Determinants for binding of a 40 kDa protein to the leaders of yeast mitochondrial mRNAs

Peter J.T.Dekker, Jeroen Stuurman, Katinka van Oosterum and Leslie A.Grivell*

Section for Molecular Biology, Department of Molecular Cell Biology, University of Amsterdam, Kruislaan 318, 1098 SM Amsterdam, The Netherlands

Received April 15, 1992; Revised and Accepted May 13, 1992

ABSTRACT

An abundant yeast mitochondrial 40 kDa protein (p40) binds with high specificity to the 5'-untranslated region of cytochrome *c* oxidase subunit II (COX2) mRNA. Using mobility shift and competition assays, we show here that purified p40 complexes with the leaders of all eight mitochondrial mRNAs of *Saccharomyces cerevisiae*. The location of the protein binding site on the different leaders is not conserved with respect to the AUG start codon. *In vitro* RNA footprint and deletion experiments have been used to define the p40-binding site on the leaders of COX1 and ATP9 mRNAs. Nucleotides at, and near, a single stranded region are protected or exposed for DEPC modification by binding of p40 to these leaders. Removal of this region from the COX1 messenger shows that it is essential for the protein-RNA interaction. While no obvious sequence similarity can be detected between the single stranded regions in different leaders, a nearby helical segment is conserved. A consensus model for p40-RNA interactions is presented and the possible biological function of p40 is discussed.

INTRODUCTION

The 5' untranslated regions of mitochondrial mRNAs in the yeast *Saccharomyces cerevisiae* are exceptional in several respects. Most leaders are extremely long and A+U rich, and are not capped at their 5'-ends. The start codon is seldom the first AUG and is not surrounded by obvious discriminatory sequences [1]. These features seem to exclude a translation initiation mechanism either by direct ribosome-mRNA recognition as in prokaryotes [2], or by a scanning model for eukaryotic cytoplasmic translation initiation [3,4]. As a consequence, recognition of the leaders by mitochondrial ribosomes and translation initiation in yeast mitochondria are, at present, poorly understood processes. Genetic studies have shown that expression of a number of mitochondrial genes is specifically controlled at the level of translation by positive regulatory proteins encoded by nuclear genes [5,6]. Although translation of at least five out of the eight mitochondrial mRNAs requires the presence of these specific

activators, their mode of action is largely unknown [7–16]. Rearrangements of mitochondrial DNA, leading to fusion of all or part of the 5'-untranslated region of one mRNA to the coding region of another, have been shown to suppress mutations in several nuclear genes that code for specific activators, localizing the site of action of these proteins to the leaders of mitochondrial mRNAs [15–17]. Some activators seem to associate with the mitochondrial small ribosomal subunit, others have been shown to interact with the inner membrane of mitochondria [18,19]. For the translation of COX3 and COB mRNA, this has led to the proposal of a model in which the activators function to target the mRNA and ribosome to the membrane where cytochrome *c* oxidase subunit III and cytochrome *b* are synthesized near their final destination [5,19].

The 5'-untranslated regions of numerous prokaryotic and eukaryotic mRNAs have been shown to be involved in the regulation of protein synthesis. While primary structure in the leaders has long been recognized as an important determinant for efficient translation initiation, higher order structures contribute significantly to the efficiency and fidelity of protein synthesis, by masking the ribosome binding site or start codon, or impeding the scanning ribosome [2,20,21]. Secondary structures are also thought to be involved in the internal initiation of translation on the leaders of picornaviral mRNAs [22]. The vast majority of leader binding regulatory proteins inhibit translation by stabilization of secondary structures. Translational repression of the synthesis of some *E.coli* ribosomal proteins [23,24], several bacteriophage encoded proteins [2] and the production of ferritin in eukaryotes [25] have been studied extensively. Apart from the mitochondrial trans-acting factors, examples of positive translational regulation by specific leader binding proteins are surprisingly scarce [26,27].

Previously we identified a protein with an apparent molecular weight of 40 kDa (p40) that interacts with the 5' untranslated region of the yeast mitochondrial COX2 gene [28]. The significance for binding of a single stranded area near the AUG-start codon was revealed by site directed mutagenesis and DEPC-footprinting. Furthermore, Bal31 deletion experiments indicated the importance of secondary structure for the formation of the correct binding site [28]. The protein was demonstrated not to be any of several

* To whom correspondence should be addressed

genetically identified mRNA specific activators, and is abundant in yeast mitochondria (~25,000 copies per cell [29]).

In this study, we present evidence for the interaction of p40 with the 5' untranslated regions of mRNAs of all protein coding genes in mitochondria. In a number of leaders, the binding site is localized in different positions with respect to the start codon, but always adopts a similar secondary structure. A consensus model for p40 binding to the different leaders of mitochondrial mRNAs is presented. The possible role of p40 in mitochondria is discussed.

