Initiation of *rrn* transcription in chloroplasts of *Euglena gracilis bacillaris*

Peter McGarvey[1], Robert B. Helling[1], Jae-Yong Lee[2], David R. Engelke[2], and M. Raafat El-Gewely[1]

[1] Department of Biology, University of Michigan, Ann Arbor, MI 48109, USA
[2] Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

Summary. The site of initiation of chloroplast rRNA synthesis was determined by S1-mapping and by sequencing primary rRNA transcripts specifically labeled at their 5'-end. Transcription initiates at a single site 53 nucleotides upstream of the 5'-end of the mature 16S rRNA under all growth conditions examined. The initiation site is within a DNA sequence that is highly homologous to and probably derived from a tRNA gene-region located elsewhere in the chloroplast genome. A nearly identical sequence (102 of 103 nucleotides) is present near the replication origin. The near identity of the two sequences suggests a common mode for control of transcription of the rRNA genes and initiation of chloroplast DNA replication. The related sequence in the tRNA gene-region does not appear to serve as a transcript initiation site.

Key words: *Euglena gracilis* — rRNA transcription — Chloroplasts

Introduction

The genes of the *rrn* operons, which encode the ribosomal RNAs (rRNAs), play a fundamental role in the growth of all organisms. Expression of these genes is regulated by a variety of control mechanisms. Control elements that have been identified in the *rrn* operons of *Escherichia coli* include two differently regulated promoters, antitermination sequences, and sequences that enhance transcript initiation (Morgan 1986; Nomura et al. 1984; Lindahl and Zengal 1986; Jinks-Robertson and Nomura 1987). These elements respond to the

nutritional state of the cell so as to increase the rate of *rrn* gene expression exponentially with growth rate (Gausing 1977). Chloroplast rRNA genes show marked structural similarities to prokaryotic *rrn* genes from which they are thought to have originated (Gray 1983; Palmer 1985), but relatively little is known about the control of their expression. The chloroplast genomes of the alga *Euglena gracilis* strain B (*bacillaris*) and Z each contain three complete and one incomplete *rrn* operons arranged in tandem (Koller and Delius 1982b; Koller et al. 1984). Like three of the seven *E. coli rrn* operons and the *rrn* operons of higher plants, the complete operons contain (in order of transcription) genes for 16S rRNA, tRNA^Ile, tRNA^Ala, 23S rRNA, and 5S rRNA (Fig. 1). The incomplete *rrn* operon retains complete 16S and 5S rRNA genes, but the middle has been replaced by a sequence containing an open reading frame (Roux and Stutz 1985) that is cotranscribed with the 16S gene (Koller et al. 1988). Though the sequences of the chloroplast rRNA genes of *Euglena* are very similar to those in the chloroplasts of other algae and plants (El-Gewely et al. 1984; Graf et al. 1982; Rochaix and Darlix 1982; Ohyama et al. 1986; Shinozaki et al. 1986), the sequence of the *rrn* leader region is not (El-Gewely et al. 1984). The leader regions of the *Euglena* chloroplast *rrn* operons contain a 305 basepair sequence that is similar to regions elsewhere in the chloroplast genome that encode tRNA genes, suggesting that these leader regions evolved from structural genes (El-Gewely et al. 1984; Orozco et al. 1980). However, the leader region has diverged sufficiently that it is unlikely to encode any tRNAs, and it has been proposed to have a regulatory role in *rrn* operon transcription. Confirmation that the *rrn* leader has evolved from a structural to a regulatory sequence requires a demonstration that *rrn* transcription initiates within or is otherwise controlled by that sequence. We have deter-
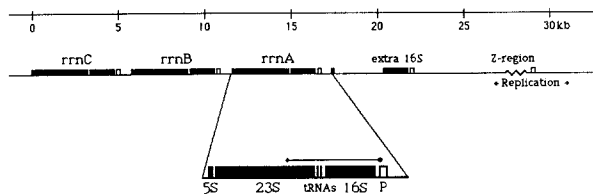
Fig. 1. The arrangement of the *rrn* operons in the B and Z strains of *Euglena gracilis* (Koller and Delius 1982b; Koller et al. 1984). The complete operons are substantially identical except for a 292 base pair deletion in the leader region of the *rrnB* operon of the B strain (El-Gewely et al. 1984; Graf et al. 1982; Orozco et al. 1980; Roux and Stutz 1985). The incomplete operon contains complete 16S and 5S rRNA genes but part of the 23S gene and the two tRNA genes have been replaced by sequences derived in part from the region between the *rrn* operons and in part from an unknown origin (Koller and Delius 1982b; Koller et al. 1984; Roux and Stutz 1985; El-Gewely et al. 1984). The Z-region is a polymorphic region containing a variable number of a 54 bp repeated sequence (Schlunegger and Stutz 1984). The bidirectional origin of replication is within or immediately to the right of the Z-region, although its exact position has not been pinpointed (Koller and Delius 1982a; Ravel-Chapuis et al. 1982) and the sequence across the origin has been determined (Schlunegger and Stutz 1984). P denotes a 103 bp sequence found upstream of all four 16S genes and adjacent to the Z-region. The *arrow* denotes the origin and direction of transcription

mined the origin of *rrn* transcription under different growth conditions and find it to be a single site within the evolved leader region and within a sequence also found both upstream of the incomplete *rrn* operon and near the replication origin of the chloroplast DNA.

