THE PRIMARY STRUCTURE OF RIBOSOMAL PROTEIN S7 FROM *ESCHERICHIA COLI* STRAINS K and B

Sequence of the C-terminal region of S7K and S7B

Denis TRITSCH and Joseph REINBOLT

Laboratoire de Biochimie, Institut de Biologie Moléculaire et Cellulaire, 15, Rue R. Descartes, 67084 Strasbourg Cédex, France

and

Brigitte WITTMANN-LIEBOLD

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

1. Introduction

Protein S7 from the small subunit of *Escherichia coli* ribosomes interacts specifically with the 16 S RNA (reviewed in reference [1]) and it has been shown that binding occurs near the 3'-end of the 16 S RNA [2–4]. Protein S7 can be cross-linked to 16 S RNA by ultraviolet irradiation of the small subunit [5]. From studies in which an in vitro complex between protein S7 and 16 S RNA was irradiated and digested with trypsin, it was suggested that four peptides were cross-linked to the RNA [6].

Knowledge of the primary structure of protein S7 is necessary both for detailed studies of the molecular mechanisms of protein–RNA interaction as well as the elucidation of the alterations of protein S7 recently found in various mutants of *E. coli* [7,8].

It has long been known that *E. coli* strains differ in their S7 proteins [9–13]. Ribosomes from strains B, C and MRE 600 contain a protein S7 which differs extensively in size, charge, amino acid composition and immunological properties from the protein S7 in strain K [13].

In this paper, we describe the sequence analysis of the C-terminal regions of proteins S7 isolated from *E. coli* strains B and K. We show that protein S7 K is longer by 24 amino acids than protein S7 B and that

this difference occurs at the C-terminal end of the protein chain.

2. Materials and methods

Protein S7 was provided by Dr H. G. Wittmann and was isolated from 30 S ribosomal subunits of *E. coli* K and B as described elsewhere [14]. Its purity was determined by two-dimensional gel electrophoresis [15].

Tryptic digestion was performed at 37°C for 3 h. The tryptic peptides and the peptides obtained by means of other endoproteases were separated on cellulose thin-layer plates and eluted [16]. Alternatively, they were isolated on a micro-column (0.3 × 10 cm) of Dowex M71 at 50°C with pyridine-formate gradients [17,18] followed, when necessary, by chromatography on cellulose thin-layer sheets. Protein S7 was also digested with pepsin and a *Staphylococcus aureus* protease which is specific for glutamoyl bonds [19]. Larger tryptic peptides were produced by first blocking the ε-amino groups of the lysine residues by exo-*cis*-3,6-endoxo-Δ4-tetrahydrophthalic acid anhydride so that trypsin would split only the arginyl bonds [20]. Carboxypeptidases A as well as B were used to establish or to confirm partial

sequences. In order to obtain large fragments, the
protein was also cleaved by treatment with cyanogen
bromide [21].

Alignment of the tryptic peptides was facilitated
by isolation and analysis of peptides obtained after
specific cleavage of the tryptophanoyl bonds of S7
with BNPS-skatole, 2-(2-nitrophenylsulphenyl)-3-methyl-
3' bromoindolenine [22]. For the determination of
tryptophan residues, proteins S7 K and S7 B were
hydrolysed with 4 N methane sulphonic acid containing
0.2% 3-(2-aminoethyl) indole [23]. The presence of
tryptophan in the peptides was established by finger-
prints on cellulose thin-layer plates, spraying with
Ehrlich's reagent.

The sequence determination of proteins S7 B and
S7 K were performed using the following three methods:

(i) Automatic Edman degradation in a solid-phase
sequenator according to the method of Laursen [24]
with attachment of the C-terminal carboxyl group to
amino-polystyrene resin [25,26];

(ii) Automatic Edman degradation [27] in an
improved Beckman sequenator [28];

(iii) Manual Edman degradation employing the
dansyl-Edman procedure [29,30].

