Volume 281, number 1,2, 64–66 FEBS 09554 April 1991

ADONIS 001457939100264E

The chloroplast gene for ribosomal protein CL23 is functional in tobacco

Fumiaki Yokoi, Minoru Tanaka*, Tatsuya Wakasugi and Masahiro Sugiura

Center for Gene Research, Nagoya University, Nagoya 464-01, Japan

Received 3 January 1991; revised version received 6 February 1991

Chloroplast *rpl23* loci potentially coding for a polypeptide homologous to the *E. coli* L23 ribosomal protein are frame-shifted in spinach and several other plants, indicating that these loci are pseudogenes. In tobacco, *rpl23* constitutes a continuous open reading frame of 93 codons and its transcript initiates at least 66 bp upstream from the initiation codon. The N-terminal amino acid sequence of a 13 kDa protein from the 50 S subunit of tobacco chloroplast ribosomes matches that derived from the tobacco *rpl23* locus. This shows that *rpl23* is a functional gene in tobacco.

Chloroplast; Ribosomal protein; *rpl23*; Pseudogene; Tobacco

1. INTRODUCTION

Chloroplast ribosomes are 70 S in size similar to *E. coli* ribosomes and their protein components are encoded in two separate genetic systems, the chloroplast and the nucleus. Determination of the entire chloroplast DNA sequences revealed the presence of 20 loci potentially coding for polypeptides homologous to *E. coli* ribosomal proteins (reviewed in [1]). Recently an additional gene for a ribosomal protein *(rp132)* was found in the tobacco chloroplast genome, indicating that the chloroplast genome contains 21 possible genes for ribosomal proteins [2].

An open reading frame (ORF) whose predicted amino acid sequence resembles that of the *E. coli* L23 protein was first found in tobacco chloroplast DNA as the starting gene *(rpl23)* of the chloroplast ribosomal protein gene cluster (the *rpl23* cluster) similar to *E. coli* S10 and *spc* operons [3]. Subsequently, *rpl23* sequences were found in a variety of plants as part of the gene cluster [4–10]. Among them *Euglena rpl23* contains two introns [7]. The cyanelle genome of *Cyanophora paradoxa*, however, lacks *rpl23* but contains *rpl3* in the corresponding gene cluster [11].

Among chloroplast *rpl23* sequences, the *rpl23* regions from spinach and six other plants have the extensive addition/deletion change and constitute no single ORF [4]. In spinach, the *rpl23* sequence is split into two overlapping ORFs of 47 and 45 codons and no gene products were detected while transcripts from them were present [6]. These observations together with the low homology to the *E. coli* L23 protein suggested either a loss of requirement of L23 in chloroplast ribosomes or a translocation of this *rpl23* gene into the nuclear genome, even in the case of *rpl23* consisting of a continuous ORF.

Here we report the presence of the gene product, CL23, from *rpl23* in chloroplast ribosomes in tobacco, confirming that *rpl23* is a functional gene at least in tobacco.

Correspondence address: M. Sugiura, Center for Gene Research, Nagoya University, Nagoya 464-01, Japan.

**Present address:* National Institute for Basic Biology, Okazaki 444, Japan

2. MATERIALS AND METHODS

Total chloroplast RNA was prepared from young tobacco *(Nicotiana tabacum* var. Bright Yellow 4) leaves as described [12]. Oligodeoxyribonucleotides of 30-mers (A, B and C) were synthesized by a DNA synthesizer (Applied Biosystems 380A). The RNA was denatured in 50% formaldehyde at 50°C for 1 h and electrophoresed on 1.1% agarose gels containing 6% formaldehyde. The fractionated RNA was blotted onto a nylon membrane (Gene Screen Plus) sheet and hybridized with [5'-^{32}P]oligonucleotide probes (A and C) as described [12]. Primer extension assay was carried out using the oligonucleotide B for a primer as described [12].

