Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 6532–6536, August 1991
Plant Biology

Molecular cloning and expression in photosynthetic bacteria of a soybean cDNA coding for phytoene desaturase, an enzyme of the carotenoid biosynthesis pathway

GLENN E. BARTLEY*, PAUL V. VIITANEN*, IRIS PECKER†, DANIEL CHAMOVITZ†, JOSEPH HIRSCHBERG†, AND PABLO A. SCOLNIK*‡

*Du Pont Central Research and Development, P. O. Box 80402, Wilmington, DE 19880-0402; and †Department of Genetics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

Communicated by Lawrence Bogorad, April 22, 1991

ABSTRACT **Carotenoids are orange, yellow, or red photoprotective pigments present in all plastids. The first carotenoid of the pathway is phytoene, a colorless compound that is converted into colored carotenoids through a series of desaturation reactions. Genes coding for carotenoid desaturases have been cloned from microbes but not from plants. We report the cloning of a cDNA for *pds1*, a soybean (*Glycine max*) gene that, based on a complementation assay using the photosynthetic bacterium *Rhodobacter capsulatus*, codes for an enzyme that catalyzes the two desaturation reactions that convert phytoene into ζ-carotene, a yellow carotenoid. The 2281-base-pair cDNA clone analyzed contains an open reading frame with the capacity to code for a 572-residue protein of predicted M_r 63,851. Alignment of the deduced Pds1 peptide sequence with the sequences of fungal and bacterial carotenoid desaturases revealed conservation of several amino acid residues, including a dinucleotide-binding motif that could mediate binding to FAD. The Pds1 protein is synthesized *in vitro* as a precursor that, upon import into isolated chloroplasts, is processed to a smaller mature form. Hybridization of the pds1 cDNA to genomic blots indicated that this gene is a member of a low-copy-number gene family. One of these loci was genetically mapped using restriction fragment length polymorphisms between *Glycine max* and *Glycine soja*. We conclude that *pds1* is a nuclear gene encoding a phytoene desaturase enzyme that, as its microbial counterparts, contains sequence motifs characteristic of flavoproteins.**

Carotenoids are photoprotective pigments present in all photosynthetic and many nonphotosynthetic organisms. Absorption of visible light and photoprotection are mediated by a chain of conjugated double bonds, the chromophore, that is formed by successive desaturations of the colorless precursor phytoene (1, 2). Deduced peptide sequences of several microbial carotenoid desaturases have been reported (3–6). A maize gene (*y1*) involved in carotenoid biosynthesis has been cloned (7), but at present it is not clear whether this gene is regulatory or structural.

Carotenoids are precursors of abscisic acid (8) and carotenoids influence chloroplast differentiation. Abscisic acid and chloroplast differentiation affect nuclear gene expression (9–11). Thus, a knowledge of the genes of the carotenoid–abscisic acid pathway is required for our understanding of the interactions between plastid differentiation and nuclear gene expression during plant development.

Norflurazon is a bleaching herbicide that inhibits the desaturation of phytoene (12). To isolate a soybean phytoene desaturase gene, we have used a heterologous probe from a gene responsible for norflurazon resistance in the cyanobacterium *Synechococcus* PCC7942 (13). We reasoned that since norflurazon inhibits the desaturation of phytoene, the *Synechococcus* gene could code for phytoene desaturase and could be used as a heterologous probe for cloning the corresponding plant genes. We report here the cloning and characterization of a soybean cDNA coding for phytoene desaturase.§

MATERIALS AND METHODS

Recombinant DNA Techniques. Cloning steps were described in ref. 14. DNA restriction endonucleases, polymerases, and ligases were used under conditions recommended by suppliers (New England Biolabs, Bethesda Research Laboratories, and Perkin–Elmer/Cetus). The PCR was used to amplify a portion of the *Synechococcus* PCC6301 genomic DNA using the primers 5′-TTCTTCGGTGCCTAC-CC(C,T)AA(C,T)AA(C,T)ATG and 5′-GAACAATTTT-TCAAT(T,C)TC(G,A)GCCAT corresponding to the protein sequences FFGAYPNM and TAEIEKLF of the *Synechococcus* PCC7942 protein that confers resistance to norflurazon (13). Hybridization conditions for library screening were as described in Berlyn *et al.* (15). The plasmid pNFPD11 was constructed by ligating a 1857-base-pair (bp) *Sst* I fragment of pds1 cDNA into the *Sst* I site of pNF3 (16).

Complementation of *Rhodobacter (Rb.) capsulatus* and Carotenoid Analysis. Plasmid pNFPD11 was mobilized by conjugation (17) into *Rb. capsulatus* BPY69. Ammonia-free minimal medium was used for induction of the *nif* promoter (16). *Rb. capsulatus* cells were grown photosynthetically in liquid medium and carotenoids were extracted (18) and saponified (19). The solvent used for thin layer chromatography was 5% (vol/vol) acetone in petroleum ether.

