Unconventional codon reading by *Mycoplasma mycoides* tRNAs as revealed by partial sequence analysis

Youssef S. GUINDY, Tore SAMUELSSON and Tor-Inge JOHANSEN

Department of Medical Biochemistry, University of Göteborg, Box 33031, S-400 33 Göteborg, Sweden

Continuing our investigation of the tRNA genes and gene products in *Mycoplasma mycoides*, we report the sequence of the gene for tRNA^{Leu} (CAA) as well as partial primary structures of the following tRNAs: Leu (CAA), Leu (UAG), Arg (UCU), Thr (AGU) and Ile (CAU). It is suggested that in *M. mycoides*, at least some of the family codon boxes are read by only one tRNA each, using an unconventional method which does not discriminate between the nucleotides in the third codon position. *M. mycoides* is the first free-living organism known to use an unconventional method of this kind.

INTRODUCTION

The mitochondrial genome contains a very limited number of tRNA genes, not enough to sustain a codon-reading scheme according to the wobble rules (Crick, 1966). Instead, the codon families (i.e. groups of four codons, having their first two nucleotides in common, that all code for the same amino acid) are read by only one tRNA each (Barrell *et al.*, 1980; Bunitz *et al.*, 1980; Heckman *et al.*, 1980). Consequently, in these cases, the reading tRNA cannot discriminate between the nucleotides that occupy the third codon position. A similar situation exists in chloroplasts, where some codon families, but not all, are read by only one tRNA each using an unconventional method without discrimination in the third codon position (Ohshima *et al.*, 1986; Shinozaki *et al.*, 1986). It has been suggested that the mitochondrion is of prokaryotic origin, and one may ask if there are prokaryotes today that rely on unconventional codon reading of this kind in order to read codon families using a minimum number of tRNAs.

Kilpatrick & Walker (1980) have reported that *Mycoplasma mycoides* contains only one glycine tRNA and that this tRNA has an unsubstituted U in the wobble position. We have tested this tRNA in a protein-synthesizing system *in vitro* and have found that it was almost as efficient in the unorthodox reading of the codons GGU and GGC as it was in conventional reading (Samuelsson *et al.*, 1983). Since this is precisely the result to be expected for a tRNA designed to read all four codons in a family, we have embarked on an inventory of the tRNA genes and their gene products in this organism (Samuelsson *et al.*, 1985, 1987, 1988). So far we have cloned and analysed 15 tRNA genes, seven of which code for tRNAs that read family codons (arginine, alanine, glycine, proline, serine, threonine and valine). In each of these cases our clones contained only one gene per codon family.

These genes all have a T in the position corresponding to the wobble nucleotide, with the exception of the arginine tRNA gene, which has an A in this position. This resembles the situation in mitochondria, and the resemblance is further strengthened by the results of our analysis of the tRNAs in *M. mycoides*. For each of the following codon families we have found only one tRNA:

alanine (GCN), leucine (CUN), proline (CCN), serine (UCN) and valine (GUN). All of these tRNAs have an unsubstituted U in the wobble position. To this list should be added the glycine tRNA reported by Kilpatrick & Walker (1980). In the present paper we report the sequence of the gene for tRNA^{Leu} (CAA) and the primary structures of the following tRNAs: tRNA^{Leu} (CAA), tRNA^{Leu} (UAG), tRNA^{Arg} (UCU), tRNA^{Thr} (AGU) and tRNA^{Ile} (CAU). It is particularly noteworthy that there are two threonine tRNAs in *M. mycoides*, one with U in the wobble position (Samuelsson *et al.*, 1987) and another with A in this position. A similar finding has been reported for *M. capricolum* by Andachi *et al.* (1987). The resemblance to mitochondrial codon reading is therefore not complete, a situation that is reminiscent of that in the chloroplast (Ohshima *et al.*, 1986; Shinozaki *et al.*, 1986).

