Proc. Natl. Acad. Sci. USA
Vol. 77, No. 6, pp. 3167–3170, June 1980
Biochemistry

Codon recognition rules in yeast mitochondria

(yeast mitochondrial code/mitochondrial tRNAs/mitochondrial decoding rules)

SUSAN G. BONITZ*, ROBERTA BERLANI*, GLORIA CORUZZI*, MAY LI*, GIUSEPPE MACINO†, FRANCISCO G. NOBREGA*, MARINA P. NOBREGA*, BARBARA E. THALENFELD*, AND ALEXANDER TZAGOLOFF*

*Department of Biological Sciences, Columbia University, New York, New York 10027; and †Istituto di Fisiologia Generale, Universita di Roma, 00100 Rome, Italy

Communicated by Cyrus Levinthal, March 11, 1980

ABSTRACT **The mitochondrial genome of *Saccharomyces cerevisiae* codes for 24 tRNAs. The nucleotide sequences of the tRNA genes suggest a unique set of rules that govern the decoding of the mitochondrial genetic code. The four codons of unmixed families are recognized by single tRNAs that always have a U in the wobble position of the anticodon. The codons of the mixed families are read by two different tRNAs. Codons terminating in a C or U are recognized by tRNAs with a G and codons terminating in a G or A are recognized by tRNAs with a U in the corresponding positions of the anticodons. There are two exceptions to these rules. In the AUN family for isoleucine and methionine, the isoleucine tRNA has a G and the methionine tRNA has a C in the wobble position. The tRNA for the arginine CGN family also has an A in the wobble position of the anticodon. It is of interest that the CGN codons have not been found in the mitochondrial genes sequenced to date. The simplified decoding system of yeast mitochondria allows all the codons to be recognized by only 24 tRNAs.**

Yeast mitochondria contain a complete set of tRNAs that function in the translation of a discrete number of mitochondrial messengers. During the past year, the sequences of most of the mitochondrial tRNA genes have been determined and virtually all of their anticodons are now known (1–4). Together with the codon usage in five mitochondrial genes, this information has permitted us to formulate the coding properties of this self-contained translational system.

In this communication we present evidence that yeast mitochondria are capable of interpreting their genetic code with less than the 32 tRNAs required by the wobble hypothesis (5). The reduction in the number of tRNAs is achieved by the use of a single tRNA to recognize the four codons in each of the eight unmixed families.‡ These families are read by the "two out of three" method first proposed by Lagerkvist (6). Furthermore, because only two tRNAs are used to discriminate among the codons of the mixed families, the entire system including the formylmethionine initiator consists of 24 tRNAs. The 24 tRNAs represent the minimum number needed to interpret a degenerate code whose amino acid assignments follow the general outline of the universal code.

Number of mitochondrial tRNAs in *Saccharomyces cerevisiae*

Several approaches have been used to estimate the number of mitochondrial tRNAs in yeast. Separation of the aminoacylated tRNAs by reverse phase chromatography on RPC-5 has provided information about the number of isoaccepting species for each of the 20 amino acids. Such analyses indicate 31–35 chromatographically distinct tRNA species (7–9). This number includes species that arise from posttranscriptional modification and therefore is an overestimate of the actual number of tRNA genes. For example, Berlani *et al.* (3) have recently shown that a point mutation in mitochondrial DNA abolishes charging of the four histidine tRNAs, suggesting that all four isoacceptors are products of a single gene.

Yeast mitochondrial tRNA genes have also been studied by hybridization of ^{3}H-labeled aminoacylated tRNAs (7, 8, 10) or of bulk ^{32}P-labeled tRNAs (11, 12) to mitochondrial DNA. The results of such hybridization experiments indicate the presence of only 22 or 23 tRNA genes that are distributed over four different regions of the genome (Fig. 1). The preponderance of tRNA genes cluster in region I which spans the quandrant between the *cap* and *oxi2* loci of the circular DNA. Wesolowski and Fukuhara (10) have mapped 17 tRNA genes in this span by "petite" deletion analysis. Borst and colleagues (11, 12) have concluded that there are 19 tRNA genes between the *cap* and *oxi2* loci based on the hybridization of bulk tRNAs to defined restriction fragments of the DNA. The second region, located between *oxi2* and *oxi3*, has been shown to code for proline, tryptophan, and formylmethionine initiator tRNA genes. The last two regions each code for a single tRNA. The glutamic tRNA gene has been mapped near the cytochrome *b* gene (10) and the serine tRNA, near the gene of subunit 9 of ATPase (12, 13).

