Use of the UGA terminator as a tryptophan codon in yeast mitochondria

(mitochondrial code/mitochondrial genes/yeast mitochondrial DNA/UGA codon/DNA sequence)

GIUSEPPE MACINO, GLORIA CORUZZI, FRANCISCO G. NOBREGA, MAY LI, AND ALEXANDER TZAGOLOFF

Department of Biological Sciences, Columbia University, New York, New York 10027

Communicated by Cyrus Levinthal, May 21, 1979

**ABSTRACT**

We propose that the UGA terminator regularly occurs as a tryptophan codon in yeast mitochondrial DNA. This conclusion is based on the sequence analysis of mitochondrial DNA regions coding for structural genes of cytochrome b, cytochrome oxidase, and the ATPase.

In the course of studying the structural gene of cytochrome b in the mitochondrial DNA of *Saccharomyces cerevisiae*, we were led to suspect that theopal terminator UGA is recognized as an amino acid codon. This suspicion arose from the observation that regions of DNA with genetic markers in cytochrome b contain multiple UGA codons in the only reading frame that could be assigned a coding function. This divergence from the universal code has been confirmed by the DNA sequences of the structural genes of a subunit of cytochrome oxidase and of the ATPase. Evidence documenting the conclusion that yeast mitochondria in fact utilize UGA as a codon for tryptophan is summarized below.

The sequence of a segment of mitochondrial DNA containing the structural gene of subunit 2 of cytochrome oxidase has recently been determined in our laboratory. An analysis of the nucleotide sequence reveals that only one of the six possible reading frames (taking into account both DNA strands) is capable of generating an amino acid sequence that is compatible with the known amino acid composition and molecular weight of the protein. The nucleotide sequence in this frame dictates a continuous protein sequence provided that UGA is assumed to code for an amino acid. The sequence of the gene begins with an AUG initiator, ends with an ochre terminator, and is flanked by sequences that are rich in A+T at the amino- and carboxyl-terminal ends of the gene. The other five frames have been excluded due to the frequent occurrence of ochre and amber, in addition to opal, terminators (Table 1). The correctness of the amino acid sequence cannot be rigorously proven at present because the primary structure of the protein is not known. It is noteworthy, however, that the deduced sequence of the yeast protein exhibits extensive homology with the primary structure of subunit 2 of bovine heart cytochrome oxidase (1, 2). As seen in Fig. 1, the yeast sequence has two long stretches of amino acids (residues 75–137 and 143–251) that have "in frame" homologies of 49% with the bovine sequence. These two amino acid stretches are especially revealing because they include four tryptophan residues that are conserved and are encoded by UGA in yeast mitochondrial DNA. The nucleotide

![FIG. 1. Amino acid homology of subunit 2 of yeast and bovine cytochrome oxidase. The positions of identical amino acids in the two sequences are marked by the dots. Shifts in the diagonal line indicate insertions of amino acids either in the bovine (upward) or yeast (downward) sequences. The four conserved tryptophans are indicated by the arrows pointing up. The position of the nonconserved tryptophan is indicated by the arrow pointing down. The bar denotes the sequence shown in Fig. 2. The yeast and bovine proteins contain 251 and 227 residues, respectively. The first 16 amino acids of the yeast subunit 2 are missing in the bovine protein.](image)

<table>
<thead>
<tr>
<th>Frame</th>
<th>Cytochrome oxidase subunit 2 gene</th>
<th>ATPase αβ2 gene</th>
<th>Cytochrome b gene</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UAA</td>
<td>UAG</td>
<td>UGA</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>17</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

Frames 1–3 and 4–6 occur in the two complementary strands. The sequence of the cytochrome b gene has not been completely determined. The region analyzed represents a continuous stretch of 750 base pairs starting from the amino-terminal end of the gene (unpublished results).
sequence coding for the amino acid sequence spanning residues 121-142 where two of the conserved tryptophans are found is shown in Fig. 2. The entire protein contains five tryptophans which is equivalent to 2 mol % and is in good agreement with the experimentally determined value of 1.5 mol % (3). The fact that all five tryptophans are encoded by UGA provides additional support for the notion that this is the preferred tryptophan codon in yeast mitochondria.

The sequences of two other yeast mitochondrial genes have been determined. The sequence of the ATPase proteolipid has no UGAs (4, 5) and is therefore not useful for evaluating the function of this codon. The absence of UGA in the proteolipid gene sequence, however, is consistent with primary structure of the protein, which is known to lack tryptophan (6). We have also determined the sequence of the gene of another ATPase subunit containing the genetic markers, ol2 (7) and pho1 (8). The identification of the structural gene sequence again depends on the assumption that UGA does not function as a terminator. The nucleotide sequence of the DNA region with the ol2 and pho1 markers has only one reading frame that is not interrupted by frequent ochre and amber terminators (Table 1). The gene starts with an AUG initiator, has a coding frame for 259 amino acids, and ends with an ochre terminator. The sequence of the ATPase gene is similar to the cytochrome oxidase gene in that each of the five tryptophans are specified by UGA. Although the gene product has not yet been identified, based on the molecular weight and amino acid composition obtained from the sequence, it is probably subunit 6 of the ATPase complex (9, 10).

Recent studies on the nucleotide sequence of the ATPase proteolipid gene indicate that yeast mitochondria utilize almost exclusively codons that have an A or U in the third position of the letter code (4, 5). The postulated role of UGA as a tryptophan codon is consistent with this principle and points to an evolutionary event that led to the adoption of an opal tRNA suppressor as a permanent feature of the mitochondrial translational machinery in S. cerevisiae. Although we cannot explain why U- and A-rich codons are favored in yeast mitochondria, several interesting points emerge from an examination of the universal code. The substitution of an A for a G in the third position does not affect the amino acid specificity of 14 out of 16 codons and in the case of UAG converts the amber to an ochre terminator. Both the ochre and the amino acid codons terminating in an A are in fact preferred in yeast mitochondria. The remaining two codons, AUG and UGG, are special because their specificity is determined by the third letter. The third letter substitution converts the AUG codon to an isoleucine codon. AUG has been retained in yeast mitochondria and this was probably influenced by the fact that it is the only codon for methionine. In the case of UGG, however, the code adapted by dispensing with the UGA terminator and using a suppressor tRNA for tryptophan. This interpretation implies that the present mitochondrial code has resulted from an evolutionary simplification. It is equally plausible, however, that mitochondrial codons have undergone little change and are therefore representative of a highly primitive code.

This research was supported by National Science Foundation Grant PCM75-16089 and National Institutes of Health Grant GM25250.