Identification of the transcription start site for the spinach chloroplast serine tRNA gene

Chih-Yu Wu, Chi-Hui Lin, Liang-Jwu Chen*

Institute of Molecular Biology, National Chung Hsing University, Taichung 40227, Taiwan, ROC

Received 2 September 1997; revised version received 20 October 1997

Abstract Deleting part of the 3′ end of the spinach chloroplast serine tRNA coding region, which destroyed the proper folding of its RNA transcript and resulted in the inhibition of tRNA processing, allowed the detection of a serine tRNA primary transcript. The transcription start site for this primary transcript, synthesized from the internal promoter, was mapped to −12 upstream from the mature tRNA coding region. Transcription analysis with various 5′ deletion mutants suggested that the AT-rich region between −31 and −11, immediately upstream of the serine tRNA transcription start site, affects the transcription efficiency, and possibly the selection of transcription start site. Identification of the transcription start site for the spinach chloroplast serine tRNA gene in this study represents the first example of 5′ end mapping of a tRNA precursor transcribed from chloroplast tRNA genes containing an internal promoter.

Key words: Internal promoter; Transcription start site; tRNA processing; Serine tRNA gene; Spinach chloroplast

1. Introduction

The 5′ end of mature tRNAs in the vast majority of cases is thought to be the result of 5′ cleavage at its precursor by a RNase P-like endonuclease [1]. This notion suggests that the transcription of precursor tRNA, from either poly- or monocistronic genes, initiates at a position upstream to the 5′ end of the mature tRNA. Prokaryotic tRNA genes, which have promoter elements upstream of their coding region, follow this model. However, exceptions can be found for eukaryotic tRNA genes, which contain internal promoters. While the transcription start sites of *Xenopus trnL* [2] and yeast *trnL3* [3] genes were mapped to positions upstream of the coding region, the start site of a *Xenopus* selenocysteine tRNA gene was mapped to the first nucleotide of the coding sequence [4,5].

In higher plants, chloroplast tRNAs are usually transcribed as monomeric precursors from their respective promoters [6–8] and subsequently processed post-transcriptionally [6,9–12] to form the mature tRNAs. The tRNA primary transcript in chloroplast possibly does not accumulate and, therefore, determination of the transcription start site for chloroplast tRNA genes using tRNA primary transcript has not been reported. Two types of promoters for chloroplast tRNA genes have been reported. One contains the prokaryote-like upstream '−10' and '−35' consensus sequences, e.g. spinach *trnM2* [13]; and the other, e.g. spinach *trnS*, *trnR* and *trnT*, requires no upstream promoter elements [7,12]. While the transcription start sites of the upstream promoter-containing tRNA genes were thought to be located 5′ to the coding region, similar to transcription start sites identified for chloroplast protein-coding genes that contain upstream '−10' and '−35' consensus promoter elements [14], the transcription start site for chloroplast genes that contain internal promoter elements was still unknown.

In this study, the spinach serine tRNA encoded by a monocistronic *trnS* gene [15] having a complete conserved intragenic element with the eukaryotic consensus A and B boxes promoter sequences ([2] and Fig. 1B) was used for transcription start site analysis. Primary transcripts of the chloroplast *trnS* gene were produced from a transcription system in which tRNA processing was inhibited, and the transcription start sites of these primary transcripts were mapped. In addition, transcription with various 5′ deletion mutants indicated that the sequences immediate 5′ of the transcription start point of *trnS* gene also played a role in determining the transcription efficiency.

