
Sequence organization of repetitive elements in the flanking regions of the chloroplast ribosomal unit of *Chlamydomonas reinhardtii*

Michel Schneider, Jean-Luc Darlix, Jeanne Erickson and Jean-David Rochaix

Departments of Molecular and Plant Biology, University of Geneva, 1211 Geneva, Switzerland

Received 26 September 1985; Revised and Accepted 8 November 1985

ABSTRACT

The flanking regions and the end of the chloroplast ribosomal unit of *Chlamydomonas reinhardtii* have been sequenced. The upstream region of the ribosomal unit contains three open reading frames coding for 111, 117 and 124 amino acids, respectively. The latter polypeptide is partially related to the ribosomal protein L16 of *E. coli*. Two of the open reading frames overlap each other and are oriented in opposite direction. The region between these open reading frames and the 5' end of the 16S rRNA gene contains numerous short direct and inverted repeats which can be folded into large stem-loop structures. Sequence elements that resemble prokaryotic promoters are found in the same region. Several of the repeated elements are distributed throughout the non-coding regions of the chloroplast inverted repeat. Sequence comparison between the 5S rRNA and its gene does not reveal any significant sequence heterogeneity between the chloroplast 5S rRNA genes.

INTRODUCTION

As in most higher plants, the chloroplast ribosomal unit of *Chlamydomonas reinhardtii* is contained within a segment of the chloroplast genome that is repeated in an inverted orientation (1). In *C. reinhardtii* the ribosomal unit consists of the 16S, 7S, 3S, 23S and 5S rRNA genes (2). Previous work has revealed a remarkable sequence homology between the chloroplast rRNA gene sequences of *C. reinhardtii* and those of higher plants and *E. coli* (3,4,5). It was also shown that a 1.5 kb restriction fragment on the 3' side of the ribosomal unit hybridizes to a large number of chloroplast DNA regions, indicating the existence of numerous repeats in this genome (7,8).

In order to gain more insights into the structure of the promoter region of the chloroplast rRNA genes and into the organization of the repeated elements, we have sequenced both

ends of the ribosomal unit of C. reinhardtii. The sequences include 1449 bp upstream of the 16S rRNA gene and the complete sequence of the 5S rRNA gene and its 3' flanking regions. Analysis of the sequences of both flanking regions reveals a large number of direct and inverted repeats that appear to be scattered throughout the inverted repeat. An unusual feature of the promoter region is the presence of three open reading frames, two of which overlap each other and are oriented in opposite direction. One of these open reading frames appears to be related to a ribosomal protein gene of E. coli.

MATERIALS AND METHODS

DNA. The 3kb chloroplast BamHI fragment Ba 1 (1) was cut with HindIII and the smaller of the two HindIII-BamHI subfragments was cloned into pBR322. The construction of the recombinant plasmid containing the chloroplast BamHI-EcoRI fragment BR1.3 was described previously (2). Plasmid DNA was isolated by published procedures (9). DNA sequencing was performed according to the chemical cleavage method of Maxam and Gilbert (10). Sequence analysis was performed on a Hewlett Packard computer, model 9845.

RNA. Cellular RNA of C. reinhardtii was prepared as described (11). The RNA was fractionated by centrifugation on a 5 to 20% linear sucrose gradient in 100mM LiCl₂, 10mM Tris-HCl pH7.8, 4mM EDTA, 0.2% SDS in a SW27 rotor for 18h at 25 000 revs/min. The top fractions of the gradient which contained mostly 5S RNA and tRNA were pooled and ethanol precipitated. This RNA was electrophoresed on a 8% polyacrylamide gel in TBE buffer (75mM Tris-borate, 1mM EDTA, pH8.3). Under these conditions the chloroplast 5S RNA migrates slightly faster than the cytoplasmic 5S RNA. After extraction from the gel, the chloroplast 5S RNA was dephosphorylated with calf intestine alkaline phosphatase (Sigma) and labelled at its 5' end using γ ^{32P}-ATP and polynucleotide kinase (PL Biochemicals). The labelled RNA was repurified by electrophoresis on a polyacrylamide gel in TBE-7M urea. RNA sequencing was performed as described (12,13). S1 nuclease mapping was done as described by Berk and Sharp (14).

