Structure and expression of the gene coding for the α -subunit of DNA-dependent RNA polymerase from the chloroplast genome of *Zea mays*

Michael Ruf and Hans Kössel

Institut für Biologie III der Universität Freiburg, Schänzlestrasse 1, D-7800 Freiburg, FRG

Received May 4, 1988; Accepted May 27, 1988

Accession no. X07810

ABSTRACT

The rpoA gene coding for the α -subunit of DNA-dependent RNA polymerase located on the DNA of *Zea mays* chloroplasts has been characterized with respect to its position on the chloroplast genome and its nucleotide sequence. The amino acid sequence derived for a 39 Kd polypeptide shows strong homology with sequences derived from the rpoA genes of other chloroplast species and with the amino acid sequence of the α -subunit from *E. coli* RNA polymerase. Transcripts of the rpoA gene were identified by Northern hybridization and characterized by S₁ mapping using total RNA isolated from maize chloroplasts. Antibodies raised against a synthetic C-terminal heptapeptide show cross reactivity with a 39 Kd polypeptide contained in the stroma fraction of maize chloroplasts. It is concluded that the rpoA gene is a functional gene and that therefore, at least the α -subunit of plastidic RNA polymerase, is expressed in chloroplasts.

INTRODUCTION

Chloroplasts are semi-autonomous organelles which possess their own DNA coding for organelle-specific mRNAs, tRNAs and rRNAs (1). The transcription of chloroplast genes is mediated by a chloroplast specific DNA-dependent RNA polymerase whose origin and diversity is a matter of dispute (2). The observations that ribosome deficient chloroplasts show continued RNA synthesis (3) and that mRNAs encoding plastidic RNA polymerase subunits are found within poly(A)-containing mRNA fractions (4) have led to the conclusion that plastidic RNA polymerase is encoded in the nucleus. Contrary to this, sequencing of chloroplast DNA from tobacco (5), liverwort (6) and spinach (7) has revealed that genes coding for polypeptides which show high degrees of sequence homology with *E. coli* RNA polymerase subunits α , β , and β' are present in the plastome of higher plants, which

supports the view of plastome encoded RNA polymerase subunits. Concerning the diversity of plastidic RNA polymerases, different enzymes have been postulated for the transcription of mRNA, tRNA and rRNA coding genes (8, 9) and for different developmental stages of plastids (10). It is, however, not clear whether the observed differences reflect different core enzymes or identical core enzymes merely containing specific σ -like or other cofactors. The latter possibility is supported by the observation of consensus promoter sequences occurring in the leader regions of many plastidic genes (11), including the rRNA genes (12).

In order to contribute to the resolution of these controversies we have started to investigate the structure and expression of maize chloroplast genes encoding subunits of plastidic RNA polymerase. In this paper we would like to report on the structural analysis of the rpoA gene and to provide evidence for its in vivo transcription. We conclude that the rpoA gene from maize chloroplasts is a functional gene and that therefore, at least the α -subunit of plastidic RNA polymerase, is expressed in chloroplasts.

MATERIALS AND METHODS

Localization and Sequence Analysis of the rpoA Gene.

Recombinant E. coli clones harboring PstI fragments from Zea mays chloroplast DNA were obtained by cloning in the vector pBR329 (13). The E. coli clone harboring the plasmid pWHsp20 (7) which contains the rpoA gene from spinach chloroplasts was kindly provided by Dr. R.G. Herrmann. Oligonucleotides of 35 and 36 nt, respectively, were selected according to conserved N-terminal and C-terminal regions of the rpoA genes from tobacco (5), liverwort (6) and spinach (7) and synthesized on a DNA synthesizer 380A from Applied Biosystems. ^{32}P -labelling of the rpoA specific probes was achieved either by random primed polymerization (14) of restriction fragments or by 5'-terminal phosphate transfer to synthetic oligonucleotides. [α - ^{32}P]dGTP for random priming and [γ - ^{32}P]ATP for 5'-terminal labelling were purchased from Amersham, U.K.

After cleavage of the maize DNA fragment PstIC, contained in

the plasmid pZmc303 (13), with restriction endonucleases BglII, EcoRI, HindIII and SalI (from Boehringer/Mannheim and Pharmacia/ Freiburg) fragments were subcloned in the M13 vectors mp8 and mp9 (15). Sequencing of the M13 clones was carried out with an Amersham sequencing kit as recommended by the supplier.

