THE PRIMARY STRUCTURE OF THE ELONGATION FACTOR G FROM *ESCHERICHIA COLI*

Amino acid sequence of the C-terminal domain

Yu. B. ALAKHOV, N. V. DOVGAS, L. P. MOTUZ, L. M. VINOKUROV and Yu. A. OVCHINNIKOV
Institute of Protein Research, USSR Academy of Sciences, 142292 Poustchino, Moscow Region, USSR

Received 25 February 1981

1. Introduction

Mild proteolytic hydrolysis of the elongation factor G splits it into a limited number of fragments [1–3]. Two fragments, T_2 and T_5 (nomenclature from [2]) encompass the whole protein polypeptide chain. Fragment T_2 represents the N-terminal part of the EF-G molecule with M_r 56 000 and fragment T_5 the C-terminal part with M_r 25 000. In terms of general packing of their polypeptide chains, these fragments represent individual structural units or domains in the EF-G molecule [4].

The study of modification of tyrosine residues in EF-G [5] has shown that modification with either tetranitromethane or iodine inactivates the protein in the ribosome-dependent GTPase reaction. Tyrosine residues subjected to modification are located in the C-terminal domain (fragment T_5) and are effectively protected against modification by EF-G binding with the 70 S ribosome or the 50 S subparticle. In the presence of guanyl nucleotides whose binding site is located in the N-terminal domain (in the N-terminal part of fragment T_2) [6] the rate of modification is considerably higher than in the free protein. Since modification of tyrosine residues does not affect the ability of EF-G to form binary complexes with guanyl nucleotides [5] and taking into account the shielding effect exerted by the ribosome on modification it has been concluded that the C-terminal domain contains one of the sites of interaction with the ribosome. It is possible that N- and C-terminal parts of EF-G interacting with GTP form a common site of interaction with the ribosome; this site disappears after GTP hydrolysis within the complex with the ribosome; as a result, EF-G loses affinity for the ribosome.

Here we report the results of the study of the primary structure of fragment T_5. The polypeptide chain of the fragment consists of 228 amino acid residues, does not contain tryptophan or cysteine residues. The C-terminal part of the fragment is two amino acid residues (Gly–Lys) shorter than the C-terminal part of EF-G.

