THE PRIMARY STRUCTURE OF THE 16 S rRNA BINDING PROTEIN S 8 FROM *ESCHERICHIA COLI* RIBOSOMES

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

Received 6 September 1974

1. Introduction

Protein S 8 from the 30 S ribosomal subunit of *Escherichia coli* was sequenced for the following reasons:
a) It binds specifically to the 16 S RNA [1–3] and a specific complex containing S 8 and 40 nucleotides of known sequence has been isolated by controlled ribonuclease digestion [4]. This small ribonucleoprotein particle, which can be dissociated and reassociated, is a promising model for investigating the chemical basis of protein–RNA interaction [5].
b) Two temperature sensitive mutants with an altered S 8 protein have been described (6). The identification of the type and location of the amino acid substitutions may also facilitate the understanding of the S 8 function in the ribosome.
c) Some general properties concerning the shape and secondary structure of the protein have been determined from hydrodynamic and low angle X-ray diffraction studies (see Discussion).

2. Materials and methods

Protein S 8 was isolated from *E. coli* K, strain AB 774, as previously described [7]. About 30 mg of protein was obtained from 1 kg of wet cells. Peptides, obtained by cyanogen bromide cleavage and by tryptic, chymotryptic and thermolysin digestion, were separated by gel-filtration, ion exchange and paper chromatography. Amino acid sequences were determined by automatic Edman degradation using a Beckman protein sequenator improved by Wittmann-Liebold [8], and a solid-phase peptide sequenator [9]. For peptide T1 manual Edman degradation and carboxypeptidase A digestion were used. PTH-amino acids were identified by thin-layer chromatography and mass spectroscopy. The methods will be later described in detail [10].

3. Results and discussion

The protein contains 109 amino acids and has a composition of Asp_6, Asn_3, Thr_5, Ser_5, Glu_7, Gln_6, Pro_5, Gly_6, Ala_{10}, Val_9, Met_4, Ile_7, Leu_9, Tyr_3, Phe_3, Lys_{12}, Arg_8, Cys_1. The 13 acidic and 20 basic amino acids are compatible with the measured isoelectric point of pH 9.2 [11].

The amino acid sequence of protein S 8 is shown in fig. 1. Details of its determination will be given elsewhere [10]. The protein sequenator gave the following results: The N-terminal sequence up to position 40, sequence 27–68 from a cyanogen bromide fragment (pos. 27–95), and the complete sequence of the tryptic peptide T16 (pos. 41–63) after modification with Braunitzer's reagent IV [12]. Solid phase Edman degradation gave the sequences of the tryptic peptides except T1, T4 and T16. The C-terminal peptide T1 was sequenced by manual Edman degradation and carboxypeptidase A treatment as well as by splitting the peptide with trypsin after modification of the cysteine residue [13].

The order of the first eight tryptic peptides was determined by the protein sequenator results. Analysis of chymotryptic, thermolysin and cyanogen bromide peptides confirmed this order and, in addition, gave the alignment of the remaining tryptic peptides. The lysine residue in position 49

