THE AMINO ACID SEQUENCE OF PROTEIN BL10 FROM THE 50 S SUBUNIT OF THE *BACILLUS STEAROTHERMOPHILUS* RIBOSOME

M. KIMURA, N. RAWLINGS and K. APPELT

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, Ihnestrasse 63–73, 1000 Berlin 33 (Dahlem), Germany

Received 23 October 1981

1. Introduction

The primary structures of almost all of the 53 proteins from the *Escherichia coli* ribosome have been determined [1], and the study of the structure of ribosomal proteins has now progressed towards the investigation of the tertiary structure. Because some of the ribosomal and related proteins (BL10, BL17, BL34, BSb) from *Bacillus stearothermophilus* have been crystallized [2], we have decided to determine their complete primary structures to facilitate the tertiary structural analyses of these proteins. We had elucidated the primary structure of BL17 [3], and have now extended this work to protein BL10.

Protein BL10 was found to be required to some extent for polyphenylalanine synthesis and peptidyltransferase activity as revealed by the single-omission reconstitution test [4]. An immunochemical cross-reaction study provided evidence that BL10 is probably homologous to protein L6 from *E. coli* [5]. This homology has also been concluded from comparison of CD and ^{1}H NMR spectra and of limited proteolysis patterns of both proteins [6].

Here, we present the complete primary structure of BL10 which has 177 amino acids. We have also predicted the secondary structural elements, e.g., α-helix, β-sheet and β-turn, by using 4 prediction programmes. Furthermore, a comparison is made between the primary structures and predicted secondary structures of BL10 and EL6 whose amino acid sequence has been determined [7].

2. Materials and methods

Protein BL10 was isolated avoiding urea and acetic acid extraction as in [8]. The identity and purity of the protein were checked by two-dimensional polyacrylamide gel electrophoresis [9].

Cyanogen bromide (CNBr) peptides were prepared by treating 150 nmol BL10 with CNBr (60 μmol) in 400 μl 70% formic acid for 20 h at room temperature. The resulting peptides were separated first on Sephadex G-75 superfine (1 × 210 cm) in 10% acetic acid and then on Sephadex G-50 superfine (1 × 140 cm) in the same solution.

Tryptic digestions of BL10 and its CNBr peptides were performed in 0.2 M *N*-methylmorpholin acetate buffer (pH 8.1) at 37°C for 3–6 h. Digestion of CNBr peptides with pepsin was carried out in 5% formic acid at 37°C for 6 h. CNBr peptide CBIII was also digested with *Staphylococcus aureus* protease in 0.1 M ammonium acetate buffer (pH 4.0) at 37°C for 20 h.

The peptides obtained by the various proteases were separated by fingerprinting using cellulose thin-layer plates as in [10]. The positions of the peptides were determined by staining with Fluram [11], and peptides were eluted with 50% acetic acid.

The amino acid sequences were determined by the DABITC/PITC double-coupling method [12] or manual solid-phase sequencing using DABITC [13]. Amino acid analyses were performed on Durrum D-500 analysers as in [10]. More details of the methods applied are in [14].

Predictions of the secondary structure were made from computerised algorithms of Scheraga, Chou and Fasman, Nagano and Robson. Details of the method are given in [15]. The prediction according to Chou and Fasman was done with the aid of a fully automated computer programme developed in [16].

Published by Elsevier/North-Holland Biomedical Press
00145793/81/0000–0000/$02.75

3. Results and discussion

3.1. *The primary structure of BL10*

The N-terminal sequence of BL10 up to position 13 was determined by the DABITC/PITC double-coupling method. The remaining amino acid sequence of BL10 was derived from analyses of peptides obtained by cleavage with cyanogen bromide, trypsin, pepsin and *Staphylococcus aureus* protease.

