Molecular genetics of the carotenoid biosynthesis pathway in plants and algae

Joseph Hirschberg, Merav Cohen, Mark Harker, Tamar Lotan, Varda Mann and Iris Pecker

Department of Genetics, The Hebrew University of Jerusalem, Givat-Ram, Jerusalem 91904, ISRAEL

Abstract: During recent years genes for more than 20 different carotenogenic enzymes have been cloned from various organisms: bacteria, cyanobacteria, fungi, algae and plants. This accomplishment has provided new molecular tools to study the enzymes and yielded new information on their structure, function and regulation. We describe here the recent progress in the molecular genetics of the carotenoid biosynthesis pathway in plants. To date, the genes for almost all the enzymes, from the early steps of the isoprenoid pathway to the predominant xanthophylls, have been cloned. Their characterization had an immense impact on our understanding of carotenoid biosynthesis at the molecular level.

INTRODUCTION

Biosynthesis of carotenoids occurs in all photosynthetic organisms - bacteria, algae and plants, as well as in some non-photosynthetic bacteria and fungi. The intermediate steps in the carotenoid biosynthesis pathway were postulated several decades ago by standard biochemical analyses using labeled precursors, specific inhibitors and characterization of mutants. However, lack of genuine *in vitro* assays for the enzymes that are involved in the conversion of phytoene to the predominant carotenes and xanthophylls has hindered a more detailed elucidation of this part of the pathway. These enzymes are usually membrane bound and labile upon purification, thus it is very difficult to isolate them. In recent years the molecular-genetic approach to studying carotenogenesis has provided a wealth of information and new perspectives of both the enzyme activities and the regulation of the pathway.

In this report we describe recent progress in the molecular genetic analysis of carotenoid biosynthesis with emphasis on plants and algae. The readers are referred to previous excellent reviews on this subject (ref. 1-4).

GENE CLONING: STRATEGIES AND ACHIEVEMENTS

For molecular biology investigations it is essential to identify and clone the genes that encode the enzymes of interest. The traditional approach to cloning novel genes starts with the identification and isolation of their polypeptide products. This has been successfully employed in the case of carotenogenic enzymes whose purification was accomplished such as the plant enzymes: geranylgeranyl pyrophosphate synthase (GGPPS) from pepper (ref. 5) and violaxanthin deepoxidase (VDE) from lettuce (ref. 6).

Alternative ways for gene isolation take advantage of various genetic methods that do not rely on protein purification. In the carotenoid biosynthesis field these methods have been utilized first in prokaryotic systems where clusters of genes for carotenoid biosynthesis enzymes were identified by transposon mutagenesis or by functional expression in the bacterium *Escherichia coli* (reviewed in ref. 2,3). Sequence analysis indicated that homologous genes are significantly conserved among all of the bacterial and fungal species studied. The implication of this conclusion is that cloning of carotenoid genes from new bacterial or fungal species can rely on sequence similarities between homologous genes in conventional "shotgun" cloning methods, polymerase chain reaction (PCR) amplification and so forth.

In plants, however, the carotenoid genes are generally not conserved with those from bacteria. Therefore,

bacterial carotenoid genes cannot be used as molecular probes for the purpose of detecting genes from plants. Cloning of the first plant-type genes took advantage of the fact that the pathway in plants is similar to that of cyanobacteria (blue-green algae), which are prokaryotes and are amenable to sophisticated genetic manipulations. In this manner the plant-type genes for phytoene desaturase (*crtP*, *Pds*) (ref. 7,8) and lycopene β -cyclase (*crtL-b*) (ref. 9) were first isolated from the cyanobacterium *Synechocystis* sp. PCC7942. For that purpose mutants which are resistant to a specific inhibitor, norflurazon in the case of phytoene desaturase and MPTA in the case of lycopene cyclase, were isolated. Once it was determined that the mutation is related to the target site of the inhibitor, the resistance gene was obtained by complementing the resistance trait in a wild-type strain of cyanobacteria with a genomic library from the mutant (ref. 7). The cyanobacterial genes were successfully used as molecular probes to clone the homologous genes from algae and higher plants (ref. 10-12).

