Title

Identification of a fourth family of lycopene cyclases in photosynthetic bacteria

Journal

Proceedings of the National Academy of Sciences of the United States of America, 104(28)

ISSN

0027-8424

Authors

Maresca, Julia A
Graham, Joel E
Wu, Martin
et al.

Publication Date

2007-07-10

DOI

10.1073/pnas.0702984104

Peer reviewed

Identification of a fourth family of lycopene cyclases in photosynthetic bacteria

Julia A. Maresca*†, Joel E. Graham*, Martin Wu‡§, Jonathan A. Eisen‡§, and Donald A. Bryant*¶

*Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802; and ‡The Institute for Genomic Research (TIGR), Rockville, MD 20850

Edited by Robert Haselkorn, University of Chicago, Chicago, IL, and approved June 1, 2007 (received for review March 30, 2007)

A fourth and large family of lycopene cyclases was identified in photosynthetic prokaryotes. The first member of this family, encoded by the *cruA* gene of the green sulfur bacterium *Chlorobium tepidum*, was identified in a complementation assay with a lycopene-producing strain of *Escherichia coli*. Orthologs of *cruA* are found in all available green sulfur bacterial genomes and in all cyanobacterial genomes that lack genes encoding CrtL- or CrtY-type lycopene cyclases. The cyanobacterium *Synechococcus* sp. PCC 7002 has two homologs of CruA, denoted CruA and CruP, and both were shown to have lycopene cyclase activity. Although all characterized lycopene cyclases in plants are CrtL-type proteins, genes orthologous to *cruP* also occur in plant genomes. The CruA- and CruP-type carotenoid cyclases are members of the FixC dehydrogenase superfamily and are distantly related to CrtL- and CrtY-type lycopene cyclases. Identification of these cyclases fills a major gap in the carotenoid biosynthetic pathways of green sulfur bacteria and cyanobacteria.

carotenogenesis | carotenoids | cyanobacteria | green sulfur bacteria | photosynthesis

Colored carotenoids are found in nearly all photosynthetic species and in a variety of nonphotosynthetic organisms and play roles in light-harvesting, photoprotection, structural maintenance of pigment–protein complexes (1), and membrane structure and fluidity (2). The cyclization of the linear compound lycopene to produce α-, β-, γ-, or ε-carotene is a branch point in carotenoid biosynthetic pathways in many species of bacteria, plants, and fungi (3, 4). The monocyclic γ-carotene is a precursor to myxoxanthophylls in cyanobacteria (5, 6), is both an intermediate and an end product of carotenogenesis in orange and yellow flowers and fruits (7–10), and is an intermediate in chlorobactene biosynthesis in green sulfur bacteria (GSB) (11–13). Dicyclic β-carotene is a major component of photosystems (PS) I and II in cyanobacteria and plants (14, 15) and is modified to isorenieratene in brown-colored GSB and some actinomycetes (16, 17).

Three classes of lycopene cyclases have previously been identified in bacteria: the CrtY-type β-cyclases that are found in many carotenogenic proteobacteria (e.g., refs. 18 and 19), *Streptomyces* spp. (17), and the *Chloroflexi*; the CrtL family, which includes the β- and ε-cyclases in some cyanobacteria and plants (20, 21); and the heterodimeric cyclases of some Gram-positive bacteria (22, 23), which are related to the lycopene cyclases of archaea (24, 25) and halophilic bacteria (26). These classes are distantly related to each other and share only a few conserved motifs, including an N-terminal flavin-binding domain that is found in the first two groups but appears to be missing in the third (27). Carotenoid monocyclases are found in both the CrtY and CrtL families (28–30), and there are no obvious sequence differences between mono- and dicyclases (28).

The cyclization of lycopene to γ- or β-carotene is an isomerization reaction, which produces no net change in mass or redox state of the substrate. At least one CrtY-type lycopene cyclase requires NADPH for activity (31); however, there is no experimental evidence about the role of the flavin predicted to be a cofactor. Although the *in vivo* substrate of most carotenoid cyclases is lycopene, some cyclases can act on neurosporene as well. At least one species of GSB has an enzyme that can cyclize both the ψ- and 7,8-dihydro-ψ ends of neurosporene (ref. 32; J.A.M., S. P. Romberger, and D.A.B., unpublished data), as can the CrtL-type cyclase from the plant *Capsicum annuum* and CrtY from *Erwinia uredovora* (33, 34). The lycopene monocyclases from *Rhodococcus erythropolis* and *Deinococcus radiodurans* R1 can cyclize only the ψ-end of neurosporene (30), whereas the dicyclase CrtL from *Synechococcus* sp. PCC 7942 cyclizes both ends of lycopene but only the ψ-end of neurosporene (20).

Previous studies have shown that the carotenoid biosynthetic pathways in cyanobacteria and the GSB *Chlorobium tepidum* are very similar (12). In GSB, cyanobacteria, and plants, the enzymes CrtP and CrtQ successively desaturate phytoene to 7,9,7′,9′tetra-*cis*-lycopene (12, 35–42), and an isomerase, CrtH, converts *cis*-lycopene to *all-trans*-lycopene (3, 43–46), instead of using CrtI to to convert phytoene to neurosporene or *all-trans*-lycopene (40, 47–49). Genes homologous to known lycopene cyclases were not found in the completed genomes of *C. tepidum*, *Gloeobacter violaceus* sp. PCC 7421, *Synechocystis* sp. PCC 6803, *Synechococcus* sp. PCC 7002, *Thermosynechococcus elongatus*, and *Nostoc* sp. PCC 7120. Because all of these species synthesize carotenoids with cyclic end groups and because their carotenogenic pathways are so similar, it seemed likely that the lycopene cyclase in GSB and that in the above-mentioned cyanobacteria would both be members of a new family of lycopene cyclases. In this study we describe the identification of a new family of lycopene cyclases found in GSB, cyanobacteria, and plants and the characterization of these enzymes from the GSB *C. tepidum* and the cyanobacterium *Synechococcus* sp. PCC 7002. A preliminary description of some results for *C. tepidum* was presented at the XIIIth International Congress on Photosynthesis (13).

