THE PRIMARY STRUCTURE OF RIBOSOMAL PROTEIN S4 FROM *ESCHERICHIA COLI*

J. REINBOLT and E. SCHILTZ

Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Berlin-Dahlem, Germany

Received 6 August 1973
Originals received 9 September 1973

1. Introduction

Protein S4 from 30 S ribosomal subunits of *E. coli* is for two reasons one of the most studied and interesting ribosomal proteins: a) About 25 *E. coli* mutants with altered S4 proteins are known [1–8]. The protein chains of many mutant S4 proteins are shorter or longer than wild type S4; b) Protein S4 binds specifically to *E. coli* 16 S rRNA [9–18] and the protein–RNA interaction has been studied in some detail [14, 15, 17, 18].

In order to get more information about the molecular mechanism of both the mutations in S4 and the protein–RNA interaction, the knowledge of the primary structure of this protein is essential. Therefore we have determined the amino acid sequence of protein S4.

2. Materials and methods

S4 protein was isolated from *E. coli* K strain AB774 as previously described [19]. Separation of peptides by SE-cellulose column and paper chromatography followed the procedure by Wittmann-Liebold [20]. Amino acid sequence studies were done with a solid-phase peptide sequenator [21, 22], with an improved Beckman protein sequencer [23] and with enzymatic degradation using leucine aminopeptidase and carboxypeptidases A and B. Trypsin, chymotrypsin, thermolysin and a staphylococcal protease which is specific for glutamoyl bonds [24] were used as endoproteases. Lysine residues were blocked at their ε-amino groups by ETPA (exo-cis-3,6-endoxo-Δ^4-tetra-hydrophthalic acid anhydride) [25]; after blockage of the lysine residues trypsin splits only at the arginines. Peptides obtained after cleavage of protein S4 with BrCN [26] were separated on Sephadex G100.

Paper No. 79 on 'Ribosomal Proteins'. Preceding paper is by B. Wittmann-Liebold, FEBS Letters 36, 247–249.

3. Results and discussion

The primary structure of protein S4 as determined by the methods mentioned in 'Materials and methods' is given in table 1. The determination of the amino acid sequence of S4 will be described elsewhere in detail [27]. The protein has a molecular weight of 22 550 and consists of 203 residues. Its amino acid composition is Asp_8, Asn_8, Thr_6, Ser_{10}, Glu_{17}, Gln_9, Pro_6, Gly_{15}, Ala_{18}, Val_{16}, Cys_1, Met_3, Ile_8, Leu_{20}, Tyr_8, Phe_4, His_3, Lys_{20}, Arg_{22} and Trp_1. The composition with 25 acidic and 45 basic amino acids is in good agreement with the high isoelectric point of this protein [28].

