

Cloning and nucleotide sequence analysis of transfer RNA genes from *Mycoplasma mycoides*

Tore SAMUELSSON, Per ELIAS, Florentyna LUSTIG and Youssef S. GUINDY

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

As part of an investigation of the tRNA genes of *Mycoplasma mycoides*, two *Hind*III fragments of mycoplasma DNA comprising 0.4 and 2.5 kilobases (kb), respectively, were cloned in pBR322 and their nucleotide sequences determined. Only one tRNA gene was found in the 0.4 kb fragment, the gene for tRNA^{Arg} with the anticodon TCT, while the 2.5 kb fragment contained nine different tRNA genes arranged in a cluster which presumably constitutes a transcriptional unit. The clustered tRNA genes, with their respective anticodons, were as follows: Arg (ACG), Pro (TGG), Ala (TGC), Met (CAT), Ile (CAT), Ser (TGA), fMet (CAT), Asp (GTC), and Phe (GAA).

INTRODUCTION

While the recognition of the first two codon positions by the anticodon appears to be fairly straightforward (with the exception of initiation) and to involve only Watson–Crick base pairs, the reading of the third codon nucleotide is more complicated and poses a number of unresolved problems. The wobble rules (Crick, 1966) prohibit readings where mispairs are formed between this position and the reading anticodon since such readings could result in translational errors. On the other hand, half of the codons belong to codon families, i.e. groups of four codons specifying the same amino acid, which have their first two nucleotides in common. In such families it makes no difference, as far as translational fidelity is concerned, how the third nucleotide is read since the first two are enough to specify the amino acid. In view of this we have sought information on unconventional methods in the reading of family codons *in vitro* and have observed a type of unconventional reading whereby a codon may be read by an anticodon that cannot, according to the wobble rules, form a stable base pair with the third codon nucleotide (Mittra *et al.*, 1977, 1979; Samuelsson *et al.*, 1980, 1983; Lustig *et al.*, 1981). An operationally similar mode of reading has been observed in mitochondria where the codon families are each read by only one tRNA which has U in the wobble position (Bonitz *et al.*, 1980; Heckman *et al.*, 1980; Andersson *et al.*, 1981). Furthermore, *Mycoplasma mycoides* has recently been reported to contain only one glycine tRNA and here also the wobble nucleotide is U (Kilpatrick & Walker, 1980). We have tested the mycoplasma tRNA^{Gly} in our *in vitro* protein synthesizing system and found it to have the reading properties to be expected for a tRNA designed to read all four glycine codons (Samuelsson *et al.*, 1983).

The codon usage in *M. mycoides* has not yet been determined but, with this caveat, it would seem that this organism might be an interesting model for unconventional codon reading. The genome of *M. mycoides*, approximately 800 kb, is relatively small compared with that of most bacteria (Razin, 1980) and one could

speculate that the limited coding space might have led to a paucity of tRNA genes. This idea is consistent with the results of preliminary studies of mycoplasma tRNAs which suggest that there are relatively few tRNA species in this organism (Walker, 1976). In order to elucidate this question we have undertaken a study of the tRNA genes in *M. mycoides*. Screening of a library of *Hind*III fragments of mycoplasma DNA by using radioactive mycoplasma tRNA^{Arg} as a probe selected two different recombinant clones and sequence analysis of the cloned fragments revealed a total of ten different tRNA genes comprising a solitary gene on one of the fragments and a cluster of nine on the other.

MATERIALS AND METHODS

Enzymes and chemicals

Restriction enzymes, T4 DNA ligase, DNA polymerase (Klenow fragment) and calf intestine alkaline phosphatase were from Boehringer-Mannheim. Polynucleotide kinase and RNA ligase was purchased from P.L. Biochemicals, and [³²P]ATP and deoxyadenosine 5'- α -[³⁵S]thiotriphosphate was obtained from Amersham International.

