Volume 296, number 3, 305–310 FEBS 10617 January 1992

Molecular cloning and expression in *Escherichia coli* of a cyanobacterial gene coding for phytoene synthase, a carotenoid biosynthesis enzyme

Daniel Chamovitz[a], Norihiko Misawa[b,*], Gerhard Sandmann[b] and Joseph Hirschberg[a]

[a]*Department of Genetics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel* and [b]*Lehrstuhl für Physiologie und Biochemie der Pflanzen, Universität Konstanz, Postfach 5560, D-7750 Konstanz, Germany.*

The first committed step in the biosynthetic pathway of carotenoids in plants and algae is the conversion of geranylgeranyl pyrophosphate (GGPP) to prephytoene pyrophosphate (PPPP), which is converted to phytoene. We have cloned the gene *pys* that encodes the enzyme phytoene synthase in the cyanobacterium *Synechococcus* PCC7942. The co-expression of *pys* in cells of *Escherichia coli* together with the gene *crtE* from *Erwinia uredovora*, which encodes geranylgeranyl pyrophosphate synthase, resulted in accumulation of phytoene. This result indicates that phytoene synthase is a single polypeptide enzyme that catalyzes the 2-step reaction from GGPP to phytoene. The deduced amino acid sequence of *pys* is highly conserved with that of pTOM5, a tomato cDNA that is differentially expressed during fruit ripening. These findings suggest that pTOM5 encodes phytoene synthase in tomato.

Carotenoid; *Synechococcus* PCC7942; *Erwinia uredovora*; Phytoene; Tomato

1. INTRODUCTION

Carotenoids form a family of pigments widely distributed in nature. They are essential components of the photosynthetic apparatus where they protect against photooxidative damage and also contribute to light harvesting for photosynthesis [1]. Specific carotenoid pigments are found in flowers and fruits of many species. In addition, carotenoids serve as precursors for abscisic acid in plants and for vitamin A in animals, and they play a role in preventing cancer in humans [2–4].

Carotenoids are synthesized by all photosynthetic organisms as well as by several non-photosynthetic bacteria and fungi [5]. The early steps in carotenoid biosynthesis from mevalonic acid are those of the central isoprenoid pathway, resulting in the formation of the C_{20} compound geranylgeranyl pyrophosphate (GGPP). The first reaction specific to carotenogenesis is the dimerization of 2 molecules of GGPP to form the colorless carotenoid phytoene, with prephytoene pyrophosphate (PPPP) as an intermediate. Stepwise desaturation reactions convert phytoene to lycopene and 2 cyclization reactions yield β-carotene. Most xanthophylls are oxygenated derivatives of β-carotene (Fig. 1).

It has been previously suggested that 2 enzymes are involved in the conversion of GGPP to phytoene in the non-sulfur photosynthetic bacterium *Rhodobacter capsulatus* – *crtB*, encoding PPPP synthase, and *crtE*, encoding phytoene synthase [6]. However, it has recently been demonstrated that the *crtE* gene from the non-photosynthetic bacteria *Erwinia uredovora* catalyzes the formation of GGPP when expressed in *E. coli* while the *crtB* gene product converts GGPP to phytoene [7].

A single monomeric protein has been isolated from *Capsicum* chromoplasts that catalyzes the dimerization of GGPP to produce PPPP and then converts it to phytoene [8]. The gene for GGPP synthase has been recently cloned from *Capsicum* [9].

The cyanobacterium *Synechococcus* PCC7942 provides an excellent model system for studying carotenoid biosynthesis. The pathway in cyanobacteria is similar to that of higher plants, while their prokaryotic nature allows for easy genetic manipulation [10]. Previously we have cloned the gene for phytoene desaturase from both *Synechococcus* PCC7942 and tomato, and found that the amino acid sequences of these enzymes are highly conserved ([11] and unpublished results). In order to further study carotenogenesis in oxygenic photosynthetic organisms, we are attempting to clone other genes involved in carotenoid biosynthesis from cyanobacteria. We present here the cloning and sequencing of the gene *pys* from the cyanobacterium *Synechococcus* PCC7942 and show, by expression in *Escherichia coli*, that it encodes phytoene synthase which converts GGPP to phytoene.

