Identity determinants of *E. coli* tryptophan tRNA

Hyouta Himeno, Tsunemi Hasegawa, Haruichi Asahara, Koji Tamura and Mikio Shimizu
Institute of Space and Astronautical Science, 3-1-1 Yoshinodai, Sagamihara, Kanagawa 229, Japan

Received September 27, 1991; Revised and Accepted November 13, 1991

ABSTRACT

The first base pair of the acceptor stem A₁-U₇₂ and the discriminator base G₇₃, as well as the anticodon nucleotides, characterize the tryptophan tRNA in *E. coli*. To determine the contribution of these nucleotides to the tryptophan acceptor activity, various transcripts of *E. coli* tryptophan tRNA mutants were constructed. Substitutions of the discriminator base G₇₃, which is conserved within prokaryotic tryptophan tRNAs, impaired aminoacylation with tryptophan. Substitutions of other purine-pyrimidine pairs for A₁-U₇₂ revealed that only U₇₂ weakly contributed to recognition by tryptophanyl-tRNA synthetase. The *E. coli* aspartic acid tRNA transcript introducing the tryptophan anticodon CCA showed almost the same tryptophan charging activity as the tryptophan tRNA transcript possessing a G₁-C₇₂ base pair. Only a low activity was detected in the mutant tryptophan tRNA transcript possessing a set of G₁-C₇₂ and A₇₃, which is observed in eukaryotic tryptophan tRNAs. These results indicate that the anticodon and G₇₃ are major identity determinants of tryptophan tRNA in *E. coli*, whereas the A₁-U₇₂ base pair is only a weak recognition element.

INTRODUCTION

The discrimination among tRNAs by cognate aminoacyl-tRNA synthetases is responsible for translational fidelity (1,2). Most tRNA molecules have apparently indistinguishable L-shaped tertiary structures, giving rise to difficulties in discrimination by synthetases. The anticodon sequence has been identified as a major acceptor identity element in several tRNAs (3—12). The region around the CCA 3' end is also most likely as a recognition element for aminoacyl-tRNA synthetase, since it is in proximity to the catalytic site. The fourth base from the 3' end (discriminator base G₇₃) is phylogenetically well conserved within every amino acid-specific tRNA isoacceptor species (13), and is essential for recognition by cognate synthetase in several tRNAs (5—12). The first base pair of the acceptor stem next to the discriminator base, and A₁-U₇₂ base pair appears only in two amino acid specific tRNAs including tRNA^Trp^ (13). Previous studies using suppressor tRNAs have suggested that the anticodon of *E. coli* tRNA^Trp^ is responsible for discriminating it from tRNA^Gln^ (18—20). In this study, we investigated the identity elements of tRNA^Trp^, especially focusing on the anticodon, discriminator base and A₁-U₇₂ base pair, using various in vitro transcript tRNAs.

MATERIALS AND METHODS

Preparation of template DNAs and in vitro transcripts

Synthetic DNA oligomers carrying the T7 promoter and tRNA genes were ligated into pUC19 and transformed into *E. coli* strain JM109 (15,21). The template DNA sequences were confirmed by dideoxy sequencing (22). Each template DNA of the base substituted mutant at position 1, 72 or 73 was prepared from a plasmid carrying the wild type tRNA sequence and two synthetic primers by mutation by polymerase chain reaction (PCR) (9,23). Transcripts of the tRNA genes were prepared in a reaction mixture containing 40 mM Tris·HCl (pH 8.1), 5 mM dithiothreitol, 2 mM spermidine, 10 mM magnesium chloride, bovine serum albumin (50 mg/ml), 2.0 mM NTPs, 20 mM 5' GMP, BstNI-digested template DNA (0.2 mg/ml), 2 units of inorganic pyrophosphatase (Sigma) and pure T7 RNA polymerase (50 μg/ml) (9,21,24). Transcripts initiated with an A were prepared in a reaction mixture containing 20 mM 5' AMP instead of 5' GMP (15). The transcripts were purified by 20% polyacrylamide gel electrophoresis (PAGE).

Aminoacylation assay

The aminoacylation reaction was carried out at 37°C in 40 μl of reaction mixture containing 60 mM Tris·HCl (pH 7.5), 10 mM magnesium chloride, 1 mM dithiothreitol, 0.1 mg/ml bovine serum albumin, 2.5 mM ATP, 8.2 μM L-[U-¹⁴C]tryptophan (611 mCi/mmol, New England Nuclear), 0.375 mM transcript RNA and various concentrations of tryptophanyl-tRNA synthetase partially purified from *E. coli* strain Q13. Eight microliter aliquots were spotted on Whatman 3MM filter paper, and cold 5% trichloroacetic acid insoluble fractions were counted in a liquid scintillation counter. One unit of the enzyme activity was defined as the amount of enzyme which catalyzes the incorporation of 1 nmol of tryptophan into tryptophanyl-tRNA in 10 min. Native tryptophan tRNA (1650 pmol/A²₆₀) was obtained from Subrinden RNA.