3. Results and discussion

It is known that the molecular weight of S7 K is
higher than that of S7 B by approximately 15% [13].
Thus the question arose as to whether the two
proteins differ in their N-terminal or C-terminal regions.
We therefore determined the amino acid sequences
of the tryptic peptides in the N-terminal regions of
both proteins and no differences were detected. On
the other hand, we found that proteins S7 B and S7 K
differed strongly in the amino acids released by treat-
ment with carboxypeptidases A and B. The results show
that both proteins differ in their C-terminal, rather
than in their N-terminal region.

The next question was to differentiate between
the following two possibilities:

(i) The sequence of protein S7 B is identical with
that of S7 K, with the exception that approximately
25 amino acids at the C-terminal of S7 K are absent
in S7 B. This can be caused by a mutation from a
sense codon to a termination codon, leading to the
shorter S7 B protein chain; or vice versa, by a muta-

tion in the termination codon of S7 B to a sense
codon, resulting in the longer chain.

(ii) The different lengths of the two S7 proteins
can be caused by a frame-shift mutation in a similar
way to that found in *E. coli* mutants with altered
S4 proteins [31].

The consequences of these possibilities are that,
in the first case, all peptides of the shorter protein
(S7 B) should also be present in the longer protein
(S7 K), whereas, in the second case, the shorter
protein should also contain peptides which are
different from that of the longer protein.

In order to distinguish between these possibilities,
proteins S7 K and S7 B were digested with trypsin
and the tryptic peptides were separated by the finger-
print method. These were eluted and analysed after
hydrolysis. The finding that all the peptides present in
S7 B were also found in S7 K and that no new peptide
was detected can be easily explained by the first but
not by the second possibility mentioned above.

Further proof was obtained by treatment of the
two protein chains with the reagent BNPS-skatole
which cleaves after tryptophan residues. It was found
that treatment of protein S7 K resulted in one peptide
more than treatment of S7 B. This extra peptide
(BNPS-KII) from S7 K was isolated by gel-filtration on
Sephadex G-100 and sequenced in a solid-phase
sequenator. It was also digested with trypsin and, after
isolation of the resulting three tryptic peptides by
the fingerprint method, their sequences were deter-
mined by the manual dansyl-Edman method and
by the solid-phase sequenator. The results from all
methods were consistent and led to the sequence of
peptide BNPS-KII as shown in fig.1. The location of
the peptide BNPS-KII at the C-terminal end of protein
S7 K is also in full agreement with the experiments in
which peptide BNPS-KII and intact protein S7 K were
treated with carboxypeptidases A and B. The same
amino acids were liberated in both cases. This result
shows that BNPS-KII is situated at the C-terminal end
of protein S7 K.

The peptide at the C-terminus of protein S7 B was
identified in the following manner. After treatment
of S7 B with BNPS-skatole, a peptide, BNPS-BI, was
isolated by gel-filtration and was sequenced by a
combination of chemical and enzymatic methods as
illustrated in fig.1. Treatment of this peptide and of
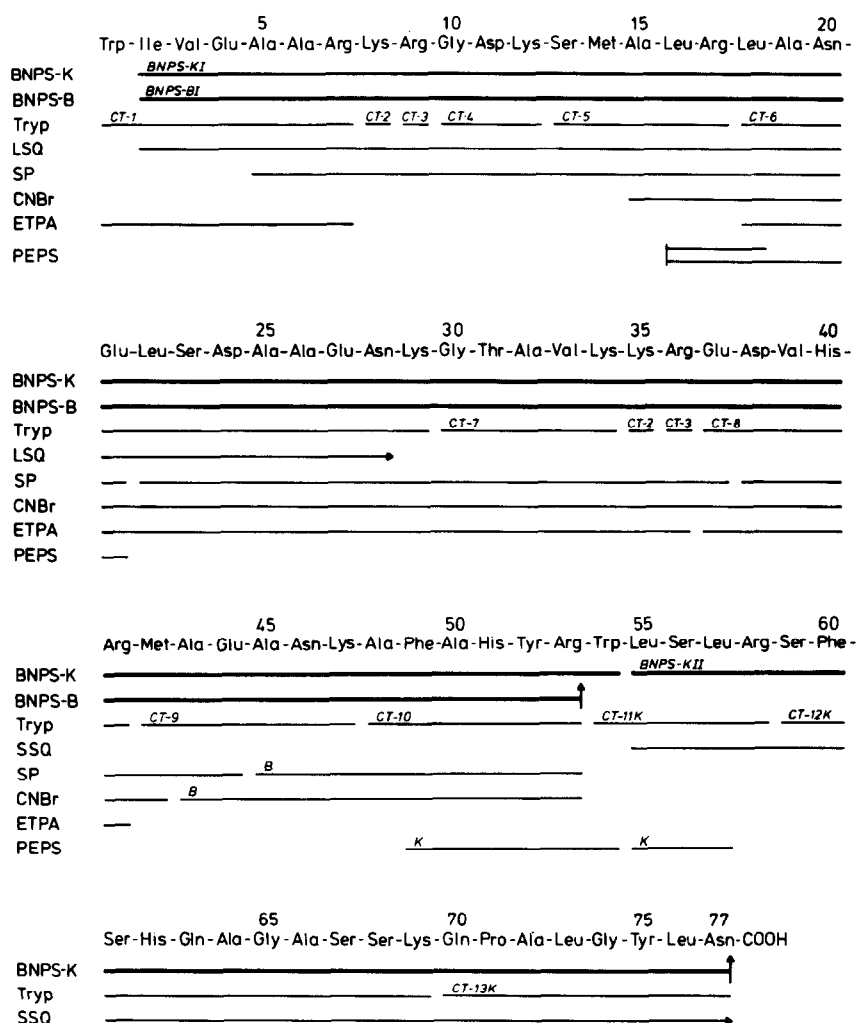the intact S7 B protein with carboxypeptidases A