2. Materials and methods

2.1. Construction of plasmid clones

Seven 5′ upstream sequence deletion mutants of the *trnS* gene, including pSU86 and pSU47, and two 3′ coding region deletion mutants, pSU86Δ and pSU47Δ, were constructed (Fig. 1). Plasmid pTZ19RTa, a derivative of pTZ19R [16], containing the terminator (Ta) from the threonine attenuator region [17], provided the framework for the construction of various clones. A 199-bp *Sspl*/*Xmn*I DNA fragment, consisting of a 86-bp 5′ non-coding sequence, the coding region (93 bp), and a 20-bp 3′ non-coding sequence of the *trnS* gene, was inserted into the *Hin*cII site of pTZ19RTa to get plasmid pSU86. Plasmid pSU47, containing 47 bp of the *trnS* 5′ non-coding sequence, is essentially the same as plasmid pSU86 except that the region between 86 and 47 were replaced with the vector sequence (Fig. 1B, underlined). The 3′ sequences in front of the terminator (Ta) are 150 and 140 bp for pSU86 and pSU47, respectively. Plasmids pSU86Δ and pSU47Δ were derived from pSU86 and pSU47 by deleting the 20-bp 3′ non-coding region and the last 8 bp of the 3′ mature tRNA coding region. The last eight nucleotides of the tRNA are presumably involved in base pairing with the 5′ end coding sequence to form the aminoacyl stem of serine tRNA (Fig. 1B, in bold and boxed). The above deletion was achieved by PCR using a 29-nt primer (5′-TGAGGATCCGAGGGATTCGAACCCTCGAT-3′), which contains a *Bam*HI site at the 5′ end and the sequences spanning +85 to +66 of the *trnS* coding region, to leave out the last eight nucleotides of the 3′ coding region and all of the 3′ non-coding region (Fig. 1B). After these steps of manipulation, the serine tRNA coding region of these two clones became 85 bp and the length from the 3′ end to the termination site was 66 bp (Fig. 1).

2.2. Preparation of the spinach chloroplast enzyme extract

Spinach chloroplast enzyme extracts, also called high salt extract (HSX), were prepared from Percoll gradient-purified spinach chloroplasts according to procedures described previously [18].

*Corresponding author. Fax: (886) (4) 287-4879.
E-mail: ljchen@dragon.nchu.edu.tw

PII S0014-5793(97)01373-2

2.3. In vitro transcription/processing assay

The in vitro transcription/processing assay was essentially as described previously [12]. Each reaction (40 μl) contained 12 mM HEPES-KOH (pH 8), 40 mM KCl, 10 mM $MgCl_2$, 10 mM DTT, 1 U/μl RNase inhibitor, 500 μM each ATP, GTP and CTP and 10 μM UTP containing 100 μCi [α-^{32}P]GTP (800 Ci/mmol, 10 μCi/μl; NEN DuPont) and the following components: one-tenth of the volume (4 μl) was supercoiled plasmid DNA (0.1 μg/μl) in 10 mM Tris-HCl (pH 8), 0.1 mM EDTA and two-fifths of the volume (16 μl) was spinach chloroplast enzyme extract (~1 μg/μl) in 'DEAE buffer' (50 mM Tricine-KOH (pH 8), 50 mM KCl, 0.5 mM EDTA, 0.5 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride, 1 mM benzamidine, 5 mM ε-amino-*n*-caproic acid, 5% (v/v) glycerol). The reaction was incubated at 30°C for 60 min, then the reaction mixture was extracted with RNA extraction buffer and precipitated. The nucleic acid pellet was resuspended in 80% formamide, 0.1% bromophenol blue and 0.1% xylene cyanol and analyzed by electrophoresis in 8% polyacrylamide-8.3 M urea DNA sequencing gels, and followed by autoradiography.

2.4. In vitro capping assay

Both the conventional capping assay using [^{32}P]GTP and a method developed by us, in which the transcript was body-labeled and then capped with cold GTP, were performed. However, the body-labeling method was found to give better results, and therefore the body-labeled transcripts from the standard in vitro transcription reaction with T7 RNA polymerase or spinach chloroplast enzyme extract [17] were used for the capping assay. Each capping reaction mixture (40 μl) contained a half amount of labeled transcripts (resuspended in 10 μl) from the transcription reaction, and 50 mM Tris-HCl (pH 8.0), 1.2 mM $MgCl_2$, 6 mM KCl, 2.5 mM DTT, 40 U RNase inhibitor, 50 μM GTP, and 4 U guanylyltransferase (BRL). After an incubation time of 30 min at 37°C, the reaction mixture was extracted and treated as described in Section 2.3.

