THE PRIMARY STRUCTURE OF PROTEIN S17 FROM THE SMALL RIBOSOMAL SUBUNIT OF *ESCHERICHIA COLI*

Makoto YAGUCHI
Division of Biological Sciences, National Research Council of Canada, Ottawa, Canada

and

H. G. WITTMANN
Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem, Germány

Received 23 December 1977

1. Introduction

Protein S17 is a basic protein of the *E. coli* 30 S ribosomal subunit [1]. It binds specifically to the 5'-terminus of the 16 S RNA [2] and it is located in the 'head' of the small subunit as found by immuno-electron microscopy [3]. Protein S17 is the product of gene *nea*A, and mutation in this gene confers resistance to the antibiotic neamine [4,5]. The amino acid replacement in protein S17 of a mutant resistant to neamine is located in residue 30 [6]. The gene for S17 has been found in the *Spc–Str* region of the *E. coli* chromosome [5,7].

The sequence of the N-terminal region of S17 has been determined [8]. We report here the complete primary structure of this protein which consists of 83 residues and has mol wt 9573.

2. Materials and methods

Protein S17 was isolated from *E. coli* K as described [9]. The identity and purity of the protein were checked by two-dimensional polyacrylamide gel electrophoresis [10]. Performic acid oxidation [11] was done at 0°C for 1 h. Tryptic and chymotryptic digestions were carried out at pH 8.0 and 37°C for 4 h or 24 h, respectively. Thermolytic digestion was at pH 8.0 and 55°C for 4 h. Digestion with *Staphylococcus aureus* protease kindly supplied by Dr G. R. Drapeau, University of Montreal, was in 50 mM acetic acid, pH 4.0, for 16 h [12]. Large fragments were produced by cyanogen bromide [13] and *N*-bromosuccinimide [14].

The isolation of large peptides was achieved by gel filtration of various digests (about 5 mg each) on Sephadex G-50 and G-75 (superfine) columns (250 × 1.5 cm). Acetic acid (5%) or 0.01 N HCl was used for the elution. Smaller peptides in the various fractions eluted from the Sephadex columns were further separated by fingerprinting on cellulose thin-layer plates [15]. Amino acid analyses were performed on a Durrum D-500 amino acid analyzer. The presence of tryptophan in the peptides was established by fingerprints on cellulose thin-layer plates sprayed with Ehrlich's reagent.

Automatic Edman degradation [16] of the larger peptides was made in a Beckman model 890C sequenator. The amino acid sequence of smaller peptides was determined by a manual micro Edman technique [17] without dansylation. The thiazolinone or PTH derivatives were hydrolysed with 6 N HCl in the presence or absence of 0.1% $SnCl_2$ [18] at 130°C for 20 h, and the amino acids formed were analyzed on a Durrum analyzer. The identification

