
Detection of multiple, unspliced precursor mRNA transcripts for the M_r 32,000 thylakoid membrane protein from *Euglena gracilis* chloroplasts

Margaret J. Hollingsworth, Udo Johannngmeier, Gerald D. Karabin, Gary L. Stiegler* and Richard B. Hallick +

Department of Chemistry, University of Colorado, Boulder, CO 80309, USA

Received 11 October 1983; Revised and Accepted 24 January 1984

ABSTRACT

The *psbA* gene is the coding locus for a polypeptide of 32 kilodaltons that is involved in electron transport through photosystem II. The 4.9 kilobasepair (kbp) *EcoRI* restriction endonuclease fragment *EcoI* from the 145 kbp *Euglena gracilis* chloroplast DNA was shown to encode *psbA*. Five transcripts of size 3.1, 2.8, 2.3, 1.8, and 1.2 kilobases were detected by hybridization of *psbA* probes to nitrocellulose filter blots of electrophoretically separated RNAs. This same pattern was observed when the hybridization probe consisted of only exon sequences from this split gene. A synthetic, intron specific probe hybridized to all RNA precursors except the 1.2 kb mature RNA. These results and *psbA* DNA sequence data lead to the conclusion that the four higher molecular weight transcripts are unprocessed precursors of the 1.2 kilobase RNA, some of which contain unspliced intervening sequences. There is an increase in *psbA* transcripts during light induced maturation of the chloroplasts.

INTRODUCTION

The "32 kilodalton" polypeptide of chloroplast thylakoid membranes is a rapidly metabolized polypeptide thought to be a "proteinacious shield" associated with photosystem II.^{1,2,3} It appears to be involved in the transport of electrons from photosystem II, and may confer sensitivity to the herbicides diuron and atrazine in PS II.^{2,4} The gene coding locus (*psbA*) for the 32 kilodalton (kd) polypeptide is found in the chloroplast genome, and has been localized and identified in maize,⁵ spinach,⁶ *Nicotiana debneyi*⁷ and *Euglena*^{8,9} chloroplasts.

The *psbA* gene polypeptide product has been studied in many organisms. Comparisons have been made between the photogene protein from maize,¹⁰ peak D from peas,¹¹ protein D-1 from *Chlamydomonas*,¹² and the 32 kd polypeptide from *Spirodela*.² They were found to be virtually identical. Significant homologies have been found in the structural properties of the 32 kd protein of all angiosperms studied to date.¹³

The complete DNA sequence for the *psbA* gene is known in spinach, *N.*

debneyi,⁷ and *Amaranthus hybridus*.¹⁴ The amino acid sequence for the 32 kd protein derived from the gene sequence of spinach and *N. debneyi* is 100% conserved. There is one amino acid substitution in *A. hybridus*, and one additional change in the atrazine resistant biotype.¹⁴ In *Spirodela*, the 32 kd protein is synthesized in amounts comparable to the large subunit of ribulose-bisphosphate carboxylase. The protein has an extremely rapid turnover rate, 50 to 60 times more rapid than that of the large subunit.¹⁵ It appears that there is an abundance of the psbA mRNA, but comparatively low levels of the protein itself. The 32 kd polypeptide levels appear to be induced and controlled by light.^{2,16} This makes the psbA gene an especially good candidate for studies on gene regulation during light induced chloroplast development.

Euglena is a particularly suitable organism for the study of chloroplast encoded mRNAs. It may be grown in the absence of light, subsisting solely upon organic carbon sources. When grown in the dark, the chloroplasts revert to a proplastid state. Upon exposure to light, these proplastids develop into functional chloroplasts over a period of 72 hours.^{17,18} Changes in *Euglena* thylakoid polypeptides during chloroplast development have been described.^{19,20} *Euglena* RNA transcripts may be isolated at varying periods during this greening process. It has been shown that some transcripts decrease in abundance during development while others increase.^{21,22}

In this paper, we will report on our studies on the 4.9 kbp *EcoRI* restriction endonuclease fragment of *EcoI* of the 145 kbp *Euglena gracilis* chloroplast DNA. We have identified three gene loci within *EcoI*. One of these loci, the psbA gene, was chosen for extensive study. Through experiments involving the hybridization of radiolabeled psbA probes to cellulose nitrate filter blots of *Euglena* RNAs, we have detected at least five psbA transcripts. Studies were undertaken to determine the effect of light induced chloroplast development on these psbA gene transcripts.

MATERIALS AND METHODS

Isolation and Identification of Nucleic Acids

Chloroplast DNA was isolated in the intact superhelical form from *Euglena gracilis* Klebs, Strain Z Pringsheim as previously described.^{23,24} Spinach (*S. oleracea*) recombinant chloroplast PstI clones and total chloroplast DNA were generously provided by Drs. W. F. Thompson and J. D. Palmer.

Euglena chloroplast DNA restriction fragments were ligated into plasmid vector restriction sites by standard procedures.²⁵ pEZC514 consists of the

EcoI fragment from *Euglena* chloroplast DNA ligated into pBR322. pE2C25 consists of the *HindIII* 10 fragment, which is a 3.6 kilobasepair (kbp) internal fragment of *EcoI*, ligated into pBR325. Individual fragments from this plasmid were purified from polyacrylamide gel slices by a modification of a crush and soak procedure.²⁶

Whole cell *Euglena* RNA was purified, as previously described,²⁷ with aurintricarboxylic acid employed as a nuclease inhibitor. Prior to RNA extraction, aliquots of cells from the different developmental stages were stored at -70°C.

Restriction Endonuclease Mapping and Localization of psbA

Localization of psbA within *EcoI* was accomplished through a heterologous membrane filter hybridization approach. Restriction nuclease fragments were transferred from agarose gels to nitrocellulose filters.²⁸ The solution for heterologous hybridization reactions contained 0.75 M NaCl, 0.075 M sodium citrate, 15% formamide, 0.075M NaPO₄ (pH 7.0), 200 µg/ml degraded Herring Sperm DNA, 1 X Denhardt's solution,²⁹ and one or more psbA specific radio-labeled DNA probes. Hybridizations were carried out in a minimal volume in a Seal-a-Meal bag at 41°C for 24 hours. The filters were washed in the bag with three 50 ml washes of 0.75 M NaCl, 0.075 M sodium citrate, 0.1% sodium dodecyl sulfate, for 20 minutes each at room temperature.