RESULTS

Several variants of *E. coli* tRNA^Trp^ transcripts were prepared, in order to examine the recognition elements of tRNA^Trp^ (Fig. 1). The unmodified transcript possessing a wild-type
tRNA\textsuperscript{Trp} sequence showed similar tryptophan charging activity to the native tRNA\textsuperscript{Trp} with only a slightly lesser initial velocity (Fig. 2).

\textit{E. coli} tRNA\textsuperscript{Trp} possesses G at the discriminator position, and A-U at the first base pair of the acceptor stem (Fig. 1). The effects of base substitutions at the discriminator position and the first base pair of the acceptor stem are shown in Figures 2 and 3, respectively. Substitutions of any other bases for G\textsubscript{73} impaired aminoacylation with tryptophan. In particular, substitution of either A or U resulted in a marked decrease in tryptophan charging activity. By contrast, the effects of base (pair) substitutions at the first base pair were apparently small, and therefore the initial velocities of the mutants were measured under the assay condition of five-fold lower enzyme concentration (Fig. 3). Substitution of G-U for A\textsubscript{72} caused little damage in tryptophan charging activity, whereas that of either A-C or G-C resulted in an approximately 2.5-fold decrease in the activity.

Although the mutations at the first base pair were confined to purine-pyrimidine pairs, these results indicate that G\textsubscript{73} is an identity element of tryptophan tRNA in \textit{E. coli}, and that U\textsubscript{72}, but not G\textsubscript{1}, is a weak recognition element.

Unlike the prokaryotic tRNA\textsuperscript{Trp}, eukaryotic tRNA\textsuperscript{Trp} possesses A at the discriminator position and G-C at the first base pair of the acceptor stem (13). As expected from the above results, the mutant \textit{E. coli} tRNA\textsuperscript{Trp} transcript possessing a set of A\textsubscript{73} and G\textsubscript{1}-C\textsubscript{72} was not a good substrate for \textit{E. coli} tryptophanyl-tRNA synthetase (Fig. 3).

To determine the contribution of the anticodon to the tryptophan acceptor identity, a transcript of the \textit{E. coli} tRNA\textsuperscript{Asp} mutant possessing the tryptophan anticodon CCA was prepared (Fig. 1), which had almost the same tryptophan charging activity as the tRNA\textsuperscript{Trp} transcript possessing G\textsubscript{1}-C\textsubscript{72} (Fig. 4).

Figure 1. The tRNA\textsuperscript{Trp} and tRNA\textsuperscript{Asp} transcripts folded into the cloverleaf structure. Arrows indicate the substitutions made in this study. Numbering is according to ref. 13.

Figure 2. Aminoacylation of native \textit{E. coli} tRNA\textsuperscript{Trp}, the tRNA\textsuperscript{Trp} transcripts possessing the wild type sequence (G73), and possessing A (A73), U (U73) and C (C73) at the discriminator position. The reaction mixture contains 0.375 \textmu M transcript RNA and 0.01 unit of tryptophanyl-tRNA synthetase fraction.

Figure 3. Aminoacylation of the tRNA\textsuperscript{Trp} transcripts possessing the wild type sequence (A1U72), possessing G-U (G1U72), A-C (A1C72) and G-C (G1C72) at the first base pair of the acceptor stem, and the mutant tRNA\textsuperscript{Trp} transcript simultaneously possessing G\textsubscript{1}-C\textsubscript{72} base pair and A\textsubscript{73} (G1C72A73). The reaction mixture contains 0.375 \textmu M transcript RNA and 0.002 units of tryptophanyl-tRNA synthetase fraction. The inset shows the aminoacylation with 0.01 unit of tryptophanyl-tRNA synthetase fraction.