Purification of arginine tRNA

Cells of *Mycoplasma mycoides* sp. *capri* were grown and harvested as previously described (Samuelsson *et al.*, 1983). A crude tRNA fraction from *M. mycoides* was prepared as described previously (Samuelsson *et al.*, 1983) and was subjected to chromatography on benzoylated DEAE-cellulose (Gillam *et al.*, 1967; Mittra *et al.*, 1977). The tRNA^{Arg} fraction was further purified by rechromatography on benzoylated DEAE-cellulose after phenoxycetylation of the tRNA^{Arg} (Gillam *et al.*, 1968) to give a final acceptor activity of approx. 2 nmol of arginine/A₂₆₀ unit. When analysed by chromatography on an h.p.l.c. column according to Bischoff *et al.* (1983) it gave a single, apparently homogeneous peak.

For use as a hybridization probe the tRNA was labelled at the 3' end with ³²P essentially as described by

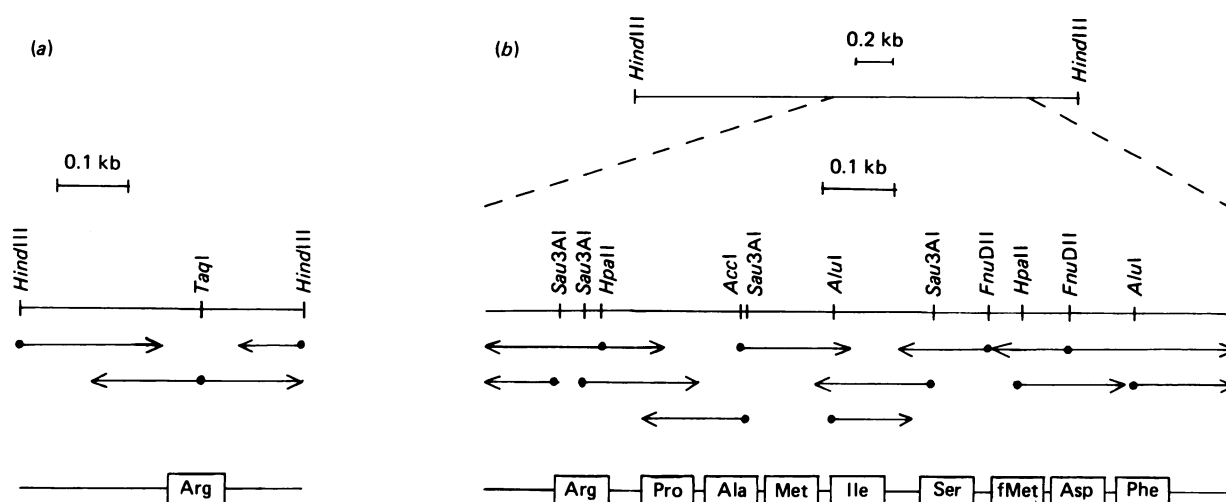


Fig. 1. Physical map of the *Hind*III inserts of recombinant plasmids pMM22 (a) and pMM25 (b)

The points and arrows represent the origin and extent of each M13 clone sequence analysis. Below is a schematic drawing showing the positions of the tRNA genes in the respective sequences.

England *et al.* (1980). The labelled tRNA was analysed by h.p.l.c. in order to check the efficiency of labelling. The specific radioactivity was approx. 2×10^7 c.p.m./ μ g of RNA.

Cloning procedures

DNA was prepared from cells of *M. mycoides* by the Marmur (1961) procedure. It was digested with *Hind*III and ligated with dephosphorylated, *Hind*III-cut pBR322. The ligation mixture was used to transform cells of *Escherichia coli* FLO1 (a *recA* derivative of MM294) which had been made competent by using the method of Hanahan (1983). Transformed cells were plated onto nitrocellulose filters on LB medium agar plates containing 40 μ g of ampicillin/ml as described by Maniatis *et al.* (1982). The resulting colonies were copied to another set of nitrocellulose filters and analysed by hybridization essentially as described by Maniatis *et al.* (1982) using incubation at 68 °C overnight in a solution containing 5 \times SSC, 1 \times Denhardt's solution, 100 μ g of salmon sperm DNA/ml, 0.2% sodium dodecyl sulphate, 1 mM-ATP, and the radioactive probe at a concentration of approx. 20 ng/ml.

