Codon reading patterns in *Saccharomyces cerevisiae* mitochondria based on sequences of mitochondrial tRNAs

Annie-Paule Sibler, Guy Dirheimer and Robert P. Martin

Laboratoire de Biochimie II, Institut de Biologie Moléculaire et Cellulaire du CNRS, Université Louis Pasteur, 15 Rue Descartes, 67084 Strasbourg, France

Received 7 November 1985

The sequences of *Saccharomyces cerevisiae* mitochondrial $tRNA_1^{Arg}$, $tRNA_2^{Arg}$, $tRNA^{Gly}$, $tRNA_2^{Lys}$, $tRNA^{Leu}$ and $tRNA^{Pro}$ are reported. Special structural features were found in $tRNA^{Pro}$, which has A_8, C_{21}, A_{48} instead of the constant residues U_8, A_{21} and pyrimidine 48, and in $tRNA_2^{Lys}$, which has a U excluded from base-pairing and bulging out from the TΨC stem. The $tRNA_1^{Arg}$, $tRNA_2^{Lys}$ and $tRNA^{Leu}$, which belong to two-codon families ending in a purine, have a modified uridine in the wobble position, which prevents misreading of C and U. It is likely to be 5-carboxymethylaminomethyluridine. $tRNA^{Gly}$ and $tRNA^{Pro}$ have an unmodified uridine in the wobble position allowing the reading of all four codons of a four-codon family. However, $tRNA_2^{Arg}$, which is a minor species and belongs to the CGN four-codon family, has an unmodified A in the wobble position. This very unusual feature raises the problem of the mechanism by which the codons CGA, CGG and CGC are recognized.

(Yeast) Mitochondria tRNA Nucleotide sequence Codon recognition Genetic code

1. INTRODUCTION

Mitochondrial protein synthesis uses a restricted number of mitochondrial (mt) DNA-coded tRNAs which is far below the minimal number of tRNAs, i.e. 32, necessary to translate all the codons in the genetic code according to the wobble hypothesis [1]. This deficit is not overcome by the import into the organelle of nuclear-coded tRNA species [2,3]. From the sequence determination of 6 *Neurospora crassa* mitochondrial tRNAs [4], the following codon recognition rules have been proposed: the tRNAs recognizing a 4-codon family have an unmodified U in the wobble position of their anticodon, whilst tRNAs which recognize a 2-codon family ending in a purine have a modified U. This allows discrimination in the mixed families between codons ending in a purine and codons ending in a pyrimidine. From the determination of the gene sequences of yeast mt tRNAs [5] and from our determinations of the primary structures of $tRNA_2^{Ser}$, $tRNA_1^{Thr}$ and $tRNA^{Trp}$ [6–8] it appears that the same rules are also operative in yeast mitochondria. However, exceptions may exist. This is the case for the decoding of codons AUG and AUA by the mitochondrial $tRNA_m^{Met}$, which has a CAU anticodon [9]. Another question was to know whether the modified U found in $tRNA^{Trp}$ [8] is also present in the other tRNAs recognizing 2-codon families ending in a purine, or if several types of U modification exist, as is the case for tRNAs in the cytoplasm of eucaryotes or in procaryotes [10]. Therefore, we have determined the sequence of $tRNA_2^{Arg}$, $tRNA^{Gly}$ and $tRNA^{Pro}$ that belong to 4-codon families and of $tRNA_1^{Arg}$, $tRNA_2^{Lys}$ and $tRNA^{Leu}$ that belong to 2-codon families ending in a purine.

This paper is dedicated to Professor S.P. Datta

2. MATERIALS AND METHODS

2.1. *Purification of tRNAs*

Preparation of yeast mitochondria and extrac-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

tion of total mt tRNA were as previously described [2]. The isolation of mt $tRNA^{Gly}$, $tRNA^{Leu}$ and $tRNA^{Pro}$ was achieved in a single step, using 2-dimensional polyacrylamide gel electrophoresis [2,11]. The two arginine isoaccepting species and the $tRNA_2^{Lys}$ isoacceptor were purified in two steps, an RPC5 column chromatography [12], followed by 2-dimensional polyacrylamide gel electrophoresis. Amino acid acceptor activities were determined using [^{3}H]amino acids (15–25 Ci/mmol; CEA, Saclay) and a preparation of yeast mitochondrial aminoacyl-tRNA synthetases [13].

