The EMBO Journal vol.9 no.12 pp.4135-4143, 1990

A 40 kd protein binds specifically to the 5'-untranslated regions of yeast mitochondrial mRNAs

B.Papadopoulou, P.Dekker, J.Blom and L.A.Grivell

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

Communicated by L.A.Grivell

Using a gel mobility shift assay we show that a 40 kd protein (p40), present in extracts of yeast mitochondria, binds specifically to the 5'-untranslated leader of cytochrome *c* oxidase subunit II mRNA. Binding of p40 to coxII RNA protects an 8-10 nucleotide segment from diethylpyrocarbonate modification, indicating that the protein interacts with only a restricted region of the 5'-leader. This segment is located at position -12 with respect to the initiation AUG. Deletion of 10 nucleotides encompassing this site completely abolishes protein binding. Nevertheless, Bal31 deletion analysis within the coxII leader shows that a major part of the leader is essential for p40 binding, suggesting that binding of the protein is also dependent on secondary structural features. p40 binds to other mitochondrial leader mRNAs including those for coxI, coxIII and cyt *b*. p40 is present in a cytoplasmic (rho^0) petite mutant lacking mitochondrial protein synthesis. It is therefore presumably nuclear encoded. The possible biological function of the protein is discussed.

Key words: mitochondrial mRNAs/5'-untranslated leaders/RNA-protein interaction/RNA secondary structure/mobility-shift

Introduction

mRNAs in yeast mitochondria have unusual properties. They are not capped at their 5'-end; they lack poly(A) tails and they often contain long, extremely AU-rich, untranslated 5'-leaders and 3'-trailer sequences (Groot *et al.*, 1974; Grivell, 1989). The initiator AUG is seldom the first in the mRNA and there is no obvious consensus sequence in the immediate surroundings that might distinguish this codon from other potential initiation sites. Genetic studies have shown that 5'-leader sequences of several mitochondrial transcripts play an important role in regulating initiation of translation. These studies have indicated that several nuclear encoded proteins act on site(s) within the 5'-leaders and positively regulate translation of specific mitochondrial transcripts (Fox, 1986). The most likely function for these proteins is that of mRNA specific initiation factors. Support for this idea is given by sequence similarity found between certain of these proteins and eukaryotic initiation factors (Grivell, 1989). Mutations within the 5'-untranslated leader of a mitochondrial mRNA have also been shown to affect its translatability (Ooi *et al.*, 1981), demonstrating the importance of these sequences for translation. So far, however, the limited number of changes studied hampers the formulation of any general conclusions regarding sequence or structure requirements essential for mRNA-protein interactions in yeast mitochondria.

Secondary structures in both prokaryotic and eukaryotic mRNAs influence both the efficiency of translation and the fidelity of initiation (Pelletier and Sonenberg, 1985; Kozak, 1986; Schimmel, 1989). Apart from the cap, however, little is known about which features of the 5'-regions of eukaryotic mRNAs are recognized by the translational machinery, since such regions are highly variable and apparently do not possess a uniform consensus sequence that regulates ribosome binding (Kozak, 1987). RNA molecules are capable of forming a variety of structures, including helices, loops, bulges, pseudoknots, that probably play an important role in the specific recognition and binding of proteins (Draper, 1989).

Our goal is to identify *trans*-acting factors that specifically interact with the 5'-leaders of mitochondrial transcripts and to study their potential role in regulating translation initiation. In this report we describe the purification and the RNA binding characteristics of a 40 kd protein (p40) that binds specifically to the 5'-untranslated sequences of coxII mRNA *in vitro*, thus providing the first direct biochemical evidence for a specific protein-RNA interaction in a yeast mitochondrial mRNA. The binding site for p40 has been located 12 residues upstream of the AUG start codon. The p40 protein also interacts with the 5'-leaders of other mitochondrial mRNAs suggesting that it is a general factor which may recognize a particular structural motif or a combination of sequence and structure within these leaders. As such, it clearly differs from the products of various PET genes characterized so far (Fox, 1986).

Results

Specific binding of a protein extracted from yeast mitochondria to coxII mRNA

To determine whether specific regions of the 5'-leader of coxII mRNA are recognized by proteins present in mitochondrial extracts, RNAs synthesized *in vitro* were incubated with mitochondrial extracts made from *Saccharomyces carlsbergensis*. Gel retardation assays were used to detect the products of specific RNA-protein interactions. After incubation of a 140 nucleotide long ^{32}P-labeled RNA transcribed from pEP40COII9, which contains the complete 5'-untranslated region and part of the coding sequence as far as the *Rsa*I site (Figure 1) with a protein fraction eluted from heparin-Sepharose at 0.2 M KCl, a distinct band migrating with slower mobility was observed (Figure 2A, lane 2). Similar experiments performed with the other fractions, eluted with different salt concentrations, did not result in any protein-RNA complex formation (data not shown).

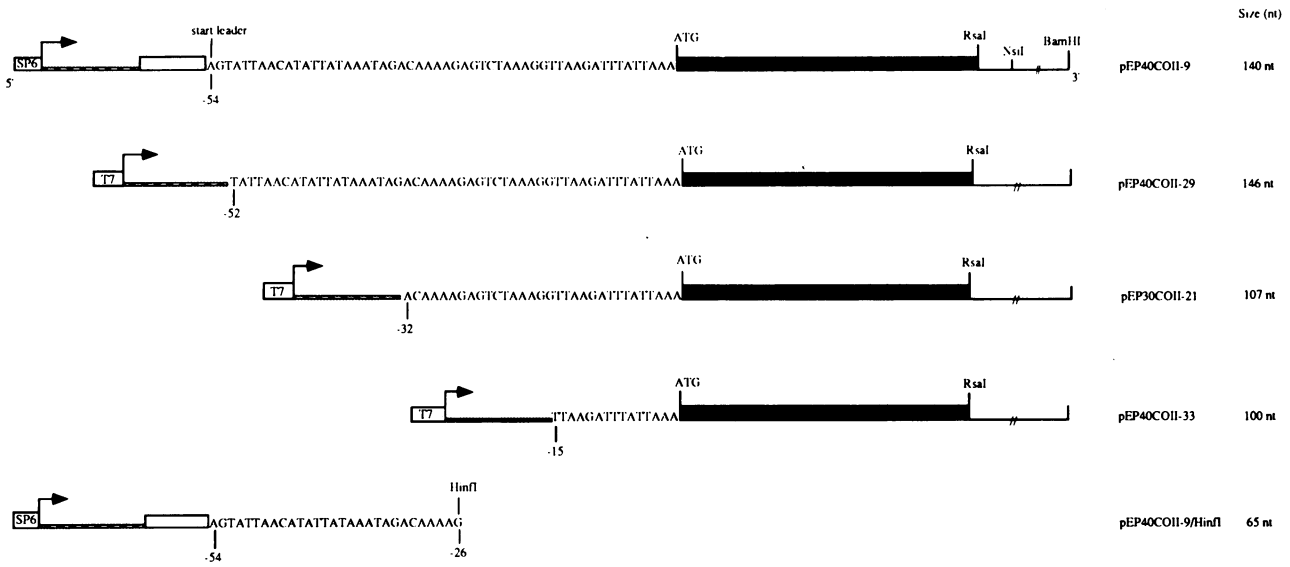

Fig. 1. DNA templates used for transcription of the coxII gene *in vitro*: right-angled arrows denote the start and direction of transcription; hatched lines indicate the polylinker sequences (25 nucleotides); the white bar indicates the 11 nucleotides upstream of the transcriptional start of the coxII mRNA in clone 9; ATG, initiation codon for coxII protein synthesis. The coxII 5'-leader is 54 nucleotide long and the sequence is shown here. The filled box represents the first 50 nucleotides of the coxII coding region as far as the *Rsa*I site. The *Bam*HI construct contains the complete coxII gene. *Rsa*I, *Nsi*I, *Bam*HI and *Hin*fI, sites used to linearize templates. The numbers denote the 5'-ends of the wild type and Bal31 deletion mutants. For further details of plasmids, see both Materials and methods section and text.

