

## Processing of chloroplast ribosomal RNA transcripts in *Euglena gracilis bacillaris*

Peter McGarvey and Robert B. Helling

Department of Biology, University of Michigan, Ann Arbor, MI 48109, USA

**Summary.** The ribosomal RNA operons (*rrn* operons) of *Euglena gracilis* chloroplasts contain genes for (in order) 16S rRNA, tRNA<sup>Ile</sup>, tRNA<sup>Ala</sup>, 23S rRNA and 5S rRNA. Major sites of cleavage of the primary *rrn* transcript were identified by Northern blot hybridization and S1-mapping. The presumptive termini of all of the mature products have now been identified. During initial processing in the chloroplast, the primary transcript is cleaved between the two tRNAs and between the 23S and 5S rRNAs so as to separate the sequences found in the different mature rRNAs. Subsequently the tRNAs are separated from the rRNAs, further trimming provides the remaining proper ends, and the 3'-ends of the tRNAs are added.

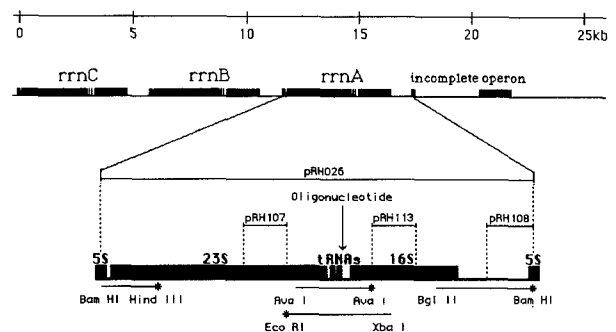
**Key words:** *Euglena gracilis* – rRNA transcription – Chloroplasts

### Introduction

The ribosomal RNAs (rRNAs) are of fundamental importance in the growth of all cells and organelles. The rRNAs are usually made from genes located in a common transcriptional unit. Transcripts of these *rrn* operons are synthesized as long precursors which are processed into the smaller mature rRNAs and tRNAs. In prokaryotes and organelles these cleavages occur so rapidly that full-length transcripts are present in low concentration if at all (Gegenheimer and Apirion 1981). However partially processed transcripts have been observed in a variety of organisms.

The chloroplast genomes of *E. gracilis* strains B (*bacillaris*) and Z each contain three complete and one incomplete *rrn* operons arranged in tandem (Koller et

al. 1984). Like three of the seven *E. coli* *rrn* operons, the complete operons of *Euglena* contain (in order of transcription) genes for 16S rRNA, tRNA<sup>Ile</sup>, tRNA<sup>Ala</sup>, 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 The incomplete *rrn* operon retains complete 16S and 5S rRNA genes, but the middle has been replaced by a sequence containing an open reading frame that is cotranscribed with the extra 16S RNA (Koller et al. 1988; Roux et al. 1983; Roux and Stutz 1985). Transcription of the chloroplast *rrn* operons of *Euglena* originates from a single site 53 nucleotides upstream of the 16S rRNA and within a DNA sequence likely to have been derived from tRNA genes located elsewhere in the chloroplast genome (McGarvey et al. 1988). A 103 base-pair sequence, identical except for a single nucleotide to that flanking



**Fig. 1.** The arrangement of the *rrn* operons in the B and Z strains of *E. gracilis* (Koller et al. 1984). The complete operons are nearly 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. 1980; Orozco et al. 1980; Roux and Stutz 1985). The incomplete operon contains complete 16S and 5S rRNA genes but only 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 source. The bars above the expanded *rrn* operon indicate clones used as probes in the Northern blots discussed in the text. The position corresponding to the oligonucleotide used as a probe is also indicated. The lines below the expanded operon indicate the end-labeled probes used in S1-mapping of the ends of *rrn* transcripts

the *rrn* transcript's origin, is present near the origin of replication, suggesting that the sequence may be involved in the initiation of replication (McGarvey et al. 1988).

The 5'- and 3'-ends of the *Euglena* chloroplast 16S rRNA (El-Gewely et al. 1984; Steege et al. 1982) and the 5' end of the 5S rRNA (Karabin et al. 1983) have been determined directly. Possible end sequences of the other mature products of the *rrn* operons and their genes have been proposed (see below). However, these proposals are not in agreement and were made on the basis of sequence similarity with the corresponding *Escherichia coli* genes rather than by direct determination.

The *Euglena* chloroplast *rrn* operons are of exceptional interest because of their unusual organization, the inclusion in the incomplete operon of a gene that may encode a protein, the wealth of knowledge on the origin of the leader sequence, and the possibility that a similar sequence may play a role in replication. We should like to understand in detail the control of expression and the evolution of the *rrn* operons and related sequences. Therefore, we have mapped the ends of the mature rRNA molecules and have identified several other points at which precursor RNA is cleaved during maturation.

## Materials and methods

**Preparation of *Euglena* RNA.** Total cell RNA was prepared from *E. gracilis bacillaris* cells grown in continuous light or continuous dark (referred to as light-grown and dark-grown RNA, respectively) and from light-grown cells treated with chloramphenicol or cycloheximide as previously described (McGarvey et al. 1988).

**Preparation of rRNA from chloroplast ribosomes.** Chloroplast ribosomes were prepared as described previously (Graves et al. 1980; McGarvey et al. 1988). The ribosomal proteins were removed by extraction twice with hot phenol/Sevag (1:1) (Sevag = chloroform/isoamylalcohol [24:1]) and twice with Sevag, and the rRNA was precipitated with 0.3 sodium acetate, 2 volumes 95% ethanol and stored at  $-70^{\circ}\text{C}$ .

