Curr Genet (1991) 20:245–251

Current Genetics

Sequence analysis of wheat mitochondrial transcripts capped in vitro: definitive identification of transcription initiation sites

Patrick S. Covello and Michael W. Gray

Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia, B3H 4H7, Canada

Received March 4, 1991

Summary. To identify transcription initiation sites in wheat mitochondria, the nascent 5′-ends of transcripts were specifically labeled by incubation of wheat mitochondrial RNA with [α-^{32}P]GTP in the presence of the enzyme guanylyltransferase. After separation of the resulting capped transcripts by electrophoresis in polyacrylamide gels, individual RNAs were recovered and directly sequenced. Four RNA sequences obtained in this way were localized upstream of the protein-coding genes *atpA, coxII, coxIII* and *orf25*. Comparison of mRNA and gene sequences allowed precise positioning of transcription initiation sites for these four genes. Sequence similarities immediately upstream of these sites define a conserved motif that we suggest as a candidate regulatory element in wheat mtDNA. The relationship between this motif and putative mitochondrial promoters in other plant species is discussed.

Key words: Mitochondrion – Transcription – Guanylyltransferase – *Triticum aestivum*

Introduction

One of the most striking features of plant mtDNA is the exceptionally high degree of conservation of coding sequences despite a highly variable gene order (Levings and Brown 1989; Gray 1989, 1990). This variability appears to be a consequence of frequent rearrangement, the main mode of evolution of the plant mitochondrial genome (Gray 1989). One result of such organizational volatility is that sequences upstream and downstream of homologous, highly conserved coding regions are often completely different in different plant species. Because flanking sequences are expected to contain elements important in directing the initiation and termination of transcription, the observed sequence divergence raises questions about how gene expression is regulated in the evolutionarily fluid plant mitochondrial genome.

Previous efforts to define promoter motifs in plant mtDNA have involved mapping the 5′-ends of a number of plant mitochondrial transcripts by primer extension and nuclease protection methods, with the aim of identifying conserved DNA sequences in the vicinity of the positions to which these 5′-ends map (Lonsdale 1989; Schuster et al 1987; Young et al. 1986). However, a limitation of this approach is that some of the 5′-ends determined in this way may actually represent sites of RNA processing rather than of transcription initiation. Furthermore, inclusion of data from different plant species in such alignments could obscure any inter-species variation in regulatory elements that may exist. For the above reasons, it would seem preferable, at least initially, to restrict comparisons to genuine transcription initiation sites within a single species. This has been done in the case of maize mitochondria in studies using guanylyltransferase to specifically label the polyphosphorylated 5′-ends of primary transcripts (Mulligan et al. 1988a, b; Kennell and Pring 1989).

In this report, we present an analysis of transcription initiation sites in wheat mtDNA. While maize and wheat are members of the same family, major differences have been noted in the pattern of transcription of the respective mtDNAs: wheat has a relatively simple mitochondrial transcription pattern, with single major transcripts for most genes, rather than the multiplicity of transcripts often seen in maize (e.g., for *coxII*, Fox and Leaver 1981; Bonen et al. 1984; *cob*, Dawson et al. 1984; Boer et al. 1985; and *atp6*, Dewey et al. 1985; Bonen and Bird 1988). It is possible that these contrasting patterns reflect differences in the regulation of transcription in the two species. To examine this possibility, and to provide further reliable data relevant to the identification of plant mitochondrial promoter sequences, we have mapped a number of transcription initiation sites in wheat mtDNA. The approach adopted, which allows very accurate determination of such sites, involves direct sequencing of the 5′-ends of RNAs "capped" by guanylyltransferase (Chris-

Offprint requests to: M. W. Gray

tianson and Rabinowitz 1983). The results suggest the presence of a loosely conserved primary sequence motif immediately adjacent to these transcription start sites.

Materials and methods

Capping of wheat mitochondrial RNA. Mitochondrial RNA (mtRNA) was extracted from embryos of wheat (*Triticum aestivum* cv. Katepwa) as described previously (Bonen et al. 1984; Bonen and Gray 1980) and precipitated overnight from 1 M NaCl at 4 °C. Up to 160 µg of wheat mtRNA was capped in a reaction volume of 100 µl, essentially as described by Auchincloss and Brown (1989) {500 µCi [α-^{32}P]GTP (3000 Ci/mmol, NEN/DuPont, Mississauga, Ontario, Canada), 15 U vaccinia virus guanylyltransferase (GIBCO/BRL, Burlington, Ontario, Canada), 50 mM Tris-HCl (pH 7.9), 1.25 mM $MgCl_2$, 6 mM KCl, 2.5 mM dithiothreitol, 0.1 mM *S*-adenosylmethionine (AdoMet), 0.6 U inorganic pyrophosphatase (Boehringer Mannheim FRG), 80 U human placental ribonuclease inhibitor (Bethesda Research Labs)}. In some cases, AdoMet was omitted from the reaction in order to generate unmethylated capped RNA.

