Directed Evolution of Biosynthetic Pathways to Carotenoids with Unnatural Carbon Backbones

Thesis by

Alexander Vincent Tobias

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2006

(Defended August 19, 2005)

ACKNOWLEDGMENTS

The opportunity to attend graduate school at Caltech has been a precious gift that I will treasure for the rest of my life. Working at this world-class institute with such talented, intelligent, motivated, and conscientious colleagues has contributed to my professional and personal development in ways I could never have imagined at the outset.

To my advisor, Frances Arnold, thank you for giving me the guidance and opportunities to develop and mature as a scientist. I have never encountered anyone who can cut directly through to the essence of an idea or concept the way you can, and I hope that some of your keen scientific vision and insight has rubbed off on me. Thank you also for always being truthful. It wasn't easy knowing that something I said, did, or wrote didn't satisfy you or come across the right way, but on the other hand, when you told me that my work was "excellent" or "just right," I knew you meant it and that really made me feel like I had succeeded.

To my other committee members, Anand Asthagiri, Linda Hsieh-Wilson, and Dianne Newman, thank you all for serving on my candidacy and thesis committees and for your insightful questions and comments.

I would like to thank the Natural Sciences and Engineering Research Council of Canada for the two postgraduate scholarships I was awarded, and the United States National Science Foundation for a grant supporting this research. I am appreciative not only of the financial contributions from these organizations, but also for the associated expressions of belief in the value of this research.

I would not have gotten very far in characterizing the carotenoids my cultures have synthesized without the dedicated assistance of Mona Shahgholi and Nathan

Dalleska. Thank you both for the hours of time you spent helping me with mass spectrometry, for answering my questions and critiquing my hypotheses, and most of all, for helping me to collect key data for this thesis.

Manish Raizada and Hyman Hartman contributed valuable comments that helped us to make the review paper reproduced in Chapter 1 significantly more focused and lucid. Gerhard Sandmann deserves many thanks for supplying me with valuable experimental guidance and eminently appropriate literature at just the right moments. I am also grateful to George Britton for his expert advice on the extraction and analysis of carotenoids from biological material. Shinichi Takaichi provided helpful assistance with the naming of the new carotenoids reported in Chapter 3.

Most of my fondest memories of Caltech involve the many wonderful students, postdocs, and visitors who have spent time in the Arnold group during my tenure here. I want to thank every one of these group members for contributing to the Arnold lab's wonderful atmosphere of interdisciplinary collaboration, for advice and ideas, and for good times at parties and other social functions.

In particular, I am indebted to Daisuke Umeno, who was my mentor and collaborator on the carotenoids project and who taught me how to clone genes, build plasmids, and "do" directed evolution. Daisuke, I am ever so grateful for your friendship, instruction, camaraderie, supportiveness, and truly unique sense of humor.

I wish to specially acknowledge several other members of the Arnold group: to Adam Hartwick, thank you for your friendship and for your ability to discuss carotenoids and their biosynthesis with me as only someone with your experience could. To Michelle Giron and Brenda Parker, thank you both for your contributions to the carotenoids

subgroup and for being such great students to mentor. To Joff Silberg, thank you for your friendship, advice, and the numerous conversations we had about everything from science and politics to family and food. To Michelle Meyer, thank you for being a wonderful officemate, for your friendship, and for all our conversations over the years. To Matt Peters, thank you for being so willing to answer (or attempt to answer) my almost incessant barrage of chemistry-related questions. To Jorge Rodriguez, thank you for your friendship and our many discussions of scientific and personal import. To Allan Drummond, thank you for your contagious passion for science and our fruitful and enlightening discussions. And to Geethani Bandara, thank you for keeping the lab running smoothly, for all your friendly and helpful advice, and for hosting such great parties at your home.

In contrast to my final year, my first four years at Caltech were not spent entirely in the lab. I am grateful to the members of the Caltech Beavers hockey team for their camaraderie and constant supply of memorable moments on and off the ice. In particular, I thank Adam Olsen and David Jenkins for their friendship and hundreds of hours of stimulating banter on the way to and from Sylmar. I would also like to thank my colleagues who served with me on the Caltech Graduate Student Council. My “GSC education” was highly complementary to my research education at Caltech, and I am appreciative of my fellow board members for helping me to improve my “soft” and leadership skills.

When I think about the path of my professional life, two of my teachers stand out as having profoundly influenced my choice to pursue a career in engineering. Charles Ledger and Fraser Simpson, my junior high school mathematics teachers, were truly in a

class of their own. They ignited my passion for math and problem-solving with their contagious enthusiasm and love for the subject as well as their innovative teaching methods (dissatisfied with the textbooks of the day, they designed their own curriculum and somehow convinced the school board to let them teach it). I might have pursued the same career path had I not been taught by Mr. Simpson and Mr. Ledger, but I doubt I would have done so with quite the same vigor.

To J.R., I am much obliged to you for your sage guidance and for improving my self-awareness. Your attentive ear and knowledgeable advice have assisted me through many of life's trials and have helped me to tame my worries. I am no doubt a better person to myself and others because of you.

My parents, Lili and David Tobias, deserve special recognition for their love and guidance and for encouraging me to explore my curiosity about the world in myriad ways. Mom and Dad, the lessons you worked so hard to teach me have served me well: finish your homework before you play, do it right the first time, respect other people and their property, help others in need, and wait for things to go on sale. Thank you for giving me just the right amount of freedom and responsibility when I was growing up so that I could learn to be independent but not get into too much trouble.

To Mikael, thank you for being the best brother anyone could ask for. I think part of our special bond is rooted in our choice of careers, which are quite different but also appealing to both of us. This allows us to live vicariously through each other, so that I get a taste of the life of a musician and you can experience a bit of life as a scientist. Although we haven't lived in the same city for some time now, I can't imagine life without you and I am grateful for our e-mails, telephone calls, and visits.

To my dearest Jill, thank you for your boundless love, compassion, understanding, and support. I feel the bond of our marriage grow stronger every day, and to be your husband is the ultimate privilege. Thank you for cheering me up when my experiments weren't working, for bringing me dinner at the lab when I could take only a short break, and for waiting for me all those times when I needed "just another 15 minutes" before I could come home. Thank you also for celebrating my successes and milestones with me, but more importantly, for insisting that these moments be celebrated when I didn't feel up to it at first. For all you have contributed, I feel like we have completed this thesis together. As we embark on the next chapter of our lives, I look forward to all the new experiences, moments, and memories that I will have the honor of sharing with you.

ABSTRACT

Over the course of evolution, nature continually discovers new small molecules through the alteration of biosynthetic enzymes and pathways by mutation and gene transfer. Hundreds of these natural products have proven indispensable to medicine, culture, and technology, greatly contributing to increases in the length and quality of human lives. Chemists have found that the "chemical space" surrounding natural products is especially rich in functional molecules, and synthesis of natural product analogs has uncovered many with new or improved properties.

Inspired by nature's search algorithm, we and others have conducted our own evolution of carotenoid biosynthetic pathways in the laboratory. Chapter 1 comprehensively reviews the motivations, accomplishments, and challenges of this research area as of early 2005, and describes in detail how biosynthetic routes to dozens of new carotenoids have been established.

To expand the number of carotenoid backbones beyond the C_{30} and C_{40} carbon scaffolds that give rise to the ~700 known natural carotenoids, we subjected a carotenoid synthase, the enzyme responsible for carotenoid backbone synthesis, to directed evolution. Chapter 2 describes the evolution of the C_{30} carotenoid synthase CrtM from *Staphylococcus aureus* for the ability to synthesize C_{40} carotenoids. This work also resulted in novel carotenoids with C_{35} backbones. We later found that some of the CrtM mutants generated in this laboratory evolution experiment, as well as several second-generation variants, are also capable of synthesizing unnatural C_{45} and C_{50} carotenoid backbones when supplied with appropriate prenyl diphosphate precursors.

Chapter 3 describes the creation of full-fledged pathways to carotenoid pigments based on the C_{45} and C_{50} scaffolds. Coexpression of the carotenoid desaturase CrtI from *Erwinia uredovora* resulted in the biosynthesis of at least 10 new C_{45} and C_{50} carotenoids with different systems of conjugated double bonds. We also present evidence of an unnatural asymmetric C_{40} carotenoid pathway beginning with the condensation of farnesyl diphosphate (FPP, C_{15}PP) and farnesylgeranyl diphosphate (FGPP, C_{25}PP). In addition to clarifying how CrtM and CrtI achieve their product specificities, this work also sheds light on the molecular mechanisms used by evolution to access new chemical diversity and the selective pressures that have shaped natural product biosynthesis.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER 1

Diversifying Carotenoid Biosynthetic Pathways in the Laboratory

Material from this chapter appears in *Diversifying Carotenoid Biosynthetic Pathways by Directed Evolution*, Daisuke Umeno‡, Alexander V. Tobias‡, and Frances H. Arnold, Microbiology and Molecular Biology Reviews, 69(1): 51-78 (2005) and is reprinted with permission from the American Society for Microbiology.

‡ Co-first authors

SUMMARY

Microorganisms and plants synthesize a diverse array of natural products, many of which have proven indispensable to human health and well-being. Although many thousands of these have been characterized, the space of possible natural products—those that could be made biosynthetically—remains largely unexplored. For decades, this space has largely been the domain of chemists, who have synthesized scores of natural product analogs and have found many with improved or novel functions. New natural products have also been made in recombinant organisms via engineered biosynthetic pathways. Recently, methods inspired by natural evolution have begun to be applied to the search for new natural products. These methods force pathways to evolve in convenient laboratory organisms, where the products of new pathways can be identified and characterized in high-throughput screening programs. Carotenoid biosynthetic pathways have served as a convenient experimental system with which to demonstrate these ideas. Researchers have mixed, matched, and mutated carotenoid biosynthetic enzymes and screened libraries of these "evolved" pathways for the emergence of new carotenoid products. This has led to dozens of new pathway products not previously known to be made by the assembled enzymes. These new products include whole families of carotenoids built from backbones not found in nature. This chapter details the strategies and specific methods that have been employed to generate new carotenoid biosynthetic pathways in the laboratory. The potential application of laboratory evolution to other biosynthetic pathways is also discussed.

INTRODUCTION

Nature is a masterful and prolific chemist. The ever-expanding catalog of natural products from plants and microbes currently stands at more than 170,000 different compounds (35). This collection has profoundly influenced human well-being. Natural products have been used to treat pain, infections, and disease for thousands of years; their equally ancient and diverse uses as colorants, spices, fragrances, aphrodisiacs, cosmetics, and toxins have been fundamental to human culture and development. This trend shows no signs of subsiding. Of the 877 new small-molecule pharmacological entities introduced worldwide between 1981 and 2002, 61% can be traced to natural products (142). A similar proportion (59%) of the 137 small molecules in phase II or III clinical trials for cancer treatment (as of June 2003) are natural products or closely related compounds (172). Natural products and their derivatives therefore remain an essential component of the pharmaceutical industry (172). Given the wealth of functional small molecules in nature, it is perhaps not surprising that so much effort today goes into prospecting for new ones or modifying existing ones. Academic and industrial laboratories, often in massive screening programs, continue to isolate and characterize promising new compounds from biological extracts. In addition, molecular scaffolds inspired by natural products are increasingly being used as the basis for synthesizing combinatorial chemical libraries that can be screened for protein-binding or other biochemical activities (30).

Although impressive in number, the known products of natural biosynthetic pathways account for but a tiny fraction of the structures that could be produced. This essentially infinite space of possible functional molecules represents an irresistible

frontier for those seeking new, bioactive compounds. For example, only 1/10 of the 877 new small-molecule pharmacological entities mentioned above are bona fide natural products—most are derivatives of natural products, synthetic compounds with natural product-derived moieties, or natural-product mimics (142). The chemical space surrounding natural products is particularly rich in biologically functional molecules, and structurally related compounds can serve human purposes even better than natural products themselves. Researchers are now beginning to explore some of these "nearly natural" products by mixing, matching, and mutating biosynthetic genes from diverse sources in organisms such as *Escherichia coli* and screening for production of novel metabolites. In this chapter we illustrate how pathways can be "evolved" rapidly in the laboratory to generate new natural products. Used in this way, directed laboratory evolution is a powerful tool for discovering new pathways and for *in vivo* combinatorial chemistry of natural products, i.e., "combinatorial biosynthesis."

In addition to providing access to novel metabolites, laboratory evolution studies can offer insights into the natural evolution of metabolic pathways and their constituent enzymes. Comparative studies of homologous enzymes can readily identify similarities such as conserved residues and folds. However, identifying the exact genetic changes responsible for differences in the specificity, stability, or other properties of related enzymes is much more difficult. When comparing sequences that have diverged over millions of years, it is almost impossible to distinguish the handful of adaptive mutations from the plethora of differences that reflect evolutionary drift. In contrast, evolution in the laboratory allows us to capture new metabolic pathways in the act of emerging and permits the identification and analysis of the molecular events that gave rise to the new

product distributions.

With carotenoid biosynthesis as our model, we describe how pathway engineers have used laboratory evolution experiments to diversify biosynthetic pathways and what they have learned from those efforts. This work has demonstrated the remarkable ability of carotenoid biosynthetic enzymes to evolve and acquire new specificities and has illuminated pathway features that facilitate the creation of new natural products based on "unnatural" molecular scaffolds. Although evolutionary pathway engineering is in its infancy, we show that it is capable of generating whole new families of natural product analogs. Some of these compounds may await discovery in nature, while others will probably never be found by natural evolution.

Overview of carotenoid biosynthetic pathways

The carotenoids are a subfamily of the isoprenoids and are among the most widespread of all natural products. Carotenoids are responsible for many of the colors of animals, plants, and microorganisms and play important biological roles as accessory light-harvesting components of photosynthetic systems, photo-protecting antioxidants, and regulators of membrane fluidity. As a whole, the natural C_{30} and C_{40} carotenoid biosynthetic pathways can be organized in a tree-like hierarchy (**Figure 1.1**), with the trunk representing the initial, universal step of backbone synthesis catalyzed by a carotenoid synthase and the outward-emanating branches and sub-branches collectively denoting the various downstream modification steps seen in different species (8). The hundreds of known carotenoid biosynthetic enzymes fall into relatively few classes, based on the types of transformation they catalyze (isoprenyl diphosphate synthases, carotenoid synthases, desaturases, cyclases, etc.). The enzymes responsible for various

versions of a basic transformation are related by evolution (111, 163, 176): the pathways diverged upon accumulation of mutations that altered enzyme specificity. Thus, a small number of basic transformations, subtly modified in different species, collectively lead to a large number of different products, although only a handful of carotenoids (usually a few and almost always fewer than ~10) are made by any one organism.

C_{40} carotenoids are made in thousands of plant and microbial species, starting with the synthase-catalyzed condensation of two molecules of geranylgeranyl diphosphate (C_{20}PP) to form phytoene (**Figure 1.1**). (Most phytoene synthases produce the 15*Z* isomer of phytoene (87).) Different types and levels of modification of this C_{40} backbone account for the majority of known carotenoids. C_{30} carotenoid pathways starting with the condensation of two molecules of farnesyl diphosphate (C_{15}PP) to form (15*Z*)-4,4'-diapophytoene (also called dehydrosqualene) are much less widespread, having been found only in a small number of bacteria such as *Staphylococcus*, *Streptococcus*, *Heliobacterium*, and *Methylobacterium* species (56, 100, 101, 129, 130, 198, 205, 206). These initial C_{30} and C_{40} condensation products are dehydrogenated in a stepwise manner by desaturase enzymes, which represent the next important branch points for pathway diversification (70). In bacteria and fungi, a single desaturase catalyzes the entire sequence of carotenoid desaturation steps (32). Carotenoid desaturation in plants, algae, and cyanobacteria (not covered in this chapter) is accomplished by two distinct desaturases and a carotenoid isomerase (7, 31, 72, 86, 131). Desaturases from bacteria such as *Pantoea* spp. (referred to in this chapter by their former *Erwinia* species names, which are more commonly used than the current approved names in the carotenoid field) also catalyze isomerization of the carotenoid's central

double bond, giving all-*E* *(trans)* desaturated products (65). With a few notable exceptions (157, 200), the desaturation level of the carotenoid backbone is well defined for each organism. In many organisms, desaturation is followed by cyclization, catalyzed by a β- or ε-cyclase and leading to carotenoids with one or two cyclized ends. A variety of further enzyme-catalyzed transformations that can include ketolation, hydroxylation, glycosylation, and oxidative cleavage act on substrates derived from the C_{30} or C_{40} backbones to produce the catalog of more than 700 known carotenoids (82).

Some organisms synthesize modified C_{40} carotenoids with additional or fewer isoprenyl units. For example, *Corynebacterium glutamicum* synthesizes "homocarotenoids" such as flavuxanthin (C_{45}) and decaprenoxanthin (C_{50}) via enzymatic prenylation of lycopene (**Figure 1.1**) (109, 112). In contrast, several fungi produce C_{35} "apocarotenoids" such as neurosporaxanthin via oxidative cleavage of monocyclic C_{40} carotenoids (**Figure 1.1**) (12, 14, 164, 215). Despite their non-C_{40} structures, these natural homo- and apocarotenoids are, from a biosynthetic perspective, part of the C_{40} family, since their biosynthesis proceeds via the C_{40} carotenoid backbone phytoene.

Isoprenoids and isoprenyl diphosphate building blocks are found in all organisms. *E. coli* has three isoprenyl diphosphate synthases: a farnesyl diphosphate (C_{15}PP) synthase (67, 68), an octaprenyl diphosphate (C_{40}PP) synthase (154), and a *Z*-type (see below) undecaprenyl diphosphate (C_{55}PP) synthase (94). A C_{30} carotenoid pathway can be built in *E. coli* branching directly from the organism's endogenous C_{15}PP supply (226). However, since *E. coli* lacks a C_{20}PP synthase function, a geranylgeranyl diphosphate (C_{20}PP) synthase gene must be heterologously expressed along with the other carotenoid biosynthetic genes in order to make C_{40} carotenoids in this organism.