DNA sequence analysis

Restriction maps were constructed using the method of Smith & Birnstiel (1976). Using the strategy indicated in Fig. 1, restriction fragments were cloned in M13 vectors and their DNA sequence determined by the chain termination method of Sanger *et al.* (1977), using deoxyadenosine 5'- α -[35 S]thiotriphosphate as radioactive substrate.

RESULTS AND DISCUSSION

Cloning of tRNA genes

To be able to screen for a large number of tRNA genes in *M. mycoides* we have in our laboratory extensively purified several tRNA species from this organism

(T. Samuelsson & Y. Guindy, unpublished work). For the present study we chose tRNA^{Arg}, purified to homogeneity, as hybridization probe.

Total *M. mycoides* DNA was digested to completion with *Hind*III, the resulting fragments separated by electrophoresis, transferred to nitrocellulose and hybridized to purified 32 P-labelled tRNA^{Arg}. In this experiment the probe hybridized with two fragments of 0.4 kb and 2.5 kb respectively (results not shown). In order to clone these fragments a library of *Hind*III fragments in pBR322 was constructed and screened by the colony hybridization technique using the tRNA^{Arg} probe. Plasmid DNA was prepared from the hybridizing clones, the mycoplasma DNA fragment was retrieved by cleaving with *Hind*III, and analysed by electrophoresis, blotting, and hybridization to the tRNA probe. *Hind*III digests of total genomic DNA were run simultaneously and for most of the clones the length of the insert agreed with either the 0.4 or 2.5 kb *Hind*III fragment in the genomic DNA lane. Two clones, designated pMM22 and pMM25, containing the 0.4 kb and the 2.5 kb insert respectively, were selected for sequence analysis.

Nucleotide sequence of the pMM22 fragment

The nucleotide sequence of the 0.4 kb insert of pMM22 was determined as outlined in Fig. 1(a). The sequence, shown in Fig. 2, reveals a region which can be folded into the characteristic cloverleaf stem and loop structure of a tRNA (Fig. 3). It has the anticodon TCT, corresponding in the gene product to NCU, where N is either U or a derivative of U, which defines it as the gene for an arginine tRNA that reads the codons AGA and AGG. Preliminary data from a sequence analysis of arginine tRNAs from mycoplasma in our laboratory suggest that in the tRNA corresponding to the tRNA^{Arg} gene the uridine in the wobble position is modified. The sequences flanking the tRNA^{Arg} gene are relatively rich in A and T and no other tRNA genes are located there. A putative Pribnow box, TATGAT, was encountered at a position 28 nucleotides upstream of the tRNA^{Arg} gene as

```

AAGCTTTATATGATAAAAAACTATTAATGAAATTACTTATAAAATATCTAACTTTTATCAAATTTTAACATGTAAAAATAAGATGACTATAGTCA
                                                                                                     100
GACTTTATAATTATTCACAAATTAGAAAAATTTTAAAAATTAATTAAGATTGTTTCTTTTCATATATAAATATGATAAATTTATATGATTGTTATTGC
                                                                                                     200
                                                                                                     *****
                                                                                                     *****
CCATGTAGCTCAGTAGGATAGAGCACGCGCCTTCTAAGCGTGAGGTGGAAAGTTTCGAGCCTTCTCGTGGGCACCATCTAGAAAACTAAATCGTGCAAAAG
-----Arg-----
                                                                                                     300
CAGGTTCTTTTATTTTAAAGAGAACCATCATGAAAACTACTATGAACAAACACTAGATCAAATTAGAGATCTAATTGATAATAATAAATTTGACCAAGC
=====>          <=====
                                                                                                     400
TT
    
```

Fig. 2. Nucleotide sequence of the pMM22 fragment

The tRNA^{Arg} gene sequence as well as putative promoter and terminator signals are indicated in the Figure.

