Volume 13 Number 4 1985 Nucleic Acids Research

Structure and cotranscription of tobacco chloroplast genes for $tRNA^{Glu}$(UUC), $tRNA^{Tyr}$(GUA) and $tRNA^{Asp}$(GUC)

Masaru Ohme, Takashi Kamogashira*, Kazuo Shinozaki and Masahiro Sugiura**

Department of Biology, Faculty of Science, Nagoya University, Chikusa, Nagoya 464, Japan

Received 3 January 1985; Accepted 25 January 1985

ABSTRACT

The location and nucleotide sequences of tobacco chloroplast genes for $tRNA^{Glu}$(UUC), $tRNA^{Tyr}$(GUA) and $tRNA^{Asp}$(GUC) have been determined. These genes lie midway between the genes for α and β/ε subunits of H^+-ATPase on the large single-copy region of the chloroplast DNA. The gene organization is $tRNA^{Glu}$ - 59bp spacer - $tRNA^{Tyr}$ - 108bp spacer - $tRNA^{Asp}$ on the same DNA strand. Northern blot hybridization studies revealed that these three tRNA genes are cotranscribed. The transcription initiation site was localized at 24 bp upstream from the $tRNA^{Glu}$ coding region and its termination site at 90 bp downstream from the $tRNA^{Asp}$ coding region by S1 mapping. The tricistronic tRNA precursor is thus calculated to be 512 bases long. Its processing was also studied by S1 mapping.

INTRODUCTION

Chloroplasts contain their own tRNAs which are believed to be coded for by their chloroplast genomes (1). Chloroplast tRNAs show high sequence homology with prokaryotic tRNAs (2), and sequences similar to E. coli "Pribnow box" and "-35 region" are found in the upstream regions of chloroplast tRNA genes. On the other hand, chloroplast tRNA genes do not code for the 3' CCA ends and some of them contain long introns (3). Like prokaryotic tRNA genes, Euglena chloroplast tRNA genes tend to be clustered (4). In contrast, higher plant chloroplast tRNA genes are not usually clustered except for maize and tobacco $tRNA^{Ile}$ and $tRNA^{Ala}$ genes located in the spacer region of 16S-23S rRNA genes (5,6) and broad bean $tRNA^{Glu}$ and $tRNA^{Tyr}$ genes (7).

Here we present the nucleotide sequences of tobacco chloroplast genes for $tRNA^{Glu}$(UUC), $tRNA^{Tyr}$(GUA) and $tRNA^{Asp}$(GUC) which are clustered within a 400 bp region. We then show that these three genes are cotranscribed. Processing of the tRNA precursor is also discussed.

MATERIALS AND METHODS

The recombinant plasmid pTP7 which contains a 6.6 kbp PstI fragment of

Nicotiana tabacum (var. Bright Yellow 4) chloroplast DNA has been constructed as described using pBR322 (8). pTP7 DNA was digested with PstI and the 6.6 kbp fragment was separated by electrophoresis in a 5% polyacrylamide gel. DNA sequence was determined by the method of Maxam and Gilbert (9).

Total tobacco chloroplast RNA was prepared from chloramphenicol-treated leaves as described (6). For Northern blot hybridization, the chloroplast RNA was denatured and electrophoresed as described (10) and transferred to a nylon membrane (Biodyne A, Pall). The membrane was immersed for 8 hr at 42°C in the hybridization buffer (50% formamide, 0.9 M NaCl, 50 mM sodium phosphate buffer (pH 7.7), 5 mM EDTA, 0.3% SDS, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin and 250 µg/ml denatured salmon sperm DNA). Hybridization was carried out at 42°C for 20 hr in the hybridization buffer containing a ^{32}P-labeled DNA prove.

