Mol Gen Genet (1995) 249:58–64

ORIGINAL PAPER

Pascale Gaubier · Hui-Ju Wu · Michèle Laudié
Michel Delseny · Françoise Grellet

A chlorophyll synthetase gene from *Arabidopsis thaliana*

Received: 28 April 1995 / Accepted: 15 June 1995

Abstract During the course of an *Arabidopsis thaliana* genome sequencing project, we identified a gene, *G4*, with a derived amino acid sequence showing homology to the product of the *Rhodobacter capsulatus bchG* locus which is involved in the esterification of bacteriochlorophyllide with geranylgeraniol. The relationship between this gene and *bchG* was confirmed by the isolation and analysis of a corresponding full-length cDNA. Comparison of genomic and cDNA sequences indicated that the gene is made up of 14 exons, some of them being very short. Southern and Northern analyses showed that this sequence represents a single-copy gene and its transcript is detected only in green or greening tissues. Both homologies and expression data suggest that this gene encodes a chlorophyll synthetase, one of the last enzymes of chlorophyll biosynthesis, and thus represents a new example of a nuclear gene encoding an enzyme of this pathway in higher plants.

Key words *Arabidopsis thaliana* · Chlorophyll synthetase

Introduction

In studies on chlorophyll biosynthesis in higher plants, the term "chlorophyll synthetase" was used first to describe an enzyme different from chlorophyllase, which was associated with etioplast membrane fractions isolated from etiolated oat seedlings, and which was responsible for the light-independent esterification of chlorophyllide *a* to chlorophyll *a*. (Rüdiger et al. 1980). These authors showed that chlorophyll synthetase preferentially used as substrates the diphosphate derivatives of geranylgeraniol (GGPP) or phytol (PPP). This final step in chlorophyll biosynthesis was further characterized in spinach chloroplasts (Soll and Schultz 1981; Soll et al. 1983). It was shown that chlorophyll synthetase is firmly bound to the thylakoid membrane, where the enzyme catalyzes the esterification of chlorophyllide *a* with GGPP, yielding Chl_{GG}, which is converted to Chl_{Ph} by a subsequent reduction of geranylgeraniol to phytol.

In photosynthetic bacteria of the genus *Rhodobacter*, the essential loci involved in bacteriochlorophyll *a* biosynthesis are clustered in a 46 kb region of the chromosome (for a review, see Bauer et al. 1993). Sequence analyses and studies of insertional or directed mutations in a number of the open reading frames present in this region allowed the assigment of several of these ORFs to specific enzymes in the bacteriochlorophyll biosynthetic pathway (for a review, see Bollivar et al. 1994a). ORF304 (Alberti 1991; EMBL sequence submission, accession number Z11165), also termed *bchG* (Taylor et al. 1993), was demonstrated to encode the chlorophyll synthase required for the esterification of bacteriochlorophyllide *a* with GG (Bollivar et al. 1994b). Although the catalytic activity of *BchG* protein remains to be demonstrated in vitro, this polypeptide represents the prokaryotic homolog of chlorophyll synthetase, the enzymatic activity which was described in higher plants by Rüdiger and coworkers (1980). To date, only nuclear genes involved in the early steps of chlorophyll biosynthesis in higher plants have been isolated (Koncz et al. 1990, 1992; Hudson et al. 1993). They encode proteins that show homology to the products of the *Rhodobacter bchI* and *bchH* genes necessary for chelation of magnesium by protoporphyrin IX, a key step in bacteriochlorophyll biosynthesis.

As part of the European genome sequencing effort to identify new genes in *Arabidopsis thaliana*, we have

Sequence described in this paper has been deposited in the Genbank databank under the accession number U19382

Communicated by A. Kondorosi

P. Gaubier · H.-J. Wu · M. Laudié · M. Delseny · F. Grellet (✉)
Laboratoire de Physiologie et Biologie Moléculaire des Plantes, URA 565 du CNRS, Université de Perpignan, Avenue de Villeneuve, F-66860 Perpignan-Cédex, France

sequenced the regions surrounding the gene encoding the late embryogenesis protein Em1 (Gaubier et al. 1993) which has been mapped on chromosome 3. The *Em1* gene was located in the 13.7 kb insert of a lambda phage genomic clone, λGEA1, which was completely sequenced. By sequence comparison we found that the gene immediately downstream of *Em1*, on the opposite strand, provisionally named *G4*, presented some similarity to the ORF304 (or *bchG* gene) located within the *Rhodobacter capsulatus* photosynthesis gene cluster. A fragment of the plant gene was used to isolate a full-length *A. thaliana* cDNA clone. Amino acid comparison of the encoded polypeptide with that derived from the *R. capsulatus bchG* gene reveals highly conserved motifs which strongly suggest that the *A. thaliana G4* gene described here encodes a chlorophyll synthetase.

