Plant Physiol. (1994) 105: 1167–1177

Ribosomes Pause during the Expression of the Large ATP Synthase Gene Cluster in Spinach Chloroplasts[1]

Neil E. Stollar[2], Jeong-Kook Kim[3], and Margaret J. Hollingsworth*

Department of Biological Sciences, State University of New York at Buffalo, Buffalo, New York 14260

The large ATP synthase gene cluster from spinach (*Spinacia oleracea*) chloroplasts encodes five genes, the last four of which encode subunits of the ATP synthase complex. In preliminary experiments (J.K. Kim, M.J. Hollingsworth [1992] Anal Biochem 206: 183–188) it was shown that ribosomes pause during translation of these open reading frames. We have examined ribosome pausing in the four ATP synthase open reading frames of this gene cluster to determine whether it could affect the final ratio of the ATP synthase polypeptides derived from the cluster. Ribosome pauses were mapped and found to be distributed in a nonuniform manner. We have quantitated the relative extent of ribosome pausing within each open reading frame. There is a general but not absolute correlation between the extent of ribosomal pausing and the protein levels found within the ATP synthase complex. We conclude that although it is not the sole factor, ribosome pausing may be a significant posttranscriptional mechanism affecting the expression of the large ATP synthase gene cluster in spinach chloroplasts.

Chloroplast function requires the activity of several multicomponent protein complexes, one being the proton-translocating ATP synthase/ATPase complex associated with the thylakoid membranes. The chloroplast ATP synthase complex, which is analogous to the mitochondrial and prokaryotic complexes, uses the energy generated by the proton gradient formed during the light reactions of photosynthesis to synthesize ATP (Strotmann and Bickel-Sandkotter, 1984; Senior, 1985). The chloroplast ATP synthase complex is composed of two subcomplexes, CF_0 and CF_1, consisting of a total of nine unique subunits. Of these nine subunits, six are encoded within the chloroplast genome (Herrmann et al., 1983; Hudson and Mason, 1988). These six genes are localized in the small (dicistronic) (Zurawski et al., 1982) and the large (five genes, four of which are ATP synthase) (Hennig and Herrmann, 1986; Hudson et al., 1987) ATP synthase gene clusters in the chloroplast genome. The remaining three subunits are encoded in the nucleus and must be imported from the cytoplasm into the chloroplast (Herrmann et al., 1983).

The large ATP synthase gene cluster from spinach (*Spinacia oleracea*) chloroplasts encodes five proteins: the small ribosomal subunit protein 2 (*rps*2) followed by (in order) four ATP synthase polypeptides, CF_0-IV (*atp*I), CF_0-III (*atp*H), CF_0-I (*atp*F), and CF_1-α (*atp*A) (Hennig and Herrmann, 1986; Hudson et al., 1987; nomenclature: Hallick and Bottomley, 1983). This cluster is transcribed from two promoters, resulting in two large polycistronic RNAs consisting of either the coding sequences for all five genes or the 3′-most three (Stollar and Hollingsworth, 1994). The two primary transcripts undergo extensive processing that involves endonucleolytic and exonucleolytic cleavages as well as splicing (Hudson et al., 1987; Kim and Hollingsworth, 1993) to remove the group II intron (Michel and Dujon, 1983; Michel et al., 1989) that interrupts the *atp*F coding region. As a result of the processing events, more than 30 RNA species are produced. Even though processing of the primary transcript is extensive, we have observed that most multicistronic as well as monocistronic RNA species can be found on polysomes and are therefore presumed to be translated (N.E. Stollar, unpublished data).

The four proteins derived from the large ATP synthase gene cluster are found within the ATP synthase complex at levels of 1:5–12:1:3 (CF_0-IV:CF_0-III:CF_0-I:CF_1-α, respectively) (McCarty et al., 1988). As mentioned above, only two transcription initiation regions have been detected for this cluster (Stollar and Hollingsworth, 1994). Although the relative strength of the two promoters is unknown, the initial ratio of open reading frames in the primary transcripts from these promoters cannot correlate to the final ratio of the polypeptides in the ATP synthase complex. Therefore, the final ratio of the polypeptides encoded by this cluster must be influenced by posttranscriptional mechanisms.

There are several mechanisms through which expression can be affected posttranscriptionally (reviewed by Gruissem and Tonkyn, 1993). Differential mRNA stability, which may be conferred in the chloroplast by an inverted repeat region downstream of the coding sequences (Stern et al., 1989; Adams and Stern, 1990; Stern et al., 1991), and RNA processing could have a significant influence on gene expression

[1] This research was funded through the National Science Foundation, grant No. MCB91–05726.

[2] Present address: Department of Microbiology and Molecular Genetics, University of California at Irvine, Irvine, CA 92715.

[3] Present address: Department of Biology, Korea University, Seoul 136–701, Korea.

* Corresponding author; fax 1–716–645–2891.

Abbreviations: Chloroplast gene designations: *atp*A, ATP synthase subunit CF_1-α; *atp*B/E, ATP synthase subunits CF_0-ϵ/β; *atp*F, ATP synthase subunit CF_0-I; *atp*H, ATP synthase subunit CF_0-III; *atp*I, ATP synthase subunit CF_0-IV; *psa*A, PSII subunit A1; *psa*B, PSII subunit A2; *psb*A, PSII protein D1; *rbc*L, large subunit of ribulose-1,5-bisphosphate carboxylase; *rps*2, chloroplast small ribosomal subunit 2; DEPC, diethyl pyrocarbonate; RT-PCR, reverse transcriptase-polymerase chain reaction.

by controlling the steady-state level of translatable transcripts present in the chloroplast. Additional posttranscriptional factors could be found at the translation level. Transcripts that have a higher relative rate of translation initiation or that are more efficient during translation elongation would produce more protein over time compared to transcripts that are not as efficient. Finally, the levels of a particular protein would also be affected by its stability compared to the others.

We have examined ribosome pausing as a possible posttranscriptional mechanism affecting the final protein ratio from the large ATP synthase gene cluster. Ribosome pausing has been shown to occur in several chloroplast systems (Berry et al., 1988; Klein et al., 1988; Kim et al., 1991). Extensive ribosome pausing could affect the rate of translation elongation to the point that it becomes the rate-limiting step in translation, rather than initiation or termination, which are more commonly observed to be rate limiting (Lodish and Jacobsen, 1972; Bergmann and Lodish, 1979). Wolin and Walter (1988) have shown that translating ribosomes can be mapped at positions where they have stalled or paused in vitro. Kim and Hollingsworth (1992) have taken this "heel-printing" method a step further by showing that in vivo ribosome pause sites can also be mapped with precision. Here, we show that many ribosome pause sites are distributed in a nonuniform fashion along the entire length of the large ATP synthase gene cluster from spinach chloroplasts. Quantitation of the total extent of pausing within each ATP synthase mRNA from the cluster reveals a general, but not absolute, correlation to the protein ratios within the ATP synthase complex. We conclude that ribosome pausing may exert a significant influence on the final ratios of the ATP synthase polypeptides derived from this cluster.

MATERIALS AND METHODS

Plant Material

Spinach (*Spinacia oleracea* cv Bloomsdale Longstanding) seeds were germinated in vermiculite and then transferred and grown hydroponically under greenhouse conditions in Peter's Professional Hydro-Sol supplemented with calcium nitrate. Young spinach leaves used in these experiments were approximately 3 to 5 cm in length.