2. MATERIALS AND METHODS

2.1. *Plasmids and strains*

Plasmid pPDSdel35, which contains the *pds* gene, has been previously described [12]. In this plasmid the gene *pds*, which encodes phytoene desaturase from *Synechococcus* PCC7942, is fused to the first 17 amino acids of *lacZ* in the vector pBLUESCRIPT KS+, and is expressed in cells of *E. coli* giving an active enzyme. Plasmid pPDSdel35PYS was constructed by inserting a 580 bp *Bam*HI frag-

*On leave from Central Laboratories for Key Technology, Kirin Brewery Ltd., Gunma 370-12, Japan.

Correspondence address: J. Hirschberg, Department of Genetics, The Hebrew University of Jerusalem, Jerusalem, 91904 Israel. Fax: (972) (2) 633066.

ment, which is found downstream from *pds* in the *Synechococcus* PCC7942 chromosome, into the *Bam*HI site of pPDSdel35 so as to recreate the original genomic sequence. Plasmid pPDPYS was derived from pPDSdel35PYS by deleting 2 internal *Pst*I fragments of 0.63 and 0.62 kb. As a result an in-frame fusion was created between the first 87 codons of *pds* and the second codon of the open reading frame sequence (*pys*) that was identified downstream from *pds*.

Plasmid pACCRT-E was constructed by inserting a 1.73 kb *Asp*718 (*Kpn*I)–*Hin*dIII fragment from pCAR16 [13], which contains the *crtE* gene from *E. uredovora* coding for GGPP synthase, into a 4.09 kb *Eco*RV–*Hin*dIII digested plasmid pACYC184. Plasmid pACCAR25ΔcrtB, which carries the *E. uredovora* carotenoid biosynthetic genes *crtE* (GGPP synthase), *crtI* (phytoene desaturase), *crtY* (lycopene cyclase), *crtZ* (β-carotene hydroxylase), and *crtX* (zeaxanthin glucosidase), was constructed by inserting a 6.51 kb filled-in *Asp*718 (*Kpn*I)–*Eco*R1 fragment of pACCAR25delE [13] into the *Eco*RV site of pACYC184.

E. coli strain JM101 was used as a host for the plasmids. The corresponding transformants were grown in the dark according to established protocols. Ampicillin (100 μg/ml), chloramphenicol (30 μg/ml) and isopropyl-β-D-galactopyranoside (0.5 mM) were added as required.

2.1. *Nucleotide sequence determination and sequence analysis*

The 580 bp *Bam*HI fragment was cloned in both orientations in pBLUESCRIPT KS+. Sequencing was performed by the dideoxy nucleotide chain-termination reaction using [α-^{35}S]dATP. Polymerization reactions using T7 DNA polymerase (Promega) were primed by either the reverse or universal primers. An additional oligonucleotide was synthesized and used as a sequencing primer so that all DNA regions were sequenced at least twice in each orientation.

Amino acid sequences were aligned with the Pileup program of the UWGCG software package, version 7 [14]. Hydropathic pattern was calculated over a window of 9 residues [15].

2.3. *Carotenoid extraction and HPLC analysis*

Carotenoids from the different *E. coli* transformants were isolated from the pelleted freeze-dried cells by extraction with methanol containing 6% KOH for 20 min at 60°C and partitioning into diethylether/petrol (b.p. 35–80°C) (1:9, v/v). The diethylether/petrol phase was evaporated to dryness and resuspended in methanol. The carotenoids were separated by HPLC on a Spherisorb ODS-1 5 μm column with acetonitrile/methanol/2-propanol (85:10:5, v/v/v) as eluent. The carotenoids were separated and detected with a Waters 994 diode array detector and spectra were directly recorded from the elution peaks. Standard carotenoids were analyzed as described [16].

3. RESULTS

3.1. *Nucleotide sequence of phytoene synthase*

Previously we have published the nucleotide sequence of the gene encoding phytoene desaturase (*pds*) from *Synechococcus* PCC7942 [11]. In the course of sequencing the genomic region downstream from *pds* we have found an open reading frame (ORF) comprising 307 codons. This ORF begins 80 bp from the stop codon of *pds* (Fig. 2A). The nucleotide sequence of this gene, designated *pys* is shown in Fig. 2B. The deduced translation product of *pys* predicts a polypeptide of a calculated mol. wt. of 35.8 kDa and a hydrophobic index of −0.4.

