Cloning and Functional Analysis of the \(\beta\)-Carotene Hydroxylase of Arabidopsis thaliana*

(Received for publication, July 15, 1996, and in revised form, August 12, 1996)

Zairen Sun†, Elisabeth Gantt‡,§, and Francis X. Cunningham, Jr.*†§

From the ‡Department of Plant Biology, University of Maryland and the §Maryland Agricultural Experiment Station, College Park, Maryland 20742

An Arabidopsis thaliana cDNA encoding the enzyme \(\beta\)-carotene hydroxylase was identified by functional complementation in Escherichia coli. The product of this cDNA adds hydroxyl groups to both \(\beta\) rings of the symmetrical \(\beta\)-carotene (\(\beta,\beta\)-carotene) to form zeaxanthin (\(\beta,\beta\)-carotene-3,3′-diol) and converts the monomeric \(\beta\)-zeacarotene (7′,8′-dihydro-\(\beta\),\(\psi\)-carotene) to hydroxy-\(\beta\)-zeacarotene (7′,8′-dihydro-\(\beta\),\(\psi\)-carotene-3-ol). The \(\epsilon\) rings of \(\delta\)-carotene (\(\epsilon,\psi\)-carotene) and \(\alpha\)-zeacarotene (7′,8′-dihydro-\(\epsilon\),\(\psi\)-carotene) are poor substrates for the enzyme. The predicted amino acid sequence of the A. thaliana enzyme resembles the four known bacterial \(\beta\)-carotene hydroxylase enzymes (31–37% identity) but is much longer, with an N-terminal extension of more than 130 amino acids. Truncation of the cDNA to produce a polypeptide lacking the first 69 amino acids does not impair enzyme activity in E. coli. Truncation to yield a polypeptide of a length comparable with the bacterial enzymes (lacking 129 N-terminal amino acids) resulted in the accumulation of the monohydroxy intermediate \(\beta\)-cryptoxanthin (\(\beta,\beta\)-carotene-3-ol), predominantly, when \(\beta\)-carotene was provided as the substrate. It is suggested that amino acid residues 70–129 of the A. thaliana enzyme may play a role in formation of a functional homodimer.

Carotenoids are essential components of the photosynthetic apparatus in plants, algae, and cyanobacteria. These yellow, orange, and red pigments protect against photodestruction, harvest light for photosynthesis, and serve a number of other important functions. Most of the carotenoids important in photosynthetic organisms are xanthophylls or oxygenated carotenoids (1). The dihydroxy carotenoid zeaxanthin (\(\beta,\beta\)-carotene-3,3′-diole) is thought to play a central role in the nonradiative dissipation of light energy under conditions of excessive photon capture by the photosynthetic light-harvesting apparatus (2). Zeaxanthin is formed from \(\beta\)-carotene (\(\beta,\beta\)-carotene) by hydroxylation at position 3 in both rings of this symmetrical precursor (Fig. 1). Zeaxanthin, in turn, serves as the substrate for biosynthesis of many other important xanthophylls.

Bacterial genes encoding enzymes that convert \(\beta\)-carotene to zeaxanthin have been described (3–6), but a gene or cDNA encoding an enzyme with this activity has not yet been identified in any photosynthetic organism. We employed a functional complementation approach to identify a cDNA that encodes \(\beta\)-carotene hydroxylase in the higher plant Arabidopsis thaliana. The polypeptide predicted by this cDNA resembles the bacterial enzymes (31–37% amino acid identities), lending strong support to the concept of a common origin for the pathways of carotenoid biosynthesis in photosynthetic and nonphotosynthetic organisms. The A. thaliana enzyme, however, is much longer than the bacterial enzymes and has an N-terminal region that is essential for proper enzyme function and yet has no counterpart in the bacterial enzymes.

EXPERIMENTAL PROCEDURES

Plasmid Construction—A cDNA encoding an isopenteny1yl pyrophosphate (IPP) isomerase in the green alga Haematococcus pluvialis§ was excised from the library cloning vector (pBluescript SK+) with BamHI and KpnI, and cloned, in frame, in the BamHI and KpnI sites of pTrcHisA (Invitrogen). Digestion of the resulting plasmid with EcoRV and KpnI produced a fragment of ~2.0 kb containing the H. pluvialis IPP isomerase downstream of the strong bacterial trc promoter. This fragment was cloned in the blunt HindIII site of pAC-BETA (7) to produce pAC-BETA-O4. A 1.9-kb Smal-EcoRV fragment containing an A. thaliana lycopene \(e\) cyclase cDNA was excised from plasmid y2 (7) and cloned in the blunt BamHI site of pTrcHisA to give pTrc-Ate. A 2.5-kb EcoRV-KpnI fragment containing the trc promoter and the A. thaliana lycopene \(e\) cyclase was excised from pTrc-Ate and cloned in the blunt HindIII site of pAC-LYC (8) to produce pAC-DLETA. The plasmid pAC-ZEAX was constructed from pAC-EHER (8) by deletion of a 1.1-kb Sall-Sall fragment containing the Erwinia herbicola gene encoding zeaxanthin glucosyltransferase.

