

The primary structure of protein L9 from the *Escherichia coli* ribosome

Roza Maria Kamp and Brigitte Wittmann-Liebold

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, D-1000 Berlin 33 (Dahlem), Germany

Received 14 September 1982

The primary structure of protein L9 from the large subunit of the *E. coli* ribosome has completely been determined. Automatic sequencing of the intact protein by means of an improved Beckman sequencer and by manual sequencing of diverse peptides deriving from cleavages with trypsin, thermolysin, chymotrypsin and pepsin by the DABITC/PITC double coupling method [FEBS Lett. (1978) 93, 205–214] have been performed to establish the amino acid sequence. Protein L9 contains 148 amino acid residues and has a molecular mass of 15 696. Predictions of secondary structural elements for this protein have been made. A strong homology exists between the primary structure of the *E. coli* protein L9 and a *B. stearothermophilus* ribosomal protein previously crystallized [FEBS Lett. (1979) 103, 66–70].

<i>E. coli</i> ribosome	<i>Large subunit protein L9</i>	<i>Primary structure determination</i>
<i>Secondary structure prediction</i>		<i>Homologies to other ribosomal proteins</i>

1. INTRODUCTION

Protein L9 is a component of the large subunit of the *E. coli* ribosome [1]. It has an isoelectric point of 6.4 and is less basic than most of the other ribosomal proteins [2]. Its shape is slightly elongated with an axial ratio of ca. 5:1 [3]. As found by crosslinking experiments protein L9 is a neighbor of proteins L2 [4,5], L10 and/or L11 [6] and L28 [7] within the ribosomal particle. It binds specifically to the 23 S RNA [8], and is present together with protein L1 [9,10] and 4 other proteins [10] in small RNP fragments isolated after controlled nuclease digestion of the intact ribosomal subunit. The gene coding for protein L9 is located at 94 min on the *E. coli* chromosome [11,12].

Here the complete primary structure of protein L9 is described. The protein consists of 148 amino acid residues and has M_r 15 696. Predictions of secondary structural elements are made employing 4 different prediction programmes. Further, the amino acid sequence of this protein is compared with those of all other ribosomal proteins.

2. MATERIALS AND METHODS

Protein L9 was isolated from 50 S subunits of *E. coli* K12, strain A19, as in [13] or [14]. The protein samples were provided by Drs H.G. Wittmann, J. Dognin and J. Dijk.

For the sequencing studies peptides were isolated applying the following enzymatic hydrolyses:

- (i) Digestion with trypsin at pH 8.1 (in 0.1 M *N*-methylmorpholine acetate buffer, for 4 h at 37°C; enzyme:substrate ratio of 1 : 50 with TPCK-trypsin from Worthington).
- (ii) Digestion with thermolysin of the protein and some tryptic peptides (same buffer as above, for 2 h at 52°C; enzyme:substrate ratio of 1 : 100 with thermolysin from Serva).
- (iii) Digestion with chymotrypsin of the protein and some tryptic peptides (same buffer for 4 h at 37°C; enzyme:substrate ratio of 1 : 100 with α -chymotrypsin, treated with TLCK from Merck).