The specificity of the complex of coxII RNA with protein(s) in the 0.2 M KCl fraction was demonstrated by competition experiments with homologous or heterologous RNA sequences. An approximately 50-fold molar excess of each unlabeled competitor RNA was preincubated with the protein extracts prior to the addition of the labeled coxII probe. The effect of this preincubation on radiolabeled protein–RNA complex formation was then determined by gel retardation analysis. Addition of increasing amounts of specific competitor, i.e. pEP40COII-9 *Rsa*I or *Bam*HI RNAs to the reaction mixture resulted in a concentration dependent reduction of complex formation. (Figure 2A, lanes 3–5). Antisense coxII RNA sequences bind protein with ~10-fold lower affinity than sense coxII RNA sequences as shown by competition in lane 9. In contrast, preincubation with non-specific competitors such as *Escherichia coli* RNA, poly(A), (U), (G) and some yeast mitochondrial and cytosolic RNAs had no significant effect on the level of protein interaction with the coxII probe (Figure 2B, lanes 3–5 and 8–10).

We also examined the ability of heterologous ^{32}P-labeled RNA substrates to bind protein. As Figure 2C shows, RNA derived from the yeast mitochondrial intron aI5 (coxI gene) or plasmid pEP30 do not form specific complexes with the 0.2 M KCl protein fraction (lanes 2–3 and 5). From these experiments we conclude that there is a specific interaction between the protein(s), present in this fraction and that portion of the coxII mRNA that contains the 5'-untranslated leader and a small part of the coding region.

The coxII RNA–protein complex is rather stable, having a dissociation half-life of ~22 min as judged from the binding assays in which dissociation of the complex is followed after addition of an excess of specific cold competitor RNA (data not shown). The coxII RNA–protein complex can also be formed at 4°C. The formation of this complex requires at least 20 mM salt, but does not need Mg^{2+} (data not shown).

Purification of the coxII RNA binding protein

In order to characterize further the coxII RNA binding factor, 0.2 M KCl heparin–Sepharose extracts were further fractionated using HPLC gel filtration column. Binding activity was localized in a fraction eluted immediately after the void volume, containing two prominent proteins of ~38 and 40 kd (Figures 3A and C, lanes A). No other HPLC fractions were able to bind coxII sequences (Figure 3C). Attempts to separate these two proteins further on DEAE ion exchange chromatography and hydrophobic interaction columns failed. Proteolytic digestion of these two proteins, independently isolated from a SDS–PAGE, using V8 protease, papain and α-chymotrypsin revealed that they are closely related in sequence. The cleavage patterns obtained with papain are shown in Figure 3B (compare lanes 3 and 4). We suggest that the 40 kd band represents a modified version of the 38 kd band or alternatively that the 38 kd band is a breakdown product of the 40 kd protein. From hereon the protein that specifically binds to the coxII RNA will be referred to as p40.

To establish further the identity of the factor binding to the coxII transcript and to see whether additional proteins contribute to binding activity, a UV crosslinking experiment was performed. Radiolabeled pEP40COII9 RNA cut with *Rsa*I was incubated with the 0.2 M KCl heparin–Sepharose fraction followed by UV irradiation and RNase A digestion. Separation of the proteins covalently linked to the labeled coxII RNA by SDS–PAGE led again to the identification of a protein of ~40 kd as the prominent RNA binding protein (Figure 4b, lane 6). Use of a protein fraction eluted by 0.1 M KCl from heparin–Sepharose fraction in the assay did not result in the crosslinking of any detectable protein bands to the RNA (Figure 4a, lane 2). Similarly, a heterologous RNA i.e. pEP30 does not crosslink with the protein (lane 8). The UV crosslinking results demonstrate that only one protein has a high binding affinity and specificity for the coxII RNA. Complex formation does not

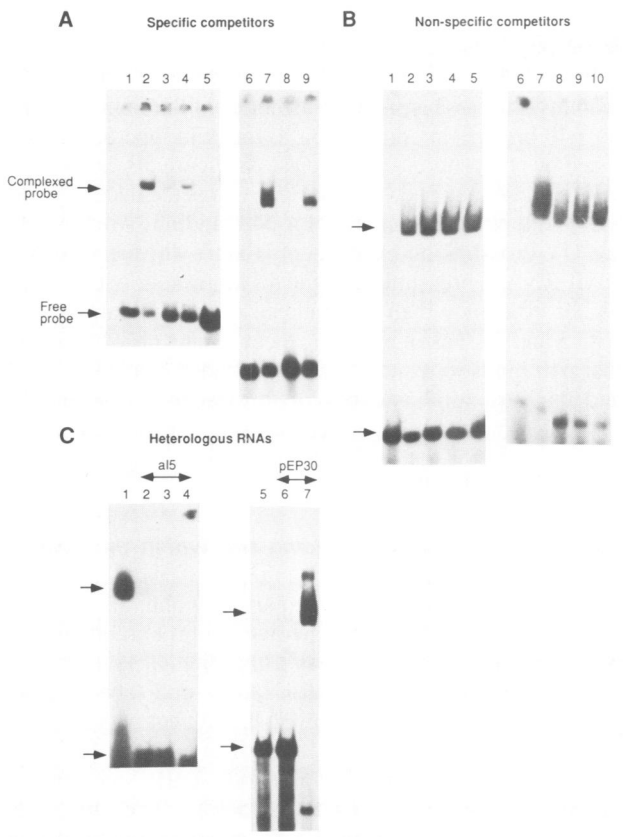

Fig. 2. Gel mobility shift and competition analysis of p40-coxII RNA binding *in vitro*. Binding assays were performed by using ^{32}P-labeled RNAs (5 ng; 10 000–15 000 c.p.m.) made *in vitro* by a T7 or SP6 transcription system, and various amounts (0.3–2 μg) of the mitochondrial protein extracts eluted with 0.2 M KCl from a heparin–Sepharose column as described in Materials and methods. Similar results were obtained using the purified 40 kd protein (data not shown). For competition experiments, an excess of cold RNA was added to the reaction prior to the addition of the labeled probe. (**A**) Mobility shift of coxII–p40 and competition using specific competitor RNAs. Lane 1 and 6, input probe pEP40COII9–*Rsa*I; lanes 2 and 7, pEP40COII9–*Rsa*I RNA with protein and without competition; lanes 3–5, competition with various concentrations of pEP40COII–*Rsa*I RNA (in order 30, 20 and 50-fold molar excess of competitor); lane 8, competition with pEP40COII9–*Bam*HI RNA (5′-leader with the whole coxII gene); lane 9, competition with coxII antisense RNA. (**B**) Competition with non-specific RNAs. Lanes 1 and 6, as A; lanes 2 and 7, as A; lane 3, competition with pEP30 RNA; lane 4, competition with aI5–*Dra*I RNA; lane 5, competition with 11k–*Mbo*II RNA; lanes 8–10, competition with poly(A), poly(U) and poly(G) respectively. (**C**) Gel mobility shift assay using radiolabeled heterologous RNA substrates. Lanes 1 and 7, as lane 2 in panel A; lanes 2–3, aI5–*Dra*I RNA (115 nt) with two different concentrations of protein; lane 4, free aI5–*Dra*I RNA; lane 5, pEP30–*Dde*I RNA (270 nt) with protein; lane 6, free pEP30–*Dde*I RNA.

