Nucleotide sequence of the 5 S rRNA gene and flanking regions in the cyanobacterium, *Anacystis nidulans*

Susan E. Douglas and W.F. Doolittle

Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia, B3J 4H7, Canada

Received 21 November 1983

The nucleotide sequence of the 5 S rRNA gene and flanking regions (including the terminal 232 nucleotides of the 23 S rRNA gene) of *Anacystis nidulans* has been determined. The spacer between the 23 S and 5 S rRNAs is 70 nucleotides long and contains no detectable transcription termination or initiation signals similar to those found in the comparable region of plastid genomes. A stable stem-and-loop structure can be formed 16 nucleotides downstream from the 5 S rRNA gene, and this is probably the transcription termination signal for the rRNA gene cluster. Homologies between cyanobacterial and plastid genes have been calculated and their evolutionary significance discussed.

Chloroplast DNA *Ribosomal RNA* Anacystis nidulans *Transcription termination* *Promoter* *Inverted repeat*

1. INTRODUCTION

The idea that plastids arose from eubacteria by endosymbiosis is an old one [1], and in light of current information is hard not to accept (review [2]). Similarities in the rRNA gene complement and organization between cyanobacteria and plastids have been demonstrated, and sequence data now available concerning these genes substantiate these findings.

In addition to the 16 S, 23 S and 5 S rRNAs found in eubacteria, the plastids of some, but not all, plants contain a 4.5 S rRNA [3]. This has been shown to be homologous to the 3′-terminus of eubacterial 23 S rRNA [4], and may be associated with the 23 S rRNA molecule by base-pairing with the 5′-end of the latter. It is derived from the same precursor molecule as the 23 S rRNA [5] and sequence analysis shows a potential transcription termination signal in the spacer between the 4.5 S and 5 S rRNA genes in tobacco [6] and maize [7] chloroplasts. A potential promoter sequence for the 5 S rRNA gene has also been found in this region, although this does not necessarily mean that it is functional.

In light of recent evidence demonstrating the similarities between cyanobacterial and chloroplast rRNAs and tRNAs ([8,9], in preparation), we have sequenced the region in *Anacystis nidulans* corresponding to the 4.5 S and 5 S rRNAs of plastids, and analysed them for possible termination or promoter signals.

2. MATERIALS AND METHODS

The recombinant plasmid containing the rRNA gene cluster was propagated in *Escherichia coli* JF1754, and the DNA purified as in [8]. Fragments to be sequenced were cleaved with restriction enzymes according to the sequencing strategy shown in fig. 1, resolved on 1% low melting agarose or 5% non-denaturing polyacrylamide gels, and purified from the gel. Sequencing was done by the dideoxynucleotide chain termination method using the phage M13 system developed in [10]. DNA sequencing gels were 33 × 40 × 0.3 cm and 8% in polyacrylamide (19:1, acrylamide:bis) and contained 8 M urea, 50 mM each of Tris and boric acid, 1 mM EDTA (pH 8.3).

Published by Elsevier Science Publishers B.V.
001457593/84/$3.00

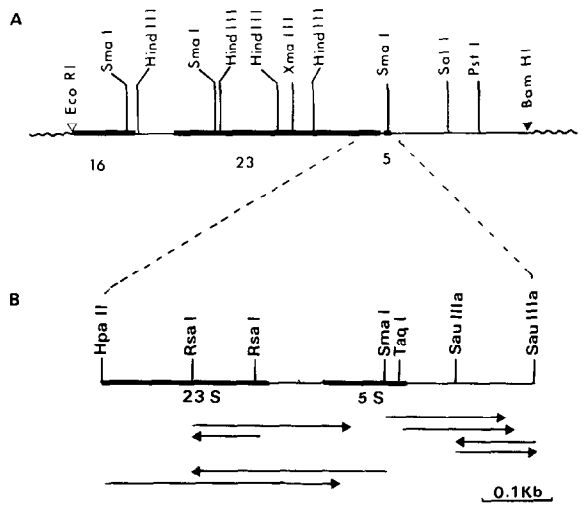

Fig. 1. (A) Restriction map of the 6.3 kb rDNA fragment cloned into pBR322. (B) Strategy used in sequencing the 23 S–5 S rRNA region. Arrows indicate the direction and extent of sequencing. The 23 S rRNA and 5 S rRNA genes are indicated by heavy bars.

