Mol Gen Genet (1989) 217:77–84

MGG

The plastid *rpo*A gene encoding a protein homologous to the bacterial RNA polymerase alpha subunit is expressed in pea chloroplasts

Saul Purton* and John C. Gray
Botany School, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK

Summary. The gene *rpo*A, encoding a protein homologous to the alpha subunit of RNA polymerase from *Escherichia coli* has been located in pea chloroplast DNA downstream of the *pet*D gene for subunit IV of the cytochrome *b-f* complex. Nucleotide sequence analysis has revealed that *rpo*A encodes a polypeptide of 334 amino acid residues with a molecular weight of 38916. Northern blot analysis has shown that *rpo*A is co-transcribed with the gene for ribosomal protein S11. A *lac*Z–*rpo*A gene-fusion has been constructed and expressed in *E. coli*. Antibodies raised against the fusion protein have been employed to demonstrate the synthesis of the *rpo*A gene product in isolated pea chloroplasts. Western blot analysis using these antibodies and antibodies against the RNA polymerase core enzyme from the cyanobacterium, *Anabaena* 7120, has revealed the presence of the gene product in a crude RNA polymerase preparation from pea chloroplasts.

Key words: Chloroplast – RNA polymerase – *rpo*A gene

Introduction

The synthesis of multisubunit complexes within the chloroplast requires the co-ordinated expression of plastid and nuclear genes. Consequently, the study of the chloroplast transcription apparatus is of major importance for the understanding of chloroplast biogenesis. However, major uncertainties remain with regard to the number of RNA polymerases present in the chloroplast, the subunit structure of the enzyme(s) and the sites of synthesis of the subunits.

The recent discovery that the chloroplast genome contains genes encoding homologues of the core components of *Escherichia coli* RNA polymerase (Sijben-Müller et al. 1986; Shinozaki et al. 1986; Ohyama et al. 1986) has led to the suggestion that the plastid RNA polymerase has a similar subunit structure to that of eubacteria. This idea is supported by antibody studies demonstrating that the *E. coli* and plastid enzymes are immunochemically related (Briat et al. 1987) and by the identification, in chloroplast extracts, of sigma factors which confer specificity to the *E. coli* RNA polymerase core complex (Surzycki and Shellenbarger 1976; Bülow and Link 1988).

* *Present address:* Departément de Biologie Moléculaire, 30 quai Ernest Ansermet, CH-1211 Genève 4, Switzerland

Offprint requests to: J.C. Gray

The suggestion that the core components of the plastid RNA polymerase are chloroplast-encoded contradicts earlier evidence that the RNA polymerase is nuclear-encoded (Ellis and Hartley 1971; Bünger and Feierabend 1980; Siemenroth et al. 1981; Lerbs et al. 1985). Furthermore, the isolation of different polymerase activities specific for either ribosomal RNA genes or transfer RNA genes (Greenberg et al. 1985) has led several workers to propose that the chloroplast contains more that one RNA polymerase (Greenberg et al. 1984; Lerbs et al. 1988).

Consequently, we have undertaken an investigation of the role of the pea chloroplast genome in encoding subunits of the plastid RNA polymerase. Here we present the sequence of the pea *rpo*A gene and demonstrate its transcription in vivo together with the upstream ribosomal protein genes *rpl*36 and *rps*11. Antibodies have been raised against the *rpo*A gene product synthesised in *E. coli* as a β-galactosidase fusion protein. Using these antibodies, the synthesis of the *rpo*A gene product in isolated chloroplasts and its presence in a crude chloroplast RNA polymerase preparation have been demonstrated.

Materials and methods

Plant material. Peas (*Pisum sativum* var. Feltham First) were grown in Fisons Levington Compost with 16 h illumination daily. Leaf material was harvested after 7 days.

Recombinant DNA. Plasmid pPscHS5 contains a 3.5 kb *Hae*II–*Sal*I fragment of pea plastid DNA (Phillips and Gray 1984) inserted into the *Sal*I site of pUC8. Plasmid pPscEK2 was prepared by inserting a 2.0 kb *Eco*RI–*Kpn*I sub-fragment of pPscHS5 into pUC18. Plasmid pPscT920 was made by inserting a 920 bp *Taq* I fragment of pPscEK2 into pUC18. Recombinant DNA was introduced into *E. coli* strain TG1 by transformation (Cohen et al. 1972) and subsequently isolated by the alkaline-SDS method of Birnboim and Doly (1979). The expression plasmid pEXT920 was prepared by inserting a 940 bp *Bam*HI–*Pst*I fragment of pPscT920 into pEX3 (Stanley and Luzio 1984) and was introduced into the *E. coli* strain pop2136.

DNA sequence determination. Restriction fragments isolated from agarose and polyacrylamide gels were inserted into the appropriate sites in M13 strains tg130 and tg131 (Kieny et al. 1983) and the hybrid DNA used to transform *E. coli*

strain TG1. Isolation of single-stranded DNA and sequencing by the dideoxynucleotide chain termination method were carried out as described by Sanger et al. (1980) with the modifications of Biggin et al. (1983).

Preparation of chloroplast RNA. RNA was isolated from chloroplasts prepared from 7-day-old pea leaves as described by Blair and Ellis (1973) except that sorbitol was substituted for sucrose in the isolation medium to avoid the problem of RNase-contaminated sucrose. RNA was extracted as described by Covey and Hull (1981) except that a chloroplast pellet rather than leaf material was used as the RNA source.

Northern blot analysis. Chloroplast RNA (20 μg) was glyoxylated and electrophoresed on 1.1% agarose gels as described by McMaster and Carmichael (1977). RNA was transferred to Hybond-N nylon filters (Amersham International, Amersham, UK) and hybridised with labelled DNA fragments as described by Newman and Gray (1988). Single-stranded DNA probes for each of the three genes were prepared by the 'prime-cut' method (Biggin et al. 1984).