RESULTS AND DISCUSSION

The BamHI-EcoRI BR1.3 fragment contains 1449 bp of the

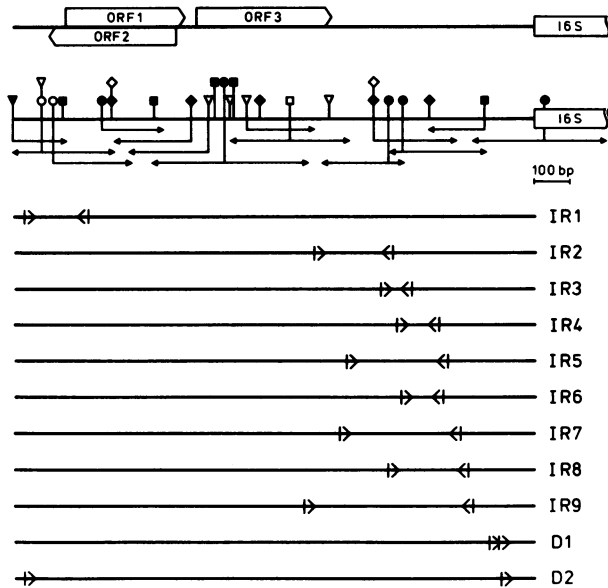


Fig. 1. 5'upstream region of the 16S rRNA gene of *C. reinhardtii*. The restriction sites *Sau3A* ○, *AluI* ■, *RsaI* ◆, *KpnI* ◇, *TaqI* ▽, *DdeI* ●, *HpaII* □, *BamHI* ▾ and the Maxam-Gilbert sequencing strategy are indicated. ORF 1 to 3 represent the three open reading frames found in this region (arrows indicate orientation). The lower part of the figure indicates the location of the inverted repeats (IR1 to IR10) and direct repeats (D1, D2) containing at least 10 contiguous nucleotides. The positions of these repeats in fig. 2 are: IR1 (32-54; 210-198), IR2 (843-861; 1056-1037), IR3 (1026-1043; 1106-1083), IR4 (1070-1081; 1175-1165), IR5 (930-968; 1204-1166), IR6 1085-1105; 1194-1174), IR7 (905-935; 1245-1215), IR8 (1148-1164; 1263-1247), IR9 (810-827; 1279-1260), D1 (1329-1346; 1350-1368), D2 (30-43; 1358-1371).

upstream region of the 16 S rRNA gene plus the 5'terminal end of this gene. The restriction map of this region and the sequencing strategy are shown in fig. 1. The nucleotide sequence of the upstream region is displayed in fig. 2. The corresponding region has been well conserved in mustard (15), maize (16), tobacco (17), spinach (18) and *Spirodela oligorhiza* (19), but the *C. reinhardtii* sequence differs considerably from its higher plant counterpart. While in all higher plants examined a tRNA val gene is located about 300 bp upstream of the 5'end of the 16S rRNA, the comparable *C. reinhardtii* region does not contain this tRNA

