Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 3166–3170, May 1987
Biochemistry

Apparent lack of discrimination in the reading of certain codons in *Mycoplasma mycoides*

(translation/tRNA/anticodon)

TORE SAMUELSSON, YOUSSEF S. GUINDY, FLORENTYNA LUSTIG, TOMAS BORÉN, AND ULF LAGERKVIST

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

Communicated by Viktor Mutt, January 27, 1987 (received for review June 19, 1986)

ABSTRACT **We report a cluster of four tRNA genes from *Mycoplasma mycoides* as well as the sequence of the alanine, proline, and valine tRNAs and the serine tRNA reading the UCN codons (where N stands for G, A, C, or U). This brings the total number of tRNA genes that we have so far characterized in this organism to 14, 6 of which code for tRNAs that read the codons of family boxes. In each of these latter cases, we found only one gene per family box, and the gene sequence contains a thymidine in the position corresponding to the wobble nucleotide, with the exception of the arginine tRNA gene that has an adenosine in this position. Furthermore, all of the tRNA structures reported here have an unsubstituted uridine in the wobble position. These findings are similar to those reported for mitochondria, especially yeast mitochondria, that contain an arginine tRNA with the anticodon ACG. However, the resemblance is not complete since we have demonstrated the presence of two isoacceptor tRNAs for threonine having uridine and adenosine, respectively, in the wobble position. It is suggested that in the *M. mycoides* at least some of the family codon boxes are read by only one tRNA each, using an unconventional method without discrimination between the nucleotides in the third codon position.**

The genetic code can be thought of as being made up of sixteen boxes with four codons in each box. All codons in a box have the same nucleotides in the first two positions; the variation between the codons is in the third position. In half of the boxes all four codons specify the same amino acid, and in what follows we will refer to such boxes as family boxes (Fig. 1). Regarding the codon families it makes no difference, as far as translational fidelity is concerned, how the third codon nucleotide is read since the first two are enough to specify the amino acid. In the rest of the boxes the codons have been divided up between amino acids or between amino acids and stop words. When reading the codons of these boxes, it is obviously necessary to be able to discriminate among nucleotides in the third codon position to avoid translational errors. To explain such discrimination, Crick in his wobble rules (1) excludes a number of pairings between the third nucleotide of the codon and the wobble nucleotide of the anticodon on structural grounds.

Most of the accumulated data concerning codon reading *in vivo* indicate that codons are read in the cell according to the wobble rules. Nevertheless, there is evidence in the literature for codon readings *in vivo* involving almost all of the forbidden base pairs. These exceptions to the rules are of two different types, (*i*) unconventional reading with discrimination among the nucleotides in the third codon position, albeit using principles other than those laid down in the wobble rules, and (*ii*) reading without discrimination in this position (ref. 2 and references therein).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

UUU Phe	UCU Ser	UAU Tyr	UGU Cys
UUC "	UCC "	UAC "	UGC "
UUA Leu	UCA "	UAA Stop	UGA Stop
UUG "	UCG "	UAG "	UGG Trp
CUU Leu	CCU Pro	CAU His	CGU Arg
CUC "	CCC "	CAC "	CGC "
CUA "	CCA "	CAA Gln	CGA "
CUG "	CCG "	CAG "	CGG "
AUU Ile	ACU Thr	AAU Asn	AGU Ser
AUC "	ACC "	AAC "	AGC "
AUA "	ACA "	AAA Lys	AGA Arg
AUG Met	ACG "	AAG "	AGG "
GUU Val	GCU Ala	GAU Asp	GGU Gly
GUC "	GCC "	GAC "	GGC "
GUA "	GCA "	GAA Glu	GGA "
GUG "	GCG "	GAG "	GGG "

FIG. 1. The genetic code. Family codon boxes are indicated by bold lines.

