Overlapping reading frames in *Oenothera* mitochondria

Rudolf Hiesel and Axel Brennicke

Lehrstuhl für Spezielle Botanik der Universität Tübingen, Auf der Morgenstelle 1, D-74 Tübingen, FRG

Received 4 October 1985

An open reading frame (ORF) preceding the cytochrome oxidase subunit II (CO II) gene in *Oenothera* mitochondria has four nucleotides in common with this gene. The last two nucleotides of the CO II initiation codon ATG are the first two nucleotides of the TGA termination codon in the upstream ORF. Both reading frames are cotranscribed in a bicistronic mRNA species of 1250 nucleotides in length in *Oenothera*. The open reading frame codes for a protein of 58 amino acids with structural homology to the ATPase subunit 8 genes in fungal and mammalian mitochondria. Using coding space optimally though overlapping genes appears to be without economical reason considering the large size of higher plant mitochondrial genomes.

(Oenothera) *Mitochondria* *Overlapping reading frame* *Cytochrome oxidase subunit II* *ATPase subunit 8* *Promotor sequence*

1. INTRODUCTION

Gene order is generally well conserved in the mitochondrial genome of mammalian species [1,2]. The organisation of mitochondrial DNA (mtDNA) in insects or fungi can be related through a few rearrangements of DNA segments [3,4]. The position of unidentified reading frames (URFs) relative to known genes has been useful in the identification of new coding regions, for example, the analogy between aapl and URFA6L reading frames coding for the ATPase subunit 8 [5,6].

Very little information is available about the relative orientation of genes in higher plant mitochondrial DNA, where only the ribosomal genes have been mapped in detail in wheat [7] and maize [8]. The coding regions for the ribosomal genes and cytochrome oxidase subunit II (CO II) have been associated with sequences involved in recombination processes [7,9]. Rearrangements of gene order in higher plants are to be expected from the frequent recombination events in comparison to the organisation of mammalian and fungal mtDNAs. The present data show an overlapping arrangement of 2 open reading frames in the mitochondrial genome of *Oenothera*. The cytochrome oxidase subunit II gene has been identified previously in *Oenothera*, the other open reading frame could encode a polypeptide of 58 amino acids in length.

2. MATERIALS AND METHODS

Mitochondrial nucleic acids from *Oenothera* tissue cultures were isolated as described [10]. Restriction and gel techniques followed standard procedures. Nuclease digestions for S_1 protection [11] experiments were sized along sequencing ladders obtained through controlled chemical modification [12].

