PRIMARY STRUCTURE OF RIBOSOMAL PROTEIN S6 FROM THE WILD TYPE AND A MUTANT OF *ESCHERICHIA COLI*

Hans HITZ, Doris SCHÄFER and Brigitte WITTMANN-LIEBOLD
Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem, West Germany

Received 16 June 1975

1. Introduction

S6 is the most acidic protein of the *E. coli* 30S ribosomal subunit [1]. Interacting with the proteins S18, S8 and S15 it protects part of the 16S RNA from nuclease attack [2]. Functionally it seems to be involved in initiation [3].

A simple procedure for the isolation of this protein has been developed resulting in the separation of five forms differing in length and charge [4]. Determination of their primary structures and the comparison to the amino acid sequence of protein S6 from an *E. coli* mutant are reported in this paper.

2. Materials and methods

Ribosomes were isolated from *E. coli*, strain K, as previously described [5]. After extraction of ribosomal proteins by acetic acid protein S6 was separated from the other proteins by a two step elution procedure followed by a gradient elution on DEAE-cellulose. This resulted in several forms of S6 as described below and elsewhere in detail [4]. Amino acid analyses were done on a Durrum D-500 analyzer and peptide mapping on cellulose thin-layer plates. Peptides obtained after cleavage with trypsin, chymotrypsin, thermolysin and staphylococcus protease [6] digestion were separated on sulfoethyl-cellulose followed by paper chromatography [7]. The two peptides obtained by specific cleavage of the protein at its single tryptophanyl bond using BNPS-skatol (2 - (2 -nitrophenyl-sulfenyl) -3 - methyl - 3' - bromoindolenine [8] were separated by gel filtration on Sephadex G100.

Amino acid sequence studies on the intact protein and a BNPS-fragment were done by automatic Edman-degradation [9] using an improved liquid phase Beckman sequenator [10]. Sequence determination of peptides, especially of the C-terminal tryptic peptides from the different forms of S6, were performed in a solid-phase sequenator [11–13]. Furthermore, some tryptic peptides were sequenced by the Dansyl-Edman-technique [14]. PTH-amino acids were identified by thin-layer chromatography and mass spectrometry [15,16]. Further details of the methods will be described elsewhere [4].

3. Results and discussion

3.1. *Primary structure of the wild type protein*

Recently Held et al. [17] separated protein S6 on DEAE-cellulose into two forms. However under our conditions for isolation of the protein [4] wild type protein S6 could be isolated in five different forms: S6-2, S6-3, S6-4, S6-5, S6-6. It remains to be seen how many of these forms are present in vivo. Comparison of their amino acid composition showed no difference among them except for the number of glutamic acid residues: S6-2 contains sixteen, S6-3 seventeen, S6-4 eighteen, S6-5 nineteen and S6-6 twenty glutamic acid residues. The amino acid sequence of the longest form, namely S6-6, has been determined as briefly described below and in fig.1. The total number of residues varies from 131 for S6-2 to 135 for S6-6. The 26 to 30 acidic and 18 basic amino acids are compatible with the isoelectric point of pH 4.8 [18]. The wild type protein S6 has an amino acid composition of Asp_{10}, Asn_4, Thr_7, Ser_4, Glu_{16-20} Gln_4, Pro_5, Gly_5, Ala_{14}, Val_{10}, Met_7,