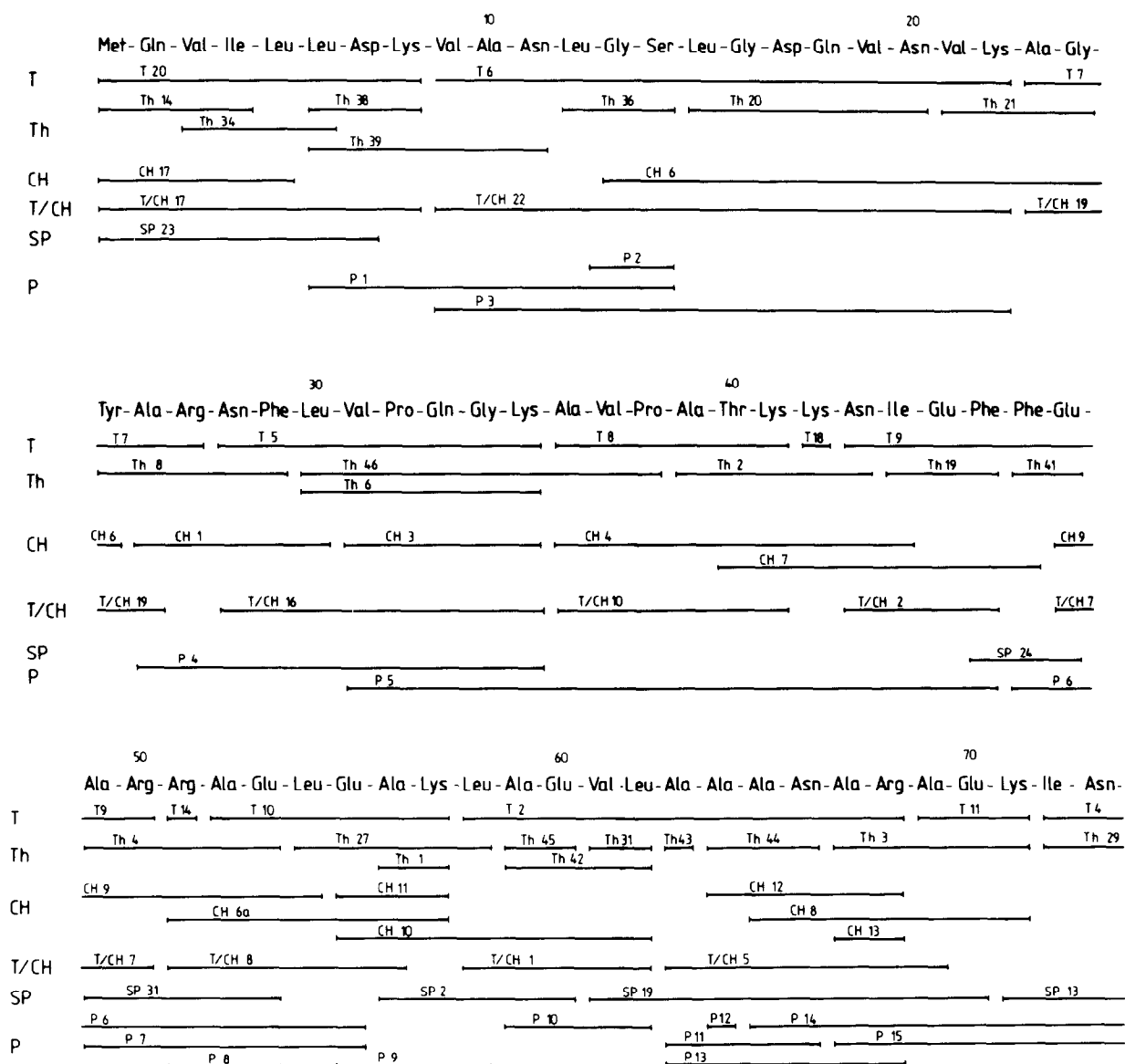


Fig. 1. (continued)

Fig. 1. Sequence determination of protein L9 from *E. coli* ribosomes: T, cleavage of protein with trypsin; Th, thermolysin; CH, α -chymotrypsin; P, pepsin; T/CH, digestion of tryptic peptides with chymotrypsin; T/SP, digestion of tryptic peptides with *Staphylococcus aureus* protease.

- (iv) Digestion with *Staphylococcus aureus* protease (SP-enzyme) at pH 8.0 (in 50 mM *N*-methylmorpholine acetate buffer for 26 h at 37°C; enzyme:substrate ratio of 1 : 50 with protease V-8 from Miles).
- (v) Digestion with pepsin at pH 2.0 (in 0.05 M acetic acid for 2 h at 37°C; enzyme:substrate ratio of 1 : 100 with pepsin from Serva).

The peptides were purified employing the following methods:

- (i) Thin-layer fingerprinting [15] applied for peptides cleaved with trypsin, thermolysin and chymotrypsin as well as for tryptic peptides further digested with thermolysin, chymotrypsin and SP-enzyme.

