Initiation of transcription in yeast mitochondria: analysis of origins of replication and of genes coding for a messenger RNA and a transfer RNA

K.A. Osinga, E. De Vries, G.T.J. Van der Horst and H.F. Tabak

Section for Molecular Biology, Laboratory of Biochemistry, University of Amsterdam, Kruislaan 318, 1098 SM Amsterdam, The Netherlands

Received 1 December 1983; Revised and Accepted 27 January 1984

ABSTRACT

The initiation of transcription of the yeast mitochondrial genes coding for subunit I of cytochrome c oxidase (COX1) and for tRNA^Thr has been examined. COX1 messenger RNA synthesis is initiated in a conserved nonanucleotide sequence (ATATAAGTA) which we have previously found immediately upstream of ribosomal RNA genes at positions at which RNA synthesis starts. The 5'-end of the precursor of tRNA^Thr is located in a variant nonanucleotide motif (TTATAAGTA), which may be characteristic for tRNA genes. Using a partially purified fraction of mtRNA polymerase, we demonstrate that RNA synthesis is precisely initiated in vitro in nonanucleotide sequences preceding both ribosomal RNA-, tRNA- and messenger RNA-encoding genes and origins of replication.

INTRODUCTION

We are studying the initiation of transcription in mitochondria of the yeast Saccharomyces cerevisiae. The mitochondrial genome encodes genes whose products are part of the respiratory chain, function in mitochondrial protein synthesis or are involved in RNA processing (see for recent reviews refs 1 and 2). Using the guanylyl transferase capping assay, which labels RNAs containing a 5'-tri- or diphosphate end, Levens et al. [3] demonstrated the existence of multiple initiation sites of transcription. We have determined the DNA sequences around the map position of 5'-ends of mature large ribosomal RNA (rRNA) and of the precursor for small rRNA. Comparison reveals an identical nonanucleotide sequence (ATATAAGTA) preceding both genes [4]. This motif has been conserved in evolution since it is also present at the same positions in front of the rRNA genes of the distantly related yeast Kluyveromyces lactis [5]. The 5' terminal sequence deduced from sequence analysis of rRNAs capped with [32P]GTP establishes that in the four rRNA
genes of *S. cerevisiae* and *K. lactis* initiation of transcription occurs at the last nucleotide of the nonanucleotide box [3,5,6]. This element also occurs in putative origins of replication and - twice - in front of the gene coding for ATPase subunit 9 [5,7]. In the latter case, the RNA synthesis is initiated mainly in the upstream nonanucleotide sequence, both in vivo and in vitro [8].

In this article we describe the analysis of transcriptional initiation of the genes coding for subunit I of cytochrome c oxidase (COX1) and for tRNA\[^{\text{Thr}}\]\(_1\), and of two origins of replication.

**MATERIALS AND METHODS**

**Yeast strains and growth conditions** - *S. cerevisiae* KL14-4A [9] was grown at 28°C in a semi-synthetic lactate medium, supplemented with 0.3% (w/v) glucose [10]. The petite mutants LH26D7, described by Hensgens et al. [11], and DS-146 (a gift from Dr. A. Tzagoloff) were grown in 2% (w/v) glucose, 1% (w/v) yeast extract and 1% (w/v) Bacto peptone at 23°C and 28°C, resp.

**RNA and DNA preparations** - mtDNA and mtRNA were isolated as described in ref. 12. Recombinant plasmid DNA and the replicative form of phage M13 clones were isolated according to a modified procedure [13] of the alkaline lysis method of Birnboim and Doly [14]. After equilibrium centrifugation in a CsCl-ethidium bromide density gradient the DNA was further purified on a Sepharose CL-2B column for complete removal of residual short nucleic acids.

**Construction and physical mapping of recombinant clones** - The insert of clone pKL11 is a 2100-bp MboI fragment containing the 5'-end of the COX1 gene [11]. Digestion of pKL11 with BstNI indicates that the yeast insert has only one BstNI site whose position in exon AI is in agreement with the sequence results of Bonitz et al. [15].