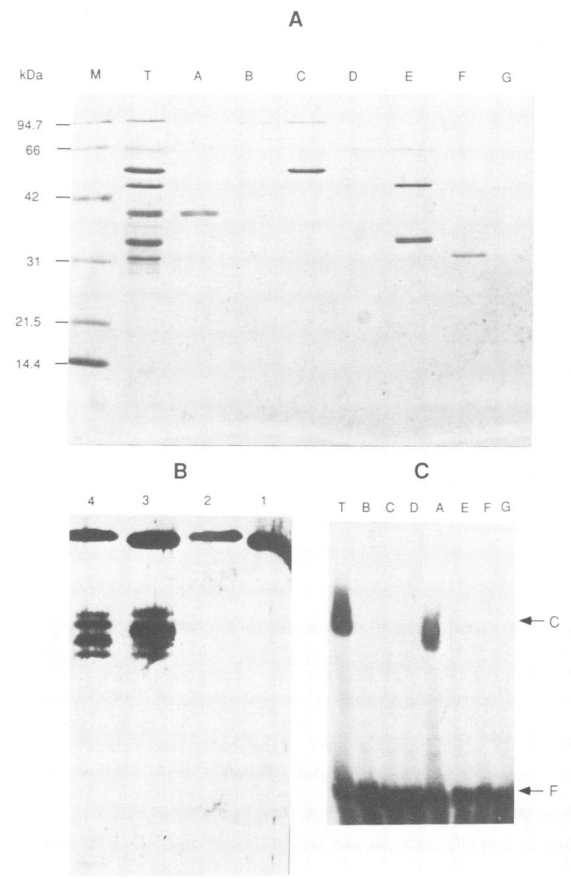

Fig. 3. Purification of the RNA binding protein from yeast mitochondria. (**A**) Fractionation by HPLC gel filtration. Lane T, protein fraction eluted from heparin–Sepharose column using 0.2 M KCl. Lanes A–G correspond to the different fractions obtained after HPLC TSK125 gel filtration of the 0.2 M heparin–Sepharose fraction separated on a 15% SDS polyacrylamide gel and stained with Coomassie blue; lane M, protein mol. wt markers in kd. (**B**) Peptide mapping of the two gel-excised proteins (38 and 40 kd) according to Cleveland *et al.* (1977). Lanes 1 and 2, untreated proteins; lanes 3 and 4, partial digestion profile of the two separate proteins obtained with papain. (**C**) Gel mobility shift assay with the 0.2 M KCl heparin–Sepharose (lane T) and the different HPLC protein fractions from A to G (lanes A–G). C, RNA–protein complex; F, free RNA.

require ATP hydrolysis since binding also occurs when ADPCP, a non-hydrolyzable analogue of ATP, is present in the reaction (Figure 4b, compare lanes 9 and 10).

To examine whether p40 is a mitochondrial or nuclear encoded protein we made mitochondrial extracts of a cytoplasmic (rho^0) petite yeast strain in which no mitochondrial protein synthesis occurs and performed a gel mobility shift assay using labeled coxII substrate. Since these extracts give rise to the same retarded coxII RNA band as detected with the wild type extracts (data not shown), we assume that p40 is indeed nuclear encoded. Whether p40 is exclusively located in mitochondria is as yet unclear, since the presence of high levels of RNase activity in other cellular fractions have so far precluded assay of binding activity.

p40 binds to the coxII 5′-untranslated region and requires the major part of the leader for its binding

To delineate which parts of the coxII transcript that yield the formation of a specific complex with p40, are required for protein binding we constructed a series of Bal31 mutants (Figure 1) and analyzed them by gel retardation assays. Removal of the first two nucleotides of the leader (pEP40COII29) did not have any effect on p40 binding (Figure 5A, lane 4). The pEP30COII21 mutant which lacks 22 nucleotides from the 5′-end of the leader shows an ~30 to 40-fold lower affinity for protein binding in comparison with the wild-type (Figure 5A, compare lanes 5 and 6 with 9). A deletion of 39 nucleotides from the 5′-end of the leader (pEP40COII33) abolishes p40 binding (Figure 5A, lane 11). When using an RNA substrate that contains the first part of the 5′-leader as far as the *Hin*fI site (position −26 within

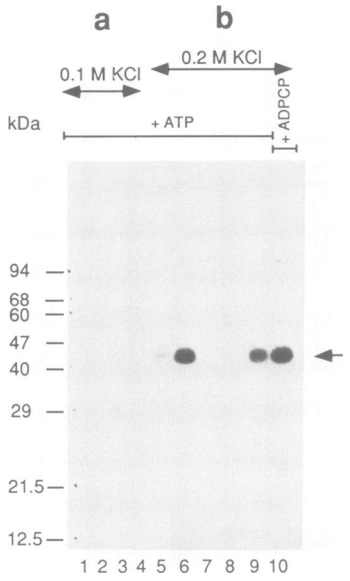

Fig. 4. UV crosslinking of p40 bound to 5′-coxII leader RNA. Crosslinking was carried out as described in Materials and methods (**a**) in the presence of 0.1 M KCl heparin−Sepharose protein fraction. Lanes 1−4 correspond to the ^{32}P-labeled RNAs pEP30COII21−*Rsa*I (mutated leader at the 5′-end), pEP40COII9−*Rsa*I (wild type), pEP40COII9−*Hin*fI (only 5′ part of the leader) and pEP30−*Dde*I respectively. (**b**) Lanes 5−8, RNA fragments in the same order as (a) (1−4) with the 0.2 M KCl heparin−Sepharose fraction which contain the p40 RNA binding protein; lane 9, pEP40COII9−*Rsa*I in the presence of ATP; lane 10, the same probe with ADPCP, a non-hydrolyzable analogue of ATP. The arrow indicates the crosslinked protein of ~40 kd.

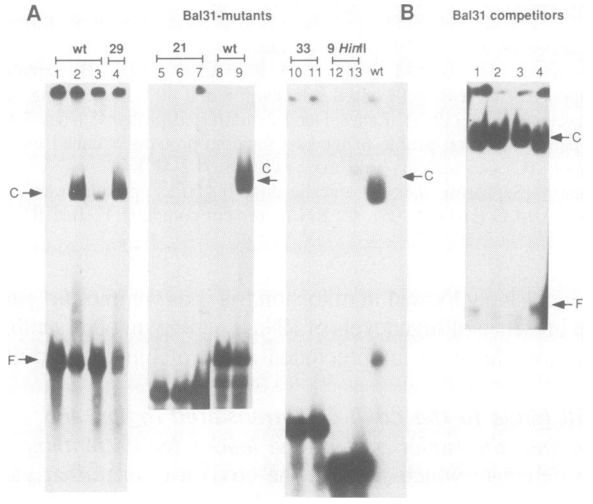

Fig. 5. Binding of p40 to truncated 5′-coxII leader mRNAs. (**A**) Gel mobility shift assays. Lanes 1 and 8, free pEP40COII9−*Rsa*I RNA (wild type); lanes 2 and 9, specific mobility shift with p40; lane 4, pEP40COII29−*Rsa*I RNA with p40; lanes 5 and 6, pEP30COII21−*Rsa*I with p40 in two different concentrations; lane 7, pEP30COII21−*Rsa*I free RNA; lane 10, pEP40COII33−*Rsa*I free RNA; lane 11, with p40, lane 12, pEP40COII9−*Hin*fI free RNA; lane 13, with p40. (**B**) Competition of coxII RNA−p40 specific complex by an excess of cold mutant RNAs. Lane 1, with pEP30COII21−*Rsa*I competitor RNA; lane 2, with pEP40COII9−*Hin*fI competitor; lane 3, with pEP40COII33−*Rsa*I competitor; lane 4, specific coxII RNA−p40 complex without competition.

the leader, Figure 1) we also observed a significant reduction of protein binding (lane 13).