Materials and methods

Plant material

Arabidopsis thaliana (L.) Heynh seeds, ecotype Columbia (Col-0), were obtained from Dr. Jerôme Giraudat (ISV, Gif-sur-Yvette, France). The plants were grown in a growth chamber on sterile peat. Organs for DNA or RNA extraction were collected, frozen in liquid nitrogen and stored at − 80° C.

Isolation and sequencing of genomic and cDNA clones

The genomic clone λGEA1 was isolated from an *A. thaliana* (Columbia) library constructed in EMBL4 and kindly provided by Dr Bernard Lescure (CNRS-INRA, Toulouse, France). This library was constructed by insertion of partial *Sau*3A digests of genomic DNA into the *Bam*HI sites of the vector (Axelos et al. 1989). The nucleotide sequence of the insert was determined on the first strand by the dideoxynucleotide chain termination method on double-stranded DNA (Chen and Seeburg 1985), using nested *Exo*III deletions (Henikoff 1984). Sequence of the complementary strand was established using selected oligonucleotides spaced every 250 bp.

A 2.4 kbp *Aat*II-*Eco*RI restriction fragment overlapping a large part of the region containing the putative chlorophyll synthetase gene was isolated, labelled by random primer extension (Feinberg and Vogelstein, 1983) and used as probe to screen an *A. thaliana* (Columbia) green shoot cDNA library constructed in vector λ Zap II and kindly provided by Dr Martin Kreis (Orsay, France). Some 300,000 plaques were screened and positive plaques were selected by three rounds of replating and hybridization. Fifteen positive clones were finally recovered. One of these clones, containing the largest insert corresponding to the expected length, was sequenced in its entirety on both strands.

Southern and Northern blot analyses

Genomic DNA was isolated from *A. thaliana* (Columbia) plantlets as described earlier (Dellaporta et al. 1983). Restricted DNA (2 μg per digest) was resolved on an 0.8% agarose gel and transferred to Hybond N membrane (Amersham) according to the supplier's specifications, then hybridized at 65° C with recommended solutions, using as a probe the *G4* cDNA labelled by random primer extension.

Subsequent washing of the membrane was performed at room temperature in 2 × SSC, 0.1% w/v SDS, then at 65° C in 1 × SSC, 0.1% w/v SDS. For DNA digests, restriction enzymes were chosen that have no cleavage site in the *G4* gene (*Bam*HI, *Hin*dIII, *Xho*I), one site (*Eco*RI) or more than one site (*Eco*RV).

Total RNA was extracted (Lee Downing et al. 1992) from various *A. thaliana* (Columbia) organs: flower buds, flowers, immature (small green), mid-stage (large green) and mature (yellow) siliques, dry seeds, stems, mature leaves (rosette) from adult plants, leaves from young plantlets, and seeds allowed to germinate for 5, 10, 30, 50 and 72 h in water in petri dishes under continuous light. RNA was electrophoresed (10 μg per sample) on 1% agarose-formaldehyde gel (Sambrook et al. 1989) then transferred to Hybond N membrane and hybridized with the labelled *G4* cDNA at 42° C in the presence of 50% v/v formamide. Washing of the membrane was performed as for Southern blotting.

Sequence data analyses

Nucleotide and amino acid sequence analyses were done with the BLAST program (Altschul et al. 1990) using the Blast Network service at the NCBI. Polypeptide sequence alignments were obtained using the CLUSTAL V software (Higgins and Sharp, 1988) with its default parameters (gap penalty: 3, K-tuple: 1, no. of top diagonals: 5, window size: 5, protein weight matrix: PAM 250).

Results and discussion

Characterization of the chlorophyll synthetase gene, isolation and sequencing of the corresponding cDNA

The insert of the genomic clone λGEA1 is 13 738 bp in length and contains the *Em1* gene previously described

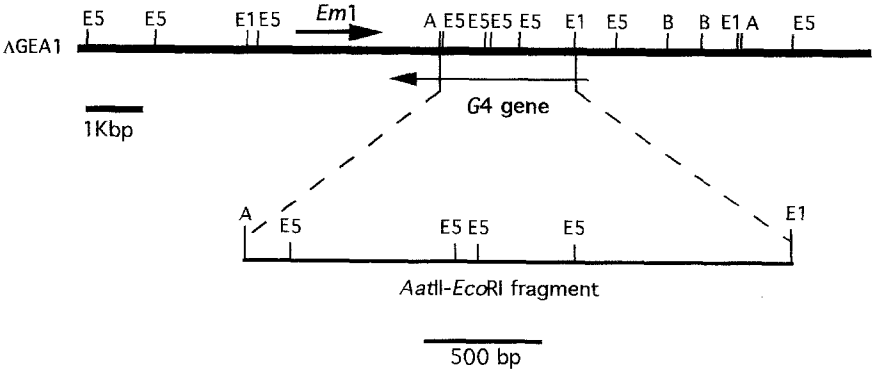


Fig. 1 Map of the λGEA1 genomic clone. Representative restriction sites: E1, *Eco*RI; E5, *Eco*RV; A, *Aat*II; B, *Bam*HI. Positions of the *Em1* and *G4* genes are indicated by *arrows*. The enlarged portion of the clone shows the *Aat*II- *Eco*RI fragment used as the probe for isolation of cDNAs corresponding to *G4*