Chloroplast Import. Linearized plasmid templates were transcribed using T7 RNA polymerase (20). *In vitro* translations were performed using [^{35}S]labeled methionine (1100 Ci/mmol; 1 Ci = 37 GBq) and rabbit reticulocyte lysate (Promega) according to the vendor's protocol. Translation reactions were terminated with 2× import buffer (20), containing 60 mM unlabeled methionine (20). Chloroplasts were obtained (21) from 10- to 12-day-old pea seedlings (*Pisum sativum*). Import was as described in ref. 21, but ice instead of $HgCl_2$ was used to stop reactions. To distinguish between bound and imported polypeptides, plastids were treated with thermolysin (20). Controls were treated identically but received no protease. Intact plastids were repurified by centrifugation through Percoll (22), washed once with import buffer (containing 5 mM EDTA), and then subjected to lysis by resuspending pellets in 250 μl of 10 mM Hepes·KOH, pH 8/5 mM EDTA. Total membranes (thylakoids plus envelopes) were recovered by centrifugation at 47,000 × *g* for 30 min. The membranes were washed once and resuspended in gel sample buffer (20). The stromal fraction was recentrifuged

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: nt, nucleotide(s).

‡To whom reprint requests should be addressed.

§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M64704).

```
   1 GAATTCCTTCTACGTACTGCCGTGGTGCTTTCACCACTGCTTACCACTAA
  51 CCTTCCTCTCTCTCTCTGCCGCTGCAAGCTTGGTACTCTCAACTCAATTC
 101 TCCACCTTATTCTTTTCACTTCTTCAGCTCTTGTTTTTTCCCAAATCTAC
 151 TTTCAAAGTGCCTGAATTCTGCAACAGTAATATTAACACTCCTCTCTTTT
 201 GTTCAGGCTTTATTTCCCCAATGGCCGCTTGTGGCTATATATCTGCTGCC
   1                   M  A  A  C  G  Y  I  S  A  A
 251 AACTTCAATTATCTCGTTGGCGCCAGAAACATATCCAAATTCGCTTCTTC
  11 N  F  N  Y  L  V  G  A  R  N  I  S  K  F  A  S  S
 301 AGACGCCACAATTTCGTTTTCATTTGGCGGGAGCGACTCAATGGGTCTTA
  28 D  A  T  I  S  F  S  F  G  G  S  D  S  M  G  L  T
 351 CTTTGCGACCCGCTCCGATTCGTGCTCCTAAGAGGAACCATTTCTCTCCC
  45  L  R  P  A  P  I  R  A  P  K  R  N  H  F  S  P
 401 TTGCGTGTCGTTTGCGTCGATTATCCACGCCCAGAGCTCGAAAACACCGT
  61 L  R  V  V  C  V  D  Y  P  R  P  E  L  E  N  T  V
 451 TAATTTCGTTGAAGCTGCTTACTTGTCTTCCACCTTTCGTGCTTCTCCGC
  78  N  F  V  E  A  A  Y  L  S  S  T  F  R  A  S  P  R
 501 GTCCTCTAAAACCCTTGAACATCGTTATTGCCGGTGCAGGATTGGCTGGT
  95   P  L  K  P  L  N  I  V  I  A  G  A  G  L  A  G
 551 TTATCAACTGCAAAATATTTGGCTGATGCTGGGCATAAACCTATATTGCT
 111 L  S  T  A  K  Y  L  A  D  A  G  H  K  P  I  L  L
 601 GGAAGCAAGAGACGTTCTAGGTGGAAAGGTTGCTGCATGGAAAGACAAGG
 128   E  A  R  D  V  L  G  G  K  V  A  A  W  K  D  K  D
 651 ATGGAGACTGGTACGAGACAGGCCTACACATCTTTTTTGGGGCTTACCCT
 145    G  D  W  Y  E  T  G  L  H  I  F  F  G  A  Y  P
 701 TATGTGCAGAACCTTTTTGGAGAACTTGGCATTAATGATCGGTTACAATG
 161  Y  V  Q  N  L  F  G  E  L  G  I  N  D  R  L  Q  W
 751 GAAAGAGCATTCTATGATTTTTGCTATGCCAAATAAGCCTGGAGAGTTTA
 178 K  E  H  S  M  I  F  A  M  P  N  K  P  G  E  F  S
