Plant Molecular Biology **27**: 607–617, 1995.

Structure and expression of the *Chlamydomonas reinhardtii alad* gene encoding the chlorophyll biosynthetic enzyme, δ-aminolevulinic acid dehydratase (porphobilinogen synthase)

Gail L. Matters and Samuel I. Beale*
*Division of Biology and Medicine, Brown University, Providence, Rhode Island 02912, USA (*author for correspondence)*

Received 1 September; accepted in revised form 9 January 1995

Key words: δ-aminolevulinic acid dehydratase, chlorophyll biosynthesis, light-regulation, porphobilinogen

Abstract

cDNA clones for the *alad* gene encoding the chlorophyll biosynthetic enzyme ALA dehydratase (ALAD) from *Chlamydomonas reinhardtii* were isolated by complementation of an *Escherichia coli* ALAD mutant (*hemB*). The *C. reinhardtii alad* gene encodes a protein that has 50 to 60% sequence identity with higher plant ALADs, and includes a putative Mg^{2+}-binding domain characteristic of plant ALADs. Multiple classes of ALAD cDNAs were identified which varied in the length of their 3′- untranslated region. Genomic Southern analysis, using an ALAD cDNA as a probe, indicates that it is a single-copy gene. This suggests that the differently sized ALAD cDNAS are not the products of separate genes, but that a primary ALAD transcript is polyadenylated at multiple sites. A time course determination of ALAD mRNA levels in 12-h light:12-h dark synchronized cultures shows a 7-fold increase in ALAD mRNA at 2 h into the light phase. The ALAD mRNA level gradually declines but continues to be detectable up to the beginning of the dark phase. ALAD enzyme activity increases 3-fold by 6 h into the light phase and remains high through 10 h. Thus, there is an increase in both ALAD mRNA level and ALAD enzyme activity during the light phase, corresponding to the previously observed increase in the rate of chlorophyll accumulation.

Introduction

Chlorophyll is the dominant tetrapyrrole pigment produced in the chloroplasts of plants and algae. The first committed chlorophyll precursor, δ-aminolevulinic acid (ALA), is produced from glutamate using a glutamyl-tRNA intermediate [1]. The enzymatic reactions from ALA to protoporphyrin IX are common to all aerobic organisms. The first step in this part of the pathway, the conversion of ALA to porphobilinogen (PBG), is catalyzed by ALA dehydratase (ALAD), which is

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers U19876 and U19877.

also called PBG synthase. ALAD has been characterized in several plant species, and the activity of this enzyme increases during greening of etiolated pea, radish and mustard seedling cotyledons [6, 11, 19]. The gene encoding ALAD has been cloned from pea, spinach, soybean and the cyanobacterium *Anacystis nidulans* [2, 9, 10, 16]. ALAD-encoding genes have also been cloned from nonphotosynthetic eukaryotes and prokaryotes such as *Escherichia coli* [12], organisms in which heme is the major tetrapyrrole end product. A soybean ALAD-encoding gene has been designated *alad* [10], whereas the homologous *E. coli* gene is *hemB* [12].

Light is required to initiate and maintain chlorophyll synthesis in angiosperms and some green algae. The effect of light on *alad* gene expression is varied. In contrast to the increase in ALAD activity during greening, no change in ALAD mRNA level was observed in greening pea leaves, suggesting light may regulate ALAD levels post-transcriptionally [2]. However in soybean, although a modest increase in ALAD mRNA occurs during greening, no change in ALAD activity or protein levels were detected [10]. In contrast, the pea enzyme PBG deaminase, which catalyzes the reaction following that catalyzed by ALAD in the tetrapyrrole biosynthetic pathway, is increased in both activity and mRNA level during greening [23].

We have used light:dark synchronized cultures of the unicellular green alga *Chlamydomonas reinhardtii* to study the effects of light on the expression of chlorophyll biosynthetic enzymes. These cells rapidly synthesize chlorophyll during the light phase, and show a light-induced increase in the expression of the chlorophyll biosynthetic gene *gsa*, which encodes glutamate-1-semialdehyde (GSA) aminotransferase (GSAT) [13]. GSAT is the enzyme just prior to ALAD in the chlorophyll biosynthetic pathway, and it converts GSA to ALA. In our continuing examination of chlorophyll biosynthetic genes from *C. reinhardtii*, we report the cloning and structure of cDNAs and a genomic DNA region that encode ALAD, and examine the cellular levels of ALAD mRNA and enzyme activity in light:dark synchronized cells. We have used the designation *alad*, as was used for the soybean ALAD-encoding gene [10], to indicate the ALAD-encoding gene in *C. reinhardtii*.

Materials and methods

Cloning and sequence analysis of C. reinhardtii *ALAD-encoding cDNAs*

E. coli hemB (RP523) [12] was obtained from the *E. coli* Genetic Stock Center (Yale University, New Haven, Conn., CGSC #7199) and was maintained on LB media supplemented with 4 μg/ml hemin. Competent *hemB* cells were transformed with a *C. reinhardtii* cDNA expression library [22] as previously described [13]. Colonies were selected on LB-ampicillin plates for heme-independent growth. DNA sequencing was done using a Sequenase kit (U.S. Biochemical) according to the manufacturer's protocol, and sequence information was compiled with The DNA Inspector IIe (Textco, West Lebanon, NH) sequence analysis program. *C. reinhardtii* strains were obtained from the *Chlamydomonas* Genetics Center (Duke University, Durham, NC). For genomic Southern analysis, restriction digests of 10 μg of *C. reinhardtii* DNA (strain CC-124) were separated on a 0.8% (w/v) agarose gel, blotted onto a nylon membrane (Nytran+, Schleicher and Schuell), and UV-crosslinked. Blots were probed with a nick-translated ALAD cDNA insert in 50% formamide/5× SSPE/2× Denhardt's reagent/5% sodium dodecylsulfate/50 μg/ml salmon testes DNA at 45 °C and washed in 2× SSPE/0.1% sodium dodecylsulfate at 50 °C followed by a final wash in 0.2× SSPE/0.1% sodium dodecylsulfate at 55 °C.

