Proc. Natl. Acad. Sci. USA
Vol. 77, No. 6, pp. 3159–3163, June 1980
Biochemistry

Novel features in the genetic code and codon reading patterns in *Neurospora crassa* mitochondria based on sequences of six mitochondrial tRNAs

(UGA codon/mitochondrial evolution/codon-reading patterns)

JOYCE E. HECKMAN, JOSHUA SARNOFF, BIRGIT ALZNER-DEWEERD, SAMUEL YIN, AND UTTAM L. RAJBHANDARY

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Communicated by H. Gobind Khorana, February 11, 1980

ABSTRACT **We report the sequences of *Neurospora crassa* mitochondrial alanine, leucine$_1$, leucine$_2$, threonine, tryptophan, and valine tRNAs. On the basis of the anticodon sequences of these tRNAs and of a glutamine tRNA, whose sequence analysis is nearly complete, we infer the following: (*i*) The *N. crassa* mitochondrial tRNA species for alanine, leucine$_2$, threonine, and valine, amino acids that belong to four-codon families (GCN, CUN, ACN, and GUN, respectively; N = U, C, A, or G) all contain an unmodified U in the first position of the anticodon. In contrast, tRNA species for glutamine, leucine$_1$, and tryptophan, amino acids that use codons ending in purines (CA^A_G, UU^A_G, and UG^A_G, respectively) contain a modified U derivative in the same position. These findings and the fact that we have not detected any other isoacceptor tRNAs for these amino acids suggest that *N. crassa* mitochondrial tRNAs containing U in the first position of the anticodon are capable of reading all four codons of a four-codon family whereas those containing a modified U are restricted to reading codons ending in A or G. Such an expanded codon-reading ability of certain mitochondrial tRNAs will explain how the mitochondrial protein-synthesizing system operates with a much lower number of tRNA species than do systems present in prokaryotes or in eukaryotic cytoplasm. (*ii*) The anticodon sequence of the *N. crassa* mitochondrial tryptophan tRNA is U*CA and not CCA or CmCA as is the case with tryptophan tRNAs from prokaryotes or from eukaryotic cytoplasm. Because a tRNA with U*CA in the anticodon would be expected to read the codon UGA, as well as the normal tryptophan codon UGG, this suggests that in *N. crassa* mitochondria, as in yeast and in human mitochondria, UGA is a codon for tryptophan and not a signal for chain termination. (*iii*) The anticodon sequences of the two leucine tRNAs indicate that *N. crassa* mitochondria use both families of leucine codons (UU^C_A and CUN; N = U, C, A, or G) for leucine, in contrast to yeast mitochondria [Li, M. & Tzagoloff, A. (1979) *Cell* 18, 47–53] in which the CUA leucine codon and possibly the entire CUN family of leucine codons may be translated as threonine.**

Mitochondria exist within the cytoplasm of eukaryotic cells (1). They contain a DNA genome and a protein-synthesizing system that is distinct from the system in the cytoplasm (2, 3). Virtually all of the protein components of the mitochondrial protein biosynthetic machinery are coded for by the nuclear DNA, made in the cytoplasm, and imported into the mitochondria. In contrast, it appears that all the RNAs necessary for mitochondrial protein synthesis (ribosomal, transfer, messenger) are made inside the mitochondrion. Although the sizes of mitochondrial DNAs from different sources vary, in all cases the mitochondrial DNA codes for 8–12 proteins, at least two ribosomal RNAs, and several tRNAs (2, 3).

A puzzling observation until now has been that the number of different tRNA species present in fungal (4, 5), amphibian (6), and mammalian (7) mitochondria is less than the minimum, 32, needed to read all the codons of the genetic code using the codon/anticodon pairing rules proposed by Crick (8). How then does the mitochondrial protein-synthesizing system read all the codons of the genetic code? Several possibilities may be considered: (*i*) import of certain tRNAs into mitochondria from cytoplasm (9), (*ii*) restricted codon usage in mitochondrial mRNAs, and (*iii*) unusual codon-reading abilities of mitochondrial tRNAs (2, 10). According to this last possibility, certain mitochondrial tRNAs could read more than the two or three codons (11) allowed by the wobble hypothesis of Crick (8). In this paper we show that this last possibility is most likely the correct explanation in *Neurospora crassa* mitochondria and we propose a molecular basis for this based on the tRNA sequences that we have determined.