Oligonucleotide Primers

Oligonucleotide primers were designed and synthesized to be identical to specific regions of the coding strand of the large ATP synthase gene cluster. Sequences and their location along the gene cluster, consistent with the bp numbering system used by Hudson et al. (1987), of each primer used in these experiments are:
rps2–3′-RNAL: GTGAAGGCCGCTCTAGC: 1194–1210,
RPI-1: CTAGGTCCTTATAACTTCTT: 1543–1562,
RPI-2: AATCATCACGTTACCTCATG: 1786–1805,
RPI-3: TAGTACCTTTAGTGGTTCCT: 2053–2072,
RPI-4: GCAGATTCGATGTTTCGAAG: 2400–2419,
RatpH: TCAACTGGATAAATATTAGC: 2765–2784,
RPH-1: GTCTTGAACTTGCAGGTTGC: 3160–3179,
RatpF: TTATTCTTATTGGGAATCTC: 3406–3425,
RPF-1: CGAGTGAATGGATATTCTGA: 4560–4579,
RPF-2: CGGAAAGGCCATTGAACAGC: 4487–4506,
RatpA: CATTTACGGACCATCAATGC: 4767–4789,
RPA-1: ATACCCGTGAGTGAGGCTTA: 5185–5204,
RPA-2: GGCGCAGGTAGTTACTAACT: 5511–5530,
RPA-3: TCTTTTATTTGCATTCACGC: 5765–5784,
RPA-4: AACTGGCGCAATTTGCAGAA: 6050–6069.

The oligonucleotide primers used in primer extension reactions to localize the positions of ribosome pause sites were 5′ end labeled with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP (4500 Ci/mM, ICN) according to the supplier's instructions.

Additional oligonucleotide primers that were used in PCR are:
atpI-COOH: AAATCAATGATGACCTTCCAGGGA: 2180–2157,
atpH-COOH: GGATTAAACAAAAGGATTCGCAAAT: 3118–3094,
atpF-COOH: ATCAGTTATTTCGTTCATCGCACCA: 4826–4802.

Plasmids and Probe Generation

A 6.4-kb *Cla*I, *Xho*I fragment of spinach chloroplast DNA, containing the entire large ATP synthase gene cluster, beginning 517 bp upstream of the *rps*2 coding region and ending 9 bp downstream of the *atp*A coding region, was cloned into pBluescript SKII− (Stratagene). The resulting plasmid, pNSGC6, was utilized for the additional cloning of individual open reading frame regions of ATP synthase genes.

To isolate individual gene regions, PCRs were performed using AmpliTaq DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT), primers rps2–3′-RNAL and atpI-COOH for *atp*I and RatpH and atpH-COOH for *atp*H, and pNSGC6 as template, according to the supplier's instructions. To produce the *atp*F-specific region, RT-PCR was performed using the GeneAmp RNA PCR kit (Perkin-Elmer Cetus) and primers RatpF and atpF-COOH and total chloroplast RNA as template, according to the supplier's instructions. The RT-PCR reactions produced two amplified pieces of DNA corresponding to spliced and unspliced *atp*F RNAs. For the purpose of these experiments, the DNA corresponding to the unspliced RNA was cloned. These amplified DNAs were cloned with the use of the pT7Blue T-Vector kit (Novagen, Madison, WI), resulting in the clones pT7Blue-I6 (*atp*I), pT7Blue-H2 (*atp*H), and pT7Blue-Fun1 (*atp*F). The *atp*A region was subcloned from pNSGC6 by first digesting it with *Ava*II and filling in the 5′ overhang that was generated with T4 DNA polymerase (New England Biolabs). The resulting DNA was additionally digested with *Cla*I. A band corresponding to the *atp*A gene was excised from a low-Tm agarose (Seaplaque, FMC, Rockland, ME) gel. The *atp*A DNA fragment was ligated into pBluescript SKII− vector that was digested with the restriction enzymes *Cla*I and *Eco*RV. The resulting plasmid was named pSKMA5.

Plasmids were linearized with appropriate restriction enzymes and used in in vitro transcription reactions with T7 RNA polymerase (Epicentre Technologies, Madison, WI). After acrylamide/urea gel electrophoresis of the reaction products, RNAs were eluted from the gel by crushing and soaking in TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA). The

specific activity of each probe was quantitated by HCl precipitation.

Chloroplast RNA Isolation

Spinach chloroplasts were isolated according to Orozco et al. (1986), with the exception that the Percoll gradients were subjected to centrifugation at 8100*g* for 20 min. The intact chloroplasts were lysed by the addition of one-third volume of 50 mM sodium acetate, 1% SDS, with 2.5 mM aurintricarboxylic acid (added as an RNase inhibitor [Hallick et al., 1977]). The RNA was extracted twice with equal volumes of phenol and chloroform:isoamyl alcohol (24:1), followed by ethanol precipitation. The chloroplast RNA pellet was dissolved in DEPC-treated H_2O, then passed over a P-30 Bio-Gel (Bio-Rad) spin column to remove the aurintricarboxylic acid.

Isolation of Ribosome-Protected RNA Fragments

Ribosome-protected RNA fragments were produced from young spinach leaves essentially as described by Kim and Hollingsworth (1992), with the exception that cycloheximide was also included in the polysome extraction buffer to 50 μg/mL. In addition, each polysome pellet was resuspended in 100 μL, rather than 50 μL, of resuspension buffer. This preparation resulted in essentially the same polysome profile that was reported by Kim and Hollingsworth (1992). After the presence of polysomes was confirmed, the polysomes in the remaining pellets were used to isolate ribosome-protected RNA fragments.

Micrococcal nuclease digestion reactions were employed to produce ribosome-protected RNA fragments from polysomal RNAs as described by Kim and Hollingsworth (1992), with the exceptions that each polysomal RNA aliquot (100 μL) was initially divided in two, and 200 units of micrococcal nuclease were used in these reactions. In addition, after the micrococcal nuclease digestion reactions, samples were not divided into two portions but were used in their entirety for the remainder of the procedure.

As a control for the activity of the micrococcal nuclease, 50 μL of purified polysomes were extracted twice with phenol and chloroform:isoamyl alcohol (24:1), precipitated with ethanol, resuspended in 50 μL of DEPC-H_2O, and subjected to micrococcal nuclease digestion.

To map 5′ ends produced in vivo, either total chloroplast RNA or 50 μL of purified polysomal RNAs were extracted twice with phenol and chloroform:isoamyl alcohol (24:1), precipitated with ethanol, and resuspended in 50 μL of DEPC-H_2O. These samples were not incubated with micrococcal nuclease but were otherwise treated the same as the nuclease-digested samples.

Primer Extension Assay

Nuclease-resistant RNA fragments, the nuclease-digested RNA control, and the undigested RNA samples were each separately annealed to single-stranded DNA complementary to the large ATP synthase gene cluster under conditions described by Kim and Hollingsworth (1992). In addition, a 5′ end-labeled oligonucleotide primer complementary to a region upstream of the area of interest was annealed to the same DNA molecule. After the annealing of the RNA and primers to the single-stranded DNA, the labeled oligonucleotide primers were extended with T4 DNA polymerase (New England Biolabs) (Kim and Hollingsworth, 1992). A sequencing ladder was generated using the Sequenase 2.0 kit (United States Biochemical) and the same oligonucleotide primer and single-stranded DNA used in the primer-extension reaction. The extension products and sequencing reactions were subjected to electrophoresis through an 8% sequencing gel (Sequagel-8, National Diagnostics, Manville, NJ). The gels were fixed, dried, and subjected to autoradiography.