In addition to the analysis of complex formation by mobility shift assays, relative affinities of the Bal31 mutants for p40 binding in comparison to the wild type RNA were tested by competition experiments. As Figure 5B (lanes 2−4) shows, these mutants were unable to compete effectively with the wild type coxII construct for p40 binding. As shown by the UV crosslinking experiment (Figure 4b) the wild type coxII RNA is crosslinked to the 40 kd protein with a much higher efficiency than the mutated ones (compare lanes 6 with 5 and 7). Taken together, these results clearly indicate that p40 binds to a site(s) in the 5′-leader of coxII mRNA and that structural integrity of the leader is essential for optimal protein binding since deletions either from the 5′ or the 3′ end of the leader dramatically affect the affinity of p40 to bind RNA.

Localization of the p40 binding site within the coxII 5′-leader sequence

To define more precisely the sequences within the 5′-leader of the coxII mRNA that are involved in protein binding we performed footprinting studies using diethylpyrocarbonate (DEPC) modification of the RNA. 5′-end labeled RNA produced by transcription of pEP40COII9, cut with *Rsa*I was first incubated with purified p40 and then modified by DEPC as described in Materials and methods. DEPC modifies A, G and U residues (A>G>>U). The modified ribonucleotides were cleaved with aniline and separated on a polyacrylamide gel. Figure 6 shows a comparison of the modification profile of coxII leader sequences under native conditions (30°C) in the presence (lane 1) and absence (lane 2) of p40. Four major differences can be seen. In the presence of p40, two regions are protected against DEPC modification, while two A residues at positions −10 and +18 display hypersensitivity to reaction. The first protected region consists of a 8−10 nucleotide fragment located between positions −14 and −21 with respect to the AUG codon. This fragment is situated just upstream of, and overlaps with, the first three nucleotides of a putative ribosome binding site (RBS) (5′-UUAAGAU-3′) complementary to the 3′-end of the 15S rRNA (Li *et al.*, 1982). Results of site-directed mutagenesis (see below) strongly suggest that this region indeed represents at least part of the binding site for p40. The second protected region is a 6 nucleotide fragment located between positions −30 and −35. Also this site may form part of the binding site of p40, since RNA folding studies (next section) indicate that it forms part of a characteristic stem−loop structure via complementarity to sequences immediately flanking the first protected region. The significance of the hyper-modified nucleotides at positions −10 and +18 is at present unknown.

The p40 binding site maps to a bulge structure within the 5′-coxII leader mRNA

To address the question of whether the binding site for p40 in the 5′-leader of coxII mRNA would correlate with a particular structure in the vicinity of the initiator AUG codon, studies on the secondary structure of the 5′-untranslated region of coxII mRNA were carried out. These involved computer analysis, chemical RNA modification and site-directed mutagenesis. For these studies the pEP40COII9−*Rsa*I construct including the whole 5′-leader

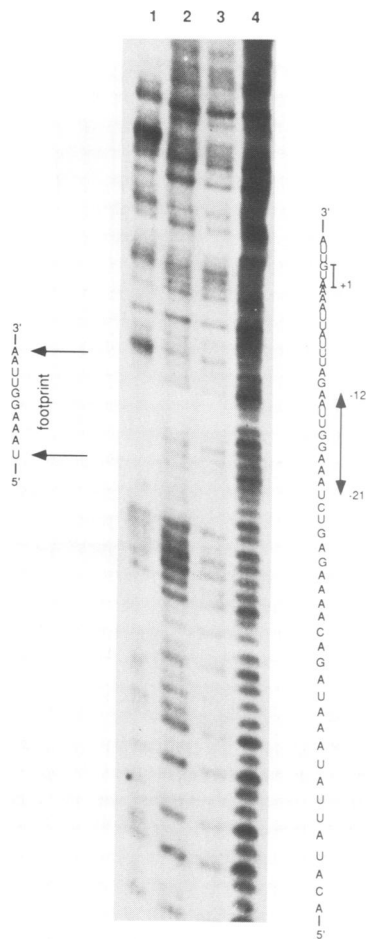

Fig. 6. Diethylpyrocarbonate (DEPC) modification footprint of p40 bound to the 5′-coxII leader mRNA. For the footprint assay p40 was mixed first with radiolabeled pEP40COII9−*Rsa*I RNA, incubated at 30°C for 30 min and then modified by DEPC as described in Materials and methods. Lane 1, RNA modified at 30°C in the presence of p40 protein. The 10 nucleotides protected from DEPC carboxyethylation lie between the arrows: lane 2, modified RNA at 30°C (native conditions); lane 3, modified RNA at 90°C (denaturing conditions); lane 4, RNA probe subjected to partial alkaline hydrolysis.

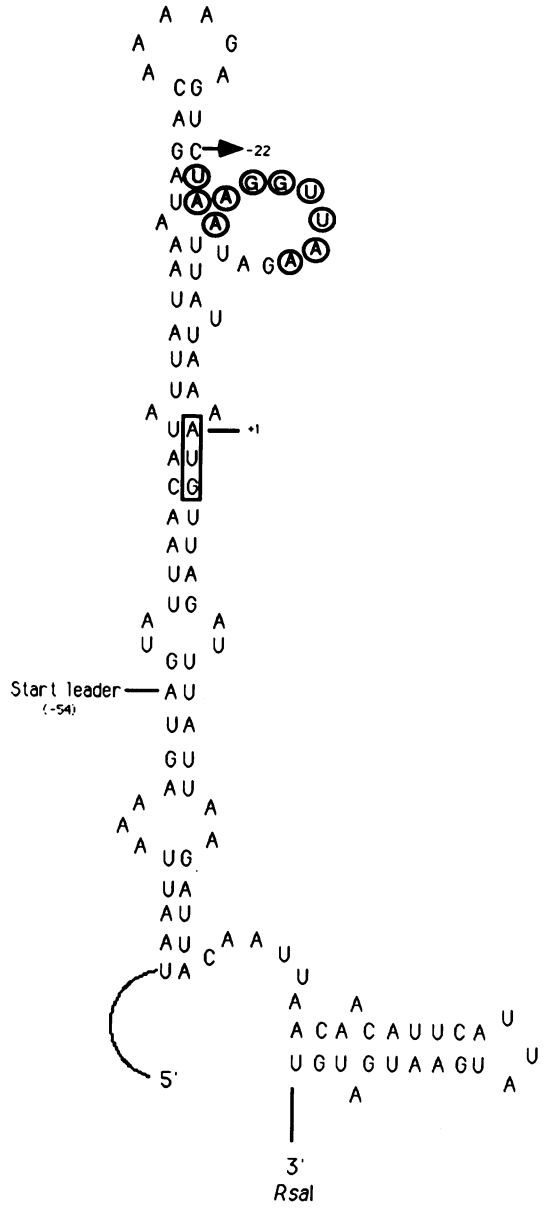

Fig. 7. Predicted secondary structure of the pEP40COII9−*Rsa*I RNA containing the 5′-coxII leader and part of the coxII coding region as far as the *Rsa*I site (140 nt), folded using the algorithm of Zuker and Stiegler (1981). The computer programme SQUIGGLES was used to generate this RNA structure. An estimate of the free energy of formation of this structure, calculated using the system of Frier *et al.* (1986), is approximately −13.7 kcal/mol. The binding decamer sequence (5′-UAAAGGUUAA-3′) revealed after footprint analysis is indicated by circles. The AUG is boxed. The arrow at position −22 indicates the only C residue protected against DMS modification.

and part of the coxII coding sequence as far as the *Rsa* site (140 nt) was used.

