Nucleic Acids Research, 1997, Vol. 25, No. 20 **4055–4060**

A conserved core element is functionally important for maize mitochondrial promoter activity *in vitro*

Angelina G. F. S. Caoile and David B. Stern*

Boyce Thompson Institute for Plant Research, Cornell University, Tower Road, Ithaca, NY 14853-1801, USA

Received June 30, 1997; Revised and Accepted August 26, 1997

ABSTRACT

We have previously used a homologous *in vitro* transcription system to define functional elements of the maize mitochondrial *atpA* promoter. These elements comprise a central domain extending from –7 to +5, relative to the transcription start site, and an upstream domain of 1–3 bp that is purine rich and centered around positions –11 to –12. Within the central domain lies an essential 5 bp core element. These elements are conserved in many mitochondrial promoters, but their functionality has only been tested for *atpA*. In this study we have introduced mutations into the corresponding elements of two *cox3* promoters and show that while the core element is essential for *cox3* promoter activity, upstream element mutations have little or no effect. To define the minimal sequence required for *in vitro* promoter activity a series of short cloned oligonucleotides corresponding to the *atpA* promoter was used. While some activity was seen with a 14 bp sequence, full activity required 26 bp, suggesting that elements other than the core and upstream region can influence promoter strength. Another series of clones showed that altered spacing between the upstream and core elements of *atpA* had a significant effect on promoter activity. These results further define important features of the plant mitochondrial transcriptional machinery.

INTRODUCTION

Plant mitochondrial genomes are multipartite and complex, yet contain a relatively small number of genes encoding mainly components of the translational apparatus and the electron transport chain. In the *Arabidopsis* mitochondrial genome 57 genes are dispersed throughout 367 kb of sequence (1). This contrasts strongly with the compact mitochondrial genomes of animals and fungi. Like genome size and content, the transcriptional strategies of plant mitochondria are distinct. Rather than having a single promoter for each strand, like mammals, or a small number of nearly identical promoters, like *Saccharomyces cerevisiae*, plant mitochondria feature a degenerate promoter sequence which is poorly conserved between monocotyledonous and dicotyledenous plants and which contains only a short consensus sequence, which is itself not universal, especially for rRNA and tRNA genes (2–7; reviewed in 8). In addition, many plant mitochondrial genes have multiple promoters (see for example 4,9), leading to complex mRNA patterns. These aspects of plant mitochondrial transcription raise the question as to how closely the plant transcriptional machinery resembles its counterparts in other organisms and whether unifying features of promoter structure can be found.

Plant mitochondrial promoter structure has been deduced from inspection of sequences surrounding known initiation sites and from *in vitro* transcription. Sequence comparisons showed that weak consensus motifs could be derived for maize (10), *Oenothera* (11), rice (12), sorghum (7), soybean (13) and wheat (14), with the most common element being YRTA (Y = C or T; R = A or G) found immediately upstream of the initiation site. To functionally test promoter elements *in vitro* transcription systems were developed from wheat (15), maize (16) and pea (17) mitochondria. In the case of maize, mutagenesis of the *atpA* promoter revealed that the YRTA 'core' sequence was indeed an essential element and that other sequences both upstream and downstream of the core influenced, but were not absolutely required for, promoter activity *in vitro* (18). These other sequences included a purine-rich motif ~10 bp upstream of the core and other bases surrounding the core in a central domain. The pea extract was shown to transcribe both a mRNA and tRNA promoter, as well as the *atp9* promoter of soybean, suggesting that transcription initiation mechanisms can be conserved across species boundaries in spite of low primary sequence homology. In addition, it was shown that the pea *atp9* promoter contained an essential element between –7 and –25 relative to the transcription initiation site, upstream of the core sequence. This promoter contains a purine-rich domain which could correspond to the upstream domain of maize *atpA*.

Apart from the examples of maize *atpA* and pea *atp9*, no other mutagenic analyses of plant mitochondrial promoters have been reported. This raises the question of whether conserved sequence elements are indeed functional in each promoter in which they are found. Although the core, central domain and upstream purine-rich domain which had been demonstrated to form parts of the maize mitochondrial *atpA* promoter are also found in other maize mitochondrial promoters, for example those of *cox3* (9,10), they have not been functionally analyzed. Here we report that the core, but not the upstream elements, of two maize *cox3* promoters are required for *in vitro* transcription initiation. We also use oligonucleotide sequences to define a minimal promoter for maize *in vitro* transcription.