In the second category of unconventional codon reading, the four codons of a family box are read without nucleotide discrimination in the third position, reminiscent of the two-out-of-three reading that we have observed in protein synthesis *in vitro* using the phage MS2 RNA as message (3–7). The analysis of mitochondrial tRNA genes and their gene products in the laboratories of Sanger (8), Tzagoloff (9), and RajBhandary (10) has established beyond doubt that in mammalian, *Neurospora*, and yeast mitochondria the family boxes are read by only one tRNA each. In all these cases the tRNA gene has a thymidine in the position corresponding to the wobble nucleotide, and, whenever the primary structure of the tRNA itself is known, the wobble nucleotide has turned out to be an unsubstituted uridine (11). The only exception to this rule is the tRNA reading the arginine family box in yeast that has adenosine in this position (12). Thus, in unconventional codon reading of the second type, a new general principle has been established as illustrated by the fact that in yeast, *Neurospora*, and mammalian mitochondria all family codons are read without discrimination in the third position.

It has been suggested that mitochondria are of prokaryotic origin, and one may consequently ask if there are prokaryotes today that rely on unconventional codon reading without discrimination in the third position to read family boxes using a minimum number of tRNAs? Kilpatrick and Walker (13) have reported that *Mycoplasma mycoides* ssp. *capri* contains only one glycine tRNA and that this tRNA has an unsubstituted uridine in the wobble position. Nothing is known about codon usage in this *Mycoplasma*; but, in the very closely related *Mycoplasma capricolum* (14), the glycine codon GGU is used frequently, and, unless there is an absolute bias against GGU and GGC in *M. mycoides*, its

single glycine tRNA must be able to read these codons. We have tested this tRNA in an *in vitro* protein synthesizing system 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 (7). This is, of course, the result to be expected for a tRNA that had been specifically designed to read all four codons in a family, suggesting that this microorganism might at least for glycine use undiscriminating reading of the type outlined above.

To elucidate further such reading in *M. mycoides*, we have begun to inventory the tRNA genes and their gene products in this organism. We have reported (15) the cloning and sequencing of the following 10 genes: 1 isolated gene for arginine tRNA and a cluster of 9 genes for the arginine, proline, alanine, methionine, isoleucine, serine, formylmethionine, aspartate, and phenylalanine tRNAs. For each family box represented in the cluster only one tRNA gene was found, and these genes had thymidine in the position corresponding to the first anticodon nucleotide with the exception of the gene for arginine tRNA that had adenosine in this position. In the present paper we report another cluster of tRNA genes in *M. mycoides* containing the genes for the asparagine, glutamate, valine, and threonine tRNAs and present the primary structure of the alanine, proline, serine, threonine, and valine tRNAs. The genes that code for the tRNAs that read the threonine and valine codon boxes have thymidine in the position corresponding to the wobble nucleotide and, in the tRNA primary structures, this nucleotide is an unsubstituted uridine. Furthermore, the alanine and proline tRNAs and the serine tRNA reading the UCN codons (where N stands for G, C, A, or U) also have an unsubstituted uridine in the wobble position. Along with the previous results (15), the emerging picture of the tRNA genes and their gene products in *M. mycoides* thus shows a striking resemblance to the situation in mitochondria, particularly yeast mitochondria (8–12). However, the resemblance to mitochondrial codon reading is not complete since our analysis of the threonine tRNAs has demonstrated an additional isoacceptor tRNA for threonine with adenosine in the wobble position.

MATERIALS AND METHODS

Cloning Procedures and DNA Sequence Analysis. DNA of the phage vector Charon 28 was obtained from Bethesda Research Laboratories, and DNA from *M. mycoides* ssp. *capri* was prepared by the Marmur procedure (16). It was digested to completion with *Hin*dIII and inserted into the *Hin*dIII site of the Charon 28 vector. Plaques were analyzed by hybridization as described (15), using purified ^{32}P-labeled tRNA^{Val} from *M. mycoides* as a probe. Insert DNA was isolated from one of the hybridizing clones and part of this fragment was sequenced by the chain-termination method of Sanger *et al.* (17).