RNA Electrophoresis, Transfer to Nitrocellulose, and Hybridization Conditions

RNA electrophoresis was carried out in 0.8% or 1% agarose, 2.2 M formaldehyde gels essentially as described by Rave et al.³⁰ Before transfer to nitrocellulose, one lane containing purified *Euglena* ct RNA was trimmed away from the rest of the gel, ethidium bromide stained, and photographed. The remainder of the gel was left unstained and transferred to nitrocellulose as described by Thomas.³¹ Hybridization was carried out in either 50% (for homologous reactions) or 15% (for heterologous reactions) formamide, 0.75 M NaCl, 0.075 M sodium citrate, 50 mM NaPO₄ (pH 7.0), 1 X Denhardts and 250 µg/ml degraded Herring Sperm DNA. Homologous reactions were carried out at 42°C. Heterologous reactions were at 37°C. Hybridization reactions involving ³²P-labeled synthetic oligonucleotides were for 3 days at 22°C in 0.6 M NaCl, 0.06 M Na citrate, 1 X Denhardts, and 200 µg/ml yeast RNA (Sigma, Type XI).

Preparation and Purification of ³²P-Labeled DNA Probes

Hybridization probes were labeled by the nick translation activity of DNA polymerase I³² or with polynucleotide kinase using [γ -³²P]-ATP as substrate.²⁶ Purification was accomplished by first running a sample containing the fragment of interest on a 3.5% polyacrylamide gel. The fragment was visualized

and cut out with as little acrylamide as possible. This piece was then placed directly into a well of another 3.5% polyacrylamide gel. After electrophoresis and visualization, the fragment was once again cut out. The DNA was purified from the acrylamide by the crush and soak procedure, as referenced above. Nick translated fragments may be obtained in greater than 95% purity using this method.

RESULTS

Localization of the Gene for the 32 kd Thylakoid Membrane Protein in *Euglena* Chloroplast DNA

One approach we have taken for the localization of *Euglena* chloroplast protein genes involves heterologous hybridization. Previously characterized higher plant^{8,33} or bacterial³⁴ protein gene segments are radiolabeled and hybridized to membrane filter blots of *Euglena* chloroplast DNA fragments. The methods and strategy for these studies have been described for locating the intron-containing gene for the large subunit of ribulose-1,5-bisphosphate carboxylase (rbcL),³⁵ and elongation factor EF-Tu (tufA).³⁴ The *Euglena* chloroplast psbA locus was determined by using specific psbA DNA probes derived from the spinach psbA gene. The spinach chloroplast psbA gene, which is located in the large single copy DNA region adjacent to one of the inverted repeat DNAs,⁶ has been completely sequenced.⁷

An 8.4 kbp Pst I restriction fragment of spinach chloroplast DNA which contains the 3'-terminal 805 base pair (bp) of the 1062 bp psbA coding sequence was radioactively labeled via nick translation and used as a hybridization probe for the initial screening of *Euglena* chloroplast DNA. A subclone of the 8.4 kbp fragment containing the psbA gene was also constructed. The spinach subclone contains a 1300 bp Sal I-Pst I fragment which has, in order, the 805 bp 3'-terminal psbA coding sequence, a 141 bp intergenic spacer, tRNA^{His} gene, and approximately 200 bp of spinach inverted repeat DNA⁷ (Gerard Zurawski, personal communication).

A gel photo of an EcoRI digest of total *Euglena* chloroplast DNA is shown in Figure 1, Lane A. The result of hybridization with the ³²P-labeled spinach 1300 bp psbA probe is shown in the corresponding autoradiogram. The 4.9 kbp EcoRI fragment EcoI hybridizes specifically to the 1300 bp probe. Also shown in the gel photo (Lane B) is a HindIII digest of total *Euglena* chloroplast DNA. In the corresponding autoradiogram, the HindIII fragment HindIII 10 is shown to be complementary to the spinach psbA probe. HindIII 10 is located entirely within EcoI. Therefore, it was concluded that the *Euglena* psbA gene

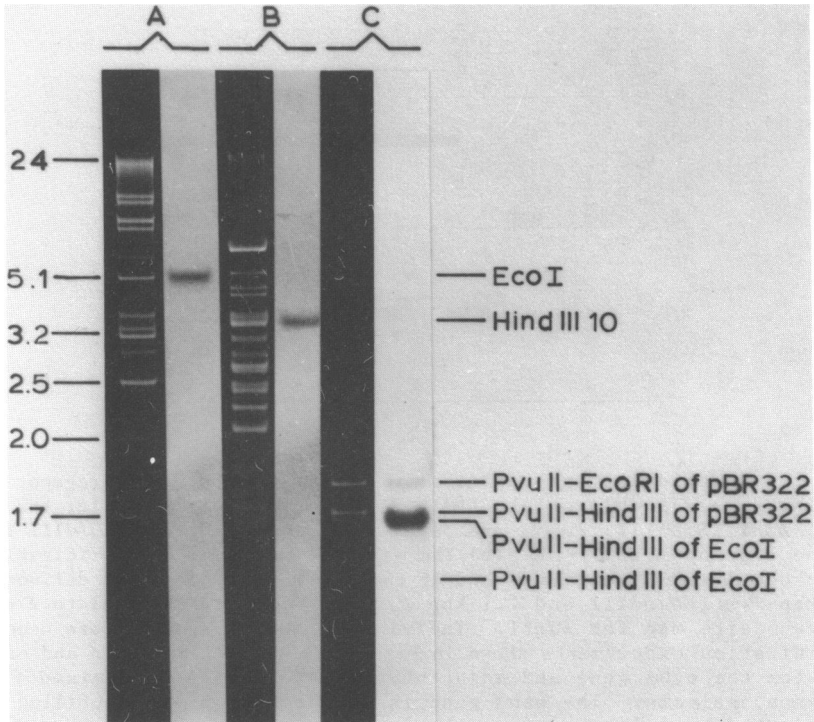


Fig. 1. Location of the *psbA* gene on the *Euglena* chloroplast genome. The gel photos are as follows: A. *Euglena* chloroplast DNA digested with *EcoRI*. B. *Euglena* chloroplast DNA digested with *HindIII*. C. Plasmid pE2C514 triply digested with the restriction endonucleases *PvuII*, *HindIII*, and *EcoRI*. The corresponding autoradiograms are of hybridization with the ^{32}P -radiolabeled spinach *psbA* probe. Agarose gel fragment size markers to the left of the figure are in kilobasepairs. Actual fragment sizes from DNA sequence analysis may vary from estimates based on agarose gel mobility. Fragments are named to the right of the figure.