2.5. Primer extension analysis

Primary tRNA transcripts synthesized in vitro were extracted from the reaction mixture and used for 5′ end mapping by primer extension. The 29-nt primer described above was 5′ end labeled with [γ-^{32}P]ATP by T4 polynucleotide kinase. Labeled primer (~10^5 cpm) was added to the tRNA primary transcripts in a mixture (30 μl) containing 50 mM Tris-HCl (pH 8.3), 60 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 1 mM each dNTP and 200 U SuperScript RNase H$^-$ reverse transcriptase (BRL). After incubating for 60 min at 37°C, the primer extension products were extracted with phenol, precipitated with ethanol and fractionated by electrophoresis in 8% polyacrylamide-8.3 M urea DNA sequencing gels. Sequencing reactions using the same primers were used as markers.

3. Results and discussion

3.1. The 5′ flanking sequence of trnS gene is required for efficient transcription

The spinach chloroplast *trnS* is a monocistronic gene and its transcription requires no upstream promoter elements [7]. Coupled transcription/processing reaction using templates pSU86 and pSU47 produced comparable amounts of tRNA+3′ trailer, 3′ trailer, and mature serine tRNA (Fig. 2A, lanes 2 and 4, Fig. 2B). Yet, reactions with mutants −19 and −11, i.e. mutants having 5′ upstream sequences deleted to −19 and −11, respectively, exhibited a transcription activity about 15% that of pSU86 (Fig. 2B). Moreover, the transcription activity of the +1 mutant which contained no 5′ upstream sequences was only about 8% of that of pSU86 (Fig. 2B). These results are similar to the observations made by Gruissem et al. [7] in which the 5′ deletion mutant −34 produced mature tRNA at an amount comparable to that of the wild-type but the amount of tRNA produced by mutant −10 was significantly reduced. In other words, both studies suggest that the sequence between −34 and −19 is important for efficient transcription and the basal transcription activity of *trnS* gene required no upstream promoter elements.

