PROCARYOTIC RIBOSOMAL PROTEINS: N-TERMINAL SEQUENCE HOMOLOGIES AND STRUCTURAL CORRESPONDENCE OF 30 S RIBOSOMAL PROTEINS FROM *ESCHERICHIA COLI* AND *BACILLUS STEAROTHERMOPHILUS*

Makoto YAGUCHI, Alastair Taylor MATHESON and Louis Peter VISENTIN*
Division of Biological Sciences, National Research Council of Canada, Ottawa, Canada K1A 0R6

Received 8 July 1974

1. Introduction

As a catalytic unit, the ribosome is a rather novel and complex organelle. Its protein constituents consist of a heterogenous collection of entities integrated into a multi-protein–RNA complex as either fixed elements of structure and/or temporally as units of function. In attempting to evidence the evolution and the structure–function relationships of the ribosome in procaryotes we have undertaken a comparative amino acid sequence analysis of ribosomal proteins from *Escherichia coli* [1], and *Bacillus stearothermophilus* [1–3] and *Halobacterium cutirubrum* [4], organisms which differ substantially in their physiological tolerances and taxonomic relationships. Our previous structural studies have indicated a high degree of homology in some of the 30 S ribosomal proteins from *E. coli* and *B. stearothermophilus* [1] suggesting as did other functional studies [5, 6] a substantial level of evolutionary conservation. We report in this paper a summary of our results on the study of the amino terminal regions of 19 ribosomal proteins *E. coli* strain Q13 and 21 from *B. stearothermophilus* strain 10. During the course of this study, Wittmann-Liebold [7] reported N-terminal amino acid sequences of 18 proteins of the 30 S subunit from *E. coli* strain K. Geisser et al. [8] compared the 30 S ribosomal proteins from several species of bacilli with those of *E. coli* by two-dimensional electrophoresis and immunological methods. Isono et al. [3] correlated the purified 30 S proteins of *B. stearothermophilus* strain 799 with those of *E. coli* by comparing two dimensional electrophoretic mobility, immunological cross-reaction, molecular weight, and amino acid composition. Most recently, Higo and Loertscher [9] reported the amino-terminal sequences of five *E. coli* 30 S ribosomal proteins and their functional equivalents from *B. stearothermophilus* strain 799.

2. Materials and methods

Ribosomal proteins of the 30 S subunit from *E. coli* strain Q13 were fractionated on carboxymethylcellulose column chromatography [10] and those from *B. stearothermophilus* strain 10 on phosphocellulose column chromatography [11]. They were further purified by gel filtration chromatography on a Sephadex G-100 column (2.5 × 270 cm) with 15% acetic as the eluant. Isolated proteins were identified by their amino acid compositions, their positions on two dimensional electrophoresis, their molecular weights and, in certain cases, their immunological cross-reactivities [3, 11–15].

The amino-terminal sequence of proteins was determined by automatic Edman degradation [16] using Beckman Model 890C sequencer with the quadrol program. The protein (3–8 mg) was dissolved in 0.5 ml of 50% formic acid containing 1 mg of dithioerythritol. The thiazolinone derivatives (or PTH-derivatives) were hydrolyzed separately with 6 N HCl and HI [17] at 130°C for 20 hr, and the amino acid formed was analyzed with a Durrum D-500 amino acid analyzer or Technicon TSM amino acid analyzer. When necessary, identification of the PTH-derivatives

Issued as N.R.C.C. Number 14138.
* Please address reprint requests to this author.