The RNA folding programme of Zuker and Stiegler (1981) revealed the existence of a rather unstable stem−loop structure followed by a bulge in the proximity of the AUG with a calculated ΔG of −13.7 kcal/ml (Figure 7). It may be significant that the footprint discussed in the previous section is in fact located in the bulge. Further studies using the most recent energy values for RNA foldings confirmed the Zuker structure and revealed in addition a large number of possible suboptimal RNA foldings (data not shown). Thus, the 5′-coxII leader may have the characteristics of a 'breathing' structure capable of adopting several foldings in solution. It is likely that a dynamic equilibrium between these alternative conformations exists under our experimental conditions.

These secondary structure predictions were supported by a partial secondary structure map of the pEP40COII9−*Nsi*I RNA by dimethylsulfate (DMS) modification technique and S1 nuclease digestion experiments. Only one C residue at position −22 within the coxII leader was protected from DMS modification (data not shown). This C to G base pairing contributes to the formation of the upper stem structure of the leader as predicted also by the computer data (see Figure 7). The information derived from such analysis is, however, not sufficient to draw any conclusions about the secondary structure of the coxII leader.

In order to characterize better the region of the coxII leader essential for p40 interaction, we introduced a 10 nucleotide deletion from −9 to −18 (5′-AAGGUUAAGA-3′) by site-directed mutagenesis in the region that was protected by footprint analysis (Figure 8A). Introduction of this deletion has not only an effect on the primary structure of the mRNA

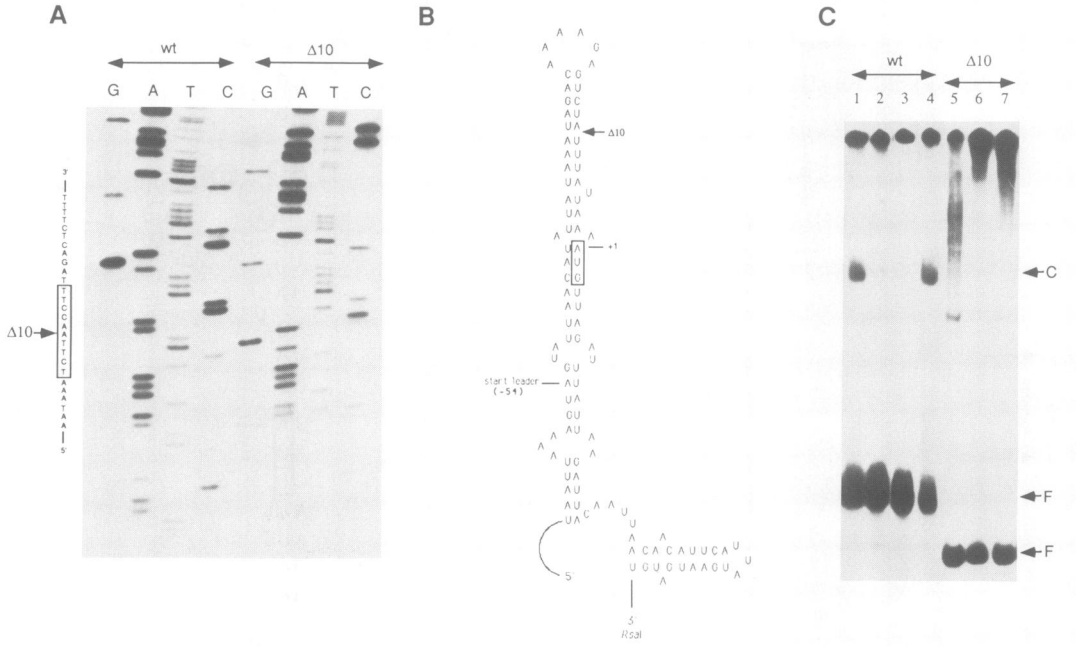

Fig. 8. Effects of a 10 nucleotide deletion within the 5′-coxII leader on its secondary structure and the binding affinity for p40. (**A**) Sequence analysis of the deletion mutant (Δ10) and comparison with the wild type. The deleted sequence is boxed. (**B**) Predicted secondary structure of the Δ10 mutant RNA (AAGGTTAAGA), deduced by the FOLD programme of Zuker (ΔG = −19.9 kcal/mol). The bulge structure and the protein binding site are removed. The arrow indicates the change introduced after mutagenesis. (**C**) Affinity of the Δ10 mutant RNA for p40 demonstrated by gel mobility-shift assays. Lane 1, competition with an excess of cold unlabeled Δ10 mutant RNA; lane 2, competition with the wild type pEP40COII9−*Rsa*I RNA; lane 3, free wt RNA; lane 4, pEP40COII9−*Rsa*I RNA with p40 protein; lane 5, free pEP40COII9 Δ10 mutant RNA; lanes 6 and 7, Δ10 mutant RNA with two different concentrations of p40.

by removing the footprint site but affects also its secondary structure, as predicted by computer studies (compare Figure 7 and Figure 8B). This deleted coxII 5′-leader was incapable of p40 binding, localizing sequences necessary for the recognition by p40 in the putative bulge structure of the leader (Figure 8C, lanes 6 and 7). Similarly, no competition of p40 binding to wild type coxII leader was observed with a 50-fold excess of cold deletion mutant RNA (Figure 8C, lane 1). Our results thus demonstrate that a putative bulge structure contains the binding site for p40 protein, but more mutagenesis and chemical modification experiments are required to determine precisely the primary and secondary structural features of that region of the coxII RNA involved in p40 recognition and binding.

p40 binds with high specificity to other mitochondrial 5′-leader mRNAs

We have asked whether p40 binds specifically to coxII mRNA or can interact generally with 5′-non-coding regions of mitochondrial mRNAs. To address this question we extended our studies to other 5′-untranslated leader sequences.