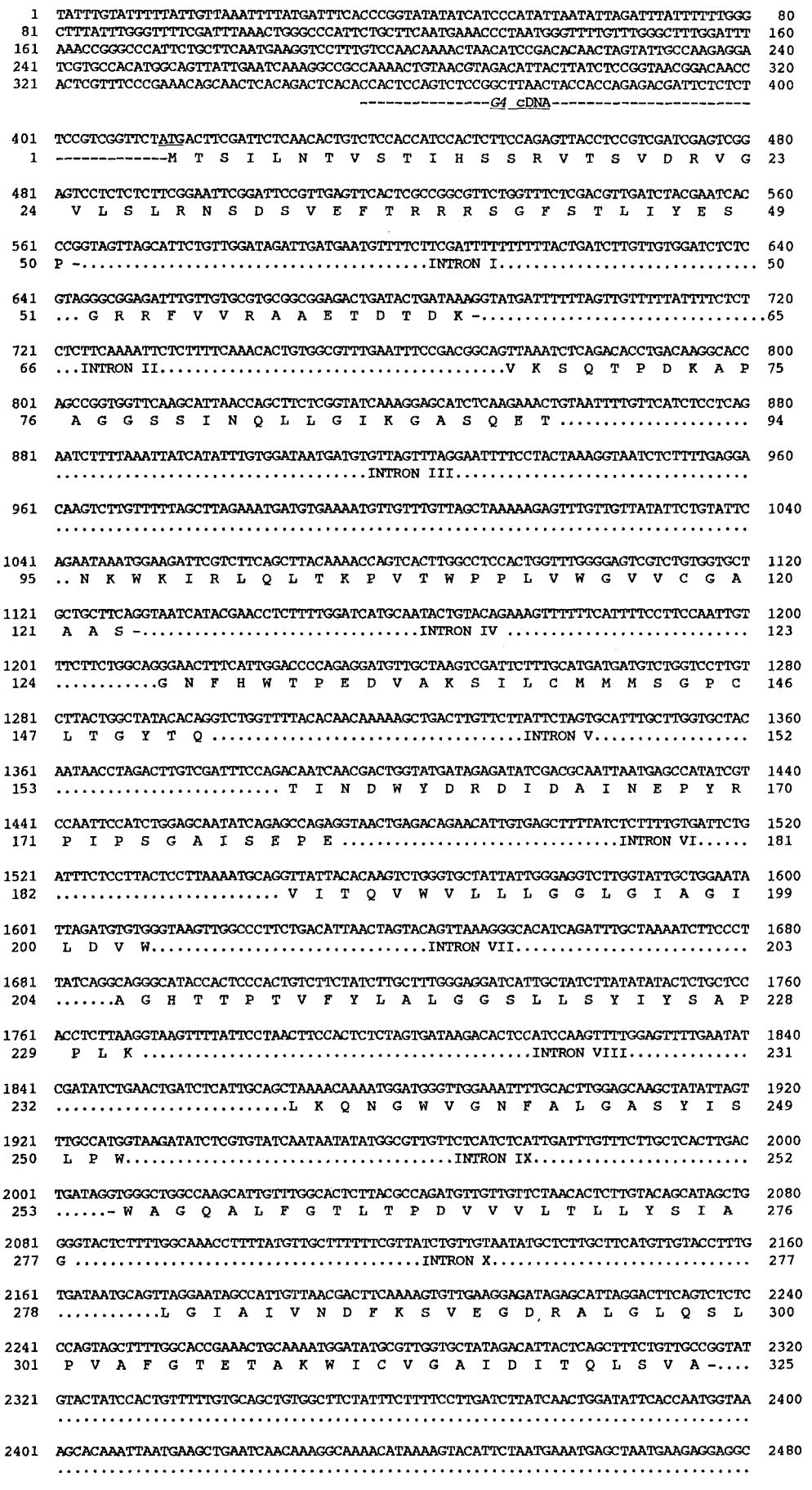


Fig. 2 Continued

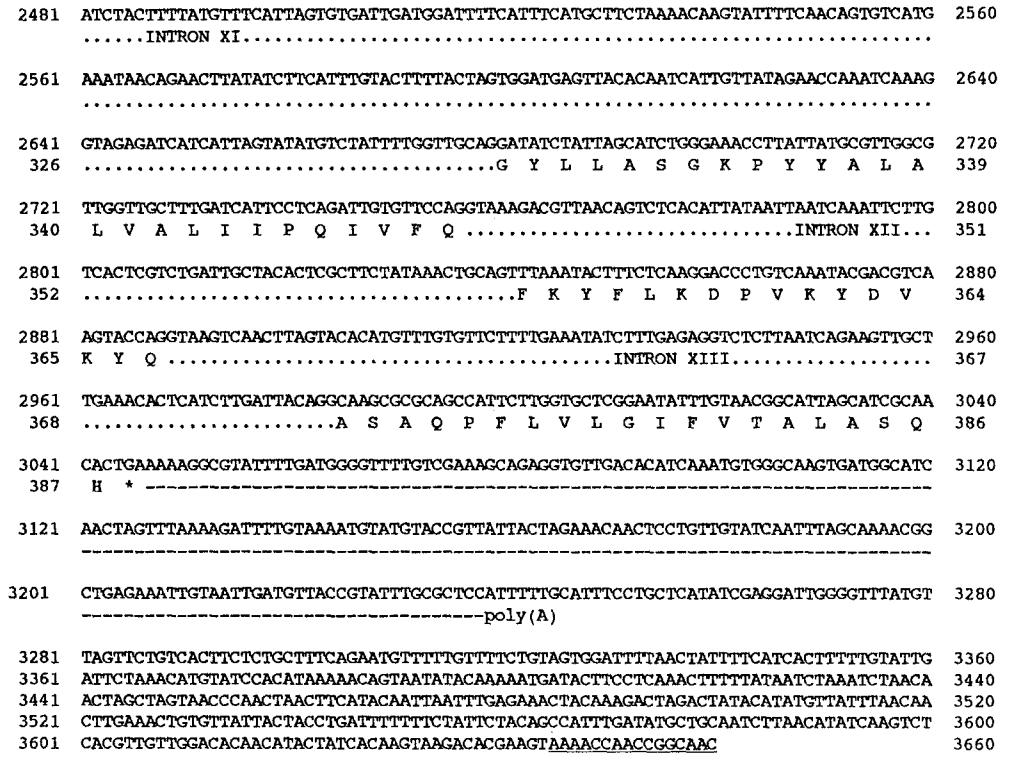


Fig. 2 Nucleotide sequence of the *G4* gene and its cognate cDNA. The DNA sequence extends to the 3′ end (*underlined*) of mRNA encoded by the *Eml* gene indicated in Fig. 1. *Dashed lines* below the genomic DNA sequence indicate the ends of the *G4* cDNA clone. The translation start of the G4 coding region is *underlined*. The amino acid sequence obtained by translation of the cDNA sequence is shown below the DNA sequence. The stop codon is marked by an *asterisk*

(Gaubier et al. 1993). Downstream of *Eml*, on the opposite strand, sequential ORFs located on four contiguous *Eco*RV fragments (Fig. 1) were compared to the non-redundant peptide sequence databases, and were found to match the polypeptide encoded by the *Rhodobacter capsulatus bchG* gene, suggesting that each corresponded to a different exon of the same gene.

In order to identify the entire coding sequence and to delineate the exon and intron positions, the 2.4 kb *Aat*II-*Eco*RI restriction fragment overlapping a large part of this region (Fig. 1) was used as probe to isolate a full-length cognate cDNA clone.

