STRUCTURE–FUNCTION RELATIONSHIPS IN *ESCHERICHIA COLI* INITIATION FACTORS

II. Elucidation of the primary structure of initiation factor IF-1*

Cynthia L. PON, B. WITTMANN-LIEBOLD and Claudio GUALERZI
Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Ihnestrasse 63–73, 1000 Berlin 33 (Dahlem), Germany

Received 5 March 1979

1. Introduction

Initiation factor IF-1 is the smallest of the three initiation factors [1] and also the one whose functional role remains so far least clear [2,3]. In fact, while no specific function has so far been attributed to this factor, IF-1 seems to stimulate more or less all the activities of the other two initiation factors [2,3] through mechanism(s) which remain obscure. It has been reported that IF-1 binds to both 30 S ribosomal subunits, preferentially in the presence of IF-2 and IF-3 [4–6] and to 70 S ribosomes [4]. Binding to the 30 S subunit produces an increase in the affinity of these particles for initiation factor IF-2 [7]. IF-1 also stimulates the ribosome dissociation activity of IF-3 by increasing both 'on' and 'off' rate constants for the dissociation of 70 S ribosomes [8]. The reports that neither *Bacillus stearothermophilus* [9] nor *Caulobacter crescentus* [10] contain IF-1-like activities are puzzling. However, the most intriguing aspect of the IF-1 functioning concerns the mechanism by which it supposedly helps the recycling of IF-2 at the 70 S level [11–13] and the mechanism and the timing of IF-1 binding and release from the ribosomes. Finally, although a number of ribosomal proteins neighbouring IF-1 on the 30 S ribosomal subunits have been identified by use of protein–protein crosslinking reagents [5], nothing is known about the molecular nature of the interaction between IF-1 and ribosomes.

In this paper, as the first step toward a better understanding of the function and structure–function relationship of IF-1, we present the complete amino acid sequence of this factor and a prediction of its secondary structure based on this sequence.

2. Materials and methods

Initiation factor IF-1 was purified to electrophoretic homogeneity from frozen *Escherichia coli* MRE600 cell paste (Microbiological Research Establishment, Porton Down) according to the procedure described elsewhere (C.G., C.P., in preparation). Purity was checked by two-dimensional polyacrylamide gel electrophoresis [14] and sodium dodecyl sulfate (SDS)–slab gel electrophoresis [15].

Enzymatic digestions were performed as follows: TPCK-trypsin (Worthington Biochemicals Corp., Freehold, NJ) at 37°C for 4 h in 0.2 M *N*-methylmorpholine acetate (pH 8.0); α-chymotrypsin (Worthington Biochemicals Corp., Freehold, NJ) at 37°C for 45 min in 0.2 M *N*-methylmorpholine acetate (pH 8.0); *Staphylococcus aureus* protease (Miles, Slough) at 37°C for 24 h in 0.05 M ammonium acetate buffer (pH 4.0). In some cases, prior to digestion, IF-1 was oxidized with performic acid according to [16].

The resulting peptides were separated by fingerprinting on cellulose thin-layer plates (Polygram CEL 300, Macherey and Nagel, Düren) essentially as in [17,18]. Peptide spots, identified by lightly spraying the plates with ninhydrin, were scraped from the plates and extracted with 5.7 N HCl containing 0.02% 2-mercaptoethanol for determination of amino acid composition or with 50% acetic acid for amino acid sequence determination.

* Paper I in this series is [29]

Specific cleavage at methionine residues was performed by treatment with cyanogen bromide for 20 h at room temperature [19]. The peptides were then separated by gel filtration on a Sephadex G-50 (superfine) column in 5% acetic acid. Purity of the peptides was checked by SDS–slab gel electrophoresis using 20% acrylamide.

The amino acid composition of total protein and peptides was determined in a Durrum D-500 amino acid analyzer.

Amino acid sequences of peptides were determined manually using the 4-*N*,*N*-dimethylaminoazobenzene 4′-isothiocyanate double coupling method [20]. The first 30 residues from the N-terminus were also determined by automatic Edman degradation in a modified Beckman sequenator [21].

