THE PRIMARY STRUCTURE OF PROTEIN S10 FROM THE SMALL RIBOSOMAL SUBUNIT OF *ESCHERICHIA COLI*

Makoto YAGUCHI, Camille ROY and H. G. WITTMANN+

Division of Biological Sciences, National Research Council of Canada, Ottawa, Ontario K1A OR6, Canada and *+Max-Planck-Institut für Molekulare Genetik, D-1000 Berlin 33 (Dahlem), Germany*

Received 15 September 1980

1. Introduction

As revealed by immuno electron microscopy studies [1,2] protein S10 is located on the 'head' of the 30 S subunit of the *Escherichia coli* ribosome. It can be crosslinked to proteins S1 and S3 by bifunctional reagents [3] suggesting the proteins are neighbours. Neutron scattering studies indicate [4] protein S10 is also in close proximity to protein S9.

Interaction between protein S10 and proteins S3 and S9 (as well as S14) has been observed during the in vitro assembly process. S10 is one of the proteins which is incorporated into the intermediate reconstitution particles at a rather late stage. Single component omission experiments indicated that protein S10 is involved in tRNA-binding to the ribosomes (reviewed in [5]).

Here, we report the complete primary structure of protein S10 and present its secondary structure based on 4 prediction programmes. Furthermore, a comparison of the primary structure of protein S10 has been made with that of other ribosomal proteins.

2. Materials and methods

Protein S10 was isolated from *E. coli* strain K as in [6]. The identity and purity of the protein were checked by two-dimensional polyacrylamide gel electrophoresis [7].

Performic acid oxidation [8] was done at 0°C for 16 h. Tryptophan was estimated as in [9]. Amino acid compositions were determined with a Durrum D-500 amino acid analyzer.

Large fragments of protein S10 were produced by cyanogen bromide [10] and *N*-bromosuccinimide cleavage [11]. The isolation of the large peptides was achieved by gel filtration on a Sephadex G-50 (superfine) column (200 × 1.5 cm) and 44% formic acid was used for elution.

Tryptic and chymotryptic digestions were done at pH 8.0 and 37°C for 4 or 24 h. Digestion with thermolysin was done at pH 8.0 and 55°C for 24 h. Digestion with *Staphylococcus aureus* protease was done in 50 mM acetic acid, pH 4.0 for 16 h [12].

Gel filtration of various enzymatic digests (~5 mg) was performed on a Sephadex G-50 (superfine) column (250 × 1.5 cm). Small peptides in the various pooled fractions eluted from the columns were further separated by fingerprinting on cellulose thin-layer plates [13].

The carboxyl-terminal end of protein S10 was digested by carboxypeptidases A and B, or carboxypeptidase Y (Pierce Chemical Co.).

Automatic Edman degradation [14] of protein S10 and its large peptides was made in a Beckman model 890C sequenator utilizing 0.5 M quadrol protein programme 122974 or DMAA peptide programme 102974. Polybrene [15] was used as carrier for peptides. The amino acid sequence of small peptides was determined by a manual micro-Edman technique [16] without dansylation.

The thiazolinone or PTH derivatives were hydrolysed with 6 N HCl in the presence or absence of 0.1% $SnCl_2$ [17] at 130°C for 20 h, and the amino acid formed was analyzed with a Durrum D-500 amino acid analyzer. The identification of some PTH derivatives such as Asp, Asn, Glu and Gln was made by thin-layer chromatography on silica gel plates [18,19].

