Structure and transcription of the tRNA\textsubscript{Pro} gene from Escherichia coli

Yoshiyuki Kuchino, Fumiko Mori and Susumu Nishimura

Biology Division, National Cancer Center Research Institute, 5-1-1, Tsukiji, Chuo-ku, Tokyo 104, Japan

Received 5 February 1985; Revised and Accepted 10 April 1985

ABSTRACT

A 5 kbp DNA fragment containing the tRNA\textsubscript{Pro} gene from Escherichia coli was cloned into Charon 21A phage and sequenced by the M13 DNA sequencing technique. When the cloned DNA fragment was used as a template for in vitro transcription with E. coli RNA polymerase, a tRNA\textsubscript{Pro} precursor of 120 nucleotide residues was obtained. The tRNA\textsubscript{Pro} gene transcribed as a single transcription unit was followed by two unusual repeating sequences, both of 108 bp. These two repeating sequences were separated by a 60 bp spacer sequence. The 5'-portion of each repeating sequence overlapped 19 bp of the 3'-terminal region of the tRNA\textsubscript{Pro} gene just like the repeating sequences in the E. coli tRNA\textsubscript{ tyr } gene. A rho-independent termination was present in the first repeating unit.

INTRODUCTION

We previously reported that Salmonella typhimurium and Escherichia coli K12 contain three isoaccepting species of proline tRNA (1). The proline tRNA species from the two bacteria have identical nucleotide sequences but differ in the extents of their post-transcriptional modification. Riddle and Roth reported that sufA and sufB mutations in the Salmonella genome induced by ICR-191 suppress a mutation caused by a single base insertion in the histidine operon (2). These suppressive mutations are dominant and affect the formations of Salmonella ProI and ProII tRNAs, respectively (3). Judging from the reported column chromatographic profile of Salmonella proline tRNAs on BD-cellulose (3) and the results of our sequence analyses of these proline tRNAs (1), the ProI and ProII tRNAs that are products of the Salmonella sufA and sufB genes, seem to correspond to tRNA\textsubscript{Pro}\textsubscript{1} and tRNA\textsubscript{Pro}\textsubscript{2} containing 5'CGG3' and 5'GGG3', respectively, as their anticodon sequence.

Since the elution profile of ProIII tRNA on BD-cellulose column chromatography is not affected by the sufA and sufB mutations, it must correspond to our tRNA\textsubscript{Pro}\textsubscript{3} containing the anticodon sequence 5'VGG3'. Hsu et
al. (4) cloned and sequenced the \textit{E. coli} tRNA gene corresponding to tRNA$_{3}^{\text{Pro}}$. The \textit{E. coli} tRNA$_{3}^{\text{Pro}}$ gene is a member of the arginine tRNA gene operon and is transcribed together with three other tRNA genes as a multicistronic transcript.

In this communication, we report the nucleotide sequence and the structural features of the \textit{E. coli} tRNA$_{1}^{\text{Pro}}$ gene. These data are necessary for future studies on the mechanism of the suppressive effect of the \textsc{sufA} mutation.

**MATERIALS AND METHODS**

\textit{E. coli} K12 tRNA$_{1}^{\text{Pro}}$, used as a probe in nucleic acid hybridization experiments, was prepared and purified as described previously (1). Its tRNA 3’-terminus was labeled using T4 RNA ligase and [5'-$^{32}$P]pCP by the procedure of Bruce and Uhlenbeck (5). Chromosomal DNA from \textit{E. coli} K12 was prepared by the method of Berns and Thomas (6). Sucrose density gradient centrifugation, for fractionation of DNA fragments in digests with restriction endonucleases, was performed at 26,000 rpm for 24 hr at 15°C in a Hitachi RPS27 rotor in a gradient of 10-40% sucrose in 20 mM Tris-HCl buffer (pH 8.0) containing 1M NaCl and 5 mM EDTA. Nucleic acid hybridization was carried out in 5xSSC (1xSSC is 0.015 M sodium citrate (pH 7.0), 0.15 M sodium chloride) and 50% formamide at 42°C for 16 hr. Labeled nucleic acids were used at a radioactivity of $10^{-5} - 10^{-6}$ cpm per ml of hybridization solution. After hybridization, the nitrocellulose filters were washed successively at room temperature for 30 min each with 6xSSC and 2xSSC, and then air-dried and autoradiographed. DNA sequence analysis was performed by the dideoxy termination method (7). In vitro transcription experiments were performed in reaction mixture (30 µl) containing 5 µg of DNA fragments, 0.02 M Tris-acetate (pH 7.9), 0.1 mM EDTA, 4 mM magnesium acetate, 0.1 mM dithiothreitol and 0.1 M KCl. Transcripts were labeled with 5 µCi [$\alpha$-$^{32}$P]GTP (400 Ci/mmol), 3 µCi [$\gamma$-$^{32}$P]GTP (14.9 Ci/mmol) or 10 µCi [$\gamma$-$^{32}$P]ATP (8000 Ci/mmol). The remaining three nucleotide triphosphates (0.2 mM each) and \textit{E. coli} RNA polymerase (1 µg) were then added to the reaction mixture, and incubation was carried out for 20 min at 37°C. After incubation, the reaction mixture was chased for 20 min at 37°C by addition of 1 µg of \textit{E. coli} RNA polymerase and 0.15 mM nucleotide triphosphate of the same species as the labeled compound. Transcription was stopped by addition of 1/10 volume of 3 M sodium acetate and 50 µg of carrier tRNA. The RNA transcripts were extracted with phenol, precipitated with 3 volumes of

3214
ethanol, and separated by electrophoresis on 6% polyacrylamide slab gel (20x40x0.05 cm) containing 7 M urea at 30 mA in TEB buffer (90 mM Trizma base, 90 mM boric acid, 4 mM EDTA Na₂).