3. Results and discussion

The primary structure of initiation factor IF-1 is presented in fig.1. Automatic sequence analysis of IF-1 established the first 30 residues of the amino-terminal end. This sequence was confirmed by manual sequence determination performed on tryptic peptides T1 and T2. The other 6 tryptic peptides were also sequenced manually.

Sequence determination of the chymotryptic and *Staphylococcus aureus* protease peptides reconfirmed the tryptic sequences and established additional alignments: the sequences of CH3a and CH4 and CH5 established the alignment T2–T3a–T4 and SP6 aligned T5–T6–T7–T8. The remaining bridge between T4 and T5 could not be established from CH or SP peptides. The corresponding CH peptide was not stained by the usual ninhydrin spraying of the fingerprint and SP peptide (position 27–55) was recovered from the fingerprints in extremely low amounts which were insufficient for sequence determination. Thus, cyanogen bromide cleavage was undertaken and the isolated peptides were completely sequenced manually with the exception of CNBr4 which was sequenced only to the 29th residue. The results reconfirmed all the previous alignments and established the bridge between T4 and T5.

The manual determination of the primary structure of all peptides was performed solely through the use of the 4-*N*,*N*-dimethylaminoazobenzene 4′-isothiocyanate double coupling method [20]. All peptides (0.5–6 nmol) recovered from the thin-layer plate (1–3 plates) following enzymatic digestion, as well as CNBr peptides purified on a Sephadex G-50 column, were successfully sequenced using this technique.

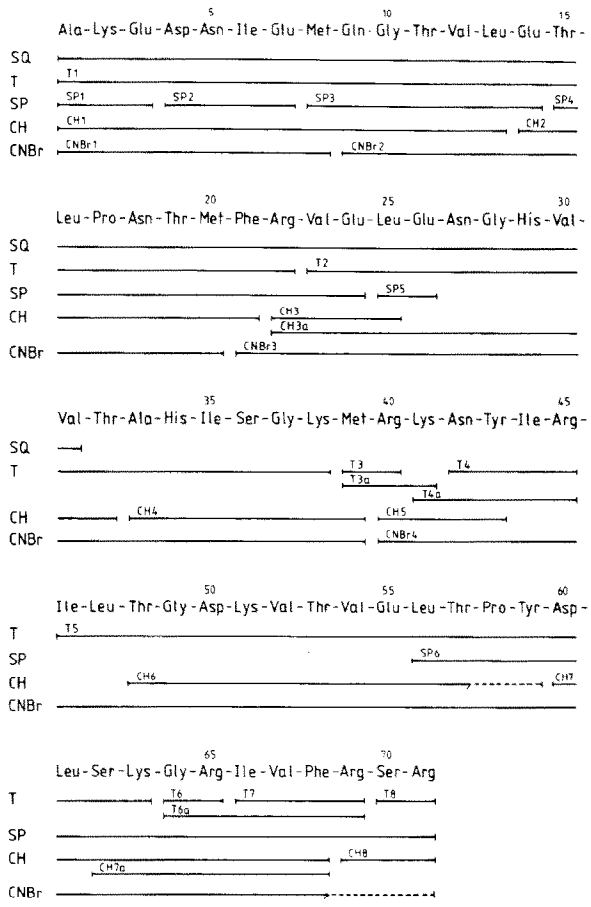

Fig.1. Amino acid sequence of *E. coli* initiation factor IF-1. SQ, derived from automatic sequenator; T, tryptic peptides; SP, *Staphylococcus aureus* protease peptides; CH, chymotryptic peptides; CNBr, peptides produced by cyanogen bromide cleavage. (- - -) sequence not determined but confirmed by amino acid composition of peptide.

