The primary structure of initiation factor IF3 from *Bacillus stearothermophilus*

Makoto Kimura, Helmut Ernst and Krzysztof Appelt

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

Received 27 June 1983

The complete amino acid sequence of the initiation factor IF3 from *Bacillus stearothermophilus* has been elucidated. This was achieved by splitting the protein with trypsin, *Staphylococcus* protease or cyanogen bromide. The amino acid sequence was determined by manual Edman degradation, using the DABITC/PITC double-coupling method. The IF3 molecule contains 171 amino acids and has an M_r of 19 677. The sequence was compared to the homologous molecule from *Escherichia coli*; about 50% of the amino acid residues were found to be identical.

Bacillus stearothermophilus *Initiation factor 3* *Protein primary structure determination* *Protein biosynthesis* *Sequence comparison*

1. INTRODUCTION

The initiation factor IF3 has been found to promote the formation of the ternary complex consisting of 70 S ribosomes, fMet-tRNA and mRNA during the initiation of protein synthesis in prokaryotes [1,2]. Studies of the molecular properties of IF3 and of the mechanism by which the factor interacts with 30 S ribosomal subunits have been extensively carried out on *Escherichia coli* IF3 [2]. Based on these studies, it has been proposed that IF3 contains two active sites and interacts intermittently with the head and large lobe of the 30 S subunit thereby influencing the rate of formation and dissociation of the codon–anticodon complex as well as other properties of the small ribosomal subunit [3].

Here, we present the complete amino acid sequence of IF3 from *Bacillus stearothermophilus* (B-IF3) and compare it to that of *E. coli* IF3 [4]. The initiation factors of both organisms are interchangeable [5]. The comparison of the homologous proteins from various organisms will identify the conserved regions within the protein chain and help in the elucidation of the functionally active sites.

2. MATERIALS AND METHODS

Initiation factor from *Bacillus stearothermophilus* NCA 1503 (obtained from CAMR, Porton) was isolated and identified as in [6].

Tryptic digestion was carried out in 0.2 M *N*-methylmorpholine acetate buffer (pH 8.1) for 3 h at 37°C using an enzyme:substrate ratio of 1:50 (w/w). The resulting peptides (T-peptides) were separated by thin-layer fingerprinting.

Carboxymethylated IF3 (1.5 mg) was digested with 30 μg *Staphylococcus aureus* protease in 0.1 M ammonium acetate buffer (pH 4.0) for 24 h at 37°C. The resulting peptides (SP-peptides) were separated by gel filtration on Sephadex G50 s.f. (1 × 140 cm) in 10% acetic acid followed by thin-layer fingerprinting.

Cleavage with CNBr (6 mg) was carried out on a 2 mg aliquot of the protein in 300 μl 70% formic acid for 20 h at room temperature in the dark. The peptides were separated on Sephadex G75 s.f. (1 × 210 cm) in 10% acetic acid.

The amino acid sequence was determined by the manual method with the 4-*N*,*N*′-dimethylamino-azobenzene-4′-isothiocyanate/phenylisothiocyanate (DABITC/PITC) double-coupling procedure

Published by Elsevier Science Publishers B.V.