Previously, we described the sequences of *N. crassa* mitochondrial formylmethionine and tyrosine tRNAs (12, 13) and discussed some of the interesting and unusual structural features that we found in these tRNAs. Here, we report the sequences of six more of the *N. crassa* mitochondrial tRNAs. Based on the unusual nature of the anticodon sequences of some of these tRNAs, we suggest that some *N. crassa* mitochondrial tRNAs are capable of reading all the codons of a four-codon family. Similar conclusions have been arrived at by Barrell *et al.* (14) based on the DNA sequences corresponding to tRNAs in human mitochondria. In addition, the anticodon sequence of *N. crassa* mitochondrial tryptophan tRNA suggests that UGA is a codon for tryptophan and not for chain termination in *N. crassa* mitochondria.

MATERIALS AND METHODS

Purification of tRNAs. This was carried out essentially as described (12, 13) using two steps of RPC-5 column chromatography except purification of tryptophan tRNA required an additional step involving electrophoresis on a 15% polyacrylamide slab gel. Threonine tRNA was also purified in two steps, a single RPC-5 chromatography (13) step followed by electrophoresis on 15% polyacrylamide gel.

Sequence Analysis of tRNAs. The procedures used were as described for the *N. crassa* mitochondrial initiator and tyrosine tRNAs (12, 13) and have been described in detail in a recent review (15). All of the oligonucleotides produced by complete TI RNase or pancreatic RNase digestions of the tRNAs were labeled at their 5′ ends with ^{32}P (16) and their sequences were determined. Various methods involving *in vitro* ^{32}P labeling were then used to obtain the overlap data necessary to align these oligonucleotides into a unique sequence for the tRNAs (13, 15).

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.

