REPEATED SEQUENCES IN METHIONYL-tRNA SYNTHETASE FROM *E. COLI*

C. J. BRUTON, R. JAKES and G. L. E. KOCH
Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

Received 3 June 1974

1. Introduction

It has been argued that some of the diversity of the subunit structures and protomeric molecular weights of aminoacyl-tRNA synthetases is explained by repeated sequences in the larger protomeric units [1]. The methionyl-, valyl- and leucyl-tRNA synthetases* from *Bacillus stearothermophilus* with protomeric molecular weights of 66 000, 110 000 and 110 000 respectively were shown to contain long repeating sequences while the smaller tyrosyl-tRNA synthetase did not. Other reports, based on fingerprinting and peptide counting only, indicate that the isoleucyl- [2] and leucyl- [3] synthetases from *E. coli* with protomeric molecular weights of 102 000 and 100 000 respectively could have similar structural features.

The methionyl-tRNA synthetase from *E. coli* has a native mol. wt. of about 175 000 [4,5] and a subunit size of 85 000 [6]. An active fragment of mol. wt. 65 000 can be made by proteolytic digestion [7] and this fragment has been shown to be a single polypeptide chain with the same amino-terminal sequence as the native enzyme [6]. Hence it must be made by the removal of about two hundred amino acids from the carboxyl-terminus of the native molecule which clearly results in the dissociation of the subunits. In this paper we show that, like the enzyme from *B. stearothermophilus*, MTS from *E. coli* contains a significiant amount of repeated sequences.

* *Abbreviations:* MTS, methionyl-tRNA synthetase; VTS, valyl-tRNA synthetase; LTS, leucyl-tRNA synthetase.

2. Materials and methods

MTS was purified to homogeneity from *E. coli* EM 20031 as described previously [8].

Trypsin (2 × recrystallised) was purchased from Worthington Biochemical Corporation.

2.1. *Tryptic digestion of carboxymethyl-MTS*

Carboxymethyl-MTS was prepared by the method of Milstein [9]. The product was dialysed exhaustively against 20 mM ammonium bicarbonate and lyophilised. About 20 mg were dissolved in 20 ml 20 mM ammonium bicarbonate and 2.75 μmoles norleucine added as internal standard. A sample was taken for amino acid analysis to estimate accurately the amount of protein taken for the digestion. 0.1 ml of a solution of trypsin which had been further purified by absorption and elution from a soybean trypsin inhibitor–Sepharose affinity column [10] and had an E_{280} of 0.70 (equivalent to 0.5 mg/ml) was added. The mixture was incubated at 37°C for 8 hr and then lyophilised.

2.2. *Purification of some tryptic peptides*

The freeze-dried material was dissolved in 0.2 M pyridine acetate, pH 3.1 and the soluble material applied to a column of Locarte amino acid analyser resin L1/5 equilibrated with the same buffer. The column was washed with 60 ml of the same buffer and then with a gradient of pyridine acetate–156 ml 0.2 M pyridine acetate, pH 3.1 to 312 ml 2.0 M pyridine acetate, pH 5.0 [11]. One sixth of the effluent was monitored automatically by the ninhydrin reaction following alkaline hydrolysis and the remainder collected in fractions. These were pooled on the basis of the elution profile. Each of the pooled fractions

North-Holland Publishing Company – Amsterdam

were analysed by high voltage paper electrophoresis at pH 6.5 and pH 2.1 and by chromatography in butanol:acetic acid:water:pyridine [12]. Those fractions which appeared to contain a single peptide were hydrolysed with 6 M HCl overnight and subjected to amino acid analysis.

2.3. *Preparation of the tryptic fragment and the peptides released*

MTS (400 mg) was dialysed overnight against 0.1 M Tris–HCl, pH 7.5, containing 10 mM 2-mercaptoethanol. Trypsin (0.25 mg) was added and incubated at 37°C for 30 min. Soybean trypsin inhibitor (0.5 mg) was added and the material cooled rapidly to 4°C. The solution was concentrated by ultrafiltration and the filtrate collected. One-hundred ml 10 mM potassium phosphate pH 6.8 was added to the concentrate and ultrafiltration continued, again collecting the filtrate. The trypsin-modified enzyme was purified as described previously [7]. The filtrate was concentrated and equilibrated with 20 mM ammonia by passage through a column of Sephadex G-10. After lyophilisation the material was oxidised with performic acid as described by Hirs [13]. These peptides were separated exactly as described above.

3. Results and discussion

Table 1 shows the yields of some of the tryptic peptides obtained from native MTS corrected for the total recovery of norleucine. The yields were calculated from the content of the basic amino acid present in each case. That all these fractions did, indeed, contain single pure peptides was confirmed by establishing the amino acid sequences of the peptides which are quoted in table 1. Amide assignments were made on the basis of electrophoretic mobility [14]. Several other peptides were obtained in yields greater than one mole/85 000 daltons but only the most significant are quoted.