Results obtained from gel mobility shift assays demonstrated that the 40 kd protein can also bind to coxIII (489 nt) and coxI (538 nt) 5′-leader mRNAs (B.Papadopoulou, P.Dekker, K.van Oostrum and L.Grivell, unpublished observations). An excess of cold coxI, coxIII and cyt *b* 5′-leader mRNAs compete efficiently with the coxII leader RNA for binding of p40 (Figure 9, lanes a4, b3 and c4 respectively). These RNAs compete as well as the coxII leader for p40 binding, strongly indicating that the interaction between the protein and these sequences is specific. The antisense transcripts of the same leader RNA sequences which are

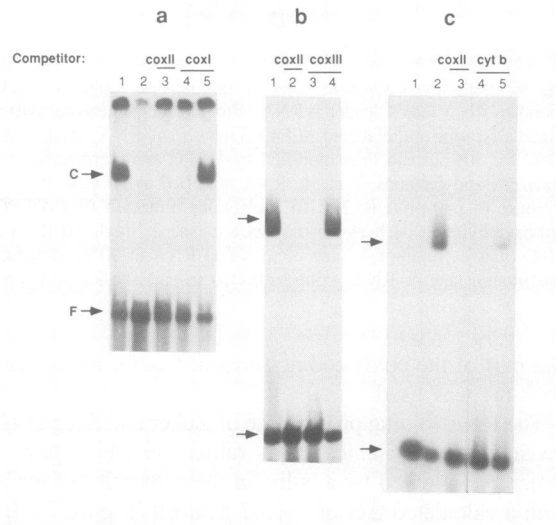

Fig. 9. Specific binding of p40 to the other mitochondrial 5′-leader mRNAs demonstrated by competition experiments. An excess of cold competitor RNA (50-fold molar excess) was mixed with p40 prior to the addition of coxII labeled probe. (**a**) Lane 1, specific complex between pEP40COII9−*Rsa*I RNA and p40; lane 2, free probe; lane 3, competition with coxII RNA; lane 4, competition with coxI RNA; lane 5, competition with the antisense coxI probe. (**b**) Lane 1, coxII−p40 complex; lane 2, competition with coxII RNA; lane 3, competition with coxIII RNA; lane 4, with antisense coxIII RNA. (**c**) Lane 1, free pEP40COII9 RNA, lane 2, with p40; lane 3, competition with coxII RNA; lane 4, competition with cyt *b* RNA. C: RNA−protein complex; F: free RNA.

expected to have extensive secondary structures, had a lower affinity for p40 binding than the sense 5′-leader mRNAs. (Figure 9, compare lanes a4 with a5, b3 with b4 and c4 with

c5), suggesting that p40 is highly sequence or secondary structure specific in its ability to bind RNA. Since all these leaders have different sizes (from 54 nt for coxII until 954 nt for cyt *b*) and can presumably fold in different secondary structures, additional studies are needed to elucidate the sequence/structure requirement for p40 binding. Footprint analysis within these leaders is in progress.

Discussion

We have used RNA mobility shift and competition assays to demonstrate the specific binding of a 40 kd protein, purified from yeast mitochondrial extracts, to the coxII 5′-leader RNA. This 40 kd protein binds to a restricted region (10 nt) of the coxII 5′-leader RNA that can potentially adopt a bulge structure. Deletion of this sequence completely abolishes protein binding. This decamer binding domain itself however, is not sufficient for p40 binding, as shown by Bal31 deletion studies within the 5′-coxII leader sequences. For example, the pEP30COII21 construct which contains the intact binding site, but lacks the first 22 nucleotides of the leader can no longer bind p40 efficiently (Figure 5A). Mutants lacking either the 5′ or the 3′ ends of the leader showed a drastically reduced binding affinity of p40 to the RNA. This suggests that the bulge predicted in Figure 7 can only form or stably exist if most of the leader sequences are present. Computer predictions indicate that all Bal31 generated mutants that failed to bind p40 had different secondary structures in comparison with the wild type leader (results not shown).

p40 binds specifically to all mitochondrial leader sequences so far tested i.e. coxI, coxII, coxIII and cyt *b* (B.Papadopoulou, unpublished results). This protein is highly selective in its ability to bind RNA since antisense transcripts made from these leaders have only a low affinity for the protein (Figure 9). Not all these leaders possess a decamer sequence identical to that revealed by footprint studies for the coxII leader, indicating that some sequence degeneracy may be tolerated. However, just as for the coxII leader, the integrity of the 5′-leaders is also important for p40 binding, since mutants produced by Bal31 deletion display only a very low affinity for the protein (B.Papadopoulou, unpublished results). We therefore cannot exclude the possibility that a secondary structure motif that is common to all yeast mitochondrial leader sequences is essential for p40 recognition and binding.

Primary and secondary structure features within the 5′-leaders of mRNAs are often associated with specific sites or regions of biological function such as translational initiation regions in several prokaryotic and eukaryotic systems. In prokaryotes, all ribosomal proteins and bacteriophage proteins known to be translational repressors bind directly to hairpin structures within the 5′-untranslated regions of the mRNAs (Thomas and Nomura, 1987; Romaniuk *et al.*, 1987). In eukaryotes, the iron-responsive hairpin structure in the 5′-non-coding region of ferritin mRNA is an example of a translational regulatory signal (Aziz and Munro, 1987); in *Drosophila* translation of the heat-shock *hsp22* mRNA is regulated by a region within 26 bases of the cap (Hultmark *et al.*, 1986).

We are currently investigating the possibility that p40 is a general initiation factor that regulates translation by binding to the 5′-untranslated regions of yeast mitochondrial mRNAs. Fox and co-workers have previously shown (Fox, 1986) that nuclear gene products (PET factors) act on the 5′-non-coding regions of specific mitochondrial mRNAs to promote translation. For example, specific activation of coxIII translation by PET54, 494 and 122 occurs by the action of these three gene products at a site or sites in a region of the 5′-leader at least 172 nt upstream of the initiation codon (Constanzo and Fox, 1988). Another mitochondrial translational activator, the product of *CBS1* gene, also acts on the 5′-leader of the cytochrome *b* mRNA and appears to be able to function at varying distances from the AUG codon (Roedel *et al.*, 1985). All these factors activate translation of specific mitochondrial transcripts while p40 seems to be a general factor which recognizes several leader sequences in yeast mitochondria. p40 does not display any similarity with the known PET factors as shown by Western blot experiments carried out with specific antibodies against these proteins (unpublished observations).

Mitochondrial mRNAs are not capped and most of them possess extremely long leaders with multiple AUG codons. It therefore seems reasonable to suppose that they are translated by a cap-independent mechanism. Internal ribosome binding has been shown to occur in the long 5′-untranslated leaders of several uncapped mRNAs of the picornavirus family (Pelletier and Sonenberg, 1988). p40 may be involved in a similar mechanism. In prokaryotes, ribosomes also bind internally by direct interaction with the Shine−Dalgarno sequence (Shine and Dalgarno, 1974). The yeast mitochondrial mRNAs lack the equivalent of a Shine−Dalgarno sequence, but Li and co-workers (1982) have proposed a putative ribosome binding site (RBS) capable of base pairing with at least seven consecutive nucleotides of the 3′-end of the yeast 15S rRNA, present in the 5′-leaders of mitochondrial mRNAs at variable distance from the initiation codon. The finding that the binding domain for p40 in the coxII leader is located just upstream of a putative RBS and at the proximity of the AUG codon, makes it possible that this protein is involved in ribosome binding. Thus, p40 could facilitate the binding of ribosomal subunits or initiation factors in the vicinity of the AUG codon by changing the structure of the leader in that region, for example.

Finally, p40 could be involved in mRNA stability. Sequences with the potential to form a stem−loop structure, localized more often at the 3′-end of both bacterial and eukaryotic cytoplasmic mRNAs appear to protect mRNA chains against exonucleolytic attack. Sequences at the 5′-end of mRNA may also be involved in the decay process (Brawerman, 1987). *CBP1*, one of the best characterized nuclear genes affecting mRNA stability in yeast mitochondria, has a target site in the 5′-leader of cyt *b* mRNA (Dieckmann and Mittelmeir, 1987).

