# PRIMARY STRUCTURE OF PROTEIN L24 FROM THE *ESCHERICHIA COLI* RIBOSOME

Brigitte WITTMANN-LIEBOLD
*Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, 1000 Berlin 33 (Dahlem), Germany*

Received 2 October 1979

## 1. Introduction

Protein L24 from the large subunit of the *Escherichia coli* ribosome belongs to a group of 6 proteins (L3, L13, L17, L21, L22 and L24) which are present in a ribonucleoprotein core obtained after controlled trypsin and ribonuclease digestion of 50 S subunits [1]. A similar group of 6 proteins (L3, L4, L13, L20, L22 and L24) are important for the formation of an early intermediate in the course of the assembly of the 50 S subunit [2]. A detailed analysis demonstrated that L24 plays no role in the late assembly and is not involved in the functions of the mature 50 S subunit [3].

The binding of protein L24 to 23 S RNA [4] has been studied in some detail [5], and it has been found that it binds to the first 550 nucleotides at the 5′-end of the 23 S RNA [6,7].

As determined by hydrodynamic studies protein L24 has a globular shape [8]. According to immunoelectron microscopy studies protein L24 is situated in the angle formed by the central and the right lateral protuberances of the large subunit in close proximity to the binding sites of proteins L32 and L33 [9].

In this paper the complete primary structure of protein L24 is given together with predictions of its secondary structure. The methods which led to the sequence elucidation of this protein are briefly summarized. Furthermore, a comparison of the L24 sequence with known structures of other ribosomal proteins is made. The knowledge of the primary structure of protein L24 is the basis for further protein chemical analyses, e.g., on mutants which show alterations in protein L24 [10,11].

## 2. Materials and methods

Protein L24 was isolated from 50 S subunits of *E. coli* strain K12 as in [12]. The protein samples were provided by Dr H. G. Wittmann. The purity of the protein was checked by two-dimensional gel electrophoresis [13].

Tryptic digestions of this protein were performed in 0.1 M *N*-methyl-morpholine acetate buffer (pH 8.1) at 37°C for 12–24 h. The lyophilized digest was dissolved in 10% acetic acid and the soluble peptides were separated by:

(i) Two-dimensional fingerprint technique on thin-layer sheets [14];
(ii) By column chromatography on SE-cellulose (30 × 1.5 cm) [15];
(iii) On Sephadex G-50, superfine (120 × 1 cm) in 10% acetic acid [16];
(iv) On a Dowex 50 (M71) micro column (90 × 2 mm) at 50°C in dilute pyridine formate gradients [16,17].

The peptides in the fractions of the columns were detected by thin-layer chromatography of appropriate amounts. The insoluble peptide moiety was dissolved in 70% formic acid and purified by gel filtration on Sephadex G-50, superfine (120 × 1 cm) in 50% formic acid. Detection of the peptides in these fractions was achieved by amino acid analysis.

The cleavage of protein L24 by *Staphylococcus aureus* protease was performed in 0.1 M ammonium acetate buffer (pH 4.0) for 48 h at 37°C. The peptides were chromatographed on Sephadex G-50, superfine (120 × 1 cm) in 10% acetic acid and purified by one-dimensional thin-layer chromatography. The large pep-

tides SP-2 and SP-3 were further digested with trypsin and thermolysin (5 h at 37°C and 2 h at 45°C, respectively) and the fragments were isolated by thin-layer fingerprint technique [14,16].

Digestion of protein L24 with pepsin was made in 0.05 N hydrochloric acid at 37°C for 4 h. The peptides were purified by a combination of column chromatography on a Dowex 50 micro-column (90 × 2 mm) at 50°C and one-dimensional preparative thin-layer chromatography or thin-layer fingerprints.

Sequencing of the diverse peptides was performed by 4 different Edman degradation techniques:

(i) Subtractive Edman method;
(ii) Micro dansyl-Edman technique [18];
(iii) Solid-phase degradation on amino polystyrene [19] as described in [20];
(iv) Peptides isolated in only 5–10 nmol amounts were sequenced by the DABITC/PITC (4-*NN*-dimethylaminoazobenzene-4′-isothiocyanate/phenyl-isothiocyanate) double-coupling method [21].

