Pure & Appl. Chem., Vol. 61, No. 3, pp. 573–576, 1989.
Printed in Great Britain.

A novel lysine-substituted nucleoside in the first position of the anticodon of minor isoleucine tRNA from *Escherichia coli*

Shigeyuki Yokoyama, Tomonari Muramatsu, Nobuyuki Horie, Kazuya Nishikawa*, Akira Matsuda**, Tohru Ueda**, Ziro Yamaizumi***, Yoshiyuki Kuchino***, Susumu Nishimura*** and Tatsuo Miyazawa+

Department of Biophysics and Biochemistry, Faculty of Science, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, *Department of Molecular Biology, Faculty of Science, Nagoya University, Chikusa-ku, Nagoya 464, **Faculty of Pharmaceutical Sciences, Hokkaido University, Kitaku, Sapporo 060, ***Biology Division, National Cancer Center Research Institute, Tsukiji, Chuo-ku, Tokyo 104, +Faculty of Engineering, Yokohama National University, Tokiwadai, Hodogaya-ku, Yokohama 240, Japan

Abstract - A minor species of isoleucine tRNA ($tRNA^{Ile}_{minor}$) specific to the codon AUA was purified from *Escherichia coli*, and a novel modified nucleoside N^+ in the first position of the anticodon was prepared. From the NMR analysis, the mass spectrometry and chemical synthesis, the structure of nucleoside N^+ was determined as 4-amino-2-(N^6-L-lysino)-1-(β-D-ribofuranosyl)pyrimidinium (lysidine). The gene for $tRNA^{Ile}_{minor}$ (*ileX*) was isolated and lysidine was found to be coded by cytidine. Lysidine (L) is a novel type of modified nucleoside, lysine-substituted cytidine. Because of this unique structure, lysidine in the first position of anticodon recognizes adenosine but not guanosine in the third position of codon. Lysidine in $tRNA^{Ile}_{minor}$ was replaced with unmodified cytidine, which resulted in a remarkable reduction of the isoleucine-accepting activity and an unexpected appearance of the methionine-accepting activity. The modification from cytidine to lysidine in the anticodon concurrently converts the amino acid specificity and codon specificity of $tRNA^{Ile}_{minor}$. This finding is important for the discussion on the evolution of assignment of the codon AUA to isoleucine or methionine in mitochondria.

INTRODUCTION

In protein biosynthesis, certain tRNA species recognize more than one codons and the number of tRNA species required for translating genetic codes on mRNA is appreciably smaller than the number of amino acid codons. In contrast to mRNA, tRNA species are modified at several specific sites after being transcribed from DNA. In particular, uridine in the first position of anticodon is almost always modified (ref. 1). We have found, from proton NMR analyses of modified pyrimidine nucleosides and nucleotides, that those two types of modifications remarkably affect the conformational characteristics of the anticodon moiety, and contribute to the correct translation of the codons of Gln, Lys and Glu and to the efficient translation of the codons of Val, Ser, Pro, Thr and Ala (ref. 2).

In the hope of the complete elucidation of the molecular mechanism in the regulation of the codon recognition, we have taken up the problem of the recognition of the codons for Ile and Met. Only there in the genetic code table, a codon box is divided into three codons for one amino acid (Ile) and one codon for the other amino acid (Met). In *Escherichia coli*, there are two tRNA species known to be specific to isoleucine. The codons AUU and AUC are recognized by $tRNA^{Ile}_{major}$ that has guanosine in the first position of the anticodon (ref. 3). On the other hand, the codon AUA is recognized by $tRNA^{Ile}_{minor}$ (ref. 4), the primary structure of which has been tentatively determined (ref. 5), although the modified nucleoside N^+ in the first position of anticodon has not been identified.