Analysis of alad *transcripts*

RNA was isolated from mid-exponential growth phase synchronous *C. reinhardtii* cells (strain CC-124) grown in TAP medium [5]. Cells were centrifugally sedimented and extracted in buffer

(50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM EDTA, 40 μg/ml Proteinase K) at room temperature for 20 min. Phenol:chloroform extractions were done until the interface was clear, and the total nucleic acids were precipitated with ethanol. After redissolving in H_2O, total RNA was precipitated by LiCl overnight at 4 °C. Oligo(dT) cellulose (BRL) chromatography was used to select poly(A^+) RNA for northern blots. Glyoxal-denatured poly(A^+) RNA (2 μg) was separated on a 1% (w/v) agarose gel in 10 mM Na-phosphate buffer, pH 7.0, blotted onto a nylon membrane and UV-crosslinked. Blots were hybridized and washed as described for Southern analysis. A constitutively expressed *C. reinhardtii* G protein β subunit-like probe, which hybridizes to a 1.2 Kb mRNA, was used to show approximately equal loadings of RNA [17]. Autoradiographs of multiple exposures were scanned using an LKB Ultroscan XL laser densitometer for quantitation within the linear response range of the film.

ALAD enzyme activity assays

Mid-exponential growth phase synchronous cultures (2×10^9 total cells) were harvested by centrifugation (3000 × ***g*** for 5 min). After resuspending in 10 ml of ALAD assay buffer (300 mM glycerol, 100 mM Tricine-NaOH pH 8.0, 10 mM $MgCl_2$, 1 mM dithiothreitol, 0.004% [w/v] phenylmethylsulfonyl fluoride), cells were broken by one pass through a French pressure cell at 20000 p.s.i., the homogenate was centrifuged for 30 min at 27000 × ***g***, and the supernatant was used for ALAD assays. The assay was begun by the addition of 5 mM ALA to the extracts (10 μl) in assay buffer, and incubated at 30 °C for 30 min. The reaction was stopped by cooling on ice and an equal volume of modified Ehrlich-Hg reagent [21] was added. Precipitated proteins and Hg-dithiothreitol were removed by centrifugation, and the absorbance of the supernatant at 555 nm was measured and used to calculate PBG production [9].

alad *genomic clones*

An ALAD cDNA was used to probe a genomic DNA library (strain CC-124) that was constructed as previously described [13]. Two overlapping clones hybridizing to the cDNA were isolated and restriction mapped. Restriction fragments containing the 5′ end of the cDNA and approximately 1.1 Kbp of upstream DNA were subcloned into pBluescript and sequenced. The transcription start site was defined by primer extension, using an oligonucleotide primer corresponding to the 5′ end of the longest cDNA.