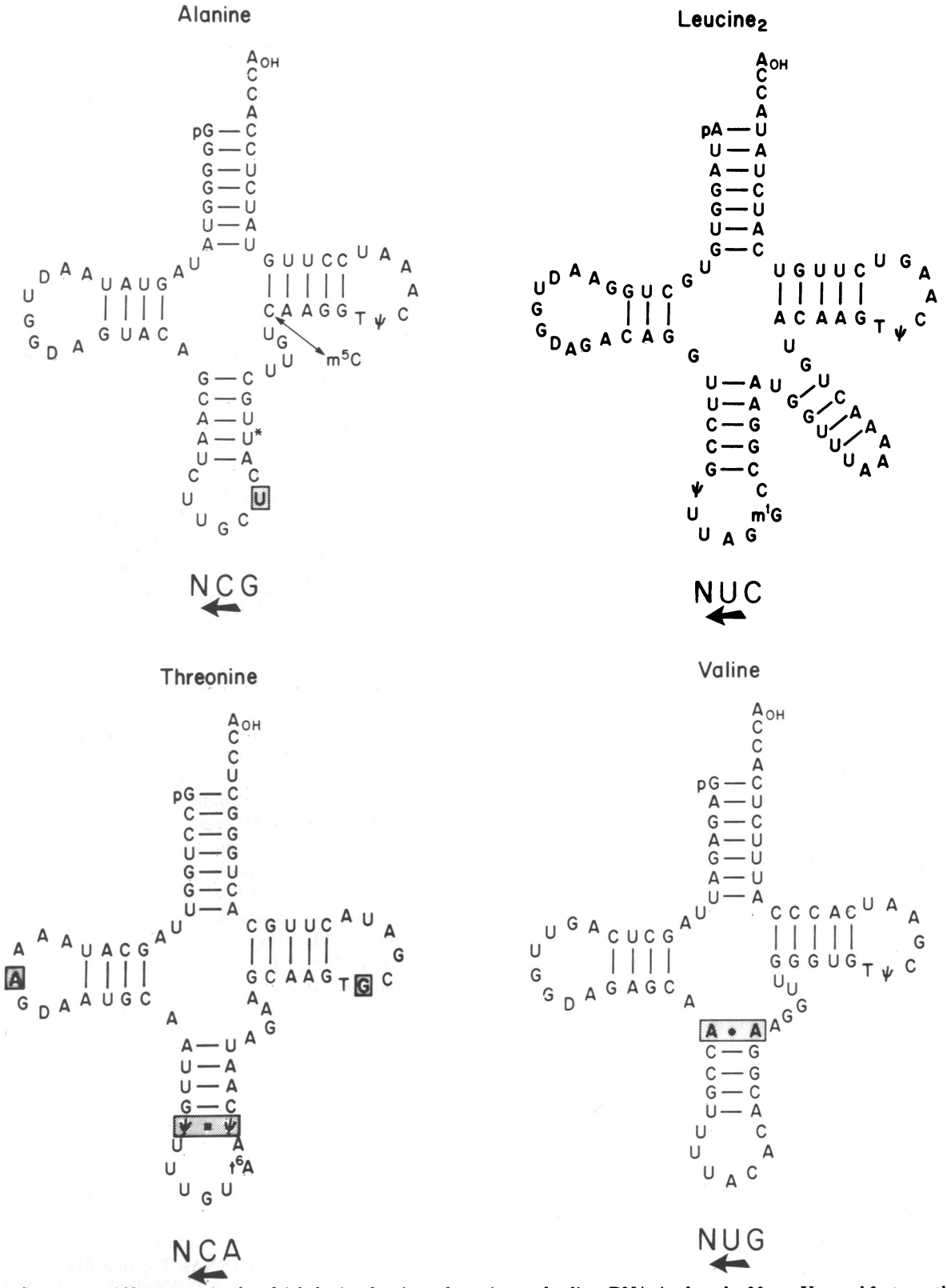

FIG. 1. Sequences of *N. crassa* mitochondrial alanine, leucine$_2$, threonine, and valine tRNAs in cloverleaf form. Unusual features that distinguish these mitochondrial tRNAs from most prokaryotic or eukaryotic cytoplasmic tRNAs are shown in boxed regions. The threonine tRNA shows the same unusual feature in the D loop and TΨC loop noted previously in *N. crassa* mitochondrial initiator tRNA (12) and also in yeast mitochondrial aspartic acid tRNA (A. Tzagoloff, personal communication). Also shown underneath each tRNA sequence are the potential codons (N = A, C, U, or G) that could be read by these tRNAs.

RESULTS AND DISCUSSION

Codon-Reading Patterns in Mitochondria. Figs. 1 and 2 show in cloverleaf form the six *N. crassa* mitochondrial tRNAs whose sequences are being reported here.

The most striking result is that, among the tRNAs analyzed, the alanine, leucine$_2$, threonine, and valine tRNA species, whose amino acids use four-codon families (GCN, CUN, ACN, and GUN, respectively) all contained an unmodified U in the first

position of the anticodon (Fig. 1), whereas the glutamine (unpublished result), leucine$_1$, and tryptophan tRNAs (Fig. 2), whose amino acids use codons ending in purines (CA^A_G, UU^A_G, and UG^A_G, respectively) contained a modified U at the same position. The presence of an unmodified U residue in the first position of an anticodon sequence is unusual and contrasts with the situation in tRNAs from prokaryotes and from eukaryotic cytoplasm in which a U residue at this position is always modified (17). The only exception to this is a yeast leucine tRNA (18). The fact that an unmodified U is present in the anticodon sequence of only those tRNAs corresponding to amino acids with four-codon families combined with the fact that we have so far not detected any other isoacceptor species for these tRNAs suggests that mitochondrial tRNAs with an unmodified U in the first position of the anticodon are capable of reading all the codons of a four-codon family (11), whereas those having a modified U at this position, such as glutamine, leucine$_1$, and tryptophan tRNAs, are restricted to reading codons ending in A or G.

If this hypothesis is correct, there would be no need to have multiple isoacceptor tRNAs for the eight four-codon families in the genetic code. This would reduce the number of tRNA species needed to read all the codons of the genetic code from 32 to 24, which is about the number of tRNA species that have been found in most fungal, amphibian, and mammalian mitochondria (4–7). Our results do not exclude the possibility that restricted codon usage also accounts in part for the reduced number of tRNA species in some mitochondria, but the available data on codon utilization patterns in yeast (19, 20) and human (21) mitochondria indicate that this cannot be the sole explanation.