MATERIALS AND METHODS

Plasmid constructions and strains

All plasmids were constructed by standard techniques [30] in the transcription vectors pEP30 or pEP40 [31], harbouring the T7 and SP6 promoters on either side of the polylinker of pUC18. 5' deletions of the *COX1* leader were made by digestion of pKL41, containing a 2 kb *Sau3A* fragment spanning the entire *COX1* leader from *S.cerevisiae* KL14-4A mitochondrial DNA, with *EcoRI*, subsequent *Bal31* exonuclease treatment, digestion with *BamHI* and blunt ligation into the *SmaI* site of pEP40. The extent of the deletion of the resultant subclones pCOX1Δ28, Δ38, Δ13, Δ35, Δ31 and Δ43 was checked by sequence analysis. *COX2* subclones have been described previously [28]. Subclones of pSP65XAI (kindly provided by Dr. T.D.Fox), containing the *COX3* leader from the *XbaI* site at -610 until the *AccI* site at +20, were constructed by digestion with *MboII* or *DraI* plus *HindIII*, with subsequent blunt ligation in the *SmaI* site of pEP40, yielding plasmids pCOX3MA and pCOX3DA. Constructs of the *COB* leader were generated by digestion of pBS/2-2 (gift of Dr. C.Dieckmann), encompassing a 1 kb *COB* leader fragment, with *BamHI*, *DraI* and *EcoRI* and subcloning of the fragments in the *BamHI/EcoRI*, *BamHI/HincII*, *SmaI* or *SmaI/EcoRI* sites of pEP30, resulting in plasmids pCOBBE, pCOBBD, pCOBDD and pCOBDE. Plasmid pCOBHE was made by digestion of pCOBDE with *HincII*, and back ligation. Leaders of *ATP6* and *ATP8* were obtained by cloning of a 2.5 kb *EcoRI* fragment (RR7), encompassing both leaders from *S.cerevisiae* KL14-4A, in the *EcoRI* site of pEP30 (pATP8-6). This plasmid was further digested with *XbaI* or *MspI*, and fragments were subcloned in respectively the *XbaI* or *AccI* site of pEP30, yielding the plasmids pATP6XE, pATP6M and pATP8M. The leaders of *ATP9* and *VAR1* were amplified by PCR on mitochondrial DNA of *S.cerevisiae* D273-10B, with primers complementary to the transcription start (5'-AATATATATAGTTTATCATA-3'), and the 3' end of the coding sequence of the *ATP9* gene (5'-TTATACACCGAATAATAATA-3'), or complementary to the *VAR1* 5'-processing site (5'-ATAATGTTATTAATAACTGC-3') and internal in the *VAR1* coding sequence (5'-CCAATCCGGTGAA-CAACC-3'). Subsequent to amplification, fragments were treated with Klenow DNA polymerase, 5' phosphorylated and then cloned in the *SmaI* site of pEP30, resulting in plasmids pATP9 and pVAR1. All subclones were created in *E.coli* strains HB101 or JM101 and routinely checked with restriction enzymes and sequence analysis of the junctions.

In vitro transcription

Templates for run-off *in vitro* transcription were produced by digestion of subclones with suitable restriction endonucleases (schematically represented in Figure 3). After purification of digested DNA, run-off transcripts were made by incubation of

1 μg of template DNA for one hour at 37°C in the presence (for direct binding studies) or absence (for competition assays) of [α -³²P] UTP, with T7 (pCOX1Δ28, pCOX1Δ31, pCOX1Δ35, pCOX2Δ21, pCOX2Δ29, pCOX2Δ33, pCOX3MA, pCOX3DA, pCOBBD, pATP8-6, pATP8M) or SP6 RNA polymerase (pCOX1Δ13, pCOX1Δ38, pCOX1Δ43, pCOX2Δ9, pSP65XAI, pCOBBE, pCOBDE, pCOBDD, pCOBHE, pATP6XE, pATP6M, pATP9, pVAR1) according to standard procedures [32]. Unincorporated nucleotides were removed by purification over a 1 ml Sephadex-G50 column. Yields and quality of transcripts were routinely checked on agarose or polyacrylamide gels.

Analysis of RNA protein interactions

Analysis of p40 binding to the mitochondrial leader RNAs was essentially according to [28]. In effect approximately 2 ng of [α -³²P] UTP-labeled RNA (50–150 cps) was mixed with 50–200 ng of purified p40 and 0.5–1.0 μg of either total purified *E.coli* or calf liver RNA (Sigma) to minimize non-specific p40-RNA interactions in a final volume of 30 μl binding buffer (10 mM Tris-HCl pH7.5, 1 mM EDTA, 3 mM MgCl₂, 50 mM NaCl, 5% glycerol). For competition experiments p40 was first incubated with a 25–100 fold molar excess of cold leader RNA for 15 min. at 30°C before addition of the radioactive labeled probe. Protein-RNA contacts were established in 15 min. at 30°C, after which complexes were resolved by electrophoresis through native polyacrylamide gels.

DEPC-footprinting analysis was essentially performed according to [28]. Approximately 500 cps (~20 ng) of 5'-end labeled COX1Δ13**DraI* and ATP9**AseI* RNA (gel purified) was mixed with, 0.5–1.0 μg of purified p40 (or with an equivalent volume of protein storage solution) and 1 μg of calf liver RNA in 50 μl 50 mM sodium cacodylate (pH7.5), and incubated for 15 min. at 30°C (these conditions are sufficient to complex all *COX1* or *ATP9* RNA, as assayed by gel retardation). 2 μl of DEPC (Sigma) was added to the incubation mixture and after thorough mixing the incubation was continued for 15 min. at 30°C. Analysis of modified RNA was essentially according to [33]. RNA was purified by phenolization and ethanol precipitation, and chain scission was induced by incubation in 20 μl freshly prepared 1 M aniline (Merck) in 0.3 M sodium acetate (final pH 4.5) at 60°C for 20 min. Two *n*-butanol precipitations were performed to remove the aniline, after which RNA pellets were carefully rinsed with ethanol, dried and resuspended in layer mix (0.03% bromophenol blue, 0.03% xylene cyanol, 25 mM EDTA in formamide). After boiling for 2 min., care was taken to layer equal amounts of radioactive material onto a 12% polyacrylamide–7 M urea gel. After fixation in 10% acetic acid, gels were dried and autoradiographed.