is within *EcoI*. This result has been presented as a preliminary report,^{15,33} and was confirmed by Keller et al.⁹

In order to study the *psbA* locus and its transcripts in more detail, *EcoI* DNA was purified and cloned in pBR322. This recombinant plasmid is designated pE2C514. Further hybridizations were done to localize the gene within *EcoI*. In lane C of Figure 1 the resulting *psbA* probe hybridization to *EcoRI*-*PvuII*-*HindIII* digestion products of pE2C14 DNA is shown. Shown in the autoradiogram is the *psbA* probe hybridization to a *PvuII*-*HindIII* fragment and the adjacent 1.1 kbp *HindIII* fragment which are both internal to *EcoI*. A detailed restric-

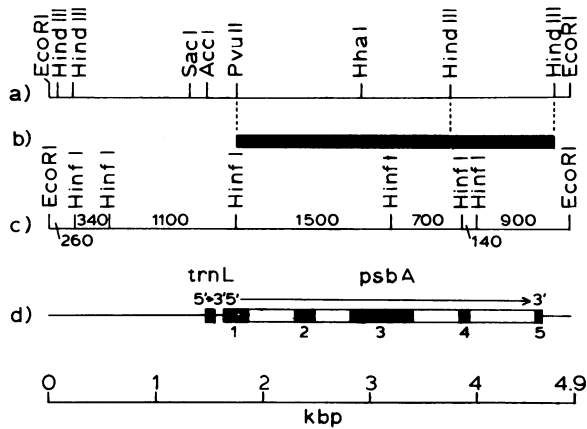


Fig. 2. Restriction endonuclease map of the 4.9 kbp *EcoI* fragment *EcoI* of *Euglena gracilis* chloroplast DNA. (a) Cleavage site maps for the enzymes *AccI*, *HhaI*, *HindIII*, *PvuII*, and *SacI*. The largest internal *HindIII* fragment is the 3.5 kbp *HindIII*-10. (b) The maximum boundaries of hybridization with the probe derived from the 3'-end of the spinach *psbA* gene are defined by the 2.0 kbp *PvuII*-*HindIII* and 1.1 kbp *HindIII* fragments internal to *EcoI*. (c) Cleavage site map for *HinfI*. Individual *HinfI* fragments were used in the hybridization experiments shown in Figure 5. (d) The location and organization for the *psbA* gene and adjacent *trnL* gene as determined from DNA sequence analysis. The *psbA* gene is composed of 5 exons (filled boxes), interrupted by 4 introns (open boxes). (G. D. Karabin, M. Farley, J. O. Narita, and R. B. Hallick, manuscript in preparation.)

tion map of the *EcoI* fragment and the maximum boundaries of the region hybridizing to the *psbA* probe are shown in Figure 2. A circular map of *Euglena* chloroplast DNA showing the location of *EcoI* and the *psbA* gene in relation to previously mapped genes is in Figure 3.

In order to confirm the identification of the *psbA* locus in *EcoI*, the complete DNA sequence of *EcoI* has been determined. This work will be described elsewhere (G. D. Karabin, M. Farley, J. O. Narita, and R. B. Hallick, manuscript in preparation). The results most germane to the gene expression studies to be described below are as follows: (1) The *Euglena* chloroplast *psbA* amino acid sequence derived from the DNA sequence has 87% homology to that of spinach and *N. debrynei*.⁷ (2) The *Euglena* *psbA* locus is interrupted by four introns ranging in size from 434-617 bp. The organization of exons and introns, and the gene polarity are illustrated in Figure 2d. (3) The minimum length of the gene from the AUG-start codon to the UAA-termination codon is 2967 bp. (4) Exons 4 and 5, which are encoded on the 1.1 kbp *HindIII* fragment internal to *EcoI*, are small (33 and 13 codons, respectively),

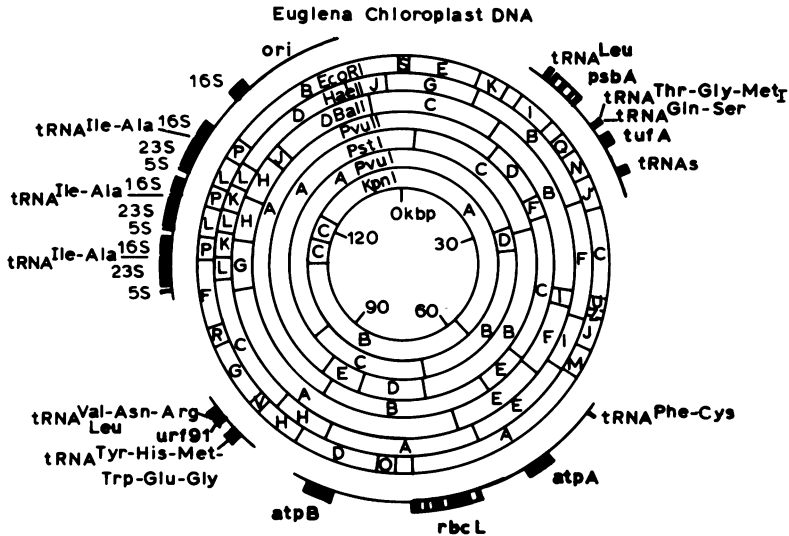


Fig. 3. Circular restriction endonuclease map of *Euglena gracilis* chloroplast DNA. The 4.9 kbp *EcoRI* fragment *EcoI* is located at approximately one o'clock in the outermost circular map. Also noted on this map are the locations of other known protein coding loci, known tRNA genes, and the three tandemly repeated rRNA operons.

and gave only weak hybridization signals with the spinach *psbA* probe compared to the *HindIII* 10.