Isolation of fusion protein and preparation of antisera. The fusion protein resulting from the expression of pEXT920 in *E. coli* was excised from a 7.5% polyacrylamide gel after electrophoresis of a crude cell extract (Stanley and Luzio 1984) in the presence of sodium dodecyl sulphate (SDS). Protein was electroeluted from the gel strips and dialysed against several changes of 0.01% (w/v) SDS. Antibodies were raised by intramuscular injections of approximately 50 μg of eluted protein emulsified with Freunds complete adjuvant into a New Zealand White rabbit at 14-day intervals. Antibodies to the RNA polymerase core enzyme of *Anabaena* 7120 were a gift from R. Haselkorn (University of Chicago, USA).

Coupled transcription-translation. Coupled transcription-translation of plasmid pPscEK2 in a cell-free system from *E. coli* strain PR7, subsequent immunoprecipitation using antibodies to the fusion protein and protein A-Sepharose, and electrophoresis in a 15% polyacrylamide gel in the presence of SDS were carried out as described by Howe et al. (1982).

Protein synthesis in isolated chloroplasts. A chloroplast preparation capable of light-driven protein synthesis was prepared from 15 g of 7-day-old pea leaves as described by Blair and Ellis (1973). The chloroplast preparation was diluted to a chlorophyll concentration of 200 μg/ml and incubated in the light at 22° C in the presence of 150 μCi L-[^{35}S]methionine. Illumination was provided by a 500 W tungsten halogen lamp at a distance of 20 cm. Protein was extracted by lysis of the chloroplasts in 5 volumes of 'lysis buffer' (Orozco et al. 1985) followed by the addition of NaCl to a final concentration of 1 M. Membrane material was removed by centrifugation. Immunochemically related proteins were precipitated using the anti-fusion protein antiserum and protein A-Sepharose essentially as described by Howe et al. (1982). Radiolabelled proteins were electrophoresed in a 10% polyacrylamide gel in the presence of SDS (Laemmli 1970) and detected by fluorography (Bonner and Laskey 1974).

Preparation of extracts from intact pea chloroplasts and from Anacystis nidulans. A crude RNA polymerase extract active in transcription was prepared from 7-day-old pea chloroplasts by the method of Orozco et al. (1985) with the modifications of Lam and Chua (1987). A crude protein extract from the cyanobacterium *Anacystis nidulans* was prepared by sonication in 'lysis buffer' (Orozco et al. 1985). SDS was added to the suspension to a final concentration of 2% (v/v) and the mixture heated at 95° C for 5 min. Undissolved cell debris was removed by centrifugation.

Western blotting. Proteins electrophoresed in a 15% polyacrylamide gel in the presence of SDS were electro-blotted on to nitrocellulose and treated with antiserum and [^{125}I]protein A according to the method of Yen and Webster (1981).

Results

Sequence of the rpo*A gene*

Nucleotide sequence analysis of a 2.1 kb *Sal*I – *Kpn*I fragment of pea chloroplast DNA containing the 3′ end of the gene (*pet*D) for subunit IV of the cytochrome *b–f* complex has revealed the presence of two open reading frames (Figs. 1, 2). The ORFs are transcribed on the opposite strand to *pet*D and encode polypeptides of 138 and 334 amino acid residues, respectively. The smaller ORF and one of 37 codons, the 3′ half of which is present on the *Sal*I – *Kpn*I fragment, have been identified as the genes for ribosomal proteins S11 and L36, respectively (Purton and Gray 1987a, b). The larger ORF encodes a product of 38916 dalton (including the initiating formyl-methionine) having significant homology to the alpha subunit of the *E. coli* RNA polymerase; the two protein sequences have a 29% identity over a 298 residue overlap. The ORF corresponds to that found immediately downstream of *pet*D in the chloroplast genomes of spinach (Sijben-Müller et al. 1986), liverwort (Ohyama et al. 1986), tobacco (Shinozaki et al. 1986), maize (Ruf and Kössel 1988) and wheat (Hird, Dyer and Gray, unpublished data) and has been called *rpo*A. A comparison of the deduced amino acid sequences from the chloroplast *rpo*A genes is presented in Fig. 3 and shows a high degree of conservation between the species; the percentage of identical residues ranges from 79% for tobacco and spinach to 54% for liverwort. Several deletions/insertions have occurred within the pea sequence relative to the other species although the reading frame is maintained, suggesting that *rpo*A is a functional gene. There are also several regions within the sequence where homology is weak, most notably at the carboxy-terminus of the protein.

Transcript analysis

To determine whether *rpo*A and the two upstream ribosomal protein genes are expressed in vivo, Northern blot analysis of pea chloroplast RNA was carried out using single-stranded radiolabelled DNA probes for each of the three genes (Fig. 4). A major transcript of 1.8 kb was detected by the *rpo*A and *rps*11 probes, together with other minor transcripts. Analysis of the sequence data presented in Fig. 2 reveals that the major inverted repeat sequences located upstream of *rps*11 and downstream of *rpo*A are