Purification and Sequence Analysis of tRNA. *M. mycoides* ssp. *capri* was grown and harvested, and crude tRNA was prepared from the cells as described (15). Individual tRNAs were purified by chromatography on benzoylated DEAE-cellulose (18). Further purification was achieved by rechromatography on benzoylated DEAE-cellulose after phenoxyacetylation of the tRNAs (19) followed by two-dimensional PAGE as described by Garel *et al.* (20). Purified tRNAs were sequenced using the Gupta and Randerath direct read out method (21). The wobble nucleotides were in some cases further authenticated by two-dimensional TLC (22).

RESULTS

Cloning and Nucleotide Sequence of *M. mycoides* tRNA Genes. We have described (15) the cloning and nucleotide sequence analysis of 10 tRNA genes from *M. mycoides* isolated from a library of *Hin*dIII fragments cloned in pBR322. In the present investigation, we constructed a Charon 28 library of *Hin*dIII fragments and screened it by plaque hybridization using purified ^{32}P-labeled tRNA^{Val} from *M. mycoides* as a probe. Several of the most strongly hybridizing clones were further analyzed and found to have a 4.5-kilobase insert. The nucleotide sequence of part of this fragment is shown in Fig. 2.

In addition to the expected tRNA^{Val} gene (anticodon TAC), the sequence revealed the following three additional tRNA genes: tRNA^{Asn} (GTT), tRNA^{Glu} (TTC), and tRNA^{Thr} (TGT). A likely candidate for a Pribnow box, TATTAT, is encountered at a position 45 nucleotides upstream of the gene for tRNA^{Asn} as indicated in Fig. 2. Further upstream is the putative −35 region consensus sequence TTGAAA. Downstream of the tRNA gene cluster there is a hairpin loop that may be an RNA polymerase terminator site. In the spacer regions between the genes, there are no such promoter or

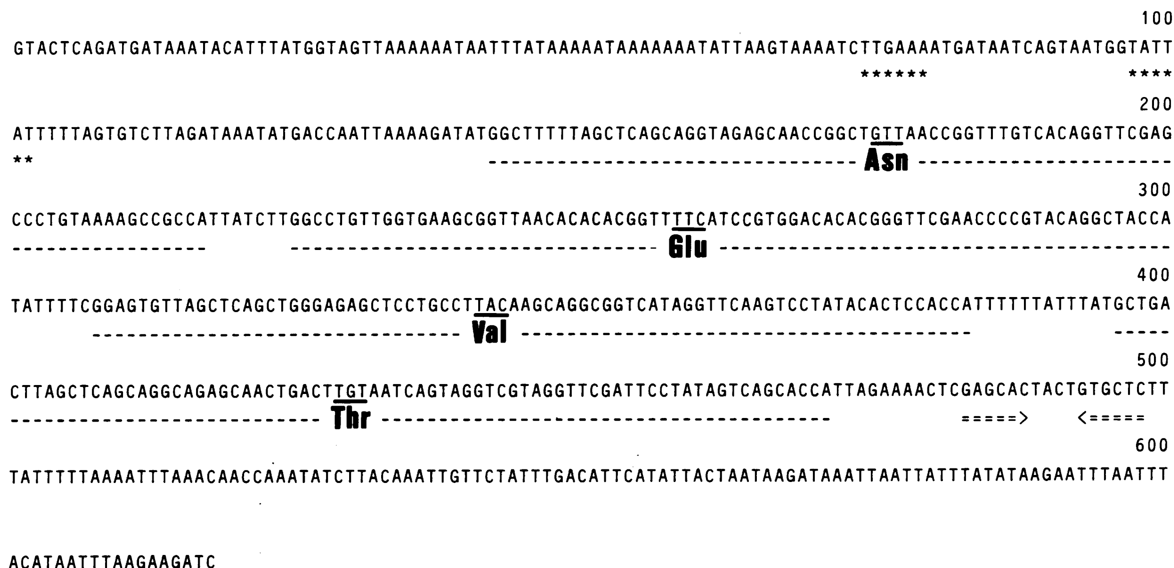

FIG. 2. The nucleotide sequence of the portion of the *Hin*dIII insert that contains the tRNA gene cluster. In the tRNA gene sequences shown the anticodons are underlined and the putative promoter and terminator sites are indicated by the symbols *** and =====><=====, respectively.

terminator sites, and it is, therefore, reasonable to assume that the tRNA gene cluster is a single transcriptional unit.