CNBr treatment of BL10 cleaved after the 2 methionine residues (pos. 40, 74) and produced 3 peptides, CBI, CBII and CBIII. Amino acid compositions and N-terminal sequences of these peptides revealed the alignment of peptides to be CBI–CBII–CBIII: the N-terminal region of CBI was identical to the N-terminal of the whole protein, CBIII contained no homoserine and must therefore be the C-terminal peptide.

The amino acid sequences of the peptides CBI (res. 1–40) and CBII (res. 41–74) were determined by sequencing their tryptic and peptic peptides. Six peptides (CBIT1–CBIT6) and 4 peptides (CBIIT1–CBIIT4) were isolated from tryptic digestions of peptides CBI and CBII, respectively, by fingerprinting. The complete sequences of all peptides except CBIT3 and CBIIT1 were determined by the DABITC/PITC double-coupling method. The peptides CBIT3 (res. 6–26) and CBIIT1 (res. 41–58) were completely sequenced by manual solid-phase sequencing. From peptic digestions of CBI and CBII, respectively, 3 peptides (CBIP1–CBIP3) and 1 peptide (CBIIP1) were isolated and then sequenced. From these results, the amino acid sequences of peptides CBI and CBII were established.

The sequence of peptide CBIII (res. 75–177), which is the C-terminal peptide comprising 103 residues, was determined by isolating and sequencing tryptic, peptic and *Staphylococcus aureus* protease peptides. Tryptic digestion of CBIII gave 20 peptides. The amino acid sequences of these peptides were completely determined by the DABITC/PITC double-coupling method or by the manual solid phase sequence technique. Seven peptides (CBIIIP1–CBIIIP7) were obtained from the peptic digestion and 8 peptides (CBIIISp1–CBIIISp8) from the *Staphylococcus aureus* protease digestion of CBIII, which were partially sequenced. From the combination of these results, the amino acid sequence of CBIII was determined (fig.1).

To confirm the alignment of the cyanogen peptides CBI, CBII and CBIII, protein BL10 was directly digested with trypsin and the resulting peptides were isolated by fingerprinting. Two peptides (TM1, res. 35–58, and TM2, res. 69–80) containing methionine residues were obtained and sequenced. The amino acid sequence of peptide TM1 provided the overlap CBI–CBII and that of peptide TM2 gave the overlap CBII–CBIII. In this way, the complete amino acid sequence of BL10 was established (fig.1).

3.2. *Characteristics of the sequence*

The amino acid composition derived from the sequence of BL10 is: $Asp_3Asn_5Thr_{14}Ser_8Glu_{18}Gln_3$-$Pro_{11}Gly_{21}Ala_9Val_{21}Met_2Ile_{10}Leu_{12}Tyr_5Phe_1His_5$-$Lys_{18}Arg_{11}$, which is in fairly good agreement with that obtained from the total hydrolysis of BL10. Tryptophan and cysteine are absent in BL10. The M_r calculated from the sequence analysis is 19 168.

Investigation of the distribution of the amino acids indicates that basic residues are generally located in the middle (res. 58–101) and C-terminal (res. 129–177) region, and there is a cluster of acidic amino acids at positions 113–123. All 5 tyrosine residues are located in the middle (pos. 82, 93, 108) and C-terminal part (pos. 156, 163).

3.3. *Comparison of the primary structures of BL10 and EL6*

The amino acid sequences of the 2 homologous proteins from *B. stearothermophilus* and *E. coli* are compared in fig.2. BL10 is one residue longer than EL6, and the proteins could be easily aligned as there are no deletions or insertions; 85 amino acids in BL10 are identical with respect to kind and position to those in EL6, which is 48% identity (fig.2). Of the non-identical residues, 65 (37%) of the substitutions could have arisen from single base mutations.

The distribution of conserved and substituted regions is quite different between the N-terminal half (res. 1–64) and the C-terminal half (res. 65–177). The N-terminal half shows relatively low homology (36%), especially in the region 36–64 where only 3 residues are identical. On the other hand, the C-terminal half shows a great similarity (76%). These results might indicate that the C-terminal half is more important for the function or maintaining the correct 3-dimensional structure of both proteins.