Our hypothesis suggests that U in the first position of the anticodon of a mitochondrial tRNA can pair not only with A and G as allowed for by the wobble hypothesis (8) but also with U and C. Although U·U and U·C pairs can form two hydrogen bonds each, in the wobble hypothesis (8) such pairings were considered possible but unlikely because their formation would necessitate bringing the glycosidic bonds much closer together than in the standard base pairs. More recently, on the basis of a study of complex lifetimes between tRNAs with complementary anticodons, Grosjean *et al.* (22) concluded that U·U and U·C base pairs do in fact contribute substantially to complex stability and proposed a conformational change involving a concerted rotation of the C(4′)-C(5′) and O(5′)-P bonds of the 3′-terminal nucleotide of the codon to bring the two bases close enough for pairing. It is likely that ribosomes play an important role in either facilitating or hindering such a conformational change. If such a conformational change is necessary for U·U or U·C pairing to occur, it is probable that mitochondrial ribosomes do facilitate it. Whether *Escherichia coli* and other ribosomes would also accommodate such a conformational change is unknown. It would be of interest to study the coding properties of the purified mitochondrial tRNAs containing U in the first position of the anticodon in both *E. coli* and mitochondrial protein-synthesizing systems to examine whether they can read all four codons in both systems or whether they would be more restricted in their coding properties in the *E. coli* system.

Our suggestion that U in the first position of the anticodon sequence of mitochondrial tRNAs can read all four nucleotides might at first glance appear to be the same as the "two out of three" codon-reading hypothesis of Lagerkvist (11) for amino acids belonging to four-codon families. The difference is that the four *N. crassa* mitochondrial tRNAs, reported here, which correspond to amino acids with four-codon families, and probably also the human mitochondrial tRNAs, contain U in the first position of the anticodon, whereas Mitra, Lagerkvist, and coworkers found that valine tRNAs containing I, G, or V (uridine-5-oxyacetic acid) in the first position of the anticodon can be made to read all four valine codons in an *E. coli in vitro* protein-synthesizing system (reviewed in ref. 11). Thus, the "two out of three" codon-reading hypothesis alone will not explain why these mitochondrial tRNAs contain an unmodified U as the first nucleotide of the anticodon. Also, experiments that compare the relative efficiency of valine tRNAs for reading valine codons by using either the base-pairing schemes proposed in the wobble hypothesis (8) or the "two out of three" codon-reading hypothesis (11) show that the tRNA using the base-