Two of the genes in the cluster code for tRNAs that read codon families, i.e., the genes for the valine and threonine tRNAs. These tRNAs have thymidine in the position corresponding to the wobble nucleotide.

Primary Structure of Five tRNAs from *M. mycoides*. As part of an inventory of the tRNAs of *M. mycoides*, we have determined the nucleotide sequence of five tRNAs that read the following family codon boxes: the alanine, proline, serine, threonine, and valine tRNAs. In these sequences, shown in cloverleaf form in Fig. 3, the 3′ and 5′ termini have been deduced from the gene sequences. The purification of the alanine, proline, and valine tRNAs as well as the serine tRNA that reads the codons UCN yielded only one isoacceptor of each species, consistent with our finding of only one gene in each case as described above and in ref. 15. For threonine tRNA, however, chromatography on benzoylated DEAE-cellulose revealed two peaks indicating the presence of more than one isoacceptor. When the primary structure of these tRNAs was determined, they were indeed found to be isoacceptors for threonine, one with an unsubstituted uridine in the wobble position (Fig. 3) and another with an adenosine in this position (unpublished results). The isoacceptor tRNA$^{Thr}_{UGU}$ corresponds to the threonine tRNA gene in the cluster shown in Fig. 2. The other isoacceptor tRNA$^{Thr}_{AGU}$ would be the first tRNA of nonmitochondrial origin that has been shown to have adenosine instead of inosine in the