Purification and sequencing of capped RNA. Capped RNA was initially purified by chromatography through a Sephadex G-50 spin-column, followed by phenol extraction and ethanol precipitation. RNA species were separated in a 2.5% polyacrylamide gel (20 × 20 × 0.15 cm) containing 0.25% *N,N'*-methylenebisacrylamide, 7 M urea, 50 mM Tris-borate, 1 mM EDTA at pH 8.3. Radioactive bands were located by autoradiography and excised. RNA was eluted overnight from gel slices which had been homogenized in equal volumes of elution buffer [0.5 M NaCl, 100 mM Tris-HCl, 10 mM EDTA (pH 9.1)] and phenol/*m*-cresol/8-hydroxyquinoline (1000:145:1) (see Rubin 1973). Eluted RNAs were precipitated from ethanol, repurified in a 6% sequencing gel, recovered from gel slices (see above) and reprecipitated. Purified RNAs were sequenced by the partial enzymatic degradation procedure using RNases T1 (Sigma Chemical Co., St. Louis, MO, U.S.A.), U2 and Phy M (Pharmacia (Canada) Inc., Baie d'Urfé, Québec, Canada), and in some cases M1 (Pharmacia) (Donis-Keller et al. 1977; MacDonnell et al. 1987).

Synthesis of G$^{5'}$pppU and terminal nucleotide analysis of capped RNAs. The uridine 5'→5'-phosphoester of guanosine 5'-triphosphate (G$^{5'}$pppU) was synthesized by reacting 0.25 mmol of the tributylammonium salt of UDP with 0.1 mmol of the dicyclohexylguanidinium salt of GMP-morpholidate (Sigma) in 5 ml of dry pyridine for 4 days at room temperature (Tazawa and Inoue 1983; Roseman et al. 1961). The product was purified by thin-layer chromatography on cellulose plates using isobutyric acid/0.5 M NH_4OH (5:3). The identity of the product as G$^{5'}$pppU was confirmed by paper electrophoresis in 20 mM sodium acetate (pH 5.0; Tazawa and Inoue 1983). Purified capped RNAs were treated with 0.4 U nuclease P1 (Sigma) in 20 µl of 50 mM ammonium acetate (pH 5.2) for 30 min, lyophilized and subjected to thin-layer chromatography (see above) using G$^{5'}$pppG, G$^{5'}$pppC, G$^{5'}$pppA (Pharmacia) and synthetic G$^{5'}$pppU as standards.

DNA sequencing. A recombinant DNA clone containing sequence upstream of the *Bam*H1 site internal to the wheat mitochondrial *atpA* gene (Schulte et al. 1989; Bégu et al. 1989a) was sequenced on both strands, as described by Zhang et al. 1988. The oligodeoxyribonucleotides used were AGGCTCTCGTTAAGGAATCG (primer A-1; synthesized using a Cyclone Plus DNA Synthesizer, Milligen/Biosearch, Burlington, MA, U.S.A.) and ACTCAGCTAGGCACTTACAAA (primer A-2; purchased from the University of Calgary). The clone pB100 (a 4.5 kbp *Bam*H1 fragment in pUC19; Schulte et al. 1989) was kindly provided by U. Kück.

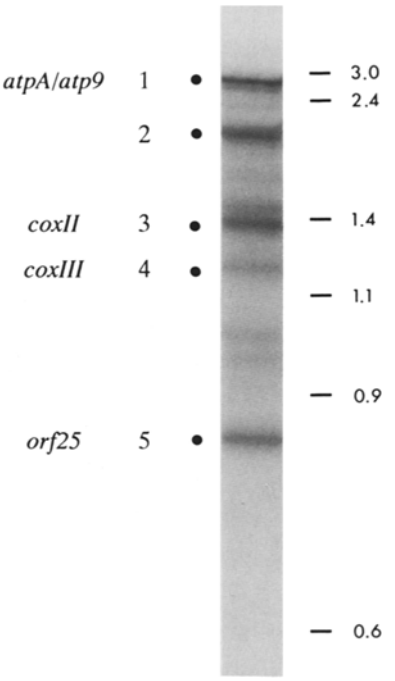

Fig. 1. Electrophoretic analysis of wheat mitochondrial transcripts capped in vitro. Autoradiogram of 1 M NaCl-insoluble wheat mtRNA labeled by incubation with [α-^{32}P]GTP in the presence of guanylyltransferase and separated by electrophoresis in a 2.5% polyacrylamide gel. Sizes of end-labelled *Acc*I and *Hae*III restriction fragments of ΦX174 are indicated in kbp on the right. Major radioactive bands are labelled 1–5 and identified as determined by RNA sequencing (see Fig. 2)

Results

Fractionation of capped wheat mitochondrial transcripts

Wheat mtRNA (1 M NaCl-insoluble fraction) was labeled with [a-^{32}P]GTP in the presence of guanylyltransferase, and in the absence or presence of AdoMet, to a specific activity of approximately 0.7×10^5 and 1.5×10^5 cpm/µg, respectively. Figure 1 shows a profile of the resulting high-molecular-weight capped RNAs, separated in a 2.5% polyacrylamide gel and visualized by autoradiography. This procedure resolved five major discrete bands, numbered 1 through 5 in Fig. 1, with sizes of approximately 2.7, 1.9, 1.3, 1.2 and 0.8 kb, respectively (based on DNA markers). In addition, a number of less abundant bands were reproducibly observed. An essentially identical high-molecular-weight pattern was found for total wheat mtRNA (not subjected to NaCl precipitation), which in addition showed bands in the 50–200 nt range (data not shown).

