Plant Physiol. (1996) 111: 61–71

A Putative Mg Chelatase Subunit from *Arabidopsis thaliana* cv C24

Sequence and Transcript Analysis of the Gene, Import of the Protein into Chloroplasts, and in Situ Localization of the Transcript and Protein

Lucien C. D. Gibson[1], Joanne L. Marrison[1], Rachel M. Leech[1], Poul E. Jensen[2,3], Diane C. Bassham[4], Marie Gibson[1], and C. Neil Hunter[1]*

Robert Hill Institute for Photosynthesis and Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2UH, United Kingdom (L.C.D.G., M.G., C.N.H.); Department of Biology, University of York, Heslington, York YO1 5DD, United Kingdom (J.L.M., R.M.L.); Department of Ecology and Molecular Biology, Royal Veterinary and Agricultural University, Bülowsvej 13, DK-1870 Frederiksberg C, Denmark (P.E.J.); and Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, United Kingdom (D.C.B.)

We have isolated and sequenced a cDNA from *Arabidopsis thaliana* cv C24 that encodes a putative Mg chelatase subunit. The deduced amino acid sequence shows a very high level of identity to a gene previously characterized from *Antirrhinum majus* (*olive*) and also high similarity to *bchH*, a bacterial gene involved in the Mg chelatase reaction of bacteriochlorophyll biosynthesis. We suggest that this gene be called *CHL H*. Northern blot analyses were used to investigate the expression of *CHL H*, another putative Mg chelatase gene, *ch*-42, and ferrochelatase. The *CHL H* transcript was observed to undergo a dramatic diurnal variation, rising almost to its maximum level by the end of the dark period, then increasing slightly at the onset of the light and declining steadily to a minimum by the end of the light period; in contrast, transcripts for *ch*-42 and ferrochelatase remained constant. A model is proposed in which the CHL H protein plays a role in regulating the levels of chlorophyll during this cycle. In situ hybridization revealed that the transcripts are located over the surface of the chloroplasts, a feature in common with transcripts for the *ch*-42 gene. The CHL H protein was imported into the stromal compartment of the chloroplast and processed in an in vitro assay. Immunoblotting showed that the distribution of CHL H protein between the stroma and chloroplast membranes varies depending on the concentration of Mg^{2+}. In situ immunofluorescence was used to establish that the CHL H and CH-42 proteins are localized within the chloroplast in vivo.

[1] Supported by the Biotechnology and Biological Sciences Research Council of the United Kingdom.

[2] Supported by the Danish Agricultural and Veterinary Research Council (grant no. 13–5005–1).

[3] Present address: Robert Hill Institute for Photosynthesis and Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2UH, UK.

[4] Supported by a Biotechnology and Biological Sciences Research Council of the United Kingdom studentship.

* Corresponding author; e-mail c.n.hunter@sheffield.ac.uk; fax 44–114–272–8697.

Chlorophyll pigments harvest light energy, which is used to drive the reductive fixation of carbon dioxide. These pigments are the most abundant on earth, yet relatively little is known about the enzymology of the committed steps of chlorophyll biosynthesis. Still less is known about the genetic control of this pathway. One important enzyme in this pathway is Mg-protoporphyrin chelatase, the enzyme that catalyzes the insertion of Mg into protoporphyrin IX. This lies at the branch point of the chlorophyll and heme biosynthetic pathways and is considered to be the first committed step of chlorophyll synthesis (see Fig. 1). Accordingly, this enzyme has been the subject of much study in plants (Fuesler et al., 1981; Walker and Weinstein 1991a, 1991b; Walker and Weinstein, 1994), although purification of the enzyme has proved difficult. At present, it has been possible to establish a continuous assay for the enzyme and to show that two components probably participate in the chelation reaction (Walker and Weinstein, 1991b; Walker and Weinstein, 1995). It is also known that there is an ATP requirement for the activation of these components (Walker and Weinstein, 1994). However, the identities of these components and the genes that encode them have not been established.

The purple photosynthetic bacteria *Rhodobacter capsulatus* and *Rhodobacter sphaeroides* have been used to study the molecular genetics of bacteriochlorophyll biosynthesis, which has many intermediates in common with the chlorophyll biosynthetic pathway (Bauer et al., 1993). These bacteria provide a convenient model system for the study of these pathways partly because the bacteriochlorophyll biosynthesis genes are clustered on a small region of the chromosome, 45 kb in length (Marrs, 1981; Coomber and Hunter, 1989). Production of bacteriochlorophyll-minus mutants has allowed the correlation of mutant phenotypes with lesions in specific genes (Marrs, 1981; Zsebo and

Abbreviation: FITC, fluorescein isothiocyanate.

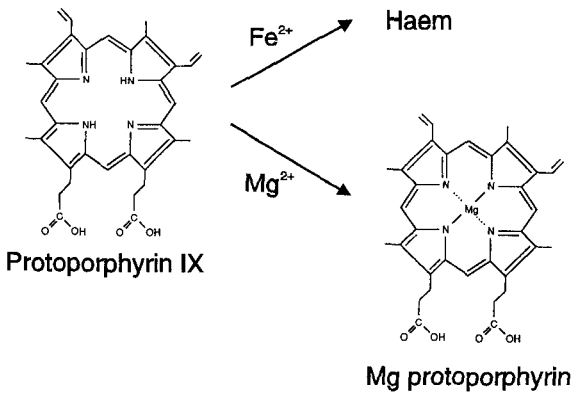

Figure 1. The branch point of chlorophyll and heme biosynthesis. The steps leading to protoporphyrin are common to both pathways, whereupon Mg is inserted to form chlorophyll or iron is chelated to form heme.

Hearst, 1984; Coomber et al., 1990; Bollivar et al., 1994). These genes have been assigned to steps in the pathway on the basis of bacteriochlorophyll intermediates accumulated by the mutants. Recently, there has been a step forward in the analysis of this pathway, as a result of overexpression of some *bch* genes in *Escherichia coli*, followed by studies of enzyme activity. The *bchM* genes of *R. capsulatus* and *R. sphaeroides* have been overexpressed in *E. coli*, and extracts from the resulting strains are able to catalyze the methylation of added Mg protoporphyrin, resulting in the formation of Mg protoporphyrin monomethyl ester. *bchM* has therefore been assigned to the methyltransferase step in bacteriochlorophyll biosynthesis (Bollivar et al., 1994; Gibson and Hunter, 1994). A similar approach has been used to demonstrate that the genes *bchI*, *bchD*, and *bchH* encode components of the Mg chelatase. *bchI* and *bchD* were expressed in one *E. coli* strain and *bchH* in another, extracts of these strains were combined, and the mixture was able to catalyze the insertion of Mg into protoporphyrin IX (Gibson et al., 1995).