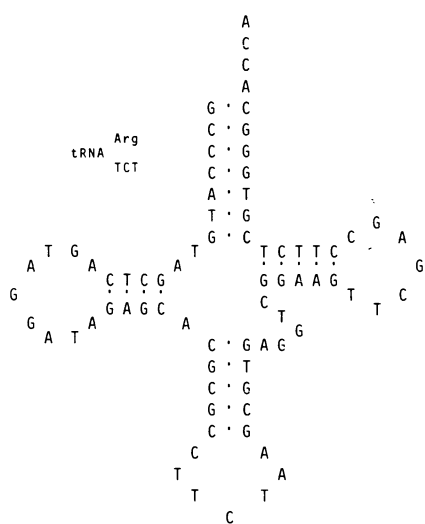


Fig. 3. Cloverleaf structure of the tRNA^{Arg} gene in the pMM22 fragment

indicated in Fig. 2. Upstream of this box is the putative -35 region consensus sequence TTGAAA. Downstream of the tRNA gene there is a hairpin loop which may be an RNA polymerase terminator site. It is conceivable that the tRNA^{Arg} gene forms a transcriptional unit that contains no other RNA or protein genes.

At present we have no information on the codon usage in *M. mycoides*, but Muto *et al.* (1984) have recently shown that in *Mycoplasma capricolum* there is a preference for using codons rich in A and T. For instance, of the arginine codons AGA is much more frequently used than all the CGN codons together. Assuming that this is true also in *M. mycoides* and provided that codon usage is correlated to the availability of the tRNAs (Ikemura, 1981), the tRNA corresponding to the tRNA^{Arg} gene should be the most abundant arginine tRNA in this organism. This is in contrast to the situation in *E. coli*, where the CGN codons are the arginine codons most frequently used and the tRNA^{Arg} with the anticodon ICG is the dominant arginine tRNA.

```

TAAAAGTATTAATATTTTTTCTTATTTATGTTGTAATGGAATGATATATCTATAATAATAAAGTCATTTAATATGATAAGATAATAACAAAGATAA
                                                                                                     100
                                                                                                     *****
                                                                                                     *****
ATTTTATATGCGCCCGTAGATCAATTGGATAGATCGCTTGACTACGGATCAAAAGGTTGGGGTTCGAGTCCCTCCGGGCGCACCATTTAGAAAAATAAT
-----Arg-----
                                                                                                     200
TCTTATTATCTTTAAACATAATAACACGATTCGGGAAGTGGCTCAGTTTGGTAGAGCATTGGTTTGGGACCGAAGGGTCGAGGTTCAAATCCTGTC
-----Pro-----
                                                                                                     300
TTCCCGACCATTATAATAATGGGCCCTTAGCTCAGCTGGGAGAGCACCTGCCTTGACGCGAGGGGTTCGACGGTTCGATCCCGTTAGGGTCCACCATTAT
-----Ala-----
                                                                                                     400
GGCGGGGTAGCTCAGTTGGTTAGAGCGTTCGGTTCATACCCGAAAGGTCGAGAGTTCAACTCTCTCCCCGCTACCATATTTTATGGACCTTTAGCT
-----Met-----
                                                                                                     500
CAGTTGGTTAGAGCATCCGGCTCATAACCGGACGGTTCATTGGTTCAGTCCAATAAGGTCACCATATTAATAAATAAATATTTTTAGTATGAAAGAAATG
-----Ile-----
                                                                                                     600
GTAGTACGGAGATTACCCAAGTCCGGCTGAAGGGATCGGTCTTGAACCGAGAGTTCGGGGAAACCCGAGCGGGGGTTCGAATCCCTCATCTTCCGCCA
-----Ser-----
                                                                                                     700
AATATTAATAATTTAAAAACAACGCGGGGTAGAGCAGTTGGTAGCTCGCCGGGCTCATAACCCGGAGGCCGAGGTTTCGAGTCTGCCCCGCAACCAAT
-----fMet-----
                                                                                                     800
GGCCCATAGCGAACGTTGGTTATCGCGCTCCCTGTACGGAGGAGATCACGGGTTTCGAGTCCCGTTGGGGTTCGCCATTTTTTATGGTCGTGTAGCTCA
-----Asp-----
                                                                                                     900
GTCGGTAGAGCAGCAGACTGAAGCTCTCGTGTGGCGGTTCAATTCGTCACGACACCACCTTGAAACAAAAATTATAAGTAGTCATAAGACTACTT
-----Phe-----
                                                                                                     1000
TTTTTTATTTTAAAAACTAATAAGTATATAATTACTTTATTTTTACTAGTAATTGGAGCAATAATTATGAAC
    
```