Results

Identification of Lycopene Cyclase in *C. tepidum*. *Escherichia coli* strain DH5α harboring plasmid pACLYC forms pink colonies

Author contributions: J.A.M. and J.E.G. contributed equally to this work; J.A.M. and J.E.G. designed research; J.A.M., J.E.G., and D.A.B. performed research; M.W. and J.A.E. contributed new reagents/analytic tools; J.A.M., J.E.G., M.W., and J.A.E. analyzed data; and J.A.M., J.E.G., and D.A.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Abbreviations: CHP, conserved hypothetical protein; GSB, green sulfur bacteria.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. EF529626 and EF529627).

†Present address: Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

§Present address: Section of Evolution and Ecology, Department of Medical Microbiology and Immunology, University of California Davis Genome Center, University of California, Davis, CA 95616.

¶To whom correspondence should be addressed at: Department of Biochemistry and Molecular Biology, Pennsylvania State University, S-235 Frear Building, University Park, PA 16802. E-mail: dab14@psu.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0702984104/DC1.

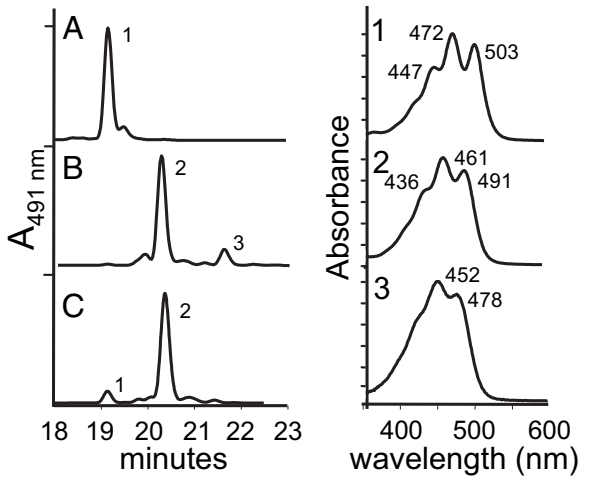

Fig. 1. Pigments produced in *E. coli* lycopene complementation strains. The absorption spectra of numbered peaks are shown on the right. Pigments extracted from *E. coli* BL21 (DE3) cells harboring plasmids pAC-LYC and pET16b, the empty vector (*A*); plasmids pAC-LYC and pCTLY, which encodes *C. tepidum cruA* (*B*); and plasmids pAC-LYC and p3-SLP, which encodes *Synechococcus* sp. PCC 7002 *cruP* (*C*). Peak 1 is lycopene; peak 2 is γ-carotene; and peak 3 is β-carotene.

containing lycopene (Fig. 1*A*). This strain was transformed with a genomic library from *C. tepidum*. Any cells expressing the *C. tepidum* lycopene cyclase should convert lycopene into γ-carotene and form orange colonies. One orange transformant was obtained, and the orange cells produced γ-carotene and small amounts of β-carotene and torulene (3′,4′-didehydro-γ-carotene) (Fig. 1*B*). The library plasmid from the orange colony, pCTLY, had a 2.7-kb insert, which encoded one partial gene (DNA ligase, ORF *CT0457*) and one complete gene, *CT0456*, annotated as a conserved hypothetical protein (CHP) and now designated *cruA*.

Phylogenetic Analyses of Lycopene Cyclases. Phylogenetic profiling analysis was performed to compare the *C. tepidum* genome to the genomes of other completely sequenced bacteria [supporting information (SI) Fig. 6]. Three *C. tepidum* genes (*CT0456*, *CT1196*, and *CT1748*), encoding CHPs had the correct phylogenetic distribution, with homologs in the genomes of *Nostoc* sp. PCC 7120, *T. elongatus*, *G. violaceus*, and *Synechocystis* sp. PCC 6803, but without homologs in the genomes of *Prochlorococcus* spp., *Synechococcus* sp. WH8501, and *Chloroflexus aurantiacus* (SI Fig. 6). Only *CT0456* was also predicted to have an FAD-binding motif. Thus, phylogenetic profiling correctly identified the targeted gene.

A blastp search (50) using the deduced amino acid sequence of the *C. tepidum* CruA identified only two groups of homologous proteins in the National Center for Biotechnology Information (NCBI) database (www.ncbi.nlm.nih.gov/BLAST), all of which are CHPs in the FixC superfamily of dehydrogenases. Genes encoding orthologous proteins are found in all available GSB genomes and in all cyanobacterial genomes that lack *crtL* or *crtY* homologs. Paralogs of CruA, denoted CruP, are found in the same cyanobacterial strains and in plants. Cyanobacterial CruA differs from CruA of GSB by having a large N-terminal extension. Cyanobacteria that have both CruA and CruP do not have a CrtL-type lycopene cyclase. *Synechococcus* sp. strains PCC 7942 and PCC 6301, which have CrtL and CruP, are the only cyanobacteria that have members of both the CrtL and CruA families. The CruA family forms a clade distinct from the CrtY-, CrtL-, and heterodimeric-type lycopene cyclase families (Fig. 2).

Inactivation of Lycopene Cyclases. After transformation with linearized pCTLY::*aadA* (Fig. 3*A*) and one passage on solid selective medium, homozygous mutants of *CT0456*::*aadA* (*cruA*::*aadA*) were isolated (Fig. 3*D*). Under standard growth conditions, wild-type *C. tepidum* and the *cruA* mutant grow at the same rate. Wild-type *C. tepidum* produces mainly chlorobactene, some γ-carotene, and small amounts of glycosylated and acylated derivatives of these two compounds (refs. 11, 12, and 51; Fig. 4*A*). The *cruA*::*aadA* mutant produced mainly lycopene (Fig. 4*A*, peak 5) as well as smaller amounts of lycopene precursors and acyl-glucoside esters of hydroxylycopene (data not shown). These data demonstrate that the *cruA*::*aadA* mutant does not produce carotenoids with cyclic end groups.