Figure 4. Aminoacylation of the wild type tRNA\textsuperscript{Trp} transcript and the tRNA\textsuperscript{Asp} transcript possessing the tryptophan anticodon. The reaction mixture contains 0.375 \textmu M transcript RNA and 0.01 unit of tryptophanyl-tRNA synthetase fraction.
DISCUSSION

The unmodified transcript with the wild type sequence of tRNA\textsuperscript{Glu} was a good substrate for \textit{E. coli} tryptophanyl-tRNA synthetase, although \textit{E. coli} tRNA\textsuperscript{Glu} possesses nine modified nucleotides, 4-thiouridine (S\textsuperscript{4}U) at position 8, dihydrouridine (D) at 16, 17 and 20, 2'-O-methylcytidine (Cm) at 32, 2-methylthio-N6-isopentenyl-adenosine (m\textsuperscript{2}iP\textsuperscript{A}) at 37, 7-methylguanosine (m\textsuperscript{7}G) at 46, 5-methyluridine (T) at 54 and pseudouridine (Y) at 55 (Fig. 1) (25). Therefore, it is considered that these modifications are not critical for recognition by tryptophanyl-tRNA synthetase as in other systems reported (5-12, 15-17, 21). Modifications sometimes have the potential to exclude non-specific recognition by other non-cognate aminoacyl-tRNA synthetases as shown in the unmodified yeast tRNA\textsuperscript{Asp} transcript with yeast arginyl-tRNA synthetase (26). In the present study, the \textit{E. coli} tRNA\textsuperscript{Asp} transcript showed no detectable tryptophan charging activity (data not shown).

The anticodon, which characterizes each amino acid specificity of tRNA, is most likely to be involved in determining its acceptor identity, and has already been identified as a major identity determinant in several tRNAs (3-12). In terms of tRNA\textsuperscript{Glu}, a C to U mutation at the second position of the anticodon (su\textsuperscript{7} mutation) causes a looseness of the acceptor specificity \textit{in vivo} (18) and a dramatic decrease of the aminoacylation activity with tryptophan \textit{in vitro} (19). The present study shows that the tRNA\textsuperscript{Asp} transcript replacing the tryptophan anticodon appreciably acquires a tryptophan acceptor identity, establishing the anticodon sequence as a major identity determinant of tRNA\textsuperscript{Glu}.

Nucleotides neighboring the CCA 3' end are also likely to be involved in the aminoacylation reaction. The discriminator base distribution among twenty amino acid specific tRNAs is greatly biased in relation to the chemical property of the side chain of the charged amino acid (27). Most hydrophobic amino acid specific tRNAs possess A at this position, while G is occupied for tRNA\textsuperscript{ser}, tRNA\textsuperscript{Gln}, tRNA\textsuperscript{Asp}, tRNA\textsuperscript{Glu}, and tRNA\textsuperscript{Asp} in \textit{E. coli} (13). Among them, the G\textsubscript{73} of tRNA\textsuperscript{Gln} (12) and tRNA\textsuperscript{Asp} (16) have been shown to be important for aminoacylation. The distribution of the first base pair of the acceptor stem is also biased (13). In \textit{E. coli}, up to fifteen amino acid specific tRNAs possess a G\textsubscript{1}-C\textsubscript{2} base pair. A U\textsubscript{1}-A\textsubscript{2} base pair appears in tRNA\textsuperscript{Asp} and tRNA\textsuperscript{Glu}, C\textsubscript{1}-G\textsubscript{2} in tRNA\textsuperscript{Pro}, and C\textsubscript{1}-A\textsubscript{2} in initiator tRNA\textsuperscript{Met}. The A\textsubscript{1}-U\textsubscript{2} pair appears only in tRNA\textsuperscript{His} and tRNA\textsuperscript{Glu}, the former possessing A at the discriminator position (13). These arrangements suggest the involvement of the three nucleotides at positions 1, 72 and 73 in discriminating tRNA\textsuperscript{Glu} from other tRNA species. Our results show that the G\textsubscript{73} is an identity determinant of tryptophan tRNA in \textit{E. coli}, whereas the A\textsubscript{1}-U\textsubscript{2} base pair is only a weak recognition element.