RESULTS

When E. coli chromosomal DNA was cleaved with EcoRI, BamHI and/or HindIII, only a single DNA fragment was obtained that hybridized to E. coli tRNA<sub>Pro</sub> (data not shown). The EcoRI-HindIII fragments of 5 kbp containing DNA complementary to tRNA<sub>Pro</sub> were purified by sucrose density gradient centrifugation and then inserted into phage vector Charon 21A DNA. The recombinant DNA was packaged <em>in vitro</em> into phage capsids using packaging lysates purchased from Promega Biotec. E. coli K802 cells were transfected with the recombinant phages obtained by <em>in vitro</em> packaging. By screening about 5000 transformants by plaque hybridization with 3'-end labeled E. coli tRNA<sub>Pro</sub> as a probe, four clones carrying the tRNA gene were obtained.

The phage DNAs extracted from these four transformants exhibited the same restriction endonuclease fragment profile. The sequence corresponding to tRNA<sub>Pro</sub> in the DNA was identified by Southern hybridization analysis as the 0.9 kbp Sau3A fragment and the 0.3 kbp HpaII fragment (Fig. 1). For determination of the nucleotide sequence of the tRNA<sub>Pro</sub> gene, fine restriction mapping was performed with several restriction endonucleases. The restriction map and the strategy for sequence analysis of the Sau3A

![Restriction enzyme map and sequencing strategy for the 1.3 kbp DNA fragment containing the E. coli tRNA<sub>Pro</sub> gene.]

The shaded portion of the map indicates the tRNA<sub>Pro</sub> gene region. The arrow above the tRNA<sub>Pro</sub> gene indicates the direction of transcription. The cleavage sites of restriction endonucleases are indicated by vertical bars in the map. The arrows in the lower portion of the figure represent the length and direction of the sequence analysis of each fragment.
Figure 2. Nucleotide sequence of the antisense strand of the 1.3 kbp DNA fragment containing the tRNA-pro gene.

The DNA sequence was determined by the dideoxy termination method using M13mp 8 and M13mp 9 as cloning vectors. The tRNA-pro gene region is boxed. The wavy lines indicate the -35 domain and Pribnow sequence (-10 domain). The region exhibiting dyad symmetry is indicated by horizontal arrows. The vertical arrow at nucleotide 243 and the first T cluster overlined with dots represent the transcription initiation and termination sites of the tRNA-pro gene, respectively. The repeating sequences are underlined. The region between nucleotides 373 and 720, indicated by asterisks, shows the possible open reading frame predicted from the DNA sequence.

The complete nucleotide sequence of the 1.3 kbp Sau3A DNA fragment is shown in Fig. 2. The sequence corresponding to tRNA-pro was found between nucleotides 248 and 324. TTGACG and TAAGATG sequences were found 37 and 14 nucleotides upstream from the 5'-terminus of the mature tRNA sequence. These sequences are both consistent with the consensus sequences that function as promoters found in the regions 35 and 10 upstream of eubacterial...
For identification of the transcription and termination sites of the tRNA\textsuperscript{Pro} gene, \textit{in vitro} transcription was carried out with the 5 kbp DNA insert as a template for E. coli RNA polymerase as described in the Methods. Products were separated by polyacrylamide gel electrophoresis. As shown in Fig. 3 (1,2), transcripts of about 120 and 1500 nucleotides were detected. Transcription of the smaller RNA species was initiated with GTP, because the
Fig. 4. Fingerprint of RNase T1 digests of $^{32}$P-labeled small RNA synthesized in an in vitro transcription reaction (A), and the predicted fingerprint from the DNA sequence corresponding to the region of nucleotides 243 to 352 (B). Labeled small RNA separated on polyacrylamide gel as shown in Fig. 3 was extracted with buffer containing 0.5 M ammonium acetate, 0.01 M magnesium acetate, 0.1 mM EDTA and 0.1% SDS, then precipitated with carrier tRNAs and ethanol, and finally subjected to RNase T1 digestion. Two dimensional ionophoresis was performed by the procedure of Barrell (15). Separation in the first dimension, from right to left, was on cellulose acetate at pH 3.5 and in the second, from top to bottom, was on DEAE-paper in 7% formic acid.

