THE PRIMARY STRUCTURE OF THE 5 S RNA BINDING PROTEIN L25 FROM *ESCHERICHIA COLI* RIBOSOMES

N. V. DOVGAS, L. F. MARKOVA, T. A. MEDNIKOVA, L. M. VINOKUROV, Yu. B. ALAKHOV and Yu. A. OVCHINNIKOV

Institute of Protein Research, Academy of Sciences of the USSR, Poustchino, Moscow Region, USSR

Received 19 March 1975

1. Introduction

We have recently suggested a new preparative procedure for isolation of specific proteins from 70 S ribosomes of *Escherichia coli* [1]. In developing these investigations we have undertaken a study of the primary structure of protein L25 included in the 5 S RNA–protein complex [2].

2. Materials and methods

Protein L25 was isolated from *E. coli* MRE-600 70 S ribosomes as described earlier [1] with a yield of about 80 mg from 10 g of the total 70 S ribosome protein. The peptides derived by cleavage of the protein molecule with trypsin, with cyanogen bromide, of the modified protein with trypsin only at the arginine residues [3], as well as by cleavage of cyanogen bromide fragments with bromosuccinimide (at the tyrosine residues [4]) were separated by gel-filtration on Sephadex G-50, chromatography on Aminex-A-5 resin, paper chromatography and paper high-voltage electrophoresis. The amino acid composition was determined on the amino acid analyzer BC-201 (Bio-Cal, BRD). The amino acid sequence of the isolated peptides was determined by Edman's phenyl-isothiocyanate method, the dansyl Edman method [5] and by automated Edman degradation on a sequenser (model 890C, Beckman, USA) using fast and slow peptide programs. The amino acid sequence of peptides T-10 and BS-2 was determined by automated degradation with a modification of lysine residues according to Braunitzer [6]. PTH-derivatives of amino acids were identified by gas-liquid and thin-layer chromatography, the DNS-derivatives by two-dimensional thin-layer chromatography on silica gel [7,8]. The C-terminal amino acid residues of protein and peptides were determined by carboxypeptidases A and B [9].

3. Results and discussion

Protein L25 was reported to have a mol. wt of 11 500–12 000 according to the data of SDS-polyacrylamide gel electrophoresis [10]. It does not contain cystein residues and its amino acid composition was determined as follows (in moles/%):
Asp – 8.6, Thr – 2.4, Ser – 2.8, Glu – 11.7, Pro – 4.7, Gly – 6.7, Ala – 11.6, Val – 8.5, Met – 3.0, Leu – 6.4, Ile – 4.9, Tyr – 3.5, Phe – 4.0, His – 3.0, Lys – 11.6, Arg – 6.7 [11].

We have determined that the N-terminal amino acid residue is methionine and that the C-terminal sequence is –Phe–Val–Arg–Ala.

Analysis of the protein L25 tryptic hydrolyzate by peptide mapping revealed twenty two spots. After preliminary separation of tryptic hydrolyzate products on Sephadex G-50 with a following fractionation of the obtained fractions on Aminex-A-5 ion-exchange resin, and also by paper chromatography and high-voltage electrophoresis on paper, 17 peptides were isolated as well as free alanine, arginine and lysine (see table 1).

To obtain overlapping peptides and a reconstruction of the polypeptide chain, the N-terminal protein sequence was determined by automated Edman

Table 1
Tryptic peptides (T-)

Peptide	Amino acid sequence
T-1	Met–Phe–Thr–Ile–Asn–Ala–Glu–Val–Arg
T-2	Lys–Glu–Gln–Gly–Lys
T-3	Glu–Gln–Gly–Lys
T-4	Gly–Ala–Ser–Arg
T-5	Gly–Ala–Ser–Arg–Arg
T-6	Arg–Leu–Arg
T-7	Leu–Arg
T-8	Ala–Ala–Asn–Lys–Phe–Pro–(Ala,Gly_2,Ile_2,Lys,Tyr)
T-9	Glu–Ala–Pro–Leu–Ala–Ile–Glu–Leu–Asp–His–Asp–(Ala,Asn,Gln,Lys_2,Met_2,Val)
T-10	Ala–Glu–Phe–Tyr–Ser–Glu–Val–Leu–Thr–Ile–Val–(Asp,Gly,Lys,Val)
T-11	Glu–Ile–Lys
T-12	Val–Lys
T-13	Ala–Gln–Asp–Val–Gln–Arg–His–Pro–Tyr–(Lys_2,Pro)
T-14	Ala–Gln–Asp–Val–Gln–Arg
T-15	Glu–Ile–Lys–Val–Lys
T-16	Leu–Gln–His–Ile–Asp–Phe–Val–Arg–Ala
T-17	Leu–Gln–His–Ile–Asp–Phe–Val–Arg
	Ala
	Arg
	Lys

degradation as well as cleavage of the protein with cyanogen bromide and a limited tryptic digest of the protein modified at the lysine residues by citraconic anhydride.