Recent accumulation of data on the sequences and three-dimensional structures of aminoacyl-tRNA synthetases has revealed that the catalytic domains of the synthetases are mainly divided into two classes (28). The crystal structures of \textit{E. coli} tRNA\textsuperscript{Gln} (29) and yeast tRNA\textsuperscript{Asp} (30) complexed with their cognate synthetases belonging to the two different classes (class I and class II, respectively) have been determined at 2.8-3.0 Å resolutions. Both tRNAs share G\textsubscript{73} and U\textsubscript{1}-A\textsubscript{2} base pair, whereas the modes of their contributions are dissimilar between the two complexes. In the \textit{E. coli} tRNA\textsuperscript{Gln} complex, the base stacking between G\textsubscript{73} and C\textsubscript{35}, and the intramolecular hydrogen bond formation between 2-amino group of G\textsubscript{73} and the phosphate of A\textsubscript{72} cause a hairpin turn of the 3' CCA tail towards the active pocket of the enzyme in which U\textsubscript{1} and A\textsubscript{2} make intermolecular hydrogen bonds with several residues on the enzyme instead of base pairing with each other (29). In contrast, the yeast tRNA\textsuperscript{Asp} complex shows no such hairpin turn (30), although substitution of any of the other three bases for G\textsubscript{73} impairs the aminoacylation activity (11). In addition, no apparent disruption of the first base pair has been observed, although the enzyme accesses the major groove of the acceptor helix (30). In the former system, base substitution of the other purine A for G\textsubscript{73} does not affect the activity as much as that of pyrimidines which are considered to prevent formation of a hairpin turn (12). However, this type of logic does not apply to the tRNA\textsuperscript{Glu} system in this study, because the mutant transcript possessing A\textsubscript{73} was not a good substrate for tryptophanyl-tRNA synthetase which is also classified as a class I synthetase. The present study also suggests that tryptophanyl-tRNA synthetase interacts with the 3' but not the 5' side of the terminal region of the acceptor helix.

The discriminator base is widely conserved within every amino acid charged among prokaryotes and eukaryotes, although with some exceptions including tRNA\textsuperscript{Glu} (13). The eukaryotic tRNA\textsuperscript{Glu} possesses A at the discriminator position and G-C at the first base pair of the acceptor stem, of which neither were preferable for recognition by \textit{E. coli} synthetase (Figs. 2 and 3). Histidine tRNA, in which the discriminator base is faced with a 5' terminal nucleotide (G\textsubscript{1}) exceptionally longer than other tRNA species, also possesses different types of discriminator bases in prokaryotes and eukaryotes (13). The eukaryotic type of this extra base pair, G\textsubscript{-1}-A\textsubscript{2}, is a less appropriate substrate for \textit{E. coli} histidyl-tRNA synthetase than the original G\textsubscript{-1}-C\textsubscript{2} base pair, but is somewhat better than the G\textsubscript{-1}-G\textsubscript{73} or G\textsubscript{-1}-U\textsubscript{73} pairs (15). The present study reveals that the eukaryotic type of discriminator base A\textsubscript{73} is less favorable for \textit{E. coli} tryptophanyl-synthetase than C\textsubscript{73}. Besides, the transcript simultaneously introducing both eukaryotic types of the discriminator base and the first base pair was also not a good substrate for the \textit{E. coli} enzyme (Fig. 3). These results indicate that the mode of tRNA\textsuperscript{Glu} recognition at the terminal region of the acceptor stem including the discriminator base has significantly diverged during evolution into eukaryotes and prokaryotes. Such a divergence of recognition site is also observed in N\textsubscript{23} of tRNA\textsuperscript{Pro} (6, 21). These contrast sharply with the discriminator base G\textsubscript{73} of tRNA\textsuperscript{Asp} (11,16) and the G\textsubscript{1}-U\textsubscript{2} base pair in the acceptor stem of tRNA\textsuperscript{Ala} (31,32), both are conserved as an identity determinant among prokaryotes and eukaryotes.

Class I aminoacyl-tRNA synthetases contain two major HI(L)GH and KMSKS consensus sequences (33). The latter sequence, which has been believed to be involved in 3' CCA binding (34,35) and/or in aminoacyl-adenylate formation (36), is modified to KMSAS in mammalian tryptophanyl-tRNA synthetase (37), and the first and last histidine residues of the HI(L)GH sequence, which is involved in aminoacyl-adenylate formation (33,38,39), are respectively replaced by threonine and asparagine in synthetases from prokaryotic sources (40-42). In addition, the N-terminal region leading to this sequence in the eukaryotic enzyme is about 160 amino acids longer than that in the prokaryotic counterpart (37,40-42). Such differences around the active pocket between prokaryotes and eukaryotes seem likely to be closely related to the difference in recognition of the 3' CCA vicinity. It is worth mentioning that the positional specificity...
for prokaryotic tRNA\textsuperscript{Tm} in initial attachment with tryptophan to 2'- or 3'-hydroxyl group of the ribose ring of the terminal adenosine residue is the reverse of that for eukaryotic tRNA\textsuperscript{Tm} (43). Further studies are required to better understand this aspect.

ACKNOWLEDGEMENTS

Gratitude is extended to the Radio Isotope Center at the University of Tokyo for the use of their facility. This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science and Culture, Japan.

REFERENCES