3.2. pys *encodes phytoene synthase that converts GGPP to phytoene*

In order to elucidate the function of the *pys* gene product, this gene was expressed in cells of *E. coli* that also carried carotenoid biosynthesis genes from *E. uredovora* (see Materials and Methods). Two plasmids that express *pys* were constructed (Fig. 3). Plasmid pPDPYS

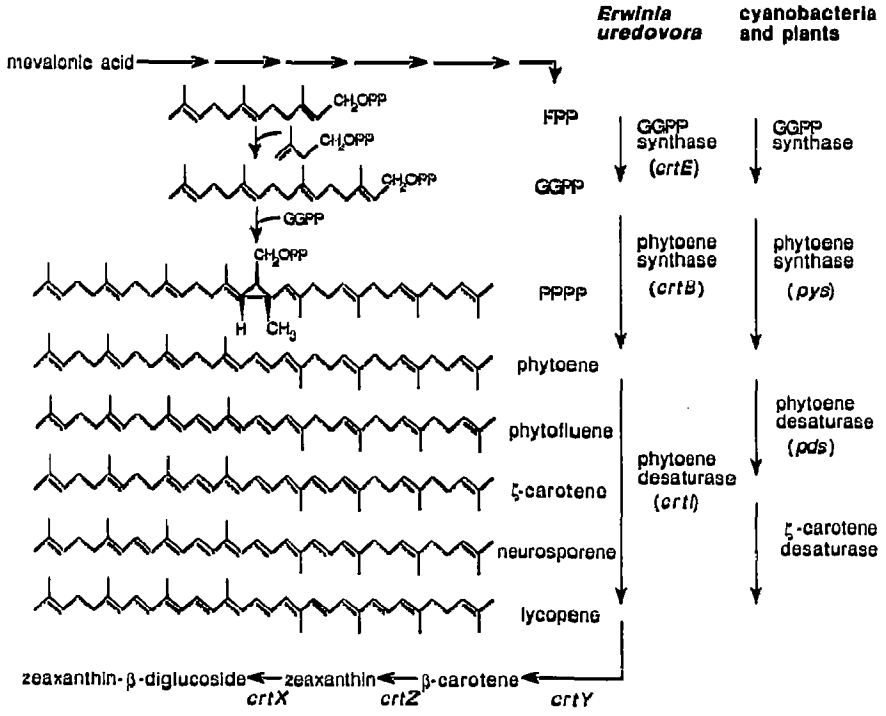
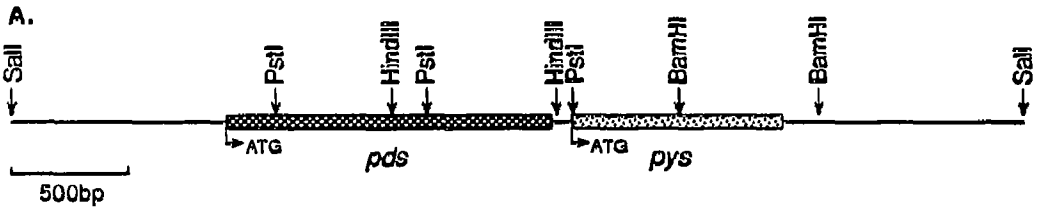


Fig. 1. Pathways of carotenoid biosynthesis in *E. uredovora* and in cyanobacteria and plants.