2.2. *Sequencing techniques*

Various procedures involving in vitro ^{32}P labeling [6,14,15] were used in parallel for the sequence analyses of the 6 mt tRNAs. Most of the sequence data were obtained using the method of Stanley and Vassilenko [16], with modifications [15]. Modified residue was characterized by 2-dimensional thin-layer chromatography using the solvent systems quoted in [15]. To confirm the nucleotide assignments and to determine the 5′- and 3′-end sequences, either read-off sequencing gels or mobility shift analyses were run on ^{32}P end-labeled samples [6,14].

3. RESULTS AND DISCUSSION

3.1. *tRNA isoacceptors*

Analyses of yeast mt tRNA isoacceptors by either RPC5 column chromatography or 2-dimensional polyacrylamide gel electrophoresis [2,11,12,17] revealed the presence of only one species specific to each glycine, leucine and proline. In contrast, two isoacceptors were found for arginine, and lysine, respectively. Of the two arginine isoacceptors, the $tRNA_2^{Arg}$ is a minor species representing only 5–10% of the major $tRNA_1^{Arg}$ [17]. This difference in the levels of expression of these two isoacceptors may be related to the difference in their respective utilization in mitochondrial protein synthesis (for further discussion, see below). For the two mitochondrial $tRNA^{Lys}$ isoacceptors, one, i.e. $tRNA_1^{Lys}$, is a cytoplasmic species which is imported into mitochondria but does not seem to participate in mitochondrial protein synthesis [3], whereas the other isoacceptor, i.e. $tRNA_2^{Lys}$, is an mt DNA encoded species. Thus, in contrast to the cytoplasm, a reduced number of tRNA isoacceptors is used in mitochondrial protein synthesis.

3.2. *Structural features*

The 6 *S. cerevisiae* mt tRNAs whose sequences are reported here are shown in the cloverleaf form in figs 1 and 2.

A characteristic of mt tRNAs is their low G + C content. For the 6 tRNAs shown, the G + C content varies from 20.2% in $tRNA_2^{Arg}$ to 45.3% in $tRNA^{Pro}$, which are respectively the lowest and highest value found for the complete set of yeast mt tRNAs.

The mt $tRNA_1^{Arg}$, $tRNA_2^{Arg}$, $tRNA^{Gly}$ and $tRNA^{Leu}$ contain all the structural characteristics, including invariant and semi-invariant residues, which are constant in the standard cloverleaf [18]. This is not the case for mt $tRNA^{Pro}$ and $tRNA_2^{Lys}$, in which important exceptions are found:

In the mt $tRNA^{Pro}$ (fig.1), the 'universal' residues U_8, A_{21} and Py_{48} are replaced by the residues A_8, C_{21} and A_{48}, respectively. The replacement of residue U_8 by another nucleoside has not been found in any other mt tRNA species from yeast, and it occurs only in some other mt tRNAs of fungal and animal origins and in T_5 coded $tRNA^{Asp}$ [19]. In the tertiary structure of yeast cytoplasmic $tRNA^{Phe}$ [20] and $tRNA^{Asp}$ [21], A_{21} is involved in an interaction with U_8 by base-backbone interactions, whereas Py_{48} interacts with Pu_{15} by 'transpairing'. The base changes observed in $tRNA^{Pro}$ may affect the stability of the tertiary structure of this tRNA.

The mt $tRNA_2^{Lys}$ (fig.2) has an original structural feature: it contains an extra uridine residue (position 51 bis) which is excluded from base pairing and bulges out from the TΨC stem. Such an outloop has also been found in the TΨC stem of 3 other tRNAs. In yeast mt $tRNA^{Phe}$ [14], it is located in the same position (between residues 51 and 52), whereas in yeast mt $tRNA_m^{Met}$ [9], it is located between positions 50 and 51. In the highly unusual UGA suppressor $tRNA^{Ser}$ from animal liver [22–24], the only non-organellar tRNA known to have an outloop, this feature is also found in the TΨC stem, but on its 3′-side (between positions 63 and 64). Finally, in the