Results

Structure of a C. reinhardtii *ALAD-encoding cDNA*

Several *C. reinhardtii* cDNA clones encoding ALAD were identified by complementation of *E. coli hemB* mutant RP523 [12]. The longest cDNA clone contains a 46-bp 5′-untranslated region and a 498-bp 3′-untranslated region (Fig. 1A). The encoded protein contains 390 amino acids and includes a presumed chloroplast transit peptide of 24 amino acids. The putative chloroplast transit peptide is similar in size and base composition to other *C. reinhardtii* stroma-targeting chloroplast transit peptides [4]. The initiation Met is assumed to be the first Met in the open reading frame, where it is preceded by stop codons in all three reading frames and lies within a favorable translation initiation context common to *C. reinhardtii* nuclear genes [14]. The deduced protein has a molecular weight of 42997 and is between 50 and 60% identical to higher plant ALADs. There is a well-conserved active-site Lys (residue 312), and a highly conserved (17 of 19 identical residues) Mg^{2+}-binding domain (residues 171 to 188) (Fig. 1B) characteristic of plant ALADs [2, 10]. An unusual feature of the *C. reinhardtii* ALAD protein sequence is an 8-amino-acid insertion not previously reported in any other prokaryotic or eukaryotic ALAD proteins (residues 273 to 280).

```
                                                    ATACAATTGACCGCGAAGATCTAGGGATTCACTAAGAACCGTAACA   46

ATGCAGATGATGCAGCGCAACGTTGTGGGCCAGCGCCCCGTCGCTGGCTCCCGCCGCTCGCTGGTGGTTGCCAAC  121
 M  Q  M  M  Q  R  N  V  V  G  Q  R  P  V  A  G  S  R  R  S  L  V  V  A↑ N   25

GTTGCGGAGGTGACCCGCCCCGCGGTCAGCACCAACGGCAAGCACCGGACTGGTGTGCCGGAGGGAACTCCCATC  196
 V  A  E  V  T  R  P  A  V  S  T  N  G  K  H  R  T  G  V  P  E  G  T  P  I   50

GTCACCCCTCAGGACCTGCCCTCGCGCCCTCGCCGCAACCGCCGCAGCGAGAGCTTCCGTGCTTCCGTTCGTGAG  271
 V  T  P  Q  D  L  P  S  R  P  R  R  N  R  R  S  E  S  F  R  A  S  V  R  E   75

GTGAACGTGTCGCCGGCCAACTTCATCCTGCCGATCTTCATCCACGAGGAGAGCAACCAGAACGTGCCCATCGCC  346
 V  N  V  S  P  A  N  F  I  L  P  I  F  I  H  E  E  S  N  Q  N  V  P  I  A   100

TCCATGCCTGGCATCAACCGCCTGGCGTATGGCAAGAACGTGATTGACTACGTTGCTGAGCCTCGCTCTTACGGT  421
 S  M  P  G  I  N  R  L  A  Y  G  K  N  V  I  D  Y  V  A  E  P  R  S  Y  G   125

GTCAACCAGGTCGTGGTTTTCCCCAAGACGCCCGACCACCTGAAGACGCAAACCGCGGAGGAGGCGTTCAACAAG  496
 V  N  Q  V  V  V  F  P  K  T  P  D  H  L  K  T  Q  T  A  E  E  A  F  N  K   150

AACGGCCTCAGCCAGCGCACGATCCGCCTGCTGAAGGACTCTTTCCCTGACCTGGAGGTGTACACGGACGTGGCT  571
 N  G  L  S  Q  R  T  I  R  L  L  K  D  S  F  P  D  L  E  V  Y  T  D  V  A   175

CTGGACCCCTACAACTCGGACGGCCACGACGGTATCGTGTCGGACGCCGGTGTGATCCTGAACGACGAGACCATC  646
 L  D  P  Y  N  S  D  G  H  D  G  I  V  S  D  A  G  V  I  L  N  D  E  T  I   200

GAGTACCTGTGCCGCCAGGCCGTGAGCCAGGCCGAGGCCGGTGCCGACGTGGTGTCGCCCTCTGACATGATGGAC  721
 E  Y  L  C  R  Q  A  V  S  Q  A  E  A  G  A  D  V  V  S  P  S  D  M  M  D   225

GGCCGCGTGGGCGCCATCCGCCGCGCCCTGGACCGCGAGGGCTTCACCAACGTGTCCATCATGTCCTACACCGCC  796
 G  R  V  G  A  I  R  R  A  L  D  R  E  G  F  T  N  V  S  I  M  S  Y  T  A   250

AAGTACGCCTCCGCCTACTACGGCCCCTTCCGTGACGCCCTGGCGTCCGCGCCCAAGCCCGGCCAGGCGCACCGG  871
 K  Y  A  S  A  Y  Y  G  P  F  R  D  A  L  A  S  A  P  K  P  G  Q  A  H  R   275

CGCATCCCCCCCCAACAAGAAGACCTACCAGATGGACCCCGCCAACTACCGCGAGGCCATCCGCGAGGCCAAGGCC  946
 R  I  P  P  N  K  K  T  Y  Q  M  D  P  A  N  Y  R  E  A  I  R  E  A  K  A   300

GACGAGGCCGAGGGCGCTGACATCATGATGGTCAAGCCCGGCATGCCGTACCTGGACGTGGTACGCCTGCTGCGT 1021
 D  E  A  E  G  A  D  I  M  M  V  K  P  G  M  P  Y  L  D  V  V  R  L  L  R   325

GAGACCAGCCCGCTGCCCGTGGCCGTGTACCACGTGTCGGGCGAGTACGCCATGCTCAAGGCGGCGGCGGAGCGC 1096
 E  T  S  P  L  P  V  A  V  Y  H  V  S  G  E  Y  A  M  L  K  A  A  A  E  R   350

GGCTGGCTGAACGAGAAGGATGCCGTGCTTGAGGCCATGACCTGCTTCCGCCGCGCCGGCGGTGACCTCATCCTC 1171
 G  W  L  N  E  K  D  A  V  L  E  A  M  T  C  F  R  R  A  G  G  D  L  I  L   375

ACCTACTACGGCATTGAGGCCTCCAAGTGGCTGGCGGGCGAGAAGTAAGCGGTCTGGTGGGCGGCTGCATGGGCG 1246
 T  Y  Y  G  I  E  A  S  K  W  L  A  G  E  K  *                              390

GCAGTGCCGGCAGATATGGAGTTGAGGAGCGGAGCTGCGGCCGGCGTGGACTGCGGAGTAAAGCGGTTGTATGGG 1321
AACCGGTGTGCGGTGCTTCGGGCGGTGAGAGAGCGTGGATGATACGAGTGGTGGAATTGCAGGTATGCCTTGTGC 1396
TTGCGTGTGGTAACGTGGTCTGTATGTTTATGTATAACCGGTGTGAGGACAAGGGACAACGAGCGGCAGGAGGGG 1471
TGCATACTTCCGCTTGGCACGCATTTTTGGGCGTCTCGGGTGGCGGATTGACTTGCACCCGGGGAATAGTTCAAA 1546
GACAGATGGCTGAAGTTGCCTTGCGGCGGCTACGTGACACGGAGACGTGAGAGAGGTGCGGCGATACCCGCGAGC 1621
AGGGGGGCAGGGGCATTGTTGCCAACGCAGCAAGACACGCGCGGCTGGCAAATCACAGGGTGTGCCCTCCCAAAG 1696
ATTGTTAGCTTCAAAGCTTTG                                                         1717
```