Identification of capped RNAs 3, 4 and 5

In order to determine the 5'-ends of the capped transcripts, capped RNAs 1–5 were subjected to enzymatic sequencing reactions. Sequencing ladders obtained for RNAs 3–5 are shown in Fig. 2. Comparison of RNA sequences determined in this way with available DNA

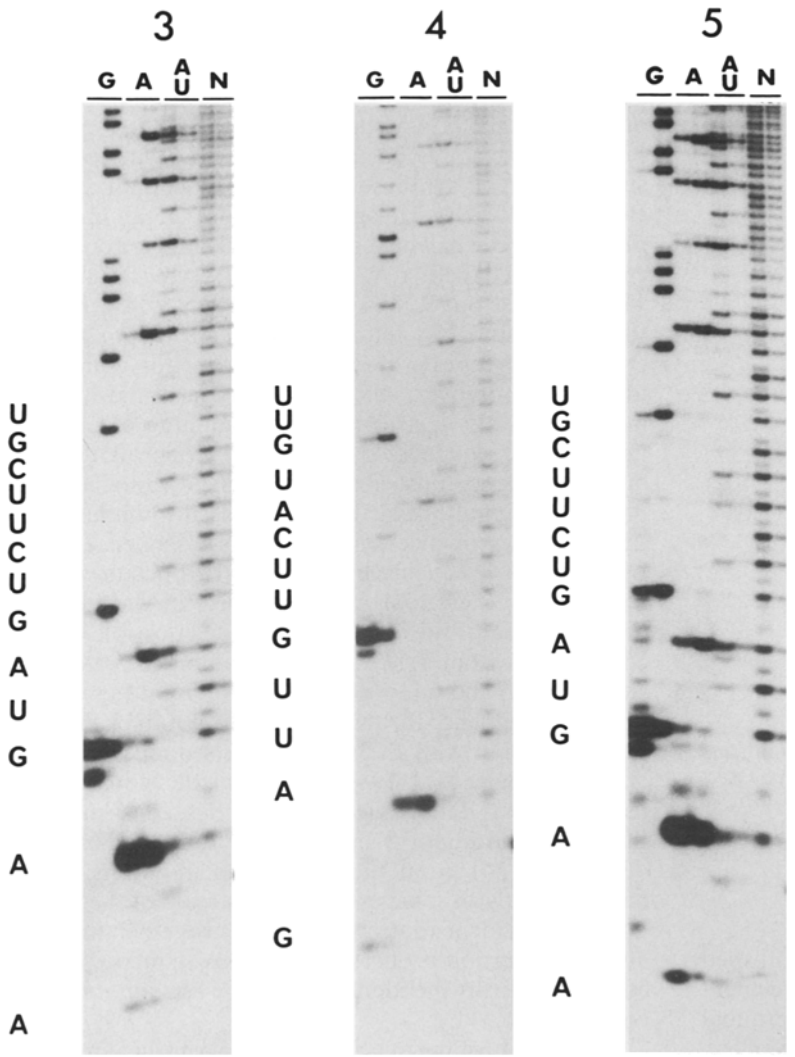

Fig. 2. 5′-terminal sequence analysis of capped wheat mtRNAs. RNA was eluted from bands *3* (*coxII*), *4* (*coxIII*) and *5* (*orf25*) of a 2.5% polyacrylamide gel (see Fig. 1), further purified by electrophoresis in a 6% polyacrylamide gel, and subjected to partial hydrolysis with RNases T1 (*lanes labeled G*), U2 (*A*) and Phy M (*A/U*) and with 0.15 M NH_4OH (*N*). Sequence is 5′ to 3′, bottom to top. In the sequencing ladder for RNA 4, bands in the right-hand G lane at positions 8 and 14 were not seen in other sequencing gels of the same RNA. These positions were, therefore, read as C

sequences identified these RNAs as transcripts of the *coxII* (Bonen et al. 1984), *coxIII* (Gualberto et al. 1990) and *orf25* (Bonen et al. 1990) genes (see also below). These assignments are supported by the results of hybridization to available wheat mtDNA clones (data not shown) and by the correspondence in transcript sizes with those determined by Northern analysis: 1.3 kb (*coxII*), 1.2 kb (*coxIII*; Gualberto et al. 1990), and 0,8 kb (*orf25*; Bonen et al. 1990). (Recent work indicates that the predominant wheat mitochondrial *coxII* mRNA is 1.3 kb in size, not 1.5 kb as originally estimated in Bonen et al. 1984; L. Bonen, personal communication.)