NRCC Publication No. 16161

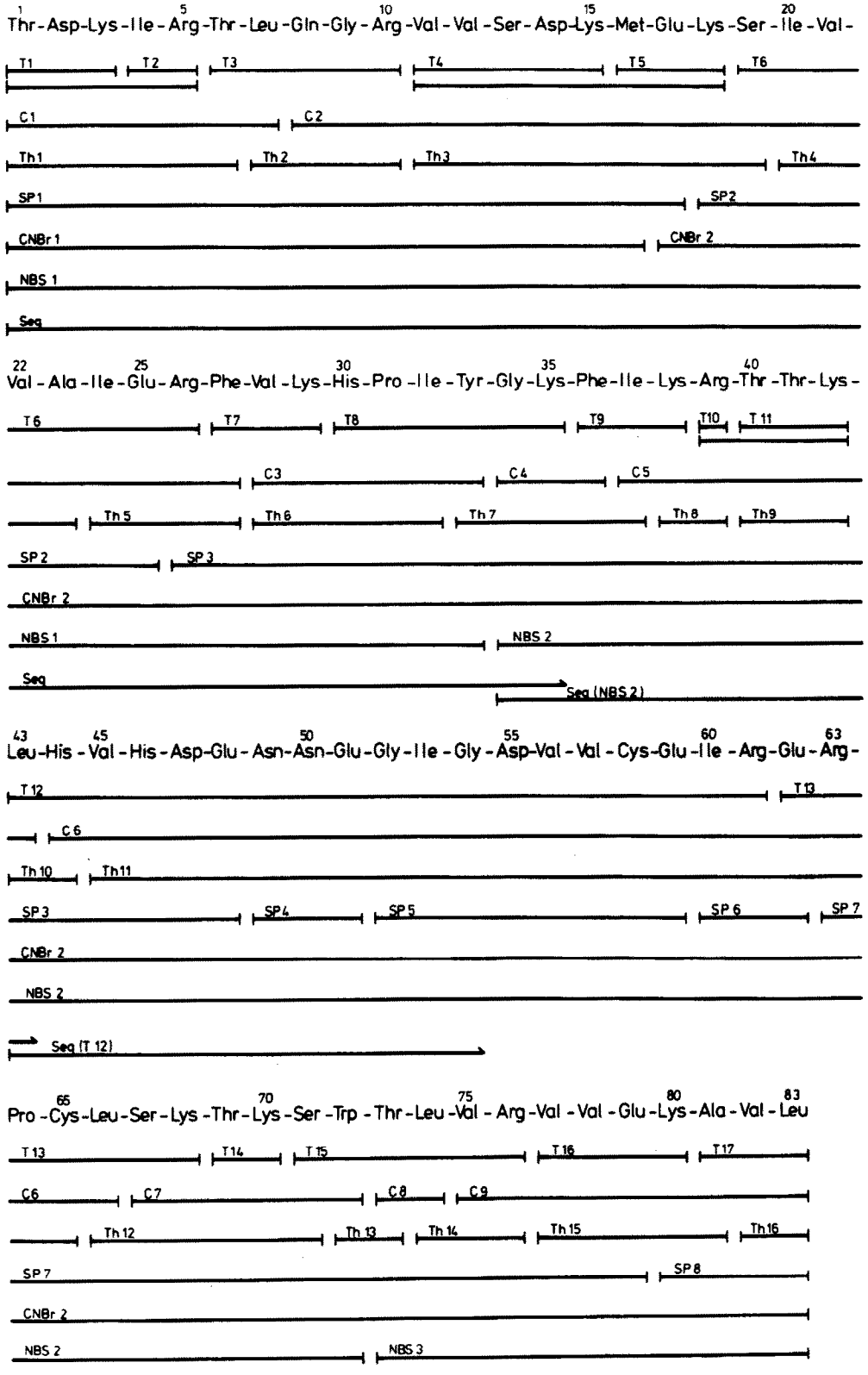
1 5 10 15 20
Thr-Asp-Lys-Ile-Arg-Thr-Leu-Gln-Gly-Arg-Val-Val-Ser-Asp-Lys-Met-Glu-Lys-Ser-Ile-Val-
T1 T2 T3 T4 T5 T6
C1 C2
Th1 Th2 Th3 Th4
SP1 SP2
CNBr 1 CNBr 2
NBS 1
Seq
22 25 30 35 40
Val-Ala-Ile-Glu-Arg-Phe-Val-Lys-His-Pro-Ile-Tyr-Gly-Lys-Phe-Ile-Lys-Arg-Thr-Thr-Lys-
T6 T7 T8 T9 T10 T11
C3 C4 C5
Th5 Th6 Th7 Th8 Th9
SP2 SP3
CNBr 2
NBS 1 NBS 2
Seq Seq (NBS 2)
43 45 50 55 60 63
Leu-His-Val-His-Asp-Glu-Asn-Asn-Glu-Gly-Ile-Gly-Asp-Val-Val-Cys-Glu-Ile-Arg-Glu-Arg-
T12 T13
C6
Th10 Th11
SP3 SP4 SP5 SP6 SP7
CNBr 2
NBS 2
Seq (T 12)
65 70 75 80 83
Pro-Cys-Leu-Ser-Lys-Thr-Lys-Ser-Trp-Thr-Leu-Val-Arg-Val-Val-Glu-Lys-Ala-Val-Leu
T13 T14 T15 T16 T17
C6 C7 C8 C9
Th12 Th13 Th14 Th15 Th16
SP7 SP8
CNBr 2
NBS 2 NBS 3

Fig.1

of PTH derivatives was made by thin-layer chromatography on silica-gel plates [19,20].

3. Results and discussion

The automatic sequence determination of protein S17 [8] as well as of the peptides NBS2 and T12 established the sequence up to residue 54 as shown in fig.1. Treatment of protein S17 with trypsin resulted in 17 peptides. Some peptides, namely T1+T2 (residues 1–5), T4+T5 (11–18), T10+T11 (39–42) were isolated as joint peptides due to partial cleavage. The trypsin-susceptible bond of Asp–Lys–Ile–Arg (positions 2–5) and Asp–Lys–Met–Glu [14–17] were not cleaved after 4 h but only after 24 h hydrolysis. It is known that the presence of charged acidic group near to the trypsin-susceptible bond reduces the rate of hydrolysis very significantly [21]. All tryptic peptides except the C-terminal portion of T12 were sequenced manually.