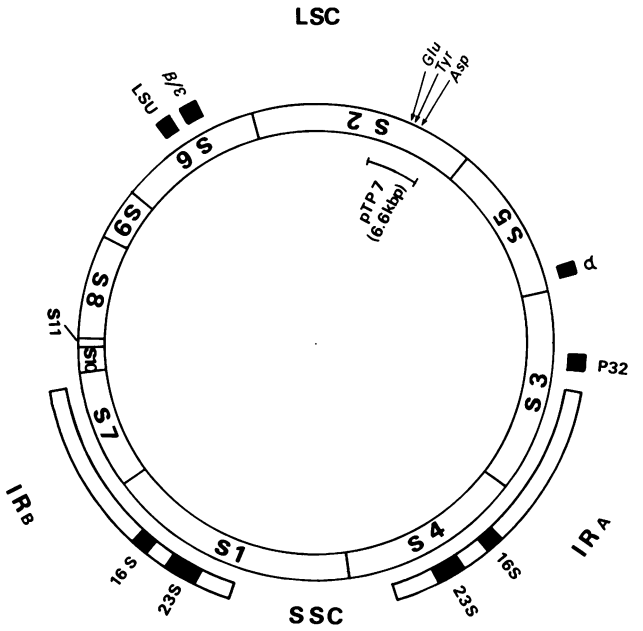


Fig. 1. Position of the cloned fragment (in pTP7) and the genes for $tRNA^{Glu}$, $tRNA^{Tyr}$ and $tRNA^{Asp}$ on the SalI cleavage map of tobacco chloroplast DNA (12,21). Arrows, tRNA genes; IR_A and IR_B, inverted repeated sequences (22); LSC and SSC, large and small single-copy regions; α, β/ε, H^+-ATPase subunit genes; LSU, ribulose bisphosphate carboxylase large subunit gene; P32, 32,000 dalton thylakoid membrane protein gene.

S1 mapping studies were carried out as described (11), except for the hybridization temperature; 57°C for the 124 base TaqI fragment, 56°C for the 212 base TaqI fragment, 52°C for 746 bp AvaII-BamHI fragment and 56°C for the 587 bp EcoRI-HhaI fragment.

RESULTS AND DISCUSSINON

Structure

The 6.6 kbp PstI fragment of tobacco chloroplast DNA is located within the SalI fragment 2 (S2) on the chloroplast DNA map (12) shown in Fig. 1. On digestion with EcoRI + EcoRV, the 6.6 kbp fragment yields 3.7, 0.5, 0.6 and 1.8 kbp sub-fragments in this order (Fig. 2). The 0.5 kbp EcoRV-EcoRI and 0.6 kbp EcoRI sub-fragments were hybridized with total tobacco chloroplast tRNA (data not shown). We then sequenced the region containing the 0.5 and 0.6 kbp sub-fragments by the sequence strategy shown in Fig. 2. Figure 3 shows the nucleotide sequence of the 1072 bp portion (from the EcoRV site to the second EcoRI site) of the 6.6 kbp PstI fragment.