Fig. 4. Nucleotide sequence of a portion of the pMM25 fragment

The tRNA gene sequences as well as putative promoter and terminator signals are indicated in the Figure.

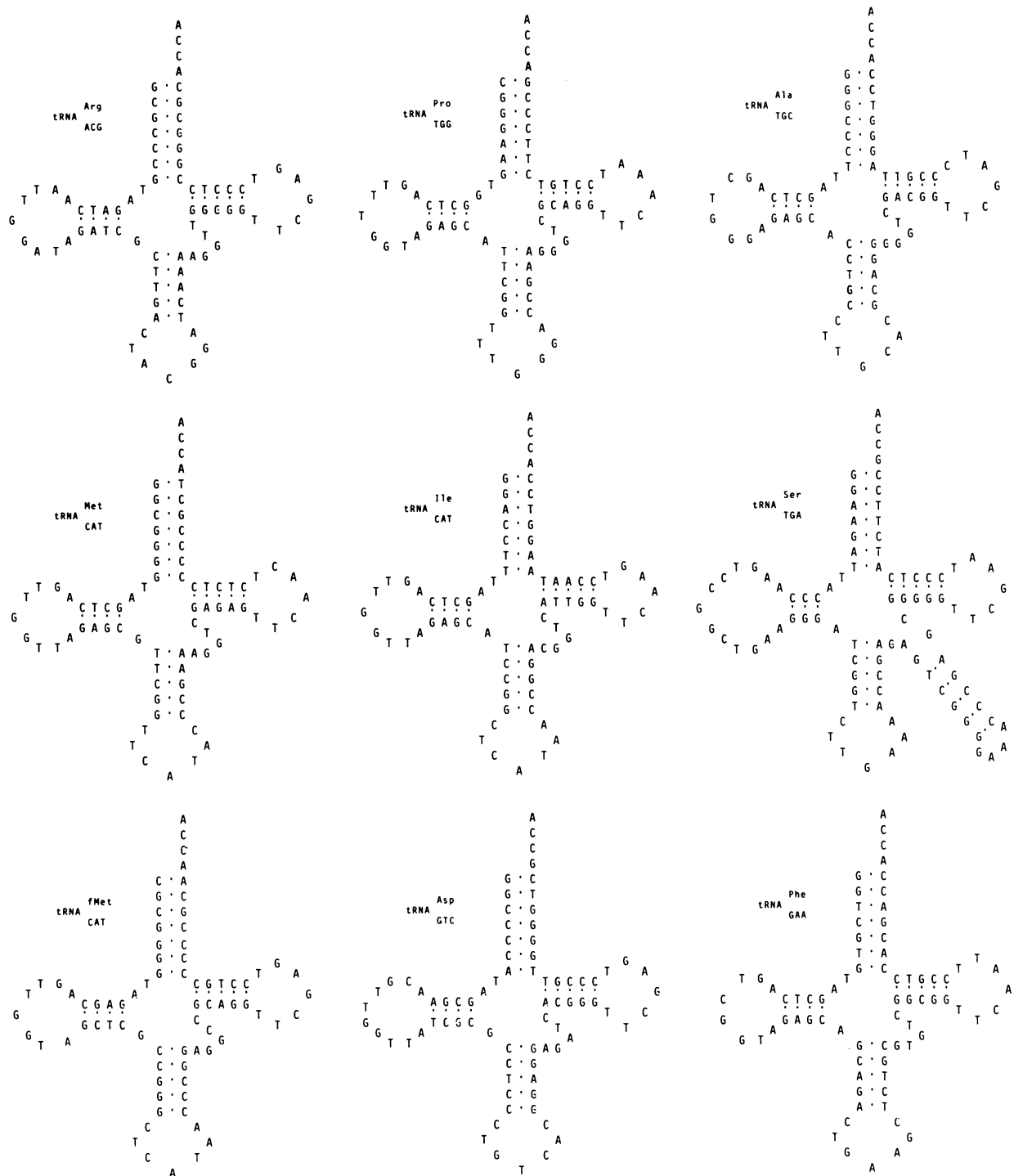


Fig. 5. Cloverleaf structures of the tRNA genes in the pMM25 fragment

Nucleotide sequence of the pMM25 fragment

Part of the 2.5 kb fragment of pMM25 was sequenced using the strategy indicated in Fig. 1(b). The nucleotide sequence, shown in Fig. 4, revealed the expected tRNA^{Arg} gene and in addition to this another eight tRNA genes. The arrangement of these genes is shown schematically in Fig. 1(b) and the cloverleaf structures corresponding to the individual genes are given in Fig. 5. The sequences that extend 350 bp upstream and 200 bp

downstream of the sequence shown in Fig. 4 were also determined (results not shown) but no tRNA genes were found in these parts of the fragment. A tentative promoter site is indicated in Fig. 4 as well as a possible terminator hairpin loop. In the spacer regions between the individual tRNA genes there are no such obvious promoter or terminator sites. It is therefore reasonable to assume that the tRNA gene cluster is a single transcriptional unit.

Clustering of tRNA genes is a common phenomenon

in many organisms. Particularly striking is the organization of the tRNA genes in *Bacillus subtilis* where most of these genes are concentrated to a few regions of the genome and are linked to ribosomal RNA operons. Green & Vold (1983) and Wawrousek (1984) have analysed a tRNA cluster from *B. subtilis* which contains 16 different tRNA genes and has some