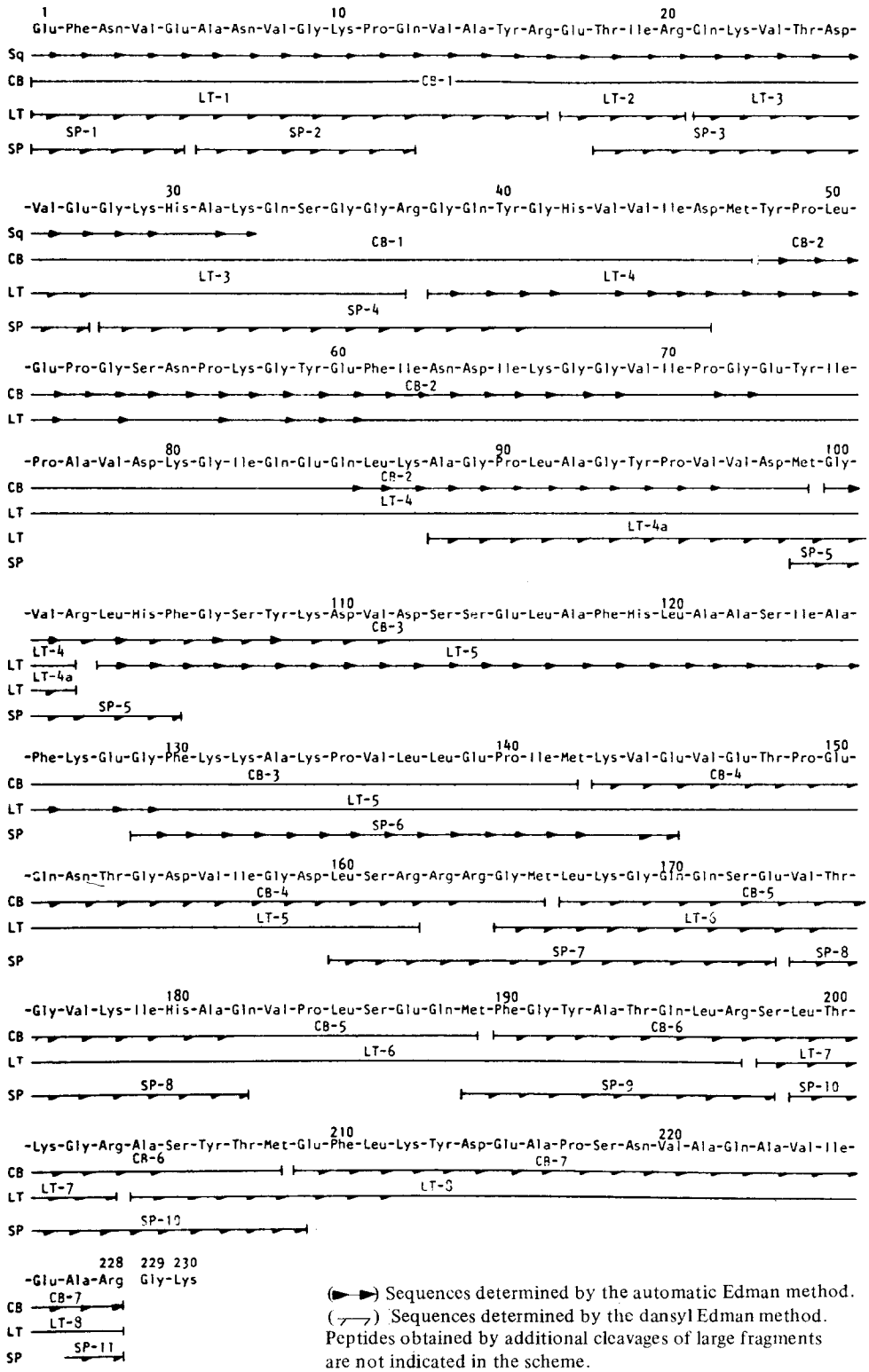
2. Materials and methods

EF-G and fragment T_5 were isolated as in [2]. Homogeneity of the protein and fragment was checked by SDS–polyacrylamide gel electrophoresis and by determining the N-terminal amino acid sequences by automatic Edman degradation.

We have studied the products of fragment T_5 cleavage by cyanogen bromide, *Staphylococcus aureus* protease (Miles Lab.) and trypsin (TPCK-treated, Worthington) after modification of the fragment with maleic anhydride. To separate the mixtures obtained we used gel filtration on Biogel P-10 or Sephadex G-25 with a following rechromatography of the fractions obtained on QAE–Sephadex A-25. To clarify the structure of large peptides of cyanogen bromide and tryptic hydrolysates, they were additionally digested with trypsin, chymotrypsin and *Staphylococcus aureus* protease. The mixtures obtained after hydrolysis were separared by the method of peptide mapping on cellulose thin-layer plates (Merck). The first dimension was electrophoresis for 1 h, 800 V (pH 3.5); the second dimension was chromatography in the system of solvents, *n*-butanol–pyridine–acetate–water (1:1:1:1, by vol.), (pH 5.4). Paper chromatography (Whatman

Scheme 1

Complete amino acid sequence of the C-terminal domain of EF-G



(►►) Sequences determined by the automatic Edman method.
(⌐¬) Sequences determined by the dansyl Edman method.
Peptides obtained by additional cleavages of large fragments are not indicated in the scheme.

3 MM) in the same solvent system as for peptide mapping was also used to separate tryptic peptides. The peptides were localized with a 0.001% fluorescamine solution in acetone and eluted by 10% acetic or 85% formic acid. Reverse-phase chromatography on LiChrosorb RP-8 (10 μm) in an acetonitryl gradient was equally effective for separation of cyanogen bromide peptides. The column dimensions were 4.6 × 250 mm: solution (A) 0.1 M ammonium acetate (pH 6.0), 5% acetonitryl; solution (B) 0.1 M ammonium acetate (pH 6.0), 50% acetonitryl.

The C-terminal amino acid residues of fragment T_5 and peptide CB-7 were determined using carboxypeptidases A and B [7].

The amino acid composition was assessed on a D-500 amino acid analyzer (Durrum). The amount of valine and isoleucine was determined from the results of a 72-hour hydrolysis.

Automatic determination of the amino acid sequence was carried out on a sequencer (model 980 C, Beckman) using the dimethylallylamine program 102974 according to the company handbook and on a solid-phase sequencer APS-240 (Rank-Hilger). PTH-derivatives of amino acids were identified by thin-layer chromatography on silica gel and HPLC (HP 1080 A chromatograph) on a Spherisorb RP-18 column (5 μm) in sodium phosphate buffer using a methanol gradient.

The amino acid sequence of isolated peptides was determined by the dansyl Edman method [8]. Amino acid dansyl derivatives were identified by thin-layer chromatography on silica gel plates. Asparagine and glutamine were determined according to [9].

3. Results and discussion

According to the data of SDS–polyacrylamide gel electrophoresis, fragment T_5 has a M_r ~25 000 and its amino acid composition was determined as follows (in mol/%):

Asp, 7.33; Thr, 3.59; Ser, 4.66; Glu, 13.51; Pro, 5.66; Gly, 10.64; Ala, 8.60; Val, 10.00; Met, 2.47; Ile, 4.62; Leu, 6.54; Tyr, 3.88; Phe, 3.11; His, 1.95; Lys, 8.93; Arg, 4.31.

