Chemistry & Biology, Vol. 12, 453–460, April, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.chembiol.2005.02.010

Identification of a Carotenoid Oxygenase Synthesizing Acyclic Xanthophylls: Combinatorial Biosynthesis and Directed Evolution

Benjamin N. Mijts, Pyung Cheon Lee, and Claudia Schmidt-Dannert*
Department of Biochemistry, Molecular Biology, and Biophysics
1479 Gortner Avenue
University of Minnesota
St. Paul, Minnesota 55108

Summary

A carotenoid desaturase homolog from *Staphylococcus aureus* (CrtOx) was identified. When expressed in engineered *E. coli* cells synthesizing linear C_{30} carotenoids, polar carotenoid products were generated, identified as aldehyde and carboxylic acid C_{30} carotenoid derivatives. The major product in this engineered pathway is the fully desaturated C_{30} dialdehyde carotenoid 4,4′-diapolycopen-4,4′-dial. Very low carotenoid yields were observed when CrtOx was complemented with the C_{40} carotenoid lycopene pathway. But extension of an in vitro evolved pathway of the fully desaturated 2,4,2′,4′-tetradehydrolycopene produced the structurally novel fully desaturated C_{40} dialdehyde carotenoid 2,4,2′,4′-tetradehydrolycopendial. Directed evolution of CrtOx by error-prone PCR resulted in a number of variants with higher activity on C_{40} carotenoid substrates and improved product profiles. These findings may provide new biosynthetic routes to highly polar carotenoids with unique spectral properties desirable for a number of industrial and pharmaceutical applications.

Introduction

Carotenoids are a structurally diverse group of pigmented biomolecules of emerging importance as food supplements or colorants and in nutraceutical and pharmaceutical applications. Carotenoids are structurally classified based on the number of backbone carbon molecules, usually C_{30}, C_{40}, or C_{50}. Carotenoid biosynthesis occurs via a head-to-head condensation reaction of isoprenoid precursors followed by a desaturation reaction to increase the number of conjugated double bonds generating the distinctive carotenoid chromophore. Generally well-conserved carotenoid synthase and desaturase enzymes catalyze these reactions [1].

Further modification reactions, catalyzed by a diverse array of modifying enzymes, generate the considerable structural diversity observed in natural carotenoids. C_{30} carotenoids are relatively uncommon linear molecules found primarily in Gram-positive microorganisms such as *Staphylococcus aureus* strains [2], *Methylobacterium rhodinum* (formerly *Pseudomonas rhodos*) [3], and *Heliobacteria* [4]. The C_{30} carotenoids described from these organisms are often modified at one or both ends by glycosylation or acylation, and synthesis is thought to occur via an oxygenated intermediate [5, 6]. Although the genes responsible for C_{30} carotenoid backbone synthesis are known (dehydrosqualene synthase *crtM* and dehydrosqualene desaturase *crtN*) [7], the modifying genes from these organisms have not been characterized.

The major carotenoid of the pathogen *Staphylococcus aureus* is staphyloxanthin (Figure 1, structure 13). Structural studies have indicated that it is a derivative of the C_{30} carotenoid diaponeurosporene modified by glycosylation and acylation at one terminal methyl group. Mutagenesis and color screening of *S. aureus* indicated that the biosynthesis proceeds by the introduction of terminal oxygen functions to form a carboxylic acid via an aldehyde intermediate. This is followed by the esterification of glucose to this carboxyl terminus and esterification of a fatty acid to the glucose moiety [5]. Similar studies on the bacterium *Methylobacterium rhodinum* suggested that diapolycopene derivatives observed in this organism with one or both termini modified by glycosylation and acylation were synthesized via similar intermediates [6].

The biosynthesis of short-chain, polar carotenoid molecules has considerable potential in industrial and biomedical applications. The recently reported plant biosynthetic pathways for crocetin [8] and bixin [9] proceed via enzymatic cleavage of C_{40} carotenoid substrates to produce oxygenated short-chain apo-carotenoids that are further modified by methylation or glycosylation. The products of these reactions are often water soluble and are important natural food additives with potential as cancer therapeutic agents [10].

By DNA shuffling of carotenoid desaturases (CrtI) from *Erwinia* species, a variant enzyme ($CrtI_{14}$) was previously reported that catalyzes the extended desaturation of the linear C_{40} carotenoid pathway to 3,4,3′4′-tetradehydrolycopene and when combined with an evolved carotenoid cyclase, produces the monocyclic carotenoid torulene [11]. This pathway was extended by combinatorial engineering with a number of carotenoid modifying enzymes, to produce an array of structurally novel products [12]. In this paper we extend this methodology to new carotenoid products by engineering and directed evolution of a *S. aureus* C_{30} carotenoid oxygenase for the production of oxygenated, linear C_{30} and C_{40} carotenoids in *E. coli*.

Results

Identification and Cloning of *S. aureus* C_{30} Carotenoid Biosynthetic Genes

NCBI BLAST homology searches [13] using the known *S. aureus* diapophytoene synthase (CrtN) amino acid sequence (GenBank accession number B55548) revealed a homologous open reading frame upstream of CrtN within the genomic DNA sequences of *S. aureus* strains. Analysis of this region by the TIGR CMR database [14] suggested that the known *S. aureus* carot-