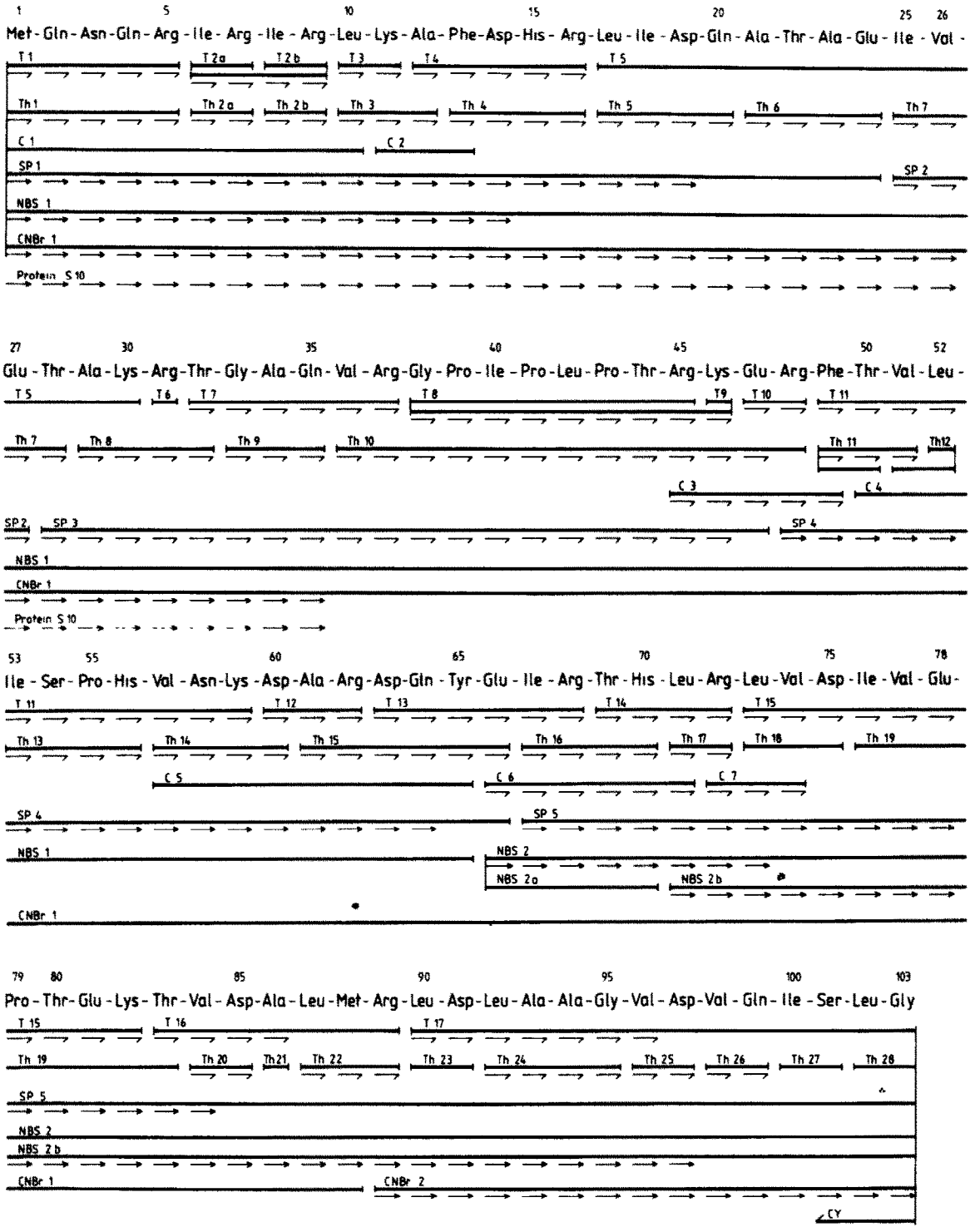

Fig.1. Amino acid sequence of protein S10 from *Escherichia coli* ribosomes: T, tryptic peptide; Th, thermolytic peptide; C, chymotryptic peptide; SP, peptide from digestion with *Staphylococcus aureus* protease; NBS, peptide cleaved with N-bromosuccinimide; CNBr, peptide cleaved with cyanogen bromide; CY, portion digested by carboxypeptidase Y; →, identified residue with Beckman sequenator; ⟶, identified residue with manual Edman degradation.

3. Results and discussion

The automatic sequence determination of the amino-terminal region of protein S10 and peptides SP1, NBS1 and CNBr1 provided the amino acid sequence of the first 35 residues as shown in fig.1, and confirmed the report [20] about the first 30 residues.

The sequence of the central region (res. 28–66) was provided mainly by two peptides (SP3, SP4). Chymotrypsin peptide C3 aligned peptide SP4 to SP3. Tryptic peptide T13 gave the overlap among SP4, SP5 and NBS2.

The three peptides (SP5, NBS2b, CNBr2) containing the carboxyl-terminal end of protein S10 were sequenced automatically. Using polybrene it was possible to completely sequence CNBr2. Though carboxypeptidase A and B did not release any amino acids, carboxypeptidase Y released Gly, Leu and Ser. These data established the amino acid sequence of the carboxyl-terminal region (res. 67–103) of protein S10.

All tryptic and thermolytic peptides expected from the amino acid sequence of protein S10 were isolated and most of them were manually sequenced as shown in fig.1.

The combination of these results gave the alignment of all peptides examined and the complete amino acid sequence of protein S10 as illustrated in fig.1.

The amino acid composition derived from the sequence is: Asp_8, Asn_2, Thr_8, Ser_2, Glu_6, Gln_6, Pro_5, Gly_4, Ala_9, Val_9, Met_2, Ile_9, Leu_{10}, Tyr_1, Phe_2, His_3, Lys_5 and Arg_{12}. Cysteine and tryptophan are absent. The amino acid composition obtained from the sequence is in excellent agreement with the data determined from the hydrolysis of the whole protein. The protein consists of 103 amino acid residues and, calculated from the above composition, the M_r is 11 736. The amino acid sequence given in fig.1 is in complete agreement with the DNA sequence of the gene for protein S10 (P. Olins and M. Nomura, personal communication).