interesting features in common with the tRNA gene cluster shown in Fig. 4. Thus, exactly the same sequence of tRNA genes that is found in the pMM25 fragment (Arg-Pro-Ala-Met-Ile-Ser-Met-Asp-Phe) is also present in the *B. subtilis* cluster, although here it forms the middle part of a longer sequence. Furthermore, a comparison between the two organisms with respect to the nucleotide sequences of the individual tRNA genes shows a substantial homology. The corresponding genes all have the same anticodon and the overall homology is on average 85%. For comparison it may be noted that the homology between the tRNA genes of the pMM25 cluster and the primary structures of the pertinent *E. coli* tRNAs now available is only 75%. This finding is consistent with the view that *Mycoplasma* is more closely related to Gram-positive than to Gram-negative bacteria (Maniloff, 1983). On the other hand, the spacer regions in the pMM25 fragment do not appear to be at all homologous to the corresponding spacers in the *B. subtilis* cluster.

Rogers *et al.* (1984) have analysed a tRNA gene cluster from *Spiroplasma*, an organism closely related to *M. mycoides*. It is interesting to note that the order of tRNAs in the 5'-terminal part (the region sequenced so far) is Cys-Arg-Pro-Ala-Met-Ile-Ser. Thus, the *Spiroplasma* cluster also appears to be homologous to the one in *M. mycoides* with the exception that in *Spiroplasma* tRNA^{Cys} occurs at the 5'-terminal end, upstream of the tRNA^{Arg}. The homology between the two organisms with respect to the nucleotide sequences of the individual tRNA genes is on average 92%.