This assignment has direct relevance to two studies carried out on higher plants. Two genes have been isolated from plants that show significant similarity to subunits of the *Rhodobacter* Mg chelatase; the nuclear gene cs/*ch*-42 from *Arabidopsis thaliana* (Koncz et al., 1990) is similar to *bchI* and *olive* from *Antirrhinum majus* is similar to *bchH* (Hudson et al., 1993). These genes have also been correlated with lesions in chlorophyll synthesis as a result of the chlorophyll-deficient phenotypes that arise from insertion of T-DNA or the Tam3 transposon into Arabidopsis and *Antirrhinum*, respectively. In view of these similarities and the demonstrated involvement of the bacterial homologs of these genes in bacteriochlorophyll synthesis, it seems likely that both *ch*-42 and *olive* encode subunits of the Mg chelatase.

In this paper, we have extended our previous bacterial studies to the model plant *A. thaliana* (cv C24); the *olive* homolog, which we propose to call *CHL H*, has been cloned and sequenced. Expression of *CHL H* was investigated at the level of transcription. The subcellular location of the *CHL H* and *ch*-42 transcripts were investigated using in situ hybridization. We show that the protein is imported into chloroplasts, and that both the CHL H and CH-42 proteins were colocalized within the chloroplasts by immunofluorescence.

MATERIALS AND METHODS

Cloning of the *CHL H* cDNA

A λZAPII library of cDNA prepared from green leaf tissue of *Arabidopsis thaliana* cv C24 was a gift from Dr. Christine Raines (University of Essex, UK). Approximately 10^7 plaque-forming units were used as the template in a PCR, using the oligonucleotides 5′-CGGAATTC(GAT)GA(GCAT)C(GT)(GCAT)C(GT)(CT)TT(GCAT)GC(GAT)AT(GCAT)GT-3′ and 5′-CGGAAGCTT(CT)GG(GCAT)TA(CT)GA(GA)GG(GCAT)GA(CT)CC(GCAT)AT-3′. One-fiftieth of this PCR was used as the template in a nested PCR, using the oligonucleotides 5′-ATTCGAATTCGA(GA)TT(CT)ATGCC(GCAT)GG(GCAT)AA(GA)CA-3′ and 5′-ACTGCTGCAGTT(GA)TT(GCAT)GC(GCAT)GC(GA)TA(GA)TA(GA)TA-3′. A 120-bp product was obtained from the nested PCR, which was used to screen 2.5×10^5 clones of the λZAPII library. Two clones were isolated, one of which was determined to contain the entire open reading frame. The pBluescript plasmid containing this cDNA was excised using the Ex-Assist helper phage supplied by Stratagene. A nested set of deletions of this plasmid was generated using exonuclease III (Henikoff, 1984) as described by Sambrook et al. (1989). These were used to determine the nucleotide sequence of the cDNA on both strands by the method of Sanger et al. (1977), using the Sequenase kit supplied by United States Biochemical. All DNA manipulations were carried out by standard methods (Sambrook et al., 1989).

RNA Preparation and Northern Analysis

Seeds of *A. thaliana* cv C24 were planted in moist peat. For diurnal rhythm experiments, seedlings were grown under a 16-h day/8-h night cycle at 24 ± 2°C and samples (3 g) from 25-d-old plants were harvested at 2-h intervals during a 24-h time period. The plants were watered with Hoagland mineral solution on alternate days. During harvesting of plants in darkness, a dim, green safelight was used. Total RNA was prepared from leaves as described by Schuler and Zielinski (1989). Samples (25 μg/lane) were electrophoresed on a 1% formaldehyde gel and blotted onto Hybond N (Amersham) by standard methods (Sambrook et al., 1989). Prehybridization and hybridization of the filters were carried out at 66°C in 5× SSPE (1× SSPE = 0.18 M NaCl, 10 mM sodium phosphate, pH 7.7, 1 mM EDTA), 5× Denhardt's solution (1× Denhardt's solution = 0.2 g/L BSA, 0.2 g/L Ficoll, 0.2 g/L PVP), 0.1% SDS, and 50 μg carrier DNA/mL. Radiolabeling of probes with ^{32}P was performed using the Megaprime kit (Amersham). Washes were performed at 66°C with 0.1× SSPE and 0.1% SDS in the final wash. Equal loading of RNA samples was confirmed by probing an identical blot with a soybean tubulin cDNA (Guiltinan et al., 1897). Sizes of transcripts were estimated by comparison with the 0.24- to 9.5-kb RNA ladder from GIBCO-BRL.

Import Studies Using Pea Chloroplasts

The precursor of the CHL H protein was synthesised in vitro by transcription of the cDNA clone followed by translation of capped transcripts in a reticulocyte lysate (Amersham) in the presence of [^{35}S]Met. Import of the labeled precursor protein into intact pea chloroplasts was carried out as described by Robinson and Barnett (1988). After the import reaction was completed, the chloroplasts were washed and fractionated as previously described (Robinson and Barnett, 1988). Equivalent amounts of each sample were separated on 7% SDS polyacrylamide gels (Laemmli, 1970), after which the gels were fixed, treated with the fluorofor Amplify (Amersham), dried, and exposed to x-ray film. Assays for cleavage of the precursor by partially purified pea stromal processing peptidase were as described by Musgrove et al. (1989).

Expression of *CHL H* in *Escherichia coli*

The primers 5′-GATCTAGACATATGCACGGTGCTG-TATCTGGAAACGGCCT-3′ and 5′-ACCAGAATTCAAGT-GGATGCCAAATTTCG-3′ were used to amplify the 5′ end of the *CHL H* cDNA. The product, of approximately 600 bp, was digested with *Nde*I and *Bam*HI and cloned into pET3a (Rosenberg et al., 1987). The resulting plasmid was digested with *Bam*HI and the remainder of *CHL H* was inserted, giving rise to a plasmid encoding a version of CHL H that we predicted would approximate the mature protein. This protein was produced in *Escherichia coli* BL21 (DE3) (Studier and Moffat, 1986) and used as a control in immunoblotting.

Preparation of Antibodies

Antibodies against the OLI protein were a gift from Dr. Andrew Hudson (University of Edinburgh, UK). Antibodies against the CH-42 protein were obtained by amplifying part of the *ch*-42 gene (Koncz et al., 1990) by PCR using the oligonucleotides 5′-AGTTCTGCAGGGAACAAGTAGT-AGGGAAGTTTG-3′ and 5′-TTAGGAATTCTCAGCTGA-AAATCTCGGCGAA-3′. The PCR fragment obtained was inserted into pET14b (Novagen, Madison, WI) and the CH-42 protein was overexpressed in *E. coli*. The protein was purified using the His tag system of Novagen. Immunization of rabbits and preparation of CH-42 antiserum were carried out at the Serum Institute (Copenhagen, Denmark).