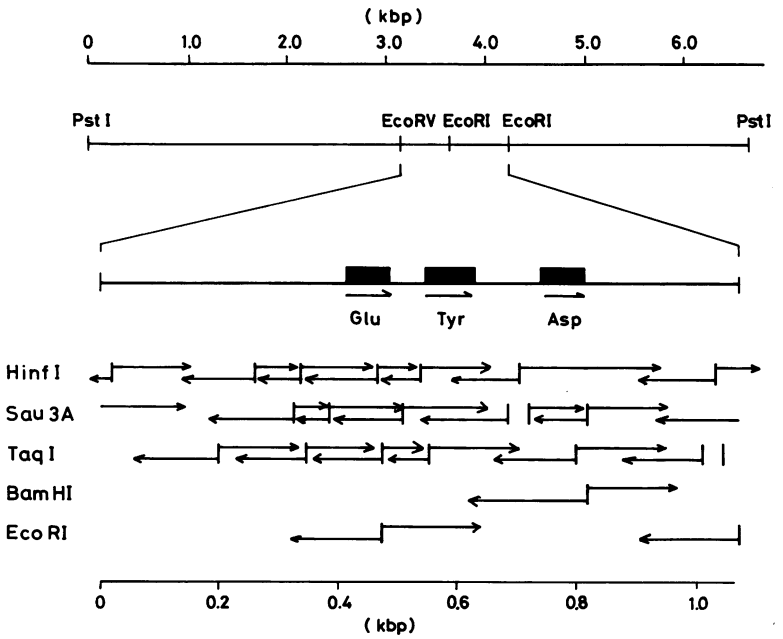
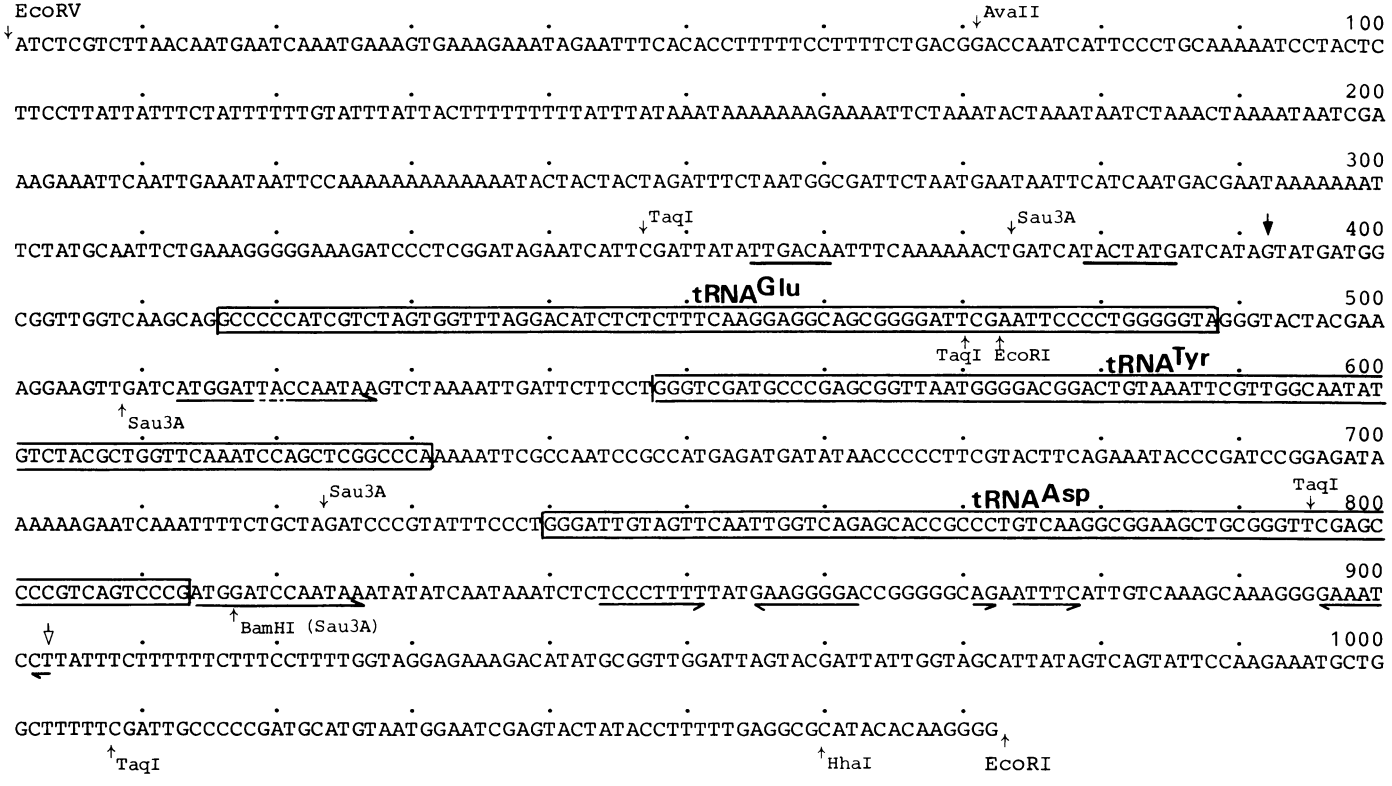


Fig. 2. Physical map of the cloned 6.6 kbp PstI fragment from tobacco chloroplast DNA and the strategy for sequencing part of it. Coding regions are shown by thick lines and arrows (⟶) below them indicate the direction of transcription.


EcoRV
AvaII
100
ATCTCGTCTTAACAATGAATCAAATGAAAGTGAAAGAAATAGAATTTCACACCTTTTTCCTTTTCTGACGGACCAATCATTCCCTGCAAAAATCCTACTC
200
TTCCTTATTATTTCTATTTTTTGTATTTATTACTTTTTTTTTTATTTATAAATAAAAAAAGAAAATTCTAAATACTAAATAATCTAAACTAAAATAATCGA
300
AAGAAATTCAATTGAAATAATTCCAAAAAAAAAAAAAATACTACTACTAGATTTCTAATGGCGATTCTAATGAATAATTCATCAATGACGAATAAAAAAAT
TaqI
Sau3A
400
TCTATGCAATTCTGAAAGGGGGAAAGATCCCTCGGATAGAATCATTCGATTATATTGACAATTTCAAAAAACTGATCATACTATGATCATAGTATGATGG
tRNAGlu
500
CGGTTGGTCAAGCAGGCCCCCATCGTCTAGTGGTTTAGGACATCTCTCTTTCAAGGAGGCAGCGGGGATTCGAATTCCCCTGGGGGTAGGGTACTACGAA
TaqI
EcoRI
tRNATyr
600
AGGAAGTTGATCATGGATTACCAATAAGTCTAAAATTGATTCTTCCTGGGTCGATGCCCGAGCGGTTAATGGGGACGGACTGTAAATTCGTTGGCAATAT
Sau3A
700
GTCTACGCTGGTTCAAATCCAGCTCGGCCCAAAAATTCGCCAATCCGCCATGAGATGATATAACCCCCTTCGTACTTCAGAAATACCCGATCCGGAGATA
Sau3A
tRNAAsp
TaqI
800
AAAAAGAATCAAATTTTCTGCTAGATCCCGTATTTCCCTGGGATTGTAGTTCAATTGGTCAGAGCACCGCCCTGTCAAGGCGGAAGCTGCGGGTTCGAGC
900
CCCGTCAGTCCCGATGGATCCAATAAATATATCAATAAATCTCTCCCTTTTTATGAAGGGGACCGGGGGCAGAATTTCATTGTCAAAGCAAAGGGGAAAT
BamHI (Sau3A)
1000
CCTTATTTCTTTTTTCTTTCCTTTTGGTAGGAGAAAGACATATGCGGTTGGATTAGTACGATTATTGGTAGCATTATAGTCAGTATTCCAAGAAATGCTG
GCTTTTTCGATTGCCCCCGATGCATGTAATGGAATCGAGTACTATACCTTTTTGAGGCGCATACACAAGGGG
TaqI
HhaI
EcoRI