Mass Excision and Screening of an A. thaliana cDNA Library—A size-fractionated 0.5–1 kb cDNA library of A. thaliana in ZAPII (9) was treated to cause a mass excision and thereby produce a phagemid library (7). The phagemid library was introduced into E. coli strain XL1-Blue, amplified on agar plates, and recovered as a plasmid library. E. coli strain DH10BZIP containing the plasmid pAC-BETA-O4 was the host for library screening. Cultures were grown in Luria-Bertani (LB) medium supplemented with chloramphenicol (CAP; 50 \(\mu\)g/ml). Cells were made competent for transformation using a chemical method (10). The A. thaliana plasmid library was transfected into DH10BZIP/pAC-BETA-O4 cells, and cells were spread on LB agar plates containing CAP and ampicillin (AMP; 150 \(\mu\)g/ml) to produce a density of about 3,000 colonies/Large Petri plate (150 × 15 mm). Colonies were examined after 5-day incubation at room temperature.

Subcloning, Truncations, DNA Sequencing, and Sequence Analysis—Plasmid DNA was purified as described by DeSal et al. (11). The A. thaliana \(\beta\)-hydroxylase cDNA was sequenced completely on both strands using an automated sequencer. The sequence of this cDNA has the GenBank accession number U58919.

For N-terminal truncation, the library plasmid containing the A. thaliana \(\beta\)-carotene hydroxylase cDNA was digested with XhoI plus

* This work was supported primarily by a grant from the United States Department of Agriculture (Project MD-9302172) to F. X. Cunningham and F. X. Cunningham, Jr. Support was also provided by the Binaural Agricultural Development Fund (Project number US-2006-91R) and by the Maryland Agricultural Experiment Station. This manuscript is scientific article number 7905 and contribution number 9240 of the Maryland Agricultural Experiment Station. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 301-405-1649; Fax: 301-314-9082; E-mail: fc18@umail.umd.edu.

The common names of carotenoids are used in the text of this manuscript. The more descriptive systematic or semisystematic names (15) are given in parentheses when first used.

The abbreviations used are: IPP, isopenteny1yl pyrophosphate; AMP, ampicillin; bp, base pairs; CAP, chloramphenicol; dbEST, data base of expressed sequence tags; HPLC, high performance liquid chromatography; kb, kilobase; LB, Luria-Bertani.

† Z. Sun, F. X. Cunningham, Jr., and E. Gantt, unpublished data.
either BglII or SacI (partial digest). The resulting truncated cDNA fragments were directionally cloned, in frame, in BamHI-XhoI- or SacI-XhoI-digested pTrcHis plasmid vectors (Invitrogen). Methods for DNA and amino acid sequence analysis and prediction of transmembrane helices are described elsewhere (7).

Carotenoid Pigment Analysis and Mass Spectrometry—Cultures of E. coli DH10BZIP containing the various plasmids were grown in 50 ml of LB medium containing AMP and CAP in 250-ml flasks and were incubated at 28 °C for 36 h with gentle shaking. Cell harvest, pigment extraction, and HPLC analysis of the carotenoid pigments are described in Cunningham et al. (8).

For mass spectrometry, the carotenoids were separated on silica gel TLC plates developed with hexane:diethyl ether (1:1, by volume). The pigments were eluted from the adsorbent with acetone, the acetone was evaporated under a stream of nitrogen gas, and the pigments were stored under nitrogen at 20 °C until analysis. Electron impact mass spectral data were obtained with a mass spectrometer model VG7070E.

