PRIMARY STRUCTURE OF PROTEIN S12 FROM THE SMALL *ESCHERICHIA COLI* RIBOSOMAL SUBUNIT

Gunki FUNATSU

Laboratory of Biochemistry, Faculty of Agriculture, Kyushu University, Fukuoka, Japan

and

Makoto YAGUCHI*

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

and

Brigitte WITTMANN-LIEBOLD

Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem, Germany

Received 4 December 1976

1. Introduction

Protein S12 is a very basic protein of the *E. coli* 30 S ribsomal subunit [1,2]. It controls the fidelity of translation [3] and plays an important role in the initiation of natural messenger RNA translation [4]. Protein S12 is the product of gene strA [5] and mutation in this protein confers resistance to [6] and dependence on streptomycin [7]. The amino acid replacements in mutants with altered S12 proteins are clustered in two regions: tryptic peptide T6, or T15 [8–11]. Some mutants resistant to neamine have two altered ribosomal proteins, S12 and S5 [12], and the amino acid replacements in S12 are also located in peptide T 15 [13].

The tryptic peptides of protein S12 have been isolated and their amino acid compositions reported [8,10]. The amino acid sequence of the N-terminal region of S12 has been determined [14,15]. Protein S12 consists of 123 amino acids and this report summarizes the determination of the amino acid sequence.

*NRCC No. 15580

2. Materials and methods

Protein S12 was isolated from *E. coli* K strain 19 as previously described [16] and provided by Dr H. G. Wittmann. The identity and purity of the protein was checked by two-dimensional polyacrylamide gel electrophoresis [17]. Performic acid oxidation [18] was performed at 0°C for 1 h. Acetylation of the protein was carried out with acetic acid anhydride in half-saturated solution of sodium acetate at 0°C for 1 h [19]. Tryptic and chymotryptic digestion was performed at pH 8.0 at 37°C for 2, 4, or 20 h. Digestion with thermolysin was at pH 8.0 at 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 [20].

Separation of the peptides was performed by SE-cellulose column [21] and paper chromatography. The fingerprint of the peptides was made on Whatman paper 3 MM by electrophoresis at 3 kV in pyridine/acetic acid/H_2O/acetone (2:4:79:15) buffer, pH 4.4, for the first dimension and paper chromatography using *n*BuOH/H_2O/acetic acid/pyridine (30:24:6:20) for the second dimension. The isolation of larger peptides such

North-Holland Publishing Company – Amsterdam

as T24, SP1, SP2 and SP5 was achieved by gel filtration of the various digests (about 5 mg each) on Sephadex G-50 and G-25 superfine columns (250 X 1.5 cm) and 15% acetic acid was used for the elution. Smaller peptides in the various fractions eluted from the Sephadex columns were further separated by fingerprinting on thin-layer cellulose [13]. Amino acid analyses were performed with Durrum, Biocal or Jeol analysers.

The amino acid sequence of peptides was determined by a manual micro Edman technique [22–24] without dansylation. The thiazolinone or PTH derivatives were hydrolyzed with 6 N HCl in the presence or absence of 0.1% $SnCl_2$ [25] at 130°C for 20 h, and the amino acids formed were analysed with the Durrum analyser. The identification of PTH derivatives was made by thin-layer chromatography on silica gel plates [26]. Automatic Edman degradation [27] of the protein was made in an improved liquid phase sequenator [28] equipped with an automatic conversion device [29] or a Beckman model 890C sequenator. Edman degradations were also performed by the solid phase technique [30] with attachment of the lysine or arginine containing peptides to aminopolystyrene resin using *p*-phenylenediisothiocyanate directly or after conversion of arginine to ornithine by treating with 20% hydrazine at 105°C for 30 min [31,32] or by attachment of the C-terminal carboxyl group to aminopolystyrene resin with a water soluble carbodiimide [31,33].

3. Results and discussion

By treatment of protein S12 with trypsin 24 peptides were obtained as well as free alanine and lysine. Some of the tryptic peptides, namely T18 (14–17), T23 (18–35), T16 (110–113) and T20 (114–119) are joint peptides due to partial cleavage. All these peptides were sequenced by a manual micro Edman technique or by the solid-phase sequence method as described under Materials and methods. Tryptic peptide T24 (56–82) was sequenced by liquid-phase Edman degradation up to position 71; the C-terminal part of T24 was determined by the analyses of other fragments (see below). Amino acid compositions of the peptides obtained from the tryptic digestion of acetylated S12 revealed the alignments of peptides T18–T23 (positions 14–35), T6–T19 (36–49), T4–T17 (99–110) and T20–T21b–T1 (114–123).