RNase T1 fractionation analysis

2000 cps (~80 ng) of [α -³²P] UTP-labeled RNA transcribed from the appropriate templates (see figure legends) was mixed in a final volume of 50 μl binding buffer containing 2 μg of calf liver RNA, 1.5 μg of purified p40, and incubated for 15 min. at 30°C. Subsequently, the mixture was treated with 20 U of RNase T1 (Boehringer) for 15 min. at 30°C. 5 μl of the reaction mixture was phenolized and precipitated (total digest fraction), the rest was spotted slowly on nitrocellulose filters and washed 5 times with 50 μl TBE buffer. Filters were washed with 1% SDS, to recover protein-bound RNA and purified by phenol extraction and ethanol precipitation. An alternative approach consisted of the same treatment, but instead the reaction mixture

was directly loaded onto a native 4% polyacrylamide gel in TBE and run for two hours at 30 mA at 4°C. Gels were directly autoradiographed and retarded material was eluted for 16 hours at room temperature in elution buffer [30]. Eluted material was purified, resuspended in layer mix, boiled for 2 min. and separated on 12% polyacrylamide–7 M urea gels.

Truncations by ribonucleases

2600 cps (~100 ng) [α - 32 P] UTP-labeled COX1 Δ 35**Dra*I RNA was digested in 20 μ l binding buffer with 50 U of RNase T1 for 15 min. at 37°C. RNA was purified and separated on a 12% polyacrylamide–7M urea gel, and directly autoradiographed. The appropriate RNase T1 fragment (COX1 Δ 35*T1) was eluted and purified by phenol extraction and ethanol precipitation. RNase CL3 deletion was done in the same manner, but ~50 ng of COX1 Δ 35*T1 was digested with 0.05 U RNase CL3 (Boehringer) in 4 μ l 8 M urea, 20 mM Tris-HCl (pH8.0) for 15 min. at 50°C, and purified (yielding COX1 Δ 35*T1*CL3). The ribonuclease is under these conditions C-specific.

RESULTS

p40 binds to the leaders of all mitochondrial mRNAs

Previously we demonstrated that full length leader RNAs of the *COX1*, *COX3* and *COB* genes compete efficiently for binding of p40 to the leader of the *COX2* gene. The antisense transcripts of the same constructs were unable to compete for this interaction, suggesting that p40 binds specifically to the leaders of these genes [28]. Since yeast mitochondrial DNA harbours only seven major protein coding genes, beside *COX2*, we asked whether all 5'-untranslated regions of mitochondrial genes have the same effect. Mitochondrial DNA fragments accommodating the full length leaders of the *ATP6*, *ATP8*, *ATP9* and *VAR1* genes were cloned in the transcription vector pEP30 as described in Materials and methods. *In vitro* transcripts of these constructs were assayed for competition of the *COX2* leader-p40 interaction. As shown in Figure 1, an excess of sense transcripts from these constructs

Figure 1. Competition for p40-*COX2* RNA binding *in vitro*. A 25-fold molar excess of cold competitor RNA was incubated with 50 ng of purified p40, prior to the addition of 32 P-labeled COX2 Δ 9**Rsa*I RNA. After mobility shift assay, radioactive material was visualized by autoradiography. Lane 1, input *COX2* RNA; lane 2, *COX2* RNA with p40 without competition; lanes 3–9, competition with COX2 Δ 9**Rsa*I, COBDE**Eco*RI, COBDE**Eco*RI, ATP6XE**Dra*I, ATP8-6**Xba*I, ATP9**Eco*RI and VAR1**Eco*RI respectively; C, p40-*COX2* complex; F, free *COX2* RNA.

competes efficiently for binding of p40 to the leader of *COX2* (lanes 6–9). Competition by these leaders was as efficient as competition by *COX2* or *COB* leaders (compare to lanes 3–5). Since this protein-RNA interaction is not competed by transcripts lacking leader sequences [28], even when present in high excess, it is likely that all leaders of mitochondrial mRNAs contain one or more specific binding site(s) for p40.

Positioning of the p40 binding site within the leaders of mitochondrial mRNAs

Except for the leader of the *COX2* gene, the leaders of mitochondrial genes are extremely large (>500 nt) and therefore not suitable for direct protein-RNA interaction studies by band-shift assays, as performed with the *COX2* leader (Figure 1, lane 2). To investigate the binding of p40 to the leaders directly, they had to be truncated to more suitable sizes (<200 nt). Shorter run-off transcripts of several leaders were produced by digestion of the DNA with suitable restriction endonucleases. DNA fragments encoding the leaders were shortened at the 5'-ends by Bal31 exonuclease, or digested with restriction endonucleases, and subcloned in the transcription vectors pEP30 and pEP40 (see Materials and methods). This approach yielded a set of subclones containing relatively small fragments of all mitochondrial leaders (see Figure 3). Transcripts produced from these fragments were tested for their ability to interact with purified p40 by band-shift analysis. A large excess (>250 fold) of total *E. coli* or calf liver RNA, to compete for nonspecific binding, was always included in the binding reactions. Figure 2 shows a result of such an analysis. Only the smallest transcripts, of subclones of the leaders of seven mitochondrial mRNAs, that were still able to interact with p40 are indicated. Direct p40 binding to the *VAR1* leader is not examined yet. At higher protein concentrations the RNAs sometimes complexed with p40 in two bands (see Figure 2). This might be due to binding of more than one p40 molecule to the leaders, or to an alternative conformation of the RNA. The migration of both COX3DA**Eco*RI (lane 3) and ATP8M**Nde*I RNA (lane 7) as two separate bands on this native gel is also likely to be a conformational effect since all RNAs used in binding experiments migrated as single bands on denaturing polyacrylamide gels. The capacity of all RNAs to interact with p40 is represented schematically in Figure 3. Since transcripts lacking leader sequences were not able to interact with p40 [28],