Our DNA sequence analysis of the various regions described above has failed to uncover tRNA genes other than those already mapped by hybridization and petite deletion analysis. A continuous DNA sequence has been obtained in region I (ref. 3; unpublished data). The number and identity of the tRNA genes found in the DNA sequence agree well with the species previously mapped in this span by Wesolowski and Fukuhara (10). Based on the DNA sequences of the four tRNA regions, we find that there are 24 mitochondrial tRNA genes in yeast. Of these, the sequence of only the tryptophan gene is not known.

Codon utilization

The nucleotide sequences of five mitochondrial genes have yielded a fairly comprehensive list of the codons used in yeast mitochondria. A tabulation of the codons found in the genes of cytochrome *b* (unpublished data), subunits 2 and 3 of cytochrome oxidase (14, 15), and subunits 6 and 9 of ATPase (16–18)

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.

‡ The unmixed families consist of four codons that specify a single amino acid (e.g., UCN for serine). The mixed families code for two amino acids or a single amino acid and a terminator (e.g., UUN for phenylalanine and leucine, UAN for tyrosine and the ochre and amber terminators.)

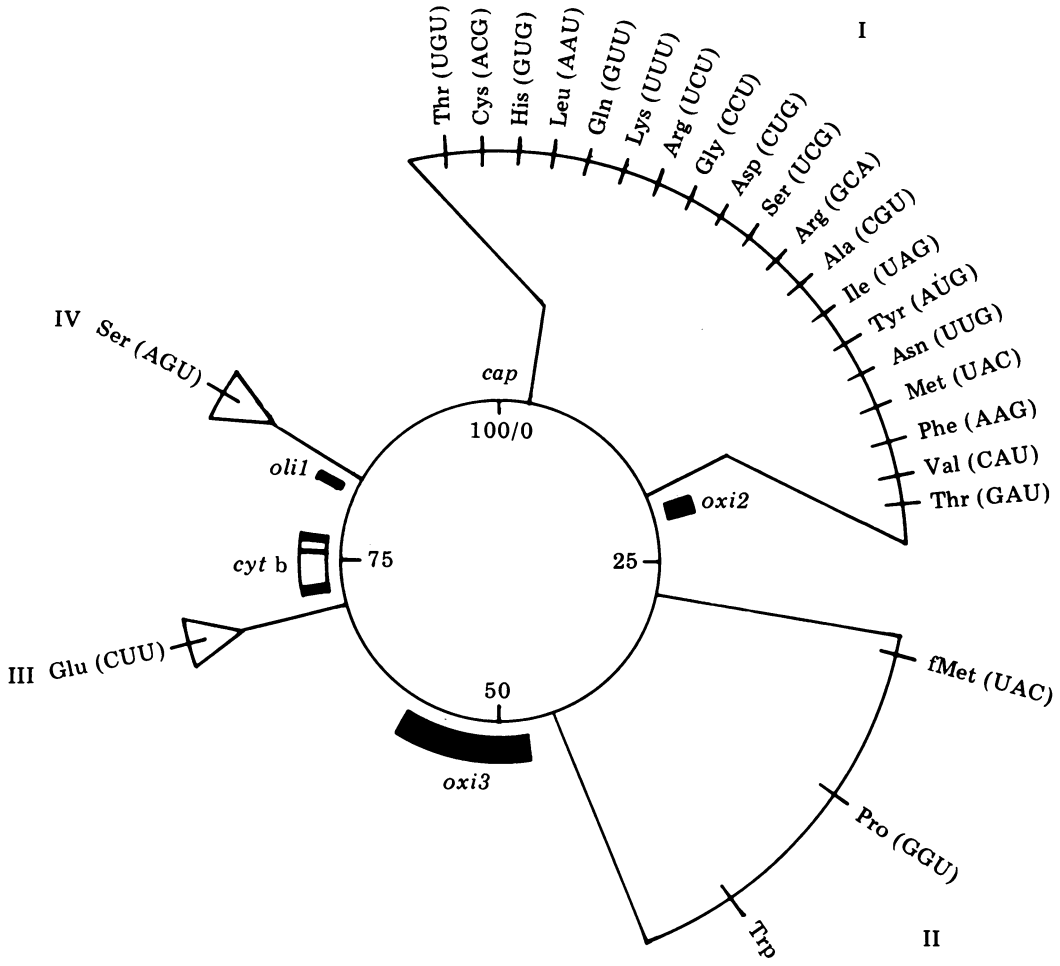

FIG. 1. tRNA genes of the mitochondrial genome of *S. cerevisiae*. The positions and relative orders of the tRNA genes are based on DNA sequences (2–4, 13) and "petite" deletion mapping (10). The anticodons (3′ → 5′) of the genes sequenced are indicated in brackets next to the cognate amino acids. *oxi2* codes for subunit 3 of cytochrome oxidase, *oxi3* for subunit 1 of cytochrome oxidase, and *oli1* for subunit 9 of the ATPase. The cytochrome *b* gene is composed of three exons (unpublished data). The *cap* locus is a marker in the 21S ribosomal RNA gene. The map units are shown on the inner circle.

is presented in Table 1. As noted previously (14–18), yeast mitochondrial genes preferentially select codons terminating in an A or U. This is evident both in families of four codons (e.g., Ala, Pro, Val) and in the mixed families (e.g., AGN series for Arg and Ser). Despite the bias for A and U in the third position, most of the codons appear to be used. The most notable exceptions are the four codons of the CGN series for Arg, none of which occur in any of the genes sequenced. Also absent from the current list are 11 other codons. These include AGG (Arg), GGC (Gly), CAC (His), CCG (Pro), UCC, UCG and AGC (Ser), ACG, CUC and CUG (Thr), and UGG (Trp). It is not excluded that further additions to the list will be made as more gene sequences are obtained.

Anticodons in yeast mitochondrial tRNAs