Total digestion of the isolated 2100-bp insert with MspI yields fragments of the following sizes: 810, 780 (the MspI-MboI border fragment containing exon AI [11]), 240, 180, 75 (the MboI-MspI border fragment), and 66 bp. A partial MspI digestion of the insert results - among others - in a fragment of 1600 bp, which positions the 810-bp MspI product immediately upstream of the 780-bp border fragment. All the MspI fragments have been sequenced after ligation into the AccI site of M13mp7 [16]. We did not sequence across the MspI sites. The sequences are extremely AT-rich (more than 90%) and contain no other restriction sites apart from some AhalIII sites. We used this enzyme to confirm our mapping and to orientate the 810-bp MspI fragment relative to exon AI. MCEV-7 is the recombinant phage with the 810-bp MspI fragment as insert.

**mtDNA of petite mutant** DS-146 containing the tRNA\[^{\text{Thr}}\]\(_1\) gene [17] was linearized with TaqI and ligated into the AccI site of M13mp7 [16]. This recombinant phage (McGH-4) was used to confirm the presence of the variant nonanucleotide box by sequence analysis. PoGH-4 is a pUR222 clone containing the BamHI insert of McGH-4.

**End-labelling of fragments and S. nuclease analysis** - DNA fragments were labelled at the 5'-ends with polynucleotide kinase after dephosphorylation with calf-intestine phosphatase [18]. Unincorporated label was removed by centrifugation through a 1-ml syringe column filled with Sephadex G-50 [19].

1890
S\textsubscript{1} nuclease mapping [20] was performed using the modification of Weaver and Weissmann [21] under conditions as described previously [4]. S\textsubscript{1} nuclease-resistant hybrids were analysed either on a 6% polyacrylamide/7 M urea gel or on a 2% alkaline agarose gel (30 mM NaOH, 1 mM EDTA) [22].

**DNA sequencing** - Single-stranded template DNA for use in the dideoxy-sequencing method of Sanger et al. [23] was prepared from recombinant M13 plaques [24] and complementary DNA synthesis started with a 18-mer synthetic primer, complementary to the sequence of phage M13 flanking the insert.

**In vitro transcription reaction** - mtRNA polymerase was isolated from an industrial *S. cerevisiae* strain ('Koningsgist', obtained from Gist Brocades, Delft) and purified as described by Levens et al. [25], except that all buffers contained 0.2 mM dithiothreitol. At the end of the purification, the polymerase was quickly frozen in liquid nitrogen and stored at \(-70^\circ\text{C}\). Transcription reactions were carried out at 20°C for 40 min in a volume of 25 μl. The final concentrations of components in the reaction mixture were 50 μg/ml DNA, 10 mM Tris-HCl (pH 7.9), 10 mM MgCl\textsubscript{2}, 5% glycerol, 40 mM KCl, 0.5 mg/ml bovine serum albumin, 125 μM GTP, ATP and CTP, 10 μM [α\textsuperscript{32}P]UTP (spec. act. 50 Ci/mmol). In some experiments, [γ\textsuperscript{32}P]ATP was used as radioactive substrate (2.5 μM; spec. act. 1500 Ci/mmol). The reactions were stopped by addition of 25 μl of a buffer containing 10 mM Tris-HCl (pH 7.9), 0.5 M NaCl, 0.1% sodium dodecylsulphate and deproteinized with phenol (equilibrated with the same buffer). Free label was removed by centrifugation through a 1-ml syringe column filled with Sephadex G-50 [19]. The in vitro transcription products were collected by ethanol precipitation and dissolved in de-ionized formamide. Samples were heated for 90 sec at 70°C before electrophoresis.