Dot-Blot Hybridization and Quantitation

Ribosome-protected RNA fragments were spotted onto dry strips of Zeta Probe membrane (Bio-Rad) at concentrations of 0.1, 0.3, and 1.0 μg in a total volume of 5 μL. As a control, 1 μg of in vitro-transcribed RNA encompassing the entire large ATP synthase gene cluster was also spotted onto each filter strip in a volume of 5 μL. This RNA was transcribed using T7 RNA polymerase (Epicentre Technologies, Madison, WI) and pNSGC6 that was linearized with *Cla*I. All spots were air dried overnight before being used for hybridization.

Filter strips were prehybridized for at least 4 h at 56°C in 6× SSC, 0.5% SDS, 0.2% ficoll, 0.2% polyvinylpolypyrrolidone, 0.2% BSA, and 200 μg/mL boiled herring sperm DNA. The prehybridization solution was removed and replaced with fresh solution. Fifty femtomoles of a radiolabeled RNA probe specific for a particular gene region were added to each filter. Hybridization was carried out at 56°C overnight. The hybridized filters were washed successively in 2× SSC; 2× SSC, 0.1% SDS; 0.5× SSC, 0.1% SDS; and 0.1× SSC, 0.1% SDS, each for 15 min at room temperature. The filters were exposed to x-ray film.

To determine the relative abundance of transcripts derived from the four ATP synthase genes of the cluster, additional dot blots were made, as described above, that consisted of 0.1, 0.5, and 2.5 μg of total chloroplast RNA, plus 1.0 μg of in vitro-transcribed RNA derived from linearized pNSGC6 as control. These membrane filters were prehybridized for 4 h at 42°C in 5× SSC, 50% deionized formamide, 0.5% SDS, 0.02% ficoll, 0.02% polyvinylpolypyrrolidone, 0.02% BSA, and 200 μg/mL herring sperm DNA. Prior to hybridization, the prehybridization solution was removed and replaced with solution that was freshly prepared. A single radiolabeled RNA probe, specific for one of the ATP synthase genes, was added to each filter and allowed to hybridize overnight at 42°C. The filters were washed as described for the protected fragments blots.

For quantitation, each spot was excised from the filter and counted by liquid scintillation. The extent of hybridization to each probe was found to be directly proportional to the quantity of RNA in the dot. This confirms that the hybridization kinetics were driven by the chloroplast RNA fragments, not by the quantity of radiolabel added to the reaction. To control for the length of each probe, the counts derived from each set of ribosome-protected RNA fragments or total

chloroplast RNAs were normalized to the counts obtained from the control spot on the same filter. The ratio of the relative number of ribosome pause sites or relative transcript abundance was determined by dividing the normalized results from each mRNA by the normalized results from the mRNA that had the lowest hybridization signal. The SE was calculated using the results from three independent nuclease-resistant RNA blots or two total chloroplast RNA blots.

RESULTS

Production of Ribosome-Protected Fragments

The first step in the production of ribosome-protected RNA fragments is the isolation of polysomal RNA. Polysomes were isolated from whole spinach cells. The presence of polysomes in the sample was confirmed by Suc gradient analysis (Kim and Hollingsworth, 1992). Polysome samples were treated with micrococcal nuclease, yielding ribosomes bound to small RNA fragments.

Localization of Ribosome Pauses

Micrococcal nuclease-resistant RNA fragments were examined to determine the ribosome pause sites within the four ATP synthase genes in the large ATP synthase gene cluster. Ribosomes protect between 30 and 35 nucleotides of RNA from digestion by nuclease (reviewed by Steitz, 1979). Therefore, if ribosomes move along the RNA at a constant rate during translation, all possible RNA fragments of 30 to 35 nucleotides should be equally represented. If ribosomes pause during translation, there will be an overrepresentation of the RNA fragments corresponding to the pause sites.

To map the position of the ribosome pauses, nuclease-resistant RNA fragments and a 5′ end-labeled oligonucleotide primer encoding the complement of the RNA of interest were annealed to single-stranded DNA. These hybrids were used in a primer extension assay with T4 DNA polymerase. T4 DNA polymerase is unable to precess through the RNA/DNA hybrids created by annealing the protected RNA fragments to the DNA (Alberts, 1984). Consequently, the RNA fragments will act as barriers preventing the polymerase from continuing the extension reaction. The result, as visualized by autoradiography, is a series of labeled DNA molecules corresponding to an extension product that starts with the 5′-labeled DNA oligonucleotide and ends at the 5′ end of an annealed RNA fragment. The extension products were subjected to denaturing gel electrophoresis adjacent to a sequencing ladder generated with the same DNA and primer. The base corresponding to the 5′ end of the protected RNA fragment can be determined by comparison to the adjacent sequencing ladder.

Figures 1 to 4 are autoradiograms showing the detection of numerous ribosome pause sites along each of the four ATP synthase genes. Although the entire sequence of the four ATP synthase genes was examined for pauses, only areas where ribosome pauses were found are represented in the

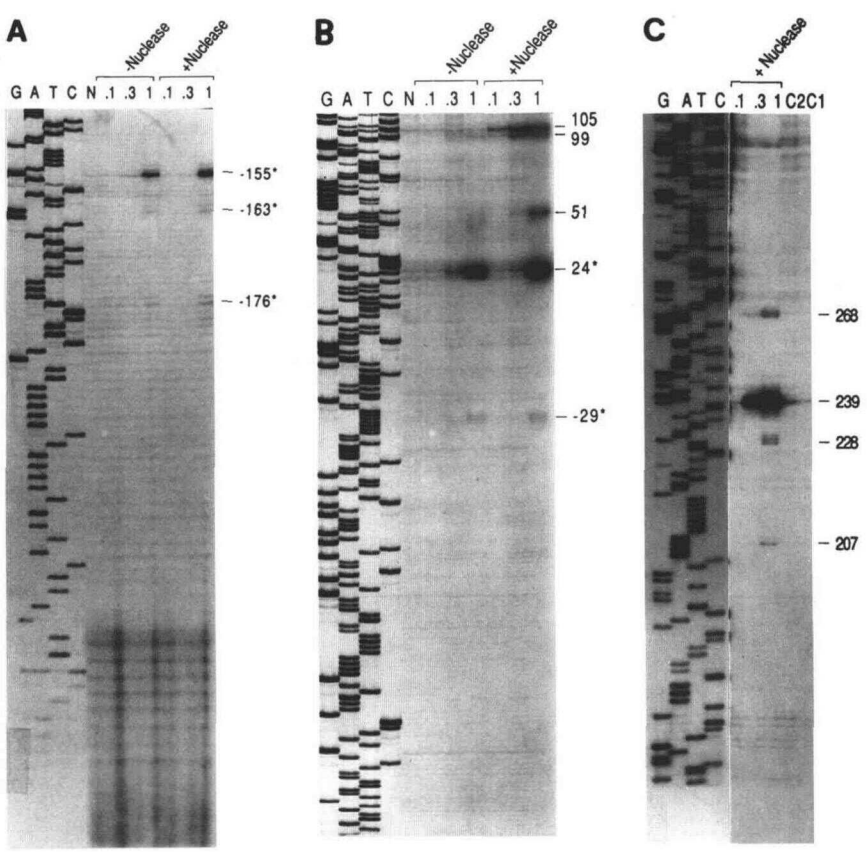

Figure 1. Mapping of ribosome pauses upstream and within the *atpI* open reading frame. A, Bases −216 through −141; B, bases −117 through +117; C, bases 171 through 412. G, A, T, and C indicate DNA sequence generated using the same primer and single-stranded DNA as in the extension reaction. N indicates extension reactions performed with the same primer and single-stranded DNA without nuclease-resistant RNA fragments. −Nuclease or C2 indicates primer extension performed in the presence of extracted, mock nuclease-treated RNAs. +Nuclease lanes contain extension products generated in the presence of RNA fragments isolated from nuclease-treated polysomes. C1 indicates primer extension reactions that included phenol/chloroform-extracted and nuclease-treated polysomal RNA. Micrograms of protected RNA fragments or total chloroplast are indicated above the lanes. 5′ ends are indicated with asterisks (*). Numbers are relative to the first base of the *atpI* open reading frame.