Sequences of the 3.6 kb genomic region containing the putative gene, of its corresponding cDNA and of the predicted translation product are shown in Fig. 2. The cDNA sequence is 1413 nucleotides (nt) long. It contains a methionine-initiated ORF of 1161 nt, encoding a protein of 387 amino acids, of predicted molecular weight 42 kDa. A stop codon is found 36 nt upstream of this methionine, indicating that the cDNA contains the full-length coding sequence. The 3′ untranslated region is relatively short (195 nt) and no classical polyadenylation signal can be found before the poly (A) tract. Downstream from the 3′ end of the cDNA sequence, at 406 nt, is found the 3′ end (on the opposite strand) of the *Eml* gene transcript. In the 5′ non-transcribed region of the gene, no typical TATA box was found at the usual distance upstream from the begining of the cDNA sequence. The closest one is located 370 bp upstream of the ATG. The mRNA corresponding to the cDNA sequence comprises 14 exon sequences, the shortest one contains 45 nt (exon 2) and the longest 208 nt (exon 1). Exon sequences share complete identity with the cDNA sequence. The intronic sequences are generally short, the smallest (intron VI) is 73 nt long, close to the minimal length requirement of between 70 and 73 nt for efficient splicing of plant introns (Goodall and Filipowicz 1990).