Although the primary structures of both proteins show relatively high homology as described above, the distribution of the charged (Asp,Glu,His,Lys,Arg)

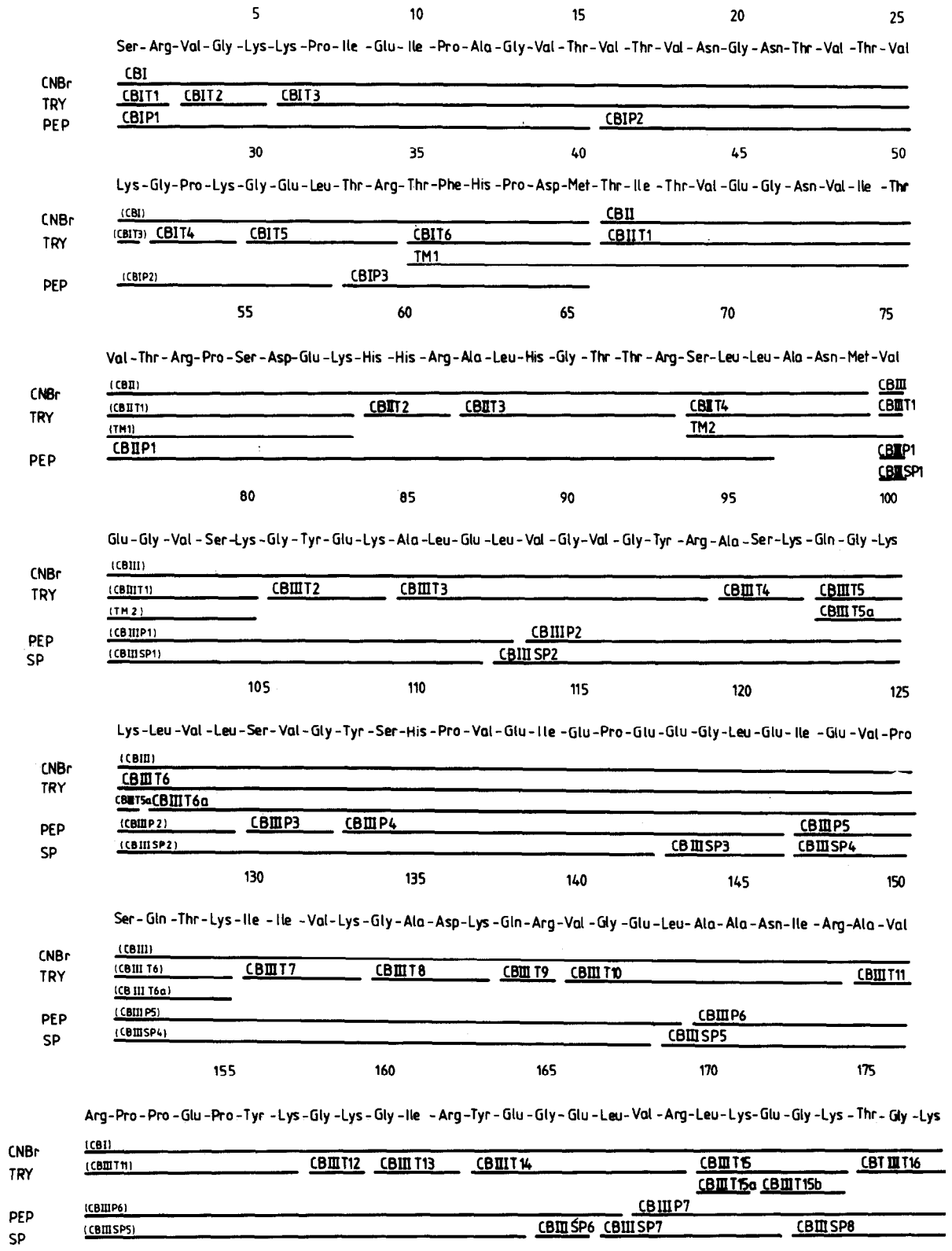


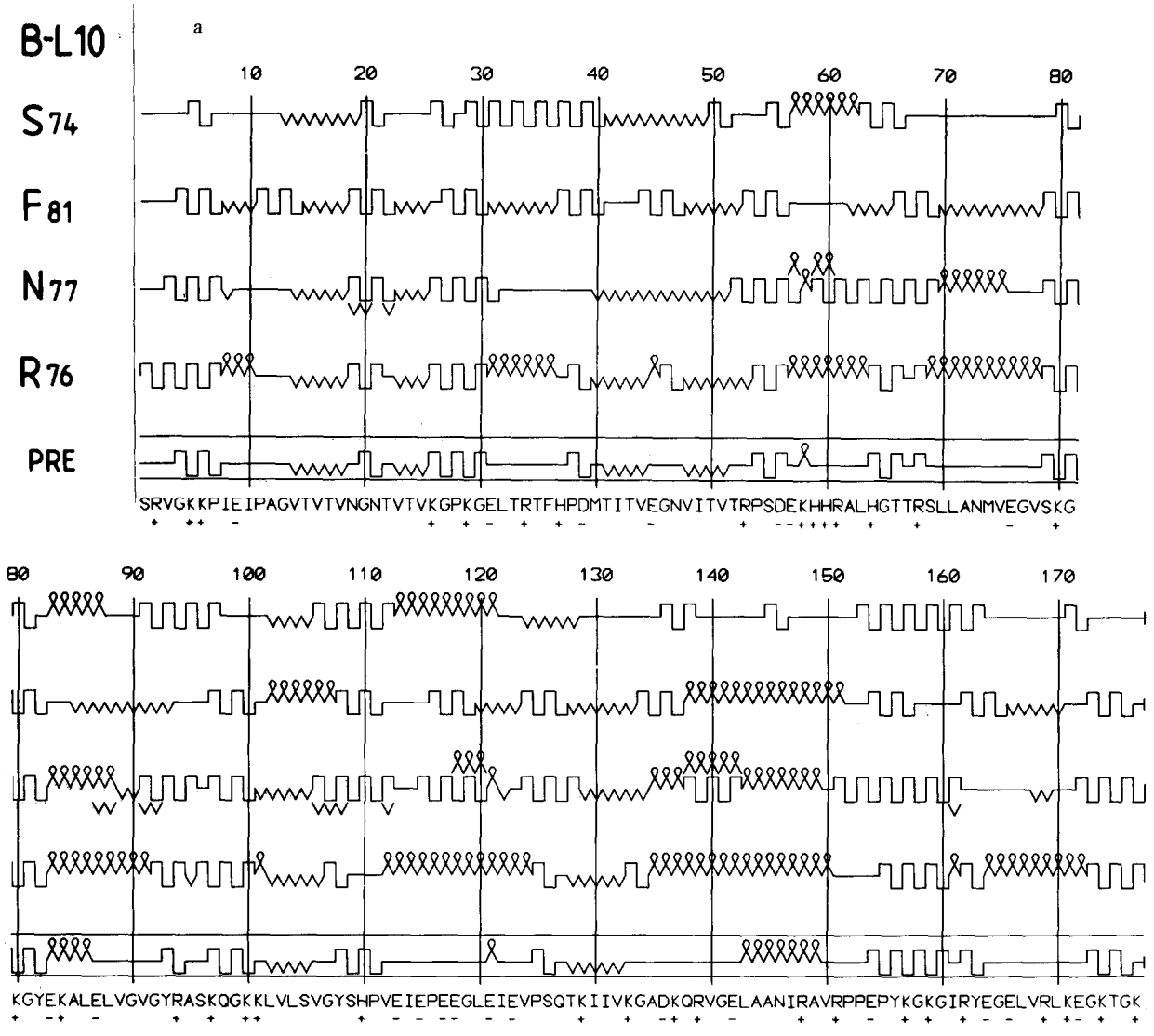
Fig.1. The primary structure of protein BL10. *Abbreviations:* CNBr, cyanogen bromide cleavage (peptides CB); TRY, trypsin digestion (peptides T); PEP, digestion with pepsin (peptides P); Sp, digestion with *Staphylococcus aureus* protease (peptides Sp).

```
                                                  10                                                  20
EL6  SER-ARG-VAL-ALA-LYS-ALA-PRO-VAL-VAL-VAL-PRO-ALA-GLY-VAL-ASP-VAL-LYS-ILE-ASN-GLY-
BL10 SER-ARG-VAL-GLY-LYS-LYS-PRO-ILE-GLU-ILE-PRO-ALA-GLY-VAL-THR-VAL-THR-VAL-ASN-GLY-
     +++ +++ +++ *** +++      +++ *** *** *** +++ +++ +++ +++      +++ *** *** +++ +++

                                                  30                                                  40
EL6  GLN-VAL-ILE-THR-ILE-LYS-GLY-LYS-ASN-GLY-GLU-LEU-THR-ARG-THR-LEU-ASN-ASP-ALA-VAL-
BL10 ASN-THR-VAL-THR-VAL-LYS-GLY-PRO-LYS-GLY-GLU-LEU-THR-ARG-THR-PHE-HIS-PRO-ASP-MET-
          *** +++ *** +++ +++      *** +++ +++ +++ +++ +++ +++ *** ***      *** ***

