THE PRIMARY STRUCTURE OF RIBOSOMAL PROTEIN L3 FROM *ESCHERICHIA COLI* 70 S RIBOSOMES

T. A. MURANOVA, A. V. MURANOV, L. F. MARKOVA and Yu. A. OVCHINNIKOV
Institute of Protein Research, USSR Academy of Sciences, 142292 Poustchino, Moscow Region, USSR

Received 27 September 1978

1. Introduction

In our laboratory the primary structures of some ribosomal proteins have been studied for several years, in particular, amino acid sequences of proteins L25, L10, L32 have been determined [1–3]. In extending these investigations we have undertaken the study of one of the large proteins of 70 S ribosomes, protein L3. This is one of the RNA-binding proteins of the 50 S subparticle [4] that has such a strong affinity to RNA that it remains bound to 23 S RNA even after the treatment of 50 S subparticles with 6 M LiCl [5]. It has also been established that in the ribosome protein L3 interacts specifically and independently of other proteins with the 3′-terminal 11 S fragment of 23 S RNA [6,7]. It is also important to note that protein L3 is one of several proteins which are found associated with the elongation factor Tu on treatment with bifunctional reagents [8]. Experiments of Hampl and Nierhaus suggest that protein L3 participates in the formation of the peptidyltransferase center of 70 S ribosomes [9]. Chang and Chang have recently found in protein L3 a methylated amino acid [10] which has been identified as N^5-methylglutamine by Lhoest and Colson [11].

The present paper concerns the determination of the complete amino acid sequence of the ribosomal protein L3, including identification and localization of the N^5-methylglutamine residue. Earlier we have reported data on peptides of tryptic, chymotryptic and thermolytic hydrolyzates of protein L3 [12,13].

2. Materials and methods

The protein was isolated from *Escherichia coli* MRE-600 by the method in [14] with a yield of 50 mg from 10 g total ribosomal protein. Two-dimensional electrophoresis in polyacrylamide gel [15] was used for identification and determination of the protein purity. The protein was hydrolyzed by trypsin (Worthington Biochem. Corp., USA) at 37°C, pH 8.0 for 70 h; by chymotrypsin (Worthington Biochem. Corp., USA) at 37°C, pH 8.0 for 24 h; by thermolysin (Serva, FRG) at 37°C, pH 8.0 for 3.5 h; by *Staphylococcus aureus* V8 protease (Miles Labs., England) in 0.1 M CH_3COONH_4 at 37°C, pH 8.0 for 35 h. Limited hydrolysis with trypsin at arginine residues was carried out for 24 h after modification of the lysine residues in the protein with maleic anhydride [16]. The protein modified in this way was also subjected to chymotryptic hydrolysis at 37°C, pH 8.0 for 3 h. Thermolysin hydrolysis of the fragments of the maleyl protein was done at 37°C, pH 8.0 for 2 h. Limited trypsin hydrolysis at lysine residues was carried out for 12 h after modification of arginine residues in the protein with malonic dialdehyde [17].

For the preparative separation of peptides we used either ion-exchange chromatography on the ion-exchange resin AG 50 × 4, -400 mesh (Bio-Rad, USA) (0.9 × 60 cm) in a system of linear gradients of concentration and pH of pyridine–acetate buffer, or gel filtration on Sephadex G-50 Superfine (Pharmacia, Sweden) (1.5 × 200 cm) accompanied by a high-voltage electrophoresis on Whatman 3 MM paper (Whatman, England).

The amino acid analysis of the protein and peptides was performed on a D-500 amino acid analyzer (Durrum, USA). To determine the tryptophan content the protein was hydrolyzed with *p*-toluenesulfonic acid [18] before amino acid analysis; the presence of

tryptophan residues in the peptides was determined with the help of Ehrlich's reagent [19]. The amino acid sequences were determined by modified Edman method [20], the N-terminal amino acid sequence of the fragment 84–124 was determined by automatic Edman degradation on a model 890C Beckman sequencer (USA) using dimethylallylamine program 102974. The C-terminal amino acid residues of the protein and peptides were studied with the help of carboxypeptidases A and B [21].