Preparation of Arabidopsis Chloroplasts and Immunoblotting

Chloroplasts were prepared from 21-d-old Arabidopsis leaves by the method of Kunst et al. (1988), resuspended in 10 mM Tricine, pH 7.6, 2 mM EDTA containing either 1 or 5 mM $MgCl_2$, and subjected to two freeze-thaw cycles. Membranes were pelleted by centrifugation in a microfuge for 1 min, and the supernatants, referred to as the stromal fractions, were retained. The pelleted membranes were washed and resuspended in the appropriate buffer. Equivalent amounts of whole chloroplasts and stromal and membrane fractions were electrophoresed in a 10% SDS-polyacrylamide gel (Laemmli, 1970). Transfer of proteins to Hybond C (Amersham) and incubation of the blot with antibodies were done by standard methods (Sambrook et al., 1989). CHL H was detected using horseradish peroxidase-conjugated goat anti-mouse antisera and the enhanced chemiluminescence detection reagents supplied by Amersham.

In Situ Labeling of Transcripts and Proteins

A. thaliana (L.) cv Landsberg erecta seedlings were grown as described by Pyke et al. (1991). Whole first leaves were harvested 16 d after sowing and fixed in 3% (w/v) paraformaldehyde, 50% (v/v) ethanol, and 5% (v/v) acetic acid and embedded in PEG 1500 as previously described (Marrison and Leech, 1994). Transverse tissue sections (7 μm) were placed onto dampened polysine slides (Merck, Poole, UK) and left to dry on a hot plate at 40°C. In situ hybridization was performed with antisense and sense digoxigenen-labeled RNA probes synthesized by an in vitro transcription reaction using 1 μg of linearized template DNA, digoxigenin-11-UTP, and T3 or T7 RNA polymerase, according to the manufacturer's protocol (Boehringer Mannheim). 25S rRNA transcripts were localized using the 0.9-kb *Bam*HI-*Eco*RI insert from pTA71 subcloned into pBS-(KS+) (Gerlach and Bedbrook, 1979; Rode et al., 1987). Hybridized probe was detected using anti-digoxigenin-alkaline phosphatase conjugate and overnight color development with nitroblue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl-phosphate as described by Marrison and Leech (1994). Starch was stained in rehydrated tissue sections using Lugol's iodine (Sigma) (6% [w/v] KI and 4% [w/v] iodine, diluted 4-fold before use).

Immunolocalization was carried out as described by Marrison and Leech (1992). Antiserum raised against the OLI protein was used to detect the CHL H protein. Sections were incubated overnight at 4°C with a 1:200 dilution of OLI or CH-42 antiserum in 0.5% (w/v) BSA/PBS or in 0.5% (w/v) BSA/PBS alone (control). CHL H was visualized using FITC-conjugated goat anti-mouse antiserum and CH-42 was visualized using FITC-conjugated anti-rabbit antiserum (Sigma) diluted in BSA/PBS as recommended by the supplier. Sections were viewed using a Nikon FXA microscope with epifluorescence attachment, high-pressure mercury lamp and filter combination dichroic mirror 510, 450- to 490-nm excitation filter, and barrier filter 515IF. Photomicrographs were taken using Kodak Ektachrome 400 color slide film with automatic exposure setting.

RESULTS

Sequence Analysis

The sequence of the *CHL H* cDNA was determined, and the data are shown in Figure 2A. The gene contains 1382 codons, which corresponds to a predicted molecular weight of 153,507 for the CHL H preprotein. The overall level of identity with the *Antirrhinum* OLI protein is 85%, and CHL H also has 41 and 40% identity with the BchH proteins from *R. sphaeroides* and *R. capsulatus*, re-

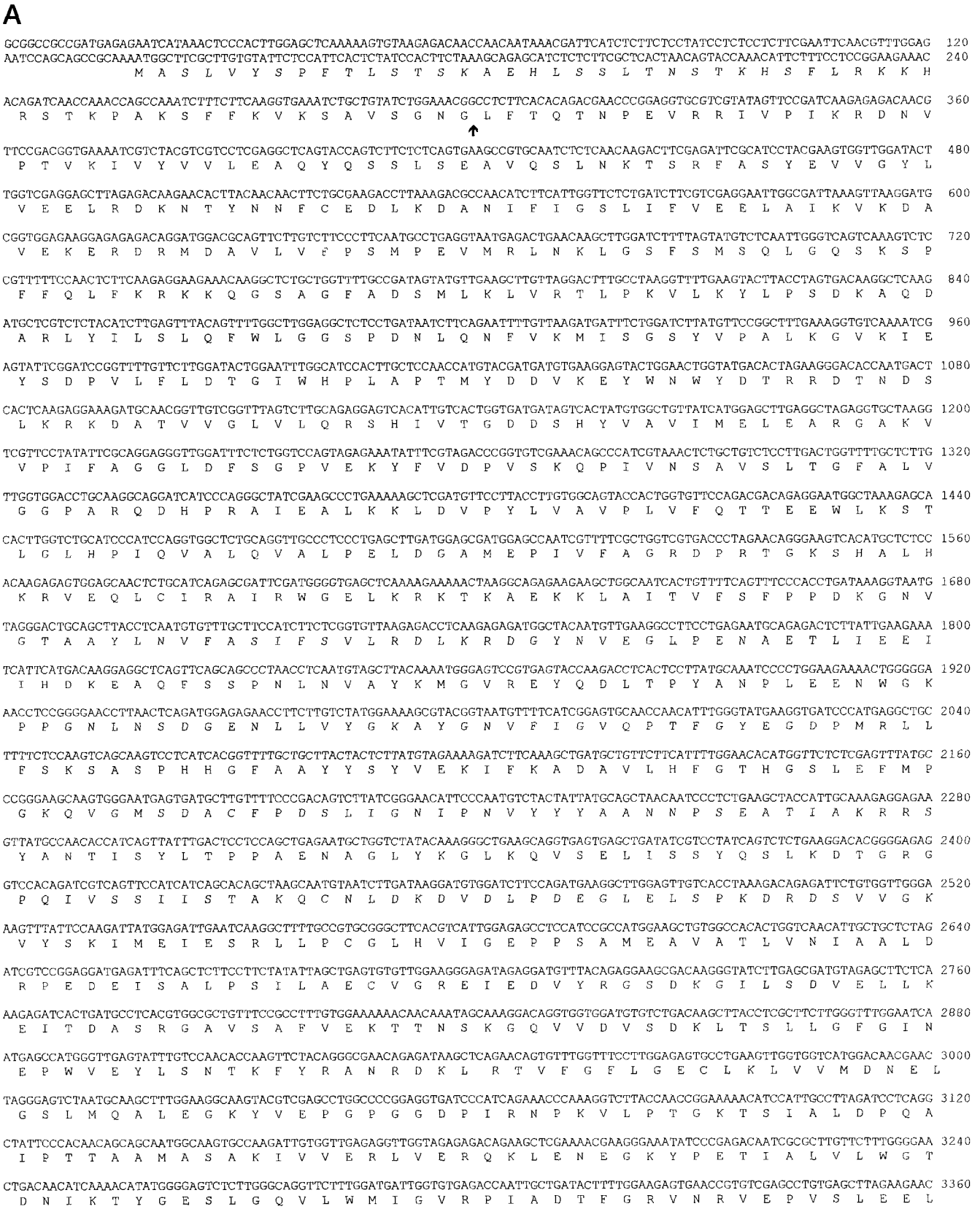

Figure 2. (Continued on facing page.)