Most of the enzymes involved in carotenoid biosynthesis are membrane associated, and the hydrophobic intermediates and products of carotenoid pathways partition to cytoplasmic or organelle membranes, depending on the type of organism (29). It is not known whether the carotenoid biosynthetic enzymes associate into complexes. Over the years, evidence for the existence of multienzyme complexes that function as "assembly lines" for carotenoid synthesis has been presented (summarized in reference (32)). Some of the most convincing evidence has come from experiments performed in the laboratory of Cerdá-Olmedo with the fungus *Phycomyces blakesleeanus* (6, 40, 58, 138). On the other hand, the fact that carotenoid enzymes from different organisms are readily combined to generate functional pathways in heterologous hosts is evidence that the formation of a specific enzyme complex is not a prerequisite for carotenoid biosynthesis. Throughout this chapter, we refer to carotenoid-synthesizing enzyme complexes when they provide a possible alternative explanation for particular observations from our laboratory and others'.

Carotenoids as a model system for laboratory pathway evolution

Carotenoid biosynthesis is particularly well suited as an experimental system for studying pathway evolution in the laboratory. Claudia Schmidt-Dannert recognized that the following features would facilitate such experiments when she first proposed to evolve these pathways as a visiting researcher in this laboratory in 1998.

Carotenoids are diverse, and their biosynthetic pathways are representative of other secondary metabolic pathways. Hundreds of carotenoids are known in nature. They are made using a few basic transformations, the earliest of which are highly conserved. The reactions involved in these early steps, such as backbone formation by carotenoid

synthases and cyclization by carotenoid cyclases, are mechanistically similar to other reactions in isoprenoid biosynthesis. Modification steps later in the carotenoid pathways cover a wide range of biotransformations, many of which appear in other natural products. Thus, experience with carotenoid pathway evolution can be applied to diversifying other biosynthetic pathways.

Carotenoid biosynthetic enzymes are highly portable. The carotenoid pathway emerges directly from the central isoprenoid pathway that exists in all organisms. Thus, in principle, any organism can supply the precursors required for an engineered carotenoid pathway. Carotenoid biosynthetic genes can be expressed in a wide range of organisms to extend or redirect an existing pathway (69, 70, 105). For instance, a bacterial phytoene desaturase (CrtI) introduced into cyanobacteria (227) and tobacco plants (136) elevated their resistance to the herbicide norflurazon (known to block plant-type phytoene desaturases). Plastid-targeted introduction of an algal β,β-carotene ketolase resulted in transgenic tobacco plants that accumulated astaxanthin and changed the color of the plant's nectory tissue from yellow to red (128). The highly publicized introduction of the β,β-carotene pathway into rice resulted in substantial β,β-carotene levels in the normally carotenoid-free endosperm and therefore greater nutritional value for this food staple (234). Phytoene-producing human cells (HeLa, NIH 3T3, etc.) were created by the addition of a bacterial phytoene synthase and showed increased tolerance to oxidative shock as well as a slower rate of H-ras-induced carcinogenesis (143, 144, 179). Almost all carotenoid biosynthetic genes cloned to date, including those from plants, can be functionally expressed in *E. coli* (119, 175, 177, 180), as can animal oxidative cleavage enzymes (98, 123, 158, 220, 231). The "portability" of these enzymes greatly facilitates

the assembly and evolution of novel biosynthetic pathways.

Evolution of carotenoid pathways can be tracked visually. Carotenoids are natural pigments, and their characteristic colors are ideal for product-based high-throughput screening. Scientists have made good use of this feature throughout the history of carotenoid research. Experiments involving the generation and analysis of mutant strains exhibiting altered colors have elucidated biosynthetic routes in various organisms (12, 15, 18, 41, 66, 73, 113, 120, 130, 155, 162, 173, 178, 196, 201, 216). Many carotenoid biosynthetic genes were cloned based on their ability to confer or alter color development in *E. coli* (53, 55, 133, 134, 160, 197). Carotenoid titers are an indicator of precursor levels and can be used to evaluate and tune upstream (isoprenoid) pathways (189, 223, 224). The colors generated in *E. coli* colonies provide a facile screen for new pathway products in a laboratory evolution experiment.

Carotenoids are valuable in their own right. The polyene chromophores of carotenoids, which absorb light in the 400-550 nm range, provide the basis for their characteristic yellow-to-red colors and their ability to quench singlet oxygen (5, 74, 79, 95, 103, 137, 218, 230). Carotenoids act as antioxidants *in vivo* (170, 183, 210), and many beneficial effects of carotenoids on human health have been reported, ranging from cancer prevention and tumor suppression to upregulation of immune function, reduction of the risk of coronary heart disease or age-related degeneration, and cataract prevention (19, 21, 43, 48, 83, 108). As natural pigments, carotenoids are used as colorants in foods, cosmetics, and flowers (117, 128, 209). Given these diverse and important properties of known carotenoids, we can assume that novel compounds of this family will also possess interesting biological or chemical functions.

Evolvable pathways and enzymes

The diversity of carotenoids and biosynthetic enzymes that exists in nature demonstrates the success of evolution in discovering and optimizing new pathways. Carotenoid pathways are evolvable and understanding the features that lead to this evolvability will allow us to design directed evolution experiments to accomplish the same goals. We define an evolvable pathway as one that can produce many new metabolites after limited genetic change. New metabolites are made following recruitment of new enzymes or functional alteration of existing ones. Such events are rare, and evolvable pathways must be particularly effective at exploiting these rare events.

In general, evolvable pathways appear to contain enzymes that are "locally specific" (32). These enzymes are not unspecific; rather, they recognize a particular structural motif common to a variety of possible substrates. Thus, if a change upstream of a locally specific enzyme generates a novel compound, there is a good chance that the enzyme will metabolize this compound further (as long as it retains the required motif) and generate a new derivative. More such enzymes further downstream will accept the new derivative(s) and produce yet more new metabolites. A carotenoid desaturase, for example, might need at most a few mutations in order to catalyze double-bond formation on a new carotenoid-like backbone. Locally specific modifying enzymes further downstream would, with high probability, accept these desaturated substrates, generating still more new compounds.

Pathway structure also contributes to evolvability: highly branched pathways can propagate discoveries along many routes. Pathway branches occur where one substrate is potentially converted to multiple products by the action of one or more enzymes. The

existence of nested branches geometrically increases the number of new products that can be synthesized from a new metabolite produced upstream, provided that the downstream enzymes accept the new substrates. Evolvable and consequently diverse, carotenoid pathways are "bushy," with multiple products possible from each basic transformation. Product possibilities are dictated by the nature of the chemical reaction and substrate, and ultimately by the specificity of the enzyme that controls that branch point. For example, desaturation of a carotenoid backbone can produce a number of different products—the one(s) produced by any given organism reflect the specificity of its particular desaturase in its environment. That specificity can change upon mutation and can open new pathway branches.

Pathways constructed entirely of locally specific enzymes would not serve a host organism well in terms of regulation and production of tried-and-true metabolites. Thus, we expect evolvable pathways to contain more specific enzymes as well. To maximize the ease of exploring new pathways while preserving key metabolic processes, we would expect these specific enzymes to be located at the earliest steps of a pathway. Here, they can serve as molecular gatekeepers, allowing only certain primary metabolites to enter but not impeding the discovery of new structures within the pathway. Manipulating these gatekeeper enzymes to admit different substrates should be an effective way to generate whole new families of natural products.

Thus, features contributing to the evolvability of a biosynthetic pathway include the use of locally specific enzymes and biotransformations that contribute to pathway branching. More-specific gatekeeper enzymes limit what flows through the pathway, providing necessary insurance against chemical chaos. In the end, however, pathways are

evolvable because their component enzymes are themselves evolvable, changing properties such as substrate and product specificity readily upon mutation. Evolvable enzymes should also exhibit multiple mutational routes to a given altered phenotype and may be comparatively robust to mutation. The former property increases the chances that new phenotypes will emerge, while the latter property allows "scanning" of similar sequences without loss of function. Perhaps surprisingly, it is still an open question whether promiscuous or locally specific enzymes are more evolvable than enzymes that are highly specific. As we detail in this chapter, as little as a single amino acid substitution can change the specificity of highly specific carotenoid biosynthetic enzymes.

An evolvable pathway, then, would consist of an evolvable gatekeeper enzyme (or set of enzymes) coupled with locally specific downstream enzymes. Such an arrangement maximally exploits mutations in the gatekeeper enzyme, converting a single newly discovered molecule into a potentially large number of new metabolites. Several lines of evidence support the view that the enzymes involved in carotenoid biosynthesis are highly evolvable and are embedded in pathways with such evolvable structures. This chapter represents the first attempt at gathering together this evidence, which collectively paints a picture of carotenoid biosynthetic pathways as dynamic systems capable of exploring a diversity of product structures much greater than is seen in nature. We now examine the various ways in which this evolvability has been exploited in the laboratory to explore and extend the product diversity of these pathways.

APPROACHES TO ENGINEERING NOVEL BIOSYNTHETIC PATHWAYS

New biosynthetic pathways can be created by recruiting foreign catalytic machinery that adds a new branch; such enzyme recruitment happened frequently during the evolution of natural-product pathways (20, 27, 116). A pathway can also diversify when the function of an existing enzyme is altered by mutation or other genetic change. The wide functional diversity that exists within the evolutionarily related carotenoid pathways is evidence that enzyme evolution has played a central role in creating chemical diversity, primarily through modification of substrate and product specificity. In the laboratory, researchers can now employ the same strategies of enzyme recruitment and modification to create new biosynthetic pathways. In this chapter, we refer to coexpressing genes from different sources in a recombinant organism as gene assembly. This approach can be distinguished from directed evolution of the component enzymes (or their regulation), where libraries of mutant alleles are expressed along with other pathway genes in host cells, which are then screened or selected for desired properties. These two approaches generate novel compounds by mimicking the major modes of natural pathway evolution: gene transfer and divergence. Thus, the two approaches complement one another, and we consider both to fall under the umbrella of laboratory pathway evolution. In this section, we discuss efforts to create new carotenoid biosynthetic pathways by using gene assembly, directed enzyme evolution, and combined approaches.

Engineering pathways by gene assembly

Assembling carotenoid biosynthetic genes from different organisms and expressing these “hybrid” pathways in a recombinant host has generated functional routes

to rare or novel carotenoids (**Figure 1.2**) (4, 5, 13, 69, 105, 118, 192, 199, 211). Where gene assembly alone successfully yielded a new pathway branch or extension, the recruited enzyme(s) usually catalyzed transformations later in the pathway. These downstream enzymes are locally specific and can accept a range of substrates. Various unusual acyclic carotenoids were obtained by assembling carotenogenic genes from *Pantoea ananatis* (*Erwinia uredovora*) and *Rhodobacter capsulatus* (4, 5). Coexpression of the C_{20}PP synthase CrtE and the phytoene synthase CrtB from *E. uredovora* along with various combinations of three different carotenoid desaturases, a carotenoid hydratase, a β-cyclase, and a hydroxylase resulted in the production of four previously unidentified carotenoids, including 1-OH-3',4'-didehydrolycopene with potent antioxidant activity (5). In another example, addition of CrtW, a β-end ketolase from *Paracoccus* sp. strain MBIC1143 ("*Agrobacterium aurantiacum*"), extended the zeaxanthin β-D-diglucoside pathway from *Erwinia*, leading to synthesis of novel compounds, astaxanthin β-D-diglucoside and adonixanthin 3'-β-D-glucoside (**Figure 1.2**) (235). As we show in later sections, other downstream carotenoid-modifying enzymes are also locally specific and should be similarly useful for "combinatorial biosynthesis" of novel carotenoids by gene assembly. We predict that systematic combinatorial assembly of these genes will result in the biosynthesis of many hundreds of new carotenoids.

Assembling locally specific enzymes in a recombinant organism can sometimes result in an interconnected "matrix" pathway that leads to numerous carotenoid structures, as illustrated in **Figure 1.3** (135). However, many of the intermediates or end products may not be accessible due to imbalances in enzyme expression, activity, or specificity. Thus, a few dominant compounds may be produced, while other (possible) ones remain

hidden (but poised to appear under different conditions). Furthermore, it can be difficult to predict the relative abundances of the different metabolites. For example, various attempts to construct a pathway for astaxanthin synthesis in a heterologous host resulted in quite different proportions of astaxanthin and intermediates (76, 135). In another example, introduction of the carotenoid biosynthetic enzymes CrtB, CrtI, CrtY, and CrtZ from *Pantoea agglomerans* (*Erwinia herbicola*) into a desaturase-deleted mutant of *Rhodobacter sphaeroides* merely restored desaturation activity and failed to alter the range of carotenoids produced (85). This occurred despite the introduction of a β-cyclase (CrtY) and a β-end hydroxylase (CrtZ) into a host lacking these catalytic functions. Only when the endogenous neurosporene hydroxylase gene (*crtC*) was deleted was a functional pathway to cyclized and hydroxylated carotenoids realized (69, 70). In this case, the endogenous CrtC in *R. sphaeroides* may have been so active that the desaturase CrtI, cyclase CrtY, and β-end hydroxylase CrtZ introduced from *E. herbicola* could not compete with it for neurosporene. Another possible explanation is that incorporation of CrtC into a carotenoid biosynthetic enzyme complex blocked the additional inclusion of CrtY, CrtZ, and more than three desaturase subunits. These examples illustrate the importance of properly coordinating pathway enzymes in order to create a specific pathway to a desired product. This problem of miscoordination makes it exceedingly difficult to systematically explore all the possible metabolites of a given set of enzymes solely by gene assembly. A major task of pathway engineers will be to coordinate assembled components in order to unmask hidden metabolites. This can be accomplished by directed evolution.

Engineering pathways by directed enzyme evolution

Nature has generated a vast number of biosynthetic genes by mutation, recombination, and natural selection. Variations in properties such as desaturation step number or synthase specificity contribute enormously to the diversity of carotenoid structures seen in nature. In the previous section, we presented examples of how gene assembly alone can lead to novel or unusual carotenoids. However, we can also consider creating enzyme variants with desired specificities (or other properties such as expression level or stability) by evolving them in the laboratory. This approach circumvents the need to find a specific function in nature and reproduce that function in a heterologous host. Instead, enzymes with the desired properties are found by screening libraries of mutants of an enzyme that performs the same type of chemical transformation.

Some enzyme specificities not available in nature become readily available by evolution in the laboratory, opening pathways to new metabolites. Decoupled from the biological constraints imposed by the need to fulfill a particular function throughout the course of evolution, a pathway can in principle explore all chemically accessible products, not just those that are biologically relevant (10). Thus, directed evolution gives us the ability to explore the space of possible chemical products much more rapidly and also more thoroughly than does natural evolution. By evolving biosynthetic enzymes, we can anticipate the discovery of large numbers of new natural products, on the benchtop in convenient laboratory hosts.

Although each directed evolution project is unique, all involve two main steps: making a library of mutants and searching that library for desired properties. The mutant library is typically generated in an error-prone PCR amplification reaction (39, 124)

tuned to generate a certain average point mutation rate or by DNA shuffling, a method that makes new point mutations and recombines existing mutations (193) (many methods are available for making libraries of mutant genes for directed evolution; useful protocols are given in reference (9)). The library of mutated DNA molecules is then subcloned into an appropriate vector for expression with other pathway genes in a convenient host organism such as *E. coli*. Hundreds or thousands of clones are screened in search of (typically) rare ones expressing favorably altered enzymes. A powerful aspect of this process is that it can be iterated—improved variants can be subjected to further rounds of mutagenesis and screening, often leading to further improvements. A sensitive, reliable screen is key to a successful directed evolution experiment: the screen allows the researcher to identify the typically rare mutants with new and interesting properties.

Different strategies exist for evolving metabolic pathways in the laboratory. In an early study, multiple genes in an arsenate degradation pathway were subjected to mutagenesis and recombination all at once (50). This "fully blind" approach makes the most sense when little is known about which genes or non-coding regions (such as ribosomal binding sites) should be mutated to obtain a desired phenotype. However, a significant disadvantage of this approach is that the combinatorial complexity (the size of the library) increases concomitantly with the length of DNA targeted for mutation. Consequently, more clones must be screened to find the rare ones with improved properties. Therefore, when possible, it is attractive to target individual genes for mutation and then screen those mutants in the context of the whole pathway; this can be thought of as a "partially blind" approach. Carotenoid biosynthetic pathways are sufficiently well characterized that investigators generally know which enzyme to target

for mutagenesis in order to achieve a new product distribution; however, the specific mutation(s) needed is unknown. Accordingly, researchers have thus far chosen to extend carotenoid pathway branches in a stepwise manner by evolving one gene at a time and expressing the library of variant alleles together with other wild-type or previously laboratory-evolved pathway genes. Schmidt-Dannert et al. demonstrated the effectiveness of this approach for diversifying the range of carotenoid structures accessible from a given set of enzymes (181).

PROBING THE EVOLVABILITY OF CAROTENOID BIOSYNTHETIC ENZYMES AND PATHWAYS

In the Introduction, we reasoned that highly evolvable pathways would be assembled from locally specific downstream enzymes and (a small number of) more specific "gatekeeper" enzymes. Furthermore, the enzymes themselves should readily alter specificity upon mutation. Modifying a gatekeeper enzyme, accompanied by fine-tuning of downstream enzymes where needed, could allow whole new families of natural products to emerge. In this section, we assess the evolvability of carotenoid biosynthetic pathways and enzymes in view of the (laboratory) data now available.