The cyanobacterial gene for phytoene synthase was discovered due to its localization downstream of crtP on the same operon and its ability to be functionally expressed in *E. coli* (ref. 13). This sequence in cyanobacteria (*crtB*) was found to share some sequence similarity with the tomato cDNA pTOM5, which was originally cloned randomly as a gene whose expression is specific to fruit ripening (ref. 14,15). pTOM5 is indeed the gene for phytoene synthase as was confirmed by anti-sense silencing of the gene in transgenic tomato plants (ref. 16) and by functional complementation in bacteria (ref. 17). A second *Psy* cDNA, which is constitutively expressed in the leaves, was also found in tomato (ref. 18).

The gene for ζ -carotene desaturase was identified among the randomly sequenced cDNA clones (ESTs) of *Arabidopsis*, (Accession No. T46272) due to its sequence similarity to *Pds* and then cloned from pepper (ref. 19). Functional expression of this cDNA in *E. coli* has confirmed its identity. The homologous gene from the cyanobacterium *Synechocystis* PCC6803, *crtQ*, has been identified during random sequencing of the whole genome (gene slr0940 in Accession No. D90914). It is interesting to note that a gene, termed *zds*, coding for an enzyme with ζ -carotene desaturase activity, was cloned from the cyanobacterium *Anabaena* PCC7120 (ref. 20). However, its amino acid sequence is similar to the bacterial-type phytoene desaturase (*crtQ*).

A novel method, termed "color complementation", to clone eukaryotic genes for carotenoid biosynthesis enzymes has been developed (ref. 21) (Fig.1). It takes advantage of the ability of the carotenogenic enzymes to function in the bacterium E. coli. In this method a recombinant pACYC184 plasmid is constructed with previously cloned genes from different species in such a way that they are functionally expressed in E. coli. Cells of E. coli that carry this plasmid produce a specific carotenoid which serves as a precursor for the enzyme under investigation. The carotenoid accumulated in the bacteria imparts a typical color that is visible to the naked eye in the colonies of this strain when grown on petri plates.



cDNA library from red cells of H. pluvialis

Fig. 1. Schematic illustration of the "color complementation" screening method for cloning cDNA which encodes a carotenoid biosyntesis enzyme. In this case, the cloning of *crtO*, the cDNA for β -C-4-oxygenase from *Haematococcus pluvialis*, is depicted.

© 1997 IUPAC, Pure and Applied Chemistry 69, 2151–2158

To clone the cDNA for an enzyme that uses this carotenoid as a precursor, a cDNA library of the appropriate tissue is constructed in the Lambda ZAP II phage vector (Stratagene). Plasmids that represent the entire mRNA repertoire are excised from the library and transfected into the *E. coli* cells that contain the recombinant pACYC184 plasmid and produce the colored carotenoid precursor. Following incubation under proper selection conditions colonies of cells that carry the two plasmids emerge on the plates. The screening for the carotenoid gene is based on color visualization of colonies of a size 3 mm in diameter. Change in color or hue of a colony can be easily detected after a couple of days. This method has been successfully used for cloning of *crtO*, a cDNA from *Haematococcus pluvialis* that encodes β -C4-oxygenase (ref. 21), and the *Arabidopsis thaliana* cDNAs for lycopene ϵ -cyclase (*CrtL-e*) (ref. 22) and β -carotene hydroxylase (ref. 23).

Another method for cloning plant genes whose product is unidentified is transposon tagging. This method was used to clone the gene for zeaxanthin epoxidase from *Nicotiana plumbaginifolia* (ref. 24). A mutant, *aba2*, with a phenotype of abscisic acid (ABA) deficiency, was generated using the maize transposon Activator. The gene was then cloned by isolating the Activator sequence from the genome of the *aba2* plant.

A future possibility for cloning genes that are involved in carotenoid biosynthesis, either encoding enzymes of the pathway or other genes that have regulatory functions, is through the technique of positional cloning. The method exploits genetic mapping of mutations that affect carotenogenesis and utilizes DNA markers in a procedure known as "chromosome walking". This strategy has been successfully employed in *Arabidopsis thaliana* for cloning various genes and recently was applied also to tomato, where many color mutants are available.

So far only two genomic sequences of carotenoid genes have been cloned and sequenced- *Psy* (ref. 15,25) and *Pds*, (ref. 26) (accession No. x71023, x78271) both from tomato.