The genes for $tRNA^{Glu}$(UUC), $tRNA^{Tyr}$(GUA) and $tRNA^{Asp}$(GUC) were found between positions 416-488, 548-631 and 740-813, respectively (Figs. 3 & 4). The $tRNA^{Glu}$ deduced from the DNA sequence shows 72, 92, 93 and 97% sequence homologies with E. coli (2), Euglena chloroplast (13), spinach chloroplast (14) and broad bean chloroplast (7) $tRNA^{Glu}$(UUC)s, respectively. The $tRNA^{Tyr}$ shows 61 and 74% sequence homologies with E. coli (2) and Euglena chloroplast (13) $tRNA^{Tyr}$(GUA)s and is identical with spinach chloroplast (14) and broad bean chloroplast (7) $tRNA^{Tyr}$(GUA)s, respectively. The $tRNA^{Asp}$ shows 73 and 99% sequence homologies with E. coli (2) and spinach chloroplast (15) $tRNA^{Asp}$(GUC)s, respectively. The 3'-CCA termini were not found in these tRNA genes and none of them contains introns.

One unusual feature is that the $tRNA^{Glu}$ has an A_{53}-U_{61} pair rather than an invariant G_{53}-C_{61} pair at the TψC stem (Fig. 4a). Such an A-U pair in the TψC stem has been reported in Euglena chloroplast (13), spinach chloroplast (14) and broad bean chloroplast (7) $tRNA^{Glu}$(UUC)s.

Based on the fine physical mapping of the PstI fragment and its neighboring regions (data not shown), these three tRNAs are coded for by the same strand as the α and β/ε subunits of H^+-ATPase (16,17, strand B) of tobacco chloroplast DNA (Fig. 1). The location of $tRNA^{Tyr}$ and $tRNA^{Asp}$ genes are consistent with the tRNA gene map obtained by the tRNA/DNA hybridization studies (18). Although $tRNA^{Glu}$ genes have not been mapped by the hybridization studies (18), the $tRNA^{Glu}$(UUC) gene was found by our DNA sequencing.

Transcription

As the spacers between $tRNA^{Glu}$ and $tRNA^{Tyr}$ genes, and $tRNA^{Tyr}$ and $tRNA^{Asp}$ genes are only 59 and 108 bp long, respectively, it can be postulated that these three genes are cotranscribed. To examine this, Northern blot hybridization studies were carried out. Generally, detection of a primary transcript for tRNA is not easy because of its rapid processing. This sometimes makes it difficult to demonstrate whether tRNA genes are cotranscribed or not. We then prepared total tobacco chloroplast RNA from chloramphenicol-treated leaves for possible accumulation of precursor RNAs (6). The RNA was immobilized in a nylon membrane and

Fig. 3. Nucleotide sequence of the region containing $tRNA^{Glu}$, $tRNA^{Tyr}$ and $tRNA^{Asp}$ genes. Coding regions are boxed. "Pribnow box"-like and "-35 region"-like sequences are underlined. Large arrows indicate the transcription initiation (↓) and termination (↓) sites, small arrows cleavage sites and horizontal arrows direct repeated sequences (⟶) and inverted repeated sequences (⟶).