Synechococcus sp. PCC 7002 strains transformed with *cruA*::*aacC1* and *cruP*::*aphA-2* constructs (Fig. 3 *B* and *C*) were grown on plates supplemented with the appropriate antibiotics. Homozygous *cruP* mutants were obtained after several passages on selective medium (Fig. 3*E*). The *cruA*::*aacC1* mutant only grew on solid media supplemented with glycerol; segregation

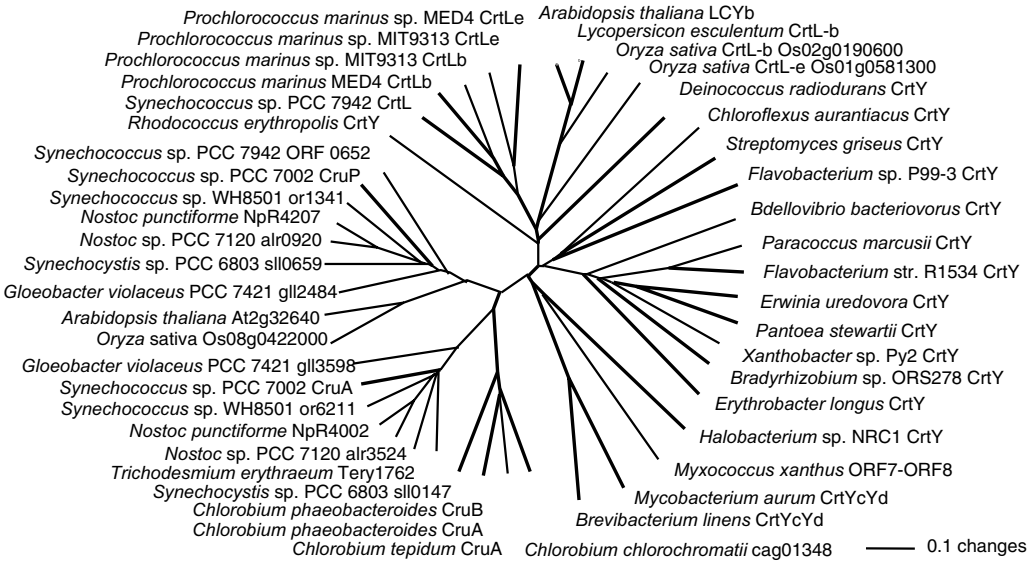

Fig. 2. Neighbor-joining phylogenetic tree of four classes of lycopene cyclases. Bold lines indicate genes or gene products that have been either genetically or biochemically characterized.

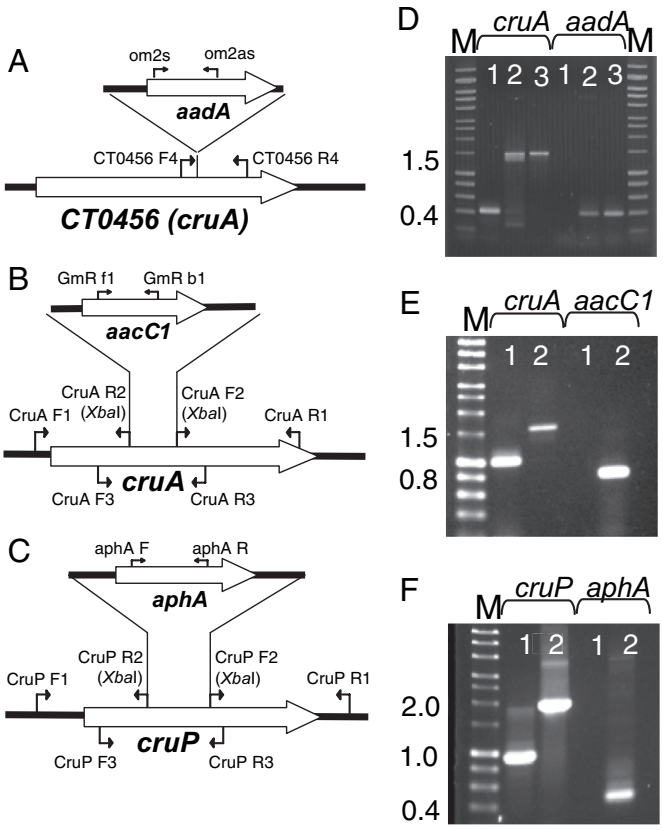

Fig. 3. Insertional inactivation of *cruA* and *cruP*. Constructs for insertional inactivation of *cruA* in *C. tepidum* (*A*) *cruA* in *Synechococcus* sp. PCC 7002 (*B*) *cruP* in *Synechococcus* sp. PCC 7002 (*C*). All three constructs produced homozygous mutants, as shown by PCR across the insertions. The locus amplified is indicated above the figure. (*D*) Lanes 1, wild-type *C. tepidum*; lanes 2 and 3, two isolates of the *C. tepidum cruA*::*aadA* mutant. (*E*) Lanes 1, wild-type *Synechococcus* sp. PCC 7002; lanes 2, *Synechococcus cruA*::*aacC1* mutant. (*F*) Lanes 1, wild-type *Synechococcus* sp. PCC 7002; lanes 2, *Synechococcus cruP*::*aphA* mutant. M, DNA size markers. The numbers at the left are sizes in kilobases.