ENZYMES AND GENES IN THE CAROTENOID BIOSYNTHESIS PATHWAY IN PLANTS

Most of the reactions in the carotenoid biosynthesis pathway have been previously proposed based on biochemical and classical genetic data. However, only the molecular characterizations of the cloned genes have revealed the existence of specific enzymes and unequivocally assigned each gene product to a specific reaction in the pathway. Thus, it was established that in plants the minimal number of enzymes that convert phytoene to β -carotene is three, in contrast to two in bacteria and fungi (Fig. 2).

Fig. 2. Comparision of the carotenoid biosynthesis pathways in various organisms. The product of gene *crtI* from *Rhodobacter* converts phytoene to neurosporene whereas in *Erwinia* it produces lycopene. The *al-1* gene product from *Neurospora* converts phytoene, following five dehydrogenation steps, to 3,4-didehydrolycopene.



© 1997 IUPAC, Pure and Applied Chemistry 69, 2151-2158

Another important conclusion of the molecular analysis is that all the enzymes that have been studied so far are single-gene products and are not heteromeric complexes. This is evident from the ability of each one of them to catalyze its specific reaction in heterologous systems, such as *E. coli*. It is also apparent that a multi-enzyme carotenogenic complex is not imperative for executing the whole pathway from phytoene to zeaxanthin or from phytoene to α -carotene. This inference is supported by the observation that various combinations of genes from phylogenically distant species, i.e. bacteria, fungi and plants, successfully function cooperatively in transgenic *E. coli* (ref. 9,13,21,27-31). The current suggestion for the carotenoid biosynthesis pathway in plants and algae is depicted in Fig. 3.



Fig. 3. Enzymes and genes in the carotenoid biosynthesis pathway in plants and algae. Genes that have been cloned are indicated in parentheses.

© 1997 IUPAC, Pure and Applied Chemistry 69, 2151–2158

PHYLOGENY AND EVOLUTION OF CAROTENOID BIOSYNTHESIS GENES

Sequence analysis of the cloned genes for phytoene synthase, phytoene desaturase and lycopene cyclase revealed a high degree of similarity between homologous genes from various species of plants, algae and cyanobacteria. Similar conservation is found amongst the bacterial and fungal genes for these enzymes. However, very little or no similarity exists between the plant-type and the bacterial-type enzymes. It is possible to postulate a phylogeny of phytoene synthases in different organisms based on their amino acid sequence. We have used the parsimony method of PHYLIP (Phylogeny Interface Package) version 3.5c (Distributed by J. Felsenstein) and the results are depicted as an unrooted phylogenic tree (Fig. 4). The small but significant similarity between the plant-type phytoene synthases (*Psy*) and the bacterial type (*crtB*) allow their assembly in a single tree (Fig. 4).



Fig. 4. Unrooted phylogenic tree of phytoene synthase (*Psy* and *crtB*) as calculated from the amino acid sequence of the proteins from various species (data from GeneBank).

In contrast, two distinct types of phytoene desaturase enzymes are found in nature - in plants, algae and cyanobacteria (*Pds*, *crtP*) and in other bacteria and fungi (*crtI*, *al-1*). The conservation in amino acid sequence is 81% among various plants and 34-60% among different bacteria. However comparison of *Pds/crtP*-type to *crtI*-type enzymes revealed less than 22% sequence conservation, most of which is in the FAD/NAD(P) binding motif found in the amino termini of the two types of enzymes. The cyanobacterial enzyme is clearly related to the plant PDS as it is 64% identical to the mature PDS of higher plants and algae. An unrooted phylogenic tree of the two types of phytoene desaturases is shown in Fig. 5. It is interesting to note that the plant-type ζ -carotene desaturase, *crtQ*, can be grouped in the *Pds* tree, suggesting that these two genes are evolutionary related. These findings support the hypothesis of a common origin of the carotenoid biosynthesis genes in green photosynthetic organisms, which is different from the one found in bacteria and fungi (ref. 11).