Figure 2. Mobility shift assay with various 32 P-labeled RNA fragments of the leaders of mitochondrial mRNAs. Lanes 1–7 contain COX1 Δ 35**Dra*I, COX2 Δ 29**Rsa*I, COX3DA**Eco*RI, COBHE**Eco*RI, ATP6M**Dde*I, ATP9**Asel* and ATP8M**Nde*I RNA respectively; –, without addition of p40; +, with addition of 50 ng purified p40.

Figure 3. Lengths and p40 binding characteristics, assayed by mobility shift and/or competition assays, of RNAs transcribed from different subclones of the leaders of all yeast mitochondrial mRNAs. Lines indicate *in vivo* untranscribed regions; open bars, *in vivo* untranslated regions; black bars, translated regions; arrows, *in vivo* transcription starts; circles, putative processing sites; +, RNA does bind p40; -, no p40 binding detected; ±, marginal p40 binding. Cotranscribed polylinker sequences are not indicated.

the binding of this protein to the leaders of mitochondrial mRNAs must be considered specific. Transcripts of larger subclones, containing all sequences required for binding, sometimes exhibited no binding activity (see Figure 3, *COX1*). Computer-aided folding, with the RNA-FOLD program [34] predicted, however, a different secondary structure for these RNAs than for the full length leader transcripts or the truncated RNAs that were still able to bind to p40 (data not shown). For instance, *COX1* constructs COX1Δ38*BamHI and COX1Δ28*MapI, which do not exhibit p40 binding activity (see Figure 3), are predicted to fold entirely differently from COX1Δ28*EcoRI or COX1Δ13*DraI RNAs, which do show p40 binding. This indicates that secondary structure in the RNAs is an important parameter for binding of p40.

The RNA fragments of the different leaders that were still able to interact with p40, varied surprisingly in their position with respect to the AUG start codons of the different mRNAs (Figure 4). The binding site ranges from the 5'-end of the *ATP9* leader, the middle of the *COX1* leader to the 3'-end of the *COX3* leader. The regions in the mitochondrial leaders sufficient for binding of p40 (hatched boxes in Figure 4) do not show a significant sequence match to the single stranded 5'-UAAA-GGUAA-3' sequence motif shown to be essential for p40

binding to the *COX2* leader [28]. All leaders compete, however, for binding to the *COX2* leader (see Figure 1 and [28]), so they all probably encompass a similar binding site for p40. This again indicates that higher order structures are important in forming the binding site for p40.

Demarcation of the p40 binding sites on the leaders of *COX1* and *COX3*

To further define the binding site of p40 on the leaders of mitochondrial mRNAs, we designed a technique that is based on the G-deficient character of the mitochondrial leaders. If the binding site contains G-residues, as with the *COX2* leader, we expect to disturb the interaction with p40 by digestion with RNase T1. However, if the binding site does not contain any G-residues, it should be possible to isolate an RNase T1 fragment containing the binding site in the form of an RNA-protein complex. The protein-bound RNase T1 fragment can be specifically retained by isolation from a retardation gel or by direct filtration through nitrocellulose and further analysed.

We performed this experiment on a transcript that contained the entire *COX1* leader (COX1Δ28*EcoRI), a shorter form of this (COX1Δ13*DraI), the entire *COX3* leader (SP65XAI*HindIII) and a truncated form (COX3DA*EcoRI), that were all

Figure 4. Localization of the p40 binding site on the leaders of all mitochondrial mRNAs. Only the *in vivo* 5'-untranslated regions, as far as the AUG initiation codon, are indicated. RNA fragments sufficient for p40 binding are hatched. Arrows indicate *in vivo* transcription starts, circles indicate putative processing sites.

shown to bind p40 (see Figure 3). In contrast to protein binding to the leader of *COX2*, complexes of the leaders of *COX1* or *COX3* with p40 were shown not to dissociate upon digestion with RNase T1, as judged by the presence of a retarded RNase T1-fragment in gel mobility shift assays (results not shown). Only a 78 nt RNA fragment is specifically bound by p40 with both *COX1* transcripts (Figure 5, lanes 2 and 6). The isolation of the same RNase T1 fragment from the smaller RNA molecule (*COX1Δ13*DraI*) as from the full length *COX1* leader indicates that a bona fide p40 binding site is present on *COX1Δ13*DraI*. The *COX1* leader therefore only contains one efficient p40 binding site, located within this 78 nt RNase T1 fragment, that is situated between positions -385 and -307 relative to the initiation codon.