MATERIALS AND METHODS

Purification and sequence analysis of tRNA

M. mycoides sp. *capri* was grown and harvested and crude tRNA was prepared from the cells as previously described (Samuelsson *et al.*, 1985). Individual tRNAs were purified by chromatography on benzoylated DEAE-cellulose (Gillam *et al.*, 1967; Mitra *et al.*, 1977). Further purification was achieved by rechromatography on benzoylated DEAE-cellulose after phenoxylacetylation of the tRNAs (Gillam *et al.*, 1968). The individual aminoacylated tRNAs were further purified by two-dimensional polyacrylamide-gel electrophoresis as described by Garel *et al.* (1977).

The purified tRNAs were sequenced using the direct read-out method (Gupta & Randerath, 1979) with some modifications. The procedure we used was as follows. The isolated tRNA (1–10 µg) was lyophilized and dissolved in 10 µl of water. The tRNA sample was then partially hydrolysed at 100 °C for 45–60 s and then re-lyophilized. The sample was dissolved in 30 µl of 40 mM-Tris/HCl, pH 8.7/10 mM-MgCl₂/10 mM-dithiothreitol, and added to a 1.5 ml Eppendorf tube containing 0.1 mCi of [γ-³²P]ATP and unlabelled ATP at a final concentration of 30 mM. Polynucleotide kinase (5–7 units) was added to the sample, which was incubated at

37 °C for 45 min. The reaction mixture was lyophilized and dissolved in 15 µl of a loading buffer (90 mM-Tris/90 mM-boric acid / 1 mM-Na₂EDTA / 0.0075 % Xylene Cyanol). The radioactively labelled tRNA fragments were divided into three equal aliquots and loaded on a 12 % polyacrylamide gel (4 mm × 220 mm × 550 mm) at different times. The maximum resolution of the fragments was observed when samples of equal volumes were electrophoresed for 2.5 h, 4.5 h and 7.5 h respectively, at 2500 V and 20 mA. The gel was maintained at constant temperature by continuous circulation of water at 30 °C in the thermostated plate (LKB Macrophor). The resulting three ladders were visualized by autoradiography for 1–2 h. The sequence ladders were transferred by direct contact to three poly(ethyleneimine)–cellulose t.l.c. plates (Machery–Nagel and Co., 20 cm × 40 cm t.l.c. Cel 300) that had been washed previously in water for 30 min. The t.l.c. plates were positioned in such a way as to allow the whole sequence ladder to be transferred 1.5–2.0 cm from the edge of the plate. After the overnight transfer, the plates were immediately washed with ethanol to remove the urea and then washed in 0.15 M-sodium acetate, pH 4.5, for about 7 min. The plates were then dried in a 37 °C oven, and with the help of a 5–6 h autoradiogram the exact position of each sequence ladder was located. RNAase T₂ was then applied onto these ladders (0.5 unit of enzyme to each 1 cm of ladder) and after covering with Parafilm they were incubated at 37 °C overnight. The plates were subsequently washed in ethanol for 10 min and dried in a 37 °C oven. A 5 cm Whatman 1 wick was attached at the top of each PEI–cellulose sheet. The chromatogram was developed first in water to the origin line and then in 0.55 M-ammonium sulphate or 4 M-lithium-formate to 3–4 cm on the wick. After chromatography, the PEI–cellulose sheets were dried and autoradiographed over night. Modified nucleotides were identified using the *R_F* values given in Randerath & Randerath (1983). In some cases wobble nucleotides were further authenticated by two-dimensional t.l.c. (Silberklang *et al.*, 1977).

Cloning procedures and DNA sequence analysis

A Charon 28 library of *M. mycoides* HindIII fragments was screened by hybridization using purified ³²P-labelled tRNA^{Leu} from *M. mycoides* as a probe. The insert (3.0 kb) from one of the hybridizing clones was inserted in the HindIII site of the plasmid pTZ19R. Insert DNA was isolated from the resulting plasmid and digested with *Sau*3AI. One of the resulting fragments that hybridized to the probe was cloned in M13 vectors and the DNA sequence was determined by the chain-termination method described by Sanger *et al.* (1977). The sequence shown in Fig. 1 was obtained from analysis of contiguous sequence of both strands of the DNA.