*To whom correspondence should be addressed. Tel: +1 607 254 1306; Fax: +1 607 254 1242; Email: ds28@cornell.edu

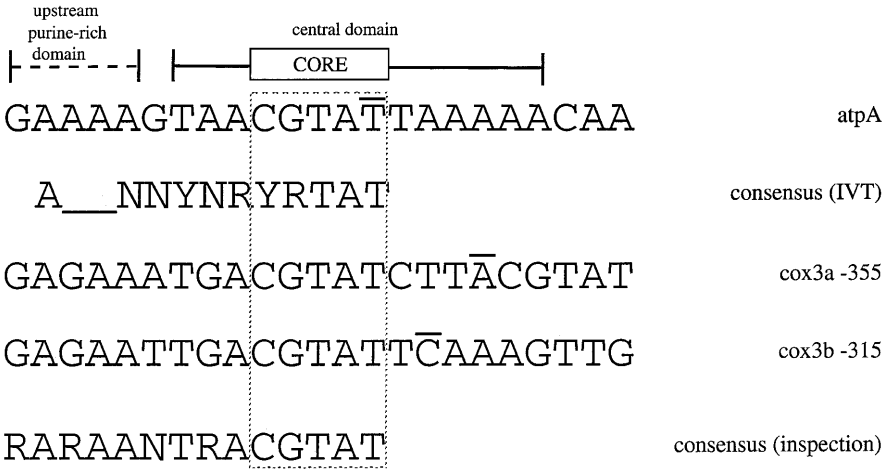

Figure 1. Maize mitochondrial promoter structure. The upper line shows the domains of the *atpA* promoter as determined by *in vitro* analysis, while the second line shows a consensus sequence based on mutagenesis (16,18). These sequences are aligned with two *cox3* promoters and at the bottom a consensus sequence that can be derived by inspection of the *cox3* and *atpA* promoters. The initiation position in each promoter is overlined.

MATERIALS AND METHODS

Transcription extracts and *in vitro* transcription

Mitochondrial protein extracts and *in vitro* transcription assays were carried out as previously described (18), except that the final chromatography step was through a Pharmacia 1 ml Resource-Q column. Templates for *in vitro* transcription were linearized plasmids or plasmid inserts gel purified using GeneClean II and a GeneSpin column (Bio 101, La Jolla, CA).

Template DNA

Cloning and site-directed mutagenesis were carried out by standard procedures. A plasmid containing the *cox3* promoters (9) was obtained from Dr R.M.Mulligan (University of California, Irvine, CA). Minimal promoter constructs were made by annealing complementary oligonucleotides with the sequences shown in the figures, repairing them with the Klenow fragment of DNA polymerase and ligating them into pBluescript linearized with *Eco*RV or *Sma*I. All constructs were verified by DNA sequencing.

RESULTS

Conservation of core promoter elements in the maize *cox3* gene

The maize mitochondrial *cox3* gene has three upstream promoters, at –316, –355 and –1050 relative to the translation initiation codon (9). Sequences in the –316 and –355 promoters that could be aligned with those in maize *atpA* included the core sequence and upstream purine-rich domains, as shown in Figure 1 and discussed in Mulligan *et al.* (10). To test the importance of these elements for *in vitro* transcription of *cox3*, site-directed mutagenesis was carried out. Figure 2 shows that five different mutations in the –355 core element each reduced transcription by 85–100% relative to the wild-type sequence (Fig. 2B and C, constructs 1–6) and, similarly, four core mutations in the –315 promoter reduced transcription by 85–100% relative to the control (constructs 8–12). Therefore, we conclude that the core element CRTAT is essential for *in vitro* transcription of multiple maize mitochondrial promoters.

In the *atpA* core element CGTAT the final T is the major transcription initiation site (10,16), however, this is not the case in the *cox3* promoters (9). This raised the question of whether the identity of the initiating nucleotide is also important for transcription initiation. To test this possibility for *cox3* –355, the initiating A was changed to C (Fig. 2, construct 7). *In vitro* transcription showed that this change completely eliminated initiation. Therefore, at least for this promoter, the first transcribed nucleotide is an essential promoter element. For all the *cox3* constructs tested we noted that mutations in one promoter had no influence on initiation at the non-mutated promoter. This suggests that under our *in vitro* conditions these promoters are not in competition for the transcription machinery.