Multiple RNA Transcripts of the *Euglena* Chloroplast *psbA* Locus

In order to study RNA transcripts from the *psbA* locus during light-induced chloroplast development, RNA was isolated from dark-adapted *Euglena gracilis* cells (0-hour), and from cells following 12, 24, 36, 48, 60, and 72 hours of light induced chloroplast development. The 0-hour cells contain immature proplastids, and the 72-hour cells contain mature chloroplasts.¹⁷ Five micrograms of each RNA were electrophoresed in adjacent lanes in a formaldehyde/agarose gel. Electrophoretically separated RNAs were transferred to cellulose nitrate filters by blotting.³¹ Putative transcripts of the *psbA* locus were initially detected by hybridization with pEZC514 DNA, ³²P-labeled by nick translation. Results are shown in Figure 4. The expected size of the *psbA* polypeptide mRNA is 1.2 kilobases (kb).^{7,9,36} There is a major transcript found upon hybridization with *EcoI* of 1.2 kbp. This RNA shows distinct developmental regulation, with a marked increase in hybridization signal during chloroplast maturation. In addition to the major 1.2 kb RNA transcript, at

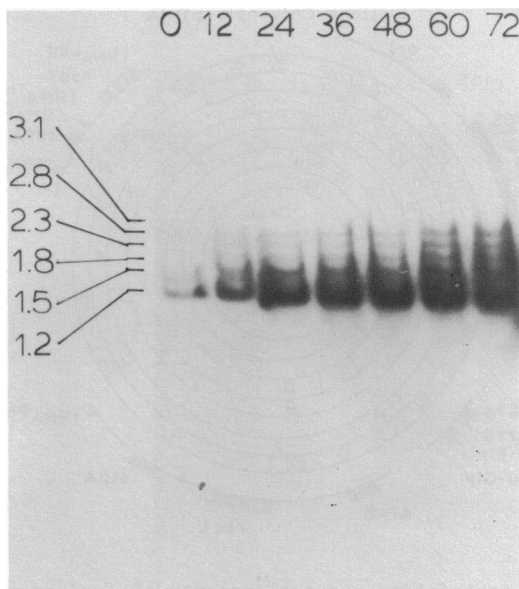


Fig. 4. Hybridization of a radiolabeled *psbA* probe to a cellulose nitrate filter blot containing electrophoretically separated RNAs from different stages of *Euglena* chloroplast development. The probe was nick translated pEZC514, which consists of *Euglena* chloroplast DNA fragment *EcoI* ligated into the vector pBR322. The lanes in the blot contain, from left to right, 5 micrograms each of whole cell RNAs from the 0, 12, 24, 36, 48, 60, and 72 hour stage of *Euglena* chloroplast development. The sizes of the RNAs detected are given to the left in kilobases.

least 5 additional, distinct RNAs are evident, with sizes 3.1, 2.8, 2.3, 1.8, and 1.5 kb. The sum of these RNA lengths, 12.7 kb, is larger than the double strand complexity of the 4.9 kbp *EcoI* DNA.

Several explanations for these multiple RNAs were considered. First, some or all of the hybridization could be due to a false positive signal due to the plasmid vector DNA present in the ^{32}P -pEZC514 DNA probe. To test this, *EcoI* DNA and the pBR322 vector DNA were individually purified, radiolabeled, and hybridized to identical RNA blots. The hybridization pattern with the *EcoI* DNA probe was identical to that found with pEZC514 DNA. None of these RNAs were detected when ^{32}P -pBR322 DNA was used as the probe (data not shown). Second, some of the RNAs detected with the *EcoI* probe might be transcripts of other regions of the chloroplast genome that fortuitously hybridized to *EcoI* DNA, despite the high stringency hybridization conditions. This could not be rigorously excluded for every region of chloroplast DNA. However,

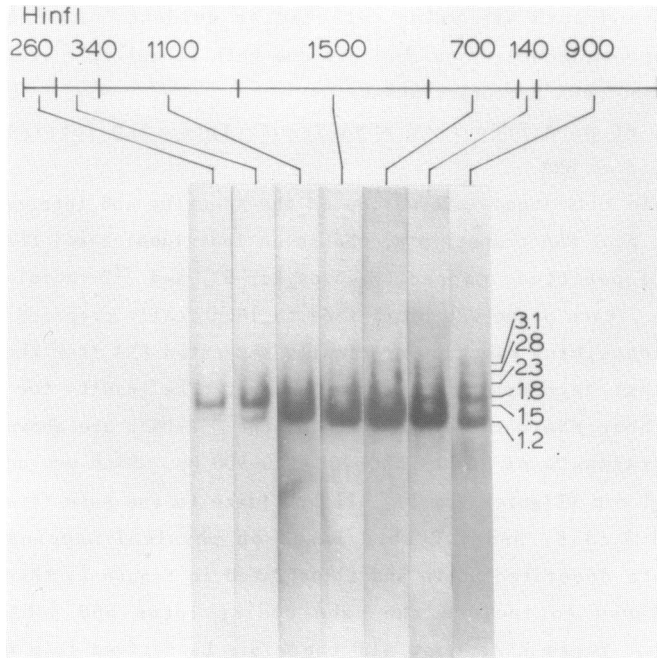


Fig. 5. Correlation of *EcoI* hybridization patterns to the internal *HinfI* fragments of *EcoI*. Shown above is the *HinfI* restriction endonuclease map of the *EcoI* restriction fragment from *Euglena gracilis* chloroplast DNA. Below are single lanes from the hybridization of each *HinfI* fragment to a cellulose nitrate filter blot of electrophoretically separated RNAs from seven stages of *Euglena* chloroplast development. The individual lanes are from either the 60 or 72 hour developmental stage. Sizes of the transcripts are given to the right, in kilobases. Sizes of the restriction fragments are given in base pairs.