RESULTS

Cloning and nucleotide sequence of the *M. mycoides* tRNA^{Leu} gene

We have previously reported the cloning and nucleotide sequence analysis of 15 tRNA genes in *M. mycoides* (Samuelsson *et al.*, 1985, 1987, 1988). In the present investigation we screened a Charon 28 library of HindIII fragments by plaque hybridization using purified ³²P-labelled tRNA^{Leu} from *M. mycoides* as a probe. A

```

50
GATCACACGC ATTAAATGCA ACTAAAAGAA AATGAACTT AAATCTACAA
100
AAGGTTAGAG TAATGGATGA AAATGGTAGT GTATTTAATA TCAAAGTATC
150
TGCTAGAACA TTAAGAACTC TAAAAAACA AGAAAAATC GTTTAATAAG
200
TAAATAATA ATAGTATGAC ATTTGTCATA TTTTTTTTAT ATAATAATTT
250
ATTATTAGCC CTTTGGCGG AATTGGCAGA CGCATTAGAC TCAAAATCTA
-----LEU-----
300
ACGAAGAAAT TCGTATCGGT TCGACCCCGA TAAAGGCAC CAATCGATTA
-----
350
ATAAACTTGG ATTATTCAA GTTTTTTTAT TTTATATTAA TTAATATAAA
400
ATAATAATAT GAATTGGTGT TATTATGAAC TGAAGCATAA AAAAAGTTAG

```

Fig. 1. Nucleotide sequence of a region containing the tRNA^{Leu} gene

fragment of 3.0 kb from one of the hybridizing clones was further analysed. The nucleotide sequence of a part of this fragment (Fig. 1) reveals a gene encoding a tRNA^{Leu} with the anticodon CAA. As mentioned in the section below, we have also determined the nucleotide sequence of the tRNA that presumably corresponds to this gene.

Primary structure of five tRNAs from *M. mycoides*

In the previous inventory of tRNAs of *M. mycoides* we determined the nucleotide sequence of five tRNAs: tRNA^{Ala} (UGC), tRNA^{Pro} (UGG), tRNA^{Ser} (UGA), tRNA^{Thr} (UGU) and tRNA^{Val} (UAC) (Samuelsson *et al.*, 1987). Continuing this study we report five more tRNA nucleotide sequences: tRNA^{Leu} (CAA), tRNA^{Leu} (UAG), tRNA^{Arg} (UCU), tRNA^{Thr} (AGU) and tRNA^{Ile} (CAU). The clover-leaf structure for tRNA^{Leu} (CAA), tRNA^{Leu} (UAG), tRNA^{Ile} (CAU) and tRNA^{Arg} (UCU) is shown in Fig. 2. The direct read-out method (Gupta & Randerath, 1979) that we used to determine the tRNA sequences gives no information on the 5'- and 3'-terminal nucleotides of the tRNA. We did not take measures to determine these specific features of the tRNA as they are not relevant to the conclusions drawn in this paper. The termini shown in Fig. 2 have instead been deduced from the corresponding gene sequences.

Chromatography on a benzoylated DEAE-cellulose column (Fig. 3) gave a single peak of tRNA^{Thr} which was further purified by rechromatography after phenoxyacetylation. The two-dimensional polyacrylamide-gel electrophoresis of an aliquot of this fraction showed two predominating spots. Both spots were eluted and their sequences determined. The first spot proved to be the isoacceptor tRNA^{Thr} (UGU) which was reported in our previous paper (Samuelsson *et al.*, 1987). The second spot is tRNA^{Thr} (AGU) with the partial primary structure shown in Fig. 4.

