Maize chloroplast RNA polymerase: the 78-kilodalton polypeptide is encoded by the plastid *rpoC1* gene

Jim Hu+, Robert F.Troxler§ and Lawrence Bogorad*
Department of Cellular and Developmental Biology, Harvard University, Cambridge, MA 02138, USA

Received December 13, 1990; Revised and Accepted March 28, 1991 Genbank accession no. M31207

ABSTRACT

The 180-, 120- and 38-kDa polypeptides found in highly purified maize plastid RNA polymerase preparations are encoded by the maize plastid genes *rpoC2*, *rpoB*, and *rpoA*, respectively [Hu, J. and Bogorad, L. (1990) Proc. Natl. Acad. Sci. USA. 87, pp. 1531 – 1535]. These genes have segments that specify amino acid sequences homologous to those of *E. coli* RNA polymerase subunits. The plastid gene products are designated b″, b and a, respectively. We report here that the amino-terminal amino acid sequence of a 78-kDa polypeptide also found in highly purified maize plastid RNA polymerase preparations matches precisely the sequence deduced from the maize plastid *rpoC1* gene which has segments homologous to the 5′ end of the *E. coli rpoC* gene. Thus, the 78-kDa polypeptide is likely to be a functional component of maize plastid DNA-dependent RNA polymerase. This polypeptide is designated subunit b′. Three polypeptides unrelated to RNA polymerase have also been identified in this preparation.

INTRODUCTION

Knowledge of the transcriptional apparatus, including the plastid RNA polymerases, is essential to understanding how plastid genes are transcribed and how plastid transcription may be regulated. In the past few years, sequences with regions homologous to *E. coli* genes *rpoA, rpoB* and *rpoC* have been identified in plastid genomes of tobacco (2), *Marchantia polymorpha* (3), spinach (4, 5), maize (1, 6, 7), rice (8) and Euglena (9). Interestingly, plastid sequences that have regions homologous to the *E. coli rpoC* gene is split into two genes, *rpoC1* and *rpoC2*. This feature probably evolved very early during the evolution of plants inasmuch as it also occurs in cyanobacteria (10). Because none of the plastid RNA polymerases has been reconstituted, it has not been possible to determine which of the polypeptides even in highly purified preparations are components of the enzymes. We approached this problem by micro-sequencing the N-termini of polypeptides whose presence is correlated with transcriptional activity in maize plastid RNA polymerase preparations. We found that the maize plastid genes *rpoA, rpoB* and *rpoC2* encode three polypeptides in maize plastid RNA polymerase preparations. These subunits are designated a, b, and b″ (1).

The experiments described here were designed to identify polypeptides in gel electrophoresis bands of approximately 78, 61 and 55 kDa that are also present in highly purified maize plastid RNA polymerase preparations. We report here that the sequence of amino acids two through ten deduced from the 5′-end of the coding region of the maize plastid gene *rpoC1* is identical to that in a 78-kDa polypeptide designated the b′-subunit. We also report that the 61 kDa band contains two polypeptides with amino-terminal sequences nearly identical to those of the a and b subunits of ribulose bisphosphate carboxylase subunit binding protein of pea (19). In addition, we have determined by immunoblotting that the 55-kDa band is the ribulose bisphosphate carboxylase large subunit.

MATERIALS AND METHODS

Maize chloroplast PF (DEAE Peak Fraction) RNA polymerase was extracted, purified and analyzed by SDS/polyacrylamide gel electrophoresis as described by Hu and Bogorad (1990). For amino-terminal sequencing, proteins in the major RNA polymerase activity peak fractions eluted from a Protein-Pak glass DEAE-5PW anion-exchange column (8.0 mm×7.5 cm; Nihon Waters, Tokyo) by a 100−500 mM KCl gradient were precipitated with 8% (V/V) trichloroacetic acid and separated on an SDS/polyacrylamide gel without further purification. After the proteins were blotted to a polyvinylidene difluoride (PVDF) membrane (11) and stained with Coomassie Brilliant Blue, protein bands containing polypeptides of 78- and 61-kDa were excised for amino terminal protein sequencing at the Harvard Biological Laboratories Microchemistry Facility. PF enzyme purified further by glycerol gradient centrifugation (highly purified RNA polymerase) was used for SDS/polyacrylamide gel electrophoresis and silver staining. *E. coli* RNA polymerase was purified according to Burgess (1976). Protein molecular weight standards were purchased from Sigma and silver staining reagents were purchased from Stratagene.