The *crtL-b* gene for lycopene β -cyclase was cloned from *Synechococcus* sp. PCC7942 (ref. 9,32). This gene was used to clone the cDNA of *CrtL-b* from tomato, tobacco (ref. 12), pepper (ref. 33), daffodil (Accession No. X98769) and *Arabidopsis* (Accession numbers L40176 and U50739). This enzyme catalyzes two cyclization reactions that convert lycopene to β -carotene. A comparison between the lycopene cyclases from tomato and the cyanobacterium *Synechococcus* PCC7942 revealed 35% identities and 55% similarities in their amino acid sequence. The *Erwinia* lycopene cyclase is 23% conserved to tomato. Three short regions of sequence similarity between the plant lycopene β -cyclase (*CrtL-b*) and the bacterial enzyme may define domains in the enzymes that play essential roles in substrate recognition or catalytic activity. Recently the gene for lycopene ϵ -cyclase (*CrtL-e*) was cloned from *Arabidopsis* (ref. 22) and tomato (our unpublished

data). This enzyme is highly conserved (37% similar and 59% identical) with the lycopene β -cyclase of other higher plants. Surprisingly, the enzyme capsanthin-capsorubin synthase (CCS) (ref. 34) is also highly conserved (58% similarities, 75% identities) with the *CrtL-b* product. The unrooted phylogenic tree of the various cyclases, shown in Fig. 6, illustrates the significant degree of conservation that exists in these enzymes and is indicative of the close phylogenic relationship between them. The data in Fig. 6 further suggest that *CrtL-e* probably evolved from *CrtL-b* in eukaryotic algae, and that *CCS* evolved in higher plants from one of the preexisting lycopene cyclases.



Fig. 5. Unrooted phylogenic tree of phytoene desaturases from bacteria and fungi (*crtI*-type) (left-top) and from plants and algae (*Pds*-type) (right-bottom).





© 1997 IUPAC, Pure and Applied Chemistry 69, 2151–2158

GENETIC REGULATION OF CAROTENOID BIOSYNTHESIS

Carotenoid formation is a highly regulated process. Concentration and composition of leaf xanthophylls are affected by light intensity (ref. 35) and the accumulation of specific carotenoids in chromoplasts of fruits and flowers is developmentally regulated (ref. 36,37).

Carotenoid accumulation during fruit ripening in tomato serves as a model system to elucidate the regulation of the process. In tomato, carotenoid concentration increases between 10 to 15-fold during fruit ripening. This change is due mainly to a 500-fold increase in the concentration of lycopene (ref. 38). Accumulation of lycopene begins at the "breaker" stage of fruit ripening after the fruit has reached the "mature green" stage. It has been shown that the mRNA levels of the genes *Psy* and *Pds* increase significantly during the breaker stage (ref. 11,38,39). In contrast, the mRNA of *CrtL-b* and *CrtL-e* decrease at this stage (ref. 12) and unpublished results (Fig. 7). A similar increase in mRNA was found in the genes *GGPPS* (for GGPP synthase), *Psy* and *Pds* during fruit ripening of bell pepper (ref. 5,40).



Fig. 7. Levels of mRNA in tomato fruits during ripening (relative units). Ripening stages: G, green; MG, mature green; B, breaker; O, orange; P, pink; R, red (ref. 11,12,38,39 and unpublished data).

Evidence for transcriptional regulation of carotenoid genes in flowers has been described. The steady-state levels of mRNA for Psy, Pds and CrtL-b increase dramatically in the petals of tomato flowers (ref. 12,39,41 and unpublished data). The changes in the steady-state levels of mRNA of Pds have been attributed to transcriptional regulation (ref. 41).

In the alga *Haematococcus pluvialis* accumulation of astaxanthin is induced by various environmental stresses, including high light, high salinity and nutrient starvation. We have recently observed an increase in the mRNA levels of the genes *Pds* and *crtO* following exposure of the algal cells to an environmental stress (ref. T. Lotan and J. Hirschberg, submitted for publication).

All the above data indicate that control of gene expression, most likely at the transcriptional level is the key regulatory mechanism that controls carotenogenesis *in vivo*. However, very little is known about the molecular mechanisms that are involved.