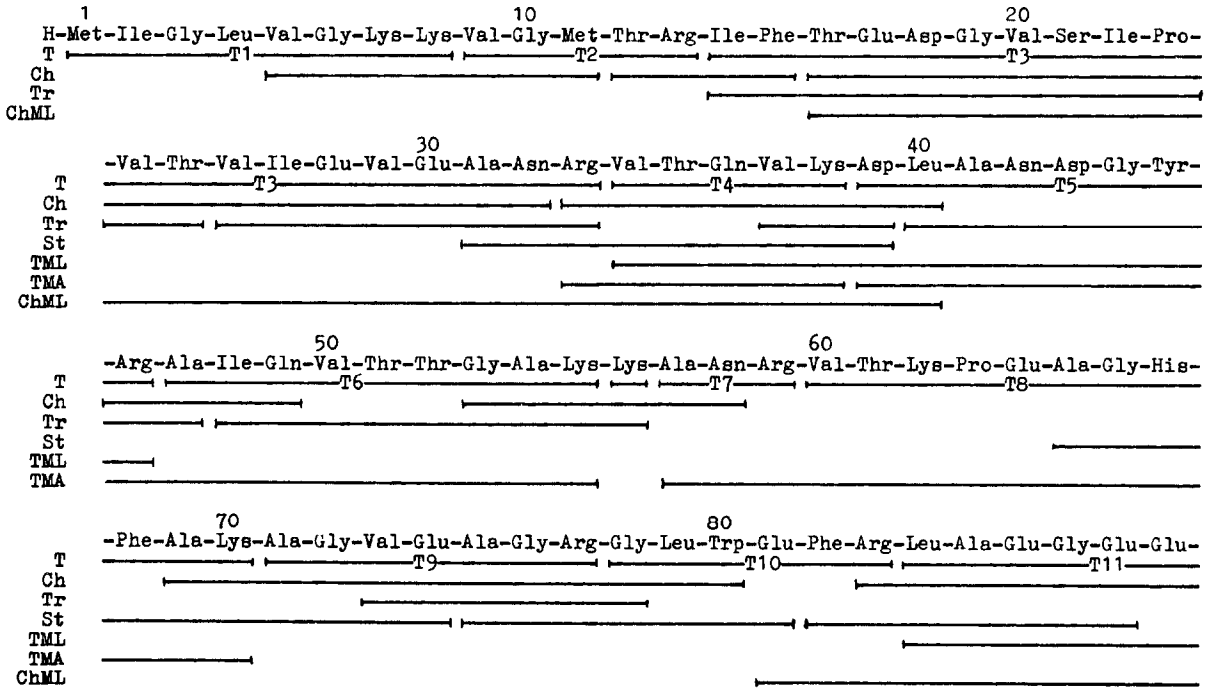
To identify the N^5-methylglutamine residue (MGln) we used ethyl acetate extracts obtained at the corresponding stages of the determination of the amino acid sequence of the peptides displaying a methylamine peak on amino acid analysis. Acid hydrolyzates of amino acid thiazolinones contained in these extracts were studied by an amino acid analysis; when methylamine was present, thiazolinones were converted into the corresponding PTH-derivatives and identified using as a standard the PTH-derivative of N^5-methylglutamine synthesized by the method in [22].