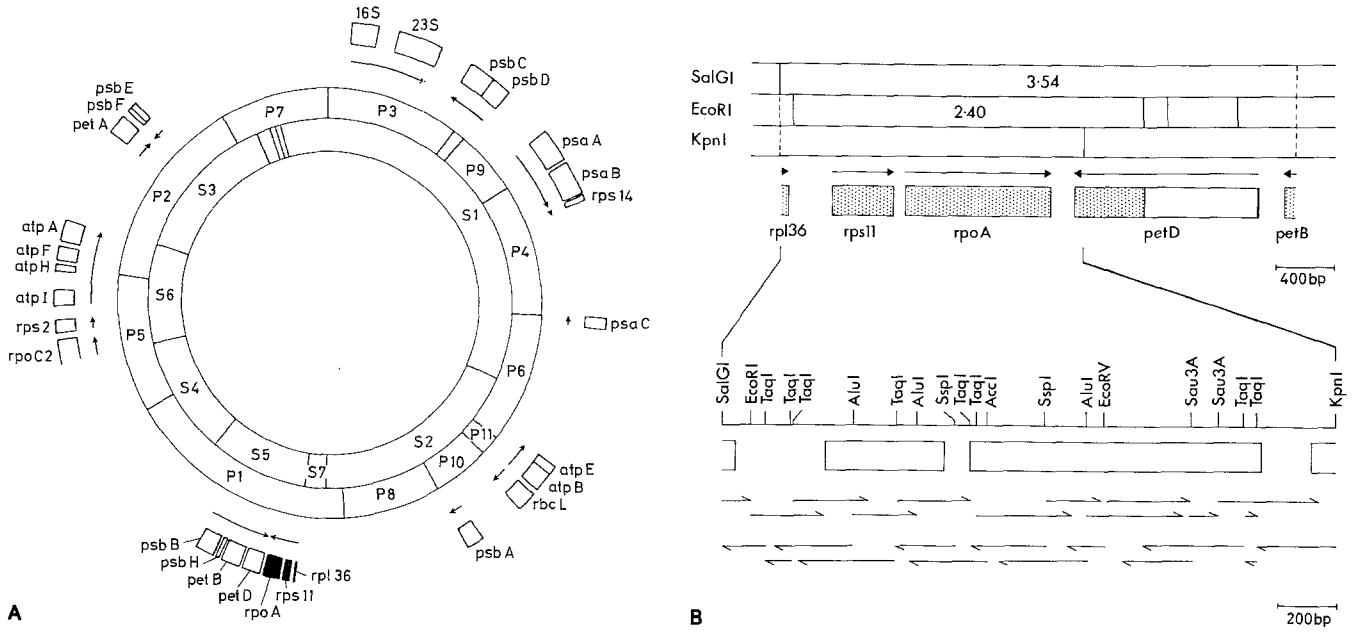

Fig. 1A and B. Restriction map of pea chloroplast DNA showing the location of the *rpo*A and *rps*11 genes. **A** Restriction sites for *Pst*I (P) and *Sal*I (S), as determined by Palmer and Thompson (1981), are shown and the restriction fragments are numbered in decreasing order of size. **B** Detailed restriction map and sequencing strategy of the region of pea chloroplast DNA containing *rpo*A and *rps*11. *Full arrows* indicate the direction of transcription of the genes; *half-barbed arrows* indicate the direction and extent of sequencing

1.8 kb apart. Furthermore, no identifiable promoter sequences are present between *rps*11 and *rpo*A. It is therefore concluded that the two genes are co-transcribed and that the two inverted repeats delimit the transcript. The largest transcript common to both *rps*11 and *rpo*A (approximately 2.4 kb) is not detected by the *rpl*36 probe. It would appear that this transcript covers a region of the DNA extending well into the *pet*D coding region. This observation supports the suggestion of Stern and Gruissem (1987) that chloroplast transcription continues beyond the inverted repeat sequences found downstream of a number of genes. This transcript is then presumed to be processed by nuclease activity. Smaller transcripts common to both *rps*11 and *rpo*A are too small to be full-length transcripts for both genes and may represent processing within the *rpo*A region. The *rpl*36 probe detected a single transcript of 0.8 kb. Whether this represents a primary transcript or the product of rapid mRNA processing of a larger transcript covering other genes is not known.

Antibodies to the rpo*A gene product*

To investigate the presence of the *rpo*A gene product within chloroplasts, antibodies were raised to a fusion protein synthesised in *E. coli* from an expression plasmid. The plasmid, pEXT920, was constructed by the insertion of 307 codons of the *rpo*A gene into the multiple cloning site of the expression vector pEX3 to generate an in-frame *lac*Z—*rpo*A gene-fusion. Fusion protein synthesised in the *E. coli* host cells was recovered by electroelution from SDS-polyacrylamide gels and used in the production of antibodies. The presence of antibodies in the antiserum against the *rpo*A moiety of the fusion protein was demonstrated by their specific immunoprecipitation of the *rpo*A gene product synthesised in a cell-free coupled transcription-translation system from *E. coli* (Howe et al. 1982) programmed with pPscEK2 (Fig. 5A).

Expression of the rpo*A gene in organello*

The light-driven synthesis of the *rpo*A gene product in isolated chloroplasts was demonstrated by immunoprecipitation from a crude protein extract prepared from pea chloroplasts after incubation in the presence of L-[^{35}S]methionine. Immunochemically related proteins were precipitated using the fusion protein antiserum and protein A-Sepharose and analysed by SDS-PAGE together with the total extracted proteins and those immunoprecipitated in the presence of a pre-immune serum (Fig. 5B). The fluorograph shows that the vast majority of the radiolabel is incorporated into the large subunit of ribulose bisphosphate carboxylase and this results in its contamination of the pre-immune and immune tracks. However, a labelled polypeptide of approximately 41 kDa is specifically immunoprecipitated using the anti-fusion protein antibodies. It is concluded that this is the *rpo*A gene product and that the *rpo*A gene is actively expressed in isolated chloroplasts.