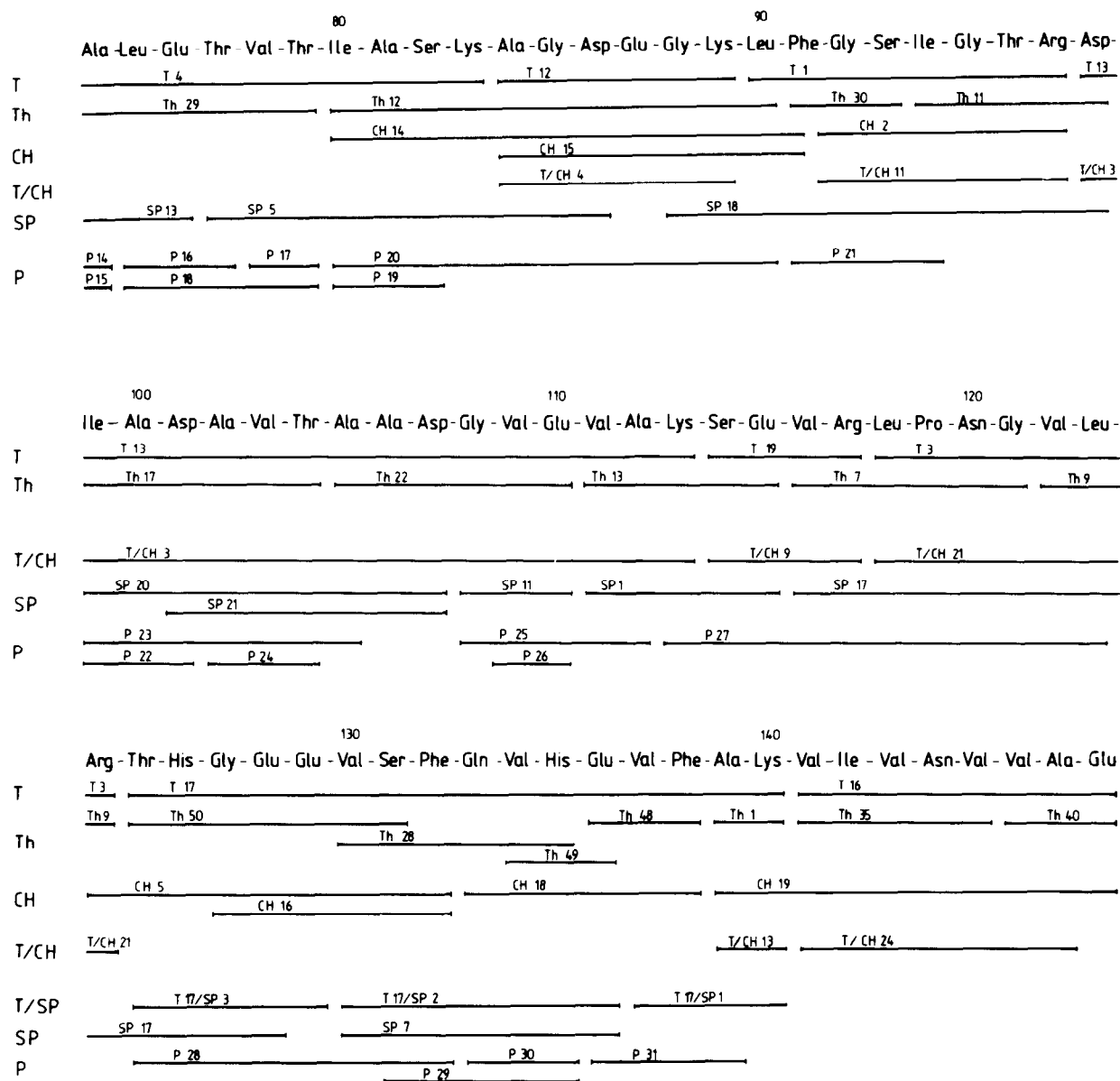


Fig. 1. (concluded)

- (ii) Ion-exchange chromatography on a Dowex 50 (M71) micro-column (90 × 2 mm) at 50°C in dilute pyridine/formate gradients for tryptic peptides.
- (iii) Gel filtration on Sephadex G-50, s.f. (1 × 200 cm) in 0.07% ammonia in water for tryptic and chymotryptic peptides and on Sephadex G-25 (same column) in 10% acetic acid for tryptic peptides as in [16].