Amino acid analyses were made in an Unichrom (Fa. Beckman) and in a Durrum D-500 analyser. The strategy and methods applied for the sequence elucidation of protein L24 are summarized in [22].

## 3. Results and discussion

### 3.1. *Sequence determination*

The N-terminal region of protein L24 up to position 55 was established [23,24] by liquid-phase Edman degradation [25] employing a considerably modified Beckman sequencer as in [23,24,26]. This N-terminal sequence was confirmed by the isolation of peptides obtained from digestions with trypsin, pepsin and *Staphylococcus aureus* protease, as shown in fig.1. In addition, the small peptides were sequenced by a modified dansyl-Edman technique [18], and the following sequences were established by the solid-phase degradation technique [19,20,22]: P7 (pos. 1–10), T9 (pos. 26–32), P17 (pos. 29–36), and T10 (pos. 33–42). The recently developed DABITC/PITC double-coupling method [21] was used for confirming the sequence of peptide T9 and T10 and for sequencing peptide T11a (pos. 44–60).

The C-terminal part of protein L24 consists of the tryptic peptides T12 (pos. 61–78), T13 (pos. 79–81), T14 (pos. 82–85), T17 (pos. 86–91), T18 (pos. 92–93), T19 (pos. 94–96) and T20 (pos. 97–103). Their sequences were obtained by dansyl-Edman degradation or solid-phase sequencing with the exception of peptide T12 whose sequence was derived from the DABITC/PITC double-coupling method [21], see fig.1. The order of these C-terminal tryptic peptides was established as shown in fig.1 by sequencing the peptic peptides P8 (60–63, solid-phase method), P2a, P10, P2 (pos. 64–72, dansyl-Edman method), P5 (pos. 95–103, solid phase method) and by partially sequencing peptide P9 (pos. 73–84) and P4 (86–94) with the dansyl-Edman method. For further confirmation, two thermolysin peptides, pos. 74–81 and 84–91, were isolated and sequenced by the DABITC/PITC method.

By treatment of protein L24 with *Staphylococcus aureus* protease, cleavage occurs at positions 9, 36, 61 and 87. The similar sized peptides SP-2 (pos. 10–36, 27 residues), SP-3 (37–61, 25 residues) and SP-4 (62–87, 26 residues) were purified for amino acid analysis and for allowing further cleavage of SP-2 by trypsin and of SP-3 by trypsin and thermolysin. The resulting fragments were isolated by thin-layer fingerprints and analysed. The thermolysin peptides, pos. 51–56 and pos. 57–61 were sequenced by the dansyl-Edman-method.

Recently, a thermosensitive mutant having an altered L24 ribosomal protein was analysed and the replacement of one glycine residue by aspartic acid was reported [29]. In comparing the sequenced area of the mutant, namely Ala–Asp–Arg–Val–Asp–Phe–Arg with the completed sequence presented in fig.1 it becomes obvious that the amino acid exchange is located at pos. 83 of the protein chain.

### 3.2. *Characterization of the sequence*

Three independent sets of peptides deriving from trypsin, pepsin and *Staphylococcus aureus* protease were isolated in order to span the entire sequence of protein L24, as shown in fig.1. The amino acid composition according to this sequence is $Asp_5$ $Asn_6$ $Thr_3$ $Ser_5$ $Glu_6$ $Gln_3$ $Pro_3$ $Gly_{10}$ $Ala_9$ $Val_{13}$ $Met_0$ $Ile_8$ $Leu_4$ $Tyr_0$ $Phe_5$ $His_1$ $Lys_{16}$ $Arg_6$ $Trp_0$ which is in agreement with recent results obtained from amino acid analysis of the entire protein. Methionine and tyrosin, present in small amounts in early analyses of this protein [27], could not be found in the peptides described