Although RNA 3 (*coxII*) and RNA 5 (*orf25*) share the same sequence at the 5′-end, the extended RNA sequence obtained from 6% polyacrylamide sequencing gels (Covello and Gray 1989; and data not shown) clearly revealed sequence divergence between these two transcripts starting 106 nt from the 5′-end, in agreement with the DNA sequence (Bonen et al. 1984, 1990). A C-to-U RNA editing site shared by the *coxII* and *orf25* transcripts (Covello and Gray 1989) was also found. The 5′-terminal sequence of RNA 4 was identical with the published DNA sequence of the *coxIII* gene (Gualberto et al. 1990) up to at least 120 nt (as far as the RNA sequence could be determined reliably; data not shown).

Terminal nucleotide analysis of RNAs 3, 4 and 5

A number of steps were taken to positively identify the terminal nucleotide of the capped RNA bands. Sequencing of RNA capped in the presence of AdoMet (which is required for methylation of the 5′-terminal G of the cap by guanylyltransferase preparations; Barbosa and Moss 1978) tended to yield sequencing ladders that were difficult to interpret near the bottom (i.e., in the vicinity of the 5′-terminus). This appeared to be due to partial inhibition of cleavage 3′ to the terminal native nucleotide, resulting in a relatively weak band representing methylated $G^{5'}$pppN (Fig. 2). For this reason, RNAs 3–5, labeled in the absence of AdoMet, were sequenced with the inclusion of an RNase M1 reaction (RNase M1 cleaves 3′ to A, G and U, producing 3′-OH-terminated products) and a complete nuclease P1 reaction (to generate $G^{5'}pppN^{3'}$-OH; data not shown). The terminal nucleotide in the M1 lane was identified by comparison with the P1 lane. The M1 lane could be brought into register with the T1, U2, Phy M and alkali reactions (which yield 3′-phosphates) by matching Cs in the M1 lane (gaps) with Cs in the alkali lanes (not cleaved by T1, U2 or Phy M). In this way, the terminal native nucleotide (a band corresponding to $G^{5'}$pppN) could be read.

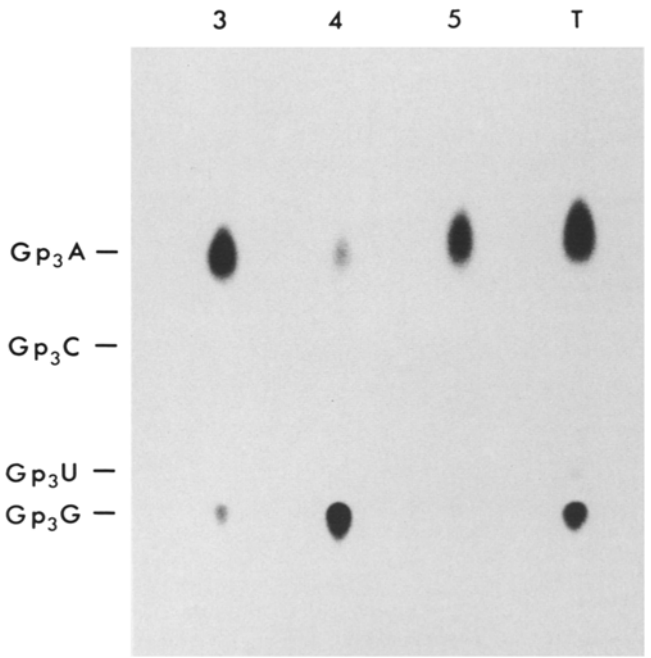

Fig. 3. 5′-terminal nucleotide analysis of capped RNA. Total salt-insoluble wheat mtRNA capped in the absence of AdoMet (*T*) and bands eluted from a 2.5% polyacrylamide gel and purified by electrophoresis in a 6% polyacrylamide gel (*bands 3, 4 and 5*) were treated with nuclease P1, and the digest subjected to thin-layer chromatography followed by autoradiography. Migration of standards ($G^{5'}$ pppN) is indicated

The assignment of terminal nucleotides was confirmed by thin-layer chromatography (Fig. 3). In agreement with the sequences shown in Fig. 2, the major 5′-terminal nucleotides in RNAs 3–5 were A, G and A, respectively. Also, our identification of the 5′-terminal nucleotide of the wheat mitochondrial *coxIII* mRNA agrees precisely with the results of primer extension analysis (Gualberto et al. 1990).

Thin-layer chromatography of total capped RNA digested with nuclease P1 (Fig. 3) indicates that most have a purine at the 5′-terminus, in agreement with data from soybean mitochondria (Auchincloss and Brown 1989) and the preference shown by many genetic systems for transcription initiation with purines (cf. Bucher and Trifonov 1986). However, it should be borne in mind that vaccinia virus guanylyltransferase itself may show a preference for certain 5′-terminal nucleotides; thus, the relative proportions of labeled cap structures may not precisely reflect the actual proportions of primary 5′-ends.