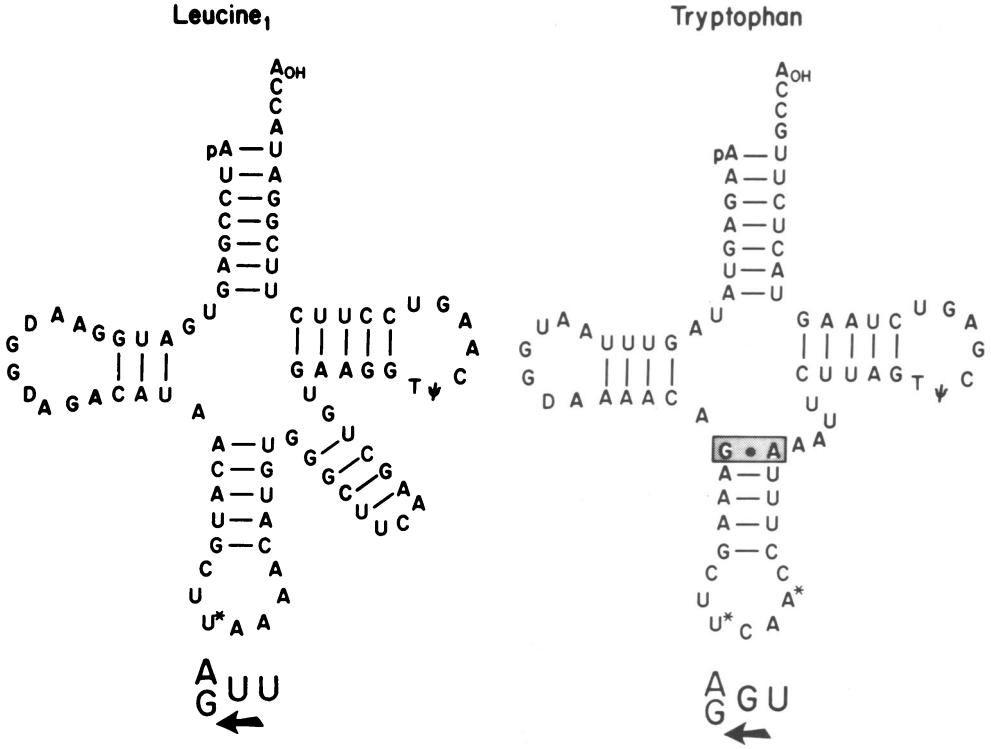

FIG. 2. Sequences of *N. crassa* mitochondrial leucine$_1$ and tryptophan tRNAs in cloverleaf form. Boxed region shows the unusual feature that distinguishes the tryptophan tRNA from other prokaryotic or eukaryotic cytoplasmic tRNAs. U*, modified derivative of U; A*, modified derivative of A. Also shown underneath each tRNA sequence are the potential codons that could be read by these tRNAs.

pairing scheme in the wobble hypothesis is more efficient than the latter by about an order of magnitude. This would explain why *E. coli* contains at least two tRNAs for reading all the codons of a four-codon family. In contrast, it appears that mitochondria in general contain only a single isoacceptor tRNA for reading all the codons of a four-codon family.

Evidence that a codon-reading pattern similar to that noted in *N. crassa* mitochondria might also operate in human mitochondria is provided by the work of Barrell *et al.* (14) who have determined the sequence of most of the human mitochondrial DNA. Their findings provide further support for the earlier work of Attardi, Davidson, and coworkers who showed that HeLa mitochondria contain a limited number of tRNA genes (7). Barrell *et al.* have shown that all the tRNAs corresponding to amino acids that belong to four-codon families contain T in the DNA at a site that corresponds to the first position of the anticodon sequence in the tRNA. Although the DNA sequence work does not allow one to predict whether the corresponding human mitochondrial tRNAs contain U or U* at this position, the results with *N. crassa* mitochondrial tRNAs suggest that they are more likely to contain U than U*.

The nature of the modification in U* present in the first position of the anticodon of *N. crassa* mitochondrial glutamine, leucine$_1$, and tryptophan tRNAs is not yet known. Based on its susceptibility to pancreatic RNase and its other properties, U* is not one of the known 2-thiouridine derivatives found at this position of some prokaryotic or eukaryotic cytoplasmic tRNAs (23). The modified nucleotide is relatively unstable and gives rise to double spots during analysis of partial digests of oligonucleotides containing it (Fig. 3). Besides the sequences of the tRNAs, we have also determined sequences of the DNA coding for $tRNA_1^{Leu}$ (unpublished data) and have confirmed that U* is derived from U and not from other nucleotides.

UGA as a Codon for Tryptophan in *N. crassa* Mitochondria. Another interesting result of this work concerns the anticodon sequence of the tryptophan tRNA. The anticodon sequence is U*CA and not CCA or CmCA as is the case with tryptophan tRNAs from prokaryotes or from eukaryotic cytoplasm (17). Because a tRNA with U*CA in the anticodon would be expected to read the codon UGA as well as the normal tryptophan codon UGG, this would imply that UGA is a codon for tryptophan and not a signal for chain termination in *N. crassa* mitochondria. In addition to providing support for recent inferences (based on comparison of human and yeast mitochondrial DNAs to beef heart cytochrome oxidase subunit II protein sequence) that UGA is a codon for tryptophan in human (21) and in yeast mitochondria (19, 20, 24), our results extend this novel feature in the genetic code of yeast and human mitochondria also to *N. crassa* mitochondria. Thus, the use of UGA as a sense codon for tryptophan instead of a termination codon may be a feature common to most mitochondrial protein-synthesizing systems.