Directed evolution of key carotenoid biosynthetic enzymes

Isoprenyl diphosphate synthases. Isoprenyl diphosphate synthases (IDSs) catalyze the consecutive condensation of five-carbon isopentenyl diphosphate with allylic diphosphates to generate the precursors to all isoprenoid compounds. A variety of IDSs supply the building blocks for the >33,000 known isoprenoids (89) ranging in size from C_5 to $C_{\sim 2,500}$ (49, 146). IDSs catalyze the 1'–4 condensation of an allylic diphosphate

(C_{5n}PP) with isopentenyl diphosphate, producing the next incremental isoprenyl diphosphate ($C_{5(n+1)}$PP) (**Figure 1.4**). This process continues until the growing isoprenoid chain reaches a certain length, at which point the reaction terminates and the product is released from the enzyme. There exist *E*-type IDSs that synthesize products with precise all-*E (trans)* double-bond stereochemistry as well as *Z*-type IDSs that catalyze the formation of *cis*-double bonds in the growing isoprenyl chain. IDS enzymes are quite specific in their product size and are often classified on this basis. Elucidating the molecular mechanisms of chain length control has been a major scientific interest in the study of this class of enzymes; for reviews, see references (97, 122, 145, 225).

Although IDSs make specific products, minor modifications can change this specificity. Ohnuma et al. randomly mutated the farnesyl diphosphate (C_{15}PP) synthase gene from *Bacillus stearothermophilus* (*bstFPS*), coexpressed this mutant library along with phytoene synthase (*crtB*) and phytoene desaturase (*crtI*) genes in *E. coli*, and looked for colonies that produced the red C_{40} carotenoid, lycopene (152). Most of the colonies were colorless because they lacked the C_{20}PP precursor needed by CrtB to make carotenoids. However, a small fraction of the colonies were pink due to lycopene production; these harbored mutant C_{15}PP synthases capable of synthesizing C_{20}PP. In further work, the same group generated additional variants of BstFPS that made even longer products (150). Some of the double and triple mutants of this enzyme yielded products as large as $C_{>100}$PP. In another study, Ohnuma et al. converted an archaebacterial C_{20}PP synthase into C_{25}PP and C_{30}PP synthases by mutagenesis coupled with complementation screening in a yeast strain deficient in C_{30}PP synthesis (148). In all the above cases, analysis of mutants making larger products revealed substitutions for

smaller amino acids at a position 5 residues upstream of the first aspartate-rich motif (FARM) of the enzyme (151). Aided by the crystal structure of *Gallus gallus* (chicken) C_{15}PP synthase (204), Tarshis et al. independently found a similar result with this enzyme (203). Site-directed mutagenesis at the fifth (F112) and fourth (F113) residues upstream of the FARM showed that smaller amino acids at these positions yielded enzymes that produce longer products. Subsequent work also demonstrated the converse: replacement of chain-length determining residues just upstream of the FARM with larger amino acids leads to shorter products. For example, *Sulfolobus acidocaldarius* C_{20}PP synthase was rationally converted into a C_{15}PP synthase using this approach (149). The C_{15}PP synthases from *B. stearothermophilus* and *G. gallus* were converted to geranyl diphosphate (C_{10}PP) synthases by the same strategy (61, 141).

This strategy of substituting residues just upstream of the FARM with smaller or larger amino acids has also been applied to a plant short-chain (C_{20}PP) IDS (104), heterodimeric medium-chain (C_{30}PP and C_{35}PP) IDSs (78, 81, 238), as well as long-chain (C_{40}PP-C_{50}PP) synthases (75, 153) to alter the product chain length. In several cases, substitution of residues lining the substrate binding pocket but distant from the FARM (or even located on the opposite subunit in heterodimeric enzymes) has also modified the product chain length (75, 77, 80, 92, 96, 148, 152, 238, 239). Many of these substitutions enhance the effects of mutations upstream of the FARM, leading to even greater changes in the product chain length.

Several C_{15}PP synthases have some degree of substrate tolerance (57, 107) and can be used for stereospecific synthesis of various insect hormones and related analogs (102, 106, 126). BstFPS was found to accept a number of nonnatural substrates (140),

and its substrate range was further broadened by substitutions at Y81, previously described as a chain length-determining residue (151) (see above). The Y81A variant of BstFPS, for example, accepts various isoprenyl diphosphate analogs with ω-oxygen atoms in their prenyl chains as substrates, leading to the enzymatic synthesis of butterfly hair pencil pheromone and analogs (125). In theory, all of these isoprenoid diphosphate analogs could serve as building blocks for the enzymatic synthesis of carotenoid analogs.

IDSs from diverse species have exhibited evolvability in the laboratory. Although the wild-type IDSs are quite specific with respect to product size, minor genetic changes were found to profoundly alter the number of elongation steps and therefore the size of the products formed by these enzymes. The residues that were found to change the specificity of these IDSs upon mutation form a key part of the reaction pocket that accommodates the elongating isoprenoid chain, and it appears that product specificity is largely dependent on the size of this pocket (61, 203). Inspection of the data from the reports describing the modification of product specificity reveals that, in general, the mutant IDSs have a broadened rather than a shifted product range. In most cases, mutants synthesize multiple new products as well as the old one. In contrast, wild-type IDSs, although related by evolution, are very specific with respect to the length of their products. We now know that a single amino acid substitution can broaden the product range of an IDS. However, we do not know the degree of genetic change required to completely shift IDS product specificity such that it synthesizes only isoprenyl diphosphates of a new length.

Carotenoid synthases. Carotenoid synthases catalyze the synthesis of a carotenoid backbone from isoprenyl diphosphate precursors. Various phytoene (C_{40})

synthases have been isolated from organisms ranging from bacteria to higher plants. Only one C_{30} carotenoid synthase has been cloned and sequenced to date: *crtM* from *Staphylococcus aureus* (226).

It has long been recognized that the enzyme-catalyzed reaction leading to formation of the carotenoid backbone is very similar to the reaction catalyzed by squalene synthase (SqS), the first committed enzyme in cholesterol biosynthesis (32). Carotenoid synthases and SqS employ the same mechanism, and SqS produces the C_{30} carotenoid backbone if deprived of NADPH (88, 202, 237). Carotenoid synthases and SqS also share several highly conserved domains, and it is likely that the two have a common evolutionary origin (176).

The biosynthesis of carotenoid backbones (and squalene) has proven to be a complex process (22, 88, 89) (**Figure 1.5a**). The reaction proceeds in two distinct steps. The first consists of abstraction of a diphosphate group from a prenyl donor followed by 1-1' condensation of the donor and acceptor and loss of a proton to form a stable cyclopropyl intermediate. In the second step, the cyclopropyl intermediate is rearranged, the second diphosphate group is lost, and the resultant carbocation is quenched after the loss of another proton (**Figure 1.5a**). Biochemical (156) and structural (159) studies have shown that these two subreactions occur in physically distinct sites in the enzyme (**Figure 1.5b**).

Carotenoid synthases are quite specific with respect to product size, which reflects a strong preference for prenyl diphosphate substrates of a particular length. The C_{40} carotenoid synthase CrtB does not detectably accept $C_{15}PP$ as a substrate, while the C_{30} synthase CrtM accepts two molecules of $C_{20}PP$ to form C_{40} carotenoids much less

efficiently than it accepts two molecules of C_{15}PP to form C_{30} carotenoids (163, 214). Even though various longer and shorter isoprenyl diphosphates are made by different organisms (146), no known carotenoids are derived from them. Terpene synthases also utilize specific precursors (42), although product specificity can be quite modest (190).

Carotenoid synthases, positioned early in the biosynthetic pathway and possessing the ability to discriminate between isoprenyl diphosphate substrates with different numbers of isoprene units, appear to play the role of pathway gatekeeper, particularly in organisms in which multiple potential substrates are available. Indeed, as we illustrate, downstream carotenoid biosynthetic enzymes are less specific than the synthases and appear to recognize only a particular motif of the substrate. In the Introduction, we reasoned that gatekeeper enzymes should be attractive targets for pathway diversification. If carotenoid synthases could be engineered to accept different prenyl substrates and synthesize carotenoid backbones from them, the locally specific downstream enzymes might accept these new backbones and generate the corresponding metabolites. This would give rise to whole new pathway families. Compared to the IDSs, however, little is known about the molecular basis for the substrate and product specificity of carotenoid synthases.

Following the lead of those who studied the IDSs (96, 122, 125, 141, 147, 148, 150-152, 203, 238, 239), we probed the evolvability of size specificity in the carotenoid synthases by looking for mutants that complement a non-native pathway. Using error-prone PCR to perform random mutagenesis, expressing the mutant genes with the remaining genes necessary to produce lycopene in *E. coli*, and screening to find red colonies, we identified single amino acid substitutions in the *S. aureus* C_{30} carotenoid

synthase CrtM (F26L or F26S) that confer the ability to make the C_{40} backbone just as efficiently as CrtB (Chapter 2, (214)). Repeating this experiment using a PCR method designed to guarantee that the libraries would be free from mutation at F26 allowed us to uncover two additional amino acid substitutions, W38C and E180G, that also confer C_{40} function on CrtM (213). Mapped onto the crystal structure of human squalene synthase (SqS, see **Figure 1.5b**), F26 and W38 appear in helices B and C, respectively, and the side chains of both residues point into the pocket that accommodates the second half-reaction (rearrangement of the cyclopropyl intermediate). Replacement of these amino acids with smaller ones significantly increased the C_{40} synthase activity of CrtM (**Figure 1.5c**). Conversely, the C_{30} synthase performance was the highest for wild-type CrtM and decreased with decreasing size of the amino acid residues at these positions. This analysis led to our proposal that wild-type CrtM is able to perform the rate-limiting first half-reaction of phytoene synthesis—condensation of two molecules of C_{20}PP to form the cyclopropyl intermediate, prephytoene diphosphate—but the second half-reaction is prevented from going to completion by steric inhibition of the normally fast rearrangement step. Unable to convert into phytoene in wild-type CrtM, prephytoene diphosphate either remains stuck in the enzyme, departs, or undergoes other types of rearrangement to yield noncarotenoid products. (A similar phenomenon is well known for SqS; in the absence of NADPH, which is required to convert presqualene diphosphate to squalene, SqS produces a complex mixture of non-squalene compounds including rillingol, 10-hydroxybotryococcene, and 12-hydroxysqualene (22, 89).) However, in CrtM mutants where F26 or W38 is replaced with a smaller or more flexible amino acid, the prephytoene diphosphate formed from two molecules of C_{20}PP is efficiently

rearranged to form phytoene. Thus, the F26 and W38 mutations apparently modify the product specificity, not the substrate specificity of CrtM: they do not act by allowing the enzyme to bind and accept GGPP as a substrate but, rather, allow the intermediate prephytoene diphosphate to be converted to phytoene.

Gain of C_{40} function by mutagenesis of CrtM usually came at a cost to the original C_{30} synthase activity. For example, some enzymes doubly mutated at F26 and W38 showed only negligible C_{30} synthase activity. However, other modes of obtaining C_{40} synthase activity were possible. The E180G substitution increased performance in both the C_{30} and C_{40} contexts (212). Mapped onto the SqS structure, E180G is positioned outside the reaction pocket but close to the site of the first half-reaction. We think this mutation accelerates the rate-limiting first half-reaction. Several such activating mutations were found by random mutagenesis of $CrtM_{F26S}$ in a successful search for improved or restored C_{30} function (D. Umeno, unpublished results). All of these activating mutations were rather far from the reaction center (**Figure 1.5b**, indicated in yellow) and enhanced both C_{30} and C_{40} synthase activity. These mutations could be increasing expression level, stability (half-life *in vivo*), or specific activity.

Like IDSs, carotenoid synthases appear to be specific, but at least in the case of CrtM, that specificity is readily altered by mutation. Furthermore, there are multiple mutational routes to the same altered phenotype. In contrast, in similar experiments with CrtB, we were unable to find any mutations that conferred C_{30} carotenoid synthase function. We are tempted to speculate that CrtM, by virtue of the fact that it does not need to select C_{15}PP from C_{20}PP in its natural host (C_{20}PP is not made in C_{30} carotenoid-producing organisms), is inherently more evolvable toward accepting C_{20}PP than is CrtB

toward accepting C_{15}PP because CrtB has always had to select C_{20}PP over C_{15}PP (the latter is present in all organisms). CrtM can in fact accept C_{20}PP in a hybrid reaction: when supplied with both C_{15}PP and C_{20}PP, wild-type CrtM can condense one molecule of each to synthesize a C_{35} carotenoid backbone (see below). CrtM has thus proven to be an evolvable enzyme, while CrtB has not, at least with respect to the specific tasks of accepting a larger (CrtM) or smaller (CrtB) substrate. As was the case with the IDSs, a single amino acid substitution was sufficient to increase the size of the products synthesized by CrtM. Furthermore, as we describe below, the mutants of CrtM that function in a C_{40} carotenoid biosynthetic pathway are also capable of synthesizing even larger carotenoid backbones when supplied with the appropriate precursors.

Carotenoid desaturases. Phytoene and 4,4'-diapophytoene, the (colorless) products of the natural carotenoid synthases, have three conjugated double bonds in the center of their backbones. Formation of carotenoid pigments requires extension of this conjugated double-bond system. The photochemical properties of a carotenoid (including its color) depend strongly on the size of the chromophore and therefore on the number of desaturation steps catalyzed by the desaturase enzyme(s). A C_{40} backbone can accommodate up to 15 conjugated double bonds, corresponding to six sequential desaturation steps. C_{30} carotenoids can undergo at most four desaturation steps (11 conjugated double bonds).

In general, bacterial C_{40} desaturases are functional on C_{30} carotenoids and *vice versa*. It is proposed that carotenoid desaturases recognize only a portion of the substrate molecule common to both C_{30} and C_{40} carotenoid backbones (163, 174). Subsequent carotenoid-modifying enzymes are also often locally specific (see below). Because each

of the desaturation intermediates represents a branch point for further diversification by downstream enzymes, altering and controlling the desaturation step number is key for creating extensive molecular diversity (70). Nature has done this: many different carotenoid desaturases are known in C_{40} pathways, each with its specific step number. Carotenoid desaturases that primarily catalyze two, three, four, and five desaturation steps are known; one-step and six-step desaturated carotenoids have been reported only as minor products in natural carotenogenic organisms. The ability of carotenoid desaturases to accept different substrates was demonstrated when C_{30} and C_{40} desaturases were tested in each other's pathway (163, 214). C_{40} desaturases from *Erwinia* (four-step enzymes), *Rhodobacter* (three-step), and *Anabaena* (two-step) are all active on C_{30} substrates. Similarly, the C_{30} desaturase CrtN showed measurable activity in a C_{40} pathway. The localized specificity of these desaturases has been exploited in engineered pathways (211) (see below).

The desaturase step number can be inferred from the color (or color change) of the carotenoids produced *in vivo*. This provides an excellent basis for screening mutant desaturases in the laboratory for the ability to accept new substrates or for changes in product specificity. Using this principle, Schmidt-Dannert et al., working in this laboratory, readily isolated desaturase variants with altered step number in a C_{40} pathway (181). Starting from a library made by DNA shuffling of two closely related four-step phytoene desaturases (*crtI* from *Pantoea agglomerans* [*E. herbicola*] and *crtI* from *Pantoea* ananatis [*E. uredovora*]), they isolated one four- to six-step variant, $CrtI_{14}$ (**Figure 1.6**), as well as 20 variants that catalyze fewer than four desaturation steps on phytoene. Wang and Liao conducted similar experiments with the three-step desaturase

from *Rhodobacter sphaeroides* (222). Two rounds of random mutagenesis by error-prone PCR and color screening resulted in variants that accumulated the four-step product, lycopene. At least five different mutations increased the step number of this desaturase. Furthermore, some combinations of mutations resulted in yet higher production of lycopene (222).

We recently investigated the ability of the C_{30} desaturase CrtN from *S. aureus* to alter its product specificity upon mutation. CrtN is a three-step desaturase in *S. aureus* (226). *E. coli* colonies expressing this enzyme together with the C_{30} synthase CrtM develop a yellow-orange color due to the production of 4,4'-diaponeurosporene (three steps) and 4,4 '-diapolycopene (four steps). When *E. coli* were transformed with a library of *crtN* mutants made by error-prone PCR, the variation in the colors of the colonies (**Figure 1.7**), from pale to weakly fluorescent, lemon, yellow, orange, and red, was striking. To our surprise, approximately 30% of the colonies were lemon or yellow in color, different from the wild type. A random subset of these clones was shown to produce the rare C_{30} carotenoids 4,4'-diapophytofluene (one-step desaturation) and 4,4'-diapo-ζ-carotene (two steps). We also isolated a large number of red mutants, which were found by high-performance liquid chromatography (HPLC) analysis to produce mainly 4,4'-diapolycopene (four desaturation steps) (**Figure 1.7**).

Multiple experiments have thus shown that bacterial carotenoid desaturases are evolvable. The high frequency of color variants in the mutant libraries suggests that there are multiple mechanisms by which the step number can change upon mutation. It is possible that most (if not all) mutants with altered desaturation step number have acquired this change in product specificity by up- or down-regulating total desaturase

activity rather than by altering intrinsic enzyme specificity, which we would expect to occur only rarely. The higher frequency of mutants with a lower desaturation number compared to mutants with an elevated desaturation number supports this idea because there are many more paths downward with respect to a property such as activity or expression level. Many mutant desaturases with elevated step numbers turned out to have mutations near a putative flavin adenine dinucleotide (FAD)-binding domain (181, 222), indicating that FAD access might be a source of step number control.