- (iv) High-performance liquid chromatography on reversed phase employing:
 - (a) RP-C18, 5 μm (ODS Hypersil from Shandon, London) in a linear gradient of dilute ammonium formate buffer at pH 7.8 and methanol as organic modifier (for tryptic peptides); and
 - (b) RP-C18, 10 μm (from Merck, Darmstadt) in a similar gradient at pH 4.4 (for SP-

and peptic peptides). Some of the SP-peptides were purified on RP-C18, 10 μ m at pH 2.0 in a 0.05% TFA/water-methanol gradient.

All HPLC-separations were performed in the Liquid Chromatograph 850 (from Du PONT Instruments, Wilmington DW) at 40°C and flowrates of 1 ml/min. The columns were packed in the laboratory by the aid of a Shandon filling apparatus, and the peptides were detected by UV absorbance at 220 nm. The buffers employed contained only traces of salts which enabled short dryings of the peptide fractions (3–4 h in a Speed Vac Concentrator) and direct micro-sequencing without problems with solubilities of hydrophobic peptides. Details of the peptide separations by HPLC-techniques are given in [17].

Sequencing of peptides was performed by the following techniques:

- (i) Micro dansyl-Edman method [18];
- (ii) DABITC/PITC double-coupling method [19];
- (iii) Solid-phase degradation using the DABITC-reagent [20].

Amino acid analyses were performed in a Durum D-500 analyzer with amounts of 1–2 nmol peptides; protein analyses were carried out with 20–50 μ g. Details of the conventional procedures are reviewed in [21,22].

3. RESULTS AND DISCUSSION

3.1. Sequence determination

The N-terminal sequence of protein L9 up to position 43 was established by liquid-phase Edman degradation performed in a modified sequencer as detailed [22,23]. This N-terminal sequence was confirmed by the isolation and analysis of peptides obtained from digestions with trypsin, thermolysin and chymotrypsin, as shown in fig. 1.

The middle region of the molecule was difficult to sequence due to a repetitive area (pos. 48–76) which contained 10 Ala, 6 Glu, 4 Leu and 3 Arg residues out of 29 amino acids. The sequence was established by sequencing various smaller-sized peptides as shown in fig. 1. The alignment of the tryptic peptides T9–T14–T10–T2–T11–T4–T12–T1 was confirmed by sequencing the thermolysin peptides Th4 (pos. 49–53), Th27 (54–58),

Th3 (67–71), and Th12 (80–90) as well as one SP-fragment SP13 (71–76).

The remainder of the tryptic peptides were aligned in the C-terminal region T1–T13–T19–T3–T17–T16 by sequencing the following fragments: Th11 (94–98), Th13 (111–115), SP17 (116–128) and CH19 (139–148).

Sequencing of the peptides T16, CH19 and T/CH24 gave the residues up to position 147 with Val–Val–Ala. The C-terminal glutamic acid was identified by amino acid analysis of these peptides. Carboxypeptidase digestion of the intact protein released more alanine and valine than glutamic acid. The sequence Val–Val–Ala–Glu at the C-terminus is in agreement with that deduced from the nucleotide sequence analysis of the gene coding for this protein (J. Schnier and K. Isono, personal communication).

3.2. Characterization of the primary structure

According to the sequence determination, protein L9 has 148 residues and the amino acid composition: Asp₆, Asn₈, Thr₆, Ser₅, Glu₁₄, Gln₄, Pro₃, Gly₁₁, Ala₂₄, Val₂₀, Met₁, Ile₇, Leu₁₂, Tyr₁, Phe₆, His₂, Lys₁₁, Arg₇. Cysteine and tryptophan are absent. The M_r , as calculated on the basis of the sequence is 15 696.

Protein L9 contains 20 acidic residues, 14 of which are located in the middle region (pos. 45–115). The N- and C-terminal parts (pos. 1–11 and 130–148) are rich in hydrophobic residues. The C-terminus contains both histidines whereas the single methionine and tyrosine residues are present in the N-terminal region.