Amino acid sequence comparison of the *A. thaliana G4 sequence to the Rhodobacter capsulatus* ORF304 (*bchG* gene)

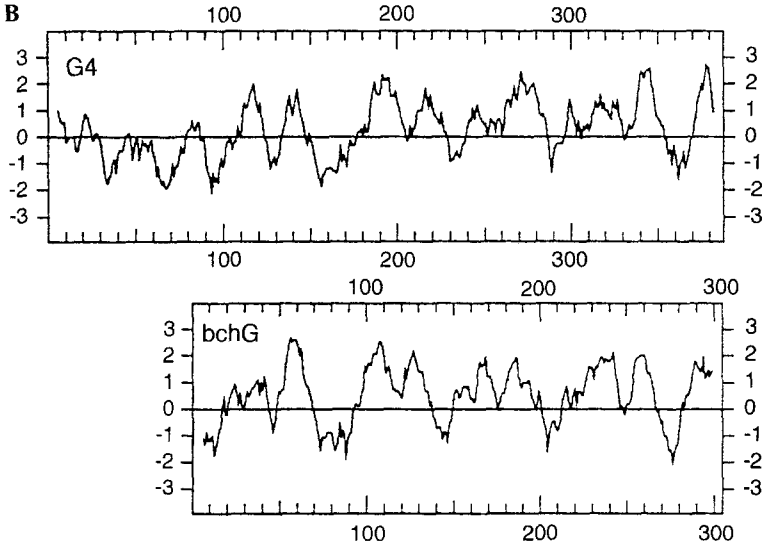
An alignment of the *G4*-encoded polypeptide with the polypeptide encoded by *bchG* is shown in Fig. 3A. Positions of introns are indicated in order to delineate exons in the G4 protein sequence. Alignment of the whole *R. capsulatus* protein, which contains 304 amino acids, begins at residue 90 of the G4 protein, after which 33.6% amino acid sequence identity and 61% amino acid sequence similarity are found between both proteins.

Three domains are particularly well conserved, with 60% (domains I and III) to 75% (domain II) amino acid identity. The existence of these very closely related domains in both proteins, together with the existence, between them, of common stretches of conservative

A

```
Ath  MTSILNTVSTIHSSRVTSVDRVGVLSLRNSDSVEFTRRRSGFSTLIYESPGRRFVVRAAETDTD
Rca  ............................................................

                                                  ---------- Domain I ----------
Ath  KVKSQTPDKAPAGGSSINQLLGIKGASQET-------NKWKIRLQLTKPVTWPPLVWGVVCGAA
Rca  ........................MSAQDLSPSRRSIPEPRAMLELIKPVTWFPPMWAYLCGAV
                              +++*           +  *** ***** * +** +***

                                        ---------- Domain II ----------
Ath  ASGNFHWTPEDVAKSILCMMMSGPCLTGYTQTINDWYDRDIDAINEPYRPIPSGAISEPEVITQ
Rca  SSNVPIW--ENKGVVVLGIVLAGPIVCGMSQAANDWCDRHVDAINEPHRPIPSGRIPG------
     ++    *  ** +  ** ++**** + * +** *** **+******** ****** *+

Ath  VWVLLLG-GLGIAGILDVWAGHTTPTVFYLALGGSLLSYIYSAPPLKLKQNGWVGNFALGASYI
Rca  LWGLYIAIAMSLLSLVVGWQL--GSWGFVATLLGVAAAWAYSVEPIRLKRSGWWGPGLVGLAYE
     ** * ++ ++++ +++  *     +  *  ** *   +  **  *+******** *   ** **

                                        ---------- Domain III ----------
ATH  SLPWWAGQALF---GTLTPD--VVVLTLLYSIAGLGIAIVNDFKSVEGDRALGLQSLPVAFGTE
Rca  GLPWITGAAVLLATADTSPGFPIVMMATLYALGAHGIMTINDFKAIEGDRKLGIKSLPAVYGPE
     +*** ** *++   +  +**  +**** ****** **  +********** *******  ** *

Ath  TAKWICVGAIDITQLSVAGYLLASGKPYYALALVALIIPQIVFQFKYFLKDPVKYDVKYQASAQ
Rca  VAAKIACTVMGLAQALVITMLYLFSKPYHATAVLVLLCGQF-WAMSVWMRDPEGKAPWYNGTGV
      *  *    ++++*  *   *   +*** * *++ *+  *+       +**       *++++

Ath  PFLVLGIFVTALASQH
Rca  VMYVSGMMITAFAIRGFTV
        * *+ +***+ +
```