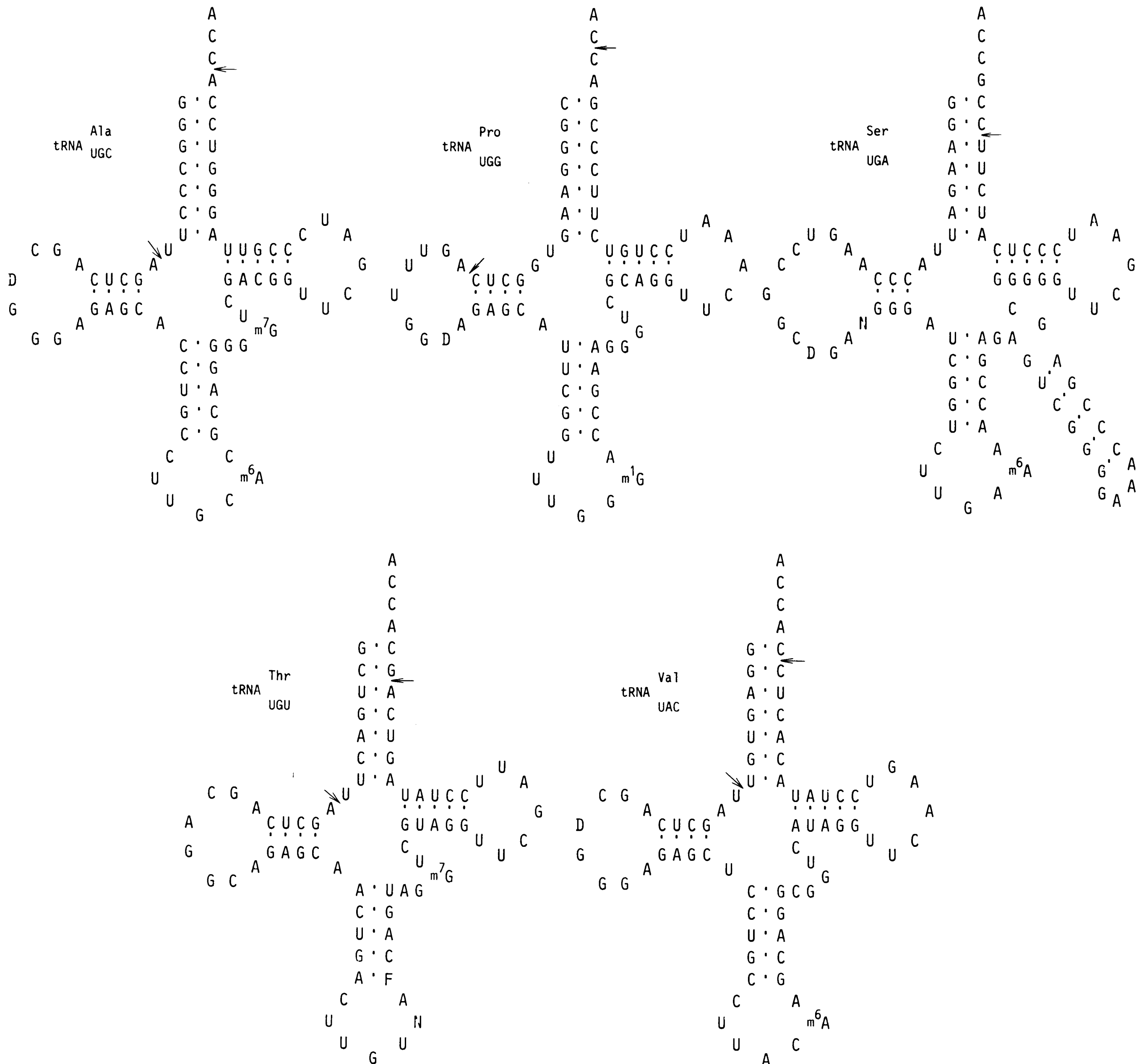

FIG. 3. Nucleotide sequences of tRNA^{Ala}, tRNA^{Pro}, tRNA^{Ser}, tRNA^{Thr}, and tRNA^{Val} shown in the cloverleaf form. The 5′- and 3′-terminal sequences indicated by arrows have been deduced from the gene sequences in Fig. 2 and ref. 15. Modified nucleotides could, therefore, occur in these terminal sequences. Furthermore, we do not know if uridine in position 8 of the tRNA^{Val} sequence obtained by the Gupta and Randerath method (21) is in fact 4-thiouridine (s^4U) in view of the extreme instability of this derivative. N in the tRNA^{Ser} and tRNA^{Thr} sequences is an unknown derivative of adenosine. m^6A, N^6-methyladenosine; D, dihydrouridine; F, pseudouridine; m^1G, 1-methylguanosine; m^7G, 7-methylguanosine.

Table 1. Anticodons of tRNAs that read family codons in *M. mycoides*

Codon family	Wobble nucleotide in gene sequence	Wobble nucleotide in tRNAs
Alanine (GCN)	T	U
Arginine (CGN)	A	
Glycine (GGN)		U
Leucine (CUN)		U
Proline (CCN)	T	U
Serine (UCN)	T	U
Threonine (ACN)	T	U A
Valine (GUN)	T	U

The table summarizes data presented in the present paper and refs. 13 and 15 as well as some unpublished results.

wobble position. Furthermore, an investigation of the leucine tRNAs in *M. mycoides* (unpublished results) has so far yielded only one tRNA for the leucine codon family CUN, and this tRNA has uridine in the wobble position. Thus, our own inventory of the *Mycoplasma* tRNAs that read family codon boxes and the results of Kilpatrick and Walker (13) would indicate that there is only one tRNA each for the alanine (GCN), glycine (GGN), leucine (CUN), proline (CCN), serine (UCN), and valine (GUN) codon boxes and that these tRNAs all have an unsubstituted uridine in the wobble position. The threonine codon box ACN, on the other hand, is read by two isoacceptor tRNAs, one with uridine and the other with adenosine in this position. These results are summarized in Table 1.