Identification of the rpo*A gene product in a transcriptionally active chloroplast extract*

To investigate the relationship between the *rpo*A gene product and the chloroplast RNA polymerase, a transcriptionally active extract was prepared from pea chloroplasts as described by Orozco et al. (1985). Western blot analysis of this extract was carried out using the antibodies against the fusion protein and also antibodies against the RNA polymerase core complex of the cyanobacterium, *Anabaena* 7120 (Schneider et al. 1987). As a positive control for these antibodies, a crude cell extract was prepared from the closely related cyanobacterium, *Anacystis nidulans*, by sonication

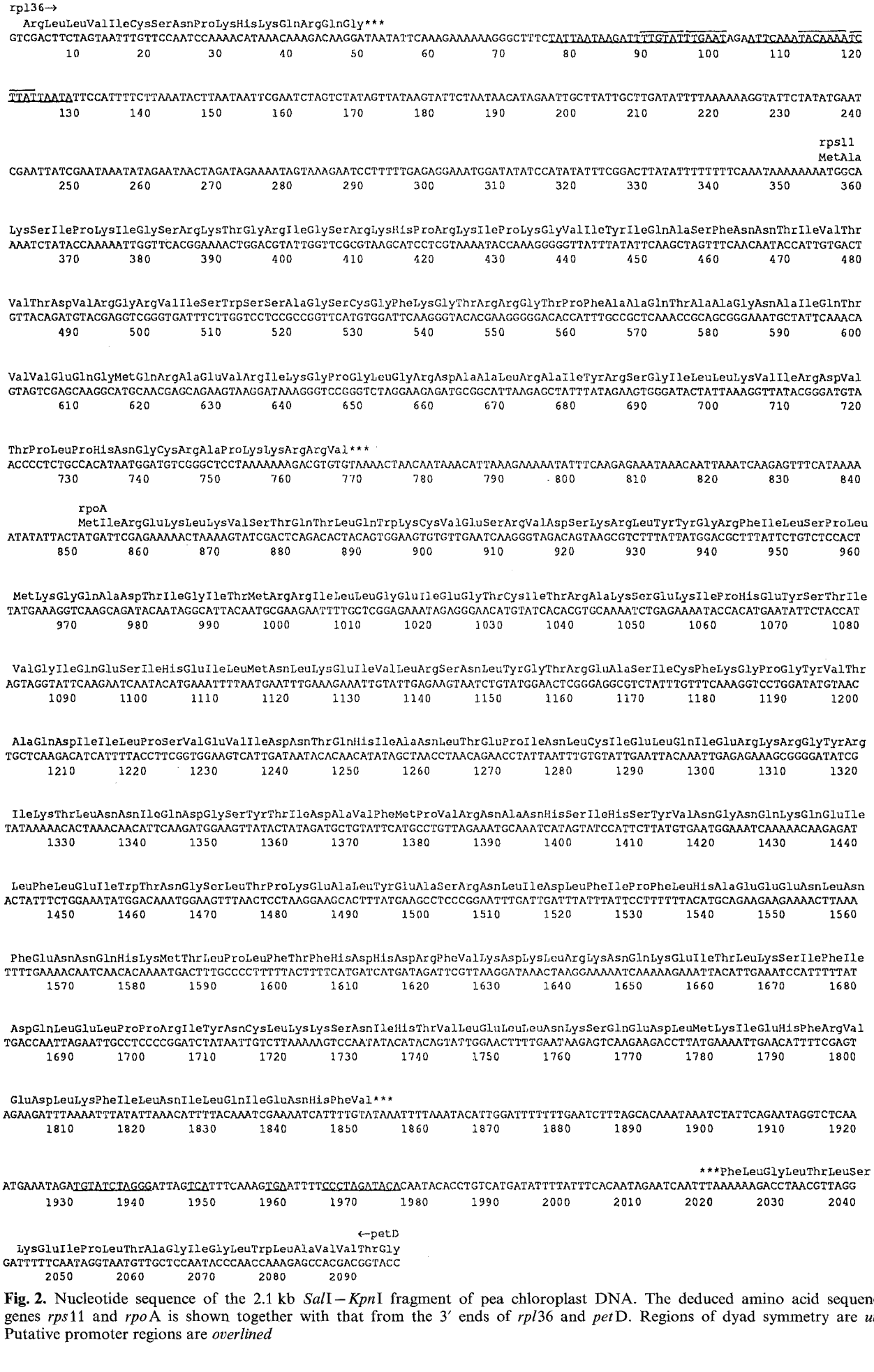

Fig. 2. Nucleotide sequence of the 2.1 kb *Sal*I – *Kpn*I fragment of pea chloroplast DNA. The deduced amino acid sequence of the genes *rps*11 and *rpo*A is shown together with that from the 3′ ends of *rpl*36 and *pet*D. Regions of dyad symmetry are *underlined.* Putative promoter regions are *overlined*

```
                .         .         .         .         .         .         .         .         .         .
pea         MIREKLKVSTQTLQWKCVESRVDSKRLYYGRFILSPLMKGQADTIGITMRRILLGEIEGTCITRAKSEKIPHEYSTIVGIQESIHEILMNLKEIVLRSNL
wheat       :V::VVAG:::::::::::::::::::::::::::::::R::::::V::AL::A::::::::::::::FGSV:::::::A::E::VQ:::L::::::::::
spinach     :V:::IR:::::::::::::::T:::C:H:::::::::::::::::::A:::A:::::::::::::::::::::::::::L:::::V:::::::::::::::
tobacco     :V:::VT:::R:::::::::::T:::::::::::::::::::::::::A:::A:::::::::::V::::V:::::::T:::::V:::::::::::I:::::
maize       .V..EITG:::::E:::::::::::::::::::::::::::::::R::::::V::AL::A::::::::::::::FGNV:::::::::::E:::Q:::L::::::::::
liverwort   ::QDEI::::::::::::I::KIE::::L:S::AI::FR::::N:V::A:::A::N::::AS::Y::IK:VK::::::I:L:::::D::I:::::::K:ES
                .         .         .         .         .         .         .         .         .         .
            YGTREASICFKGPGYVTAQDIILP-SVEVIDNTQHIANLTEPINLCIELQIERKRGYRIKTLNNIQDGSYTIDAVFMPVRNANHSIHSYVNGNQKQEILF
            ::V:D::::V:::R:I::::::::P:::IV:TA:P::::::::DF::D:::K:D:::QTELRK:Y:::::P::::S:::::::Y::F:CG:::E:L::::
            ::AC:::::VR::RG:::::::::PY::IV:::::::S:::::D:::G:L::N:::H::AP::F::::FP:::L:::::::V:::::::G:::E:::::::
            :::SD::::V::::S:::::::::PY::IV:::::::S:::::DF::G:::::N:::L:::PH:F:::::P:::::::::::::::::::G:::E:::::::
            ::V:D::::V:::R:I::C:::::P:::ID:T::D::::R::VDF:::::::K:D:A:HTELRK:S:::::P::::::::::V:Y::F:CG:::E:H::::
            FEPQK:Y:SVL::KKI:::::KG:SCIKIMIIA:Y::T:NKD:L:E:::N::KD:::::EN:QKY:E:LFPV::::::I::::Y:V::FESEKKIK::::
                .         .         .         .         .         .         .         .         .         .
            LEIWTNGSLTPKE-ALYEASRNLIDLFIPFLHAEEENLNFENNQHKMTLPL-FTFHDHDRFVKDKLRKNQKEITLKSIFIDQLELPPRIYNCLKKSNIHT
            ::YGS::::::::-:::::::::::::L:::::::::GAS::E:KNRF:P::-:::QK--:L--TN:K::K:G:P:NC:::::::::TS:T:::::RA::::
            :::::::::::::-:::::::::::L:::::::::-:V:L:D::::VS:::-::::N--::L--AEI:::K:K:A::F:::::::::::::::::::::::