B



C

```
                    domain I
Ath  101 LQLTKPVTWPPLVWGVVCGAAASG 124
Rca   20 LELIKPVTWFPPMWAYLCGAVSSN  43
Cau   18 IELADPVTWISPVLVCFCGALASG  41
           +++  **** + +   +*** +*

                    domain II
Ath  154 INDWYDRDIDAINEPYRPIPSGAI 177
Rca   71 ANDWCDRHVDAINEPHRPIPSGRI  94
Cau   75 INDYFDRDLDAINDPQRPIPAGII  98
           **  ***+******+* *****+* *

                    domain III
Ath  282 VNDFKSVEGDRALGLQSLPVAFGTE 307
Rca  198 INDFKAIEGDRKLGIKSLPAVYGPE 222
Cau  198 LNDIKSVEGDRKLGLQSLTVAIGVK 222
           +***+*+***** *****+++* *
```

amino acid replacements, strongly suggest that the *A. thaliana* G4 protein is homologous to the bacterial bchG protein sequence.

The first 89 N-terminal residues of the *A. thaliana* protein are without conterpart in the bacterial protein. These residues correspond to the two first exons and the major part of the third exon. They could represent a composite targeting sequence comprising sequentially cleavable peptides necessary for protein translocation from the cytoplasm to the thylakoid membrane, as expected for a nuclear encoded, plastid chlorophyll synthetase (Soll et al. 1983; Lindsten et al. 1990). This possibility was evaluated by analyzing the G4 amino

Fig. 3A Protein sequence alignment between the *G4*-encoded protein of *Arabidopsis thaliana* (Ath) and the *bchG*-encoded protein of *Rhodobacter capsulatus* (Rca). Identical residues are indicated by *asterisks*, conservative replacements by *crosses*, and deletions (gaps) by *dashes*. Positions of introns in the *A. thaliana* gene sequence are shown with *arrowheads*. The three conserved domains are indicated by *dashed lines* above the protein sequences. **B** Comparison of hydropathic plots of G4 and bchG proteins. **C** Alignment of the three conserved domains found in the *A. thaliana* G4 protein (Ath), the *R. capsulatus* BchG protein (Rca) and in a protein of unknown function (Niedermeier 1994; Genbank accession number Z34000, ORF 2) of *Chloroflexus aurantiacus* (Cau). Identical residues are indicated by *asterisks* and conservative replacements by *crosses*. Residues at the begining and at the end of each domain are numbered according to their position in the corresponding protein sequence