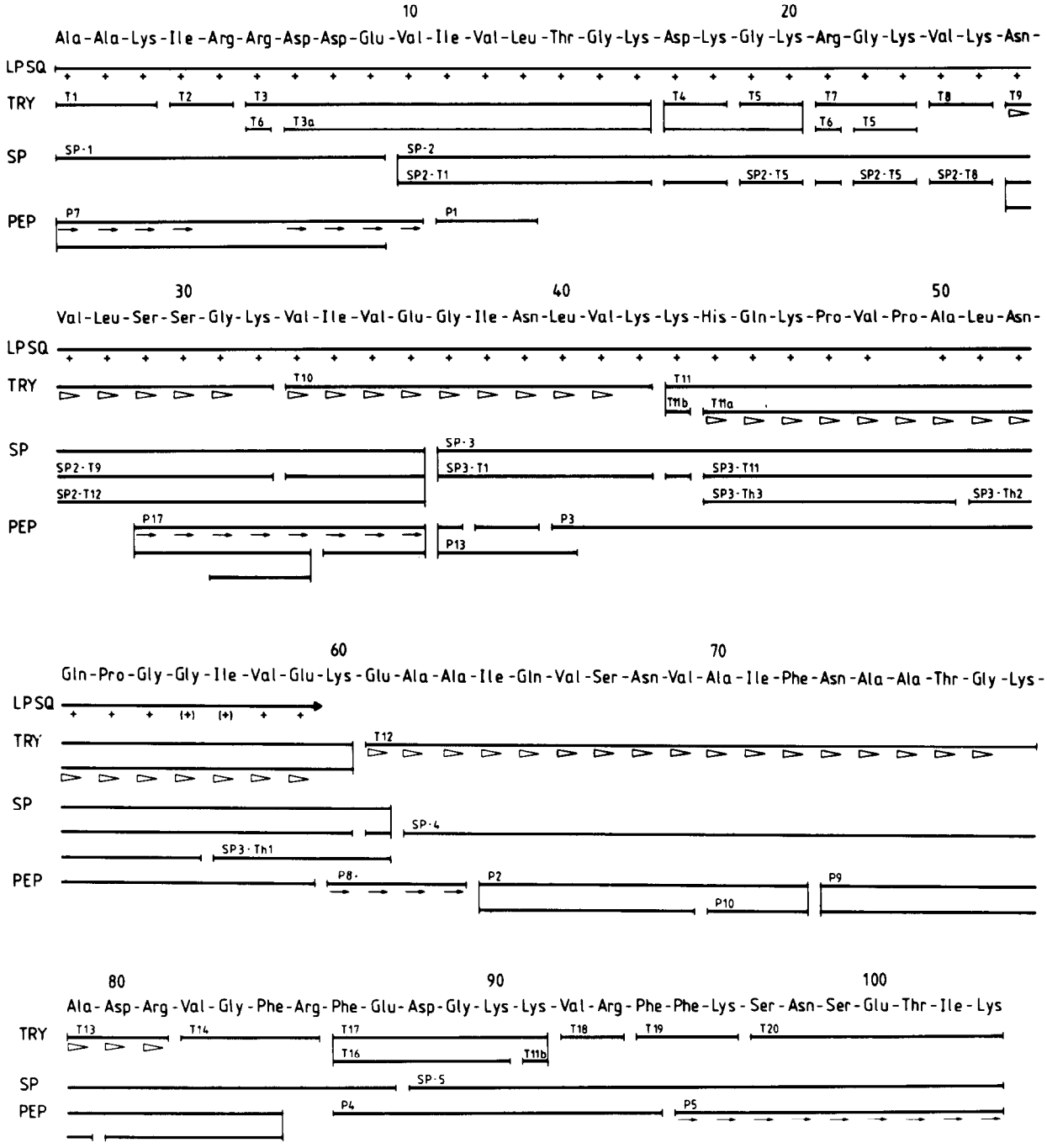


Fig.1. Primary structure of protein L24 from *Escherichia coli* ribosomes. TRY, trypsin digestion; SP, digestion with *Staphylococcus aureus* protease; PEP, pepsin digestion; thermolysin digestion, THE. LPSQ: liquid-phase Edman degradation in a modified Beckman sequencer; (+) means positively identified by thin-layer chromatography; (+) means weak spot; >, DABITC/PITC double-coupling method; →, automated solid-phase sequencing.

above. The molecular weight based on the sequence of protein L24 is 11 185 which is considerably lower than the value derived by earlier molecular weight determination by SDS–gel electrophoresis [28]. Such a discrepancy has also been experienced for other basic proteins. In recent sedimentation equilibrium experiments a value of 12 000 has been obtained [8] which is in full agreement with the molecular weight derived from the sequence.