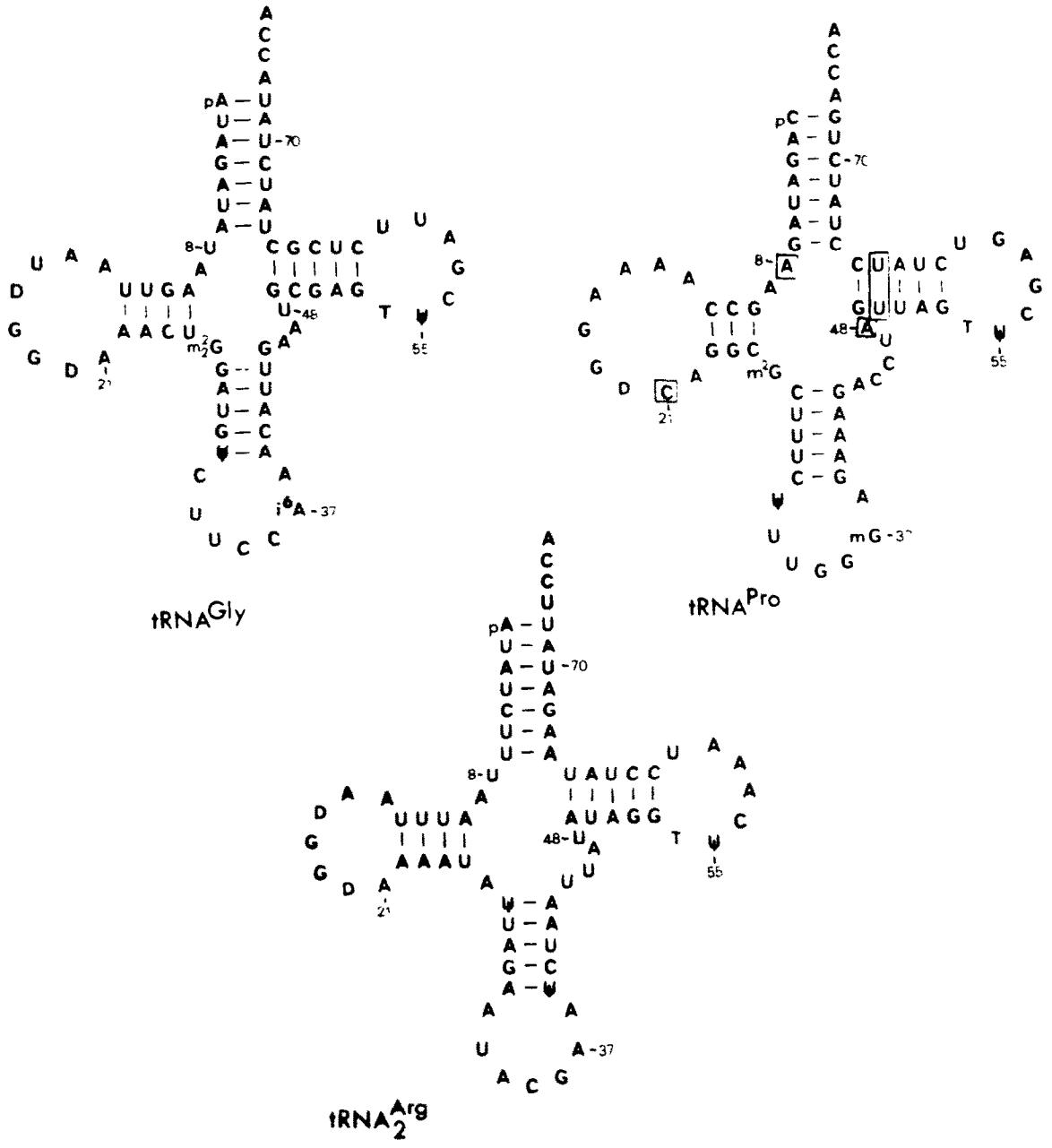

Fig.1. Nucleotide sequences of *S. cerevisiae* mitochondrial glycine, proline, and arginine-2 tRNAs in cloverleaf form. Unusual structural features are boxed.