* To whom correspondence should be addressed

Present addresses: +Research Institute, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, M5G 1X8, Canada and §Department of Biochemistry, Boston University, School of Medicine, 80 East Concord Street, Boston, MA 02118, USA

RESULTS

Maize plastid PF RNA polymerase (12) has been purified extensively from sucrose density gradient-purified chloroplasts through a number of steps including PEG 8000 precipitation, DEAE chromatography and glycerol density gradient centrifugation (1). Enzymes purified through all of these steps are referred to as 'highly purified' in this paper. The 180-, 120- and 38-kDa polypeptides have been identified as the products of maize chloroplast genes *rpoC2*, *rpoB*, and *rpoA*, respectively (1). The only known chloroplast *rpo* gene whose product has been not identified is *rpoC1*. It is predicted to encode a 78-kDa protein (2, 3, 5, 7, 8). Possibly because of the low intensity of staining with Coomassie Brilliant Blue, a 78-kDa band was seen only very faintly in protein gels of highly purified RNA polymerase preparations (1). However, a silver stained polyacrylamide electrophoresis gel of the highly purified maize plastid PF RNA polymerase clearly shows bands of apparent molecular masses of 180,000, 120,000, 85,000, 78,000, 61,000, 55,000, and 38,000. A band is also discernible at the 64,000 position (Fig. 1). Because there appeared to be so little protein in the 78-kDa band after glycerol density gradient centrifugation, we used protein in the corresponding band from a preparation purified only after the DEAE chromatography step for micro-sequencing. The amino acid sequence obtained from this material was for a single protein (Fig. 3).

We had determined the nucleotide sequence of the 5′ end of the maize plastid *rpoC1* gene (GENBANK accession # M31207) earlier and the entire *rpoC1* gene from maize has now been sequenced independently by Igloi et al. (7). Fig. 3 shows that the amino-terminal sequence of the second through tenth amino acid residues of the 78-kDa protein revealed by protein sequencing matches precisely the sequence deduced from the 5′ end of the *rpoC1* gene. The protein product of the maize plastid *rpoC1* is predicted to have regions of homology to the 5′ end of the *E. coli* RNA polymerase b′-subunit which is encoded by the *rpoC* gene (7), suggesting that the 78-kDa protein is also a functional subunit of the maize plastid PF RNA polymerase. As shown in Fig. 3, the amino-terminal sequences deduced from the *rpoC1* genes of rice and maize are almost identical although a few variations can be found between the sequences of two monocots (rice and maize) and two dicots (tobacco and spinach); the corresponding region of the liverwort gene is somewhat more divergent. Since the overall sequence homology among the

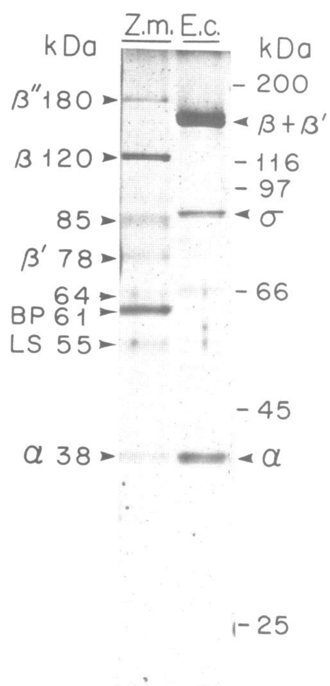

Fig. 1. A silver-stained SDS/polyacrylamide gel showing the electrophoretic separation of polypeptides present in highly purified *Zea mays (Z.m.)* chloroplast RNA polymerase PF enzyme (12) and in purified *E. coli* RNA polymerase *(E.c.)* holoenzyme. The 180-, 120-, and 38-kDa polypeptides (in *Z.m.*) had been identified previously (1) and are designated subunits b″, b and a, respectively. The *E. coli* RNA polymerase subunits b′, b, s, and a are marked. The sizes of molecular weight standards (not shown) are indicated by numbers at the right-hand side of the figure. The numbers at the left-hand side represent the sizes of the polypeptides in the *Z.m.* RNA polymerase preparation based on the positions of the molecular weight markers. The 61 kDa band labeled BP contains polypeptides with amino-terminal sequences homologous to those of the a and b subunits of pea ribulose bisphosphate carboxylase subunit binding protein (Fig. 2B, ref. 19). The 55 kDa band marked LS is shown to contain material that reacts with an antibody against tobacco ribulose bisphosphate carboxylase large subunit (data not shown).