*Correspondence: schmi232@tc.umn.edu

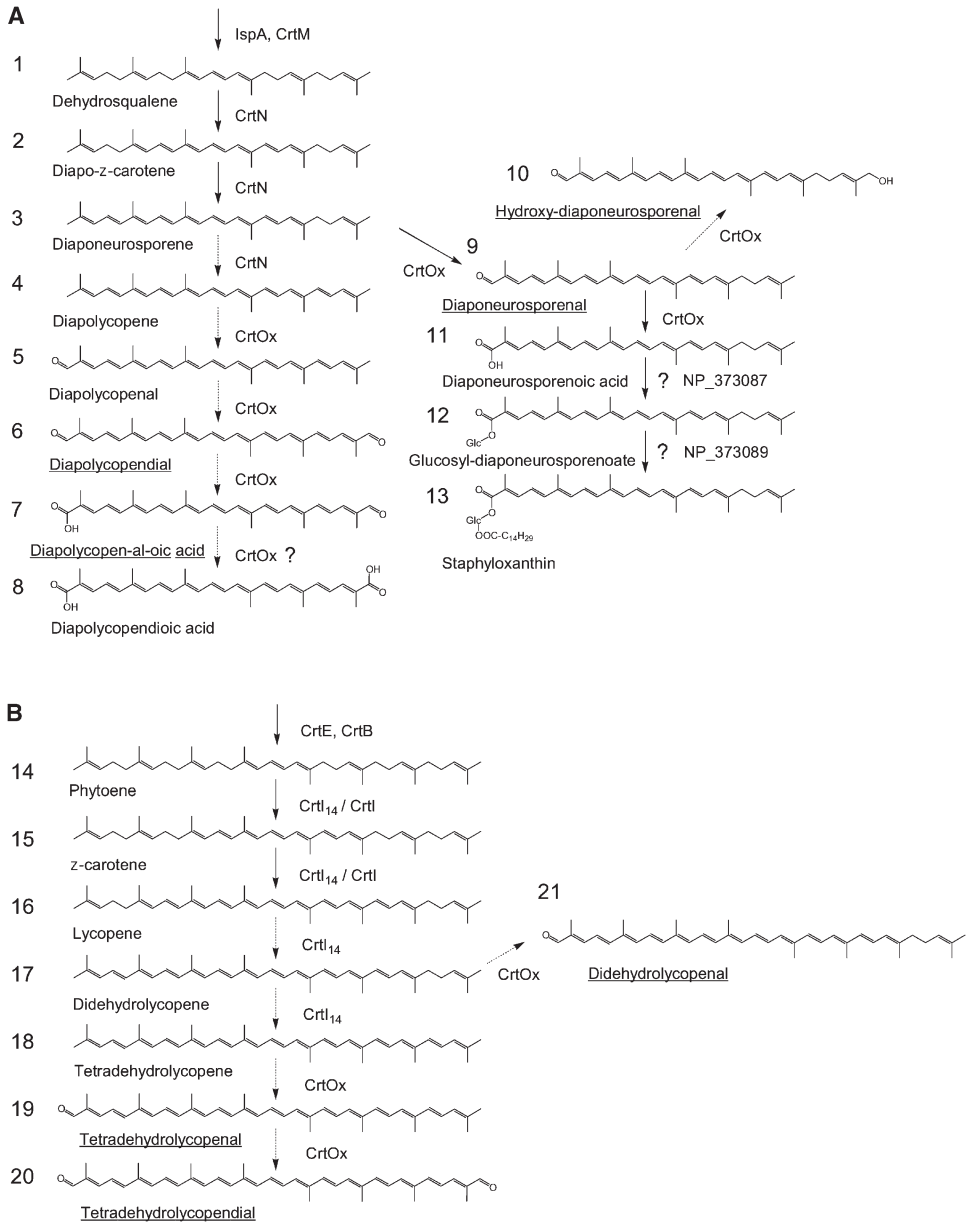

Figure 1. Biosynthesis of Oxygenated C_{30} and C_{40} Carotenoids

Biosynthetic pathways to oxygenated, linear C_{30} (A) and C_{40} (B) carotenoids. Solid arrows represent natural biosynthetic pathways suggested for staphyloxanthin in *S. aureus* (A) and lycopene (B). Biosynthetic pathway steps in engineered recombinant *E. coli* are indicated by dashed arrows. Carotenoids identified in recombinant *E. coli* are underlined.

enoid biosynthetic genes were preceded by three additional genes in a typical bacterial operon structure (Figure 2). These consisted sequentially of a small ORF with no known functional homology (GenBank accession number NP_373089), a homolog of *crtN* (GenBank accession number NP_373088, 25% identity, 48% similarity), an ORF with a conserved glycosyl transferase motif (GenBank accession number NP_373087), followed by the known genes *crtN* and dehydrosqualene synthase *crtM* (GenBank accession number A55548). The structure of this operon suggested these genes may be involved in the biosynthesis of staphyloxanthin (Figure 1, structure 13), the glycosylated, acylated major carotenoid of *S. aureus* [2]. In order to functionally characterize these genes, they were subcloned into the overexpression vector pUCMod for expression in *E. coli* (see Table 1 for plasmid and gene names). However, cloning of the genes NP_373089 and NP_373087 proved extremely difficult, and although clones with correct sequences were obtained, the *E. coli* strains harboring these plasmids grew extremely slowly with an unusual, transparent colony morphology on agar plates. *E. coli* JM109 harboring pUCMod-NP_373088 did not demonstrate this growth inhibition. When pUCMod-

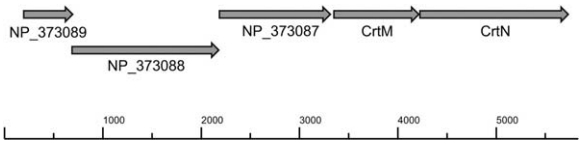

Figure 2. *S. aureus* Carotenoid Operon

Region of the *S. aureus* genome sequence showing open reading frames upstream of the known *S. aureus* carotenoid genes *crtM* (dehydrosqualene synthase) and *crtN* (dehydrosqualene desaturase) in a typical bacterial operon structure. NP_373089, *crt*OX; NP_373089, no known functional homology; NP_373087, conserved glycosyl transferase motif.

NP_373088 was electrotransformed into strain JM109 harboring the plasmid pAC-*ispA-crtM-crtN* (producing the C_{30} carotenoid diapolycopene) [11], cells with a deep red phenotype were produced (Figure 3A), indicating a change in carotenoid production when compared to the yellow-orange cells of the background strain.