The N-terminal and C-terminal regions of protein L24 are rather rich in charged amino acids leaving two neutral stretches with a higher number of hydrophobic amino acids for the middle part of the protein chain (in pos. 47–58 and 62–75). The 3 proline residues present in this protein are clustered in one of these areas (in pos. 47–54). The 5 phenylalanine residues are located in the C-terminal one-third of protein chain. Notable are furthermore the highly charged regions Lys–Ile–Arg–Arg–Asp–Asp–Glu (pos. 3–9), Lys–Asp–Lys–Gly–Lys–Arg–Gly–Lys–Val–Lys (pos. 16–25), Lys–Lys–His–Gln–Lys (pos. 42–46) and Arg–Phe–Glu–Asp–Gly–Lys–Lys–Val–Arg (pos. 85–93).

### 3.3. *Secondary structure predictions of protein L24*

Based on the primary sequence, predictions of the secondary structure were made for this protein employing 4 different algorithms as in [30]. More recent parameters were used in the calculations according to Chou and Fasman: the occurrence of $\beta$-turns was predicted according to [31], and for helix and $\beta$-sheet the parameters of [32] were taken. The prediction results for protein L24 are presented in fig.2. Under the condition that at least 3 out of 4 predictions are in agreement, an average prediction was made, as given in the line 'PRE' of fig.2. This indicates the most probable conformational state for the individual residues. According to this prediction protein L24 has $\geqslant 9\%$ helix, 8% extended structure and 21% $\beta$-turn conformation. One helix region is strongly predicted for pos. 59–65, and two areas of extended structure are likely for pos. 10–14 and 32/33–35/37. Pronounced turn areas were calculated for positions 15–23 and 96–99. Starting points for turns are strongly predicted according to the rules of Chou and Fasman [31] for positions 15, 17, 19, 28–29, 52–54 and 86–87. The N-terminal part of the protein chain as well as the C-terminal region are likely to form a pleated sheet region.

### 3.4. *Comparison with known ribosomal sequences*

By means of computer programmes the sequence of protein L24 was compared with 48 known *E. coli* ribosomal proteins (reviewed in [33]) and with 5

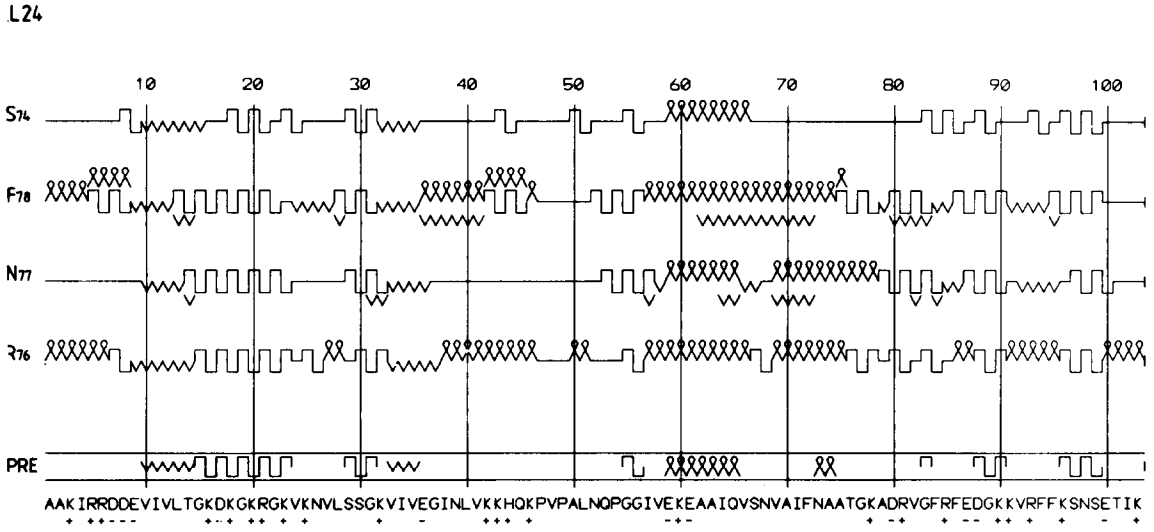


Fig.2. Predictions of the secondary structure of protein L24 according to 4 different methods. For explanations see text.