	Putative Metal-Binding Region	Bound Metal
Chlamydomonas reinhardtii	VYT**D**VAL**D**PYNS**D**GH**D**GIV	Mg^{2+}
Selaginella martensii	IYT**D**VAL**D**PYSS**D**GH**D**GIV	Mg^{2+}
Pea	IYT**D**VAL**D**PYSS**D**GH**D**GIV	Mg^{2+}
Soybean	IYT**D**VAL**D**PYSS**D**GH**D**GIV	Mg^{2+}
Tomato	IYT**D**VAL**D**PYSS**D**GH**D**GIV	Mg^{2+}
Spinach	IYT**D**VAL**D**PYYY**D**GH**D**GIV	Mg^{2+}
Bradyrhizobium japonicum	VLC**D**VAL**D**PFTS**H**G**HD**GLI	
Rhodobacter capsulatus	IMT**D**IAL**D**PYNANG**HD**GLV	
Anacystis nidulans R2	VAADT**C**L**C**EYTP**H**G**HC**GYL	Zn^{2+}
Methanothermus sociabilis	VITDV**C**L**C**QYTE**H**G**HC**GIV	Zn^{2+}
Escherichia coli	VMSDT**C**F**C**EYTS**H**G**HC**GVL	Zn^{2+}
Clostridium josui	VITDI**C**M**C**EYTS**H**G**HC**GIL	Zn^{2+}
Staphylococcus aureus	IVADT**C**L**C**EYTD**H**G**HC**GVI	Zn^{2+}
Bacillus subtilis	VVADT**C**L**C**EYTD**H**G**HC**GLV	Zn^{2+}
Yeast	IICDV**C**L**C**EYTS**H**G**HC**GVL	Zn^{2+}
Mouse	VACDV**C**L**C**PYTS**H**G**HC**GLL	Zn^{2+}
Rat	VACDV**C**L**C**PYTS**H**G**HC**GLL	Zn^{2+}
Human	VACDV**C**L**C**PYTS**H**G**HC**GLL	Zn^{2+}
Invariant residues	D GH G	

Fig. 1. A. Nucleotide sequence and deduced amino acid sequence of a *C. reinhardtii* cDNA encoding ALAD. The end of a potential chloroplast transit peptide is marked by the upward arrow. The singly underlined peptide has significant identity with a Mg^{2+}-binding domain common to plant ALADs but is unlike the Zn^{2+}-binding domain of prokaryotic and nonplant eukaryotic ALADs. An 8-amino-acid, in-frame insertion unique to the *C. reinhardtii* ALAD cDNA is doubly underlined. B. Comparison of the deduced amino acid sequence of the metal-binding domains of ALAD proteins from various sources. Residues proposed to coordinate the metal are shown in bold typeface.

Heterogeneity in ALAD cDNA 3′-untranslated regions

Several of the *C. reinhardtii* ALAD cDNAs isolated by functional complementation were found to differ in the size of their 3′-untranslated region (3′ UTR). Three independent classes of ALAD cDNAs were identified whose 3′ UTRs contain 337, 466, or 498 bp (Fig. 2A). Each class of ALAD cDNAs is polyadenylated, and the 3′ UTR sequence up to the polyadenylation site is identical in all the clones. However, the putative polyadenylation signal for *C. reinhardtii* nuclear genes (TGTAA), normally found 14 to 17 bp upstream of the poly(A) addition site, is not perfectly present in any of the cDNA classes and not present elsewhere in the 3′ UTR. To determine whether these cDNAs are the result of multiple polyadenylation sites of the ALAD primary transcript or alternatively whether they are the products of multiple *alad* genes, genomic Southern analysis was done. The longest 3′ UTR contains

