Nucleotide sequences of tobacco chloroplast genes for elongator tRNA\textsuperscript{Met} and tRNA\textsuperscript{Val} (UAC): the tRNA\textsuperscript{Val} (UAC) gene contains a long intron

Hiroshi Deno, Akira Kato, Kazuo Shinozaki and Masahiro Sugiura*  
National Institute of Genetics, Mishima, Shizuoka-ken 411, Japan  
Received 20 October 1982; Accepted 11 November 1982

ABSTRACT  
The nucleotide sequences of tobacco chloroplast genes for elongator tRNA\textsuperscript{Met} and tRNA\textsuperscript{Val} (UAC) have been determined. The tRNA\textsuperscript{Val} gene contains a 571 base pairs intron located in the anticodon loop. The tRNA\textsuperscript{Val} gene is transcribed as a 750 bases precursor RNA molecule. Both tRNAs deduced from the DNA sequences show 97% sequence homologies with those of spinach chloroplasts.

INTRODUCTION  
Chloroplasts contain their own tRNAs which are believed to be coded in chloroplast genomes (1). Chloroplast tRNA genes so far analyzed have both prokaryotic and eukaryotic features. Sequences similar to E. coli "Pribnow box" and "-35 region" are found in regions upstream from chloroplast tRNA genes (2-6) and the sequences of chloroplast tRNAs show higher homologies with those of bacterial tRNAs than those of eukaryotic cytoplasmic tRNAs (7). On the other hand, the 3' terminal CCA sequences of chloroplast tRNAs are not encoded by chloroplast DNA (2-6, 8-13) and introns have been found in some chloroplast tRNA genes (8-10). However, the introns of chloroplast tRNA genes are very long (458 to 949 bp), compared with those of eukaryotic nuclear tRNA genes (13 to 60 bp).

Spinach chloroplast tRNA\textsuperscript{Met} and tRNA\textsuperscript{Val} genes are reported to be located in the vicinity of the genes for \(\beta\) and \(\varepsilon\) subunits of ATPase and the large subunit (LS) of ribulose-1,5-bisphosphate carboxylase (1, 14). Recently we determined the nucleotide sequences of the LS (15, 16), \(\beta\) and \(\varepsilon\) genes (17) in tobacco chloroplast DNA. In this paper, we report the nucleotide sequences of tobacco chloroplast tRNA\textsuperscript{Met} and tRNA\textsuperscript{Val} genes in
the region downstream from the e gene and the presence of a long intron in the tRNA$^{Val}$ gene.

MATERIALS AND METHODS

Preparation of DNA fragment and DNA sequencing

Recombinant plasmid pTB17 containing a 1.6 Md BamHI fragment of Nicotiana tabacum (var. Bright Yellow 4) chloroplast DNA had been constructed (17). The plasmid DNA was digested with BamHI and the 1.6 Md fragment was separated from vector pBR322 by electrophoresis in a 1% agarose gel. DNA sequence was determined by the method of Maxam and Gilbert (18).

SI nuclease mapping

The 5' ends of a 190 bp HinfI fragment from the 0.46 Md Clal-BamHI fragment was labeled with $^{32}$P. After strand separation by 5% polyacrylamide gel electrophoresis, the $^{32}$P-labeled coding strand was hybridized with total tobacco chloroplast tRNAs (0.5 µg) at 58°C for 4 hr in 20 µl of 60% formamide, 0.04 M PIPES-NaOH (pH 6.8), 0.5 M NaCl and 1 mM EDTA (19). The hybridization mixture was diluted with 280 µl of 30 mM Na-acetate (pH 4.6), 1 mM ZnSO$_4$, 0.25 M NaCl and 20 µg/ml denatured calf thymus DNA, and digested with 4.5 units of SI nuclease (Sigma) at 37°C for 1 hr. The protected fragment was electrophoresed in a 12% polyacrylamide gel containing 7 M urea in parallel with the sequencing ladder of the coding strand.