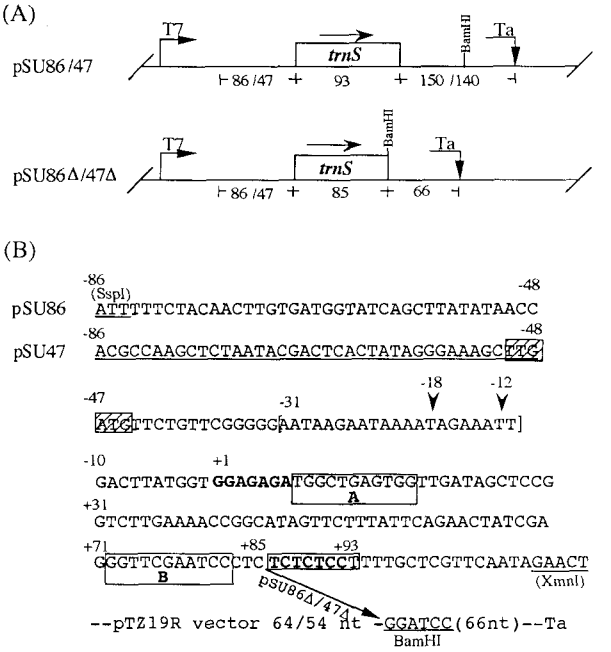

Fig. 1. Schematic diagram of plasmids pSU86, pSU86Δ, pSU47, pSU47Δ and sequences of *trnS* gene. A: Each construct contains the coding region of the *trnS* gene and various lengths of the 5′ non-coding sequence. Plasmid pSU86 contains 86 bp and pSU47 contains 47 bp of the 5′ non-coding sequence, whereas the 3′ flanking sequences to the termination site (Ta, with vertical arrows) are 150 bp (including 64 bp of pTZ19R vector DNA) for pSU86 and 140 bp (including 54 bp of pTZ19R vector DNA) for pSU47. Plasmids pSU86Δ and pSU47Δ each contain 85 bp of the mature tRNA coding region and 66 bp 3′ flanking sequences to the termination site. Horizontal arrows indicate the direction of transcription. B: Sequence of the 199-bp *Ssp*I/*Xmn*I DNA fragment and the upstream vector sequences of pSU47 (underlined) are shown. A possible −35 consensus promoter element for transcription starts at −18 is hatchboxed. The coding region of *trnS* is numbered from +1 to +93 and the aminoacyl stem forming sequences are in bold. An AT-rich region between −31 and −11 is bracketed. Boxed A and B are the two highly conserved intragenic promoter elements. The eight nucleotides at the 3′ end aminoacyl stem forming sequence (boxed) and the downstream sequences before the *Bam*HI site were deleted from plasmids pSU86 and pSU47 to generate plasmids pSU86Δ and pSU47Δ, respectively (indicated by the arrow). The transcription start sites determined in this study are indicated by arrowheads.

Several lines of evidence indicate that the 5′ flanking sequences, in addition to the two highly conserved internal promoter elements, also contributed to the promoter function of eukaryotic tRNA genes ([19–22], and references in [22]). An AT-rich sequence between −34 and −15 of silkworm alanine tRNA gene has been reported to affect the transcription activity significantly [22]. It was noted that the region between −31 and −11 of spinach chloroplast *trnS* gene is extremely AT-rich (Fig. 1B, in brackets), similar to that of the silkworm alanine tRNA gene. This similarity and our deletion analyses and those of Gruissem et al. [7] led us to believe that the AT-rich upstream sequence of the spinach chloroplast serine tRNA gene is indeed required for efficient transcription.

3.2. 3′ Deletion of the trnS coding region results in the production of an unprocessed primary transcript