At present, sequences have been determined for 23 of the estimated 24 mitochondrial tRNA genes. The codons for the 20 amino acids as well as the anticodons determined from the sequences of the genes are listed in Fig. 2. The codon assignments differ from the universal code in two respects. Both human (19) and yeast (20) mitochondria have been shown to translate the UGA terminator as a tryptophan codon. Yeast mitochondria have also been found to recognize CUA as a threonine rather than a leucine codon (2). For reasons mentioned below, the four codons of the CUN family have been assigned to threonine. There is reasonably good evidence that the UGA and CUN codons are the only exceptions to the universal code in yeast mitochondria (14–18).

Several remarkable facts emerge from an examination of the matrix in Fig. 2. First is the presence of a U in the wobble position of the anticodon in all the tRNAs belonging to the unmixed codon families (UCN, CCN, GCN, etc.).[§] This suggests that a single tRNA can read all four codons in any given family, provided that the nucleotide in the wobble position is a U. It is of interest that the only exception to this rule occurs in the tRNA for the arginine CGN series whose codons have not been detected in any of the yeast mitochondrial genes analyzed.

As has been pointed out by Lagerkvist (6), in a two out of three reading method, there must be sufficiently strong base pairing between the first two nucleotides to ensure translational fidelity. This is achieved in part by having at least one G·C base pair in the codon/anticodon interaction. The special properties of yeast mitochondrial tRNAs may also facilitate hydrogen bonding between the U in the wobble position of the anticodon and the pyrimidine bases in the third position of the codons. The low G + C content of mitochondrial tRNAs (18–35%) (1–4)

[§] Because the threonine tRNA with the 3′-GAU-5′ anticodon (2) fits into the general pattern, we believe that all four codons of the CUN family belong to threonine.

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

FIG. 2. Codons and anticodons of the yeast mitochondrial genetic code. The codons (5′ → 3′) are at the left and the anticodons (3′ → 5′) are at the right in each box. The wobble nucleotides of the anticodons are underlined. The sequence of the tryptophan tRNA gene has not been determined; the predicted anticodon is marked with an asterisk.

could introduce sufficient structural flexibility in the tRNA backbone to allow formation of the shorter U·U and U·C base pairs. Grosjean *et al.* (21) have shown that stable short interactions are permitted on stereochemical grounds but are normally prevented by the ribosome. It is conceivable that mitochondrial ribosomes lack this discriminatory property.