 801 GTCGATTTGATTTTCCTGAAGTTCTTCCCTCCCCATTGAATGGAATATGG
 195   R  F  D  F  P  E  V  L  P  S  P  L  N  G  I  W
 851 GCAATATTGAGGAACAATGAGATGCTTACATGGCCAGAGAAAGTAAAATT
 211 A  I  L  R  N  N  E  M  L  T  W  P  E  K  V  K  F
 901 TGCAATTGGGCTTCTCCCAGCTATGCTTGGCGGACAGCCATATGTTGAGG
 228 A  I  G  L  L  P  A  M  L  G  G  Q  P  Y  V  E  A
 951 CTCAAGATGGTCTTTCTGTTCAAGAATGGATGAAAAAGCAGGGCGTACCT
 245   Q  D  G  L  S  Y  Q  E  W  M  K  K  Q  G  V  P
1001 GAACGGGTAGCTGATGAGGTGTTCATAGCAATGTCAAAGGCACTAAACTT
 261  E  R  V  A  D  E  V  F  I  A  M  S  K  A  L  N  F
1051 CATCAATCCTGATGAACTTTCAATGCAATGTATATTGATTGCTTTAAACC
 278  I  N  P  D  E  L  S  M  Q  C  I  L  I  A  L  N  R
1101 GATTTCTTCAGGAGAAACATGGTTCTAAGATGGCCTTTTTGGATGGCAAT
 295   F  L  Q  E  K  H  G  S  K  M  A  F  L  D  G  N
1151 CCACCCGAAAGACTTTGTATGCCAATAGTTGATTATATTCAGTCCTTGGG
 311 P  P  E  R  L  C  M  P  I  V  D  Y  I  Q  S  L  G
1201 TGGTGAAGTTCATCTAAATTCGCGCATTCAAAAAATTGAGCTAAATGATG
 328  G  E  V  H  L  N  S  R  I  Q  K  I  E  L  N  D  D
1251 ATGGAACGGTGAAGAGCTTCTTACTAAATAATGGGAAAGTGATGGAAGGG
 345   G  T  V  K  S  F  L  L  N  N  G  K  V  M  E  G
1301 GATGCTTATGTGTTTGCAACTCCAGTGGATATTCTGAAGCTTCTTCTACC
 361  D  A  Y  V  F  A  T  P  V  D  I  L  K  L  L  L  P
1351 AGATAACTGGAAAGGGATTCCATATTTCCAGAGATTGGATAAATTAGTTG
 378  D  N  W  K  G  I  P  Y  F  Q  R  L  D  K  L  V  G
1401 GCGTCCCAGTCATAAATGTTCACATATGGTTTGACAGAAAACTGAAGAAC
 395    V  P  V  I  N  V  H  I  W  F  D  R  K  L  K  N
1451 ACATATGATCACCTTCTCTTTAGCAGAAGTCCCCTTCTGAGTGTATATGC
 411 T  Y  D  H  L  L  F  S  R  S  P  L  L  S  V  Y  A
1501 TGACATGTCAGTAACTTGCAAGGAATATTATAGCCCAAACCAGTCAATGT
 428 D  M  S  V  T  C  K  E  Y  Y  S  P  N  Q  S  M  L
1551 TAGAGTTGGTTTTTGCACCAGCCGAAGAATGGATTTCACGTAGTGATGAT
 445    E  L  V  F  A  P  A  E  E  W  I  S  R  S  D  D
1601 GATATTATTCAAGCCACGATGACTGAGCTTGCCAAACTCTTTCCTGATGA
 461  D  I  I  Q  A  T  M  T  E  L  A  K  L  F  P  D  E
1651 AATTTCTGCAGACCAAAGCAAAGCAAAGATTCTCAAGTACCATGTTGTTA
 478  I  S  A  D  Q  S  K  A  K  I  L  K  Y  H  V  V  K
1701 AAACACCAAGGTCGGTTTACAAAACTGTTCCAAATTGTGAACCTTGTCGA
 495   T  P  R  S  V  Y  K  T  V  P  N  C  E  P  C  R
1751 CCCATTCAAAGATCTCCTATAGAAGGTTTCTATTTAGCTGGAGATTACAC
 511  R  I  Q  R  S  P  I  E  G  F  Y  L  A  G  D  Y  T
1801 AAAACAAAAATATTTAGCTTCAATGGAAGGCGCTGTTCTTTCTGGGAAGC
 528   K  Q  K  Y  L  A  S  M  E  G  A  V  L  S  G  K  L
1851 TTTGTGCACAGGCTATTGTACAGGATTCTGAGCTACTAGCTACTCGGGGC
 545    C  A  Q  A  I  V  Q  D  S  E  L  L  A  T  R  G
1901 CAGAAAAGAATGGCTAAAGCAAGTGTTGTGTAACAAAAACAAGAATTGAA
 561   Q  K  R  M  A  K  A  S  V  V  *
1951 AGAGTCATGGTAGAGTACAGGAGCATCATTTCAACTTTGGCATTCTTTGT
2001 CTGTGGTCAGGACTCAGGAGACCTTCAACTTTATTAGTTCATACGAATAA
2051 AGAAAGGCTCAGCTTCTGAAATTTAGCTGCACCGTCGTCAACTGTGTGCA
2101 ATAAGCTATACGGAACAAACGACATGTGTCAACTTTAAGTCAGCCCATTG
2151 TTTTGTTATCCTCCAATTTTCTGGATCAATGTTTGTATTGGAAAGAAATA
2201 TGTCATTATTCAAACTTGTTTATATCCACTTTTTTATTTATCAACATTT
2251 GTCACAACCTTTCGTTGAGTAAAAAAAAAAAAAAAAAAGAATTC
```