**Preparation and labeling of DNA.** Figure 1 shows a map of the *rrn* operons of *Euglena* chloroplast DNA and the location of the cloned DNA segments used as probes. Recombinant plasmids containing chloroplast DNA were constructed by ligation of chloroplast DNA restriction fragments into the plasmid vector RSF2124 or pBR322 (Lomax et al. 1977; Helling et al. 1979). The plasmids pRH026 and pRH022 contain respectively the BamE and BamF DNA segments corresponding to the *rrnA* and *rrnB* operons cloned into the BamHI site of pBR322. Plasmids pRH107 and pRH108 were produced by subcloning portions of pRH022 into pBR322. Plasmid pRH113 was produced by subcloning a segment of *rrnC* from plasmid pML12 (Lomax et al. 1977) into pBR322. Large- and small-scale plasmid DNA preparations were made as described (El-Gewely and Helling 1980; Maniatis et al. 1982). Enzymes were obtained from Bethesda Research Laboratories (BRL), Boehringer Mannheim Biochemicals, International Biochemicals Inc. (New Haven, Conn.) and New

England Biolabs (Beverly, Mass.), and used according to the manufacturer's directions. Plasmid DNA and DNA fragments used as size standards were radiolabeled by nick-translation (Rigby et al. 1977) using [ $\alpha$ - $^{32}\text{P}$ ]dCTP and a kit (BRL) according to the manufacturer's directions.

**End-labeling of DNA.** 5'-Ends were labeled at specific endonuclease cleavage sites using [ $\gamma$ - $^{32}\text{P}$ ]ATP and T4 polynucleotide kinase as described elsewhere (McGarvey et al. 1988). 3'-Ends were labeled as follows: Ten to fifty micrograms of plasmid DNA were cut with a restriction enzyme, leaving a recessed 3'-OH. After the DNA was cut to completion, 10–50  $\mu\text{Ci}$  of [ $\alpha$ - $^{32}\text{P}$ ]dCTP or [ $\alpha$ - $^{32}\text{P}$ ]dATP and 0.2 mM each of dGTP, dTTP, and dATP or dCTP were added to the reaction mixture. Five units of DNA polymerase I (Klenow fragment, BRL) were added and the mixture was incubated at room temperature for 40 min. The reaction was stopped with 2  $\mu\text{l}$  of 0.5 M EDTA. After labeling, the 5'- or 3'-end-labeled fragments were cut with a second restriction endonuclease. The end-labeled DNA fragments were separated by electrophoresis through an agarose gel and the appropriate fragment was removed either by electroelution into a dialysis membrane (Girvitz et al. 1980), or extraction from low melting temperature agarose (Weislander 1979).

**Oligonucleotide preparation and labeling.** A 20 base deoxyoligonucleotide (5'-CCTCTTGGAAAAATAAGAGG-3'), identical in sequence to the first two nucleotides of the antisense strand of the tRNA<sup>Leu</sup> gene found in each complete *rrn* operon and the 18 nucleotides preceding it, was prepared using the phosphoramidite method in an Applied Biosystems 380A synthesizer at the oligonucleotide synthesis facility of the Center for Molecular Genetics at The University of Michigan. The oligonucleotide was 5'-end-labeled using T4 polynucleotide kinase (BRL) and separated from unincorporated [ $\gamma$ - $^{32}\text{P}$ ]ATP by electrophoresis through a 20% polyacrylamide gel. The radioactive bands was cut from the gel and the oligonucleotide allowed to diffuse out of the gel slice into 1–2 ml of water overnight at  $4^{\circ}\text{C}$ .

**S1-mapping.** Hybridization and S1-mapping were carried out as described previously (McGarvey et al. 1988).

**Preparation and hybridization of Northern blots.** RNA samples containing 10–20  $\mu\text{g}$  RNA were heated for 10 min at  $60^{\circ}\text{C}$  in 20  $\mu\text{l}$  2.2 M formaldehyde, 50% formamide, 0.018 M  $\text{Na}_2\text{HPO}_4$ , and 0.002 M  $\text{NaH}_2\text{PO}_4$  (pH 7.0). The samples were mixed with 2  $\mu\text{l}$  loading buffer (50% glycerol, 1 mM EDTA, 0.4% bromophenol blue, 0.4% xylene cyanol) and loaded on a horizontal 20 cm  $\times$  12 cm 1.5% agarose gel prepared as described previously (Lehrach et al. 1977). Electrophoresis was carried out at 50 volts in buffer containing 2.2 M formaldehyde, 0.018 M  $\text{Na}_2\text{HPO}_4$ , 0.002 M  $\text{NaH}_2\text{PO}_4$  (pH 7.0) until the bromophenol blue dye had moved at least 75% of the length of the gel. HindIII restriction fragments of lambda DNA (BRL) used as molecular weight standards were treated as above. After electrophoresis, the gel was rinsed three times in deionized water to remove excess formaldehyde. The RNA was transferred to a Hybond-N membrane (Amersham) by a capillary blot (Southern 1975) using  $10\times$  SSC (SSC = 150 mM NaCl, 15 mM sodium citrate). The membrane was air-dried, and the nucleic acids were covalently bound to the membrane by irradiation with long wavelength ultraviolet light (340 nm) for 5 min (Church and Gilbert 1984). The membrane was then baked at  $80^{\circ}\text{C}$  for 2 h in a vacuum oven to remove any remaining formaldehyde. Membranes were pre-incubated in hybridization buffer (1 M NaCl, 1% sodium dodecylsulfate, 10% dextran sulfate) for 2–6 h at  $60^{\circ}\text{C}$ . The nick-translated probes (approximately  $1\times 10^6$  cpm each) were denatured by boiling for 10 min in 2–5 ml hybridization buffer containing enough salmon sperm DNA to give a final concentration of 200  $\mu\text{g}/\text{ml}$  when added to the prehybridization mixture. Hybridization was performed at  $65^{\circ}\text{C}$