Nucleic Acids Research

```

                .30                .60
GGATCCTAACAAARATGTGGCTGTCTGGTATAAAAAAGTAAAGAAACTACACATTTTAGTGTGTTTATATGATGATATTTG
                .90                .120
ATCGAGCATCTGAGATACTCCATCAGAGGAT CAT CTA CTA TCG TTT AAC CAT GAA CGA GCT TTG CGA
AGT ABA TGA TAG CAA ATT GGT ACT TGC TCG AAA CGC
*Op Arg Ser Asp Asn Leu Trp Ser Arg Ala Pro Arg

                .150                .180
Met Ile Cys Phe Lys Gln Gln Lys Ile Thr Ser Leu Phe Val Ser Leu Ile Leu Leu
GAT ATG ATA TGC TTC AAA CAA CAA AAA ATT ACT AGC TTA TTT GTT TCG CTT ATT TTA CTT
TCT ATA CTA TAC GAA GTT TGT TGT TTT TTA ATG ATC GAA TAA ACA AAG CGA ATA AAA TGA
Ser Ile Ile His Lys Leu Cys Cys Phe Ile Val Leu Lys Asn Thr Glu Ser Ile Lys Ser

                .210                .240
Leu Thr Thr Ile Leu Ile Cys Trp Phe Cys Val Leu Ser Pro Lys Val Asn Gly Thr Asn
TRA ACA ACG ATT TTA ATT TGT TGG TTT TGT GTT TTA AGT CCT ARG GTA AAC GGT ACC AAT
AAA TTG TTG CTA AAA TTA AAC AAC CAA AAC ACA AAA TTC AGG ATT CCA TTT GCC ATG GTT
Lys Val Val Ile Lys Ile Gln Gln Asn Gln Thr Lys Leu Gly Leu Thr Phe Pro Val Leu

                .270                .300
Ser Pro Ile Glu Val Asn Val Tyr Asn Pro Ile Pro Leu Ala Val Met Arg Gly Gly Ile
AGT CCA ATT GAA GTA AAT GTA TAT AAT CCA ATT CCA TTS GCT GTA ATG ABA GAA GGT ATT
ATC AAG TTA ACT TCA TTT ACA TAT ATT AAG TTA AAG TAA CCG ACA TTA CTC TCC TCC ATA
Leu Gly Ile Ser Thr Phe Thr Tyr Leu Gly Ile Gly Asn Ala Thr Ile Leu Pro Arg

                .330                .360
Pro Leu Pro Gly Met Pro Pro Val Pro Thr Ala Thr Pro Ser Leu Pro Arg Ser Gly Phe
CCT CTA CCA GGT ATG CCT CCA GTT CCT ACC ACA ACA CCT TCA TTA CCT CAC TCG GGG TTT
AAG ABA TGG TCC ATA CAG AAG TCA AAG ATG GCG TTG TGG ARG TAA TGG ACG GAG CCC CAA
Gly Arg Gly Pro Ile Gly Gly Thr Gly Val Ala Val Gly Asn Gly Arg Glu Pro Asn

                .390                .420
Thr Ser Ser Ala Lys Lys Ile Lys Glu Ser Arg Lys Gln Lys Ser Thr Ala Leu Gln Ala
ACG TCG TCA GCT AAA AAA ATT AAA GAA TCG CGT AAA CAA AAA AGC ACT GCA CTA CAA GCG
ATG CAG CAG TCG ATT TTT TTA ATT TCT TAG CCG ATT TGT TTT TTC GTC ACG TGA TGT TCG
Val Asp Asp Ala Leu Phe Ile Leu Ser Asp Arg Leu Cys Phe Leu Val Ala Ser Cys Ala

                .450                .480
Val Lys Asp Gln Tyr Ile Leu Arg Val Ala Arg Leu *Op
GTT AAA GAC CAA TAC ATT TTA CGC GTG CCG CGA TTA TGA TTGTACAAAGTGAACCCARAGGGTTTAAAT
CCA ATT TCT GGT TAT GTA
Thr Leu Ser Trp Tyr Met