61 kDa band	Residue (1 … 5 … 10)
Maize chap. alpha:	Ala-Ala-Lys-Asp-Ile-Ala-Phe-Asp-Gln- X -
Pea chap. alpha:	Ala-Ala-Lys-Asp-Ile-Ala-Phe-Asp-Gln-His-
Maize chap. beta:	Ala-Lys-Glu-Leu-Tyr-Phe-Asn-Lys-Asp-Gly-
Pea chap. beta:	Ala-Lys-Glu-Leu-His-Phe-Asn-Lys-Asp-Gly-

Fig. 2. Comparison of amino-terminal sequences of maize and pea (19) ribulose bisphosphate carboxylase large subunit binding protein. X represents a residue not identified positively. Abbreviation: Chap., chaperonin or ribulose bisphosphate carboxylase subunit binding protein.

```
M: ATG ATT GAC CAA TAT AAA CAT AAA CAA CTT CAA ATT GGA CTC GTT TCC CCT CAA
   Met Ile Asp Gln Tyr Lys His Lys Gln Leu Gln Ile Gly Leu Val Ser Pro Gln
   -----------------------------------------------------------------------
R:  :   :   :   :   :   :   :  Gln  :   :   :   :   :   :   :   :   :   :
T:  :   :   :  Arg  :   :   :  Gln  :   :  Arg  :   :  Ser  :   :   :   :
S:  :   :   :   :   :   :   :  Gln  :   :  Arg  :   :  Ser  :   :   :   :
L:  :  Thr Tyr  :  Lys  :   :  Gln His  :  Arg  :  Glu Leu Ala  :   :  Glu

M: CAA ATA AAG GCT TGG GCT AAA AAA ATC CTA CCT AAT GGG GAA GTC GTT GGC GAA
   Gln Ile Lys Ala Trp Ala Lys Lys Ile Leu Pro Asn Gly Glu Val Val Gly Glu
   -----------------------------------------------------------------------
R:  :   :   :   :   :   :  Asn  :  Thr  :   :   :   :   :   :   :   :   :
T:  :   :  Ser  :   :   :  Thr  :   :   :   :   :   :   :  Ile  :   :   :
S:  :   :  Ser  :   :   :  Thr  :   :   :   :   :   :   :  Ile  :   :   :
L:  :   :  Arg Asn  :   :  Glu Arg Val  :   :   :   :   :  Ile  :   :  Gln

M: GTC ACA AGG CCC TCC ACT TTT CAT TAT AAA ACC GAT AAA CCA GAA AAA GAT GGA
   Val Thr Arg Pro Ser Thr Phe His Tyr Lys Thr Asp Lys Pro Glu Lys Asp Gly
   -----------------------------------------------------------------------
R:  :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
T:  :   :  Lys  :  Tyr  :   :   :   :   :   :  Asn  :   :   :   :  Met Asp
S:  :   :  Lys  :  Tyr  :   :   :   :   :   :  Asn  :   :   :   :   :   :
L:  :   :  Lys  :  Tyr  :  Leu  :   :   :   :  His  :   :   :   :   :   :
```

Fig. 3. Nucleotide sequence of the 5′ end of the maize chloroplast *rpoC1* gene and its deduced amino acid sequence. The sequence of residues 2 through 10 (underlined) at the amino-terminus of the 78-kDa polypeptide was determined by micro-protein sequencing. This sequence matches precisely the amino-terminal sequence deduced from the maize chloroplast *rpoC1* gene. The amino-terminal amino acid sequences deduced from the *rpoC1* genes of rice (8), tobacco (2), spinach (5), and liverwort (3) are aligned under the maize sequence. Abbreviations: M, maize; R, rice; T, tobacco; S, spinach; L, liverwort.

identified plastid *rpoC1* genes is 70 to 96% (7), it is likely that all the plastid *rpoC1* genes encode functional RNA polymerase subunits.