AUA specific $tRNA^{Ile}$ from spinach chloroplasts [25], an outloop occurs on the 3′-side of the anticodon stem. The existence of such an unusual structural feature may have functional implications for the corresponding tRNAs. In animal liver $tRNA^{Ser}$, the outloop is supposed to be responsible for the suppressor function of the tRNA [22]. In chloroplastic $tRNA^{Ile}$, it has been suggested that the outloop allows enhancement of the specificity of AUA recognition [25]. Concerning the outloop in the yeast mt $tRNA_m^{Met}$, we have proposed that it enhances C·A wobble,

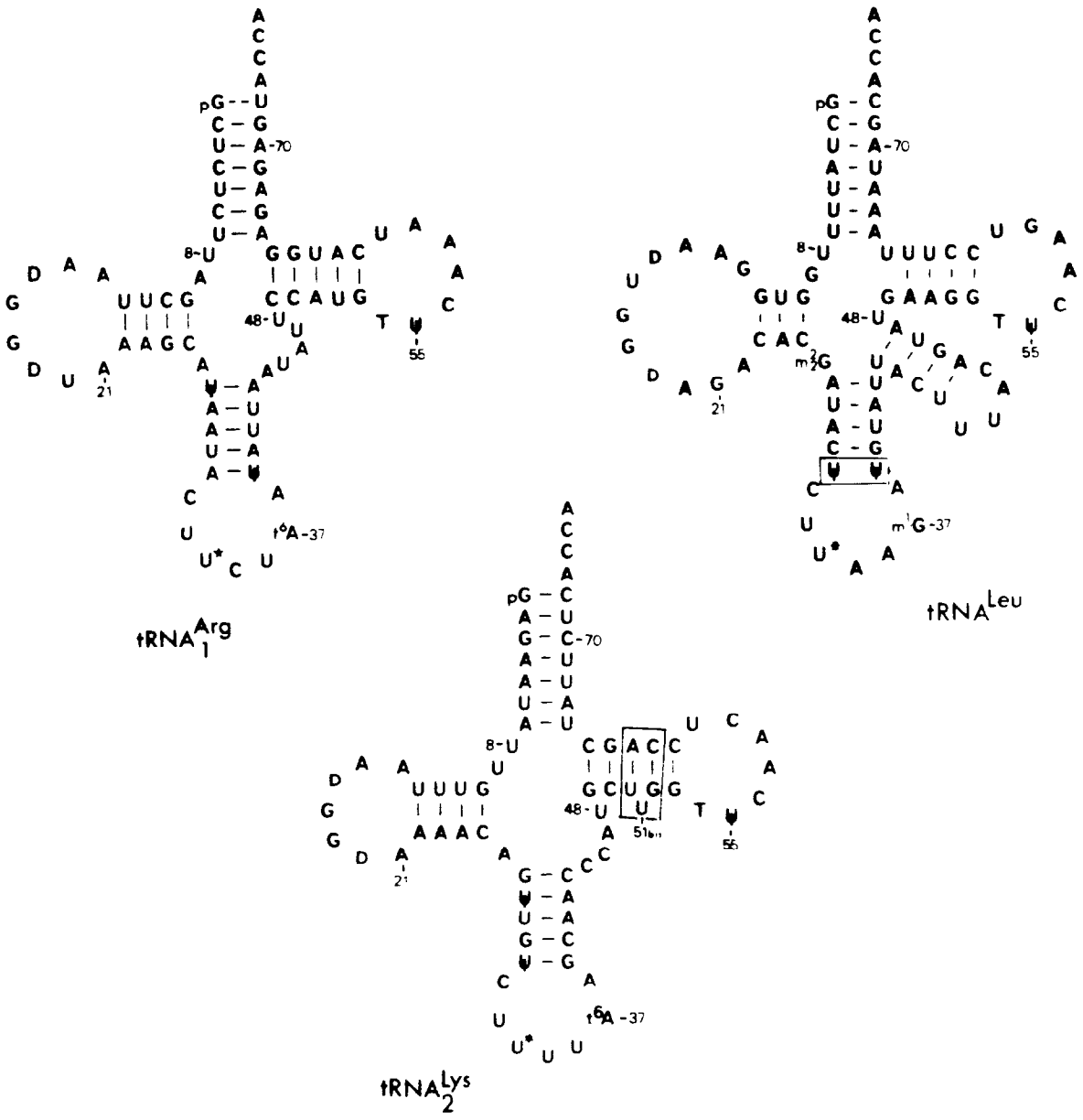

Fig.2. Sequences of *S. cerevisiae* mitochondrial arginine-1, leucine, and lysine-2 tRNAs in cloverleaf form. Boxed regions show unusual structural features.