RNA Preparation, Northern Blot Analysis and S₁ Mapping

Total chloroplast RNA was isolated as described (12). It was usually not treated with DNase, because we observed no influence of the very low chloroplast DNA content in our RNA preparation. RNA was fractionated on formaldehyde containing gels (16) and, after transfer to nitrocellulose (Schleicher and Schüll, Dassel), hybridized with ³²P-labelled restriction fragments (17). S₁ mapping was carried out as described earlier (12). RNA markers and mAP paper for specific binding of poly (A)-containing RNA were purchased from Amersham, U.K.

Preparation of Peptide Antisera and Immune blotting

Synthesis of the C-terminal heptapeptide ProLysAsnLysPheSerLeu as derived from the DNA sequence (see Fig.2) was kindly performed by Dr. Brüller, University of Marburg, B.R.D. It was conjugated to carrier protein as follows: 16 mg of ovalbumin or 25 mg of BSA (both purchased from Sigma) and 12 mg of the heptapeptide were dissolved in 2 ml 0.1 M sodium phosphate buffer (pH 7.4). 1 ml of 20 mM aqueous glutaraldehyde solution was then added slowly in an ice bath. After 30 min incubation in the cold, incubation was continued for additional 30 min at room temperature.

Antisera were obtained by intramuscular injection into rabbits of ovalbumin-heptapeptide conjugate (2.5 mg/ml): complete Freud's adjuvant suspensions (1:1); boosting was carried out after two weeks (18). The specificity of the sera against the heptapeptide was tested by immune-blotting with BSA-heptapeptide conjugate prepared in the same way as the ovalbumin-heptapeptide conjugate (19).

Western blotting of proteins from maize chloroplast S-30 extracts and immunodetection were carried out according to published standard procedures (17, 20, 21). Antisera were added at dilutions of 1:500; the phosphatase conjugated second anti (rabbit antibody) (obtained from Sigma) at a dilution of

1:1000. The phosphatase assay was carried out according to the recommendations of the supplier.

Isolation of a chloroplast S-30 extract from 10 days old maize seedlings was carried out according to a published procedure (22).

RESULTS AND DISCUSSION

Structural Analysis of the *rpoA* Gene

The *E. coli* clone pZmc303 (13) was selected for the structural analysis of the maize chloroplast *rpoA* gene, as plasmid DNA from this clone showed cross hybridization with plasmid DNA from *E. coli* clone pWHsp20 which contains the *rpoA* gene from spinach chloroplasts (7). In order to narrow down the position of the maize *rpoA* gene on the 15.3 Kb PstIC fragment, two synthetic oligonucleotides were used as hybridization probes, the sequences of which were selected according to highly conserved regions of the *rpoA* genes from tobacco (5), liverwort (6) and spinach (7) chloroplasts. The oligonucleotide 5'TTTAC-TATCTGTTCTTGATTCAACACACTTCCACTG was complementary to a 5'-terminal region - as determined subsequently by sequencing to the positions 380 to 350 (see Fig. 2) - whereas the oligonucleotide 5'ATAAGATCTTCTTGACTGTTATTCAAAAGGTCCAA shows complementarity to the 3'-terminal region - as later evident from sequencing to positions 1237 to 1202 (see Fig. 2). Though the complementarity of these oligonucleotides with the maize *rpoA* sequence is not 100% (as their sequences were selected according to the heterologous *rpoA* genes) hybridization with the subfragments of the PstIC fragment provided significant signal strength to allow fine mapping of the *rpoA* gene as depicted in Fig. 1. Final proof of the *rpoA* gene as being positioned close to the right hand end of fragment PstIC, which contains the two *SalI* sites (see Fig. 1 top) came from sequencing. The sequence of the *rpoA* gene as determined in M13 clones (Fig. 1, bottom) is depicted in Fig. 2 together with the flanking regions and the C-terminal region of the preceding *rps11* gene and the position of restriction sites shown in the map of Fig. 1. The position and orientation of the *rpoA* gene on the physical map of the maize plastome (13, 23) is homologous to tobacco,

Fig. 1: Position of restriction fragment *Pst*IC on the maize chloroplast DNA (top) and physical map and sequencing strategy of the *rpoA* containing region (bottom). The *Pst*IC fragment marked by striation in the top part was cloned in the *Pst*I site of vector *pBR329* to yield the plasmid *pZmc303* as outlined in the middle part. From this plasmid the 2200 bp long *Ava*I/*Pst*I restriction fragment was used for cloning and sequencing the subfragments covering the *rpoA* gene and the C-terminal part of the *rps11* gene as indicated in the lower part. The short arrows below the restriction map of the *Ava*I/*Pst*I fragment mark the regions sequenced from individual M13 subclones. The long arrow below the gene symbols indicates the direction of the transcript