Digestion of protein S17 with chymotrypsin and thermolysin gave 9 and 16 peptides, respectively. Those of the C-terminal region of the protein chain (peptides C7, C8, C9, Th12, Th13, Th14, Th15, Th16) were manually sequenced.

Staphylococcus aureus protease specifically produced 8 peptides of which SP4, SP5, SP6 and SP8 were sequenced manually. This resulted in the sequence of the C-terminal region of T12 and the overlap of T12 and T13. The sequence of the C-terminal end of S17 was confirmed by the analysis with carboxypeptidase A.

The combination of the described results gave the alignment of all peptides and the complete amino acid sequence as illustrated in fig.1. The amino acid composition derived from the sequence of S17 is:

Asp_4, Asn_2, Thr_6, Ser_4, Glu_7, Gln_1, Pro_2, Gly_4, Ala_2, Val_{12}, Met_1, Ile_7, Leu_5, Tyr_1, Phe_2, His_3, Lys_{10}, Arg_7, Cys_2 and Trp_1.

This composition is in very good agreement with the values obtained after hydrolysis of the protein. Protein S17 has mol. wt 9573.

Comparison of the sequence of protein S17 with other ribosomal proteins of known structure [3,22] revealed several identical regions which are shown in table 1.

Based on the amino acid sequence of protein S17, predictions for a possible secondary structure of this protein were made using different predictive methods [23–26]. According to this approach which has also been used for other ribosomal proteins [27–30] the following structural features are expected for protein S17: three helical regions for positions 15–18, 23/25-27/28 and 79–82; turns or loops for positions 9–10, 33–35, 47–51, 64–65 and 70–71; extended structure or β-sheet for positions 4–7, 57–58 and around position 73. In addition extended structures might be possible for the regions in positions 19–24, 52–56 and 74–78.

The amino acid replacement in the altered protein

Table 1
Regions of protein S17 identical with regions of other *E. coli* ribosomal proteins

Peptide	Protein	Positions	Protein	Positions
Val–Glu–Lys–Ala–Val	S17	78–82	S19	18– 22
Leu–Gln–Gly–Arg	S17	7–10	S4	92– 95
Met–Glu–Lys–Ser	S17	16–19	L23	24– 27
Gly–Arg–Val–Val	S17	9–12	S4	125–128
Tyr–Gly–Lys–Phe	S17	33–36	IF-3	71– 74

Fig.1. Amino acid sequences of protein S17 from *E. coli* ribosomes. T, tryptic peptide; C, chymotryptic peptide; Th, thermolytic peptide; SP, peptide from digestion with *Staphylococcus aureus* protease; CNBr, peptide cleaved with cyanogen bromide; NBS, peptide cleaved with *N*-bromosuccinimide; Seq, automatic sequencing by liquid-phase degradation in a sequenator.

S17 from a neamine-resistant mutant altered in the fidelity of translation [5] is located at position 30 [6]. In a conditionally lethal mutant, which is defective in 30 S subunit assembly, the amino acid replacement in S17 was found at position 67 [31]. Furthermore, an altered protein S17 was identified in strain AT2472 [32] and in a streptomycin-independent revertant of mutant VT [33].

Comparative studies on 30 S ribosomal proteins from *E. coli* and *B. stearothermophilus* show a structural and functional correspondence between the two sets of proteins [34–36]. There is also a high degree of sequence homology between the N-terminal sequence of protein S17 from both bacteria [37].

Acknowledgements

We thank Dr Wittmann-Liebold for many stimulating discussions and for provision of computer data including homologies and secondary structure, and J. Krauss and C. Roy for excellent technical assistance.

