The Sequence of Nucleotides in $tRNA^{Ile}$ from *E. coli* B

By Michael Yarus
Department of Chemistry
University of Colorado
Boulder, Colorado 80302
USA

and B.G. Barrell
MRC Laboratory of Molecular Biology
Cambridge
England

Received March 23, 1971

Two major species of isoleucine acceptor RNA are found when *E. coli* B tRNA is fractionated on benzoylated DEAE[1] (unpublished data of M.Y.) or by column partition chromatography.[2] Here we present the nucleotide sequence of a pool of pure $tRNA^{Ile}$, in which two sequence variants were detected. Because of poor yields when small amounts of ^{32}P-$tRNA^{Ile}$ were chromatographed on long columns of benzoylated DEAE, it was not possible to separately purify these peaks. This gave rise to no ambiguities in the sequences, but means that the two sequence variants cannot be definitely identified with the two major peaks usually observed, though this seems a reasonable surmise.

A culture of *E. coli* B derived from a single cell was grown for several generations in a low phosphate medium[3] containing ^{32}P-phosphate of specific activity 6.1 curies/milliatom P. Cell tRNA + 5S RNA was extracted by treatment of cells with cold phenol[4], and then passed over an 0.8 x 1 cm benzoylated DEAE column in 1 M NaCl to remove large RNA and DNA fragments present as contaminants. Endogenous aminoacyl-tRNA was discharged and subsequent fractionation of phenoxyacetylisoleucyl-tRNA[5] on

a small amount of benzoylated DEAE and reverse phase chromatography (RPC-3[6]) gave a product in which, presumably, both major species were combined and which was >98% pure, as judged from the presence of contaminating oligonucleotides in nuclease digests. The structure of most pancreatic and T_1-RNAse-produced oligonucleotides was determined by methods previously published[7-9], and a total of 48 partial pancreatic and T_1RNAse digestion products, isolated by thin layer homochromotography[10], were degraded by the same methods to complete the structure and to order the products of complete digestion. All residues in the oligonucleotides from both the pancreatic and T_1RNAse digests were isolated after complete enzymatic degradation to nucleotides and characterized by paper electrophoresis at pH 3.5, chromatography on thin layers of cellulose in isopropanol-HCl-H_2O, and in isopropanol-H_2O-NH_3.[7] A coeval summary of these methods is in press.[11]

The sequence, containing 77 nucleotides, is shown in Figure 1. The structure of the 3' terminus has previously been studied[12], with results in agreement with those presented here. Of the two sequences detected, the more abundant has the D, rather than U, at position 17 from the 5' end. No other variations in the sequence were found. We therefore surmise that the existence of these two sequences does not necessarily imply two genes for $tRNA^{Ile}$, but that the rate of modification of this point in the D loop must lag the rate of production of $tRNA^{Ile}$.

The structure may easily be folded into a typical planar cloverleaf with an aminoacyl stem 7 basepairs long, two 5 base-paired stems for the TψC loop and anticodon and a 4 base pair stem for the dihydrouracil loop. The TψC and anticodon loops contain 7 residues, as in other known tRNA structures. (See figure) In spite of its local similarity in sequence to other

E. coli tRNA's, e.g., $tRNA_M^{Met}$ [13], the structure has uridylic acid rather than 4-thio-uridylic acid at the 8th position from the 5' end, and is therefore not subject to photochemical cross linking between 4-thio-uridylic acid and cytidylic acid at position 13.[14] It also has a minor nucleotide at a position at which a modification has not been seen before and closer to the 3' end than previously observed in any tRNA: the ψ at position 66 at the base of the TψC stem. The 5' terminus is pA rather than pG, as is more usual.

The nucleotide symbolized X in the figure is most probably a derivative of U which is multiply modified; it is degraded by both mild acid and standard alkaline hydrolysis. The native form probably has an extra negative charge, judging from the low mobility of oligonucleotides containing it during ionophoresis on DEAE paper at pH 3.5; after exposure to 7% formic acid, the resulting nucleotide is U-like in ionophoresis. However, it exhibits very high chromatographic mobility in ethanolic solvents, which suggests a very hydrophobic substituent, perhaps an ethyl or several methyl groups. It is resistant to pancreatic RNAse in both forms, which is a quality of N-3 substituted pyrimidines, but not usually those having C-4, N^4, or C-5 modifications, e.g., [15]. Alkaline hydrolysis gives a partial yield of a very fast product in ionophoresis, reminiscent of the product of ring cleavage in dihydropyrimidines.[16] These properties are singly those of previously known minor nuelcotides, but together they suggest a new compound whose structure must be complex; its elucidation will have to await direct study of larger amounts of material.

The nucleotide symbolized A* resembles an aminoacyl-adenylic acid[17]; its behavior in ionophoresis and chromatography is definitely inconsistent with a methyl-, isopentenyl-, or thiomethyl

6 to 8 min. This group is represented by bands 1a and 1b. Synthesis of early proteins, represented by bands 2a and 2b, begins between 4 and 6 min after infection and is substantially arrested by about 20 min. Synthesis of late proteins, represented by bands 3a, 3b, and 3c, begins between 10 and 12 min after infection and continues until lysis. This wild-type pattern of T5 protein synthesis is also found in D9·amH18a-infected cells (Plate 2), in which no detectable DNA synthesis takes place. The most striking result is, therefore, that late proteins are synthesized when no phage DNA synthesis occurs. It should also be noted that the shut-off of early protein synthesis (bands 2a and 2b in Plate 2) appears normal. Since the intensities of corresponding bands in Plates 1 and 2 are very similar, the differences in the amounts of protein synthesized in $T5^+$- and in mutant-infected cells is probably rather small.