3.3. Secondary structure predictions of protein L9

Based on the primary structure, prediction of secondary structural elements were made for this protein employing 4 different computer assisted algorithms as in [24]. More recent parameters were used in the calculations according to Chou and Fasman, namely for turns [25] and for α -helices and β -sheets [26], employing a new FORTRAN programme [27]. The prediction results are shown in fig. 2. Agreement among the results with the 4 algorithms was found for pos. 42–71/76 (one strong helix) and at positions 12–17, 22–25, 32–35, 86–89, 91–97 and 119–121 (turns). The N- and C-termini of protein L9 are likely to be folded in β -sheets.

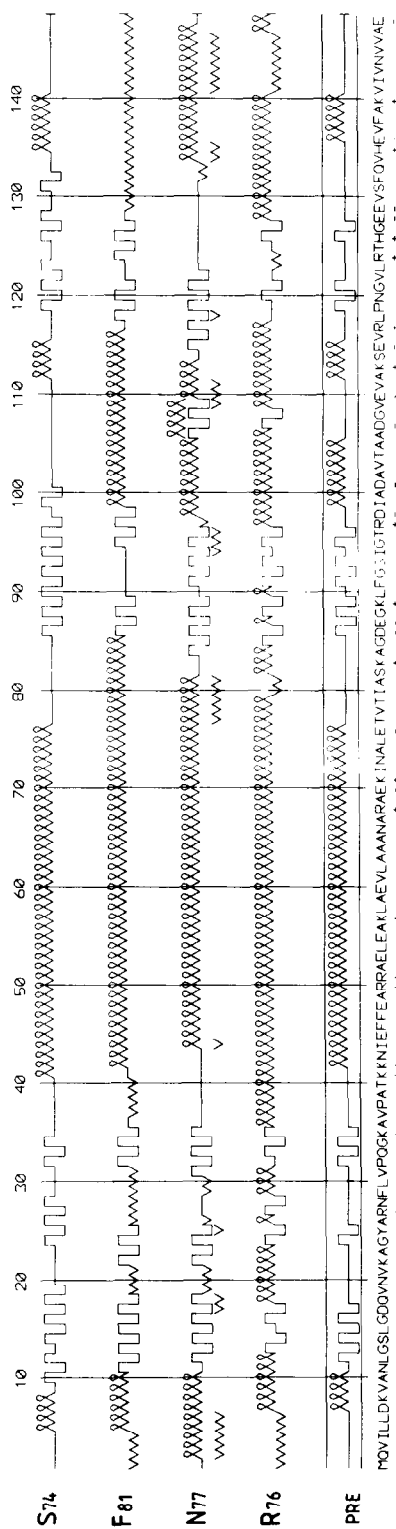


Fig. 2. Secondary structure predictions of protein L9. The predictions were made according to [29], S; [25-27], F; [30], N; [31], R; as detailed in [24]; (xxx) α -helix; (vvv) β -sheet; (---) random coil conformation.

The percentages of predicted α -helix, β -sheet and turn structures are given in table 1. According to the prediction of Chou and Fasman 35–45% α -

Table 1

Percentage values of predicted conformations in protein L9

Prediction according to:	α -Helix (%)	β -Sheet (%)	Turn (%)	Unpredicted (%)
[29]	34	0	28	39
[25,26] ^a	35–46	25–36	22	7
[30]	35–52	6–25	20–27	18
[31]	67	11	18	5

^a Computed by a programme according to [27]