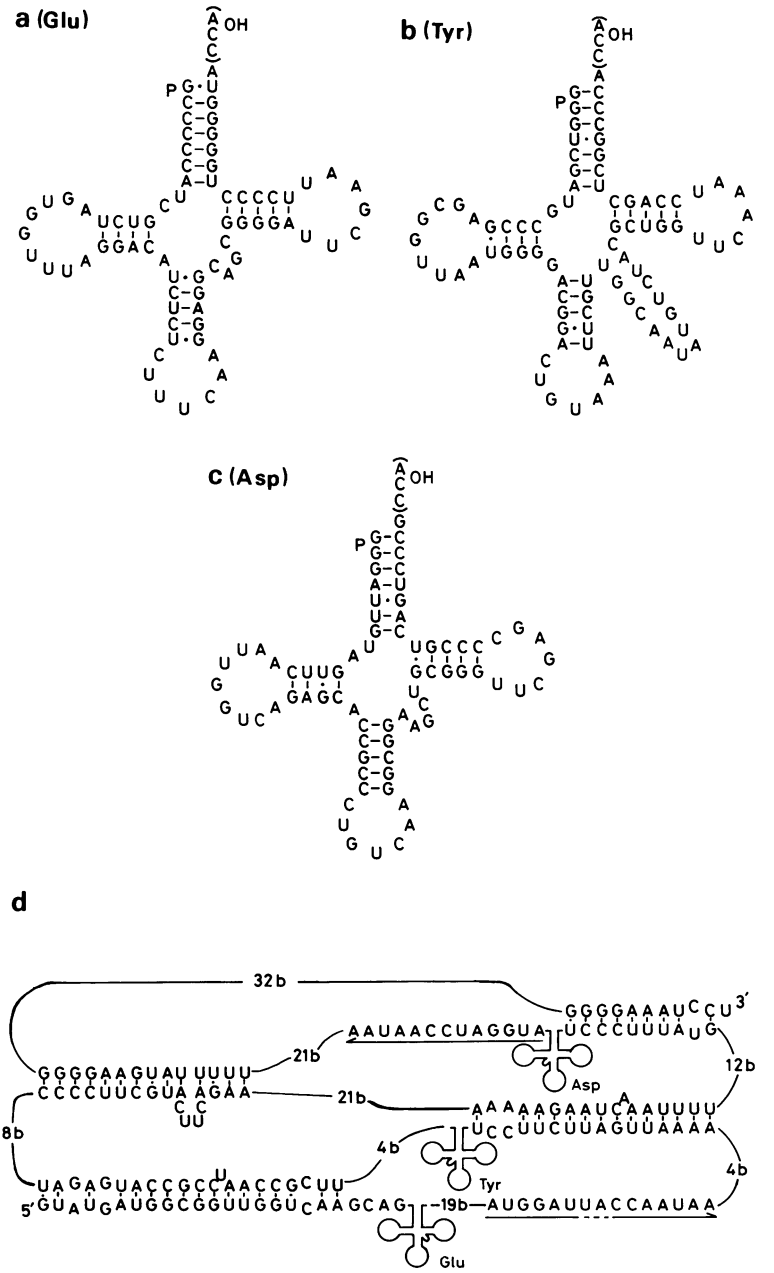


Fig. 4. Sequences of unmodified bases and cloverleaf structures of tobacco chloroplast $tRNA^{Glu}$ (a), $tRNA^{Tyr}$ (b) and $tRNA^{Asp}$ (c) predicted from the DNA sequences. (d) Possible secondary structure of the tRNA precursor. Arrows show direct repeated sequences.

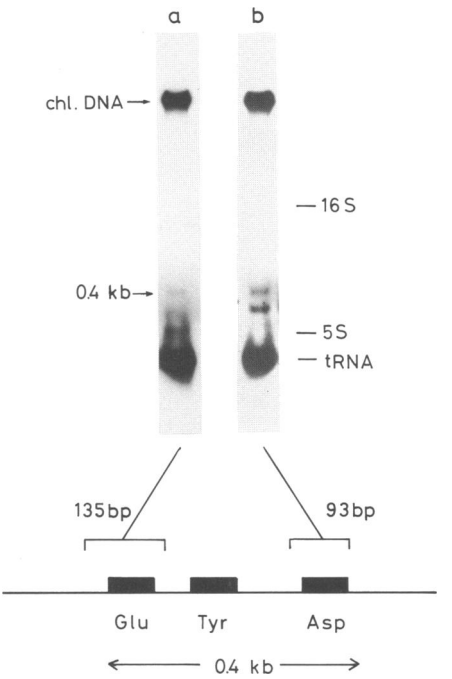


Fig. 5. Detection of tRNA precursors. Autoradiographs of Northern blots of tobacco chloroplast RNA hybridized to ^{32}P-labeled 135 bp Sau3A fragment (lane a) and 93 bp Sau3A fragment (lane b). Size markers are E. coli 16S RNA, tobacco chloroplast 5S RNA and E. coli $tRNA^{Val}$.