Table 1

The N-terminal sequence homologies between ribosomal proteins from the 30 S subunit of *Escherichia coli* Q13 (E) and *Bacillus stearothermophilus* IO (B)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ES2	ALA	THR	VAL	SER	MET	ARG	ASP	MET	LEU	LYS	ALA	GLY	VAL	HIS	PHE
BS2a	SER	VAL	ILE	SER	MET	LYS	GLN	LEU	LEU	GLU	SER	GLY	VAL	HIS	PHE
BS2b	MET	MET	TYR	ASN	VAL	ARG	GLY	GLU	ASN	LEU	GLU	VAL	THR	PRO	ALA
ES3	GLY	GLN	LYS	VAL	HIS	PRO	ASN	GLY	ILE	ARG	LEU	GLY	ILE	VAL	LYS
BS3/BS4	GLY	GLN	LYS	VAL	ASN	PRO	ILE	GLY	LEU	ARG	ILE	GLY	ILE	ILE	ARG
ES4	ALA	ARG	TYR	LEU	GLY	PRO	LYS	LEU	LYS	LEU	SER	ARG	ARG	GLU	GLY
BS5	ALA	ARG	TYR	THR	GLY	PRO	MET	TRP	LYS	ILE	SER	ARG	ARG	LEU	GLY
ES5	"Blocked" N-terminus														
BS6	MET	ARG	ARG	ILE	ASP	PRO	ASN	LYS	LEU	GLU	LEU	GLU	GLU	ARG	VAL
ES6	MET	ARG	HIS	TYR	GLU	ILE	VAL	PHE	MET	VAL	HIS	PRO	ASP	GLN	SER
BS9	MET	ARG	LYS	TYR	GLU	ILE	MET	TYR	ILE	ILE	ARG	PRO	ASP	MET	ASP
ES7	PRO	LYS	PHE	GLY	VAL	GLN	GLY	LEU	ARG	LYS	PHE	LEU	PRO	ASP	PRO
BS7	PRO	ARG	ARG	GLY	PRO	VAL	ALA	LYS	ARG	ASP	VAL	LEU	PRO	ASP	PRO
ES8	SER	MET	GLN	ASP	PRO	ILE	ALA	ASP	MET	LEU	THR	ARG	ILE	ARG	ASN
BS8	VAL	MET	THR	ASP	PRO	ILE	ALA	ASP	MET	LEU	THR	ARG	ILE	ARG	ASN
ES9	ALA	GLU	ASN	GLN	TYR	TYR	GLY	THR	GLY	ARG	ARG	LYS	SER	SER	ALA
BS10	ALA	GLN	VAL	GLN	TYR	TYR	GLY	THR	GLY	ARG	ARG	LYS	SER	SER	VAL
ES10	MET	GLN	ASN	GLN	ARG	ILE	ARG	ILE	ARG	LEU	LYS	ALA	PHE	ASP	HIS
BS13	ALA	LYS	GLU	LYS	ILE	ARG	ILE	ARG	LEU	LYS	ALA	TYR	ASP	HIS	ARG