                .510                .540
Met Gln Gly Leu Glu Tyr Thr Leu Glu Pro Glu Gln Ala Lys Lys Cys Lys Thr His
TTT GTG CAG GGT TTA GAT TAT ACC CTC GAA CCC GAG CAA GCT AAA AAA TGC AAA ACT CAT

                .570                .600
Gln Leu Val Asp Leu Ala Lys Ile Asp Leu Asn Ser Trp Ser Gln Glu Asn Val Arg Pro
CAA CTT GTT GAC TTA GCG AAA ATC GAT TTA AAT AGC TGG TCA CAA GAA AAT GTA ABA CCT

                .630                .660
Lys Thr Leu Leu Gly Arg Thr Ser Lys Arg Leu Thr Ser Gln Phe Thr Tyr Glu Phe
AAA ACC TTG TTA GGT CBA ACG TCA AAA AAA CGT TTA ACA AGT CAA TTT TGG TAC GAG TTT

                .690                .720
Thr Lys Lys Gln Lys Leu Lys Asn Ala Leu Tyr Tyr Phe Val Ile Tyr Asp Ile Thr Asp
ACT AAA AAA CAA AAA CTG AAA AAC GCT CTT TAT TAT TTT GTC ATT TAT GAC ATC ACA GAC

                .750                .780
Asn Ser Ile Leu Pro Val Glu Thr Leu Asn Gln Lys Asn Pro Leu Ser Thr Ser Phe Pro
AAT TCT ATA CTG CCG GTA GAG ACA CTA AAC CAA AAA AAC CCT CTA TCA ACC TCC TTC CCC

                .810                .840
Phe Pro Phe Gly Arg Pro Leu Arg Gly Ile Asn Phe Ser Gly Ser Cys Leu His Cys Leu
TTC CCC TTC GGG CGT CCC CTT CGT GGG ATA AAT TTT AAT GGC AGT TCG CTG CAG TGC CTC

                .870                .900
Leu Arg Val Leu Lys *Am
CTT CAA GTG TTA AAA TAG ACBACAGTGCAGTTCBCTGCCCACTGCCTATATTTATATACTGCBATAAATCTTA
                IR3                IR2                .930
                .960                .990                IR7                .1020
GTCCCAAGGGGTTTACATATCCGAAGGAGGABACBAGCAGTBCBGTACCACGCCACTGBCGCTCTAATATAAATATTTG
                IR5                .1050                .1080
GBCAABTAARCTTAGAATAAATTTATTTGCTBCGTTAGCAGGTTTACATACTCCCTARGTTTACTTGCCCGAAGGGGAG
                .1110                IR2                .1140                .1170
GAGGACGTCCTCCACGGAAATATAAATATTTAGTGCAGTGBSTACATAAATAAATTTATGTAAACCCCTTCGGCAACT
                .1200                .1230                IR5                .1260
AAGTTTATCGAGTATTAACTCCTAGTATATAAATATCGCAGTTBGCAGCACAATTTATTTATTTGTCCTTARG
                .1290                .1320                IR7                .1350
GGGAGGGGAARACAAATTTATTTTACTCGGAGCAGCTGTTATTTGAAATTTTATTAARAAAAAATAAATAATTTGA
                .1380                .1410
                IR9
CAAAAAAATAAAAAATTAATTAARACACTGGGATGTTCTACATCAAAAATCAAAAGGGTTTAAATCCCBAC
                .1440
AAAAATTTAACTTTAAGAGT