hybridized with either the 135 bp Sau3A fragment (374-508) containing the $tRNA^{Glu}$ gene or the 93 bp Sau3A fragment (724-816) containing $tRNA^{Asp}$ gene. Each DNA fragment hybridized to an about 0.4 kb RNA band and several other smaller RNA bands (Fig. 5). The 0.4 kb RNA is just large enough for these three tRNA genes, indicating that the three tRNA genes are transcribed as a single unit. To our knowledge, this is the first case to detect a polycistronic tRNA precursor in chloroplasts. Kuntz et al. have postulated that $tRNA^{Glu}$ and $tRNA^{Tyr}$ genes from broad bean chloroplasts may constitute an operon by their sequence analysis (7).

The position corresponding to the 5' end of the precursor RNA was determined by S1 mapping as shown in Fig. 6a. The [5' ^{32}P]-coding strand of the 124 bp TaqI fragment (347-470) that contains most of the $tRNA^{Glu}$ gene and its upstream region was hybridized with total tobacco chloroplast RNA and digested with S1 nuclease. The S1-resistant DNA fragments were

a)
TaqI 347
124 b
tRNAGlu
TaqI 470
Precursor RNA
S1 nuclease
5' G T A T G
b)
G A>C T±C c S1 a S1 b
3' 5'
5' (-24)
c)
TaqI 796
tRNAAsp
212 b
TaqI 1007
Precursor RNA
S1 nuclease
A T C C T 3'
d)
G A>C T±C c S1 a S1 b
5' 3'
3' (+90)

electrophoresed in parallel with the sequencing ladder generated from the coding strand (Fig. 6b). A correction of 1.5 nucleotides for the position of the 5' end was applied. The 5' end of the precursor RNA was mapped at position 392 (G), 24 bp upstream from the $tRNA^{Glu}$ coding region (Fig. 3). A "Pribnow box"-like sequence TACTATG (379-385) and a "-35 region"-like sequence TTGACA (355-360) were found 7 bp and 32 bp upstream from the initiation site, respectively.

The position corresponding to the 3' end of the precursor RNA was also determined as shown in Fig. 6c,d. The [3' ^{32}P]-coding strand of the 212 bp TaqI fragment (796-1007) that contains the 3'-terminus of the $tRNA^{Asp}$ and its downstream region was used. A correction of 0.5 nucleotide was made for the position of the 3' end. The 3' end of the precursor RNA was mapped at position 903 (T), 90 bp downstream from the $tRNA^{Asp}$ coding region (Fig. 3). In the region around the termination site, inverted repeated sequences (844-851 and 855-862; 871-878 and 896-903) followed by a T-rich sequence (903-925) were found. These sequences resemble rho-factor independent transcriptional termination sites of prokaryotes. The precursor RNA for these three genes is thus calculated to be 512 nucleotides long (392-903). In chloroplast tRNA genes, to our knowledge, this is the first report to determine both 5' and 3' ends of a precursor RNA.

Processing

To study processing of the tricistronic tRNA precursor, the [5' ^{32}P BamHI site]-coding strand of the 746 bp AvaII-BamHI fragment (71-816), that contains the three tRNA genes but lacks most of the 3'-extension of the transcribed region, was annealed to total tobacco chloroplast RNA and digested with S1 nuclease (Fig. 7a). The S1-resistant DNA fragments were loaded on a sequencing gel with HaeIII-digested ΦX174 RF-DNA fragments and a 405 bp DNA fragment as size markers. We could find two processing intermediates besides the 430 base primary transcript fragment (the 5' start site to the BamHI site)(Fig. 7b). One is a 405 base band which is likely produced by cleavage at the 5' end of $tRNA^{Glu}$ and another is a 275 base band produced by cleavage at the 5' end of $tRNA^{Tyr}$. No other bands were detected.