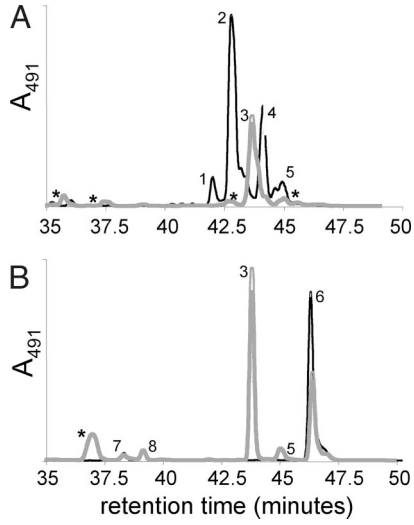

Fig. 4. HPLC elution profiles, monitored at 491 nm, of carotenoids extracted from wild-type (black lines) and *cruA* mutant strains (gray lines). (*A*) Carotenoids extracted from *C. tepidum* strains and normalized to BChl *c* content. Peak 1, OH-chlorobactene glucoside laurate; peak 2, chlorobactene; peak 3, lycopene; peak 4, 1′,2′-dihydrochlorobactene; peak 5, γ-carotene. (*B*) Carotenoids from *Synechococcus* sp. PCC 7002 strains, normalized to Chl *a* content. Peak 3, lycopene; peak 5, γ-carotene; peak 6, β-carotene; peak 7, cryptoxanthin; peak 8, Chl *a*. *, unidentified carotenoids with lycopene chromophores.

was achieved after many rounds of growth at reduced light intensity and after transfer to liquid medium that was also supplemented with glycerol (Fig. 3*F*). Although the strain in which *cruP* was inactivated was phenotypically identical to the wild type, the *cruA* mutant showed a severe growth defect. The *cruA* locus did not segregate completely in strains in which both *cruA* and *cruP* were insertionally inactivated. Wild-type *Synechococcus* sp. PCC 7002 produces large amounts of the dicyclic carotenoids β-carotene and zeaxanthin as well as the monocyclic myxoxanthophyll but does not accumulate lycopene or γ-carotene (Fig. 4*B*). The *cruA* mutant strain accumulates significant amounts of lycopene and γ-carotene (Fig. 4*B*).

Expression of Cyanobacterial *cruA* and *cruP* in *E. coli*. The *E. coli* strain BL21(DE3) pACLYC was transformed with expression plasmids encoding *cruA* or *cruP* or with the corresponding empty vector, pET3aTr (52). The plasmids p3-SLA and p3-SLP encode the *Synechococcus* sp. PCC 7002 *cruA* and *cruP* genes, respectively. Cells harboring pAC-LYC and pET3aTr or p3-SLA produced only lycopene (data not shown). However, cells with pAC-LYC and p3-SLP produced primarily γ-carotene (Fig. 1*C*). Because the CruP homologs are the most divergent sequences from the GSB CruA family, this result implies that all members of the CruA/CruP family in bacteria are probably carotenoid cyclases.

Discussion

We have identified a fourth family of enzymes, CruA and CruP, which cyclize lycopene in photosynthetic prokaryotes (Fig. 5). The characterization of the activities of three members of the CruA family enables the assignment of a function to a large family of CHPs, which are found in all available GSB genomes, in some cyanobacterial genomes, in plants, and in *Herpetosiphon aurantiacus*, a nonphotosynthetic organism in the phylum *Chloroflexi*. These proteins are part of the FixC superfamily of dehydrogenases and, like the CrtL- and CrtY-type lycopene cyclases, are predicted to be integral membrane proteins with an N-terminal flavin-binding domain.

Orthologs of *cruA* are found in all cyanobacterial genomes that lack CrtL-type lycopene cyclases and a paralogous gene, *cruP*, is found in the same species as well as in *Synechococcus* spp. PCC 7942 and 6301. The characterization of CruA and CruP in *Synechococcus* sp. PCC 7002 resolves a question that had existed since the completion of the genome of *Synechocystis* sp. PCC 6803, because it did not contain homologs of known lycopene cyclases (53). With the identification of the lycopene cyclases in

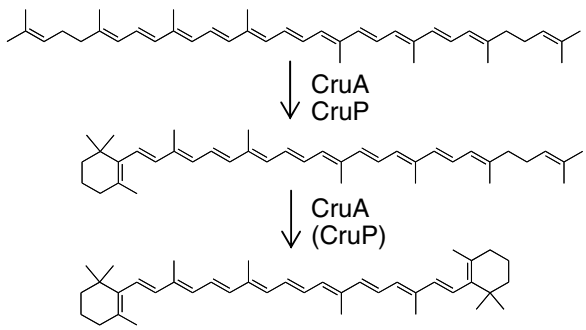

Fig. 5. Reaction(s) catalyzed by CruA and CruP. CruA and CruP from *Synechococcus* sp. PCC 7002 have both activities. However, the overall activity of CruP is lower in *Synechococcus* sp. PCC 7002, and CruP may also have less activity in converting γ-carotene to β-carotene.

these species, all gene products necessary for the production of β-carotene in cyanobacteria have now been identified (SI Fig. 7).

In addition, with the identification of CruA as the sole lycopene cyclase in *C. tepidum*, all genes encoding enzymes required for the synthesis of chlorobactene in *C. tepidum* have also been identified (SI Fig. 8; refs. 12, 13, and 51). The other genes in the chlorobactene biosynthetic pathway were identified on the basis of their similarity to genes involved in carotenoid biosynthesis in other organisms, and their functions were confirmed by targeted gene inactivation (12, 51). However, this approach failed to identify the lycopene cyclase in *C. tepidum* (12), and, consequently, a complementation assay was used to identify the gene encoding the lycopene cyclase in GSB. The ORF *CT0456* encoding the lycopene cyclase is now named *cruA*.

The distribution of CruA is unique, because it appears almost exclusively in genomes of organisms that synthesize lycopene from phytoene via the CrtP–CrtQ–CrtH pathway and have type I photosynthetic reaction centers. *H. aurantiacus* is the only nonphototroph whose genome encodes *cruA*. Because horizontal transfer of carotenoid biosynthetic genes appears to occur frequently between unrelated microorganisms (54, 55), this distribution leads one to question why *cruA*, as well as *crtP*, *crtQ*, and *crtH*, have not been transferred more broadly. The genes encoding phytoene synthases, CrtI-type carotenoid desaturases, and carotenoid glycosyltransferases are found not only in photosynthetic bacteria but also in *Deinococcus radiodurans*, the planctomycete *Rhodopirellula baltica*, actinomycetes, and archaea (17, 24, 25, 51). The restricted distribution of CruA might indicate that some CruA cyclases require another protein subunit or a cofactor that is not commonly synthesized by many bacteria for optimal activity. Alternatively, although CruA can use the *all-trans* lycopene made in *E. coli* as a substrate, its preferred substrate *in vivo* may be an intermediate specifically produced by the CrtP–CrtQ–CrtH pathway. The fact that *crtP*, *crtQ*, *crtH*, and *cruA* are not clustered in any of the genomes in which they are found (56, 57) may have resulted in their limited distribution: it is possible that transfer of only one of the genes does not provide a competitive advantage to the organism that receives it.