North-Holland Publishing Company – Amsterdam

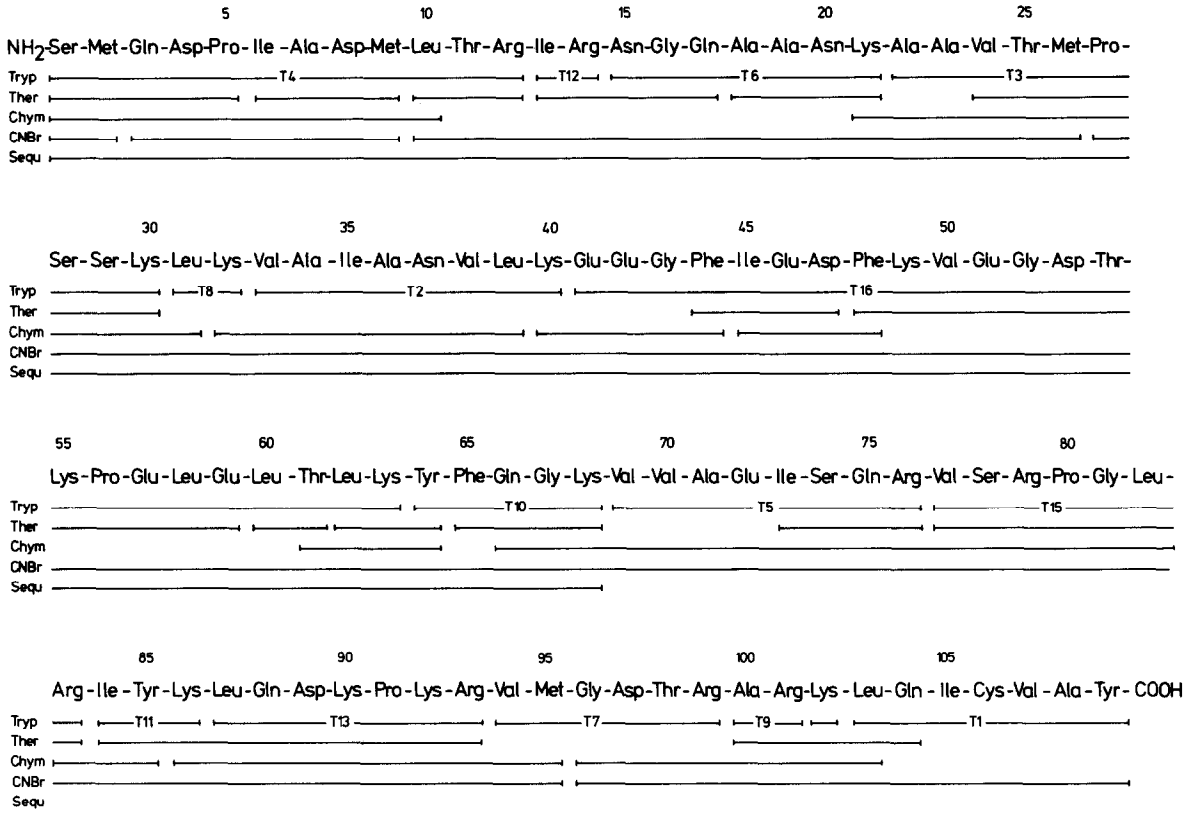

Fig. 1. Sequence of the ribosomal protein S 8 of *E. coli*. Tryp. = peptides from trypsin digestion; Sequ. = automatic sequencing by the Beckman-sequenator; CNBr = peptides resulting from CNBr cleavage; Chym. = peptides from chymotrypsin digestion; Ther = peptides from thermolysin digestion.

was not split by trypsin, although it is followed by a valine. The reason for this peculiarity is not yet known.

A comparison with the amino acid sequence of the rRNA binding protein S 4 [14] of *E. coli* shows some similarities: a) Four tripeptides have identical sequences (see table 1). b) In protein S 4, four of the six prolines present neighbour basic amino acids and in protein S 8 three of the five prolines are also adjacent to arginine or lysine residues.

Other noteworthy features of the S 8 sequence are: a) Basic amino acids are often separated by a single hydrophobic amino acid residue, for example: Arg–Ile–Arg; Arg–Ala–Arg; Lys–Leu–Lys; Lys–Pro–Lys. b) There is an unequal distribution of basic, acidic and hydrophobic residues. The N-terminal part of the protein is mainly hydrophobic with scattered basic amino acids, the middle portion shows clustered acidic residues and the C-terminal region is more basic.

X-ray diffraction studies of S 8 fibers and low angle X-ray scattering in solution indicate S 8 to be a molecule of high axial ratio and 50–55% α-helical structure at acidic pH [15]. The helix content agrees

Table 1
Identical tripeptides in ribosomal proteins S 4 and S 8

Protein	Positions	Sequence	Protein	Positions
S 4	7–9	Lys–Leu–Lys	S 8	30–32
S 4	42–44	Ala–Arg–Lys	S 8	100–102
S 4	62–64	Arg–Ile–Tyr	S 8	83–85
S 4	106–109	Thr–Arg–Ala	S 8	96–100

with Chou and Fasman's rules to predict protein secondary structure [16,17]. Using their parameters, high probabilities for α-helical structures were obtained for positions 6–10, 17–26, 30–51, 57–63, 68–73 and 86–90. A cyanogen bromide peptide (pos. 27–95) isolated from S 8 gave very well orientated fibers which showed an α-helical arrangement. The elucidation of the S 8 sequence will be useful for further physical and physical-chemical studies especially on the complex of the protein with its small RNA binding site as well as for chemical studies on the amino acid replacements in the mutants mentioned in the introduction.

Acknowledgement

I would like to thank Dr B. Wittmann-Liebold for her help and introduction to sequencing methodology.

References

[1] Mizushima, S., Nomura, M. (1970) Nature 226, 1214.
[2] Schaup, H. W., Green, M., Kurland, C. G. (1970) Molec. Gen. Genet. 109, 193.
[3] Garrett, R. A., Rak, K. H., Daya, L. and Stöffler, G. (1971) Molec. Gen. Genet. 114, 112.
[4] Schaup, H. W., Sogin, M. L., Kurland, C. G. and Woese, C. R. (1973) J. Bacteriol. 115, 82.
[5] Ungewickell, E., Garrett, R. A., Ehresmann, C., Stiegler, P. and Fellner, P., submitted to Eur. J. Biochem.
[6] Zubke, W., Stöffler, G., Wittmann, H. G. and Apirion, D. (1974) Molec. Gen. Genet., submitted.
[7] Hindennach, I., Stöffler, G. and Wittmann, H. G. (1971) Eur. J. Biochem. 23, 7.
[8] Wittmann-Liebold, B. (1973) Hoppe-Seyler's Z. Physiol. Chem. 354, 1415.
[9] Laursen, R. A. (1971) Eur. J. Biochem. 20, 89.
[10] Stadler, H., manuscript in preparation.
[11] Kaltschmidt, E. (1971) Analytical Biochem. 43, 25.
[12] Braunitzer, G., Schrank, B., Ruhfus, A., Petersen, S. and Petersen, U. (1971) Hoppe-Seyler's Z. Physiol. Chem. 352, 1730.
[13] Raftery, M. A. and Cole, R. O. (1963) Biochem. Biophys. Res. Commun. 10, 467.
[14] Reinboldt, J. and Schiltz, E. (1973) FEBS Letters 36, 250.
[15] Stadler, H. and Paradies, H. H., manuscript in preparation.
[16] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 211.
[17] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 222.
[18] Paradies, H. H., manuscript submitted.