Fig. 2: Nucleotide sequence of the *rpoA* gene and its flanking region from maize chloroplasts. The derived amino acid sequences of the *rpoA* gene and of the C-terminal region of the *rps11* gene are also presented. Direct and inverted repeat sequences in the trailer region (see also Fig. 5b) are marked by pairs of arrows. A possible ribosomal binding site positioned about 17 nt upstream of the putative translational start of the *rpoA* gene is marked by framing.

liverwort and spinach (5, 6, 7). As judged from its neighborhood to the *rps11* gene and from the transcript size determined by Northern blotting (see below), it appears likely that the

Fig. 3: Comparison of amino acid sequences derived from the rpoA genes of E. coli (C), maize (M), spinach (S), tobacco (T) and liverwort (L). Identical amino acids in the E. coli and maize sequences are marked by framing. Exchanges of a conservative nature according to the criteria proposed by Feng et al. (24) are marked by dots. Positions of the spinach, tobacco and liverwort sequences identical with the maize sequence are left blank. Deletions are symbolized by ▲ (single amino acids) and △ (larger portions). Numbering refers to the left most amino acid of each row.

maize rpoA gene also represents the most distal cistron of a gene cluster located on the large single copy region close to its border with the inverted repeat region IRII, as is observed for the other plastome species.

The sequence homology with the rpoA genes from the other spe-

cies is 78.4% (maize/tobacco), 77.8% (maize/spinach) and 66.4% (maize/liverwort) at the nucleotide level, which is in agreement with a roughly equal and, as compared to the more distantly related liverwort, a closer phylogenetic distance to both the dicotyledon species (tobacco and spinach). In Fig. 3 the relationship of rpoA genes at the level of the derived amino acid sequences is shown. Between the chloroplast species homologies of 68.4% (maize/spinach), 68.5% (maize/tobacco) and 51.8% (maize/liverwort) are observed. In addition to this, many of the amino acid substitutions are of the neutral type as judged by the criteria proposed by Feng et al. (24). The reading frame is conserved for all of the four plastidic rpoA genes in spite of occasional deletions/insertions. We interpret this as being due to a strong selectional pressure for maintaining the plastidic rpoA genes as functional genes. Fig. 3 includes also the alignment with the amino acid sequence derived from the E. coli rpoA gene (25). This shows 26.8% homology with the maize sequence but several larger insertions/deletions between the two sequences are evident.

The N termini from maize, spinach and tobacco show conservation of the sequence MetValArgGlu, whereas a different sequence and a deletion is observed in liverwort and in the E. coli sequence, respectively. It should, however, be pointed out, that the N termini, as deduced from nucleotide sequences, may not represent the true N termini, since mRNA splicing may cause a shift of sequences positioned further upstream in the primary transcript, as has been observed in the psbB operon with sequences encoding short N-terminal peptides (26, 27).

Analysis of in vivo Transcripts and S₁ Mapping

Northern blot analysis of total RNA from maize chloroplasts with an rpoA gene specific probe leads to the patterns shown in Fig. 4. A series of transcripts of 0.9 to 10 Kb is observed. The most likely explanation of this series of bands is that the largest species represent the primary transcript of the gene cluster starting with the trnI or rpl23 gene as is suggested from the respective clusters in the plastomes from tobacco (5) and liverwort (6), and that the ladder of smaller RNA species represents processing intermediates similar to the products

Fig. 4: Northern hybridization of total maize chloroplast RNA using the 900 bp restriction fragment BamHI/BglIII, which encodes most of the rpoA gene, as a probe (see Figs. 1 and 2). In both lanes the same amount (17 μ g) of RNA was used for gel electrophoresis in 1% agarose in the presence of 20 % formaldehyde. The RNA applied for lane 1 was first hybridized in 10-fold excess (170 μ g) to mAP-paper, which specifically binds poly(A)-containing RNA. After washing off the unbound RNA, bound chloroplast RNA was eluted with water at 70°C. This "poly(A)"-enriched chloroplast RNA fraction represented about one tenth of the RNA originally applied and was subsequently used for gel electrophoresis. RNA separated in lane 2 was not fractionated by absorption to mAP-paper. In lane 1 the position of a DNase sensitive band (chloroplast DNA) is also marked. The marker positions given alongside correspond to RNA markers whose relative positions were calibrated in several individual runs.

observed for the psbB operon from spinach (27). It is not clear, whether all the RNA species can function as mRNAs for rpoA translation or whether only the smallest species represents a translatable rpoA message.