Characterization of C_{30} Carotenoids

For analysis of carotenoids, *E. coli* strain JM109 harboring pAC-*ispA-crtM-crtN* and pUC-NP_373088 was cultured in LB medium supplemented with glycerol at 30°C, as previous results have indicated that low temperatures and glycerol supplementation enhance carotenoid production [15]. Initial analysis of carotenoid extracts by TLC (Figure 3) indicated that a number of carotenoids were present compared to a control strain JM109 harboring pAC-*ispA-crtM-crtN* and pUCMod vector without insert DNA. HPLC analysis indicated the presence of a number of additional polar peaks (Figure 4A). These were analyzed by mass spectrometry and

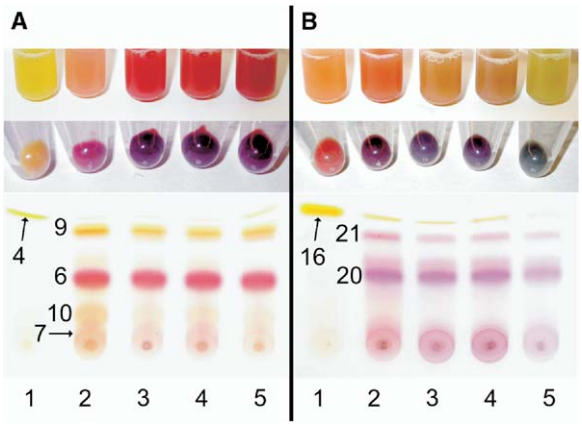

Figure 3. Carotenoid Production in Recombinant *E. coli* Cells

LB media cultures (top), cell pellets (center), and TLC analysis (bottom) of recombinant C_{30} (A) and C_{40} (B) carotenoid-producing *E. coli* strains. Background plasmids strains are JM109 pAC-*ispA-crtM-crtN* (A) and pAC-*crtE-crtB-crtI*$_{14}$ (B) cotransformed with 1. pUCMod, 2. pUC-*crtOx*, 3. pUC-*crtOx*$_{mut1}$, 4. pUC-*crtOx*$_{mut2}$, and 5. pUC-*crtOx*$_{mut3}$. Identified compounds from Figure 1 are indicated.

by a combination of HPLC retention times, UV-Vis fine spectra, and Mass-spectra the major peaks were assigned structures in Figure 1. The properties of the carotenoids are summarized in Table 2. The major product is a violet compound with a $[M]^+$ of *m/z* 429.0 assigned as 4,4′-diapolycopen-4,4′-dial. Characteristic mass fragments of two aldehyde functions were observed (M-18, M-28, M-18-18, M-18-28) as well as characteristic carotenoid extrusion losses of toluene (M-92) and xylene (M-106). The major violet carotenoid reacted very rapidly with $NaBH_4$, producing a more polar, yellow-

Table 1. Genes and Plasmids Used in This Study

Gene	Enzyme	Typical Reaction Catalyzed	Accession Number or Reference
ispA	FDP synthase	Head-to-head condensation of 2 IDP + 1 DMADP	AB025791
crtM	Dehydrosqualene synthase	Head-to-head condensation of 2 FDP	X73889
crtN	Diapophytoene synthase	Introduction of three desaturations in dehydrosqualene	X73889
crtE	GGDP synthase	Head-to-head condensation of IDP + FDP	D90087
crtB	Phytoene synthase	Head-to-head condensation of 2 GGDP	D90087
crtI	Phytoene desaturase	Introduction of four desaturations in phytoene	D90087
crtI$_{14}$	*In vitro* evolved phytoene desaturase	Introduction of six desaturations in phytoene	[11]
crtOx	Diapocarotenal synthase	Oxygenation of acyclic carotenoid end groups	This study
Plasmid		**Properties**	**Reference**
pUCMod		Cloning vector, high copy-number, constitutive *lac* promoter, Amp^R	[11]
pACMod		Cloning vector, low copy-number, constitutive *lac* promoter, Cm^R	[11]
pAC-*ispA-crtM-crtN*		Constitutively expressed *E. coli* IspA, *S. aureus* CrtN, and *S. aureus* CrtM producing diapolycopene	[12]
pAC-*crtE-crtB-crtI*		Constitutively expressed *E. uredovora* CrtE, CrtB, and CrtI producing lycopene	[11]
pAC-*crtE-crtB-crtI*$_{14}$		Constitutively expressed *E. uredovora* CrtE, CrtB, and in vitro-evolved desaturase $CrtI_{14}$ producing tetradehydrolycopene	[11]
pUC-*crtOx*		Constitutively expressed *S. aureus* diapocarotenal synthase CrtOx	This study
pUC-*crtOx*$_{mut1-3}$		Constitutively expressed *S. aureus* diapocarotenal synthase $CrtOx_{mut1-3}$ obtained by error-prone PCR mutagenesis	This study