a number of other cloned *Euglena* chloroplast DNA fragments containing known genes gave hybridization patterns clearly distinct from that of *EcoI*. These included subclones of *EcoA* (*rbcl*), *EcoD* (*atpB*), *EcoN* (*tufA*), *EcoG* (an unidentified open reading frame, *urf 91*), *EcoP* (16S rRNA), and tRNA coding loci of *EcoV*, *EcoH*, and *EcoG* (M. J. Hollingsworth, unpublished observations). We therefore conclude that all of the RNAs are wholly or in part encoded in *EcoI*. Third, it is possible that some of the six RNAs are products of a different gene or genes besides *psbA*, which are located in *EcoI*. Finally, some of the RNAs in the 1.5-3.1 kbp size range might be unprocessed precursor mRNA transcripts of *psbA* or of other loci in *EcoI*. With the characterization of *psbA* as an intron-containing gene, the interesting possibility that unspliced

psbA mRNA precursors are being detected is evident. As shown below, the multiple RNAs complementary to *EcoI* are due both to multiple genes within *EcoI* and unprocessed psbA precursor mRNAs.

Correlation of psbA Precursor mRNA Transcripts with Internal Restriction Fragments of *EcoI* DNA

To obtain additional information on the identity and interrelationship of the multiple *EcoI* RNA transcripts, the seven individual *HinfI* DNA subfragments of *EcoI* were purified, mapped (Figures 2c, 5), and ³²P-radiolabeled by nick translation. Each probe was hybridized to identically prepared membrane filter blots containing electrophoretically separated RNA from the seven stages of chloroplast development described above. The results for either the 60 hour or 72 hour RNA sample with each of the ³²P-DNAs are shown in Figure 5. The *HinfI* fragments of 1100-1500-700-140-900 bp, which are adjacent on the *EcoI* physical map (Figures 2c, 5), all hybridize to the same five RNAs of size 3.1, 2.8, 2.3, 1.8, and 1.2 kb. Based on physical mapping and the DNA sequence data described above and illustrated in Figure 2, these *HinfI* fragments are known to include the psbA coding locus and psbA intervening sequence(s). These RNAs must all therefore be derived from the psbA transcription unit. The primary transcript is at least 3.1 kb in length. The 2.8, 2.3, 1.8, and perhaps the 1.5 kb RNAs must be partially processed precursors of the mature 1.2 kb psbA mRNA. At least some of these precursor transcripts, including the 3.1 kb pre-mRNA, contain unspliced introns.

A second *EcoI* gene transcript was detected by hybridization to the adjacent 260-340-1100 bp *HinfI* fragments (Figure 5). This gene is separated from the psbA locus by a tRNA coding region containing a tRNA^{Leu} gene in the 1100 bp *HinfI* fragment (G. D. Karabin and R. B. Hallick, submitted for publication). The major transcript detected by membrane filter hybridization is 1.5 kbp (Figure 5). This RNA was also detected with pEZC514 DNA as probe. In addition, a weak 1.2 kb RNA hybridization signal is obtained with the 260 and 340 bp ³²P-labeled *HinfI* DNA probes. This may be a consequence of minor psbA sequence contamination in the 260 and 340 bp DNA preparations, breakdown of the 1.5 kb RNA transcript, or a third gene within *EcoI*.

Hybridization of a psbA Exon-specific DNA Probe to Unspliced psbA Precursor and Mature mRNAs

To obtain further evidence that the transcripts of *EcoI* are at least in part unprocessed psbA mRNA precursors, a hybridization probe of defined exon sequence was prepared. With reference to the restriction map of *EcoI* (Figure 2), the 270 bp *HinfI*-*HhaI* fragment contains exon sequence homologous to codons

140-231 of the spinach *psbA* locus,⁷ but lacks intron sequences (G. D. Karabin and R. B. Hallick, unpublished observations). These restriction nuclease sites are also present in spinach chloroplast *psbA*. This fragment was radio-labeled in two different strand specific reactions. The 1500 bp *HinfI* fragment was first 5'-³²P end labeled with polynucleotide kinase, and then treated with *HhaI* to yield 5'-end label on the DNA template strand. The 5'-end labeled 270 bp *HinfI-HhaI* DNA fragment was subsequently purified via polyacrylamide gel electrophoresis. Alternatively, the 5'-end labeling was initially at the *HhaI* site, followed by *HinfI* digestion and fragment purification, to yield, as a control, a probe with the "mRNA-like" sequence radio-labeled.

The two ³²P-DNAs were then hybridized to membrane filter blots of RNAs from various chloroplast development stages as described above. The DNA probe 5'-end labeled on the template strand, and thus complementary to *psbA* mRNA, hybridized to RNAs of 3.1, 2.8, 2.3, 1.8, and 1.2 kb as expected for the *psbA* transcripts. The gel region containing the 1.5 kb RNAs was not clear in this experiment. Furthermore, the changes in hybridization signal during development paralleled those described above (Figure 6). The control probe, 5'-end labeled on the RNA-like strand, gave a very weak signal at 1.2 kb. It took six times longer for this very faint signal to appear than it did for the very strong DNA template signal. We conclude that the weak signal is due to slight contamination of the RNA-like probe by DNA template specific *psbA* sequences. This confirms the conclusion that the 3.1, 2.8, 2.3, and 1.8 kb *EcoI* transcripts are all *psbA* mRNA precursors, some of which must contain unspliced intervening sequences.