RESULTS AND DISCUSSION

A Visual Screen for Identification of Genes or cDNAs Encoding β-Carotene Hydroxylase—Using carotenoid biosynthetic genes from the bacterium E. herbicola, we constructed the plasmids pAC-BETA and pAC-ZEAX such that cells of E. coli that contain them accumulate β-carotene and zeaxanthin, respectively, and thereby form colonies that are yellow in color. E. coli colonies containing pAC-ZEAX could be distinguished, with some difficulty, from those containing pAC-BETA by a subtle difference in the color. The difference in colony color (and the amount of carotenoid pigment) could be enhanced by the appropriate choice of the host E. coli strain, by incubation of Petri plates at room temperature for 5 days rather than overnight at 37 °C and by introduction of a cDNA encoding the enzyme IPP isomerase. Under optimal conditions the color difference was readily apparent: E. coli colonies containing β-carotene were a deep orange-yellow in color and those containing zeaxanthin were a bright yellow color. The ability to visually distinguish colonies containing zeaxanthin from those that contain β-carotene provided a method for screening for genes or cDNAs that produce a functional β-carotene hydroxylase enzyme when introduced into E. coli.

Identification of an A. thaliana cDNA Encoding β-Carotene Hydroxylase by Functional Complementation in E. coli—An A. thaliana cDNA library was transfected into E. coli DH10BZIP containing the plasmid pAC-BETA-O4 (pAC-BETA with the addition of a cDNA encoding IPP isomerase), and the cells were spread on agar plates containing the requisite antibiotics to maintain pAC-BETA-O4 and select for library plasmids. A single bright yellow colony was observed in a background of the ~250,000 deep orange-yellow colonies that were examined.
DNA and Amino Acid Sequence Analysis of the *A. thaliana* β-Carotene Hydroxylase cDNA—The *A. thaliana* cDNA encoding the β-carotene hydroxylase enzyme is 956 base pairs (bp) in length with an open reading frame that extends from the beginning of the cDNA (starting with base 3) to a termination codon (TGA) at bases 885–887. A search of the data base of expressed sequence tags (dbEST) identified an *A. thaliana* cDNA (clone VBVPH03; GenBank accession number F13822 for the N-terminal sequence and F13851 for the C-terminal sequence) that appears to be identical to the β-carotene hydroxylase cDNA described in this work. The dbEST sequence provides an additional 100 bp of N-terminal sequence (it is shorter at the C-terminus) and suggests that the full-length polypeptide comprises with a methionine residue that begins 46 bp upstream of the first nucleotide in the sequence of the cDNA described in this work. The polypeptide predicted by the combination of the dbEST sequence and the cDNA sequence determined in this work contains 310 amino acids and has a molecular weight of 34,400 and a pI of 9.4.

An alignment of the predicted amino acid sequence of the *A. thaliana* β-hydroxylase with those of the four known bacterial β-carotene hydroxylases is displayed in Fig. 4. There are few gaps in the alignment, and a number of highly conserved regions are evident, the most notable being that labeled “Motif I.” A search of the nucleotide and protein sequence data bases for polypeptides containing the sequence HDGLVH(Q,K,R,W,F)P identified only the known bacterial β-carotene hydroxylase enzymes even when mismatches were allowed. The predicted *A. thaliana* polypeptide does not contain the sequence FLYA XX- XXXXXXXLXE (beginning at position 274 in the alignment of Fig. 4), which was suggested as a signature sequence for enzymes that act upon carotenoid β rings (14).

The identity between the predicted sequences of the *A. thaliana* and the bacterial hydroxylases ranges from 31 to 37% with more than one-fourth of the identically conserved residues being histidines. Plots of amino acid hydropathy (not shown) and other analyses (see “Experimental Procedures”) suggest three probable transmembrane helical regions in corresponding locations in the bacterial and plant enzymes. Interestingly, a fourth transmembrane region was predicted for the higher plant sequence in a region (residues 82–102, Fig. 4) that aligns upstream of where the bacterial sequences begin.

**A Portion of the *A. thaliana* N Terminus with No Bacterial Counterpart Is Important to Enzyme Function**—The β-carotene hydroxylase of *A. thaliana* is presumed to be located in the thylakoid membranes of the chloroplasts in this plant. Therefore, it is not unexpected that the sequence of the predicted

---

**Fig. 3.** HPLC elution profiles of carotenoid pigments extracted from cultures of *E. coli* containing pAC-DELTA and the *A. thaliana* cDNA encoding β-carotene hydroxylase. Left panel, *E. coli* cells contained plasmid pAC-DELTA and the empty cloning vector pBluescript SK+ (control). Right panel, *E. coli* cells contained plasmid pAC-DELTA and the *A. thaliana* cDNA encoding β-carotene hydroxylase.

