Hypothesis

Is transcription of higher plant chloroplast ribosomal operons regulated by premature termination?

Jean François Briat, Michel Dron* and Régis Mache

Université I de Grenoble, Laboratoire de Physiologie Cellulaire Végétale, CNRS ERA 488, BP 68, F-38402 Saint Martin d'Hères, Cédex and **Université de Paris XI, Centre d'Orsay, Laboratoire de Biologie Moléculaire Végétale, Bât. 430, F-91405 Orsay, France*

Received 3 August 1983

Three stem-loop structures (H1, H2, H3) can be formed in the leader transcript of the chloroplast rDNA of spinach, tobacco and maize. H2 and H3 partially overlap and cannot exist simultaneously. These potential hairpins lead us to postulate that higher plant chloroplast rDNA is regulated by premature termination. This mechanism could be controlled by the presence or absence of a ribosome translating an hypothetical leader peptide encoded in the rDNA leader sequence of these 3 higher plants.

Higher plant *Chloroplast* *rDNA* *Transcription* *Stem-loop* *Attenuation*

1. INTRODUCTION

It is well known that many events involved in the regulation of genetic expression act at the transcription level especially at the initiation and termination steps of this reaction (review [1–3]).

A particular regulatory mechanism named attenuation has been established which controls the transcription of some bacterial operons involved in amino acid biosynthesis [4]. In this model, 3 regions of dyad symmetry can form two secondary structures in the leader transcript, prior to the structural genes. One structure forms a stem and loop transcription terminator called attenuator which prevents readthrough of the structural genes. Termination at the attenuator is modulated by two mutually exclusive alternate secondary structures. Formation of these alternate structures is regulated by the translation of a small peptide encoded in the leader transcript.

Formation of a stem and loop transcription attenuator has also been reported for the rplJL–rpoBC operon [5] and in the leader region of the S10 ribosomal protein operon of *Escherichia coli* [6]. Such a mechanism does not seem to be related only to prokaryotes since attenuation and pausing of RNA polymerase have recently been described for SV40 DNA [7,8].

Furthermore this regulatory mechanism is also suggested for the transcription of the ribosomal operon rrnB of *E. coli* since pausing and attenuation has been reported in the leader region of this operon [9]. It has to be pointed out that in this case the *nus* A gene protein and the guanosine tetraphosphate (ppGpp) activate the premature termination at the tL attenuator.

The nucleotide sequence and initiation site of transcription upstream the 16 S rRNA gene of spinach chloroplast DNA have been recently determined [10]. Corresponding sequences have also been established for tobacco [11] and maize [12] chloroplast DNA. We report the existence of putative secondary structures in the transcripts of the leader region of ribosomal gene of these plants (two dicotyledons and one monocotyledon). The possible role of these structures in the regulation of chloroplast rDNA transcription is presented as an hypothesis.

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

2. METHODS

Cloning, sequence data, 16 S rRNA gene mapping and initiation site of transcription of ribosomal operon have been described for spinach chloroplast DNA [10] and for tobacco chloroplast DNA [11]. A schematic organization of the leader region of the spinach chloroplast rDNA is shown on fig.1. For maize chloroplast DNA the sequence data of the promoter region has been published in [12]. A comparison was made of the sequence, upstream the nucleotide +1 of the 16 S rRNA gene, of the chloroplast DNA of the 3 higher plants. We have looked for stem-loop structures in the leader region of the 16 S rRNA gene by research of G–C rich inverted repeats. Calculation of free energy of hairpins was made as in [13].

3. RESULTS

In spinach chloroplasts, the initiation site could start at position −113 from the nucleotide +1 of the 16 S rRNA gene [10]. This point corresponds to the F3 region in maize [12] which is a DNA sequence protected against DNase 1 digestion after binding of *E. coli* RNA polymerase. The promoter region we described in spinach is very conserved in the case of tobacco. However, authors in [12,14] have suggested another possible initiation site upstream the starting position we described for spinach. Nevertheless, the stem-loop structures reported below are localized within the leader transcript of the chloroplast ribosomal gene of the 3 plants.