Hybridization of a *psbA* Intron-specific DNA Probe to Unspliced *psbA* Precursor mRNAs

As a final confirmation that the 3.1, 2.8, 2.3, 1.8, and 1.5 kb RNA transcripts of *EcoI* are unspliced *psbA* pre-mRNAs, a defined intron-specific sequence was employed. The synthetic oligonucleotide 5'-GATGTTATTCTTTC was constructed to complement transcripts of intron 2 (G. D. Karabin, M. Farley, J. Narita, and R. B. Hallick, in preparation). The tetradecamer was radio-labeled at the 5'-end with polynucleotide kinase, and hybridized to a blot of electrophoretically separated RNAs. In order to determine the position of pre-mRNAs, this blot had previously been hybridized to a *psbA* probe containing DNA from the 5'-leader region through exon 3. The results of this experiment are shown in Fig. 7. The probe that contains sequences from exon 1-3, introns 1-2, and some 5'-leader sequence hybridizes to RNAs of 3.1, 2.8, 2.3,

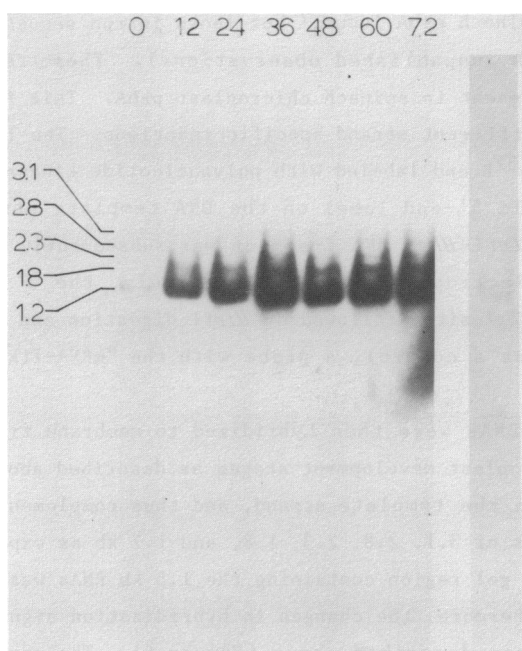


Fig. 6. Hybridization of an exon specific probe for *psbA* to a cellulose nitrate blot of electrophoretically separated RNA from seven stages of *Euglena gracilis* chloroplast development. The probe was a 270 base pair *HhaI-HinfI* restriction endonuclease fragment that has been shown by DNA sequence analysis (G. D. Karabin, unpublished observations) to contain only sequences within the coding region of *psbA*. The autoradiogram to the left results when only the *HinfI* restriction endonuclease cleavage site is radiolabeled, resulting in only the DNA template strand of the fragment having radioactivity. The lanes are marked above with the developmental stage from which the RNA was isolated. Sizes of the transcripts are given to the left in kilobases. The autoradiogram to the right is the result of hybridization of the *HhaI-HinfI* fragment when it has been radiolabeled only at the *HhaI* restriction endonuclease site, resulting in only the "RNA-like" strand being radioactive. The cellulose nitrate filter blot for this probe contained electrophoretically separated *Euglena gracilis* chloroplast RNA. Although the specific activity of the two probes was the same, the "RNA-like" probe gave only this faint signal, even after an exposure to x-ray film that was six times longer than that of the DNA-template probed autoradiogram. We suggest that the signal from the "RNA-like" strand of the *psbA* exon probe was due to minor contamination from a DNA template probe.

1.8, 1.5 (doublet), and 1.2 kb. The major signal at 1.2 kb is the mature *psbA* mRNA. The intron 2 specific oligonucleotide did not hybridize to the 1.2 kb mature mRNA, but did hybridize to the same precursors of 3.1, 2.8, 2.3 and 1.5 kb. In addition, a faint signal at 0.4 kb is also evident. This may be the excised intervening sequence from intron 2. In a control experiment (data not

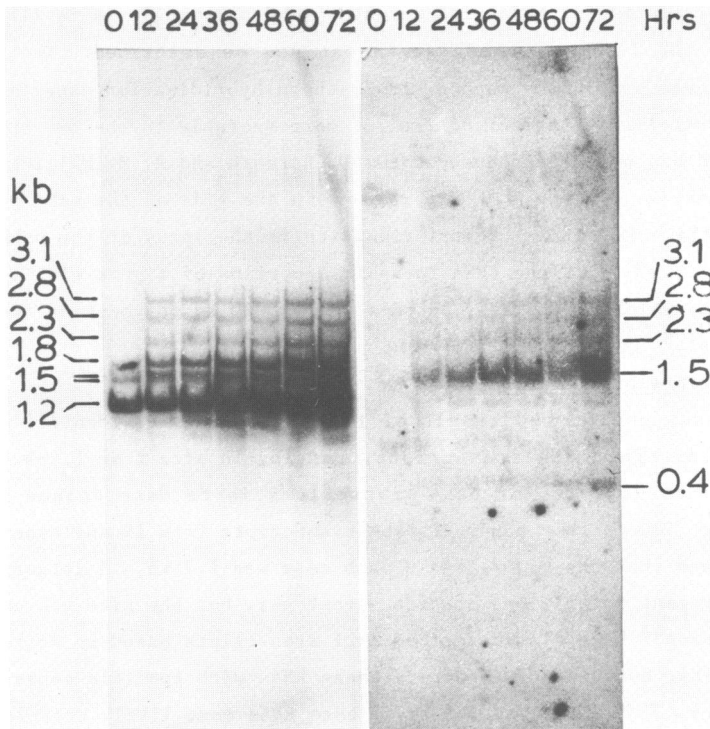


Fig. 7. Hybridization of a radiolabeled *psbA* probe and a chemically synthesized intron 2 specific probe to a filter blot of RNAs from different stages of *Euglena* chloroplast development separated on 1% formaldehyde/agarose gels. Left, the probe was a nick translated 1337 bp *ThaI-HhaI* fragment, which contains DNA sequences from the *trnL-psbA* spacer (*ThaI* site) to an internal site in exon 3 (*HhaI* site). Right, the same blot was treated at 95°C in 0.1 X SSC (1 min) and then washed in 20 X SSC at 22°C to remove the probe, re-exposed to X-ray film to determine that all radioactivity had been removed, and then re-hybridized with 5'-³²P-GATGTTATCTTTC.

shown) an intron 1 specific synthetic oligonucleotide, complementary to the opposite strand of DNA as the intron 2 probe, gave no detectable hybridization to any chloroplast RNA.