```

Fig. 2. Nucleotide sequence of the 5'upstream region of the 16S rRNA gene of *C. reinhardtii* as shown in fig. 1. The sequence of the non-coding ribosomal strand is shown. The protein sequences corresponding to the three open reading frames are indicated

between positions 152 and 487, 465 and 112, 519 and 893. The non-coding strand of ORF2 is also shown between nucleotides 465 and 112. Underlined sequences indicate the inverted repeat regions (fig. 1) that participate in the stem-loop structure shown in fig. 4. The arrows mark the start site of transcription; regions homologous to the -10 and -35 boxes are framed.

val gene, nor does it contain any pseudo tRNA genes, as has been reported for Euglena (20).

Three open reading frames are present in the 5'upstream region of the chloroplast ribosomal unit of C. reinhardtii (figs. 1,2). Two of them, ORF1 and ORF3, have the same orientation as the 16S rRNA gene and they are located between positions 152 to 487 and 519 to 893, respectively (fig. 2). They could possibly code for polypeptides of 111 and 124 amino acids. A computer search for related proteins has revealed a 20% amino acid sequence homology over a 55 amino acid stretch between the ORF3 polypeptide and the ribosomal protein L16 of E. coli (21). It has been shown that 5-6 out of 33 polypeptides of the large chloroplast ribosomal subunit of C. reinhardtii are synthesized within the chloroplast (22). However the homologue of the E. coli L16 protein has not yet been identified among the ribosomal proteins of C. reinhardtii. It remains to be explored whether the ORF3 is indeed a ribosomal protein gene. Another possibility is that ORF3 is a pseudogene: we have been unable to detect a homologous transcript. The open reading frame ORF2, (position 465 to 112 in fig. 2) is oriented in the opposite direction. It overlaps nearly all of the first open reading frame and encodes a putative polypeptide of 117 amino acids. An open reading frame of opposite polarity to the rRNA genes has also been found in higher plants in the ribosomal upstream region. In mustard and maize the homologous open reading frames, of 44 and 55 codons respectively, share 30 equivalent codons (15). A similar open reading frame is also present in tobacco except that the 15 amino terminal codons are deleted. In spinach, the first 38 codons of an open reading frame have been determined which specifies an amino acid sequence that is different from the corresponding maize sequence (18). Comparison of the higher plant open reading frames with those of C. reinhardtii does not reveal any significant amino acid sequence homology. However, the hydropathic

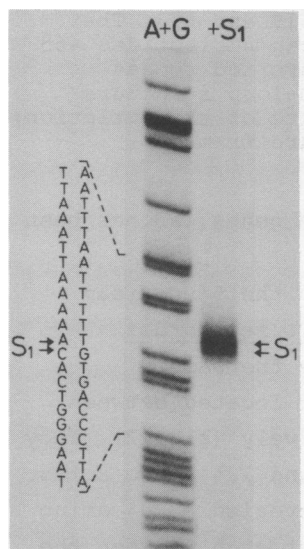


Fig. 3. Mapping of the start site of transcription of the ribosomal unit. Left: Autoradiogram of A+G sequence ladder of the 5'end-labelled coding strand of the 391 DdeI fragment that spans the 5'end of the 16S rRNA gene (cf. fig. 1). Right: size of the protected fragment after hybridization with *C. reinhardii* RNA and subsequent S1 nuclease digestion.

profiles of the *C. reinhardii* polypeptide encoded by ORF 2 (fig. 1) and of its maize counterpart are similar, (data not shown). Whether this observation has any functional significance remains an open question.

In contrast to spinach where a transcript of the ribosomal open reading frame has been detected (18), none could be found in *C. reinhardii* for ORF 1, 2 or 3. It is noteworthy that while the chloroplast protein genes of *C. reinhardii* sequenced to date (23,24,25) contain the TAA stop codon, this is not observed for the three ribosomal open reading frames which end with TGA and TAG. Although there is no evidence that these open reading frames are expressed, it cannot be excluded that they are transcribed at a very low level and/or that their transcripts are unstable.

The 5'end of the ribosomal transcript (fig. 2) was determined by S1 nuclease digestion (fig. 3) of hybrids formed between *C. reinhardii* RNA and a 5'end-labelled 391 bp DdeI fragment (fig. 1) that spans the 5'end of the 16S rRNA gene. Although this analysis does not allow one to discriminate between authentic transcriptional start sites and 5' termini of processed transcripts (26), transcription appears to start 72 bp upstream of the 5'end of the coding sequence of the mature 16S rRNA gene. The presumed start site is preceded by the sequences TAAATT