Function of the upstream domain

Most maize mitochondrial promoters contain a purine-rich domain 10–20 bp upstream of the transcription initiation site (see fig. 11 in ref. 10). Site-directed mutagenesis showed that for *atpA* transversions in position –13 or –12 (–9 and –8 relative to the core sequence) caused a 50–70% decrease in transcription activity, thus defining an upstream domain required for optimal activity. To test whether similar sequences in the *cox3* promoters also played a role in transcription initiation, GA motifs at positions –9 and –8 relative to the core sequence were modified by site-directed mutagenesis, as shown in Figure 3. When these templates were used for *in vitro* transcription, however, only small changes in activity were seen relative to the wild-type control, with residual transcription ranging from 62 to 100%. These results indicate that the mutated nucleotides play only a minor role in *cox3* transcription initiation *in vitro* and that either an upstream domain is not present as defined for *atpA* or that the upstream domains of *cox3* comprise other nearby purines or other sequences.

In *atpA* the upstream domain is located 10 bp from the core promoter element (see Fig. 1). Since this distance corresponds to one helical turn, it was of interest to determine whether this

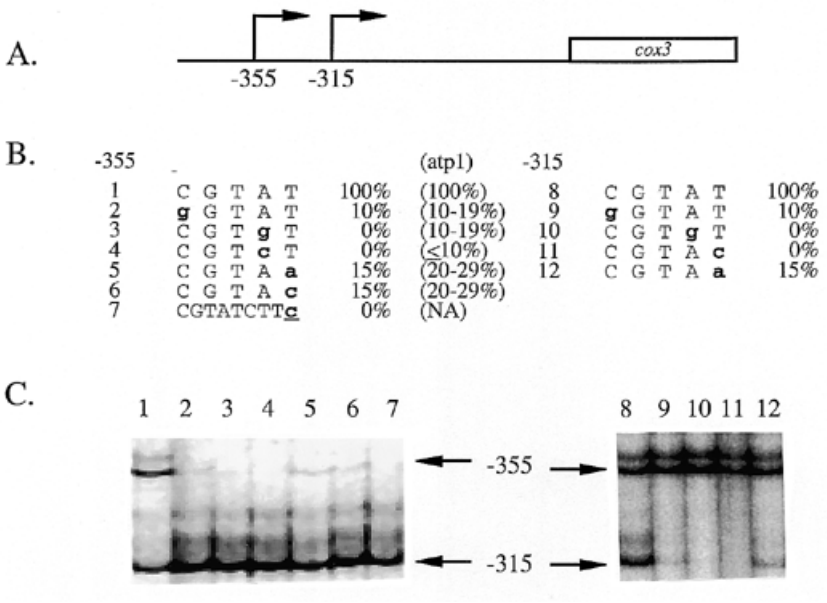

Figure 2. Site-directed mutagenesis in the core regions of the *cox3* –355 and –315 promoters. (**A**) Transcription initiation sites are indicated by arrows relative to the coding region. (**B**) Transcription activities of point mutants. To correct for unequal loading, activity of the mutated promoter was normalized to the wild-type promoter in the same construct and then compared to 100% wild-type activity as shown in lanes 1 and 8. Values shown are the averages of several experiments. (**C**) Representative 5% sequencing gel analysis of ^{32}P-labeled *in vitro* transcripts. The shadow band above the –355 transcript in the right panel is a gel artifact.