Characterization of capped RNAs 1 and 2

RNA from bands 1 and 2 yielded sequencing ladders that indicated a mixture of capped 5′-ends. However, the sequence ladder of RNA 1 could be interpreted as the superimposition of two RNA sequences that are identical apart from the presence in one of a single extra nucleotide at the 5′-end. The sequence determined was 5′-(A)GGGAUUCUUGCAAUUUUACUGCAGGCAGCGYAGCA-3′ (where Y is a pyrimidine nucleoside), with the longer sequence predominating. Knowing the size of RNA 1 (2.7 kb) and given its hybridization to a clone containing the *atp9* gene (data not shown), the 5′-terminal sequence of RNA 1 could be matched closely to the DNA sequence reported for the region upstream of the gene for the α subunit of ATP synthase of wheat mitochondria (*atpA*; Bégu et al. 1989b). This gene lies about 0.5 kb upstream of the *atp9* gene in wheat mitochondria and is apparently co-transcribed with it (Schulte et al. 1989; Bégu et al. 1989a). While our RNA sequence for RNA 1 overlapped the published DNA sequence upstream of *atpA* (Bégu et al. 1989b), the two sequences were not completely identical; moreover, the transcription initiation site mapped slightly upstream of the beginning of the published DNA sequence. Consequently clone pB100, containing sequence upstream of the wheat mitochondrial *atpA* gene (Schulte et al. 1989), was used to obtain additional DNA sequence data. This DNA sequence agrees with the RNA sequence except for a T at position 3 (relative to the 5′-terminal A), where the RNA sequence is difficult to read, and a C at position 18 (a possible C-to-U editing site within a TCR motif common to many such sites; Covello and Gray 1990). We note five differences in our DNA sequence for the 5′-flanking region of the *atpA* gene compared with the same DNA sequence reported elsewhere (Bégu et al. 1989b). Primer extension analysis, using primer A-1 and dideoxynucleotides, confirmed that the major and minor 5′-ends of hybridizing RNAs map −342 and −341 from the translation initiation codon, respectively, with some extension upstream of these positions also evident (data not shown). Together with size and hybridization properties, the RNA sequence reported here confirms the identity of RNA 1 as an *atpA/atp9* co-transcript.

Based on probing of a Southern blot of wheat mtDNA with primer A-1, it appears that the *atpA* gene is either single copy or is contained within a repeat that is >9 kbp in size (data not shown). Seven enzymes with 6-nt recognition sequences generated single hybridizing fragments of <12 kbp, whereas *Kpn*I generated two fragments, both >15 kbp. Consequently, there is a possibility that the RNA sequenced here and used as a template for primer extension is not transcribed from the *atpA* copy that has been cloned and sequenced. Hybridization and mapping data reported elsewhere (Bonen et al. 1984, 1990; Gualberto et al. 1990), as well as Southern analysis of wheat mtDNA using the oligodeoxyribonucleotide AGAAAAACACACCCAGACTC (corresponding to the anti-sense sequence upstream of the *coxIII* gene; data not shown), have indicated that the *coxII, coxIII* and *orf25* genes are each single copy.

An unambiguous RNA sequence could not be obtained for RNA 2 (1.9 kb), apparently due to substantial 5′-end heterogeneity. This species may correspond to one of the other protein-coding transcripts previously identified in wheat mitochondria, although in estimated size it does not precisely match any of these (*nad3/rps12*, 3.0 kb, Gualberto et al. 1988; *cob*, 2.4 kb, Boer et al. 1985; *coxI*, 2.2 kb, Bonen et al. 1987; *atp6*, 1.4 kb, Bonen and Bird 1988). Alternatively, it might represent a primary transcript of wheat mitochondrial 18S rRNA (1955 nt); however, primer extension analyses have failed to provide

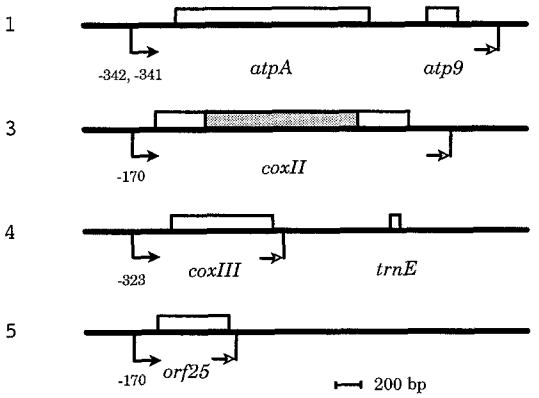

Fig. 4. Maps of mature cappable transcripts of wheat mitochondrial genes. Numbers at left correspond to capped RNA species (see Fig. 1). *Filled arrows* indicate transcription initiation sites with nucleotide positions indicated relative to translation initiation codons. *Open arrows* indicate 3′-termini estimated from S1 nuclease mapping in the case of *coxIII* (Gualberto et al. 1990), or deduced from 5′-termini and transcript sizes. *Open boxes* denote coding regions, with the single intron in the *coxII* gene (Bonen et al. 1984) indicated by *stippling*