The amino acid composition of initiation factor IF-1 derived from the sequence is: Asp_3 Asn_4 Thr_7 Ser_3 Glu_6 Gln_1 Pro_2 Gly_5 Ala_2 Val_7 Met_3 Ile_5 Leu_6 Tyr_2 Phe_2 His_2 Lys_5 Arg_6. These results are in good agreement with the total amino acid composition of the protein, but are quite different from the com-

position of IF-1 from *E. coli* Q13 reported in [22]. The most striking differences are the high content of lysine (10 residues) and arginine (8 residues), and the presence of 1 cysteine and 1 tryptophan reported in [22] compared to lysine (5 residues) and arginine (6 residues) and the complete absence of cysteine and tryptophan in the protein of the present report. Our results, on the other hand, are in good agreement with the amino acid composition of IF-1 reported in [23].

The mol. wt 8119 calculated from the sequence is also in good agreement with the values determined by SDS–polyacrylamide gel electrophoresis (mol. wt 8200–8900) and by estimation from the amino acid composition (mol. wt 8700) reported in [23]. These values are somewhat lower than that found by gel filtration (mol. wt 9300) [1,23].

The comparison of the sequence of IF-1 with known sequences from ribosomal proteins and factors showed only two tetrapeptides and one tripeptide to be identical. These are: IF-1 positions 13–16 with S7 positions 46–49 and equidistant away in both proteins IF-1 positions 24–26 with S7 positions 57–59; IF-1 positions 41–44 with S18 positions 29–32. This low degree of sequence homology is probably not statistically significant.

The secondary structure of IF-1 was predicted by 4 different methods [24–28] (fig.2). The final prediction (see fig.2, PRE) is based on the agreement of at least 3 out of the 4 methods. According to this prediction IF-1 contains a minimum of 11% α-helix, 17% turns (or loops) and 7% β-sheet and thus consists primarily of β-turns and random coil.

Acknowledgements

We wish to thank Professor H. G. Wittmann for his constant interest and encouragement as well as for the gift of some IF-1. We are also extremely grateful to Dr J. Y. Chang for teaching us the micromethod for manual sequence determination and for providing us with the DABITC reagent.

References

[1] Thach, R. E., Hershey, J. W. B., Kolakofsky, D., Dewey, K. F. and Remold-O'Donnell, E. (1969) Cold Spring Harbor Symp. Quant. Biol. 34, 277–284.

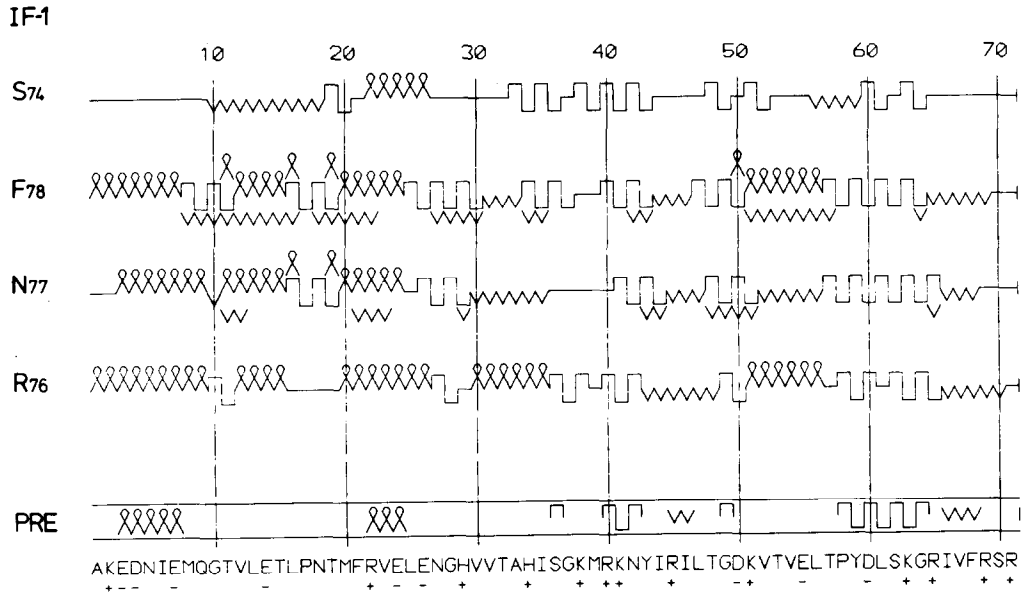

Fig.2. Prediction of the secondary structure of IF-1. Symbols: ( ) α-helix; ( ) turns; ( ) β-sheet; PRE, prediction. The calculation of the secondary structure was taken from references S74 [24]; F78 [25,26]; N77 [27]; R76 [28].