Carotenoid desaturases tend to exhibit more well-defined product specificity in their natural hosts than when (over)expressed in a heterologous host. For instance, CrtN appears to be a three-step enzyme and produces almost exclusively 4,4'-diaponeurosporene in *S. aureus* (226). In a heterologous host, however, its desaturation step number is less distinct. Several groups have reported that CrtN produces the four-step 4,4'-diapolycopene as a major product in an *E. coli* system (118, 163, 211, 214). Desaturase step number can also be altered by manipulation of enzymes further downstream in the carotenoid pathway (69). It is not surprising that altering the environment of a desaturase can alter its product specificity. If downstream enzymes that normally remove certain desaturation products are not present, the desaturase may have an opportunity to catalyze further desaturation steps. In addition, if these enzymes associate in complexes, expression of a desaturase with other carotenoid biosynthetic genes from different organisms could yield complexes with altered or suboptimal substrate transfer properties, which could affect the desaturation step number. Further research should shed light on the source of desaturase specificity and on the ways in which specificity can be altered.

Carotenoid cyclases. Cyclization is another important branch point for carotenoid diversification. Cyclic carotenoids are produced in plants, algae, and photosynthetic bacteria, and enzymes catalyzing the formation of different cyclic products have been characterized (11, 53, 55, 84, 109-111, 194). A given cyclase usually produces only one kind of ring structure (32). However, an exception was recently reported by Stickforth et al., who showed that the enzyme CrtL-*e* from the cyanobacterium *Prochlorococcus marinus* is both a β- and an ε-cyclase (194). The cyclases share nearly identical mechanisms, differing only in the final rearrangement step (**Figure 1.8a**).

Although lycopene, with its two ψ-ends, is the natural substrate for most known cyclases, studies have shown that some cyclases can act on a wider range of substrates (**Figure 1.8b**). An ability to convert different substrates was reported for *Arabidopsis* ε-end lycopene cyclase, which can cyclize the ψ-end of neurosporene to form α-zeacarotene (54, 55). Britton noted that the only apparent requirement for recognition and catalysis by carotenoid cyclases is that the substrate have a ψ-end (**Figure 1.8b**) (32). However, Takaichi et al. showed that bacterial and plant lycopene cyclases can cyclize the 7,8-dihydro-ψ-end of ζ-carotene and neurosporene (199). Recently, we showed that *E. uredovora* β-cyclase and a plant ε-cyclase can efficiently cyclize nonnatural carotenoids with a C_{35} backbone (211). Similarly, Lee et al. reported that the C_{30} carotenoid 4,4'-diaponeurosporene can be cyclized by *Erwinia* β-cyclases, leading to the novel cyclic C_{30} carotenoid diapotorulene (118). Another interesting enzyme is capsanthin-capsorubin synthase (CCS) from *Capsicum annuum*. The primary natural function of CCS is to rearrange the epoxidized cyclic carotenoids antheraxanthin and violaxanthin into the cyclopentyl κ-end products capsanthin and capsorubin. When expressed in *E. coli*,

however, CCS was shown to also possess lycopene β-cyclase activity, cyclizing both ends of lycopene to yield β,β-carotene (84).

All known β-cyclases, except one recently discovered in a marine bacterium (207), create rings at both ends of lycopene. In contrast, known ε-cyclases except those from *Lactuca sativa* (lettuce) and the flower *Adonis aestivalis* generate monocyclic carotenoids (53). In an attempt to manipulate the ring number and understand how it is determined, Cunningham and Gantt constructed a series of single-crossover chimeras of one-step (*Arabidopsis thaliana*) and two-step (*Lactuca sativa*) ε-cyclases (53). Analysis revealed a region of six amino acids involved in ring number determination. Further mutagenesis experiments identified single-amino-acid substitutions that alter ring number specificity: in the lettuce two-step cyclase Dy4, substitution of H457 with leucine converted the enzyme into a monocyclase; the corresponding L448H mutation in Y2, the *Arabidopsis* monocyclase, resulted in an enzyme that forms two rings. Thus, ring number specificity can be modulated with a single amino acid substitution. The authors of this study suggested that the residues they identified may play a role in dimer formation (53). Lycopene may be oriented in the plasma membrane such that one of its ends is more accessible to the cyclase than the other. Only when cyclase dimers are formed does binding of the more accessible end bring the less accessible end close enough to the other subunit for it to be cyclized as well. An alternative but similar explanation for the above observations is that the mutations alter the cyclase oligomerization state within a carotenoid-synthesizing complex. For example, when one cyclase subunit is present in each complex, carotenoid products are cyclized on only one end whereas complexes with two cyclase subunits would cyclize carotenoids at both ends. Association of cyclase

monomers in a complex would depend on their interactions with each other and the other constituents of the complex, and a single amino acid substitution could be sufficient to disrupt (or promote) this association. It is also possible that the single amino acid substitutions described above alter the enzyme's intrinsic preference for the carotenoid substrate that has already been cyclized on one end. The H457L mutation may reduce the preference of Dy4 for the monocyclic substrate, while the L448H mutation increases the ability of Y2 to cyclize it.

$CrtI_{14}$ is a four- to six-step desaturase, and *E. coli* cells expressing this variant along with CrtE (a C_{20}PP synthase) and CrtB (a C_{40} carotenoid synthase) can synthesize 3,4,3',4'-tetradehydrolycopene (**Figure 1.6**) (181). When CrtY, the β-cyclase from *Erwinia*, was coexpressed with these enzymes, the cells synthesized exclusively β,β-carotene—the same product made by cells expressing wild-type CrtE, CrtB, CrtI, and CrtY. Directed evolution, however, could create a pathway leading to a new cyclized product: after performing DNA shuffling to make a library of CrtY variants and coexpressing these with $CrtI_{14}$ and CrtE, Schmidt-Dannert et al. identified clones that accumulated the monocyclic carotenoid torulene (**Figure 1.6**) (181). Pigment analysis from cells harboring the CrtY mutants including $CrtY_2$ (**Figure 1.6**) revealed torulene together with lycopene, 3,4,3',4'-tetradehydrolycopene, β,ψ-carotene, and β,β-carotene. Torulene-producing variants were also discovered when the CrtY library was created by error-prone PCR (D. Umeno, unpublished results). These mutants made up 2 to 5% of the library. This high frequency of torulene-producing mutants suggests that the torulene pathway emerged by down-regulation of cyclase activity. Mutants of CrtY with decreased catalytic activity would compete less efficiently with $CrtI_{14}$ for the ends of the carotenoid

substrate, with the result that only one end is cyclized while the other is desaturated, leading to torulene. Alternatively, the CrtY mutants found in the torulene-producing clones may be compromised in their ability to form dimers, either as part of larger carotenoid enzyme complexes or not.

Although the underlying mechanisms for the changes in cyclase step number remain obscure, it is clear that carotenoid cyclases are evolvable: their phenotype can change dramatically with a small number of mutations, and there appear to be multiple pathways to a particular phenotype. As we show in the next section, combining evolved carotenoid desaturases and cyclases with locally specific enzymes further downstream allows an even larger natural product space to be sampled. To date, carotenoid cyclases have not been evolved in the laboratory for altered product specificity, for example to convert a β-cyclase to an ε-cyclase or vice-versa. These cyclases have very similar chemical mechanisms (**Figure 1.8a**), and we predict that cyclase product specificity will be easily modified by mutation.

Enzymes catalyzing further modifications. Numerous enzymes catalyzing hydroxylation, epoxidation, glycosylation, O-methylation, acyl transfer, prenyl transfer, oxidative cleavage, and other reactions on cyclic and acyclic carotenoids have been cloned and expressed in *E. coli* (for reviews, see references (175) and (180)). Misawa et al. made an early demonstration of the localized specificity of these modifying enzymes (135). Introduction of the β-end 3-hydroxylase CrtZ and the β-end 4-ketolase CrtW into *E. coli* harboring a β,β-carotene pathway led to the accumulation of as many as nine carotenoids (**Figure 1.3**). Recent work has revealed that many carotenoid-modifying enzymes act on a range of substrates (5, 16, 118, 191, 197).

Lee et al. exploited the catalytic promiscuity of a number of carotenoid-modifying enzymes in an *E. coli* system to extend laboratory-evolved pathways to 3,4,3',4'-tetradehydrolycopene and torulene, generating biosynthetic routes to several carotenoids not identified in nature (118) (**Figure 1.6**). CrtA, an oxygenase from *Rhodobacter capsulatus* that normally acts on the methoxy carotenoid spheroidene, was found to insert one keto group into ζ-carotene, neurosporene, and lycopene, and to insert two keto groups into tetradehydrolycopene. Monocyclic torulene, synthesized via the pathway constructed using the laboratory-evolved desaturase and cyclase (181), served as a substrate for CrtO, a β,β-carotene ketolase from *Synechocystis* sp. strain PCC 6803; CrtU, a β,β-carotene ring desaturase from *Brevibacterium linens*; and CrtZ, a β,β-carotene hydroxylase from *E. herbicola*. In addition, 3-hydroxytorulene, the product of the action of CrtZ on torulene, was further metabolized by CrtX, a zeaxanthin glycosylase from *E. herbicola*, leading to the synthesis of torulene-3-β-D-glucoside (**Figure 1.6**). In this work, the pathway to torulene generated by directed evolution of two upstream carotenoid biosynthetic enzymes was extended in several different directions by adding genes for further transformations. Owing to their intrinsic localized specificity, coexpression of these downstream modifying enzymes with the foreign torulene pathway resulted in a series of novel torulene derivatives.

There are no reports yet of directed evolution of carotenoid oxygenases or post-cyclase carotenoid-modifying enzymes. Work by Sun et al. (197), however, provides an interesting example of how modifying the sequence of a carotenoid hydroxylase can alter the enzyme's product distribution. The β,β-carotene hydroxylase from *A. thaliana* catalyzes two hydroxylation steps, converting β,β-carotene to zeaxanthin (>90%) in *A.*

thaliana and when expressed in *E. coli* (**Figure 1.9**). Noticing that this enzyme had an N-terminal extension of more than 130 amino acids compared with other known β,β-carotene hydroxylases, Sun et al. expressed in *E. coli* a truncated version of the *Arabidopsis* hydroxylase lacking the N-terminal 129 amino acids. The truncated enzyme primarily catalyzed only one hydroxylation step, and β-cryptoxanthin (**Figure 1.9**) accumulated as the main product (>75%) (197). The molecular basis for this altered hydroxylation step number is not known; the authors speculated that truncation yielded an enzyme deficient in the ability to form dimers.

Carotenoid cleavage enzymes comprise an important class of carotenoid-modifying enzymes. Enzymatic oxidative cleavage of carotenoids in both carotenogenic and non-carotenogenic organisms creates diverse "apocarotenoid" structures, many of which are highly biologically significant (71). Well-known products of carotenoid cleavage enzymes include retinol and retinoic acids (mammalian hormones) (220), abscisic acid (a plant growth regulator) (186, 236), and the plant pigment bixin (28). The localized specificity of several carotenoid cleavage enzymes has recently been demonstrated. One important cleavage enzyme is β,β-carotene-15,15'-monooxygenase (βCM, formerly named β,β-carotene-15,15'-dioxygenase before the enzyme was shown to employ a monooxygenase mechanism (121)), which converts β,β-carotene into retinal (**Figure 1.10a**) (158, 165, 219, 231, 232). βCM cleaves substrates other than β,β-carotene, such as α-carotene (228) and β-cryptoxanthin (**Figure 1.10a**) (123, 228). The enzyme has a strong preference for substrates with at least one unsubstituted β-end (123, 228) and appears to always cleave a substrate's central 15-15' double bond (115). Recently discovered homologs of βCM catalyze the asymmetric cleavage of carotenoids

(28, 98, 186). One murine oxygenase cleaves β,β-carotene at the 9'-10' double bond (**Figure 1.10b**), yielding β-apo-10'-carotenal (C_{27}) and β-ionone (C_{13}) (98). This enzyme also catalyzes the cleavage of lycopene, resulting in the formation of apolycopenals (**Figure 1.10b**) (98). Another carotenoid cleavage enzyme from *A. thaliana* was shown to cleave an impressive array of different substrates *in vitro* (**Figure 1.10c**) (185). Thus, carotenoid cleavage enzymes are quite versatile and their localized specificity may be exploited to generate diverse new carotenoid derivatives in the laboratory.

Creation of pathways to carotenoids with new carbon scaffolds

Given the proven evolvability of different enzymes involved in carotenoid biosynthesis, we investigated the possibility of creating routes to whole new families of carotenoids built on scaffolds other than the symmetric C_{30} and C_{40} backbones seen in nature. We reasoned that if carotenoid synthases could be engineered to accept different substrates and synthesize unnatural carotenoid backbones, locally specific downstream enzymes might accept and modify these backbones, leading to whole new families of carotenoids. The downstream enzymes could also be evolved to produce still more novel products or to function more efficiently on their new substrates. Compared to increasing the diversity of known C_{30} or C_{40} carotenoids by altering the specificity of downstream pathway enzymes, generating entire new pathway families by focusing on upstream enzymes has a much greater capability to explore new carotenoid structures. The entire space of carotenoids built on non-C_{30} or non-C_{40} backbones, for example, is uncharted. Generating novel chain length carotenoids can also be viewed as a rigorous test of carotenoid pathway evolvability.

A particularly exciting goal is to generate carotenoids with chromophores or

physico-chemical properties not possible in carotenoids based on natural backbones. Millions of years of evolution gave rise to the ~700 carotenoids known in nature. Laboratory evolution of new carotenoid pathways based on unnatural backbones has the potential to at least double or triple this diversity. If significant numbers of new products can be generated, laboratory pathway evolution would be validated as an important tool for preparing large natural product libraries from which compounds with valuable properties and biological activities can be discovered.

A C_{35} carotenoid pathway. We recently discovered that the C_{30} synthase CrtM could assemble a 35-carbon backbone when supplied with C_{15}PP and C_{20}PP precursors (211). When transformed with *crtE* from *E. uredovora* (encoding a C_{20}PP synthase) and *crtM* from *S. aureus*, *E. coli* cells accumulate 4-apophytoene, an asymmetrical C_{35} carotenoid formed via heterocondensation of C_{15}PP and C_{20}PP (**Figure 1.11**). This novel backbone comprised 40 to 60% of total carotenoids, with C_{30} 4,4'-diapophytoene and C_{40} phytoene as minor components.

We found that the C_{35} backbone was further metabolized by various C_{40}- and C_{30}-carotenoid biosynthetic enzymes, creating a family of new C_{35} carotenoids (**Figure 1.11**). Carotenoid desaturases from C_{40} (*Erwinia* CrtI) or C_{30} (*Staphylococcus* CrtN) pathways efficiently converted the C_{35} backbone to yield acyclic carotenoids with different levels of desaturation (211). It is interesting to note that the C_{30} desaturase CrtN showed a higher step number (four or five steps) in the C_{35} pathway than in its native C_{30} pathway (three or four steps) or in a C_{40} pathway (two or three steps). Thus, *E. coli* cells expressing CrtN develop an intense red color only when they produce the C_{35} substrate. The larger number of desaturation steps catalyzed on a C_{35} substrate may reflect a

tradeoff between the enzyme's substrate preference and the number of desaturation steps the substrate can possibly undergo. CrtN most probably prefers its native C_{30} substrate, but the C_{40} substrate can undergo more desaturations. With a C_{35} substrate, a balance of these two factors can yield the high observed step number if, for example, the enzyme (or enzyme complex) binds the "C_{30}-like" end of the C_{35} substrate strongly and catalyzes one or more additional desaturation steps on the "C_{40}-like" end compared with a C_{30} substrate. Other scenarios can also plausibly lead to the same result.

Assembling the enzymes for the new backbone (CrtE and CrtM) with wild-type desaturases provided some of the possible C_{35} desaturation products. The remaining desaturation products were obtained by directed evolution. Random mutagenesis of carotenoid desaturases expressed in *E. coli* synthesizing the C_{35} carotenoid backbone generated colonies having a spectrum of colors, which reflected desaturase mutants with altered step numbers (211). In fact, individual clones that produced each of the possible acyclic C_{35} carotenoids as the main product were identified (**Figure 1.11**). Addition of the β-cyclase CrtY from *E. uredovora* or the ε-cyclase Dy4 (*L. sativa*) further diversified the C_{35} pathway, yielding at least four new monocyclic C_{35} carotenoids (**Figure 1.11**).

As discussed in the Introduction, some fungi naturally accumulate carotenoids with 35 carbon atoms (12, 14, 164, 215). All of these are monocyclic oxygenated carotenoids (xanthophylls) and are believed to result from the action of an unidentified oxidative cleavage enzyme on C_{40} carotenoids such as torulene. Disruption of cyclase activity in *Neurospora crassa* resulted in the biosynthesis of an acyclic C_{35} xanthophyll (12). These C_{35} xanthophylls and the biosynthetic routes leading to their formation are qualitatively different from laboratory-generated C_{35} carotenoids biosynthesized via

heterocondensation of C_{15}PP and C_{20}PP. Notably, the direct route to the C_{35} carotenoid backbone 4-apophytoene opens many possibilities for further pathway diversification by downstream enzymes (desaturases, cyclases, etc.) whereas many fewer options exist for diversifying the already desaturated, cyclized, and oxygenated C_{35} carotenoids found in fungi.

Because C_{35} carotenoids have asymmetric backbones, each desaturation step can potentially yield more than one product, depending on the direction of the desaturation step (and the previous steps). Thus, acyclic C_{35} carotenoids with 3, 5, 7, 9, 11, and 13 conjugated double bonds can assume 1, 2, 3, 3, 2, and 1 (a total of 12) possible structures, respectively. In this sense, the C_{35} pathway is inherently more complex and explores a much larger "structure space" than do pathways based on the symmetric C_{30} and C_{40} carotenoid backbones. Cyclization and other modifications of the backbone can further increase the number of possible C_{35} carotenoids. Considering their broad substrate tolerances, we expect that other carotenoid-modifying enzymes will also be functional in this C_{35} pathway and that one could make many hundreds of new carotenoids. The generation of a new, full-fledged C_{35} carotenoid biosynthetic pathway in the laboratory is a convincing demonstration of the evolvable nature of carotenoid biosynthetic enzymes and pathways and serves as an excellent illustration of how gene assembly coupled with directed evolution can rapidly access diverse chemical structures.