EL 9	Met-Gln-Val-Ile-Leu-Leu-Asp-Lys-Val ¹⁰ -Asn-Leu-Gly-Ser-Leu-Gly-	
BL 17	Met-Lys-Val-Ile-Phe-Leu-Lys-Asp-Val-Lys-Gly-Lys-Gly-Lys-Lys-Gly-	
	... *** ... **	
EL 9	Asp-Gln-Val ²⁰ -Asn-Val-Lys-Ala-Gly-Tyr-Ala-Arg-Asn-Phe-Leu-Val-Pro-	
BL 17	Glu-Ile-Lys-Asn-Val-Ala-Asp-Gly-Tyr-Ala- - -Asn-Phe-Leu-Phe-Lys-	
	... **	
EL 9	Gln-Gly-Lys-Ala-Val-Pro-Ala ⁴⁰ -Thr-Lys-Lys-Asn-Ile-Glu-Phe-Phe-Glu-	
BL 17	Gln-Gly-Leu-Ala-Ile-Glu-Ala-Thr-Pro-Ala-Asn-Leu-Lys-Ala-Leu-Glu-	
	... **	
EL 9	Ala-Arg-Arg-Ala-Glu-Ile-Glu-Ala-Lys-Leu-Ala ⁶⁰ -Glu-Val-Leu-Ala-Ala-	
BL 17	Ala-Gln-Lys- - -Gln-Lys-Glu-Gln-Arg-Gln-Ala-Ala-Glu-Leu-Ala-Asn-	
	... **	
EL 9	Ala-Asn-Ala-Arg-Ala- - - -Glu-Lys-Ile-Asn-Ala-Leu-Glu-Thr-	
BL 17	Ala-Lys-Lys-Leu-Lys-Glu-Gln-Leu-Lys-Leu- - - - -Thr-	
	... **	
EL 9	Val-Thr-Ile ⁸⁰ -Ala-Ser-Lys-Ala-Gly-Asp-Glu-Gly-Lys ⁹⁰ -Leu-Phe-Gly-Ser-	
BL 17	Val-Thr-Ile-Pro-Ala-Lys-Ala-Gly-Glu-Gly-Gly-Arg-Leu-Phe-Gly-Ser-	
	... **	
EL 9	Ile-Gly-Thr-Arg-Asp-Ile ¹⁰⁰ -Ala-Asp-Ala-Val-Thr-Ala-Ala-Asp-Gly-Val-	
BL 17	Ile-Thr-Ser-Lys-Gln-Ile-Ala-Glu-Ser-Leu-Gln-Ala-Gln-His-Gly-Leu-	
	... **	
EL 9	Glu-Val-Ala-Lys-Ser-Glu-Val ¹¹⁰ -Arg-Leu-Pro-Asn-Gly-Val-Leu-Arg-Thr-	
BL 17	Lys-Leu-Asp-Lys-Arg-Lys-Ile-Glu-Leu-Ala-Asp- - -Ala-Ile-Arg-Ala-	
	... **	
EL 9	His-Gly-Glu-Glu- - -Val-Ser-Phe-Gln-Val-His- - -Glu-Val-Phe-Ala-	
BL 17	Leu-Gly-Tyr-Thr-Asn-Val-Pro-Val-Lys-Leu-His-Pro-Glu-Val-Thr-Ala-	
	... **	
EL 9	Lys-Val-Ile ¹⁴⁰ -Val-Asn-Val-Val-Ala-Glu	
BL 17	Thr-Leu-Lys-Val-His-Val-Thr-Glu-Gln-Lys	
	... **	

Fig. 3. Comparison of protein L9 from *E. coli* (EL9) and of protein L17 from *Bacillus stearothermophilus* (BL17) [28]: (+) identical residues in both proteins; (*) amino acids whose codons differ by one nucleotide.

helix and 25–35% β -sheet were calculated for protein L9. Ambiguous conformations were predicted for pos. 5–9 and 72–82 where α -helix and β -sheet were calculated with similar probability values (not shown in fig. 2).

3.4 Search for homologous regions

Comparison of the sequence of protein L9 with that of all other sequenced ribosomal proteins was made. The best homology found was to protein BL17 of *Bacillus stearothermophilus* ribosomes [28] (fig. 3). Other homologies were less pronounced and will be discussed elsewhere.

ACKNOWLEDGEMENTS

We thank Drs H.G. Wittmann, J. Dijk and J. Dognin for providing protein samples of L9. The technical assistance of Mr A. Bosserhoff and D. Kamp is greatly acknowledged.