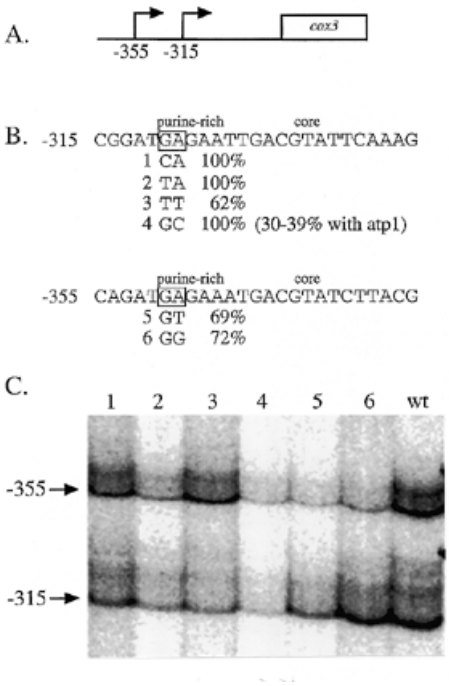

Figure 3. Site-directed mutagenesis in the putative upstream domains of the *cox3* –355 and –315 promoters. (**A**) Transcription initiation sites are indicated by arrows relative to the coding region. (**B**) Transcription activities of point mutants. To correct for unequal loading, activity of the mutated promoter was normalized to the wild-type promoter in the same construct and then compared to 100% wild-type activity as shown in lane wt. Values shown are the averages of several experiments. For construct 4 the equivalent mutation in *atpA* reduced *in vitro* transcription to 30–39% of the wild-type level (18). (**C**) Representative 5% sequencing gel analysis of ^{32}P-labeled *in vitro* transcripts.

spacing was of functional significance. To test this possibility, several promoter variants were constructed based on a 30mer sequence that gives full promoter activity (see below). In these variants, as shown in Figure 4A, the 5 bp spacer between the core and upstream elements was either replaced with unrelated sequences known not to decrease transcription activity (18) (Space1), deleted (Space0) or extended to 10 (Space2) or 15 bp (Space3). For Space2 and Space3 the extra added nucleotides were pyrimidine rich to ensure that a new purine-rich region would not be created 10 bp upstream of the core. Each of these variants was cloned in two orientations into two different plasmids, to control for modifying effects of flanking vector sequences.

In vitro transcription results for representative constructs are shown in Figure 4B. If the core and purine-rich motifs functioned only when on the same helical face, only the control and constructs Space1 and Space3 were expected to give full *in vitro* activity, whereas Space0 and Space2 should have the characteristics of upstream domain mutants, yielding 30–50% of wild-type promoter activity. The results show, however, that only Space1 gave full activity, cloned into either pBluescript or pUC18. Each of the other constructs had ~50% of the activity of Space1, suggesting that the 10 bp spacing of the core and upstream motifs was required for function of the upstream domain. It is possible that in the cases of Space0, Space2 and Space3 the reduced activity results from disturbed contacts between the transcriptional apparatus and the promoter region. For each of the spacing variants two bands were seen for each transcription reaction. We believe that the upper band, which increases in size as the spacer is extended, as would be expected for a transcript initiating upstream of the upstream domain, results from spurious initiation near the border of vector and insert sequences. It is absent from the 30mer control because this insert was placed in a polylinker of the opposite orientation (KS versus

