Chloroplast ribosomal protein L32 is encoded in the chloroplast genome

Fumiaki Yokoi, Anelia Vassileva*, Nobuaki Hayashida**, Keita Torazawa, Tatsuya Wakasugi and Masahiro Sugiura

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

Received 24 September 1990

The 50 S subunit of chloroplast ribosomes was prepared from tobacco leaves. The proteins were fractionated and the N-terminal amino acid sequence of a 14 kDa protein was determined. This sequence matches the N-terminal sequence deduced from ORF55 located between *ndhF* and *trnL* on the small single-copy region of tobacco chloroplast DNA. The deduced protein shows homology to *E. coli* and *B. stearothermophilus* L32 proteins, and it has been named as CL32 and ORF55 as *rpl32*. The tobacco chloroplast genome therefore contains 21 different ribosomal protein genes.

Chloroplast; Ribosomal protein; *rpl32*; Tobacco

1. INTRODUCTION

Chloroplasts of green plants contain their own DNA, replication, transcription and translation apparatus. Chloroplast ribosomes are 70 S in size similar to prokaryotic ribosomes and contain 3-5 rRNAs and about 60 ribosomal proteins (reviewed in [1]). All the rRNAs are known to be encoded in the chloroplast genome (reviewed in [1]). Analyses of the entire sequence of tobacco, rice and liverwort chloroplast genomes have shown that the chloroplast genomes contain 20 sequences potentially coding for polypeptides homologous to *E. coli* ribosomal proteins [2-4]. The rest of chloroplast ribosomal proteins were thought to be encoded in the nuclear genome.

In tobacco chloroplasts all the 20 putative genes for ribosomal proteins are transcribed [5]. In spinach the translation products from *rpl2, rpl14, rpl16, rpl22, rpl33, rpl36, rps12* and *rps19* have so far been identified [6-8], indicating that these genes are functional in spinach. In order to confirm whether the tobacco 20 genes are functional, we started to analyze the ribosomal proteins from tobacco chloroplast 50 S subunit. During the course of the analysis we found a new ribosomal protein gene, *rpl32*, in the chloroplast genome. Our finding indicates the presence of 21 different chloroplast genes potentially coding for ribosomal proteins.

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

**Present address:* Institute of Genetic Engineering, 2232 Kostinbrod-2, Bulgaria

***Present address:* RIKEN Institute, Tsukuba 305, Japan

2. MATERIALS AND METHODS

The 50 S subunit of chloroplast ribosomes was prepared from tobacco leaves (*Nicotiana tabacum* var. Bright Yellow 4) essentially according to the method of Capel and Bourque [9]. The ribosomal proteins were fractionated by reverse-phase chromatography (Pharmacia column PRO-RPC HR5/10) followed by SDS-polyacrylamide gel electrophoresis (details will be published elsewhere). The protein band was transferred to a polyvinylidene difluoride membrane and its N-terminal amino acid sequence was determined by a gas-phase protein sequencer (Applied Biosystems 470A-120A). Computer-assisted analysis was carried out using the GENETYX program on a NEC PC-98XL personal computer.

3. RESULTS AND DISCUSSION

The reverse-phase chromatography of 50 S ribosomal proteins from tobacco chloroplasts produced at least 12 sharp peaks (details will be published elsewhere). The 13th fraction was resolved further into two bands of 20 kDa and 14 kDa (13A and 13B, respectively) by SDS-polyacrylamide gel electrophoresis. After blotting to a polyvinylidene difluoride membrane, the 14 kDa band was subjected to protein sequencing.

The N-terminal amino acid sequence of the 14 kDa ribosomal protein (13B) matches the sequence of the first 18 amino acids deduced from an open reading frame of 55 codons (ORF55) present in the tobacco chloroplast genome [2] as shown in Fig. 1. ORF55 is located between *ndhF* and *trnL* on the small single-copy region of the genome (Fig. 2). Computer-assisted search

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

```
EL32   AVQQNKPTRSKRGMRRSHDALTAVTSLSVDKTSGEKHLRHHIT-ADGYYRGR-KVIAK  ( 56 )
       **       **   *        *      * * *          *   * *
CL32   AVPKKRTSTSKKRIRKNIWKRK-GYSIALKAFSLAKSLSTGNSK---SFFVRQTKINK  ( 54 )
       AVPKKRTSTSKKRIRKNI
       ***  ***   ** *    *                *     *       *    **
BL32   AVPFRRTSKTRKRLRRTHFKLQ-VPGMVQCPNCGEWKLAHRVCKACGTYKGR-DVVNK  ( 56 )
```

Fig. 1. Comparison of the amino acid sequence of tobacco chloroplast ribosomal protein CL32 (CL32) predicted form ORF55 with those of the *E. coli* L32 (EL32) and *B. stearothermophilus* L32 (BL32). Asterisks indicate identical amino acids. The determined N-terminal amino acid sequence of the 14 kDa protein is shown below CL32. The initial methionines are omitted. Parentheses indicate residue numbers. Dashes denote gaps introduced to optimize sequence alignment.

revealed that the translation product deduced from ORF55 has similarity to ribosomal protein L32s of *E. coli* [10] and *B. stearothermophilus* [11] (20% and 26% homology, respectively) as shown in Fig. 1. Hydropathy profiles for respective proteins were further compared and found to be highly similar to each other (Fig. 3). Our preliminary Northern blot analysis revealed the presence of a major transcript of 1.5 kb from *rpl32* (unpublished result). We have therefore concluded that the 14 kDa protein is a homologue to the *E. coli* L32 protein, and designated it as CL32 and ORF55 as *rpl32*. Tobacco CL32 has a calculated molecular mass of 6185 Da (the first methionine is excluded) and is a basic protein rich in lysine residues (12 residues out of 54).

rpl32 is also present in rice, broad bean and liverwort chloroplast genomes [3,4,12]. As shown in Fig. 4, homologies among these CL32 proteins are high in their N-terminal halves (up to the first 45 residues) but not in their C-terminal halves, and the size of CL32s differs from species to species. This feature is sometimes observed in other chloroplast ribosomal proteins deduced from the DNA sequences, e.g. tobacco CL22 of 155 residues versus spinach CL22 of 199 residues [5,7]. The extreme case is the *E. coli* S7 protein, 177 residues in strain K12 and 154 residues in strain B [13]. These observations suggest that the C-terminal parts of some of the ribosomal proteins are not important for their functions.