                                                  50                                                  60
EL6  GLU-VAL-LYS-HIS-ALA-ASP-ASN-THR-LEU-THR-PHE-GLY-PRO-ARG-ASP-GLY-TYR-ALA-ASP-GLY-
BL10 THR-ILE-THR-VAL-GLU-GLY-ASN-VAL-ILE-THR-VAL-THR-ARG-PRO-SER-ASP-GLU-LYS-HIS-HIS-
         *** ***      *** *** +++      *** +++ ***      *** ***      ***          ***

                                                  70                                                  80
EL6  TRP-ALA-GLN-ALA-GLY-THR-ALA-ARG-ALA-LEU-LEU-ASN-SER-MET-VAL-ILE-GLY-VAL-THR-GLU-
BL10 ARG-ALA-LEU-HIS-GLY-THR-THR-ARG-SER-LEU-LEU-ALA-ASN-MET-VAL-GLU-GLY-VAL-SER-LYS-
     *** +++ ***      +++ +++ *** +++ *** +++ +++      *** +++ +++      +++ +++ *** ***

                                                  90                                                  100
EL6  ASP-PHE-THR-LYS-LYS-LEU-GLU-LEU-VAL-GLY-VAL-GLY-TYR-ARG-ALA-ALA-VAL-LYS-GLY-ASN-
BL10 GLY-TYR-GLU-LYS-ALA-LEU-GLU-LEU-VAL-GLY-VAL-GLY-TYR-ARG-ALA-SER-LYS-GLN-GLY-LYS-
     *** ***      +++      +++ +++ +++ +++ +++ +++ +++ +++ +++ +++ +++ ***      *** +++ ***

                                                  110                                                 120