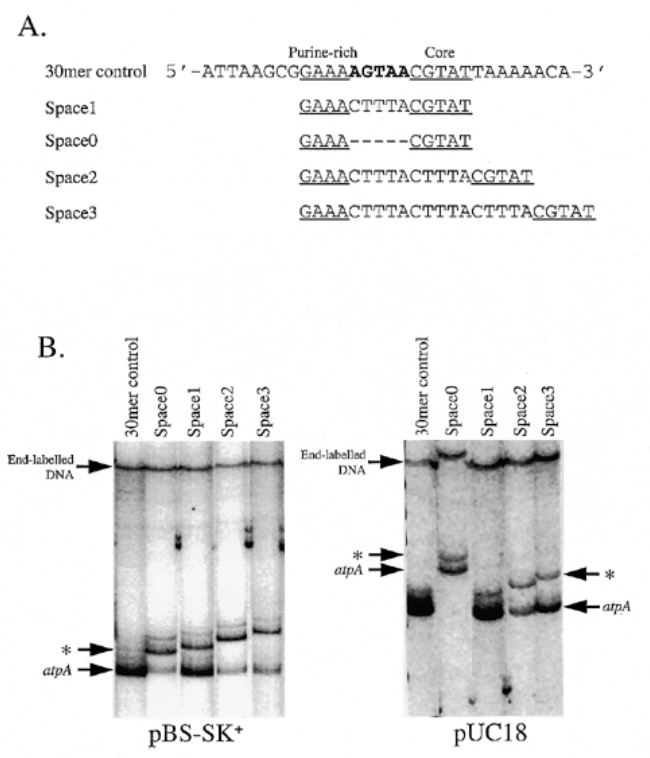

Figure 4. Analysis of spacing between the purine-rich and the core regions of an *atpA* 30 bp minimal promoter. (**A**) Sequences of test constructs. The five bases between the two regions were changed from AGTAA (shown in bold) to CTTTA so that a purine-rich region was not recreated in the different constructs. The resultant insert sizes were: Space1, 30 bp; Space0, 25 bp; Space2, 35 bp; Space3, 40 bp. Each of these oligonucleotides was inserted into the *Eco*RV site of pBluescript in the minus orientation (B, left) or into the *Sma*I and *Kpn*I sites of pUC18 (B, right). (**B**) *Pvu*II fragments containing the insert were gel purified and used as templates for *in vitro* transcription. In both panels *atpA* indicates the test transcript. The asterisk indicates a transcript that initiates near the vector/insert border. The mitochondrial transcription extract contains a terminal UTP transferase activity which end-labels the DNA fragment and serves as a control for the amount of template in the reaction; such activities have previously been found associated with RNA polymerase preparations in plant systems (see for example 31,32). In the pUC18 panel Space0 has a short unknown insert in the vector downstream of the promoter sequence, which results in a slightly longer *Pvu*II fragment and run-off transcript. Approximate transcription activities relative to the 30mer, shown for these gels but based on this and other gels, were, for the left panel: Space0, 45%; Space1, 100%; Space2, 55%; Space3, 55%. In the right panel the values were: Space0, 20%; Space1, 45%; Space2, 20%; Space3, 25%.

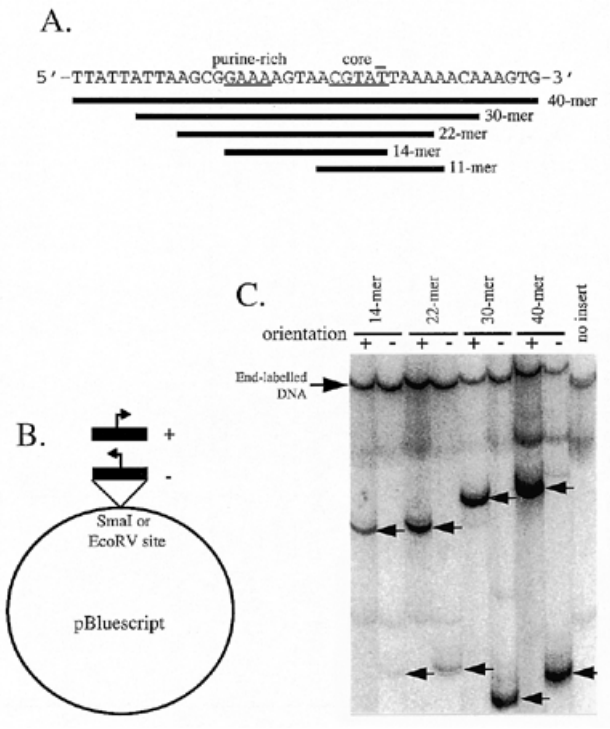

Figure 5. Analysis of minimal maize mitochondrial promoters. (**A**) Sequence surrounding the *atpA* transcription initiation site (overlined) and extent of cloned test promoters. (**B**) The oligonucleotides shown in (A) were cloned into the *Sma*I (14mer and 22mer) or *Eco*RV (30mer and 40mer) sites of pBluescript in each of two orientations. (**C**) Gel-purified *Pvu*II fragments from each plasmid were transcribed *in vitro* and a representative gel is shown. Arrows point to transcripts initiated at the *atpA* promoter. Approximate transcription activities relative to the 40mer, based on this and other gels, were: 30mer, 75%; 22mer, 40%; 14mer, 25%.

SK) and thus is in a different vector sequence context. This additional transcript is a useful confirmation, in addition to the amount of labeled template visible in each lane, that similar amounts of DNA were added to each reaction.