Materials and methods

Growth conditions. Cultures were grown in Hutner's acidic organotrophic medium modified to contain 0.02 µg/ml Vitamin B_{12} (Schiff et al. 1971). Four liters of medium in a 5 l flat sided Pyrex bottle were inoculated with 50 ml of a fully grown 200 ml culture of *E. gracilis bacillaris* and grown 4–7 days with aeration at 27–29 °C. Light-grown cultures were exposed to continuous light from four 20 W fluorescent bulbs placed 10–15 cm away. Dark-grown cultures were wrapped in foil and kept in a darkened room. In order to enhance the recovery of unprocessed rRNA, and possibly perturb the regulation of the *rrn* operon some cultures were treated with d(−)threo-chloramphenicol (CAM) or cycloheximide (CHI) prior to harvesting. Treatment with chloramphenicol inhibits translation on plastid ribosomes and increases the yield of unprocessed or partially processed rRNA from the chloroplasts of *Spirodela oligorhiza* and spinach (Hartley and Head 1979; Keus et al. 1984). Cycloheximide used at low concentration has been shown to increase the amount of rRNA in *Euglena* chloroplasts (Heizmann 1974). Cultures treated with CAM or CHI were first grown as single 4 l cultures to late log phase, split into two 2 l cultures each diluted with an additional 1 l of media containing sufficient antibiotic to give a final concentration of 2 mg/ml CAM or 1 µg/ml CHI, grown for an additional 12 h and harvested. Growth was monitored daily by

counting heat-killed cells in a hemocytometer and by measuring chlorophyll concentration as described (Arnon 1949). Cell cultures were harvested during late log phase.

Preparation of total cell RNA. The cells were harvested by centrifugation at 5,000 rpm for 1 min at 4 °C in a Sorvall GSA rotor. The pellets were combined into one 250 ml bottle and washed with cold buffer I (0.33 M sorbitol, 0.05 M *tris*-hydroxymethylamino methane (TRIS; pH 7.5)), lysed with 3 vol./ gm wet wt. of lysis buffer (0.25 M NaCl, 0.1 M EDTA, 0.05 M TRIS (pH 7.5), 0.1 M β-mercaptoethanol, 4 M guanidine isothiocyanate, 2% Triton X-100, 2% Sarcosyl), extracted 3 times at room temperature with phenol/Sevag (1:1) (Sevag = chloroform:isoamylalcohol [24:1]) and 3 times with Sevag. The nucleic acids were ethanol-precipitated, washed with 70% ethanol, dried with vacuum and dissolved in 0.1 x SSC (SSC = 15 mM NaCl, 1.5 mM sodium citrate). RNA was separated from DNA by two precipitations from 2 M LiCl at 4 °C for 4–12 h (Baltimore 1966). The purified RNA was washed with 70% and 95% ethanol, dissolved in 0.1 x SSC and precipitated with 1/10 vol. 7.5 M ammonium acetate, 2 vol. 95% ethanol, and stored at −70 °C. RNA prepared in this way will be referred to as total RNA. Light-grown total RNA will refer to RNA from untreated cells, dark-grown total RNA to RNA from cells grown in the dark, CAM- or CHI-treated total RNA to RNA from light-grown cells treated with CAM or CHI.

Preparation of rRNA from chloroplast ribosomes. Chloroplast ribosomes were prepared as described (Graves et al. 1980). The ribosomal proteins were removed by extraction twice with hot phenol/Sevag and twice with Sevag and the rRNA was precipitated with 0.3 M sodium acetate, 2 vol. 95% ethanol and stored at −70 °C.

Preparation of plasmid DNA. DNA fragments were prepared from plasmids pRH026 and pRH022 which contain, respectively, the *rrnA* and *rrnB* operons (Fig. 1) cloned into the BamH1 site of pBR322 (Helling et al. 1979). The DNA fragments used for S1 experiments were: a 979 bp AvaI to BglII fragment and a 1,359 bp BamH1 to BglII fragment, both 5′-end-labeled at a BglII site 185 nucleotides inside the 16S gene. Large and small scale plasmid DNA preparations were performed as described (El-Gewely and Helling 1980; Maniatis et al. 1982). Restriction enzymes and nucleic acid modifying enzymes were obtained from Bethesda Research Laboratories (BRL), Boehringer Mannheim Biochemicals (BMB), International Biochemicals Inc. (New Haven, Conn.) and New England Biolabs (Beverly, Mass.) and used according to manufacturer's directions.

5′-end labeling of DNA. Ten to fifty µg of plasmid pRH026 or pRH022 were cut to completion with the restriction endonuclease BglII. The mixture of DNA fragments was treated with calf intestinal phosphatase and labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase as described (Maniatis et al. 1982). The labeled DNA fragments were then cut with either BamHI or AvaI restriction endonuclease. The end-labeled DNA fragments were separated by electrophoresis through an agarose gel and the appropriate fragment was collected either by electroelution into a dialysis membrane (Girvitz et al. 1980) or by extraction from low melting temperature agarose (Weislander 1979).

In vitro capping of RNA. Each reaction contained 50 to 200 µg *Euglena* RNA, 50 mM TRIS-HCl (pH 8.0), 2 mM $MgCl_2$, 5 mM dithiothreitol, 6 mM KCl, 3 to 6 µM [α-^{32}P]GTP (> 3,000 Ci/mM), 5 units guanylyltransferase (BRL) in a total of

50 μl. The reaction mixture was incubated at 37 °C for 1 h, stopped with 5 μl of 0.5 M EDTA, adjusted to 100 μl with STE buffer (0.1 M NaCl, 0.01 M TRIS [pH 7.5] and 1 mM EDTA), extracted once with phenol/Sevag (1:1), and passed through a Sephadex G-50 spun column equilibrated with STE buffer (Maniatis et al. 1982). The capped RNA was frozen and stored at −70 °C until use.