DETERMINATION OF CHEMICAL STRUCTURE OF NOVEL MODIFIED NUCLEOSIDE

Purification of E. coli $tRNA^{Ile}_{minor}$. In the present study, a large amount of $tRNA^{Ile}_{minor}$ sufficient for the determination of the chemical structure of N^+ was purified from unfractionated tRNA (160,000 A_{260} unit) from *E. coli* A19. $tRNA^{Ile}_{minor}$ was purified by successive chromatography on columns of DEAE-Sephadex A-50 at pH 7.5, DEAE-Sephadex A-50 at pH 4.0, benzoylated DEAE-cellulose, Sepharose 4B, and benzoylated DEAE-cellulose. Thus, a total of 90 A_{260} units of the purified preparation of $tRNA^{Ile}_{minor}$ was obtained (ref. 6).

Nucleotide sequence of E. coli $tRNA^{Ile}_{minor}$. The sequencing of $tRNA^{Ile}_{minor}$ was carried out by the method described by Kuchino *et al* (ref. 5,7) and was confirmed by the Donis-Keller method (ref. 8). The nucleotide sequence of this $tRNA^{Ile}_{minor}$ (Fig. 1) was identical with that reported previously (ref. 5) except that the nucleoside in position 27 was now found to be G rather than A (ref. 6).

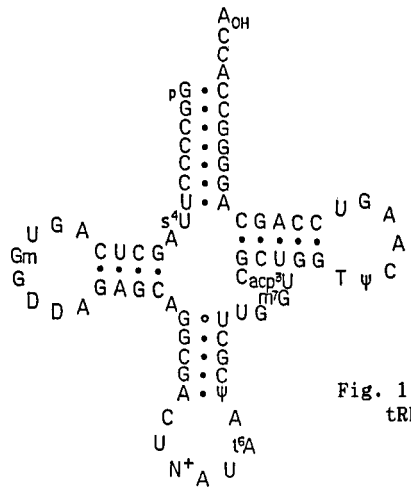

Fig. 1. Nucleotide sequence of $tRNA^{Ile}_{minor}$ from *E. coli* A19

Preparation of Nucleoside N^+ *from* $tRNA^{Ile}_{minor}$. Purified preparation of $tRNA^{Ile}_{minor}$ (90 A_{260} units) was digested to nucleosides with ribonuclease A, cobra venom phosphodiesterase, and *E. coli* phosphomonoesterase. This digest was subjected to HPLC with an ODS column, and finally 0.60 A_{260} units of purified nucleoside N^+ was obtained (ref. 6).

^{1}H-NMR spectroscopy of nucleoside N^+. The 400-MHz ^{1}H-NMR spectrum of nucleoside N^+ in ^2H_2O solution was analyzed. All the ribose proton resonances (H1', H2', H3', H4', H5' and H5'') and the H5 and H6 proton resonances of pyrimidine ring were observed. In addition, the proton resonances of spin-coupled series of one methine and four methylene groups were observed, suggesting the presence of lysine moiety. The pH dependence of chemical shifts indicated that the ε-amino group of lysine is substituted in the pyrimidine ring (ref. 6).

Mass spectrum of trimethylsilylated nucleoside N^+. The mass spectrum of trimethylsilylated nucleoside N^+ was measured. As for the ion of *m/z* 731, the exact mass was found to be 731.3835, and the composition was determined as $C_{30}H_{65}O_6N_5Si_5$ with a calculated mass of 731.3781. By subtracting the contribution of five trimethylsilyl groups, the molecular formula of nucleoside N^+ was determined as $C_{15}H_{25}O_6N_5$ (ref. 6).

Chemical structure of nucleoside N^+. From the NMR spectroscopy and mass spectrometry, two possible structures of the neutral form of nucleoside N^+ were derived. One is a cytidine derivative, in which the oxygen atom in position 2 is replaced by the ε nitrogen atom of lysine, 4-imino-2-(N^6-L-lysino)-1-(β-D-ribofuranosyl)pyrimidine (ref. 6). Such a nucleoside is probably protonated at neutral pH to form 4-amino-2-(N^6-L-lysino)-1-(β-D-ribofuranosyl)pyrimidinium (designated as k^2C) as in the case of 2,4-diamino-1-methylpyrimidine. The alternative one is 2-imino-4-(N^6-L-lysino)-1-(β-D-ribofuranosyl)pyrimidine.