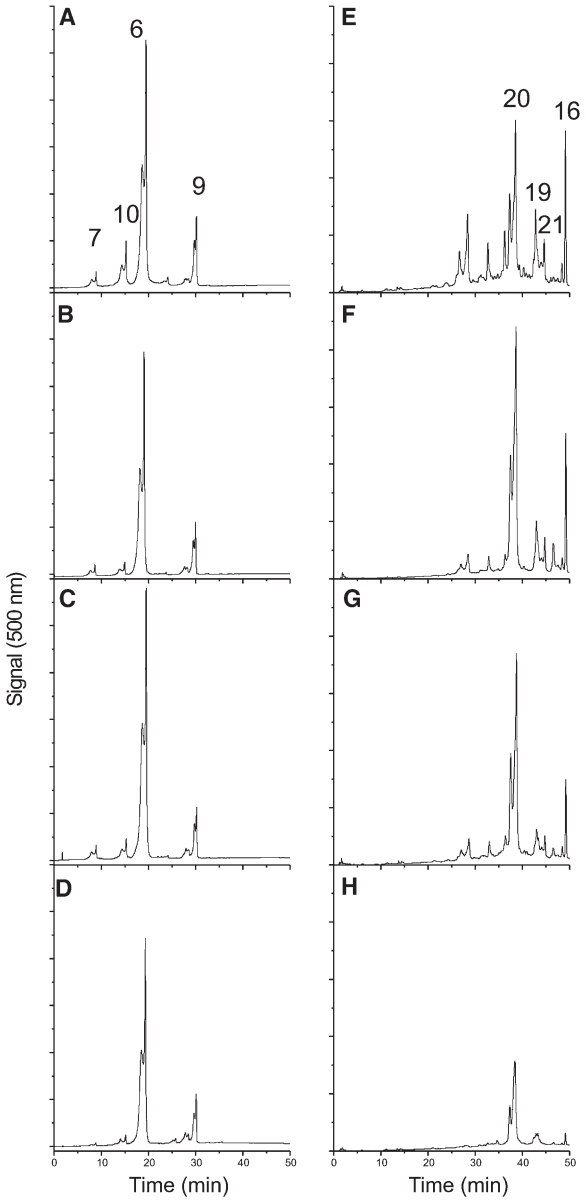

Figure 4. Analysis of *E. coli* Cells Producing Oxygenated C_{30} Carotenoids

HPLC profiles of recombinant *E. coli* expressing the C_{30} carotenoid background plasmid pAC-*ispA*-*crtN*-*crtM* with (A) pUC-*crtOx*, (B) pUC-$crtOx_{mut1}$, (C) pUC-$crtOx_{mut2}$, (D) pUC-$crtOx_{mut3}$, and *E. coli* expressing the C_{40} carotenoid background plasmid pAC-*crtE*-*crtB*-$crtI_{14}$ with (E) pUC-*crtOx*, (F) pUC-$crtOx_{mut1}$, (G) pUC-$crtOx_{mut2}$, and (H) pUC-$crtOx_{mut3}$.

orange compound by TLC analysis as a result of the reduction of terminal aldehyde groups to hydroxyl groups (data not shown). A less polar peak with a $[M]^+$ of *m/z* 417.1 consistent with a mono-aldehyde derivative of diaponeurosporene (Figure 4A, peak 9) was also observed with mass fragments characteristic of a single aldehyde group (M-18, M-28). Reaction with $NaBH_4$ rapidly generated a more polar, yellow product by TLC with an Rf consistent with a carotenoid monoalcohol. These results indicate the enzyme encoded by ORF NP_373088 catalyzes the addition of one or more aldehyde groups to C_{30} carotenoid terminal methyl groups and is likely responsible for the synthesis of the monoaldehyde intermediate observed in the biosynthesis of staphyloxanthin [5]. As a result we have designated this oxygenase gene CrtOx, diapocarotenal synthase.

In addition to the major product 4,4′-diapolycopen-4,4′-dial, two more polar peaks were observed by HPLC analysis, suggesting products with additional oxygen functions are present. Although present in relatively low yields, parent masses could be obtained for both compounds and structures putatively assigned. The least polar of these (Figure 4A, peak 10) has a parent mass $[M]^+$ of *m/z* 433.1 and characteristic fragments of M-2, and M-18, and was assigned as hydroxy-4,4′-diaponeurosporenal. Reaction with $NaBH_4$ was rapid and produced a more polar, yellow product by TLC, indicating reduction to a carotenoid dialcohol. The lack of UV-Vis fine structure observed for this compound (Table 2) suggests the aldehyde function is located adjacent to the conjugated double bond system, and this compound could therefore be assigned as 4′-hydroxyl-4,4′-diaponeurosporen-4-al (Figure 1, compound 10). The most polar peak (Figure 4A, peak 7) with a $[M]^+$ of *m/z* 445.0 and a very strong fragment at M-18 and an additional fragment at M-36 was assigned as 4,4′-diapolycopen-4-al-4′-oic acid. Although present in relatively low yields, the presence of this compound suggests that CrtOx catalyses both the oxidation of C_{30} carotenoids to aldehydes and the further oxidation of these aldehyde groups to carboxylic acids. The expected final product of this engineered pathway, 4,4′-diapolycopen-4,4′-dioic acid was not observed in acetone extracts and increasing culture times of 48, 64, and 96 hr yielded only similar product profiles by HPLC analysis (data not shown). However, cell pellets of these cultures after acetone extraction demonstrated that levels of a red pigment not accessible by acetone extraction increased over time. This pigment could not by extracted with a range of organic solvents but could be solubilized by the addition of 1% aqueous KOH to cell pellets followed by stirring at room temperature for 2 hr. This is consistent with the chemical properties of the plant C24 dicarboxylic acid carotenoid norbixin [9] that forms a soluble potassium salt in aqueous KOH. On the addition of acetic acid to pH 5, an insoluble precipitate formed that was not soluble in a number of organic solvents tested with the exception of DMSO. The lack of solubility of this compound in most organic solvents renders purification and analysis difficult and no clear analytical data could be obtained. However, the known biochemical pathway and physical properties of the pigment strongly suggest a diapocarotene-dioic acid.

C_{40} Carotenoids Produced by Complementation of CrtOx with the In Vitro-Evolved Tetradehydrolycopene Pathway

When *E. coli* JM109 harboring the plasmid pAC-*crtE*-*crtB*-*crtI* necessary for C_{40} carotenoid lycopene synthesis was electrotransformed with pUC-*crtOx*, very low levels of more polar carotenoids were observed by

Table 2. Properties of Carotenoids

Structure (Figure 1)	Compound	Exact Mass	Observed $[M]^+$ (m/z)	Observed Fragments	UV-Vis Maxima [Shoulder] (in Acetonitrile)
6	Diapolycopenedial	428.27	429.0	M-18, M-28, M-36, M-46, M-92, M-106	508 [536]
7	Diapolycopen-al-oic acid	444.27	445.0	M-18, M-36, M-110	515 [539]
9	Diaponeurosporenal	416.31	417.1	M-18, M-28, M-92, M-106	469 [490]
10	Hydroxy-Diaponeurosporenal	432.3	433.1	M-2, M-18, M-28, M-36, M-92, M-106	480 [500]
19	Tetradehydrolycopenal	546.39	547.2	M-18, M-92, M-106	521 [552]
20	Tetradehydrolycopendial	560.37	561.1	M-18, M-28, M-36, M-92, M-106	537 [563]
21	Didehydrolycopenal	548.4	549.2	M-18, M-28, M-92, M-106	513 [541]

TLC when compared to a control strain transformed with pUCMod. HPLC analysis of carotenoid extracts of *E. coli* strain JM109 harboring pAC-*crtE-crtB-crtI* with pUCMod or pUC-*crtOx* (Figure 5) indicated that the presence of CrtOx significantly reduced the yield of lycopene and increased the accumulation of the precursor molecule phytoene. Although very low levels of more polar products were observed, these results indicate that lycopene is a poor substrate for CrtOx and that an enzyme interaction may be disrupting desaturation by CrtI. We previously described the complementation of