3. RESULTS AND DISCUSSION

One of the rRNA gene clusters of *A.nidulans* has been cloned into pBR322 and a restriction map constructed, as in [8]. The strategy used in sequencing the region encompassing the 3′-terminus of the 23 S rRNA gene, spacer, 5 S rRNA gene and 5′-flanking region is shown in fig. 1. The nucleotide sequence is shown in fig. 2. The terminus of the 23 S rRNA gene was determined on the basis of homology to the 23 S rRNAs of *E.coli* and *Z.mays* [11,12]. The termini of the 5 S rRNA gene were located by comparison to the rRNA sequence [13]. The 5 S rRNA gene is 120 nucleotides long and the spacer between it and the 23 S rRNA gene is 70 nucleotides long. The 3′-terminus of the 23 S rRNA gene shows 65% homology to that of *E.coli* and an average homology of 67.3% to the 4.5 S rRNA of the plastids (table 1) when aligned as shown in fig. 3. The 23 S rRNA of *E.coli* shows only 59% homology to plastid 4.5 S rRNAs.

```
        10          20          30          40          50
CCGGAGGAAC GCACCGCTGG TGTACCAGTT ATCGTGCCAA CGGTAAACGC
        60          70          80          90         100
TGGGTAGCTA CGTGTGGAGT GGATAACCGC TGAAAGCATC TAAGTGGGAA  23 S
       110         120         130         140         150
GCCCACCTCA AGATGAGTAC TCTCATGGCA TAAGCCAGTA AGGTCACGGG
       160         170         180         190         200
TAGAACACCC GTTAATAGGC GCTATGTGGA AGTTCAGCAA TGGATGAAGC
       210         220         230         240         250
TGAGGCGTAC TAATAGACCG AGGGCTTGAC CTCTAACACT TTGATATCGG
       260         270         280         290         300
CACTCTCCTC TATGCAGCCT TCAAGGCTCT AATCTCCATA GAGTCACACC
       310         320         330         340         350
CTTCCTGGTG TCTATGGCGG TATGGAACCA CTCTGACCCC ATCCCGAACT
       360         370         380         390         400  5 S
CAGTTGTGAA ACATACCTGC GGCAACGATA GCTCCCGGGT AGCCGGTCGC
       410         420         430         440         450
TAAAATAGCT CGACGCCAGG TCCCTAATCA AAACATCCTA ACCCCCCTCT
       460         470         480         490         500
GCCTCACTAC AACAGCGGAG GGGGGTTAGT CTTTCGATCC CTATCTATCC
       510         520         530         540         550
ATAATCAGTG CTCATGCTTC TTTGCCCCAA TTAGCTCAGC CTAGAAAGCG
       560         570         580         590         600
TTAGTTATTG GGAAAGGCGT TACTATCCCG CCTAATCTTC AACAATGATC
```

Fig. 2. Nucleotide sequence of the 23 S–5 S rRNA region of *A. nidulans*. The 5 S rRNA gene and the 3′-terminus of the 23 S rRNA gene are boxed.

The secondary structure of this region, shown in fig. 4, can be folded into a structure very similar to that proposed for chloroplast 4.5 S rRNA [14], and can base-pair with the 5′-terminus of the 23 S rRNA, as can 4.5 S rRNA. The 23 S–4.5 S rRNA spacer region found in chloroplasts is found at position 138 of the *A.nidulans* sequence. The spacer is 78 nucleotides long in maize [12] and 101 nucleotides long in tobacco [6] chloroplasts. Although this spacer is absent in *A.nidulans*, a sequence homologous to that between another small rRNA (3 S) and the 23 S rRNA in *Chlamydomonas* [16] is present in the 23 S rRNA of *A.nidulans* (in preparation) and *Z.mays* [12].

The 23 S–5 S rRNA spacer of *A.nidulans* is much smaller (70 nucleotides) than that of tobacco plastid (256 nucleotides), and as such does not have room for typical eubacterial promoter and terminator signals [17]. A 6 bp stem-and-loop structure can be formed 34 nucleotides downstream from the terminus of the 23 S rRNA gene, but this leaves insufficient room in the 23 remaining nucleotides of the spacer for a promoter signal. Likewise, the sequence TTG found in the '-35 region' of all *E.coli* rRNA operons, is found

Table 1

Percent homologies between 4.5 S rRNAs of plastids and the 3′-terminus of the 23 S rRNAs of *A. nidulans* and *E. coli*

	E	A	Z	N	W
E. coli		65	60	59	59
A. nidulans			66	71	65
Z. mays chloroplast				82	99
N. tabacum chloroplast					82

```
E.  GAAGGA  ACGTTGAAGACGACGACGTTG   ATAGGCCGGGTGTGTAAGCGCAGCGATGCGTTG AGCTAACCGGTACTAATGAACCGTGAGGCTT AACCT
A.  T....T C...GGT.GA.. .. C....A   ......GCTA....G...TT....A...GA. .A....G.GGC........AG.... AG.....G ....
Z.  T....T AG..GC.. ..... .C....G AA.....TGTCAA...G...T....TG...TA. .C....G.GGCA.C....C....GAAC .AT..G.....
N.  .....T C...GC.. ..... .C....T*CG.....TGTCAA...G...T....TG...TA. .C....G.GGCA.C....CAG....GT..A...G.....
W.  T....TGAG..GC.. ..... .C....T AA.....TGTCAA...G...T....TG...TA. .C....G.GGCA.C....C....GAAC .AT..G.....
```