Both *C. tepidum* CruA and *Synechococcus* sp. PCC 7002 CruP are active in an *E. coli* complementation assay, and both are primarily monocyclases in this assay system. The failure to observe any cyclase activity when the *Synechococcus* sp. PCC 7002 *cruA* gene was inserted into *E. coli* is similar to results obtained with the orthologous gene from *G. violaceus*, which failed to cyclize lycopene in a similar assay (58). This finding may reflect a requirement for another protein subunit, a specific cofactor not produced in *E. coli*, or processing of the N-terminal extension that makes the cyanobacterial CruA homologs unique within this family. Because inactivation of *cruA* in *Synechococcus* sp. PCC 7002 resulted in reduced levels of cyclic carotenoids and the accumulation of lycopene and γ-carotene, and because a mutant lacking both CruA and CruP is apparently inviable, *cruA* most likely encodes a lycopene cyclase. The presence of active CruP in the *cruA* mutant explains the continued production, albeit at a greatly reduced rate, of both β-carotene and myxoxanthophyll. Cyanobacteria may need two cyclases to regulate production of mono- and dicyclic carotenoids independently. Myxoxanthophyll has been shown to accumulate in cyanobacteria in response to UV radiation while β-carotene levels remain constant (59, 60). Independent regulation of these genes may also explain the phenotypes of the *cruP* and *cruA* mutants. The major cyclase, CruA, can easily complement the absence of CruP. However, the accessory cyclase, CruP, which apparently has lower activity or expression, is unable to complement fully the absence of CruA. Because both lycopene and γ-carotene accumulate in the *cruA* mutant, it seems likely that CruP has lower overall activity than CruA rather than specifically reduced dicyclase activity.

All CrtL-, CrtY-, and CruA/CruP-type lycopene cyclases are all predicted to be integral membrane proteins with flavin-binding domains. Given that there is no net redox change during the cyclization of lycopene, there is no obvious role for the flavin. Some flavoproteins can catalyze reactions with no net change in redox state (see ref. 61 for examples), and perhaps this reaction requires this type of catalysis; however, CruA does not resemble any of these proteins in sequence. It is possible that its catalytic function is most similar to that of the type II isopentenyl diphosphate (IPP)/dimethylallyl diphosphate (DMAPP) isomerase, which converts IPP to DMAPP by isomerizing a double bond (62, 63). Both type I and type II IPP/DMAPP isomerases appear to catalyze isomerization by stereoselective protonation and deprotonation of the substrate (64, 65), and the type II IPP isomerase, like CrtY, requires a reduced flavin and NADPH for activity (62, 63, 64). In this case, although there is no net oxidation or reduction, both reduction and oxidation of the substrate occur during the reaction, and the flavin may have a role in these half-reactions.

A recent study suggested that an ortholog of the γ- and β-carotene desaturase/methyltransferase CrtU, Sll0254, in *Synechocystis* sp. PCC 6803 is a bifunctional lycopene cyclase/dioxygenase in that species and in the other myxoxanthophyll-producing species lacking CrtL- or CrtY-type lycopene cyclases (66). However, the authors of that study failed to construct a mutant that accumulated lycopene and did not demonstrate accumulation of either β- or γ-carotene in an *E. coli* strain that produced lycopene. In our hands, complete inactivation of *sll0254* did not affect either β-carotene or myxoxanthophyll biosynthesis (SI Fig. 9). Inactivation of *cruA* in *Synechococcus* sp. PCC 7002 reduced the levels of both of these carotenoids and resulted in the accumulation of lycopene and γ-carotene (Fig. 5*B*). Finally, although all cyanobacteria synthesize β-carotene, orthologs of *sll0254* are not found in all cyanobacterial genomes lacking CrtL- or CrtY- homologs (ref. 56; SI Fig. 10), but orthologs of both *cruA* and *cruP* are found in all of those genomes.

Genes apparently orthologous to *cruP* are also found in several plant genomes, including *Oryza sativa* and *Arabidopsis thaliana*. CruP from *Synechococcus* sp. PCC 7002 is more similar to its homolog in rice (Fig. 3; Os08g0422000; 44% identical and 62% similar) than it is to CruA of the same cyanobacterium (33% identical and 51% similar). The other predicted lycopene cyclases in rice, CrtL-b and CrtL-e (Os02g0190600 and Os01g581300, respectively) are only ≈34% identical and 50% similar to the functionally characterized CrtL from *Synechococcus* sp. PCC 7942 (20, 67). Although the plant CruP orthologs have not been characterized biochemically, their genes are expressed in plant tissues that accumulate cyclic carotenoids. An analysis of the tissue distribution of the *O. sativa* CruP homolog by ESTs reveals transcripts for this protein in leaves, stems, and inflorescences. Both the numbers of ESTs and their tissue distribution are similar to those for CrtL-e, the predicted rice ε-cyclase. However, the ESTs for CrtL-b, the predicted β-cyclase, are less numerous and are mostly reported for stems, inflorescences and flowers, with only a single EST being reported from leaves. These data suggest that, in rice, the CruP homolog could be the principal lycopene β-cyclase in photosynthetic tissues and that the cyanobacterial ancestor of chloroplasts may have had both types of cyclases, as *Synechococcus* spp. PCC 7942 and PCC 6301 do today. Considering both the high sequence similarity of cyanobacterial CruP to homologous proteins in plants and the expression profile of the rice CruP homolog, we find it reasonable to speculate that the plant homologs of CruP are also lycopene cyclases. Should this prove to be the case, there will be some fascinating work to be done in comparing the function and distribution of these two classes of lycopene cyclases in plants.