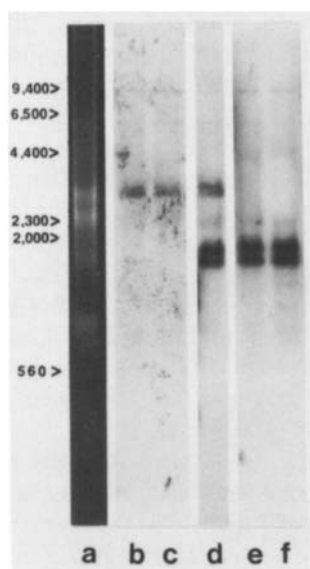


Fig. 2. Northern blot hybridization using probes specific for the *rrn* operons. Lanes a, ethidium-bromide-stained agarose gel of light-grown total RNA; lanes b and c, Northern blot of 10 µg dark-grown total RNA probed with pRH107 (23S rRNA); Lane d, Northern blot of 10 µg dark-grown total RNA probed with pRH026 (*rrnA*); lanes e and f, Northern blot of 10 µg dark-grown total RNA probed with pRH113 (16S rRNA)

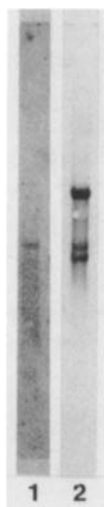


Fig. 3. Northern blot hybridization using an *rrn* spacer region oligonucleotide probe. Lane 1, 10 µg light-grown total cell RNA probed with oligonucleotide; lane 2, the same membrane washed and reprobed with pRH022 (*rrnB*)

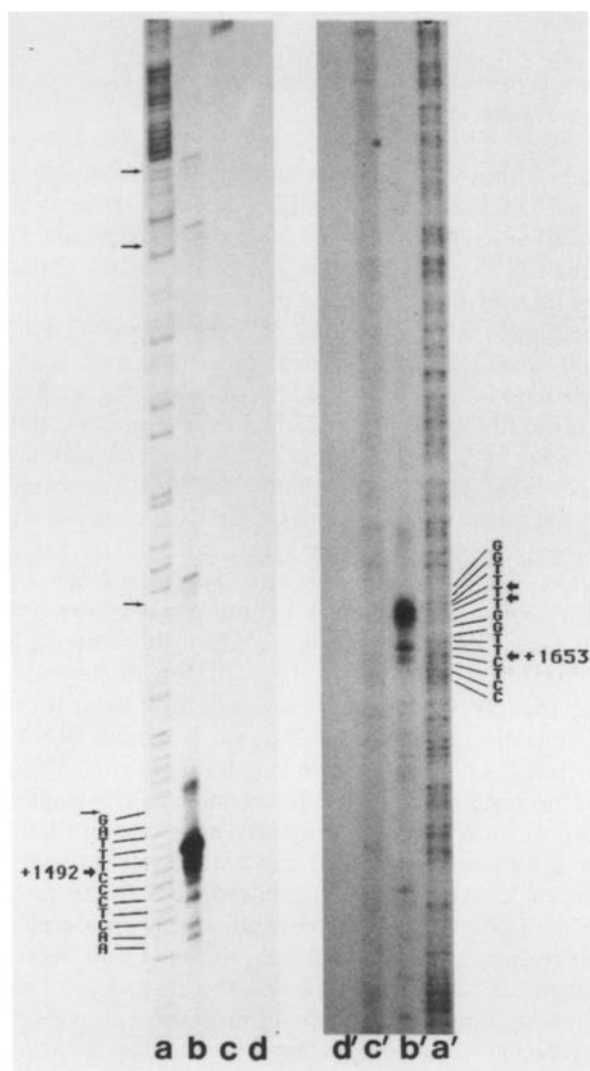
for 12–16 h, after which the samples were washed for 2 × 5 min in 2 × SSC at room temperature, 2 × 30 min in 2 × SSC, 1% SDS at 65°C and 2 × 30 min in 0.1 × SSC at 65°C. Hybridization with the 5′-end-labeled oligonucleotide ( $2\text{--}3 \times 10^6$  cpm) was performed at 40°C as described previously (Dalbadie-McFarland et al. 1982). The membranes were air-dried and hybridization was visualized by autoradiography using Kodak XAR-5 film and Cronex intensifying screens (Dupont, Wilmington, DE) at  $-70^\circ\text{C}$ .