REFERENCES

- [1] Kaltschmidt, E. and Wittmann, H.G. (1970) Proc. Natl. Acad. Sci. USA 67, 1276–1282.
- [2] Kaltschmidt, E. (1971) Analyt. Biochem. 43, 25–31.
- [3] Giri, L., Franz, A. and Dijk, J. (1979) Biochem. 18, 2520–2525.
- [4] Expert-Bezancon, A., Barritault, D., Clegg, J.C.S., Milet, M., Kourine, X. and Hayes, D. (1975) FEBS Lett. 59, 64–69.
- [5] Kenny, J.W. and Traut, R.R. (1979) J. Mol. Biol. 127, 243–263.
- [6] Clegg, C. and Hayes, D. (1974) Eur. J. Biochem. 42, 21–28.
- [7] Traut, R.R., Lambert, J.M., Boileau, G. and Kenny, J.W. (1979) in: Ribosomes (Chambliss G. et al. eds) pp. 89–110, Univ. Park Press, Baltimore, MD.
- [8] Marquardt, O., Roth, H.E., Wystup, G. and Nierhaus, K.H. (1979) Nucleic Acids Res. 6, 3641–3650.
- [9] Newton, I. and Brimacombe, R. (1974) Eur. J. Biochem. 48, 513–518.
- [10] Roth, H.E. and Nierhaus, K.H. (1975) J. Mol. Biol. 94, 111–121.
- [11] Isono, K. and Kitakawa, M. (1977) Mol. Gen. Genet. 153, 115–120.
- [12] Isono, K. and Kitakawa, M. (1978) Proc. Natl. Acad. Sci. USA 75, 6163–6167.

- [13] Hindennach, I., Kaltschmidt, E. and Wittmann, H.G. (1971) *Eur. J. Biochem.* 23, 12–16.
- [14] Dijk, J. and Littlechild, J. (1979) *Methods Enzymol.* 59, 481–502.
- [15] Hitz, H., Schäfer, D. and Wittmann-Liebold, B. (1977) *Eur. J. Biochem.* 75, 497–512.
- [16] Heiland, I., Brauer, D. and Wittmann-Liebold (1976) *Hoppe-Seyler's Z. Physiol. Chem.* 357, 1751–1770.
- [17] Kamp, R., Yao, Z.J. and Wittmann-Liebold, B. (1982) submitted.
- [18] Chen, R. (1976) *Hoppe-Seyler's Z. Physiol. Chem.* 357, 873–886.
- [19] Chang, J.Y., Brauer, D. and Wittmann-Liebold, B. (1978) *FEBS Lett.* 93, 205–214.
- [20] Salnikow, J., Lehmann, A. and Wittmann-Liebold, B. (1981) *Analyt. Biochem.* 117, 433–442.
- [21] Wittmann-Liebold, B. and Lehmann, A. (1980) in: *Methods in Peptide and Protein Sequence Analysis* (Birrr, Ch. ed) pp. 49–72, Elsevier Biomedical, Amsterdam, New York.
- [22] Wittmann-Liebold, B. (1981) in: *Chemical Synthesis and Sequencing of Peptides and Proteins* (Liu, T.Y. et al. eds) pp. 75–110, Elsevier Biomedical, Amsterdam, New York.
- [23] Wittmann-Liebold, B., Geissler, A.W. and Marzinzig, E. (1975) *J. Supramol. Struct.* 3, 426–447.
- [24] Dzionara, M., Robinson, S.M.L. and Wittmann-Liebold, B. (1977) *Hoppe-Seyler's Z. Physiol. Chem.* 358, 1003–1019.
- [25] Chou, P.Y. and Fasman, G.D. (1977) *J. Mol. Biol.* 115, 135–145.
- [26] Chou, P.Y. and Fasman, G.D. (1978) *Annu. Rev. Biochem.* 47, 251–276.
- [27] Rawlings, N., Ashman, K. and Wittmann-Liebold, B. (1982) submitted.
- [28] Kimura, M., Dijk, J. and Heiland, I. (1980) *FEBS Lett.* 121, 323–326.
- [29] Burgess, A.W., Ponnuswamy, P.K. and Scheraga, H.A. (1974) *Israel J. Chem.* 12, 239–286.
- [30] Nagano, K. (1977) *J. Mol. Biol.* 109, 251–274.
- [31] Robson, B. and Suzuki, E. (1976) *J. Mol. Biol.* 107, 327–356.