DISCUSSION

Gene Mapping on *Euglena* Chloroplast DNA Fragment *EcoI*

It has previously been reported that the 4.9 kbp *EcoRI* fragment *EcoI* of *Euglena* chloroplast DNA contains genes for one or more tRNAs,^{37,38,39} including a tRNA^{Leu}.⁴⁰ *EcoI* has also been reported to contain genes for the 32 kd polypeptide associated with photosystem II, designated *psbA*,^{8,9,33} and a

second polypeptide of molecular weight 46,000 daltons.⁹ As is illustrated in Figure 2, the locus of these genes can now be described. The tRNA coding locus has been previously mapped via Southern hybridization experiments to the 1.8 kbp *EcoRI-PvuII* fragment of *EcoI*,³⁹ more recently to the 500 bp *SacI-PvuII* region, and has now been sequenced (G. D. Karabin and R. B. Hallick, submitted for publication). The 3.0 kbp region to the side of the tRNA gene(s) that includes the unique *HhaI* restriction site is the locus of the *psbA* gene. On the opposite side of the tRNA locus is a portion of a gene with a 1.5 kb RNA transcript, which most likely encodes the 46,000 kd polypeptide identified by Keller et al.⁹

RNA Transcripts of the *psbA* Locus

The most unexpected result of this study was the identification of at least six distinct RNA transcripts, ranging in size from 1.2-3.1 kb. This hybridization pattern for *psbA* transcripts is to date unique for *Euglena* chloroplast DNA. In a study of *psbA* transcripts from 19 angiosperms, Palmer et al³⁶ found that the major RNA in each case was 1.2 kb. A larger transcript was also present in only one species, sweet pea, but the size was only 1.4 kb. Dix and Rawson⁴¹ have also described *EcoI* transcripts based on Northern hybridization data, but did not associate these RNAs with specific genes. The RNAs were of 1.4, 2.0, 2.5, and 3.2 kb. These RNAs most likely correspond to the 1.2, 1.8, 2.3, and 3.1 kb RNAs described in the present study. Keller et al⁹ also described major *EcoI* transcripts of 14S and 17S which we believe correspond to our 1.2 and 1.5 kb RNAs. They also noted the appearance of faint bands larger than 17S, but did not further characterize these RNAs. It is apparent from a comparison of the hybridization data of Keller et al⁹ to the present results that our RNA is significantly less degraded. We believe that this is due to our use of aurintricarboxylic acid as an RNase inhibitor during RNA isolation.⁴²

We conclude that the larger transcripts of *psbA* in *Euglena* chloroplasts represent unprocessed precursors of the mature mRNA for the 32 kd photosystem II polypeptide. The mature mRNA size, which is the major transcript in *Euglena*, is 1.2 kb, the same as that found in higher plants. At least five processing events are required to produce the mature mRNA if a sequential processing pathway is assumed. Processing would proceed 3.1 kb → 2.8 kb → 2.3 kb → 1.8 kb → 1.5 kb → 1.2 kb.

The processing of the primary *psbA* transcript could involve three types of events. These are cleavage of a 5'-nontranslated leader sequence, removal of a 3'-nontranslated trailer sequence, and splicing of intervening sequences

in the mRNA. From partial sequence analysis of the gene, there is evidence that psbA contains four intervening sequences. Therefore four of the psbA mRNA processing events are presumed to involve splicing reactions. We predict that most of the precursor transcripts detected are partially spliced mRNAs. A more detailed analysis of psbA mRNA precursor processing reactions is in progress. We hope to more precisely define mRNA splicing sites and the processing pathway. The detection of multiple, unspliced pre-mRNAs with an intron 2 probe is an initial indication that processing may not simply involve sequential removal of intervening sequences in the primary RNA transcript. It has already been possible to demonstrate that the intron 2 probe will serve as a primer for reverse transcription of intron 2 cDNA using total chloroplast RNA as template (U. Johannigmeier and R. B. Hallick, unpublished observations).

This is the second example of analysis of RNA transcripts of a split chloroplast protein gene. The gene for the large subunit of ribulose-1,5-bisphosphate carboxylase (rbcL) of *Euglena* has at least nine introns.^{35,43} Unlike the psbA RNA, unspliced rbcL transcripts are not readily detected.

Developmental Regulation of psbA mRNA

There is a significant increase in the level of psbA mRNA transcription during the course of light induced chloroplast development. The increase has not been rigorously quantitated. Based on densitometry of X-ray films from various Northern hybridization experiments, we estimate approximately a five-fold increase in RNA during the 72 hour period of chloroplast development. The largest increase in mRNA occurs during the zero hour to twelve hour time interval. There is no major shift during chloroplast development in the proportion of transcripts of the various precursor size classes.

Rawson et al.²² have used a solution hybridization approach to study developmental changes in *EcoI* transcripts in *Euglena*. They find that the abundance of *EcoI* RNAs, which would include both psbA and the 1.5 kb mRNA, increases from 50 to 350 copies of transcript per cell during 48 hours of greening. This sevenfold increase is in good agreement with the present findings. We can therefore conclude that there is a temporal change in the cellular content of psbA mRNA levels in *Euglena gracilis* during light induced chloroplast development.

ACKNOWLEDGMENTS

This work was supported by Grant GM21351 from the National Institutes of Health to R.B.H. U. J. was supported by the Deutsche Forschungsgemeinschaft (123-3). The authors gratefully acknowledge Drs. Ellen Reardon and Vivian

MacKay for their helpful advice on the blotting of RNA gels. We are also grateful to Drs. W. F. Thompson and J. D. Palmer for the generous gift of spinach recombinant *Pst*I clones and total spinach chloroplast DNA, and to Dr. Scott E. Bingham for the construction of pEZC514.