After cleavage of the protein by cyanogen bromide, the peptide mixture was separated by gel filtration on Sephadex G-50 with the isolation of three fragments, CB-1, CB-2 and CB-3 as well as homoserine, a product of the methionine N-terminal residue transformation. Fragment CB-2 represents an Asn-Hse dipeptide contained in the tryptic peptide T-9. Determination of the C-terminal sequence with carboxypeptidases A and B showed that fragment CB-3 has a structure identical to the C-terminal sequence of the protein (i.e. is C-terminal), while fragment CB-1 has a C-terminal sequence –Lys–Val–Hse. Determination of the N-terminal sequence of fragment CB-1 showed that it coincides with the N-terminal sequence of protein. It follows from the above that fragment CB-2 is between fragments CB-1 and CB-3. This is also corroborated by the structure of peptide T-9.

Determination of the amino acid sequence by automated Edman degradation gave the following results. The N-terminal sequence up to the amino acid residue in position 35 was determined on a whole protein molecule. The sequence 2-35 was corroborated by the automated analysis of the CB-1 fragment sequence. The N-terminal sequence of thirty two amino acid residues was determined on the CB-3 fragment. Thus, the obtained data permitted to join the following tryptic peptides into one polypeptide chain: T-1, T-2, T-3, T-4, T-5, T-6, T-7, T-8, and also T-9, T-10, T-11, T-12, T-13. Overlapping of fragments T-8, T-9, T-10 was done by determinating peptide structures obtained at a limited tryptic digest of the protein molecule as well as by determination of peptide structures formed with the cleavage of cyanogen bromide fragments at the tyrosine residues.

The protein molecule contains six arginine residues, four of them are in the N-terminal fragment of CB-1 and two in the C-terminal fragment of CB-3, the position of these residues being determined by automated sequence analysis of the corresponding peptide fragments.

Limited tryptic digest of the protein molecule at

Table 2
Limited tryptic peptides (LT-)

Peptide	Amino acid sequence
LT-1	Met–Phe–Thr–Ile–Asn–Ala–Glu–Val–Arg
LT-2	Lys–Glu–Gln–Gly–Lys–Gly–Ala–Ser–Arg
LT-3	Lys–Glu–Gln–Gly–Lys–Gly–Ala–Ser–Arg–Arg
LT-4	Ala–Ala–Asn–Lys–Phe–Pro–Ala–Ile–Ile–Tyr–Gly–Gly– –Lys–Glu–Ala–Pro–Leu–Ala–Ile–Glu–Leu–Asp–His–Asp– –Lys–Val–Met–Asn–Met–Gln–Ala–Lys–Ala–Glu–Phe–Tyr– –(Ala_1,Arg_1,Asx_2,Glx_4,Gly_1,Ile_2,Leu_1,Lys_3,Ser_1, Thr_1,Val_5)
LT-5	His–Pro–Tyr–Lys–Pro–Lys–Leu–Gln–His–Ile–Asp–Phe– –Val–Arg–Ala

the arginine residue resulted in the isolation of 5 peptides (see table 2), including peptide LT-4 (22–79) where the sequence corresponding to the sequence 22–57 in the protein molecule was determined by automated Edman degradation.

Thus the sequence of amino acid residues 1–82 was determined by the automated analysis method. The sequence 83–94 in peptides LT-5 and T-16 was determined by the Edman dansyl method and also by determination of the four C-terminal amino acids in the protein molecule, the cyanogen bromide fragment CB-3 and the peptide T-16. The obtained complete amino acid sequence of protein L25 is given in scheme 1.

The exact mol. wt of protein L25 is 10 912; the polypeptide chain consists of 94 amino acid residues and has the following amino acid composition: Asp_6, Asn_3, Thr_2, Ser_2, Glu_6, Gln_5, Pro_4, Gly_5, Ala_{11}, Val_8, Met_3, Ile_7, Leu_5, Tyr_3, Phe_4, His_3, Lys_{11}, Arg_6. The increased isoleucine residue content and the decreased content of leucine residues in comparsion with the literature data [11] can be explained by the difference in the strains of *E. coli* used to prepare the protein.

