PRIMARY STRUCTURE OF PROTEIN S13 FROM THE SMALL RIBOSOMAL SUBUNIT OF *ESCHERICHIA COLI*

Harald LINDEMANN and Brigitte WITTMANN-LIEBOLD
Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Berlin-Dahlem, Germany

Received 4 October 1976

1. Introduction

Protein S13 is a rather basic component of the small subunit of the *Escherichia coli* ribosome with an isoelectric point of pH > 12 [1]. Binding of S13 to the 3′-terminal region of the 16 S ribosomal RNA of *E. coli* has been shown recently [2]. For more intensive study of this RNA–protein interaction, knowledge of the primary structure will be very helpful. Protein S13 is located at the top of the small head of the 30 S ribosomal subunit, in close proximity to protein S14, as shown by immuno-electronmicroscopy [3–5]. Other neighbours of S13, namely S4 and S19, have been detected by cross-linking methods [6–8]. Furthermore, S13 has been shown to be located near to the binding site of IF2 and IF3 [9,10]. These results are supported by the finding that S13 plays some role in initiation [11].

In this paper we present the complete structure of protein S13. Predictions of the secondary structure of protein S13 are made and comparison of the sequence of S13 with other ribosomal proteins of *E. coli* of known primary structure is performed.

2. Materials and methods

Protein S13, isolated from *E. coli* K strain A19 as described earlier [12], was provided by Dr H. G. Wittmann. The purity of the protein was checked by two-dimensional polyacrylamide gel electrophoresis [13].

Enzymatic digesiton with trypsin treated with 1-chloro-4-phenyl-3-tosylamidobutan-2-one (TPCK) from Merck (Darmstadt) was performed at pH 7.8 at 37°C for 4 h. Digestion with thermolysin was done at pH 8.0 at 37°C for 1 h and with *Staphylococcus aureus* protease [14] in 0.3% acetic acid, pH 4.3, at 37°C for 20 h. Details of the enzymatic digestion procedures have been described previously [15,16].

The resultant peptides were isolated by the fingerprint technique on cellulose thin-layer plates [15,16] in order to obtain information about the number of peptides and their amino acid compositions. The peptides were then separated by gel filtration on Sephadex G-50 (superfine, 140 × 1 cm) in 10% acetic acid [17] and, when necessary, by rechromatography on a micro-column of Dowex M 71 (0.3 × 10 cm) at 55°C with pyridine–formate gradients [17,18]. The fractions were examined by spotting portions of every second fraction onto cellulose thin-layer plates. Fractions containing more than one peptide were further purified by chromatography on cellulose thin-layer sheets.

Amino acid analyses were performed on a Durrum D 500 analyser and with an LKB-Biocal 3201 analyser, as described previously [19]. Performic acid oxidation was carried out according to the method of Hirs [20]. For the determination of tryptophan, protein S13 was either hydrolysed with 4 N methane sulphonic acid containing 0.2% tryptamine for 24 h at 115°C, or a tryptic fingerprint of S13 was developed with Ehrlich's reagent.

The amino acid sequences of the peptides were determined by various methods:

(1) Automated Edman degradation [21] of intact protein in an improved Beckman sequenator [22].

(2) Edman degradation using the solid-phase technique [23] with attachment of the C-terminal carboxyl groups to amino–polystyrene resins [24,25].

(3) Manual Edman degradation combined with dansylation of the free N-terminal amino acid residue after each degradation step [26,27].

The released 2-anilino-5-thiazolinone derivatives of glutamic acid, aspartic acid and their amides were converted to the phenylthiohydantoin derivatives in order to distinguish the free acids from their amides by thin-layer chromatography [27]. In case of uncertainties after dansylation, the free amino acids were liberated from their 2-anilino-5-thiazolinones or phenylthiohydantoin derivatives by hydrolysis and analysed on an amino acid analyser.