We have often found that polypeptides of around 85, 64, 61 and 55 kDa are copurified with chloroplast RNA polymerase activity (1, 15, Fig. 1). The material in the 61 kDa band was blotted to a PVDF membrane and subjected to microsequencing. It was found (Fig. 2) to contain polypeptides with amino-terminal sequences nearly identical to those of the a and b subunits of pea ribulose bisphosphate carboxylase subunit binding protein (19).

Since ribulose bisphosphate carboxylase may be the most abundant protein complex in plants and the large subunit of this enzyme is also about 55-kDa, we investigated the possibility that the 55-kDa protein in maize plastid RNA polymerase preparations might be the ribulose bisphosphate carboxylase large subunit (1). We found that polypeptides in the 55-kDa band in maize plastid extracts as well as in PF RNA polymerase preparations (see Fig. 1) cross-react with a polyclonal antibody against the tobacco ribulose bisphosphate carboxylase large subunit (data not shown). Western blotting was done as described by Pluskal et al. (1986) and the antibodies was kindly provided by Dr. Steven Rodermel. The identities of the 85 and 64 kDa polypeptides are under investigation.

DISCUSSION

Since four *rpo* genes (i.e. genes with regions of homology to those encoding RNA polymerase subunits in *E. coli*) were found on plastid genomes (2−4), the *rpo* gene products have been sought. We have approached the problem of identifying products of *rpo* genes by first correlating transcriptional activity with the presence of polypeptides in maize chloroplast RNA polymerase fractions at different purification stages and subsequently microsequencing the amino termini of the polypeptides found in RNA polymerase preparations (1). We have reported that the 180-, 120- and 38-kilodalton polypeptides in the maize plastid RNA polymerase fractions are encoded by the plastid genes *rpoC2*, *rpoB*, and *rpoA*, respectively (1). The maize *rpoC1* gene is located downstream of the *rpoB* gene (1, 7). It has the potential to encode a 78-kDa polypeptide with 62−96% amino acid sequence identity to the deduced amino acid sequences from other plastid *rpoC1* genes and 36% amino acid identity to the first 600 amino acid residues of the *E. coli* RNA polymerase b′-subunit(7). We have shown here that the sequence of the second through tenth amino acid residues of the 78-kilodalton RNA polymerase subunit (b′-subunit) is identical to the sequences deduced from the corresponding portion of the maize plastid *rpoC1* gene. It seems most likely that the maize plastid *rpoC1* gene encodes the b′ subunit but it is possible that a nuclear copy of the gene occurs and that the plastid gene does not encode a functional subunit or that there are two proteins of the same size which are identical in their sequence of amino acids two through ten but are otherwise different proteins encoded by two different genes and *rpoC1* does not, in fact, encode the b′ subunit.

Thus, if the b′ subunit is encoded by *rpoC1*, the products of all the *rpo* genes that have been mapped on the maize plastid chromosome have been identified directly as proteins in maize plastid RNA preparations. Table 1 summarizes features of the maize plastid *rpo* genes and their gene products. Since plastid *rpo* genes are highly conserved among different plant species (7), the plastid *rpo* genes of other plants should also encode functional RNA polymerase subunits.

Comparisons with bacterial and other RNA polymerases make it seem likely that other subunits of plastid RNA polymerase — e.g. the equivalent of sigma factors—exist. They could be encoded either in nuclear genes or in plastid genes. If they are encoded in plastid chromosomes their sequences must be so different from those of known RNA polymerase polypeptides that they have not been revealed in comparisons of total chloroplast DNA sequences with sequences in available data banks.

The b′-subunit (78 kDa) of the maize chloroplast RNA polymerase was not stained as intensely as might be expected relative to other subunits with either silver (Fig. 1) or Coomassie Brilliant Blue (1). We do not know whether this reflects differences in the intrinsic staining characteristic of this protein or a lower ratio of the b′-subunit to other subunits in the polymerase fraction. The 180-kDa b″-subunit is also not stained with silver as well as the 120-kDa b-subunit, but it shows a comparable level of intensity when stained with Coomassie Brilliant Blue (1). Although the *E. coli* RNA polymerase subunits, a, b, and b′, form one type of core enzyme, a_2bb', it is not clear whether the plastid RNA polymerase subunits, a, b, b′ and b″, form one or more types of core enzymes. The plastid RNA polymerase core subunits might form a variety of core enzymes, for example, $a_2bb'b''$ or a_2bb' and a_2bb''. With the chloroplast transformation systems now available (16−18), it should be feasible to probe this problem *in vivo*.