## Results

### *Analysis of rrn transcription using Northern blot hybridization*

Figure 2 shows dark-grown total RNA probed with a cloned DNA segment containing an entire *rrn* operon (lane d) or a segment specific for the 16S (lanes e and f) or the 23S (b and c) rRNAs. The probe for the entire operon hybridizes in three prominent bands corresponding to RNA molecules of approximately 2,800, 1,700, and 1,500 nucleotides. The 2,800 and 1,500 nucleotide RNAs correspond in size to the mature 23S and 16S rRNAs (Graf et al. 1982; Koller et al. 1984) and hybridize with the 23S and 16S rRNA-specific probes, respectively. The band containing the 2,800 nucleotide RNA is a doublet that includes 23S rRNA and a closely following longer precursor (see farther). The 1,700 nucleotide RNA hybridizes with the probe from the 16S rRNA gene (lanes e and f) but not with the 23S rRNA-specific probe (lanes b and c). In addition, a 2,300 nucleotide RNA hybridizes faintly with the probe from the 16S rRNA gene. This results most likely from hybridization to the cytoplasmic small subunit rRNA, which is 2,305 nucleotides in length (Sogin et al. 1986) and has been shown to hybridize with the chloroplast 16S gene on Southern blots (Curtis and Rawson 1982). The 2,300 nucleotide band was not seen in Northern blots of RNA from isolated chloroplasts (McGarvey 1988). Long exposures showed higher molecular weight bands with all probes up to about 6,000 nucleotides, the length expected of a transcript of an entire *rrn* operon. The faint hybridization at about 9 kb may result from a small amount of contaminating DNA and was not usually seen. Northern blots using RNA prepared from cells grown in the light (Fig. 3) and from cells treated with cycloheximide or chloramphenicol showed similar patterns (McGarvey 1988). Although these experiments were carried out with the *bacillaris* strain, we have shown that the Z strain gives a similar Northern blot profile, including a 1,700 nucleotide band that hybridizes with the 16S rRNA specific probe (not shown).

Previous experiments using S1-mapping and 5′-capping with radiolabelled GTP (McGarvey et al. 1988) revealed that *rrn* transcripts originated only 53 bases upstream of the mature 16S rRNA 5′-end, and thus a precursor extending from the transcript origin to the 3′-end of the 16S rRNA alone could not account for the size of the RNA in the 1,700 nucleotide band. A longer 5′-leader on a transcript of the extra 16S gene was also unlikely, since its leader region DNA is almost identical with that of the complete *rrn* operons for the first several hundred bases (Roux et al. 1983; El-Gewely et al. 1984). In order to determine whether,



**Fig. 4.** Determination of the 3'-ends of transcripts containing 16S rRNA. Autoradiogram of a 6% polyacrylamide gel from an S1-experiment in which an approximately 800 bp fragment 3'-end-labeled at an *Ava*I site in the 16S gene (nucleotide +1339; Graf et al. 1982) was used as a probe. Lanes contained: *a* and *a'*, A+G sequencing reaction products from labeled probe (Maxam and Gilbert 1980); *b* and *b'*, complete reaction mixture with 1.0  $\mu$ g of light-grown total RNA; *c* and *c'*, control, complete hybridization mixture not digested with S1; *d* and *d'*, control without RNA. The sequence shown is that of the RNA-like strand. Arrows point to the bonds cleaved by S1. The ends created by S1-cleavage and chemical cleavage are not identical (Maxam and Gilbert 1980; Green and Roeder 1980), and a 1.5 nucleotide downshift of the S1-band is required to align it with the corresponding A+G band (McGarvey et al. 1988). The *C* at +1492 (lanes *a*-*d*) corresponds to the 3'-end of the 16S rRNA (Steege et al. 1982). The *C* at +1653 (lanes *a'*-*d'*) corresponds to the 3'-end of the tRNA<sup>Ile</sup>, 161 nucleotides beyond the 3'-end of the 16S rRNA gene

instead, the extra 200 nucleotides present in the 1,700 nucleotide band were from the 16S to 23S spacer region 3' of the 16S rRNA, a 20-base oligonucleotide complementary to a portion of this region and beginning 69 nucleotides beyond the 16S rRNA coding sequence (see Fig. 1) was synthesized and hybridized to a Northern blot of total cell RNA. The oligonucleotide hybridized to the 1,700 base band, showing that the band represents 16S rRNA with a 3'-extension into the spacer region (Fig. 3). Possibly two or more families of rRNA which differ in their 3'-ends are found in this band, as the sequence following the extra 16S gene diverges sharply from that following the other 16S rRNA genes (Roux et al. 1983; El-Gewely et al. 1984).

A 3.6-kb transcript including the extra 16S rRNA and adjacent open reading frame has been demonstrated but is at such low frequency relative to the mature rRNA that it cannot be detected at the low film exposure of this figure (Koller et al. 1988). The same workers were unable to detect smaller transcripts containing the open reading frame.

The 5S rRNA was not seen by Northern blot hybridization of any of the RNA preparations. It has been reported that small stable RNAs such as tRNAs and 5S rRNAs are lost in isolation procedures that employ 4 M guanidine isothiocyanate (Bird 1986). However preparations of RNA in which guanidine isothiocyanate was not used were also depleted of 5S rRNA and tRNAs as judged by ethidium bromide fluorescence and Northern blots (not shown). Use of LiCl, which selectively precipitates single-stranded nucleic acids, may have selected against small RNAs that contain substantial secondary structure. Nevertheless 5S rRNA was detected in S1-mapping experiments using the same RNA preparations (see below). This RNA may not have been seen in blot hybridization because small RNAs may not have bound well to the membrane. Northern blot hybridization with a probe containing most of the 5S rRNA gene showed no hybridization with high molecular weight RNA, thereby demonstrating that few if any transcripts contained both 5S and 23S rRNA (not shown).