There are three tRNA genes in the pMM25 cluster that have the same anticodon, CAT, corresponding to NAU in the gene products where N is C or a derivative of C. The seventh tRNA gene in the cluster codes for a tRNA identical to the tRNA^{Met} that has been sequenced by Walker & RajBhandary (1978). The fourth tRNA gene is analogous to the tRNA^{Met} gene in *B. subtilis* mentioned above and we therefore assume that it is, indeed, a methionine tRNA gene. However, the amino acid assignment is less obvious for the remaining tRNA gene with the anticodon CAT, gene number five. Green & Vold (1983) conclude that the corresponding gene in the *B. subtilis* cluster codes for a tRNA specific for isoleucine rather than methionine. They reason that in *B. subtilis* only one tRNA^{Met}, apart from the initiator tRNA^{Met}, has been identified so that one of the putative tRNA^{Met} genes would be redundant. Furthermore, both *E. coli* (Kuchino *et al.*, 1980), T₄ (Fukuda & Abelson, 1980) and chloroplasts (Kashdan & Dudock, 1982) have isoleucine tRNAs with a modified C in the wobble position. Finally, the presence of an A·U base pair as the third pair in the acceptor stem of this tRNA may prevent it from being recognized by the methionine: tRNA ligase (Schulman, 1979): This base pair is present also in the corresponding *M. mycoides* sequence. A definite amino acid assignment for the tRNA encoded by this mycoplasma tRNA gene must await further studies, but in the present paper we have, nevertheless, tentatively identified it as an isoleucine tRNA gene.

It is interesting to note that the tRNA^{Phe} gene that appears at the 3'-terminus of the pMM25 cluster codes for a tRNA which is identical with the tRNA^{Phe} from *Mycoplasma capricolum* sequenced by Kimball *et al.* (1974). This finding supports the view that these two organisms are very closely related (Rogers *et al.*, 1985).

Four of the genes present in the cluster code for tRNAs that read codon families, as defined above, i.e. the genes for the arginine, alanine, proline and serine tRNAs. The tRNA^{Arg} gene has A in the position corresponding to the anticodon wobble nucleotide and its gene product therefore probably has I in this position. Since the two fragments sequenced in this paper are the only one that hybridize to the tRNA^{Arg} probe, *M. mycoides* would seem to contain only two arginine tRNAs, one which reads the CGN codons and one which reads the AGA and AGG codons. The genes for tRNA^{Ala}, tRNA^{Pro} and tRNA^{Ser} all have T in the position corresponding to the wobble nucleotide. Further investigation will hopefully reveal if any of their gene products, which should have U or a modified U in the wobble position, can read all four codons in a family as the glycine tRNA presumably can.

We thank Professor Ulf Lagerkvist for his comments on the manuscript. We are indebted to Anne-Marie von Essen and Tor Inge Johansen, for expert technical assistance. This investigation was supported by a grant from the Swedish Medical Research Council.