thus permitting the decoding of both AUG and AUA [9]. It would be interesting to know whether the outloop in the mt $tRNA^{Phe}$ and $tRNA_2^{Lys}$ also confers unusual decoding properties. An argument in favor of a possible frameshift suppressor function for these two tRNAs is provided by the existence of frameshift mutations which are leaky in yeast mitochondrial protein coding genes and which affect a UUU codon for phenylalanine [26] and an AAA codon for lysine [27]. The leakiness of the mutations could be due to the ability of the $tRNA^{Phe}$ and $tRNA_2^{Lys}$ to restore, at least to a certain extent, the correct reading frame. Such a possible

frameshift suppressor function could be due to the outloop which is present in the TΨC stem of both tRNAs. The outloop may influence the positioning of the tRNA in the P site of the mitochondrial ribosome, in such a way as to induce slippage.

3.3. *Modified nucleotides*

The 6 tRNAs reported here contain a low number (from 6 to 9) of modified residues (figs 1 and 2), which is a characteristic of mitochondrial tRNAs, whatever their origin. Ribothymidine (in position 54), pseudouridine (positions 27, 28, 31, 39 and 55) and dihydrouridine (one or two residues in the D loop) are found in each sequence. Only 3 other positions in the tRNA cloverleaf are sometimes modified. These are:

(i) residue 26 which is m^2G in $tRNA^{Pro}$ and m_2^2G in $tRNA^{Gly}$ and $tRNA^{Leu}$. This is a eucaryotic feature [10].

(ii) residue 37 which is m^1G in $tRNA^{Leu}$ and $tRNA^{Pro}$, t^6A in $tRNA_1^{Arg}$ and $tRNA_2^{Lys}$, and i^6A in $tRNA^{Gly}$. In contrast, residue 37 of the minor $tRNA_2^{Arg}$ is an unmodified A.

(iii) the wobble position of the anticodon. Analysis by 2-dimensional thin-layer chromatography using different solvent systems revealed the presence in $tRNA_1^{Arg}$, $tRNA_2^{Lys}$ and $tRNA^{Leu}$ of a modified uridine showing migration properties identical with those of 5-carboxymethylaminomethyluridine ($cmnm^5U$) (fig.3), which has been identified in $tRNA^{Gly}$ of *Bacillus subtilis* [28]. The analysis of the exact nature of this modified uridine is currently in progress.

3.4. *Codon recognition patterns*

A striking result is that, among the yeast mt tRNAs sequenced, glycine and proline (fig.1) as well as serine-2 [6] and threonine-1 [7] tRNA species, whose amino acids use 4-codon families (GGN, CCN, UCN and CUN, respectively) all contain an unmodified U in the first position of the

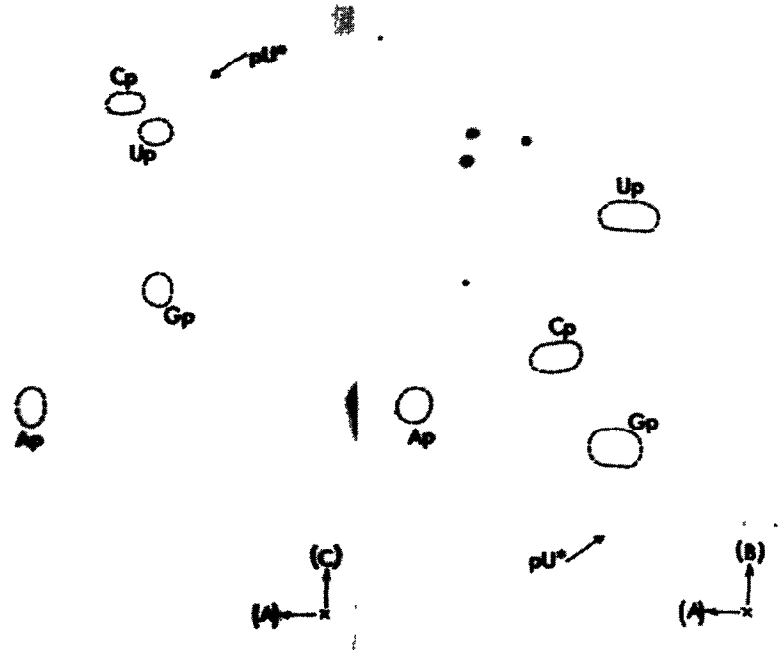

Fig.3. Analysis of the 5′-labeled modified uridine in position 34 (first position of the anticodon) of mitochondrial $tRNA^{Leu}$ by 2-dimensional thin-layer chromatography, using the following solvent systems: (A) isobutyric acid/H_2O/25% NH_4OH (66:33:1, by vol.); (B) 2-propanol/conc. HCl/H_2O (68:17.6:14.4, by vol.); (C) 0.1 M sodium phosphate, pH 6.8/ammonium sulfate/*n*-propanol (100:60:2, v/w/v). The modified uridine in the first position of the anticodon in mitochondrial $tRNA_1^{Arg}$ and $tRNA_2^{Lys}$ shows identical migration properties.