B.

```
1    CGCAGCAGCGGACGACCGCTGGCGGCGAGTCAGGCATAGACACTGTTACCCATGTTGCAGTTGCTAGCTCTGAAGCTTCACTGGTCACCCTCTCACCCCC
101  GACTCGCACCGTCTGTCTGATGCTGCAGATGATACCGGCACGGCCTTCACGGGCGCTCGCCCTCTCTGTCGGAGGCCTACGAGGAGTGCCGACAAATTACC
201  GCTCGCTACGCTAAGACCTTTTACCTCGGGACGCTTTTGATGCCCGAGGCCAAGCGTCAGGCCATCTGGGCCATCTACGTCTGGTGTCGTCGTACCGATG
301  AACTAGTCGATGGCCCTCAAGCAGCGCAGACTAACTTCGCGACCCTGGATGCTTGGGAGCGGCGGTTGGAACGACTGTTTGCTGGGGAACCCGAAGACGA
401  CTGCGATGTTGCCTTGGTGGATACCTTGGCTCGCTATCCTCTCGATATCCAGCCCTTCCGCGACATGATCGAAGGGCAGCGGATGGATCTGCTGCAAAAC
501  CGCTATTCCACTTTTGAGGATCTCTACACCTACTGCTACCGTGTCGCCGGTACTGTGGGACTGATGTCACAGCCGGTGATGGGCATTGAATCGACCAATA
601  GTCGTGCCCCTTGGGATCCGACAACTCCGCCCGACCCAACGGCAAGAAGCGCTAGCACTTGGCATTGCCAACCAGCTCACCAATATTCTGCGCGATGTCGG
701  TGAAGATGCCCGTCGAGGCCGCATCTATCTCCCCCAAGAGGAGCTAGCGCAGTTCAACTACAGCGAACAAGACTTGTTCAATGGCGTGATTGACGATCGC
801  TGGCGGGCCTTTATGCAGTTCCAGCTTGATCGCGCTCGTGATTACTTTGAGCAAGCCGAGCGGTGCATTGAACTCAGCCATGATGCTCGCTGGCCAGTCT
901  GGGCCTCCCTGATGCTCTATCGCGAAATTCTCGATGTGATTGAGCACAATAACTACGACGTCTTCCGAAAGCGGGCCTACGTACCGACTTGGCGGAAACT
1001 CTGTAGTCTGCCGGTGGCAATGCTACGAGCTACGGTTCTTTAGCTACGATCGCGGCCGCTTTTCTCGACCATCGATCAACGGTTCGTTGCCAGAAGGCGA
1101 TCAGTCGTTGAGCCTGGTGGACATAGCCCCCGCCAAACAAGTTGACGTGATTGAGAACGTGATAGAGGTTGTAGAGATCGCGGCGGCTGCGATAGCCCGG
```

Fig. 2. (A) Restriction map of *pds* and *pys* in the *Synechococcus* PCC7942 genome. (B) Nucleotide sequence of *pys*. The underlined sequence is the 3' end of the coding sequence of *pds*. The initiation and termination codons of *pys* are doubly underlined.

contains a chimeric gene made up of the first 87 codons of *pds* fused to the *pys* gene starting from the third codon. This gene was transcribed from the *lac* promoter. Plasmid pPDSdel35PYS is essentially pPDSdel35 to which the contiguous *Bam*HI fragment downstream of *pds* was inserted, completing the entire coding sequence of *pys*. Consequently, this plasmid carries *pds* and *pys* that are co-transcribed from the *lac* promoter. Two plasmids that carry carotenoid biosynthetic enzymes from *E. uredovora* were constructed in the plasmid vector pACYC184: plasmid pACCRT-E contains the *crtE* gene that codes for GGPP synthase and plasmid pACCAR25ΔcrtB contains all the carotenoid biosynthesis genes from *E. uredovora* except for *crtB* (Fig. 3). These genes are transcribed in *E. coli* from their endogenous promotors.

