INTERNAL SEQUENCE HOMOLOGY IN *E. COLI* ISOLEUCYL- AND VALYL-tRNA SYNTHETASES

Frantisek KALOUSEK and William KONIGSBERG

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Conn. 06510, USA

Received 3 November 1975

1. Introduction

Aminoacyl-tRNA synthetases can be classified into four groups based on their subunit structure [1]. Five of the fifteen enzymes purified from *E. coli* consist of single polypeptide chains (α) with mol. wts around 100 000, approximately double the subunit mol. wt of synthetases composed of dimers (α_2) [2]. In view of the functional relationship among enzymes in this class it might not be unreasonable to expect that they could have evolved from a common ancestral gene. This notion was considered unlikely, however, because of the widely differing mol. wts among the protomers. Nevertheless we have recently found that leucyl-tRNA synthetase from *E. coli* had about half the number of lysine containing peptides expected on the basis of its mol. wt and amino acid composition. This finding was interpreted as suggestive evidence that leucyl-tRNA synthetase might have arisen via gene duplication and fusion [3]. In addition, Koch et al. [4] have demonstrated that internally repeated sequences occur in three aminoacyl-tRNA synthetases from *B. stearothemophilus.* We have extended these studies to valyl- and isoleucyl-tRNA synthetases from *E. coli* and have found that they also appear to have internal sequence repetition based on observed yields of peptides isolated after digestion with trypsin and chymotrypsin.

2. Material and methods

Isoleucyl-tRNA synthetase from *E. coli* K12 was purified as described by Durekovic et al. [5]. Valyl-tRNA synthetase was purified to homogeneity using polyethyleneimine precipitation of nucleic acids, DEAE-cellulose, hydroxylapatite chromatography and preparative gel electrophoresis. Seryl-tRNA synthetase was purified by a modified method of Katze and Konigsberg [6].

Carboxymethylation and succinylation of the proteins were carried out as described earlier [3]. Iodination was performed as described by Greenwood et al. [7] except that 6 M guanidinium chloride was used in the reaction mixture. Tryptic and chymotryptic digests were prepared by incubation of modified proteins with either enzyme in 100 mM ammonium bicarbonate at 25°C, pH 8.5 for 16 h at an enzyme to substrate ratio of 1 : 50. The chymotryptic digests of ^{14}C-succinylated seryl-tRNA synthetase was not suitable for the peptide mapping because of streaking and insolubility, so it was treated with thermolysin (1 : 100, 10 h at 20°C) to obtain smaller fragments. The total number of peptides was estimated by two-dimensional peptide mapping [3] using either silica gel or cellulose thin layer plates 20 × 20 cm. In a typical experiment 50 μg of ^{14}C-succinylated or 1 μg of iodinated protein was used and the peptides were detected by autoradiography. To determine yields, lysine containing peptides were isolated from 200 nmol of ^{14}C-succinylated protein. The chymotryptic digests were subjected to paper electrophoresis at pH 6.5. No detectable amount of insoluble material was found at the origin. After elution from paper, peptides were purified further by paper electrophoresis at pH 1.9 and pH 5.0 on silica-gel thin layer plates using electrophoresis at pH 6.5 in one direction and chromatography (butanol–acetic acid–water 3 : 1 : 1, v/v/v) in the other. The peptides were eluted with 2% formic acid and the yields calculated after amino

acid analysis. The purity of the peptides was evaluated by amino acid composition and in most cases by dansylation. The radioactive label was useful in correcting for losses during purification and also for the detection of very small amounts of material.

3. Results and discussion

We have previously reported evidence for internal sequence repetition in leucyl-tRNA synthetase from *E. coli* [3]. This conclusion was based on finding of only half the number of expected lysine containing peptides. Although the estimate of the number of peptides was consistent with the idea that the structural gene for this enzyme evolved by duplication and fusion, we wished to obtain more extensive evidence to support this conclusion. Koch et al. [4] concluded that valyl-tRNA synthetase and some others from *B. stearothemophilus* arose via gene duplication. Their inference was based on molar yields exceeding 100% of several peptides obtained from a tryptic digest of this enzyme. We also wished to determine the molar yields of peptides that could be obtained from chymotryptic and tryptic digests of two more (100 K dalton) aminoacyl-tRNA synthetases from *E. coli,* those which are specific for charging valine and isoleucine. Since only very small amounts of the highly purified enzymes could be conveniently prepared, the method of choice required that radioactive peptides be used for identification. Thus we reacted valyl- and isoleucyl-tRNA synthetases with carrier free ^{14}C-succinic anhydrydride which acylated the ε-amino groups of lysines and the NH_2-terminal amino groups, and with ^{125}I which iodinated the tyrosine residues. Chymotryptic digestion of these aminoacyl-tRNA synthetases was followed by separation of the peptides by thin layer chromatography. The number of ^{14}C-succinylated lysine and ^{125}I-tyrosine peptides from isoleucyl- and valyl-tRNA synthetases were compared with those obtained from seryl-tRNA synthetase. This enzyme, which is a dimer composed of identical 50 K dalton subunits, served as a control. The results presented in table 1 show that the expected number of tyrosine and lysine containing peptides were obtained from seryl-tRNA synthetase but that only half the expected number (based on the amino acid compositions of the 110 000 mol. wt. polypeptides) were obtained from valyl- and isoleucyl-tRNA synthetases. The amino acid compositions, NH_2-termini and molar yields of peptides which were isolated are given in table 2. Only ^{14}C-succinylated lysine containing peptides, where the yield exceeded one mole per mole of protomer, are listed although several other ^{14}C-succinylated-lysine containing peptides were isolated in reasonable (0.6–0.9 mol per protomer) yield. These may, in fact, have been present in quantities greater than one mole per mole of protomer but because of mechanical losses (trailing on the plates or incomplete digestion) appear in less than molar amounts. Since no corrections were applied except for losses during elution, the yields represent a minimum figure. When this procedure was applied to ^{14}C-succinylated lysozyme as a control, the yields ranged between 30 and 70 percent. Thus there are undoubtedly many other peptides from valyl- and isoleucyl-tRNA synthetase not listed in table 2 which are present as 2 copies per protomer.