The amino acid sequence determined on the sequenser was corroborated by determination of structures of peptides isolated from tryptic and limited tryptic digests, and, moreover, by determination of structures of peptides formed by cleavage of cyanogen bromide fragments CB-1 and CB-3 at tyrosine residues 31, 57, 82 (see table 3). The structure of peptides BS-2 (32–48) and T-10 (54–68) was determined on the sequenser after their modification with the Braunitzer reagent III. At tryptic digestion the cleavage of the following bonds proved to be incomplete: –Lys–Phe–(25–26), –Lys–Val–(46–47), –Lys–Val–(71–72), –Arg–His–(79–80), –Lys–Pro–(83–84), which permitted in a number of cases to obtain an additional overlapping of tryptic fragments.

It can be expected that the determination of the primary structure of protein L25 will provide further

Scheme 1
Complete Amino Acid Sequence of Protein L25

1 10 20
Met–Phe–Thr–Ile–Asn–Ala–Glu–Val–Arg–Lys–Glu–Gln–Gly–Lys–Gly–Ala–Ser–Arg–Leu–Arg–Ala–Ala–Asn–Lys–Phe–
30 40 50
Pro–Ala–Ile–Ile–Tyr–Gly–Gly–Lys–Glu–Ala–Pro–Leu–Ala–Ile–Glu–Leu–Asp–His–Asp–Lys–Val–Met–Asn–Met–
60 70
Gln–Ala–Lys–Ala–Glu–Phe–Tyr–Ser–Glu–Val–Leu–Thr–Ile–Val–Val–Asp–Gly–Lys–Glu–Ile–Lys–Val–Lys–Ala–
80 90 94
Gln–Asp–Val–Gln–Arg–His–Pro–Tyr–Lys–Pro–Lys–Leu–Gln–His–Ile–Asp–Phe–Val–Arg–Ala

Table 3
Bromosuccinimide peptides (BS-)

Peptide	Amino acid sequence
BS-1	Phe–Thr–Ile–Asn–Ala–Glu–Val–Arg–Lys–Glu–Gln–Gly– –Lys–Gly–Ala–Ser–Arg–Arg–Leu–Arg–Ala–Ala–Asn–Lys– –Phe–Pro–Ala–Ile–Ile–Tyr
BS-2	Gly–Gly–Lys–Glu–Ala–Pro–Leu–Ala–Ile–Glu–Leu–Asp– –His–Asp–(Lys,Met,Val)
BS-4	Ser–Glu–Val–Leu–Thr–Ile–Val–(Ala,Arg,Asp_2,Gln_2, Glu,Gly,His,Ile,Lys_3,Pro,Tyr,Val_3)
BS-5	Lys–Pro–Lys–Leu–Gln–His–Ile–Asp–Phe–Val–Arg–Ala

insight on protein–RNA interaction since protein L25 forms a stable functionally active complex with 5 S RNA.

References

[1] Alakhov, Yu. B., Mednikova, T. A., Motuz, L. P., Markova, L. F., Dovgas, N. V., Kashparov, I. A. and Ovchinnikov, Yu. A. (1975) Bioorganischeskaya Khimiya 1, in press.

[2] Horne, J. R. and Erdmann, V. A. (1972) Molec. Gen. Genet. 119, 337–344.

[3] Kolb, E., Harris, J. J. and Bridgen, J. (1974) Biochem. J. 137, 185–197.

[4] Ramachandran, L. K. and Witkop, B. (1967) in: Methods in Enzymology, XI, pp.283–299, Acad. Press, N.Y., London.

[5] Gray, W. R. (1967) in: Methods in Enzymology, XI, pp. 469–475, Acad. Press, N.Y., London.

[6] Braunitzer, G., Schrank, B. and Ruhfus, A. (1971) Hoppe-Seyler's Z. Physiol. Chem. 352, 1730–1733.

[7] Belenki, B. G., Gankina, E. S., Pryanishnikova, S. R. and Erastov, D.P. (1967) Molekul. Biol. USSR 1, 184–189.

[8] Belenki, B. G., Gankina, E. S. and Nesterov, V. V. (1967) Dokl. Akad. Nauk SSSR, 172, 91–93.

[9] Ambler, R. (1967) in: Methods in Enzymology, XI, pp. 155–166, Acad. Press, N.Y., London.

[10] Dzionara, M., Kaltschmidt, E. and Wittman, H. G. (1970) Proc. Nat. Acad. Sci. US. 67, 1909–1913.

[11] Kaltschmidt, E., Dzionara, M. and Wittmann, H. G. (1970) Molec. Gen. Genet. 109, 292–297.