PRIMARY STRUCTURE OF PROTEIN L28 FROM THE LARGE SUBUNIT OF *ESCHERICHIA COLI* RIBOSOMES

Brigitte WITTMANN-LIEBOLD and Elke MARZINZIG
Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, 1 Berlin 33 (Dahlem), Germany

Received 27 June 1977

1. Introduction

Elucidation of the amino acid sequences of ribosomal proteins is a necessary step towards an understanding of the ribosomal structure and function at the molecular level. So far the primary structures of about two thirds of the 54 *Escherichia coli* ribosomal proteins have been determined in our laboratory or in collaboration with other groups. In continuing these studies we have recently accomplished the determination of the amino acid sequence of protein L28 isolated from the 50 S ribosomal subunit.

In this paper we present the complete primary structure of this protein and predict its secondary structure according to four methods. Furthermore, a search was made for peptides which are identical in protein L28 and in any other of the *E. coli* ribosomal proteins whose sequences are already known.

2. Materials and methods

Protein L28 was isolated from *E. coli*, strain K, as described earlier [1]. The purity of the protein samples, which were provided by Dr H. G. Wittmann, was tested by two-dimensional polyacrylamide gel electrophoresis [2].

Enzymatic digestions were performed with trypsin, treated with 1-chloro-4-phenyl-3-tosylamidobutan-2-one (Merck, Darmstadt) and with thermolysin (Serva, Heidelberg) at pH 8.1 and 37°C for 4 h and 1 h, respectively. Digestions with *Staphylococcus aureus* protease [3] were made at pH 4.0, for 48 h at 37°C. The intact protein was treated with carboxypeptidase A and a mixture of A and B (Boehringer, Mannheim) at pH 8.1, for 1 h at 37°C. Details of the enzymatic digestions are given elsewhere [4–6].

Cyanogen bromide cleavage [7] of the protein was performed in 70% formic acid (in the presence of 50 μg dithioerythritol/mg protein) at 37°C for 16 h with 5 mg reagent/mg protein. After lyophilisation the procedure was repeated.

The trypsin and thermolysin peptides were isolated by the fingerprint technique on thin-layer [5,6,8] Cel 300 and Cel 400 sheets (Macherey and Nagel, Düren). In addition, the tryptic peptides were chromatographed on a micro-column of Dowex M71 (0.3 × 10 cm) [9] at 55°C, developed with the following pyridine–formate gradients: 0.1 M, pH 2.7, to 1.0 M, pH 6.5 and 1.0 M, pH 6.5, to 2.0 M, pH 6.5, as described in ref. [5]. The SP-peptides, derived from digestion with *Staphylococcus aureus* protease, were separated on Cel 400 thin-layer sheets by fingerprinting. The cyanogen bromide peptides were isolated by gel filtration on a Sephadex G-25 (superfine) column (190 × 0.5 cm) in 5% acetic acid [5].

Amino acid analyses of the intact protein and of the peptides were performed on Durrum analyzers as described elsewhere [6,10]. Performic acid oxidation was carried out according to the method of Hirs [11]. For the detection of tryptophan, a tryptic fingerprint of L28 was developed with Ehrlich's reagent.

Sequence determinations of the intact protein were performed using liquid-phase Edman degradation [12] in a modified Beckman sequenator [13] equipped with an automatic conversion device [14]. The overlapping peptides were sequenced by the solid-phase method according to Laursen [15] with

 North-Holland Publishing Company – Amsterdam

attachment of the C-terminal carboxyl group to amino–polystyrene resin [16,17]. Identification of the amino acid derivatives released was performed both by thin-layer chromatography and by analysis of the free amino acids after hydrolysis, as previously described [18]. Small tryptic and thermolysin peptides were degraded manually by the dansyl-Edman procedure [19,20].