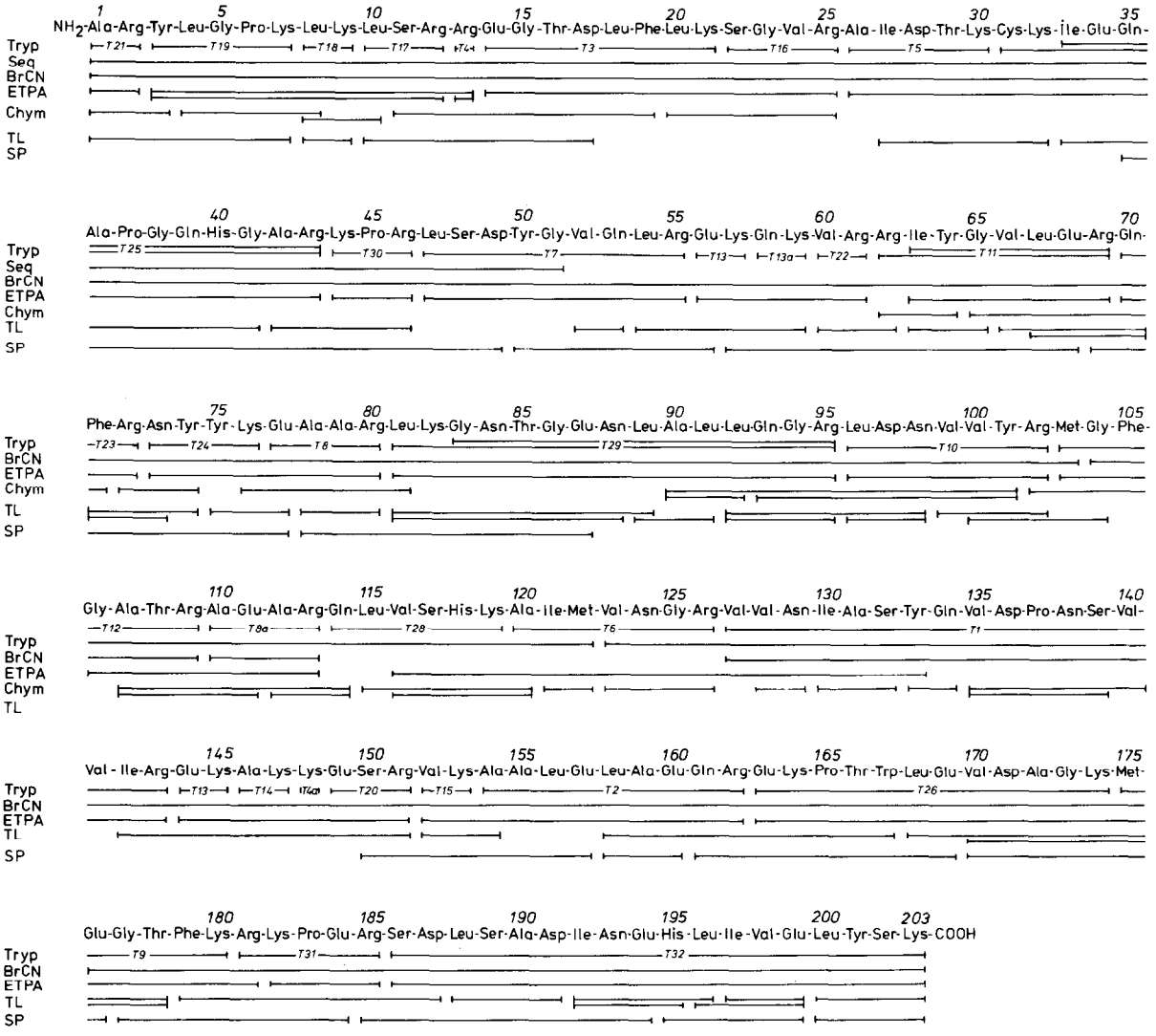
Two tripeptides occur twice in the primary structure, namely Tyr–Gly–Val in positions 50–52 and 64–66, Arg–Lys–Pro in 43–45 and 181–183, one tripeptide even occurs three times Arg–Glu–Lys in 55–57, 143–145 and 162–164. Fifteen dipeptides are present between two and four times, namely seven dipeptides twice, six dipeptides three times and two dipeptides four times: Arg–Glu in positions 13–14, 55–56, 143–144 and 163–164, Ala–Arg in positions 1–2, 42–43, 79–80 and 112–113.

Because of the relatively high number of basic resi-

 North-Holland Publishing Company – Amsterdam

Table 1

Sequence of the ribosomal protein S4 of *E. coli*.



Tryp. = peptides from trypsin digestion; Seq. = automatic sequencing by the Beckman-sequenator; BrCN = peptides resulting from BrCN cleavage; ETPA = peptides obtained from a tryptic digest after blocking the amino groups of the protein by ETPA; Chym. = peptides from chymotrypsin digestion; TL = peptides from thermolysin digestion; SP = peptides from staphylococcal protease digestion.

dues (45 out of 203) it is not surprising to find accumulation of basic amino acids, e.g. in positions 7–13, 43–46, 55–62, 143–153 and 180–185. It remains to be seen whether these regions are involved in protein–RNA interaction. Studies with chemical modification of protein S4, when bound to RNA or free, are in progress.

Knowledge of the primary sequence of protein S4 is also necessary in order to determine the molecular mechanism by which S4 protein chains of many revertants from streptomycin dependence to independence become shorter or longer than S4 from the wild type [6]. Comparison of the S4 proteins from such a mutant and wild type revealed a combination of deletion and frame shift as the mechanism for premature chain termination (E. Schiltz, unpublished results).