**RESULTS**

**Initiation of transcription of the COX1 gene**

The mature messenger RNA (mRNA) of the split gene coding for COX1 (genetic locus OXI3) has a length of 2100 nucleotides [11]. This 18S RNA has a 5' untranslated leader of approximately 400 nucleotides and a 3' extension of about 100 nucleotides, as determined by electron microscopy of RNA/DNA heteroduplexes [11]. We have used \(S\textsubscript{1}\) nuclease mapping to localize the 5'-end of this mRNA more precisely (Fig. 1). Clone pKL41 was digested with BstNI, large fragments were labelled at the 5'-end with polynucleotide kinase as indicated and hybridized with total mtRNA from wild-type KL14-4A or the petite mutant LH26D7. After treatment with \(S\textsubscript{1}\) nuclease, the protected DNA was analysed by alkaline agarose gel electrophoresis. Based on the length of the protected DNA (670 nucleotides), the 5'-end of the 18S RNA can be mapped at a distance of about 540 bp from the AUG start codon, placing it in a 810-bpMspI fragment. This fragment was cloned into phage M13mp7 (resulting in clone McEV-7) and
Fig. 1. Initiation of transcription of the COX1 gene. A. S nuclease mapping of the 5'-end of COX1 mRNA. A 5' 32P-labelled digest of pKL41 x BstNI was hybridized with total mtRNA and treated with S nuclease. The protected DNA was analysed on a 2% alkaline agarose gel, blotted onto nitrocellulose filters and visualized by autoradiography. Lane a, pKL41 x BstNI. Before labelling the vector fragments of 383, 121 and 13 bp have been removed by Sepharose CL-2B column chromatography. Lanes b and c, S nuclease analysis (65 U/ml) after hybridization with total mtRNA from KL14-HA and LH26D7, resp.; m, 32P-labelled DNA marker from pBR322 x Hinfl. B. In vitro transcription using [α-32P]UTP of the 810-bp MspI fragment cloned into phage M13 (McEV-7). To obtain a run-off transcript McEV-7 was restricted with EcoRI, which cuts next to the AccI cloning site of phage M13. The products were fractionated on a 5% denaturing polyacrylamide gel. Lane a, run-off transcription products; m, 32P-labelled DNA marker from pBR322 x Hinfl. C. Restriction map at the 5'-end of the COX1 gene. For construction, see Materials and Methods. Black box, exon A1; open box, part of first intron. S1 nuclease-resistant DNA and run-off transcription product are diagrammed. * 32P-labelled end. The lengths are indicated in nucleotides.

sequenced (see Fig. 2). A perfect nonanucleotide box is present in the region containing the 5'-end of the mRNA. We have used McEV-7 as template in an homologous in vitro transcription system described by Edwards et al. [26]. Digestion of McEV-7 with EcoRI results in a specific run-off transcript of 290 nucleotides (Fig. 1B) whereas when the reaction is performed in the absence of CTP (the first C residue to be incorporated is at position +80) a transcript of 79 nucleotides is made (not shown). These results demonstrate that initiation of transcrip-
Fig. 2. DNA sequence around the start of COX1 mRNA synthesis. The non-coding strand is shown. The nonanucleotide sequence is boxed.

Initiation of transcription of the gene for tRNA\textsubscript{Thr}.

Inspection of published mtDNA sequences reveals that some tRNA genes are preceded by a variant of the nonanucleotide box with the sequence TTATAAGTA. In order to test whether transcription is initiated in this variant nonanucleotide sequence, we have selected the gene coding for tRNA\textsubscript{Thr}. This gene is of additional interest since it is located on the DNA strand of opposite polarity to that encoding the other genes [17]. If transcription of this gene starts at the box, the mature 5'-end of the tRNA must be generated by processing since the nonanucleotide sequence is located 135 nucleotides upstream of the gene. Martin and Underbrink-Lyon [27] have identified a gene on the mitochondrial genome which codes for an RNA product involved in processing of the 5'-ends of tRNAs. Petite deletion mutants, which do not contain this gene, accumulate precursor tRNA with 5' extensions. In order to map the 5'-end of the tRNA\textsubscript{Thr} precursor, we performed an S\textsubscript{1} nuclease experiment using total mtRNA from petite mutant DS-146, which lacks the gene involved in tRNA processing. mtDNA from this mutant was restricted with
Fig. 3. Initiation of transcription of the tRNA\textsuperscript{Thr} gene. A. S\textsubscript{1} nuclease mapping of the 5'-end of the tRNA\textsuperscript{Thr} precursor. A 5'\textsuperscript{32}P-labelled digest of petite mutant DS-146 x DdeI was hybridized with total mtRNA of DS-146 and treated with S\textsubscript{1} nuclease. The protected DNA was analysed on a 6% denaturing polyacrylamide gel. Lane a, input DNA; b and c, S\textsubscript{1} nuclease analysis with 50 U/ml or 100 U/ml of S\textsubscript{1} nuclease, resp.; m, \textsuperscript{32}P-labelled DNA marker from pBR322 x HinfI. B. In vitro transcription using [\gamma\textsuperscript{32}P]ATP on DS-146 cloned into pUR222 (PcGH-4). The run-off transcription products from PcGH-4 template, cleaved with HinfI (lane a), were fractionated on a 5% denaturing polyacrylamide gel. Lane m, \textsuperscript{32}P-labelled DNA marker from pBR322 x MspI. C. Diagram of the yeast insert in PcGH-4. The tRNA\textsuperscript{Thr} gene is indicated by the black box. The run-off transcription product and S\textsubscript{1} nuclease-resistant DNA are depicted. The lengths are indicated in nucleotides.