3.1. *Evidence of stem-loop structures in the leader region of chloroplast higher plants rDNA*

Three stem and loop structures are predictable (G–O) from the nucleotide sequence of the 5′-end of the 16 S rRNA gene of these 3 plants (fig.2). It has to be noticed that these hairpins could occur in the ribosomal precursor transcript during its elongation and only before transcription of the 3′-end of the 16 S rRNA gene since it has been suggested [12,14] that 5′- and 3′-ends of 16 S rRNA from maize and tobacco chloroplast can anneal to make typical RNase 3 processing signals as in *E. coli* [12,14].

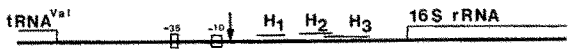

Fig.1. Schematic organization of the leader region of the spinach chloroplast rDNA. Vertical arrow indicates the start site of the transcript [10]. Probnow box and −35 region are represented as small boxes. H1, H2 and H3 correspond to the sequences which can form the base-paired structures described in the text.

The H1 structure (fig.2) is located between positions −106 and −82 in spinach (G = −12.4 kcal) and tobacco (G = −10.2 kcal) and between positions −111 and −84 in maize (G = −6.4 kcal). An interesting point about H1 is the existence of the sequence 5′ AUACAA 3′ immediately following the 3′-end of the stem. This sequence is present in the 3 plants analyzed within a completely conserved 10 bases sequence. An identical sequence also exists in the mRNA terminator sequence of the spinach large subunit of the ribulose bisphosphate carboxylase (fig.2d) for which it has been shown that the termination of transcription effectively occurs in vivo [16].

The H2 structure (between −68 and −48 in spinach, −68 and −47 in tobacco and −65 and −49 in maize) and H3 (between −52 and −23 in spinach, −52 and −22 in tobacco and −55 and −25 in maize) cannot coexist because they partially overlap (fig.1a,b,c). They are less stable than H1 (G between −5.6 and −2.4 kcal) but are more stable than the rrnBtL attenuator (G = −2.4 kcal) described in *E. coli* [9].

It has to be noticed that the sequence following the 3′-end of the H2 and H3 stem are conserved in the 3 plants analyzed for 9 bases in the case of H2 and a long stretch in the case of H3. These sequences as the one following H1 are AU-rich but no poly(U) residues appeared clearly. Only 3 consecutive U's exist after the H2 structure (fig.2a,b,c). By analogy with what is known about the termination of transcription in prokaryote organisms [2], such structures as H1, H2 and H3 can be related to pause sites, more than to terminators. However, by analogy with bacterial factor-dependent terminators, these structures could eventually participate in a premature termination of the rDNA transcription, and if they did they would need unknown factors.