Carotenoids with longer backbones. The C_{25}PP isoprenoid precursor farnesylgeranyl diphosphate and the C_{30}PP hexaprenyl diphosphate serve as components of archaebacterial membrane lipids and quinone side chains, respectively. They have never been found to be incorporated into carotenoids. However, these longer analogs of

the C_{40} carotenoid precursor C_{20}PP could in principle be used in the biosynthesis of carotenoid backbones having 45, 50, 55, or 60 carbon atoms. Similarly, the monoterpene precursor geranyl diphosphate (C_{10}PP) could, in theory, be used to construct C_{20}, C_{25}, or asymmetric C_{30} carotenoids. All these backbone structures would be possible, provided that there were synthases capable of condensing these substrates. Although carotenoid synthases with these activities are not known in nature, we reasoned that an existing synthase could be engineered to make at least some of these new backbones.

As discussed above, we had probed carotenoid synthase evolvability by screening for mutants of CrtM that function as a C_{40} synthase and mutants of CrtB that function as a C_{30} synthase. CrtM and CrtB diverged long ago and have only 30% amino acid sequence identity. Nonetheless, CrtM was readily converted to a C_{40} synthase (Chapter 2, (214)). Sequencing confirmed that multiple single-substitution pathways leading to this phenotype were possible (213). We then tested whether one or more of these CrtM mutants might be able to synthesize even larger backbones.

To supply a C_{25}PP substrate in *E. coli*, we constructed and expressed the Y81A variant of the C_{15}PP synthase from *B. stearothermophilus* ($BstFPS_{Y81A}$). This variant was shown by Ohnuma's group to produce C_{25}PP as the main product *in vitro* (150, 151). When coexpressed with $BstFPS_{Y81A}$, the single mutants $CrtM_{F26L}$ and $CrtM_{F26S}$ discovered in our original C_{40} pathway screening experiment (Chapter 2, (214)) synthesized 16-isopentenylphytoene (the C_{45} carotenoid backbone, C_{20}+C_{25}) and 16,16'-diisopentenylphytoene (the C_{50} backbone, C_{25}+C_{25}) (**Figure 1.12**). Site saturation mutagenesis at residues 26, 38, and 180 (also discovered to play a role in conferring C_{40} function on CrtM (213)) resulted in several double and triple mutants with improved

ability to synthesize C_{45} and C_{50} carotenoid backbones (212). These carotenoid backbones have never been reported in nature.

Interestingly, some of the new CrtM-based C_{45} and C_{50} synthases showed almost no activity in the original C_{30} pathway (212). Thus, carotenoid synthases, although highly substrate-specific in their natural pathways, can both broaden and shift specificity with minimal genetic change. It would be interesting to see whether these variants can be engineered further to become producers of C_{50} carotenoids exclusively (i.e., no smaller side products) or to generate even larger carotenoids, with 60- or even 70-carbon chains. Although such long structures may not be biologically relevant (and may even be toxic), it might be possible to make them in an engineered pathway in *E. coli*.

When co-transformed with *crtI* (phytoene desaturase), *E. coli* harboring the C_{45} and C_{50} pathways accumulated several novel desaturated C_{45} and C_{50} carotenoids (see Chapter 3). Given the proven evolvability of carotene desaturases and cyclases and the localized specificity of the carotenoid-modifying enzymes further downstream, we envision that it will be possible to generate many new carotenoids based on these backbones.

Whereas the fully desaturated, red C_{40} carotenoid 3,4,3',4'-tetradehydrolycopene has a chromophore consisting of 15 conjugated double bonds, C_{45} and C_{50} carotenoid backbones can accommodate 17 and 19 conjugated double bonds, respectively. Extension of a conjugated system increases the wavelength(s) of light absorbed by a molecule and predictably changes the molecule's color as observed under white light (33). Dodecapreno-β-carotene, a chemically synthesized analog of β,β-carotene with 19 conjugated double bonds, was reported to absorb light at wavelengths longer than 600 nm

and to possess bordeaux red to blue-violet pigmentation, depending on the solvent (93). Two of this molecule's conjugated double bonds are contributed by the two sterically-hindered β-ionone groups. Therefore, just as lycopene with 11 all-*trans*-conjugated double bonds absorbs longer-wavelength light than β,β-carotene with 9 *trans* plus 2 beta double bonds, a linear, all-*trans* carotenoid with 19 conjugated double bonds would absorb even longer-wavelength light and would possess even further blue-shifted pigmentation than dodecapreno-β-carotene. Carotenoids with 19 all-*trans* double bonds would have possible uses as colorants and might possess interesting or unique antioxidant properties.

DISCUSSION AND FUTURE DIRECTIONS

Revised view of carotenoid biosynthetic pathways

Carotenoid biosynthetic pathways have been described as having a tree-like organization, as depicted in **Figure 1.1** (8). The conserved upstream reactions leading to backbone synthesis form the trunk, which supports many branches (and subbranches) representing increasingly diverse and species-specific downstream steps. However, many examples outlined in this chapter have shown that enzymes from one branch can often convert substrates from a different branch. Thus, in light of the ease with which it is now possible to combine carotenoid biosynthetic genes from diverse sources in recombinant organisms, pathway engineers might be better served by replacing the carotenoid "tree" with a more web-like model that emphasizes the interconnectivity of enzymes and metabolites and de-emphasizes the natural circumstances in which they are found. In addition, the tree model, with its conserved trunk and variable branches, implies that the

major mode of diversity generation is through the addition of new branches and sub-branches. However, we now know that an important source of new carotenoid diversity is the generation of new trunks (new backbones), which can quickly be filled with many branches of their own by using enzymes taken from nature's extant pathway branches and evolving them in their new pathway context.

Future challenge: specific pathways

To date, most laboratory-evolved and many engineered pathways accumulate novel carotenoids as components of a complex mixture. This may be fine and sufficient for the discovery of new compounds, but to be practically useful as a synthetic route to a specific compound, for example, the pathway must undergo further engineering to become more specific. Natural carotenoid pathways, which tend to produce only a small number of end products, probably followed a similar course of "pruning" many products down to a chosen few. As laboratory evolution experiments demonstrate, invention is an overnight event, but shaping newly discovered pathways into mature, specific ones can require much subsequent engineering effort. In the laboratory thus far, we have succeeded only in the first step, and real work remains. We list several possible approaches to converting a newborn, nonspecific pathway into one that makes a single novel metabolite.

Evolving enzyme specificity. The most straightforward (but maybe the most labor-intensive) way to control pathway specificity is to engineer more specific biosynthetic enzymes. Most laboratory-evolved enzymes, however, show broadened rather than shifted specificity, and therefore the pathways containing them accumulate a combination of old and new products. In contrast, nature often does a perfect job, such

that two related enzymes (with a common ancestor but now sitting in different pathways) show no overlap in their substrate or product ranges.

We think that pathways specific for newly discovered metabolites can be engineered, possibly after accumulating neutral mutations (in the context of the desired pathway) and certainly by directly screening for mutants that produce less of the old compounds while maintaining (or even improving) their ability to synthesize the new ones. The laboratory-discovered C_{35}, C_{45}, and C_{50} carotenoid pathways provide sophisticated experimental systems for testing this approach to dealing with what is a central, and challenging, problem in metabolic-pathway engineering. In addition, such experiments may contribute to our understanding of the early evolution of metabolism. The increasingly favored "patchwork" model of pathway evolution states that the specific metabolic pathways we observe today evolved from leaky assemblages of nonspecific enzymes (90, 171, 233) but provides no molecular explanation for how this happened. Experiments aimed at evolving pathway specificity in the laboratory would permit researchers to track the impact of each genetic change as the pathways acquire greater specificity, and they should therefore provide molecular insight into this poorly understood process.

Designed channeling. Channeling of intermediates between successive enzymes in a pathway can eliminate undesirable fluxes and byproducts. As mentioned above, there is evidence supporting the existence of carotenoid biosynthetic enzyme complexes, although we do not know what role such complexes might play in pathways assembled in a recombinant organism. Designing enzyme complexes that can channel or shuttle intermediates from one enzyme or catalytic site to the next in a particular sequence may

find a use in directing the pathway toward efficient production of desired metabolites. Nature discovered this strategy for other metabolites, giving rise to the marvelous enzymatic assembly lines that synthesize polyketides and non-ribosomal peptides (reviewed in reference (169)).

In an effort to engineer substrate channeling between successive enzymes in a biosynthetic pathway, Brodelius et al. fused a C_{15}PP synthase with a sesquiterpene cyclase (34). The resultant fusion enzyme was more efficient at converting isopentenyl diphosphate into the cyclized C_{15} product than was the corresponding quantity of mixed unfused enzymes (34). Efforts in this direction might better regulate recombinant carotenoid pathways in *E. coli*, although significant structural information, which is currently unavailable for carotenoid enzymes, may be required.

Specificity by flux balance. Even pathways harboring individually promiscuous enzymes can achieve reasonable specificity as a whole if all steps are properly balanced. If the activity or expression level of each enzyme is properly tuned, production of the desired product can be enhanced and the proportion of unwanted side products can be diminished. In addition to modifying the sequences of promoter or ribosomal binding sites, Smolke et al. have shown that the individual expression levels of multiple enzymes can be tuned to some degree by altering the transcriptional order of the genes in an operon and varying the stability of mRNA toward nuclease degradation by using engineered hairpin structures or RNase recognition sites (189). Although adjusting the expression levels of enzymes with broad substrate or product specificity is unlikely to eliminate undesired metabolites outright, it should permit some degree of modulation of the ratio of desired to undesired pathway products.

Specificity by diversion of unwanted intermediates. Balancing enzyme activities or expression levels cannot eliminate side products if an enzyme converts a single substrate into multiple products, only one of which is desired. If such enzymes cannot be engineered to narrow their product range, one strategy for achieving pathway specificity is to selectively divert unwanted intermediates away from downstream enzymes in the pathway. Instead, these intermediates could be catabolized or incorporated into products that are easily separated from the desired one(s) by introducing appropriately selective enzymes. In terms of the conversion yield of precursors into products, this approach is more wasteful than is engineering enzymes to be more specific. However, because product recovery costs can vastly exceed fermentation expenses in industrial processes, product purity can be more important than product titer (23). Thus, if a tradeoff is unavoidable, the better choice may be to accept lower product yield in favor of increased purity.

Prospects and challenges for diversifying other pathways by laboratory evolution

Although the best examples of diversifying metabolic pathways by directed evolution come from work on the carotenoids, this approach can be applied to other natural product pathways as well. There are very good indications that other pathways and biosynthetic enzymes will be as readily evolvable in the laboratory as the carotenoid pathways. Jones and Firn hypothesize that many natural product pathways are in fact already conducting their own combinatorial biosynthesis and high-throughput screening programs, at least on evolutionary timescales (62-64, 91). Because the probability that any particular small molecule will have potent biological activity is inherently low, Jones and Firn argue that organisms increase the odds of discovering new metabolites that

provide a selective advantage by exploring a large number of different structures. Thus, the argument continues, natural product pathways have evolved traits that maximize the production and retention of product diversity while minimizing the costs of synthesizing different metabolites, most of which are not beneficial to the producer. Among these traits are branched pathway structures, which enable the discovery of new metabolites as well as the retention of existing ones; matrix pathways or metabolic grids, where a small number of enzymes convert structurally similar precursors to a large number of products by multiple routes; and promiscuous enzymes, which can accept a variety of substrates and convert newly discovered metabolites into additional new products. As we have shown, carotenoid pathways possess these traits as well as others (e.g., individual enzymes are evolvable) that can be exploited for forward evolution. On the other hand, because carotenoids already fulfill important biological roles and no organism is known to synthesize a large number of different carotenoids, it is not apparent that nature is conducting massive searches for new and improved carotenoids. Other natural product pathways, however, do appear to be in "search" mode (for example, the bracken fern makes at least 27 different sesquiterpene indan-1-ones (91)), and these pathways may be even more evolvable than carotenoid pathways.

Isoprenoid pathways are an obvious target for laboratory evolution. Many isoprenoid biosynthetic enzymes accept a variety of natural and unnatural substrates (139, 140, 147). (91). The gibberellins, a class of diterpenoid hormones with more than 120 known members, are synthesized by a metabolic grid composed of enzymes with broad substrate tolerance (59, 64). We have shown how the localized specificity of downstream carotenoid biosynthetic enzymes allowed pathway branches leading to whole sets of

novel compounds to be opened when the substrate or product preferences of key upstream enzymes were altered. In a similar example demonstrating the localized specificity of other isoprenoid biosynthetic enzymes, a mutated hydroxylase located early in the monoterpene pathway of spearmint resulted in a plant with a completely shifted monoterpene profile resembling that of peppermint (52). Other isoprenoid biosynthetic enzymes have very broad product specificity and synthesize multiple compounds (in one case, more than 50) from a single substrate (25, 47, 51, 60, 190, 229). Pathways harboring such enzymes are therefore highly branched and provide many possible avenues for further product diversification by directed evolution of downstream enzymes. Enzymes with broad specificity, a key feature of evolvable pathways, are found in the biosynthetic pathways of many other natural products, including alkaloids, polyketides, shikimate derivatives, nonribosomal peptides, and volatile esters (3, 26, 36, 169, 184, 188). It is likely that all these pathways will be readily diversified by laboratory evolution.

Laboratory evolution will also be useful in constructing hybrid pathways for new natural products. Nature parsimoniously achieves great increases in small-molecule product diversity through the biosynthesis of hybrid natural products. Mycophenolic acid, tocotrienols and tocopherols (vitamin E), cannabinoids, isoprenoid quinones, furanocoumarins, furanoquinolines, alizarin anthraquinones, and certain flavonoids and alkaloids are examples of natural products formed by the fusion of products from two or more distinct biosynthetic pathways (36, 127). Chemical synthesis of unnatural hybrids of natural products has given rise to molecules with potent biological activity, many of which possess combined properties of their individual natural product constituents (reviewed in references (208) and (132)). The number of possible natural product hybrids

is astronomical, and, like nonhybrid natural products, most have yet to be made. Accordingly, altering the specificity of enzymes catalyzing the fusion reactions that lead to hybrid products in nature may be a promising way to explore natural product diversity.

While laboratory evolution is likely in principle to be effective at exploring natural product diversity in a number of pathways, in practice, some experiments will be far easier than others. Thus, the above discussion is not complete without consideration of the technical issues involved in making and interrogating large libraries of mutant pathways. The first challenge is to create and express libraries of mutant enzymes in the context of a functional pathway. By virtue of their high transformation efficiency, fast growth, and the large number of available tools for genetic manipulation, *E. coli* cells are attractive for laboratory evolution experiments (and many other metabolic engineering experiments). Also, because *E. coli* does not synthesize carotenoids on its own, there is no background pool of carotenoids to contend with for laboratory evolution or combinatorial biosynthesis experiments with carotenoid biosynthetic pathways. *E. coli,* however, is not a universal protein expression machine, nor is it an ideal production host for all compounds. For example, it lacks the ability to glycosylate proteins. In addition, the precursors required for the biosynthesis of certain natural products, including many polyketides and isoprenoids, are present at low levels or not at all in *E. coli*. The use of engineered strains of *E. coli* with increased capacity for isoprenoid production (reviewed in reference (17)) and the ability to synthesize complex polyketides (reviewed in reference (161)) will overcome this limitation. Alternatively, libraries can be expressed and even constructed in other organisms such as *Saccharomyces cerevisiae,* for which a host of laboratory evolution tools exist (1, 2, 37, 38, 45, 46). Recently, an approach to the

laboratory evolution of polyketide biosynthesis was demonstrated in which *E. coli* was used as a host for generating a plasmid library by *in vivo* recombination. This library was subsequently transformed into a special strain of *Streptomyces venezuelae* for expression and screening (99). This approach combines the ease of genetic manipulation of *E. coli* with the ability of *Streptomyces* to generate bioactive glycosylated polyketides. Many directed enzyme evolution experiments have used an analogous approach in which the mutant library was generated in *E. coli,* but expressed in another, more suitable host such as *Bacillus subtilis* (44, 187, 240). This approach is then limited by the efficiency of transformation (often poor, especially for eukaryotic cells) and the ability to make cell lines that stably and reproducibly express the mutant sequences (and their pathway products).

Once the library of mutant pathways is constructed and expressed, it must be screened to identify those that produce a new or desired metabolite(s). As is the case for directed protein evolution, this part of the experiment is usually the most challenging and labor-intensive. In the examples discussed in this chapter, it was relatively easy to monitor the characteristic pigmentation of carotenoids (e.g., using simple, color-based visual screening of bacterial colonies on plates). It is possible that color-based screens will similarly prove useful for evolving flavonoid (195) or porphyrin (114) biosynthetic pathways, since these products are also pigmented. Nonetheless, most natural products do not generate fluorescent or UV-visible spectra that are characteristic and distinguishable in the presence of cell debris, and other strategies must be employed for screening. Screening becomes much more challenging when there are no characteristic spectral changes that can be measured *in situ*, or when the desired products must be isolated and

analyzed before they can be identified. (Even among visibly colored families of compounds, new structures are not identifiable by colony color screening unless their pigmentation differs markedly from that of the products of the native pathway. Additionally, cell color reflects the total product mixture, and evolved pathways must therefore generate a significantly different product spectrum from the native pathway in order to be apparent in a simple color-based screen.) Thus, more general methods for rapidly screening microbial clones for the production of a wide range of natural products are desirable. Ideally, the screens would require minimal purification or work-up of cells or lysates, would be highly sensitive and reproducible, and would be amenable to automation. Mass spectrometry (MS) can be used when a target product has a different mass or fragmentation pattern from the product(s) of the wild-type enzyme(s). Using MS, it is possible to screen up to 10,000 samples per day (182), and MS screening has been used to for the directed evolution of enzymes with altered product profiles or enantioselectivity (166, 168). However, due to the large capital investment required (>$1 million), automated, high-throughput MS-based screening equipment is found primarily in well-funded industrial research laboratories. Screening by HPLC or thin-layer chromatography is also feasible but similarly requires sample pretreatment and is only high-throughput with parallel, automated instrumentation. Screening using nuclear magnetic resonance spectroscopy (NMR) has also been demonstrated (167). Up to 1400 samples per day could be analyzed by an NMR spectrometer equipped with a commercially available flow-through NMR probe head. Product concentrations, however, must be quite high for detection by NMR, and pretreatment is required. In some cases, pooling of samples increases the number of clones that can be screened, provided that

there is sufficient sensitivity to detect the novel compounds (which are often present at a lower concentration). Screens based on biological properties (e.g., antimicrobial activity or protein binding), widely used in drug discovery (24, 217, 221), could also be adapted to laboratory evolution studies.