Figure 5 shows that a 79 nt (from -76 to +3) and a 166 nt (-479 to -313) RNase T1 fragment of the complete *COX3* leader are most prominently retained by nitrocellulose filtration (lane 4). The same 79 nt fragment is also retarded by p40 from the smaller *COX3DA*EcoRI* transcript (lane 8). These observations leave two possible interpretations. First, the binding site for p40 could be formed by bases belonging to both the 79 nt and the 166 nt RNase T1 fragment. If this is the case, the binding activity of *COX3DA*EcoRI* must be ascribed to spontaneous formation of an alternative binding site within the 79 nt RNase T1 fragment, since this molecule lacks the information of the 166 nt fragment. While p40 is highly selective in its ability to bind RNA [28], it is not clear why an alternative binding site should be formed with such ease. The second and more likely interpretation is that *COX3DA*EcoRI* adopts a wild type secondary structure, that is recognized by p40, explaining the isolation of the 79 nt fragment in both RNAs. If this is the case, the *COX3* leader possibly contains more than one p40 binding site. We have located, by deletion experiments, fragments sufficient for p40 binding (see Figure 4), which means that they contain all parameters essential for binding. This does not, however, mean that they are the only p40 binding sites on the leaders of mitochondrial mRNAs. As found for *COX3*, it is possible that more than one p40 binding site is present on the leaders of other mitochondrial mRNAs.

The p40 binding site on the *COX1* leader maps to a single stranded region

To determine more precisely the binding site for p40 within the 78nt RNase T1 fragment of *COX1* leader RNA, studies on

Figure 5. RNase T1 fractionation analysis of the leaders of *COX1* and *COX3*. Total RNase T1 digests of *COX1Δ28*EcoRI* (lane 1), *SP65XAI*HindIII* (lane 3), *COX1Δ13*DraI* (lane 5) and *COX3DA*EcoRI* (lane 7), and RNase T1 fragments interacting with p40 were isolated by nitrocellulose filtration (lanes 2 and 4), or elution from retardation gels (lanes 6 and 8), and separated on a 12% polyacrylamide/7M urea gel. Sizes in nucleotides of the RNase T1 fragments are indicated.

transcripts that overlap with this fragment were carried out. Predicted secondary structure folding of the *COX1Δ13*DraI* RNA (from -424 until -288), that accommodates the complete 78 nt RNase T1 fragment, and was shown to interact with p40 (Figure 3), revealed that this molecule has the capacity to form two separate hairpin structures (Figure 6A). Cotranscribed polylinker sequences fold into a separate domain and are not depicted in this figure. The existence of both hairpins was confirmed by nuclease protection experiments, although it is probable that the base of each helix folds into one large single stranded loop since these nucleotides were more exposed to RNase A than nucleotides in the helices (data not shown). To localize the p40 binding site to either the 5'- or the 3'-half of the molecule, we produced shorter transcripts of *COX1Δ13* by digestion of the DNA template with *AseI* (*COX1Δ13*AseI*) and by digestion of *Bal31* deletion clone *COX1Δ35* with *DraI* (*COX1Δ35*DraI*). These molecules accommodate helix 1 and 2 of *COX1Δ13*DraI* respectively (Figure 6A), with conservation of wild type secondary structures predicted by computer modelling. Gel mobility shift experiments with both molecules and p40 shows that with the 5'-arm (*COX1Δ13*AseI*) only a minute amount of complex was observed, compared to the 3'-arm (*COX1Δ35*DraI*, compare lanes 2 and 3 to 5 and 6 of Figure 6B), so the main p40 binding site is located in the 3'-arm of *COX1Δ13*DraI* (helix 2). We investigated the *COX1Δ35*DraI* molecule further in two steps, making use of base-specific RNases. First, the 3' 17 nt were removed by RNase T1, leaving a molecule of 42 nt (*COX1Δ35*T1*) that consists entirely of bases belonging to the 78nt RNase T1 fragment isolated from the wild type leader (Figure 5, lane 2). This

Figure 6. A, Predicted secondary structure of RNA fragments of the COX1Δ13*DraI subclone. The 5' helix (helix 1), the base of the hairpins, and the 3' helix (helix 2) are indicated. Sequences belonging to the 78 nt. RNase T1 fragment of *COX1* that interacts with p40, are outlined. B, Gel mobility shift assay with these RNA molecules; lanes 1–3, ~3 ng of COX1Δ13*AseI; lanes 4–6, ~3 ng of COX1Δ35*DraI; lanes 7–8, ~2 ng of COX1Δ35*T1; lanes 9–10, ~2 ng of COX1Δ35*T1*CL3. RNAs were incubated without (lanes 1, 4, 7 and 9), with 65 ng (lanes 2 and 5) or 125 ng (lanes 3, 6, 8 and 10) of purified p40.

molecule was purified on a 12% polyacrylamide gel (7M urea) and is still predicted to form helix 2 and a 5' single stranded region (Figure 6A). As can be seen in Figure 6B this deletion has no effect on the affinity of the molecule for p40 (lane 8). In the second step we also removed the 5' 11 nt unpaired stretch of COX1Δ35*T1 with the C-specific RNase CL3 (COX1Δ35*T1*CL3). Removal of these bases completely abolishes the ability of the molecule to bind p40 (Figure 6B, lane 10). Since the only difference between COX1Δ35*T1 and COX1Δ35*T1*CL3 is the absence of the predicted single stranded stretch of nucleotides of the former molecule, the difference in p40 binding properties must be ascribed to this stretch. Therefore, we conclude that this single stranded region is an essential determinant in the formation of the recognition site for p40 in the wild type leader of *COX1* mRNA.