A striking feature of the tRNA sequences shown in Fig. 3 is that they all contain a comparatively small number of modified nucleotides, something that has been noted (13, 23) also for the glycine and formylmethionine tRNAs from this organism. In this context the presence of an unmodified uridine in position 55 of our sequences is particularly noteworthy. All bacterial tRNAs so far analyzed, including the glycine and formylmethionine tRNAs from *M. mycoides* sequenced by Kilpatrick and Walker (13) and Walker and RajBhandary (23), have pseudouridine in this position. However, when we sequenced (unpublished results) *Mycoplasma* glycine tRNA, we could not demonstrate any pseudouridine in position 55. The reason for this discrepancy is not clear, but one explanation could be that the strain of *M. mycoides* that we have used differs in this respect from the one used by Kilpatrick and Walker (13). On the other hand, our strain has at least one enzyme that can convert uridine into pseudouridine since the latter nucleotide is present in the anticodon stem of the threonine tRNA (Fig. 3).

DISCUSSION

The organization of the tRNA genes in the present cluster corresponds closely to that found in the cluster of nine genes that we have reported (15). There is a putative promoter in the expected position upstream of the first tRNA gene and a possible terminator site downstream of the gene cluster. An examination of the spacer regions between the genes reveals no likely candidates for additional promoter or terminator sites. It would, therefore, seem that this tRNA gene cluster is a single transcriptional unit. As in the reported cases (15), the 3′-terminal CCA is encoded in the tRNA genes.

We now have sequence information for a total of 14 tRNA genes in *M. mycoides*, 6 of which code for tRNAs that read family box codons. In each of these latter cases, we have found only 1 gene per family box. These genes contain a thymidine in the position corresponding to the wobble nucleotide, with the exception of the arginine tRNA gene that has an adenosine in this position. This is reminiscent of the mitochondrial situation, and it is particularly noteworthy that in the yeast mitochondrion there is an arginine tRNA with the anticodon ACG (8–12). The resemblance is further strengthened by our finding that the *Mycoplasma* tRNAs that read the family box codons for alanine, proline, serine, and valine all have an unsubstituted uridine in the wobble position, as has the glycine tRNA sequenced by Kilpatrick and Walker (13). Furthermore, our analysis of the leucine tRNA from *M. mycoides*, that reads the leucine (CUN) family box, has so far revealed only one isoacceptor and this tRNA has an unsubstituted uridine in the wobble position (unpublished results). On the other hand, we have also demonstrated that there are two tRNA isoacceptors for threonine in this organism, one that has an unsubstituted uridine in the wobble position (Fig. 3) and another that has an adenosine in this position (unpublished results and Table 1). The resemblance to the mitochondrial situation is, therefore, not complete. Nevertheless, it is in our opinion strong enough to warrant the suggestion that in the *M. mycoides* at least some of the family codon boxes are read by only one tRNA each, using an unconventional method without discrimination between the nucleotides in the third codon position. If this is so, it would represent the first case so far encountered of a free-living organism using this type of unconventional codon reading. To explain how such reading is possible, we must either assume that pyrimidine·pyrimidine oppositions can in these instances result in a stable base pair (10) or, alternatively, that codons may occasionally be read in spite of the fact that the anticodon cannot form a stable base pair with the third codon nucleotide, i.e., using a two-out-of-three mechanism (24).

We do not know yet what mechanism operates in unconventional codon reading without discrimination in the third codon position, whether it is of the two-out-of-three or the "uridine-reads-all" type. However, regardless of what the answer to this question may be, it seems obvious that the structural basis of the canonical principles of codon reading should be reexamined. In this context it is pertinent to ask what structural features, other than the anticodon itself, might influence the ability of a tRNA to read a certain codon. An obvious answer is of course that the ribosomes of for instance mitochondria might have unusual properties that could stabilize an unconventional codon·anticodon interaction. Another possibility would be that the structural environment provided for the anticodon by the tRNA molecule could affect its ability to read. This suggestion is consistent with our finding that the *Mycoplasma* glycine tRNA showed efficient codon reading without discrimination between the nucleotides in the third position when it was tested in an *in vitro* protein synthesizing system derived from *Escherichia coli* (7). Furthermore, Yarus and coworkers (25, 26) have proposed that the ability of an anticodon to suppress a nonsense mutation is influenced by the structure of the tRNA molecule.