References

[1] Wittmann, H. G. (1974) in: Ribosomes (Nomura, M., Tissières, A. and Lengyel, P. eds) pp. 93–114, Cold Spring Harbor Lab., Cold Spring Harbor, New York.
[2] Zimmermann, R. A. (1974) in: Ribosomes (Nomura, M., Tissières, A. and Lengyel, P. eds) pp. 225–269, Cold Spring Harbor Lab., Cold Spring Harbor, New York.
[3] Stöffler, G. and Wittmann, H. G. (1977) in: Molecular Mechanism of Protein Biosynthesis (Weissbach, H. and Pestka, S. eds) pp. 117–202, Academic Press, New York.
[4] Cannon, M., Cabezón, T. and Bollen, A. (1974) Molec. Gen. Genet. 130, 321–326.
[5] Bollen, A., Cabezón, T., DeWilde, M., Villarroel, R. and Herzog, A. (1975) J. Mol. Biol. 99, 795–806.
[6] Yaguchi, M., Wittmann, H. G., Cabezón, T., DeWilde, M., Villarroel, R., Herzog, A. and Bollen, A. (1976) J. Mol. Biol. 104, 617–620.
[7] Dekio, S. (1971) Molec. Gen. Genet. 113, 20–30.
[8] Wittmann-Liebold, B. (1973) FEBS Lett. 36, 247–249.
[9] Hindennach, I., Stöffler, G. and Wittmann, H. G. (1971) Eur. J. Biochem. 23, 7–11.
[10] Kaltschmidt, E. and Wittmann, H. G. (1970) Analyt. Biochem. 36, 401–412.
[11] Hirs, C. H. W. (1956) J. Biol. Chem. 219, 611–621.
[12] Houmard, J. and Drapeau, G. R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.
[13] Gross, E. (1967) in: Methods in Enzymology, Vol XI (Hirs, C. H. W. ed) pp. 238–255, Academic Press, New York.
[14] Ramachandran, L. K. and Witkop, B. (1967) in: Methods in Enzymology, Vol XI (Hirs, C. H. W. ed) pp. 283–299, Academic Press, New York.
[15] Yaguchi, M., Wittmann, H. G., Cabezón, T., DeWilde, M., Villarroel, R., Herzog, A. and Bollen, A. (1975) Molec. Gen. Genet. 142, 35–43.
[16] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[17] Bruton, C. J. and Hartley, B. S. (1970) J. Mol. Biol. 52, 165–178.
[18] Mendez, E. and Lai, S. Y. (1975) Analyt. Biochem. 68, 47–53.
[19] Wittmann-Liebold, B., Geissler, A.-W., Marzinzig, E. (1975) J. Supramol. Struc. 3, 426–447.
[20] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.
[21] Kasper, C. B. (1975) in: Protein Sequence Determination (Needleman, S. B. ed) p. 155, Springer-Verlag, Berlin.
[22] Wittmann-Liebold, B. and Dzionara, M. (1976) FEBS Lett. 61, 14–19.
[23] Burgess, A. W., Ponnuswamy, P. K. and Scheraga, H. A. (1974) Israel J. Chem. 12, 239–286.
[24] Chou, P. Y., Adler, A. J. and Fasman, G. D. (1975) J. Mol. Biol. 96, 29–45.
[25] Nagano, K. (1977) J. Mol. Biol. 109, 251–274.
[26] Robson, B. and Suzuki, E. (1976) J. Mol. Biol. 107, 327–356.
[27] Wittmann-Liebold, B., Robinson, S. M. L. and Dzionara, M. (1977) FEBS Lett. 81, 204–213.
[28] Wittmann-Liebold, B., Robinson, S. M. L. and Dzionara, M. (1977) FEBS Lett. 77, 301–307.
[29] Dzionara, M., Robinson, S. M. L. and Wittmann-Liebold, B. (1977) Hoppe-Seyler's Z. Physiol. Chem. 358, 1003–1019.
[30] Dzionara, M., Robinson, S. M. L. and Wittmann-Liebold, B. (1978) J. Supramol. Struct. in press.
[31] Herzog, A., Cabezón, T., Petre, J., Yaguchi, M. and Bollen, A. (1978) submitted.
[32] Muto, A., Otaka, E., Itoh, T., Osawa, S. and Wittmann, H. G. (1975) Molec. Gen. Genet. 140, 1–5.
[33] Dabbs, E. R. and Wittmann, H. G. (1976) Molec. Gen. Genet. 149, 303–309.
[34] Isono, K., Isono, S., Stöffler, G., Visentin, L. P., Yaguchi, M. and Matheson, A. T. (1973) Molec. Gen. Genet. 127, 191–195.
[35] Higo, K., Held, W., Kahan, L. and Nomura, M. (1973) Proc. Natl. Acad. Sci. USA 70, 944–948.
[36] Isono, S. and Isono, K. (1975) Eur. J. Biochem. 50, 482–488.
[37] Yaguchi, M., Matheson, A. T. and Visentin, L. P. (1974) FEBS Lett. 46, 296–300.