

Plant Gene Register

Nucleotide Sequence of an *Arabidopsis* cDNA for Geranylgeranyl Pyrophosphate Synthase

Pablo A. Scolnik and Glenn E. Bartley*

Central Research and Development, E.I. Du Pont de Nemours and Company, Experimental Station, Wilmington, Delaware 19880-0402

GGPS is an enzyme involved in the early steps of isoprenoid synthesis in plants. The bell pepper (*Capsicum annuum*) GGPS catalyzes the stepwise addition of isopentenyl PPI, first to dimethylallyl PPI, second to geranyl PPI, and finally to farnesyl PPI to form geranylgeranyl PPI (Kuntz et al., 1992). Geranylgeranyl PPI is utilized by plants for many compounds such as carotenoids, Chls, and phyloquinones and as a precursor for GA₃. A plant GGPS cDNA from *Capsicum* ripening fruit has been cloned and sequenced, but only the deduced amino acid sequence has been published (Kuntz et al., 1992).

To obtain a GGPS cDNA, we have cloned, by amplification using the PCR, a gene fragment for GGPS from *Arabidopsis* using degenerate oligonucleotides based on conservation of amino acids among several GGPSs (Carattoli et al., 1991; Michalowski et al., 1991; Kuntz et al., 1992; Math et al., 1992). An upstream oligonucleotide, 5'-GCIGCITG(T/C)GCIGTIGA(A/G)ATG-3' (I = inosine), corresponding to the peptide sequence AACAVEM, and a downstream oligonucleotide, 5'-GGTTACATCCAAGAT(A/G)TCGTC-3', corresponding to the peptide sequence DDILDVT, were used to amplify *Arabidopsis* genomic DNA. Upon electrophoresis of the resulting PCR products, a single band of about 450 bp was observed; it was purified and subcloned into *Sma*I-cut pBluescript using a procedure developed in our laboratory. This fragment was used to screen a cDNA library constructed from norflurazon-treated *Arabidopsis* seedlings (Table I). One plaque hybridizing to the probe was purified and the cDNA insert sequenced. Analysis of the sequence revealed an open reading frame encoding a protein of 371 amino acids. Alignment of this protein to GGPSs from *Erwinia uredovora* and *Rhodobacter capsulatus* suggests an N-terminal transit peptide of 76 amino acids. Comparison of this protein sequence to the protein sequence of GGPS from pepper showed 84.5% similarity and 73.5% identity between the putative mature peptides and, unsurprisingly, very little similarity or identity between the putative transit peptides. Sequence differences between the cDNA and the genomic fragment indicate the existence of more than one GGPS gene in *Arabidopsis*. We propose naming this gene GGPS1. The genomic fragment has been mapped to the lower arm of chromosome 2, whereas

Table I. Characteristics of an *A. thaliana* GGPS cDNA

Organism:	<i>Arabidopsis thaliana</i> (L.) Heynh., WS ecotype, Brassicaceae.
Enzyme, Function:	GGPS, isoprenoid biosynthesis.
Source:	Custom cDNA library in λ ZAPII constructed by Clontech from norflurazon-treated seedlings and screened by DNA hybridization.
Cloning Strategy:	PCR of genomic DNA using degenerate oligonucleotides based on conserved amino acid sequences of GGPS and subsequent hybridization to the cDNA library.
Techniques:	Genomic DNA was amplified in a Perkin-Elmer 9600 thermocycler for 35 cycles using components of a PCR kit (Perkin-Elmer Cetus, Norwalk, CT) and the following program: denaturation, 92°C, 0.5 min, ramp time 0.5 min; annealing, 50°C, 0.5 min, ramp time 2 min; extension, 72°C, 3 min, ramp time 1 min; final extension, 72°C, 8 min. The resulting PCR product was treated with the large fragment of T4 polymerase and with polynucleotide kinase, gel purified using Spin-X (Costar, Cambridge, MA), and ligated into <i>Sma</i> I-cut pBluescript KS ⁻ . The sequence of the subcloned insert and cDNA was determined by dideoxy sequencing of double-stranded DNA using a Sequenase kit (United States Biochemical).
Characteristics of cDNA:	cDNA of 1242 nucleotides containing a 5' untranslated region of 29 bp and a 3' untranslated region of 100 bp. No poly(A) tail observed.
Structural Features of the Deduced Amino Acid Sequence:	The open reading frame indicates a protein of 371 amino acids with a mol wt of 40,206. The putative mature protein shows 84.5% similarity to and 73.5% identity with the <i>Capsicum</i> mature peptide of GGPS. The putative transit peptide containing amino acid residues 1 to 76 shows very little homology to the putative transit peptide from <i>Capsicum</i> . The sequence analysis software of the Genetics Computer Group, University of Wisconsin, was used to analyze the cDNA and deduced peptide sequence.

the cDNA has shown no polymorphisms with identical genomic blots (Reiter et al., 1992).

Abbreviations: GGPS, geranylgeranyl pyrophosphate synthase; GGPS, gene encoding GGPS.

* Corresponding author; fax 1-302-695-8480.

ACKNOWLEDGMENTS

We would like to thank Leslie J. Dannenberger for her excellent work in mapping of the genomic fragment and Mary Bailey for the synthesis of the oligonucleotides.

Received November 23, 1993; accepted December 8, 1993.

Copyright Clearance Center: 0032-0889/94/104/1469/02.

The GenBank accession number for the cDNA sequence reported in this article is L25813.

LITERATURE CITED

- Carattoli A, Romano N, Ballario P, Morelli G, Macino G** (1991) The *Neurospora crassa* carotenoid biosynthetic gene (Albino 3) reveals highly conserved regions among prenyltransferases. *J Biol Chem* **266**: 5854–5859
- Kuntz M, Römer S, Suire C, Hugueney P, Weil JH, Schantz R, Camara B** (1992) Identification of a cDNA for the plastid-located geranylgeranyl pyrophosphate synthase from *Capsicum annuum*: correlative increase in enzyme activity and transcript level during fruit ripening. *Plant J* **2**: 25–34
- Math SK, Hearst JE, Poulter CD** (1992) The *crtE* gene in *Erwinia herbicola* encodes geranylgeranyl diphosphate synthase. *Proc Natl Acad Sci USA* **89**: 6761–6764
- Michalowski CB, Löffelhardt W, Bohnert HJ** (1991) An ORF323 with homology to *crtE*, specifying prephytoene pyrophosphate dehydrogenase, is encoded by cyanelle DNA in the eukaryotic alga *Cyanophora paradoxa*. *J Biol Chem* **266**: 11866–11870
- Reiter RS, Williams JGK, Feldmann KA, Rafalski JA, Tingey SV, Scolnik PA** (1992) Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proc Natl Acad Sci USA* **89**: 1477–1481