FIG. 1. Nucleotide sequence of the pds1 cDNA coding strand and deduced amino acid sequence of the phytoene desaturase protein. Nucleotides are numbered in the 5′ to 3′ direction. The coding region starts at nt 221. *Eco*RI sites (GAATTC) flanking the pds1 cDNA are italicized. Amino acid residues with a match score of at least 3 in the phytoene desaturase pattern (see Fig. 2) are underlined. Residues showing perfect conservation (match score = 5) are shown in boldface type. Putative FAD-binding residues (24) are boxed. Amino acid residues are shown in the standard one-letter code.

at 140,000 × *g* for 30 min and the resultant supernatant was removed for SDS/PAGE. This second centrifugation step greatly improved the analysis of "stromal" carotenoid desaturase, presumably by the partial removal of ribulose 1,5-bisphosphate carboxylase. Samples were subjected to SDS/PAGE on 10% gels. Gel sample solubilization, electrophoresis, and fluorography were carried out as described (20).

RESULTS

cDNA Cloning, Nucleotide Sequence, and Predicted Protein Sequence. A radiolabeled *Synechococcus* PCC6301 probe was used to screen 5 × 10^5 plaques of a soybean immature cotyledon bacteriophage λ cDNA library. Two positive clones isolated showed inserts of similar size (data not shown). A plasmid (pSoysk) derived from one λ clone was purified, and both strands were sequenced by the dideoxynucleotide termination method (23). The 2281-bp insert contains a single open reading frame that stops at nucleotide (nt) 1931, a TAA codon (Fig. 1). The first initiation codon is located at nt 221. Thus, if this ATG is the initiation codon, the cDNA has 200 bp of 5′ untranslated sequence, a 1710-bp coding region, and 322 bp of untranslated 3′ sequence, followed by a poly(A) tail.

The assignment of the ATG codon at nt 221 as the start codon of *pds1* is supported by three lines of evidence. (*i*) This codon occurs within the sequence CCCAATGGC, which approximates the plant ribosome-binding consensus sequence (25). (*ii*) As shown below, this cDNA encodes a precursor protein (prPds1) that is active in an *in vitro* chloroplast import reaction, suggesting that all *pds1* 5′ sequences coding for the prPds1 N-terminal region required for import are present in this cDNA. (*iii*) Based on primary sequence conservation (see below), we calculate that the prPds1 transit peptide has about 94 amino acid residues. If the ATG at nt 221 is not the start codon, the transit peptide would exceed 170 residues, which seems unlikely (26).

The *pds1* open reading frame encodes a 572-amino acid protein, with an estimated molecular weight of 63,851. The 322-bp untranslated 3′ region contains two putative poly(A) addition signals at nt 2046 (AATAAA) and nt 2100 (AATAA). However, these sequences are located further upstream of the polyadenylylation site than the usual 15–40 nt found in many plant genes (27).

Primary Sequence Conservation Among Prokaryotic and Eukaryotic Carotenoid Desaturases. To identify amino acid residues that, by virtue of being conserved among different carotenoid desaturases, could be important in structure and function (28), we applied a local alignment algorithm (ref. 29; Fig. 2*A*) to the deduced peptide sequences of four carotenoid desaturases, *Rb. capsulatus* CrtI and CrtD, *Neurospora crassa* Al-1, and soybean Pds1. The resulting pattern for the family contains 501 characters with 17 conserved residues (Fig. 2*B*). The alignment "tree" (Fig. 2*C*) shows that CrtI and Al-1 are the most closely related sequences (match score = 215.27), followed by CrtD (match score = 142.29) and Pds1 (match score = 80.13). Pds1 amino acid residues conserved with a match score of 3 or higher are shown in Fig. 1. From the N terminus the first region of conservation has the sequence Gly-Xaa-Gly-Xaa_2-Gly-Xaa_3-Ala-Xaa_6-Gly. This is characteristic of dinucleotide-binding folds that mediate binding of FAD or NAD(P) (31). This motif has been observed (5, 24) at the N-terminal region of microbial carotenoid desaturases. Extensive conservation has been reported between carotenoid desaturase residues that mediate FAD binding in glutathione reductase and dihydrolipoamide de-