E. coli cells were co-transformed with the above plasmids in various combinations. Carotenoids were ex-

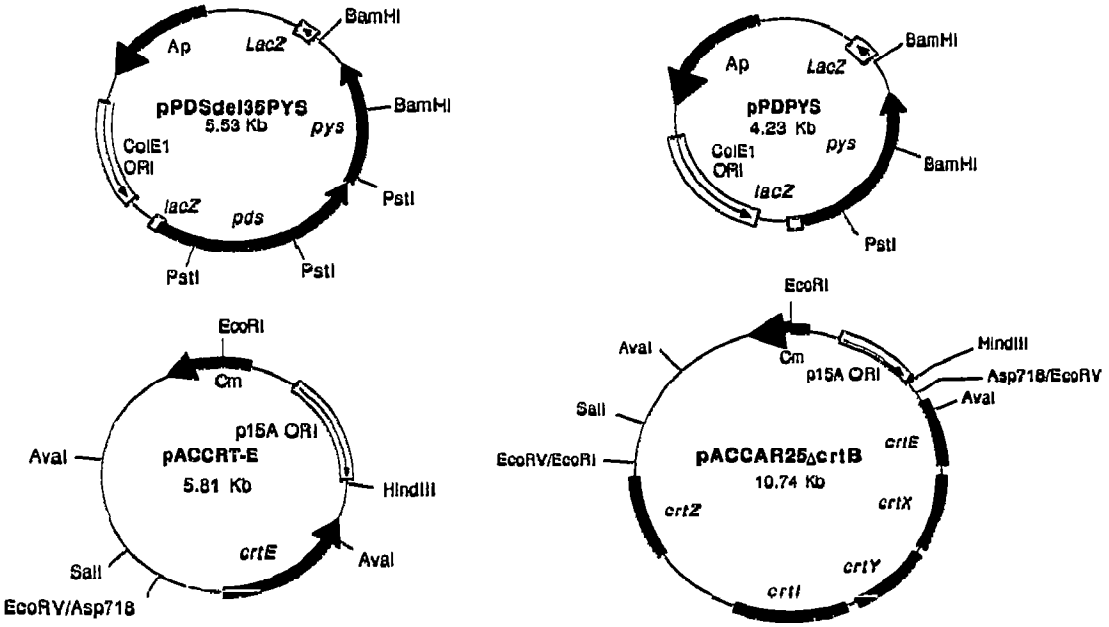


Fig. 3. Plasmids containing carotenoid biosynthesis genes that were constructed for transformation of *E.coli* (see Materials and Methods for details).

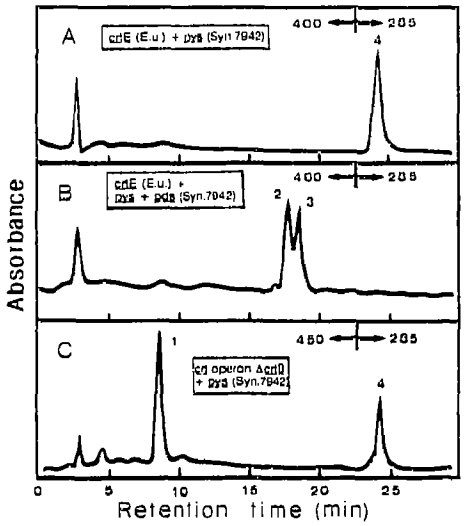


Fig. 4. HPLC analysis of carotenoids extracted from *E. coli* cells carrying plasmids: (A) pACCRT-E and pPDPYS. (B) pACCRT-E and pPDSdel35PYS. (C) pACCAR25ΔcrtB and pPDPYS. (1) zeaxanthin; (2,3) *cis*-isomeres of ζ-carotene; (4) phytoene.

tracted from the *E. coli* transformant cells and separated by HPLC. Cells of *E. coli* that contained plasmids pACCRT-E and pPDPYS, carrying genes *crtE* and *pys*, respectively, accumulated phytoene (Fig. 4A). Cells that contained pACCRT-E and pPDSdel35PYS, carrying *crtE*, *pys* and *pds*, accumulated *cis* ζ-carotene (Fig. 4B), and cells that contained pACCAR25ΔcrtB, carrying the entire carotenoid operon from *Erwinia* without *crtB*, and pPDPYS, accumulated phytoene and zeaxanthin (Fig. 4C). Carotenoid peaks in the HPLC profile were identified by their typical retention time and their absorbtion spectra (Fig. 5) [16]. Co-expression of pACCRT-E with an antisense construct of pPDSdel35PYS, or the expression of any of these plasmids alone, did not result in the accumulation of any carotenoids (data not shown). These results establish that *pys* encodes the enzyme phytoene synthase which catalyzes the conversion of GGPP to phytoene.

4. DISCUSSION

Cells of *E.coli* do not normally synthesize phytoene or other carotenoids, but they do produce trace amounts of GGPP. In contrast, *E. coli* cells carrying the *crtE* gene from *E. uredovora* accumulate larger amounts of GGPP [7]. The results presented in this work show that the gene *pys*, which is found downstream from *pds* in the *Synechococcus* PCC7942 genome, codes for phytoene synthase. In addition, it is apparent that PYS is a single polypeptide enzyme that mediates the conversion of GGPP to phytoene. Since this is a 2-step reaction in which PPPP is an intermediate, we postulate that PYS catalyzes both reactions.