APPLICATION OF CAROTENOID GENES IN BIOTECHNOLOGY

As mentioned above, the carotenoid biosynthesis genes (or cDNA) are functional when properly expressed in bacteria. Therefore, it is likely that the carotenoid pathway in plants can be manipulated by gene-transfer technology. Constitutive expression of pTOM5, the tomato cDNA for phytoene synthase, in transgenic tobacco has led to dwarfism due to redirecting metabolites from the gibberellin pathway (ref. 42). Transformation of tobacco with the *Erwinia crt1* gene for phytoene desaturase conferred tolerance to the herbicide norflurazon and caused a small increase of β -carotene and minor changes in the composition of xanthophylls in the leaves (ref. 43). Recently, the daffodil cDNA for phytoene synthase, *Psy*, was transferred to rice under the regulation of an endosperm-specific promoter in order to achieve carotenoid biosynthesis in this tissue (Burkhardt et al., presented in the 11th International Symposium on Carotenoids, Leiden, 1996). As a result phytoene accumulated in the seeds of the transgenic plants. We have expressed the gene *crtO*, encoding β -C-4-oxygenase which converts β -carotene to canthaxanthin, from the green alga *Haematococcus pluvialis* in the cyanobacterium *Synechococcus* PCC7942. The genetically engineered cyanobacteria produced astaxanthin as well as other ketocarotenoids (Harker and Hirschberg, submitted for publication). The results confirm that *crtO* can function in cyanobacteria in conjunction with the intrinsic β -carotene hydroxylase to produce astaxanthin. Transformation of the same cDNA to tobacco plants induced accumulation of a high concentration of ketocarotenoids, including astaxanthin, in the chromoplasts of the nectary tissue (Mann et al. submitted for publication). However, only low concentrations of ketocarotenoids were detected in the leaves of these plants. These results provide the first evidence for the future possibilities of genetic engineering of the carotenoid biosynthesis pathway towards the production of novel carotenoids such as astaxanthin in plants. It is obvious that this possibility will have a major impact on future crop breeding and on the commercial production of carotenoids.

Acknowledgement: This research was supported by Research Grant Award No. IS-2192-92R from BARD - The United States-Israel Binational Agricultural Research and Development Fund.