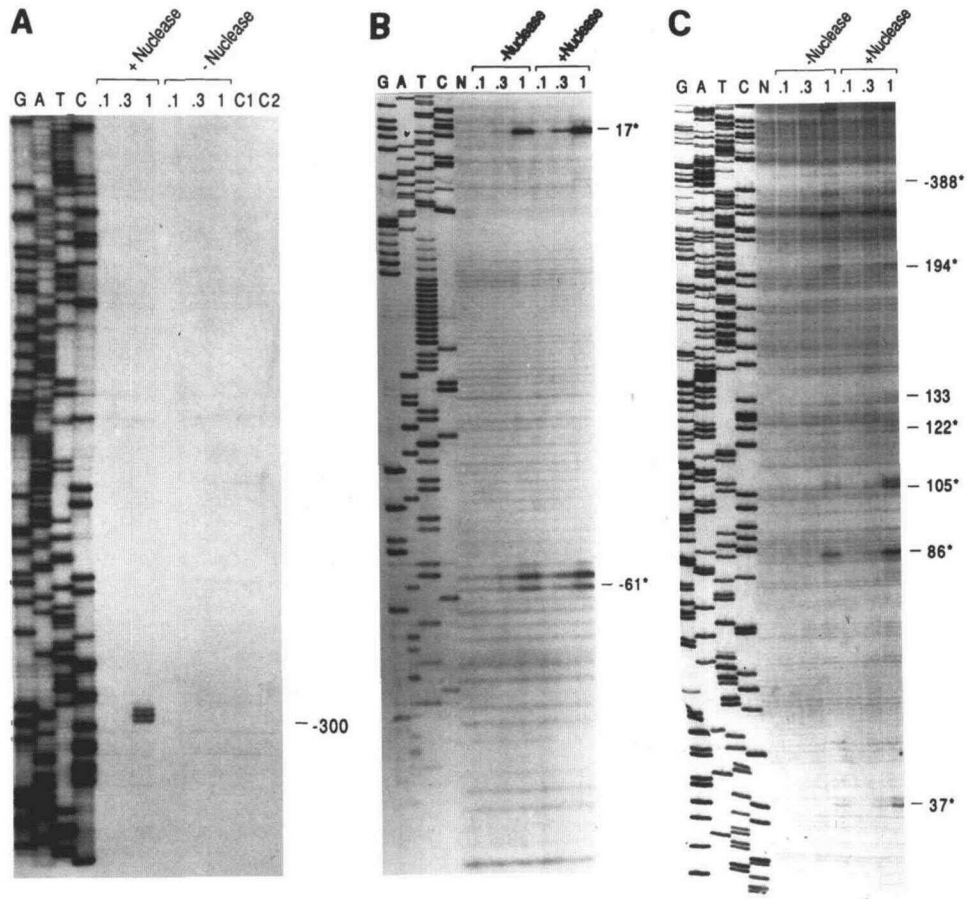

Figure 2. Mapping of ribosome pauses upstream and within the *atp*H open reading frame. A, −328 through −98; B, −77 through 32; C, 20 (*atp*H) through −304 (*atp*F). Lane designations are defined in the legend for Figure 1. 5′ ends are indicated with asterisks (*). Numbers are relative to the first base of the *atp*H open reading frame except −388, which is relative to the start of the *atp*F open reading frame.

figures. The majority of the sites were located within the open reading frames rather than in the intergenic spaces. The sites from the 5′ upstream region through the first 300 bases of *atp*A have been previously reported, and those autoradiograms are not shown here (Kim and Hollingsworth, 1992). In some cases (for example, base −176 in Fig. 1A), multiple bands were detected immediately adjacent to each other. These multiple bands are assigned as a single stop, because we believe that they are caused by small variations in the extent of micrococcal nuclease digestion. Figure 5 is a composite diagram showing the location of each ribosome pause site relative to the open reading frames of the large ATP synthase gene cluster and to each other. Bands that were consistently too faint to photographically reproduce were omitted from the composite figure.

5′ End Mapping

The primary transcripts from the large ATP synthase gene cluster undergo extensive processing, resulting in approximately 40 unique RNA species (Hudson et al., 1987; Stollar and Hollingsworth, 1994). The composition of each of these RNA species has been determined (Hudson et al., 1987; Stollar and Hollingsworth, 1994), but the 5′ and 3′ ends of these transcripts have not been completely localized. In addition to being able to map positions where ribosomes pause in vivo, this primer extension assay has enabled us to map several of the 5′ ends from the large ATP synthase gene cluster found in vivo in spinach chloroplasts. The 5′ end mapping assay was performed in the same manner as the protected fragment localization with the exception that undigested RNA (rather than protected fragments) was annealed to the DNA template. Therefore, extension products terminated at positions that correspond to the in vivo 5′ ends of the RNA molecules. There were several 5′ ends located among and within the four ATP synthase genes (Figs. 1–4). The positions of the 5′ ends are shown in the composite diagram (Fig. 5).

It was surprising to discover that at every location where a 5′ end was mapped in this gene cluster, a nuclease-resistant fragment was also observed at the same position (Figs. 1–4). Thus, it appears that these 5′ ends are bound by proteins in vivo. The proteins that protect the 5′ ends of the transcripts from this cluster are not necessarily ribosomes. However, the protected fragments were isolated from the pellet remaining after centrifugation through two Suc cushions. If the protein(s) protecting the 5′ ends of these transcripts are not