Based on 4 programmes for the prediction of the secondary structure of proteins (details in [21]) a diagram has been drawn for the secondary structure of protein S10 (fig.2). There is a long α-helix consisting of ~20 amino acids at the N-terminus (pos. 10–30). Turns and β-sheet structure alternate from pos. 30 to the C-terminus. A possible helix region in the C-terminal region of the protein chain is much less probable than that at the N-terminus. The prediction programmes indicate protein S10 contains 19% α-helix, 5% β-sheet and 12% turns.

A computer search for homologous regions between protein S10 and other ribosomal proteins

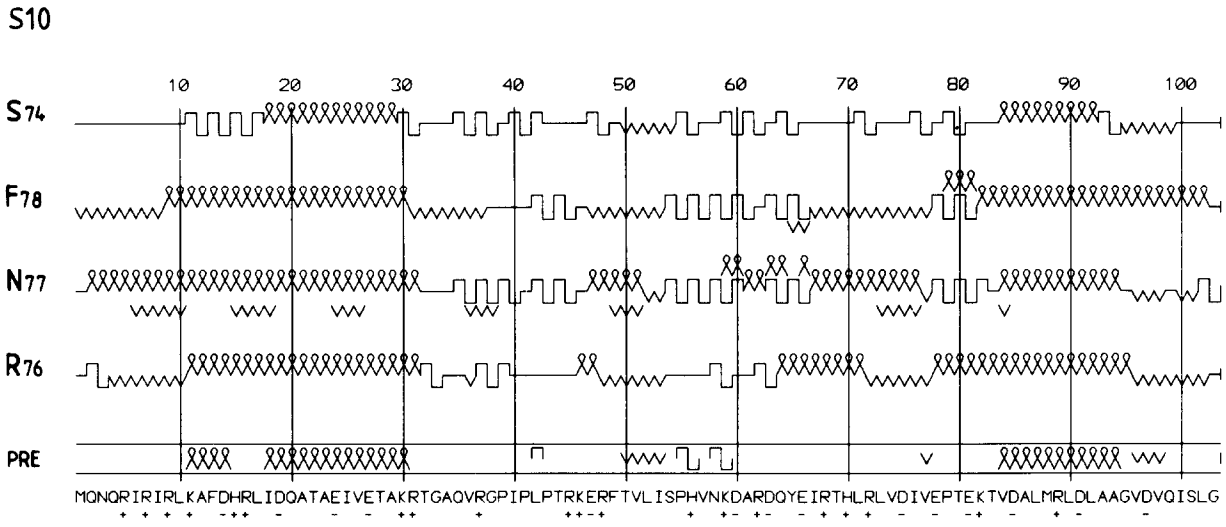

Fig.2. Secondary structure of protein S10 as predicted according to 4 different methods (S, F, N, R). The line 'PRE' summarizes the secondary structure obtained when at least 3 out of the 4 predictions are in agreement. The amino acid sequence of protein S10 is shown in the bottom line in the one-letter code (details in [21]).

Table 1
Homologous regions of protein S10 and other ribosomal proteins

S10	(pos. 4- 9):	Gln-Arg-Ile-Arg-Ile-Arg
S4	(pos. 58- 63):	Gln-Lys-Val-Arg-Arg-Ile
S10	(pos. 90- 93):	Leu Asp-Leu-Ala
S4	(pos.156-159).	Leu-Glu-Leu-Ala
S10	(pos. 5- 12):	Arg-Ile-Arg-Ile-Arg-Leu-Lys-Ala
S14	(pos. 80- 87):	Arg-Ile-Lys-Val-Arg-Glu-Ala-Ala
S10	(pos. 47- 56):	Glu-Arg-Phe-Thr-Val-Leu-Ile-Ser-Pro-His
L4	(pos.122-130):	Glu-Lys-Phe-Ser-Val-Glu-Ala———Pro-Lys
S10	(pos. 94-100):	Ala-Gly-Val-Asp-Val-Glu-Ile
L6	(pos. 12- 18):	Ala-Gly-Val-Asp-Val-Lys-Ile
S10	(pos. 59- 63):	Lys-Asp-Ala-Arg-Asp
L10	(pos.159-163):	Arg-Asp-Ala-Lys-Glu
L7/12	(pos. 81- 85):	Lys-Glu-Ala-Lys-Asp
S10	(pos. 4- 8):	Gln-Arg-Ile-Arg-Ile
L22	(pos. 15- 19):	Gln-Lys-Val-Arg-Leu
S10	(pos. 81- 92):	Glu-Lys-Thr-Val-Asp-Ala-Leu-Met-Arg-Leu-Asp-Leu
L29	(pos. 8- 18):	Glu-Lys-Ser-Val-Glu-Glu-Leu-Asn-Thr———Glu-Leu
S10	(pos. 5- 10):	Arg-Ile-Arg-Ile-Arg-Leu
L33	(pos. 5- 10):	Arg-Glu-Lys-Ile-Lys-Leu
S10	(pos. 4- 8):	Gln-Arg-Ile-Arg-Ile
B-L9	(pos. 64- 68)	Gln-Lys-Ile-Lys-Val
S10	(pos. 59- 66):	Lys-Asp-Ala-Arg-Asp-Gln-Tyr-Glu
B-L9	(pos. 81- 88):	Lys-Glu-Ala-Lys-Glu-Leu-Val-Asp