As an additional gene coding for a ribosomal protein was found, the chloroplast genomes from tobacco, rice and liverwort contain 21 different genes potentially coding for ribosomal proteins (*rps16* in absent but *rpl21* is present in liverwort). Recently nuclear-encoded ribosomal proteins showing no sequence similarity with any of the *E. coli* ribosomal proteins were found in spinach chloroplasts (R. Mache and A.R. Subramanian, personal communication). We cannot therefore exclude the possibility that there exist additional chloroplast genes for ribosomal proteins which have no homology with any of the *E. coli* ribosomal proteins.

Acknowledgements: We thank Dr M. Sugita, Dr T. Matsubayashi and Mr Y. Li for their valuable suggestions. We thank Dr K. Kobayashi and Miss Y. Kato for protein sequencing. A.V. was supported by fellowships from the Ministry of Economy and Planning of Bulgaria and the Japan Society for the Promotion of Science. This

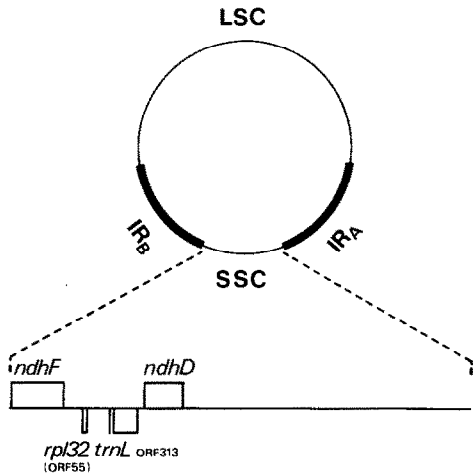

Fig. 2. Location of *rpl32* (ORF55) on the tobacco chloroplast genome. The lower part shows an enlarged map of the small-single copy region (SSC).

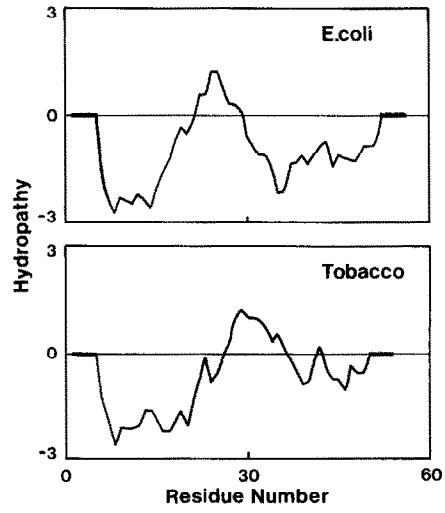

Fig. 3. Hydropathy plots of tobacco CL32 and *E. coli* L32 proteins. Hydropathy was calculated with a 11-point window according to Kyte and Doolittle [14].

```
tobacco     AVPKKRTSTSKKRIRKNIWKRKGYSIALKAFSLAKSLSTGNSKSFFVRQTKINK                ( 54 )
rice        --------M--------L--K-T-FSIVQSY-----R-FSGVSEHPKPKGFSRQQTNNRVLG        ( 62 )
broad bean  P-------I---K----F--K---KA--------D-IL--T--VIVL                      ( 47 )
liverwort   --------K--T----A---N-ANKS--R-------IL-NR----YYTINDKLLNSSKSISTSKLDES ( 68 )
```

Fig. 4. Comparison of the amino acid sequence of tobacco CL32 with those predicted from rice ORF63, broad bean ORF48 and liverwort ORF69. Dashes indicate identical residues with those of tobacco. Parentheses show residue number (the initial methionines are omitted).

work was supported by a Grant-in-Aid from the Minsitry of Eduction, Science and Culture.

REFERENCES

[1] Sugiura, M. (1989) Annu. Rev. Cell Biol. 5, 51–70.
[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] 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.
[4] Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Nature 322, 572–574.
[5] Ohto, C., Torazawa, K., Tanaka, M., Shinozaki, K. and Sugiura, M. (1988) Plant Mol. Biol. 11. 589–600.
[6] Thomas, F., Massenet, O., Dorne, A.M., Briat, J.F. and Mache, R. (1989) Nucleic Acids Res. 16, 2461–2472.
[7] Zhou, D.-X., Quigley, F., Massenet, O. and Mache, R. (1989) Mol. Gen. Genet. 216, 439–445.
[8] Markmann-Mulisch, U., von Knoblauch, K., Lehmann, A. and Subramanian, A.R. (1987) Biochem. Int. 15, 1057–1067.
[9] Capel, M.S. and Bourque, D.P. (1982) J. Biol. Chem. 257, 7746–7755.
[10] Wittmann, L.B., Greuer, B. and Pannenbecker, R. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1977–1979.
[11] Tanaka, I., Kimura, M., Kimura, J. and Dijk, J. (1984) FEBS Lett. 166, 343–346.
[12] Herdenberger, F., Weil, J.H. and Steinmetz, A. (1988) Curr. Genet. 14, 609–615.
[13] Reinbolt, J., Tritsch, D. and Wittmann, L.B. (1979) Biochimie 61, 501–522.
[14] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.