The benzoylated DEAE-cellulose chromatogram showed three peaks of leucine tRNA isoacceptors (Fig. 3). The first peak was further purified by chromatography after phenoxyacetylation and then subjected to two-dimensional gel electrophoresis followed by sequence

Fig. 2. Nucleotide sequences in clover-leaf form of *M. mycoides* tRNA^{Leu} (CAA) (a), tRNA^{Leu} (UAG) (b), tRNA^{Arg} (UCU) (c), and tRNA^{Ile} (CAU) (d)

The 5'- and 3'-terminal sequences indicated by arrows have been deduced from the gene sequences in Fig. 1, and from Samuelsson *et al.* (1985, 1987). Modified nucleotides could therefore occur in these terminal sequences. N denotes a nucleotide which has not been determined. U* and C* denote unknown modifications of U and C respectively. Abbreviations: m⁶A, N⁶-methyladenosine; m¹G, 1-methylguanosine; m¹A, 1-methyladenosine; m⁷G, 7-methylguanosine; t⁶A, N-[(9-β-D-ribofuranosylpurin-6-yl)carbamoyl]threonine; D, dihydrouridine; Ψ, pseudouridine.

Fig. 3. Chromatography of *M. mycoides* tRNAs on BD-cellulose

Crude tRNA from *M. mycoides* (92 mg) was chromatographed on a column of BD-cellulose (100 cm × 1.5 cm). Elution was carried out using a linear NaCl gradient. The mixing chamber contained 2 litres of 0.3 M-NaCl/0.01 M-MgSO₄, and the reservoir contained 2 litres of 1.0 M-NaCl/0.01 M-MgSO₄. —, Absorbance at 260 nm; ●, tRNA^{Leu}; ○, tRNA^{Thr}.

determination. The result showed that this peak contained the isoacceptor tRNA^{Leu} (CAA), corresponding to the gene sequence reported in Fig. 1, as well as tRNA^{Leu} (UAG). Further fractionation of the Leu II and III peaks (which were poorly resolved on the chromatogram) followed by sequence determination indicated that they contained the same isoacceptor. This tRNA corresponds to the product of the tRNA^{Leu} (TAA) gene in *M. capricolum* (Andachi *et al.*, 1987).
Following the same purification procedures we have also isolated and sequenced the tRNA^{Arg} (U*CU), where U* is a modification of U, as well as tRNA^{Ile} (C*AU), where C* is a modified C (Fig. 2).

DISCUSSION

In Fig. 5 we have summarized the results of our inventory of tRNA genes and their gene products in *M. mycoides* comprising both the present paper as well as previous publications (Samuelsson *et al.*, 1985, 1987, 1988). The most striking feature of the translational apparatus in this microorganism is the fact that with one exception the codon families are read by only one tRNA

8

AGCUCAGUDGGDAGAGCAAUUGACUAGUNAΨCAAUAGm⁷GUCGAAGGUUCAAAUCCUUUAGUCAG

71

Fig. 4. Partial nucleotide sequence of *M. mycoides* tRNA^{Thr} (AGU)

The anticodon is underlined. N denotes an unknown modified nucleotide, D, Ψ and m⁷G are as in Fig. 2.

	U	C	A	G	
U	Phe G Phe Leu Leu C C	Ser T U	Tyr Tyr Stop Stop	Cys Cys Stop Trp	U C A G
C	Leu U	Pro T U	His His Gln Gln	Arg A	U C A G
A	Ile Ile C* C Ile Met C	Thr T U A	Asn G Asn Lys Lys	Ser Ser Arg T* U Arg	U C A G
G	Val T U	Ala T U	Asp G Asp Glu T Glu	Gly T U	U C A G

Fig. 5. Summary of inventory of tRNA genes and tRNAs in *M. mycoides*

Family codon boxes are indicated by bold lines. The bold letters in the boxes denote the wobble nucleotide in the tRNA gene sequence and the fine letters denote the wobble nucleotide in the tRNA sequence. *U and *C are as in Fig. 2.

each. These tRNAs have an unsubstituted U in the wobble position except for the tRNA^{Arg} which has A (or possibly I) in this position. The threonine tRNAs, however, are different. Our results demonstrate the presence of two isoacceptors, one with A and the other with U in the wobble position. In this context, it is worth noting that the tRNA^{Thr} (AGU) is the first tRNA of non-mitochondrial origin that has been shown to have adenosine instead of inosine in the wobble position. A similar tRNA^{Thr} was reported in *M. capricolum* by Andachi *et al.* (1987).