#### *Analysis of 3'-ends of transcripts containing 16S rRNA*

The positions in DNA corresponding to the ends of RNA molecules can be determined by nuclease S1 mapping (see McGarvey et al. 1988). S1-Mapping involves hybridizing RNA with complementary DNA and determining the extent to which the RNA or DNA is protected from degradation by a single-strand-specific nuclease such as S1. An approximately 800 bp DNA fragment 3'-end-labeled at an *Ava*I site in the 16S rRNA gene and extending within the 23S rRNA gene

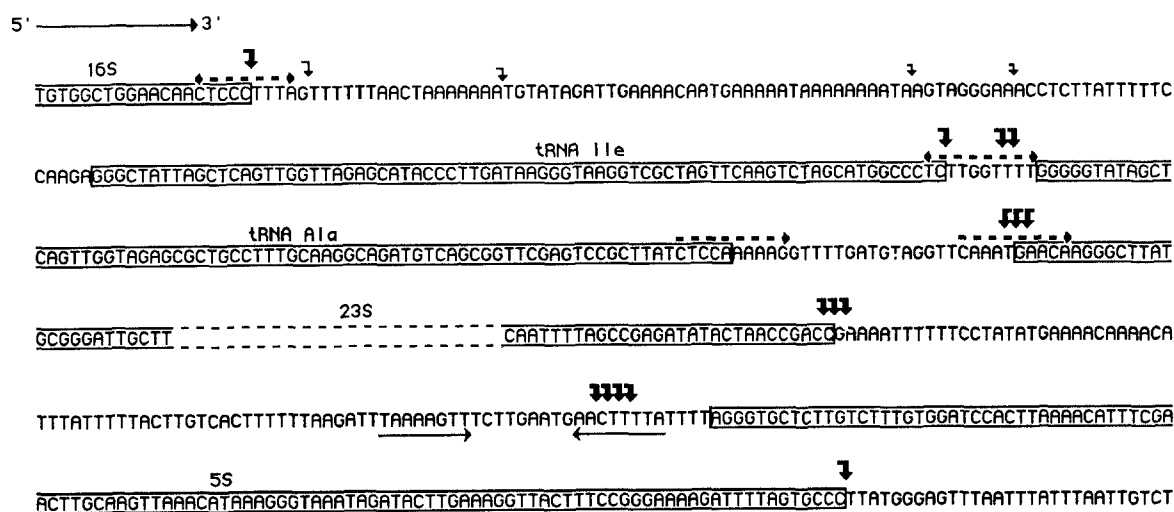


Fig. 5. Points at which the rRNA transcript is processed. The DNA sequence shown includes the 16S-23S spacer, the 23S-5S spacer and the sequence downstream of the 5S gene, and is that of the RNA-like strand. The 16S to 23S rRNA segment is from the Z strain (Graf et al. 1980; Orozco et al. 1980), and the rest is from the B strain *rrnA* operon (El-Gewely et al. 1984). The *thick arrows* designate cleavage points represented by prominent bands in S1-experiments. *Arrows with tails extending left* correspond to 3'-ends; *arrows with tails extending right*, to 5'-ends. The *thin arrows* correspond to less intense bands. *Dashes* above the sequence with *arrows pointing right* show regions corresponding to bands seen on low-resolution gels when mapping 3'-ends; *arrows pointing left* correspond to 5'-ends. *Arrows under* the sequence designate inverted repeat sequences. Only S1-cleavages that occurred regularly are represented. S1-cleavages that did not occur reproducibly are assumed to be artifacts

(see Fig. 1) was hybridized with total RNA. After treatment with nuclease S1 the protected DNA fragments were sized by electrophoresis through a polyacrylamide gel. S1 digestion often resulted in a set of adjacent DNA bands clustered around the end of the RNA:DNA hybrid molecule. In experiments using RNAs whose ends have been determined directly, we found that the most prominent band or the longer of two equally prominent adjacent bands corresponded to the true end of the RNA and we assigned the ends of mature rRNAs accordingly.

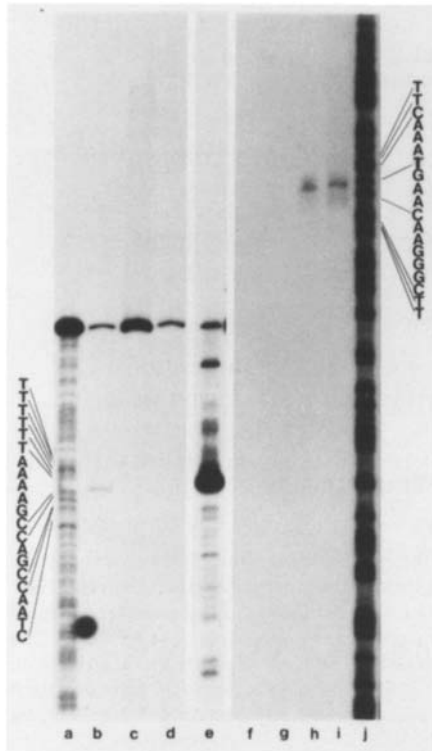
Several bands may be seen in Fig. 4. The prominent band in lane b indicates that a DNA fragment of 158 nucleotides was protected and thus places the 3'-end of the 16S rRNA transcript at nucleotide 1492 (Graf et al. 1982; Steege et al. 1982; McGarvey et al. 1988). Experiments using RNA purified from chloroplast ribosomes showed a similar band but no bands farther downstream (McGarvey 1988). A second set of prominent bands was seen using total cell RNA and indicated RNAs terminating 161, 166, and 167 nucleotides past the 16S rRNA (Fig. 4, lane b'). These ends occurred at the 3'-end of tRNA<sup>Ile</sup> (after nucleotide 1653) and within the eight nucleotide spacer region between tRNA<sup>Ile</sup> and tRNA<sup>Ala</sup> (Fig. 5). Other faint but reproducible bands shown by thin arrows (Fig. 4) indicated RNA with 3'-ends in the 16S rRNA to tRNA<sup>Ile</sup> spacer region.