RNA was labeled with [$\gamma$-$^{32}$P]GTP, but not with [$\gamma$-$^{32}$P]ATP (Fig. 3 (3-6)). On the contrary, the larger transcript was not labeled with either [$\gamma$-$^{32}$P]GTP or [$\gamma$-$^{32}$P]ATP. Since transcription of almost all E. coli genes is initiated at either a G or A residue (8), the small RNA transcript was analysed further. Fig. 4a shows a fingerprint of an RNase T1 digest of the small RNA labeled with [$\alpha$-$^{32}$P]GTP. This fingerprint profile coincided completely with that predicted from the DNA sequence between nucleotides 243 and 352 (Fig. 4b). Nearest neighbor analysis of $^{32}$P-labeled RNase T1 digests (data not shown) supported the conclusion that the small RNA is derived from nucleotides 243 to 359. These results indicate that transcription of the tRNA$^{Pro}$ gene was initiated at the G residue at nucleotide 243 and the terminated within the T cluster preceded by the dyad symmetry of nucleotides 354 to 359.
DISCUSSION

In this work we cloned and sequenced the proline I tRNA gene from E. coli K12. The DNA sequence of the tRNA structural gene was found to be identical with that of tRNA\(^{\text{Pro}}\) reported previously (1). In \textit{in vitro} transcription experiments with \textit{E. coli} RNA polymerase, the cloned 5 kbp DNA fragment produced two RNA transcripts of 120 and 1500 nucleotides (Fig. 3). Fingerprint analysis of RNase T\(_1\) digests showed that the RNA of 120 nucleotides was the product of the tRNA\(^{\text{Pro}}\) gene. The transcription initiation site of the tRNA\(^{\text{Pro}}\) gene was shown to be 5 bp before the 5'-terminus of the mature tRNA sequence, and its termination site to be within a T cluster located 30 bp downstream of the 3' end of the mature tRNA sequence.

The tRNA\(^{\text{Pro}}\) gene was transcribed as a single transcription unit. In general, \textit{E. coli} tRNA genes form parts of multimeric transcription units containing several tRNA genes, although there are a few exceptions, such as the tRNA genes for arginine, methionine and serine tRNA (9-11). These exceptional single tRNA genes comprise operons with other genes, such as tRNA\(^{\text{Met}}\) with \textit{nusA} (10) or tRNA\(^{\text{Ser}}\) with \textit{divE} genes (11). From this point of view, it is noteworthy that the tRNA\(^{\text{Pro}}\) gene is also followed by an open reading frame consisting of 115 amino acids (nucleotides 373 to 720). If this sequence is actually transcribed \textit{in vivo}, the size of its transcript could be more than 400 bp. As shown in Fig. 3, a transcript of 1500 nucleotides was synthesized in an \textit{in vitro} transcription reaction with the 5 kbp cloned DNA fragment as a template (Fig. 3). However, this transcription product was not labeled with [\(\gamma\)\(^{\text{32}}\)P]GTP or [\(\gamma\)\(^{\text{32}}\)P]ATP. From these results, this 1500 nucleotide transcript could be the product of a gene located in a different transcription unit or be an artifact synthesized \textit{in vitro}.

Another interesting feature of the tRNA\(^{\text{Pro}}\) gene is the presence of the two repeating sequences of 108 bp containing a rho-independent termination signal. These sequences are arranged tandemly in the 3'-distal portion of the tRNA gene, and separated by a spacer sequence of 60 bp. The first 19 bp of each repeating sequence overlaps the 3'-portion of the tRNA structural gene. The significance of these repeating sequences in the tRNA\(^{\text{Pro}}\) gene is unknown, but it is interesting that a similar structural feature has been found in the \textit{E. coli} tRNA\(^{\text{Tyr}}\) gene (12), namely, four repeating sequences of 178 bp arranged tandemly with no spacer, and with a rho-dependent transcription termination in the second repeating unit.

A stem and loop structure was found in the 5'-control region of the
tRNA\textsubscript{Pro} gene (Fig. 2). A similar dyad symmetrical structure located in the region between the 35 domain and the transcription initiation site has been found in some other E. coli genes, such as the arginine tRNA gene operon and the alaS gene (13). In the alaS gene, the symmetrical structure acts as the binding site of the repressor protein. Therefore, the dyad symmetrical structure in the 5'-control region of the tRNA\textsubscript{Pro} gene may also be important in expression of the tRNA gene.

Riddle and Roth reported that the sufA gene of S. typhimurium is the structural gene for proline I tRNA and is located near the xyl\textsubscript{1} gene at 79 min (14). Frameshift suppressor proline I tRNA, which is the product of the mutated sufA gene, suppresses a mutation produced by insertion of an extra C into the Salmonella histidine operon (2). Since the anticodon sequence of normal proline I tRNA is $5'$CGG$3'$, the anticodon sequence of the frameshift suppressor proline I tRNA that suppresses the four base codon with an extra C might be $5'$CGGG$3'$.

For clarification of the molecular mechanism of frameshift suppression in sufA mutation, we are now determining the nucleotide sequences of suppressor proline I tRNA and the corresponding tRNA gene.

ACKNOWLEDGEMENTS

This work was partly supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

REFERENCES