anticodon. In contrast, arginine-1, leucine, lysine-2 (fig.2) as well as tryptophan tRNA [8], which recognize 2-codon families ending in a purine (AGR, UUR, AAR and UGR, respectively) all contain a modified U in the wobble position. Moreover, this modified uridine appears to be the same in the 4 tRNAs. It is most likely to be cmnm^{5}U. Murao and Ishikura [28] have shown that in ribosome binding experiments, this modified residue recognizes both A and G and to a lesser extent, also U. However, ^{1}H-NMR analyses of different modified uridines [29] suggest that cmnm^{5}U in the first position of the anticodon should be preferentially in the C_3^1-endo conformation. This enhances the rigidity of the anticodon, so as to prohibit the misrecognition of codons terminating in a pyrimidine. Such a rigidity of cmnm^{5}U would also not favor the pairing with G [30]. This could explain the strong bias against utilization of G in the 3rd position of mixed codon families ending in a purine in yeast mitochondrial protein coding genes (see table 1). An exception for

Table 1

Codon usage in *S. cerevisiae* mitochondrial genes

Amino acid	Codon	Genes[a]	ORF[b]
Ala	GCA	53	33
	U	66	65
	C	6	11
	G	4	9
Arg	AGA	42	126
	G	0	7
	CGA	0	1
	U	1	14
	C	0	2
	G	1	1
Asn	AAU	187	423
	C	11	28
Asp	GAU	51	155
	C	2	10
Cys	UGU	13	56
	C	1	1
Gln	CAA	36	75
	G	4	7
Glu	GAA	39	121
	G	3	17
Gly	GGA	28	57
	U	97	129
	C	1	6
	G	7	16
His	CAU	51	75
	C	4	3
Ile	AUU	192	319
	C	29	23
Leu	UUA	277	386
	G	2	15
Lys	AAA	67	40
	G	3	3
Met	AUA	38	160
	G	71	51
Phe	UUU	80	161
	C	63	21

Table 1

(continued)

Amino acid	Codon	Genes[a]	ORF[b]
Pro	CCA	37	27
	U	44	73
	C	4	6
	G	1	9
Ser	UCA	75	77
	U	40	80
	C	1	11
	G	0	3
	AGU	24	76
	C	0	3
Thr	ACA	53	78
	U	41	72
	C	1	12
	G	0	17
	CUA	17	30
	U	5	33
	C	0	2
	G	2	0
Trp	UGA	39	54
	G	0	6
Tyr	UAU	104	244
	C	13	18
Val	GUA	78	99
	U	48	70
	C	4	4
	G	6	11

[a] Data for genes are for cytochrome oxidase subunits I, II and III, apocytochrome *b*, ATPase subunits 6, 8 and 9, and ribosomal protein Var 1 (quoted in [35,36])

[b] Data for open reading frames are from 7 intron-encoded reading frames in the 21 S rRNA, oxi3 and cob, and 3 free-standing unassigned reading frames, near oxi1, oxi2 and 15 S RNA genes (quoted in [35,37,38])

the decoding of 2-codon families ending in a purine by tRNAs having a modified U in the wobble position is provided by the mt $tRNA_m^{Met}$, which despite its CAU anticodon, is able to read not only AUG but also AUA [9].