3. Results

Protein L28 is a rather basic component of the large subunit of *E. coli* ribosomes. It consists of 77 amino acid residues with the amino acid composition of Asp_2, Thr_5, Ser_5, Glu_4, Pro_2, Gly_5, Ala_5, Val_8, Met_1, Ile_2, Leu_6, Tyr_1, Phe_3, His_3, Lys_7, Arg_{11}, Cys_1, Trp_1, Gln_1, and Asn_4. The cysteine residue was determined as cysteic acid by amino acid analysis after oxidation of L28 and its appropriate peptides with performic acid. Tryptophan was detected in the appropriate peptides deriving from trypsin and thermolysin digestion of L28 on fingerprints which were sprayed with Ehrlich's reagent. The amino acid composition, derived from the sequence as shown in fig.1, is in good agreement with the results of amino acid analyses performed on the intact protein. According to this composition the molecular weight of L28 amounts to 8877.

Digestions with trypsin, thermolysin and *Staphylococcus aureus* protease were employed for the isolation and analyses of the peptides which are indicated in fig.1. The complete sequence determination was achieved:

(a) By automatic liquid-phase Edman degradation of protein L28 up to position 45.
(b) By solid phase sequencing of cyanogen bromide fragment CNBr II (56–77).
(c) By manual dansyl Edman degradations of the tryptic peptides T_1 (45–49), T10 (50–53), T9 (54–56), T17 (74–76) and of SP III (70–77), as given in fig.1.

In addition, partial N-terminal sequences of the other peptides were determined manually.

The alignment of the tryptic peptides in the N-terminal region is given by the sequenator results and this is supported by peptides, such as Th20 (6–20), Th14 (21–31), Th21 (32–36) and Th6 (37–45). The sequence T1–T10–T9–T28–T26 was established by peptides Th13 (48–54), Th17 (55–62), Th10 (63–65) and Th4 (66–69) and additionally by analysis of the long fragment SP II (41–69). Furthermore, CNBr I (1–55) was digested with *Staphyloccocus aureus* protease; from this hydrolysis SP IV (43–55) was obtained. CNBr II (56–77) served as an overlapping fragment for the C-terminal region of protein L28. The alignment T13–T17–T5 was covered by the sequence of Sp III (70–77) and the analysis of Th 15 (70–77). Carboxypeptidase A and B treatment of protein L28 liberated tyrosine and lysine from the C-terminus of the protein chain. More details of the elucidation of the primary structure of protein L28 will be presented elsewhere.

Protein L28 has 21 basic amino acids compared with only 6 acidic residues. The N-terminal part of the protein up to position 36 contains 11 of these basic residues and acidic amino acids are absent from this region. This sequence can be compared in character with the long basic area found at the C-terminus of the protein S13 [21].

Based on the amino acid sequence of protein L28, predictions for the possible secondary structure of this protein were made [22] using four different predictive methods (for details see ref. [23]). On the basis of three predictions in agreement the following structural features are expected:

(a) Three helical regions (for positions 19–23, 39–44 and 67–72).
(b) Turns or loops (for positions 7–8, 13–17, 30–31, 53–54 and 60–61.

Fig.1. Primary structure of ribosomal protein L28 from *E. coli*. Symbols: (LP-SEQ) liquid-phase Edman degradation of intact protein in an improved Beckman sequenator including automatic conversion reaction; ($\overline{+}$) identification of released PTH derivatives by thin-layer chromatography; ($\overline{(+)}$) identified weakly, ($\oplus$) additional analysis of liberated amino acid after hydrolysis (5.7 N HCl, 0.02% mercaptoethanol, 130°C) in the analyzer; (TRYP) trypsin digestion; (THER) thermolysin digestion; (SP) digestion with *Staphylococcus aureus* protease; (CNBr) cleavage by cyanogen bromide; (SP-SEQ) solid-phase Edman degradation, (→) identification of PTH derivative by thin-layer chromatography, (. . .) not identified, ((→)) identified weakly; (⟶) dansyl-Edman degradation.