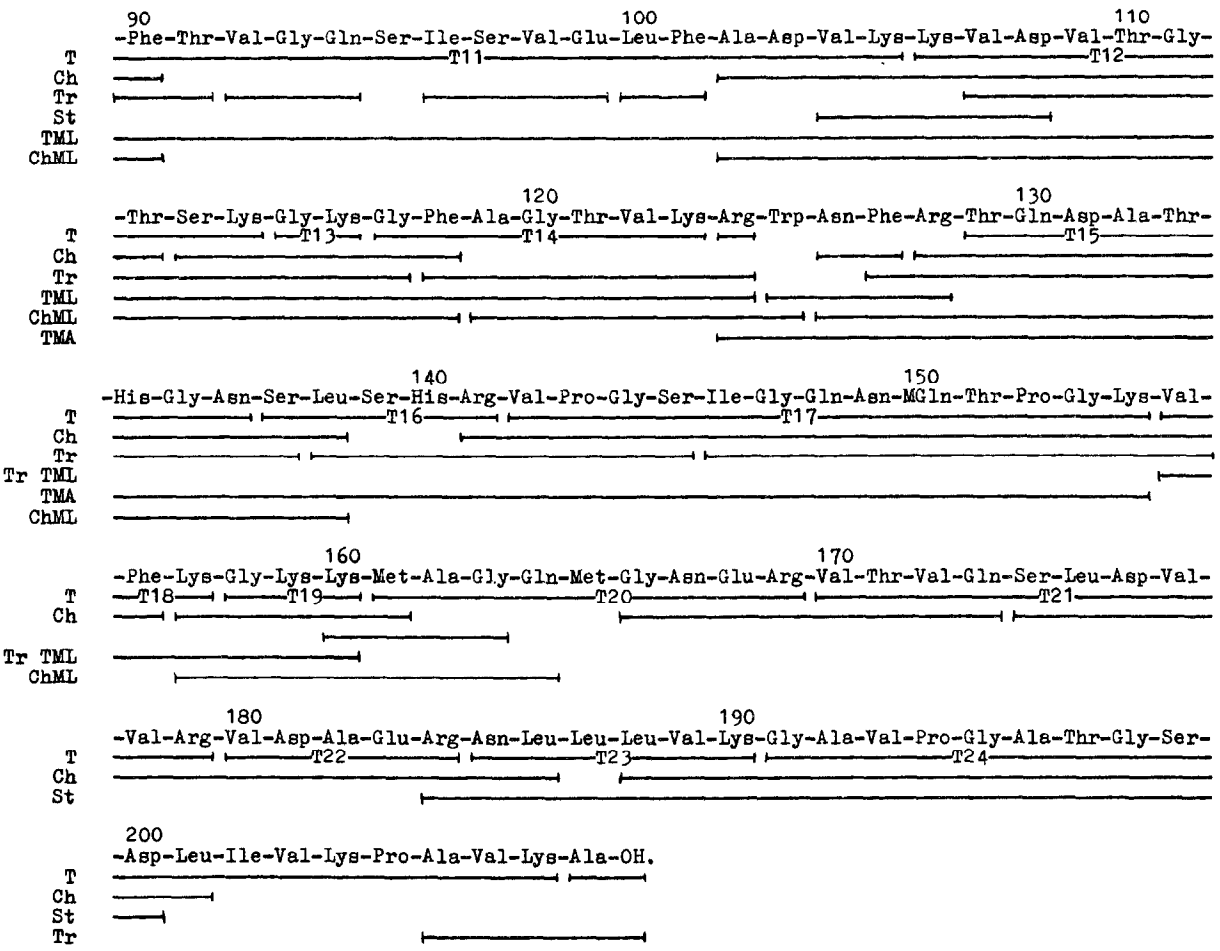
3. Results and discussion

Protein L3 was isolated by the method in [14]. It has been noted that on ion-exchange chromatography of the total 70 S ribosomal protein under the conditions described protein L3 always elutes from the column in a mixture with proteins L11 and S2. Proteins L3 and L11 have the same isoelectric point at 9.7 [23] and the presence of protein S2 (isoelectric point at 6.7) [23] in the fractions containing proteins L3 and L11 possibly indicates that either protein L3 or L11 or both form with protein S2 a complex stable under separation conditions.

The complete amino acid sequence of the ribosomal protein L3 is represented in Scheme 1. The polypeptide chain of the protein consists of 209 amino acid residues and has the following amino acid composition: Asp 9, Asn 8, Thr 16, Ser 9, Glu 12, Gln 7, Pro 6, Gly 26, Ala 19, Val 28, Met 4, Ile 8, Leu 11, Tyr 1, Phe 8, His 3, Lys 19, Trp 2, Arg 12, MGln 1. This is in agreement with the data obtained by amino acid analysis of the protein [12]. The

Scheme 1





The complete amino acid sequence of *E. coli* ribosomal protein L3. Non-overlapping peptides of the tryptic hydrolyzate are indicated. For the other types of hydrolysis only peptides giving overlaps between tryptic fragments or providing additional information on the structure are indicated. MGln, N^5-methylglutamine; T, tryptic peptides; Ch, chymotryptic peptides; Tr, thermolytic peptides; St, fragments resulting from the protein digestion by *Staphylococcus aureus* protease; TML, tryptic peptides of protein L3 modified at lysine residues; TMA, tryptic peptides of protein L3 modified at arginine residues; ChML, chymotryptic peptides of protein L3 modified at lysine residues; TrTML, thermolytic peptide obtained on digestion of one of the fragments produced on limited trypsin hydrolysis of the maleyl protein.

molecular weight calculated from the primary structure is 22 232 corresponding to the value determined by high-speed equilibrium sedimentation [24].