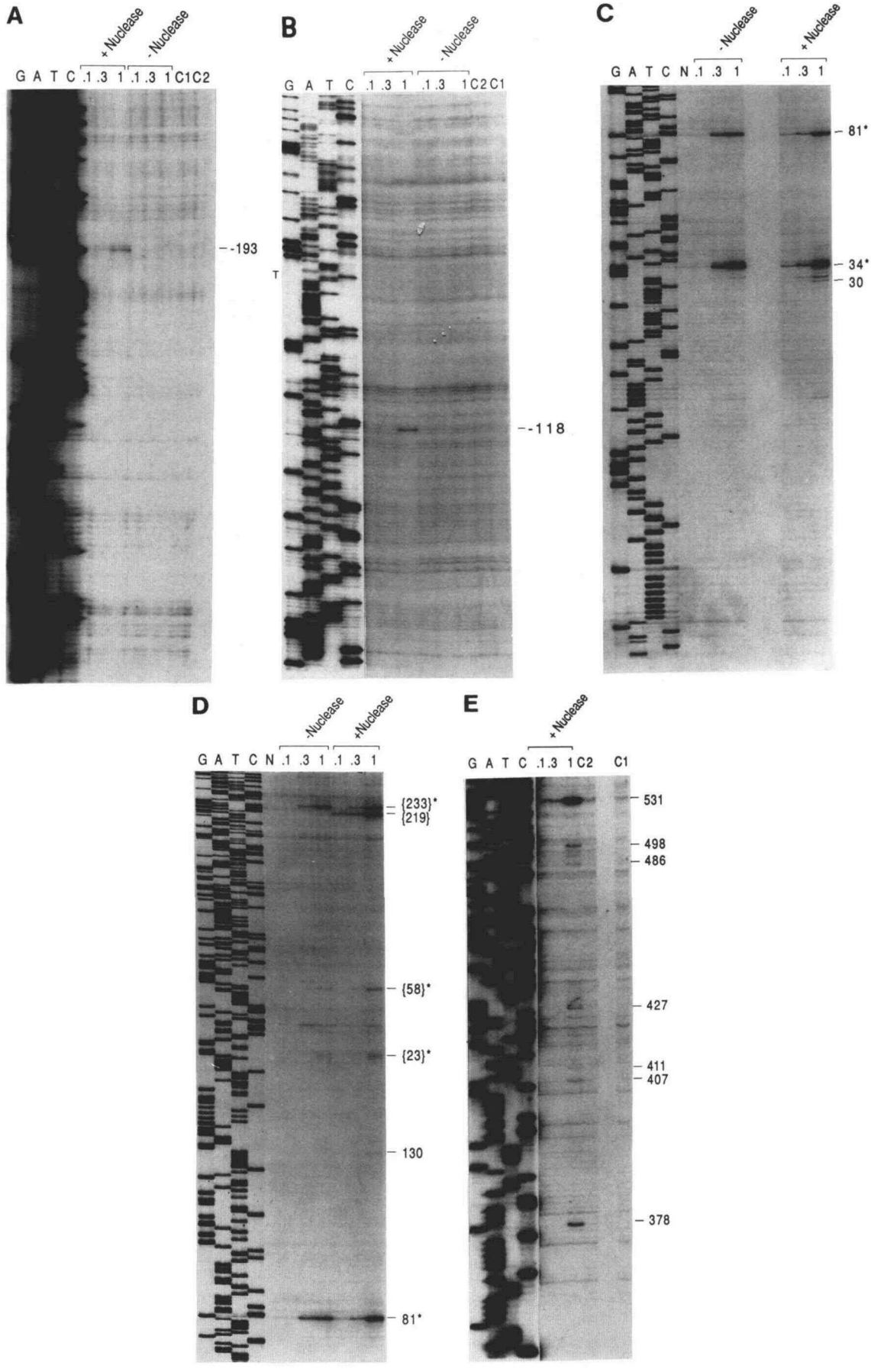

Figure 3. Legend on facing page.

ribosomes, they must be involved in another complex of significant size and density.

Quantitation of Ribosome Pauses within Each Gene of the Cluster

To ascertain whether there is any correlation between ribosome pausing and the final ATP synthase protein levels, the relative number of ribosome pauses within each open reading frame was determined. The ribosome mapping experiments revealed that the primer extension reaction products corresponding to protected RNA fragments reproducibly vary in intensity. Therefore, the overall degree of pausing within any given open reading frame is not directly proportional to the number of pauses. The intensity of pausing at a given site must also be taken into account.

To quantitate the ratio of the extent of pausing along each ATP synthase open reading frame, dot-blot hybridizations were performed. Four sets of nuclease-resistant RNA fragments, each at three different concentrations, were spotted onto strips of membrane filters. Each set was hybridized to a single radiolabeled RNA probe complementary to an individual open reading frame. In addition to the ribosome-protected RNA fragments spotted onto the filters, an in vitro-transcribed RNA corresponding to the entire length of the gene cluster was also applied to the filters and used for normalization of the hybridizations.

Figure 6 shows the results of the dot-blot hybridizations. Each spot from the hybridized filters was excised and its radioactivity was determined by liquid scintillation. The experimental counts were normalized to those from the control RNA hybridized to the same probe. This normalization served to factor out variations in net hybridization due to differing specific activities of the probes and differing lengths of the four open reading frames. Protected fragments from the *atp*H open reading frame were determined to have the lowest extent of normalized overall hybridization. The ratio of the number of protected fragments within each of the ATP synthase genes of the cluster was therefore calculated with respect to the *atp*H gene, yielding a ratio of hybridization to each of the four open reading frames relative to *atp*H. The average values from three independent experiments are shown in Figure 6. The ratio of the relative number of protected fragments within each open reading frame was determined to be 3:1:3:5 (*atp*I:*atp*H:*atp*F:*atp*A, respectively). The polypeptide ratio of these four open reading frames in the ATP synthase complex is 1:5–12:1:3 (McCarty et al., 1988).

Two control analyses were performed to investigate whether the pausing ratios could be caused by factors extraneous to the ribosomes themselves. For example, if the abundance of *atp*H transcripts is very low compared to the other genes, this could explain the relatively low number of ribosome-protected fragments observed within the coding region of *atp*H. The relative abundance of transcripts from each of the four ATP synthase genes was determined in a manner similar to the quantitation of the relative number of ribosome pause sites, except that the dot blots contained total chloroplast RNA rather than micrococcal nuclease-resistant polysome fragments. Values normalized to the lengths of the probes indicated that transcripts were present in a ratio of 1:3:3:6 (*atp*I:*atp*H:*atp*F:*atp*A, respectively) (data not shown). Since this is significantly different from the ribosome pausing ratio (3:1:3:5), we conclude that ribosome pausing is not correlated with the abundance of a given RNA. Another mechanism that might extraneously affect the overall pausing would be if ribosome binding was proportional to the length of the transcripts. If that were true, the number of pauses per nucleotide would be equivalent for all four genes. However, the *atp*H and *atp*F open reading frames, which are the shortest of the four genes, were determined to have the most pauses per nucleotide, 0.03, whereas *atp*I and *atp*A, the longer two open reading frames, had 0.01 pauses per nucleotide. Therefore, the observed degree of pausing cannot be attributed to the length of the open reading frame. We conclude that the ratios generated in the dot-blot analysis represent real differences in the numbers of protected fragments isolated from these open reading frames and are not due to intrinsic differences in the transcripts themselves.

DISCUSSION

The five genes of the large ATP synthase gene cluster are co-transcribed from two locations (Stollar and Hollingsworth, 1994). Although the relative strengths of these two promoters are unknown, transcription from them would result in an initial ratio of transcripts significantly different from the final ratio of the polypeptides. Therefore, posttranscriptional mechanisms are necessary to control the final ratios of the ATP synthase polypeptides from this cluster. In this set of experiments, we have investigated ribosome pausing as one posttranscriptional mechanism that could affect the expression of this cluster. We have shown that there are numerous sites where ribosomes pause, that the majority of these sites are within the open reading frames, and that the pauses are distributed in a nonuniform fashion.

If ribosome pausing were the primary mechanism controlling expression of the ATP synthase genes from this cluster, then based on the protein ratios it would be expected that: (a) the *atp*H gene would have the least amount of pausing along its open reading frame; (b) the degree of ribosome pausing within the *atp*I and *atp*F genes would be equivalent and greater than those found in *atp*H; and (c) the *atp*A gene would have an amount of pausing intermediate to the amount

Figure 3. (On facing page). Mapping of ribosome pauses upstream and within the *atp*F gene. A, −281 through −129; B, −178 through 25; C, −37 through 107; D, 64 (exon 1) through {288} (intron); E, 357 through 550 (exon 2). Lane designations are defined in the legend for Figure 1. 5′ ends are indicated with asterisks (*). Numbers are relative to the first base of the *atp*F open reading frame. Protected RNA fragments localized within the intron are indicated by braces (e.g. {233}) and are relative to the first base of the intron. Bases in the *atp*F open reading frame were numbered as they would be after splicing. The intron sequence was numbered starting with {1} as the first base of the intron.