Preparation and hybridization of southern blots using capped RNA. Restriction fragments were separated by electrophoresis through a 1% agarose gel containing 0.5 μg/ml ethidium bromide (Maniatis et al. 1982). The DNA was transferred to a Gene Screen Plus membrane (Dupont, Wilmington, Del.) by capillary action using 10 x SSC (Southern 1975). The membrane was dried and preincubated in hybridization buffer (1 M NaCl, 1% sodium dodecyl-sulfate (SDS), 10% dextran sulfate) for 6 h at 60 °C. Sufficient salmon sperm DNA in 5 ml hybridization buffer to give a final concentration of 200 μg/ml when added to the prehybridization mixture was denatured by boiling for 10 min. Capped, light-grown total RNA (approximately 4 x 10^6 cpm) was added to the denatured DNA and immediately mixed with the membrane and hybridization mixture. Hybridization was carried out at 60 °C for 12−16 h. The membrane was washed 2 x 15 min. in 2 x SSC with 1 μg/ml RNase A (Sigma, St. Louis, Mo.) at room temperature, 2 x 30 min. in 2 x SSC, 1% SDS at 60 °C and 2 x 30 min in 0.1 x SSC at 60 °C. The membrane was air-dried and the hybridization bands visualized by autoradiography with Kodak XAR-5 film and Cronex intensifying screens (Dupont) at −70 °C.

S1-mapping. 20 to 50 μg of [α-^32P]GTP-capped RNA (1−5 x 10^5 cpm Cerenkov) or unlabeled RNA and 0.05 to 0.5 μg of unlabeled or 5'-end-labeled (5−10 x 10^3 cpm Cerenkov) DNA were combined and ethanol-precipitated with 30 μg of yeast tRNA (BMB) as carrier. The precipitated nucleic acids were washed with 70% ethanol, dried under vacuum, and dissolved in 10 μl of hybridization buffer (0.04 M PIPES buffer (Sigma), 0.4 M NaCl, 0.001 M EDTA, 80% formamide, pH 6.4) (Berk and Sharp 1977). The mixture was denatured by heating at 80 °C for 20 min and hybridization was carried out at 40 °C for 3 h. Ninety μl cold S1 buffer (0.28 M NaCl, 0.05 M sodium acetate, 4.5 mM ZnSO₄, pH 6.4, 20 μg/ml salmon sperm DNA, Sigma) was added to each tube, 150 units S1 nuclease (BRL) was added to tubes containing capped RNA and 300 units S1 was added to tubes containing labeled DNA and unlabeled RNA. All tubes were incubated for 1 h at 30 °C. Reactions were stopped with 50 μl of 4 M ammonium acetate-0.1 M EDTA, extracted with phenol/Sevag (1:1) and precipitated with ethanol and 30 μg carrier tRNA. The reaction mixtures were washed and dried as above, dissolved in 10−15 μl of loading buffer (80% formamide, 50 mM TRIS-borate [pH 8.0] 1 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue), denatured at 90 °C for 1 min, subjected to electrophoresis through a 6% polyacrylamide-7 M urea gel in TRIS-borate buffer (Maniatis et al. 1982), and visualized by autoradiography. Size standards were prepared by subjecting 10^5 cpm of the end-labeled DNA to A+G chemical cleavage reactions (Maxam and Gilbert 1980).

5'-end sequencing of capped RNA. After hybridization with complementary DNA, treatment with S1 nuclease, and electrophoresis through a denaturing polyacrylamide gel (Fig. 4 lane g), the (α-^32P)GTP-capped RNA was removed from the gel by diffusion into water and precipitated with ethanol. Aliquots were subjected to sequence analysis as described (Boguski et al. 1980; Donis-Keller et al. 1977; Donis-Keller 1980). Reac-

tion products were separated on a denaturing 20% polyacrylamide gel and visualized by autoradiography.

Results

### Identification of primary RNA transcripts by in vitro capping

In order to identify the 5'-end(s) of chloroplast primary transcripts, *Euglena* RNA preparations were "capped" in vitro using [α-^32P]GTP and guanylyltransferase. Guanylyltransferase labels triphosphate 5'-ends and not monophosphate ends as found on processed RNAs (Moss 1981). The 5'-end-labeled RNA produced by this enzyme was then hybridized to *rrn* DNA on southern blots and used in nuclease S1 mapping experiments with labeled and unlabeled DNA from the *rrn* leader region.

Figure 2 shows the electrophoretic profiles of RNA preparations capped in vitro after purification from organisms grown under different conditions. Numerous RNA bands ranging in size from over 2,000 nucleotides to fewer than 50 can be seen. Most of the prominent bands were also seen using capped RNA from isolated chloroplasts (indicated by stars). However, most did not appear to be products of the *rrn* operons. Northern blot hybridization using labelled *rrn* DNA revealed that the chloroplast *rrn*-specific transcripts were 1,500 nucleotides or larger, corresponding to the mature rRNAs and their precursors (not shown). However, *rrn*-specific transcripts smaller than about 200 nucleotides would not have been observed in those experiments.