A

```
TAGGAAGGCCGAGGATCGATGTAGTTGTTAACTGCTCAGGGGTCTTCCGTGATCTCTTTATCAACCAGATGAACCTTCTTGACCGAGCTATCAAGATGGTGGCGGACGTAGATGAGCCCG 3480
  G  R  P  R  I  D  V  V  V  N  C  S  G  V  F  R  D  L  F  I  N  Q  M  N  L  L  D  R  A  I  K  M  V  A  D  V  D  E  P  V

TAGAGCAAAATTTTGTAAGGAAACACGCGTTGGAACAAGCAGAGGCGCTTGGCATTGATATTAGAGAGGCAGCGACAAGAGTTTTCTCAAACGCTTCAGGGTCATACTCAGCCAACATCA 3600
  E  Q  N  F  V  R  K  H  A  L  E  Q  A  E  A  L  G  I  D  I  R  E  A  A  T  R  V  F  S  N  A  S  G  S  Y  S  A  N  I  S

GTCTTGCTGTTGAAAACTCGTCATGGAACGATGAGAAACAGCTTCAGGACATGTACTTGAGCCGCAAATCGTTTGCGTTTGATAGTGATGCTCCTGGAGCAGGAATGGCTGAGAAGAAGC 3720
  L  A  V  E  N  S  S  W  N  D  E  K  Q  L  Q  D  M  Y  L  S  R  K  S  F  A  F  D  S  D  A  P  G  A  G  M  A  E  K  K  Q

AGGTCTTTGAGATGGCTCTTAGCACTGCAGAAGTCACCTTCCAGAACCTGGATTCTTCAGAGATTTCTTTGACTGATGTGAGCCACTACTTCGATTCTGACCCTACAAATCTAGTTCAGA 3840
  V  F  E  M  A  L  S  T  A  E  V  T  F  Q  N  L  D  S  S  E  I  S  L  T  D  V  S  H  Y  F  D  S  D  P  T  N  L  V  Q  S

GTTTGAGGAAGGATAAGAAGAAACCAAGCTCTTACATTGCTGACACTACAACTGCAAACGCGCAGGTGAGGACACTATCTGAGACAGTGAGGCTGGACGCAAGAACAAAGCTGCTGAATC 3960
  L  R  K  D  K  K  K  P  S  S  Y  I  A  D  T  T  T  A  N  A  Q  V  R  T  L  S  E  T  V  R  L  D  A  R  T  K  L  L  N  P

CAAAGTGGTACGAAGGAATGATGTCAAGTGGATATGAAGGAGTTCGTGAGATAGAGAAGAGACTGTCCAACACTGTGGGATGGAGTGCAACGTCAGGTCAAGTAGACAATTGGGTCTACG 4080
  K  W  Y  E  G  M  M  S  S  G  Y  E  G  V  R  E  I  E  K  R  L  S  N  T  V  G  W  S  A  T  S  G  Q  V  D  N  W  V  Y  E

AGGAGGCCAACTCAACTTTCATCCAAGACGAGGAGATGCTGAACCGTCTCATGAACCACCAATCCCAACTCCTTCAGGAAAATGCTTCAGACTTTCTTGGAGGCCAATGGTCGTGGCTACT 4200
  E  A  N  S  T  F  I  Q  D  E  E  M  L  N  R  L  M  N  T  N  P  N  S  F  R  K  M  L  Q  T  F  L  E  A  N  G  R  G  Y  W

GGGACACTTCCGCTGAAAACATAGAGAAGCTCAAGGAATTGTACTCGCAGGTGGAAGACAAGATCGAAGGGATCGATCGATAAACAATGGGATATAAGCCTTTCTTCTTGTAAATGAACT 4320
  D  T  S  A  E  N  I  E  K  L  K  E  L  Y  S  Q  V  E  D  K  I  E  G  I  D  R  *

TAAAACTTTATTTTTGTGTATTACAGTATTCTACCAAACGACATCATCTCTCTGTGAATAAGAGAATGTTTTGGCTGACATTTTCTGACTTTAAAAAAAAAAAAAAAAA 4428
```

B

```
At   ·MASLVYSPFT LSTSKAEHLS SLTNSTKHSF LRKKHRSTKP AKSFFKVKSA  50
Syn  .......... .......... .......... .......... ..........
Rs   .......... .......... .......... .......... ..........
Rc   .......... .......... .......... .......... ..........

At   VSGNGLFTQT NPEVRRIVPI KRDNVPTVKI VYVVLEAQYQ SSLSEAVQSL 100
Syn  .....MFTNV KSTIPRVDPE ALNGRQLLKV VYVVLESQYQ SALSAAVRNI  45
Rs   .........MH GE.VSGPAGT P.....GYSI AIVTLDAHAA GPVARIAPRL  36
Rc   .........MH DESMSGTMPL PPHRPGGYNV VIITLDQHAA GPAARALPRL  42
```

Figure 2. A, DNA sequence of the *CHL H* cDNA (upper line) and deduced amino acid sequence (lower line). The arrow shows the putative processing site of the transit peptide. B, N-terminal portion of an alignment between Arabidopsis CHL H (At), *Synechocystsis* ChlH (Syn), *R. sphaeroides* BchH (Rs), and *R. capsulatus* BchH (Rc). Conserved amino acid residues are in boldface.

spectively. CHL H also exhibits 65% identity with ChlH, a putative Mg chelatase subunit from *Synechocystis* PCC 6803 (Jensen et al., 1996a). Hudson et al. (1993) noted the strong similarity between the OLI protein and CobN, which is a subunit of the cobaltochelatase enzyme from *Pseudomonas dentrificans* (Crouzet et al., 1991). Similarly, the *Arabidopsis* CHL H protein shows 27% identity with CobN. The significance of these homologies to the role played by the CHL H protein will be commented on in "Discussion."

Figure 2B shows an alignment between the N termini of CHL H from Arabidopsis, the *Synechocystis* ChlH protein, and the *Rhodobacter* BchH proteins. Given that the overall level of similarity between all of these proteins is quite high, the CHL H protein appears to have a 55-amino acid extension at the N terminus that may constitute the transit peptide. Indeed, this sequence does have typical features of a stroma-targeting signal because it has a high proportion of Ser/Thr and an overall positive charge (von Heijne et al., 1989) and the motif MASLV at the N terminus of the protein is similar to that found in several other chloroplast transit peptides (Karlin-Neumann and Tobin, 1986). There does not appear to be a motif at the N terminus that fits with the cleavage-site motif suggested by Gavel and von Heijne (1990). Therefore, we suggest that the processing site may be at the point where the N terminus of the *Synechocystis* protein aligns with CHL H; this is marked by an arrow in Figure 2A.