EL6  VAL-ILE-ASN-LEU-SER-LEU-GLY-PHE-SER-HIS-PRO-VAL-ASP-HIS-GLN-LEU-PRO-ALA-GLY-ILE-
BL10 LYS-LEU-VAL-LEU-SER-VAL-GLY-TYR-SER-HIS-PRO-VAL-GLU-ILE-GLU-PRO-GLU-GLU-GLY-LEU-
         ***      +++ +++ *** +++ *** +++ +++ +++ +++ ***      *** ***      *** +++ ***

                                                  130                                                 140
EL6  THR-ALA-GLU-CYS-PRO-THR-GLN-THR-GLU-ILE-VAL-LEU-LYS-GLY-ALA-ASP-LYS-GLN-VAL-ILE-
BL10 GLU-ILE-GLU-VAL-PRO-SER-GLN-THR-LYS-ILE-ILE-VAL-LYS-GLY-ALA-ASP-LYS-GLN-ARG-VAL-
          +++      +++ *** +++ +++ *** +++ *** *** +++ +++ +++ +++ +++ +++      ***

                                                  150                                                 160
EL6  GLY-GLN-VAL-ALA-ALA-ASP-LEU-ARG-ALA-TYR-ARG-ARG-PRO-GLU-PRO-TYR-LYS-GLY-LYS-GLY-
BL10 GLY-GLU-LEU-ALA-ALA-ASN-ILE-ARG-ALA-VAL-ARG-PRO-PRO-GLU-PRO-TYR-LYS-GLY-LYS-GLY-
     +++ *** *** +++ +++ *** *** +++ +++      +++ *** +++ +++ +++ +++ +++ +++ +++ +++

                                                  170
EL6  VAL-ARG-TYR-ALA-ASP-GLU-VAL-VAL-ARG-THR-LYS-GLU-ALA-LYS-LYS-LYS
BL10 ILE-ARG-TYR-GLU-GLY-GLU-LEU-VAL-ARG-LEU-LYS-GLU-GLY-LYS-THR-GLY-LYS
     *** +++ +++ *** *** +++ *** +++ +++      +++ +++ *** +++ ***
```