We have not pursued the identities of the non-*rrn* transcripts, but it is worthwhile noting that several such RNAs appear to be developmentally regulated (Fig. 2). A band at 125 nucleotides was prominent using total RNA from light grown organisms (Fig. 2, lanes b, d, e), but was reduced greatly with dark-grown total RNA (Fig. 2, lane c) and appeared to be absent using RNA from isolated chloroplasts (not shown). Other bands of approximately 150 and 190 nucleotides were more intense using dark-grown total RNA (Fig. 2, lane c) than using light-grown RNA (Fig. 2, lanes b, d, e), and also appeared to be absent using RNA from isolated chloroplasts.
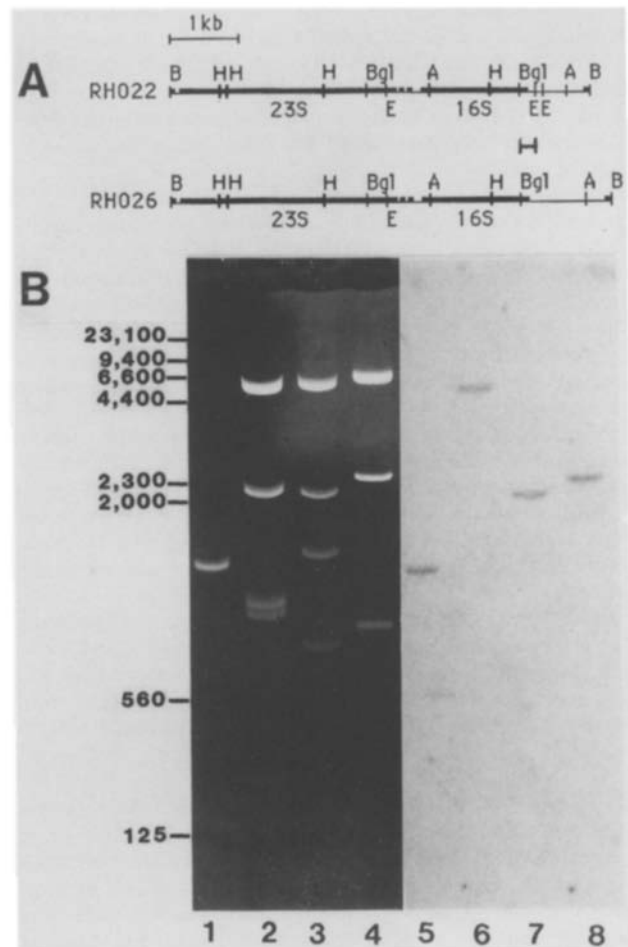
In all cases RNAs with triphosphate 5'-ends available for capping were increased in the presence of the drugs. A reasonable explanation for the increased number of cappable ends is that both drugs inhibit normal processing and degradation of RNA. Some new bands were seen in the drug-treated preparations (Fig. 2). These may represent normally unseen RNA processing intermediates.

In order to determine whether some of the cappable RNAs originated from the *rrn* region, [α-^32P]GTP-cap-
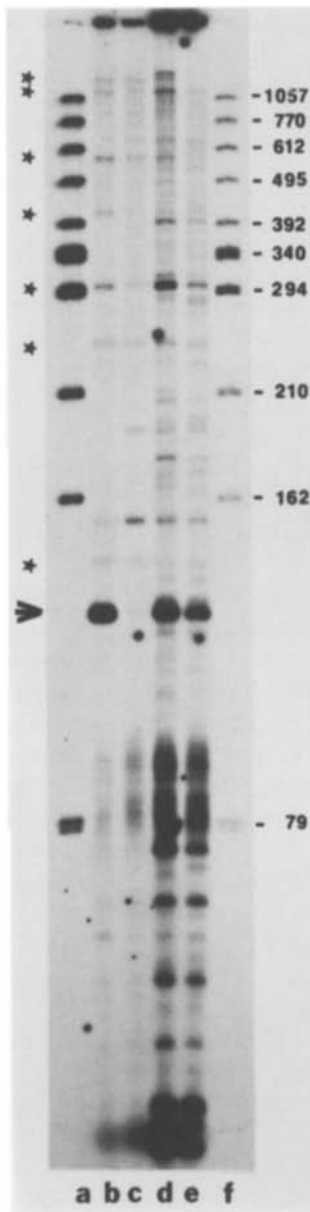
Fig. 2. In vitro capping of RNA. Autoradiogram of [α-^{32}P]GTP-capped RNA electrophoresed through a 6% polyacrylamide gel. Lanes contained a) 5'-end-labeled φX174 HincII size standards; b) 6 μg light-grown CHI-treated total RNA; c) 6 μg dark-grown total RNA; d) 6 μg CAM-treated total RNA; e) 6 μg CHI-treated total RNA; f) 5'-end-labeled φX174 HincII size standards. The *stars* indicate bands containing chloroplast RNA as shown by comparison to [α-^{32}]GTP-capped RNA from isolated chloroplasts (not shown). The *arrow* designates a transcript prominent in light-grown organisms, but rare in dark-grown cells and not seen in RNA from isolated chloroplasts

Fig. 3A, B. *rrn*-specific transcripts can be capped in vitro. A A restriction map of the chloroplast DNA from the *rrnA* (pRH026) and *rrnB* (pRH022) operons cloned into the BamH1 site of pBR322. A = AvaI, B = BamH1, Bgl = BglII, E = EcoR1, and H = HindIII. The only segment common to all four hybridizing bands shown in Fig. 3B is depicted between the maps for *rrnA* and *rrnB*, and corresponds to the DNA between the Bgl site in the 16S rRNA gene and the nearest Eco site. B Hybridization of capped light-grown total RNA with cloned *rrnA* or *rrnB* chloroplast DNA. *Lanes 1–4* are from an ethidium bromide stained 1% agarose gel containing *1*, a 1,359 bp BamH1 to BglII fragment from pRH026 containing the entire *rrnA* leader region; *2*, pRH026 digested with HindIII; *3*, pRH026 digested with AvaI; *4*, pRH022 digested with EcoR1. *Lanes 5–8* are from an autoradiogram of a southern blot of the same gel as *lanes 1–4* using [α-^{32}]GTP-capped light-grown total RNA as probe

ped RNA was hybridized with restriction fragments from the *rrn* operons. Figure 3b shows an ethidium bromide stained agarose gel of the separated DNA fragments, and the corresponding autoradiograph after blotting and hybr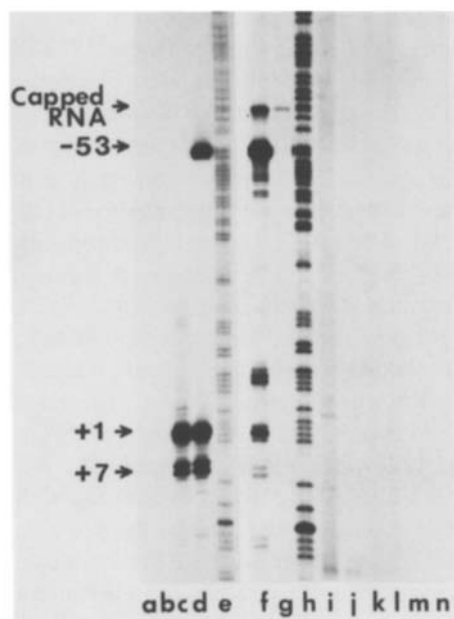idization. In all four lanes the hybridizing band includes the leader segment 5' to the 16S rRNA gene (Fig. 3A). We conclude that the RNA preparations contained *rrn*-specific transcripts with the original 5'-end still intact.