E. coli cells that carry the plasmid pACCAR25, which contains all of the carotenoid biosynthetic genes from *E. uredovora*, accumulate zeaxanthin [13]. However, when *pys* complements the deficiency of *crtB* in pACCAR25ΔcrtB, both zeaxanthin and phytoene accumulate (Fig. 4C). The accumulation of phytoene is explained by the excess in phytoene synthase activity relative to the *Erwinia* enzymes that convert phytoene to zeaxanthin. This is because *pys* is carried by a high copy number plasmid and is overexpressed from a strong *E. coli* promoter (induced p*lac*), while the genes from *E. uredovora* are expressed from their endogenous promotors on a low copy number plasmid. When *pds* is overexpressed together with *pys*, no phytoene accumulates and only ζ-carotene is found (Fig. 4B).

The deduced amino acid sequence of *pys* is highly conserved (59% identities and 72% similarities) to the protein product of pTOM5, a tomato cDNA clone differentially expressed during fruit ripening (Fig. 6) [17]. However, pTOM5 codes for a larger polypeptide and therefore, the homology to *pys* begins only after a leader region of over 100 amino acid residues in the eukaryotic protein. Since pTOM5 is nuclear encoded and PYS is a plastid enzyme, we postulate that this region serves as a signal peptide which facilitates the import of PYS into plastids. The calculated molecular weight of PYS in cyanobacteria is 35.7 kDa. This is approximately the expected size of polypeptide which would be obtained if the pTOM5 protein product is cleaved around residue 100 during its import to plastids. In that case, the mature functional enzyme in tomato has the same size as

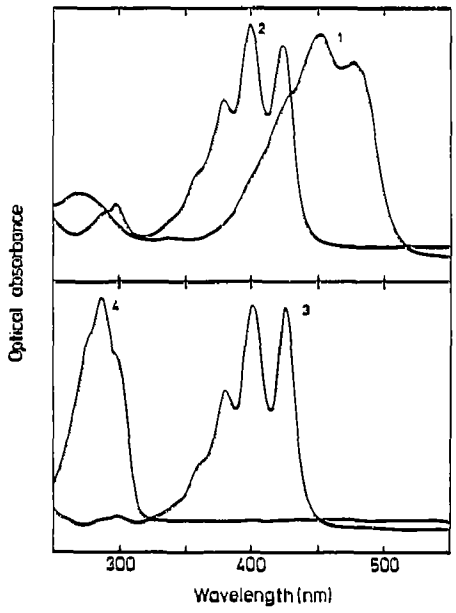


Fig. 5. Absorbtion spectra of the carotenoids isolated following HPLC. Numbers refer to the peaks of the HPLC separation in Fig. 4.