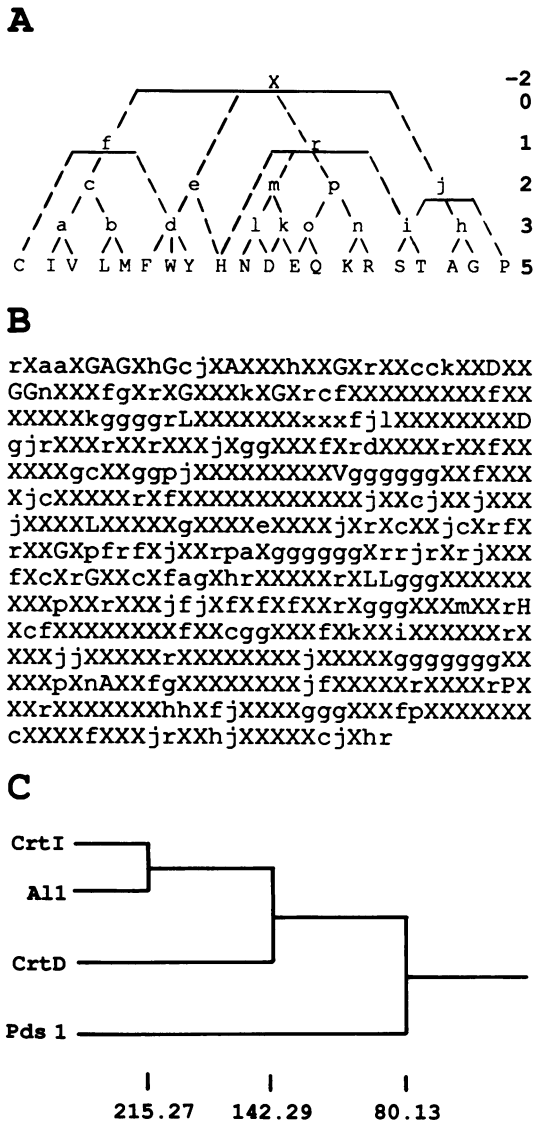

FIG. 2. Primary protein sequence pattern from microbial and plant carotenoid desaturases. (*A*) Amino acid class hierarchy used to construct the pattern. Uppercase characters indicate standard one-letter code; lowercase characters, amino acid classes; X, any amino acid. Match scores are indicated on the right (modified from ref. 29). (*B*) Carotenoid desaturase similarity pattern, based on *Rb. capsulatus* CrtI (3) and CrtD (30), *N. crassa* Al-1 (4), and soybean Pds1. g, Gap character. (*C*) Dendrogram generated, including overall matching scores. For this alignment we changed CrtD at position 13 from R to G assuming a *crtD223* mutation at the corresponding codon (24).

hydrogenase, two disulfide oxidoreductases (24). Based on similar comparison of Pds1 to glutathione reductase and dihydrolipoamide dehydrogenase, we identify 19 Pds1 residues that could play a role in FAD binding (Fig. 1).

***In Vitro* Import of Pds1 into Pea Chloroplasts.** In all microbial carotenoid desaturases analyzed, the dinucleotide binding fold is located near the N terminus (5, 24). This plus the facts that these proteins are encoded by nuclear genes and are targeted to plastid envelopes (2, 32) suggest that the soybean Pds1 protein is initially synthesized as a higher molecular weight precursor, possessing a transit peptide that targets the protein to plastids. To test this we conducted *in vitro* chloroplast import experiments. Translation of a pds1 transcript resulted in the synthesis of a polypeptide (Fig. 3, lane T) whose apparent molecular mass in SDS/PAGE gels (≈59 kDa) is in reasonable agreement with the size of the precursor predicted from the DNA sequence.

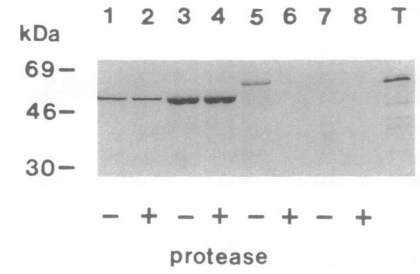

FIG. 3. Import of Pds1 into chloroplasts. Import reactions were carried out in the presence (lanes 1–4) or absence (lanes 5–8) of 5 mM ATP. The latter reactions also contained nigericin (0.26 μM) and valinomycin (0.88 μM) to completely inhibit the production of ATP by photophosphorylation (33). After import, the plastids with (+) or without (−) protease treatment were fractionated to yield total membranes (lanes 1, 2, 5, and 6) and membrane-free stroma (lanes 3, 4, 7, and 8). A photograph of the SDS/PAGE fluorogram is shown. All lanes received an equivalent amount of chloroplasts. Lane T contains translation products used for import.

The radiolabeled polypeptide was then used for *in vitro* import into isolated pea chloroplasts. After import, intact chloroplasts were recovered and treated with protease, and membrane and soluble plastid fractions were isolated and analyzed by SDS/PAGE (Fig. 3). When import experiments were conducted in the presence of ATP (lanes 1–4), ^{35}S-labeled prPds1 entered the chloroplasts (as judged by protease resistance) and was processed to a smaller mature form. Surprisingly, imported mature Pds1 was detected in both soluble and membrane fractions. The former could represent either stromal material or soluble material liberated from the envelope intermembrane space. Although the mobility in SDS/PAGE is similar for soluble and membrane-associated Pds1, at present we do not know whether any differences exist in primary sequence between these forms. Regardless, no protease-resistant polypeptides were observed in any chloroplast fraction when import experiments were conducted in the complete absence of ATP (Fig. 3, lanes 5–8). Thus, like other precursor polypeptides targeted to chloroplasts, the import of prPds1 is energy dependent (33).