3.2. *A peptide leader could be encoded in the chloroplast leader transcript*

Evidence of a peptide leader has been reported

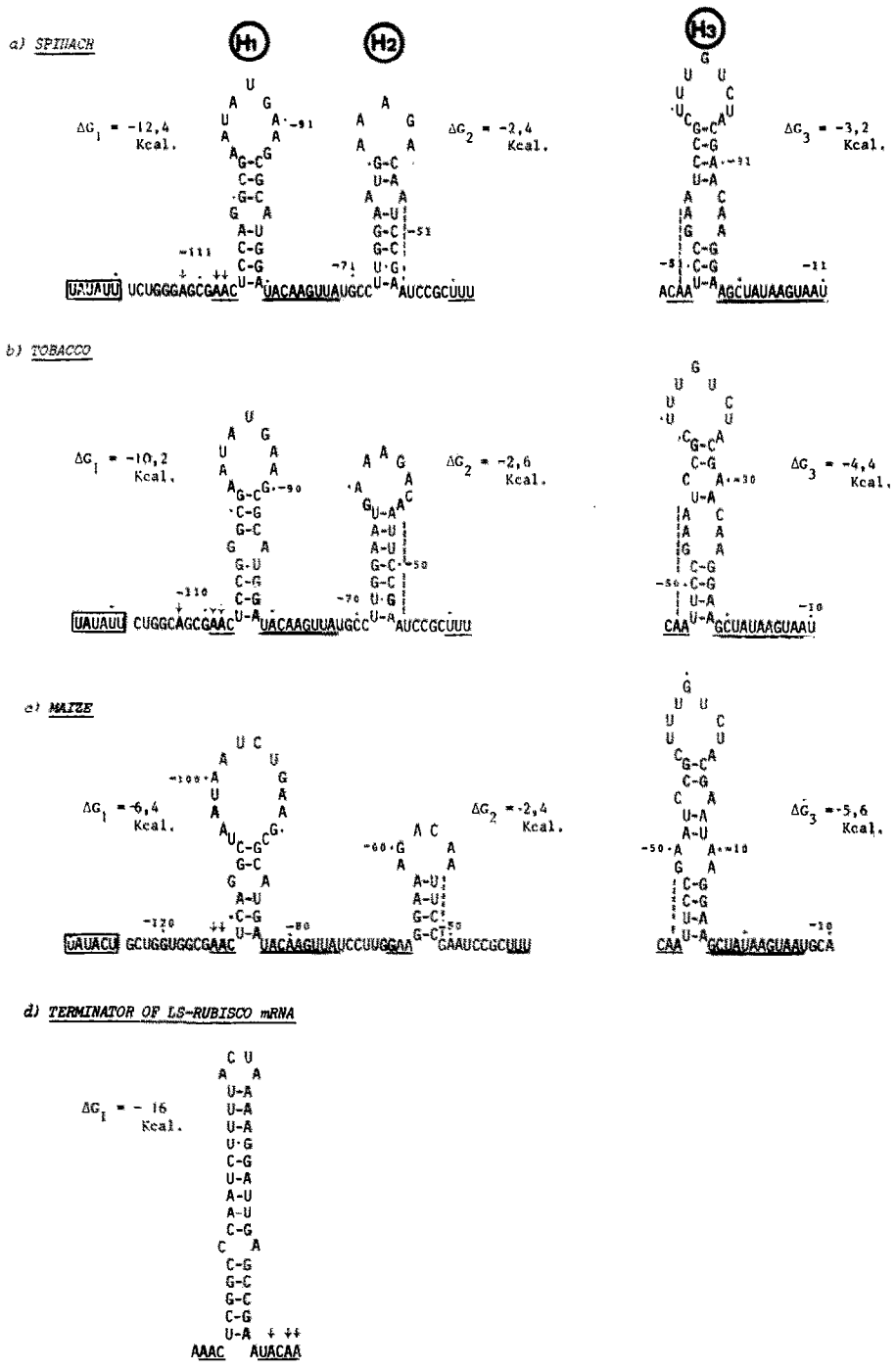

Fig.2. Stem and loop structures in the leader transcript of the chloroplast ribosomal genes. (a) Spinach; (b) tobacco; (c) maize. These sequences are deduced from the DNA sequence data and are aligned from the nucleotide +1 of the 16 S rRNA. Boxed letters indicate −10 Probnow box. Vertical arrows show the possible initiation sites. Underlined sequences are conserved in the 3 plants. Dashed lines show the overlapping of H2 and H3. (d) Terminator of the mRNA of the large subunit of the ribulose bisphosphate carboxylase from spinach [15]. Arrows indicate termination sites of transcription.