Two different situations arise in the mixed families. With the exception of the AUN series for isoleucine and methionine, which is considered separately, all the mixed families are read by two tRNAs. The tRNAs that recognize codons terminating in a U or C have a G in the corresponding position of the anticodon. This allows for normal G·C or wobble G·U base pairing. On the other hand, those tRNAs designed to read codons with an A or G in the third position always have a U in the wobble position of the anticodon. The obvious question may be raised as to why this group of tRNAs does not read mixed family codons terminating in a U or C. Two mechanisms may be invoked to explain how misreading is prevented in the mixed families. Lagerkvist (6) has pointed out that, in most of the mixed families, the second position of the codon is a purine. In the case of phenylalanine and leucine (UUN series), in which this does not apply, the first two nucleotides of the codon form weak U·A base pairs. The interactions involving the shorter U·U and U·C base pairs in such codons may be insufficiently stable and require a three out of three reading method. Reading fidelity could also be ensured by modification of the U at the wobble position of the anticodon. This has been shown for the leucine and several other mitochondrial tRNAs of *Neurospora crassa* (22). In each case the modified U occurs in the tRNAs that recognize mixed family codons terminating with a U or C (22).

There is one mixed family whose tRNAs do not fit the general pattern described above. The nucleotide sequence of the isoleucine tRNA indicates that the wobble position of the anticodon is occupied by a G. The G is probably modified to a base that can base pair with the A, U, or C of the standard isoleucine codons. In contrast to human mitochondria which have been proposed to recognize AUA as a code word for methionine (19), in yeast this codon appears to be translated as isoleucine (14). The fourth member in the AUN family is the methionine codon AUG. The methionine tRNA has a C in the anticodon and is consequently restricted to read only the AUG codon. The sequence of the formylmethionine tRNA has in the wobble position of the anticodon a C as does the mitochondrial initiator in *N. crassa* (23).

Concluding remarks

The single most significant feature of the yeast mitochondrial code is the recognition of all the codons by use of only 24 different tRNAs. This number of tRNAs corresponds to the absolute minimum required to read all the three-letter codons, given the usual degeneracies. The simplified decoding rules described for yeast mitochondria appear to operate in other mitochondria as well. Based on the sequences of the mitochondrial tRNAs in *N. crassa*, Heckman *et al.* (22) reached conclusions similar to those reported here. Barrell *et al.* (24) have determined sequences of the tRNA genes of human and bovine mitochondria; both systems also make use of a limited number of tRNAs with anticodons similar to those of the yeast and *N. crassa* tRNAs.

There is insufficient evidence at present to permit a decision as to whether the mitochondrial genetic code is more primitive or more highly evolved than has been described for other prokaryotic and eukaryotic systems. We would like to speculate, however, that mitochondria represent an evolutionary simplification in which a minimum number of tRNAs have been conserved without compromising functional efficiency. Most translational systems make use of a complex array of tRNAs, some having important secondary roles such as the regulation

Table 1. Yeast mitochondrial codons

Amino acid	Codon	No.	Amino acid	Codon	No.
Ala	GCA	32	Lys	AAA	23
	G	1		G	1
	U	47			
	C	2	Met	AUG	45
Arg	AGA	28	Phe	UUU	34
	G	0		C	38
	CGA	0			
	G	0	Pro	CCA	24
	U	0		G	0
	C	0		U	26
				C	1
Asn	AAU	49			
	C	4	Ser	AGU	12
				C	0
Asp	GAU	31		UCA	46
	C	1		G	0
				U	23
Cys	UGU	11		C	0
	C	1			
			Thr	ACA	34
Gln	CAA	22		G	0
	G	3		U	27
				C	1
Glu	GAA	31		CUA	12
	G	2		G	0
				U	2
Gly	GGA	15		C	0
	G	3			
	U	64	Trp	UGA	26
	C	0		G	0
His	CAU	36	Tyr	UAU	56
	C	0		C	6
Ile	AUA	6	Val	GUA	54
	U	108		G	3
	C	21		U	38
				C	1
Leu	UUA	164			
	G	2			

This table lists the sum of all the codons present in the genes of cytochrome *b* (unpublished data), subunits 6 (18) and 9 (16, 17) of ATPase, and subunits 2 (14) and 3 (15) of cytochrome oxidase.

of synthesis of specific gene products (25, 26). Mitochondria synthesize eight or nine proteins that are subunits of constitutive enzymes of the organelle (27, 28). The expression of mitochondrial genes in yeast and probably in other eukaryotes is controlled by the nuclear genome (28). This rather narrow role of mitochondrial protein synthesis in the production of a select class of constitutive proteins may have been a crucial factor that led to the reduced number of tRNAs found in present-day mitochondria.