Minimal sequences required for *in vitro* promoter activity

Since only mutations in the core promoter sequence have pronounced effects on transcription activity, this raises the question of whether the core element itself is sufficient to direct accurate initiation. We had previously (16) tested an 11 bp sequence surrounding the *atpA* initiation site for autonomous initiation activity by cloning it into three different plasmids. However, in only one case was the plasmid transcriptionally active, suggesting that essential sequences were absent. This 11mer extended from 1 bp upstream of the core element to the +6 position, as shown in Figure 5A. To define the minimal active promoter, a variety of sequences ranging from 14 to 40 bp were cloned in both orientations into two different restriction sites in the pBluescript plasmid and tested for *in vitro* transcription activity. Figure 5C shows that all of these sequences were active in both orientations, but that the activities were variable. When corrected for uridine content (the transcripts were labeled with [^{32}P]UTP) and gel loading the 14mer, 22mer and 30mer gave ~25, 40 and 75% of the activity of the 40mer respectively. Since the 14mer spans only the core and upstream domains, we conclude that these elements alone are sufficient to promote accurate transcription initiation *in vitro*.

The 22mer was slightly more active than the 14mer, but more significant increases were seen with the 30mer and 40mer. Because of the way in which the constructs were made (Fig. 5A), the enhancing activity could come from upstream and/or downstream of the 22mer sequence, i.e. either within the transcribed region or upstream of the upstream domain. To attempt to resolve this ambiguity, a series of 26mer constructs were made, as shown in Figure 6A. Our initial results, shown in Figure 6B, suggested that the sequence AACA in the transcribed region could positively influence *in vitro* transcription activity, since the 30mer and 26mer B had significantly higher activity

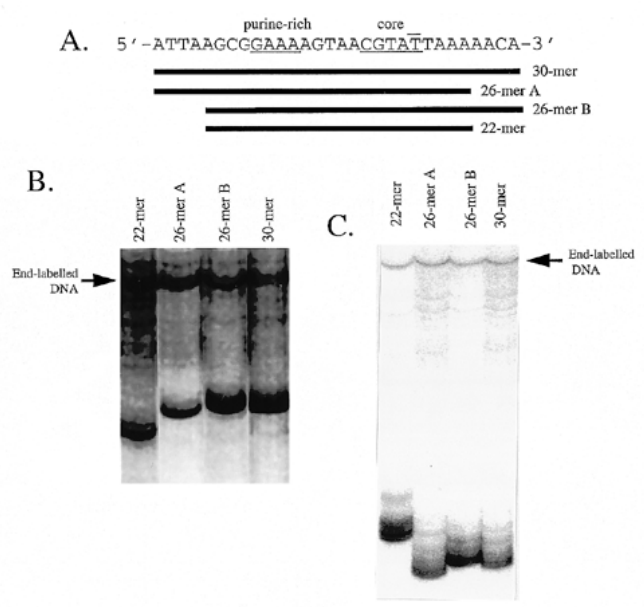

Figure 6. Analysis of 26 bp promoter constructs. (**A**) Sequence surrounding the *atpA* transcription initiation site (overlined) and extent of cloned test promoters. The 26mer oligonucleotides shown in (A) were inserted in the *Eco*RV site of pBluescript in either the plus (**B**) or minus (**C**) orientation and transcribed *in vitro* as described in the legend to Figure 5, with the 22mer and 30mer constructs from Figure 5 as controls. Approximate transcription activities relative to the 30mer, based on this and other gels, were, for (B): 26mer B, 100%; 26mer A, 50%; 22mer, 35%; for (C): 26mer B, >100%; 26mer A, 100%, 22mer, 100%. It is likely that the 30mer reaction in (C) is anomalously low; the relative activities of 26mers A and B are similar, however, in (B) and (C).

than the 22mer or 26mer A. However, when inserted into the test vector in the opposite orientation (Fig. 6C) the differences in activity were less pronounced, although 26mer B still had a slightly higher activity than 26mer A. It is likely that subtle contextual effects from the plasmid vector are responsible for this inconsistency. Taken together, our results shown in Figures 5 and 6 suggest that there are not discrete elements, apart from the core and upstream regions previously defined, that regulate transcription initiation. Instead, it appears that gradual increases in activity can be associated with longer 'minimal promoter' inserts. Because the transcription template was in all cases the same length (a gel-purified *Pvu*II fragment), the differences cannot be ascribed to template size. In conclusion, further definition of promoter elements awaits a more purified RNA polymerase preparation.