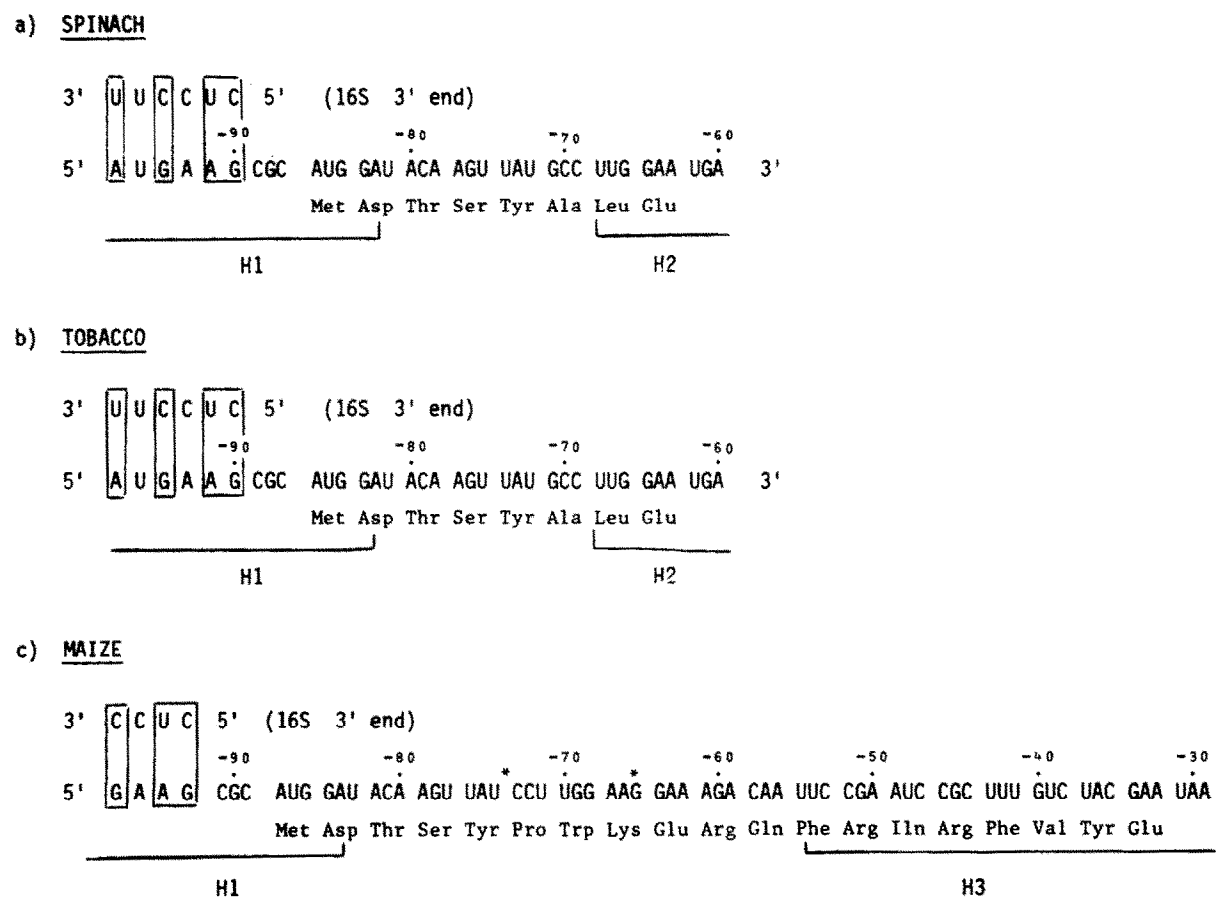

Fig.3. Structure of the hypothetical leader peptide encoded in the chloroplast rRNA leader transcript. (a) Spinach; (b) tobacco; (c) maize. Boxed letters indicate Shine–Dalgarno sequence. * Indicates sequence modifications in maize as compared to spinach and tobacco.

in the case of some bacterial operons involved in amino acids biosynthesis [4] but not in the case of other leader regions [6,9]. We have looked for the peptide coding capacity of the chloroplast rDNA leader transcript and results are shown in fig.3.