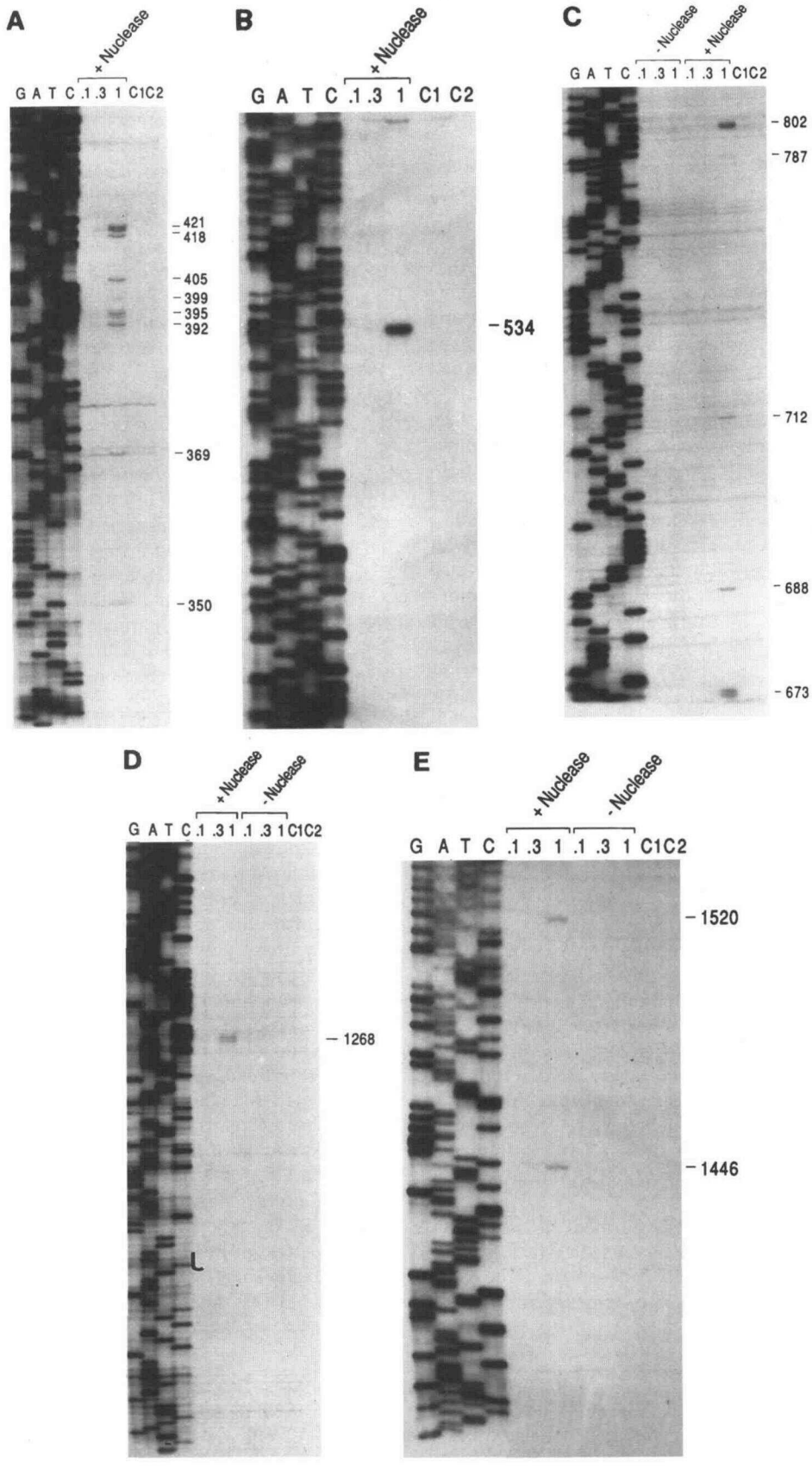

Figure 4. Legend appears on facing page.

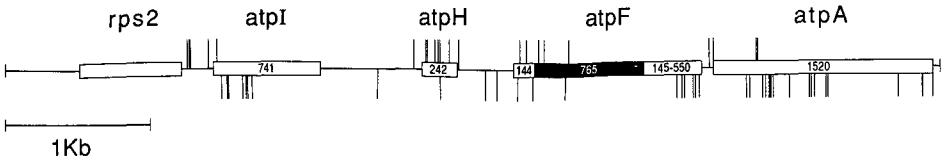

Figure 5. Location of mapped ribosome pause sites and 5′ ends along the large ATP synthase gene cluster. The lines below the gene cluster indicate the positions of mapped pause sites, and the lines above indicate the positions of nuclease-resistant fragments that also correspond to 5′ ends. Open boxes represent open reading frames, and the filled box represents the group II intron within the *atp*F gene. The number within each box represents the number of bp encoded by that portion of DNA. Ribosome pauses were not mapped for *rps*2.

found in *atp*H and the amount found in both *atp*I and *atp*F. Consistent with this hypothesis, the dot-blot data indicate that *atp*H has the least amount of pausing, and the *atp*I and *atp*F genes have an equivalent degree of pausing, which is 3-fold higher than that found in *atp*H. In contrast, pausing within the coding region of the *atp*A gene was found to be the greatest among the four ATP synthase genes of the cluster. These results suggest that although ribosome pausing may be a significant mechanism affecting the expression of these ATP synthase genes, there must be other factors that also affect the final ratios of the polypeptides in the complex.

In the process of mapping ribosome pause sites, it became evident that several of the nuclease-resistant RNA fragments corresponded to the location of in vivo 5′ ends, indicating that the 5′ ends were bound by either ribosomes or some other factor(s). Recently, a set of nuclear-encoded factors have been found that interact with the 5′ leaders of chloroplast mRNA from *Chlamydomonas reinhardtii*. They have been suggested to regulate translation (Danon and Mayfield, 1991). If the 5′ end protection observed here is not due to ribosomes, it must be due to a complex of similar or greater size/density, because the protected fragments are analyzed after centrifugation through two Suc cushions. Relatively small protein-RNA complexes (roughly 6S) will not pass through these cushions (C.T. Baumann and P.D. Gollnick, personal communication). The exact nature of the source of the 5′ end protection within spinach chloroplasts is currently under investigation.

As is the case with the transcript 5′ ends, it is possible that some of the protected fragments mapped along this gene cluster are due to factors other than ribosome pausing. For example, several protected fragments are located in intergenic regions, where ribosomes are typically not expected to be bound. However, translational coupling, where a ribosome does not dissociate after finishing a polypeptide but continues onto a downstream open reading frame and initiates translation there, is a well-known phenomenon in prokaryotic gene expression. In particular, translational coupling is a critical factor in the control of expression of the *atp* operon in *Escherichia coli* (McCarthy, 1990). Translational coupling has also been proposed to be involved in the expression of the spinach chloroplast *atp*B/E gene cluster (Zurawski et al., 1982). Additional protected sites that might not be caused by ribosomes include regions near and within the group II intron of *atp*F. Protection of these sites could be provided by putative splicing factors binding to the pre-mRNA, if a complex of these factors were large and dense enough to pass through the two Suc cushions.

Despite the possibility that a few of the protected fragments may be bound by nonribosomal complexes, the most logical explanation for the majority of the protected fragments is that they are caused by paused ribosomes. Translational control, such as that described here, has been shown to occur in several chloroplast systems in vivo. Translation of the large subunit protein of the *rbc*L in amaranth seedlings is observed only when the seedlings are grown in the light. When plants are transferred to the dark, no new *rbc*L protein is synthesized, despite the fact that the *rbc*L mRNA is still bound to polysomes (Berry et al., 1988). A similar observation was made by Klein et al. (1988), who noted that the polypeptide products of the *psa*A and *psa*B genes were not detected in plastids of dark-grown barley seedlings even though transcripts for these proteins were present and associated with membrane-bound polysomes. This suggests that control of expression of these three genes occurs at the translation elongation step. In addition, Kim et al. (1991) have shown that ribosomes pause during the translation of the chloroplast reaction center protein D1, the product of *psb*A, and they hypothesize that the pausing may facilitate the co-translational insertion into the thylakoid membrane or to the binding of cofactors such as Chl.