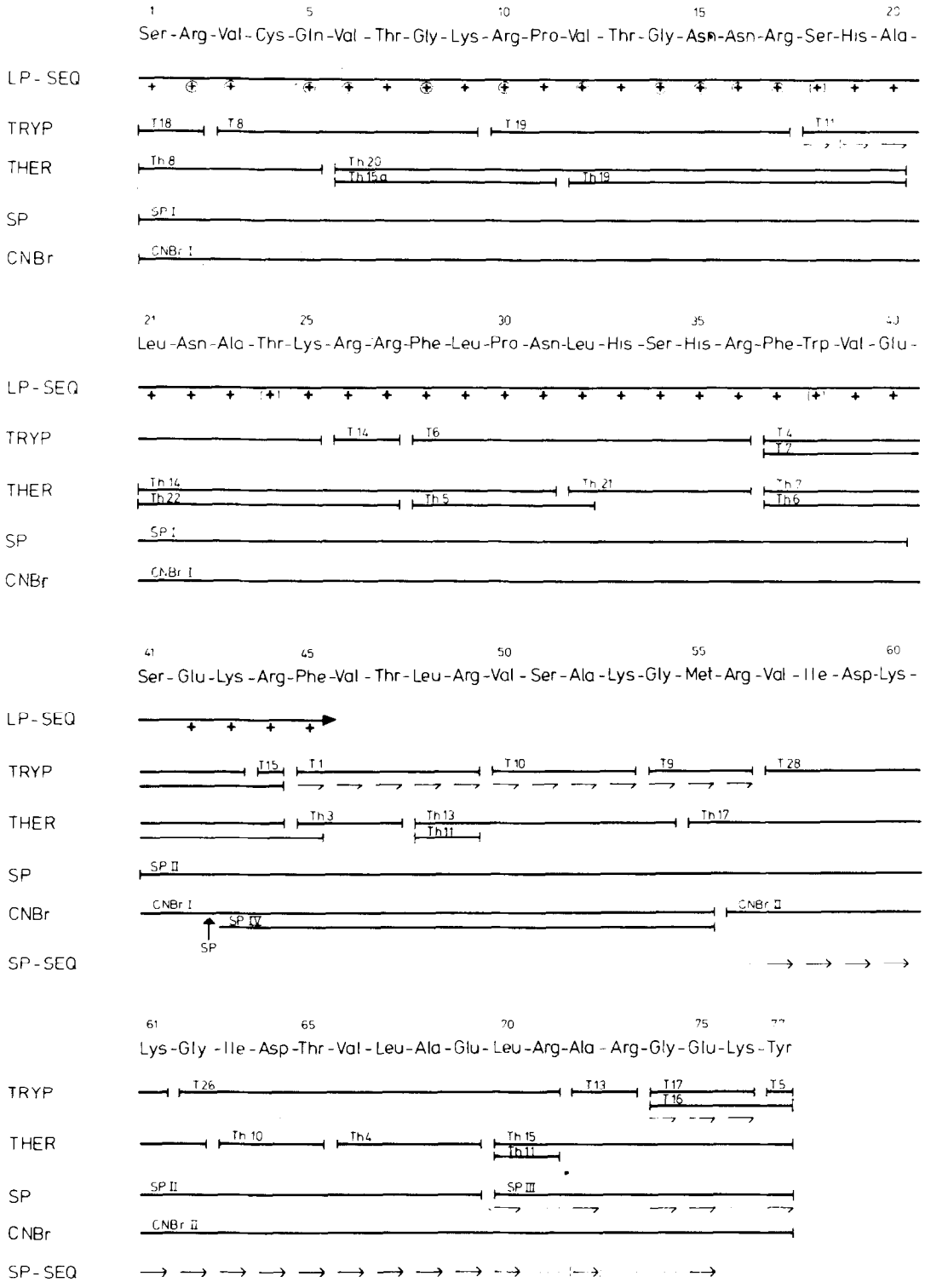
1 5 10 15 20
Ser-Arg-Val-Cys-Gln-Val-Thr-Gly-Lys-Arg-Pro-Val-Thr-Gly-Asn-Asn-Arg-Ser-His-Ala-
LP-SEQ
TRYP
T 18
T 8
T 19
T 11
THER
Th 8
Th 20
Th 15a
Th 19
SP
SP I
CNBr
CNBr I
21 25 30 35 40
Leu-Asn-Ala-Thr-Lys-Arg-Arg-Phe-Leu-Pro-Asn-Leu-His-Ser-His-Arg-Phe-Trp-Val-Glu-
LP-SEQ
TRYP
T 14
T6
T 4
T 7
THER
Th 14
Th 22
Th 5
Th 21
Th 7
Th 6
SP
SP I
CNBr
CNBr I
41 45 50 55 60
Ser-Glu-Lys-Arg-Phe-Val-Thr-Leu-Arg-Val-Ser-Ala-Lys-Gly-Met-Arg-Val-Ile-Asp-Lys-
LP-SEQ
TRYP
T 15
T 1
T 10
T9
T 28
THER
Th 3
Th 13
Th 11
Th 17
SP
SP II
CNBr
CNBr I
SP IV
CNBr II
SP
SP-SEQ
61 65 70 75 77
Lys-Gly-Ile-Asp-Thr-Val-Leu-Ala-Glu-Leu-Arg-Ala-Arg-Gly-Glu-Lys-Tyr
TRYP
T 26
T 13
T 17
T 16
T 5
THER
Th 10
Th4
Th 15
Th 11
SP
SP II
SP III
CNBr
CNBr II
SP-SEQ