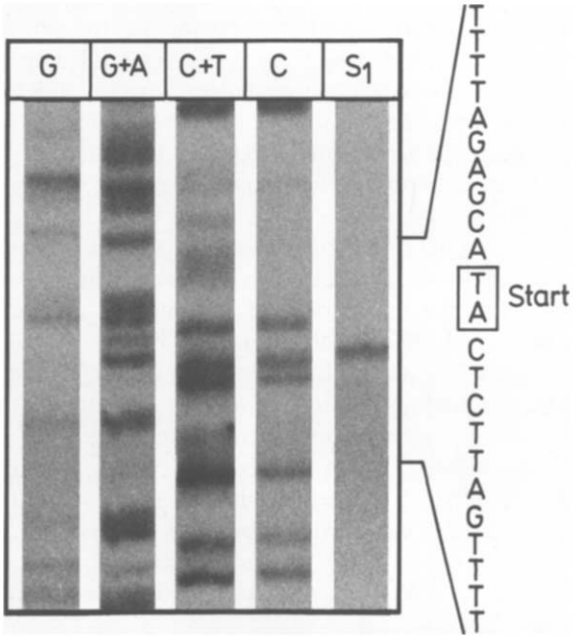
3. RESULTS AND DISCUSSION

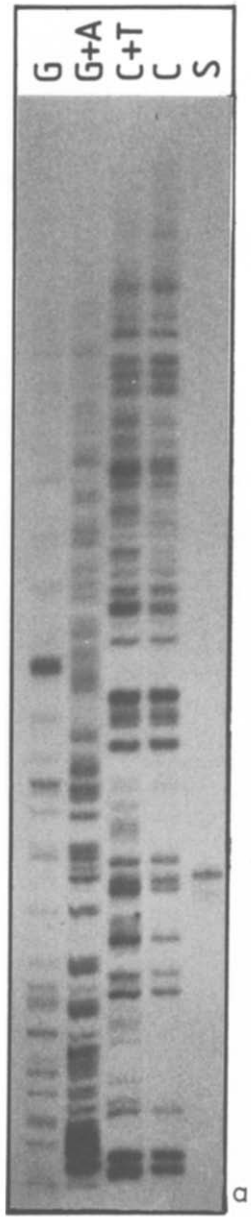
The CO II gene in *Oenothera* mitochondria is transcribed in a single mRNA of 1250 nucleotides [10]. The homogenous 5'-end of this transcript has been mapped at the nucleotide level to the T/A at positions $-210/-211$ upstream of the CO II gene (fig.1). An open reading frame (ORF) starts with an ATG initiation codon 173 base pairs (bp) upstream from the CO II gene and extends for 4 nucleotides into the CO II reading frame with a

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

shifted triplet assignment. The TGA termination codon of the preceding ORF has the TG dinucleotide in common with the initiation ATG of the CO II gene. In contrast to all other mitochondrial translation systems analysed so far only higher plant mitochondria use the triplet TGA as a termination codon [13,14]. No CGG tryptophan codon is found within this open reading frame; the single Trp residue is encoded by the classic TGG.

Both ORFs (CO II and the upstream ORF) are preceded in *Oenothera* by consensus sequences proposed to be involved in ribosome binding and initiation of translation [15,16]. In the putative ribosome binding recognition sites 5 out of 8 nucleotides are possibly involved in binding to the 3′-terminus of the 18 S rRNA as indicated by





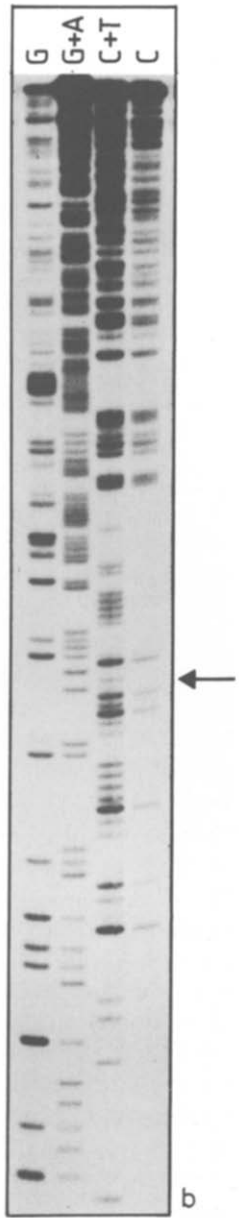


Fig.1. S_1 protection maps the 5′-end of the single 1250 nucleotides to the T/A at positions −210/−211 upstream of the CO II initiation codon. (a) S_1-protected DNA sized along a sequencing ladder and (b) a sequencing gel of this region, where the position of the S_1-protected fragment is indicated by the arrow.

asterisks in fig.2. The sequence block 5′-AAAAGAA(A)-3′, found repeatedly in plant mitochondrial protein coding presequences between the proposed ribosome binding site and the initiation codon, is present in front of both ORFs in *Oenothera* with comparable spacing indicating functional initiation sites for translation of both ORFs. The CO II signal sequences are entirely within the preceding ORF. A long mRNA leader 5′- to the protein coding regions does not appear to be generally essential for accurate translation, as there are only 22 nucleotides for the preceding