Fig.2. Comparison of the protein BL10 from *B. stearothermophilus* with protein EL6 from *E. coli*. The symbols (+++) represent that the residues are identical in the two proteins, those (***) are substitutions that could have arisen by single base mutations.

and the hydrophobic (Pro,Val,Met,Ile,Leu,Phe) amino acids seems to be somewhat different throughout both protein chains. There are some clusters (res. 26–45, 53–68, 113–123) of charged residues in BL10, whereas in EL6 they are evenly distributed throughout. With regard to the hydrophobic residues, their distribution is even denser in EL6 than BL10.

The properties of the ribosomes from *B. stearothermophilus* and *E. coli* are different in several respects [17]. One of them is that the ribosome from *B. stearothermophilus* is much more heat resistant than that from *E. coli*. Also, for free proteins, it has been found that BL10 is heat denatured around 60°C whereas EL6 around 40°C [6]. From comparisons of amino acid sequences of thermophilic and mesophilic molecules (ferredoxin, glyceraldehyde-3-phosphate dehydrogenase and lactate dehydrogenase), it was concluded [18] that in thermophiles Gly is generally substituted by Ala; Ser by Thr or Ala; Lys by Arg; and Asp by Glu. Accordingly, we have examined the substitutions of amino acids in BL10 and EL6, but the result is not in agreement with this hypothesis; there is only one such substitution (Asp by Glu at pos. 113). We find the reverse of the suggested thermophilic substitutions to occur more frequently (Ala by Gly, Ala by Ser and Thr by Ser, all occurring twice).



3.4. *Prediction of secondary structure of BL10 and comparison of predicted secondary structures of BL10 with EL6*

The predicted secondary structure of BL10 is shown in fig.3a. The line labelled 'PRE' indicates regions of the sequence where 3 out of the 4 predictions agree. All predictions suggest many β-turns, indicating that BL10 is a tightly coiled protein; and all methods, except Robson's, predict rather few α-helices (confined mainly to the center and C-terminal).