Complementation of a Phytoene-Accumulating *Rb. capsulatus* Mutant by *pds1*. To determine whether the cloned cDNA indeed codes for a phytoene desaturase, we constructed a plasmid that incorporates the *pds1* region coding for the mature protein under the control of bacterial transcription and translation signals. The resulting plasmid pNFPD11 (Fig. 4) contains the promoter, ribosome-binding sequence, and first 25 codons of *Rb. capsulatus nifH* fused in-frame to *pds1* codon 73. The *nif* (*ni*trogen *f*ixation) promoter can be induced by depriving cells of a fixed nitrogen source (16). Plasmid pNFPD11 was mobilized by conjugation into *Rb. capsulatus* BPY69, a strain with a genetic lesion in *crtI*, the gene that codes for phytoene desaturase (18). When BPY69(pNFPD11) cells were grown in minimal medium without fixed nitrogen, cultures turned yellow and, when cells were grown in ammonia-supplemented medium, they remained blue-green, the color of BPY69 cells lacking the soybean cDNA clone (data not shown). Pigments were extracted and analyzed by absorption spectroscopy (Fig. 4*B*). The yellow pigment in BPY69(pNFPD11) cells grown in ammonia-free medium was identified as 15-*cis*-ζ-carotene, indicating that Pds1 mediates two successive desaturations from phytoene (Fig. 4*D*). The identity of the compound was confirmed by TLC analysis, where it was shown to comigrate with tomato fruit ζ-carotene (data not shown).

Based on its structure and function, we assign this gene the name *pds1* (*p*hytoene *des*aturase 1).

Genetic Mapping of *pds1*. The pds1 cDNA was hybridized to blots containing *Glycine max* and *Glycine soja* DNA digested with 12 restriction enzymes (data not shown). The

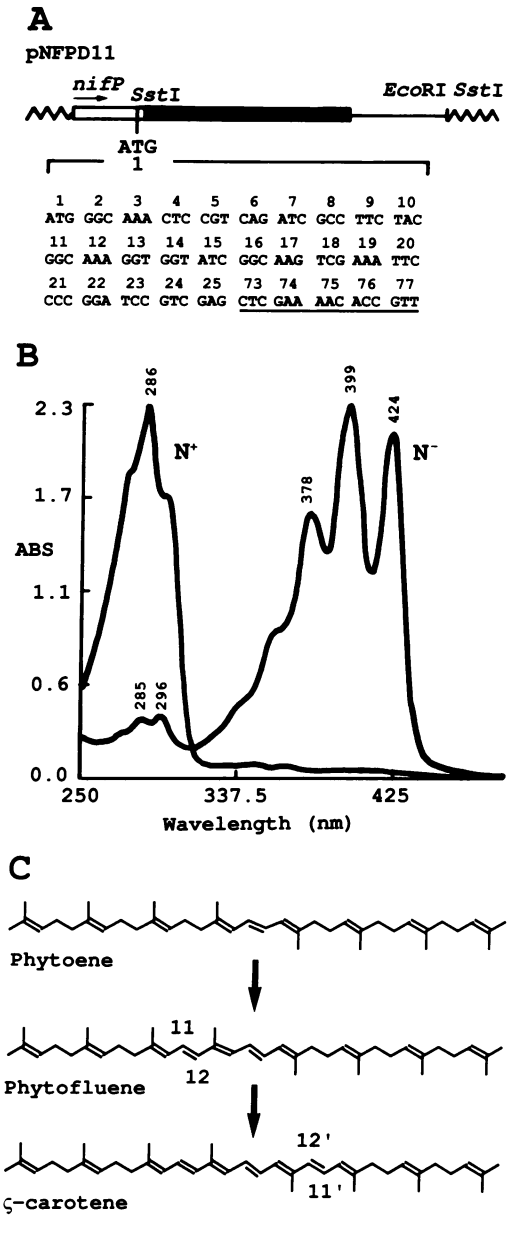

FIG. 4. Complementation of a *Rb. capsulatus* carotenoid mutant by the pds1 cDNA. (*A*) Construction of the *nif* expression plasmid pNFPD11. Open rectangle, *Rb. capsulatus nif* sequences up to *nifH* codon 25 and including the *nif* promoter (nifP) and translation initiation signal of pNF3 (16); solid boxes, pds1 coding sequences fused to *nif* DNA at the *Sst* I site (GAGCTC, codon 73). The first five *pds1* codons are underlined. Wavy lines represent vector DNA. The *Eco*RI site is located at the 3′ end of the pds1 cDNA (see Fig. 1). (*B*) Absorption spectra in petroleum ether of carotenoids accumulated by *Rb. capsulatus* BPY69 (pNFPD11) grown in minimal medium with (N^+) or without (N^-) ammonia. Absorption maxima are indicated for each peak. Phytoene and ζ-carotene accumulate under N^+ and N^- conditions, respectively. (*C*) Desaturations mediated by Pds1. Carbon atoms at which desaturation takes place are numbered.

patterns obtained with *Eco*RI, *Bam*HI, and *Hin*dIII suggest that *pds1* is a member of a low-copy-number nuclear gene family (Fig. 5). Polymorphisms detected with *Eco*RV (Fig. 5) allowed mapping of one allele to linkage group 9 in a soybean restriction fragment length polymorphism map (A. Rafalski and S. Tingey, personal communication).