            ::::::::::::::-::H:::::::::::::::M::D::YLQD:::TVP:SP-:::::--KL--A::I::K:K:A::::::::S:::S:::::::M:::Y:
            :::::::K:::::GGI:::::::::::L::::T:::GTS::E:KNRL:P::-L::QK--::--TN:K::K:G:P:NC:::::::::S:T:::::RA::PS
            :::::D::::::::-:::::::::::::::::LINS:KKEK::GIEKTNESNMSY:P:QS----:SLDIE:MT:DVAF:H:::::::::A:A::::::V:V::
                .         .         .         .
            -VLELLNKSQEDLMKIEHFRVEDLKFILNILQIENHFV              334(1)  100%(2)
            -L:D::S:TE:::LR:DS::M::R:H:WDT:--:K:LPIDLLKNKLCI    339     69%
            -L:D:::N:::::I:MK:::I::V:Q:FGT:--:K:::IDL-KNKR      335     79%
            -L:D:::N:::::::::::::S::V:R::G::--:KY::IDLAKNKF     337     79%
            AL:D::V:TE:::I::NS:GM::G:L:WDT:--:K:LPIDLPKNKFSL    341     69%
            -IAD::HY:ED::I::KN:GKKSVEQV:EA:--KKR:SIQLPKNKNYL    340     54%
```

Fig. 3. Comparison of the deduced amino acid sequence of *rpo*A from pea with that from wheat, spinach, tobacco, maize and liverwort. A *colon* indicates a matched residue. A *dash* indicates a gap introduced into the sequence to maximise alignment. (1), Number of amino acids; (2), percentage of identical residues with respect to pea

followed by SDS-extraction. Both protein extracts were separated on a polyacrylamide gel in the presence of SDS, transferred to nitrocellulose and probed with each of the antisera. The results are shown in Fig. 6 and show that both antisera recognise a 43 kDa protein in the chloroplast preparation. It is concluded that this is the *rpo*A gene product. In the control track (track C) four proteins are detected. The three largest bands are the β, γ and α subunits, respectively, of the cyanobacterial RNA polymerase. The smallest band probably represents a breakdown product. A second protein of approximately 130 kDa is detected in the chloroplast extract using the anti-*Anabaena* RNA polymerase antibodies (track D). This is possibly the product of the chloroplast *rpo*B gene. This gene has recently been sequenced in spinach chloroplast DNA (Hudson et al. 1988) and shown to encode a 128 kDa product which is 50% homologous to the β-subunit of *E. coli* RNA polymerase. The protein detected in the cyanobacterial extract by the anti-fusion protein antibodies does not correlate with any of the known polymerase subunits and possibly represents a protein immunochemically related to β-galactosidase.

Several attempts have been made to inhibit the transcriptional activity of the chloroplast extract using the anti-fusion protein antibodies but without success (data not shown). This may be due to a lack of antibody binding under the conditions used or, possibly, a lack of inhibition of activity by bound antibodies.

Discussion

The gene, *rpo*A, encoding a polypeptide homologous to the α-subunit of *E. coli* RNA polymerase, has been located and sequenced in pea chloroplast DNA. Comparison of the deduced amino acid sequence of *rpo*A from pea, spinach, tobacco, liverwort and wheat reveals a high level of homology although variations between species are seen at the carboxy-terminus and in the middle of the protein.