Fig. 6. Determination of the 5' and 3' ends of the tRNA precursor by S1 mapping. (a,c) Schematic presentation of the procedure. Filled boxes indicate tRNA coding regions and open boxes non-coding regions. Triangles indicate [5' ^{32}P]-ends (▽) and [3' ^{32}P]-ends (▽). (b,d) Autoradiographs of S1-protected fragments in parallel with sequencing ladders of the coding strands (G, A>C, T+C, C). 5 units (S1 a) and 15 units (S1 b) of S1 nuclease were used. An arrow in (b) indicates the 5' end of mature $tRNA^{Glu}$.

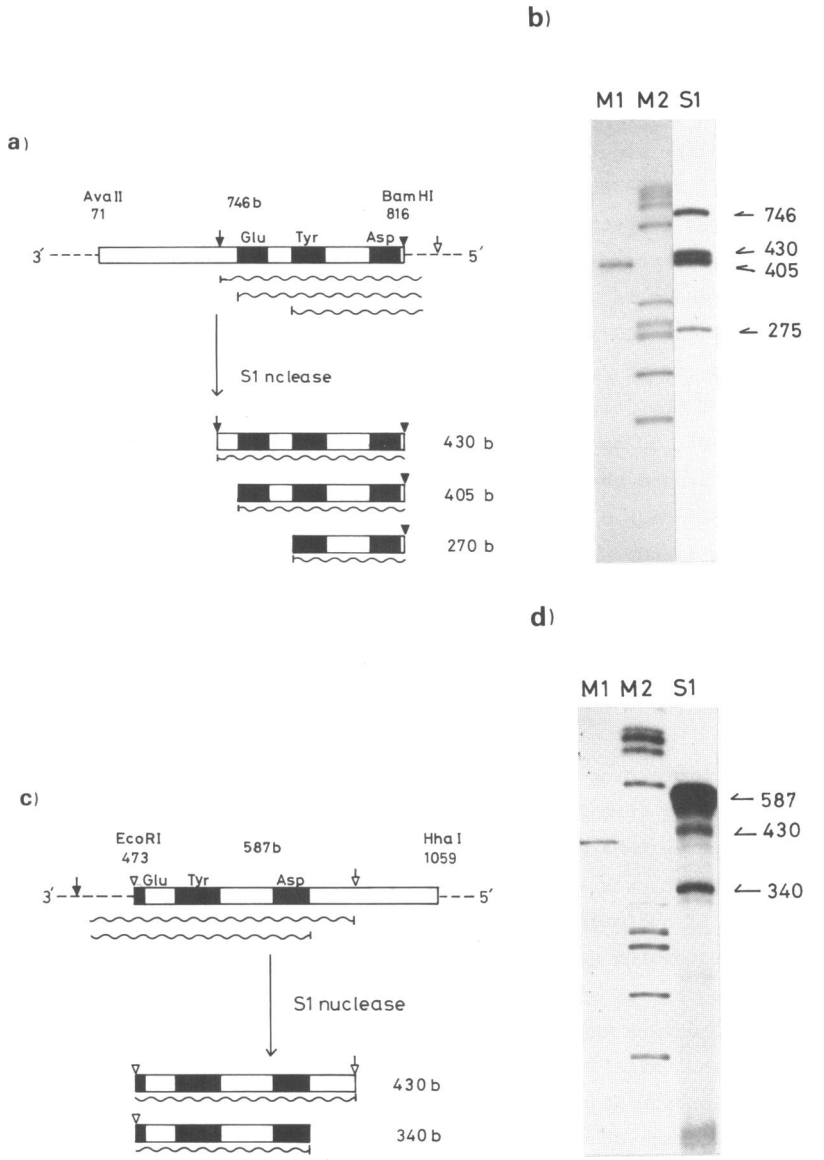


Fig. 7. Detection of processing intermediates by S1 mapping. (a,c) Schematic presentation of procedure. Filled boxes indicate tRNA coding regions and open boxes non-coding regions. Triangles indicate [5' ^{32}P]-ends (▼) and [3' ^{32}P]-ends (▽). Arrows indicate transcription initiation (↓) and termination (↓) sites. Wavy lines show the tRNA precursor and its processing intermediates. (b,d) Autoradiographs of S1-protected fragments in parallel with size markers, a 405 bp DNA fragment (M1) and ΦX174 RF-DNA HaeIII digest (M2).

Processing at the 3' end was also studied. The [3' ^{32}P EcoRI site]-coding strand of 587 bp EcoRI-HhaI fragment (473-1059) was subjected to S1 mapping as above (Fig. 7c). We could obtain a 430 base band representing the primary transcript fragment (the EcoRI site to the 3' termination site) and a 340 base band which is likely produced by cleavage at the 3' end of $tRNA^{Asp}$ (Fig. 7d). A ladder was also observed at the bottom of the S1 autoradiograph (data not shown).