gave the results listed in table 1. In addition to the data given in table 1 there are some homologies between protein S10 (pos. 52–102) and protein L4 (pos. 127–182) as well as between protein S10 (pos. 80–101) from *E. coli* and protein L20 (pos. 15–36) from *Halobacterium cutirubrum*. There is considerable homology between S10 from *E. coli* and the corresponding ribosomal proteins from *Bacillus subtilis* and *Bacillus stearothermophilus* (summarized in [22]).

Acknowledgements

We thank Dr Wittmann-Liebold for provision of computer data including homologies and secondary structure, and Dr R. A. F. Reithmeier, Joan Logan, and Sussan MacMillan for technical assistance. This is NRCC publication no. 18496.

References

[1] Tischendorf, G. W., Zeichhardt, H. and Stöffler, G. (1975) Proc. Natl. Acad. Sci. USA 72, 4820–4824.
[2] Lake, J. A. (1980) in: Ribosomes (Chambliss, G. et al. eds) pp. 207–236, University Park Press, Baltimore, MD.
[3] Traut, R. R., Lambert, J. M., Boileau, G. and Kenny, J. W. (1980) in: Ribosomes (Chambliss, G. et al. eds) pp. 89–110, University Park Press, Baltimore, MD.
[4] Moore, P. B. (1980) in: Ribosomes (Chambliss, G. et al. eds) pp. 111–133, University Park Press, Baltimore, MD.
[5] Nomura, M. and Held, W. A. (1974) in: Ribosomes (Nomura, M. et al. eds) pp. 193–223, Cold Spring Harbor Laboratory Press, Long Island, NY.
[6] Hindennach, I., Stöffler, G. and Wittmann, H. G. (1971) Eur. J. Biochem. 23, 7–11.
[7] Kaltschmidt, E. and Wittmann, H. G. (1970) Analyt. Biochem. 36, 401–412.
[8] Hirs, C. H. W. (1956) J. Biol. Chem. 219, 611–621.
[9] Liu, T. Y. and Chang, Y. H. (1971) J. Biol. Chem. 248, 2842–2848.
[10] Gross, E. (1967) in: Methods in Enzymology (Hirs, C. H. W. ed) vol. 11, pp. 238–255, Academic Press, New York.
[11] Ramachandran, L. K. and Witkop, B. (1967) in: Methods in Enzymology (Hirs, C. H. W. ed) vol. 11, pp. 283–299, Academic Press, New York.
[12] Houmard, J. and Drapeau, G. R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.
[13] Yaguchi, M., Wittmann, H. G., Cabezon, T., De Wilde, M., Villarroel, R., Herzog, A. and Bollen, A. (1975) Mol. Gen. Genet. 142, 35–43.
[14] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[15] Klapper, D. G., Wilde, G. E. and Capra, J. D. (1978) Analyt. Biochem. 85, 126–131.
[16] Bruton, C. J. and Hartley, B. S. (1970) J. Mol. Biol. 52, 165–178.
[17] Mendez, E. and Lai, S. Y. (1975) Analyt. Biochem. 68, 47–53.
[18] Wittmann-Liebold, B., Geissler, A.-W. and Marzinzig, E. (1975) J. Supramol. Struct. 3, 426–447.
[19] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.
[20] Higo, K. and Loertscher, K. (1974) J. Bacteriol. 118, 180–186.
[21] Dzionara, M., Robinson, S. M. L. and Wittmann-Liebold, B. (1977) Hoppe-Seyler's Z. Physiol. Chem. 358, 1019–1103.
[22] Matheson, A. T., Möller, W., Amons, R. and Yaguchi, M. (1980) in: Ribosomes (Chambliss, G. et al. eds) pp. 297–332, University Park Press, Baltimore, MD.