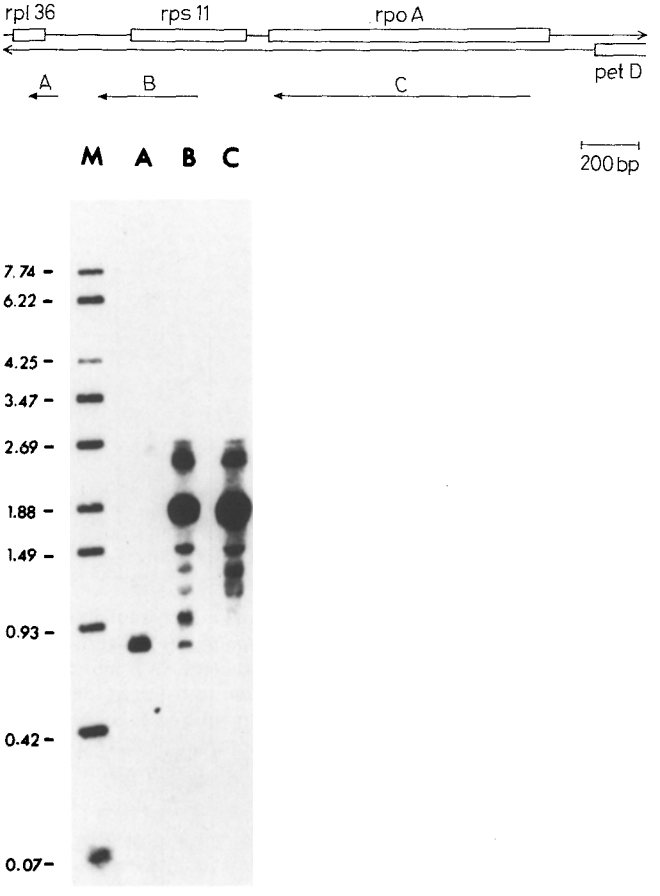

Fig. 4. Northern blot analysis of transcripts of the *rpo*A, *rps*11 and *rpl*36 genes. Glyoxylated chloroplast RNA fractionated by gel electrophoresis and immobilised on nylon membranes was hybridised with single-stranded probes shown in the *upper* part of the figure. A, 101 bp *Sal*I–*Eco*RI fragment; B, 357 bp *Taq*I fragment; D, 920 bp *Taq*I fragment. Lane M shows glyoxylated *Sty*I fragments of λ DNA. Sizes shown are in kilobases

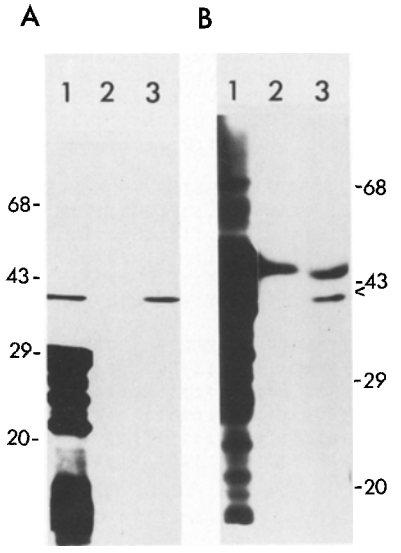

Fig. 5. A and B. Immunoprecipitation of [^{35}S]methionine-labelled proteins synthesised in the *Escherichia coli* transcription-translation system and in isolated pea chloroplasts. **A** Analysis of the products of pPscEK2 in the *E. coli* cell-free system. Lane 1, total products; lane 2, immunoprecipitation with pre-immune serum; lane 3, immunoprecipitation with antibodies to the fusion-protein. **B** Analysis of the products of protein synthesis by isolated pea chloroplasts. Lane 1, total products; lane 2, immunoprecipitation with pre-immune serum; lane 3, immunoprecipitation with antibodies to the fusion-protein. The location of the *rpo*A gene product is indicated by an *arrow*. Size markers are bovine serum albumin (68 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa) and trypsin inhibitor (20 kDa)

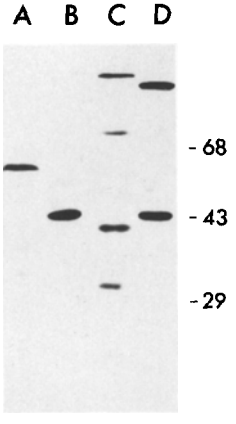

Fig. 6. Western blot analysis of chloroplast and cyanobacterial extracts. A crude extract of *Anacystis nidulans* (tracks A and C) and a high-salt extract of pea chloroplasts (tracks B and D) were fractionated by SDS-gel electrophoresis and transferred to nitrocellulose. The blots were then incubated with antibodies to the fusion-protein (tracks A and B) or to the core complex of RNA polymerase from *Anabaena* 7120 (tracks C and D). Size markers were as for Fig. 5

As is the case for all chloroplast genomes so far studied, with the possible exception of *Euglena* (Cushman et al. 1988) and geranium (J.D. Palmer, personal communication), the pea *rpo*A gene is located immediately downstream of the *pet*D gene for subunit IV of the cytochrome complex and is encoded on the opposite strand. The complete sequence of the tobacco and liverwort genomes has revealed that *rpo*A is the most distal member of a cluster of twelve 'house-keeping' genes encoding components of the plastid transcription and translation apparatus. The gene order matches that found in the sequential *E. coli* operons S10, *spc* and α although a number of genes are absent from the chloroplast gene cluster. The gene organisation in pea appears to be essentially the same as that in tobacco and liverwort. The ribosomal protein genes *rps*11 and *rpl*36 are located immediately upstream of *rpo*A. Upstream of *rpl*36, other members of the gene cluster have been identified by heterologous hybridisation (Palmer et al. 1988) and, in the case of *rpl*16, by nucleotide sequencing (Purton and Gray, unpublished results).

The pea genes *rpl*36, *rps*11 and *rpo*A are actively transcribed in vivo. The *rps*11 and *rpo*A genes appear to be co-transcribed as is the case in the *E. coli* α-operon (*rps*13 – *rps*11 – *rps*4 – *rpo*A – *rpl*17) (Bedwell et al. 1985). Interestingly, the comparable region of the *Bacillus subtilis* genome appears to represent an intermediate between that of *E. coli* and chloroplasts (Suh et al. 1986). Here the *rps*4 gene is absent, thereby linking *rps*11 and *rpo*A, whereas *rps*13, which is not present in the plastid genome, is still present in *B. subtilis*. Whether the pea *rpl*36 gene is also co-transcribed is not known. A single transcript of approximately 0.8 kb is detected using the *rpl*36 probe. Presumably this transcript corresponds to one or other of the pea transcripts (0.65 and 0.85 kb) detected by Woodbury et al. (1988) using DNA probes from either side of the unique *Sal*1 site within *rpl*36. A major inverted repeat sequence is present between *rpl*36 and *rps*11 (Fig. 2). Whether this functions as a transcript processing site or transcriptional terminator is not known, although it should be noted that two overlapping promoter-like sequence are also present within the inverted repeat sequence (Fig. 2). S1 nuclease mapping of this region has not resolved the uncertainty (Purton and Gray, unpublished results).