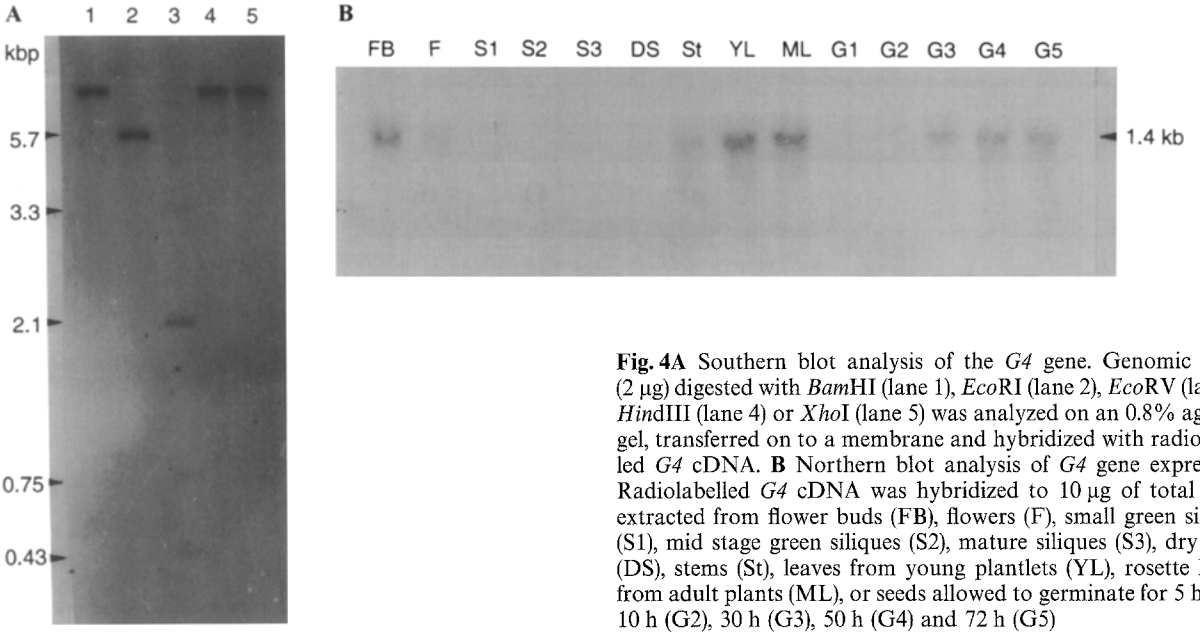


Fig. 4A Southern blot analysis of the *G4* gene. Genomic DNA (2 µg) digested with *Bam*HI (lane 1), *Eco*RI (lane 2), *Eco*RV (lane 3), *Hin*dIII (lane 4) or *Xho*I (lane 5) was analyzed on an 0.8% agarose gel, transferred on to a membrane and hybridized with radiolabelled *G4* cDNA. **B** Northern blot analysis of *G4* gene expression. Radiolabelled *G4* cDNA was hybridized to 10 µg of total RNA extracted from flower buds (FB), flowers (F), small green siliques (S1), mid stage green siliques (S2), mature siliques (S3), dry seeds (DS), stems (St), leaves from young plantlets (YL), rosette leaves from adult plants (ML), or seeds allowed to germinate for 5 h (G1), 10 h (G2), 30 h (G3), 50 h (G4) and 72 h (G5)

acid sequence with the Psort computer program (Nakai and Kanehisa 1992). Results predicted a chloroplast thylakoid membrane protein with a 73% probability.

Hydropathic plots comparison of plant and bacterial proteins (Fig. 3B) shows that the *A. thaliana* protein profile strongly resembles the bacterial protein profile (excluding the long amino-terminal sequence), with an hydrophobic domain distribution consistent with a transmembrane protein. This similarity of hydropathy profiles again argues for functional similarity between the two proteins.

In comparing the G4 protein with databases, we also found significant homologies with another complete polypeptide of unknown function encoded by a gene of the green filamentous bacterium *Chloroflexus aurantiacus* (Niedermeier 1994; Genbank accession number Z34000, ORF 2). This polypeptide also contains 304 amino acids and is described by the author as 31.5% identical and 59.8% similar to the BchG protein of *Rhodobacter capsulatus*. Its corresponding gene is located upstream from two genes encoding proteins of the chlorosome (Niedermeier et al. 1994). The amino acid sequence of this polypeptide is 37% identical and 61.7% similar to the G4 protein, and the alignment of both proteins also starts after the putative 89-amino acid targeting sequence of the plant protein (not shown). The three conserved domains are present in the *C. aurantiacus* protein, as shown in Fig. 3C, suggesting that this protein is a *C. aurantiacus* homolog of the *R. capsulatus* BchG protein and of the *A. thaliana* G4 protein.

Gene copy number and transcript analysis

In order to estimate the copy number of the *G4* gene, a genomic digest of Columbia ecotype DNA was probed with the cognate cDNA. As shown in figure 4A, autoradiography of the membrane reveals hybridizing fragments indicating the existence of a single *G4* gene copy in the genome of *A. thaliana*. The *Bam*HI, *Hin*dIII and *Xho*I digests show only one hybridizing fragment. The *Eco*RI digest shows one 5.7 kb fragment containing 90% of the *G4* gene, the 2.7 kb fragment containing the last 10% of the hybridizing sequences is not visible. The *Eco*RV digest shows four expected gene fragments, 3.3, 2.1, 0.75 and 0.43 kb in length, with labelling intensities directly related to their representation in the transcribed sequence; the smallest expected fragment (91 bp) is not visible because it has migrated out of the 0.8% agarose gel.

In order to analyze the expression pattern of the *G4* gene, total RNA was prepared from various parts of *A. thaliana* plants and used for Northern blot analysis with the *G4* cDNA as probe. Autoradiography of the membrane (Fig. 4B) showed a unique transcript with a size corresponding to that of the cDNA probe (1.4 kb). *G4* mRNA was detected, at a low level, in green tissues: flower buds and flowers, stems and leaves, and in greening cotyledons of seeds after 30 h of hydration and illumination. The transcript was also detectable in the total RNA of immature and mid-stage siliques after overexposure of the membrane (not shown) but not in the RNA of mature siliques, dry seeds, or 5 and 10-h germinated seeds. Nevertheless, expression of the *G4*

gene is not strictly light-dependent since a partial *G4* cDNA has been isolated recently from an *A. thaliana* cDNA library of 5-day-old etiolated seedlings grown in vitro (Desprez et al. Genbank accession number Z34566). The presence of such a transcript in etiolated plantlets is in agreement with the previous observations that chlorophyll synthetase is present in etioplasts as well as in chloroplasts (Rüdiger, 1992).

Definitive evidence that this gene encodes the chlorophyll synthetase awaits the production of the recombinant G4 protein and in vitro demonstration of its enzymatic activity with the appropriate substrates. This might be difficult because substrates are not commercially available and because the protein may require other proteins or a membrane environment for activity. It is also interesting to note that the *Em1* gene maps in the vicinity of the colour mutant *ch-6* (*chlorina*) (Fischerova 1975; Relichova 1976)) which was mapped on chromosome 3 (Koornneef et al. 1983) and it is tempting to speculate that the *ch-6* mutation affects the *G4* gene, although this could only be clearly demonstrated by complementation of the mutant. The detailed characterization of the cDNA and the sequence of this gene now opens up all these lines of research.