Digestions of protein S12 with chymotrypsin gave 16 peptides, most of which were sequenced manually. The sequence of chymotryptic peptide C15 (positions 95–116) was determined by the solid-phase method for positions 95–112. Peptides C4 (7–37), C8 (49–56), C9 (38–56) and C15 (95–116) were obtained by short digestion, indicating the alignments of peptides C2–C3, C6–C7, C5–C6–C7 and C13–C14. Further tryptic cleavage of these chymotryptic peptides released the tryptic peptides and their fragments derived from the corresponding regions. These were identified by fingerprinting of the hydrolysates and analyses of the peptides.

Staphylococcus aureus protease specifically produced only five peptides. The sum of the amino acid residues of these peptides agrees with the total amino acid composition of protein S12. Liquid-phase degradation of peptide SP5 (76–123) determined the sequence up to position 93 and established the sequence of the C-terminal region of tryptic peptide T24. It gave the alignment T24–T10–T15.

Furthermore, the thermolysin peptides were isolated and most of them were sequenced by the manual degradation technique. Peptides TH13 (56–59), TH14 (60–61), TH15 (62–65) and TH16 (66–72) were found useful to establish the sequence of T24.

The combination of these results with the N-terminal sequence of protein S12 determined previously by means of liquid-phase Edman degradation [14,15,28] gave the alignment of all peptides and the complete amino acid sequence as illustrated in fig.1.

The unknown residue X at position 88 is yet to be identified. Manual Edman degradation of peptide T15 (86–93) gave a gap for position 88; the same result was obtained by degrading peptide SP5 in the sequenator. It should be pointed out that the peptide bond between lysine at position 87 and X at position 88 is not cleaved by trypsin hydrolysis. Several streptomycin mutants were found to have aspartic acid in place of X at position 88 (unpublished results). The experiments for the identification of the unknown amino acid are in progress.

The amino acid composition derived from the sequence of S12 is: Asp_2, Asn_5, Thr_8, Ser_6, Glu_4, Gln_4, Pro_7, Gly_{11}, Ala_9, Val_{15}, Ile_3, Leu_8, Tyr_4, Phe_1, His_3, Lys_{13}, Arg_{15}, Cys_4 and one unknown residue X. It is in very good agreement with the data determined from the total hydrolysis of the intact or performic acid oxidized