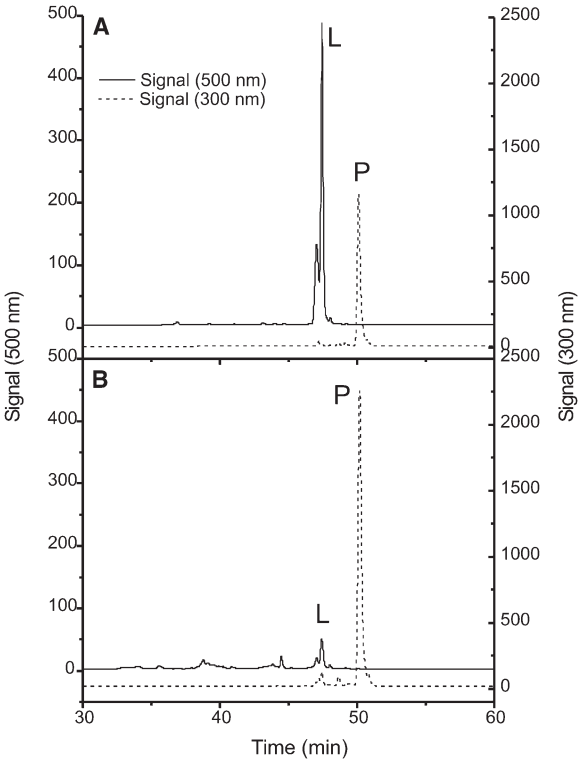

Figure 5. Carotenoid Profiles of Lycopene-Producing Recombinant *E. coli* Cells with and without CrtOx Present

HPLC profiles at wavelengths of 300 and 500 nm of recombinant *E. coli* expressing the C_{40} carotenoid lycopene background plasmid pAC-*crtE-crtB-crtI* with (A) pUCMod and (B) pUC-*crtOx*. Based on mass spectrometry, HPLC retention time, and UV-Vis spectra, the peaks were identified as L: lycopene (500 nm) and P: phytoene (300 nm).

the in vitro-evolved C_{40} carotenoid 2,4,2′,4′-tetradehydrolycopene pathway with a number of carotenoid modifying enzymes in recombinant *E. coli* [12]. When *E. coli* JM109 harboring the plasmid pAC-*crtE-crtB-crtI*$_{14}$ necessary for tetradehydrolycopene synthesis was electrotransformed with pUC-*crtOx*, cells with a deep red color phenotype were produced when compared to the pink/red color of the background strain with pUCMod vector without insert. HPLC analysis (Figure 4E) indicated that a number of new polar products were present. In order to structurally characterize these carotenoids they were analyzed by HPLC-mass spectrometry. The assigned structures of the major peaks on the HPLC chromatogram shown in Figure 1 were determined by a combination of HPLC retention times, UV-Vis spectra, and Mass spectra summarized in Table 2. The major product (Figure 4E, peak 20), with a parent mass $[M]^+$ of *m/z* 561.1 and mass fragments characteristic of two aldehyde functions (M-18, M-28, M-18-18) and characteristic carotenoid extrusion losses of toluene (M-92) and xylene (M-106) was identified as the fully desaturated C_{40} dialdehyde 2,4,2′,4′-tetradehydrolycopendial. However, considerable levels of the C_{40} biosynthesis pathway precursor lycopene were also observed (Figure 4E, peak 16). This suggests that although CrtOx is active on more desaturated C_{40} carotenoid substrates, it has little activity on pathway precursors such as lycopene. Two additional less polar molecules could be identified as mono-aldehyde derivatives. The least polar of these (Figure 4E, peak 21) with a $[M]^+$ of *m/z* 549.2 was assigned as 2,4-didehydrolycopenal and the higher yield, more polar peak (Figure 4E, peak 19) with a $[M]^+$ of *m/z* 547.2 was assigned as 2,4,2′,4′-tetradehydrolycopenal. Both had the characteristic mass fragments of one aldehyde function (M-18, M-28) and characteristic carotenoid extrusion losses of toluene (M-92) and xylene (M-106). No peaks corresponding to aldehyde derivatives of lycopene were present. Although additional highly polar compounds can be observed in the HPLC chromatogram, these were relatively low yield and molecular structures could not be positively identified by mass spectrometry. These may represent low yields of mono- or dicarboxylic acids or nonspecific pathway derivatives.

Construction of In Vitro Evolution Libraries and Isolation and Sequence of Mutants

In order to alter the product profile of the C_{40} carotenoids produced by complementation with CrtOx and

improve the relative yields of oxygenated C_{40} carotenoids, an error-prone PCR mutagenesis library of CrtOx was constructed and electrotransformed into *E. coli* strain JM109 harboring the plasmid pAC-*crtE-crtB-crtI*$_{14}$ necessary for 2,4,2′,4′-tetradehydrolycopene production. Colonies with altered carotenoid production were identified by color screening. Screening of approximately 3000 colonies by this method yielded three mutants with altered cell pigment phenotypes designated $CrtOx_{Mut1}$, $CrtOx_{Mut2}$, and $CrtOx_{Mut3}$. These colonies had deep purple ($CrtOx_{Mut1}$, $CrtOx_{Mut2}$) and blue/gray ($CrtOx_{Mut3}$) color phenotypes (Figure 3B). The DNA sequences of the inserts of these plasmids were determined and plasmids retransformed into *E. coli* cells harboring pAC-*crtE-crtB-crtI*$_{14}$, confirming the color phenotypes observed. The mutants were also electrotransformed into *E. coli* JM109 harboring pAC-*ispA-crtN-crtM* to test the activity of these mutants against C_{30} carotenoid substrates (Figure 3A). The amino acid changes in the CrtOx sequence for each mutant were as follows: $CrtOx_{Mut1}$–V355A, Q483L; $CrtOx_{Mut2}$–V355A; $CrtOx_{Mut3}$–Y62H, Y189C, K393E.