Fig. 3. Alignment of 4.5 S rRNAs of plastids with the 3′-terminus of the 23 S rRNAs of *A. nidulans* and *E. coli*. A, *A. nidulans*; E, *E. coli*; Z, *Z. mays*; N, *Nicotiana tabacum*; W, wheat. (*) Position of a 7-nucleotide insertion in the 4.5 S rRNA of *N. tabacum* relative to the other plastids.

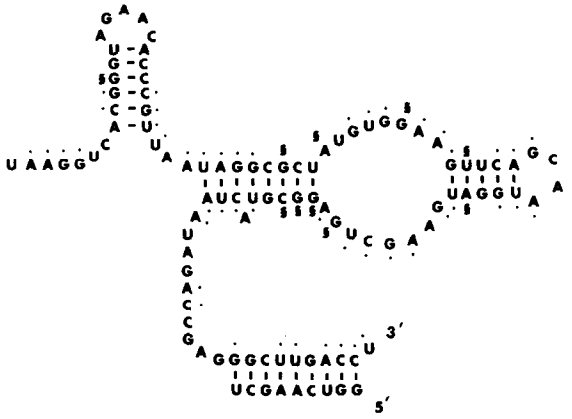

Fig. 4. Secondary structure of the 3′-terminus of *A. nidulans* 23 S rRNA, adapted from the model in [14], Nucleotides which are shared by *A. nidulans*, *E.coli* and plastids are indicated with (*). Those which are shared by *A.nidulans* and plastids only are marked with (§).

9 nucleotides beyond the terminus of the 23 S rRNA gene, but this leaves insufficient room for a termination signal. No sequence corresponding to the Pribnow box sequence (TATPuATPu) is present in the spacer. Thus, it seems unlikely that termination of transcription of the larger rRNAs, and initiation of transcription of the 5 S rRNA occurs here in *A.nidulans*.

A stem-and-loop structure typical of ϱ-independent termination signals is found downstream of the 5 S rRNA (see fig. 5). A less stable stem-and-loop may be formed further downstream. This sequence is characterised by the presence of several direct repeats, a phenomenon also reported in the *rrn*B operon of *E.coli* [11]. One of these sequences (CTAATC) is also found in the 23 S–5 S gene spacer. None of these direct repeats are present in the spacers of plastids thus

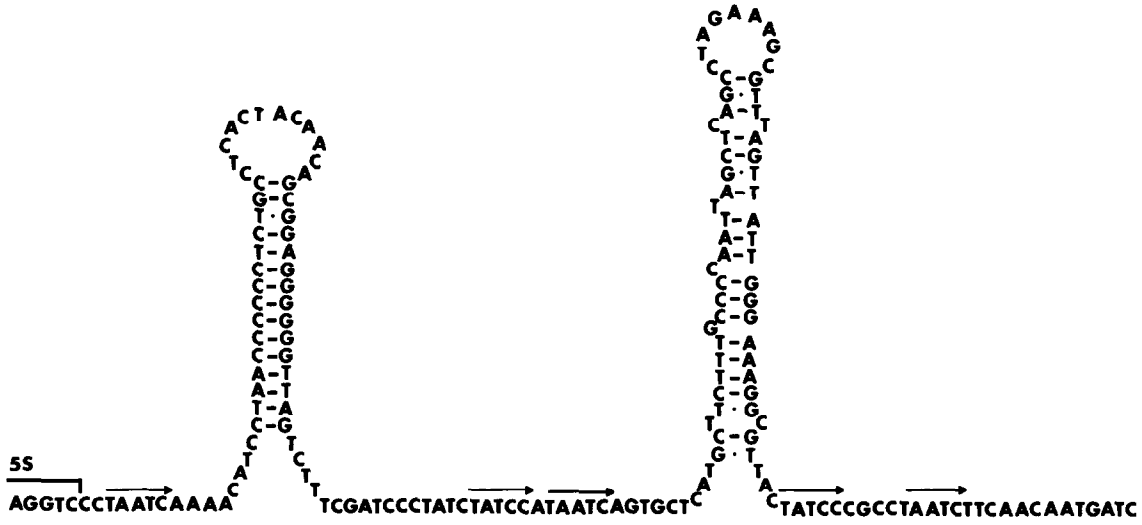

Fig. 5. Potential secondary structure of the 5′-flanking region of the 5 S rRNA gene of *A. nidulans*. Arrows mark direct repeats.

far studied and the significance of this abundance of repeats in such a small region is puzzling.