Table 1. Maize plastid *rpo* genes and their gene products

gene	protein	size predicted from gene (kDa)	size estimated on gels* (kDa)	aa identity to corresponding sequence of E. coli (%)
rpoA	α	39	38	26
rpoB	β	121.6	120	37.5
rpoC1	β′	78.3	78	36.2**
rpoC2	β″	176.1	180	27.6***

*, based on migration in SDS/(10%wt/vol)polyacrylamide gel (1).
**, Homology to first 600 amino acid residues of the *E. coli* RNA polymease β′-subunit.
***, Homology to the last 800 amino acid residues of the *E. coli* RNA polymerase β′-subunit. Abbreviation: aa, amino acid.

ACKNOWLEDGEMENTS

We would like to thank Gisele Drouin for preparing maize chloroplasts. We are grateful to William S.Lane and Ruth Davenport at the Harvard Microchemistry Facility for protein sequencing. We are also grateful to Dr. Donald Jin for helpful comments on the manuscript. We would also like to thank Dr. Steven Rodermel for providing the antibody against tobacco ribulose bisphosphate carboxylase large subunit. This work was supported in part by a research grant from the National Institute of General Medical Sciences.

REFERENCES

1. Hu, J. and Bogorad, L. (1990) Proc. Natl. Acad. Sci. USA **87**, 1531−1535.
2. Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, 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. and Sugiura, M. (1986) EMBO J. **5**, 2043−2049.

3. Ohyama, K., Fukuzawa, H., Kohchi, T., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Nature (London) **322**, 572–574.
4. Sijben-Müller, G., Hallick, R. B., Alt, J., Westhoff, P. and Herrmann, R. G. (1986) Nucl. Acids Res. **14**, 1029–1044.
5. Hudson, G. S., Holton, T. A., Whitfeld, P. R. and Bottomley, W. (1988) J. Mol. Biol. **200**, 639–654.
6. Ruf, M. and Kössel, H. (1988) Nucl. Acids Res. **16**, 5741–5745.
7. Igloi, G., Meinke, A., Dory, I. and KÖssel, H. (1990) Mol. Gen. Genet. **221**, 379–394.
8. Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K. and Sugiura, M. (1989) Mol. Gen. Genet. **217**, 185–194.
9. Little, M. C. and Hallick, R. B. (1988) J. Biol. Chem. **263**, 14302–14307.
10. Xie, W.-Q., Jager, K. and Potts, M. (1989) J. Bacteriol. **171**, 1967–1973.
11. Matsudaira, P. (1987) J. Biol. Chem. **262**, 10035–10038.
12. Zaitlin, D., Hu, J. and Bogorad, L. (1989) Proc. Natl. Acad. Sci. USA **86**, 876–880.
13. Burgess, R. R. (1976) In *RNA polymerase*, R. Losick, M. Chamberlin, eds., PP. 69–100, Cold Spring Harbor Laboratories, Cold Spring Harbor, New York.
14. Pluskal, M. G., Przekop, M. B., Kavonian, M. R., Vecoli, C. and Hicks, D. A. (1986) BioTechniques **4**, 272–282.
15. Kidd, G. H. and Bogorad, L. (1980) Biochim. Biophys. Acta **609**, 14–30.
16. Boynton, J. E., Gillham, N. W., Harris, E. H., Hosler, J. P., Johnson, A. M., Jones, A. R., Randolph-Anderson, B. L., Robertson, D., Klein, T. M., Shark, K. B. and Sanford, J. C. (1988) Science **240**, 1534–1538.
17. Blowers, A. D., Bogorad, L., Shark, K. B. and Sanford, J. C. (1989) Plant Cell **1**, 123–132.
18. Svab, Z., Hajdakiewicz, P. and Maliga, P. (1990) Proc. Natl. Acad. Sci. USA **87**, 8526–8530.
19. Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K. Dennis, D. T. Georgopoulos, C. P., Hendrix, R. W. and Ellis, J. R. (1988) Nature **333**, 330–334.