Because it was interesting to compare the secondary structure of BL10 with that of EL6, we also predicted the secondary structure of EL6 as shown in fig.3b and compared individual predictions. The result shows a considerable homology between the secondary structures of both proteins. For example, taking the result obtained from the prediction of Chou and Fasman, the following turns are identical in both proteins: 44–47, 53–56, 97–100, 116–119, 124–127, 134–137, 154–157 and 162–165. β-sheet is in agreement for the following regions: 6–10, 14–18, 23–26, 31–33, 48–51, 75–93, 120–123, 128–132 and

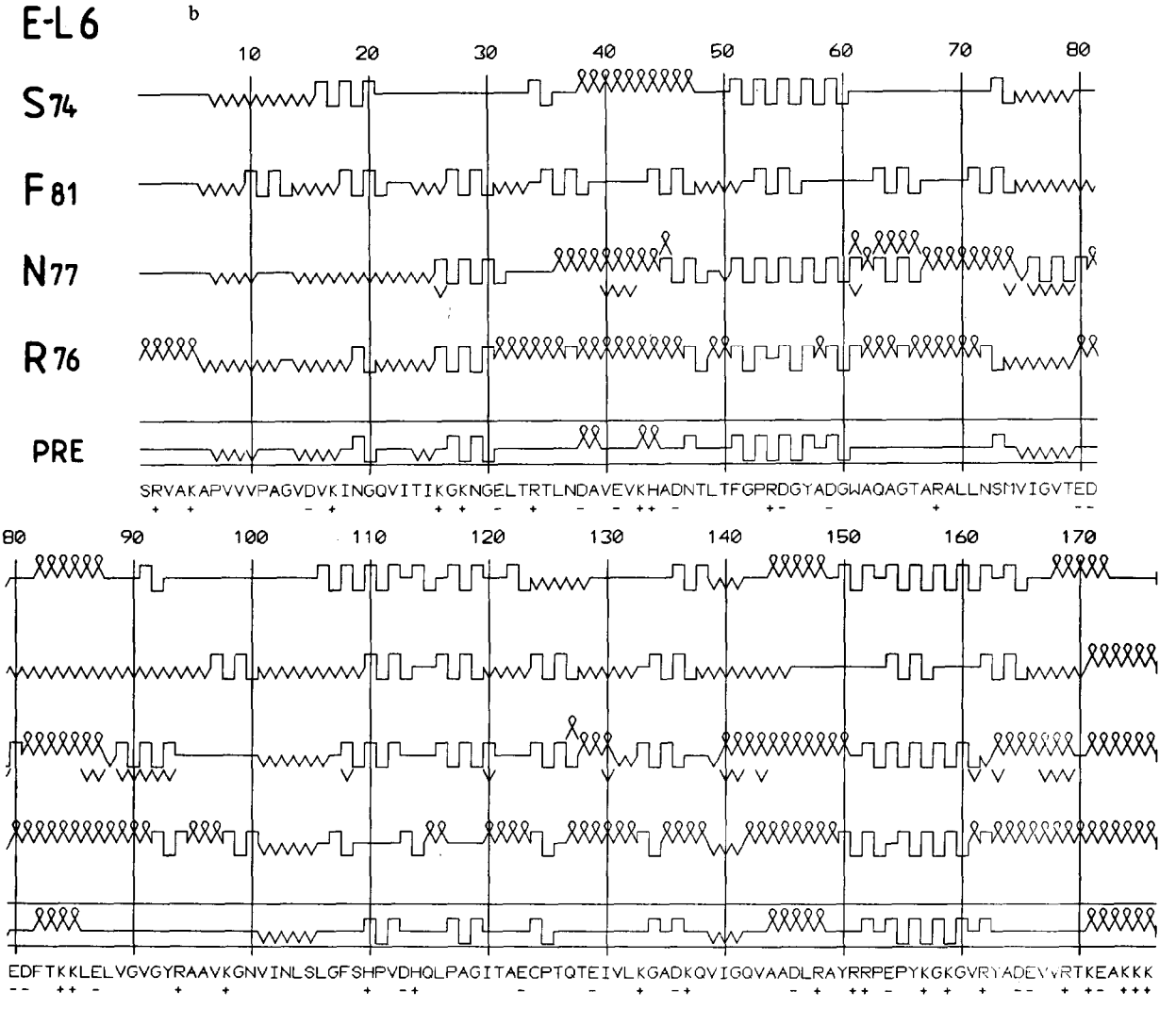


Fig.3. Secondary structure prediction of proteins, BL10 (a) and EL6 (b), according to 4 methods (see [15]). The symbols represent residues in a helical region (ΩΩΩΩΩΩ), β-sheet (ΛΛΛΛΛΛ), β-turns (⌐⌐⌐) and random coil conformation (——). The line 'PRE' summarizes the secondary structure obtained when at least 3 of the 4 predictions (S,F,N,R) are in agreement.

166–170. One more helix (pos. 171–176) is predicted for protein BL10 than for EL6. Although the primary structures agree well between the 2 proteins, there are several large differences, especially residues 37–61. It is interesting that in this region Chou and Fasman's predictions are strikingly similar: for BL10 turns at 37–40, 44–47, 53–56, β-sheet 48–52; for EL6 turns at 35–38, 44–47, 53–56, β-sheet 48–51.

Acknowledgements

We thank Dr H. G. Wittmann for his continuous interest and encouragement, and Drs J. Dijk and B. Wittmann-Liebold for many helpful discussions.

References

[1] Wittmann, H. G., Littlechild, J. A. and Wittmann-Liebold, B. (1980) in: Ribosomes (Chambliss, G. et al. eds) pp. 51–88, University Park Press, Baltimore MD.