Several factors have been previously suggested to cause ribosomes to pause. These include translation initiation sequences (Wolin and Walter, 1988; Schaefer et al., 1989), translation termination codons (Wolin and Walter, 1988), mRNA secondary structure (Jacks et al., 1988; Fu et al., 1991; Tu et al., 1992), rare codons (Varenne et al., 1984), membrane insertion (Kim et al., 1991), or cofactor binding (Kim et al., 1991). Analysis of the exact nature of factors that influence ribosome pausing at a particular site may be hampered due

Figure 4. (On facing page). Mapping of ribosome pauses upstream and within the *atp*A open reading frame. A, 336 through 481; B, 431 through 647; C, 672 through 829; D, 1201 through 1343; E, 1378 through 1535 (*atp*A ends at base 1521). Lane designations are defined in the legend for Figure 1. 5′ ends are indicated with asterisks (*). Numbers are relative to the first base of the *atp*A open reading frame. Ribosome pauses in bases −40 through 299 have been reported elsewhere (Kim and Hollingsworth, 1992).

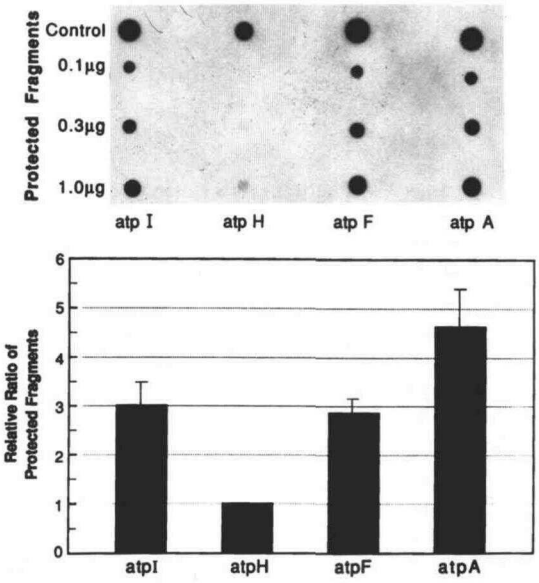

Figure 6. Quantitation of ribosomal pauses within the four ATP synthase genes of the cluster. Upper, Dot-blot hybridization. Each membrane contains nuclease-resistant RNA fragments at concentrations of 0.1, 0.3, and 1.0 µg, plus, as control, 1.0 µg of an in vitro-transcribed RNA spanning the entire gene cluster. Blots were hybridized with radiolabeled RNA probes as indicated. Lower, Representation of the extent of hybridization to the dot blots. The ratio for the extent of hybridization within each open reading frame was determined and the average from three experiments is shown.

to multiple ribosomes stacked up behind one another, thus shifting the actual location of the initial pause site upstream (Wolin and Walter, 1988). However, the fact that the pause sites detected were identical over a 10-fold range of protected fragments (and were virtually undetectable at the lower concentrations) implies that the observed stops were due to primary pause sites.

The nucleotide sequences surrounding each ribosome pause site were examined for potential causes of the pausing event. Pause sites were also examined to determine whether there was any correlation between the relative strength of a pause (i.e. darkness of a band) and particular factors that might affect pausing. *Cis*-acting factors that were investigated include the presence of rare codons and putative secondary structure. The locations of the pauses were also compared to the positions of Shine-Dalgarno sequences/translation initiation sites and to translation termination codons (Bonham-Smith and Bourque, 1989). In addition, the nucleotide sequence was examined for the possibility that pauses could correlate with co-translational insertion of the nascent polypeptide into the thylakoid membrane or by interference with hypothesized splicing factors bound near the 5′ splice site of the *atp*F intron.

Codon abundance in a given organism appears to be correlated to the abundance of the cognate tRNA (Ikemura, 1985). It has been shown in some systems that low-abundance codons can cause a ribosome to hesitate while waiting for the appearance of the appropriately charged cognate tRNA (Lizardi et al., 1979). The positions of ribosome pauses were investigated with respect to the positions of rare codons. Rare codons were assigned after generation of a codon frequency chart using the program CODONFREQUENCY (Devereux et al., 1984) on the entire population of sequenced spinach chloroplast open reading frames. Rare codons were defined as those that matched the criteria proposed by Zhang et al. (1991). Several rare codons were detected that were located within 14 to 24 nucleotides downstream of the 5′ ends of protected fragments. Considering that ribosomes would protect between 30 and 35 bases from nuclease digestion (reviewed by Steitz, 1979), this would place the rare codons somewhere within the ribosome. There were four cases in this gene cluster in which two or three rare codons were found immediately adjacent to each other (*atp*I: 111–116, 240–248, 273–279; *atp*A: 1284–1289). In all four cases, a ribosome pause was mapped at that site. However, there were many single rare codons within the open reading frames of the four ATP synthase genes that were not associated with ribosome pauses.

Although the presence or absence of secondary structure in vivo is often shown to differ from that predicted by computer analysis, potential secondary structures located just 3′ to the site of paused ribosomes were examined using the computer programs FOLD (Zucker and Stiegler, 1981) and SQUIGGLES (Devereux et al., 1984). Most of the pauses were associated with potential downstream secondary structures with a ΔG at 25°C of less than −19 kcal/mol. However, there are other predicted secondary structures that do not correlate with protected fragments. There was no obvious structural consensus among the computer-predicted secondary structures associated with protected fragments.

Kim et al. (1991) have proposed that ribosomes pause during co-translational insertion of some chloroplast polypeptides into the thylakoid membrane. The protein product of the *atp*F gene, CF_0-I, is embedded in the thylakoid membrane (Strotmann and Bickel-Sandkotter, 1984), and, when fully processed, lacks the first 17 amino acid residues (Bird et al., 1985). The pause located at +378 of the *atp*F coding region corresponds to a polypeptide chain of 128 amino acid residues. About 40 amino acid residues of the nascent polypeptide have to pass through a ribosome before its amino terminus is exposed (Yonath et al., 1987). In this case, about 88 amino acid residues of the nascent polypeptide chain would be extruded from the ribosome, and these include a putative membrane-spanning domain (Hudson et al., 1987). Therefore, it is possible that the pause at +378 is due to co-translational insertion into the thylakoid membrane.

In summary, we have shown that ribosome pausing, although not the sole factor, is a posttranscriptional mechanism that may significantly affect the expression of the large ATP synthase gene cluster in spinach chloroplasts. No single potential cause for ribosome pausing could be identified. Pausing is undoubtedly influenced by a complex set of both *cis*- and *trans*-acting factors. The identity of these factors and the extent to which they affect translational elongation within the chloroplast are currently under investigation.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Drs. Gerald Koudelka and Jim Berry for critical reading of the manuscript and many helpful

comments. We thank Christopher Baumann and Dr. Paul Gollnick for communication of results prior to publication. We also thank Jim Stamos for his excellent illustration and photography work.

Received January 24, 1994; accepted April 8, 1994.
Copyright Clearance Center: 0032-0889/94/105/1167/11.