Table 1
Identical peptides in protein L24 and other ribosomal protein (*E. coli* = Ec, rat liver = RL)

| Ec-L24 positions | Identical sequence | Protein | Positions |
|---|---|---|---|
| 21–25 | Arg–Gly–Lys–Val–Lys | Ec-L14 | 49– 53 |
| 5– 8 | Arg–Arg–Asp–Asp | Ec-S6 | 112–115 |
| 20–23 | Lys–Arg–Gly–Lys | Ec-S5 | 155–158 |
| 22–25 | Gly–Lys–Val–Lys | Ec-L27 | 64– 67 |
| 40–43 | Leu–Val–Lys–Lys | Ec-L22 | 46– 49 |
| 48–51 | Val–Pro–Ala–Leu | Ec-S12 | 20– 23 |
| 60–63 | Lys–Glu–Ala–Ala | Ec-S4 | 76– 79 |
| 60–63 | Lys–Glu–Ala–Ala | Ec-L10 | 162–165 |
| 62–65 | Ala–Ala–Ile–Gln | Ec-S5 | 57– 60 |
| 63–66 | Ala–Ile–Gln–Val | Ec-L3 | 47– 50 |
| 75–78 | Ala–Thr–Gly–Lys | Ec-L13 | 20– 23 |
| 76–79 | Thr–Gly–Lys–Ala | Ec-L19 | 103–106 |
| 81–84 | Arg–Val–Gly–Phe | Ec-S5 | 44– 47 |
| 81–84 | Arg–Val–Gly–Phe | Ec-S16 | 35– 38 |
| 89–92 | Gly–Lys–Lys–Val | Ec-L3 | 6– 9 |
| 89–92 | Gly–Lys–Lys–Val | Ec-L22 | 26– 29 |
| 89–92 | Gly–Lys–Lys–Val | RL-L7a | 8– 11 |

complete structures of proteins from other organisms (yeast, *Bacillus subtilis* and *Artemia salina*) (reviewed in [34]).

Protein L24 shares the pentapeptide Arg–Gly–Lys–Val–Lys (pos. 21–25) with protein L14 (pos. 49–53). In addition, identical tetrapeptides were found both in L24 on the one hand and 11 other proteins from the *E. coli* ribosome, on the other hand, as given in table 1. This is considered to occur by random events [35]. No homologous areas with >5 adjacent amino acids were found when conservative replacements, such as glutamic acid for aspartic acid, isoleucine for valine, arginine for lysine and serine for threonine were allowed in the search programme. Certain structural similarities were observed only for protein L24 (pos. 7–47 and 4–47) on the one hand and *E. coli* protein L14 (pos. 37–71) and protein L13 (pos. 11–46) on the other hand.

Comparing the primary structure of protein L24 with that of eukaryotic ribosomal proteins, some similarities have been found between the N-terminal region of protein L24 and the rat liver ribosomal protein RL-L27 [36] as well as between positions 18–25 and 47–50 of protein L24 and positions 2–9 and 13–16 of protein RL-L7a [36].

## Acknowledgements

I wish to thank Mr Chris Southan for skillful technical assistance during the isolation and sequencing of the peptic fragments and some of the tryptic peptides of protein L24 and Mr A. Lehmann for performing the solid-phase sequence determinations. The early sequence investigations on this protein had been done partly by Dr R. R. Crichton.