All of these data strongly support the conclusions

Table 1
^{125}I-Tyrosine and ^{14}C-succinylated lysine containing peptides of seryl-, valyl- and isoleucyl-tRNA synthetases

	Tyrosine		Lysine	
Aminoacyl-tRNA synthetase	Number of peptides found	Residues per subunit[a]	Number of peptides found	Residues per subunit
Seryl-	12–14	12	22	22
Valyl-	23	39	22–27	47
Isoleucyl-	17	32	27–32	61

[a] Residues per subunit were taken from the following sources: seryl-tRNA synthetase [6], valyl-tRNA synthetase [9] and isoleucyl-tRNA synthetase [5].

Table 2
Amino acid composition of lysine peptides of valyl- and isoleucyl-tRNA synthetases

	Valyl-tRNA synthetase							Isoleucyl-tRNA synthetase				
	Peptide number							Peptide number				
Amino Acid	1	2	3	4	5	6	7	1	2	3	4	5
Lys	1	1	2	1	1	1	1	2	1	1	1	1
His	1							2			1	
Arg	2	2	3	2	3	1		1		2		
Asp	2		2	2	2		2	6	1	2		2
Thr	1	1	1	2	1		1	1	1			1
Ser	3	1	3	1	1	2	1	4		1		1
Glu	3		4	2	3		2	5	1	1		4
Pro		1		1		1	1	1		1		
Gly	2		2	2	1	2	3	8	1		1	
Ala	1		3	4	2	1	2	1	1		2	2
Cys		1									1	
Val		1	1	2	2	3	1	2	1		1	2
Met		1		1				1				
Ile	1	1	1	1			1	3	0	2	1	1
Leu	2	1	2	2	2	2		8				2
Tyr			1			1	1	1		1	1	1
Phe	1	1	1	1	1		1			1		
N-terminal	Ser	n.a	Ala	Leu	n.a	Val	Ile	Asp	Thr	n.a	Gly	Glu
Mol of peptide/mol of protomer	1.53	1.19	1.34	1.57	1.23	1.74	1.32	1.12	1.56	1.69	1.30	1.19

n.a., not analyzed.

arrived at by Kula [8] Waterson and Konigsberg [3] and Koch et al. [4] that the monomeric aminoacyl-tRNA synthetases of mol. wt. approx. 100 000 have internally duplicated sequences which probably arose via gene duplication and fusion, although, it should be noted, there is no evidence that would rule out posttranscriptional or posttranslational joining. We were not able to split either valyl- or isoleucyl-tRNA synthetases from *E. coli* into domains of 50 000 daltons by limited proteolysis as reported by Koch et al. [4] for valyl-tRNA synthetase from *B. stearothermophilus*. Nevertheless, it seems reasonable to assume that the tandemly duplicated sequences found in the *B. stearothermophilus* enzymes would also occur in *E. coli*.

Acknowledgement

This work was supported by grant GM 12607.

References

[1] Muench, K. H. and Joseph, D. R. (1971) in: Nucleic Acid–Protein Interaction, Miami Winter Symposium (Ribbons, D. W., Woessner, J. P. and Schultz, J., eds.) pp. 172–179, North-Holland Publishing Co., Amsterdam.
[2] Kalousek, F. and Konigsberg, W. (1975) in: Synthesis of Amino Acids and Proteins (Arnstein, H. R. V. ed.), pp. 57–88, MTP International Review of Science.
[3] Waterson, R. M. and Konigsberg, W. (1974) Proc. Natl. Acad. Sci. USA 71, 376–380.
[4] Koch, G. L. E., Boulanger, Y. and Hartley, B. S. (1974) Nature 249, 316–320.
[5] Durekovic, A., Flossdorf, J. and Kula, M. R. (1973) Eur. J. Biochem. 36, 528–533.
[6] Katze, J. R. and Konigsberg, W. (1970) J. Biol. Chem. 245, 923–930.
[7] Greenwood, F. D., Hunter, W. M. and Glover, J. S. (1963) Biochem. J. 89, 114–123.
[8] Kula, M. R. (1973) FEBS Lett. 35, 299–302.
[9] Yaniv, M. and Gros, F. (1969) J. Mol. Biol. 44, 1–15.