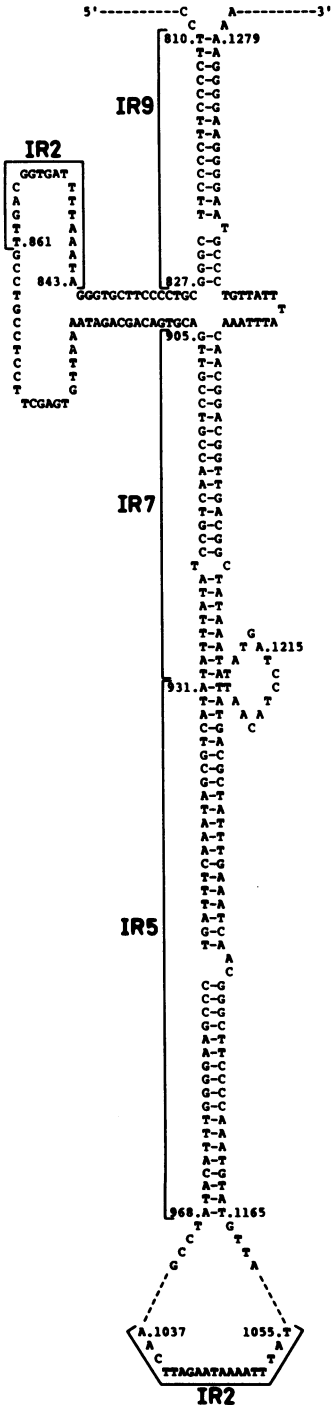


Fig. 4. Stem-loop structure present upstream of the 16S rRNA gene of *C. reinhardtii*. IR2 to IR9 correspond to the repeated elements shown in figs. 1 and 2. Key nucleotides are numbered as in fig. 2.

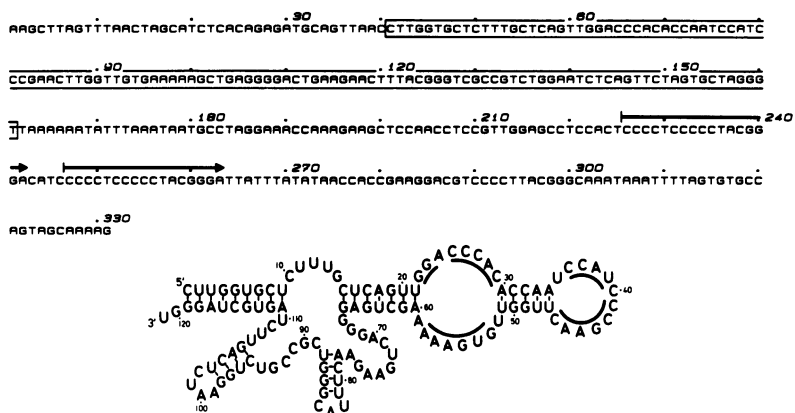


Fig. 5. Upper: Sequence of the chloroplast 5S rRNA gene and of its flanking regions. The 5S rRNA gene is framed. The two arrows indicate two 17bp direct repeats. Lower: Secondary structure of the chloroplast 5S rRNA. The dark lines indicate conserved regions between *C. reinhardtii* and *E. coli* 5S rRNA.

(positions 1368-1373) and TTGACA (positions 1345-1350) which resemble the prokaryotic -10 and -35 boxes (27).

The DNA sequence shown in fig. 2 contains numerous nearly perfect direct and inverted repeats outside of the ribosomal coding regions. The positions of these repeats containing at least 10 contiguous bases are indicated in fig. 1. It is possible to combine several of these repeats into a large palindromic structure displayed in fig. 4. Secondary structures of this sort have been observed previously in the electron microscope in the 5' and 3' flanking regions of the 16S rRNA gene (2). An extensive secondary structure was also observed downstream of the chloroplast ribosomal unit (2).

In order to gain more insights into the structure of this downstream region, its DNA sequence was established. The region sequenced includes the gene of the chloroplast 5S rRNA. RNA sequencing was also performed on the 5S rRNA, which allowed for precise mapping of the 5S rRNA gene boundaries. Since the RNA and DNA sequences are consistent, it does not appear that there is any significant sequence heterogeneity between the chloroplast 5S rRNA genes. Fig. 5 shows that the 5S rRNA contains 121 nucleotides and that it can be folded into a structure similar to