Fig.1. Comparison of the C-terminal regions of proteins S7 K and S7 B. Abbreviations: BNPS-Skatole, peptides resulting from BNPS-skatole cleavage; Tryp, tryptic peptides; LSQ, Edman degradation in an improved Beckman sequenator including automatic conversion reaction; SP, peptides from digestion with *Staphylococcus aureus* protease; CNBr, peptides resulting from CNBr cleavage; ETPA, peptides obtained from a tryptic digest after blocking the amino groups by ETPA; Peps, peptides from digestion with pepsine; SSQ, sequence elucidated by solid-phase sequenator. The peptides specific to S7 B or S7 K are mentioned as 'B' or 'K'. All the other peptides are obtained from both proteins.

and B resulted in the release of the same amino acids. This finding demonstrates the location of peptide BNPS-BI at the C-terminus of protein S7 B.

Comparison of the amino acid sequence of peptide BNPS-BI isolated from protein S7 B with that of BNPS-KI from S7 K showed that both peptides have almost identical primary structures. They differ only in one tryptophan residue which is located at the

C-terminus of BNPS-KI but is absent in BNPS-BI. As shown in fig.1, which represents the terminal part of proteins S7 B and S7 K, both protein-chains are identical in this region with the exception that S7 K is longer by 24 amino acids than S7 B. All peptides found in S7 B are also present in S7 K, and no other peptide was found which differs in the two proteins. Hence, a frame-shift is improbable as the mutational

event which leads to the different lengths of the two proteins. Whether protein S7 B derived from S7 K by a mutation from a sense- to a termination-codon, or whether the reverse mechanism occurred cannot be decided with certainty. However, it should be borne in mind that ribosomes from all *E. coli* strains and from all other Enterobacteriaceae species studied so far [13,32,33] have a S7 protein similar to S7 B and not to S7 K. This finding would suggest that S7 K is derived from S7 B and not vice versa.

3.1. Secondary structure

Using the method of Chou and Fasman [34,35] for predicting secondary structure in proteins, the following results were obtained for the C-terminal region of protein S7 K:

(i) α-Helices in positions 1–6, 13–27, 31–50 and 59–64.

(ii) β-Turns in positions 10–12, 28–31, 51–54 and 65–68.

Following the method of Burgess et al. [36], α-helical regions were predicted for positions 4–8, 15–25, 32–37 and 40–49; bends for 9–13, 38–39, 52–53, 55–61, 66–69 and 72–73.