The presence of an unmodified uridine in the wobble position of tRNAs which read 4-codon families is unusual and contrasts with the situation in both prokaryotic and eukaryotic cytoplasmic tRNAs, in which a uridine in the first position of the anticodon is always modified [19]. The only exception to this is a yeast cytoplasmic $tRNA^{Leu}$, which has been shown to read all 4 codons of the CUN leucine family [31]. Crick, in his wobble hypothesis [1], noted that U·U and U·C base pairings were possible, but discounted them on the basis that they were too close and would cause misreading in the mixed codon families. Grosjean et al. [32] have examined the stability of a number of anticodon-anticodon interactions and concluded that the only stable non-wobble pairs involve U pairing with U and C. Thus, an unmodified uridine is the only base that can possibly form a more or less stable pair with all 4 nucleotides in the wobble position of the codon. The only exception to this rule is provided by the yeast mt $tRNA_2^{Arg}$ which belongs to the 4-codon family CGN, but has an unmodified A in the first position of the anticodon. The presence of an unmodified A in this position has never been found in any sequenced tRNA. In fact, A is always modified to inosine [19], except in *E. coli* $tRNA_{NAA}^{Leu}$ where its modification is of unknown structure [33]. In the wobble hypothesis [1], inosine pairs with A, U and C, whereas an unmodified A pairs with U. This means that mt $tRNA_2^{Arg}$ recognizes only the CGU codon. An examination of the frequency of utilization of CGN codons in yeast mitochondrial protein coding genes (table 1) reveals that none of the genes which code for a structural protein of the respiratory chain uses CGN codons, but exclusively AGR codons for arginine. However, in the Var 1 ribosomal protein gene, both CGU and CGG are used (one each). In intronic open reading frames coding for 'maturases' or in free-standing unassigned reading frames the CGN codons are more represented (18 utilizations), but their frequency remains low when compared to the utilization of AGR. Nevertheless, the utilization of some CGA, CGG and CGC codons raises the question of the mechanism of their recognition by mt $tRNA_2^{Arg}$. In this case, a 'two out of three' mechanism [34] remains the most plausible one, especially since the ACG/CGN interaction contains 2 pairs C:G able to secure stability. In addition, yeast mitochondria do not contain any other tRNA able to compete with $tRNA_2^{Arg}$ for the decoding of CGN,

It should also be emphasized that the low frequency of utilization of CGN codons compared to AGR in mitochondrial genes (table 1) parallels the fact that the $tRNA_2^{Arg}$ is a minor species when compared to $tRNA_1^{Arg}$. Such a situation suggests the existence of a mechanism which allows adaptation of the levels of expression of the two isoacceptors to their respective utilization in mitochondrial protein synthesis.

In conclusion, the special rules which govern codon-anticodon recognition in yeast mitochondria explain why 24 tRNAs are sufficient to translate all the codons of the genetic code.

ACKNOWLEDGEMENTS

We thank Dr Y. Boulanger for careful reading of the manuscript. This work was supported by grants from the CNRS (ATP 2080 and 4256).

REFERENCES

[1] Crick, F.H.C. (1966) J. Mol. Biol. 19, 548–555.
[2] Martin, R.P., Schneller, J.P., Stahl, A.J.C. and Dirheimer, G. (1977) Nucleic Acids Res. 4, 3497–3510.
[3] Martin, R.P., Schneller, J.P., Stahl, A.J.C. and Dirheimer, G. (1979) Biochemistry 18, 4600–4605.
[4] Heckman, J.E., Sarnoff, J., Alzner-Deweerd, B., Yin, S. and RajBhandary, U.L. (1980) Proc. Natl. Acad. Sci. USA 77, 3159–3163.
[5] Bonitz, S.G., Berlani, R., Coruzzi, G., Li, M., Macino, G., Nobrega, F.G., Nobrega, M.P., Thalenfeld, B.E. and Tzagoloff, A. (1980) Proc. Natl. Acad. Sci. USA 77, 3167–3170.
[6] Martin, R.P., Sibler, A.P. and Dirheimer, G. (1982) Biochimie 64, 1073–1079.
[7] Sibler, A.P., Dirheimer, G. and Martin, R.P. (1981) FEBS Lett. 132, 344–348.
[8] Sibler, A.P., Bordonné, R., Dirheimer, G. and Martin, R.P. (1980) CR Acad. Sci. Paris 290, 695–698.
[9] Sibler, A.P., Dirheimer, G. and Martin, R.P. (1985) Nucleic Acids Res. 13, 1341–1345.