S1-mapping of 5'-RNA termini

The initiation point of *rrn* transcription was determined by using S1 nuclease-mapping. S1-mapping involves

Fig. 4. Determination of the 5'-end of the 16S rRNA and of the initiation point for transcription of the *rrn* operon. The autoradiograms show bands of labeled RNA or *rrnA* DNA that were protected from S1-nuclease digestion by hybridization and sized on polyacrylamide gels. Experiments with results shown in *lanes a–f, h, i, l, m* used 5'-end labeled DNA as hybridization probe. [α-^{32}P]GTP-capped RNA was used as probe in the experiments with results shown in *lanes f, g, i, j and k. Lanes a* through *n* represent the following: *a*, control without RNA; *b*, control with 1 μg mature rRNA but lacking S1-nuclease; *c*, complete reaction mixture with 1 μg mature chloroplast rRNA; *d*, complete reaction mixture with 1 μg total cell RNA; *e*, size standards resulting from a Maxam and Gilbert (A + G) reaction of the AvaI-BglII probe; *f*, complete reaction mixture using 20 μg capped total cell RNA, DNA probe, and 50 ng unlabeled AvaI-BglII DNA; *g*, complete reaction mixture using 20 μg capped total cell RNA and 50 ng unlabeled AvaI-BglII DNA; *h*, same as *lane e*; *i*, control with 20 μg capped total cell RNA, DNA probe, and no S1; *j* control with 20 μg capped total cell RNA, 50 ng unlabeled DNA, no S1; *k*, control with 20 μg capped total cell RNA and no DNA; *l*, control with DNA probe, 1 μg total cell RNA, no S1; *m*, control with DNA probe and no RNA; *n*, empty *lane*. Position numbers refer to the S1 bands and not to the corresponding band in the sequence ladder. A 1–1.5 bp downshift of the most intense S1 band is required to align it with the correct DNA sequence (Green and Roeder 1980). This can be seen in *lanes d* and *e* where a 1.5 base shift aligns the +1 band with a TG corresponding to a UG known to be at the 5' end of the 16S RNA (see text). Bands at both positions +6 and +7 may be seen, but usually the +6 band is less prominent. Oligonucleotide analysis has shown that the +7 band but not the +6 band could correspond to a mature RNA end (see text). The band at position +12 in *lanes e* and *h* is the result of a single strand nicking activity associated with digestion by AvaI during preparation of the DNA probe. The bands at about −9 in *lane f* are seen occasionally and are the result of partial S1-digestion through an AT-rich region of a longer RNA:DNA hybrid. Long exposure revealed no additional bands in *lane g*. Control *lanes* showed bands high up in the gel, as expected, corresponding to hybridization of complementary DNA strands, and, in the absence of S1 nuclease, of large labelled DNA and RNA molecules

hybridizing the RNA of interest to a homologous DNA segment and determining the extent to which the DNA or RNA is protected from degradation by a single-strand-specific nuclease such as S1 (Berk and Sharp 1977). A DNA fragment extending from nucleotide +185 (BglII site) in the 16S *rrnA* rRNA gene to nucleotide −794 (AvaI site) in the *rrnA* leader, and end-labeled with ^{32}P at the BglII site was hybridized with RNA. After treatment with nuclease S1 the protected DNA fragments were sized by electrophoresis through a polyacrylamide gel (Fig. 4).

Several bands may be seen (lanes c, d). The 5'-end of the 16S rRNA had been predicted from the DNA sequence (Graf et al. 1982; El-Gewely et al. 1984) and corresponds to the position labeled +1. The band at +1 appears following hybridization with both total light grown RNA and with rRNA from purified chloroplast ribosomes (Fig. 4, lanes c, d). An additional band corresponding to an RNA with a 5'-end at position +7 was invariably seen even when different RNA preparations and different conditions of hybridization and S1-digestion were used. Both the +1 and +7 positions are consistent with the results of other experiments that identified the 5'-end as beginning with a UG (Zablen et al. 1975; see sequence in Fig. 6). The DNA sequences of this region in the three complete and one incomplete operons are nearly identical (El-Gewely et al. 1984; Graf et al. 1982; Roux et al. 1983) so the two bands are unlikely to represent the products of different DNA sequences. We conclude that the +1 position is the 5'-end of the 16S rRNA. The +7 band probably results from S1-digestion of an RNA-DNA hybrid either poorly paired due to the AT-rich region near the 16S end, or improperly paired due to localized self-pairing of the DNA or RNA, and is unlikely to correspond to the terminus of an RNA found in vivo. However, the possibility that there are two populations of 16S rRNA with different 5'-ends cannot be ruled out.