Fig. 1

(c) Extended structure or β-sheet (for positions 2/4–6, 45/46–50/51 and 55/56–58). In comparison to the other ribosomal proteins investigated, the proportion of extended structure (14%) predicted for protein L28 is relatively high [22] (cf. S12, 13% [23] and L25, 10% [24]).

The comparison of the amino acid sequence of protein L28 with other ribosomal proteins reveals a very low degree of sequence similarity.

References

[1] Hindennach, I., Kaltschmidt, E. and Wittmann, H. G. (1971) Eur. J. Biochem. 23, 12–16.
[2] Kaltschmidt, E. and Wittmann, H. G. (1970) Analyt. Biochem. 36, 401–412.
[3] Houmard, J. and Drapeau, G. R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.
[4] Stadler, H. and Wittmann-Liebold, B. (1976) Eur. J. Biochem. 66, 49–56.
[5] Heiland, I., Brauer, D. and Wittmann-Liebold, B. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 1751–1770.
[6] Hitz, H., Schäfer, D. and Wittmann-Liebold, B. (1977) Eur. J. Biochem. 75, 497–512.
[7] Gros, E. and Witkop, B. (1961) J. Am. Chem. Soc. 83, 1510–1511.
[8] Bitar, K. G. and Wittmann-Liebold, B. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1343–1352.
[9] Chen, R. and Wittmann-Liebold, B. (1975) FEBS Lett. 52, 139–140.
[10] Wittmann-Liebold, B., Greuer, B. and Pannenbecker, R. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1977–1979.
[11] Hirs, C. H. W. (1956) J. Biol. Chem. 219, 611–621.
[12] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[13] Wittmann-Liebold, B. (1973) Hoppe-Seyler's Z. Physiol. Chem. 354, 1415–1431.
[14] Wittmann-Liebold, B., Graffunder, H. and Kohls, H. (1976) Analyt. Biochem. 75, 621–633.
[15] Laursen, R. A. (1971) Eur. J. Biochem. 20, 89–102.
[16] Previero, A., Derancourt, J., Coletti-Previero, M. A. and Laursen, R. A. (1973) FEBS Lett. 33, 135–138.
[17] 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.
[18] Wittmann-Liebold, B., Geissler, A. W. and Marzinzig, E. (1975) J. Supramol. Struct. 3, 426–447.
[19] Gray, W. R. and Hartley, B. S. (1963) Biochem. J. 89, 379–380.
[20] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.
[21] Lindemann, H. and Wittmann-Liebold, B. (1977) Hoppe-Seyler's Z. Physiol. Chem. 358, 843–863.
[22] Dzionara, M., Robinson, S. M. L. and Wittmann-Liebold, B., J. Supramol. Struct. (1977) to be submitted.
[23] Dzionara, M., Robinson, S. M. L. and Wittmann-Liebold, B. (1977) Hoppe-Seyler's Z. Physiol. Chem. 358, 1003–1019.
[24] Wittmann-Liebold, B., Robinson, S. M. L. and Dzionara, M. (1977) FEBS Lett. 81, 204–213.