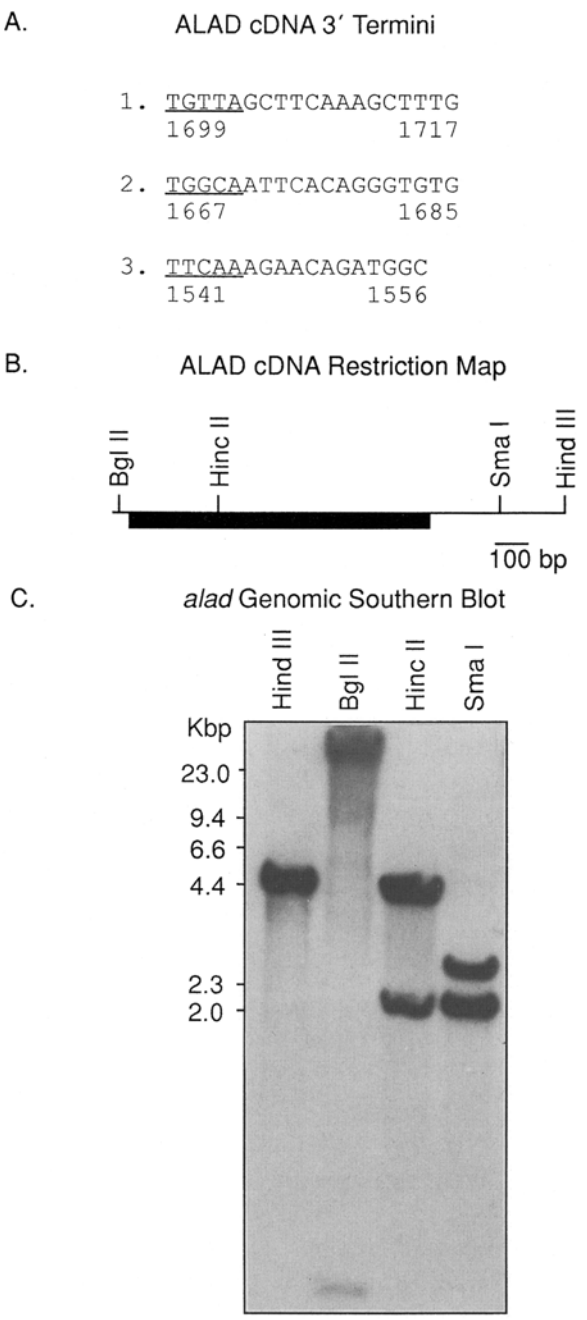


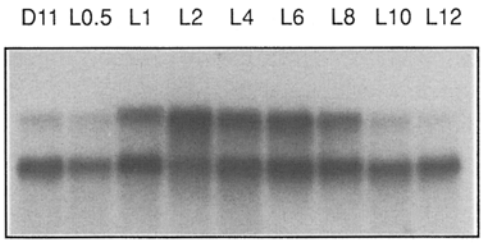
Fig. 2. A. Sequence comparison of the 3′ termini of 3 independent classes of ALAD cDNAs. A presumed polyadenylation signal for each class is underlined. B. A restriction map of the longest ALAD cDNA. C. Genomic Southern blot of *C. reinhardtii* strain CC-124 DNA digested with restriction enzymes that are known to cut the cDNA. DNA digests were separated on a 0.8% (w/v) agarose gel, blotted to a nylon membrane, and hybridized with a ^{32}P-labeled ALAD cDNA probe. Positions of DNA size markers are shown on the left.

a *Hind* III restriction site at the very end of the clone that is not present in the other cDNA classes. If the different cDNAs represent independent gene products, a Southern blot of *Hind* III digested genomic DNA probed with the ALAD cDNA should contain multiple bands. There is also one cDNA restriction site for *Bgl* II, *Hinc* II, and *Sma* I (Fig. 2B), and digests with these enzymes should show more than two bands on a Southern blot if multiple genes are present. Using the entire ALAD cDNA as a probe, a genomic Southern blot (strain CC-124 DNA) revealed a single hybridizing band in the *Hind* III-digested DNA and 2 bands in the DNA digested by *Bgl* II, *Hinc* II, or *Sma* I (Fig. 2C).

Pattern of ALAD mRNA accumulation and ALAD enzyme activity level

RNA was isolated from 12h-light:12 h-dark synchronized cultures at 1 h prior to the start of the light phase (D11), and at intervals during the light phase (L0.5 through L12). Northern analysis of these RNA samples, using the ALAD cDNA as a probe, shows that the ALAD mRNA level increases steadily from the dark level to a peak at 2 h into the light phase (L2) (Fig. 3). An increase in the ALAD mRNA level is detectable within 30 min after the light phase begins. After the peak is reached at 2 h, ALAD mRNA decreases gradually but continues to be present up to the end of the light phase. A densitometric scan of the northern blot was done to quantify the increase in the ALAD mRNA level, and each lane was normalized to the constitutively-expressed G protein signal. This analysis reveals a 7-fold increase in the ALAD mRNA level from the dark to the peak at 2 h of light.

For comparison to ALAD mRNA levels, whole cell extracts were made from a similar set of dark- and light-phase time points and were assayed for ALAD enzyme activity (Fig. 4). The ALAD activity level begins to rise by 2 h of light, and a sharp increase in ALAD activity occurs between 4 and 6 h into the light phase. The level of activity remains relatively unchanged from 6 through



	Relative Band Density									Normalized Induction Ratio	
	D11	L0.5	L1	L2	L4	L6	L8	L10	L12	L0.5/D11	L2/D11
ALAD	0.66	0.63	2.16	4.39	3.50	3.59	2.07	0.58	0.37		
G Protein	3.55	2.51	3.80	3.44	4.30	4.54	3.89	2.92	2.94		
ALAD/G Protein	0.19	0.25	0.57	1.28	0.82	0.79	0.53	0.20	0.13	1.36	6.90

Fig. 3. Northern blot of poly $(A)^+$ RNA (2 μg per lane) isolated from light:dark synchronized cultures at 1 h prior to the light phase (D11) and from various time points during the light phase (L0.5 = 0.5 h of light to L12 = 12 h of light). The blot was probed with an ALAD cDNA and with a constitutively expressed G protein β subunit-like cDNA from *C. reinhardtii* [17] to show approximately equal RNA loadings.

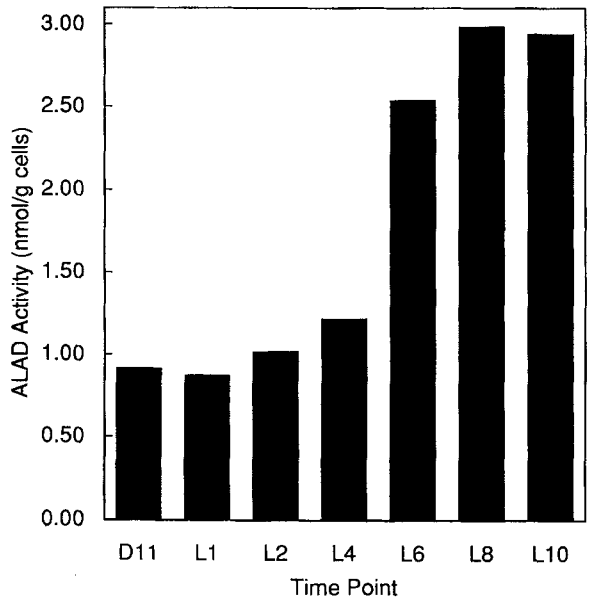


Fig. 4. Time course of ALAD enzyme activity in cell extracts made from various stages of *C. reinhardtii* synchronous cultures. Cell extracts from a dark-phase time point just prior to the light phase (D11) and six light-phase time points (L1 through L10) were assayed spectrophotometrically for the *in vitro* conversion of ALA to PBG catalyzed by ALAD.