Carotenoid Analysis of In Vitro-Evolved CrtOx Mutants Complemented with C_{30} and C_{40} Pathways

HPLC chromatograms of the carotenoid products of the in vitro-evolved CrtOx genes coexpressed with the recombinant 2,4,2′,4′-tetradehydrolycopene pathway are shown in Figures 4F–4H. Based on HPLC retention times and UV-Vis spectra, it appeared that the mutant enzymes produced no significant amounts of new products. However, considerable alterations in the product profiles were detected, consistent with the altered color phenotypes observed by screening (Figure 3B). $CrtOx_{Mut1}$ and $CrtOx_{Mut2}$, as expected from the similar amino acid sequence and colony color phenotype of these mutants, have similar overall product profiles. Both have higher yields of 2,4,2′,4′-tetradehydrolycopendial ($CrtOx_{Mut1}$: 57% increase, $CrtOx_{Mut2}$: 18% increase) and lower yields of the more polar, unidentified products. In addition, a significant decrease in the relative yield of the precursor lycopene was also observed. This is likely to be partly responsible for the altered color phenotypes observed by screening. By peak integration of spectra at 500 nm, wild-type CrtOx produces a ratio of 2.6:1 (2,4,2′,4′-tetradehydrolycopendial:lycopene) whereas $CrtOx_{Mut1}$ has a ratio of 4.8:1 and $CrtOx_{Mut2}$ 6.4:1. Finally, $CrtOx_{Mut3}$ shows the most dramatic change in product spectrum accordant with the greater number of amino acid changes observed. Although it has the lowest yield of 2,4,2′,4;-tetradehydrolycopendial (41% decrease compared to CrtOx wild-type), the precursor compound lycopene is almost completely absent compared to 2,4,2′,4;-tetradehydrolycopendial with a ratio of 21:1 and significant peaks of the more polar, unidentified compounds observed are not present. By integration of all carotenoid peaks observed at 500 nm, 2,4,2′,4′;-tetradehydrolycopendial represents 84.5% of the total carotenoids detected in this strain.

In contrast, the in vitro evolved CrtOx enzymes appeared to have a less dramatic effect on the product profiles of the C_{30} diapolycopene pathway (Figures 4B–4D). However, a considerable decrease in the accumulation of the more polar products can be observed (Figures 4B–4D, peaks 7 and 10). This suggests that the mutations may have some influence on the putative carboxylic acid synthesis function of the enzyme on C_{30} carotenoid substrates.

Discussion

The C_{30} carotenoid biosynthesis genes *crtM* and *crtN* from *S. aureus* have previously been reported to catalyze the synthesis of linear C_{30} carotenoid backbones but not the major *S. aureus* carotenoid staphyloxanthin. By analysis of the region surrounding these genes in the *S. aureus* genome, we identified an operon-like structure that may be involved in staphyloxanthin biosynthesis. Heterologous expression of the gene *crtOx* in *E. coli*, along with *crtM* and *crtN*, resulted in the synthesis of oxygenated C_{30} carotenoids. A minor peak with a parent mass consistent with mono-carboxylic acid/mono-aldehyde derivative was also observed. This strongly suggests that the single gene *crtOx* catalyzes the biosynthesis of both the aldehyde and carboxylic acid intermediates in staphyloxanthin biosynthesis.

Interestingly, the major product observed in recombinant *E. coli* was a dialdehyde derivative of the fully desaturated C_{30} carotenoid diapolycopene, which is in contrast to staphyloxanthin, a diaponeurosporene derivative oxygenated at only one terminus. This product is likely the result of the engineered *E. coli* pathway in which all enzymes are constitutively expressed and CrtOx is expressed on the high copy-number plasmid pUCMod and the remaining genes on low copy-number pACMod. This may increase the pathway flux to more oxygenated products and thus a dialdehyde carotenoid derivative is formed. The fully desaturated (diapolycopene) backbone of the final product is consistent with previously reported recombinant C_{30} carotenoid biosynthesis in which diapolycopene is produced as a major product [12]. As CrtOx is homologous to CrtN and other carotenoid desaturases, it is possible that it retains some desaturase activity but expression of CrtOx and CrtM, in the absence of CrtN, failed to produce pigmented carotenoids (data not shown). The amino acid sequence relatedness of CrtN and CrtOx suggests evolution via a gene duplication event and subsequent functional differentiation. Although there are relatively few examples of carotenoid desaturase homologs with an alternative activity, some of the many putative carotenoid desaturases annotated in genome sequences may also have alternative functions and only experimental characterization of these genes will clarify this.

Two additional genes were identified in the putative *S. aureus* carotenoid operon. ORF NP_373087 has a conserved glycosyl transferase motif, suggesting it is may catalyze the addition of the sugar moiety observed in staphyloxanthin. The remaining small ORF NP_373089 has no significant homology to known enzymes but may be a structurally novel acyl-transferase. The growth inhibition on overexpression of these genes in *E. coli* indicates they have some influence on *E. coli* metabolism and may be accepting alternative sub-

strates. Additional characterization of these enzymes in a recombinant system to confirm activity may be possible with a more modulated expression system.

As with previous carotenoid modifying enzymes investigated [12], CrtOx was a relatively promiscuous enzyme, readily able to accept C_{40} carotenoid substrates. Initial experiments by complementing CrtOx with the genes necessary for synthesis of the C_{40} carotenoid lycopene (Figure 1, structure 16) resulted in a significant reduction in carotenoid yield when compared to a control. When combined with the in vitro-evolved 2,4,2′,4′-tetradehydrolycopene pathway, CrtOx catalyzed the synthesis of highly desaturated mono- and di-aldehyde C_{40} carotenoids. The engineered 2,4,2′,4′-tetradehydrolycopene pathway also accumulates significant levels of the precursor lycopene and this was also observed with the addition of CrtOx. This accumulation, and the lack of observed oxygenated lycopene derivatives, indicates that CrtOx preferentially accepts more desaturated substrates. The lack of carotenoid production and desaturation activity observed when CrtOx was coexpressed with the lycopene biosynthesis pathway may be the result of the formation of a disrupted carotenogenic enzyme complex. The major product of CrtOx activity on the 2,4,2′,4′-tetradehydrolycopene biosynthesis pathway was identified as the deep violet dialdehyde derivative 2,4,2′,4′-tetradehydrolycopendial. Although a number of more polar peaks were observed on HPLC analysis, they could not be positively identified. Based on the results of the engineered C_{30} pathway, these may represent carboxylic acid derivatives but further characterization is necessary to confirm this.