Northern Analysis of *CHL H* Expression

Figure 3 shows an analysis of *CHL H* expression by northern blotting that demonstrates that a large transcript of approximately 4.5 kb is present in leaves but not in roots. In addition, an investigation of fluctuations in transcript levels throughout a diurnal cycle was carried out. RNA was prepared from plants harvested at 2-h intervals during an 8-h dark/16-h light cycle, blotted, and probed with the *CHL H* cDNA. An identical blot was probed with the tubulin cDNA to provide a constitutive control; the level of this transcript is constant, indicating that the amounts of different RNA samples on the blot are essentially the same. The level of the *CHL H* transcript increased strongly in the dark, almost reaching its maximum level by the end of the dark period. A further increase in the transcript level was observed slightly after the onset of the light, rising to a maximum after 1 h and then decreasing gradually during the light period. We also probed an identical blot with the

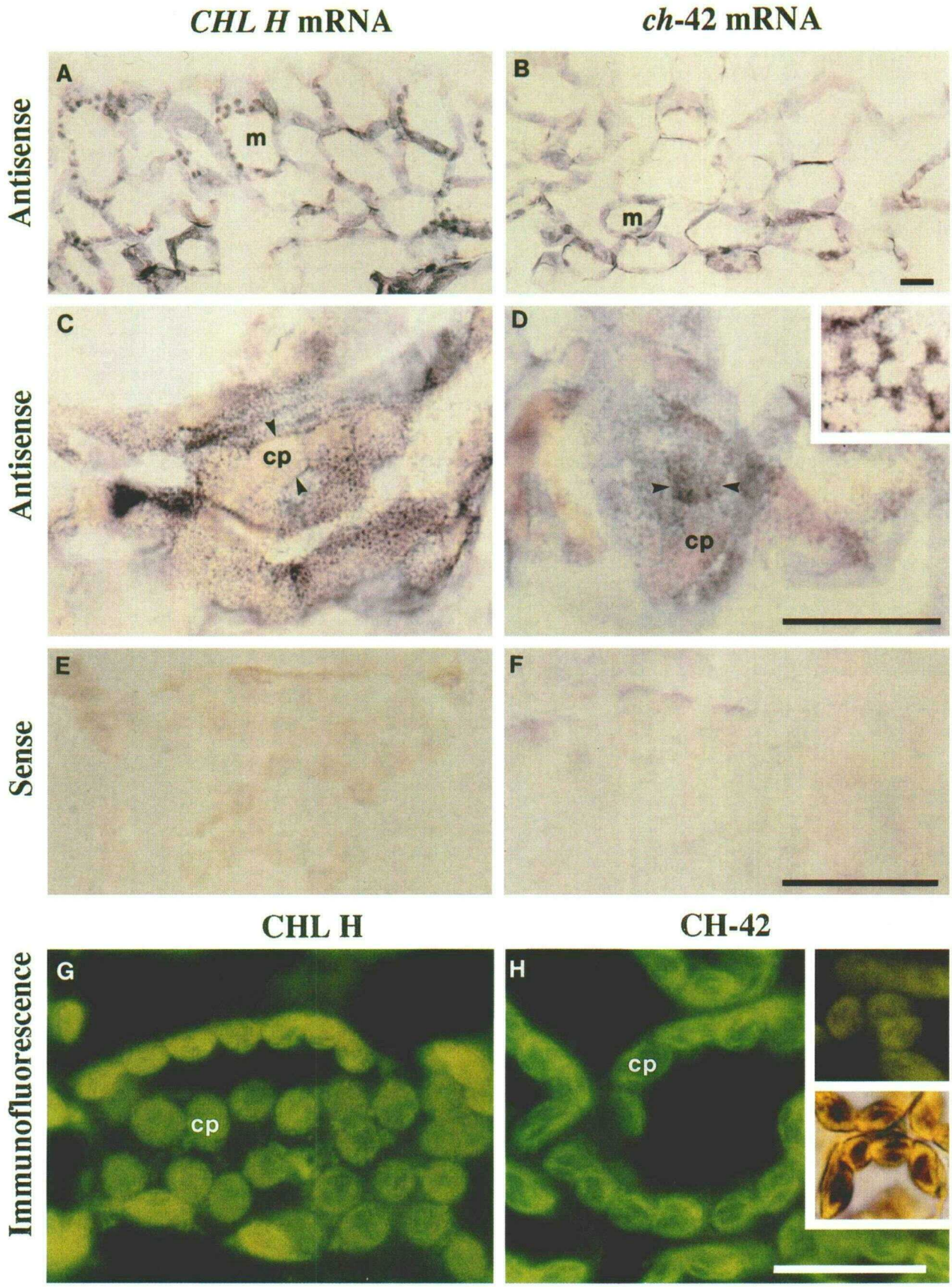

Figure 4. (Legend appears on facing page.)

CHL H genes. After hybridization with the sense strand of *CHL H* and *ch*-42 RNA probes (used as controls) and color detection, no positive mRNA staining was observed and the chloroplasts appeared pink in cross-section (Fig. 4, E and F).

The subcellular location of chlorophyll biosynthetic enzymes was also determined by probing thin leaf sections of Arabidopsis with antibodies raised against the OLI and CH-42 proteins (Fig. 4, G and H). The results show that these biosynthetic enzymes are confined to the chloroplasts in every mesophyll cell within the leaf; thus, every cell and every chloroplast has the potential to synthesize chlorophyll. The insets in Figure 4H show the tissue after incubation with 0.5% (w/v) BSA/PBS and FITC (immunolabeling control, upper inset) and after incubation with Lugol's iodine (lower inset), illustrating the starch grains within the chloroplasts (brown precipitate). The starch grains account for the characteristic striated pattern seen in Arabidopsis chloroplasts of this age.

Import of the Nascent CHL H Preprotein into Chloroplasts

The use of an import assay followed by fractionation of the chloroplast provides more information on the organellar location of proteins. An in vitro study was carried out using pea chloroplasts to demonstrate that the CHL H protein is imported into isolated chloroplasts. Using the full-length cDNA clone as a template, transcripts were produced in vitro from the T7 promoter, and this in turn was translated in a rabbit reticulocyte lysate in the presence of [^{35}S]Met to produce labeled CHL H precursor protein

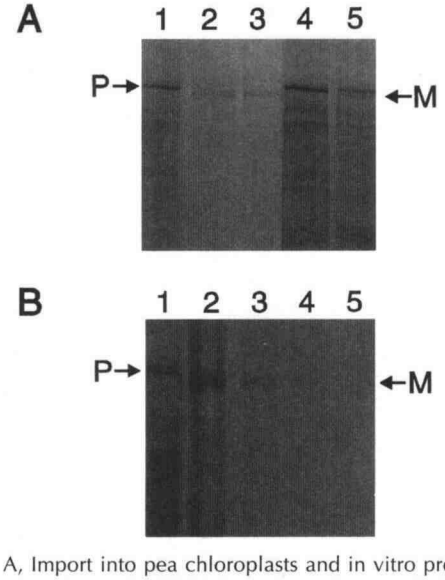

Figure 5. A, Import into pea chloroplasts and in vitro processing of the CHL H protein. Lanes 1 and 4, In vitro-translated CHL H precursor; lane 2, washed chloroplasts analyzed after import of the CHL H protein; lane 3, protease-treated chloroplasts; lane 5, CHL H protein after incubation with partially purified stromal processing peptidase. B, Localization of the CHL H protein within chloroplasts. Lane 1, CHL H precursor; lane 2, washed chloroplasts; lane 3, protease-treated chloroplasts; lane 4, stromal fraction; lane 5, membrane fraction. P, Precursor protein; M, mature protein.