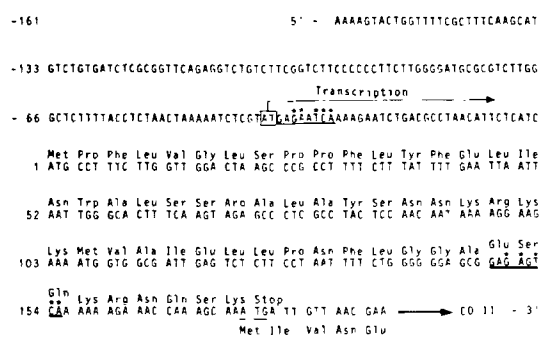


Fig.2. Nucleotide sequence of the region upstream of the cytochrome oxidase subunit II gene in *Oenothera* mitochondria. The 5'-end of the transcript is indicated by the arrow pointing in the direction of transcription. The putative ribosome binding sites are underlined and the 5 out of 8 possibly pairing nucleotides are indicated by asterisks. Amino acids for the upstream ORF are given above the nucleotide sequence, the translated cytochrome oxidase subunit II ORF underneath.

ORF and 17 for the CO II gene between the putative ribosome binding site (and the following AAAGAAA theme) and the initiation codon. The proposed ribosome recognition sequence of the preceding ORF starts immediately at the 5'-end of the single bicistronic transcript.

The unprocessed transcript covering both ORFs is probably transcribed alternatively into each of the polypeptides encoded, as the existence of the 2 putative ribosome binding sites indicates.

The sequence around the 5'-terminus of the mRNA shows strong homology to the sequences around the 5'-termini of the major transcript from the cytochrome oxidase subunit I gene in maize mitochondria [17] and the 2 transcripts from the CO II locus observed in pea mitochondria [18]. Alignment of homologies is shown in fig.3 with a possible consensus sequence of the available data.

The significance of these sequence homologies with respect to their identification as transcription initiation signals and/or processing sites can only be decided through experiments more specific to detect promotor sequences like in vitro capping or in vitro transcription assays. The primary transcript might be far longer and remain undetected in these experiments, if present at very low levels only.

The amino acid sequence deduced from the DNA level shows structural homology to the genes

Oe CO II		5'- AAATCTCGT'ATG -3'
Pea CO II	Site I	5'- AAATCACGT'AAG -3'
	Site II	5'- AAATTTACT'AAG -3'
Zm CO I		5'- AAACTA--T AAG -3'
Consensus		5'- AAATYNNNT'AAG -3'

Fig.3. A consensus sequence derived from the nucleotide identities around the 5'-termini of the *Oenothera* (Oe), the pea cytochrome oxidase subunit II and the maize (Zm) cytochrome oxidase subunit I transcripts could possibly function as a recognition signal for the mRNA beginning, i.e. promotor, if these mRNAs are indeed initiated at this position.

for the ATPase subunit 8 in fungal and mammalian mitochondria (fig.4). The overall length of the amino acid chain (58 residues) in *Oenothera* corresponds to a polypeptide of intermediate size in comparison with this polypeptide from other organisms. The human [6] (68 amino acids), mouse [18] (67 amino acids) and bovine [2] (66 amino acids) 'URFA6L' are longer, the yeast [5] 'aap 1' (48 amino acids) and the *Aspergillus* [19] 'URFx' (48 amino acids) are shorter than the *Oenothera* sequence. The *Drosophila* URFA6L [21] encodes an ATPase subunit 8 of intermediate length (53 amino acids).

Amino acid sequence homology for this subunit is much lower than for the other polypeptides identified in the mitochondrial genomes of different species, 61.8% for URFA6L and 79.2% for CO I between human and bovine mitochondria [2]. Only the terminal amino acids and an 8–10 amino acid covering sequence domain are conserved at varying sites in the polypeptides [5]. The average hydropathy of the presumptive ATPase subunit 8 polypeptides varies widely between species: −0.22 in *Oenothera*, 1.27 in *Saccharomyces*, 0.08 in bovine, −0.36 in human and 0.26 in *Drosophila* mitochondria [22].