atpA/atp9	TCGAAATAGCGTATAAAGTGATT
coxII	CAGAAAACGCGTATAGTAAGTAG
orf25	TCGAAAACGCGTATAGTAAGTAG
coxIII	GAAATTTTGCATAGAATGATTGT
Wheat consensus	rAaannGCrTAtArtragt
Young et al. 1986	AAATYTCNTAAG$^{A}_{T}$GAA
Schuster et al. 1987	AAATNTCRTAAG$^{A}_{T}$GA
Lonsdale 1989	AAATN$_{1-6}$TAAG$^{T}_{A}$GA

Fig. 5. Alignment of mtDNA sequences encompassing the transcription initiation sites (*underlined*) identified in this report: *atpA/atp9* (this report); *coxII* (Bonen et al. 1984; L. Bonen, unpublished results); *orf25* (Bonen et al. 1990); *coxIII* (Gualberto et al. 1990). Consensus sequences reported previously are shown for comparison. Partial DNA sequence for the wheat mitochondrial *atpA* gene, including and in addition to that shown, has been submitted to the EMBL database (accession no. X54387)

evidence of such a precursor (T.Y.K. Heinonen, unpublished results) and hybridization experiments do not support this possibility. Mature wheat mitochondrial 26S and 18S rRNAs are not cappable under the conditions we describe here.

The reproducible pattern of capped wheat mitochondrial RNAs includes relatively minor high-molecular-weight bands as well as low-molecular-weight capped RNAs. Our attempts to isolate RNA from some of the minor high-molecular-weight bands failed to yield samples of sufficient radioactivity and/or purity to permit sequencing.

Precise localization of transcription initiation sites

Figure 4 illustrates the relationship of the mapped transcription initiation sites to their respective genes. In each case, the transcription initiation site is within 350 bp of a translation initiation codon. An alignment of the transcription initiation sites for *atpA/atp9, coxII, orf25* and *coxIII* DNA sequences is shown in Fig. 5. Despite the small sample size, the alignment reveals striking similarities in DNA sequence around these sites, suggestive of a conserved sequence motif; in Fig. 5, this motif is termed the "wheat consensus". Of particular interest is the most highly conserved sequence block, GCrTAtA, where r is a purine nucleoside. Also apparent is a purine-rich sequence, either GAAA or rAaa (depending on whether variable spacing is allowed), which occurs further upstream in each of the four cases. Figure 5 also shows previously reported consensus motifs based on alignments of DNA sequences to which 5′-ends of plant mitochondrial transcripts map. These also contain a purine-rich (AAAT) sequence. In addition the Schuster et al. (1987) and Young et al. (1986) alignments contain CRTA (where R is a purine nucleoside) and CNTA motifs, respectively, that align with the CRTA sequence of the wheat consensus.

Discussion

A prerequisite for identification of promoter and other upstream regulatory elements is the precise identification of transcription initiation sites. We have used selective labeling with [α-^{32}P]GTP in the presence of guanylyltransferase in an attempt to identify and map the most abundant 5′-ends of primary transcripts present in wheat mitochondria. The utility of this approach is demonstrated by our detection of five major cappable species, four of which could readily be identified (by direct RNA sequencing) with previously characterized wheat mitochondrial protein-coding genes, namely *atpA/atp9, coxII orf25* and *coxIII*. These results indicate that wheat mitochondrial mRNAs, in contrast to their cytosolic counterparts (Kennedy and Lane 1975; Haffner et al. 1978; Saini and Lane 1977; Lane 1981), are uncapped in vivo.

Although the results of separate capping experiments with independently isolated preparations of wheat mtRNA have yielded consistently reproducible patterns of labeled RNAs, it is difficult to say how accurately this pattern reflects, either qualitatively or quantitatively, the actual distribution of primary transcripts in the steady state RNA population. Both in vivo and in vitro, the observed pattern may be influenced by a number of biochemical phenomena, including: modification of the 5′-termini of primary transcripts (e.g., by dephosphorylation) so as to render them uncappable; loss of primary 5′-termini due to subsequent RNA processing; modification of the 5′-termini of processed transcripts to render them cappable (cf. L'Abbé et al. 1990); differential reactivity of primary 5′-ends in the capping reaction; and contamination of mitochondrial RNA with non-mitochondrial cappable transcripts. The last possibility does not appear to be a problem in the case of wheat mtRNA preparations, at least in the high-molecular-weight range, as total wheat embryo RNA shows no evidence of major cappable species >150 nt in size (data not shown). In the case of the *coxII* transcript, we are encouraged that we have identified a genuine primary 5′-terminus by the fact

that a partially purified wheat mitochondrial extract is able to initiate transcription from a wheat *coxII* template at precisely the site predicted by the capping-sequencing results presented here (Hanic-Joyce and Gray 1991).