When S1 mapping was done with total RNA an additional band corresponding to a transcript with a 5'-end at position −53 was observed (lane d; see also Fig. 5 for proof that the band corresponds exactly to −53). Similar bands were observed with RNA extracted from purified chloroplasts and with total RNA from CHI-, CAM-treated and dark grown cells (not shown). No band corresponding to position −53 was observed with rRNA from chloroplast ribosomes (lane c). No additional bands corresponding to RNAs extending further upstream were observed. We conclude that a transcript extends from position −53 into the sequence found in the mature 16S rRNA and is presumably a precursor to the 16S rRNA. However, it remained possible that the RNA beginning at −53 was derived from a longer original transcript.
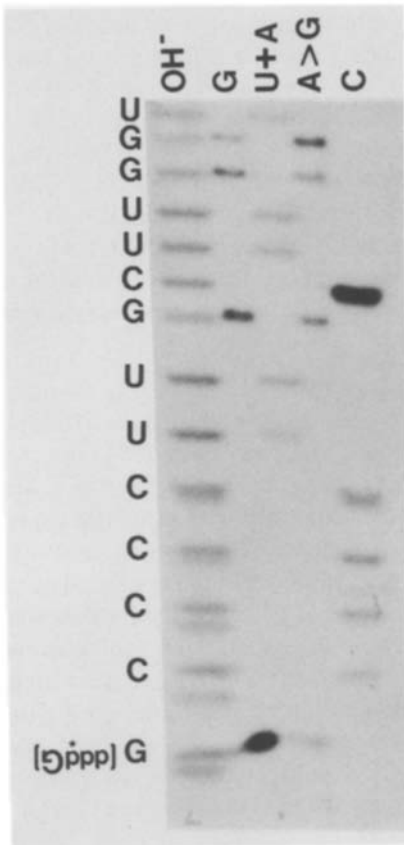
Fig. 5. Sequence analysis of the 5'-end of the capped *rrn* transcript. *Lanes* represent partial digestion with: heat at alkaline pH *("OH-"lane)*; RNase T1 *("G" lane)*; RNase Phy M *("U + A" lane)*; RNase U2 *("A > G" lane)*; and RNase CL3 *("C" lane)*. The deduced nucleotide sequence of the first 14 positions from the capped 5'-terminus is shown at the *left* of the *panel* (the cap is designated as an upside-down G). The cause of the doublets formed by alkaline hydrolysis is not certain, but we have observed it repeatedly when using such capped transcripts. It is likely to result from attack on the cap (α) phosphodiester bond, but could also be 2'- and 3'-phosphate isomers of the 3'-termini

In order to identify the true 5'-end(s) of primary transcripts we capped them in vitro with [α-^{32}P]GTP and then used S1-mapping to identify transcripts specific to the *rrn* operons. The 5'-end labeled RNA was hybridized to labeled or unlabeled DNA from the leader region, digested with the nuclease, and electrophoresed through a polyacrylamide gel under denaturing conditions (Fig. 4, lanes f and g). When labeled RNA was hybridized with unlabeled DNA a single band was observed (lane g). When the labeled RNA was hybridized with 5'-end-labeled DNA, a set of additional bands corresponding to those observed in the previous S1 experiments was seen also (lane f). The RNA band in lanes f and g migrated as if the RNA were longer than the corresponding DNA in the band at −53 (238 nucleotides in length). The difference in the mobilities of the two bands varied with the electrophoresis conditions. Because of the presence of an added GDP on the 5'-end of the RNA from capping, the difference in sequence between complementary strands, and the difference in mobility between RNA and DNA, the bands are not expected to comigrate. The capping experiments have been repeated with similar results using RNA from cells grown under a variety of conditions including complete darkness (not shown).

The nucleotide sequence at the 5'-end of the primary transcript

It seemed likely that the capped RNA transcript corresponded to the DNA band at position −53 and that transcription of the *rrn* operons initiates at this point in the DNA sequence. In order to identify unambiguously the initiating nucleotide, the capped, S1 nuclease-resistant RNA fragment (Fig. 4) was isolated and subjected to enzymatic end-sequence analysis (Fig. 5). The 5'-capped nucleotide is the G at position −53.
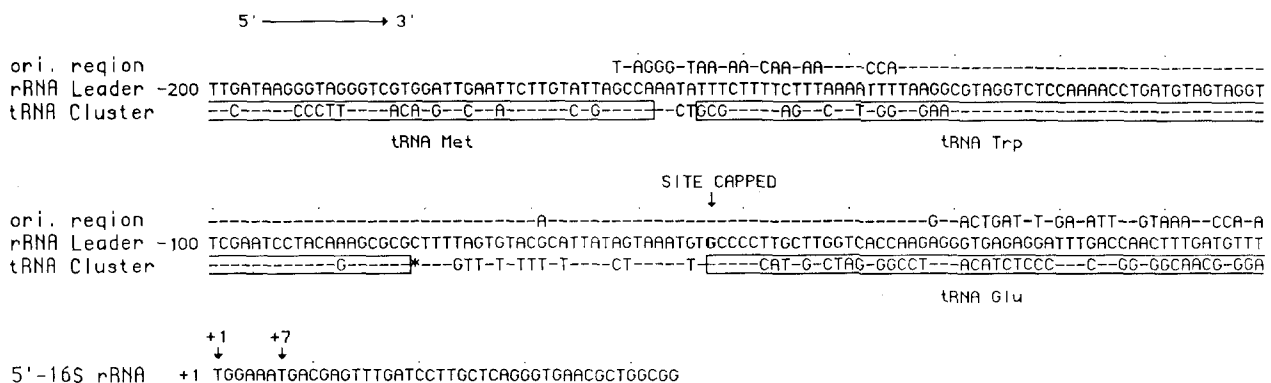
Fig. 6. Comparison of the DNA sequence of the *Euglena* chloroplast rRNA leader region (El-Gewely et al. 1984) with the corresponding sequence from a cluster of chloroplast tRNA genes (Hollingsworth and Hallick 1982) and a sequence near the chloroplast origin of replication (Schlunegger and Stutz 1984). Only the RNA-like strand is shown. *Dashes* (--) signify that the nucleotide is identical to that found in the leader region. *Asterisks* (*) signify deletions. The nucleotides corresponding to the +1 and +7 bands identified by S1-mapping and the site of transcript initiation are signified by *arrows*

Discussion

The DNA sequence of the 5'-end of the 16S rRNA plus 200 bases of the upstream leader region is shown in Fig. 6. Transcription of the *rrn* operons initiates at a single nucleotide 53 nucleotides upstream of the mature 5'-end of the 16S rRNA during growth in the light and dark as well as during treatment with the translation inhibitors CHI and CAM. Single initiation sites for *rrn* transcription have also been identified in the chloroplasts of several higher plants (Briat et al. 1987; Strittmaller et al. 1985; Sun et al. 1986). This is in contrast to the multiple initiation sites found for *rrn* genes in maize mitochondria (Mulligan et al. 1988) and in *E. coli* (Jinks-Robertson and Nomura 1987).