Northern blot hybridization

Total tobacco chloroplast RNA was prepared as described (15). RNA was incubated in 1 M glyoxal, 50% dimethyl sulfoxide and 10 mM Na-phosphate (pH 7.0) at 50°C for 1 hr and electrophoresed in a 1.2% agarose gel containing 1 mM Na-phosphate. After electrophoresis, RNA was transferred from the gel to nitrocellulose filter (Schleicher & Schuell, BA85) and immobilized by the method of Thomas (20). Filter-immobilized RNA was hybridized with the $^{32}$P-labeled 337 bp Clal-HinfI fragment from the 0.46 Md Clal-BamHI fragment.

RESULTS AND DISCUSSION

DNA sequence

We had cloned a 1.6 Md BamHI fragment of tobacco chloro-
plast DNA which codes for the 3' part of the β subunit and the entire ε subunit of ATPase (17). On digestion with Clal, the 1.6 Md BamHI fragment isolated from recombinant plasmid pTB17 yields 0.46, 0.74 and 0.4 Md sub-fragments (Fig. 1). Total tobacco chloroplast tRNA hybridized to the 0.46 Md Clal-BamHI and 0.74 Md Clal-Clal sub-fragments (data not shown). We then sequenced the DNA region downstream from the ε gene by the sequence strategy shown in Fig. 1. Fig. 2 shows the nucleotide sequence of the left-hand 1106 bp (DdeI-BamHI) of the 1.6 Md BamHI fragment.

A tRNA gene was found between positions 2419-2491 (Fig. 2) and 224 bp apart from the 3' end of ε coding region in an opposite orientation in the 0.74 Md Clal-Clal sub-fragment.

**Fig. 1** Physical map of the cloned 1.6 Md BamHI fragment from tobacco chloroplast DNA and the strategy for sequencing part of it. Coding regions are shown by thick lines and an intron is shown by a box. The lower part shows the sequencing strategy. Horizontal arrows indicate the direction and extent of DNA segments sequenced.
Fig. 2 DNA sequence of the region containing tRNA<sup>Met</sup> and tRNA<sup>Val</sup> genes. Possible "Pribnow boxes" and "-35 regions", and palindromic sequences are underlined. Boxed sequences (2419-2491; 2682-2717 and 3289-3325) represent tRNA<sup>Met</sup> and tRNA<sup>Val</sup> sequences, respectively. Boxed ATG, TGA and TAA indicate the initiation and stop codons of ORFs. Numbers are bp relative to the 5' end of β gene.
(Fig. 1). Judging by the anticodon sequence, it is a tRNA\textsuperscript{Met} gene (Fig. 3a). It should be an elongator tRNA\textsuperscript{Met} gene because it shows significantly higher sequence homology with the elongator tRNA\textsuperscript{Met} (97%) than with the initiator tRNA\textsuperscript{Met} (51%) of spinach chloroplasts (21, 22). The sequence of tobacco chloroplast elongator tRNA\textsuperscript{Met} deduced from the DNA sequence shows 71% homology with that of E. coli and below 54% homology with those of yeast, mammals and rat mitochondria (7).

tRNA\textsuperscript{Val} gene

A sequence near identical to the 5' part of spinach chloroplast tRNA\textsuperscript{Val} (UAC) (23) was found between positions 2682-2719 in an opposite orientation with the tRNA\textsuperscript{Met}. However, the sequence beyond the position 2720 was not complementary to the anticodon stem sequence (ACTCG). We then searched for a sequence (CGAGT) complementary to the anticodon stem and a GTTC sequence which is

![Diagram of tobacco chloroplast tRNA\textsuperscript{Met}](image)

Fig. 3 (a) Cloverleaf structure of tobacco chloroplast tRNA\textsuperscript{Met} predicted from the DNA sequence. (b) Possible secondary structure of a putative transcript containing a tRNA\textsuperscript{Met} sequence. Arrows indicate possible processing sites.
Fig. 4 Determination of the 5' end of 3'-half exon for the tRNA^Val by SI mapping. (a) Schematic presentation of the procedure. Thick line indicates a coding region. A triangle indicates the 5' end labeled with $^{32}$P. (b) The SI protected fragment ($S_1$) was electrophoresed in parallel with a sequence ladder of the original strand (G, A>C, T+C, C). The top band in line $S_1$ represents a incomplete digest.