Acknowledgement We thank Dr P.A. Scolnik (The Du Pont Company, Wilmington) for RFLP mapping of *Em1* gene. This work was supported by the Centre National de la Recherche Scientifique (CNRS), the EEC ESSA (European Scientists Sequencing Arabidopsis) project and the Groupe de Recherche et d'Etude des Génomes (GREG).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Axelos M, Bardet C, Liboz T, Le Van Thai A, Curie C and Lescure B (1989) The gene family encoding the *Arabidopsis thaliana* translation elongation factor EF-1α: molecular cloning, characterization and expression. Mol Gen Genet 219:106–112

Bauer CE, Bollivar DW, Suzuki JY (1993) Genetic analyses 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 (1994a) Directed mutational analysis of bacteriochlorophyll *a* biosynthesis in *Rhodobacter capsulatus*. J Mol Biol 237:622–640

Bollivar DW, Wang S, Allen JP, Bauer CE (1994b) Molecular genetic analysis of terminal steps in bacteriochlorophyll *a* biosynthesis: characterization of a *Rhodobacter capsulatus* strain that synthesizes geranylgeraniol-esterified bacteriochlorophyll *a* Biochemistry 33:12763–12768

Chen Y, Seeburg PY (1985) Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA 4:165–170

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1:19–21

Feinberg AP, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonucleases fragments to high specific activity. Anal Biochem, 132:6–13

Fischerova H (1975) Linkage relationships of recessive chlorophyll mutations in *Arabidopsis thaliana* (L.) Heynh. Biol Plantarum 17:182–188

Gaubier P, Raynal M, Hull G, Huestis G, Grellet F, Arenas C, Pagès M, Delseny M (1993) Two different *Em*-like genes are expressed in *Arabidopsis thaliana* seeds during maturation. Mol Gen Genet 238:409–418

Goodall GJ, Filipowicz W (1990) The minimum length of pre-mRNA introns in monocots and dicots. Plant Mol Biol 14: 727–733

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

Higgins DG, Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237–244

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

Koncz C, Mayerhofer R, Koncz-Kalman Z, Nawrath C, Reiss B, Redei 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

Koncz C, Németh K, Rédei GP, Schell J (1992) T-DNA insertional mutagenesis on *Arabidopsis*. Plant Mol Biol 20:963–976

Koornneef M, van Eden J, Hanhart CJ, Stam P, Braaksma FJ, Feenstra WJ (1983) Linkage map of *Arabidopsis thaliana*. J Hered 74:265–272

Lee Downing W, Mauxion F, Fauvarque MO, Reviron MP, de Vienne D, Vartanian N, Giraudat J (1992) A *Brassica napus* transcript encoding a protein related to the Künitz protease inhibitor family accumulates upon water stress, not in seeds. Plant J 2:685–693

Lindsten A, Welch CJ, Schoch S, Ryberg M, Rüdiger W, Sundqvist C (1990) Chlorophyll synthetase is latent in well preserved prolamellar bodies of etiolated wheat. Physiol Plant 80: 277–285

Nakai K, Kanehisa M (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14: 897–911

Niedermeier G, Shiozawa JA, Lottspeich F, Feick RG (1994) The primary structure of two chlorosome proteins from *Chloroflexus aurantiacus*. FEBS Lett 342:61–65

Relichova J (1976) Some new mutants. Arab Inf Serv 13:25–28

Rüdiger W (1992) Last steps in chlorophyll biosynthesis: esterification and insertion into the membrane, In: Argyroudi-Akoyunoglou JH (ed) *Regulation of chloroplast biogenesis* Plenum Press, New York, pp 183–190

Rüdiger W, Benz J, Guthoff C (1980) Detection and partial characterization of activity of chlorophyll synthetase in etioplast membranes. Eur J Biochem 109:193–200

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, (2nd edn) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Soll J, Schultz G (1981) Phytol synthesis from geranylgeraniol in spinach chloroplasts. Biochem Biophys Res Commun 99: 907–912

Soll J, Schultz G, Rüdiger W, Benz J (1983) Hydrogenation of geranylgeraniol. Two pathways exist in spinach chloroplasts. Plant Physiol 71:849–854

Taylor DP, Cohen SN, Clark WG, Marrs BL (1983) Alignment of the genetic and restriction maps of the photosynthetic region of the *Rhodopseudomonas capsulata* chromosome by a conjugation-mediated marker rescue technique. J Bacteriol 154:580–590