The techniques that we have developed and used successfully for the analysis of the specific interaction between the p40 and several 5′-leaders of mitochondrial mRNAs should be applicable to other factors that might interact with such sequences. Further studies of *trans*-acting factors capable of specific interactions with 5′-RNA sequences should thus provide us with a better understanding of the molecular basis of the mechanisms that control translational initiation in yeast mitochondria.

Materials and methods

Plasmid constructions and strains

All plasmids were constructed by standard techniques (Maniatis *et al.*, 1982). To generate plasmids pEP40COII9, pEP40COII29, pEP40COII33 and pEP30COII21, the pMT36 construct (a 2.5 kb *Hae*III fragment containing the entire coxII gene of *Saccharomyces cerevisiae* strain D273-10B cloned in pBR322, kindly provided by Dr T.D.Fox, Cornell University, NY), was digested with *Mnl*I and the 1.16 kb *Mnl*I isolated fragment containing the coxII gene sequences was treated with Bal31 exonuclease. The blunt-ended Bal31 generated fragments were subcloned afterwards either into the *Sma*I site of the pEP30 or pEP40 vectors (Laird, 1988) in the + orientation relative to the T7 promoter for pEP30COII21, pEP40COII33, pEP40COII29 and SP6 for pEP40COII9. The orientation of the coxII fragments was verified by sequence analysis using the SP6 or T7 primers (Promega). Plasmids used to make coxI and coxIII transcripts were pEP40COI28 containing the complete coxI leader cloned in the pEP40 expression vector and pSP65XAI (a generous gift from Dr T.D.Fox) containing the whole coxIII leader and a small part of the coding sequence until the *Acc*I site, respectively. The pBS/2-2 plasmid, was provided by Dr C.Dieckmann (University of Arizona). It contains a 1.17 kb fragment including the cytochrome *b* leader and part of the first exon cloned in the *Eco*RI−*Bam*HI sites of pBS vector. The construct pEP30-al5 contains part of the coxI intron from yeast mitochondria and the pEP30-11k corresponds to the whole leader and part of the coding sequence of the 11k yeast cytosolic mRNA.

The yeast *S.carlsbergensis* strain NCYC-74 was used to prepare mitochondrial extracts. *E.coli* HB101 and JM101 strains were used for transformation and mutagenesis experiments.

In vitro transcription

SP6 and T7 coxII transcripts from wild type and Bal31 mutants were made for pEP30COII and pEP40COII constructs linearized at the *Rsa*I site, 50 nucleotides downstream of the ATG codon according to standard procedures (Melton *et al.*, 1984). Transcription reaction mixtures were treated with RNase-free DNase I in the presence of RNasin (10 U) (Promega). Unincorporated nucleotides were removed by purification over a 1 ml Sephadex-G50 column. Yields of transcripts were calculated from the incorporation of [α-^{32}P]UTP into RNA. All RNAs were analyzed for integrity on 6% polyacrylamide−7 M urea gels and visualized by autoradiography. The antisense RNAs were prepared by transcription with T7 or SP6 RNA polymerases after digestion of the relevant constructs with the appropriate enzymes.

Extract preparation and protein purification

Mitochondria were isolated by a modified procedure of Grivell *et al.* (1971). Yeast cells were grown at 28°C for 20 h in lactate medium (1.5% lactic acid, 2% sodium lactate, 0.1% glucose, 8 mM $MgSO_4.7H_2O$, 45 mM $(NH_4)_2HPO_4$ and 0.5% yeast extract, pH 4.5). Cells (200 g wet weight) were washed twice with water, resuspended in 5 ml/g wet weight of 0.1 M Tris−HCl pH 9.0 and 2.5 mM dithiothreitol and incubated at 30°C for 30 min. They were washed again with water followed by suspension in 2 ml of spheroplasting buffer (1.35 M sorbitol, 0.1 mM EDTA, pH 7.4) per g of cells. Zymolyase 100 000 (0.3 mg/g cells) was added and the preparation was incubated at 30°C with gentle shaking until spheroplasts formed (about 2 h). Spheroplasts were then washed with 1.35 M sorbitol, suspended in 2 ml/g of cells of ice-cold breaking buffer (0.6 M mannitol, 0.5 mM EDTA, pH 6.7) and lysed in a Sorval Omnimixer homogenizer by agitation for 2 × 30 s. Cell debris were removed by centrifugation of the lysate at 4000 r.p.m. for 5−10 min. Mitochrondria were pelleted by several low (4000 r.p.m.) and high speed (12 000 r.p.m.) centrifugations, suspended in a small volume of breaking buffer and stored at −70°C. For mitochrondrial extract preparation, frozen mitochrondrial pellets (2 g wet weight) were suspended in 10 ml of buffer A (1 M KCl, 20 mM Tris−HCl pH 7.5, 0.1 mM EDTA, 1 mM PMSF and 6 mM β-mercaptoethanol) and broken by sonication 4 × 30 s. After centrifugation at 16 000 r.p.m. for 25 min, the supernatant was recovered as the S30 fraction and the ribosomes were removed by further centrifugation at 4°C and at 55 000 r.p.m. for 2 h 30 in a Ti-65 Beckman rotor. After dialysis against buffer B (20% glycerol, 20 mM KCl, 20 mM Tris−HCl pH 7.2, 0.1 mM EDTA, 1 mM PMSF and 6 mM β-mercaptoethanol) the supernatant was applied to a heparin−Sepharose column CL-6B (Pharmacia) which had been equilibrated with buffer B and proteins were eluted using a step gradient of 20-100-200-300-500 and 1000 mM KCl. Fractions were dialyzed against buffer B, concentrated ~10-fold by ultrafiltration and stored at −70°C. Protein concentration were determined by the method of Bradford (1976) in combination with UV absorbance measurement. The 0.2 M KCl heparin−Sepharose protein fraction containing the RNA binding factor (p40) was separated then by HPLC TKS 125 (LKB) gel filtration in 10 mM NaPi pH 7.5 and 100 mM Na_2SO_4. This resulted in the highly purified p40 protein, eluted in fraction A (see Figure 3A).

Analysis of RNA−proteins interactions

The protocol developed for the analysis of p40 binding to the mitochondrial leader mRNAs is based on the method used by Dorsman *et al.* (1988) for the study of DNA−protein interactions. A standard RNA binding experiment was performed by mixing 100−200 ng of purified p40 protein with 5 ng of [α-^{32}P]UTP labeled RNA (10 000−15 000 c.p.m.) in a final volume of 30 μl in 10 mM Tris−HCl pH 7.5, 0.1 mM EDTA, 7 mM β-mercaptoethanol, 0.8 mM PMSF, 3 mM $MgCl_2$, 50 mM NaCl and 5% glycerol and incubated for 15 min on ice or at 30°C. *E.coli* competitor RNA (1 μg) was added in the reaction mixture in order to minimize the non-specific association of p40 with RNA. Increasing the incubation period had no effect on the level of complex formation (data not shown). For competition assays, a 10 to 100-fold molar excess of unlabeled RNA probe was preincubated with the protein for 15 min at 30°C prior to the addition of the specific labeled probe and further incubation for 15 min. After incubation, the p40−RNA complexes were resolved by electrophoresis through 4% non-denaturing polyacrylamide (acrylamide:bisacrylamide ratio of 40:1) gels run at 180 V for 2 h at 4°C using 1 × TBE as the gel and running buffer. The gels were then fixed (10% acetic acid, 10% methanol), dried and visualized by autoradiography.