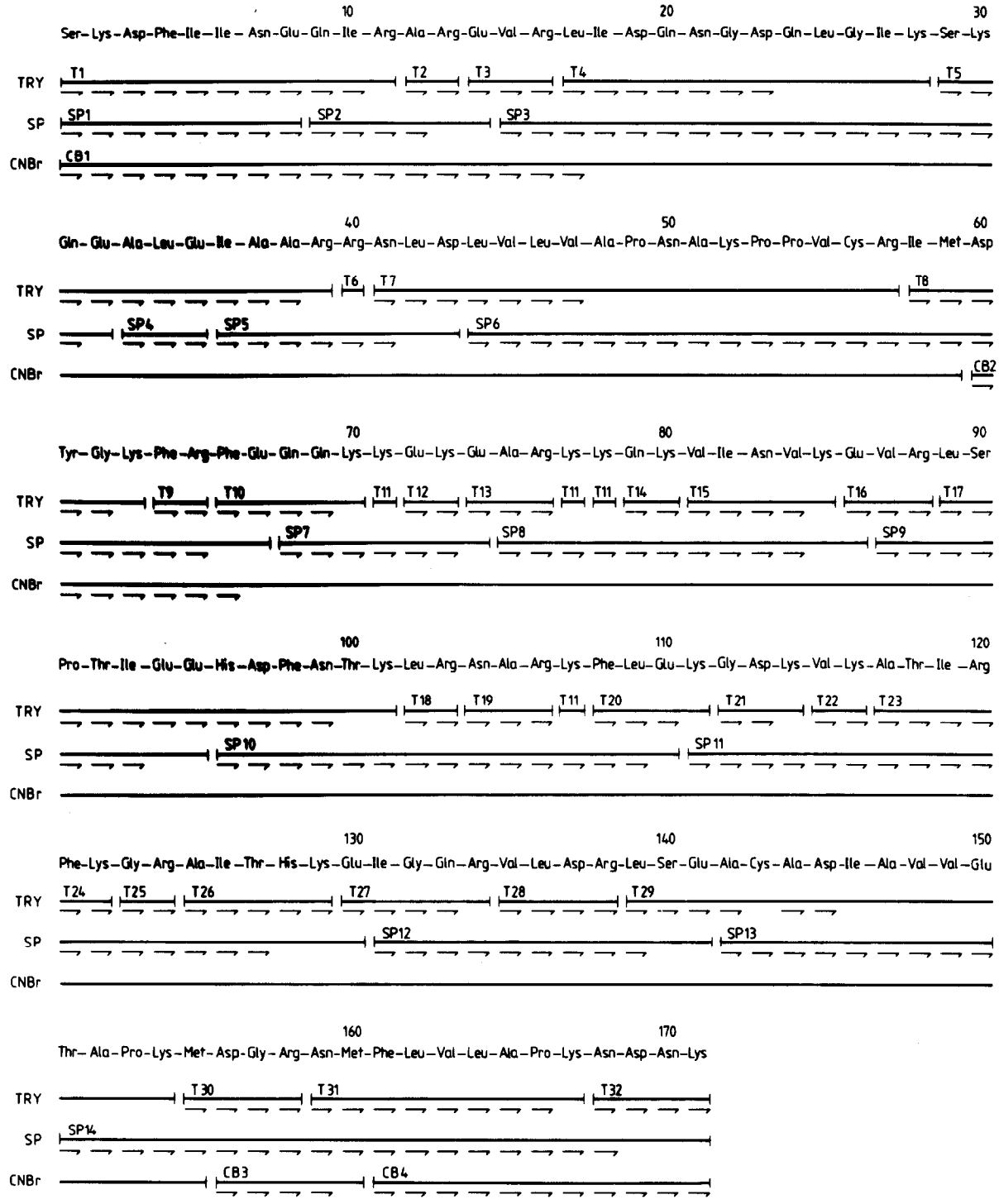

Fig.1. The amino acid sequence of initiation factor 3 from *B. stearothermophilus*. Sequence data on individual peptides are indicated as follows: (——→) sequenced by DABITC/PITC double-coupling method [7]; TRY, SP and CNBr indicate peptides derived from cleavage with trypsin, *Staphylococcus* protease, and cyanogen bromide, respectively.

cedure [7] as given in [8].

Amino acid analyses were performed on Durrum D-500 analyzers. Cysteine residues were determined as cysteic acid after performic acid oxidation [9]. Details of the experimental procedures used in sequencing are reviewed in [8].

3. RESULTS AND DISCUSSION

3.1. *Sequence determination*

All tryptic peptides were isolated by fingerprinting. They were completely sequenced with the exception of 3 peptides, namely T4 (positions 17–28), T7 (positions 41–57) and T29 (positions 139–154). Micro-sequencing of these peptides did not give the complete sequences due to their hydrophobicity and large size. For the determination of these remaining sequences and for the alignment of tryptic peptides, suitable overlap peptides were produced by digestion of the carboxymethylated IF3 protein with *Staphylococcus* protease. The resulting peptides were separated on Sephadex G50 followed by fingerprinting: 14 peptides were obtained and sequenced (fig.1). In this way the sequences of the tryptic peptides T4, T7 and T29 could be deduced from the sequences of peptides SP3, SP6 and SP14. All tryptic peptides were ordered as shown in fig.1.

To confirm unambiguously the alignment of the tryptic and SP-peptides, IF3 was cleaved with CNBr, and the resulting peptides were separated by gel filtration on Sephadex G-75.

As predicted from the presence of 3 methionine residues, 4 major peptides (CB1–CB4) were obtained and sequenced directly. The peptides CB1 and CB2 were further digested with pepsin and chymotrypsin, and the peptides thus formed were separated by fingerprinting and sequenced.

In this way the amino acid sequence of protein IF3 was completely elucidated as shown in fig.1. The amino acid composition derived from the sequences of B-IF3 is:

Asp_{11} Asn_{10} Thr_5 Ser_4 Glu_{14} Gln_8 Pro_6 Gly_7 Ala_{15} Val_{12} Met_3 Ile_{13} Leu_{13} Tyr_1 Phe_7 His_2 Lys_{22} Arg_{16} Cys_2

and the calculated M_r is 19677.

