Volume 11 Number 6 1983 Nucleic Acids Research

A putative gene of tobacco chloroplast coding for ribosomal protein similar to *E. coli* ribosomal protein S19

Mamoru Sugita* and Masahiro Sugiura**

National Institute of Genetics, Mishima 411, and Department of Biology, Nagoya University, Chikusa, Nagoya 464, Japan

Received 25 January 1983; Accepted 17 February 1983

ABSTRACT

A partial sequence of a cloned 3.2 Md BamHI fragment from tobacco chloroplast DNA revealed the occurrence of a putative gene for ribosomal protein. The putative gene is located on the left margin of the large single-copy region in the chloroplast DNA. The coding region contains 276 bp (92 codons). The amino acid sequence deduced from the DNA sequence shows 55% homology with that of E. coli S19 (91 amino acid residues).

INTRODUCTION

Chloroplasts are known to contain circular DNAs of about 10^8 daltons and their own transcriptional and translational machineries (1). Chloroplast ribosomes are 70S in size and closely related to prokaryotic ribosomes. Chloroplast ribosomes of higher plants contain 23S, 16S, 5S and 4.5S rRNAs which are coded for by the chloroplast DNA (2). The nucleotide sequences of rRNA genes from maize and tobacco chloroplasts have been determined and found to be highly homologous to those from E. coli (3-7). Chloroplast ribosomal proteins (r-protein) in higher plants have been analyzed by two-dimentional gel electrophoresis and their number in the 30S and 50S subunits (24-26 and 34-36, respectively) were found to be close to that of E. coli r-proteins (8-10). The primary structure of chloroplast r-protein L12 from spinach has recently been determined (11). Chloroplast r-proteins are believed to be encoded in both nuclear and chloroplast DNAs. Eneas-Filho et al. have reported that at least 6 r-proteins of the 30S and 5 of the 50S subunits were made in the isolated chloroplast from pea (10). However, no genes for r-proteins in chloroplast DNA have been-to our knowledge - identified and sequenced so far.

During the course of sequencing a 3.2 Md BamHI fragment of tobacco chloroplast DNA, we found an open reading frame (ORF) whose amino acid sequence was similar to that of E. coli r-protein S19. We present here the structure of a putative r-protein gene from tobacco chloroplasts.

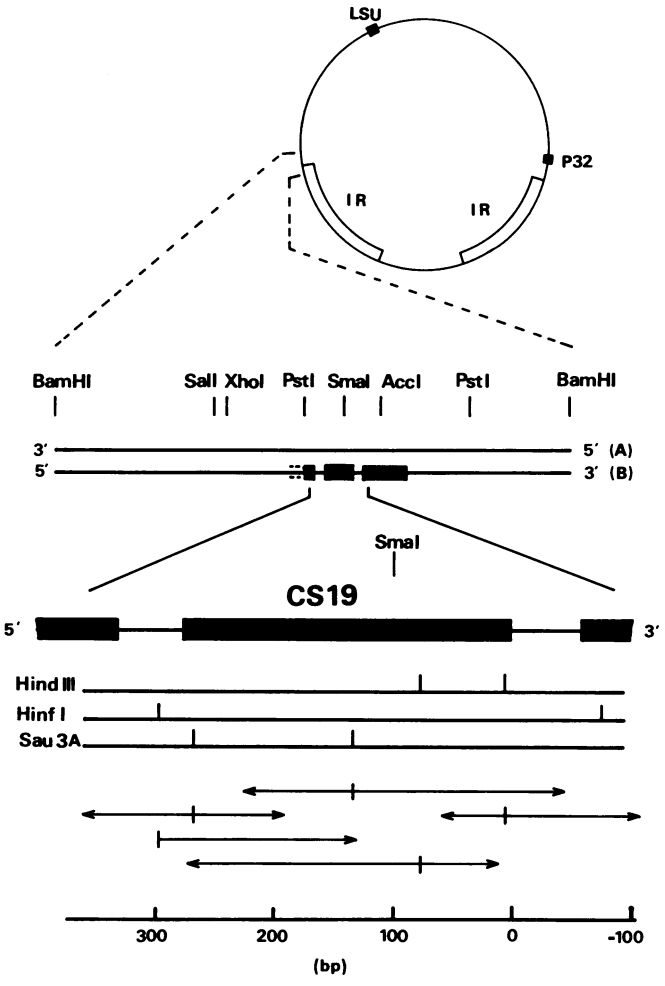

Fig. 1. Location, physical map of the cloned 3.2 Md BamHI fragment (Ba7) from tobacco chloroplast DNA and the strategy for sequencing part of it. LSU, gene for the large subunit of ribulose-1,5-bisphosphate carboxylase; P32, gene for the 32 kilodalton thylakoid protein and IR, inverted repeat. (A), DNA strand which codes for LSU and (B), its complementary strand. Coding regions are shown by thick lines.