p40 binding exposes and protects nucleotides against DEPC modification in the leaders of *COX1* and *ATP9*

In order to identify the nucleotides directly involved in binding of p40, we performed diethylpyrocarbonate (DEPC) footprinting experiments. The reliability of such an experiment is heavily dependent on the size of the RNA molecule tested. Therefore we chose two small transcripts, that still retained the capacity to interact with p40, COX1Δ13*DraI (136 nt) and ATP9*AseI (92 nt, see Figure 3). Predicted secondary structures of both molecules are depicted in Figure 7B. As for COX1Δ13*DraI, the structure of ATP9*AseI was confirmed by nuclease protection experiments (data not shown). The computer predicted structure of this molecule was similar to those of larger *ATP9* leader transcripts shown to interact with p40 and, therefore, probably folds as the wild type configuration of this part of the *ATP9* leader. Purified RNA molecules were 5'-end labeled and, after incubation with purified p40, modified with DEPC (carboxyethylation at the N7-position of purines). The modified ribonucleotides were cleaved with aniline and separated on a 12% polyacrylamide gel (7M urea). Comparison of the modified

nucleotides of COX1Δ13*DraI and ATP9*AseI RNAs in the absence or presence of p40 yields both protection and exposure of nucleotides (Figure 7A). Protected bases in the COX1Δ13*DraI RNA surround the single stranded region shown to be essential for p40 binding (see previous section). With the ATP9*AseI molecule, protection predominates in the 3' arm, near or at the loop (Figure 7A and B). The most likely explanation for this observation is that the single stranded region in the COX1Δ13*DraI molecule and the loop in the ATP9*AseI molecule are directly involved in binding of p40 and the nucleotides are thereby protected from DEPC modification by the protein. Exposed nucleotides in the COX1Δ13*DraI RNA are mainly present at the loop of the 3' hairpin, 3' of the protected nucleotides. Exposure in the ATP9*AseI molecule is oriented 5' of the protected area, on the 5' arm of the molecule (Figure 7A and B). Surprisingly, the exposure of nucleotides is, in both molecules, much clearer than the protection. Exposure of nucleotides probably originates from structural changes in the RNA molecules and could, therefore, be functionally significant. It should be noted that the exposed areas in both *COX1* as in *ATP9* have a potential to base pair with a region around their respective start codons (Figure 7C). Whether this ability is significant, or coincidental, remains to be established.

DISCUSSION

In order to obtain a definition of the p40 binding site, we have compared primary and secondary structure motifs in the leaders of the *COX1*, *COX2* and *ATP9* mRNAs. By DEPC footprinting, site-directed mutagenesis and deletion analysis we have established that in the leaders of *COX2* [28], *COX1* and *ATP9* (this study) a single stranded region appears to be required for p40 binding, but no obvious sequence similarity exists in this region of the leaders. Since p40 binds to the leaders with high specificity, an unpaired non-conserved sequence cannot be the only determinant for this interaction. Closer examination of the

Figure 7. A, DEPC footprint analysis of 5'-end labeled COX1Δ13**Dra*I and ATP9**Ase*I. Lanes with and without added p40 are shown. Protected or exposed nucleotides and areas are indicated by triangles pointing down (protected) or up (exposed). B, Predicted secondary structure folding of these RNA molecules. Protected or exposed nucleotides are indicated by triangles pointing down (protected) or up (exposed). C, Potential base pairing between exposed areas of the COX1 and ATP9 leaders with their respective initiation codons. Exposed nucleotides are depicted with triangles on the lower strands, start codons for both COX1 and ATP9 are indicated by arrows on the upper strands.

three leaders reveals an helical region with conserved sequence in the immediate vicinity of the single stranded region (Figure 8). This suggests that p40 recognizes an helical region with conserved sequence 5'-UUUAUA-3' in combination with a stretch of unpaired nucleotides that is situated at the immediate 5' side of this sequence. The COX1Δ13**Ase*I molecule, which has residual binding activity (Figure 6B, lanes 2–3), contains a 'degenerate consensus' (G-U pair instead of A-U, Figure 8), that may be responsible for the marginal binding activity of the molecule. The secondary structures proposed have been confirmed by nuclease protection experiments on the constructs COX1Δ13**Dra*I and ATP9**Ase*I. Computer modelling of large constructs of the COX1 and ATP9 leader, that have retained p40 binding activity, predicts similar local structures (data not shown). These have yet to be confirmed experimentally.

Some of the sequence similarities may be coincidental: the A+U richness of the leaders of mitochondrial mRNAs results in an almost one in two chance that nucleotides in similar positions in two RNAs will have the same base. However, the essential proximity of a single stranded region renders the binding site more specific. Studies on autoregulatory ribosomal proteins of *E. coli* revealed that often their RNA binding sites are short helical segments with unusual bulge and loop features [23]. Since the A form RNA helix has a deep, narrow major groove, which may exclude protein binding, helix sequences appear important only at base pairs adjacent to bulges or loops, where the helix is probably distorted from its regular conformation [23]. The consensus presented here for the p40 binding site shows

Figure 8. A consensus model for p40 binding to the leaders of mitochondrial mRNAs. A conserved helix (boxed) is present in the vicinity of a single stranded loop or bulge. A perfect consensus is present in the COX1, COX2 and ATP9 leaders, deduced from the structures for COX1 and ATP9 presented in Figure 7B (COX1Δ13**Dra*I and ATP9**Ase*I, rotated for clearer alignment) and for COX2, slightly modified from the structure shown in Figure 7 of [28], without significant changes in ΔG. A degenerate consensus is present in COX1Δ13**Ase*I (Figure 6A), containing a G-U base pair instead of A-U, indicated with the arrow.

something similar. The essential single stranded region is likely to have a weakening effect on the stability of the helical conformation, rendering the nucleotides more accessible to protein contacts.