CONCLUSIONS

In its short history, laboratory pathway evolution has produced more than 30 carotenoids that have never been seen in nature. With the ability to make new backbones in *E. coli* comes the potential to generate whole families of novel carotenoids through the action of wild-type or laboratory-evolved carotenoid-modifying enzymes on the new carbon scaffolds. Applied to other biosynthetic pathways, laboratory pathway evolution could allow researchers to access thousands of molecules that are difficult to produce in practical quantities by synthetic chemistry, are expensive to isolate from natural sources, or have not been found in nature.

Seeing is believing. Evolving carotenoid biosynthetic pathways in the laboratory has allowed us to witness first-hand their remarkable evolutionary potential. In the process, we have learned that carotenoid enzymes can acquire new specificities that allow them to accept different substrates, function in a foreign pathway, make a different product from the same substrate, and even give rise to new pathways. Furthermore, all of these changes can be effected in just one round of evolution and are usually brought about by a single amino acid substitution. As evolution proceeds, we will gain further, detailed information on how pathways diverge and, we hope, some elementary understanding of how the stringent specificity of enzymes and pathways we see in nature

is achieved at the molecular level.

We wonder how far laboratory pathway evolution can be taken. For example, starting with the genes that make up an arbitrary biosynthetic pathway in nature, can we evolve any chemically possible pathway made up of the same basic transformations? We have seen that the carotenoid synthase CrtM can rapidly acquire new specificities, lose much of its original function, and yet quickly recover its original function via different mutational routes. This is an encouraging lesson: evolving an enzyme with narrowed specificity does not "paint it into a corner," evolutionarily speaking. It seems that one can abolish a particular specificity with ease but cannot abolish an enzyme's inherent ability to evolve. Thus, the enzymes that exist in 2005 may be perfectly good starting points for creating widely diverse pathways.

The dramatic functional changes we see in laboratory pathway evolution experiments remind us that natural product biosynthetic pathways are not merely a snapshot of history to be observed and documented. Rather, these pathways are continually changing entities whose evolutionary dynamics we are just beginning to probe. Laboratory evolution experiments with carotenoid pathways have given us an exciting glance at these dynamics. Future experiments promise both a greater understanding of how metabolic systems evolve and the discovery and optimization of biosynthetic routes to a host of molecules with diverse beneficial impacts on human health and well-being.

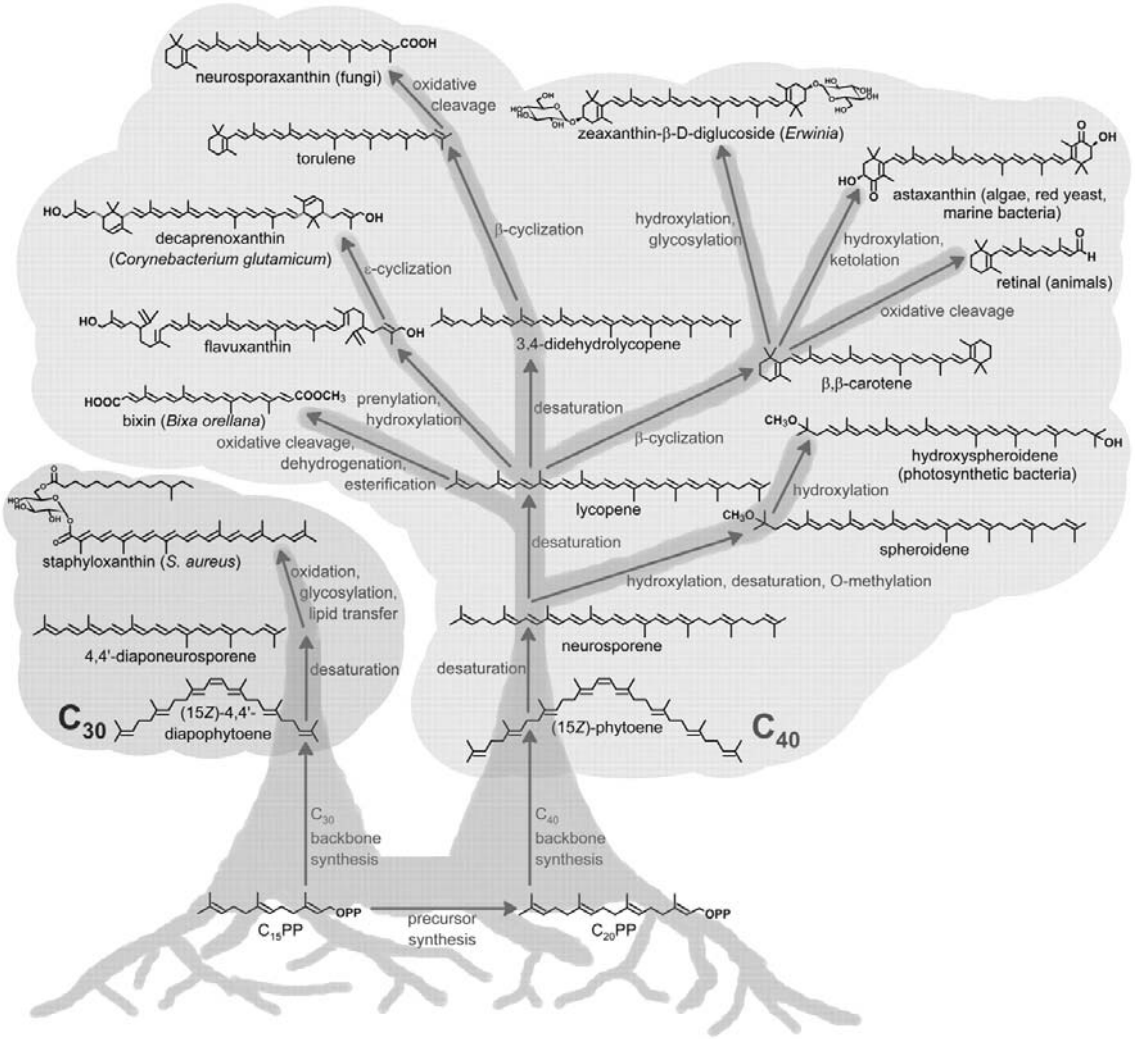

neurosporaxanthin (fungi)
oxidative cleavage
torulene
zeaxanthin-β-D-diglucoside (Erwinia)
astaxanthin (algae, red yeast, marine bacteria)
decaprenoxanthin (Corynebacterium glutamicum)
β-cyclization
ε-cyclization
hydroxylation, glycosylation
hydroxylation, ketolation
retinal (animals)
oxidative cleavage
flavuxanthin
3,4-didehydrolycopene
β,β-carotene
bixin (Bixa orellana)
prenylation, hydroxylation
oxidative cleavage, dehydrogenation, esterification
desaturation
β-cyclization
hydroxyspheroidene (photosynthetic bacteria)
lycopene
hydroxylation
spheroidene
staphyloxanthin (S. aureus)
desaturation
oxidation, glycosylation, lipid transfer
hydroxylation, desaturation, O-methylation
neurosporene
4,4'-diaponeurosporene
desaturation
desaturation
C30
(15Z)-4,4'-diapophytoene
(15Z)-phytoene
C40
C30 backbone synthesis
C40 backbone synthesis
C15PP
precursor synthesis
C20PP

Figure 1.1

Figure 1.1. Natural carotenoid biosynthetic pathways can be organized in a tree-like hierarchy. The biosynthesis of all natural carotenoids begins with the enzymatic assembly of a C_{30} or C_{40} backbone. These backbones are desaturated, cyclized, oxidized, and otherwise modified by downstream enzymes in various species-specific combinations. Shown are several common types of enzymatic transformations that occur in natural carotenoid pathways. Common carotenoids formed early in a pathway, such as lycopene, are modified in different organisms depending on the enzymes present. The multiple enzymatic routes originating from intermediates common to many end products result in extensive pathway branching (and sub-branching).

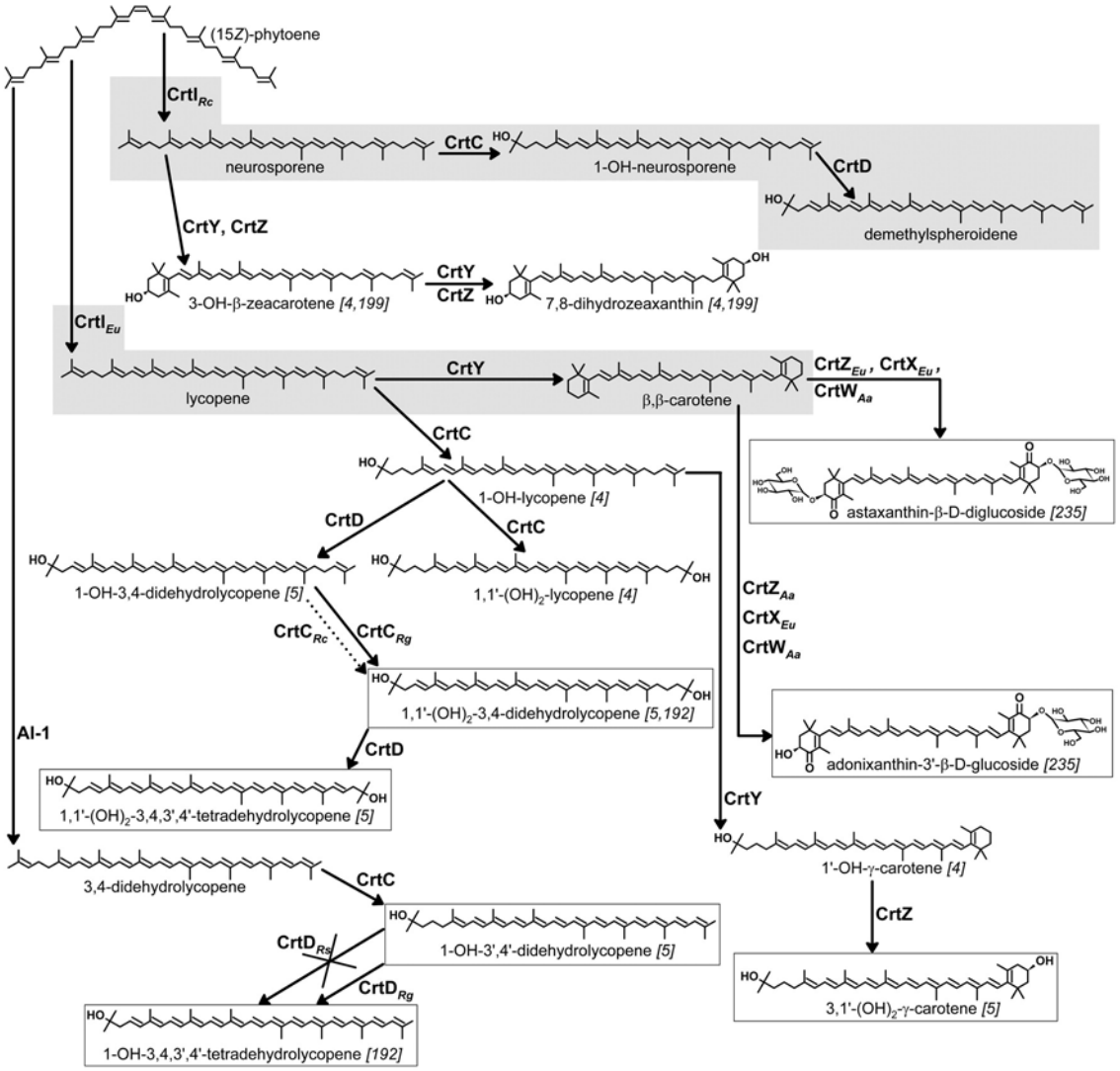

(15Z)-phytoene
CrtI$_{Rc}$
neurosporene
CrtC
1-OH-neurosporene
CrtD
demethylspheroidene
CrtY, CrtZ
3-OH-β-zeacarotene [4,199]
CrtY
CrtZ
7,8-dihydrozeaxanthin [4,199]
CrtI$_{Eu}$
lycopene
CrtY
β,β-carotene
CrtZ$_{Eu}$, CrtX$_{Eu}$, CrtW$_{Aa}$
astaxanthin-β-D-diglucoside [235]
CrtC
1-OH-lycopene [4]
CrtD
CrtC
1-OH-3,4-didehydrolycopene [5]
1,1'-(OH)$_2$-lycopene [4]
CrtZ$_{Aa}$
CrtX$_{Eu}$
CrtW$_{Aa}$
CrtC$_{Rc}$
CrtC$_{Rg}$
1,1'-(OH)$_2$-3,4-didehydrolycopene [5,192]
adonixanthin-3'-β-D-glucoside [235]
AI-1
CrtD
1,1'-(OH)$_2$-3,4,3',4'-tetradehydrolycopene [5]
CrtY
1'-OH-γ-carotene [4]
3,4-didehydrolycopene
CrtC
CrtZ
1-OH-3',4'-didehydrolycopene [5]
CrtD$_{Rs}$
CrtD$_{Rg}$
3,1'-(OH)$_2$-γ-carotene [5]
1-OH-3,4,3',4'-tetradehydrolycopene [192]

Figure 1.2

Figure 1.2. Selected examples of gene assembly leading to novel carotenoids in *E. coli*. Combining enzymes from different source organisms can lead to rare or novel carotenoids. In some cases, such as with CrtC and CrtD (4, 5), new pathway branches can be uncovered by replacing an enzyme with its counterpart from a different organism (192). Carotenoids in boxes had not been previously identified in biological material, or chemically synthesized. Enzyme abbreviations: Al-1, phytoene desaturase from *Neurospora crassa* (five-step); CrtC, neurosporene 1,2-hydratase; CrtD, 1-OH-neurosporene 3,4-desaturase; $CrtI_{Eu}$, phytoene desaturase from *Erwinia uredovora* (four-step); $CrtI_{Rc}$, phytoene desaturase from *Rhodobacter capsulatus* (three-step); CrtW, β,β-carotene 4,(4')-ketolase; CrtX, zeaxanthin glycosylase; CrtY, lycopene β-cyclase; CrtZ, β,β-carotene 3,(3')-hydroxylase. Other enzyme subscripts: *Aa*, "*Agrobacterium aurantiacum*" (current designation, *Paracoccus* sp. strain MBIC1143); *Eu*, *Erwinia uredovora* (current approved name, *Pantoea ananatis*); *Rg*, *Rubrivivax gelatinosus*; *Rs*, *Rhodobacter sphaeroides*. Natural pathways leading to the synthesis of demethylspheroidene and β,β-carotene are shaded. The dotted arrow represents trace enzymatic conversion, and the crossed-out arrow denotes undetectable enzymatic conversion. References to original publications describing the biosynthesis of various compounds are given in brackets beside compound names.

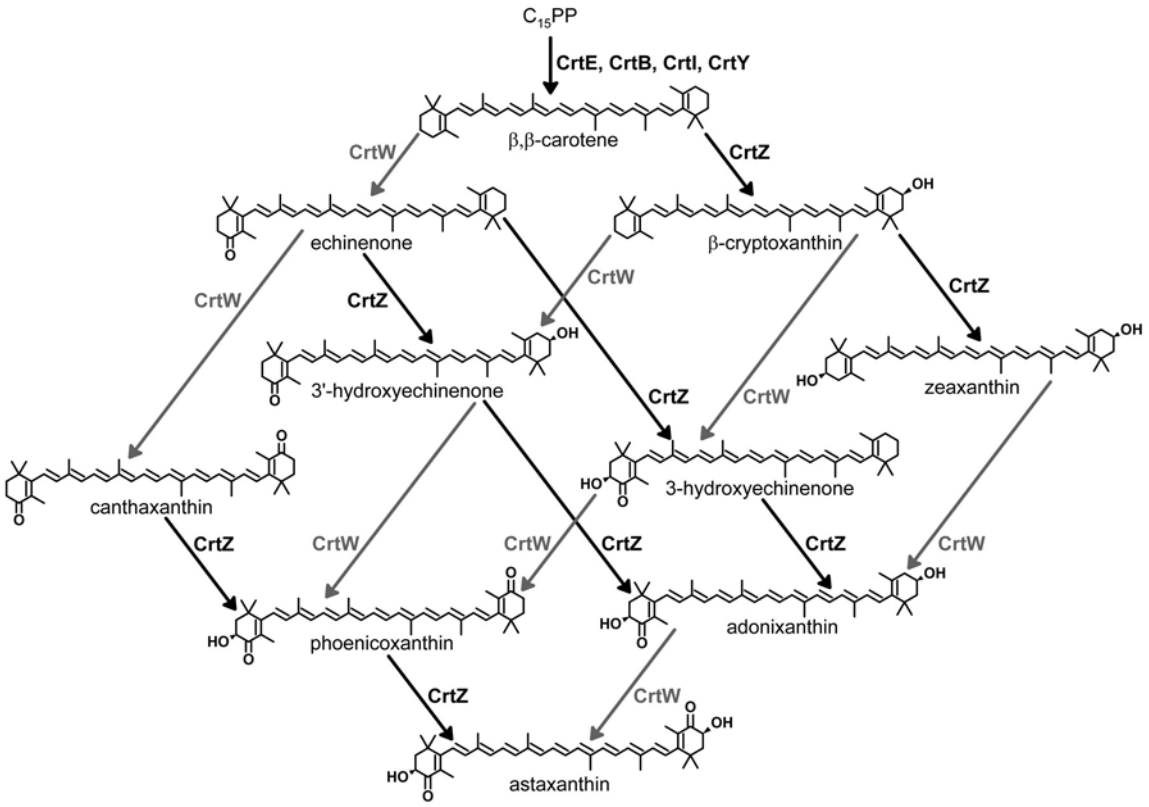


Figure 1.3. "Matrix" pathway resulting from the coexpression of carotenoid biosynthetic enzymes with broad specificity. Coexpression in *E. coli* of CrtW [β,β-carotene 4,(4')-ketolase] and CrtZ [β,β-carotene 3,(3')-hydroxylase] from "*Agrobacterium aurantiacum*" (current designation, *Paracoccus* sp. strain MBIC1143) along with the genes for β,β-carotene synthesis from *E. uredovora* (current approved name, *Pantoea ananatis*) resulted in the biosynthesis of at least nine different β,β-carotene derivatives (135). This figure was made using data from reference (135).