3.2. *Sequence comparison with* Escherichia coli *IF3 (E-IF3)*

In fig.2, the amino acid sequence of B-IF3 is aligned with that of E-IF3. B-IF3 is ten residues shorter than E-IF3 and hence gaps have to be left to enable the structures to be matched. There are identical residues at 84 positions in the 2 sequences (fig.2); i.e. at 50% of the residues compared; and additional 21 residues are conservative replacements.

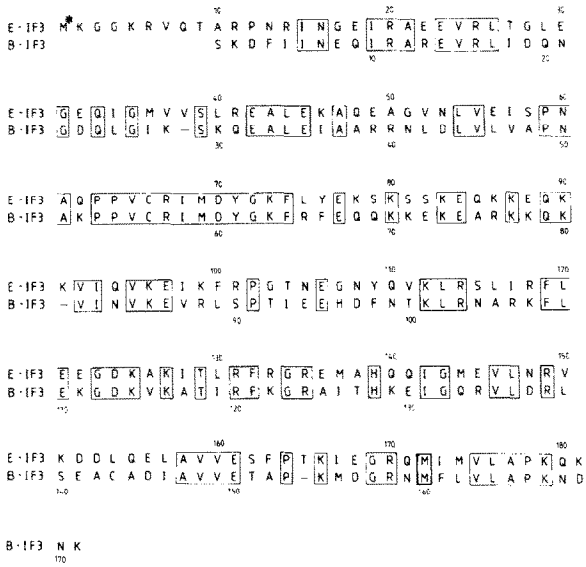

Fig.2. Comparison of the primary structures of initiation factor IF3 from *E. coli* (E-IF3) and *B. stearothermophilus* (B-IF3). Identical residues are enclosed in solid boxes. In position 108 of E-IF3 an Asp residue replaces the Asn residue originally found [10]; M*, monomethylmethionine.

From the comparison of the primary sequences of B-IF3 with E-IF3 (fig.2) it is clear that the former protein lacks 9 amino acids at the N-terminus. This is not surprising since it is known that a short form of E-IF3, lacking the N-terminal hexapeptide [4], can bind to the ribosome and is functionally active, albeit with reduced efficiency [2].

Taking into account the regions of the E-IF3 molecule investigated so far by means of various chemical and enzymatic modifications and found to be of functional relevance [2,10–13], a remarkable conservation is apparent. However, it is premature on the basis of only two sequences to draw firm conclusions concerning the extent of evolutionary conservation of specific regions of the molecule. Further sequence data from other bacterial species will therefore be required.

ACKNOWLEDGEMENTS

We are grateful to Drs C. Gualerzi, B. Wittmann-Liebold, and H.G. Wittmann for helpful discussions and critical reading of the manuscript. We also thank D. Kamp for performing amino acid analyses.

REFERENCES

[1] Grunberg-Manago, M. (1980) in: Ribosomes: Structure, Function and Genetics (Chambliss, G. et al. eds) pp.445–477, University Park Press, Baltimore MD.

[2] Gualerzi, C. and Pon, C.L. (1981) in: Structural Aspects of Recognition and Assembly in Biological Macromolecules (Balaban, M. et al. eds) pp.805–826, Int. Sci. Serv.: Rehovot; Philadelphia PA.

[3] Pon, C.L., Pawlik, R.T. and Gualerzi, C. (1982) FEBS Lett. 137, 163–167.

[4] Brauer, D. and Wittmann-Liebold, B. (1977) FEBS Lett. 79, 269–275.

[5] Pon, C.L., Friedman, S.M. and Gualerzi, C. (1972) Mol. Gen. Genet. 116, 192–198.

[6] Pawlik, R.T., Littlechild, J., Pon, C.L. and Gualerzi, C. (1981) Biochem. Internat. 2, 421–428.

[7] Chang, J.Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.

[8] Wittmann-Liebold, B. and Lehmann, A. (1980) in: Methods in Peptide and Protein Sequence Analysis (Birr, Ch. ed) pp.49–72, Elsevier, Amsterdam, New York.

[9] Needleman, S.B. (1975) in: Protein Sequence Determination: Molecular Biology, Biochemistry and Biophysics, vol.8, p.129, Springer, Berlin, New York.

[10] Bruhns, J. and Gualerzi, C. (1980) Biochemistry 19, 1670–1676.

[11] Ohsawa, H. and Gualerzi, C. (1981) J. Biol. Chem. 256, 4905–4912.

[12] Pon, C., Cannistraro, S., Giovane, A. and Gualerzi, C. (1982) Arch. Biochem. Biophys. 217, 47–57.

[13] Lammi, M., Paci, M., Pon, C.L. and Gualerzi, C. (1982) Biochem. Internat. 5, 429–436.