The 5 S rRNA of *A.nidulans* shares several features with those of plastids, such as base-pairing between nominally single-stranded regions b and d, e and g of the model proposed in [18]. Eubacterial 5 S rRNAs typically cannot form base pairs here. In addition to similarities in secondary structure, the 5 S rRNAs of *A.nidulans* show higher primary sequence homology to each other than either do to *E.coli* (table 2).

The similarities in primary sequence between plastid and cyanobacterial genes reinforces the endosymbiotic hypothesis for the origin of chloroplasts. However, the difference in transcription patterns remains perplexing. 'Lower' plant chloroplasts show the same transcription pattern as *A.nidulans*, yet in some higher plant chloroplasts the 5 S rRNA gene is presumed to be transcribed independently of the larger rRNAs. Furthermore, the ribosomes of some plant choroplasts contain small rRNAs (such as the 3 S, 7 S of *Chlamydomonas* [16] and the 4.5 S of tobacco [6], maize [7], spinach [19], wheat [20] and fern [21]) in addition to the 23 S rRNAs. This represents a unique feature in the way in which the precursor to chloroplast ribosomal RNA is processed and may reflect differences in selection pressure on the genomes of prokaryotes and chloroplasts. This idea is reinforced by the fact that plant chloroplast genes are often interrupted by intervening sequences, whereas those of the prokaryotes studied thus far are not. The introduction of post-transcriptional processing of a common precursor RNA may be a characteristic of chloroplasts acquired after endosymbiosis.

Table 2
Percent homologies between 5 S rRNAs

	E	P	A	S	W
E. coli		59	58	54	47
Prochloron			75	68	50
A. nidulans				66	47
Spinach chloroplast					41

The 5 S rRNA genes from *E.coli* (E), *Prochloron* (P), *A. nidulans* (A), spinach chloroplast (S) and wheat mitochondrion (W) were compared

ACKNOWLEDGEMENTS

Technical help by Jason D. Hofman and Jessica Boyd is gratefully acknowledged. This work was supported by a Medical Research Council grant to W.F.D. and an NSERC student fellowship to S.E.D.

REFERENCES

[1] Altman, R. (1890) Die Elementarorganismen und ihre Beziehungen zu den Zellen, Vent und Comp., Leipzig.
[2] Gray, M.W. and Doolittle, W.F. (1982) Microbiol. Rev. 46, 1-42.
[3] Bowman, C.M. and Dyer, T.A. (1979) Biochem. J. 183, 605-613.
[4] MacKay, R.M. (1981) FEBS Lett. 123, 17-18.
[5] Hartley, M.R. (1979) Eur. J. Biochem. 96, 311-320.
[6] Takaiwa, F. and Sugiura, M. (1980) Mol. Gen. Genet. 180, 1-4.
[7] Dyer, T.A. and Bedbrook, J.R. (1979) in: Genome Organization and Expression in Plants (Leaver, C.J. ed) Plenum Press, New York.
[8] Williamson, S.E. and Doolittle, W.F. (1983) Nucleic Acids Res. 11, 225-235.
[9] Tomioka, N. and Sugiura, M. (1983) Mol. Gen. Genet. 191, 46-50.
[10] Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309-321.
[11] Brosius, J., Dull, T.J., Sleeter, D.D. and Noller, H. F. (1981) J. Mol. Biol. 148, 107-127.
[12] Edwards, K. and Kossel, H. (1981) Nucleic Acids Res. 9, 2853-2869.
[13] Corry, M.J., Payne, P.I. and Dyer, T.A. (1974) FEBS Lett. 46, 63-66.
[14] Machatt, M.A., Ebel, J.-P. and Branlant, C. (1981) Nucleic Acids Res. 9, 1533-1549.
[15] Takaiwa, F. and Sugiura, M. (1982) Eur. J. Biochem. 124, 13-19.
[16] Rochaix, J.D. and Darlix, J.L. (1982) J. Mol. Biol. 159, 383-395.
[17] Gilbert, W.A. (1976) in: RNA Polymerase (Losick, R. and Chamberlin, M. eds) p. 193 Cold Spring Harbour Laboratory, New York.
[18] Garrett, R.A., Douthwaite, S. and Noller, H.F. (1981) Trends Biochem. Sci. 6, 137-139.
[19] Whitfeld, P.R., Leaver, C.J., Bottomley, W. and Atchison, B.A. (1978) Biochem. J. 15, 1103-1112.
[20] Wildeman, A.G. and Nazar, R.N. (1980) J. Biol. Chem. 255, 11896-11900.
[21] Takaiwa, F., Kusuda, M. and Sugiura, M. (1982) Nucleic Acids Res. 10, 2257-2260.