**Fig. 4.** Amino acid sequence alignment of the *A. thaliana* β-carotene hydroxylase and the known bacterial β-carotene hydroxylase enzymes. A consensus sequence is given below the alignment. A capital letter in the consensus sequence indicates that the amino acid residue is identical for all five polypeptides. A lowercase letter is used where the amino acid residue for three of the four bacterial enzymes is identical to that predicted for the *A. thaliana* enzyme. Arrows above the alignment mark the sites of truncation resulting after digestion with the indicated restriction enzymes. A dot in a sequence denotes a gap in the alignment. The first 16 amino acid residues (in lowercase type) were predicted by the N-terminal sequence of an *A. thaliana* expressed sequence tag (GenBank accession number F13822). GenBank accession numbers: *A. thaliana*, U58919; *Alcaligenes sp.*, D58422; *Agrobacterium aurantiacum*, D58420; *E. herbicola* strain Eho10, M87280; *Erwinia uredovora*, D90087. TM, transmembrane.
enzyme includes an N-terminal region that is not found in the bacterial enzymes (Fig. 4). The N-terminal extension of more than 130 amino acids, however, is much longer than typically found for chloroplast targeting sequences. We truncated the A. thaliana cDNA to examine whether and how much of this N-terminal extension was essential to enzyme function.

The enzyme SacI was used to remove that portion of the cDNA encoding the first 69 amino acids of the A. thaliana β-carotene hydroxylase (Fig. 4). The product of this truncated cDNA (fused at the N terminus to a peptide encoded by the plasmid vector) efficiently converted β-carotene to zeaxanthin (92–93% of the total carotenoid) in cells of E. coli. However, the major product (75–77% of the total) was β-cryptoxanthin rather than zeaxanthin (Fig. 5, right panel). Small amounts of zeaxanthin (16–18%) and β-carotene (7%) were also present.

The high ratio of β-cryptoxanthin to β-carotene in cells containing the BglII-truncated hydroxylase (~11 to 1) indicates that the β rings of β-carotene are hydroxylated with greater efficiency than the not yet hydroxylated β ring of β-cryptoxanthin. We speculate that the A. thaliana enzyme normally associates with a second β hydroxylase (or with an ε hydroxylase) to form a dimer and that a portion of the N terminus (e.g. between residues 69 and 130) mediates these subunit interactions. Experiments to clarify the role of the N-terminal domain and determine whether the hydroxylase does, indeed, function as a dimer are now in progress.

REFERENCES
1. Goodwin, T. W. (1980) The Biochemistry of the Carotenoids 2nd Ed., Vol. 1, Chapman and Hall, London
2. Demmig-Adams, B., and Adams, W. W., III (1993) in Carotenoids in Photosynthesis (Young, A. J., and Britton, G., eds) pp. 206–252, Chapman and Hall, London
3. Perry, K. L., Simonitch, T. A., Harrison-Lavoie, K. J., and Liu, S. (1986) J. Bacteriol. 168, 607–612
4. Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K., and Harashima, K. (1990) J. Bacteriol. 172, 6704–6712
5. Misawa, N., Kajiwara, S., Kondo, K., Yokoyama, A., Satomi, Y., Saito, T., Miki, W., and Ohnishi, T. (1995) Biochem. Biophys. Res. Commun. 209, 867–876
6. Misawa, N., Satomi, Y., Kondo, K., Yokoyama, A., Kajiwara, S., Saito, T., Ohnishi, T., and Miki, W. (1995) J. Bacteriol. 177, 6575–6584
7. Cunningham, F. X., Jr., Pogson, B., Sun, Z., McDonald, K. A., DellaPenna, D., and Gantt, E. (1996) Plant Cell, in press
8. Cunningham, F. X., Jr., Sun, Z., Chamovitz, D., Hirschberg, J., and Gantt, E. (1994) Plant Cell 6, 1107–1121
9. Kieber, J. J., Rothenberg, M., Roman, G., Feldman, K. A., and Ecker, J. R. (1993) Cell 72, 427–441
10. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
11. Del Sal, G., Manfioletti, G., and Schneider, C. (1988) Nucleic Acids Res. 16, 9878
12. Pogson, B., McDonald, K., Truong, M., Britton, G., and DellaPenna, D. (1996) Plant Cell, in press
13. Britton, G. (1990) in Carotenoids: Chemistry and Biology (Krinsky, N. I., Mathews-Roth, M., and Taylor, R. F., eds) pp. 167–184, Plenum Press, New York
14. Bouvier, F., Hugueney, P., d'Harlingue, A., Kuntz, M., and Camara, B. (1994) Plant J. 6, 45–54
15. Weeds, B. C. L., and Moss, G. P. (1995) in Carotenoids, Vol IB: Spectroscopy (Britton, G., Liaaen-Jensen, S., and Pfander, H., eds) pp. 27–70, Birkhauser Verlag, Basel, Switzerland