REFERENCES

- 1. G. Sandmann, Eur. J. Biochem. 223, 7 (1994).
- 2. G. A. Armstrong, J. Bacteriol. 176, 4795 (1994).
- 3. G. A. Armstrong, J. E. Hearst, FASEB J. 10, 228 (1996).
- 4. G. E. Bartley, P. A. Scolnik, G. Giuliano, Annu. Rev. Plant Physiol 45, 287 (1994).
- 5. M. Kuntz, S. Römer, C. Suire, P. Hugueney, J. H. Weil, R. Schantz, B. Camara, Plant J. 2, 25 (1992).
- 6. R. C. Bugos, H. Y. Yamamoto, Proc. Natl. Acad. Sci. USA 93, 6320 (1996).
- 7. D. Chamovitz, I. Pecker, G. Sandmann, P. Böger, J. Hirschberg, Z. Naturforsch. 45c, 482 (1990).
- 8. D. Chamovitz, I. Pecker, J. Hirschberg, Plant Mol. Biol. 16, 967 (1991).
- 9. F. X. Cunningham, Jr., D. Chamovitz, N. Misawa, E. Gantt, J. Hirschberg, FEBS Lett. 328, 130 (1993).
- G. E. Bartley, P. V. Viitanen, I. Pecker, D. Chamovitz, J. Hirschberg, P. A. Scolnik, Proc. Natl. Acad. Sci. USA 88, 6532 (1991).
- 11. I. Pecker, D. Chamovitz, H. Linden, G. Sandmann, J. Hirschberg, Proc. Natl. Acad. Sci. USA 89, 4962 (1992).
- 12. I. Pecker, R. Gabbay, F. X. Cunningham, J. Hirschberg, Plant Mol. Biol. 30, 807 (1996).
- 13. D. Chamovitz, N. Misawa, G. Sandmann, J. Hirschberg, FEBS Lett. 296, 305 (1992).
- 14. J. A. Ray, C. R. Bird, M. Maunders, D. Grierson, W. Schuch, Nucleic Acids Res. 15, 10587 (1987).
- J. A. Ray, P. Moureau, A. S. Bird, D. Grierson, M. Maunders, M. Truesdale, P. M. Bramley, W. Schuch, *Plant Mol. Biol.* 19, 401 (1992).
- 16. P. Bramley, C. Teulieres, I. Blain, C. Bird, W. Schuch, Plant J. 2, 343 (1992).
- 17. G. E. Bartley, P. V. Viitanen, K. O. Bacot, P. A. Scolnik, J. Biol. Chem. 267, 5036 (1992).
- 18. G. E. Bartley, P. A. Scolnik, J. Biol. Chem. 268, 25718 (1993).
- 19. M. Albrecht, A. Klein, P. Hugueney, G. Sandmann, M. Kuntz, FEBS Lett. 372, 199 (1995).
- 20. H. Linden, A. Vioque, G. Sandmann, FEMS Microbiol. Lett. 106, 99 (1993).
- 21. T. Lotan, J. Hirschberg, FEBS Lett. 364, 125 (1995).
- 22. F. X. Cunningham, B. Pogson, Z. R. Sun, K. A. Mcdonald, D. Dellapenna, E. Gantt, Plant Cell 8, 1613 (1996).
- 23. Z. R. Sun, E. Gantt, F. X. Cunningham, J. Biol. Chem. 271, 24349 (1996).
- 24. E. Marin, L. Nussaume, A. Quesada, M. Gonneau, B. Sotta, P. Hugueney, A. Frey, A. Marionpoll, *EMBO J.* 15, 2331 (1996).
- 25. R. G. Fray, D. Grierson, Plant Mol. Biol. 22, 589 (1993).
- 26. V. Mann, I. Pecker, J. Hirschberg, Plant Mol. Biol. 24, 429 (1994).
- 27. H. Linden, N. Misawa, D. Chamovitz, I. Pecker, J. Hirschberg, G. Sandmann, Z. Naturforsch. 46c, 1045 (1991).
- 28. I. Martinez-Ferez, B. Fernandez-Gonzalez, G. Sandmann, A. Vioque, Biochim. Biophys. Acta 1218, 145 (1994).
- 29. N. Misawa, M. R. Truesdale, G. Sandmann, P. D. Fraser, C. Bird, W. Schuch, P. M. Bramley, J. Biochem. Tokyo. 116, 980 (1994).
- 30. S. Kajiwara, T. Kakizono, T. Saito, K. Kondo, T. Ohtani, N. Nishio, S. Nagai, N. Misawa, Plant Mol. Biol. 29, 343 (1995).
- 31. A. Raisig, G. Bartley, P. Scolnik, G. Sandmann, J. Biochem. Tokyo. 119, 559 (1996).
- 32. F. X. Cunningham, Jr., Z. R. Sun, D. Chamovitz, J. Hirschberg, E. Gantt, Plant Cell 6, 1107 (1994).
- 33. P. Hugueney, A. Badillo, H. C. Chen, A. Klein, J. Hirschberg, B. Camara, M. Kuntz, Plant J. 8, 417 (1995).
- 34. F. Bouvier, P. Hugueney, A. Dharlingue, M. Kuntz, B. Camara, Plant J. 6, 45 (1994).
- 35. A. V. Ruban, A. J. Young, A. A. Pascal, P. Horton, Plant Physiol. 104, 227 (1994).
- 36. R. G. Fray, D. Grierson, Trends Genet. 9, 438 (1993).
- 37. G. Gillaspy, H. Ben-David, W. Gruissem, Plant Cell 5, 1439 (1993).
- 38. P. D. Fraser, M. R. Truesdale, C. R. Bird, W. Schuch, P. M. Bramley, Plant Physiol. 105, 405 (1994).
- 39. G. Giuliano, G. E. Bartley, P. Scolnik, Plant Cell 5, 379 (1993).
- 40. S. Römer, P. Hugueney, F. Bouvier, B. Camara, M. Kuntz, Biochem. Biophys. Res. Commun. 196, 1414 (1993).
- 41. V. Corona, B. Aracri, G. Kosturkova, G. E. Bartley, L. Pitto, L. Giorgetti, P. A. Scolnik, G. Giuliano, Plant J. 9, 505 (1996).
- 42. R. G. Fray, A. Wallace, P. D. Fraser, D. Valero, P. Hedden, P. M. Bramley, D. Grierson, Plant J. 8, 693 (1995).
- 43. N. Misawa, S. Yamano, H. Linden, M. R. Defelipe, M. Lucas, H. Ikenaga, G. Sandmann, Plant J. 4, 833 (1993).