Materials and Methods

Bacterial Strains and Growth Conditions. Wild-type *C. tepidum* strain is plating strain WT2321 (68), derived from strain ATCC 49652 (69). All *Chlorobium* strains were grown anaerobically in liquid CL medium or on solid CP medium as described in ref. 70.

Cells were grown on CP at 40°C in a Coy anaerobic chamber (Coy, Grass Lake, MI) with an atmosphere of 10% CO_2, 10% H_2, and 80% N_2. *C. tepidum* transformants in which the *cruA* gene was interrupted with the *aadA* cassette were selected on CP plates supplemented with 200 $\mu g \cdot ml^{-1}$ streptomycin and 400 $\mu g \cdot ml^{-1}$ spectinomycin.

Wild type and mutant strains of *Synechococcus* sp. PCC 7002 were grown in A^+ media supplemented with 1 $mg \cdot ml^{-1}$ $NaNO_3$ (71). For growth of the mutant strains, the antibiotics kanamycin (100 $\mu g \cdot ml^{-1}$) and/or gentamicin (50 $\mu g \cdot ml^{-1}$) were added to the growth media as required. Liquid cultures were grown in 30-ml tubes at 38°C and were bubbled with air supplemented with 1% (vol/vol) CO_2. Solid medium was prepared by addition of 3 g of $Na_2S_2O_3$ and 15 g of Bacto-agar per liter of A^+ medium. Due to a lethal sensitivity of the *cruA* mutant to Bacto-agar, plates for the *cruA* mutant were prepared by addition of 7.5 g of molecular biology grade agarose. Where applicable, solid or liquid media were supplemented with 20 mM glycerol. Fluorescent light at an intensity of 250 $\mu mol \cdot m^{-2} \cdot s^{-1}$ photons or reduced light intensity of 30 $\mu mol \cdot m^{-2} \cdot s^{-1}$ photons was continuously provided for cell growth.

Routine recombinant DNA procedures were carried out in electrocompetent *E. coli* DH10B or DH5α. *E. coli* grown in LB or M9 supplemented with 25 $\mu g \cdot ml^{-1}$ chloramphenicol, 100 $\mu g \cdot ml^{-1}$ ampicillin, or 100 $\mu g \cdot ml^{-1}$ spectinomycin and 50 $\mu g \cdot ml^{-1}$ streptomycin as required.

Preparation of Genomic Library and Identification of *cruA*. *C. tepidum* genomic DNA was digested with the enzyme *Sau*3AI (1.25×10^{-2} unit$\cdot \mu g^{-1}$ DNA). Fragments of 2.5–5 kb were ligated into pUC19 that had been digested with *Bam*HI. The library was introduced by electroporation into *E. coli* strain DH10B with the pAC-LYC plasmid. Transformants were plated on LB plates containing 50 $\mu g \cdot ml^{-1}$ chloramphenicol and 100 $\mu g \cdot ml^{-1}$ ampicillin and visually screened for development of orange coloration. The plasmids from the single orange transformant were extracted, and the ends of the cloned *C. tepidum* DNA insert were sequenced using the M13U and M13R primers at the Nucleic Acid Facility at Pennsylvania State University.

Inactivation of *CT0456*. The *aadA* cassette, conferring resistance to streptomycin and spectinomycin, was excised from plasmid pSRA2 (70) using *Hin*cII and inserted into the *Msc*I site of plasmid pCTLY. The resulting construct, pCTLY::*aadA*, was linearized with the enzyme ScaI and purified with the PerfectPrep Gel Cleanup Kit. Wild-type *C. tepidum* cells (100 μl of a late exponential phase culture) were transformed with the linearized plasmid as described in ref. 70. Single colonies were picked and restreaked on selective plates after 1 week. To confirm segregation of alleles, the *CT0456* and *CT0456*::*aadA* alleles were PCR-amplified using the primers CT0456 F4 and CT0456 R4 (SI Table 1 and Fig. 3 *A* and *D*). The *aadA* gene was amplified from transformed strains by using the primers om2-as and om2-s (SI Table 1 and Fig. 3 *A* and *D*).

Inactivation of *cruA* and *cruP* in *Synechococcus* sp. PCC 7002. Left and right flanks of *cruA* and *cruP* were amplified using the primers described in SI Table 1. PCR products were gel-purified using the Eppendorf (Westbury, NY) PerfectPrep Gel Cleanup kit and digested with the enzyme XbaI. Plasmids pMS266, with the *aacC1* gene encoding resistance to gentamicin, and pRL161, with the *aphA-2* cassette encoding resistance to kanamycin, were digested with *Xba*I to excise the antibiotic resistance cartridges. These cartridges were then ligated to the left and right flanks of *cruA* or *cruP* (Fig. 3 *C* and *D*). The ligation products were separated on an agarose gel and the band corresponding to the correct construct was purified using the Eppendorf PerfectPrep Gel Cleanup kit. This product was used to transform *Synechococcus* sp. PCC 7002 as described in ref. 72. Segregation was confirmed by PCR amplification of the *cruA* or *cruP* loci by using primers CruA F3/R3 or CruP F3/R3.

Cloning of *cruA*, *cruB*, and *cruP* Expression Constructs. The *cruA* and *cruP* genes were amplified by PCR from *Synechococcus* sp. PCC 7002 genomic DNA by using the primer pairs CruAX-F/CruAX-R for *cruA* and CruPX-F/CruPX-R for *cruP* (SI Table 1). The PCR products were subcloned into plasmid pCR4-TOPO by using the TOPO-TA cloning kit (catalog no. K4500-01; Invitrogen, Carlsbad, CA) and then excised from the plasmid by using the restriction sites in the primers. The fragments containing *cruA* or *cruP* were ligated into pET3atr (52) to produce p3-SLA and p3-SLP, respectively.