Color screening of CrtOx error-prone PCR libraries with the engineered 2,4,2′,4′-tetradehydrolycopene pathway yielded three clones with an altered color phenotype. Sequencing revealed each clone contained a unique pattern of mutations although $CrtOx_{Mut1}$ $CrtOx_{Mut2}$ share an amino acid change. Although HPLC analysis of the carotenoid profiles of the mutant strains indicated no novel products were observed, considerable changes in product distribution were observed. These changes are responsible for the altered color phenotypes observed. All of these mutants have relatively reduced yields of the pathway precursor lycopene, which again suggests some interaction of these heterologous enzymes is taking place. Mutant $CrtOx_{Mut3}$ has the most significant change in product profile and although overall yield is lower, 2,4,2′,4′-tetradehydrolycopendial in produced in considerable excess over other carotenoids detected. It is very difficult to speculate on the function of the amino acid changes observed in the mutants as no structural data are available for the carotenoid desaturase enzyme family. Coexpression of the CrtOx variants with the C_{30} carotenoid biosynthesis pathway yielded similar carotenoid product profiles to the wild-type CrtOx clone. However, a reduced accumulation of the more polar carboxylic acid products observed in these samples, along with the C_{40} pathway, suggests these mutations may compromise the aldehyde oxidase function of the enzyme. This also indicates that these highly polar carotenoid products are the result of enzymatic activity of wild-type CrtOx and not nonspecific in vivo activity.

Significance

Highly polar, linear carotenoid molecules frequently demonstrate unique spectral properties and oxygenated short-chain carotenoids may be water soluble, a considerable advantage in many industrial and pharmaceutical applications. By analysis of *S. aureus* genome sequences, we identified a carotenoid desaturase homolog CrtOx. Coexpression of this gene with the previously characterized, recombinant C_{30} biosynthetic pathway in *E. coli* yielded oxygenated carotenoids, including the major product identified as the fully desaturated C_{30} di-aldehyde carotenoid 4,4′-diapolycopen-4,4′-dial. These results indicate CrtOx is responsible for the biosynthesis of aldehyde and carboxylic acid carotenoid intermediates observed in *S. aureus*. Coexpression of CrtOx with genes necessary for the synthesis of the C_{40} carotenoid lycopene resulted in low carotenoid production in this strain. However, when CrtOx was combined with a previously reported, in vitro-evolved pathway for the synthesis of 2,4,2′,4′-tetradehydrolycopene, significant yields of oxygenated C_{40} carotenoids were observed, including the major product identified as the fully desaturated C_{40} di-aldehyde carotenoid 2,4,2′,4′-diapolycopendial. By directed evolution, CrtOx mutants with significantly improved production and yields of 2,4,2′,4′-diapolycopendial were generated. These mutants are ideal candidates to develop high-yield processes for the recombinant synthesis of this pink/violet, highly polar carotenoid.

Experimental Procedures

Sequence Analysis

Whole genome DNA sequences of *S. aureus* strains MW2 (NC_003923), N315 (NC_002745), and Mu50 (NC_002758) were obtained from NCBI. Protein sequences of *S. aureus* CrtN (B55548) and CrtM (A55548) were obtained from NCBI. Homology searches were performed using NCBI BLAST software [13]. Genome region analysis and ORF prediction were performed using TIGR Comprehensive Microbial Resource [14]. Sequence editing was performed using Bioedit software [16].

Plasmid Construction

The construction of the plasmids pAC-*ispA*-*crtM*-*crtN*, pAC-*crtE*-*crtB*-*crtI*, and pAC-*crtE*-*crtB*-*crtI*$_{14}$ producing diapolycopene, lycopene, and 2,4,2′,4′-tetradehydrolycopene, respectively, has been described previously [11, 12]. The C_{30} carotenoid oxygenase gene *crtOx* (GenBank accession number NP_373088) was amplified from *S. aureus* (ATCC 35556D) genomic DNA using the oligonucleotides CrtOx-f-X (5′-GCTCTAGAAGGAGGATTACAAAATGACTAAACATAT CATCG-3′) and CrtOx-r-N (5′-TTCCTTTGCGGCCGCTCACTTCCT ATTCTTCGC-3′) with added XbaI and NotI sites, respectively. The forward primer CrtOx-f-X also contains an optimized Ribosome Binding Site (underlined). This PCR product was digested with XbaI and NotI enzymes, gel purified, and ligated into similarly prepared pUCMod vector [11]. Insert containing plasmids were isolated and a number sequenced to confirm no PCR errors were present. Additional *S. aureus* genes (GenBank accession numbers NP_373087 and NP_373089) were amplified with specific primers and cloned in the same fashion.

For error-prone PCR mutagenesis [17], the *crtOx* gene in pUCMod was amplified with the PCR primers (5′-CCGACTGGAAAGC GGG-3′ and 5′-ACAAGCCCGTCAGGG-3′) flanking the gene and promoter. The PCR reaction mix consisted of 1× Promega Mg^{2+} free thermophilic buffer (Promega, Madison, WI), 10 ng/ml template plasmid, 1 μM of each primer, 5 U *Taq* DNA polymerase, and 0.3

mM dNTP mix. $MgCl_2$ and $MnCl_2$ were added to a final total salt concentration of 2 mM and separate reactions were performed with 0.2, 0.1, 0.05, and 0.025 mM final concentrations of $MnCl_2$. PCR was carried out with a program of 95°C for 4 min followed by 32 cycles of 94°C for 1 min, 50°C for 1 min, and 72°C for 1 min and finally 72°C for 7 min. The PCR products were purified using a QIAquick gel extraction kit (Qiagen, Valencia, CA), combined and digested with the restriction enzymes XbaI and NotI. The DNA fragments were ligated into the corresponding sites of the pUCmod vector [11] and electrotransformed into competent *E. coli* JM109 harboring pAC-*crtE-crtB-crtI*$_{14}$. Transformants were plated on LB agar plates supplemented with 100 μg/ml carbenicillin and 50 μg/ml chloramphenicol. After 18 hr of incubation at 30°C in the dark, colonies were replicated using a nitrocellulose membrane and transferred onto fresh LB plates containing the same antibiotics. Colonies were screened visually for color variants after an additional 24 hr incubation at room temperature. Mutations in the *crtOx* sequence were confirmed by DNA sequencing.