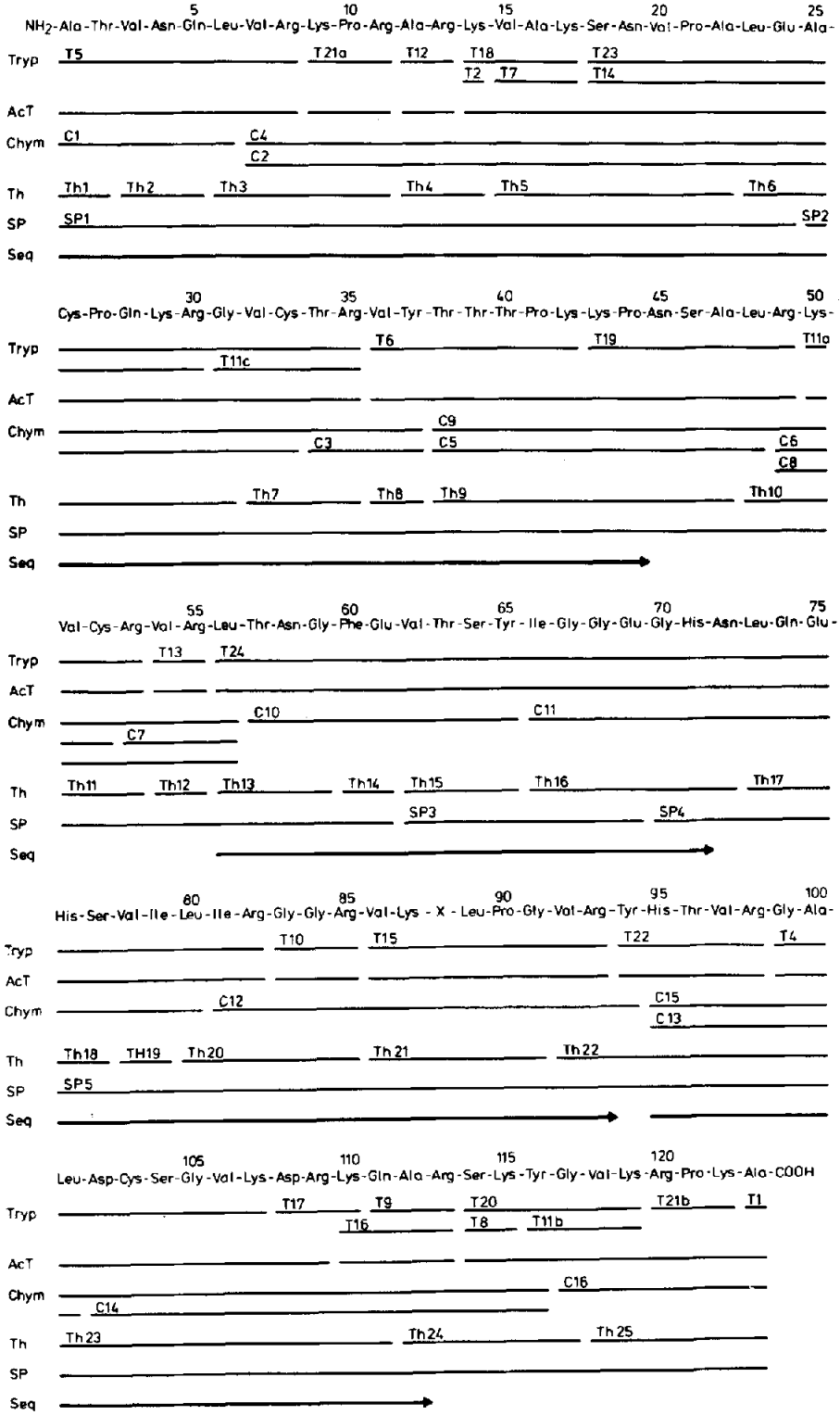
NH2-Ala-Thr-Val-Asn-Gln-Leu-Val-Arg-Lys-Pro-Arg-Ala-Arg-Lys-Val-Ala-Lys-Ser-Asn-Val-Pro-Ala-Leu-Glu-Ala-
Tryp
T5
T21a
T12
T18
T2
T7
T23
T14
AcT
Chym
C1
C4
C2
Th
Th1
Th2
Th3
Th4
Th5
Th6
SP
SP1
SP2
Seq
Cys-Pro-Gln-Lys-Arg-Gly-Val-Cys-Thr-Arg-Val-Tyr-Thr-Thr-Thr-Pro-Lys-Lys-Pro-Asn-Ser-Ala-Leu-Arg-Lys-
T6
T19
T11a
T11c
C9
C3
C5
C6
C8
Th7
Th8
Th9
Th10
Val-Cys-Arg-Val-Arg-Leu-Thr-Asn-Gly-Phe-Glu-Val-Thr-Ser-Tyr-Ile-Gly-Gly-Glu-Gly-His-Asn-Leu-Gln-Glu-
T13
T24
C10
C11
C7
Th11
Th12
Th13
Th14
Th15
Th16
Th17
SP3
SP4
His-Ser-Val-Ile-Leu-Ile-Arg-Gly-Gly-Arg-Val-Lys-X-Leu-Pro-Gly-Val-Arg-Tyr-His-Thr-Val-Arg-Gly-Ala-
T10
T15
T22
T4
C12
C15
C13
Th18
TH19
Th20
Th21
Th22
SP5
Leu-Asp-Cys-Ser-Gly-Val-Lys-Asp-Arg-Lys-Gln-Ala-Arg-Ser-Lys-Tyr-Gly-Val-Lys-Arg-Pro-Lys-Ala-COOH
T17
T9
T20
T21b
T1
T16
T8
T11b
C16
C14
Th23
Th24
Th25

Fig. 1

protein. Methionine and tryptophan are absent. The 31 basic and six acidic residues are compatible with the high isoelectric point of this protein [34]. Half of the basic residues form basic doublets such as Arg–Lys, Lys–Arg, Lys–Lys. Protein S12 has a high valine and cysteine content compared to the other ribosomal proteins in the small subunit [35,36]. If aspartic acid is used for the unknown X, as found in some mutants, the molecular weight of protein S12 is 13608.

Comparison of the sequence of S12 with other ribosomal proteins of known structures [37,38] revealed several identical regions which are shown in table 1.

Only a small amount of secondary structure can be predicted for protein S12. According to the method of Burgess et al. [39], positions 13–16 are helical and positions 26–29, 39–42, 51–54 and 99–103 have β-sheet conformation. According to the rules of Chou and Fasman [40,41] positions 10–17 and 71/73–76/77 are predicted to be α-helical structures. β-Sheet regions are predicted for positions 30–40, 48–56 and 77–81. Protein S12 must have an elongated conformation in the 30 S particle since by immuno-electronmicroscopy two antigenic determinants are found to be exposed at well-separated sites [42].