Pigment Analysis. *C. tepidum* and *Synechococcus* sp. PCC 7002 strains used for pigment analysis were grown until midexponential phase, and *E. coli* strains were grown until stationary phase. Cells were harvested by centrifugation, and pigments were extracted by sonication in acetone:methanol (7:2 vol/vol). Pigments were separated by HPLC (Agilent Model 1100) equipped with a diode array detector (model G1315B) and controlled with Agilent ChemStation software (Agilent Technologies, Palo Alto, CA) on a 25 cm by 4.6 mm Discovery 5-μm C18 column (Supelco, Bellefonte, PA). The solvent system for analysis of all strains was modified from the previously described system (73). Solvent B was 50% methanol, 30% ethyl acetate, and 20% acetonitrile, and solvent A consisted of water:methanol:acetonitrile (62.5:21:16.5) containing 10 mM ammonium acetate. The gradients used for separation of pigments produced by *E. coli*, *C. tepidum*, and *Synechococcus* sp. PCC 7002 are described in SI Table 2.

Phylogenetic Analyses. The predicted amino acid sequence of CruA was used in a blastp search against the nonredundant database at NCBI (50) to identify homologs in other organisms. For inclusion in the phylogenetic analyses, sequences of lycopene cyclases were selected that had been characterized or that were closely related to characterized proteins. Amino acid sequences were aligned using the ClustalW module in MacVector 7.1.1 (Accelrys, San Diego, CA). Phylogenies were reconstructed using Paup version 4.0b (Sinauer, Sunderland, MA).

All predicted proteins of *C. tepidum* were searched using blastp against all proteins of completed or nearly completed bacterial genomes as described in ref. 56. *C. tepidum* proteins were then clustered according to the patterns of species in which homologs were found, a method known as phylogenetic profiling (74).

Mass spectrometric analyses were performed by Dr. A. Daniel Jones (Michigan State University, East Lansing, MI). The plasmid pAC-LYC was the gift of Dr. Francis X. Cunningham, Jr. (University of Maryland, College Park, MD). We thank an anonymous reviewer for helpful suggestions that improved the manuscript. This work was supported by U.S. Department of Energy Grant DE-FG02–94ER20137 and National Science Foundation Grant MCB-0519743 (both to D.A.B.).

1. Frank HA, Cogdell RJ (1996) *Photochem Photobiol* 63:257–264.
2. Varkonyi Z, Masamoto K, Debreczeny M, Zsiros O, Ughy B, Gombos Z, Domonkos I, Farkas T, Wada H, Szalontai B (2002) *Proc Natl Acad Sci USA* 99:2410–2415.
3. Armstrong GA (1997) *Annu Rev Microbiol* 51:629–659.
4. Sieiro C, Poza M, de Miguel T, Villa TG (2003) *Int Microbiol* 6:11–16.
5. Takaichi S, Maoka T, Masamoto K (2001) *Plant Cell Phys* 42:756–762.
6. Takaichi S, Mochimaru M, Maoka T, Katoh H (2005) *Plant Cell Phys* 46:497–504.
7. Kishimoto S, Maoka T, Sumimoto K (2005) *Biosci Biotechnol Biochem* 69:2122–2128.
8. Ruiz D, Egea J, Tomás-Barberán FA, Gil MI (2005) *J Agric Food Chem* 53:6368–6374.
9. Bonora A, Pancaldi S, Gualandri R, Fasulo MP (2000) *J Exp Bot* 51:873–884.