[10] Dirheimer, G. (1983) Rec. Res. Cancer Res. 84, 193–201.
[11] Martin, R.P. and Dirheimer, G. (1983) Mol. Biol. (Moscow) 17, 1117–1125.
[12] Martin, R.P., Schneller, J.M., Stahl, A.J.C. and Dirheimer, G. (1976) in: Genetics and Biogenesis of Chloroplasts and Mitochondria (Bücher, ed.) pp.755–758, Elsevier, Amsterdam, New York.
[13] Schneller, J.M., Schneller, C., Martin, R.P. and Stahl, A.J.C. (1976) Nucleic Acids Res. 3, 1151–1165.
[14] Martin, R.P., Sibler, A.P., Schneller, J.M., Keith, G., Stahl, A.J.C. and Dirheimer, G. (1978) Nucleic Acids Res. 5, 4579–4592.
[15] Sibler, A.P., Martin, R.P. and Dirheimer, G. (1979) FEBS Lett. 107, 182–186.
[16] Stanley, J. and Vassilenko, S. (1978) Nature 274, 87–89.
[17] Martin, R.P., Sibler, A.P., Bordonné, R. and Dirheimer, G. (1983) Mol. Biol. (Moscow) 17, 1126–1146.
[18] Dirheimer, G., Keith, G., Sibler, A.P. and Martin, R.P. (1979) in: Transfer RNA: Structure, Properties and Recognition (Schimmel, P.R. et al. eds) pp.19–41, Cold Spring Harbor Laboratory, NY.
[19] Sprinzl, M., Moll, J., Meissner, F. and Hartmann, T. (1985) Nucleic Acids Res. 13, Suppl., r1–r49.
[20] Kim, S.H., Suddath, F.L., Quigley, G.L., McPherson, A., Sussman, J.L., Wang, A.H.J., Seeman, N.C. and Rich, A. (1974) Science 185, 435–440.
[21] Moras, D., Comarmond, M.B., Fischer, J., Weiss, R., Thierry, J.C., Ebel, J.P. and Giégé, R. (1980) Nature 288, 669–674.
[22] Diamond, A., Dudock, B.S. and Hatfield, D. (1981) Cell 25, 497–506.
[23] Hatfield, D., Diamond, A. and Dudock, B.S. (1982) Proc. Natl. Acad. Sci. USA 79, 6215–6219.
[24] Hatfield, D.L., Dudock, B.S. and Eden, F.C. (1983) Proc. Natl. Acad. Sci. USA 80, 4940–4944.
[25] Francis, M.A. and Dudock, B.S. (1982) J. Biol. Chem. 257, 11195–11198.
[26] Fox, T.D. and Weiss-Brummer, B. (1980) Nature 288, 60–63.
[27] Lazowska, J., Jacq, C. and Slonimski, P.P. (1980) Cell 22, 333–348.
[28] Murao, K. and Ishikura, H. (1978) Nucleic Acids Res. Special Publication 5, s333–s336.
[29] Yokoyama, S., Watanabe, T., Murao, K., Ishikura, H., Yamaizumi, Z., Nishimura, S. and Miyazawa, T. (1985) Proc. Natl. Acad. Sci. USA 82, 4905–4909.
[30] Hillen, W., Egert, E., Lindner, H.J. and Gassen, H.G. (1978) FEBS Lett. 94, 361–364.
[31] Weissenbach, J., Dirheimer, G., Falcoff, R., Sanceau, J. and Falcoff, E. (1977) FEBS Lett. 82, 71–76.
[32] Grosjean, H.J., De Henau, S. and Crothers, D.M. (1978) Proc. Natl. Acad. Sci. USA 75, 610–614.
[33] Yamaizumi, Z., Kuchino, Y., Harada, F., Nishimura, S. and McCloseky, J.A. (1980) J. Biol. Chem. 255, 2220–2225.
[34] Lagerkvist, U. (1978) Proc. Natl. Acad. Sci. USA 75, 1759–1762.
[35] Hudspeth, M.E.A., Ainley, W.M., Shumard, D.S., Butow, R.A. and Grossman, L.I. (1982) Cell 30, 617–626.
[36] Macready, I.G., Novitski, C.E., Maxwell, R.J., John, U., Ooi, B.G., McMullen, G.L., Lukins, H.B., Linnane, A.W. and Nagley, P. (1983) Nucleic Acids Res. 11, 4435–4451.
[37] Martin, R.P., Bordonné, R. and Dirheimer, G. (1982) in: Cell Function and Differentiation, part B, FEBS volume 65 (Akoyunoglou, G. et al. eds) pp.355–365, A.R. Liss, New York.
[38] Michel, F. (1984) Curr. Genet. 8, 307–317.