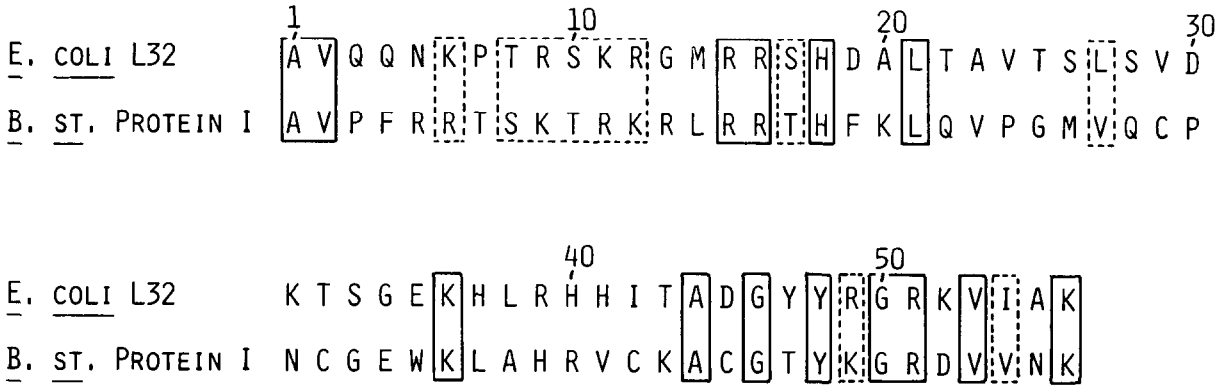

Fig.3. Comparison of the primary structure of protein I from the *B. stearothermophilus* ribosome with the homologous protein L32 from the *E. coli* ribosome. Identical residues are enclosed by solid lines and conservative replacements by dashed lines.

above the background level. Secondly, these two proteins have the same number of amino acids. Thirdly, many of the replacements are conservative (e.g. positions 8–12). We will therefore name protein I, BL32.

No clearly related protein was found for protein II among the *E. coli* ribosomal proteins, although the N-terminal region resembles that of *E. coli* L34.

Protein I: Met-Lys-Val-Arg-Pro-Ser-Val

EL34: Met-Lys-Arg-Thr-Phe-Gln-Pro-Ser-Val

As described above for protein L32, it is possible that certain ribosomal proteins, especially small ones, show very low homology to their counterparts in *E. coli*. Indeed, antibodies prepared against *E. coli* protein L34 did not react with 50 S ribosomal proteins from *B. stearothermophilus* [12], and no protein has so far been found in *B. stearothermophilus*, which corresponds to *E. coli* L34. Protein II is, therefore, a good candidate for the *B. stearothermophilus* L34 protein. However, more data, also from other organisms, are necessary to make this tentative assignment less ambiguous.

ACKNOWLEDGEMENTS

We would like to thank Professor H.G. Wittmann for his interest and encouragement, Mr Volker Brachvogel for his technical assistance in protein purification and Mr Detlev Kamp for performing the amino acid analysis. I.T. is an Alexander von Humboldt fellow.

REFERENCES

[1] Appelt, K., Dijk, J. and Epp, O. (1979) FEBS Lett. 103, 66–70.
[2] Liljas, A. and Newcomer, M.E. (1981) J. Mol. Biol. 153, 393–398.
[3] Appelt, K., Dijk, J., White, S.W., Wilson, K.S. and Bartels, K. (1983) FEBS Lett. 160, 72–74.
[4] White, S.W., Appelt, K., Dijk, J. and Wilson, K.S. (1983) FEBS Lett. 163, 73–75.
[5] Appelt, K. and Dijk, J. (1983) J. Biol. Chem., in press.
[6] 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.
[7] Chang, J.Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[8] Horne, J.R. and Erdmann, V.A. (1972) Mol. Gen. Genet. 119, 337–344.
[9] Wittmann-Liebold, B. (1983) Adv. Protein Chem., in press.
[10] Needleman, S.B. and Wunsch, C.D. (1970) J. Mol. Biol. (1970) 48, 443–453.
[11] Kimura, M. and Chow, C.K. (1983) submitted.
[12] Fahnestock, S.R., Strycharz, W.Z. and Marquis, D.M. (1981) J. Biol. Chem. 256, 10111–10116.