The transcript pattern observed for pea *rpo*A differs from that recently presented for maize (Ruf and Kössel 1988). The largest transcript observed in pea is 2.4 kb whereas in maize, transcripts as large as 10 kb were detected. It is possible that this is a consequence of faster transcript processing in pea. Alternatively, the transcriptional organisation of this region may differ in legumes relative to other plants as suggested by Palmer et al. (1988). The 2.4 kb transcript is detected by both the *rps*11 and *rpo*A probes but not by the *rpl*36 probe. This suggests that the transcript extends approximately 0.5 kb beyond the inverted repeat sequence downstream of *rpo*A, into the convergently transcribed *psb*B – *psb*H – *pet*B – *pet*D operon. This supports the findings of Stern and Gruissem (1987) that the *rpo*A – *pet*D inverted repeat functions as a processing site for longer transcripts. Interestingly, transcription from the opposite DNA strand of the *psb*B – *psb*H – *pet*B – *pet*D operon has been reported previously (Kohchi et al. 1988).

Antibodies raised against the fusion protein resulting from the expression in *E. coli* of a *lac*Z – *rpo*A gene-fusion have been employed to demonstrate the light-driven synthesis of the *rpo*A gene product in isolated pea chloroplasts and its presence in a transcriptionally active extract prepared from intact chloroplasts. The discrepancy between the size of the protein as predicted from the gene sequence (39 kDa) and that determined by gel electrophoresis (43 kDa) is in keeping with that observed for the *E. coli*

RNA polymerase alpha subunit. Here a 36.5 kDa protein (Ovchinnikov et al. 1977) migrates on gels at 39 kDa (Burgess 1976). Furthermore, the size of the pea protein matches that determined recently for the maize *rpo*A gene product by Western blotting using antibodies raised against a synthetic peptide (Ruf and Kössel 1988).

The assignment of the *rpo*A gene product as a component of the chloroplast RNA polymerase is supported by the protein's cross-reaction with antibodies raised against the purified RNA polymerase core complex (subunit structure: $\alpha\beta\beta'\gamma$) from the photosynthetic prokaryote, *Anabaena* 7120. These antibodies also detect a second polypeptide in the chloroplast extract of approximately 130 kDa. It is speculated that this is the product of the chloroplast *rpo*B gene, which has recently been located in the pea plastid genome (Palmer et al. 1988). The spinach *rpo*B gene product has a predicted size of 128 kDa and a 50% similarity with the *E. coli* beta subunit (Hudson et al. 1988). Polypeptides corresponding to the cyanobacterial γ and β' subunits are not detected in the chloroplast extract despite our observation that their sizes correlate reasonably well with the predicted sizes of the spinach *rpo*C_1 and *rpo*C_2 gene products, respectively (Hudson et al. 1988).

The results presented here demonstrate that the *rpo*A gene is transcribed and translated in chloroplasts. Furthermore, a 43 kDa polypeptide was detected in the chloroplast RNA polymerase preparation using antibodies against the *rpo*A gene product and against the cyanobacterial polymerase. This gives support to the idea that the *rpo* genes of plastid DNA encode the core components of a chloroplast RNA polymerase with a subunit structure analogous to that of eubacteria.

Acknowledgements. We are most grateful to J.P. Luzio (University of Cambridge) for the gift of the pEX vectors and the *E. coli* host pop2136; R. Haselkorn (University of Chicago) for the gift of antibodies against the RNA polymerase core enzyme of *Anabaena* 7120; and C.J. Howe (University of Cambridge) for providing the cyanobacterium *Anacystic nidulans*. S.P. was supported by an S.E.R.C. studentship.

References

Bedwell D, Davis G, Gosink M, Post L, Nomura M, Kestler H, Zengel JM, Lindahl L (1985) Nucleotide sequence of the alpha ribosomal protein operon of *Escherichia coli*. Nucleic Acids Res 13:3891–3903

Biggin M, Farrell PJ, Barrell BG (1984) Transcription and DNA sequence of the *Bam*HI L fragment of B95-8 Epstein-Barr virus. EMBO J 3:1083–1090

Biggin MD, Gibson TJ, Hong GF (1983) Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc Natl Acad Sci USA 80:3963–3965

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Blair GE, Ellis RJ (1973) Protein synthesis in chloroplasts. I. Light-driven synthesis of the large subunit of Fraction I protein by isolated pea chloroplasts. Biochim Biophys Acta 319:223–234

Bonner WM, Laskey RA (1974) A film detection method for tritium-labelled proteins and nucleic acids in polyacrylamide gels. Eur J Biochem 46:83–88

Briat J-F, Bisanz-Seyer C, Laulhère J-P, Lerbs S, Lescure A-M, Mache R (1987) The RNA polymerase from chloroplasts and its use for *in vitro* transcription of plastid genes. Plant Physiol Biochem 25:273–281

Bülow S, Link G (1988) Sigma-like activity from mustard (*Sinapis alba* L.) chloroplasts conferring DNA-binding and transcription specificity to *E. coli* RNA polymerase. Plant Mol Biol 10:349–357

Bünger W, Feierabend J (1980) Capacity for RNA synthesis in 70 *S* ribosome-deficient plastids of heat-bleached rye leaves. Planta 149:163–169

Burgess RR (1976) Purification and physical properties of *E. coli* RNA polymerase. In: Losick R, Chamberlin M (eds) RNA polymerase. Cold Spring Harbor Laboratories, Cold Spring Harbor, New York, pp 69–100

Cohen SN, Chang ACY, Hsu L (1972) Non-chromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc Natl Acad Sci USA 69:2110–2114

Covey SN, Hull R (1981) Transcription of cauliflower mosaic virus DNA. Detection of transcripts, properties and location of the gene encoding the virus inclusion body protein. Virology 111:463–474

Cushman JC, Hallick RB, Price CA (1988) The two genes for the P_{700} chlorophyll *a* apoprotein on the *Euglena gracilis* chloroplast genome contain multiple introns. Curr Genet 13:159–171

Ellis RJ, Hartley MR (1971) Sites of synthesis of chloroplast proteins. Nature New Biol 233:193–196