10 h into the light phase. The peak level of ALAD activity (L8) represents a 3.2-fold increase over the dark-phase activity level.

Sequence of the alad *genomic upstream region*

Two overlapping genomic DNA clones were selected from a *C. reinhardtii* CC-124 genomic library using an ALAD cDNA as a probe. One of the genomic *alad* clones contained the entire protein-coding sequence and untranslated regions as well as approximately 1.1 Kbp of 5′-upstream DNA, while the other clone contained only a part of the coding region, the 3′-untranslated region, and downstream DNA. The genomic DNA upstream from the 5′ end of the cDNA was sequenced, and the transcription start site was determined to be 71 bp from the beginning of the ALAD-encoding open reading frame. The upstream DNA sequence contains a general TATA-like box (bp −65 to −60) followed by a CG-rich region. When the *alad* upstream sequence was

compared to the promoter region of the *C. reinhardtii gsa* gene, no significant similarities were detected, including no evidence of the promoter sequences previously associated with light-regulated genes.

Discussion

We have isolated several *C. reinhardtii* clones which encode the chlorophyll biosynthetic enzyme ALAD. Functional complementation of an *E. coli hemB* mutant, which lacks ALAD activity, and a high degree of sequence similarity of the longest open reading frame to other ALAD genes indicates that these cDNAs encode ALAD. The encoded protein (390 residues) is similar in size to other ALADs from pea (398 residues, deduced from an incomplete cDNA sequence), soybean (412 residues), and spinach (423 residues). The presence of an apparent chloroplast transit peptide indicates the mature protein is localized to the chloroplast, in agreement with the previously determined location of this enzyme in pea [19]. Plant ALADs contain a highly conserved Mg^{2+}-binding domain which is also found in the deduced *C. reinhardtii* protein (Fig. 1A, residues 171–188). The *C. reinhardtii* metal-binding sequence is identical to those of the *Selaginella martensii*, pea, soybean, and tomato ALADs in 17 of 19 residues, and identical to that of the spinach ALAD in 16 of 19 residues (Fig. 1B). In other eukaryotic and prokaryotic ALADs, including the ALAD of the cyanobacterium *Anacystis nidulans* [9], the metal-binding domain is different in its amino acid sequence and coordinates Zn^{2+} rather than Mg^{2+}. Although all ALAD metal-binding domains contain several invariant residues that are highly conserved (Fig. 1B, asterisks), the Zn^{2+}-binding domain uses Cys and His residues as putative metal ligands, whereas the Mg^{2+}-binding domains, including that of *C. reinhardtii* ALAD, contains Asp residues to coordinate the metal (Fig. 1B, bold type). A third class of ALAD, present in members of the α-proteobacteria (*Bradyrhizobium japonicum*, *Rhodobacter capsulatus*), has a homologous region with some features of both types of metal-binding domains [7], and these sites may bind K^+ instead of Mg^{2+} or Zn^{2+} [15, 18].

Multiple ALAD cDNA clones isolated by complementation were found to differ in the length of their 3′ UTR, although all contained the same sequence up to the poly(A) site and all appeared to be polyadenylated normally. Genomic Southern analysis indicates the presence of a single *alad* gene in *C. reinhardtii*, suggesting that these cDNAs do not represent the products of multiple genes (Fig. 2C). The various cDNAs are probably the result of differently processed ALAD primary transcripts. The normal polyadenylation signal for *C. reinhardtii* mRNAs, TGTAA, is not perfectly present in the 3′ UTR of any of the ALAD cDNAs. This suggests that the multiple polyadenylation sites may have been used to generate the various length 3′ UTRs due to the lack of a strong indicator for where the poly(A) should be attached. The existence of multiple polyadenylation sites has been reported for several higher plant mRNAs. cDNAs with varying length 3′ UTRs have been found for the ribulose bisphosphate carboxylase/oxygenase small subunit and a chlorophyll *a*/*b* binding protein of petunia and the *bronze* gene product of maize [3]. Whether the length differences of the various ALAD cDNA classes are associated with separate functions for these mRNAs in *C. reinhardtii* is not clear. All of the ALAD cDNAs, when expressed in *E. coli*, rescued the *hemB* mutant phenotype, and approximately equal numbers of cDNAs from each class were found by complementation screening. Northern analysis of ALAD transcripts shows only one size of ALAD mRNA equal to the longest cDNA class. However, in our northern blot analysis system, this cDNA class probably cannot be distinguished from the cDNA class that is only 32 bases shorter. One other possible explanation for the presence of only one detectable cDNA size by the northern blot analysis is that the strain that was used for the Southern and northern analyses (CC-124) was different from the strain from which the cDNA library was constructed (CC-621, NO^-). Although we have not tested for the presence of multiple *alad* genes in

strain CC-621, strains CC-124 and CC-621 were both derived from the same progenitor, 137c, and are closely related.