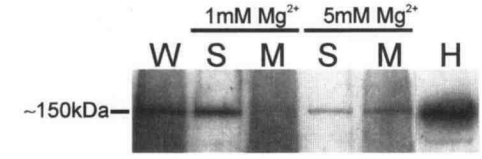

Figure 6. Immunoblot analysis of the location of the CHL H protein. W, Whole chloroplasts; S, stromal fractions; M, membrane fractions; H, *E. coli* extract containing expressed CHL H protein. The concentration of Mg^{2+} in the chloroplast lysis buffer is indicated.

(Fig. 5A, lanes 1 and 4). Figure 5A shows that upon incubation with chloroplasts the protein is processed to a smaller size (lane 2). This protein band can also be seen after protease treatment of the chloroplasts (lane 3), indicating that the protein was protected from digestion and, therefore, was imported. Furthermore, the CHL H protein was cleaved in vitro by the partially purified stromal processing protease; lane 5 shows that approximately 50% of the labeled precursor was processed to a smaller protein. In Figure 5B the location of the protein within the chloroplast was investigated; after the chloroplasts were incubated with labeled CHL H, the chloroplasts were washed, lysed, and separated into stroma and membrane fractions. A band corresponding to the mature-sized protein can be seen in the stroma fraction (lane 4), whereas no such band could be detected in the membrane fraction (lane 5).

Localization of CHL H by Immunoblotting

To confirm the location of CHL H within the chloroplast, Arabidopsis chloroplasts were prepared and fractionated into stroma and membrane fractions. Walker and Weinstein (1991b) demonstrated that Mg chelatase activity required soluble and membrane fractions of lysed pea chloroplasts. They later reported that Mg chelatase activity was rendered entirely soluble if pea chloroplasts were lysed in buffers containing low concentrations of Mg^{2+} (Walker and Weinstein, 1995). Accordingly, the Arabidopsis chloroplasts were lysed in buffer containing either 1 or 5 mM Mg^{2+}. The immunoblot in Figure 6 shows that in 1 mM Mg^{2+}, CHL H is largely present in the stroma fraction, whereas in 5 mM Mg^{2+}, a significant proportion of the CHL H protein is membrane associated.

DISCUSSION

We have cloned and sequenced a putative Mg chelatase subunit, *CHL H*, from *A. thaliana*. Hudson et al. (1993) have cloned the equivalent gene from *Antirrhinum majus* and noted its likely importance in chlorophyll synthesis, since they were able to correlate a mutation in the gene with a chlorophyll-deficient phenotype. They speculated that it could encode a Mg chelatase subunit due to homologies with two bacterial genes involved in tetrapyrrole-metal chelation. One of these bacterial genes, *bchH*, has now been positively identified as a Mg chelatase subunit (Gibson et al., 1995). Furthermore, it seems likely that BchH binds the protoporphyrin substrate during the Mg-chelation reaction (Gibson et al., 1995). The other bacterial protein, CobN from *Pseudomonas denitrificans*, is involved in the chelation

of cobalt into a porphyrin during coenzyme B_{12} biosynthesis and is also responsible for binding the porphyrin substrate (Debussche et al., 1992). The striking similarities between CHL H and BchH/CobN suggest that CHL H is a component of Mg chelatase in Arabidopsis and is presumably the protoporphyrin-binding component in this reaction. The extremely high level of identity between CHL H and OLI (85%), CHL H and *Synechocystis* ChlH (65%), and CHL H and BchH (40%) indicate that these Mg chelatase subunits are members of a highly conserved class of proteins. This is borne out by the fact that CH-42, another protein implicated in chlorophyll biosynthesis (Koncz et al., 1990; Armstrong et al., 1993), shows 50% identity to BchI, which is also a bacterial Mg chelatase subunit (Gibson et al., 1995). These proteins have evidently been highly conserved through evolution.

In Figure 3 it was shown that expression of *CHL H* is confined to photosynthetic tissues, since the transcript was detected in leaves but not in roots. This would be expected for a gene involved in the formation of the photosynthetic apparatus. In addition, the transcript undergoes a dramatic diurnal variation, rising almost to its maximum level by the end of the dark period, then increasing slightly at the onset of the light and declining steadily to a minimum by the end of the light period. This is somewhat similar to the behavior of the *olive* transcript from *Antirrhinum*, although the increase of *olive* mRNA in the dark was much more rapid. If the transcript levels reflect the amount of protein in the cell and CHL H is indeed the protoporphyrin-binding subunit, this may be a mechanism to regulate the amount of chlorophyll (see below). It is intriguing that *ch*-42 transcript levels remain almost constant. Koncz et al. (1990) demonstrated that *ch*-42 transcription was induced strongly in dark-adapted plants that were subsequently placed in the light. However, it seems that in plants that are established in a day/night cycle there is no response of this transcript to light, and we could reasonably assume that CH-42 levels remain constant.

Hudson et al. (1993) proposed that OLI could have a regulatory function. In view of the biochemical evidence that the bacterial homolog, BchH, binds protoporphyrin (Gibson et al., 1995), a mechanism is proposed below that involves such a binding reaction between CHL H and protoporphyrin in regulating flux down the tetrapyrrole biosynthetic pathways, primarily the chlorophyll branch. At the end of the dark period, the level of CHL H transcript has almost achieved its maximum, and the assumption made in this model is that the level of cognate protein also varies in the same manner as the transcript. The elevated level of CHL H is proposed to sequester protoporphyrin IX, thus depleting the pool of substrate for ferrochelatase, the level of which does not vary. The consequent drop in heme levels stimulates flux down the tetrapyrrole biosynthetic pathway, since heme has been shown to repress the activity of glutamyl-tRNA reductase (Pontoppidan and Kannangara, 1994). Therefore, at the end of the dark period the overall effect of increasing CHL H levels would be to stimulate the production of protoporphyrin, with a consequent effect on chlorophyll biosynthesis.