3. Results and discussion

Ribosomal protein S13 consists of 117 amino acid residues with the amino acid composition of Asp_6, Asn_2, Thr_6, Ser_6, Glu_6, Gln_2, Pro_4, Gly_{11}, Ala_{11}, Cys_1, Val_7, Met_2, Ile_{12}, Leu_9, Tyr_2, Phe_1, His_3, Lys_{11} and Arg_{15}. The cysteine residue was determined as cysteic acid after oxidation of S13 and appropriate peptides with performic acid followed by amino acid analysis. Analysis of S13 after hydrolysis with methane sulphonic acid [28] gave no indication of the presence of tryptophan in this ribosomal protein. In addition, the appropriate test for tryptophan on a tryptic thin-layer fingerprint gave negative results. According to this amino acid composition the molecular weight of S13 is 12 970. Its complete primary structure is shown in fig.1. Out of a total of 117 amino acid residues, there are 29 basic residues, 15 of which are located in the C-terminal third (positions 86–117). The 12 acidic residues are distributed over the N-terminal two-thirds (positions 1–85). Thus, protein S13 contains a long C-terminal region clustered with basic residues (15 out of 32 amino acids, without any acidic residue). The acidic amino acids mainly occur in the middle part of the molecule (positions 40–71), a finding which had been also observed for the RNA binding protein S8 [16].

Sequence determination was accomplished by combination of the results obtained from automated liquid-phase Edman degradation of the intact protein with the results obtained from Edman degradation of the peptides resulting from enzymatic digestions of S13 using TPCK-trypsin, thermolysin and *Staphylococcus aureus* protease.

The alignment of the peptides T1–T2–T3–T4–T5–T6–T7–T8 (positions 1–60), which were all sequenced by the dansyl-Edman degradation or the solid-phase technique, was done by employing the sequence obtained from liquid-phase Edman degradation [21] of the intact protein S13 [29]. In addition, the analysis of peptide SP-1 (positions 1–40) and its sequence determination with the solid-phase technique up to position 24, as well as the sequence of SP-1a (positions 29–40), SP-2 (positions 41–46), SP-3 (positions 47–49) and SP-4 (positions 50–58) were in accordance with these results. Furthermore, isolation of various peptides derived from cleavage of protein S13 with thermolysin as well as dansyl-Edman degradations of part of these peptides confirmed the sequence of the N-terminal region. The peptide SP-5 (positions 59–65) overlapped the tryptic peptides T8–T9; SP-6 (positions 66–71) the tryptic peptides T9–T10–T11; SP-7 (positions 72–81) the tryptic peptides T11–T10–T12 and SP-8 (positions 82–117) the tryptic peptides T12–T13–T14–T10–T15–T16–T17–T18–T4–T19–T20–T21–T19. The sequences obtained for SP-5, SP-6, SP-7 and the N-terminal sequence of SP-8 were in full agreement with the sequences of the tryptic peptides. In addition there are overlaps of the thermolysin peptides belonging to this region. (For details see figure 1.)

Cleavage of S13 at aspartic acid (position 81)

Fig.1. Primary structure of ribosomal protein S13 from *E. coli*. Symbols: SQ, Edman degradation of intact protein in an improved Beckman sequenator including automatic conversion reaction; T, tryptic digestion; SP, digestion with *Staphylococcus aureus* protease; TH, digestion with thermolysin. (⟷) Identification by thin-layer chromatography of the PTH derivatives released with Beckman sequenator (⟶) identification by thin-layer chromatography of the PTH derivatives released by solid-phase method (⟶) identification as the dansylated amino acids with combined dansyl-Edman degradation technique (⟶) identification as the PTH derivatives after combined dansyl-Edman degradation and conversion of the PTC derivatives (●⟶) identification as the free amino acids after combined dansyl-Edman degradation and hydrolysis of the PTC or PTH derivatives (▲) identification as cysteic acid after oxidation with performic acid.