It has been shown that an appropriately folded structure of

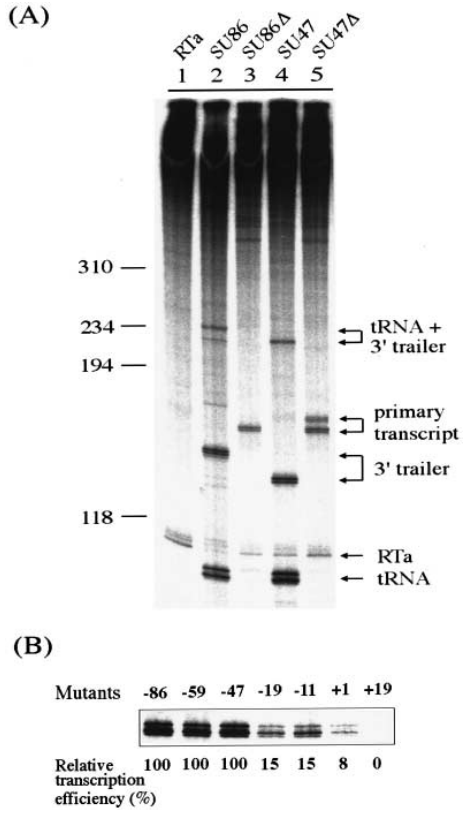

Fig. 2. In vitro transcription/processing assays for various plasmid DNA templates. A: Various plasmid templates, pSU86 (lane 2), pSU86Δ (lane 3), pSU47 (lane 4), pSU47Δ (lane 5) and pTZ19RTa (lane 1, used as control), were incubated in a chloroplast enzyme extract for 60 min and their RNA products were analyzed on a 8% polyacrylamide-8.3 M urea gel. The RNAs produced by both transcription and processing (tRNA+3′ trailer, 3′ trailer, and mature tRNA, lanes 2 and 4) or by transcription only (lanes 3 and 5, indicated as primary transcript) are shown. The processed RNA products are assigned based on their apparent size predicted from each construct and supported by comparison with the DNA size markers, similar to our previous observations for threonine tRNA processing [12]. A background product indicated as RTa appears in all lanes. The smaller products of tRNA+3′ trailer and 3′ trailer for template pSU47 resulted from a deletion in the template of 10 nucleotides at the 3′ end during construction (Fig. 1A). ϕX174-*Hae*III DNA size markers are indicated on the left of the gel. B: The mature tRNAs produced with various 5′ deletion mutants. The relative transcription efficiency was calculated based on the assumption that the processing activity was the same in all reactions and each data point is the average of at least three individual experiments. The multiple bands shown in the tRNAs are thought to result from partial tRNA 3′ CCA processing, similar to our previous observations [12].

tRNA precursor is required for processing [1,23,24]. Therefore, destroying the proper folding of tRNA structure may result in the inhibition of tRNA processing, and consequently the detection of an unprocessed primary transcript. Two plasmids, pSU86Δ and pSU47Δ, were constructed by deleting the last 8 bp (Fig. 1B, boxed) of the 3′ coding region of the *trnS* gene and the downstream sequence before the *Bam*HI site in pSU86 and pSU47, respectively, to prevent the transcription products from normal folding and thus to obtain the unprocessed transcripts. No transcripts corresponding to the mature tRNA or processed products were detected after incubating these plasmids in spinach chloroplast enzyme extract (Fig. 2A, lanes 3 and 5). Instead, a transcript with a size of about 163 nt from pSU86Δ (lane 3) and two transcripts of 163 and 169 nt from pSU47Δ (lane 4) were observed. These results suggest that the RNA molecules produced from templates pSU86Δ and pSU47Δ were not processed and thus might represent the primary transcripts of the *trnS* gene.

An in vitro capping assay was performed to identify the primary transcript. The capped RNA molecules (Fig. 3, lanes 2 and 6), with an additional GTP nucleotide added to the 5′ end of the primary transcript by the activity of guanylyltransferase, migrated one nucleotide more slowly than the uncapped form (Fig. 3, lanes 1 and 5, see also the enlarged pictures). This observation confirms that the transcript synthesized from template pSU86Δ by chloroplast enzyme extract (lane 6) was a primary transcript. As a control, a similar mobility shift was seen for the transcript synthesized from the T7 promoter by T7 RNA polymerase (lane 2). On the other hand, the RNA molecules produced from template pSU86, including the tRNA+3′ trailer and 3′ trailer, were processed products and not capped (lane 4), since their mobility was similar to that of the uncapped RNA molecules (lane 3).