Glutamine is the N-terminal amino acid and arginine is the C-terminal amino acid.

The amino acid sequence up to position 32 inclusively was determined on a whole molecule by the automated liquid-phase Edman method.

The molecule of fragment T_5 contains 6 methionine residues, therefore 7 peptides resulted from cyanogen bromide cleavage (CB-1–7, scheme 1). The structure of peptides CB-2 and CB-3 was determined by the solid phase Edman method up to positions 29 and 12, respectively. Peptides CB-3–7 were studied by the dansyl Edman method and the amino acid sequences were determined up to positions 12, 22, 15 and 20, respectively. To clarify the complete amino acid sequences of cyanogen bromide peptides, we studied additionally the hydrolysates of peptide CB-1 with trypsin and *Staphylococcus aureus* protease; those of peptide CB-2 with trypsin and chymotrypsin; of peptide CB-4 with chymotrypsin; of peptides CB-5 and CB-6 with trypsin; of peptide CB-7 with thermolysine. As a result, we determined the complete amino acid sequences of 6 cyanogen bromide peptides, except for peptide CB-3 which was isolated in insufficient quantity. The structure of this fragment of the polypeptide chain was studied on peptide LT-5 isolated from tryptic hydrolysate of T_5 modified with maleic anhydride. This peptide encompasses almost completely the structure of peptides CB-3 and CB-4. The peptide was studied by the solid phase Edman method up to position 27 and then was subjected to chymotrypsin and thermolysine hydrolysis. As a result, we determined completely the structure of peptide CB-3 and obtained overlapping with peptide CB-4. To obtain peptides overlapping other cyanogen bromide fragments, we studied the remaining peptides of limited tryptic hydrolysate and products of *S. aureus* protease hydrolysis. The structure of these peptides was mainly determined by the dansyl Edman method.

Thus, the structure study of peptides isolated after digestion of the T_5 molecule or its fragments by cyanogen bromide, trypsin, chymotrypsin, thermolysine and *S. aureus* protease gave the complete amino acid sequence of the C-terminal domain (fragment T_5) of EF-G (scheme 1). The polypeptide chain consists of 228 amino acid residues, the exact M_r is 25011.8 and the amino acid composition is as follows:

Asp_{10}, Asn_6, Thr_8, Ser_{12}, Glu_{19}, Gln_{13}, Pro_{13}, Gly_{24}, Ala_{18}, Val_{22}, Met_6, Ile_{12}, Leu_{14}, Tyr_{10}, Phe_8, His_5, Lys_{18}, Arg_{10}.

We have shown [2] that all the fragments of limited trypsinolysis of EF-G have arginine as the C-terminal amino acid while in EF-G the C-terminal amino acid is lysine. Therefore it was necessary to determine by how many residues is fragment T_5 shortened from

the C-terminus in comparison with EF-G. The study of products of cyanogen bromide cleavage of the whole EF-G molecule (to be published elsewhere) resulted in isolation of a peptide having a sequence analogous to that of the cyanogen bromide peptide (CB-7) of fragment T_5 but differing from it by 2 amino acid residues Gly-Lys. Thus it has been concluded that limited EF-G trypsinolysis results in cleavage of a dipeptide from the C-terminus of fragment T_5. The study of another EF-G cyanogen bromide peptide has shown that fragment T_5 directly adjoins fragment T_4.

References

[1] Skar, D. C., Rohrbach, M. S. and Bodley, J. W. (1975) Biochemistry 14, 3922–3926.
[2] Alakhov, Yu. B., Motuz, L. P., Stengrevics, O. A., Vinokurov, L. M. and Ovchinnikov, Yu. A. (1977) Bioorg. Khim. 3, 1333–1345.
[3] Alakhov, Yu. B., Stengrevics, O. A., Bundulis, Yu. P., Motuz, L. P. and Vinokurov, L. M. (1979) Bioorg. Khim. 5, 330–339.
[4] Alakhov, Yu. B., Stengrevics, O. A., Filimonov, V. V. and Venyaminov, S. Yu. (1979) Eur. J. Biochem. 99, 585 591.
[5] Alakhov, Yu. B., Zalite, I. K. and Kashparov, I. A. (1980) Eur. J. Biochem. 105, 531–538.
[6] Girshovich, A. S., Bochkareva, E. S., Pozdnyakov, V. A. and Ovchinnikov, Yu. A. (1978) FEBS Lett. 85, 283–286.
[7] Amber, R. P. (1967) Methods Enzymol. 11, 155–156.
[8] Gray, W. R. (1967) Methods Enzymol. 11, 469–475.
[9] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.