Codons for Leucine in *N. crassa* Mitochondria. Although human, yeast, and *N. crassa* mitochondria appear to share certain common features in their genetic code or codon-reading patterns, there are other instances in which they differ. Tzagoloff and coworkers (25, 26) have shown that the CUA codon for leucine (and possibly the entire CUN family of leucine codons) may be translated as threonine in yeast mitochondria. This conclusion is based on a comparison of the DNA sequence coding for yeast mitochondrial ATPase subunit 9 to the homologous protein and the finding that a CTA sequence in the DNA appears as threonine in the protein (25, 27). Also, yeast mitochondria contain two isoacceptor tRNA species for threonine and only one for leucine (5). The sequence has been determined for one of these threonine tRNA genes. The DNA sequence data show that this threonine tRNA has an anticodon sequence that corresponds to the CUN family of leucine codons (26). In contrast to the situation in yeast mitochondria, *N. crassa* mitochondria contain two tRNAs for leucine and one for threonine. Furthermore, the tRNA sequence data show that one of the leucine tRNAs ($tRNA_1^{Leu}$, Fig. 2) is potentially capable

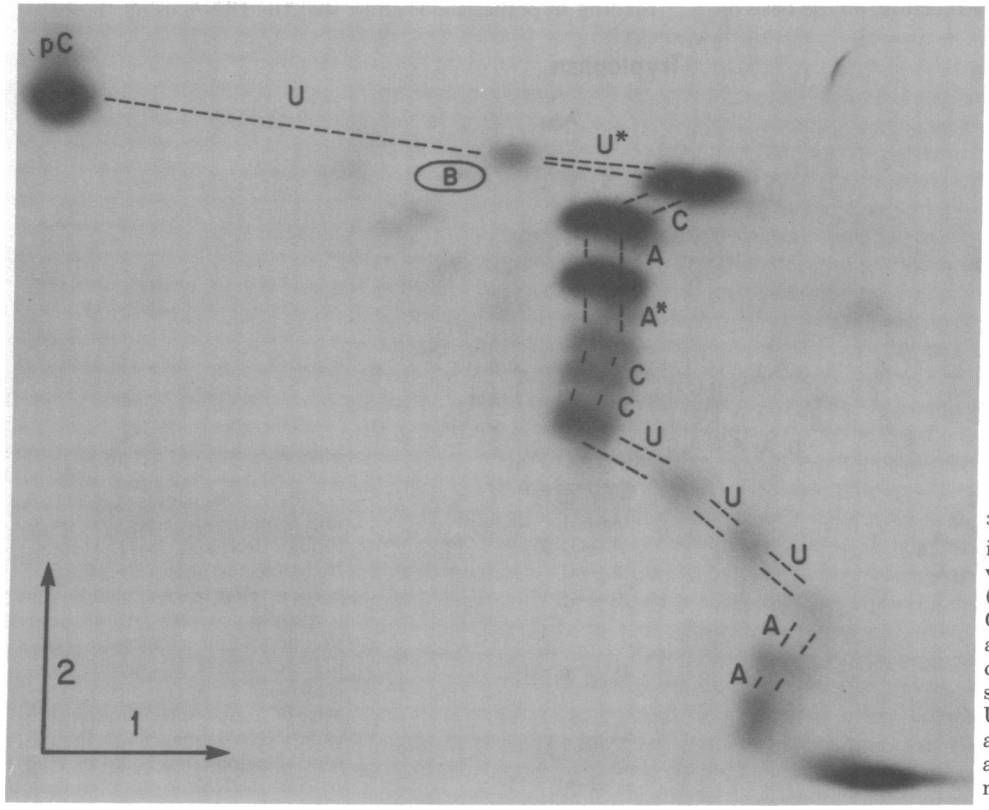

FIG. 3. Autoradiogram of 5′-^{32}P-labeled oligonucleotides present in a partial nuclease P1 and snake venom phosphodiesterase digestion (15) of 5′-[^{32}P]C-U-U*-C-A-A*-C-C-U-U-U-A-A-A-U-U-C-U-U-A-G as analyzed by two-dimensional homochromatography (12). Note the double spots in all oligonucleotides containing U* caused by the conversion of U* to another U derivative during digestion and work up. B, xylene cyanole blue marker.

of reading the UU$\frac{A}{G}$ family of leucine codons and the other one ($tRNA_2^{Leu}$, Fig. 1), the CUN family of leucine codons.