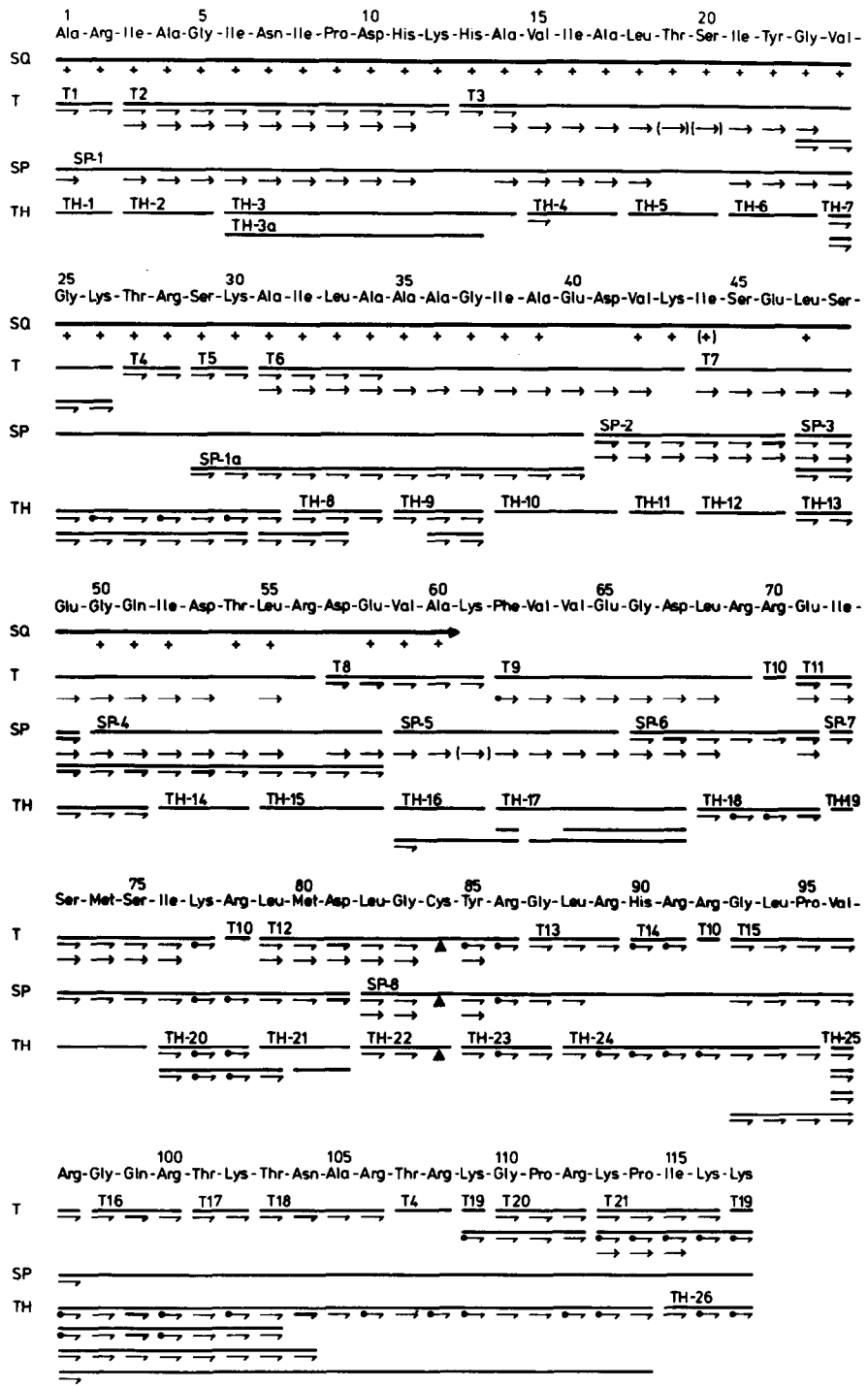
1 5 10 15 20
Ala-Arg-Ile-Ala-Gly-Ile-Asn-Ile-Pro-Asp-His-Lys-His-Ala-Val-Ile-Ala-Leu-Thr-Ser-Ile-Tyr-Gly-Val-
SQ
T
T1
T2
T3
SP
SP-1
TH
TH-1
TH-2
TH-3
TH-3a
TH-4
TH-5
TH-6
TH-7
25 30 35 40 45
Gly-Lys-Thr-Arg-Ser-Lys-Ala-Ile-Leu-Ala-Ala-Ala-Gly-Ile-Ala-Glu-Asp-Val-Lys-Ile-Ser-Glu-Leu-Ser-
SQ
(+)
T
T4
T5
T6
T7
SP
SP-2
SP-3
SP-1a
TH
TH-8
TH-9
TH-10
TH-11
TH-12
TH-13
50 55 60 65 70
Glu-Gly-Gln-Ile-Asp-Thr-Leu-Arg-Asp-Glu-Val-Ala-Lys-Phe-Val-Val-Glu-Gly-Asp-Leu-Arg-Arg-Glu-Ile-
SQ
T
T8
T9
T10
T11
SP
SP-4
SP-5
SP-6
SP-7
TH
TH-14
TH-15
TH-16
TH-17
TH-18
TH19
75 80 85 90 95
Ser-Met-Ser-Ile-Lys-Arg-Leu-Met-Asp-Leu-Gly-Cys-Tyr-Arg-Gly-Leu-Arg-His-Arg-Arg-Gly-Leu-Pro-Val-
T
T10
T12
T13
T14
T10
T15
SP
SP-8
TH
TH-20
TH-21
TH-22
TH-23
TH-24
TH-25
100 105 110 115
Arg-Gly-Gln-Arg-Thr-Lys-Thr-Asn-Ala-Arg-Thr-Arg-Lys-Gly-Pro-Arg-Lys-Pro-Ile-Lys-Lys
T
T16
T17
T18
T4
T19
T20
T21
T19
SP
TH
TH-26