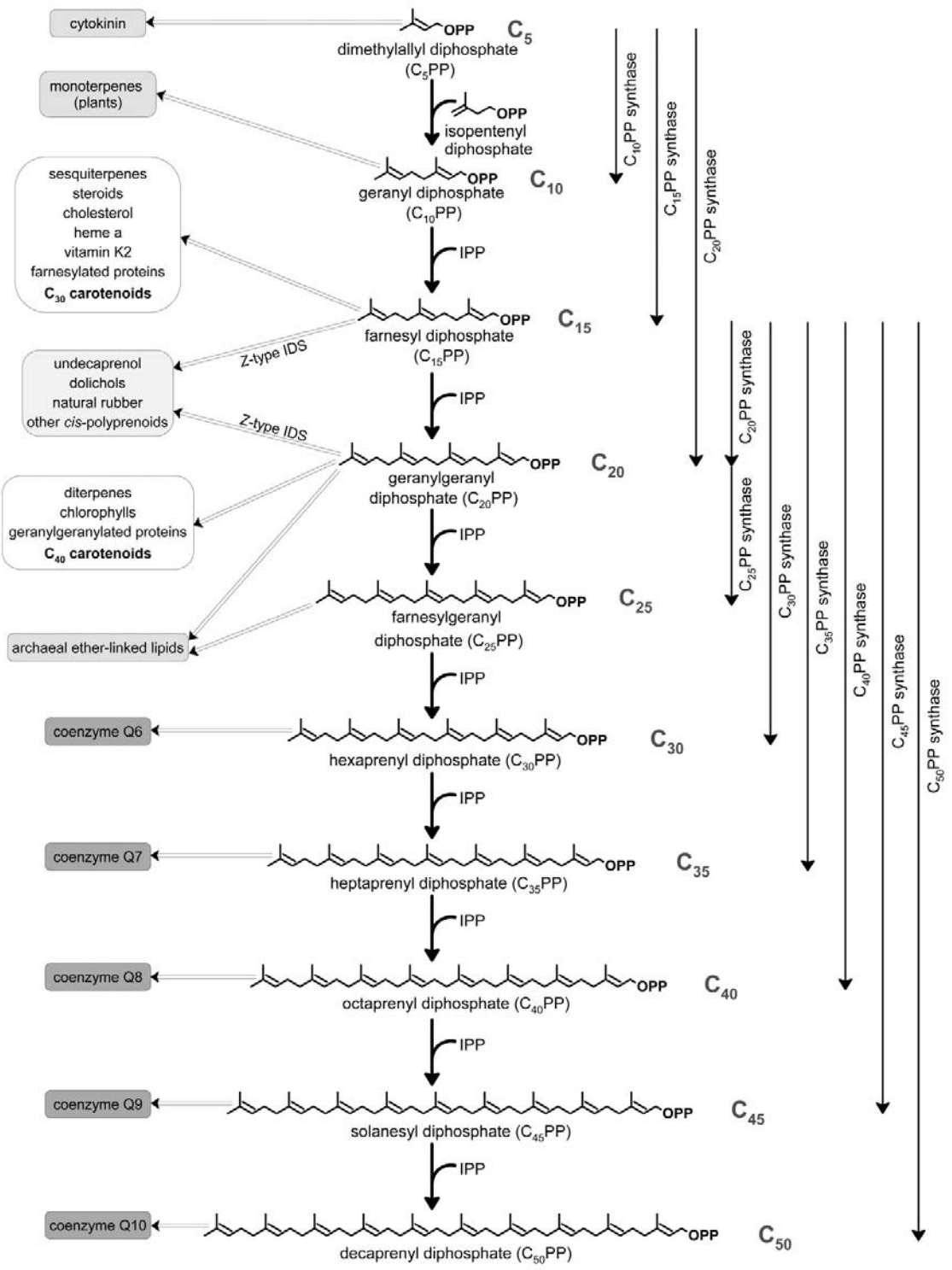

cytokinin
OPP
C5
dimethylallyl diphosphate
(C5PP)
monoterpenes
(plants)
OPP
isopentenyl
diphosphate
sesquiterpenes
steroids
cholesterol
heme a
vitamin K2
farnesylated proteins
C30 carotenoids
OPP
C10
geranyl diphosphate
(C10PP)
IPP
OPP
C15
farnesyl diphosphate
(C15PP)
Z-type IDS
undecaprenol
dolichols
natural rubber
other cis-polyprenoids
Z-type IDS
IPP
OPP
C20
geranylgeranyl
diphosphate (C20PP)
diterpenes
chlorophylls
geranylgeranylated proteins
C40 carotenoids
IPP
OPP
C25
farnesylgeranyl
diphosphate (C25PP)
archaeal ether-linked lipids
IPP
coenzyme Q6
OPP
C30
hexaprenyl diphosphate (C30PP)
IPP
coenzyme Q7
OPP
C35
heptaprenyl diphosphate (C35PP)
IPP
coenzyme Q8
OPP
C40
octaprenyl diphosphate (C40PP)
IPP
coenzyme Q9
OPP
C45
solanesyl diphosphate (C45PP)
IPP
coenzyme Q10
OPP
C50
decaprenyl diphosphate (C50PP)
C10PP synthase
C15PP synthase
C20PP synthase
C20PP synthase
C25PP synthase
C30PP synthase
C35PP synthase
C40PP synthase
C45PP synthase
C50PP synthase

Figure 1.4

Figure 1.4. Chain elongation reactions catalyzed by natural all-*E* IDSs. Each IDS adds a defined number of IPP molecules to an allylic diphosphate substrate before the final product is released. Wild-type IDSs have quite stringent product specificity; catalytic ranges of known all-*E* IDSs are shown with arrows on the right. Natural products derived from the various isoprenyl diphosphates are shown on the left. This figure was made using data from reference (225).

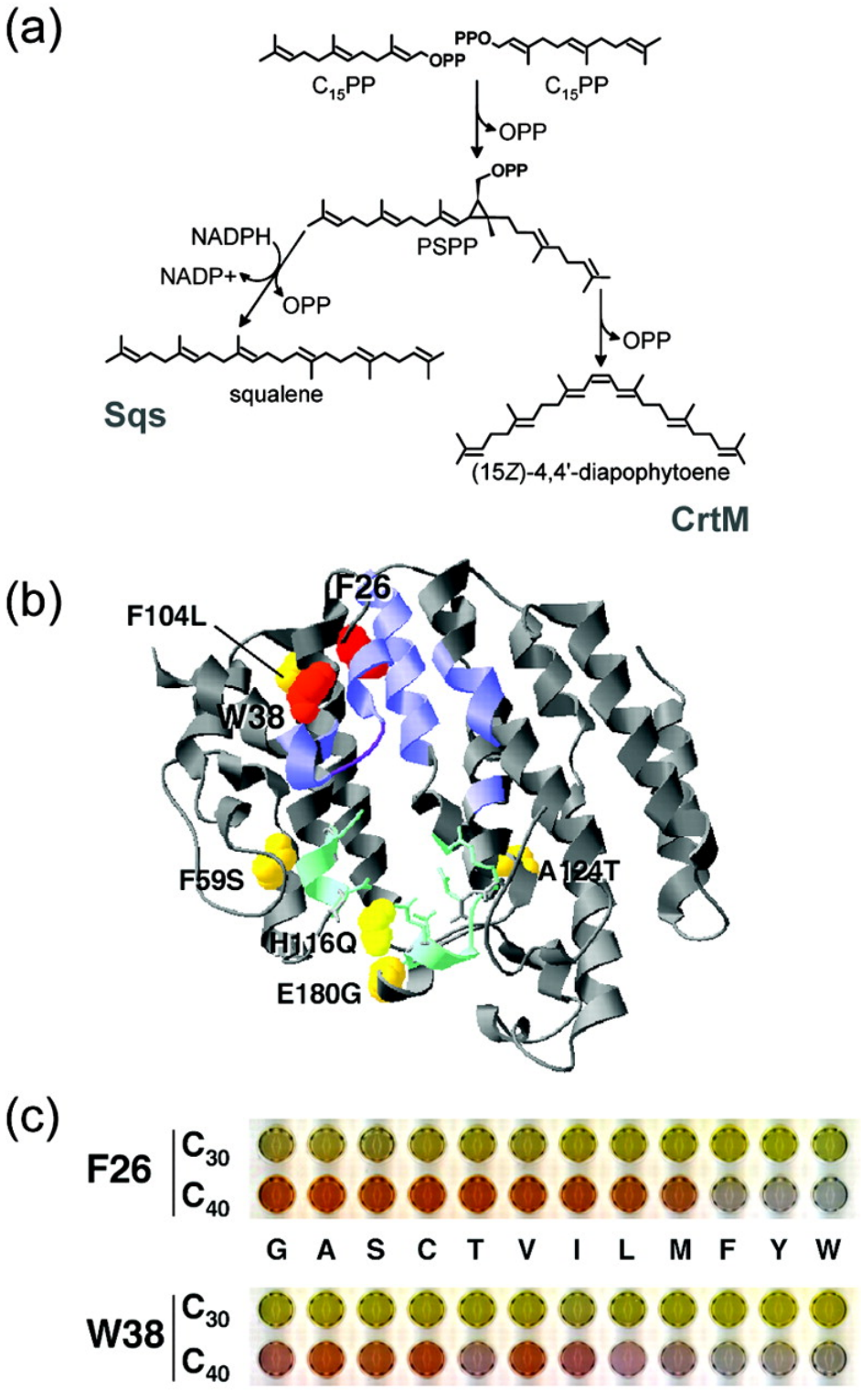

(a)
PPO
OPP
C15PP
C15PP
OPP
OPP
NADPH
NADP+
PSPP
OPP
OPP
squalene
Sqs
(15Z)-4,4'-diapophytoene
CrtM
(b)
F26
F104L
W38
F59S
A124T
H116Q
E180G
(c)
F26
C30
C40
G A S C T V I L M F Y W
W38
C30
C40

Figure 1.5

Figure 1.5. Product length determination by CrtM and mutants thereof. ***(a)*** Reaction mechanism of squalene synthase (left pathway) and CrtM (right pathway). Both enzymes catalyze the head-to-head condensation of two molecules of $C_{15}PP$ to yield linear C_{30} hydrocarbons. The enzymes employ similar mechanisms involving a common stable intermediate, presqualene diphosphate (PSPP). The mechanisms differ only in the final rearrangement step, which is followed in the case of squalene synthase by reduction of the central double bond. ***(b)*** Crystal structure of human squalene synthase (PDB ID: 1EZF). Green residues are involved in the first half-reaction (formation of PSPP), while blue residues form the pocket of the second half-reaction in which PSPP is rearranged. Red residues correspond by sequence alignment to F26 and W38 of CrtM, which were found to control the product size of that enzyme (213, 214). Shown in yellow are the residues that align with mutations in CrtM that improve both its C_{30} and C_{40} carotenoid synthase activities. ***(c)*** Mutation analysis of F26 and W38 of CrtM. Various site-directed variants of CrtM (indicated by their one-letter amino acid code) were constructed and tested for their synthase activity in both C_{30} and C_{40} carotenoid pathways. Variants are arranged from left to right in increasing order with respect to the van der Waals' volume of the substituted amino acid. The C_{30} carotenoid synthase activity of each variant was estimated from the level of yellow pigmentation when coexpressed with CrtN. Similarly, the C_{40} carotenoid synthase function of each variant was evaluated in terms of the level of red pigmentation when coexpressed with CrtE and CrtI. Panels *a* and *b* were made using data from reference (212).

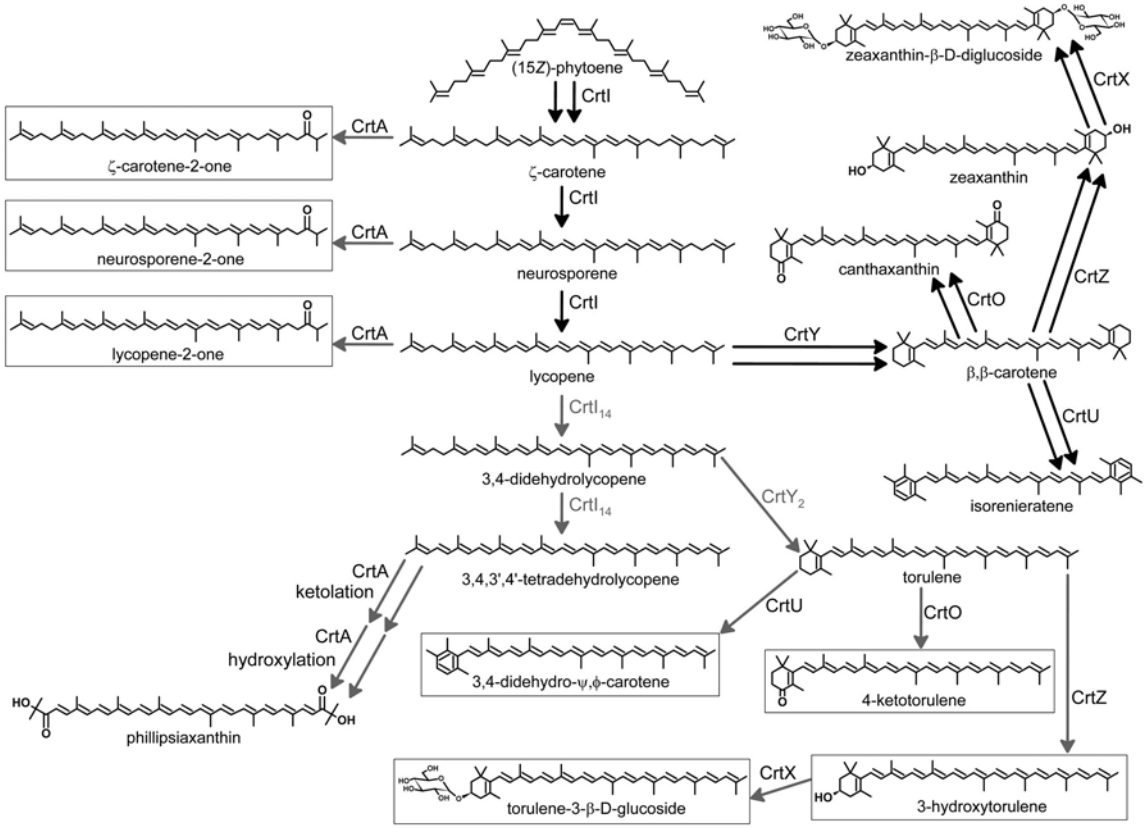
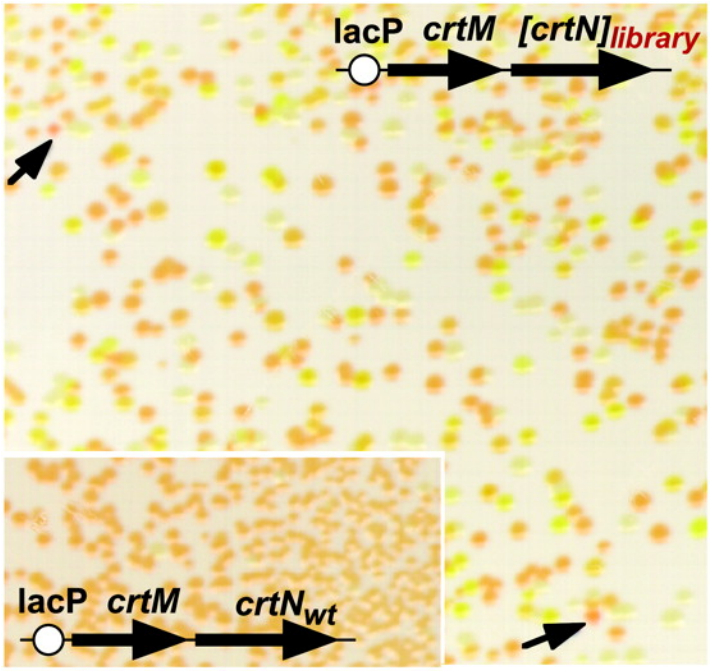