Chloroplast ribosomal subunits were prepared from mature tobacco leaves essentially by the reported method [13]. Ribosomal proteins were separated by reverse-phase chromatography followed by SDS-polyacrylamide gel electrophoresis (details will be published elsewhere). A protein band was blotted onto a polyvinylidene difluoride membrane sheet and its N-terminal amino acid sequence was determined by a gas-phase protein sequencer (Applied Biosystems 470A-120A).

Computer-aided analysis was performed using the GENETYX program on a NEC PC-98XL personal computer.

3. RESULTS AND DISCUSSION

3.1. *Transcripts from* rp123

rpl23 is the first gene in the *rpl23* cluster and present twice in the tobacco chloroplast genome. The gene organizations are *trnI* (74 bp)–165 bp spacer–*rpl23* (93 codons)–21 bp spacer–*rpl2* (274 codons + 666 bp intron) in the IR_B followed by ten other genes in the LSC and the same genes/spacers in the IR_A followed by *trnH* on the opposite strand of the LSC [3,14].

Published by Elsevier Science Publishers B. V.

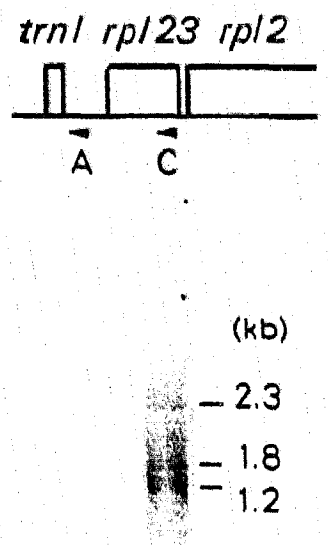

Fig. 1. Detection of transcripts from *rpl23* by Northern blot assay. The upper part shows the organization of *rpl23* and adjacent genes. Arrowheads indicate the positions of 30-mer oligonucleotide probes (A, positions −87 to −116; C, positions 270 to 241). The numbering starts from the first nucleotide of *rpl23*.

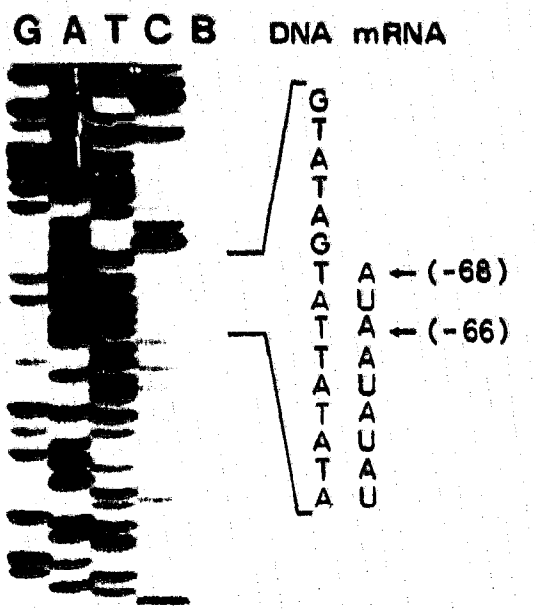

Fig. 2. Determination of the 5′ end of a transcript from *rpl23* by primer extension assay. Total tobacco chloroplast RNA was annealed with oligonucleotide B (positions 88 to 59), extended with AMV reverse transcriptase. The gel pattern shows the extended DNA (B) in parallel with the sequence ladders (G, A, T, C) of the DNA strand complementary to oligonucleotide B. The same result was obtained by using oligonucleotide C as a primer (data not shown).

Multiple transcripts from *rpl23* were reported previously [3]. Northern blot hybridization using probe C has confirmed the presence of major transcripts of 2.3 kb, 1.8 kb and 1.2 kb from *rpl23* as shown in Fig. 1. No transcript was detected by probe A, indicating that the transcription starts between probe A and probe C positions (data not shown). Primer extension analysis was then performed using oligonucleotide B as a primer. Extended fragments were observed at positions −66 and −68 as shown in Fig. 2. A Pribnow-box-like sequence TATATTC is present 6-8 bp upstream from these positions. Based on these observations, the −66 and −68 positions are most likely the initiation sites of *rpl23* transcription. The presence of multiple transcripts from *rpl23* is probably due to two copies of *rpl23* (one is in IR_B and another in IR_A) and to processing of the primary transcript.