ES11	PHE	LYS	ALA	PRO	ILE	ARG	ALA	ARG	LYS	ARG	VAL	ARG	LYS	GLN	VAL
BS11	ALA	ARG	ARG	THR	ASN	THR	ARG	LYS	ARG	ARG	VAL	ARG	LYS	ASN	ILE
ES12	ALA	THR	VAL	ASN	GLN	LEU	VAL	ARG	LYS	PRO	ARG	ALA	ARG	LYS	VAL
BS12	PRO	THR	ILE	ASN	GLN	LEU	VAL	ARG	LYS	GLY	ARG	GLU	LYS	LYS	VAL
ES13	ALA	ARG	ILE	ALA	GLY	ILE	ASN	ILE	PRO	ASP	HIS	LYS	HIS	ALA	VAL
BS14	ALA	ARG	ILE	ALA	GLY	VAL	ASP	ILE	PRO	ARG	ASP	LYS	ARG	VAL	VAL
ES14	ALA	LYS	GLN	SER	MET	LYS	ALA	ARG	GLU	VAL	LYS	ARG	VAL	SER	LEU
BS21	ALA	LYS	LYS	SER	MET	ILE	ALA	LYS	GLN	LYS	ARG	THR	PRO	LYS	PHE
ES15	SER	LEU	SER	THR	GLU	ALA	THR	ALA	LYS	ILE	VAL	SER	GLU	PHE	GLY
BS18	ALA	LEU	THR	GLN	GLU	ARG	LYS	ARG	GLU	ILE	ILE	GLY	GLU	PHE	LYS
ES16	MET	VAL	THR	ILE	ARG	LEU	ALA	ARG	HIS	GLY	ALA	LYS	LYS	ARG	PRO
BS15	ALA	VAL	LYS	ILE	ARG	LEU	LYS	ARG	MET	GLY	THR	LYS	LYS	LYS	PRO
ES17	THR	ASP	LYS	ILE	ARG	THR	LEU	GLN	GLY	ARG	VAL	VAL	SER	ASP	LYS
BS16	SER	GLN	ARG	ASN	GLN	ARG	LYS	VAL	THR	VAL	GLY	ARG	VAL	VAL	GLY
ES18	"Blocked" N-terminus														
BS19	ALA	GLY	ARG	LYS	GLY	GLY	ARG	GLY	LYS	ARG	ARG	LYS	VAL	CYS	TYR
ES19	PRO	ARG	SER	LEU	LYS	LYS	GLY	PRO	PHE	ILE	ASP	LEU			
BS17	GLY	ARG	SER	LEU	LYS	LYS	GLY	PRO	PHE	SER	ASP	GLY	HIS	ILE	MET
ES20	ALA	ASN	ILE	LYS	SER	ALA	LYS	LYS	ARG	ALA	ILE	GLN	SER	GLU	LYS
BS20	ALA	ASN	ILE	LYS	SER	ALA	ILE	LYS	ARG	ALA	LYS	THR	SER	GLU	LYS
ES21	PRO	VAL	ILE	LYS	VAL	ARG	GLY	ASN	GLU	PRO	PHE	ASP	VAL	ALA	LEU
BS6a	PRO	MET	GLU	VAL	PHE	GLU	GLN	ALA	LEU	LYS	ASN	SER			