3.3. Identification of the transcription start sites

Determination of transcription start sites was carried out by

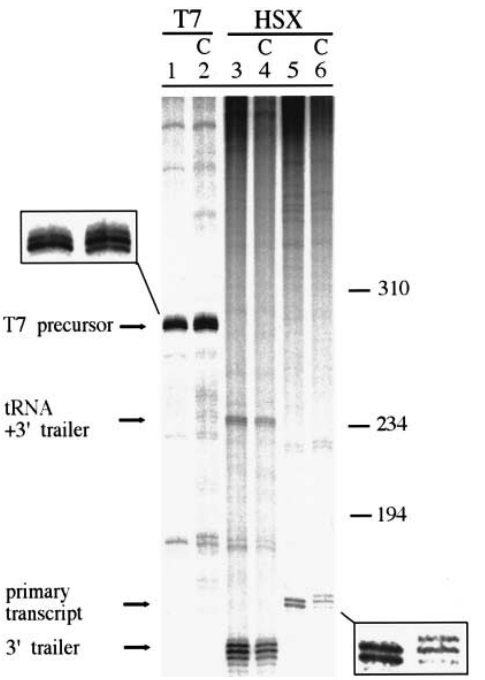

Fig. 3. In vitro capping analysis of the RNAs generated from templates pSU86 and pSU86Δ. Labeled RNAs were synthesized in pair with template pSU86Δ (lanes 1–2 and 5–6) by T7 RNA polymerase (T7) or spinach chloroplast enzyme extract (HSX). RNAs in lanes 3 and 4 were synthesized from template pSU86 with chloroplast enzyme. One half of the labeled RNAs were treated (even lanes) with guanylyltransferase in the presence of GTP. Reactions were analyzed by electrophoresis through 8% polyacrylamide-8.3 M urea DNA sequencing gels. The migrations of the treated RNAs (even lanes) were run in pair with the untreated RNAs (odd lanes). Enlarged pictures (boxed) show the details of the different migrations. The processed intermediates and products (mature tRNAs ran off the gel and are not included) of template pSU86 are shown in lanes 3 and 4. The multiple bands shown in the primary transcripts and the 3′ trailers are thought to result from the termination at two different sites in the Ta terminator [17]. ϕX174-*Hae*III DNA size markers are indicated on the right of the gel.