DISCUSSION

We have cloned a soybean cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthesis pathway.

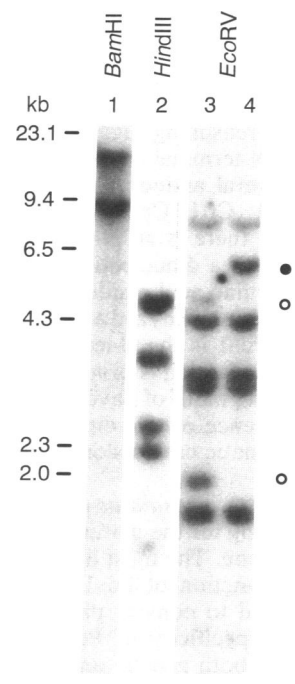

FIG. 5. Genomic blots and genetic mapping of *pds1*. The enzymes indicated above the lanes were used to digest DNA from *G. soja* (lanes 1–3) or *G. max* (lane 4). Solid and open circles indicate *G. max* and *G. soja pds1* alleles, respectively, detected with *Eco*RV.

Results indicate that Pds1 is a nuclear-encoded chloroplast protein that catalyzes two successive desaturation reactions.

The pds1 cDNA was isolated from a cDNA library prepared from immature soybean cotyledons. The low number of positive clones (2 in 5×10^5 clones) suggest that this is a low-abundance transcript. Indeed, we failed to detect the pds1 transcript in Northern blots containing up to 30 μg of total soybean RNA or 2 μg of poly(A)$^+$ RNA (data not shown). Although at present we do not know whether the cDNA isolated is a full-length clone, the complementation and chloroplast import results indicate that all of the information necessary for enzymatic activity and plastid targeting is present in the cDNA and suggest that the N terminus of the cloned protein is most likely intact.

Upon import into chloroplasts, mature Pds1 was detected in the soluble and membrane fractions. Carotenoids are membrane pigments and, based on immunological (3) and biochemical (32, 34) analyses, phytoene desaturase and all subsequent enzymes in the pathway are also membrane-bound. Thus, it seems logical to assume that the enzymatically active form of Pds1 is also associated with membranes. This raises questions as to the origin of the soluble Pds1 protein noted above. It is possible that this form is a non-physiological artifact that accumulates only during the artificial conditions used during *in vitro* chloroplast import. Alternatively, soluble Pds1 may be an intermediate in the membrane insertion process or it may play a nonenzymatic role in either the chloroplast stroma or envelope intermembrane space.

We have not yet defined the chloroplast import processing site. From a number of experiments similar to that shown in Fig. 3, it can be estimated that upon import the precursor polypeptide is shortened by 5–7 kDa. However, estimates obtained from denaturing gels are frequently associated with rather large errors. For microbial carotenoid desaturases, the dinucleotide-binding fold is located 4–9 amino acid residues downstream from the N terminus (4, 5, 24). Thus, if the situation with Pds1 is similar, the mature form would have

≈478 residues and a deduced molecular mass of 54 kDa. This would be in the same range as *N. crassa* and *Rb. capsulatus* phytoene desaturases (24) and is in general agreement with the observed mobility of mature Pds1 in SDS/PAGE gels (Fig. 3). This line of reasoning suggests that the precursor of Pds1 possesses an N-terminal transit peptide of ≈9.8 kDa.

We identified several amino acid residues that are conserved among Pds1, CrtI, CrtD, and Al-1. Among these conserved residues, there is an N-terminal region with the characteristic motif of a dinucleotide binding fold (Fig. 1). However, several amino acid residues that are conserved in all microbial desaturases analyzed are not conserved in Pds1. Notably absent in Pds1 is the C-terminal region containing the His-Pro dipeptide, which is proposed to be equivalent to the His-Pro active site motif of flavoprotein oxidoreductases (24). Thus, the presence of this dipeptide motif is not necessary for the enzymatic conversion of phytoene into ζ-carotene.

Expression in an *Rb. capsulatus* carotenoid mutant of the portion of *pds1* coding for the mature protein resulted in the synthesis of ζ-carotene. The most likely explanation for this result is that the function of Pds1 is to catalyze the two desaturations needed to convert phytoene into ζ-carotene. Alternatively, the specificity of Pds1 may be changed in bacteria. However, both genetic and biochemical evidence indicates that plants have a desaturase that converts phytoene into ζ-carotene. Mutant plants of several species blocked in carotenoid desaturations accumulate either phytoene or ζ-carotene, suggesting that two distinct genes code for two desaturases (for review, see ref. 1). Beyer *et al.* (34) showed that the conversion of phytoene into 15-*cis*-ζ-carotene and the subsequent conversion of this pigment into lycopene are two distinct enzymatic reactions. Hence, it appears that plants have two carotenoid desaturases and the simplest explanation for our results is that we have cloned a cDNA for the first of the two enzymes.