We propose that at the end of the light period the opposite situation occurs, with lowered levels of CHL H and lowered δ-aminolevulinic acid and chlorophyll biosynthesis. Thus, CHL H could play an important role in matching the requirement for chlorophyll synthesis with fluctuations in light conditions. More work will have to be performed on the patterns of *CHL H* transcripts, but it is interesting to note that the activity of Mg chelatase does exhibit a diurnal variation in barley (Jensen et al., 1996b). Furthermore, there is some similarity between *CHL H* and *cab6* in that both transcripts display significant oscillations of mRNA accumulation and both transcripts rise in the dark independently of light. This is often the case with transcripts that are under the influence of a circadian rhythm, such as the Arabidopsis PSII *cab* genes (Kay and Millar, 1992). PSI *cab* genes from Arabidopsis are less well characterized, but PSI *cab* genes are known to be under circadian control in tomato (Kellmann et al., 1993). This apparent synchronization of *CHL H* and *cab* transcript levels might reflect the fact that the LHC polypeptides are the major chlorophyll-binding proteins in the chloroplast.

In situ hybridization revealed that *CHL H* and *ch*-42 transcripts are localized over the surface of the chloroplasts. To our knowledge, this is the first visualization of transcripts encoding chloroplast proteins preferentially accumulating at the chloroplast surface (Fig. 4, C and D). By analogy with the situation in mitochondria (Verner, 1993), this could indicate that the translation of *CHL H* and *ch*-42 transcripts is coupled to the import of the nascent polypeptides. It supports the view that within the mesophyll cell a co-translational mechanism may at least partly account for the import of polypeptides into the chloroplast. However, more work is needed to determine whether this is a general rule for other nuclear-encoded chloroplast proteins or whether this is a specific feature of the import of chlorophyll biosynthetic enzymes. The immunolocalization study (Fig. 4, G and H) clearly shows that both proteins are found inside the chloroplast. Previous immunofluorescence studies demonstrated that it is possible to identify Rubisco in chloroplasts and to correlate this with the distribution of transcripts encoding this protein (Marrison and Leech, 1994). The data in Figure 4 extend this technique to provide the first definitive identification of the localization of a chlorophyll biosynthetic enzyme within the chloroplast. This kind of approach opens the way to a description of transcript and protein levels for the chlorophyll biosynthetic pathway within developing plant tissue. The large number of starch grains within the chloroplasts (Fig. 4H, lower inset) give rise to the striated staining pattern that is seen in Figure 4, G and H. This pattern does not reflect the position of the thylakoid membranes within the chloroplasts.

The chloroplast import data in Figure 5 show that the CHL H preprotein is processed following import into the chloroplast. Since this protein is so large, it is not possible to make an accurate estimate of the size of the mature protein, although by comparison with *Synechocystis* and *Rhodobacter* sequences a possible cleavage site can be deduced (Fig. 2). Indeed, it is remarkable that the transcript

encoding a preprotein of this size was translated in vitro. This experiment, using a heterologous system, suggested that the CHL H protein is located in the stroma (Fig. 5B). However, immunoblotting of fractionated Arabidopsis chloroplasts with OLI antibodies (Fig. 6) shows that, at least in vitro, the distribution of CHL H between the stromal and membrane fractions depends on the concentration of Mg^{2+}. It appears that the interaction between CHL H and the membrane is only peripheral, since CHL H is largely soluble at low Mg^{2+} concentration. It is possible that the Mg^{2+}-dependent distribution of Mg chelatase activity observed by Walker and Weinstein (1995) is due to the CHL H protein becoming membrane associated in high Mg^{2+} concentrations. In the light of our demonstration that association and dissociation of CHL H with the membrane appears to take place in the 1 to 5 mM range of Mg^{2+} concentration, it is interesting to note that the concentration of free Mg^{2+} within the chloroplast varies between 1 and 6 mM (reviewed by Leegood et al., 1985). Koncz et al. (1990) reported that CH-42 is a stromal protein, and the equivalent protein from soybean has been shown to reside in the stroma (Nakayama et al., 1995). Assuming that both CHL H and CH-42 are Mg chelatase subunits, they must presumably interact with one another. It will require further study to determine the nature of the association between CHL H and CH-42.

ACKNOWLEDGMENTS

This work benefitted from the use of the SEQNET facility of the Biotechnology and Biological Sciences Research Council at Daresbury. We would particularly like to thank Dr. Andrew Hudson for providing us with the *olive* cDNA and the OLI antibodies. We would also like to thank Dr. Alison Smith for providing us with the Arabidopsis ferrochelatase cDNA, Dr. Czaba Koncz for providing the *ch*-42 cDNA, Professor Chris Leaver for the wheat 25S rRNA gene, and Dr. Gamini Kannangara for help in production of the CH-42 antibodies.

Received October 2, 1995; accepted February 20, 1996.
Copyright Clearance Center: 0032-0889/96/111/0061/11.
The accession number for the sequence reported in this article is Z68495.

LITERATURE CITED

Armstrong GA, Cook DN, Ma D, Alberti M, Burke DH, Hearst JE (1993) Regulation of carotenoid and bacteriochlorophyll biosynthesis genes and identification of an evolutionarily conserved gene required for bacteriochlorophyll accumulation. J Gen Microbiol **139:** 897–906

Bauer CE, Bollivar DW, Suzuki JY (1993) Genetic analysis of photopigment biosynthesis in eubacteria: a guiding light for algae and plants. J Bacteriol **175:** 3919–3925

Bollivar DW, Suzuki JY, Beatty JT, Dobrowolski JM, Bauer CE (1994) Directed mutational analysis of bacteriochlorophyll a biosynthesis in *Rhodobacter capsulatus*. J Mol Biol **237:** 622–640

Coomber SA, Chaudri M, Connor A, Britton G, Hunter CN (1990) Localised transposon Tn5 mutagenesis of the photosynthetic gene cluster of *Rhodobacter sphaeroides*. Mol Microbiol **4:** 977–989

Coomber SA, Hunter CN (1989) Construction of a physical map of the 45 kb photosynthetic gene cluster of *Rhodobacter sphaeroides*. Arch Microbiol **151:** 454–458

Crouzet J, Levy-Schil S, Cameron B, Cauchois L, Rigault S, Rouyez M-C, Blanche F, Debussche L, Thibaut D (1991) Nucleotide sequence and genetic analysis of a 13.1 kilobase-pair *Pseudomonas denitrificans* DNA fragment containing five *cob* genes and identification of structural genes encoding Cob(I-)alamin adenosyl transferase, cobyric acid synthase, and bifunctional cobinamide kinase-cobinamide phosphate guanylyltransferase. J Bacteriol **173:** 6074–6087

Debussche L, Couder M, Thibaut D, Cameron B, Crouzet J, Blanche F (1992) Assay, purification and characterisation of cobaltochelatase, a unique complex enzyme catalyzing cobalt insertion in hydrogenobyrinic acid *a,c*-diamide during coenzyme B_{12} biosynthesis in *Pseudomonas denitrificans*. J Bacteriol **174:** 7445–7451

Fuesler TP, Wright LA, Castelfranco PA (1981) Properties of magnesium chelatase in greening etioplasts. Plant Physiol **67:** 246–249

Gavel Y, von Heijne G (1990) A conserved cleavage-site motif in chloroplast transit peptides. FEBS Lett **261:** 455–458