Both methods gave a very high α-helix content for the C-terminal region (positions 1–53) which is identical in proteins S7 B and S7 K: 89% according to Chou and Fasman [34,35] and 57% according to Burgess et al. [36]. Both methods predict the region from positions 15–50 to contain mainly α-helices.

Acknowledgements

References

[1] Zimmermann, R. A. (1974) in: Ribosomes, Monograph Series, pp. 225–269, Cold Spring Harbor Laboratory, Long Island, NY.

[2] Muto, A., Ehresmann, C., Fellner, P. and Zimmermann, R. A. (1974) J. Mol. Biol. 86, 411–432.

[3] 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 Acid Res. 2, 279–302.

[4] Rinke, J., Yuki, A. and Brimacombe, R. (1976) Eur. J. Biochem. 64, 77–89.

[5] Möller, K. and Brimacombe, R. (1975) Mol. Gen. Genet. 141, 343–355.

[6] Ehresmann, B., Reinbolt, J., Backendorf, C., Tritsch, D. and Ebel, J. P. (1976) FEBS Lett. 67, 316–319.

[7] Isono, K., Krauss, J. and Hirota, Y. (1976) Mol. Gen. Genet. 149, 297–302.

[8] Dabbs, E. R. and Wittmann, H. G. (1976) Mol. Gen. Genet. 149, 303–309.

[9] Leboy, P. S., Cox, E. C. and Flaks, J. G. (1964) Proc. Natl. Acad. Sci. USA 52, 1367–1374.

[10] Birge, E. A., Craven, G. R., Hardy, S. J. S., Kurland, C. G. and Voynow, P. (1969) Science 164, 1285–1286.

[11] Takata, R., Dekio, S., Otaka, E. and Osawa, S. (1969) Mol. Gen. Genet. 105, 113–121.

[12] Sypherd, P. S. (1969) J. Bact. 99, 379–382.

[13] Kaltschmidt, E., Stöffler, G., Dzionara, M. and Wittmann, H. G. (1970) Mol. Gen. Genet. 109, 303–308.

[14] Hindennach, I., Stöffler, G. and Wittmann, H. G. (1971) Eur. J. Biochem. 23, 7–11.

[15] Kaltschmidt, E. and Wittmann, H. G. (1970) Anal. Biochem. 36, 401–412.

[16] Bitar, K. G. and Wittmann-Liebold, B. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1343–1352.

[17] Chen, R. and Wittmann-Liebold, B. (1975) FEBS Lett. 52, 139–140.

[18] Heiland, I., Brauer, D. and Wittmann-Liebold, B. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 1751–1770.

[19] Houmard, J. and Drapeau, G. R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.

[20] Riley, M. and Perham, R. N. (1970) Biochem. J. 118, 733–739.

[21] Gross, E. and Witkop, B. (1962) J. Biol. Chem. 237, 1856–1860.

[22] Fontana, A. (1972) Methods in Enzymology XXV, 419–423.

[23] Liu, T.-Y. and Chang, Y. H. (1971) J. Biol. Chem. 246, 2842–2848.

[24] Laursen, R. A. (1971) Eur. J. Biochem. 20, 89–102.

[25] Previero, A., Derancourt, J., Coletti-Previero, M. A. and Laursen, R. A. (1975) FEBS Lett. 33, 135–138.

[26] 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.

[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] Gray, W. R. and Hartley, B. S. (1963) Biochem. J. 89, 379–380.

[30] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.

[31] Schiltz, E. (1974) cited by Wittmann, H. G. and Wittmann-Liebold, B. in: Ribosomes. Monograph Series, pp. 115–140. Cold Spring Harbor Laboratory, Long Island, NY.

[32] Osawa, S., Takata, R. and Dekio, S. (1970) Mol. Gen. Genet. 107, 32–38.

[33] Geisser, M., Tischendorf, G. W., Stöffler, G. and Wittmann, H. G. (1973) Mol. Gen. Genet. 127, 111–128.

[34] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 211–245.

[35] Chou, P. Y., Adler, A. J. and Fasman, G. D. (1975) J. Mol. Biol. 96, 29–45.

[36] Burgess, A. W., Ponnuswamy, P. K. and Scheraga, H. A. (1974) Israël J. Chem. 12, 239–286.