REFERENCES

- Andersson, S., Bankier, B. G., de Bruijn, M. H. L., Coulson, E. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Scheir, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1981) *Nature (London)* **290**, 457-465
- Bonitz, S. G., Berlan, R., Coruzzi, G., Li, M., Nobrega, F. G., Nobrega, M. P., Thalenfeld, B. E. & Tzagoloff, A. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3167-3170
- Bischoff, R., Graeser, E. & McLaughlin, L. W. (1983) *J. Chromatogr.* **257**, 305-315
- Crick, F. H. C. (1966) *J. Mol. Biol.* **19**, 548-555
- England, T. E., Bruce, A. G. & Uhlenbeck, O. C. (1980) *Methods Enzymol.* **65**, 65-74
- Fukuda, K. & Abelson, J. (1980) *J. Mol. Biol.* **139**, 377-391
- Gillam, I., Millward, S., Blew, D., von Tigerstrom, M., Wimmer, E. & Tener, G. M. (1967) *Biochemistry* **6**, 3043-3056
- Gillam, I., Blew, D., Warrington, R. C., von Tigerstrom, M. & Tener, G. M. (1968) *Biochemistry* **7**, 3459-3468
- Green, C. J. & Vold, B. S. (1983) *Nucleic Acids Res.* **11**, 5763-5774
- Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557-580
- Heckman, J. E., Sarnoff, J., Alzner-Deweerd, B., Yin, S. & RajBhandary, U. L. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3159-3163
- Ikemura, T. (1981) *J. Mol. Biol.* **146**, 1-21
- Kashdan, M. A. & Dudock, B. S. (1982) *J. Biol. Chem.* **257**, 11191-11194
- Kilpatrick, M. W. & Walker, R. T. (1980) *Nucleic Acids Res.* **8**, 2783-2786
- Kimball, M. E., Szeto, K. S. & Söll, D. (1974) *Nucleic Acids Res.* **1**, 1721-1732
- Kuchino, Y., Watanabe, S., Harada, F. & Nishimura, S. (1980) *Biochemistry* **19**, 2085-2089
- Lustig, F., Elias, P., Axberg, T., Samuelsson, T., Tittawella, I. & Lagerkvist, U. (1981) *J. Biol. Chem.* **256**, 2635-2643
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory, New York

- Maniloff, J. (1983) *Annu. Rev. Microbiol.* **37**, 477-499
- Marmur, J. (1961) *J. Mol. Biol.* **3**, 208-218
- Mitra, S. K., Lustig, F., Åkesson, B., Lagerkvist, U. & Strid, L. (1977) *J. Biol. Chem.* **252**, 471-478
- Mitra, S. K., Lustig, F., Åkesson, B., Axberg, T., Elias, P. & Lagerkvist, U. (1979) *J. Biol. Chem.* **254**, 6397-6401
- Muto, A., Kawauchi, Y., Yamao, F. & Osawa, S. (1984) *Nucleic Acids Res.* **12**, 8209-8217
- Razin, S. (1978) *Microbiol. Rev.* **42**, 414-470
- Rogers, M. J., Steinmetz, A. A. & Walker, R. T. (1984) *Israel J. Med. Sci.* **20**, 768-772
- Rogers, M. J., Simmons, J., Walker, R. T., Weisburg, W. G., Woese, C. R., Tanner, R. S., Robinson, I. M., Stahl, D. A., Olsen, G., Leach, R. H. & Maniloff, J. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1160-1164
- Samuelsson, T., Elias, P., Lustig, F., Axberg, T., Fölsch, G., Åkesson, B. & Lagerkvist, U. (1980) *J. Biol. Chem.* **255**, 4583-4588
- Samuelsson, T., Axberg, T., Borén, T. & Lagerkvist, U. (1983) *J. Biol. Chem.* **258**, 13178-13184
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467
- Schulman, L. (1979) in *Transfer RNA: Structure, Properties and Recognition* (Schimmel, P. R., Söll, D. & Abelson, J. N., eds.), pp. 311-324, Cold Spring Harbor Laboratory, Cold Spring Harbor
- Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3**, 2387-2398
- Walker, R. T. (1976) in *Proceedings of the International Conference on the Synthesis, Structure and Chemistry of tRNAs and Their Components*, Poznan, Poland, pp. 291-305, Dymaczewo, Poznan
- Walker, R. T. & RajBhandary, U. L. (1978) *Nucleic Acids Res.* **5**, 57-70
- Wawrousek, E. F., Narasimhan, N. & Hansen, J. N. (1984) *J. Biol. Chem.* **259**, 3694-3702

Received 8 July 1985; accepted 16 July 1985