In both prokaryotes and eukaryotes, primary and secondary structure features within the 5' leaders of mRNAs are often involved in regulation of gene expression. Secondary structures normally have an inhibitory effect on translation initiation, either

by shielding the ribosome binding site and/or the start codon in prokaryotes (reviewed in [2]), or inhibition of scanning by the eukaryotic ribosome (reviewed in [20]). Regulatory proteins usually interact with specific RNA sequences located in the 5' untranslated region and thereby stabilize secondary structure and repress translation. Activators of translation have also been identified, but are less common. Nuclear encoded factors that activate translation of specific mitochondrial mRNAs have been found in *S.cerevisiae* and are thought to interact with specific sites in the 5' leaders of the mRNAs [5]. The protein identified as p40 ([28]; this study) is capable of interacting with the leaders of all mitochondrial mRNAs (Figure 3). The steady state concentration of mitochondrial mRNAs, in derepressed conditions, is approximately 100 to 500 copies per transcript per cell [35]. Since the concentration of p40 is at least 25,000 copies per cell [29], the protein has the capacity to complex all mRNA molecules present.

The abundance and general character of p40 suggests to us it is ill-suited for fine regulation of protein synthesis. The apparent lack of conservation of the location of the p40 binding site on the different leaders (Figure 4) does not, however, exclude a direct or indirect function in the translation of the messages. For example, the L10-L7/L12 ribosomal protein complex of *E.coli* represses translation of *rplJ* mRNA by binding to a site that is more than 100 bases upstream of the affected ribosome binding site [36]. In analogy to this, virtually nothing is known about the secondary and tertiary structure of mitochondrial leaders, so the distance in bases between the p40 binding site and the initiation codon may not reflect their span in the folded RNA.

We previously suggested that p40 could activate translation in some way by interacting with mitochondrial mRNAs [28]. In two instances, however, mRNAs lacking putative p40 binding sites have been shown to be translated in yeast mitochondria. The 5' two-third of the 5' untranslated region of *COX3* mRNA (as far as position -172) fused to the *COB* structural gene, lacks the 79nt RNase T1 fragment (-76 to +3) demonstrated to bind to p40 (Figure 5, lanes 4 and 8), but is still capable of directing translation of the *COB* gene [17]. This indicates that the p40 binding site on the *COX3* leader identified by us (Figure 4) is not essential for translation initiation. However, p40 might stimulate *COX3* mRNA translation by binding to the second site isolated by RNase T1 fractionation experiments (Figure 5, lane 4, 166 nt. fragment). If this is true, the function of the binding site nearer to the initiation codon is not clear.

Chimeric mRNAs containing the 5' third of the *COB* leader attached to the 3' half of the *ATP9* leader can direct translation on the *COB* or *ATP9* structural genes [37,38]. Both chimeric RNAs are stable and competent in translation initiation in the absence of the binding site for p40 in both the *ATP9* leader (located at the 5' end), and in the *COB* leader (located at the 3' end). This appears to indicate that p40 is not essential for stability or translation of mitochondrial transcripts. However, as for the *COX3* leader, more p40 binding sites could be present on the leaders of either *COB* or *ATP9* or both, and stimulate translation of the chimeric mRNAs. In addition, the secondary structure of the fusion transcripts are likely to deviate considerably from the wild type folding of either of their constituents. The possibility that alternative p40 binding sites might be formed in these transcripts cannot be formally ruled out.

An intriguing possibility is that the function of the p40-leader interaction could be the stabilization of certain structures in the leaders, that may have a negative influence on the translatability

of the mRNAs. We have noticed a marked increase in the stability of both the *COX1* Δ 13**DraI* and *ATP9***AseI* molecules under the influence of p40, assayed by nuclease protection (data not shown). In addition, the interaction of p40 with these molecules also resulted in the exposure to DEPC of an area that has the capacity to base pair with the respective initiation codons (Figure 7C). A structural change in the leaders, imposed by p40, possibly renders this area more accessible to base pair contacts. Closer inspection of the DEPC footprint on the leader of *COX2* (including the initiation codon) published earlier [28], reveals protection by p40 of nucleotides in the vicinity of the start codon. The base pair interaction may have already occurred here, shielding the start codon for ribosome recognition. If this base pairing is indeed a common phenomenon in all leaders, the initiation codons of all mRNAs could be shielded by a secondary or tertiary interaction imposed by p40, leading to an inhibition of protein synthesis.