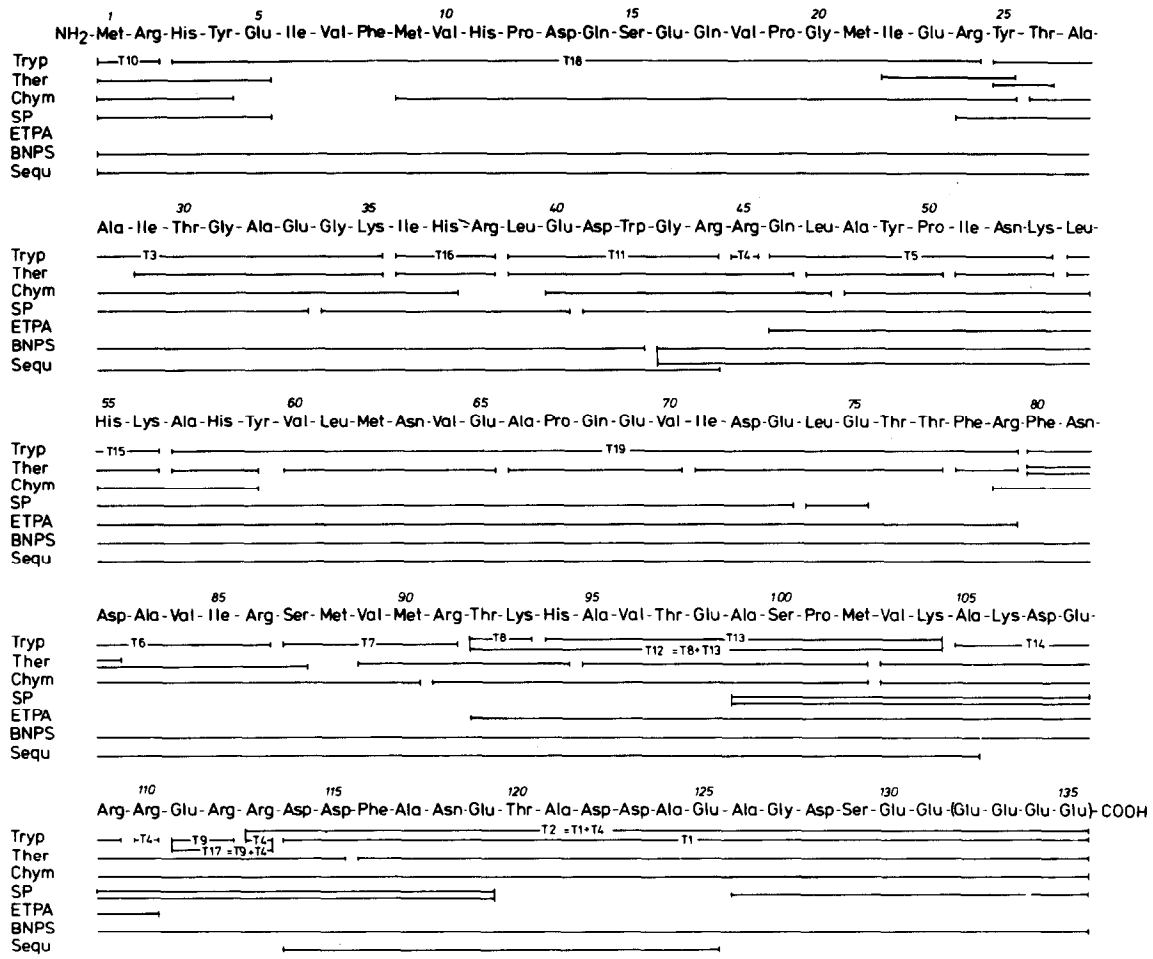

Fig.1. Amino acid sequence of protein S6 from *E. coli* ribosomes. Tryp=tryptic peptides. Ther=thermolysin peptides. Chym= chymotryptic peptides. SP= peptides from digestion with Staphylococcus protease. ETPA= tryptic peptides after blocking the lysins with ETPA. BNPS= Peptides from cleavage of the protein at its tryptophanyl bond. Sequ= sequence regions elucidated by Edman degradation in sequenators.

Ile_7, Leu_5, Tyr_4, Phe_4, His_6, Lys_6, Arg_{12}, Trp_1.

Sequenator studies on the intact protein elucidated the N-terminal sequence up to position 44 whereas the sequence of positions 43–105 were determined by sequenator studies on the BNPS-skatol fragment 2 (pos. 43–135). Solid-phase Edman degradation resulted in the sequences of the tryptic peptides T10, T18, T3, T7 and the C-terminal peptides T1. All other peptides except T19 (pos. 57–79) were sequenced by the Dansyl-Edman-technique. The alignment of the first 14 tryptic peptides was made from the protein sequenator results. The same order and the further alignment of the remaining peptides were obtained by the analyses of peptides isolated after treatment of S6 with trypsin, thermolysin, chymotrypsin or staphylococcal protease.

Comparison of all peptides from the various forms of protein S6 by thin-layer fingerprinting and by amino acid analysis showed that only the C-terminal peptides are different. They varied in their amounts of glutamic acid. This finding is in full agreement with the results obtained by the amino acid analysis of the various intact S6 forms (see above).

It was further shown by sequence determination of the tryptic peptide T1 that S6-2 ends with two, S6-3 with three, S6-4 with four, S6-5 with five and S6-6 with six succeeding glutamic acid residues at their C-termini.