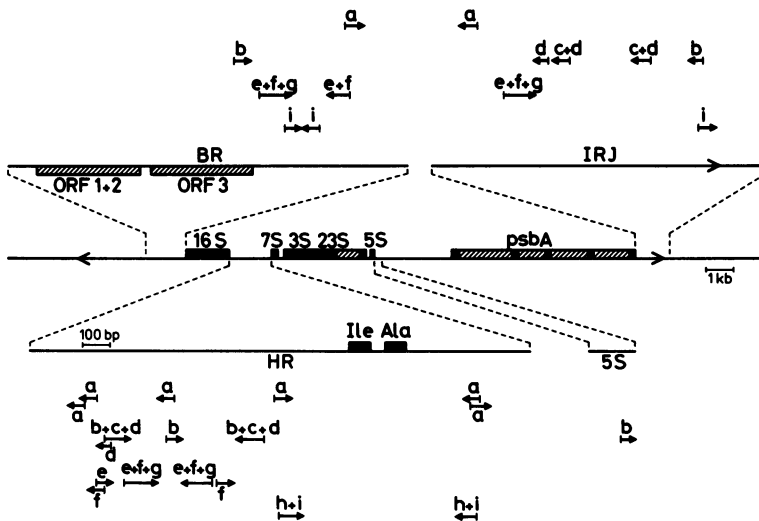


Fig. 6. Location of repeated elements in the non-coding regions of the chloroplast inverted repeat of *C. reinhardtii*. The 16S, 7S, 3S, 23S, 5S RNA genes (2), the *psbA* gene (24) and the tRNA^{Ile} and tRNA^{Ala} genes are indicated. Hatched areas in these genes represent introns. ORF 1+2 (overlapping open reading frames) and ORF3 are as shown in figs. 1,2. The different repeated elements are labelled a to i. BR refers to the 1449bp upstream region of the ribosomal unit (fig. 2); HR represents the 1805bp 16S-7S spacer (Schneider and Rochaix, unpublished results), 5S indicates the DNA region of fig. 5 and IRJ refers to the 1060bp upstream region of *psbA* (Erickson and Rochaix, unpublished results). These four regions are shown enlarged. The ends of the inverted repeat are indicated by arrow heads. Size bars of 1kb and 100bp are given for the entire repeat and for the enlarged portions, respectively.

that of the *E. coli* 5S rRNA. It can be seen that the sequences of two of the loops are nearly identical in the 5S rRNA of *C. reinhardtii* and *E. coli*. In contrast, there is no apparent homology between the chloroplast and cytoplasmic 5S rRNAs of *C. reinhardtii* (28).

A spacer of 90 bp separates the 5' end of the 5S rRNA gene from the 3' end of the 23S rRNA gene (29). The downstream region of the 5S rRNA gene contains repetitive sequence elements, some of which are related to repeated motifs found in the ribosomal promoter region (element b, fig. 6). In addition to the 5' and 3' flanking regions of the ribosomal operon of *C. reinhardtii*, other non-coding regions of the inverted repeat have been sequenced.

They include the ribosomal 16S-23S spacer region (Schneider and Rochaix, unpublished results) and the 5' flanking region of the *psbA* gene (Erickson, Rahire and Rochaix, unpublished results). A computer search for common sequences between all of these regions has revealed several repeated elements present either in the same or inverted orientation. The locations of these elements (a to i) are indicated in fig. 6. Since these sequences have homology with the IR2, IR5, IR7 and IR9 repeats (fig. 2) they may also be involved in the formation of secondary structures similar to that seen in fig. 4.

Palmer (30) has shown that the chloroplast inverted repeat of higher plants undergoes recombination which results in an inversion in the chloroplast genome of one single copy region relative to the other. A similar flipping mechanism has also been observed in *C. reinhardtii* (31,32). In this organism it has been demonstrated that flipping still occurs in chloroplast deletion mutants lacking either end of the inverted repeat (32). This implies that the recombination events are not restricted to one unique site but that they can occur in several regions of the inverted repeat. It remains to be explored whether the repetitive elements described here play a role in this process.

ACKNOWLEDGEMENTS

We thank O. Jenni for drawings and photography. This work was supported by grants 3.587.084 to J.D.Rochaix and to P.F. Spahr 3.664.080 from the Swiss National Science Foundation.