Transcripts containing the 16S rRNA and tRNA<sup>Ile</sup>, with 3'-ends in the tRNA<sup>Ile</sup>-tRNA<sup>Ala</sup> spacer (+1653 to +1659) and with 5'-ends at the 16S rRNA 5'-end (+1)

or at the transcript origin (−53), should range in length from 1653 to 1712 nucleotides. These incompletely processed transcripts together with a possible transcript intermediate from the extra 16S rRNA gene (see earlier) account for the 1700 nucleotide band seen in the northern blots (Fig. 2).

#### *Analysis of the 5'-end of transcripts containing 23S rRNA*

Figure 6f–j shows the results from an S1 experiment using a 5'-end-labeled XbaI-EcoRI DNA fragment (Fig. 1) to determine the 5'-ends of the 23S rRNA-containing transcripts. Experiments using rRNA from purified chloroplast ribosomes (lane h) and cycloheximide-treated total RNA (lane i) each show a single band representing the 5'-end of the 23S RNA. On long exposure, bands farther upstream corresponding to RNAs with 5'-ends between the two spacer tRNAs and at the 3'-end of the 16S rRNA were observed in the lane with total RNA (not shown). Figure 5 shows the DNA sequence of the 5'-end of the 23S rRNA and indicates the position of the S1 bands observed. The resolution of the gel shown and of other gels is sufficient to determine that the 5'-end of the 23S rRNA is the G designated in Fig. 5 or one of its flanking nucleotides. The G is 9 bp from the 5'-end proposed on the basis of similarity to the *E. coli* 23S rRNA sequence (Orozco et al. 1980), and 7 bp from the 5'-end corresponding to

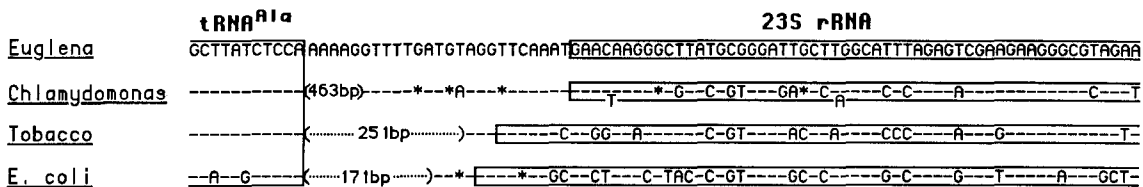


**Fig. 6.** Determination of the 3'- and 5'-ends of transcripts containing 23S rRNA. Autoradiogram of a 6% polyacrylamide gel from an S1-experiment. A HindIII-BamHI fragment approximately 530 bp long and 3'-end-labeled at the HindIII site inside the 23S gene was used as probe in lanes a-e. Lanes contained: a, A+G sequencing reaction products from labeled probe; b, complete reaction mixture containing 3 µg dark-grown total RNA; c, control, complete hybridization mixture not digested with S1; d, control without RNA; e, complete reaction mixture containing 5 µg chloramphenicol-treated total RNA. An approximately 1,100 bp XbaI-EcoRI fragment, 5'-end-labeled at the EcoRI site inside the 23S gene, was used as probe in lanes f-j. Lanes contained: f, control without RNA; g, control, complete hybridization mixture not digested with S1; h, complete reaction mixture containing 0.3 µg chloroplast rRNA; i, complete reaction mixture containing 10 µg cycloheximide-treated total RNA; j, A+G sequencing reaction products from labeled probe. The sequences shown are those of the RNA-like strand

that identified by S1-mapping in tobacco (Takaiwa and Sugiura 1982), but gives a 5'-terminus identical with that of the 5'-end of the 7S chloroplast rRNA of *Chlamydomonas reinhardtii* (Rochaix and Darlix 1982) (Fig. 7). The 7S rRNA is believed to correspond to the 5'-end of the other large subunit rRNAs.

*Analysis of the 3'-end of transcripts containing 23S rRNA*

Figure 6a-e shows the results from an S1 experiment using a 3'-end-labeled HindIII-BamHI DNA fragment (Fig. 1) to determine the 3'-end of the 23S rRNA. The resolution of this and other gels was sufficient to determine that the 3'-end is one of three nucleotides (lanes b and e). (The highest band in each lane, a-e, represents the full length DNA probe.) As with the 5'-end, the observed 3'-end of the 23S rRNA differs from the possible termini proposed on the basis of DNA sequence similarity with *E. coli* (El-Gewely et al. 1984; Karabin et al. 1983; Roux and Stutz 1985). Other bands corresponding to 3'-ends inside the 23S to 5S spacer were seen on long exposures (Fig. 6, lane e). The most prominent of these bands represents an RNA with a 3'-end about 10-12 nucleotides upstream from the 5S gene and within a region containing inverted repeat sequences that could form a stem-loop structure (see DNA sequence in Fig. 5). Such structures have been suggested as recognition structures in rRNA processing reactions (King et al. 1986). The broad band just above the overexposed band corresponding to the 3'-end of 23S rRNA (lane e) results from cleavage at a number of sites within an AT-rich sequence. We cannot tell whether these ends were formed from processing or result from the S1 procedure itself. No band corresponding to an RNA extending to the 5'-end of the 5S rRNA was observed.