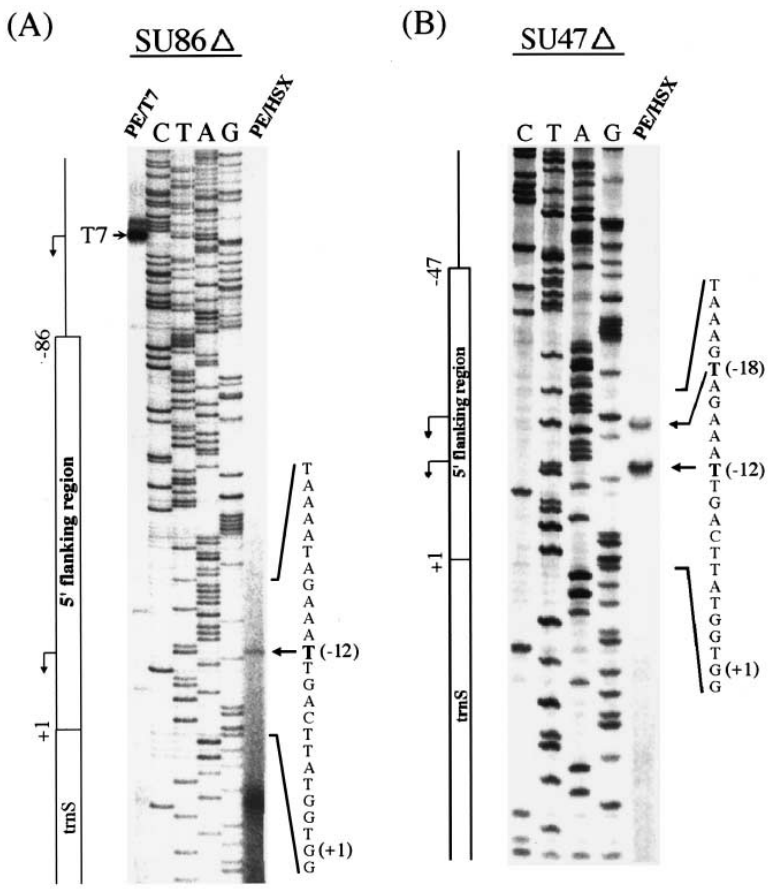

Fig. 4. Primer extension analysis of RNAs synthesized from templates pSU86Δ and pSU47Δ. A 29-nt primer (spanning from +85 to +66 of the *trnS* gene) was 5′ end labeled with [γ-^{32}P]ATP and used to synthesize cDNA fragments with RNA template from pSU86Δ by T7 RNA polymerase (A, lane PE/T7), and with RNA templates from pSU86Δ (A) or pSU47Δ (B), by spinach chloroplast enzyme extract (lanes PE/HSX). Lanes C, T, A, G show the sequence of the 5′ flanking and portion of the *trnS* gene. The discrepancy for the −12 transcription start site between A and B can be accounted for by the different radioisotope-labeled nucleotides used in the sequencing reactions (body-labeled with [α-^{35}S]dATP in A and end-labeled with [γ-^{32}P]ATP in B). The start site at position −12 in B was considered more accurate and therefore position −12 was taken as the start site of *trnS* transcription (indicated by the arrow). Schematic diagrams showing the locations of the *trnS* coding and its 5′ flanking regions relative to the sequence are provided on the left side of each panel.

primer extension using an oligonucleotide primer complementary to positions 66–85 of the *trnS* coding region (see Section 2). The transcription start site of the transcript produced by pSU86Δ was mapped to position −12 upstream of the mature coding region (Fig. 4A). The transcription start sites of the two transcripts produced by pSU47Δ were mapped to positions −12 and −18, respectively (Fig. 4B). It is interesting to note that an additional transcript starts at −18 for pSU47Δ. This transcript might have resulted from an artificial promoter created during plasmid construction. The sequences upstream of −47 were replaced by the vector sequences (Fig. 1B, underlined) after deletion and a sequence TTGATG element (Fig. 1B, hatchboxed), similar to the −35 promoter element, generated at the junction of the vector and the *trnS* gene. This element in conjunction with an AT-rich region, resembling a possible −10 promoter element, might have served as a promoter. It has been demonstrated that the spinach chloroplast enzyme extract, containing various types of RNA polymerase activities, is able to recognize both the prokaryote-type promoter [14,25] and the eukaryote-like tRNA internal promoter [7,12]. Thus, it is possible that two different types of promoters were recognized simultaneously in this enzyme extract. The transcription start site of the primary transcript synthesized by T7 RNA polymerase was mapped to its known position and used as a positive control in this assay (Fig. 4A).

3.4. *Selection of transcription start site for the chloroplast trnS gene*

The transcription start sites for most eukaryotic tRNAs have been shown to be located upstream of the coding region. For instance, the transcription start sites for most yeast tRNA genes have been mapped between −12 and −10 with a relatively constant distance (between 16 and 19 bp) to the internal promoter box A element [3]. In this study, the transcription start site for the spinach chloroplast *trnS* gene was mapped to −12, the same position as most yeast tRNA genes, therefore, a transcription initiation mechanism similar to that of the yeast tRNA is suggested.

The sequence immediately upstream of the transcription start point of a tRNA gene has been suggested to interact with protein factors, such as the transcription factor TFIIIB, which directs the RNA polymerase III to the position where the transcription start site was assumed [26–28]. In yeast tRNA genes, an upstream consensus sequence element (CSE) can work in concert with the A box element in selecting the transcription start site, if the CSE is positioned at a proper

distance from the A box [3]. No such CSE was observed in the chloroplast *trnS* gene, and instead, an extremely AT-rich sequence (19 out of 21 nucleotides between −31 and −11) was found immediately upstream of the start site. This implies that the AT-rich upstream sequence, like the binding region for TFIIIB, might serve as a site for the formation of transcription initiation complex, participating in selection of the transcription start site. Whether these AT-rich sequences, like the yeast CSE, are involved in the selection of the transcription start site will require further investigation.

Acknowledgements: We thank Drs. Y.H. Tseng, C.N. Sun and E.M. Orozco for discussion and critical reading of the manuscript. This work was supported by Grants NSC85-2311-B-005-029 and NSC86-2311-B-005-011 from the National Science Council of the Republic of China.