Interestingly an almost identical pattern of rpoA specific RNAs is obtained, when a chloroplast RNA fraction enriched in oligo A sequences ("poly(A)" plus-RNA) is used for the Northern blot analysis (Fig. 4, lane 1). It appears, that the absorption of rpoA-specific RNA species to the oligo(dT)-containing matrix (mAP paper) is due to the numerous oligo(A) tracts present in the rpoA gene such as the regions 250-260, 475-490, 885-895, 1032-1044, 1093-1108 and 1379-1398 (see Fig. 2). The fractionation of rpoA-mRNA into the "poly(A)" plus RNA caused by the oligo(A) tracts may have been one reason for the presence of rpoA specific mRNA in the poly(A)-containing mRNA fraction from spinach, which was taken as evidence for a nuclear DNA encoded α -subunit of plastidic RNA polymerase (4).

Data obtained from S_1 mapping of rpoA-specific RNA are shown in Fig. 5. The four positions (a1, a2, b1 and b2) corresponding to the S_1 generated bands of the trailer region appear within pairs of short direct repeat sequences. Surprisingly, no S_1 generated band was observed distal to the inverted repeat structure of position 1460-1510. Apparently, the hairpin structure folded from this inverted repeat does not function as a chain terminator. However, chain termination mediated by this structure, but followed by a rapid processing to yield the observed termini (a1, a2, b1 and b2) can not be excluded. The absence of RNA sequences corresponding to the 3'-terminal hairpin also excludes its function as an mRNA-stabilizing element as proposed by Stern and Grussem (28). The sequence complementary to the inverted repeat region also can form a hairpin structure (not depicted). From its sequence and positional homology with the corresponding region from the spinach plastome (7) it is likely to represent the terminator signal for the psbB operon.

We have also tried to map the 5'-terminal end of rpoA mRNA. Several S_1 resistant bands are observed (data not shown) among which one major band corresponds to the nucleotide position

Fig. 5: S₁-mapping of the termini of the *rpoA* transcripts. For mapping of 3' termini (a) the 3' labelled restriction fragment *Bgl*III/*Pst*I, containing 390 bp, was used. Lanes 1-4 show product patterns obtained by base specific chemical degradation according to the Maxam-Gilbert sequencing technique (30); lane 5 shows the S₁ generated products obtained with maize chloroplast RNA; lane 6 shows a control in which *E. coli* tRNA was used instead of chloroplast RNA. In (b) the positions of the 3' termini derived from (a) are presented relative to the sequence of the trailer region and its hairpin structure.

Fig. 6: Western blot analysis of proteins from a maize chloroplast S-30 extract after SDS-polyacrylamide gel electrophoresis. For the two lanes to the left (a) preimmune serum (lane 1) and serum from the same animal after immunization with the C-terminal heptapeptide of the maize *rpoA* gene product were used. For the two lanes to the right (b) peptide induced serum was used in the absence (1) and in the presence (2) of 500 µg of the C-terminal heptapeptide. The band of lane 1 competed out in lane 2 by the peptide is marked. Positions of marker proteins of known molecular mass are also indicated.

398. This position codes within the N-terminal region of the *rpoA* gene (see Fig. 2) and might therefore represent a splice junction. It is also noteworthy that the amino acid sequences preceeding this position (positions 1 - 30 in Fig. 3) show very little homology to the *E. coli* *rpoA* gene product. This again may be a consequence of the true amino terminal sequences being encoded by a separate small exon further proximal to the *rpoA* gene. Experiments aiming at the unequivocal identification of the region coding for the N-terminal amino acids are in progress.

Identification of the *rpoA* encoded peptide by Western Blotting

From the DNA sequence depicted in Fig. 2 a polypeptide of a molecular mass of 38,96 Kd can be deduced (neglecting the

uncertainty of the N-terminus, see above). The N-terminal and the C-terminal heptapeptides, as deduced from the DNA sequence, were synthesized and used for induction of peptide antibodies according to standard procedures (19, 20, 29). Western blot analyses of a soluble protein fraction from maize chloroplasts with antibodies raised against the C-terminal heptapeptide are shown in Fig. 6. A band of the expected size between the markers of 36 and 48 Kd is observed (lane a2) which is absent when preimmune serum is applied (lane a1). This band is specific for the C-terminal heptapeptide as its appearance can be eliminated by competition with an excess of free heptapeptide (lanes b, 1 and 2). In view of the interspecies diversity of the derived C-terminal amino acid sequences (see Fig. 3), it appears highly unlikely that a putative nuclear gene would have conserved the identical C terminus. We therefore believe, that the 39 Kd peptide visible in the Western blot is encoded by the plastidic rpoA gene.