DdeI, 5' end-labelled and hybridized with total DS-146 mtRNA. The result is shown in Fig. 3A. The length of the S\textsubscript{1} nuclease-resistant DNA (185 nucleotides) indicates that the 5'-end of the precursor tRNA maps in the TTATAAGTA sequence, suggesting that it is used \textit{in vivo} as a transcription-initiation site. We then performed an \textit{in vitro} transcription experiment with the same fraction of RNA polymerase activity as used for transcription of the COX1 template and the large rRNA gene [7]. Clone PcGH-4, in which the repeat unit of DS-146 had been inserted, was restricted with HinfI and served as template. The occurrence of a run-off product of 295 nucleotides demonstrates that also
in vitro tRNA-precursor synthesis is precisely initiated in this variant nonanucleotide box (Fig. 3B).

In vitro transcription-initiation at an origin of replication

The yeast mitochondrial genome contains three or four origins of replication, depending on the strain [28,29], which are composed of a 300-bp conserved sequence (abbreviated 'ori' (origin) or 'rep' (replicon)). We have already observed that the nonanucleotide sequence is present at the very end of such an origin, just in front of the first of the three characteristic GC clusters. RNA cappable in vitro has been detected in petite mutants consisting of repeated origins. The transcrip-
tion of this RNA is initiated at the last nucleotide of the box [5]. Baldacci and Bernardi [30] have shown by S1 nuclease analysis that the 5'-ends of RNA isolated from such petite mutants map at the position of the nonanucleotide sequence.

We have tested the replication origins rep-1 and rep-2 in the in vitro transcription system. pSCM128 (rep-2) was restricted either with EcoRV or with MspI and used as template in in vitro transcription experiments. The length of the run-off products corresponds exactly with the map position of these sites, assuming a start in the nonanucleotide sequence (Fig. 4A). These experiments have been carried out with [α-32P]UTP. Because the first nucleotide of these RNAs is an A residue [5,31], the RNA made in vitro can also be labelled with [γ-32P]ATP. The result, obtained with the same cleaved template DNAs as used in Fig. 4A, is shown in Fig. 4B (lanes a' and b') and illustrates that there is no difference between these two labelling procedures. pGT22-A (rep-1) contains an XbaI site just upstream of the nonanucleotide sequence and this enables us to investigate whether sequences upstream of position -10 are dispensable for precise initiation of transcription. Fig. 4B shows that on pGT222-A, restricted either with EcoRV or EcoRV and XbaI, the expected run-off RNAs are synthesized.

We conclude that the same RNA polymerase preparation which recognizes rRNA, mRNA and tRNA promoters is capable of initiating transcription at origins of replication. The mtRNA polymerase preparation is not pure enough to decide whether this is a property of one RNA polymerase or due to insufficient separation of an RNA polymerase and a primase.

DISCUSSION

Transcription of the rRNA genes in two distantly related yeast strains S. cerevisiae and K. lactis is initiated within a conserved nonanucleotide sequence [5]. Here we have extended this analysis and studied initiation of RNA synthesis of a mRNA and a tRNA gene and at origins of replication. A compilation of sequences around initiation sites for transcription is shown in Fig. 5. S1 nuclease analysis shows that the COX1 mRNA has an untranslated leader of 540 nucleotides. Sequence analysis
Fig. 5. DNA sequence comparison of initiation sites for transcription. The data are from: rRNA genes [4], tRNA\textsuperscript{Thr} (this study), tRNA\textsuperscript{Phe} [31]; tRNA\textsuperscript{fMet} [37], tRNA\textsuperscript{Glu} [38], ATPase subunit 9 [8], COX1 (this study), replication origins [5,31; this study]. Note in the glutamic tRNA nonanucleotide box the deviating G at position -3 instead of an A. The DNA sequence, however, in this region of the gene has been determined before the identification of the nonanucleotide motif [39] and has not been checked thereafter [38].

reveals that its 5'-end is located in a perfect nonanucleotide motif. We demonstrate that transcription indeed starts at this position using the homologous in vitro transcription system developed by Edwards et al. [26]. In support of this finding, a transcript from this region of the genome has recently been found, which can be capped in vitro with guanylyl transferase [31]. The sequence of the 5' terminus of this RNA perfectly matches the DNA sequence starting with the last A in this nonanucleotide sequence.