The level of ALAD mRNA in light:dark synchronized cells is low just prior to the beginning of the light phase, increases to a peak level at 2 h into the light phase, and then gradually diminishes. The general pattern of ALAD mRNA levels is similar to that determined for another *C. reinhardtii* chlorophyll biosynthetic gene, *gsa*, which encodes the enzyme just prior to ALAD in this pathway (GSAT) [13]. The mRNAs for these two genes begin to accumulate early in the light phase and reach peak levels at 2 h of light. However, the total increase in GSAT mRNA is greater than the increase in ALAD mRNA. Whereas the amount of GSAT mRNA increases 26-fold from D11 to L2, the ALAD mRNA level increases only about 7-fold. The pattern of decline in the levels of GSAT and ALAD mRNAs after the peak at 2 h also is different. The GSAT mRNA level decreases quickly between L2 and L4, and the mRNA is undetectable by 8 h of light. The ALAD mRNA level decreases more slowly, and ALAD mRNA is detectable even as cells enter the dark phase. This observation is consistent with the possibility that differential rates of mRNA turnover for GSAT and ALAD mRNAs may be one mode of regulation of chlorophyll biosynthesis during the light phase.

```
-1128                                    TCTTGTCATGATGCCCCAGGTGACCCCT
-1100 CCATCACCTGGTGTGTTCTGCTCCTGCCGACCGCGCAGGATCGGACCACG
-1050 GCCACATTGCCTCCAAGGACAGCGGCGCCCTGGCCTGGCATTACGATTTA
-1000 CGAGTGCAGCGGTGAGCAGTAGTGCATGCAGCTGCTGGGTTGGCAGTGCG
 -950 CCGCTCCTGCAATGGAACTGTGTCTCCCCATCTCCACAGCTGGTTTGTCT
 -900 TGTGCCAAGCTTGGGCTGAAATCCATCGGCGTGCCCTTGATGCTTCGCCG
 -850 CCTGCAGGCCAGCACAGCTGATGGTGCGCTCCATCAACCACTGGTGATGC
 -800 GCCGTACGTCCTCGGCTGGAGCGGTCAAGGCGGGGGAGCCCAATCGAATG
 -750 GACTAGTGCACCATTAATCTTAACGCTCCAGTGGACATGCATCCACTAAT
 -700 GCTTGGCAGCTATGATTCCCACGTATGATTCCATTTATTATGATTCATTA
 -650 CTCGCCCCTACATCGCGGGGCATGGGGTTTTGGGAGCGTGCTGGTGTGAA
 -600 CGTGTGTGCGCTGCCCATGCGCACAACGCGGACCGGACCATCAGACATGG
 -550 AAAGTGGCAGGTCCATTCAAAGGACGTCTGTAGATAGAACACGGGCATTA
 -500 GATGTGAGCCTGAACTGTGGTTTCATGGAGACAGCTTCCATAATAGTACA
 -450 CATGTAAGTAACTGGCGATCGTGTGTTGCTCGACAACATGGCTCCCGTCC
 -400 ACATTGCAGCCCCACTTCGCGTTCATGCCCGCCTGACATCCCCATAACGT
 -350 TCCACCGTCTCTACCTCAGCCCGTCCTGGCCCGTCCGGCTCCGCTGTTTC
 -300 ATGTCTGCAACTCGCCTCTCAAGAATGCGCCCCATGAGCAACCCCCATGA
 -250 GCGTGACGAAGCGAAGGGGAAGCCGAGGCGAGCTGGGGAAAACGCGCACG
 -200 GCACGGTAATCCTCAACGACTGTGACCACAACAGTCAAAGATGTGACAGC
 -150 GCCGAAGCGCCGAATGGTTAAGCATTTAGCGATTGGTCAGTTTTGCCAAG
 -100 ACGTTGTCAAGCGACCGCAAGAGACCACGAAATGCTTAATTGCAGCGAGC
  -50 CCATCTCGCCGATCTCGCGATTTAGGCGTTAGGGCCTCCCCAATAGGTTC
   +1 CTGTGCAAATTGTAGGTGCTTCTGCATACAATTGACCGCGAAGATCT
```

Fig. 5. Genomic DNA sequence of the region upstream of the *alad* gene in *C. reinhardtii*. The 5′ extent of the longest ALAD cDNA is indicated in bold, and a putative TATA-box is underlined.

The increase in ALAD enzyme activity occurs subsequent to the increase in ALAD mRNA level in synchronous cultures. ALAD activity is relatively low in dark phase cultures just prior to the light phase and increases sharply between 4 and 6 h into the light phase, which is after the peak in ALAD mRNA level. The 3.2-fold increase in ALAD activity from the D11 to L8 time points is less than the 7-fold increase in the ALAD mRNA level. Therefore, the increase in ALAD activity in the light can be accounted for by increases in ALAD mRNA, rather than by a post-transcriptional mechanism as was proposed for ALAD in greening pea leaves [2]. The change in ALAD mRNA and ALAD activity also is consistent with the pattern of chlorophyll accumulation in synchronized cultures. A period of rapid chlorophyll accumulation occurs between 4 and 9 h into the light phase, and level of the mRNA encoding a chlorophyll *a*/*b* binding protein is increased during this time frame as well [8]. The level of ALAD mRNA is somewhat less than 50% higher at D11 than at L12. This small increase may reflect the need for increased heme synthesis during cell division, which occurs in the dark phase.

Although the increases in ALAD mRNA and ALAD enzyme activity occur during the light phase of synchronous cultures, comparison of the *alad* genomic DNA upstream sequence to the *gsa* genomic upstream DNA failed to reveal the presence of sequence elements previously identified in the *gsa* promoter that are associated with light regulation (Fig. 5). Thus, although ALAD and GSAT mRNA levels appear to vary in a similar

pattern during the light phase, the regulation of *alad* transcription may be different from that of *gsa*. The *C. reinhardtii alad* gene is expressed differently than the *alad* genes in etiolated pea and soybean tissues transferred to light. Soybean leaf ALAD mRNA increases 2-fold after 24 h of light but the ALAD protein level and enzyme activity do not increase, and ALAD mRNA levels in etiolated pea leaves do not change during 48 h of illumination. The difference between these systems and *C. reinhardtii*, in which both ALAD mRNA levels and enzyme activity increase in the light, may reflect the difference between etiolated tissues and the light:dark synchronized cultures used in this study.