Fig.1

Table 1
Identical tetrapeptides

Peptide	Protein	Positions	Protein	Positions
Lys His Ala Val	S13	12– 15	S6	93– 96
Ile Tyr Gly Val	S13	21– 24	S4	63– 66
Thr Arg Ser Lys	S13	27– 30	L32	8– 11
Leu Ala Ala Ala	S13	33– 36	L16	119–122
Glu Val Ala Lys	S13	58– 61	L10	16– 19
Ala Lys Phe Val	S13	60– 63	L1	35– 38
Val Glu Gly Asp	S13	64– 67	S8	50– 53
Arg Arg Gly Leu	S13	91– 94	S15	51– 54
Arg Gly Gln Arg	S13	97–100	L5	91– 94

during digestion with *Staphylococcus aureus* protease resulted in high peptide yields. The extremely basic C-terminus of S13 (positions 86–117) might favour this hydrolysis. Another unspecific cleavage of S13 by *Staphylococcus* protease occurring at arginine (position 28) was found in a rather low yield. All experiments which lead to the elucidation of the primary structure of protein S13 will be presented in detail elsewhere.

The secondary structure of protein S13 was predicted by two different methods. Based on the rules of Chou and Fasman [30,31], α-helical regions were predicted for positions 10–14/15, positions 28–49 and positions 55–65. β-Sheet regions were predicted for positions 14–24, positions 50–54/55, positions 76–85 and positions 98–101. The method of Burgess et al. [32] gave the following predictions: α-helix for positions 11–17, positions 30–40 and positions 55–62, whereas a β-sheet region for positions 18–22 was indicated. This shows that the predictions for the α-helical regions in S13 are almost in accord with both methods. Similarly, one β-sheet structure in the region of positions 14/18–22/24 was postulated. These predictions result in an α-helix content of approximately 23% in protein S13. The content of β-sheet structure cannot be judged so far, as the two methods gave different results.

The comparison of the sequence of S13 with other sequences of *E. coli* ribosomal proteins reveals a series of nine identical tetrapeptides (table 1), and 47 identical tripeptides. This represents a very low degree of sequence similarity when compared to findings with some other ribosomal proteins [17,33] and can easily be explained on a random basis [34].

Acknowledgements

We thank Mr Alfred Franz, Mr Arnold Lehmann and Mr Detlev Kamp for excellent technical assistance and Dr W. N. Strickland for kindly reading the manuscript.