General Comments on the Code and Codon-Reading Patterns in Mitochondria. What are the overall implications of this modification in the genetic code and codon-reading patterns in human, yeast, and *N. crassa* mitochondria? There are only two cases in the genetic code in which an amino acid is specified by a single codon. These are UGG for tryptophan and AUG for methionine. As described above for human, yeast, and *N. crassa* mitochondria, UGA is also a codon for tryptophan. Thus, both UGA and UGG are codons for tryptophan in mitochondria. Similarly, at least in human mitochondria (21) it appears that, besides AUG, AUA also may be a codon for methionine. These findings along with the fact that a single tRNA may read all four codons of a four-codon family in mitochondria suggest that, at least in human and *N. crassa* mitochondria, the genetic code is a simple one in which codons are read by tRNAs *primarily* in sets of two (codons ending in purines comprising one set and those ending in pyrimidines the other) or in sets of four. In contrast, in prokaryotes or in eukaryotic cytoplasm, within four-codon families the four codons are read by two tRNAs (8), not usually in simple sets of two (23) but often with one of the tRNAs reading three codons ending in U, C, or A (generally in eukaryotes by tRNAs containing I in the anticodon) or in U, A, or G (generally in prokaryotes by tRNAs containing uridine-5-oxyacetic acid or 5-methoxy-U in the anticodon).

It is attractive to consider the possibility that the unusual codon-reading patterns in human and *N. crassa* mitochondria represent a simple, and hence perhaps primitive, form (28) of the genetic code. However, the data available do not allow us to rule out the alternate possibility that the genetic code and codon-reading patterns in mitochondria may have evolved to this current state as a result of unique selection pressures operating within the mitochondrion—for instance, selection pressures which (*i*) minimize the use of codons ending in G and (*ii*) minimize the number of different tRNA genes that are maintained within mitochondria. That changes in the genetic code can indeed arise in individual mitochondrial systems is emphasized by the finding that, although CUA (and possibly the CUN family of leucine codons) is translated as threonine in yeast mitochondria, CUN codes for leucine and not for threonine in *N. crassa* mitochondria. Thus, the use of CUA as a codon for threonine in yeast mitochondria probably arose well after the evolutionary divergence of these two organisms. Such a digression from the genetic code would be selected against in most systems, but it might survive in a small genome like the mitochondrion, if (*i*) the codon(s) in question was not used frequently and (*ii*) if the change did not affect in any significant manner the few proteins coded for by the mitochondrial genome. It would clearly be of interest to see how frequently the CUN series of codons are used in yeast mitochondria.

Finally, by using the purified tRNAs, whose sequences have been described, as hybridization probes we have reported the mapping and cloning of the genes for alanine, formylmethionine, $leucine_1$, $leucine_2$, threonine, tyrosine, and valine tRNAs (29, 30). The gene for glutamine tRNA has now been localized in *Hin*dIII fragment 10 of the mitochondrial DNA along with the gene for $tRNA_2^{Leu}$ and one of the genes for $tRNA^{fMet}$; that for $tRNA^{Trp}$ has been localized in *Hin*dIII fragment 7a, along with the gene for $tRNA^{Val}$.

We are grateful to Dr. W. Edgar Barnett for providing us with purified mitochondria and to Dr. P. Schimmel for suggestions on the manuscript. We thank Jan Hack for help in the purification of the tRNAs and Eric Ackerman for participating in the analysis of $tRNA^{Val}$. This work was supported by Grant NP114 from the American Cancer Society and Grant GM 17151 from the National Institutes of Health.

Note Added in Proof. Recent work indicates that a codon reading pattern similar to that described for *N. crassa* mitochondria may also be operative in yeast mitochondria (31).