Greenberg BM, Narita JO, De Luca-Flaherty C, Gruissem W, Rushlow KA, Hallick RB (1984) Evidence for two RNA polymerase activities in *Euglena gracilis* chloroplasts. J Biol Chem 259:14880–14887

Greenberg BM, Narita JO, De Luca-Flaherty C, Hallick RB (1985) Properties of chloroplast RNA polymerases. In: Steinback KE, Bonitz S, Arntzen CJ, Bogorad L (eds) Molecular biology of the photosynthetic apparatus. Cold Harbor Spring Laboratory, Cold Spring Harbor, New York, pp 303–310

Howe CJ, Bowman CM, Dyer TA, Gray JC (1982) Localisation of wheat chloroplast genes for the beta and epsilon subunits of ATP synthase. Mol Gen Genet 186:525–530

Hudson GS, Holton TA, Whitfeld PR, Bottomley W (1988) Spinach chloroplast *rpo*BC genes encode three subunits of the chloroplast RNA polymerase. J Mol Biol 200:639–654

Kieny MP, Lathe R, Lecocq JR (1983) New versatile cloning and sequencing vectors based on bacteriophage M13. Gene 26:91–99

Kohchi T, Yoshida T, Komano T, Ohyama K (1988) Divergent mRNA transcription in the chloroplast *psb*B operon. EMBO J 7:885–891

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. Nature 227:680–685

Lam E, Chua N-H (1987) Chloroplast DNA gyrase and *in vitro* regulation of transcription by template topology and novobiocin. Plant Mol Biol 8:415–424

Lerbs S, Bräutigam E, Mache R (1988) DNA-dependent RNA polymerase of spinach chloroplasts: characterization of α-like and σ-polypeptides. Mol Gen Genet 211:459–464

Lerbs S, Bräutigam E, Parthier B (1985) Polypeptides of DNA-dependent RNA polymerase of spinach chloroplasts: characterization by antibody-linked polymerase assay and determination of sites of synthesis. EMBO J 4:1661–1666

McMaster GK, Carmichael GG (1977) Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc Natl Acad Sci USA 74:4835–4838

Newman BJ, Gray JC (1988) Characterisation of a full-length cDNA clone for pea ferredoxin-$NADP^+$ reductase. Plant Mol Biol 10:511–520

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572–574

Orozco EM, Mullet JE, Chua N-H (1985) An *in vitro* system for

accurate transcription initiation of chloroplast protein genes. Nucleic Acids Res 13:1283–1302

Ovchinnikov YA, Lipkin VN, Modyanov NN, Chertov OY, Smirnov YV (1977) Primary structure of α-subunit of DNA-dependent RNA polymerase from *Escherichia coli*. FEBS Lett 76:108–111

Palmer JD, Thompson WF (1981) Rearrangements in the chloroplast genomes of mung bean and pea. Proc Natl Acad Sci USA 78:5533–5537

Palmer JD, Osorio B, Thompson WF (1988) Evolutionary significance of inversions in legume chloroplast DNAs. Curr Genet 14:65–74

Phillips AL, Gray JC (1984) Location and nucleotide sequence of the gene for the 15.2 kDa polypeptide of the cytochrome *b–f* complex from pea chloroplasts. Mol Gen Genet 194:477–484

Purton S, Gray JC (1987a) Nucleotide sequence of the gene for ribosomal protein S11 in pea chloroplast DNA. Nucleic Acids Res 15:1873

Purton S, Gray JC (1987b) Nucleotide sequence of the gene for ribosomal protein L36 in pea chloroplast DNA. Nucleic Acids Res 15:9080

Ruf M, Kössel H (1988) Structure and expression of the gene coding for the α-subunit of DNA-dependent RNA polymerase from the chloroplast genome of *Zea mays*. Nucleic Acids Res 16:5741–5754

Sanger F, Coulsen AR, Barrell BG, Smith AJH, Roe BA (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J Mol Biol 143:161–178

Schneider GJ, Tumer NE, Richaud C, Borbely G, Haselkorn R (1987) Purification and characterization of RNA polymerase from the cyanobacterium *Anabaena* 7120. J Biol Chem 262:14633–14639

Shinozaki J, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamagashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete sequence of the tobacco chloroplast genome: its organization and expression. EMBO J 5:2043–2049

Siemenroth A, Wollgiehn R, Neumann D, Börner Th (1981) Synthesis of ribosomal RNA in ribosome-deficient plastids of the mutant "albostrians" of *Hordeum vulgare* L. Planta 153:547–555

Sijben-Müller G, Hallick RB, Alt J, Westhoff P, Herrmann RG (1986) Spinach plastid genes coding for initiation factor IF-1, ribosomal protein S11 and RNA polymerase α-subunit. Nucleic Acids Res 2:1029–1044

Stanley KK, Luzio JP (1984) Construction of a new family of high efficiency bacterial expression vectors: identification of cDNA clones coding for human liver proteins. EMBO J 3:1429–1434

Stern DB, Gruissem W (1987) Control of plastid gene expression: 3′ inverted repeats act as mRNA processing and stabilizing elements but do not terminate transcription. Cell 51:1145–1157

Suh J-W, Boylan SA, Price CW (1986) Gene for the alpha subunit of *Bacillus subtilis* RNA polymerase maps in the ribosomal gene cluster. J Bacteriol 168:65–71

Surzycki SJ, Shellenbarger DL (1976) Purification and characterization of a putative sigma factor from *Chlamydomonas reinhardii*. Proc Natl Acad Sci USA 73:3961–3965

Woodbury NW, Roberts LL, Palmer JD, Thompson WF (1988) A transcription map of the pea chloroplast genome. Curr Genet 14:75–89

Yen TSB, Webster RE (1981) Bacteriophage f1 gene II and X proteins. J Biol Chem 256:11259–11265

Communicated by D.M. Lonsdale

Received October 4, 1988