The tRNA processing seems to be initiated by endonucleolytic cleavage at the 5' ends of $tRNA^{Glu}$ and $tRNA^{Tyr}$, because discrete RNA bands probably cleaved at the 5' end of $tRNA^{Glu}$ or $tRNA^{Tyr}$ were detected. The 3' extension of the precursor RNA seems to be processed endonucleolytically and prior to processing at the 5' end of $tRNA^{Asp}$ because we found a 340 base discrete band (Fig. 7d) and could not find a short RNA band corresponding only $tRNA^{Asp}$ (Fig. 7b). A ladder was observed in the 3' processing experiment, suggesting that the processing at the 3' sides of $tRNA^{Glu}$ and $tRNA^{Tyr}$ is exonucleolytic. Endo- and exonucleolytic cleavages have been reported in *in vitro* processing studies in *Euglena* chloroplast tRNA gene clusters (19,20). Figure 4d shows a possible secondary structure of the tRNA precursor. Interestingly, stem structures could be formed between the 5' upstream region of each tRNA sequence and a distant non-coding region. The processing enzymes may recognize these structures. Direct repeated sequences (513-527 and 814-826) were found downstream from the $tRNA^{Glu}$ and $tRNA^{Asp}$ regions (Fig. 3 & 4d).

ACKNOWLEDGEMENT

This work was supported in part by Grants-in-Aid from the Ministry of Education, Science and Culture, and from the Toray Science Foundation.

*Present address, Tokushima Institute, Otuka Pharmaceutical Co., Ltd., Tokushima 771-01, Japan

** To whom correspondence should be addressed.

REFERENCES

1. Weil, J.H. and Parthier, B. (1982) in Encyclopedia of Plsnt Physiology, Boulter, D. and Parthier, B. Eds. 14A, 65-112, Springer, Heidlberg.
2. Sprinzl, M. and Gauss, D.H. (1984) Nucleic Acids Res. 12, r1-r132.
3. Sugiura, M., Shinozaki, K. and Ohme, M. (1985) in Molecular Form and Function of the Plant Genome, Vloten-Doting, L. et al. Eds., Plenum, New York, in press.
4. Hallick, R.B., Hollingsworth, M.J. and Nickoloff, J.A. (1984) Plant Mol. Biol. 3, 169-175.
5. Koch, W., Edwards, K. and Kössel, H. (1981) Cell 25, 203-213.

6. Takaiwa, F. and Sugiura, M. (1982) Nucleic Acids Res. 10, 2665-2676.
7. Kuntz, M., Weil J.H. and Steinmetz, A. (1984) Nucleic Acids Res. 12, 5037-5047.
8. Sugiura, M. and Kusuda, J. (1979) Mol. Gen. Genet. 172, 137-141.
9. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
10. Deno, H., Kato, A., Shinozaki, K. and Sugiura, M. (1982) Nucleic Acids Res. 10, 7511-7520.
11. Shinozaki, K. and Sugiura, M. (1982) Gene 20, 91-102.
12. Seyer, P., Kowallik, K.V. and Herrmann, R.G. (1981) Current Genet. 3, 189-204.
13. Hollingsworth, M.J. and Hallick, R.B. (1982) J. Biol. Chem. 257, 12795-12799.
14. Holschuh, K., Bottomley, W. and Whitfeld, P.R. (1984) Plant Mol. Biol. 3, 313-317.
15. Holschuh, K., Bottomley, W. and Whitfeld, P.R. (1983) Nucleic Acides Res. 11, 8547-8554.
16. Deno, H., Shinozaki, K. and Sugiura, M. (1983) Nucleic Acids Res. 11, 2185-2191.
17. Shinozaki, K., Deno, H., Kato, A. and Sugiura, M. (1983) Gene 24, 147-155.
18. Bergmann, P., Seyer, P., Burkard, G. and Weil, J.H. (1984) Plant Mol. Biol. 3, 29-36.
19. Gruissem, W., Prescott, D.M., Greenberg, B.M. and Hallick, R.B. (1982) Cell 30, 81-92.
20. Greenberg, B.M., Gruissem, W. and Hallick, R.B. (1984) Plant. Mol. Biol. 3, 97-109.
21. Ohme, M., Kamogashira, T., Shinozaki, K. and Sugiura, M. (1984) Nucleic Acids Res. 12, 6741-6749.
22. Sugita, M., Kato, A., Shimada, H. and Sugiura, M. (1984) Mol. Gen. Genet. 194, 200-205.