1. Tedeschi, H. (1976) *Mitochondria: Structure, Biogenesis and Transducing Functions* (Springer, New York).
2. Borst, P. & Grivell, L. A. (1978) *Cell* **15**, 705–723.
3. Tzagoloff, A., Macino, G. & Sebald, W. (1979) *Annu. Rev. Biochem.* **48**, 419–441.
4. Terpstra, P., DeVries, H. & Kroon, A. M. (1977) in *Mitochondria 1977*, eds. Bandlow, W., Schweyen, R. J., Wolf, K. & Kaudewitz, F. (DeGruyter, New York), pp. 291–302.
5. Martin, N. & Rabinowitz, M. (1978) *Biochemistry* **17**, 1628–1634.
6. Dawid, I. B., Klukas, C. K., Ohi, S., Ramirez, J. L. & Upholt, W. B. (1976) in *The Genetic Function of Mitochondrial DNA*, eds. Saccone, C. & Kroon, A. M. (North-Holland, New York), pp. 3–14.
7. Angerer, L., Davidson, N., Murphy, W., Lynch, D. & Attardi, G. (1975) *Cell* **9**, 81–90.
8. Crick, F. H. C. (1966) *J. Mol. Biol.* **19**, 548–555.
9. Chiu, N., Chiu, A. & Suyama, Y. (1975) *J. Mol. Biol.* **99**, 37–50.
10. RajBhandary, U. L., Heckman, J. E., Yin, S., Alzner-DeWeerd, B. & Ackerman, E. (1979) *Cold Spring Harbor Monograph on Transfer RNA*, eds. Schimmel, P., Söll, D. & Abelson, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 3–17.
11. Lagerkvist, U. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1759–1762.
12. Heckman, J. E., Hecker, L. I., Schwartzbach, S. D., Barnett, W. E., Baumstark, B. & RajBhandary, U. L. (1978) *Cell* **13**, 83–95.
13. Heckman, J. E., Alzner-DeWeerd, B. & RajBhandary, U. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 717–721.
14. Barrell, B. G., Anderson, S., Bankier, A. T., deBruijn, 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.
15. Silberklang, M., Gillum, A. M. & RajBhandary, U. L. (1979) *Methods Enzymol.* **59**, 58–109.
16. Simsek, M., Ziegenmeyer, J., Heckman, J. E. & RajBhandary, U. L. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 1041–1045.
17. Sprinzl, M., Grueter, F., Spelzhaus, H. & Gauss, D. H. (1980) *Nucleic Acids Res.* **8**, r1–r22.
18. Randerath, K., Chia, L. S. Y., Gupta, R. C., Randerath, E., Hawkins, E. R., Brum, C. K. & Chang, S. H. (1975) *Biochem. Biophys. Res. Commun.* **63**, 157–163.
19. Coruzzi, G. & Tzagoloff, A. (1979) *J. Biol. Chem.* **254**, 9324–9330.
20. Fox, T. D. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6534–6538.
21. Barrell, B. G., Bankier, A. T. & Drouin, J. (1979) *Nature (London)* **282**, 189–194.
22. Grosjean, H., DeHenau, S. & Crothers, D. M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 610–614.
23. Nishimura, S. (1978) in *Transfer RNA*, ed. Altman, S. (MIT Press, Cambridge, MA), pp. 168–195.
24. Macino, G., Coruzzi, G., Norbrega, F. G., Li, M. & Tzagoloff, A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3784–3785.
25. Macino, G. & Tzagoloff, A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 131–135.
26. Li, M. & Tzagoloff, A. (1979) *Cell* **18**, 47–53.
27. Hensgens, L. A. M., Grivell, L. A., Borst, P. & Bos, J. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1663–1667.
28. Jukes, T. H. (1973) *Nature (London)* **246**, 22–26.
29. Heckman, J. E. & RajBhandary, U. L. (1979) *Cell* **17**, 583–595.
30. Heckman, J. E., Yin, S., Alzner-DeWeerd, B. & RajBhandary, U. L. (1979) *J. Biol. Chem.* **254**, 12694–12700.
31. Bonitz, S. G., Berlani, R., Coruzzi, G., Li, M., Nobrega, F. G., Nobrega, M. P., Thalenfeld, B. E. & Tzagoloff, A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3167–3170.