The presence of only one tRNA per codon family obviously precludes reading of these codons according to the wobble rules, and indicates that the tRNA cannot discriminate between the nucleotides that occupy the third position of family codons. The fact that there seems to be a strong bias against codons ending in C and G in the *M. capricolum* (Andachi *et al.*, 1987) and presumably also in the very similar *M. mycoides*, does not alter this general conclusion. Even if these mycoplasmas employed exclusively codons ending in U and A, the tRNAs with U in the wobble position would still have to be able to read both types of family codons instead of discriminating between them as required by the wobble rules.

It would thus seem that there are two main types of codon reading: one that is used in the cytoplasmic protein synthesis of most organisms and which involves discrimination between the nucleotides that occupy the third position of family codons; and another where such discrimination does not occur, which is found in organelles like mitochondria and chloroplasts as well as in small genome mycoplasmas like *M. mycoides*. Both mitochondria and chloroplasts are believed to be derived from microorganisms and it is an intriguing thought that when a microorganism for some reason is compelled to reduce the size of its genome, it reverts to a primitive translational method with a very simplified tRNA apparatus unable to discriminate between the third position nucleotides of family codons (Lagerkvist, 1986).

We thank Professor Ulf Lagerkvist for helpful suggestions and comments on the manuscript. We also thank Monica Olsson and Anne-Marie von Essen for expert technical assistance and Majbritt Hellquist for typing the manuscript. This investigation has been made possible by a grant from the Swedish Medical Council.

REFERENCES

- Andachi, Y., Yamao, F., Iwami, M., Muto, A. & Osawa, S. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7398–7402
- Barrell, B. G., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3164–3166
- Bonitz, S. G., Berlani, R., Coruzzi, G., Li, M., Macino, G., Nobrega, F. G., Nobrega, M. P., Thalenfeld, B. E. & Tzagoloff, A. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3167–3170
- Crick, F. H. C. (1966) *J. Mol. Biol.* **19**, 548–555
- Garel, J. P., Garner, R. L. & Siddiqui, M. A. O. (1977) *Biochemistry* **16**, 3618–3624
- Gillam, I., Millward, S., Blew, D., von Tigerström, M., Wimmer, E. & Tener, G. M. (1967) *Biochemistry* **6**, 3043–3056
- Gillam, I., Blew, D., Warrington, R. C., von Tigerström, M. & Tener, G. M. (1968) *Biochemistry* **7**, 3459–3468
- Gupta, R. C. & Randerath, R. (1979) *Nucleic Acids Res.* **6**, 3443–3458
- Heckman, J. E., Sarnoff, J., Alzner-Deweerd, B., Yin, S. & RajBhandary, U. L. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3159–3163
- Kilpatrick, M. W. & Walker, R. T. (1980) *Nucl. Acids Res.* **8**, 2783–2786
- Lagerkvist, U. (1986) *BioEssays* **4**, 223–226
- Mitra, S. K., Lustig, F., Åkesson, B., Lagerkvist, U. & Strid, L. (1977) *J. Biol. Chem.* **252**, 471–478
- Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-I., Inokuchi, H. & Ozeki, H. (1986) *Nature (London)* **322**, 572–574
- Randerath, E. & Randerath, K. (1983) in *Methods of DNA and RNA Sequencing* (Weissman, S. M., ed.), pp. 169–233, Praeger, New York
- Samuelsson, T., Axberg, T., Borén, T. & Lagerkvist, U. (1983) *J. Biol. Chem.* **258**, 13178–13184
- Samuelsson, T., Elias, P., Lustig, F. & Guindy, Y. S. (1985) *Biochem. J.* **232**, 223–228
- Samuelsson, T., Guindy, Y. S., Lustig, F., Borén, T. & Lagerkvist, U. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3166–3170
- Samuelsson, T., Borén, T., Johansen, T.-I. & Lustig, F. (1988) *J. Biol. Chem.* **263**, 13692–13699
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. & Sugiura, M. (1986) *EMBO J.* **5**, 2043–2049
- Silberklang, M., Prochiantz, A., Haenni, A.-L. & RajBhandary, U. (1977) *Eur. J. Biochem.* **72**, 465–478

Received 15 August 1988/24 October 1988; accepted 1 November 1988