Amino acid sequence analyses have been done for many altered S12 proteins isolated from streptomycin resistant mutants [8,9], from streptomycin dependent mutants [10,11] and from neamine resistant mutants [13]. The results from these studies show that the amino acid replacements are restricted to only two regions of S12: Lys at position 42 and a short region from position 85–91. Amino acid replacements in these two regions drastically affect the fidelity of translation (see ref. [2] for more details). An amino acid replacement at position 87, but not at position 42, leads to activation of the capability of the ribosome to perform spontaneous non-enzymatic translocation [43]. From proteinchemical [9] and genetic [44] studies on the mutants with altered S12 proteins it is concluded that the gene for the protein S12 is transcribed in the counter clockwise direction on the *E. coli* chromosome [45].

Ribosomal protein S12 has also been isolated from *Bacillus stearothermophilus* [46–48] and *Bacillus subtilis* [49]. There is a high degree of homology

Table 1
Regions of protein S12 identical with regions of other *E. coli* ribosomal proteins

Peptide	Protein	Positions	Protein	Positions
Arg–Lys–Pro–Arg	S12	8–11	S4	43–46
Arg–Ala–Arg–Lys	S12	11–14	S11	6–9
Lys–Val–Ala–Lys	S12	14–17	L23	33–36
Lys–Arg–Gly–Val	S12	29–32	L20	4–7
Tyr–Thr–Thr–Thr	S12	37–40	L33	20–23
Leu–Ile–Arg–Gly	S12	80–83	L22	23–26
Val–Arg–Gly–Ala	S12	97–100	L7/L12	72–75
Gly–Val–Lys–Arg	S12	117–120	L27	21–24

Fig.1. Amino acid sequences of protein S12 from *E. coli* ribosomes. Tryp = tryptic peptide, Chym = chymotryptic peptide, Th = thermolytic peptide, SP = peptide from digestion with *Staphylococcus aureus* protease, AcT = peptides from trypsin digestion of acetylated protein S12, Seq = automatic sequencing by liquid- or solid-phase degradation in a sequenator.

among the N-terminal sequences of protein S12 from *E. coli* and the two bacilli [50]. In streptomycin mutants of *Bacillus stearothermophilus* [51] and *Bacillus subtilis* [49] the altered ribosomal proteins are homologous to *E. coli* protein S12.

Acknowledgements

We should like to thank Dr H. G. Wittmann for many stimulating discussions and M. Bergmann, J. Krauss, C. Roy and D. Semple for excellent technical assistance. We thank Dr W. N. Strickland for kindly reading the manuscript.