**Fig. 7.** Comparison of the 5'-ends of the 23S rRNA genes of the *E. gracilis* chloroplast (Graf et al. 1980; Orozco et al. 1980), tobacco chloroplast (Takaiwa and Sugiura 1982) and *E. coli* (Brosius et al. 1981) and the 7S rRNA of the *Chlamydomonas reinhardtii* chloroplast (Rochaix and Darlix 1982). Only the RNA-like strand is shown. Dashes indicate bases identical to that found in *Euglena*; asterisks, deletions; brackets and subscripts, insertions



**Fig. 8.** Determination of the 3'-end of transcripts containing 5S rRNA. Autoradiogram of an 8% polyacrylamide gel from an S1-experiment. A 1,359 bp BamHI-BglII fragment 3'-end-labeled at the BamHI site in the 5S gene was used as probe. *Lanes* contained: *a*, control, complete hybridization mixture not digested with S1; *b*, complete hybridization mixture containing 5  $\mu$ g chloramphenicol-treated total RNA; *c*, A+G sequencing reaction products of labeled probe; *d*, control without RNA. The sequence shown is that of the RNA-like strand. A 1.5 nucleotide downshift of the S1-band is required to align it with the corresponding A+G band

#### *Analysis of the 3'-end of transcripts containing 5S rRNA*

Figure 8 shows the results from an S1 experiment using a 3'-end-labeled BamHI-BglII DNA fragment as probe and total chloramphenicol-treated RNA to determine the 3'-end of the 5S rRNA. Experiments with rRNA from purified chloroplast ribosomes showed identical results (not shown). Karabin et al. (1983) sequenced the 5'-end of the 5S rRNA, and the combined data (see Fig. 5) show that the mature 5S rRNA is 113 nucleotides long rather than the 116 nucleotides proposed on the basis of DNA sequence homology. We observed no reproducible band that might indicate a transcript extending downstream of the 3'-end of the 5S rRNA.

#### **Discussion**

Several laboratories have demonstrated rRNA precursors about 200 nucleotides larger than the mature 16S and 23S molecules (Carrit and Eisenstadt 1973; Heizmann 1974; Scott 1976; Schiemann et al. 1978). The rate of maturation of the 16S rRNA precursor is affected markedly by the culture conditions, becoming slow under conditions of energy depletion (Carrit and Eisenstadt 1973; Schiemann et al. 1978). Possibly rRNA maturation requires conditions for polysome formation as in *E. coli* (Srivastava and Schlessinger 1988).

Our work confirms the previous reports and extends them by mapping precisely the endpoints of many of these precursors, as well as of the mature

RNAs themselves. Using Northern blot hybridization and S1-mapping of transcripts we found that the major precursors of 16S rRNA observed had 3'-ends in the spacer region between tRNA<sup>Ile</sup> and tRNA<sup>Ala</sup>. Our Northern blots suggested the presence of a 23S rRNA precursor, although it migrated close to the 23S mature product. The existence of longer 23S rRNA-containing transcripts was revealed clearly by S1-mapping. These experiments showed such transcripts with 3'-ends resulting from cleavage in a sequence with the potential for forming a stem-loop structure and located just before the 5S rRNA (Figs. 5, 6). Transcripts containing 23S rRNA with 5'-ends in the region between the tRNAs were also found (Fig. 5). The length of a 23S rRNA-containing transcript extending between these positions would be about 180 nucleotides longer than the 23S rRNA, and most probably corresponds to the pre-23S rRNA observed by others.

This work has also defined several of the ends of the mature rRNAs and tRNAs. In each case in which the proposed end had been inferred only by comparison with the corresponding sequence of *E. coli*, the predicted end was in error. However in the case of the 5'-end of the chloroplast 23S rRNA, the end found by S1-mapping corresponds to the equivalent end point of the rRNA of *Chlamydomonas* chloroplasts (Fig. 7). The terminal 3'-CAA of tRNAs is often added posttranscriptionally, but we have shown that the first C is encoded in the *Euglena* tRNA<sup>Ile</sup> gene (Figs. 4, 5), just as it is in some tRNA genes of T4 (Pragai and Apirion 1982) and in some chloroplast tRNA genes of land plants (Wang et al. 1988). The mature tRNA is completed by the addition of single cytidylate and adenylate residues.

This work, together with the pulse and pulse-chase experiments of Carrit and Eisenstadt (1973), Heizmann (1974), and Scott (1976), suggests that the primary *rrn* transcript is cleaved rapidly in the spacers between the two tRNAs and between the 23S and 5S rRNA genes. Subsequent processing is slower, and two major rRNA precursors accumulate, one containing the 16S rRNA sequence and the other, the 23S rRNA sequence. In chloroplasts of other organisms, rapid cleavage at several sites between the large rRNA and the 5S rRNA separates the 5S rRNA-containing sequences from the other *rrn* RNAs, and the mature 5S rRNA may be released very quickly (Strittmatter and Kössel 1984; Audren et al. 1987). A similar rapid separation of the 5S and the 23S rRNA sequences seems to occur in *Euglena* chloroplasts because we were unable to detect transcripts containing both either by Northern blot hybridization or by S1-mapping. Thus the first steps in rRNA processing separate the sequences found in the different mature rRNAs from each other. Subsequently the tRNAs appear to be separated

from the rRNAs, further trimming provides the remaining proper ends, and the 3'-ends of the tRNAs are added.