We thank Du Pont personnel for the following contributions: K. Abremski (protein alignments), K. Bacot (chloroplast import), A. Rafalski (mapping *pds1*), K. Ripp (cDNA library), J. Williams (design of PCR oligonucleotides), and M. Orbach and S. Russell (critical reading of the manuscript). We also thank P. Beyer (Freiburg) for his assistance in the identification of *cis*-ζ-carotene.

1. Bramley, P. M. & Mackenzie, A. (1988) in *Current Topics in Cellular Regulation*, eds. Horecker, B. L. & Stadtman, E. R. (Academic, New York), Vol. 29, pp. 291–332.
2. Bartley, G. E., Coomber, S. A., Bartholomew, D. & Scolnik, P. A. (1991) in *Cell Culture and Somatic Cell Genetics of Plants*, eds. Bogorad, L. & Vasil, I. (Academic, New York), Vol. 7B, pp. 331–345.
3. Bartley, G. E. & Scolnik, P. A. (1989) *J. Biol. Chem.* **264**, 13109–13113.
4. Schmidhauser, T. J., Lauter, F. R., Russo, V. E. A. & Yanofsky, C. (1990) *Mol. Cell. Biol.* **10**, 5064–5070.
5. Armstrong, G. A., Alberti, M. & Hearst, J. E. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9975–9979.
6. Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K. & Harashima, K. (1990) *J. Bacteriol.* **172**, 6704–6712.
7. Buckner, B., Kelson, T. L. & Robertson, D. S. (1990) *Plant Cell* **2**, 867–876.
8. Creelman, R. A. (1989) *Physiol. Plant.* **75**, 131–136.
9. Giuliano, G. & Scolnik, P. (1988) *Plant Physiol.* **86**, 7–9.
10. Burgess, D. G. & Taylor, W. C. (1988) *Mol. Gen. Genet.* **214**, 89–96.
11. Bartholomew, D. M., Bartley, G. E. & Scolnik, P. A. (1991) *Plant Physiol.* **96**, 291–296.
12. Clarke, I. E., Bramley, P. M., Sandmann, G. & Böger, P. (1982) in *Biochemistry and Metabolism of Plant Lipids*, eds. Wintermans, J. F. & Kuiper, P. J. C. (Elsevier, Amsterdam), pp. 549–554.
13. Chamovitz, D., Pecker, I. & Hirschberg, J. (1991) *Plant Mol. Biol.* **16**, 967–974.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
15. Berlyn, M. B., Last, R. L. & Fink, G. R. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4604–4608.
16. Pollock, D., Bauer, C. E. & Scolnik, P. A. (1988) *Gene* **65**, 269–275.
17. Giuliano, G., Pollock, D., Stapp, H. & Scolnik, P. A. (1988) *Mol. Gen. Genet.* **213**, 78–83.
18. Giuliano, G., Pollock, D. & Scolnik, P. A. (1986) *J. Biol. Chem.* **261**, 12925–12929.
19. Davies, B. H. (1976) in *Chemistry and Biochemistry of Plant Pigments*, ed. Goodwin, T. W. (Academic, New York), Vol. 2, pp. 65–66.
20. Viitanen, P. V., Doran, E. R. & Dunsmuir, P. (1988) *J. Biol. Chem.* **263**, 15000–15007.
21. Reed, J. E., Cline, K., Stephens, L. C., Bacot, K. O. & Viitanen, P. V. (1990) *Eur. J. Biochem.* **194**, 33–42.
22. Cline, K., Werner-Washburne, M., Lubben, T. H. & Keegstra, K. (1985) *J. Biol. Chem.* **260**, 3691–3696.
23. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
24. Bartley, G. E., Schmidhauser, T. J., Yanofsky, C. & Scolnik, P. A. (1990) *J. Biol. Chem.* **265**, 16020–16024.
25. Lütcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F. & Scheele, G. A. (1987) *EMBO J.* **6**, 43–48.
26. Keegstra, K. & Olsen, L. J. (1989) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 471–501.
27. Mogen, B. D., MacDonald, M. H., Graybosch, R. & Hunt, A. G. (1990) *Plant Cell* **2**, 1261–1272.
28. Berg, P. & Reichardt, J. K. V. (1988) *Nucleic Acids Res.* **16**, 9017–9026.
29. Smith, R. F. & Smith, T. F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 118–122.
30. Armstrong, G. A., Alberti, M., Leach, F. & Hearst, J. E. (1989) *Mol. Gen. Genet.* **216**, 254–268.
31. Moller, W. & Amons, R. (1985) *FEBS Lett.* **186**, 1–7.
32. Lütke-Brinkhaus, F., Liedvogel, B., Kreuz, K. & Kleining, H. (1982) *Planta* **156**, 176–180.
33. Grossman, A., Bartlett, S. & Chua, N.-H. (1980) *Nature (London)* **285**, 625–628.
34. Beyer, P., Mayer, M. & Kleinig, H. (1989) *Eur. J. Biochem.* **148**, 141–150.