References

[1] Wittmann, H. G. (1974) in: Ribosomes, pp. 93–114, Cold Spring Harbor Monograph Series, Cold Spring Harbor, N.Y.
[2] Wittmann, H. G. and Wittmann-Liebold, B. (1974) in: Ribosomes, pp. 115–140, Cold Spring Harbor Monograph Series, Cold Spring Harbor, N.Y.
[3] Nomura, M., Mizushima, S., Ozaki, M., Traub, P. and Lowry, C. V. (1969) Cold Spring Harbor Symp. Quant. Biol. 34, 49–61.
[4] Held, W. A., Gette, W. R. and Nomura, M. (1974) Biochemistry 13, 2115–2122.
[5] Gorini, L. (1974) in: Ribosomes, pp. 791–803, Cold Spring Harbor Monograph Series, Cold Spring Harbor, N.Y.
[6] Ozaki, M., Mizushima, S. and Nomura, M. (1969) Nature 222, 333–339.
[7] Birge, E. A. and Kurland, C. G. (1969) Science 166, 1282–1284.
[8] Funatsu, G., Nierhaus, K. and Wittmann, H. G. (1972) Biochim. Biophys. Acta 287, 282–291.
[9] Funatsu, G. and Wittmann, H. G. (1972) J. Mol. Biol. 68, 547–550.
[10] Itoh, T. and Wittmann, H. G. (1973) Molec. Gen. Genet. 127, 19–32.
[11] Van Acken, U. (1975) Molec. Gen. Genet. 140, 61–68.
[12] De Wilde, M., Cabezon, T., Villarroel, R., Herzog, A. and Bollen, A. (1975) Molec. Gen. Genet. 142, 19–33.
[13] Yaguchi, M., Wittmann, H. G., Cabezon, T., De Wilde, M., Villarroel, R., Herzog, A. and Bollen, A. (1975) Molec. Gen. Genet. 142, 35–43.
[14] Wittmann-Liebold, B. (1973) FEBS Lett. 36, 247–249.
[15] Yaguchi, M., Roy, C., Matheson, A. T. and Visentin, L. P. (1973) Canadian J. Biochem. 51, 1215–1217.
[16] Hindennach, I., Stöffler, G. and Wittmann, H. G. (1971) Eur. J. Biochem. 23, 7–11.
[17] Kaltschmidt, E. and Wittmann, H. G. (1970) Analyt. Biochem. 36, 401–412.
[18] Hirs, C. H. W. (1956) J. Biol. Chem. 219, 611–621.
[19] Fraenkel-Conrat, H. (1957), Method in Enzym. IV, 247–269.
[20] Houmard, J. and Drapeau, G. R. (1972) Proc. Natl. Acad. Sci USA 69, 3506–3509:
[21] Wittmann-Liebold, B. and Wittmann, H. G. (1971) Biochim. Biophys. Acta 251, 44–53.
[22] Iwanaga, S., Wallen, P., Grondahl, N. J., Henschen, A. and Blombäck, B. (1969) Eur. J. Biochem. 8, 189–199.
[23] Bruton, C. J. and Hartley, B. S. (1970) J. Mol. Biol. 52, 165–178.
[24] Edman, P. and Henschen, A. (1975) in: Protein Sequence Determination, Needleman, S. B., ed) pp. 232–279, Springer-Verlag.
[25] Medez, E. and Lai, S. Y. (1975) Analyt. Biochem. 68, 47–53.
[26] Wittmann-Liebold, B., Geissler, A.-W., Marzinzig, E. (1975) J. Supramol. Struct. 3, 426–447.
[27] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[28] Wittmann-Liebold, B. (1973) Hoppe-Seyler's Z. Physiol. Chem. 354, 1415–1431.
[29] Wittmann-Liebold, B., Graffunder, H. and Kohls, H. (1976) Analyt. Biochem. 75, 621–633.
[30] Laursen, R. A. (1971) Eur. J. Biochem. 20, 89–102.
[31] Previero, A., Deroncourt, J., Coletti-Previero, M. A. and Laursen, R. A. (1973) FEBS Lett. 33, 135–138.
[32] Laursen, R. A., Horn, M. J. and Bonner, A. G. (1972) FEBS Lett. 21, 67–70.
[33] Wittmann-Liebold, B. and Lehmann, A. (1975) in: Solid Phase Methods in Protein Sequence Analysis (Laursen, R. A. ed.) pp. 81–90, Pierce Chemical Company, Rockford, USA.
[34] Kaltschmidt, E. (1971) Analyt. Biochem. 43, 25–31.
[35] Kaltschmidt, E., Dzionara, M. and Wittmann, H. G. (1970) Molec. Gen. Genet. 109, 292–297.
[36] Kahan, L., Held, W. A. and Nomura, M. (1974) J. Mol. Biol. 88, 797–808.
[37] Wittmann-Liebold, B. and Dzionara, M. (1976) FEBS Lett. 61, 14–19.
[38] Wittmann-Liebold, B. and Dzionara, M. (1976) FEBS Lett 65, 281–283.
[39] Burgess, A. W., Ponnuswamy, P. K. and Scheraga, H. A. (1974) Israel J. Chem. 12, 239–286.
[40] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 211–222.
[41] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 222–245.
[42] Tischendorf, G. W., Zeichhardt, H. and Stöffler, G. (1975) Proc. Natl. Acad. Sci. USA 72, 4820–4824.
[43] Asatryan, L. S. and Spirin, A. S. (1975) Molec. Gen. Genet. 138, 315–321.
[44] Breckenridge, L. and Gorini, L. (1970) Genetics 65, 9–25.

[45] Wittmann, H. G., Yaguchi, M., Piepersberg, W. and Böck, A. (1975) J. Mol. Biol. 98, 827–829.
[46] Higo, K., Held, W., Kahan, L. and Nomura, M. (1973) Proc. Natl. Acad. Sci. USA 70, 944–948.
[47] Isono, K., Isono, S., Stöffler, G., Visentin, L. P., Yaguchi, M. and Matheson, A. T. (1973) Molec. Gen. Genet. 127, 191–195.
[48] Isono, S. and Isono, K. (1975) Eur. J. Biochem. 50, 482–488.
[49] Itoh, T., Kosugi, H., Higo, K. and Osawa, S. (1975) Molec. Gen. Genet. 139, 293–301.
[50] Yaguchi, M., Matheson, A. T. and Visentin, L. P. (1974) FEBS Lett. 46, 296–300.
[51] Isono, K. (1974) Molec. Gen. Genet. 133, 77–86.