No evidence for transcription terminating beyond the 5S gene was observed. Though sequences analogous to rho-independent terminators have been found in the DNA 3' of many chloroplast genes, including the *Euglena* 5S rRNA genes (El-Gewely et al. 1984), termination either in vivo or in vitro has not been observed at these sites (Deng and Gruissem 1987). There is evidence that the 3'-inverted repeats found in chloroplast mRNAs function as processing and stabilizing elements and not as termination signals (Stern and Gruissem 1988). It is likely that processing of the 3'-end of the 5S rRNA is so fast that an unprocessed transcript extending across this site would not be observed by the procedures used here. Such precursors are not observed in wild-type *E. coli* cells because of the rapid processing of the 5S rRNA (King et al. 1986).

*Acknowledgements.* We thank J. Palmer and A. S. Sussman for their criticism, and Debora Hall for her assistance with manuscript preparation. Supported by NIH grant GM25565, an NIH Biomedical Support Grant, Rackham Graduate Student Support Grants and an NIH Genetics Traineeship (GM07544) to PM.

## References

- Audren H, Bisanz-Seyer C, Briat J-F, Mache R (1987) *Curr Genet* 12:263-269
- Bird RC (1986) *Biotechniques* 4:202-204
- Brosius J, Dull TJ, Steeler DD, Noller HF (1981) *J Mol Biol* 148:107-127
- Carrit B, Eisenstadt J (1973) *FEBS Lett* 36:116-119
- Church GM, Gilbert W (1984) *Proc Natl Acad Sci USA* 81:1991-1995
- Curtis SE, Rawson JRY (1982) *Plant Physiol* 69:67-71
- Dalbadie-McFarland GL, Cohen LW, Riggs AD, Morin C, Itakura K, Richards JH (1982) *Proc Natl Acad Sci USA* 79:6409-6413
- Deng X-W, Gruissem W (1987) *Cell* 49:379-387
- El-Gewely MR, Helling RB (1980) *Anal Biochem* 102:423-428
- El-Gewely MR, Helling RB, Dibbitts JGT (1984) *Mol Gen Genet* 194:432-443
- Gegenheimer P, Apirion D (1981) *Microbiological Rev* 45:502-541
- Girvitz SC, Bacchetti S, Rainbow AJ, Graham FL (1980) *Anal Biochem* 106:492-496
- Graf L, Kossel H, Stutz E (1980) *Nature* 286:908-910
- Graf L, Roux E, Stutz E, Kössel H (1982) *Nucleic Acids Res* 10:6369-6381
- Graves MC, Breitenberger CA, Spremulli LL (1980) *Arch Biochem Biophys* 204:444-454
- Green MR, Roeder RG (1980) *Cell* 22:231-242
- Heizmann P (1974) *Biochem Biophys Res Commun* 56:112-118
- Helling RB, El-Gewely MR, Lomax MI, Baumgartner JE, Schwartzbach SD, Barnett WE (1979) *Mol Gen Genet* 174:1-10
- Karabin GD, Narita JO, Dodd JR, Hallick RB (1983) *J Biol Chem* 258:14790-14796
- King T, Sirdeskumukh R, Schlessinger D (1986) *Microbiol Rev* 50:428-451
- 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
- Lehrach H, Diamond D, Wozney JM, Boedtner H (1977) *Biochemistry* 16:4743-4751
- Lomax ML, Helling RB, Hecker LI, Schwartzbach SD, Barnett WE (1977) *Science* 196:202-205
- Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular cloning*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Maxam AM, Gilbert W (1980) *Methods Enzymol* 65:499-560
- McGarvey P (1988) PhD dissertation, University of Michigan
- McGarvey P, Helling RB, Lee JY, Engelke DR, El-Gewely MR (1988) *Curr Genet* 14:493-500
- Orozco EM, Rushlow KE, Dodd JR, Hallick RB (1980) *J Biol Chem* 255:10997-11003
- Pragai B, Apirion D (1982) *J Mol Biol* 154:465-484
- Rigby PWJ, Dieckmann M, Rhodes C, Berg P (1977) *J Mol Biol* 113:237-251
- Rochaix J-D, Darlix J-L (1982) *J Mol Biol* 159:383-395
- Roux E, Stutz E (1985) *Curr Genet* 9:221-227
- Roux E, Graf L, Stutz E (1983) *Nucleic Acids Res* 11:1957-1968
- Schiemann J, Wollgiehn R, Parthier B (1978) *Biochem Physiol Pflanz* 172:507-519
- Scott NS (1976) *Phytochemistry* 15:1207-1213
- Southern EM (1975) *J Mol Biol* 98:503-517
- Sogin ML, Elwood HJ, Gunderson JH (1986) *Proc Natl Acad Sci USA* 83:1383-1387
- Srivastava AK, Schlessinger D (1988) *Proc Natl Acad Sci USA* 85:7144-7148
- Steege DA, Graves MC, Spremulli LL (1982) *J Biol Chem* 257:10430-10439
- Stern DB, Gruissem W (1987) *Cell* 51:11145-11157
- Strittmatter G, Kössel H (1984) *Nucleic Acids Res* 12:7633-7647
- Takaiwa F, Sugiura M (1982) *Eur J Biochem* 124:13-19
- Wang MJ, Davis NW, Gegenheimer P (1988) *EMBO J* 7:1567-1574
- Weislander L (1979) *Anal Biochem* 98:305-309

Communicated by W. Birky

Received December 22, 1988/February 21, 1989