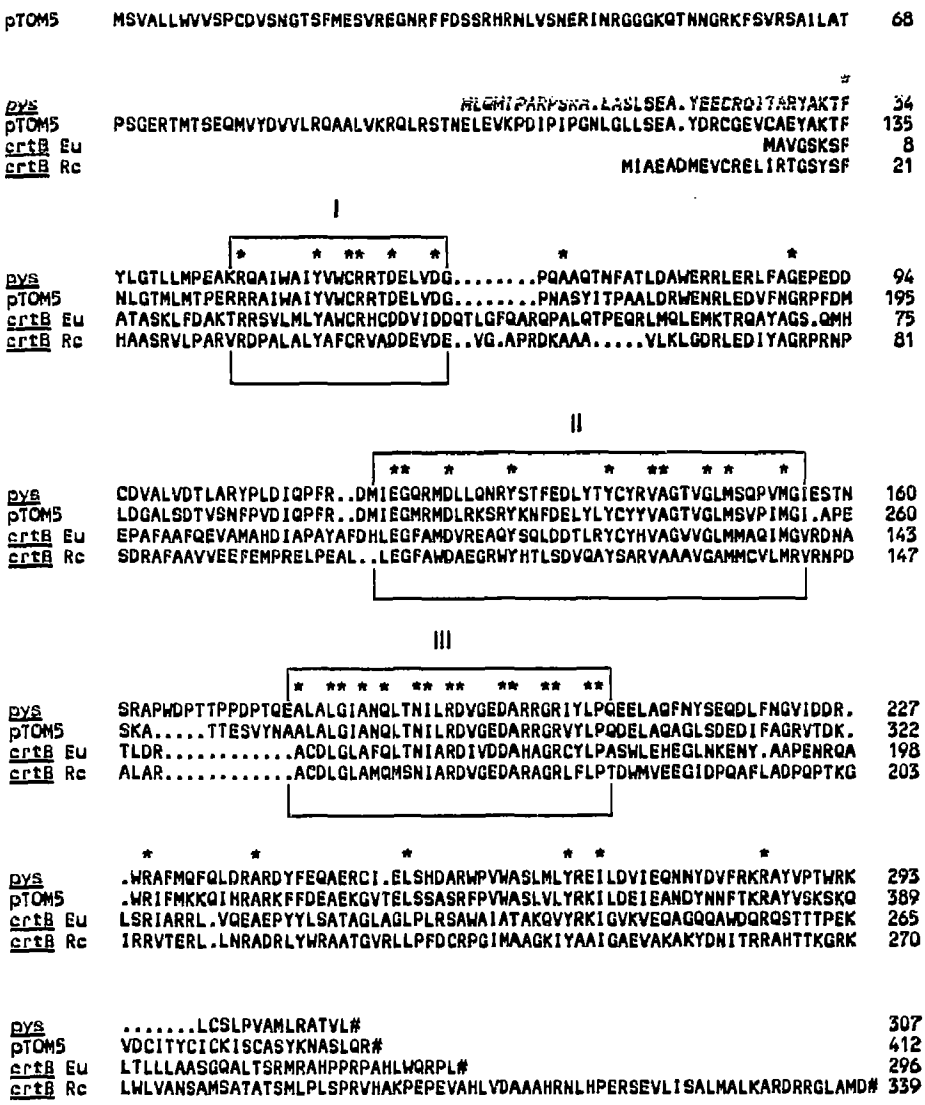


Fig. 6. Alignment of the deduced amino acid sequences of phytoene desaturase genes from various organisms. *pys*, *Synechococcus* PCC7942; pTOM5, tomato [18]; *crtB* Eu, *E. uredovora* [14]; *crtB* Rc and *Rb. capsulatus* [6]. Residues that are conserved in all 4 species are marked with asterisks; 3 regions of high homology are boxed.

its counterpart in cyanobacteria. However, there is no evidence yet for such a processing step.

The exact function of pTOM5 has yet to be elucidated. However, it has recently been implied that pTOM5 could encode an enzyme involved in the conversion of GGPP to carotenoids [18]. We suggest that pTOM5 (encoded by the tomato *pys* gene) is homologous to the cyanobacterial *pys*, both encoding the same enzyme.

A lower but significant homology is found between the cyanobacterial PYS and the *crtB* gene products of *E. uredovora* [13] (33% identities and 50% similarities) and *Rb. capsulatus* [6] (30% identities and 44% similarities). The alignment of all 4 phytoene synthase proteins reveals 3 regions that are highly conserved (Fig. 6). One of them, box III in Fig. 6, contains a short sequence of charged amino acids (residues 177–211) which show homology to GGPP synthase from *Capsicum* [9] and to farnesyl pyrophosphate (FPP) synthase from *E. coli* [19] and humans [20] (Fig. 7). The common feature to all these enzymes is their catalysis of prenyltransferase activity during condensation of allylic pyrophosphates to produce higher prenyl pyrophosphates or during condensation of 2 GGPP to produce phytoene. We speculate that this region in PYS is involved in the binding to and/or removing of the pyrophosphate.

It is not yet clear whether *pds* and *pys* are transcribed in *Synechococcus* PCC7942 independently or whether they constitute an operon. Northern hybridization analysis has not detected any *pys* transcripts, suggesting a

	190 200
pys	NILRDVGEDARRGRIYL
pTOM5	NILRDVGEDARRGRVYL
crtB (E.u.)	NIARDIVDDAHAGRCYL
crtB (R.c.)	NIARDVGEDARAGRLFL
GGPPS (C.a.)	DLPCMDNDDLRRGKPTM
crtE (E.u.)	DMPCMDDAKLRRGRPTI
FPPS (*E. coli*)	DLPAMDDDDLRRGLPFC
FPPS (human)	DDPIMDSSLTRRGQTCW

Fig. 7. Alignment of a region in PYS from *Synechococcus* PCC7942 with homologous sequences from pTOM5 [17], *E. uredovora* phytoene synthase (*crtB* E.u.) [13], *Rb. capsulatus* phytoene synthase (*crtB* R.c.) [6], bell pepper GGPP synthase [9], *E. uredovora* GGPP synthase (*crtE* E.u) [13], *E. coli* farnesyl pyrophosphate synthase (FPPS) [19] and human FPPS [20].