Acknowledgements

We thank Dr. H.G. Wittmann and Dr. B. Wittmann-Liebold for stimulating discussions and helping with the manuscript. The technical assistance of Mrs. J. Kuhlmeyer, Mrs. E. Nehls and Mr. W. Stubba is greatly appreciated. We also wish to thank Dr. G. Funatsu for providing the chymotryptic peptides.

References

[1] Deusser, E., Stöffler, G., Wittmann, H.G. and Apirion, D. (1970) Mol. Gen. Genet. 109, 298.
[2] Birge, E.A. and Kurland, C.G. (1970) Mol. Gen. Genet. 109, 356.
[3] Kreider, G. and Brownstein, B.L. (1971) J. Mol. Biol. 61, 135.
[4] Zimmermann, R.A., Carvin, R.T. and Gorini, L. (1971) Proc. Natl. Acad. Sci. U.S. 68, 2263.
[5] Donner, D. and Kurland, C.G. (1972) Mol. Gen. Genet. 115, 49.
[6] Funatsu, G., Puls, W., Schiltz, E., Reinbold, J. and Wittmann, H.G. (1972) Mol. Gen. Genet. 115, 131.
[7] Zimmermann, R.A., Ikeya, Y. and Sparling, P.F. (1973) Proc. Natl. Acad. Sci. U.S. 70, 71.
[8] Hasenbank, R., Guthrie, C., Stöffler, G., Wittmann, H.G., Rosen, L. and Apirion, D. Mol. Gen. Genet., in press.
[9] Mizushima, S. and Nomura, M. (1970) Nature 226, 1214.
[10] Schaup, H.W., Green, M. and Kurland, C.G. (1970) Mol. Gen. Genet. 109, 193.
[11] Schaup, H.W., Green, M. and Kurland, C.G. (1971) Mol. Gen. Genet. 112, 1.
[12] Garrett, R.A., Rak, K.H., Daya, L. and Stöffler, G. (1971) Mol. Gen. Genet. 114, 112.
[13] Stöffler, G., Daya, L., Rak, K.H. and Garrett, R.A. (1971) J. Mol. Biol. 62, 411.
[14] Schaup, H.W., Sogin, M., Woese, C. and Kurland, C.G. (1972) Mol. Gen. Genet. 114, 1.
[15] Zimmermann, R.A., Muto, A., Fellner, P., Ehresmann, C. and Branlant, C. (1972) Proc. Natl. Acad. Sci. U.S. 69, 1282.
[16] Nanninga, N., Garrett, R.A., Stöffler, G. and Klotz, G. (1972) Mol. Gen. Genet. 119, 175.
[17] Daya-Grosjean, L., Garrett, R.A., Pongs, O., Stöffler, G. and Wittmann, H.G. (1972) Mol. Gen. Genet. 119, 277.
[18] Schulte, C. and Garrett, R.A. (1972) Mol. Gen. Genet. 119, 345.
[19] Hindennach, I., Stöffler, G. and Wittmann, H.G. (1971) European J. Biochem. 23, 7.
[20] Wittmann-Liebold, B. (1971) Hoppe-Seyler's Z. Physiol. Chem. 352, 1705.
[21] Laursen, R.A. (1971) European J. Biochem. 20, 89.
[22] Laursen, R.A., Horn, M.J. and Bonner, A.G. (1972) FEBS Letters 21, 67.
[23] Wittmann-Liebold, B., Hoppe-Seyler's Z. Physiol. Chem., in press.
[24] Houmard, J. and Drapeau, G.R. (1972) Proc. Natl. Acad. Sci. U.S. 69, 3506.
[25] Riley, M. and Perham, R.N. (1970) Biochem. J. 118, 733.
[26] Gross, E. and Witkop, B. (1962) J. Biol. Chem. 237, 1856.
[27] Schiltz, E. and Reinbolt, J., manuscript in preparation.
[28] Kaltschmidt, E. (1971) Anal. Biochem. 43, 25.