*Present address: Syngene, Inc., P.O.Box 2211, Fort Collins, CO 80522, USA

+To whom correspondence should be addressed

REFERENCES

1. Renger, G. (1976) *Biochim. Biophys. Acta* 440, 287-300.
2. Mattoo, A. K., Pick, U., Hoffman-Falk, H. and Edelman, M. (1981) *Proc. Natl. Acad. Sci. USA* 78, 1572-1576.
3. Pfister, K., Steinback, K. E., Gardner, G. and Arntzen, C. J. (1981) *Proc. Natl. Acad. Sci. USA* 78, 981-985.
4. Steinback, K. E., McIntosh, L., Bogorad, L. and Arntzen, C. J. (1981) *Proc. Natl. Acad. Sci. USA* 78, 7463-7467.
5. Link, G., Coen, D. M. and Bogorad, L. (1978) *Cell* 15, 725-731.
6. Driesel, A. J., Speirs, J. and Bohnert, H. J. (1980) *Biochim. Biophys. Acta* 610, 297-310.
7. Zurawski, G., Bohnert, H. J., Whitfeld, P. R. and Bottomley, W. (1982) *Proc. Natl. Acad. Sci. USA* 79, 7699-7703.
8. Stiegler, G. L., Matthews, H. M., Bingham, S. E. and Hallick, R. B. (1982) *Fed. Proc. FASEB* 41, 758.
9. Keller, M., Rutti, B. and Stutz, E. (1982) *FEBS Lett.* 149, 133-137.
10. Grebanier, A. E., Coen, D. M., Rich, A. and Bogorad, L. (1978) *J. Cell Biol.* 78, 734-746.
11. Blair, G. E. and Ellis, R. J. (1974) *Biochim. Biophys. Acta* 319, 223-234.
12. Chua, N.-H. and Gillham, N. W. (1977) *J. Cell Biol.* 74, 441-452.
13. Hoffman-Falk, H., Mattoo, A. K., Marder, J. B., Edelman, M. and Ellis, R. J. (1982) *J. Biol. Chem.* 257, 4583-4587.
14. Hirschberg, J. and McIntosh, L. (1983) *Science* 222, 1346-1349.
15. Hoffman-Falk, H., Mattoo, A. K., Marder, J., Fluhr, R. and Edelman, M. (1980) *Proc. of 13th FEBS Meeting*, 22.
16. Reisfeld, A., Gressel, J., Jakob, K. M. and Edelman, M. (1978) *Photochem. Photobiol.* 27, 161-165.
17. Ben-Shaul, Y., Schiff, J. A. and Epstein, H. T. (1964) *Plant Physiol.* 39, 231-240.
18. Schiff, J. A., Zeldin, M. H. and Rubman, J. (1967) *Plant Physiol.* 42, 1716-1725.
19. Bingham, S. and Schiff, J. A. (1979) *Biochim. Biophys. Acta* 547, 531-543.
20. Gilbert, C. W. and Buetow, D. E. (1982) *Biochem. Biophys. Res. Commun.* 107, 649-655.
21. Chelm, B. K., Hallick, R. B. and Gray, P. W. (1979) *Proc. Natl. Acad. Sci. USA* 76, 2258-2262.
22. Rawson, J. R. Y., Boerma, C. L., Andrews, W. H. and Wilkerson, C. G. (1981) *Biochemistry* 20, 2639-2644.
23. Chelm, B. K., Hoben, P. J. and Hallick, R. B. (1977) *Biochemistry* 16, 776-781.

24. Hallick, R. B., Richard, O. C. and Gray, P. W. (1982) in *Methods in Chloroplast Molecular Biology* (M. Edelman, R. B. Hallick and N.-H. Chua, eds.) Elsevier Biomedical, Amsterdam, pp. 281-294.
25. Bahl, C. P., Mariani, K. J. and Wu, R. (1976) *Gene* 1, 81-96.
26. Maxam, A. M. and Gilbert, W. in *Methods in Enzymology* 65, (L. Grossman and K. Moldave, eds.) Academic Press, New York, pp. 499-560.
27. Gray, P. W. and Hallick, R. B. (1979) *Biochemistry* 18, 1820-1825.
28. Langridge, J., Langridge, P. and Bergquist, P. L. (1980) *Anal. Biochem.* 103, 264-271.
29. Denhardt, D. (1966) *Biochem. Biophys. Res. Commun.* 23, 641-645.
30. Rave, N., Crkvenjakov, R. and Boedtker, H. (1979) *Nucleic Acids Res.* 6, 3559-3567.
31. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* 77, 5201-5205.
32. Weinstock, R., Sweet, R., Weise, M., Cedar, H. and Axel, R. (1978) *Proc. Natl. Acad. Sci. USA* 75, 1299-1303.
33. Hallick, R. B., Greenberg, B. M., Gruissem, W., Hollingsworth, M. J., Karabin, G. D., Narita, J. O., Passavant, C. W. and Stiegler, G. L. (1983) in *Structure and Function of Plant Genomes* (NATO Advanced Studies Institute) (O. Cifferi and L. Dure, eds.) Plenum Press, New York, in press.
34. Passavant, C. W., Stiegler, G. L. and Hallick, R. B. (1983) *J. Biol. Chem.* 258, 693-695.
35. Stiegler, G. L., Matthews, H. M., Bingham, S. E. and Hallick, R. B. (1982) *Nucleic Acids Res.* 10, 3427-3444.
36. Palmer, J. D., Edwards, H., Horgensen, R. A. and Thompson, W. F. (1982) *Nucleic Acids Res.* 10, 6819-6832.
37. Hallick, R. B., Rushlow, K. E., Orozco, E. M., Jr., Stiegler, G. L. and Gray, P. W. (1979) *ICN/UCLA Symp. Molec. Cell. Biol.* 15, (Extrachrom. DNA), 127-141.
38. Orozco, E. M., Jr. and Hallick, R. B. (1982) *J. Biol. Chem.* 257, 3258-3264.
39. El-Gewely, M. R., Lomax, M. I., Lau, E. T., Helling, R. B., Farmerie, W. and Barnett, W. E. (1981) *Molec. and General Genetics* 181, 296-305.
40. Kuntz, M., Keller, M., Crouse, E.J., Burkard, G. and Weil, J. H. (1982) *Curr. Genetics* 6, 63-69.
41. Dix, K. P. and Rawson, J. R. Y. (1983) *Curr. Genetics* 7, 265-272.
42. Hallick, R. B., Chelm, B. K., Gray, P. W. and Orozco, E. M., Jr. (1977) *Nucleic Acids Res.* 4, 3055-3064.
43. Koller, B., Gingrich, J., Farley, M., Delius, H. and Hallick, R. B. (1984) *Cell*, in press.