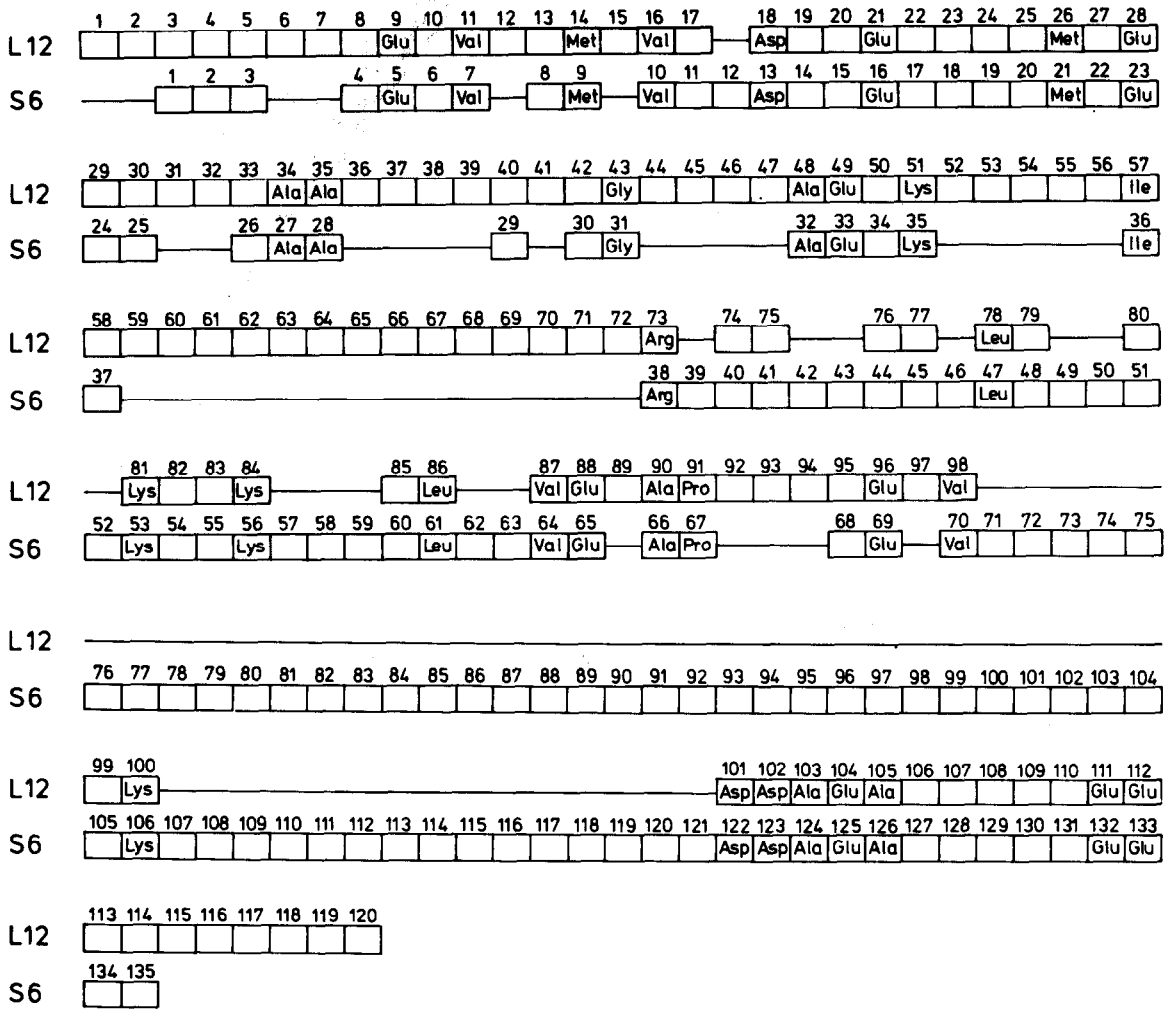

Fig.2. Comparison of the primary structures of proteins L12 and S6 from *E. coli* ribosomes. The amino acid positions are arranged in such a way as to indicate identical sequences.