The sequence TATAGT preceding the −53 initiation site resembles the consensus prokaryotic −10 promoter sequence TAtaaT (McClure 1985; Reznikoff et al. 1985). No sequence closely resembling the consensus sequence TTGaca is found at the −35 region although CCTACA is nearby. Most reported transcript origins from chloroplast genes have been inferred from S1 mapping or from in vitro studies and have not been verified by capping of transcripts found in vivo. However most proposed chloroplast promoters of protein, tRNA, and rRNA genes in higher plants fit both the prokaryotic consensus −10 and −35 sequences (Bradley and Gatenby 1985; Hanley-Bowden et al. 1985; Kung and Lin 1985; Strittmaller et al. 1985; Sun et al. 1986; Grussem and Zurawski 1985). Several chloroplast tRNA genes in spinach do not have upstream promoters and seem to use internal promoters (Gruissem et al. 1986). The *Euglena* chloroplast is reported to contain two RNA polymerases, one of which is specific for the rRNA genes (Greenberg and Hallick 1986; Narita et al. 1985). This rRNA-specific polymerase may recognize a promoter that does not conform to the consensus prokaryotic promoter.

The leader region from −49 to −181 is 75% identical to part of a cluster of six chloroplast tRNA genes (see Fig. 6). We were unable to detect initiation of transcription within this region of the tRNA cluster by southern hybridization or by S1-mapping using capped RNA (not shown). These tRNA genes have been transcribed and processed in vitro using a partially purified RNA polymerase activity from *Euglena* chloroplasts that does not transcribe the *rrn* operons (Greenberg and Hallick 1986). The *rrn*-specific RNA polymerase activity from *Euglena* chloroplasts did not transcribe the non-*rrn* encoded tRNA genes in vitro (Narita et al. 1985). It seems unlikely therefore that the homologous DNA segment in the tRNA gene cluster serves as a primary promoter for transcription of adjacent tRNA genes. This suggests that following insertion of a copy of that segment just upstream of the 16S rRNA gene it

evolved to a new regulatory role in which it serves as origin and possibly as promoter for *rrn* transcription. Nevertheless the conservation of most of the segment homologous to the tRNA^{Trp} gene (48 of 49 consecutive nucleotides) suggests some similar function at the two locations. Studies on several biosynthetic operons of enteric bacteria have suggested that the evolution of transcription-controlling sequences from tRNA genes may be of widespread occurrence (Ames et al. 1983; Williamson and Jackson 1987).

An even more remarkable homology has been observed between this portion of the *rrn* leader region and a sequence found near the origin of replication (Schlunegger and Stutz 1984). As seen in Fig. 6 an almost identical copy (102 of 103 base pairs) of a leader region sequence extending across the rRNA transcription initiation site and the sequence homologous to the tRNA^{Trp} gene is present near the origin. If transcription initiated at the nucleotide corresponding to −53 in the *rrn* leader it would progress toward the polymorphic Z-region (Fig. 1) only 61 nucleotides downstream. Little is known about DNA replication in chloroplasts. *Euglena* and *Chlamydomonas* are the only two organisms with clearly mapped origins of chloroplast DNA replication (Koller and Delius 1982a; Ravel-Chapius et al. 1982; Waddel et al. 1984). Possible origins of replication have been identified by in vitro DNA synthesis in petunia and maize (de Haas et al. 1987; Gold et al. 1987).

The evolution of the common sequence may be traced with uncommon clarity (El-Gewely et al. 1984). A segment from a tRNA gene cluster was inserted upstream of the *rrn* genes (El-Gewely et al. 1984), where it serves as initiation point for *rrn* transcription (this paper), and may have acquired a more direct role in transcription control. The right end of that same segment and a short adjoining sequence (Fig. 6) was then translocated close to the replication origin. Examination of the ends of the transferred sequences suggests that the alternative pathway in which DNA translocated from the tRNA genes to the replication origin region, thence to the *rrn* leader, is less plausible. The homology of the *rrn* leader and the tRNA genes continues leftward (see El-Gewely et al. 1984) beyond the region of similarity of either with DNA at the origin region (Schlunegger and Stutz 1984), while the homology of the *rrn* leader and the origin region extends for a short distance rightward past the region of apparent similarity of either with the tRNA region (Fig. 6). This pattern is consistent with transfer from the leader to the origin region and inconsistent with the reverse direction of transfer.

The extraordinary sequence conservation at the *rrn* operons and near the origin suggests identity in an important function at the two locations. The function involves initiation and possible control of transcription

at the *rrn* operons, and therefore possibly at its translocated position as well. What might be transcribed there? One possibility is an RNA primer for replication. It has not been shown that the conserved sequence at or close to the origin region plays a role in replication. Nevertheless, its location suggests such a relationship and that suggestion serves as a useful working model. Replication and rRNA synthesis are two basic functions of the chloroplast genome and in *Euglena* both chloroplast DNA and rRNA increase in response to light (Chelm et al. 1977). A common sequence controlling the separate processes would ensure their coordinate control.