This research was supported by Grant HL 22174 from the National Institutes of Health and Grant PCM 16089 from the National Science Foundation.

1. Miller, D. L., Martin, N. C., Pham, H. D. & Donelson, J. E. (1979) *J. Biol. Chem.* **254,** 11735–11740.
2. Li, M. & Tzagoloff, A. (1979) *Cell* **18,** 47–53.
3. Berlani, R., Pentella, C., Macino, G. & Tzagoloff, A. (1980) *J. Bacteriol.* **141,** in press.
4. Bos, J. L., Osinga, K. A., Van der Horst, G. & Borst, P. (1979) *Nucleic Acids. Res.* **6,** 3255–3266.
5. Crick, F. H. C. (1966) *J. Mol. Biol.* **19,** 548–555.
6. Lagerkvist, U. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 1759–1762.
7. Casey, J., Cohen, M., Rabinowitz, M., Fukuhara, H. & Getz, G. S. (1972) *J. Mol. Biol.* **63,** 431–440.
8. Martin, N. C. & Rabinowitz, M. (1978) *Biochemistry* **16,** 4672–4677.
9. Tzagoloff, A., Foury, F. & Macino, G. (1978) in *Biochemistry and Genetics of Yeast: Pure and Applied Aspects,* eds. Bacila, M., Horecker, B. L. & Stoppani, A. O. M. (Academic, New York), pp. 477–488.
10. Wesolowski, M. & Fukuhara, H. (1979) *Mol. Gen. Genet.* **170,** 261–275.
11. Van Ommen, G. J. B., Groot, G. S. P. & Borst, P. (1977) *Mol. Gen. Genet.* **154,** 255–262.
12. Grivell, L. A., Arnberg, A. C., Boer, P. H., Borst, P. Bos, J. L., Van Bruggen, E. F. J., Groot, G. S. P., Hecht, N. B., Hensgens, L. A. M., Van Ommen, G. J. B. & Tabak, H. F. (1979) in *Extrachromosomal DNA,* ICN-UCLA Symposia on Molecular and Cellular Biology, eds. Cummings, D. J., Borst, P., Dawid, I. B., Weissman, S. M. & Fox, C. F. (Academic, New York), Vol. 15, pp. 305–324.
13. Tzagoloff, A., Macino, G., Nobrega, M. P. & Li, M. (1979) in *Extrachromosomal DNA* ICN-UCLA Symposia on Molecular and Cellular Biology, eds. Cummings, D. J., Borst, P., Dawid, I. B., Weissman, S. M. & Fox, C. F. (Academic, New York), Vol. 15, pp. 339–355.
14. Coruzzi, G. & Tzagoloff, A. (1979) *J. Biol. Chem.* **254,** 9324–9330.
15. Thalenfeld, B. E. & Tzagoloff, A. (1979) *J. Biol. Chem.,* in press.
16. Macino, G. & Tzagoloff, A. (1979) *J. Biol. Chem.* **254,** 4617–4623.
17. Hensgens, L. A. M., Grivell, L. A., Borst, P. & Bos, J. L. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 1663–1667.
18. Macino, G. & Tzagoloff, A. (1980) *Cell,* in press.
19. Barrell, B. G., Bankier, A. T. & Drouin, J. (1979) *Nature (London)* **282,** 189–194.
20. Macino, G., Coruzzi, G., Nobrega, F. G., Li, M. & Tzagoloff, A. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3784–3785.
21. Grosjean, H. J., DeHenan, S. & Crothers, D. M. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 610–614.
22. Heckman, J. E., Sarnoff, J., Alzner-Deweerd, B., Yin, S. & RajBhandary, U. L. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 3159–3163.
23. Heckman, J. E., Hecker, L. I., Schwartzbach, S. D., Barnett, W. E., Baumstark, B. & RajBhandary, U. L. (1978) *Cell* **13,** 83–95.
24. Barrell, B. G., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drovin, 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.
25. Calhoun, D. H. & Hatfield, G. W. (1975) *Annu. Rev. Microbiol.* **29,** 275–299.
26. Brenchley, J. E. & Williams, L. S. (1975) *Annu. Rev. Microbiol.* **29,** 251–274.
27. Borst, P. & Grivell, L. A. (1978) *Cell* **15,** 705–723.
28. Tzagoloff, A., Macino, G. & Sebald, W. (1979) *Annu. Rev. Biochem.* **48,** 419–441.