An intimate proximity of the CO II and the ATPase subunit 8 genes in *Oenothera* mitochondria would be strikingly different from the organisation in mammalian and insect mitochondria, where the ATPase subunit 8 gene (URFA6L) is closely associated with the ATPase subunit 6 coding sequence [1,2,6,21,23]. The subunit 8 gene (aap 1) is located in yeast between the genes for cytochrome oxidase subunit I (oxi 3) and ATPase subunit 6 [5], in *Aspergillus* between the genes for ATPase subunit 6 and URF 4 [24].

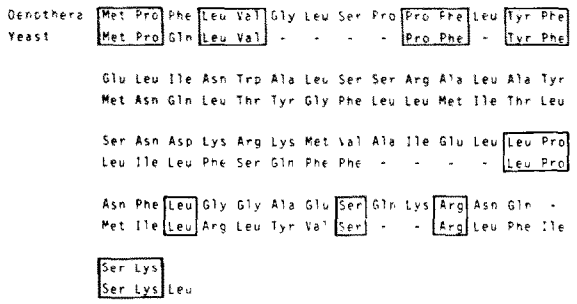


Fig.4. Alignment of the *Oenothera* ORF upstream of the cytochrome oxidase subunit II gene with the yeast ATPase subunit 8 polypeptide. Homologous amino acids between the *Oenothera* and the yeast sequence are boxed, showing the amino acids conserved at both termini. Amino acid comparison of the 2 polypeptide sequences is 26% for the alignment shown.

Economizing coding space through overlapping genes appears unnecessary in view of the 10–150-fold higher complexity of plant mitochondrial genomes [25–27] in comparison to the mammalian mtDNAs with the observation that a comparable number of polypeptides are made in isolated mitochondria [28]. The organisation of these 2 reading frames in *Oenothera* mitochondria however might represent a byproduct of extensive sequence rearrangements leading to an accidental proximity of the 2 genes in this species and to separate coding regions and transcription units in other plant mitochondria. The open reading frame upstream of the CO II gene in *Oenothera* is not encoded in a similar position in the other plants that have been analysed so far, as no homology is found in the upstream region from the CO II genes of maize [13], pea [18], wheat [29] and rice [30] to the *Oenothera* sequence.

A similar species-specific clustering has been observed for the 18 S rRNA and $tRNA^{fMet}$ genes, which are separated by just one base pair in wheat mtDNA [31,32], but far apart in the *Oenothera* mitochondrial genome [33,34]. Analysis of other gene arrangements in the mitochondrial genomes will show if rearrangement and accidental clustering of coding sequences is a common phenomenon in higher plant mitochondrial mtDNAs and might help to identify functions for the as yet undefined usefulness of the large sequence complexity of these genomes.

ACKNOWLEDGEMENTS

Homology between the described ORF and the ATPase subunit 8 gene in yeast was first pointed out to us by Dr P. Nagley at a meeting in Ladenburg, FRG. We are indebted to Drs P. Isaac and C. Leaver for computer programs and to Dr P. Blanz for his support. We thank U. Christner for assistance with the tissue cultures. This work was supported by grants from the DFG, a Graduiertenstipendium and a Heisenberg Research Career Development award.