LITERATURE CITED

Adams CC, Stern DB (1990) Control of mRNA stability in chloroplasts by 3′ inverted repeats: effects of stem and loop mutations on degradation of psbA mRNA in vitro. Nucleic Acids Res **18:** 6003–6010

Alberts BM (1984) The DNA enzymology of protein machines. Cold Spring Harbor Symp Quant Biol **49:** 1–12

Bergmann JE, Lodish H (1979) A kinetic model of protein synthesis: application to hemoglobin synthesis and translational control. J Biol Chem **254:** 11927–11937

Berry JO, Carr JP, Klessig DF (1988) mRNAs encoding ribulose-1,5-bisphosphate carboxylase remain bound to polysomes but are not translated in amaranth seedlings transferred to darkness. Proc Natl Acad Sci USA **85:** 4190–4194

Bird CR, Koller B, Auffret AD, Huttly AK, Howe CJ, Dyer TA, Gray JC (1985) The wheat chloroplast gene for CF_0 subunit I of ATP synthase contains a large intron. EMBO J **4:** 1381–1388

Bonham-Smith PC, Bourque DP (1989) Translation of chloroplast-encoded mRNA: potential initiation and termination signals. Nucleic Acids Res **17:** 2057–2080

Danon A, Mayfield SPY (1991) Light regulated translational activators: identification of chloroplast gene specific mRNA binding proteins. EMBO J **10:** 3993–4001

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res **12:** 387–395

Fu LN, Ye RQ, Browder LW, Johnston RN (1991) Translational potentiation of messenger RNA with secondary structure in *Xenopus*. Science **251:** 807–810

Gruissem W, Tonkyn JC (1993) Control mechanisms of plastid gene expression. Crit Rev Plant Sci **12:** 19–55

Hallick RB, Bottomley W (1983) Proposals for the naming of chloroplast genes. Plant Mol Biol Rep **1:** 38–43

Hallick RB, Chelm BK, Gray PW, Orozco EM Jr (1977) Use of aurintricarboxylic acid as an inhibitor of nucleases during nucleic acid isolation. Nucleic Acids Res **4:** 3055–3064

Hennig J, Herrmann RG (1986) Chloroplast ATP synthase of spinach contains nine nonidentical subunit species, six of which are encoded by plastid chromosomes in two operons in a phylogenetically conserved arrangement. Mol Gen Genet **203:** 117–128

Herrmann RG, Westhoff P, Alt J, Winter P, Tittgen J, Bisanz C, Sears BB, Nelson N, Hurt E, Hauska G, Viebrock A, Sebald W (1983) Identification and characterization of genes for polypeptides of the thylakoid membrane. *In* O Cifferi, L Dure III, eds, Structure and Function of Plant Genomes. Plenum Publishers, New York, pp 143–153

Hudson GS, Mason JG (1988) The chloroplast genes encoding subunits of the H^+-ATP synthase. Photosynth Res **18:** 205–222

Hudson GS, Mason JG, Holton TA, Koller B, Cox GB, Whitfeld PR, Bottomley W (1987) A gene cluster in the spinach and pea chloroplast genomes encoding one CF_1 and three CF_0 subunits of the H^+-ATP synthase complex and the ribosomal protein S2. J Mol Biol **196:** 283–298

Ikemura T (1985) Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol **2:** 13–34

Jacks T, Madhani HD, Masiarz FR, Varmus HE (1988) Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. Cell **55:** 447–458

Kim J, Klein PG, Mullet JE (1991) Ribosomes pause at specific sites during synthesis of membrane-bound chloroplast reaction center protein D1. J Biol Chem **266:** 14931–14938

Kim JK, Hollingsworth MJ (1992) Localization of in vivo ribosome pause sites. Anal Biochem **206:** 183–188

Kim JK, Hollingsworth MJ (1993) Splicing of group II introns in spinach chloroplasts (in vivo): analysis of lariat formation. Curr Genet **23:** 175–180

Klein RR, Mason HS, Mullet JE (1988) Light regulated translation of chloroplast proteins. I. Transcripts of psaA-psaB, psbA, and rbcL are associated with polysomes in dark-grown and illuminated barley seedlings. J Cell Biol **106:** 289–301

Lizardi PM, Mahdavi V, Shields D, Candelas G (1979) Discontinuous translation of silk fibroin in a reticulocyte cell-free system and in intact silk gland cells. Proc Natl Acad Sci USA **76:** 6211–6215

Lodish HF, Jacobsen M (1972) Regulation of hemoglobin synthesis. Equal rates of translation and termination of α- and β-globin chains. J Biol Chem **247:** 3622–3629

McCarthy JEG (1990) Post-transcriptional control in the polycistronic operon environment: studies of the atp operon of *Escherichia coli*. Mol Microbiol **4:** 1233–1240

McCarty RE, Shapiro AB, Yu F (1988) Regulation and mechanism of the chloroplast ATP synthase in relation to function. *In* SE Stevens Jr, DA Bryant, eds, Light-Energy Transduction in Photosynthesis: Higher Plant and Bacterial Models, The American Society of Plant Physiologists, Rockville, MD, pp 290–304

Michel F, Dujon B (1983) Conservation of RNA secondary structure in two intron families including mitochondrial-, chloroplast-, and nuclear-encoded members. EMBO J **2:** 33–38

Michel F, Umesono K, Ozeki H (1989) Comparative and functional anatomy of group II catalytic introns—a review. Gene **82:** 5–30

Orozco EM Jr, Mullet JE, Hanley-Bowdoin L, Chua N-H (1986) In vitro transcription of chloroplast protein genes. Methods Enzymol **118:** 232–253

Schaefer M, Hartz D, Gold L, Simoni RD (1989) Ribosome-binding sites and RNA-processing sites in the transcript of the *Escherichia coli unc* operon. J Bacteriol **171:** 3901–3908

Senior AE (1985) The proton-ATPase of *Escherichia coli*. Curr Top Membr Transp **23:** 135–151

Steitz JA (1979) RNA-RNA interactions during polypeptide chain initiation. *In* G Chambliss, GR Craven, J Davies, K Davis, L Kahan, M Nomura, eds, Ribosomes: Structure, Function, and Genetics. University Park Press, Baltimore, MD, pp 479–495

Stern DB, Jones H, Gruissem W (1989) Function of plastid mRNA 3′ inverted repeats. RNA stabilization and gene-specific protein binding. J Biol Chem **264:** 18742–18750

Stern DB, Radwanski ER, Kindle KL (1991) A 3′ stem/loop structure of the *Chlamydomonas* chloroplast atpB gene regulates messenger RNA accumulation in vivo. Plant Cell **3:** 285–297

Stollar NE, Hollingsworth MJ (1994) Expression of the large ATP synthase gene cluster from spinach chloroplasts. J Plant Physiol (in press)

Strotmann H, Bickel-Sandkotter S (1984) Structure, function and regulation of chloroplast ATPase. Annu Rev Plant Physiol **35:** 97–120

Tu C, Tzeng TH, Bruenn JA (1992) Ribosomal movement impeded at a pseudoknot required for frameshifting. Proc Natl Acad Sci USA **89:** 8636–8640

Varenne S, Buc J, Lloubes R, Lazdunski C (1984) Translation is a non-uniform process. Effects of tRNA availability on the rate of elongation of nascent polypeptide chains. J Mol Biol **180:** 549–576

Wolin SL, Walter P (1988) Ribosome pausing and stacking during translation of a eukaryotic mRNA. EMBO J **7:** 3559–3569

Yonath A, Lenard KR, Wittmann HG (1987) A tunnel in the large ribosomal subunit revealed by three-dimensional image reconstruction. Science **236:** 813–816

Zhang S, Zubay G, Goldman E (1991) Low-usage codons in *Escherichia coli*, yeast, fruit fly and primates. Gene **105:** 61–72

Zuker M, Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res **9:** 133–148

Zurawski G, Bottomley W, Whitfeld PR (1982) Structures of the genes for the beta and epsilon subunits of spinach chloroplast ATPase indicate a dicistronic mRNA and an overlapping translation stop/start signal. Proc Natl Acad Sci USA **79:** 6260–6264