References

[1] Kaltschmidt, E. (1971) Analyt. Biochem. 43, 25–31.
[2] Hochkeppel, H.-K., Spicer, E. and Craven, G. R. (1976) J. Mol. Biol. 101, 155–170.
[3] Tischendorf, G. W., Zeichhardt, H. and Stöffler, G. (1974) Molec. Gen. Genet. 134, 209–223.
[4] Tischendorf, G. W., Zeichhardt, H. and Stöffler, G. (1975) Proc. Natl. Acad. Sci. USA 72, 4820–4824.
[5] Lake, J. A. and Kahan, L. (1975) J. Mol. Biol. 99, 631–644.
[6] Sommer, A. and Traut, R. R. (1974) Proc. Natl. Acad. Sci. USA 71, 3946–3950.
[7] Lutter, L. C., Bode, U. and Kurland, C. G. (1974) Molec. Gen. Genet. 129, 167–176.
[8] Sun, T.-T., Bollen, A., Kahan, L. and Traut, R. R. (1974) Biochemistry 13, 2334–2340.
[9] Traut, R. R., Heimark, R. L., Sun, T.-T., Hershey, J. W. B. and Bollen, A. (1974) in: Ribosomes (Tissières, A., Nomura, M. and Lengyel, P., eds) pp. 271–308, Cold Spring Harbour Laboratory, New York.
[10] Bollen, A., Heimark, R. L., Cozzone, A., Traut, R. R. and Hershey, J. W. B. (1975) J. Biol. Chem. 250, 4310–4314.
[11] Lelong, J. C., Gros, D., Gros, F., Bollen, A., Maschler, R. and Stöffler, G. (1974) Proc. Natl. Acad. Sci. USA 71, 248–252.
[12] Hindennach, I., Stöffler, G. and Wittmann, H. G. (1971) Eur. J. Biochem. 23, 7–12.
[13] Kaltschmidt, E. and Wittmann, H. G. (1970) Analyt. Biochem. 36, 401–412.

[14] Houmard, J. and Drapeau, G. R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.
[15] Bitar, K. G. and Wittmann-Liebold, B. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1343–1352.
[16] Stadler, H. and Wittmann-Liebold, B. (1976) Eur. J. Biochem. 66, 49–56.
[17] Heiland, I., Brauer, D. and Wittmann-Liebold, B., (1976) Hoppe-Seyler's Z. Physiol. Chem. in press.
[18] Chen, R. and Wittmann-Liebold, B. (1975) FEBS Lett. 52, 139–140.
[19] Wittmann-Liebold, B., Greuer, B. and Pannenbecker, R. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1977–1979.
[20] Hirs, C. H. W. (1967) Meth. in Enzymol. 11, 59–62.
[21] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[22] Wittmann-Liebold, B. (1973) Hoppe-Seyler's Z. Physiol. Chem. (1973) 354, 1415–1431.
[23] Laursen, R. A. (1971) Eur. J. Biochem. 20, 89–102.
[24] Previero, A., Derancourt, J., Coletti-Previero, M. A. and Laursen, R. A. (1975) FEBS Lett. 33, 135–138.
[25] Wittmann-Liebold, B. and Lehmann, A. (1975) in: Solid Phase Methods in Protein Sequence Analysis (Laursen, R. A., ed) pp. 81–90, Pierce Chemical Company, Rockford, USA.
[26] Gray, W. R. and Hartley, B. S. (1963) Biochem. J. 89, 379–380.
[27] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.
[28] Liu, T. Y. and Chang, Y. H. (1971) J. Biol. Chem. 246, 2842–2848.
[29] Wittmann-Liebold, B. (1973) FEBS Lett. 36, 247–249.
[30] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 211–222.
[31] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 222–245.
[32] Burgess, A. W., Ponnuswamy, P. K. and Sheraga, H. A. (1974) Israel J. Chem. 12, 239–286.
[33] Wittmann-Liebold, B., Marzinzig, E. and Lehmann, A. (1976) FEBS Lett. 68, 110–114.
[34] Wittmann-Liebold, B. and Dzionara, M. (1976) FEBS Lett. 65, 281–283.