## References

[1] Kühlbrandt, W. and Garrett, R. A. (1978) FEBS Lett. 94, 207–212.
[2] Spillmann, S., Dohme, F. and Nierhaus, K. H. (1977) J. Mol. Biol. 115, 513–523.
[3] Spillmann, S. and Nierhaus, K. H. (1978) J. Biol. Chem. 253, 7047–7050.
[4] Stöffler, G., Daya, L., Rak, K. H. and Garrett, R. A. (1971) J. Mol. Biol. 62, 411–414.
[5] Schulte, C., Morrison, C. A. and Garrett, R. A. (1974) Biochemistry 13, 1032–1037.
[6] Branlant, C., Krol, A., Sriwidada, J., Fellner, P. and Crichton, R. (1973) FEBS Lett. 35, 265–272.
[7] Branlant, C., Sriwidada, J., Krol, A. and Ebel, J. P. (1977) Eur. J. Biochem. 74, 155–170.
[8] Giri, L. and Dijk, J. (1979) Arch. Biochem. Biophys. 193, 122–129.
[9] Tischendorf, G. W. and Stöffler, G. (1976) Z. Immunforsch. 152, 120.
[10] Isono, K. and Krauss, J. (1976) Mol. Gen. Genet. 149, 297–302.
[11] Dabbs, E. (1977) Mol. Gen. Genet. 165, 73–78.
[12] Hindennach, I., Kaltschmidt, E. and Wittmann, H. G. (1971) Eur. J. Biochem. 23, 12–16.
[13] Kaltschmidt, E. and Wittmann, H. G. (1970) Analyt. Biochem. 36, 401–412.
[14] Hitz, H., Schäfer, D. and Wittmann-Liebold, B. (1977) Eur. J. Biochem. 75, 497–512.
[15] Wittmann-Liebold, B. and Wittmann, H. G. (1971) Biochim. Biophys. Acta 251, 44–53.
[16] Heiland, J., Brauer, D. and Wittmann-Liebold, B. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 1751–1770.
[17] Chen, R. and Wittmann-Liebold, B. (1975) FEBS Lett. 52, 139–140.
[18] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.
[19] Laursen, R. A. (1971) Eur. J. Biochem. 20, 89–102.
[20] Wittmann-Liebold, B. and Lehmann, A. (1975) in: Solid Phase Methods in Protein Sequence Analysis, (Larusen, R. A. ed) pp. 83–90, Pierce Chemical Company, Rockford.

[21] Chang, Y. J., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[22] Wittmann-Liebold, B., Brauer, D. and Dognin, M. (1977) in: Solid Phase Methods in Protein Sequence Analysis, Inserm Symp. no. 5, (Previero, A. and Coletti-Previero, M.-A. eds) pp. 219–232, Elsevier/North-Holland, Amsterdam, New York.
[23] Wittmann-Liebold, B. (1973) Hoppe-Seyler's Z. Physiol. Chem. 354, 1415–1431.
[24] Wittmann-Liebold, B., Geissler, A. W. and Marzinzig, E. (1975) J. Supramol. Struct. 3, 426–447.
[25] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[26] Wittmann-Liebold, B., Graffunder, H. and Kohls, H. (1976) Analyt. Biochem. 75, 621–633.
[27] Kaltschmidt, E., Dzionara, M. and Wittmann, H. G. (1970) Mol. Gen. Genet. 109, 292–297.
[28] Dzionara, M., Kaltschmidt, E. and Wittmann, H. G. (1970) Proc. Natl. Acad. Sci. USA 67, 1909–1913.
[29] Cabezon, T., Herzog, A., Petre, J., Yaguchi, M. and Bollen, A. (1977) J. Mol. Biol. 116, 361–374.
[30] Dzionara, M., Robinson, S. M. L. and Wittmann-Liebold, B. (1977) Hoppe-Seyler's Z. Physiol. Chem. 358, 1003–1019.
[31] Chou, P. Y. and Fasman, G. D. (1977) J. Mol. Biol. 115, 135–175.
[32] Chou, P. Y. and Fasman, G. D. (1978) Ann. Rev. Biochem. 47, 251–276.
[33] Wittmann, H. G., Littlechild, J. and Wittmann-Liebold, B. (1979) in: Ribosomes (Davies, J. et al. eds) University Park Press, Baltimore, in press.
[34] Matheson, A. T. and Möller, W. (1979) in: Ribosomes (Davies, J. et al. eds) University Park Press, Baltimore, in press.
[35] Wittmann-Liebold, B. and Dzionara, M. (1976) FEBS Lett. 65, 281–283.
[36] Wittmann-Liebold, B., Geissler, A. W., Lin, A. and Wool, I. G. (1979) submitted.