MATERIALS AND METHODS

Recombinant plasmid pTB7 containing a 3.2 Md BamHI fragment of Nicotiana tabacum (var. Bright Yellow 4) chloroplast DNA was constructed as described (12). DNA sequence was determined by the method of Maxam and Gilbert (13).

RESULTS AND DISCUSSION

Recombinant plasmid pTB7 contains a 3.2 Md BamHI fragment (designated Ba7) of tobacco chloroplast DNA. Hybridization experiments indicated that the Ba7 fragment is contained in the Sal-7, Bgl-4, Pvu-1 and Xho-11 fragments on the physical map of tobacco chloroplast DNA constructed by Seyer et al. (14). The Ba7 fragment is therefore located on the left margin of the large single-copy region and may contain a part of the left segment of an inverted repeat.

A physical map of the cloned Ba7 fragment and the strategy for sequencing part of it are shown in Fig. 1. Fig. 2 shows the nucleotide sequence of a 480 bp portion which contains an ORF of 276 bp (92 codons). We found that the amino acid sequence

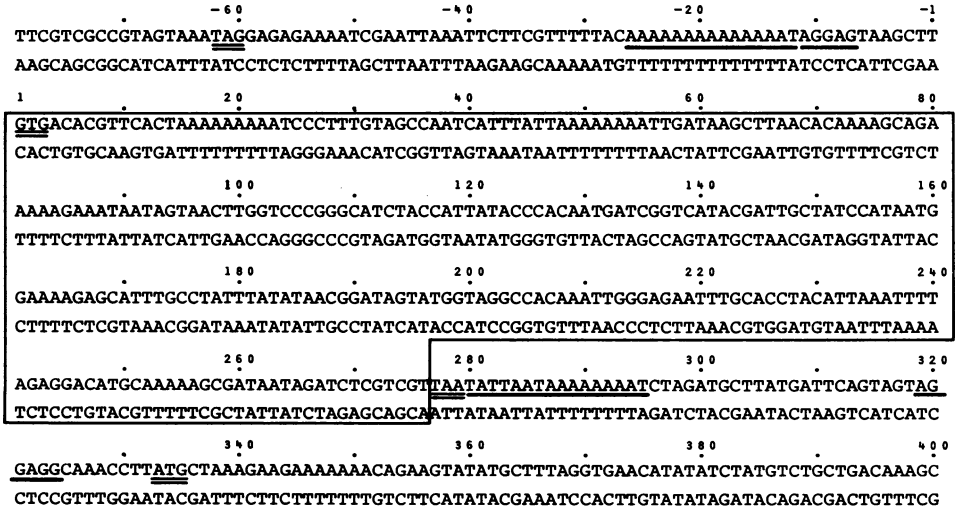

Fig. 2. Nucleotide sequence of the putative gene for CS19 and its flanking regions. The coding region is boxed. Putative ribosome binding sites and AT-pairs are underlined.

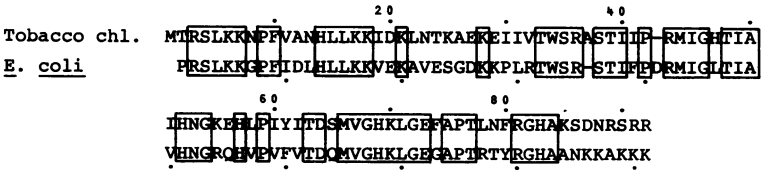

Fig. 3. Comparison between the amino acid sequence deduced from the putative gene for CS19 of tobacco chloroplast and that of E. coli S19. Homologies are indicated by boxes.

deduced from the DNA sequence shows 55% homology with that reported for E. coli S19 (15) as shown in Fig. 3. The molecular weight calculated from the deduced amino acid composition is 10443 which is quite similar to that (10299) of E. coli S19 (15). Based on the sequence homology, we tentatively identified this ORF to be a gene for a chloroplast r-protein corresponding to E. coli S19 and designated gene rpsS or r-protein CS19. A 55% homology seems to be high enough to support the above idea

Table 1 Codon usage in the putative CS19 gene of tobacco chloroplast (Tc) and in several r-protein genes of E. coli (Ec)

1st	2nd position											3rd
	U			C			A			G		
		Tc	Ec		Tc	Ec		Tc	Ec		Tc	Ec	
U	Phe	3	9	Ser	2	24	Tyr	1	4	Cys	0	1	U
	Phe	0	17	Ser	1	21	Tyr	0	13	Cys	0	6	C
	Leu	3	4	Ser	1	1	Term	1	7	Term	0	1	A
	Leu	2	3	Ser	0	2	Term	0	0	Trp	1	4	G
C	Leu	1	4	Pro	2	4	His	5	4	Arg	3	46	U
	Leu	0	3	Pro	2	1	His	1	9	Arg	0	24	C
	Leu	1	0	Pro	0	6	Gln	0	8	Arg	0	0	A
	Leu	0	67	Pro	0	32	Gln	0	26	Arg	1	1	G
A	Ile	4	16	Thr	1	32	Asn	5	4	Ser	1	4	U
	Ile	2	37	Thr	1	21	Asn	1	32	Ser	1	7	C
	Ile	4	1	Thr	4	3	Lys	9	77	Arg	2	1	A
	Met	2	24	Thr	2	2	Lys	1	28	Arg	0	0	G
G	Val	0	47	Ala	1	75	Asp	3	14	Gly	1	44	U
	Val	0	8	Ala	1	13	Asp	0	31	Gly	1	35	C
	Val	3	42	Ala	4	42	Glu	3	58	Gly	3	1	A
	Val	0	17	Ala	0	27	Glu	1	15	Gly	0	0	G