Chemical synthesis of k^2C. In the hope of determining the structure of nucleoside N^+, an unequivocal synthesis of k^2C was performed which utilized the substitution of 4-amino-2-methylthio-1-(β-D-ribofuranosyl)pyrimidinium chloride (1) with N^2-Cbz-L-lysine. Compound 1 reacted with N^2-Cbz-L-lysine in the presence of equimolar amount of sodium ethoxide in ethanol to afford 4-amino-2-(N^2-Cbz-N^6-L-lysino)-1-(β-D-ribofuranosyl)pyrimidinium (2) in 73% yield. This compound was purified by the reverse phase HPLC. De-benzyloxycarbonylation of 2 was then performed under hydrogen atmosphere with palladium hydroxide as catalyst to give 4-amino-2-(N^6-L-lysino)-1-(β-D-ribofuranosyl)pyrimidinium (k^2C) (ref. 6).

Comparison of nucleoside N^+ *and authentic* k^2C. The chemically synthesized k^2C was found to be identical with that of nucleoside N^+, from the comparison of UV absorption spectrum, the retention time in HPLC on an ODS column, ^{1}H-NMR and mass spectra. Probably, nucleoside N^+ has an L-lysine moiety rather than a D-lysine moiety, just as the chemically synthesized k^2C having an authentic L-lysine moiety (ref. 6).

Ionization state of nucleoside N^+. In paper electrophoresis at pH 8, nucleoside N^+ was found to migrate as fast as 1,6-dimethyladenosine with one positive charge. Further, the analysis of the pH dependence of proton chemical shifts of nucleoside N^+ indicates that in the pH range 3.5-8.5, nucleoside N^+ has one net positive charge in the 2,4-diaminopyrimidine moiety (Fig. 2) (ref. 6).

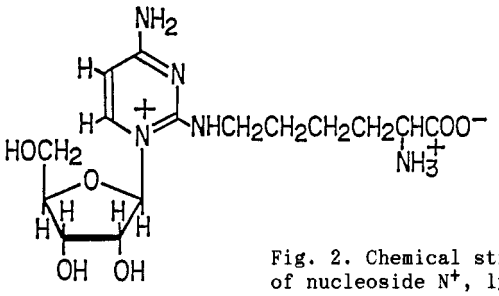

Fig. 2. Chemical structure of nucleoside N^+, lysidine

Novel nucleoside LYSIDINE. The chemical structure of the unusual modified nucleoside N^+ in the first position of the anticodon of *E. coli* $tRNA^{Ile}_{minor}$ is now determined as shown in Fig. 2. In order to determine the nucleoside precursor of N^+, the gene of $tRNA^{Ile}_{minor}$ (*ileX*) was cloned from *E. coli* and N^+ was found to be coded by cytidine (ref. 9). Nucleoside N^+ is a novel type of nucleoside; in particular, nucleoside N^+ is the first example of lysine substituted nucleoside. We propose to call nucleoside N^+ as "lysidine" (a hybrid of lysine and cytidine) with a one-letter code "L". Probably this nucleoside L is derived from cytidine by substituting the oxygen atom in position 2 with ε-nitrogen atom of L-lysine.

Lysidine specifically recognizes adenosine but not guanosine. $tRNA^{Ile}_{minor}$ recognizes the codon AUA only (4); lysidine (L) in the first position of the anticodon specifically recognizes adenosine rather than guanosine. There are two probable structures for the base pair of L and A. In one model, nucleoside L is in a tautomeric form with an NHR group in position 2 and an NH_2 group in position 4. On the other hand, in the other model, nucleoside L is in a tautomeric form with an >NH group in position 3 and an =NH group in position 4. However, the presence of the bulky group R in position 2 does not allow the formation of base bair with guanosine. Thus, the modification of C to L in the first position of the anticodon will avoid the mistranslation of the methionine codon AUG to isoleucine (ref. 6).