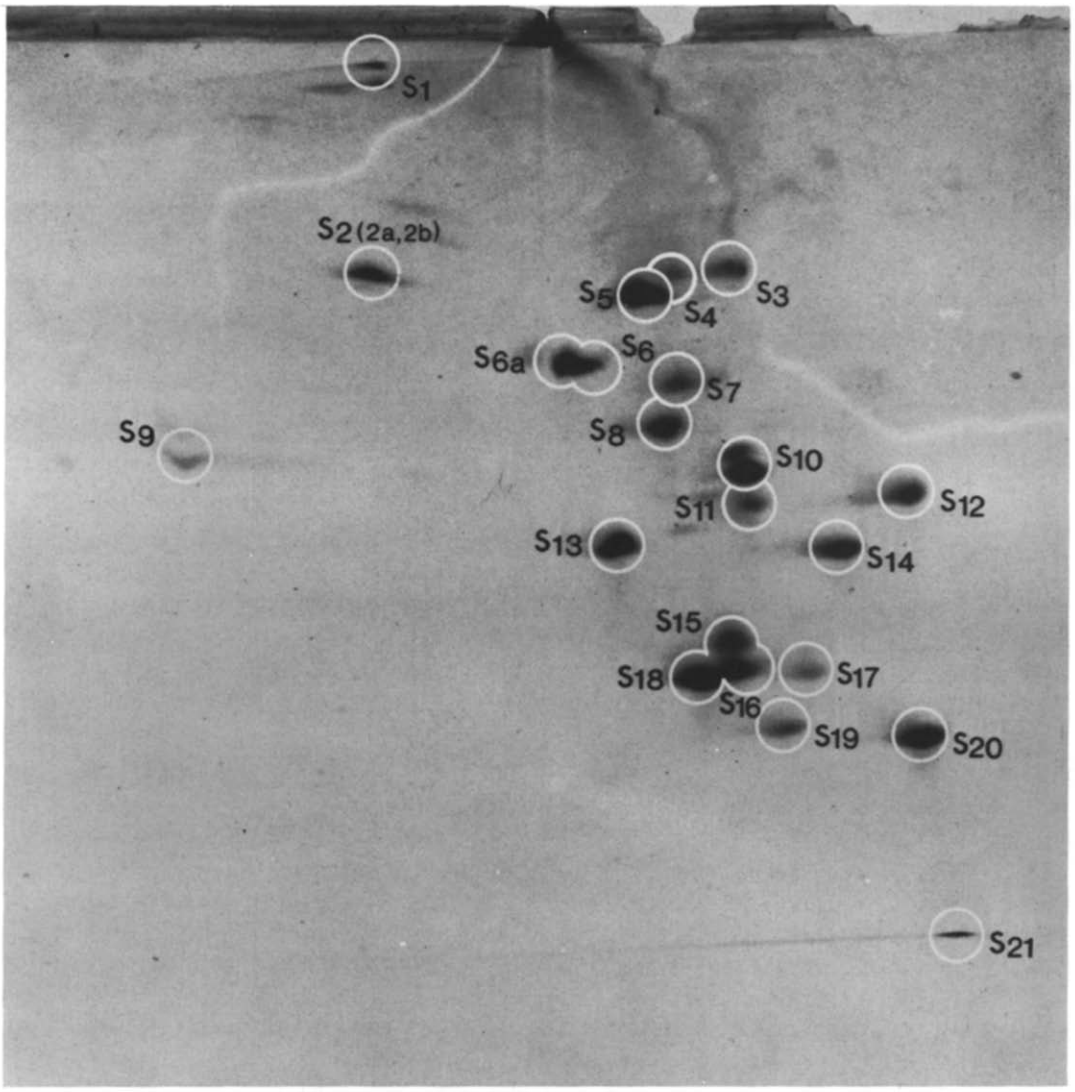

Fig. 1. Two-dimensional electropherogram of *B. stearothermophilus* 30 S proteins indicating the positions of the proteins whose amino-terminal sequence is depicted in table 1.

was made by thin layer chromatography on silica gel plates [18].

3. Results and discussion

The N-terminal amino acid sequences (12–15 residues) of 17 ribosomal proteins from the 30 S subunit of *Escherichia coli* strain Q13 (EC proteins) and 21 proteins from *Bacillus stearothermophilus* strain 10 (BS proteins) and their proposed structural equivalents based on the N-terminal sequence homologies are shown in table 1. The nomenclature based the two dimensional electrophoresis is that of Wittmann et al. [12] for EC proteins and of Isono et al. [3] for BS proteins. Fig. 1 shows the two dimensional electrophoretic pattern of BS proteins from strain 10 and indicates the positions of three proteins (BS 2a, BS 2b, and BS 6a) in addition to those given in the previous paper [3]. The position indicated

for BS S2 [3] actually consists of two spots (BS 2a and BS 2b) which are very close to each other, but they have different N-terminal amino acid sequences.

Our sequence data of EC proteins from strain Q13 (table 1) agree well with 18 EC proteins from strain K reported by Wittmann-Liebold [7] and 5 proteins from strain MRE 600 by Higo and Loertscher [9]. However, EC S10 from strain Q13 and MRE 600 has different residues from EC S10 of strain K [7] at positions 7, 8, 9, 10, and 11. EC S7 from Q13 is different from strain K [7] at positions 4, 6, 8, 9, and 11, but our preparation of EC S7 from Q13 is still heterogeneous and its sequence given in table 1 should be considered as tentative.

Among 23 BS proteins from strain 10 only one protein (BS S1) has blocked N-terminus, and the remaining have specific N-terminal sequences which are distinctly different from each other except BS S3 and BS S4 have an identical N-terminal sequence.

When the N-terminal sequences of BS proteins are compared with EC proteins, 18 pairs between EC and BS proteins can be formed. The three pairs (EC-S10 and BS S13, EC S17 and BS S16, and EC S21 and BS S6a), however, require some shifting of positions in order to align the corresponding identical residues. The number of identical residues between the 18 pairs varies from 4 to 12 residues per 12–15 residues.