10. Mercadante AZ, Steck A, Pfander H (1999) *J Agric Food Chem* 47:145–151.
11. Takaichi S, Wang ZY, Umetsu M, Nozawa T, Shimada K, Madigan MT (1997) *Arch Microbiol* 168:270–276.
12. Frigaard N-U, Maresca JA, Yunker CE, Jones AD, Bryant DA (2004) *J Bacteriol* 186:5210–5220.
13. Maresca JA, Frigaard, N-U, Bryant DA (2004) in: *Photosynthesis: Fundamental Aspects to Global Perspectives: Proceedings of the XIIIth International Congress on Photosynthesis, Montreal*, eds van der Est A, Bruce D (Allen, Lawrence, KS) pp 884–886.
14. Fromme P, Jordan P, Krauss N (2001) *Biochim Biophys Acta* 1507:5–31.
15. Loll B, Kern J, Saenger W, Zouni A, Biesiadka J (2005) *Nature* 438:1040–1044.
16. Liaaen-Jensen S (1965) *Acta Chemica Scand* 19:1025–1030.
17. Krugel H, Krubasik P, Weber K, Saluz HP, Sandmann G (1999) *Biochim Biophys Acta* 1439:57–64.
18. Misawa N, Nakagawa M, Kobayashi K, Yamano S, Izawa Y, Nakamura K, Harashima K (1990) *J Bacteriol* 172:6704–6712.
19. Matsumura H, Takeyama H, Kusakabe E, Burgess JG, Matsunaga T (1997) *Gene* 189:169–174.
20. Cunningham FX, Jr, Sun Z, Chamovitz D, Hirschberg J, Gantt E (1994) *Plant Cell* 6:1107–1121.
21. Stickforth P, Steiger S, Hess WR, Sandmann G (2003) *Arch Microbiol* 179:409–415.
22. Krubasik P, Sandmann G (2000) *Mol Gen Genet* 263:423–432.
23. Viveiros M, Krubasik P, Sandmann G, Houssaini-Iraqui M (2000) *FEMS Microbiol Lett* 187:95–101.
24. Peck RF, Johnson EA, KrebsMP (2002) *J Bacteriol* 184:2889–2897.
25. Hemmi H, Ikejiri S, Nakayama T, Nishino T (2003) *Biochem Biophys Res Commun* 305:586–591.
26. Mongodin EF, Nelson KE, Daugherty S, DeBoy RT, Wister J, Khouri H, Weidman J, Walsh DA, Papke RT, Sanchez-Perez G, *et al.* (2005) *Proc Natl Acad Sci USA* 102:18147–18152.
27. Krubasik P, Sandmann G (2000) *Biochem Soc Trans* 28:806–810.
28. Cunningham FX, Jr, Gantt E (2001) *Proc Natl Acad Sci USA* 98:2905–2910.
29. Teramoto M, Takaichi S, Inomata Y, Ikenaga H, Misawa N (2003) *FEBS Lett* 545:120–126.
30. Tao L, Picataggio S, Rouviere PE, Cheng Q (2004) *Mol Genet Genomics* 271:180–188.
31. Hornero-Mendez D, Britton G (2002) *FEBS Lett* 515:133–136.
32. Hirabayashi H, Ishii T, Takaichi S, Inoue K, Uehara K (2004) *Photochem Photobiol* 79:280–285.
33. Takaichi S, Sandmann G, Schnurr G, Satomi Y, Suzuki A, Misawa N (1996) *Eur J Biochem* 241:291–296.
34. Albrecht M, Takaichi S, Misawa N, Schnurr G, Boger P, Sandmann G (1997) *J Biotechnol* 58:177–185.
35. Albrecht M, Linden H, Sandmann G (1996) *Eur J Biochem* 236:115–120.
36. Bartley GE, Scolnik PA, Beyer P (1999) *Eur J Biochem* 259:396–403.
37. Breitenbach J, Fernandez-Gonzalez B, Vioque A, Sandmann G (1998) *Plant Mol Biol* 36:725–732.
38. Breitenbach J, Kuntz M, Takaichi S, Sandmann G (1999) *Eur J Biochem* 265:376–383.
39. Fraser PD, Linden H, Sandmann G (1993) *Biochem J* 291:687–692.
40. Linden H, Misawa N, Chamovitz D, Pecker I, Hirschberg J, Sandmann G (1991) *Z Naturforsch* 46c:1045–1051.
41. Linden H, Misawa N, Saito T, Sandmann G (1994) *Plant Mol Biol* 24:369–379.
42. Schneider C, Boeger P, Sandmann G (1997) *Protein Expr Purif* 10:175–179.
43. Breitenbach J, Vioque A, Sandmann G (2001) *Z Naturforsch* 56c:915–917.
44. Isaacson T, Ronen G, Zamir D, Hirschberg J (2002) *Plant Cell* 14:333–342.
45. Masamoto K, Wada H, Kaneko T, Takaichi S (2001) *Plant Cell Phys* 42:1398–1402.
46. Park H, Kreunen SS, Cuttriss AJ, DellaPenna D, Pogson BJ (2002) *Plant Cell* 14:321–332.
47. Fraser PD, Misawa N, Linden H, Yamano S, Kobayashi K, Sandmann G (1992) *J Biol Chem* 267:19891–19895.
48. Giuliano G, Pollock D, Scolnik PA (1986) *J Biol Chem* 261:12925–12929.
49. Raisig A, Bartley G, Scolnik PA, Sandmann G (1996) *J Biochem* 119:559–564.
50. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) *Nucleic Acids Res* 25:3389–3402.
51. Maresca JA, Bryant DA (2006) *J Bacteriol* 188:6217–6223.
52. Tan S (2001) *Protein Expr Purif* 21:224–234.
53. Kaneko T, Tanaka A, Sato S, Kotani H, Sazuki T, Miyajima N, Sugiura M, Tabata S (1995) *DNA Res* 2:153–166.
54. Giraud E, Hannibal L, Fardoux J, Jaubert M, Jourand P, Dreyfus B, Sturgis JN, Vermeglio A (2004) *J Biol Chem* 279:15076–15083.
55. Phadwal K (2005) *Gene* 345:35–43.
56. Eisen JA, Nelson KE, Paulsen IT, Heidelberg JF, Wu M, Dodson RJ, Deboy R, Gwinn ML, Nelson WC, Haft DH, *et al.* (2002) *Proc Natl Acad Sci USA* 99:9509–9514.
57. Liang C, Zhao FQ, Wei W, Wen ZX, Qin S (2006) *Int J Biol Sci* 2:197–207.
58. Steiger S, Jackisch Y, Sandmann G (2005) *Arch Microbiol* 184:207–214.
59. Ehling-Schultz M, Bilger W, Scherer S (1997) *J Bacteriol* 179:1940–1945.
60. Ivanov AG, Miskiewicz E, Clarke AK, Greenberg BM (2000) *Photochem Photobiol* 72:772–779.
61. Bornemann S (2002) *Nat Prod Rep* 19:761–772.
62. Barkley SJ, Cornish RM, Poulter CD (2004) *J Bacteriol* 186:1811–1817.
63. Kaneda K, Kuzuyama T, Takagi M, Hayakawa Y, Seto H (2001) *Proc Natl Acad Sci USA* 98:932–937.
64. Barkley SJ, Desai SB, Poulter CD (2004) *Org Lett* 6:5019–5021.
65. Wouters J, Oudjama Y, Barkley SJ, Tricot C, Stalon V, Droogmans L, Poulter CD (2003) *J Biol Chem* 278:11903–11908.
66. Mohamed H, Vermaas WFJ (2006) *J Bacteriol* 188:3337–3344.
67. Cunningham FX, Chamovitz D, Misawa N, Gantt E, Hirschberg J (1993) *FEBS Lett* 328:130–138.
68. Wahlund TM, Madigan MT (1995) *J Bacteriol* 177:2583–2588.
69. Wahlund TM, Woese CW, Castenholz RW, MadiganMT (1991) *Arch Microbiol* 156:81–90.
70. Frigaard N-U, Bryant DA (2001) *Appl Environ Microbiol* 67:2538–2544.
71. Stevens DR, Porter R (1980) *Proc Natl Acad Sci USA* 77:6052–6056.
72. Frigaard N-U, Sakuragi Y, Bryant DA (2004) *Methods Mol Biol* 274:325–340.
73. FrigaardN-U, Voigt GD, Bryant DA (2002) *J Bacteriol* 184:3368–3376.
74. Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D, Yeates TO (1999) *Proc Natl Acad Sci USA* 96:4285–4288.