Finally, it should be emphasized that, although unconventional codon reading is an interesting problem in its own right, its real significance lies in the fact that it is the mirror image of the translational fidelity problem. By studying unconventional codon reading, important information about the molecular basis for the selection of the correct codon·anticodon interaction in the reading of the mRNA might be obtained.

The expert technical assistance of A.-M. von Essen and M. Olsson is gratefully acknowledged. This investigation was supported by a grant from the Swedish Medical Research Council.

1. Crick, F. H. C. (1966) *J. Mol. Biol.* **19**, 548–555.
2. Lagerkvist, U. (1986) *BioEssays* **4**, 223–226.
3. Mitra, S. K., Lustig, F., Åkesson, B., Lagerkvist, U. & Strid, L. (1977) *J. Biol. Chem.* **252**, 471–478.
4. Mitra, S. K., Lustig, F., Åkesson, B., Axberg, T., Elias, P. &

Lagerkvist, U. (1979) *J. Biol. Chem.* **254**, 6397–6401.
5. Samuelsson, T., Elias, P., Lustig, F., Axberg, T., Fölsch, G., Åkesson, B. & Lagerkvist, U. (1980) *J. Biol. Chem.* **255**, 4583–4588.
6. Lustig, F., Elias, P., Axberg, T., Samuelsson, T., Tittawella, I. & Lagerkvist, U. (1981) *J. Biol. Chem.* **256**, 2635–2643.
7. Samuelsson, T., Axberg, T., Borén, T. & Lagerkvist, U. (1983) *J. Biol. Chem.* **258**, 13178–13184.
8. 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. USA* **77**, 3164–3166.
9. 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. USA* **77**, 3167–3170.
10. Heckman, J. E., Sarnoff, J., Alzner-Deweerd, B., Yin, S. & RajBhandary, U. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3159–3163.
11. Sprinzl, M., Moll, J., Meissner, F. & Hartmann, T. (1985) *Nucleic Acids Res.* **13** Suppl., 1–57.
12. Sibler, A.-P., Dirheimer, G. & Martin, R. P. (1986) *FEBS Lett.* **194**, 131–138.
13. Kilpatrick, M. W. & Walker, R. T. (1980) *Nucleic Acids Res.* **8**, 2783–2786.
14. Muto, A., Yamao, F., Kawauchi, Y. & Osawa, S. (1985) *Proc. Jpn. Acad.* **61**, 12–15.
15. Samuelsson, T., Elias, P., Lustig, F. & Guindy, Y. S. (1985) *Biochem. J.* **232**, 223–228.
16. Marmur, J. (1961) *J. Mol. Biol.* **3**, 208–218.
17. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
18. Gillam, I., Millward, S., Blew, D., von Tigerstrom, M., Wimmer, E. & Tener, G. M. (1967) *Biochemistry* **6**, 3043–3056.
19. Gillam, I., Blew, D., Warrington, R. C., von Tigerström, M. & Tener, G. M. (1968) *Biochemistry* **7**, 3459–3468.
20. Garel, J. P., Garber, R. L. & Siddiqui, M. A. O. (1977) *Biochemistry* **16**, 3618–3624.
21. Gupta, R. C. & Randerath, R. (1979) *Nucleic Acids Res.* **6**, 3443–3458.
22. Silberklang, M., Prochiantz, A., Haenni, A.-L. & RajBhandary, U. L. (1977) *Eur. J. Biochem.* **72**, 465–478.
23. Walker, R. T. & RajBhandary, U. L. (1978) *Nucleic Acids Res.* **5**, 57–70.
24. Lagerkvist, U. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1759–1762.
25. Yarus, M., McMillan, C., III, Cline, S., Bradley, D. & Snyder, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5092–5096.
26. Yarus, M. (1982) *Science* **218**, 646–652.