Since both EC S5 and EC S18 have blocked N-termini and their N-terminal sequences are not yet known, it is not possible to make direct sequence comparison with BS proteins. However, the similarities in positions in the two-dimensional electrophoresis, amino acid composition, and molecular weights suggest that BS S6 and BS S19 are corresponding to EC S5 and EC S18 respectively. BS S2b has so far no corresponding EC protein. It is uncertain whether this is a ribosomal protein specific to *B. stearothermophilus* or a supernatant factor which tightly bound to the ribosomal subunits.

The N-terminal amino acid sequence of each individual protein from *E. coli* and *B. stearothermophilus* are quite different from each other suggesting a unique primary structure and a specific function for each protein. It is quite evident that there is essentially one-to-one structural correspondence between each EC and BS protein, indicating a high degree of amino acid sequence conservation during evolution.

The fidelity to which these homologies are evinced at the level of primary sequence of each corresponding protein in these procaryotes is important for future assessments of at least two important questions concerning ribosome structure-function relationships: a) the chemical features or 'active sites' in the proteins which allow for overall conservation of function in the procaryotic translation apparatus, and b) whether some higher level of homology exists that is to say, in the symmetrical arrangements of the proteins or active sites within the ribosomal architecture. Sequence data of the type presented herein then becomes an important basis from which these assessments can be made.

Acknowledgements

We are grateful for the technical assistance provided by Mr J. Christ, Mr F. Rollin, Mr W. Rowsome, Mr C. Roy, Mr R. Reitmeier, Miss N. Smith and Mr R. Walker.

References

[1] Yaguchi, M., Roy, C., Matheson, A. T. and Visentin, L. P. (1973) Canadian J. Biochem. 51, 1215–1217.
[2] Visentin, L. P., Matheson, A. T. and Yaguchi, M. (1974) FEBS Letters 41, 310–314.
[3] Isono, K., Isono, S., Stöffler, G., Visentin, L. P., Yaguchi, M. and Matheson, A. T. (1973) Molec. Gen. Genet. 127, 191–195.
[4] Oda, G., Strøm, A. R., Visentin, L. P. and Yaguchi, M. FEBS Letters (in press).
[5] Nomura, M., Traub, P. and Bechmann, H. (1968) Nature 219, 793–799.
[6] Higo, K., Held, W., Kahan, L. and Nomura, M. (1973) Proc. Natl. Acad. Sci. U. S. 70, 944–948.
[7] Wittmann-Liebold, B. (1973) FEBS Letters 36, 247–249.
[8] Geisser, M., Tischendorf, G. W., Stöffler, G. (1973) Molec. Gen. Genet. 127, 129–145.
[9] Higo, K. and Loertscher, K. (1974) J. Bacteriol. 118, 180–186.
[10] Hindennach, I., Stöffler, G. and Wittmann, H. G. (1971) Eur. J. Biochem. 23, 7–11.
[11] Visentin, L. P., Matheson, A. T. and Yaguchi, M. Manuscript in preparation.
[12] Wittmann, H. G., Stöffler, G., Hinddennach, I., Kurland, C. G., Randall-Hazelbauer, L., Birge, E. A., Nomura, M., Kaltschmidt, E., Mizushima, S., Traut, R. R. and Bickle, T. A. (1971) Molec. Gen. Genet. 111, 327–333.
[13] Kaltschmidt, E., Dzionara, M. and Wittmann, H. G. (1970) Molec. Gen. Genet. 109, 292–307.

[14] Moore, P. B., Traut, R. R., Noller, H., Pearson, P. and Delius, H. (1968) J. Mol. Biol. 31, 441–461.
[15] Craven, G. R., Voynow, P., Hardy, S. J. S. and Kurland, C. G. (1969) Biochemistry 8, 2906–2915.
[16] Edamn, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[17] Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G. and Ballantyne, D. L. (1971) Biochemistry 10, 4912–4921.
[18] Wittmann-Liebold, B. (1973) Hoppe-Seyloer's Z. Physiol. Chem. 354, 1415–1431.