References

[1] Kirsebom, L.A. (1995) Mol. Microbiol. 17, 411–420.
[2] Galli, G., Hofstetter, H. and Birnstiel, M.L. (1981) Nature 294, 626–631.
[3] Fruscoloni, P., Zamboni, M., Panetta, G., Paolis, A.D. and Tocchini-Valentini, G.P. (1995) Nucleic Acids Res. 23, 2914–2918.
[4] Lee, B.J., Pena, P., Tobin, J.A., Zasloff, M. and Hatfield, D. (1987) Proc. Natl. Acad. Sci. USA 84, 6384–6388.
[5] Park, J.M., Choi, I.S., Kang, S.G., Lee, J.Y., Hatfield, D.L. and Lee, B.J. (1995) Gene 162, 13–19.
[6] Gruissem, W., Greenberg, B.M., Zurawski, G., Prescott, D.M. and Hallick, R.B. (1983) Cell 35, 815–828.
[7] Gruissem, W., Elsner-Menzel, C., Latshaw, S., Narita, J.O., Schaffer, M.A. and Zurawski, G. (1986) Nucleic Acids Res. 14, 7541–7556.
[8] Tonkyn, J.C. and Gruissem, W. (1993) Mol. Gen. Genet. 241, 141–152.
[9] Yamaguchi-Shinozaki, K., Shinozaki, K. and Sugiura, M. (1987) FEBS Lett. 215, 132–136.
[10] Wang, J.M., Davis, N.W. and Gegenheimer, P. (1988) EMBO J. 7, 1567–1574.
[11] Oommen, A., Li, X. and Gegenheimer, P. (1992) Mol. Cell. Biol. 12, 865–875.
[12] Cheng, Y.-S., Lin, C.-H. and Chen, L.-J. (1997) Biochem. Biophys. Res. Commun. 233, 380–385.
[13] Gruissem, W. and Zurawski, G. (1985) EMBO J. 4, 1637–1644.
[14] Hanley-Bowdoin, L. and Chua, N.-H. (1987) Trends Biochem. Sci. 12, 67–70.
[15] Holschuh, K., Bottomley, W. and Whitfeld, P.R. (1984) Nucleic Acids Res. 12, 8819–8834.
[16] Mead, D.A., Skorupa, E.S. and Kemper, B. (1986) Protein Eng. 1, 67–74.
[17] Chen, L.J. and Orozco, E.M.J. (1988) Nucleic Acids Res. 16, 8411–8432.
[18] Orozco Jr., E.M., Mullet, J.E., Hanley-Bowdoin, L. and Chua, N.-H. (1986) Methods Enzymol. 118, 232–253.
[19] Dingermann, T., Burke, D.J., Sharp, S., Schaack, J. and Soll, D. (1982) J. Biol. Chem. 257, 14738–14744.
[20] Sajjadi, F.G., Miller, R.C. and B., S.G. (1987) Mol. Gen. Genet. 206, 279–284.
[21] Rooney, R.J. and Harding, J.D. (1988) Nucleic Acids Res. 16, 2509–2521.
[22] Palida, F.A., Hale, C. and Sprague, K.U. (1993) Nucleic Acids Res. 21, 5875–5881.
[23] Sharp, S.J., Schaak, J., Cooley, L., Burke, D.J. and Soll, D. (1985) Crit. Rev. Biochem. 19, 107–144.
[24] Abelson, J. (1991) Harvey Lect. 85, 1–42.
[25] Chen, L.-J., Rogers, S.A., Bennett, D.C., Hu, M.-C. and Orozco Jr., E.M. (1990) Curr. Genet. 17, 55–64.
[26] Kassavetis, G.A., Braun, B.R., Nguyen, L.H. and Geiduschek, E.P. (1990) Cell 60, 235–245.
[27] Kassavetis, G.A., Bartholomew, B., Blanco, J.A., Johnson, T.E. and Geiduschek, E.P. (1991) Proc. Natl. Acad. Sci. USA 88, 7308–7312.
[28] Kassavetis, G.A., Joazeiro, C.A.P., Pisano, M., Geiduschek, E.P., Colbert, T., Hahn, S. and Blanco, J.A. (1992) Cell 71, 1055–1064.