DISCUSSION

Zweig et al. (11) reported that no empty T5 heads are produced in cells infected with $T5^+$ in the presence of 0.2 M hydroxyurea. Since phage DNA synthesis was not observed in these cells, it was suggested that the synthesis of T5 structural (i.e., late) proteins is linked to the synthesis of phage DNA. In view of our present results, we think the experiments of Zweig et al. probably indicate that the inhibition of protein synthesis and of DNA synthesis by high concentrations of hydroxyurea occur by independent mechanisms. The recent demonstration that empty T5 heads are produced in non-permissive cells infected with D9·amH18a (Zweig, personal communication) confirms the findings reported in this paper. Also, Dr. J. M. Buchanan (personal communication) has informed us that near-normal levels of T5-induced lysozyme (a late protein) accumulate in non-permissive cells infected with a polymeraseless mutant of T5. Thus, it seems quite clear that in T5-infected cells, the normal regulation of protein synthesis is not dependent upon phage DNA synthesis.

In the case of T4, the requirement of DNA synthesis for late protein synthesis can be circumvented by using multiple mutants that induce both defective

third codon for isoleucine, must be a minor species in E. coli B. This is consistent with the observed fact that AUU and AUC can easily be observed to bind isoleucyl-tRNA from mixed E. coli B tRNA's to ribosomes, but that results with AUA are negative.[21] It may also be significant that of five isoleucine codons identified in R17 RNA, all are AUU or AUC (personal communication from F. Sanger). In contrast, Torula $tRNA^{Ile}$ has the anticodon IAU^{16}, and thus can presumably respond to all three isoleucine codons. However, we observe a quantitative yield of GAU and no IAU, or I, or any other likely anticodon in E. coli $tRNA^{Ile}$. It therefore seems plausible that AUA may be a rarely used codon in E. coli B, though it apparently can be translated.[22]

Acknowledgements

I (M.Y.) would like to acknowledge the hospitality extended to me by Dr. Fred Sanger during this work. It was supported by USPHS GM-15925-03.

References

1. Gillam, I., Millward, S., Blew, D., von Tigerstrom, M., Wimmer, E. and Tener, G.M., Biochemistry, 6, 3043 (1967).

2. Meunch, K.H., and Berg, P., Biochemistry 5, 970 (1966).

3. Garen, A. and Levinthal, C., Biochim Biophys Acta 38, 470 (1960).

4. Zubay, G., J. Mol. Biol. 4, 347 (1962).

5. Gillam, I., Blew, D., Warrington, R.C., von Tigerstrom, M., and Tener, G.M., Biochemistry 7, 3459 (1967).

6. Weiss, J.F., Pearson, R.L., and Kelmers, A.D., Biochemistry 7, 3479 (1968).

7. Sanger, F., Brownlee, G.G., and Barrell, B.G., J. Mol. Biol. 13, 373 (1965).

8. Brownlee, G.G., and Sanger, F., J. Mol. Biol. 23, 337 (1967).

9. Brownlee, G.G., Sanger, F., and Barrell, B.G., J. Mol. Biol. 34, 379 (1968).

10. Brownlee, G.G., and Sanger, F., Eur. J. Biochem. 11, 395 (1969).

11. Barrell, B.G., in Procedures in Nucleic Acid Research, Cantoni, G.L. and Davies, D.R., editors, Harper and Row, N.Y. (1970).

12. Berg, P., Lagerkvist, U., and Dieckmann, M., J. Mol. Biol. 5, 159 (1962).

13. Cory, S., Marcker, K.A., Dube, S.K., and Clark, B.F.C., Nature 220, 1039 (1968).

14. Yaniv, M., Faure, A., and Barrell, B.G., Nature 223, 1331 (1969).

15. Brimacombe, R.L.C., Griffin, B.E., Haines, J.A., Haslam, W.J., and Reese, C.B., Biochem. 4, 2453 (1965); and Gassen, H.G. and Witzel, H., Eur. J. Biochem. 1, 36 (1967).

16. Cohn, W.E., and Doherty, D.G., J. Amer. Chem. Soc. 78, 2863 (1956).

17. Schweizer, M.P., Chheda, G.B., Baczynskyj, L., and Hall, R.H., Biochemistry, 8, 3283 (1969).

18. Takemura, S., Murakami, M., and Miyazaki, M., J. Biochem. (Tokoyo) 65, 553 (1969).

19. Ishikura, H., Yamada, Y., Murao, K., Saneyoshi, M., and Nishimura, S., Biochem. Biophys. Res. Commun. 37, 990 (1969).

20. Crick, F.H.C., J. Mol. Biol. 19, 548 (1966).

21. Söll, D., Ohtsuka, E., Jones, D.S., Lohrmann, R., Hayatsu, H., Nishimura, S., and Khorana, H.G., Proc. Nat. Acad. Sci. 54, 1378 (1965).

22. Gardner, R.S., Wahba, A.J., Basilio, C., Miller, R.S., Lengyel, P., and Speyer, J.F., Proc. Nat. Acad. Sci. 48, 2087 (1962).