REFERENCES

[1] Bibb, M.J., Van Etten, R.A., Wright, C.W., Walberg, M.W. and Clayton, D.A. (1981) Cell 26, 167–180.
[2] Anderson, S., Bankier, A.T., Barrell, B.G., De Brujin, A.R., 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. and Young, I.G. (1982) in: Mitochondrial Genes (Slonimski, P. et al. eds) pp.5–43, Cold Spring Harbor Laboratory Press, NY.
[3] Clary, D.O., Goddard, J.M., Martin, S.C., Fauron, C. and Wolstenholme, D.R. (1982) Nucleic Acids Res. 10, 6619–6637.
[4] Dujon, B. (1983) in: Mitochondria 1983 (Schweyen, R.J. et al. eds) pp.1–24, De Gruyter, Berlin.
[5] Novitski, C.E., Macreadie, I.G., Maxwell, R.J., Lukins, H.B., Linnane, A.W. and Nagley, P. (1984) Curr. Genet. 8, 135–146.
[6] Mariottini, P., Chomyn, A., Attardi, G., Trovato, D., Strong, D.D. and Doolittle, R.F. (1983) Cell 32, 1269–1277.
[7] Falconet, D., Lejeune, B., Quetier, F. and Gray, M.W. (1984) EMBO J. 3, 297–302.
[8] Stern, D.B. and Palmer, J.D. (1984) Nucleic Acids Res. 12, 6141–6157.
[9] Stern, D.B., Dyer, T.A. and Lonsdale, D.M. (1982) Nucleic Acids Res. 10, 3333–3340.
[10] Hiesel, R. and Brennicke, A. (1983) EMBO J. 2, 2173–2178.
[11] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbor Press, NY.
[12] Maxam, A. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[13] Fox, T.D. and Leaver, C.J. (1981) Cell 26, 315–323.

[14] Schuster, W. and Brennicke, A. (1985) Curr. Genet. 9, 157–163.
[15] Leaver, C.J., Dawson, A.J., Isaac, P.G. and Jones, V.P. (1983) in: Mitochondria 1983 (Schweyen, R.J. et al. eds) pp.269–283, De Gruyter, Berlin.
[16] Dawson, A.J., Jones, V.P. and Leaver, C.J. (1984) EMBO J. 3, 2107–2113.
[17] Isaac, P.G., Jones, V.P. and Leaver, C.J. (1985) EMBO J. 4, 1617–1623.
[18] Moon, E., Kao, T. and Wu, R. (1985) Nucleic Acids Res. 13, 3195–3212.
[19] Michael, N.L., Rothbard, J.B., Shiurba, R.A., Linke, H.K., Schoolnik, G.K. and Clayton, D.A. (1984) EMBO J. 3, 3165–3175.
[20] Scazzocchio, C., Brown, T.A., Waring, R.B., Ray, J.A. and Davies, R.W. (1983) in: Mitochondria 1983 (Schweyen, R.J. et al. eds) pp.303–312, De Gruyter, Berlin.
[21] De Brujin, M. (1983) Nature 304, 234–240.
[22] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[23] Clary, D.O. and Wolstenholme, D.R. (1985) J. Mol. Evol., in press.
[24] Grisi, E., Brown, T.A., Waring, R.B., Scazzochio, C. and Davies, R.W. (1982) Nucleic Acids Res. 10, 3531–3539.
[25] Ward, B.L., Anderson, R.S. and Bendich, A.J. (1981) Cell 25, 793–803.
[26] Lonsdale, D.M., Hodge, T.P. and Fauron, C. (1984) Nucleic Acids Res. 12, 9249–9261.
[27] Palmer, J.D. and Shields, C.R. (1984) Nature 307, 437–440.
[28] Leaver, C.J., Forde, B.G., Dixon, L.K. and Fox, T.D. (1982) in: Mitochondrial Genes (Slonimski, P. et al. eds) pp.457–470, Cold Spring Harbor Press, NY.
[29] Bonen, L., Boer, P.H. and Gray, M.W. (1984) EMBO J. 3, 2531–2536.
[30] Kao, T., Moon, E. and Wu, R. (1984) Nucleic Acids Res. 12, 7305–7315.
[31] Gray, M.W. and Spencer, D.F. (1983) FEBS Lett. 161, 323–327.
[32] Spencer, D.F., Schnare, M.N. and Gray, M.W. (1984) Proc. Natl. Acad. Sci. USA 81, 493–497.
[33] Gottschalk, M. and Brennicke, A. (1985) Curr. Genet. 9, 165–168.
[34] Brennicke, A., Möller, S. and Blanz, P. (1985) Mol. Gen. Genet. 198, 404–410.