Carotenoid Production and Extraction

For HPLC and HPLC-mass spectrometry analysis, 100 ml cultures were grown in LB medium supplemented with 0.5% glycerol, 100 μg/ml carbenicillin, and 50 μg/ml chloramphenicol at 30°C for 24 hr and cells harvested by centrifugation (30 min, 4000 × g, 4°C). Carotenoid extraction was performed as described previously [12]. In brief, 5 ml of acetone was added to cell pellets and samples incubated in a sonicating water bath at 4°C for 30 min, followed by centrifugation (20 min, 4000 × g, 4°C) to remove cell debris. Extractions with acetone were repeated until no visible pigment remained and the supernatants pooled. Pooled extracts were dried down completely under a stream of N_2 gas and resuspended in 5 ml of ethyl acetate. Carotenoids were two-phase extracted with 10 ml 5 M NaCl, the solvent phase recovered, dried down, and resuspended in hexane or ethyl acetate.

TLC Analysis

Crude and purified C_{30} and C_{40} carotenoid extracts were initially analyzed by thin-layer choromatography with Whatman normal phase silica gel 60 plates developed using hexane:ethyl acetate (3:1).

HPLC and HPLC-Mass Spectrometry Analysis of Carotenoids

HPLC separation was performed using a Zorbax 300SB-C18 column (4.6 × 150 mm, 2.5 μm; Agilent Technologies, Palo Alto, CA) at a flow rate of 1 ml min^{-1} using an Agilent 1100 HPLC system equipped with a photodiode array detector. For carotenoid separations the mobile phase consisted of dH_2O:acetonitrile (30:70) for 0–5 min followed by a gradient to 100% acetonitrile at 45 min. Mass spectrometry was performed under the same conditions as HPLC analysis. Mass spectra were monitored in a mass range of m/z 200–1000 on a LCQ mass spectrophotometer equipped with an atmosphere pressure chemical ionization (APCI) interface (Thermo Finnigan, USA) as described previously [12].

Acknowledgments

The authors gratefully acknowledge support from the Defense Advanced Research Projects Agency (DARPA-BIOS N66001-02-1-8928) and the David and Lucile Packard Foundation (grant 2001-18996).

Received: November 11, 2004
Revised: February 14, 2005
Accepted: February 15, 2005
Published: April 21, 2005

References

1. Mijts, B.N., Lee, P.C., and Schmidt-Dannert, C. (2004). Engineering carotenoid biosynthetic pathways. Methods Enzymol. *388*, 315–329.
2. Marshall, J.H., and Wilmoth, G.J. (1981). Pigments of *Staphylococcus aureus*, a series of triterpenoid carotenoids. J. Bacteriol. *147*, 900–913.
3. Kleinig, H., Schmitt, R., Meister, W., Englert, G., and Thommen, H. (1979). New C30-carotenoic acid glucosyl esters from *Pseudomonas rhodos*. Z. Naturforsch. *34c*, 181–185.
4. Takaichi, S., Inoue, K., Akaike, M., Kobayashi, M., Oh-oka, H., and Madigan, M.T. (1997). The major carotenoid in all known species of heliobacteria is the C30 carotenoid 4,4'-diaponeurosporene, not neurosporene. Arch. Microbiol. *168*, 277–281.
5. Marshall, J.H., and Wilmoth, G.J. (1981). Proposed pathway of triterpenoid carotenoid biosynthesis in *Staphylococcus aureus*: evidence from a study of mutants. J. Bacteriol. *147*, 914–919.
6. Kleinig, H. (1982). On the biosynthesis of C30 carotenoic acid glucosyl esters in *Pseudomonas rhodos*. Analysis of *car*-mutants. Z. Naturforsch. *37c*, 758–760.
7. Wieland, B., Feil, C., Gloria-Maercker, E., Thumm, G., Lechner, M., Bravo, J.M., Poralla, K., and Gotz, F. (1994). Genetic and biochemical analyses of the biosynthesis of the yellow carotenoid 4,4'-diaponeurosporene of *Staphylococcus aureus*. J. Bacteriol. *176*, 7719–7726.
8. Bouvier, F., Suire, C., Mutterer, J., and Camara, B. (2003). Oxidative remodeling of chromoplast carotenoids: identification of the carotenoid dioxygenase CsCCD and CsZCD genes involved in Crocus secondary metabolite biogenesis. Plant Cell *15*, 47–62.
9. Bouvier, F., Dogbo, O., and Camara, B. (2003). Biosynthesis of the food and cosmetic plant pigment bixin (annatto). Science *300*, 2089–2091.
10. Escribano, J., Alonso, G.L., Coca-Prados, M., and Fernandez, J.A. (1996). Crocin, safranal and picrocrocin from saffron (*Crocus sativus* L.) inhibit the growth of human cancer cells *in vitro*. Cancer Lett. *100*, 23–30.
11. Schmidt-Dannert, C., Umeno, D., and Arnold, F.H. (2000). Molecular breeding of carotenoid biosynthetic pathways. Nat. Biotechnol. *18*, 750–753.
12. Lee, P.C., Momen, A.Z., Mijts, B.N., and Schmidt-Dannert, C. (2003). Biosynthesis of structurally novel carotenoids in *Escherichia coli*. Chem. Biol. *10*, 453–462.
13. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.
14. Peterson, J.D., Umayam, L.A., Dickinson, T., Hickey, E.K., and White, O. (2001). The comprehensive microbial resource. Nucleic Acids Res. *29*, 123–125.
15. Lee, P.C., Mijts, B.N., and Schmidt-Dannert, C. (2004). Investigation of factors influencing production of the monocyclic carotenoid torulene in metabolically engineered *Escherichia coli*. Appl. Microbiol. Biotechnol. *65*, 538–546.
16. Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. *41*, 95–98.
17. Cadwell, R.C., and Joyce, G.F. (1992). Randomization of genes by PCR mutagenesis. PCR Methods Appl. *2*, 28–33.