Table 1
Yield of tryptic peptides from native MTS

Sequence	Moles/85 000 daltons
Asn–Val–Phe–Ser–Gly–Ile–Arg	1.60
Gly–Met–Phe–Leu–Pro–Asp–Arg	1.48
Val–Asp–Leu–Arg	1.54
Leu–Leu–Arg	1.50
Leu–Val–Arg	1.50
Phe–Val–Lys	1.60
Asn–Arg	1.62

This is an analogous experiment to that described for MTS from *B. stearothermophilus* [1] with comparable results. In this case more peptides were obtained, the yields are much higher and the peptides have been sequenced to confirm homogeneity. These results indicate that the MTS of *E. coli* also contains several peptides more than once in its polypeptide chain.

The conversion of the native enzyme to an active trypsin-modified enzyme with the loss of about 200 amino acid residues allowed us to study the distribution of these repeats within the chain. If short sequences are repeated in tandem or close together, some should appear twice in the quarter of the molecule digested away and few, if any, should appear both in the trypsin-modified enzyme and among those released. Whereas if the repeating units are large, we expect that one copy of some of the high yield peptides should be released and both copies of some retained within the trypsin-modified enzyme. Those released should also appear in the trypsin-modified enzyme itself.

Table 2 shows the sequences of five of the peptides released in the preparation of the active trypsin-modified enzyme which were purified in good yield. The digestion conditions were chosen to be very mild and it was expected that partial cleavages would result. Furthermore the original isolation of the released peptides by ultrafiltration was designed to give optimal yields of the intact trypsin-modified enzyme and not of the peptides split off. Hence no attempt was made to accurately quantitate the yields of the peptides released but it was of the order of 0.5 moles/85 000

Table 2
Peptides released during very mild tryptic digestion

Asn–Val–Phe–Ser–Gly–Ile–Arg
Val–Asp–Leu–Arg
Leu–Val–Arg
Leu–Leu–Arg
Val–Leu–Ala–Ala–Arg

daltons in each case. The yield of the intact trypsin-modified enzyme was greater than 90% and hence this yield of these peptides could not come from a total digest of part of the material. Four of these peptides are indeed known to be present twice in the whole molecule (see table 1). These four peptides have also been isolated and sequenced from a tryptic digest of the trypsin-modified enzyme (C. J. Bruton, unpublished results). The fifth peptide Val–Leu–Ala–Ala–Arg was not present in high yield in the original digest of the native enzyme nor has it been found in the tryptic digest of the trypsin-modified enzyme. Presumably it comes from a non-repeated part of the sequence.

These results clearly indicate that MTS from *E. coli* has repeated sequences in different portions of the polypeptide chain very analogous to MTS, VTS and LTS from *B. stearothermophilus*. The results are perfectly compatible with the gene duplication and fusion hypothesis and the existénce of a peptide in only one copy suggests that the alternative explanations of post-transcriptional or post-translational fusion are probably unlikely. The final proof of this will, of course, come from the complete sequence of MTS which is well advanced.

References

[1] Koch, G. L. E., Boulanger, Y. and Hartley, B. S. (1974) Nature (in press).
[2] Kula, M.-R. (1973) FEBS Letters 35, 299.
[3] Waterson, R. M. and Konigsberg, W. H. (1974) Proc. Natl. Acad. Sci. U.S. 71, 376.
[4] Heinrikson, R. L. and Hartley, B. S. (1967) Biochem. J. 105, 17.
[5] Lemoine, F., Waller, J.-P., van Rapenbusch, R. (1968) Eur. J. Biochem. 4, 213.
[6] Koch, G. L. E. and Bruton, C. J. (1974) FEBS Letters 40, 180.
[7] Cassio, D. and Waller, J.-P. (1971) Eur. J. Biochem. 20, 283.
[8] Atkinson, A., Bruton, C. J. and Jakes, R. (1973) Abstr. FEBS Meeting, Dublin, No. 71.
[9] Milstein, C. (1966) Biochem. J. 101, 352.
[10] Porath, J. and Sundberg, L. (1971) in: Protides of the Biological Fluids (Peters, H., ed.), Vol. 18, p. 401. Pergamon Press, Oxford.
[11] Schroeder, W. A. (1967) in: Methods in Enzymology (Hirs, C. H. W., ed.), Vol. XI, p. 351. Academic Press, New York.
[12] Waley, S. G. and Watson, J. (1953) Biochem. J. 55, 328.
[13] Hirs, C. H. W. (1966) J. Biol. Chem. 219, 611.
[14] Offord, R. E. (1966) Nature 211, 591.