RNA−protein crosslinking by UV light

RNA probe (15−20 ng/50 000 c.p.m./μl) was incubated with 5 μg of protein fractions from heparin−sepharose as described for gel retardation experiments except that 2 mM of ATP or 1 mM of ADPCP (a non-hydrolyzable analogue of ATP) were added separately to the reactions. After 10 min incubation at 30°C the reaction mixtures were irradiated 30 cm below a UV light (256 nm) (UV dose = 180 erg/mm^2, s, or 1.8 J) for 15 min at 4°C in open Petri-plates. Unprotected RNA was digested for 15 min at 37°C with 0.1 μg of RNase A, and the crosslinked protein−RNA complexes were analyzed by electrophoresis on 12.5% SDS−polyacrylamide gels according to Laemmli (1970). The gels were stained, dried and autoradiographed.

RNA footprinting analysis

Unlabeled pEP40COII9−*Rsa*I RNA template (1 μg) was dephosphorylated by calf intestinal phosphatase and 5′-end labeled with 50 μCi of [γ-^{32}P]ATP and T4 polynucleotide kinase (10 U). The 5′-end labeled RNA was gel-purified (6% polyacrylamide−7 M urea) and mixed with an excess of the p40 protein in the binding solution (without Tris−HCl) and 10 μg of carrier *E.coli* RNA for 15 min at 30°C. 2 μl of diethylpyrocarbonate (DEPC) was then added in 100 μl of 50 mM sodium cacodylate (pH 7.5) in the presence of 10 mM $MgCl_2$ and followed by further incubation for 30 min at 37°C. Only the unprotected RNA regions are exposed to modification by DEPC. The RNA was phenolized and recovered by ethanol precipitation. The modified RNA molecules were then cleaved by 1 M aniline in 20 μl of sodium acetate (pH 4.5) at 60°C for 20 min to induce chain scission. Two *n*-butanol extractions were performed to remove the aniline and the final pellet was washed with ethanol, dried and resuspended in 5 μl sterile water and 5 μl sample mix (0.03% bromophenol blue, 0.03% xylene cyanol, 25 mM EDTA in formamide). Samples were then boiled for 2 min and loaded onto a 7% polyacrylamide−8 M urea gel. Gels were then dried and autoradiographed.

RNA secondary structure analysis

RNA secondary structure was generated using the computer algorithm (FOLD) of Zuker and Stiegler (1981). The graphic representation of the secondary structure was produced using the SQUIGGLES programme. The energies used by Zuker's programme were defined by Freier *et al.* (1986).

Site-directed mutagenesis

Directed mutagenesis was performed in the M13mp10COII9 wild type construct using the bacteriophage M13 mutagenesis system according to Zoller and Smith (1987). Oligonucleotides were synthesized on a Milliger Biosearch Cyclone-plus programmable DNA synthesizer (Isogen Bioscience, Amsterdam). The nucleotide sequence 5′-CTAACATTTTAATAAATAG-ACTCTTTTGTC-3′ of the 30mer oligonucleotide is complementary to nucleotides +7 and −9 and −20 to −33 of the coxII 5′-untranslated leader. The mutation was designed to remove 10 nucleotides within the bulge structure present in coxII 5′-leader. The precise mutation was confirmed by double-strand dideoxynucleotide chain termination sequencing using the Sequenase system. The mutated coxII 5′-leader sequences were subcloned back into the *Eco*RI−*Bam*HI sites of the pEP40 expression vector in the + orientation relative to the SP6 promoter.

Acknowledgements

We wish to thank Dr T.D.Fox (Cornell University, Ithaca, NY) and Dr C.L.Dieckmann (University of Arizona, Tucson, AZ) for providing us with pMT36, pSP65XAI and pBS/2-2 plasmids respectively. We are grateful to Dr D.A.M.Konings (NIH, Frederick, MA) for performing secondary structure studies. H.van Heugten (Rijksuniversiteit Utrecht) for his helpful collaboration in the UV crosslinking experiment, M.de Haan for his contribution in the Bal31 analysis and K.van Oostrum for technical assistance. We also thank Dr M.Meijer for critical reading of the manuscript. B.P. was a postdoctoral fellow of the EEC Biotechnology Action Programme (BAP). This work was supported in part by grants from the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (NWO).

References

Aziz,N. and Munro,H.N. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8478–8482.
Brawerman,G. (1987) *Cell*, **48**, 5–6.
Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
Cleveland,D.W., Fischer,S.G., Kirschner,N.W. and Laemmli,U.K. (1977) *J. Biol. Chem.*, **252**, 1102–1106.
Constanzo,M.C. and Fox,T.D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2677–2681.
Dieckmann,C.L. and Mittelmeier,T.M. (1987) *Curr. Genet.*, **12**, 391–397.
Dorsman,J.C., van Heeswijk,W.C. and Grivell,L.A. (1988) *Nucleic Acids Res.* **16**, 7287–7301.
Draper,D.E. (1989) *Trends Biochem. Sci.*, **14**, 335–338.
Fox,T.D. (1986) *Trends Genet.*, **2**, 97–100.
Freier,S.M., Kierzek,R., Jaeger,J.A., Sugimoto,N., Caruthers,M.H., Neilson,T. and Turner,D.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9373–9377.
Grivell,L.A. (1989) *Eur. J. Biochem.*, **182**, 477–493.
Grivell,L.A., Reijnders,L. and Borst,P. (1971) *Biochim. Biophys. Acta*, **247**, 91–103.
Groot,G.S.P., Flavell,R.A., Ommen,G.J.B.van and Grivell,L.A. (1974) *Nature*, **252**, 167–169.
Hultmark,D., Klemenz,R. and Gehring,W. (1986) *Cell*, **44**, 429–438.
Kozak,M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2850–2854.
Kozak,M. (1987) *Nucleic Acids Res.* **15**, 8125–8148.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Laird,P.W. (1988) Ph.D. thesis. University of Amsterdam.
Li,M., Tzagoloff,A., Underbrink-Lyon,K. and Martin,N.C. (1982) *J. Biol. Chem.*, **257**, 5921–5928.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) In *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.* **12**, 7035–7056.
Ooi,B.G., Lukins,H.B., Linnane,A.W. and Nagley,P. (1981) *Nucleic Acids Res.* **15**, 1965–1977.
Pelletier,J. and Sonenberg,N. (1985) *Cell*, **40**, 515–526.
Pelletier,J. and Sonenberg,N. (1988) *Nature*, **334**, 320–325.
Roedel,G., Koerte,A. and Kandewitz,F. (1985) *Curr. Genet.*, **9**, 641–648.
Romaniuk,P.J., Lowary,P., Wu,H.N., Stormo,G. and Uhlenbeck,O.C. (1987) *Biochemistry*, **26**, 1563–1568.
Schimmel,P. (1989) *Cell*, **58**, 9–12.
Shine,J. and Dalgarno,L. (1974) *Proc. Natl. Acad. Sci. USA*, **74**, 1342–1346.
Thomas,M.S. and Nomura,M. (1987) *Nucleic Acids Res.* **15**, 3085–3096.
Zoller,M.J. and Smith,M. (1987) *Methods in Enzymology*, **154**, 329.
Zuker,M. and Stiegler,P. (1981) *Nucleic Acids Res.* **9**, 133–148.

Received June 21, 1990; revised September 10, 1990