[2] Grunberg-Manago, M. and Gros, F. (1977) Prog. Nucl. Acid Res. Mol. Biol. 20, 209–284.
[3] Revel, M. (1977) in: Molecular Mechanisms of Protein Biosynthesis (Weissbach, H. and Pestka, S. eds) pp. 245–321, Academic Press, New York.
[4] Van der Hofstad, G. A. J. M., Buitenhek, A., Van den Elsen, P. J., Voorma, H. O. and Bosch, L. (1978) Eur. J. Biochem. 89, 221–228.
[5] Langberg, S., Kahan, L., Traut, R. R. and Hershey, J. W. B. (1977) J. Mol. Biol. 117, 307–319.
[6] Gualerzi, C. and Pon, C. (1978) Methods Enzymol. 59, 782–795.
[7] Stringer, E. A., Sarkar, P. and Maitra, U. (1977) J. Biol. Chem. 252, 1739–1744.
[8] Noll, M., Hapke, B. and Noll, H. (1973) J. Mol. Biol. 80, 519–530.
[9] Kay, A. C., Graffe, M. and Grunberg-Manago, M. (1976) Biochimie 58, 183–199.
[10] Leffler, S. and Szer, W. (1974) J. Biol. Chem. 249, 1465–1468.
[11] Benne, R., Arentzen, R. and Voorma, H. O. (1972) Biochim. Biophys. Acta 269, 304–310.
[12] Kay, A. C. and Grunberg-Manago, M. (1972) Biochim. Biophys. Acta 277, 225–230.
[13] Sarkar, P., Stringer, E. A. and Maitra, U. (1974) Proc. Natl. Acad. Sci. USA 71, 4986–4990.
[14] Kaltschmidt, E. and Wittmann, H. G. (1970) Analyt. Biochem. 36, 401–412.
[15] Laemmli, U. K. (1970) Nature 227, 680–685.
[16] Hirs, C. H. W. (1956) J. Biol. Chem. 219, 611–621.
[17] Yaguchi, M., Wittmann, H. G., Cabézon, T., DeWilde, M., Villarroel, R., Herzog, A. and Bollen, A. (1975) Mol. Gen. Genet. 142, 35–43.
[18] Hitz, H., Schäfer, D. and Wittmann-Liebold, B. (1977) Eur. J. Biochem. 75, 497–512.
[19] Gross, E. and Witkop, B. (1961) J. Am. Chem. Soc. 83, 1510–1511.
[20] Chang, J. Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[21] Wittmann-Liebold, B. (1973) Hoppe-Seyler's Z. Physiol. Chem. 354, 1415–1431.
[22] Lee-Huang, S., Sillero, M. A. and Ochoa, S. (1971) Eur. J. Biochem. 18, 536–543.
[23] Hershey, J. W. B., Yanov, J., Johnston, K. and Fakunding, J. L. (1977) Arch. Biochem. Biophys. 182, 626–638.
[24] Burgess, A. W., Ponnuswamy, P. K. and Scheraga, H. A. (1974) Israel J. Chem. 12, 239–286.
[25] Chou, P. Y. and Fasman, G. D. (1977) J. Mol. Biol. 115, 135–175.
[26] Chou, P. Y. and Fasman, G. D. (1978) Ann. Rev. Biochem. 47, 251–276.
[27] Nagano, K. (1977) J. Mol. Biol. 109, 251–274.
[28] Robson, B. and Suzuki, E. (1976) J. Mol. Biol. 107, 327–356.
[29] Ewald, R. and Gualerzi, C. (1977) Biochem. Biophys. Res. Commun. 77, 1517–1524.