When combined with 5′-terminal nucleotide analysis, direct sequencing of the 5′-terminal regions of capped primary transcripts permits the mapping of transcription initiation sites precisely to the nucleotide. This allows accurate alignment of potential regulatory elements which, even in this relatively small data set, show considerable similarity in nucleotide sequence (Fig. 5). While attempts have been made previously to define a consensus promoter by aligning DNA sequences in the vicinity of the mapped positions of 5′-termini of a number of different plant mitochondrial RNAs (Lonsdale 1989; Schuster et al. 1987), these comparisons have suffered from the limitation that some of the sequences in question (ones that have not been further verified by capping experiments) may actually represent RNA processing sites. Young et al. (1986), for example, suggested that the motif shown in Fig. 5 is an RNA processing signal. Consequently, the inferences drawn from such comparisons should be properly qualified, and may even be invalidated by any fortuitous similarity between promoter and processing site sequences, as well as by possible species-specific differences.

In fact, the wheat consensus derived from the alignment in Fig. 5 differs distinctly from regulatory sequence motifs previously proposed on the basis of 5′-end mapping data and DNA sequence alignments. The wheat consensus derived from our capping-sequencing data contains the sequence AAA in common with previous alignments (Fig. 5) and a CRTA motif which is in agreement with the Schuster et al. (1987) and Young et al. (1986) alignments. However, apart from these similarities, there is considerable divergence between the wheat mtDNA consensus and those defined by the other alignments. The apparent discrepancy in the inferred positions of transcription start sites (see Fig. 5) may reflect the fact that the primer extension and nuclease protection techniques are inherently less precise in localizing 5′-termini than is the direct RNA sequencing approach. Furthermore, as noted earlier, previous alignments based on 5′-end mapping may have included sequences other than transcription initiation sites. On the other hand, our wheat alignment is based on a limited data set, which may be too small at present to permit an unambiguous comparison with previous alignments. However, additional data obtained recently in an independent study (L. Bonen, personal communication) confirm the differences we note here between wheat and other plant mtDNAs in the vicinity of RNA start sites.

In other mitochondrial systems, notably animal (Chang et al. 1987) and fungal (Christianson and Rabinowitz 1983; Osinga et al. 1982; Kennell and Lambowitz 1989), transcription is generally initiated within a functionally defined promoter motif. It is reasonable to suggest that the wheat consensus in Fig. 5 is (or is part of) a promoter element, although a functional assay will be required to define it convincingly. If the motifs deduced by previous alignments (Lonsdale 1989; Schuster et al. 1987; Young et al. 1986) are also promoter elements, these may be different from the promoters in wheat mtDNA. All three of the previous alignments (Fig. 5) favour dicotyledonous over monocotyledonous species; consequently, the differences noted here may reflect specific differences in mtDNA promoters betwen these two major groups of plants. In collaboration with G. Brown (McGill University) we have obtained strong support for this possibility from capping-sequencing and cloning studies that define three soybean mitochondrial transcription start sites (Brown et al. 1991).

Within the monocotyledons, it is instructive to compare the data presented here with that for maize mitochondria. Initially, little sequence similarity could be discerned in the vicinity of maize mitochondrial transcription start sites mapped by capping/nuclease protection (Mulligan et al. 1988a, b; Kennell and Pring 1989). This is in marked contrast to the situation in yeast mitochondria, where there is a highly conserved nonanucleotide promoter at mapped initiation sites (Christianson and Rabinowitz 1983; Osinga et al. 1982). However, Mulligan et al. (1991) have recently ordered mapped initiation sites on the basis of relative promoter strength. From this analysis, they have proposed the maize mitochondrial promoter consensus (A/T)CRTA(T/G)A(A/T)AAA. This is generally in good agreement with the wheat alignment (Fig. 5) except at the 3′-end. Mulligan et al. (1991) also note a $(G/C)(AT)_4$ motif, which is variable in position, upstream of the above consensus; this may correspond to the GAAA(A/T) element that occurs in the four wheat sequences. It is difficult to say whether there are subtle differences in the putative mitochondrial promoter sequences of wheat and maize, possibly related to the differences observed in transcription patterns, i.e., single (wheat) vs. multiple (maize) initiation sites for individual genes. Clearly, the development of accurate and efficient in vitro mitochondrial transcription systems from different plant species will be critical in defining more precisely what constitutes a promoter in various plant mitochondria.

Acknowledgements. This investigation was supported by a Postdoctoral Fellowship (Biotechnology Training Program) to P.S.C. and Operating Grant MT-4124 (to M.W.G.) from the Medical Research Council of Canada. M.W.G. is a fellow in the Evolutionary Biology Program of the Canadian Institute for Advanced Research. We are grateful to L. Bonen, U. Kück and R.M. Mulligan for generously providing clones and/or data prior to publication, to T.Y.K. Heinonen for carrying out preliminary work that led to the experiments reported here, and to R.W. Chambers for assistance in the synthesis of $G^{5'}pppU$. We thank M.N. Schnare, D.F. Spencer, L. Bonen, G.G.Brown, M.B. Coulthart and K. Lonergan for advice and comments.