EFFECT OF MODIFICATION OF AMINOACYLATION OF $tRNA^{Ile}_{minor}$ SPECIES

Effect of modification of cytidine to lysidine on aminoacylation of tRNA. The anticodon of *E. coli* $tRNA^{Ile}_{minor}$ has been found to be coded by CAT, that is characteristic to $tRNA^{Met}$ (ref. 9). Thus, only after the post-transcriptional modification of C(34), the anticodon (LAU) of *E. coli* $tRNA^{Ile}_{minor}$ is matured from the anticodon CAU, that should recognize the methionine codon AUG. Then, what should happen to the aminoacylation of the tRNA species, if the modification of the anticodon from C to L did not occur? If an immature tRNA species with the anticodon CAU accepts isoleucine, this should result in the mistranslation of the codon AUG to isoleucine rather than methionine. This problem may be solved by the aminoacylation experiments on a putative precursor of *E. coli* $tRNA^{Ile}_{minor}$, where L(34) is replaced by cytidine, the precursor of lysidine.

Substitution of the anticodon of $tRNA^{Ile}_{minor}$ with CAU. The putative precursor of $tRNA^{Ile}_{minor}$ was prepared by the use of a variety of enzymes. *E. coli* $tRNA^{Ile}_{minor}$ was cleaved by ribonuclease A to yield the 5'-half molecule, which was then elongated by the ligation with pUCAp after dephosphorylation of the 5'-teminus. $tRNA^{Ile}_{minor}$ was cleaved also with ribonuclease U_2 to yield the 3'-half molecule. The elongated 5'-half molecule and the 3'-half molecule were annealed and then ligated fo form $tRNA^{Ile}_{minor(CAU)}$, a putative precursor of $tRNA^{Ile}_{minor}$ with the anticodon CAU (ref. 9).

Isoleucine accepting activity of $tRNA^{Ile}_{minor(CAU)}$. $tRNA^{Ile}_{minor}$ species is charged with isoleucine as efficiently as $tRNA^{Ile}_{major}$. However, the isoleucine-accepting activity of $tRNA^{Ile}_{minor}$ was remarkably reduced by the substitution of L(34) with cytidine. This was surprising; isoleucyl-tRNA synthetase has been considered not to recognize the first letter of anticodon, since this enzyme charges isoleucine to both of $tRNA^{Ile}_{major}$ and $tRNA^{Ile}_{minor}$ which are different from each other in the first letter of the anticodon. However, in the present study, isoleucyl-tRNA synthetase was found no longer to charge $tRNA^{Ile}_{minor(CAU)}$ (ref. 9). These indicate that this enzyme discriminates against tRNA species with the methionine anticodon CAU, avoiding mistranslation of the methionine codon AUG; cytidine in the first position of anticodon constitutes a "negative" determinant for the aminoacylation by isoleucyl-tRNA synthetase.

Methionine accepting activity of $tRNA^{Ile}_{minor}$. Naturally, $tRNA^{Ile}_{major}$ or $tRNA^{Ile}_{minor}$ may not be charged with methionine. Surprisingly, however, upon the substitution of lysidine with cytidine, $tRNA^{Ile}_{minor(CAU)}$ was efficiently charged with methionine (ref. 9). This suggests the possibility that a precursor $tRNA^{Ile}_{minor(CAU)}$ with the methionine anticodon serves as a $tRNA^{Met}$. The nucleoside in position 34 of tRNAs has been found to be critical for the recognition by this enzyme (refs. 10,11). Probably, for the aminoacylation by methionyl-tRNA

synthetase, cytidine in the first position of the anticodon constitutes a "positive" determinant.

Concurrent conversion of amino acid specificity and codon specificity. In summary, in the present series of studies (ref. 6,9), the chemical structure of a novel modified nucleoside in the first position of the anticodon of $tRNA^{Ile}_{minor}$ from *E. coli* was determined. This novel nucleoside LYSIDINE is coded by cytidine in the structure gene of this tRNA species and has a lysine moiety substituted in position 2 of the cytosine ring. Even more surprising, the modification from cytidine to lysidine concurrently converts the amino acid specificity and codon specificity of $tRNA^{Ile}_{minor}$ (Fig. 3).