Premature synthesis of the mitochondrially encoded hydrophobic proteins in the mitochondrial matrix could result in misfolding and degradation of the translation products. Association of p40 (which is located in the mitochondrial matrix, unpublished results) with all mRNAs could prevent premature protein synthesis in the matrix. This block in translation might only be reversible by interaction of this complex with the mRNA-specific activators, that are thought to act on the leaders of mitochondrial mRNAs [16,17]. Several of these *trans* acting factors are largely associated with the mitochondrial inner membrane [5,19] and two of them (PET122 and CBS2) interact with the small ribosomal subunit [18,19]. These factors have, therefore, the capacity to associate mRNA and ribosomes at the inner membrane of mitochondria, where the translation products can assemble into respiratory chain complexes. Translation of chimeric mRNAs lacking the p40 binding site would, in this model, only lead to a less economic protein synthesis, not to a block in translation. Some mRNAs might be prematurely translated in the matrix, because of the lack of translation inhibition, but others might still reach the specific translational activators at the inner membrane. Consequentially, the *trans*-activators would, in this case, still be required to target the mRNA and ribosome to the membrane. In this model, the presence of p40 ensures that the assembly of an initiation complex and translation of mitochondrial mRNAs only occur near the final destination of the protein, i.e. at the inner membrane.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. B.Papadopoulou for constructing and assaying some of the *COX1*, *COX3* and *COB* subclones, W.Mulder and Dr. M.Simitsopoulou for PCR amplification and cloning of the *ATP9* and *VAR1* leaders. We are grateful to Dr. A.Bednarsz for critical reading of the manuscript, and Dr. T.D.Fox (Cornell University, Ithaca, NY) and Dr. C.L.Dieckmann (University of Arizona, Tucson, AZ) for providing us with plasmids pSP65XAI and pBS/2-2 respectively.

REFERENCES

1. Grivell, L.A. (1989) *Eur. J. Biochem.*, **182**, 477-493.
2. Gold, L. (1988) *Ann. Rev. Biochem.*, **57**, 199-233.
3. Kozak, M. (1983) *Microbiol. Rev.*, **47**, 1-45.
4. Kozak, M. (1989) *J. Cell Biol.*, **108**, 229-241.
5. Costanzo, M.C. and Fox, T.D. (1990) *Ann. Rev. Genet.*, **24**, 91-113.
6. Fox, T.D. (1986) *Trends Genet.*, **2**, 97-100.

7. Ackerman, S.H., Gatti, D.L., Gellefors, P., Douglas, M.G. and Tzagoloff, A. (1991) *FEBS Lett.*, **278**, 234–238.
8. Costanzo, M.C. and Fox, T.D. (1986) *Mol. Cell. Biol.*, **6**, 3694–3703.
9. Costanzo, M.C., Seaver, E.C. and Fox, T.D. (1986) *EMBO J.*, **5**, 3637–3641.
10. Decoster, E., Simon, M., Hatat, D. and Faye, G. (1990) *Mol. Gen. Genet.*, **224**, 111–118.
11. Dieckmann, C.L. and Tzagoloff, A. (1985) *J. Biol. Chem.*, **260**, 1513–1520.
12. Kloeckener-Gruissem, B., McEwen, J.E. and Poyton, R.O. (1988) *J. Bacteriol.*, **170**, 1399–1402.
13. Payne, M.J., Schweizer, E. and Lukins, H.B. (1991) *Curr. Genet.*, **19**, 343–351.
14. Rödel, G. and Fox, T.D. (1987) *Mol. Gen. Genet.*, **206**, 45–50.
15. Poutre, C.G. and Fox, T.D. (1987) *Genetics*, **115**, 637–647.
16. Rödel, G. (1986) *Curr. Genet.*, **11**, 41–45.
17. Costanzo, M.C. and Fox, T.D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2677–2681.
18. Haffter, P., McMullin, T.W. and Fox, T.D. (1991) *Genetics*, **127**, 319–326.
19. Michaelis, U., Körte, A. and Rödel, G. (1991) *Mol. Gen. Genet.*, **230**, 117–185.
20. Kozak, M. (1991) *J. Biol. Chem.*, **266**, 19867–19870.
21. Schimmel, P. (1989) *Cell*, **58**, 9–12.
22. Jackson, R.J., Howell, M.T. and Kaminski, A. (1990) *Trends Bioch. Sci.*, **15**, 477–483.
23. Draper, D.E. (1989) *Trends Bioch. Sci.*, **14**, 335–338.
24. Nomura, M., Gourse, R. and Baughman, G. (1984) *Ann. Rev. Biochem.*, **53**, 75–118.
25. Theil, E.C. (1990) *J. Biol. Chem.*, **265**, 4771–4774.
26. Altuvia, S., Locker-Giladi, H., Koby, S., Ben-Nun, O. and Oppenheim, A.B. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6511–6515.
27. Wulczyn, F.G. and Kahmann, R. (1991) *Cell*, **65**, 259–269.
28. Papadopoulou, B., Dekker, P., Blom, J. and Grivell, L.A. (1990) *EMBO J.*, **9**, 4135–4143.
29. Dekker, P.J.T., Papadopoulou, B. and Grivell, L.A. (1991) *Biochimie*, **73**, 1487–1492.
30. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) In *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
31. Laird, P.W. (1988) Ph.D. thesis, University of Amsterdam.
32. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
33. Conway, L. and Wickens, M. (1989) *Meth. Enzym.*, **180**, 369–379.
34. Zuker, M. and Stiegler, P. (1981) *Nucleic Acids Res.*, **9**, 133–148.
35. Mueller, D.M. and Getz, G.S. (1986) *J. Biol. Chem.*, **261**, 11816–11822.
36. Christensen, T., Johnson, M., Fiil, N.P. and Friesen, J.D. (1984) *EMBO J.*, **3**, 1609–1612.
37. Dieckmann, C.L. and Mittelmeier, T.M. (1987) *Curr. Genet.*, **12**, 391–397.
38. Mittelmeier, T.M. and Dieckmann, C.L. (1990) *Curr. Genet.*, **18**, 421–428.