DISCUSSION

In this paper we have shown that the core motif YRTAT found in most plant mitochondrial promoters is functional in the case of two maize *cox3* promoters. This extends our previous observations that deletion (16) or mutagenesis (18) of this sequence strongly reduced or abolished *in vitro* promoter activity. We conclude that this core sequence is likely to be essential for function in each promoter within which it is found. We have also shown that for *cox3* –355, mutation of the initiation position, which is located 4 bp downstream of the core sequence, abolished transcription. Interestingly, the +1 position of the yeast mitochondrial promoter can be mutated to any base with no significant effect on initiation rates (19). In contrast, mutagenesis of purine-rich elements in the *cox3* promoters corresponding in position and sequence to the enhancing upstream domain of the maize *atpA* promoter had no discernible effect on transcription. While this result is surprising given the effect of mutations on *atpA* transcription and the prevalence of purine-rich domains upstream of plant mitochondrial core motifs (10), the general A+T richness of mitochondrial non-coding regions could result in fortuitous A-rich domains around promoters.

Attempts to define a minimal functional plant mitochondrial promoter have been difficult. In yeast only seven of the eight conserved bases have constraints (19), thus defining a 7 bp minimal promoter, whereas the promoters recognized by the related (20) bacteriophage RNA polymerases T3 and T7 are homologous over 23 bp, including 6 bp of transcribed sequence (21). In the case of minimal promoters based on the maize mitochondrial *atpA* gene an 11mer spanning the core sequence and some flanking sequences, representing the most highly conserved sequences based on alignments, was unable to direct transcription initiation in two out of three cloning contexts (16), suggesting that it depended on vector sequences in the successful case. In this paper we have shown that a 14mer which includes the core and upstream domains is sufficient to direct initiation, but that longer sequences show significantly higher activities. These results are somewhat difficult to interpret, but may suggest that multiple contacts between the transcription machinery and the template can facilitate initiation and that some of these contacts are lacking in the shorter promoter versions. In the case of T7 promoters so-called initiation and binding domains could be defined, which are clustered in the –9 to +1 region of the promoter (22). Immediately upstream of this region are 3 bp that are responsible for discrimination by T3 and T7 RNA polymerases (23). Thus, like the maize mitochondrial promoter, T7 promoter elements are distributed over an ~15 bp region and many mutations within this region have only modest effects on promoter activity (see for example 24). The quantitative aspects of T7 promoter analysis have been greatly facilitated by the purified polymerase and *in vivo* assays and clearly such methods will be necessary to conduct a fine analysis of plant mitochondrial promoter function.

The ability of plant mitochondrial polymerases to recognize promoters from multiple species, as seen with our maize extract (W.D.Rapp and D.B.Stern, unpublished results) and in other work (17), suggests that there are conserved features of the transcriptional apparatus. Indeed, partial sequences encoding T7/T3/yeast mitochondria-like RNA polymerases have been found in a wide variety of eukaryotes (25), and complete sequences for nuclear genes in *Chenopodium album* (26) and *Arabidopsis thaliana* (27) have been published. The mitochondrial localization of the translation products from one of the *Arabidopsis* genes has been inferred from *in vitro* import studies, although the presence of the protein *in vivo* has not yet been verified. Taken together, the data appear to favor a unified mode of transcription initiation in fungal, vertebrate (28) and plant mitochondria, with the core polymerase derived from a bacteriophage ancestor. The variability in promoter structure most likely reflects divergence among polymerases, as in the case of T3 and T7 RNA polymerases, where single nucleotide changes in the promoter (23) or a single amino acid change in the polymerase (29) can determine promoter recognition. Other mechanisms may

operate in more primitive organisms. For example, the mitochondrial genome of the protozoan *Reclinomonas americana* contains four genes that could specify subunits of a typical eubacterial RNA polymerase (30).

ACKNOWLEDGEMENTS

We thank Shelley Lupold and Bill Rapp for helpful discussions and Mike Mulligan for providing *cox3* plasmid DNA. This work was supported by NIH grant GM52560-02 to D.B.S.