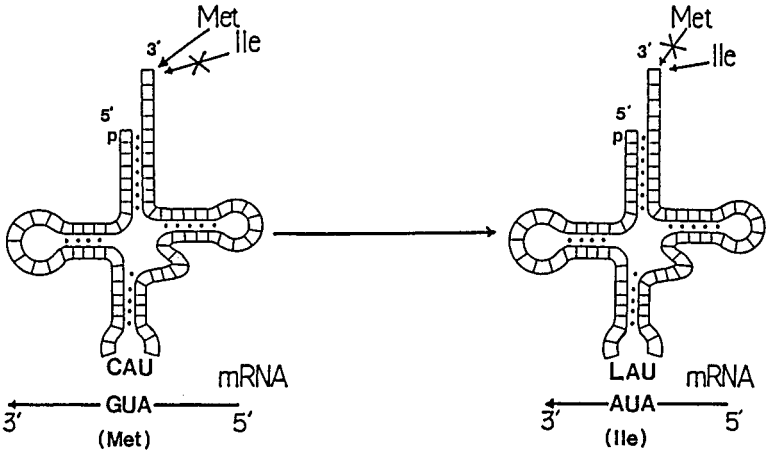

Fig. 3. Concurrent conversion of amino acid specificity and codon specificity of $tRNA^{Ile}_{minor}$

Involvement of tRNA modification enzyme in alteration of genetic codes. In mitochondria of some organisms, including mammals (ref. 12) and *Saccharomyces cerevisiae* (ref. 13), the codon AUA is used for methionine rather than for isoleucine. We suggest that, in mitochondria of those organisms, the gene of the putative enzyme for the modification from cytidine to lysidine was lost. This resulted in the occurrence of tRNA species with the unmodified anticodon CAU, which eventually evolved into $tRNA^{Met}$. Probably, the disappearance of one (set of) enzyme for the post-transcriptional modification from cytidine to lysidine is directly involved in the alternative use of the codon AUA for methionine or isoleucine.

Acknowledgement

This work was supported in part by Grant-in-Aid for Specially Promoted Research (60060004) from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. S. Nishimura, in *TRANSFER RNA: Structure, Properties, and Recognition* (P.R. Schimmel, D. Soll and J.N. Abelson, eds.), pp. 59-79, Cold Spring Harbor Laboratory, New York, 1979.
2. S. Yokoyama, T. Watanabe, K. Murao, H. Ishikura, Z. Yamaizumi, S. Nishimura and T. Miyazawa, *Proc. Natl. Acad. Sci., USA 82*, 4905-4908 (1985).
3. M. Yarus and B.G. Barrell, *Biochem. Biophys. Res. Commum. 43*, 729-734 (1971).
4. F. Harada and S. Nishimura, *Biochemistry 13*, 300-307 (1974).
5. Y. Kuchino, S. Watanabe, F. Harada and S. Nishimura, *Biochemistry 19*, 2085-2089 (1980).
6. T. Muramatsu, S. Yokoyama, N. Horie, A. Matsuda, T. Ueda, Z. Yamaizumi, Y. Kuchino, S. Nishimura and T. Miyazawa, *J. Biol. Chem.* (in press).
7. Y. Kuchino, M. Kato, H. Sugisaki and S. Nishimura, *Nucleic Acids Res. 6*, 3459-3469 (1979).
8. H. Donis-Keller, A.M. Maxam and W. Gilbert, *Nucleic Acids Res. 4*, 2527-2538 (1977).
9. T. Muramatsu, T. Miyazawa, K. Nishikawa, F. Nemoto, Y. Kuchino, S. Nishimura and S. Yokoyama (to be published).
10. L. Stern and L.H. Schulman, *J. Biol. Chem. 252*, 6403-6408 (1977).
11. L.H. Schulman, H. Pelka and M. Susani, *Nucleic Acids Res. 11*, 1439-1455 (1983).
12. S. Anderson, A.T. Bankier, B.G. Barrell, M.H.L. de Bruijn, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, F. Sanger, P.H. Schreier, A.J. Smith, R. Staden and I.G. Young, *Nature 290*, 457-465 (1981).
13. M.E.S. Hudspeth, W.M. Ainley, D.S. Shumard, R.A. Butow and L.I. Grossman, *Cell 30*, 617-626 (1982).