Tc: The initiation codon (GUG) is not included. Ec: Includes the genes for L11, L1, L10, L7/L12, and in N-and/or C-terminal portions of S7, L14, S17, S4, S11, L16 and S13 genes (19).

because the amino acid sequence of r-protein L12 from spinach chloroplast shows about 50% homology with that of E. coli L12 (11) and 25-62% sequence homologies were found in the α, β and ε subunits of H^+-ATPases between tobacco chloroplast and E. coli (16, 17).

A sequence AGGAG, which is complementary to the 3' end of tobacco chloroplast 16S rRNA (--CUCCU--3', ref. 7) and a putative ribosome binding site, is present -8 to -12 bp upstream from the initiation codon GTG (see Fig. 2). A marked feature in the flanking regions is 15 AT pairs before the putative ribosomal binding site and 16 AT pairs after the termination codon TAA (see Fig. 2).

E. coli S19 gene is in the S10 operon which is located at 72 min on the E. coli chromosome and comprises the genes for S10, L3, L4, L23, L2, L22, S19, S3, L16, L29 and S17 (18). Preliminary sequence analysis indicated that there are ORFs about 60 bp before and after the putative gene for CS19 (see Fig. 1). These two ORFs have little homology with any of the r-proteins encoded in the S10 operon of E. coli.

Table 1 shows the codon usage in the putative CS19 gene and in several genes for r-proteins in E. coli. The codon usage in the putative CS19 gene differs from that of E. coli r-protein genes. We also found differences between the codon usage of tobacco chloroplast and E. coli genes for the α, β and ε subunits of H^+-ATPases (16, 17).

ACKNOWLEDGEMENTS

We thank Mrs. Mie Kusuda for construction of a recombinant plasmid, Drs. K. Shinozaki and F. Takaiwa for valuable suggestions and Emeritus Prof. K. Sasaki for encouragement. This work was supported in part by a Grant-in-Aid from the Ministry of Education, Science and Culture.

*Present Address: Department of Botany, Hokkaido University, Sapporo 060, Japan.

**To whom correspondence should be addressed; Department of Biology, Nagoya University, Chikusa, Nagoya 464, Japan.

REFERENCES

1. Herrmann, R.G. and Possingham, J.V. (1980) in Chloroplast, Reinert, J. Eds., pp. 45-96, Springer-Verlag, Berlin.

2. Sugiura, M., Kusuda, J., Shinozaki, K. and Takaiwa, F. (1980) in Genetics and Evolution of RNA polymerase, tRNA and ribosomes, Osawa, S. et al. Eds., pp. 437-449, Univ. Tokyo Press, Tokyo.
3. Schwarts, Z. and Kössel, H. (1980) Nature 283, 739-742.
4. Edwards, K. and Kössel, H. (1981) Nucleic Acids Res. 9, 2853-2869.
5. Takaiwa, F. and Sugiura, M. (1980) Mol. Gen. Genet. 180, 1-4.
6. Takaiwa, F. and Sugiura, M. (1982) Eur. J. Biochem. 124, 13-19.
7. Tohdoh, N. and Sugiura, M. (1982) Gene 17, 213-218.
8. Mache, R., Dorne, A.-M. and Batlle, R.M. (1980) Mol. Gen. Genet. 177, 333-338.
9. Capel, M.S. and Bourque, D.P. (1982) J. Biol. Chem. 257, 7746-7755.
10. Eneas-Filho, J., Hartley, M.R. and Mache, R. (1981) Mol. Gen. Genet. 184, 484-488.
11. Bartsch, M., Kimura, M. and Subramanian, A.-R. (1982) Proc. Natl. Acad. Sci. USA 79, 6871-6875.
12. Sugiura, M. and Kusuda, J. (1979) Mol. Gen. Genet. 172, 137-141.
13. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
14. Seyer, P., Kowallik, K.V. and Herrmann, R.G. (1981) Current Genetics 3, 189-204.
15. Yaguchi, M. and Wittmann, H.G. (1978) FEBS Lett. 88, 227-230.
16. Deno, H., Shinozaki, K. and Sugiura, M. (1983) submitted.
17. Shinozaki, K., Deno, H., Kato, A. and Sugiura, M. (1983) submitted.
18. Lindahl, L., Post, L., Zengel, J., Gilbert, S.F., Strycharz, W.A. and Nomura, M. (1977) J. Biol. Chem. 252, 7365-7383.
19. Post, L.E. and Nomura, M. (1980) ibid. 255, 4660-4666.