References

Ames BN, Tsang TH, Buck M, Christman MF (1983) Proc Natl Acad Sci USA 80:5240–5242

Arnon DI (1949) Plant Physiol 24:1–15

Baltimore D (1966) J Mol Biol 18:421–428

Berk AJ, Sharp PA (1977) Cell 12:721–732

Boguski MS, Hieter PA, Levy CC (1980) J Biol Chem 255: 2160–2163

Bradley D, Gatenby AA (1985) EMBO J 4:3641–3648

Briat J-F, Bisanz-Seyer C, Lescure A-M (1987) Curr Genet 11: 259–263

Chelm BK, Hoben PJ, Hallick RB (1977) Biochem 16:776–781

Donis-Keller H, Maxam AM, Gilbert W (1977) Nucleic Acids Res 4:2527–2538

Donis-Keller H (1980) Nucleic Acids Res 8:3133–3142

El-Gewely MR, Helling RB (1980) Anal Biochem 102:423–428

El-Gewely MR, Helling RB, Dibbits JGT (1984) Mol Gen Genet 194:432–443

Gausing KJ (1977) J Mol Biol 115:335–354

Gold B, Carillo N, Tewari KK, Bogorad L (1987) Proc Natl Acad Sci USA 84:194–198

Graf L, Roux E, Stutz E, Kossel H (1982) Nucleic Acids Res 10:6369–6381

Graves MC, Breitenberger CA, Spremulli LL (1980) Arch Biochem Biophys 204:444–454

Gray WM (1983) BioScience 33:693–699

Green MR, Roeder RG (1980) Cell 22:231–242

Greenberg BM, Hallick RB (1986) Plant Mol Biol 6:89–100

Girvitz SC, Bacchetti S, Rainbow AJ, Graham FL (1980) Anal Biochem 106:492–496

Gruissm W, Zurawski G (1985) EMBO J 4:3375–3383

Gruissem W, Elsner-Menzel C, Latshaw S, Narita JO, Schaffer MA, Zurawski G (1986) Nucleic Acids Res 14:7541–7556

deHaas JM, Kool AJ, Overbeeke N, vanBrug W, Nijkamp HJJ (1987) Curr Genet 12:377–386

Hanley-Bowden L, Orozco EM Jr, Chua N-N (1985) Mol Cell Biol 5:2733–2745

Hartley MR, Head C (1979) Eur J Biochem 96:301–309

Heizmann P (1974) Biochimie 56:1357–1364

Helling RB, El-Gewely MR, Lomax MI, Baumgartner JE, Schwartzbach SD, Barnett WE (1979) Mol Gen Genet 174: 1–10

Hollingsworth MJ, Hallick RB (1982) J Biol Chem 257:12795–12799

Jinks-Robertson S, Nomura M (1987) Ribosomes and tRNA. In: Neidhardt F (ed) *Escherichia coli* and *Salmonella typhimurium.* Am Soc Microbiol, Washington DC, pp 1358–1385

Keus RA, Dekker AF, Kreuk KCJ, Groot GSP (1984) Curr Genet 9:91–97

Koller B, Delius H (1982a) EMBO J 1:995–998

Koller B, Delius H (1982b) FEBS Lett 139:86–92

Koller B, Delius H, Helling RB (1984) Plant Mol Biol 3:127–136

Koller B, Roux E, Montandon P-E, Stutz E (1988) Plant Mol Biol 10:339–347

Kung SD, Lin CM (1985) Nucleic Acids Res 13:7543–7549

Lindahl L, Zengal JM (1986) Annu Rev Genet 20:297–326

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Maxam AM, Gilbert (1980) Methods Enzymol 65:499–560

McClure WR (1985) Annu Rev Biochem 54:171–204

Morgan EA (1986) J Bacteriol 168:1–5

Moss B (1981) In: Chirikjian JG, Papas TS (eds) Gene amplification and analysis, vol 2. Elsevier/North-Holland, Amsterdam New York, pp 254–265

Mulligan RM, Malony AP, Walbot V (1988) Mol Gen Genet 211:373–380

Narita JO, Rushlow KE, Hallick RB (1985) J Biol Chem 260: 11195–11199

Nomura M, Gourse R, Baughman G (1984) Annu Rev Biochem 53:75–117

Ohyama K, Fukusawa H, Kohchi T, Schrai H, Sano S, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H (1986) Nature 322:572–574

Orozco EM, Rushlow KE, Dodd JR, Hallick RB (1980) J Biol Chem 255:10997–11003

Palmer JD (1985) Annu Rev Genet 19:325–354

Ravel-Chapius P, Heizmann P, Nigon V (1982) Nature 300:78–81

Reznikoff WS, Siesulu DA, Cowing DA, Gross CA (1985) Annu Rev Genet 19:355–387

Rochaix J-D, Darlix J-L (1982) J Mol Biol 159:383–395

Roux E, Graf L, Stutz E (1983) Nucleic Acids Res 11:1957–1968

Roux E, Stutz E (1985) Curr Genet 9:221–227

Schiff JA, Lyman H, Russell GK (1971) Methods Enzymol 23: 143–168

Schlunegger B, Stutz E (1984) Curr Genet 8:629–634

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Tamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Tamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) EMBO J 5:2043–2049

Southern E (1975) J Mol Biol 98:503–517

Strittmaller G, Godzicka-Jozefiak A, Kössel H (1985) EMBO J 4:599–604

Sun E, Shapiro PR, Wu BW, Tewari KK (1986) Plant Mol Biol 6:429–439

Waddel J, Wang XM, Wu M (1984) Nucleic Acids Res 12:3843–3856

Weislander L (1979) Anal Biochem 98:305–309

Williamson RM, Jackson JH (1987) J Bacteriol 169:2893–2895

Zablen LB, Kissil MS, Woese CR, Buetow DE (1975) Proc Natl Acad Sci USA 72:2418–2422