Acknowledgements

We thank E.H. Harris and the *Chlamydomonas* Genetics Center for providing *C. reinhardtii* cultures, B.J. Bachmann and the *E. coli* Genetic Stock Center for *E. coli* cultures, and J.P. Woessner for the *C. reinhardtii* cDNA expression library. This work was supported by National Science Foundation Grant DCB91-03253 and Department of Energy Grant DEFG02-88ER13918.

References

1. Beale SI, Weinstein JD: Tetrapyrrole metabolism in photosynthetic organisms. In: Dailey HA (ed) Biosynthesis of Heme and Chlorophylls, pp. 287–391. McGraw-Hill, New York (1990).
2. Boese QF, Spano AJ, Li J, Timko MP: Aminolevulinic acid dehyratase in pea (*Pisum sativum* L.). J Biol Chem 266: 17060–17066 (1991).
3. Dean C, Tamaki S, Dunsmuir P, Favreau M, Katayama C, Dooner H, Bedbrook J: mRNA transcripts of several plant genes are polyadenylated at multiple sites *in vivo*. Nuc Acids Res 14: 2229–2240 (1986).
4. Franzén L-G: Analysis of chloroplast and mitochondrial targeting sequences from the green alga *Chlamydomonas reinhardtii*. Biol Membr 11: 304–309 (1994).
5. Harris EH: The *Chlamydomonas* Sourcebook. Academic Press, San Diego (1989).
6. Huault C, Aoues A, Colin P: Immunochemical study of δ-aminolevulinate dehydratase in etiolated radish cotyledons. Plant Physiol Biochem 25: 723–728 (1987).
7. Jaffe EK: Predicting the Zn(II) ligands in metalloproteins: case study, porphobilinogen synthase. Comments Inorg Chem 15: 67092 (1993).
8. Johanningmeier U, Howell SH: Regulation of light-harvesting chlorophyll-binding protein mRNA accumulation in *Chlamydomonas reinhardtii*. Possible involvement of chlorophyll synthesis precursors. J Biol Chem 259: 13541–13549 (1984).
9. Jones MC, Jenkins JM, Smith AG, Howe CJ: Cloning and characterization of genes for tetrapyrrole biosynthesis from the cyanobacterium *Anacystis nidulans* R2. Plant Mol Biol 24: 435–448 (1994).
10. Kaczor CM, Smith MW, Sangwan I, O'Brian MR: Plant δ-aminolevulinic acid dehydratase. Plant Physiol 104: 1411–1417 (1994).
11. Kasemir H, Masoner M: Control of chlorophyll synthesis by phytochrome II. The effect of phytochrome on aminolevulinate dehyratase in mustard seedlings. Planta 126: 119–126 (1975).
12. Li J-M, Umanoff H, Proenca R, Russell CS, Cosloy SD: Cloning of the *Escherichia coli* K-12 *hemB* gene. J Bacteriol 170: 1021–1025 (1988).
13. Matters GL, Beale SI: Structure and light-regulated expression of the *gsa* gene encoding the chlorophyll biosynthetic enzyme, glutamate-1-semialdehyde aminotransferase, in *Chlamydomonas reinhardtii*. Plant Mol Biol 24: 617–629 (1994).
14. Merchant S, Hill K, Kim JH, Thompson J, Zaitlin D, Bogorad L: Isolation and characterization of a complementary DNA clone for an algal pre-apoplastocyanin. J Biol Chem 265: 12372–12379 (1990).
15. Nandi DL, Shemin D: δ-Aminolevulinic acid dehydratase of *Rhodopseudomonas spheroides*. I. Isolation and properties. J Biol Chem 243: 1224–1230 (1968).
16. Schaumburg A, Schneider-Poetsch HAW: Characterization of plastid 5-aminolevulinate dehyratase (ALAD; EC 4.2.1.24) from spinach (*Spinacia oleracea* L.) by sequencing and comparison with non-plant ALAD enzymes. Z Naturforsch 47c: 77–84 (1992).
17. Schloss JA: A *Chlamydomonas* gene encodes a G protein β subunit-like polypeptide. Mol Gen Genet 221: 443–452 (1990).
18. Shioi Y, Doi M: Characterization of porphobilinogen synthase from an aerobic photosynthetic bacterium, *Erythrobacter* sp. strain OCh 114. Plant Cell Physiol 29: 843–848 (1988).
19. Smith AG: Enzymes for chlorophyll synthesis in developing peas. In: Akoyunoglou G, Senger H (eds) Regulation of Chloroplast Differentiation, pp. 49–54. Alan R. Liss, New York (1986).
20. Smith AG: Subcellular localization of two porphyrin-synthesis enzymes in *Pisum sativum* (pea) and *Arum* (Cuckoo-pint) species. Biochem J 249: 423–428 (1988).
21. Urata G, Granick S (1963) Biosynthesis of α-amino-

ketones and the metabolism of aminoacetone. J Biol Chem 238: 811–820 (1963).

22. Waffenschmidt S, Woessner JP, Beer K, Goodenough UW: Isodityrosine cross-linking mediates insolubilization of cell walls in *Chlamydomonas*. Plant Cell 5: 809–820 (1993).

23. Witty M, Wallace-Cook ADM, Albrecht H, Spano AJ, Michel H, Shabanowitz J, Hunt DF, Timko MP, Smith AG: Structure and expression of chloroplast-localized porphobilinogen deaminase from pea (*Pisum sativum* L.) isolated by redundant polymerase chain reaction. Plant Physiol 103: 139–147 (1993).