In the case of spinach and tobacco (fig.3a,b) a ribosome binding site (Shine–Dalgarno sequence) exists in the loop of the H1 hairpin. This Shine–Dalgarno sequence is followed by an unidentified reading frame (URF) of 8 amino acids. A stop codon UGA is located in the left part of the stem of the H2 structure. In maize (fig.3c) a ribosome binding site in the H1 loop also exists, followed by an URF of 19 amino acids. The stop codon UAA is located in the right part of the stem of H3. The difference between this maize peptide and the tobacco and spinach peptide is due to a G deletion upstream the −73 position in maize and because of the presence of a G at −65 in maize instead of a U in spinach and tobacco (at −62).

4. HYPOTHESIS OF REGULATION

The potentially existing elements in the transcript of the leader region of the chloroplast 16 S rRNA gene lead us to propose a hypothesis concerning a regulatory mechanism for the control of the transcription of the rDNA operon. This hypothesis is based on the model of attenuation [4]. Two cases could be considered depending on whether a ribosome is bound or not to the leader transcript.

4.1. *No ribosome translates the leader peptide*

In this case H1 would exist as well as H2 which would then exclude the formation of H3. As we have discussed above, these structures are more probably related to RNA polymerase pause sites than to factor-independent terminators. So rDNA could be transcribed allowing synthesis of new ribosomes needed to repair or develop the chloroplast.

4.2. *A ribosome translates the leader peptide*

Because of a ribosome excess inside the chloroplast, ribosomes translating the leader peptide would restrain the formation of H1 and H2. In this case, H3 could form and would be necessary for the termination of transcription in the leader region of the rDNA. This possibility is less evident in the case of maize since a ribosome could move through H3. The termination on maize H3 would be possible only if there was a delay between translation and transcription. Furthermore, it has to be noticed that the sequence 5′ CU C CA 3′ located in the foot of the left part of the H1 stem is complementary to the Shine–Dalgarno sequence 5′ UG A AG 3′ in fig.3. Such a complementarity could allow only one translation round of the leader peptide by preventing other ribosomes to bind, as has been shown in the case of the trp operon [4].

Experimental works are in progress to test these different hypotheses.

ACKNOWLEDGEMENT

We thank M. Lebrun for his help to construct H1, H2 and H3 structures and for discussions.

REFERENCES

[1] Chamberlin, M.J. (1974) Annu. Rev. Biochem. 43, 721–775.
[2] Adhya, S. and Gottesman, M. (1978) Annu. Rev. Biochem. 47, 967–996.
[3] Rosenberg, M. and Court, D. (1979) Annu. Rev. Genet. 13, 319–353.
[4] Yanofski, C. (1981) Nature 289, 751–758.
[5] Barry, G., Squires, C. and Squires, C.L. (1980) Proc. Natl. Acad. Sci. USA 77, 3331–3335.
[6] Lindahl, L., Archer, R. and Zengel, J.M. (1983) Cell 33, 241–245.
[7] Skolnick-David, J. and Aloni, Y. (1983) EMBO J. 2, 179–184.
[8] Pfeiffer, P., Hay, N., Pruzan, R., Jakobovits, E.B. and Aloni, Y. (1983) EMBO J. 2, 185–191.
[9] Kingston, R.E. and Chamberlin, M.J. (1981) Cell 27, 523–531.
[10] Briat, J.F., Dron, M., Loiseaux, S. and Mache, R. (1982) Nucleic Acids Res. 10, 6865–6878.
[11] Todoh, N., Shinozaki, K. and Sugiura, M. (1981) Nucleic Acids Res. 9, 5399–5406.
[12] Schwarz, Z., Kössel, H., Schwarz, E. and Bogorad, L. (1981) Proc. Natl. Acad. Sci. USA 78, 4748–4752.
[13] Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Utilenbeck, O.C., Crothers, D.N. and Gralla, J. (1973) Nature New Biol. 246, 40–41.
[14] Todoh, N. and Sugiura, M. (1982) Gene 17, 213–218.
[15] Gegenheimer, P. and Apirion, D. (1981) Microbiol. Rev. 45, 502–541.
[16] Zurawski, G., Perrot, B., Bottomley, W. and Whitfeld, P.R. (1981) Nucleic Acids Res. 9, 3251–3270.