We have also tried, unsuccessfully, to raise antibodies against the N-terminal heptapeptide (data not shown). This failure again may reflect the fact that the N-terminal region, as depicted in Fig. 3, does not represent the true N terminus and that the latter is encoded by a separate exon as already mentioned above.

ACKNOWLEDGEMENT

The technical assistance of Mrs E. Schiefermayr and S. Krien is gratefully acknowledged. We thank Drs. E. Fritzsche and R.G. Herrmann for kindly providing us with the clones pZmc303, pWHsp20, respectively and Dr. Brüller for synthesizing the two terminal peptides of the maize rpoA gene and Dr. G.L. Igloi for the syntheses of oligonucleotides and for critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 206) and by the Fonds der Chemischen Industrie, B.R.D.

REFERENCES

1. Weil, J.H. (1987) Plant Sci. 49, 149-157.
2. Briat, J.F., Lescure, A.M., Mache, R. (1986) Biochimie 68, 981-990.
3. Bünge, W., Feierabend, J. (1980) Planta 149, 163-169.

4. Lerbs, S., Bräutigam, E., Parthier, B. (1985) *EMBO J.* 4, 1661-1666.
5. Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Haya-shida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., Sugiura, M. (1986) *EMBO J.* 5, 2043-2049.
6. Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-I., Inokuchi, H., Ozeki, H., (1986) *Nature* 322, 572-574.
7. Sijben-Müller, G., Hallick, R.B., Alt, J., Westhoff, P., Herrmann, R.G. (1986) *Nucl. Acids Res.* 14, 1029-1044.
8. Gruissem, W., Narita, J.O., Greenberg, B.M., Prescott, D.M., Hallick, R.B. (1983) *J. Cell Biochem.*, 22, 31-46.
9. Gruissem, W., Greenberg, B.M., Zurawski, G., Prescott, D.M., Hallick, R.B. (1983) *Cell* 35, 815-828.
10. Reiss, T., Link, G. (1985) *Eur. J. Biochem.* 148, 207-212.
11. Hanley-Bowdoin, L., Chua, N.-H. (1987) *TIBS* 12, 67-70.
12. Strittmatter, G., Gozdicka-Jozefiak, A., Kössel, H. (1985) *EMBO J.* 4, 599-604.
13. Fritzsche, E. (1988) Doctoral Thesis, University of Freiburg.
14. Feinberg, A.P., Vogelstein, B. (1983) *Anal. Biochem.* 132, 6-13.
15. Messing, J. (1983) *Methods Enzymol.* 101, 20-78.
16. Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) *Molecular Cloning*, Cold Spring Harbor Laboratory, NY.
17. Alwine, J.C., Kemp, D.J., Stark, G.R. (1977) *Proc. Natl. Acad. Sci.* 74, 5350-5354.
18. Mariottini, P., Chomyn, A., Attardi, G., Trovato, D., Strong, D.D., Doolittle, R.F. (1983) *Cell* 32, 1269-1277.
19. Nivison, H.T., Hanson, M.R. (1987) *Plant Mol. Biol. Reporter* 5, 295-309.
20. Towbin, H., Staehelin, T., Gordon, J. (1979) *Proc. Natl. Acad. Sci.* 76, 4350-4354.
21. Brunette, W.N. (1980) *Anal. Biochem.* 112, 195-203.
22. Bard, J., Bourque, D.P., Hildebrand, M., Zaitlin, D. (1985) *Proc. Natl. Acad. Sci.* 82, 3983-3987.
23. Larrinua, I.M., Muskavitch, K.M.T., Gubbins, E.J., Bogorad, L. (1983) *Plant Mol. Biol.* 2, 129-140.
24. Feng, D.F., Johnson, M.S., Doolittle, R.F. (1985) *J. Mol. Evol.* 21, 112-125.
25. Meek, D.W., Hayward, R.S. (1984) *Nucl. Acids Res.* 12, 5813-5821.
26. Tanaka, M., Obokata, J., Chunwongse, J., Shinozaki, K., Sugiura, M. (1987) *Mol. Gen. Genet.* 209, 427-431.
27. Westhoff, P., Herrmann, R.G. (1988) *Eur. J. Biochem.* 171, 551-564.
28. Stern, D.B., Gruissem, W. (1987) *Cell* 51, 1145-1157.
29. Walter, G. (1986) *J. Immun. Methods* 88, 149-161.
30. Maxam, A.M., Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.