REFERENCES

1 Unseld,M., Marienfeld,J.R., Brandt,P. and Brennicke,A. (1997) *Nature Genet*., **15**, 57–61.
2 Binder,S., Thalheim,C. and Brennicke,A. (1994) *Curr. Genet*., **26**, 519–523.
3 Lizama,L., Holuigue,L. and Jordana,X. (1994) *FEBS Lett*., **349**, 243–248.
4 Nakazono,M., Ishikawa,M., Yoshida,K.T., Tsutsumi,N. and Hirai,A. (1996) *Curr. Genet*., **29**, 417–422.
5 Remacle,C. and Marechal Drouard,L. (1996) *Plant Mol. Biol*., **30**, 553–563.
6 Giese,A., Thalheim,C., Brennicke,A. and Binder,S. (1996) *Mol. Gen. Genet*., **252**, 429–436.
7 Yan,B. and Pring,D.R. (1997) *Curr. Genet*., submitted for publication.
8 Tracy,R.L. and Stern,D.B. (1995) *Curr. Genet*., **28**, 205–216.
9 Mulligan,R.M., Lau,G.T. and Walbot,V. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7998–8002.
10 Mulligan,R.M., Leon,P. and Walbot,V. (1991) *Mol. Cell. Biol*., **11**, 533–543.
11 Binder,S. and Brennicke,A. (1993) *J. Biol. Chem*., **268**, 7849–7855.
12 Nakazono,M., Tsutsumi,N. and Hirai,A. (1995) *Plant Sci*., **105**, 227–234.
13 Brown,G.G., Auchincloss,A.H., Covello,P.S., Gray,M.W., Menassa,R. and Singh,M. (1991) *Mol. Gen. Genet*., **228**, 345–355.
14 Covello,P.S. and Gray,M.W. (1991) *Curr. Genet*., **20**, 245–252.
15 Hanic-Joyce,P.J. and Gray,M.W. (1991) *Mol. Cell. Biol*., **11**, 2035–2039.
16 Rapp,W.D. and Stern,D.B. (1992) *EMBO J*., **11**, 1065–1073.
17 Binder,S., Hatzack,F. and Brennicke,A. (1995) *J. Biol. Chem*., **270**, 22182–22189.
18 Rapp,W.D., Lupold,D.S., Mack,S. and Stern,D.B. (1993) *Mol. Cell. Biol*., **13**, 7232–7238.
19 Biswas,T.K., Baruch,T. and Getz,G.S. (1987) *J. Biol. Chem*., **262**, 13690–13696.
20 Masters,B.S., Stohl,L.L. and Clayton,D.A. (1987) *Cell*, **51**, 89–99.
21 Dunn,J.J. and Studier,F.W. (1983) *J. Mol. Biol*., **166**, 477–535.
22 Chapman,K.A., Gunderson,S.I., Anello,M., Wells,R.D. and Burgess,R.R. (1988) *Nucleic Acids Res*., **16**, 4511–4524.
23 Klement,J.F., Moorefield,M.B., Jorgensen,E., Brown,J.E., Risman,S. and McAllister,W.T. (1990) *J. Mol. Biol*., **215**, 21–29.
24 Chapman,K.A. and Burgess,R.R. (1987) *Nucleic Acids Res*., **15**, 5413–5432.
25 Cermakian,N., Ikeda,T.M., Cedergren,R. and Gray,M.W. (1996) *Nucleic Acids Res*., **24**, 648–654.
26 Weihe,A., Hedtke,B. and Borner,T. (1997) *Nucleic Acids Res*., **25**, 2319–2325.
27 Hedtke,B., Borner,T. and Weihe,A. (1997) *Science*, **277**, 809–811.
28 Bogenhagen,D.F. (1996) *J. Biol. Chem*., **271**, 12036–12041.
29 Raskin,C.A., Diaz,G., Joho,K. and McAllister,W.T. (1992) *J. Mol. Biol*., **228**, 506–515.
30 Lang,B.F., Burger,G., O'Kelly,C.J., Cedergren,R., Golding,G.B., Lemieux,C., Sankoff,D., Turmel,M. and Gray,M.W. (1997) *Nature*, **387**, 493–497.
31 Takanami,Y. and Fraenkel-Conrat,H. (1982) *Biochemistry*, **21**, 3161–3167.
32 Boege,F. (1982) *Biosci. Rep*., **2**, 379–389.