In spinach, Thomas et al. detected by S1 mapping transcripts starting beyond the upstream *trnI* gene and in front of and within the split *rpl23* region, while they found no corresponding gene products and described it as a pseudogene transcribed [6]. The second starting site, −66 to −67 bp upstream from *rpl23*, in spinach matches the initiation site in tobacco but no larger transcripts including the *trnI* sequence were detected in tobacco by Northern blot analysis as described above. Primer extension analysis using oligonucleotide C indicated the presence of several faint bands falling into the tobacco *rpl23* coding region (data not shown). At present we cannot tell whether these 5′ ends represent transcription initiations, processing sites or extension artifacts.

3.2. *The gene product from* rp123

The 50 S subunit proteins from tobacco chloroplast ribosomes were fractionated by reverse-phase chromatography. The 33th fraction was further resolved into three bands of 28 kDa, 26 kDa and 13 kDa (bands 33A, 33B and 33C, respectively) by SDS-polyacrylamide gel electrophoresis (details will be published elsewhere). The N-terminal amino acid sequence of the 13 kDa ribosomal protein (33C) is identical to that deduced from the tobacco *rpl23* sequence as shown in Fig. 3. Based on this correspondence and the presence of its transcripts mentioned above, we have concluded that *rpl23* is functional at least in tobacco chloroplasts.

Fig. 3 also shows comparisons of amino acid sequences deduced from other *rpl23* regions so far reported. *rpl23*s in rice, wheat and maize consist of ORFs of 93 codons and show high amino acid sequence homology with that in tobacco, suggesting that these *rpl23*s are functional. The liverwort and *Euglena rpl23* sequences contain 91 and 100 codons, respectively, and show low amino acid sequence homology with tobacco CL23. The deduced amino acid sequences from spinach