Table 1
C-terminal sequence and relative amounts of the five forms of protein S6 from *E. coli* ribosomes

Protein	C-terminal sequence	Relative amounts isolated
S6-2	–Ser–Glu–Glu	1
S6-3	–Ser–Glu–Glu–Glu	4
S6-4	–Ser–Glu–Glu–Glu–Glu	7
S6-5	–Ser–Glu–Glu–Glu–Glu–Glu	3
S6-6	–Ser–Glu–Glu–Glu–Glu–Glu–Glu	1
	130 135	

This is illustrated in table 1. The high content of aspartic and glutamic acid residues in the C-terminal tryptic peptide T1 and the succession of 2–6 glutamic acid residues at the end of the protein chain are remarkable.

A comparison of the amino acid sequence of protein S6 with those [19] of the most acidic protein, namely L12, from the 50S subunit is given in fig.2. The two proteins have some regions of similar or identical structure; e.g. positions 122–126 of S6 are identical with positions 101–105 of L12. Further comparison of protein S6 with other ribosomal proteins will be given elsewhere [4].

Following the method of Chou and Fasman [20,21] 59% α-helical regions and 21% β-sheet can be predicted for the secondary structure of protein S6.

3.2. *Primary structure of the mutant protein*

In an *E. coli* mutant isolated by its resistance to neo- and kanamycin an altered S6 protein has been detected by two-dimensional polyacrylamide gel electrophoresis and by cellulose acetate electrophoresis [22]. This mutant contains only one form of S6 which is identical with S6-2, the shortest form from wild type. This was shown by the determination of the primary structure of the mutant protein in a similar way as described above. The absence of the longer forms of S6 in this mutant can be explained by mutation in the codon for glutamic acid at position 132 which leads to an ochre or amber codon each of which terminates protein synthesis.

Acknowledgement

We thank Dr H.G. Wittmann for encouragement and support, and Mrs E. Marzinzig and Mr A. Lehmann for excellent help.

References

[1] Kaltschmidt, E. and Wittmann, H.G. (1970), Anal. Biochem. 36, 401–412.
[2] Zimmermann, R.A., Mackie, G.A., Muto, A., Garrett, R.A., Ungewickell, E., Ehresmann, C., Stiegler, P., Ebel, J.P. and Fellner, P. (1975), Nucleic Acids Res. 2, 279–302.
[3] Nomura, M., Mizushima, S., Ozaki, M., Traut, P. and Lowry, C. V. (1969), Cold Spring Harbor Symp. Quant Biol. XXXIV, 49–61.
[4] Hitz, H. and Schäfer, D., manuscript in preparation.
[5] Hindennach, I., Stöffler, G. and Wittmann, H.G. (1971), Eur. J. Biochem. 23, 7-11.
[6] Hommard, J. and Drapeau, G.R. (1972), Proc. Nat. Acad. Sci. USA 69, 3506–3509.
[7] Wittmann, H.G. and Wittmann-Liebold, B. (1971), Biochem. Biophys. Acta 251, 44–53.
[8] Fontana, A. (1972), Methods Enzym. XXV, 419–423.
[9] Edman, P. and Begg, G. (1967), Eur. J. Biochem. 1, 80–91.
[10] Wittmann-Liebold, B. (1973), Hoppe-Seyler's Z. Physiol. Chem. 354, 1415–1431.
[11] Laursen, R.A. (1971), Eur. J. Biochem. 20, 89–102.
[12] Laursen, R.A., Horn, M.J. and Bonner, A.G. (1972), FEBS Lett. 21, 67–70.
[13] Wittmann-Liebold, B. and Lehmann, A., First International Conference on Solid-Phase Methods in Protein Sequence Analysis, Boston, Mass. (1975), in press.
[14] Gray, W.R. and Hartley, B.S. (1963), Biochem. J. 89, 379–380.
[15] Wittmann-Liebold, B., Geißler, A. W. and Marzinzig, E. (1975), submitted to Biochem. Biophys. Acta.
[16] Wittmann-Liebold, B., Graffunder, H. and Schäfer, W., manuscript in preparation.
[17] Held, W.A., Mizushima, S. and Nomura, M. (1973), J. Biol. Chem. 248, 5720–5730.
[18] Kaltschmidt, E. (1971), Anal. Biochem. 43, 25–31.
[19] Terhorst, C., Möller, W., Laursen, R. and Wittmann-Liebold, B. (1973), Eur. J. Biochem. 34, 138–152.
[20] Chou, P.J. and Fasman, G.D. (1974), Biochemistry 13, 211–222.
[21] Chou, P.J. and Fasman, G.D. (1974), Biochemistry 13, 222–245.
[22] Wittmann, H.G., Stöffler, G. and Apirion, D., unpublished.