Several types of enzyme hydrolysis of the polypeptide chain were used for investigation of the primary structure. First the protein was subjected to trypsin and then to chymotrypsin and thermolysin hydrolyses. The products of hydrolysis were separated by ion-exchange chromatography and high-voltage paper electrophoresis. As a result, 36 peptides, as well as free lysine, arginine and alanine, were isolated from the tryptic hydrolvzate, 47 peptides from the chymotryptic hydrolyzate and 71 peptides from the thermolytic one [12,13]. The data obtained from this study permitted determination of the amino acid sequences of several large fragments [13]. As there was an arginine residue at the C-terminus of 4 out of the 7 fragments obtained, it was useful to carry out a

limited tryptic hydrolysis of the polypeptide chain at lysine residues with a preliminary modification of the arginine residues by malonic dialdehyde. After separation by gel filtration on Sephadex G-50 the hydrolyzate was divided into several fractions. One of these fractions contained a mixture of fragments 9–38 and 124–154 which we did not succeed in separating, therefore the amino acid sequence of the N-terminal part of the fragment 124–154 was determined using this mixture since the N-terminal amino acid sequence of the fragment 9–38 was known [13]. In this case we obtained the overlap between tryptic peptides T14 and T15 including the tryptic peptide Trp–Asn–Phe–Arg (125–128) which had not been isolated earlier. The other polypeptide chain fragments obtained after limited tryptic hydrolysis at lysine residues were isolated in a pure form by high-voltage paper electrophoresis of the corresponding fractions. The most interesting among them was peptide 57–70 permitting us to join tryptic peptides T7 and T8.

The next stage was the modification of protein L3 with maleic anhydride at lysine residues and limited tryptic hydrolysis at arginine residues. After gel filtration of the hydrolyzate on Sephadex G-50 we isolated the peptide Trp–Asn–Phe–Arg and the fragment 84–124 which is a convenient object for automated Edman degradation on a sequencer. As a result, its N-terminal amino acid sequence 84–104 was determined (scheme 1). One of the fractions which was obtained on gel filtration of the tryptic hydrolyzate of the maleyl protein and which contained a mixture of fragments 142–169 and 185–209 was subjected to thermolysin hydrolysis without separation. High-voltage paper electrophoresis of the hydrolyzate resulted in isolation of the peptide Val–Phe–Lys–Gly–Lys–Lys (155–160) which permitted us, together with chymotryptic peptides 157–161 and 160–163, to join the tryptic fragments T18 and T20.

Chymotrypsin digestion of the protein proceeded in suspension owing to the low solubility at pH 8.0. Therefore the time of hydrolysis had to be increased which in some cases led to undesirable cleavage of bonds such as Asn–Arg (32–33, 58–59) and Lys–Arg (123–124). To isolate the necessary peptides we decided to repeat the chymotrypsin digestion of the protein after its modification with maleic anhydride. As the modified protein had a good solubility at the basic pH values it was possible to decrease the time of hydrolysis to 3 h which contributed to the obtaining of larger fragments. As a result, a number of interesting peptides were isolated from the hydrolyzate, among them peptide 81–90 connecting tryptic peptides T10 and T11 and fragments 119–125 and 126–138 corroborating the sequence 124–128 determined using the mixture of tryptic fragments 9–38 and 124–154 of the protein L3 modified at arginine residues.

The peptides isolated from the protein hydrolyzate on digestion with *Staphylococcus aureus* protease gave the missing overlap between tryptic peptides T3 and T4 (fragment 31–39) and corroborated some sequences.

On determination of the amino acid composition of acid hydrolyzates of some peptides, methylamine was found. This suggested that N^5-methylglutamine is included in their composition and after a thorough study it was identified and localized in position 150 (scheme 1).

The polypeptide chain of protein L3 has some interesting features. First among these is the presence of the methylated amino acid residue. It should be also noted that there are identical or similar amino acid sequences: Ala–Asn–Arg–Val–Thr (31–35, 57–61), Trp–Glu–Phe–Arg (80–83) and Trp–Asn–Phe–Arg (125–128), Lys–Gly–Lys (114–116, 157–159), Val–Thr–Val (24–26, 170–172), Arg–Val–Thr (33–35, 59–61, 169–171), Val–Pro–Gly (142–144, 193–195), Gly–Lys–Lys (6–8, 158–160), Leu–Leu–Val–Lys (187–190) and Leu–Ile–Val–Lys (201–204), sequences Val–Thr and Val–Lys are encountered 6 times. There are symmetrical regions of the polypeptide chain Ala–Lys–Lys–Ala (54–57), Asp–Val–Lys–Lys–Val–Asp (103–108), Lys–Gly–Lys (114–116, 157–159), Val–Thr–Val (24–26, 170–172).