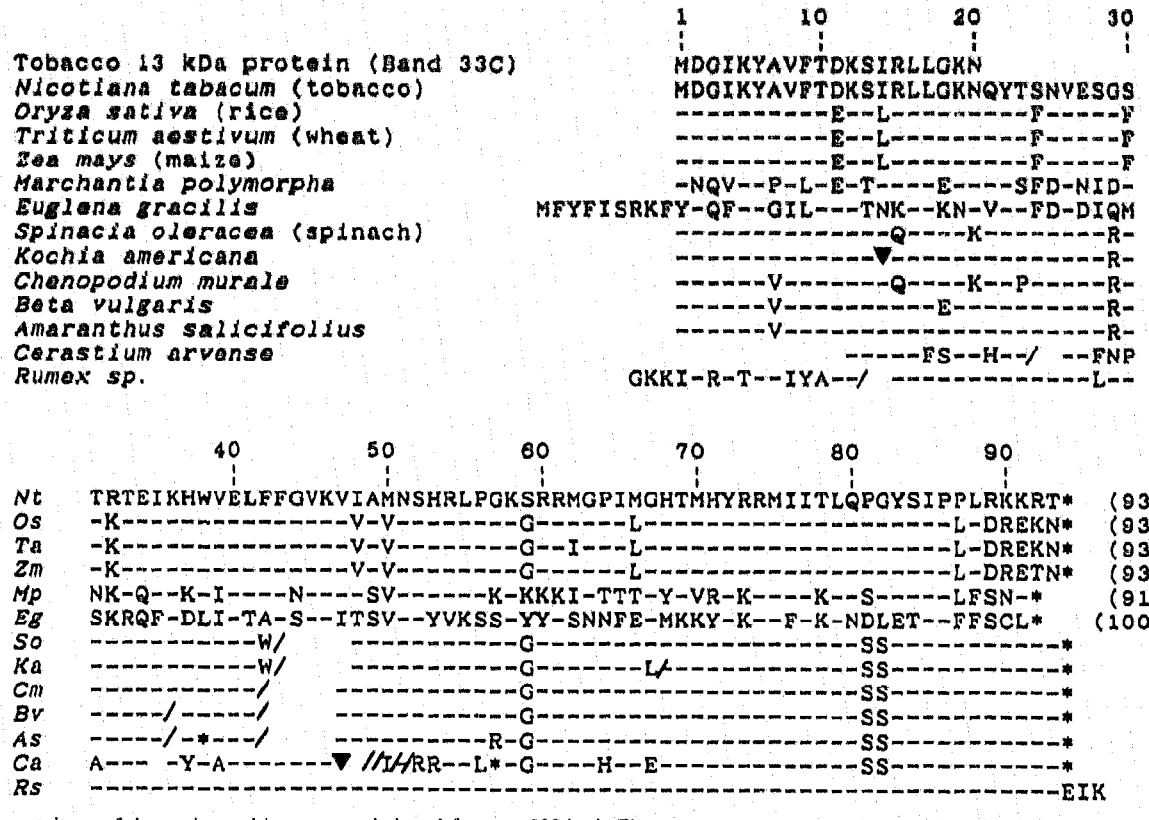

Fig. 3. Comparison of the amino acid sequence deduced from *rpl23* loci. The uppermost sequence is the determined N-terminal 20 amino acid sequence of the tobacco 13 kDa protein. Numbers of residues are in parentheses. Dashes indicate identical residues with those of tobacco. Slashes show sites of frame-shifts, and triangles indicate sites of insertions. Asterisks are termination codons.

and six other species (all belong to *Caryophylliidae*) again show high homology with that from tobacco but they are frame-shifted.

The present result suggests that CL23 protein is required in chloroplast ribosomes and *rpl23* has been transferred into the nuclear genome of some plants. If so, it remains a puzzle why apparently defective copies of *rpl23* are present in the chloroplast genome. It is necessary to isolate cDNAs for CL23 in spinach or related plants.

Acknowledgements: We thank Dr. A. Vassileva and Dr. N. Hayashida for help, and Dr. M. Sugita and Mr. K. Torazawa for suggestions and discussions. This work was supported by a Grant-in-Aid from the Ministry of Education.

REFERENCES

[1] Sugiura, M. (1989) Annu. Rev. Cell Biol. 5, 51-70.
[2] Yokoi, F., Vassileva, A., Hayashida, N., Torazawa, K., Wakasugi, T. and Sugiura, M. (1990) FEBS Lett. 276, 88-90.
[3] Tanaka, M., Wakasugi, T., Sugita, M., Shinozaki, K. and Sugiura, M. (1986) Proc. Natl. Acad. Sci. USA 83, 6030-6034.
[4] Zurawski, G. and Clegg, M.T. (1987) Annu. Rev. Plant Physiol. 38, 391-418.
[5] Fukuzawa, H., Kohchi, T., Sano, T., Shirai, H., Umesono, K., Inokuchi, H., Ozeki, H. and Ohyama, K. (1988) J. Mol. Biol. 203, 333-351.
[6] Thomas, F., Massenet, O., Dorne, A.M., Briat, J.F. and Mache, R. (1988) Nucleic Acids Res. 16, 2461-2472.
[7] Christopher, D.A., Cushman, J.C., Price, C.A. and Hallick, R.B. (1988) Curr. Genet. 14, 275-286.
[8] McLaughlin, W.E. and Larrinua, I.M. (1988) Nucleic Acids Res. 16, 8183.
[9] Hiratsuka, J., Shimada, H., Whittier, R.F., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, M., 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.
[10] Bowman, C.M., Baker, R.F. and Dyer, T.A. (1988) Curr. Genet. 14, 127-136.
[11] Evrard, J.L., Kuntz, M. and Weil, J.H. (1990) J. Mol. Evol. 30, 16-25.
[12] Yao, W.B., Meng, B.Y., Tanaka, M. and Sugiura, M. (1989) Nucleic Acids Res. 23, 9583-9591.
[13] Capel, M.S. and Bourque, D.P. (1982) J. Biol. Chem. 257, 7746-7755.
[14] 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.