Gerlach WL, Bedbrook JR (1979) Cloning and characterisation of ribosomal RNA genes from wheat and barley. Nucleic Acids Res **7:** 1869–1885

Gibson LCD, Hunter CN (1994) The bacteriochlorophyll biosynthesis gene, *bchM*, of *Rhodobacter sphaeroides* encodes S-adenosyl-L-methionine:Mg protoporphyrin methyltransferase. FEBS Lett **352:** 127–130

Gibson LCD, Willows RD, Kannangara CG, von Wettstein D, Hunter CN (1995) Magnesium-protoporphyrin chelatase of *Rhodobacter sphaeroides*: reconstitution of activity by combining the products of the *bchH, -I* and *-D* genes expressed in *Escherichia coli*. Proc Natl Acad Sci USA **92:** 1941–1944

Guiltinan MJ, Ma D-P, Barker RF, Bustos MM, Cyr RJ, Yadegari R, Fosket DE (1987) Isolation, characterisation, and sequence of two divergent β-tubulin genes from soybean (*Glycine max* L.). Plant Mol Biol **10:** 171–184

Henikoff S (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:** 351–359

Hudson A, Carpenter R, Doyle S, Coen ES (1993) *Olive*: a key gene required for chlorophyll biosynthesis in *Antirrhinum majus*. EMBO J **12:** 3711–3719

Jensen PE, Kristensen M, Hoff T, Lehmbeck J, Stummann BM, Henningsen KW (1992) Identification of a single copy gene encoding a type I chlorophyll *a/b* binding polypeptide of photosystem I in *Arabidopsis thaliana*. Physiol Plant **84:** 561–567

Jensen PE, Stummann BM, Henningsen KW (1996a) Sequence announcement for *Synechocystis* PCC6803 *chlH*, accession number U29131. Plant Mol Biol (in press)

Jensen PE, Willows RD, Petersen BL, Vothnecht UC, Stummann BM, Kannangara CG, von Wettstein D, Henningsen KW (1996b) Genes for Mg chelatase subunits in barley: *Xantha-f, -g* and *-h*. Mol Gen Genet **250:** 383–394

Karlin-Neumann GA, Tobin EM (1986) Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework. EMBO J **5:** 9–13

Kay SA, Millar AJ (1992) Circadian regulated *cab* gene transcription in higher plants. *In* M Young, ed, The Molecular Biology of Circadian Rhythms. Marcel Dekker, New York, pp 73–89

Kellmann J-W, Merforth N, Wiese M, Pichersky E, Piechulla B (1993) Concerted circadian oscillations in transcript levels of nineteen *Lha/b* (*cab*) genes in *Lycopersicon esculentum* (tomato). Mol Gen Genet **237:** 439–448

Koncz C, Mayerhofer R, Koncz-Kalman S, Nawrath C, Reiss B, Rédei GP, Schell J (1990) Isolation of a gene encoding a novel chloroplast protein by T-DNA tagging in *Arabidopsis thaliana*. EMBO J **9:** 1337–1346

Kunst K, Browse J, Somerville C (1988) Altered regulation of lipid biosynthesis in a mutant of *Arabidopsis* deficient in chloroplast glycerol-3-phosphate acyltransferase activity. Proc Natl Acad Sci USA **85:** 4143–4147

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. Nature **227:** 680–685

Leegood RC, Walker DA, Foyer CH (1985) Regulation of the Benson-Calvin cycle. *In* J Barber, NR Baker, eds, Photosynthetic

Mechanisms and the Environment. Elsevier Science Publishers, Amsterdam, The Netherlands, pp 190–258

Marrison JL, Leech RM (1994) The sub-cellular and intra-organelle recognition of nuclear and chloroplast transcripts in developing leaf cells. Plant J **6:** 605–614

Marrs B (1981) Mobilisation of the genes for photosynthesis from *Rhodopseudomonas capsulata* by a promiscuous plasmid. J Bacteriol **146:** 1003–1012

Musgrove JE, Elderfield PD, Robinson C (1989) Proteases in the stroma and thylakoids of pea chloroplasts. Plant Physiol **90:** 1616–1621

Nakayama M, Masuda T, Sato N, Yamagata H, Bowler C, Ohta H, Shioi Y, Takamiya K-i (1995) Cloning, subcellular localisation and expression of *chlI*, a subunit of magnesium chelatase in soybean. Biochem Biophys Res Commun **215:** 422–428

Pontoppidan B, Kannangara CG (1994) Purification and partial characterization of barley glutamyl-tRNA^{Glu} reductase, the enzyme that directs glutamate to chlorophyll biosynthesis. Eur J Biochem **225:** 529–537

Robinson C, Barnett LK (1988) Isolation and analysis of chloroplasts. *In* CH Shaw, ed, Plant Molecular Biology: A Practical Approach. IRL Press, Oxford, UK, pp 67–78

Rode A, Hartmann C, Benslimane A, Picard E, Quetier F (1987) Gametoclonal variation detected in the nuclear ribosomal DNA from double haploid lines of a spring wheat (*Triticum aestivum* L., cv. 'cesar'). Theor Appl Genet **74:** 31–37

Rosenberg AH, Lade BN, Chui D-s, Lin S-W, Dunn JJ, Studier FW (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56:** 125–135

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning. A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger F, Nickler S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA **74:** 5463–5467

Schuler MA, Zielinski RE (1989) Methods in Plant Molecular Biology. Academic Press, New York, pp 89–96

Smith AG, Santana MA, Wallace-Cook ADM, Roper JM, Labbe-Bois R (1994) Isolation of a cDNA encoding chloroplast ferrochelatase from *Arabidopsis thaliana* by functional complementation of a yeast mutant. J Biol Chem **269:** 13405–13413

Studier FW, Moffat BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol **189:** 113–130

Verner K (1993) Co-translational protein import into mitochondria: an alternative view. Trends Biochem Sci **18:** 366–371

von Heijne G, Steppuhn J, Hermann RG (1989) Domain structure of mitochondrial and chloroplast targetting peptides. Eur J Biochem **180:** 535–545

Walker CJ, Weinstein JD (1991a) Further characterization of the magnesium chelatase in isolated developed cucumber chloroplasts. Plant Physiol **95:** 1189–1196

Walker CJ, Weinstein JD (1991b) *In vitro* assay of the chlorophyll biosynthetic enzyme Mg-chelatase: resolution of the activity into soluble and membrane-bound fractions. Proc Natl Acad Sci USA **88:** 5789–5793

Walker CJ, Weinstein JD (1994) The magnesium insertion step of chlorophyll biosynthesis is a two-stage reaction. Biochem J **299:** 277–284

Walker CJ, Weinstein JD (1995) Re-examination of the localization of Mg-chelatase within the chloroplast. Physiol Plant **94:** 419–424

Zsebo KM, Hearst JE (1984) Genetic-physical mapping of a photosynthetic gene cluster from *Rhodopseudomonas capsulata*. Cell **37:** 937–947