References

Auchincloss AH, Brown GG (1989) Biochem Cell Biol 67:315–319
Barbosa E, Moss B (1978) J Biol Chem 253:7698–7702
Bégu D, Graves P-V, Litvak S, Araya A (1989a) Nucleic Acids Res 17:9491
Bégu D, Graves P-V, Litvak S, Araya A (1989b) Nucleic Acids Res 17:9492

Boer PH, McIntosh JE, Gray MW, Bonen L (1985) Nucleic Acids Res 13:2281–2292
Bonen L, Bird S (1988) Gene 73:47–56
Bonen L, Gray MW (1980) Nucleic Acids Res 8:319–335
Bonen L, Boer PH, Gray MW (1984) EMBO J 3:2531–2536
Bonen L, Boer PH, McIntosh JE, Gray MW (1987) Nucleic Acids Res 15:6734
Bonen L, Bird S, Belanger L (1990) Plant Mol Biol 15:793–795
Brown GG, Auchincloss A, Covello PS, Gray MW, Menassa R, Singh M (1991) Mol Gen Genet (in press)
Bucher P, Trifonov EN (1986) Nucleic Acids Res 14:10009–10026
Chang DD, Fisher RP, Clayton DA (1987) Biochim Biophys Acta 909:85–91
Christianson T, Rabinowitz M (1983) J Biol Chem 258:14025–14033
Covello PS, Gray MW (1989) Nature 341:662–666
Covello PS, Gray MW (1990) Nucleic Acids Res 18:5189–5196
Dawson AJ, Jones VP, Leaver CJ (1984) EMBO J 3:2107–2113
Dewey RE, Levings III CS, Timothy DH (1985) Plant Physiol 79:914–919
Donis-Keller H, Maxam AM, Gilbert W (1977) Nucleic Acids Res 4:2527–2538
Fox TD, Leaver CJ (1981) Cell 26:315–323
Gray MW (1989) Annu Rev Cell Biol 5:25–50
Gray MW (1990) In: Dennis DT, Turpin DH (eds) Plant physiology, biochemistry and molecular biology. Longman, London, England
Gualberto JM, Wintz H, Weil J-H, Grienenberger J-M (1988) Mol Gen Genet 215: 118–127
Gualberto JM, Domon C, Weil J-H, Grienenberger J-M (1990) Curr Genet 17:41–47
Haffner MH, Chin MB Lane BG (1978) Can J Biochem 56:729–733
Hanic-Joyce PJ, Gray MW (1991) Mol Cell Biol 11:2035–2039
Kennedy TD, Lane BG (1975) Can J Biochem 53:1346–1348
Kennell JC, Lambowitz AM (1989) Mol Cell Biol 9:3603–3613
Kennell JC, Pring DR (1989) Mol Gen Genet 216:16–24
L'Abbé D, Lang BF, Desjardins P, Morais R (1990) J Biol Chem 265:2988–2992
Lane BG (1981) Can J Biochem 59:868–870
Levings III CS, Brown GG (1989) Cell 56:171–179
Lonsdale DM (1989) In: Stumpf PK, Conn EE (eds) The biochemistry of plants, vol 15. Academic Press, New York
MacDonnell MT, Hansen JN, Ortiz-Conde BA (1987) Methods Microbiol 19:357–404
Mulligan RM, Lau GT, Walbot V (1988a) Proc Natl Acad Sci USA 85:7998–8002
Mulligan RM, Maloney AP, Walbot V (1988b) Mol Gen Genet 211:373–380
Mulligan RM, Leon P, Walbot V (1991) Mol Cell Biol 11:533–543
Osinga KA, De Haan M, Christianson T, Tabak HF (1982) Nucleic Acids Res 10:7993–8006
Roseman S, Distler JJ, Moffatt JG, Khorana HG (1961) J Am Chem Soc 83:659–663
Rubin GM (1973) J Biol Chem 248:3860–3875
Saini MS, Lane BG (1977) Can J Biochem 55:819–824
Schulte E, Staubach S, Laser B, Kück U (1989) Nucleic Acids Res 17:7531
Schuster W, Hiesel R, Wissinger B, Schobel W, Brennicke A (1987) In: von Wettstein D, Chua N-H (eds) Plant molecular biology. NATO ASI series, vol 140. Plenum Press, New York
Tazawa I, Inoue Y (1983) Nucleic Acids Res 11:2907–2915
Young EG, Hanson MR, Dierks PM (1986) Nucleic Acids Res 14:7995–8006
Zhang H, Scholl R, Browse J, Somerville C (1988) Nucleic Acids Res 16:1220

Communicated by H. Bertrand