[2] Appelt, K., Dijk, J., Reinhardt, R., Sanhuesa, S., White, S. W., Wilson, K. and Yonath, A. (1981) J. Biol. Chem. in press.
[3] Kimura, M., Dijk, J. and Heiland, I. (1980) FEBS Lett. 121, 323–326.
[4] Auron, P. E. and Fahnestock, S. R. (1981) J. Biol. Chem. 256, 10105–10110.
[5] Fahnestock, S. R., Strycharz, W. A. and Marguis, D. M. (1981) J. Biol. Chem. 256, 10111–10116.
[6] Appelt, K. and Dijk, J. (1981) in preparation.
[7] Chen, R., Arfsten, U. and Chen-Schmeisser, U. (1977) Hoppe-Seyler's Z. Physiol. Chem. 358, 531–535.
[8] Dijk, J. and Littlechild, J. A. (1979) Methods Enzymol. 59, 481–502.
[9] Kaltschmidt, E. and Wittmann, H. G. (1970) Anal. Biochem. 36, 401–412.
[10] Hitz, H., Schäfer, D. and Wittmann-Liebold, B. (1977) Eur. J. Biochem. 75, 497–512.
[11] Udenfriend, S., Stein, S., Bohlen, P., Dairman, W., Leimgruber, W. and Weigele, M. (1972) Science 178, 871–872.
[12] Chang, J. Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[13] Chang, J. Y. (1979) Biochim. Biophys. Acta 578, 188–195.
[14] Wittmann-Liebold, B., Brauer, D. and Dognin, J. (1977) in: Solid Phase Methods in Protein Sequence Analysis (Previero, A. and Coletti-Previero, M. A. eds) pp. 219–232, Elsevier/North-Holland, Amsterdam, New York.
[15] Dzionara, M., Robinson, S. M. L. and Wittmann-Liebold, B. (1977) Hoppe-Seyler's Z. Physiol. Chem. 358, 1003–1019.
[16] Rawlings, N., Ashman, K. and Wittmann-Liebold, B. (1981) in preparation.
[17] Nomura, M., Traub, P. and Bechmann, H. (1968) Nature 219, 793–799.
[18] Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H., Frank, G. and Tratschin, J. D. (1979) Biochemistry 18, 5698–5703.