REFERENCES

1. Rochaix, J.D. (1978) *J. Mol. Biol.* 126, 597-617.
2. Rochaix, J.D. and Malnoe, P.M. (1978). *Cell* 15, 661-670.
3. Allet, B. and Rochaix, J.D. (1979). *Cell* 18, 55-60.
4. Dron, M. Rahire, M. and Rochaix, J.D. (1982) *Nucl. Acids Res.* 10, 7609-7620.
5. Rochaix, J.D. and Darlix, J.L. (1982). *J. Mol. Biol.* 159, 383-395.
6. Rochaix, J.D., Rahire, M. and Michel, F. (1985) *Nucl. Acids Res.* 13, 975-984.
7. Rochaix, J.D. (1981) *Experientia* 37, 323-332.
8. Gelvin, S. and Howell, S. (1979) *Mol. Gen. Genet.* 173, 315-322.
9. Katz, L., Kingsbury, D.T. and Helinski, D.R. (1973). *J. Bacteriol.* 114, 577-591.

10. Maxam, A.M. and Gilbert, W. (1980). *Methods in Enzymol.* vol. 65, 499-560. Grossman, L. and Moldave, K. eds., Academic Press, New York.
11. Rochaix, J.D. and Malnoe, P.M. (1978) *Molec. Gen. Genet.* 166, 269-275.
12. Donis-Keller, M., Maxam, A. and Gilbert, W. (1977) *Nucl. Acids Res.* 8, 2527-2538.
13. Simoncsits, A., Brownlee, G.G., Brown, R.S., Rubin, J.R. and Guilly, H. (1977) *Nature* 269, 833-836.
14. Berk, A.J. and Sharp, P.A. (1977) *Cell* 12, 721-732.
15. Przybyl, D., Fritzsche, E., Edwards, K., Kössel, H., Falk, H., Thompson, J.A. and Link, G. (1984) *Plant Molec. Biol* 3, 147-158.
16. Schwarz, Zs., Kössel, H., Schwarz, E. and Bogorad, L. (1981) *Proc. Natl. Acad. Sci USA* 78, 4748-4752.
17. Tohdoh, N., Shinozaki, K. and Sugiura, M. (1981) *Nucl. Acids Res* 9, 5399-5406.
18. Briat, J.F., Dron, M., Loiseaux, S. and Mache, R. (1982) *Nucl. Acids Res* 10, 6805-6878.
19. Keus, R.J.A., Dekker, A.F., Van Roon, M.A. and Groot, G.S.P. (1983) *Nucl. Acids Res.* 12, 6537-6546.
20. Orozco, E.M., Rushlow, K.E., Dodd, J.R. and Hallick, R.B. (1980) *J. Biol. Chem.* 255, 10997-11003.
21. Brosius, J. and Chen, R. (1976) *FEBS Lett.* 68, 105-109.
22. Schmidt, R.J., Richardson, C.B., Gillham, N.W. and Boynton, J.E. (1983) *J. Cell Biol.* 96, 1451-1463.
23. Dron, M., Rahire, M. and Rochaix, J.D. (1982) *J. Mol. Biol.* 162, 775-783.
24. Erickson, J., Rahire, M. and Rochaix, J.D. (1984) *EMBO J.* 3, 2753-2762.
25. Rochaix, J.D., Dron, M., Rahire, M. and Malnoe, P. (1984) *Plant Molec. Biol* 3, 363-370.
26. Strittmatter, G., Gozdzicka-Jozefiak, A. and Kössel, H. (1985) *EMBO J.* 4, 599-604.
27. Rosenberg, M. and Court, D. (1979) *Ann. Rev. Genet.* 13, 319-353.
28. Darlix, J.L. and Rochaix, J.D. (1981) *Nucl. Acids Res.* 9, 1291-1299.
29. Rochaix, J.D., Rahire, M. and Michel, F. (1985) *Nucl. Acids Res.* 13, 975-984.
30. Palmer, J.D. (1983) *Nature* 301, 92-93.
31. Aldrich, J., Cherney, B., Merlin, E., Williams, C. and Mets, L. (1985) *Current Genet* 9, 233-238.
32. Palmer, J.D. (1985) in *Molecular Biology of the photosynthetic apparatus*. Arntzen, C.J., Bonitz, S. Bogorad, L. and Steinback, K. eds. Cold Spring Harbor, New York, in press.
33. Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.* 157, 105-132.