low abundance *pys* mRNA in cyanobacteria (data not shown). Interestingly, no transcripts of pTOM5 and *pds* could be detected in foliar tissues of tomato and soybean ([18,21], Pecker et al., submitted for publication). Since no consensus promoter sequences are found between *pds* and *pys*, the expression of both genes in the plasmid pPDSdel35PYS which is evident by the synthesis of ζ-carotene in *E.coli*, indicates that transcription can proceed from *pds* through to *pys*. The arrangement of *pds* followed by *pys* in the same operon is similar to that of *crtI* and *crtB* in *Rb. capsulatus* and *E. uredovora*. This is especially interesting since *pds* and *crtI*, both encoding phytoene desaturase, are very dissimilar in their primary structure and are thought to have arisen through convergent evolution ([11], Pecker et al., submitted for publication).

pys is the second gene cloned for the multi-enzyme pathway of carotenogenesis in oxygenic photosynthetic organisms. Genetic manipulation of *pys* and *pds* may enable the regulation of carotenoid content and composition in agronomically important crops. Studies in this direction are underway.

Acknowledgements: This research was supported by the German-Israeli Foundation for Scientific Research and Development GIF, Grant I-13-089, and by the Deutsches Bundesministerium für Forschung und Technology.

REFERENCES

[1] Goodwin, T.W. (1980) The Biochemistry of Carotenoids, vol. I, Chapman-Hall, New York.
[2] Zeevart, J.A.D. and Creelman, R.A. (1988) Annu. Rev. Plant Physiol. Plant Mol. Biol. 39, 439–473.
[3] Lechance, P. (1988) Clin. Nutr. 7, 118–122.
[4] Mathews-Roth, M.M., Lausen, N., Drouin, G., Richter, A. and Krinsky, N.I. (1991) Oncology 48, 177–179.
[5] Sandmann, G. (1991) Physiol. Plant. 83, 186–193.
[6] Armstrong, G.A., Alberti, M., Leach, F. and Hearst, J.E. (1989) Mol. Gen. Genet. 216, 254–268.
[7] Sandmann, G. and Misawa, N. (1991) FEMS Microbiol. Lett. (in press).
[8] Dogbo, O., Laferriere, A., d'Harlingue, A. and Camara, B. (1988) Proc. Natl. Acad. Sci. USA 85, 7054–7058.
[9] Kuntz, M., Römer, S., Weil, J.-H., Schantz, R. and Camara, B. (1991) Plant J. (in press).
[10] Chamovitz, D., Pecker, I., Sandmann, G., Boger, P. and Hirschberg, J. (1990) Z. Naturforsch. 45c, 482–486.
[11] Chamovitz, D., Pecker, I. and Hirschberg, J. (1991) Plant Mol. Biol. 16, 967–974.
[12] Linden, H., Misawa, N., Chamovitz, D., Pecker, I., Hirschberg, J. and Sandmann, G. (1991) Z. Naturforsch. 46c, 160–166.
[13] Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, I., Nakamura, K. and Harashima, K. (1990) J. Bacteriol. 172, 6704–6712.
[14] Devereux, J., Haeberli, H. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[15] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[16] Ernst, S. and Sandmann, G. (1988) Arch. Microbiol. 150, 590–594.
[17] Ray, J., Bird, C., Maunders, M., Grierson, D. and Schuch, W. (1987) Nucleic Acids Res. 15, 10587–10589.
[18] Bird, C.R., Ray, J.A., Fletcher, J.D., Boniwell, J.M., Bird, A.S., Teulieres, C., Blain, I., Bramley, P.M. and Schuch, W. (1991) Biotechnology 9, 635–639.
[19] Fujisaki, S., Hara, H., Nishimura, Y., Horiuchi, K. and Nishino, T. (1990) J. Biochem. 108, 995–1000.
[20] Wilkin, D.J., Kutsunai, S.Y. and Edwards, P.A. (1990) J. Biol. Chem. 265, 4607–4614.
[21] Bartley, G.E., Viitanen, P.V., Pecker, I., Chamovitz, D., Hirschberg, J. and Scolnik, P.A. (1991) Proc. Natl. Acad. Sci. USA 88, 6532–6536.