Finally, it should be noted that the sequence 1–32 agrees well with the N-terminal amino acid sequence of protein L3 determined by Wittmann-Liebold et al. [25] using a sequencer.

Acknowledgements

The authors would like to express their gratitude to L. A. Krivosheeva for the help in isolation of the

ribosomal protein L3 and to L. M. Vinokurov for the determination of the N-terminal amino acid sequence of the fragment 84–124 in a sequencer.

References

[1] Dovgas, N. V., Markova, L. F., Mednikova, T. A., Vinokurov, L. M., Alakhov, Yu. B. and Ovchinnikov, Yu. A. (1975) FEBS Lett. 53, 351–354.
[2] Dovgas, N. V., Vinokurov, L. M., Velmoga, I. S., Alakhov, Yu. B. and Ovchinnikov, Yu. A. (1976) FEBS Lett. 67, 58–61.
[3] Vinokurov, L. M., Alakhov, Yu. B., Golov, E. A. and Ovchinnikov, Yu. A. (1976) Bioorgan. Khim. 2, 1013–1017.
[4] Garrett, R. A., Müller, S., Spierer, P. and Zimmermann, R. A. (1974) J. Mol. Biol. 68, 553–557.
[5] Spierer, P., Zimmermann, R. A. and Mackie, G. A. (1975) Eur. J. Biochem. 52, 459–468.
[6] Chen-Schmeisser, U. and Garrett, R. (1976) Eur. J. Biochem. 69, 401–410.
[7] Giri, L., Tam, M. F. and Hill, W. E. (1976) Biochemistry 15, 5188–5192.
[8] Fabian, U. (1976) FEBS Lett. 71, 256–260.
[9] Stöffler, G. and Wittmann, H. G. (1977) in: Molecular Mechanisms of Protein Biosynthesis (Weissbach, H. and Pestka, S. eds) pp. 181–182, Academic Press, London, New York, San Francisco.
[10] Chang, C. N. and Chang, F. N. (1975) Biochemistry 14, 468–477.
[11] Lhoest, J. and Colson, Ch. (1977) Molec. Gen. Genet. 154, 175–180.
[12] Muranov, A. V., Muranova, T. A. and Markova, L. F. (1978) Bioorgan. Khim. 4, 293–301.
[13] Muranova, T. A., Muranov, A. V. and Markova, L. F. (1978) Bioorgan. Khim. 4, 1197–1202.
[14] Alakhov, Yu. B., Mednikova, T. A., Motuz, L. P., Markova, L. F., Dovgas, N. V., Kashparov, I. A. and Ovchinnikov, Yu. A. (1975) Bioorgan. Khim. 1, 581–587.
[15] Kaltschmidt, E. and Wittmann, H. G. (1970) Anal. Biochem. 36, 401–412.
[16] Butler, P. J. G. and Hartley, B. S. (1972) Methods Enzymol. 25, 191.
[17] Muranov, A. V. and Modyanov, N. N. (1978) Bioorgan. Khim. 5/2, in press.
[18] Liu, T. Y. and Chang, Y. H. (1971) J. Biol. Chem. 246, 2842–2848.
[19] Scoffone, E. and Fontana, A. (1970) in: Protein Structure Determination, Molecular Biology, Biochemistry and Biophysics (Needleman, S. B. ed) vol. 8, p. 208, Springer-Verlag, Berlin, Heidelberg, New York.
[20] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.
[21] Ambler, R. P. (1972) Methods Enzymol. 25, 143–154; 262–272.
[22] Lichtenstein, N. (1942) J. Am. Chem. Soc. 64, 1021–1022.
[23] Kaltschmidt, E. (1971) Anal. Biochem. 43, 25–31.
[24] Dzionara, M., Kaltschmidt, E. and Wittmann, H. G. (1970) Proc. Natl. Acad. Sci. USA, 67, 1909–1913.
[25] Wittmann-Liebold, B., Geissler, A. W. and Marzinzig, E. (1975) J. Supramol. Struct. 3, 426–447.