The transcription of the tRNA\textsuperscript{Thr}\textsubscript{1} gene was studied as a representative of the tRNA genes. This gene is the only one encoded by the DNA strand of opposite polarity to that on which the remaining genes are located [17]. A variant nonanucleotide sequence (TTATAAGTA) is present upstream of this and some other tRNA genes (Fig. 5). The S\textsubscript{1} nuclease mapping of mtRNA, combined with the precise run-off transcripts in the in vitro transcrip-
igation assay, clearly illustrates that initiation of transcription occurs in the box. The presence of such a variant box in front of tRNA genes only may put these genes in a separate category. We have noticed the presence of such a 'T-box' immediately upstream of the gene coding for tRNA\textsuperscript{Cys} [32]. However, S\textsubscript{1} nuclease experiments by Frontali et al. [33] do not give any evidence that RNA synthesis starts in the vicinity of this box. Instead, longer transcripts are observed, whose 5'-ends map in the proximity of the upstream tRNA\textsuperscript{Thr} gene [34]. The tRNA\textsuperscript{Cys} gene may, therefore, belong to the same transcription unit as the 21S rRNA and tRNA\textsuperscript{Thr} gene [35]. This may suggest that in this case either the variant nonanucleotide box is used as initiation site only under certain physiological conditions, or that essential promoter elements are absent from this region.

Finally, we have obtained evidence that the partially-purified RNA polymerase preparation is capable of specifically initiating RNA synthesis at origins of replication. Again, the start is within the box. Most likely the transcription in these origins can be explained as RNA priming of DNA replication. Some petite mutants, characterised by a short repeat unit length and the presence of an origin of replication, synthesize mtRNA with the length of the repeat unit or larger [30], whose 5'-ends are localized in the nonanucleotide sequence [5,30]. In one such a petite mutant, Baldacci and Bernardi [36] have recently also demonstrated the presence of short RNA segments covalently attached to the 5'-end of DNA, thereby substantiating the priming function of this RNA. The 5'-ends of these short RNA segments map at the same position as those of longer molecules [36]. Such long RNAs may result from a shortage of DNA replication enzymes compared to the highly elevated number of replication origins in these petite mutants, so preventing the switch from RNA to DNA synthesis. The RNAs made \textit{in vitro} by run-off transcription thus resemble the class of long RNAs synthesized in these petite mutants.

It is interesting that the same fraction of RNA polymerase is capable of transcribing all classes of genes. However, the mtRNA polymerase which we have used has not been purified to homogeneity. It remains to be shown, therefore, whether one or
more types of polymerase (e.g. a polymerase and a primase) are involved in the transcription of yeast mtDNA.

ACKNOWLEDGEMENTS

We are indebted to Mrs H.Blanc and B.Dujon (Centre de Génétique Moléculaire du C.N.R.S., 91-Gif-sur-Yvette, France) for gifts of the plasmids pSCM128 and pGT22 and providing us with the sequence of rep-2 prior to publication, to Dr. A.Tzagoloff (Fairchild Center for Life Sciences, Columbia University, New York, N.Y. 10027, USA) for a gift of DS-146, and to Gist Brocades N.V. (Delft) for gifts of 'Koningsgist'. We thank Mrs M.J.A.Groot Koerkamp for expert technical assistance and Profs L.A.Grivell and P.Borst and Dr. R.Benne for their comments on the manuscript. This work was supported in part by a grant to P.Borst-L.A.Grivell from The Netherlands Foundation for Chemical Research (SON) with financial aid from The Netherlands Organization for the Advancement of Pure Research (ZWO).

Abbreviations: bp, base pair(s); COX1, subunit I of cytochrome c oxidase; mRNA, messenger RNA; rep, replicon; rRNA, ribosomal RNA; tRNA, transfer RNA.

REFERENCES

Nucleic Acids Research