indicative of the GT'C sequence. We could find the sequence between positions 3291-3325 which was able to form a cloverleaf structure with the above sequence (positions 2682-2719). This sequence (positions 3291-3325) is near identical to the 3'-half of spinach tRNA$^{Val}$ (UAC). Based on the sequence homology (97%), they are a tRNA$^{Val}$ (UAC) gene. Because the two sequences are
The location of intron in the tRNA gene was determined by...
SI mapping. The \([5'-^{32}P]\) coding strand (positions 3123-3312) of a 190 bp HinfI fragment which contains a portion of the 3'-half exon was annealed with total tobacco chloroplast tRNAs and treated with SI nuclease. The major SI-resistant DNA fragments were determined to be 5'-ACCG--- and 5'-CCG--- (Fig. 4). This indicates that the intron is located at position either 36/37 or 37/38 of the tRNA (Fig. 5a). The sequence of tRNA\textsubscript{Val}(UAC) is unchanged, even if the intron site varies between positions 36 and 39. The precise location of the intron needs the analysis of the ends of a RNA fragment excised out.

Open reading frames (ORFs) have been reported in the introns of maize chloroplast tRNA\textsubscript{Ile}, tRNA\textsubscript{Ala} and tRNA\textsubscript{Leu} genes (8, 10) and tobacco chloroplast tRNA\textsubscript{Ile} and tRNA\textsubscript{Ala} genes (9). No such ORFs could be detected in the case of tobacco chloroplast tRNA\textsubscript{Val}(UAC) gene. However, an ORF (positions 2880-2689, 64 codons) which starts in the intron and ends in the 5'-half exon was found on the tRNA coding strand. On the opposite strand, another ORF (positions 2612-2797, 62 codons) which starts in the region upstream from the tRNA\textsubscript{Val}(UAC) gene and ends within the intron was also found. It is not clear whether these ORFs are expressed \textit{in vivo}.

There is no complementary sequence between both ends of the intron in the tRNA\textsubscript{Val}(UAC) gene. When the introns in tobacco chloroplast tRNA\textsubscript{Val}(UAC), tRNA\textsubscript{Ile} and tRNA\textsubscript{Ala} genes were compared, common sequences GNGYCRNTG near the 5' ends and TGNTTTA near the 3' ends were found. These sequences may play a role for the recognition of splicing sites.

Expression of tRNA\textsubscript{Met} gene and tRNA\textsubscript{Val}(UAC) gene

Hybridization experiments using total tobacco chloroplast tRNAs indicated that these tRNA genes were expressed \textit{in vivo} (data not shown). Two "Pribnow box"-like sequences TATGATA (2553-2547) and TATAATC (2544-2538) and two "-35 region"-like sequences TAGAT (2578-2574) and TTG (2568-2566) were found in the region upstream from the tRNA\textsubscript{Met} gene. Likewise, in the region upstream from the tRNA\textsubscript{Val}(UAC) gene, two "Pribnow boxes" TATCTTG (2632-2638) and TAAATA (2659-2665) and two "-35 regions" GTTGA (2608-2612) and TTGACA (2636-2641) were found. Possible secondary structures of putative transcripts contain-
ing tRNA$^{Met}$ and tRNA$^{Val}(UAC)$ sequences can be constructed as shown in Fig. 3b and Fig. 5b.

To detect a precursor RNA molecule for the tRNA$^{Val}(UAC)$, the following hybridization experiment was carried out. Total tobacco chloroplast RNA extracted from young tobacco leaves was loaded on a 1.2% agarose gel. After electrophoresis, the RNA was transferred to nitrocellulose filter and hybridized with $^{32}$P-labeled 337 bp ClaI-HinfI fragment (positions 2783-3119) containing a part of the intron. The DNA fragment hybridized to an RNA band of about 750 bases (Fig. 6). Judging by its size, this RNA should be a precursor RNA for the tRNA$^{Val}(UAC)$. From the codon usage of tobacco $\beta$ and $\epsilon$ genes (16, 17), the tRNA$^{Val}(UAC)$ is one of the frequently used tRNAs. Therefore it is suggested that the presence of a large intron does not interfere effective expression of the tRNA$^{Val}(UAC)$ gene.

ACKNOWLEDGEMENTS

We thank Mrs. Mie Kusuda and Miss Chieko Yamada for technical assistance. This work was supported in part by Grants-in-Aid from the Ministry of Education, Science and Culture.
REFERENCES