(15Z)-phytoene
CrtI
ζ-carotene
CrtA
ζ-carotene-2-one
CrtI
neurosporene
CrtA
neurosporene-2-one
CrtI
lycopene
CrtA
lycopene-2-one
CrtI14
3,4-didehydrolycopene
CrtI14
3,4,3',4'-tetradehydrolycopene
CrtA
ketolation
CrtA
hydroxylation
phillipsiaxanthin
CrtY
β,β-carotene
CrtY2
torulene
CrtU
3,4-didehydro-ψ,ϕ-carotene
CrtO
4-ketotorulene
CrtZ
3-hydroxytorulene
CrtX
torulene-3-β-D-glucoside
CrtU
isorenieratene
CrtO
canthaxanthin
CrtZ
zeaxanthin
CrtX
zeaxanthin-β-D-diglucoside

Figure 1.6

Figure 1.6. Extension of laboratory-evolved pathways to 3,4,3',4'-tetradehydrolycopene and torulene by coexpression of downstream carotenoid modifying enzymes (118). Boxed structures had not been previously identified in biological material. Native enzyme reactions are shown with black arrows. Gray arrows depict reactions not seen in natural organisms, and laboratory-evolved enzymes are written in gray lettering. Double arrows indicate enzymatic modification of both ends of a carotenoid substrate. Enzyme abbreviations: CrtA, spheroidene monooxygenase; CrtI, phytoene desaturase from *E. uredovora* (current approved name, *Pantoea ananatis*) (four-step), $CrtI_{14}$, laboratory-evolved mutant of CrtI (four- to six-step) (180); CrtO, β,β-carotene 4,(4')-ketolase; CrtU, β,β-carotene desaturase; CrtX, zeaxanthin glycosylase; CrtY, lycopene β-cyclase; $CrtY_2$, laboratory-evolved mutant of CrtY (180); CrtZ, β,β-carotene 3,(3')-hydroxylase. Biosynthesis of oxygenated derivatives of ζ-carotene and neurosporene was attributed to early termination by CrtA of the desaturation sequence of CrtI, which is normally a four-step desaturase (118). Although CrtA is primarily a ketolase, it is thought to catalyze the introduction of both hydroxy groups of phillipsiaxanthin (C. Schmidt-Dannert, personal communication). This figure was made using data from reference (118).


lacP crtM [crtN]library
lacP crtM crtNwt

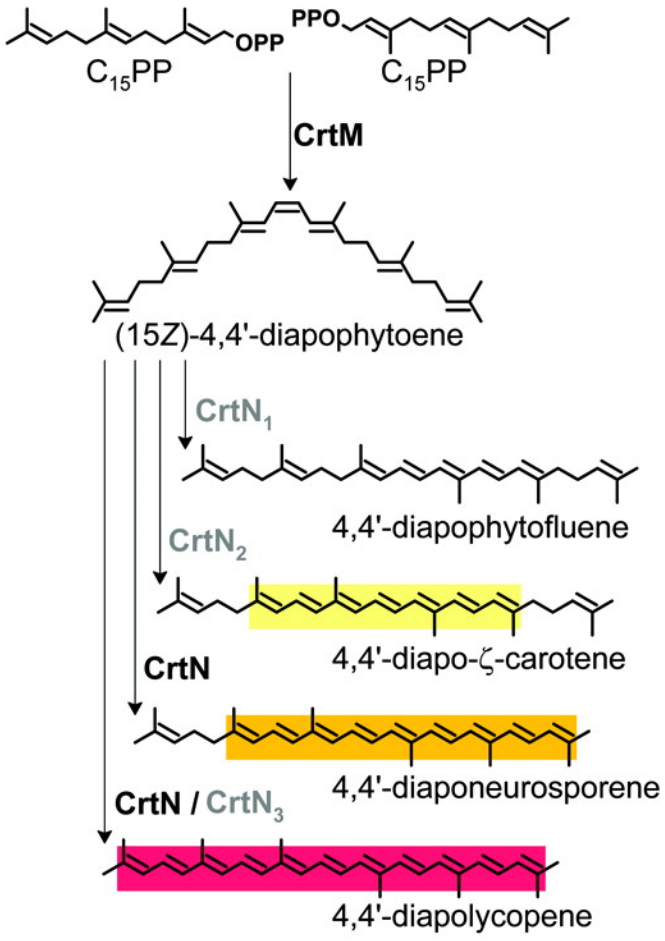

OPP
C15PP
PPO
C15PP
CrtM
(15Z)-4,4'-diapophytoene
CrtN1
4,4'-diapophytofluene
CrtN2
4,4'-diapo-ζ-carotene
CrtN
4,4'-diaponeurosporene
CrtN / CrtN3
4,4'-diapolycopene

Figure 1.7

Figure 1.7. The step number of the C_{30} desaturase CrtN is readily modified by mutation. A library of mutant *crtN* alleles was coexpressed with the C_{30} carotenoid synthase *crtM* in *E. coli*. Wild-type CrtN is a three- to four-step desaturase when expressed in *E. coli*, causing colonies to appear orange (inset). Colonies harboring CrtN mutants with increased step number were deep orange and red (highlighted by arrows), reflecting the longer carotenoid chromophores formed by these mutant desaturases. Colonies harboring mutants with decreased step number appear yellow/lemon, while colonies with non-functional or single-step desaturases are pale since these products are colorless. Shown below are the possible C_{30} desaturation products and the desaturases that synthesize them. $CrtN_{1\text{-}3}$ are discovered mutants of CrtN with altered desaturation step number. Each colored box depicts the approximate color of the carotenoid in white light. Molecules without colored boxes are colorless. Abbreviations: lacP, *lac* promoter; wt, wild-type.

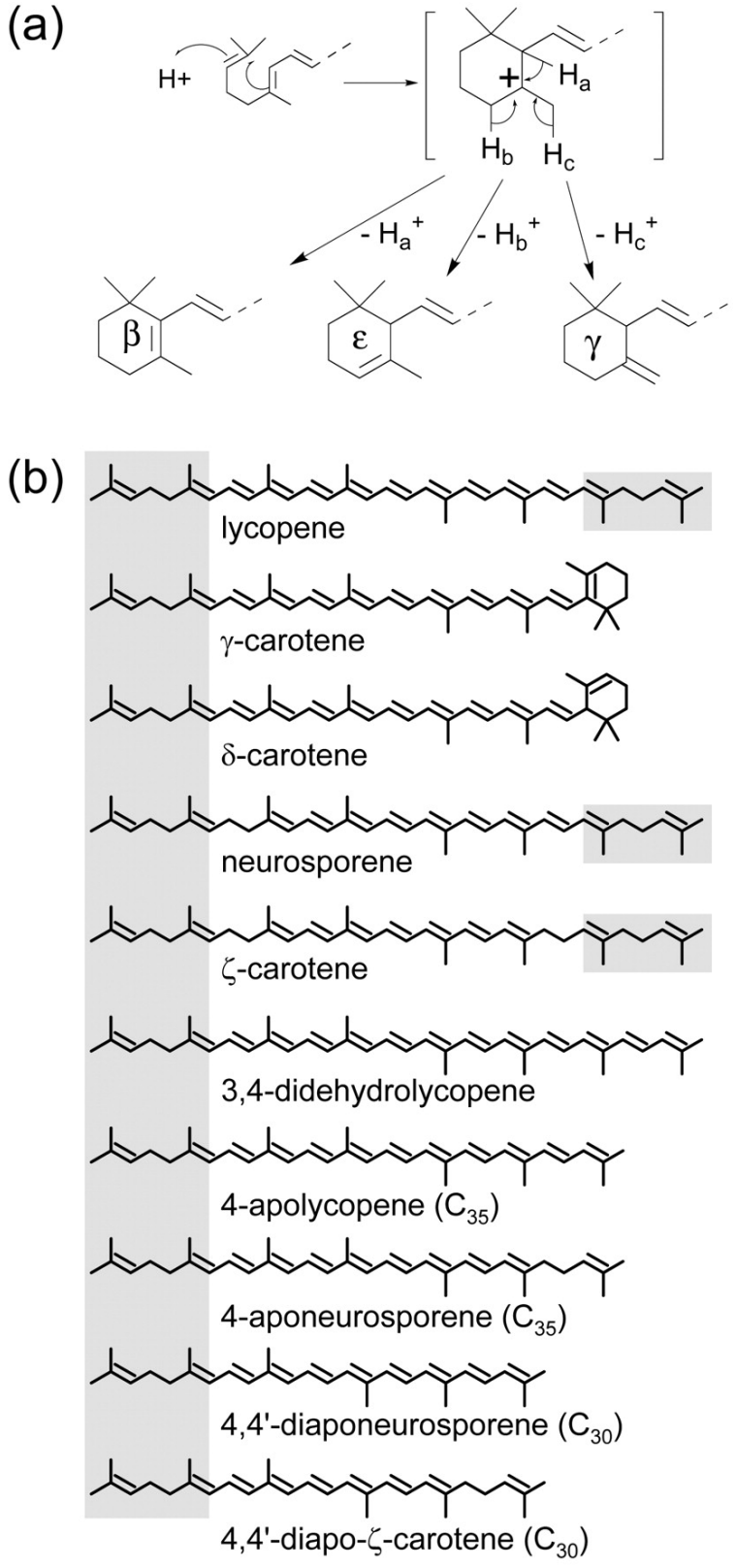

(a)
H+
+
Ha
Hb
Hc
- Ha+
- Hb+
- Hc+
β
ε
γ
(b)
lycopene
γ-carotene
δ-carotene
neurosporene
ζ-carotene
3,4-didehydrolycopene
4-apolycopene (C35)
4-aponeurosporene (C35)
4,4'-diaponeurosporene (C30)
4,4'-diapo-ζ-carotene (C30)

Figure 1.8

Figure 1.8. Cyclization of carotenoids. ***(a)*** Mechanism of carotenoid cyclization. Lycopene with its two ψ-end groups is the preferred substrate for all known carotenoid cyclases. All cyclases produce the same carbocationic intermediate on protonation of C-2 of the substrate. Product ring type (β, ε, or γ) is determined by which proton (a, b, or c in the figure) is eliminated in the subsequent rearrangement step. This panel was made using data from reference (32). ***(b)*** Substrates accepted by lycopene cyclases. The shaded portion(s) of each structure is cyclized. The first three compounds, with their ψ-end groups, have long been known as natural substrates for lycopene cyclases. The remaining seven structures, including neurosporene with its 7,8-dihydro-ψ-end (left side of molecule) and ζ-carotene with two 7,8-dihydro-ψ-ends, have more recently been shown to be cyclized.

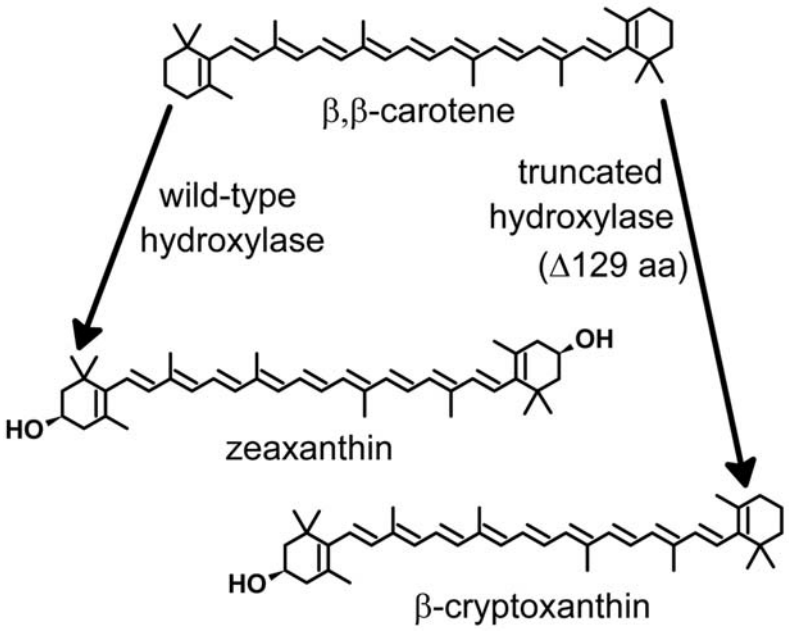


Figure 1.9. Alteration of product specificity of the β,β-carotene hydroxylase from *Arabidopsis thaliana.* The wild-type enzyme catalyzes two hydroxylation steps, converting β,β-carotene to zeaxanthin when expressed in either *A. thaliana* or *E. coli*. When the N-terminal 129 amino acids (aa) were removed, the truncated enzyme expressed in *E. coli* catalyzed only one hydroxylation step, leading primarily to synthesis of β-cryptoxanthin (197).

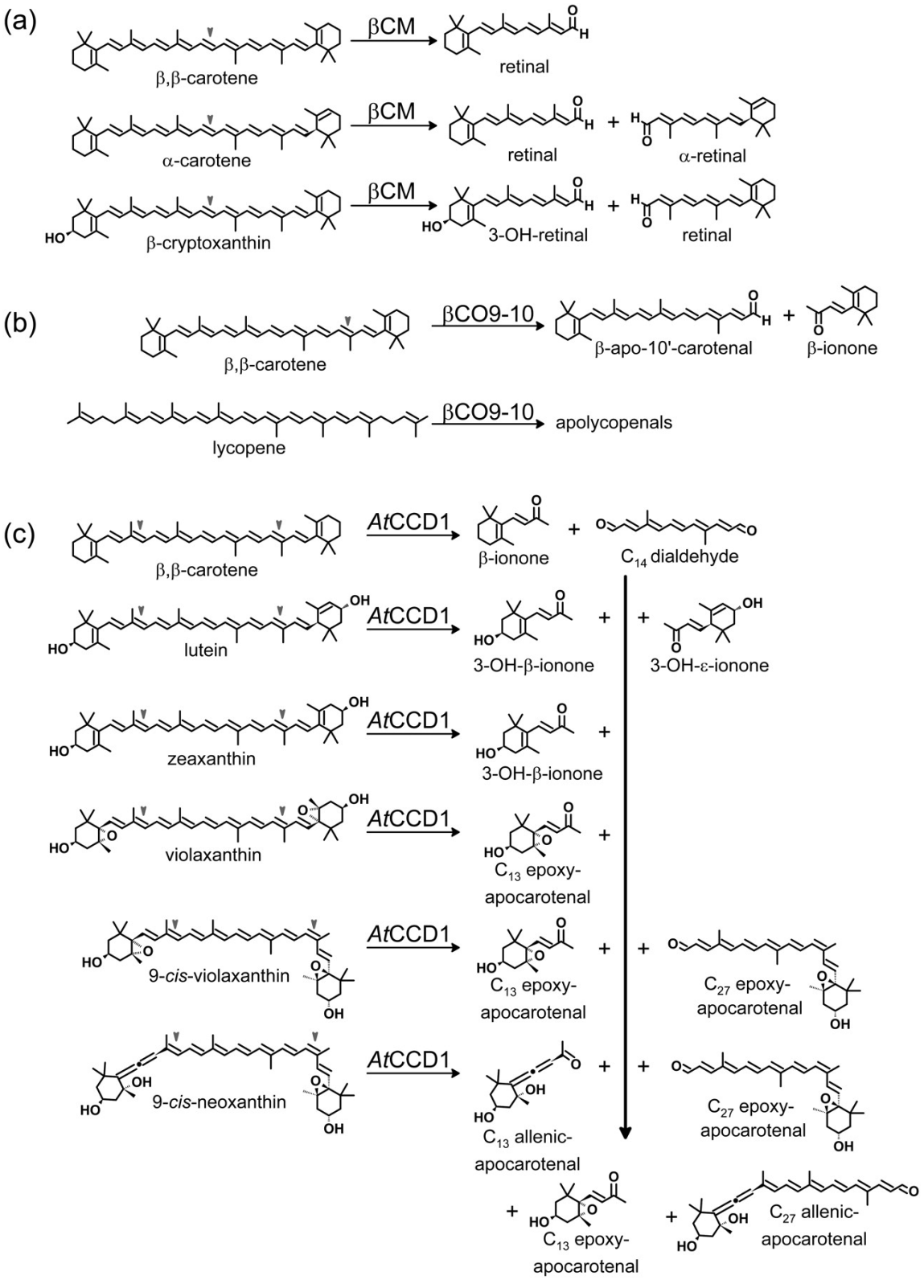

(a)
βCM
β,β-carotene
retinal
βCM
α-carotene
retinal
α-retinal
βCM
β-cryptoxanthin
3-OH-retinal
retinal
(b)
βCO9-10
β,β-carotene
β-apo-10'-carotenal
β-ionone
βCO9-10
lycopene
apolycopenals
(c)
AtCCD1
β,β-carotene
β-ionone
C14 dialdehyde
AtCCD1
lutein
3-OH-β-ionone
3-OH-ε-ionone
AtCCD1
zeaxanthin
3-OH-β-ionone
AtCCD1
violaxanthin
C13 epoxy-apocarotenal
AtCCD1
9-cis-violaxanthin
C13 epoxy-apocarotenal
C27 epoxy-apocarotenal
AtCCD1
9-cis-neoxanthin
C13 allenic-apocarotenal
C27 epoxy-apocarotenal
C13 epoxy-apocarotenal
C27 allenic-apocarotenal

Figure 1.10

Figure 1.10. Several carotenoid cleavage enzymes display localized specificity. Gray arrowheads show bonds cleaved. ***(a)*** β,β-carotene-15,15'-monooxygenase (βCM) can cleave substrates with an unsubstituted β-end (123, 228). ***(b)*** Murine β,β-carotene-9',10'-oxygenase (βCO9-10) can also cleave lycopene, leading to uncharacterized apolycopenals (98). ***(c)*** One carotenoid cleavage dioxygenase from *A. thaliana* (*At*CCD1) cleaves at least six different carotenoid substrates, leading to a variety of products (185). The downward arrow from C_{14} dialdehyde indicates that this product was detected from all six substrates shown.

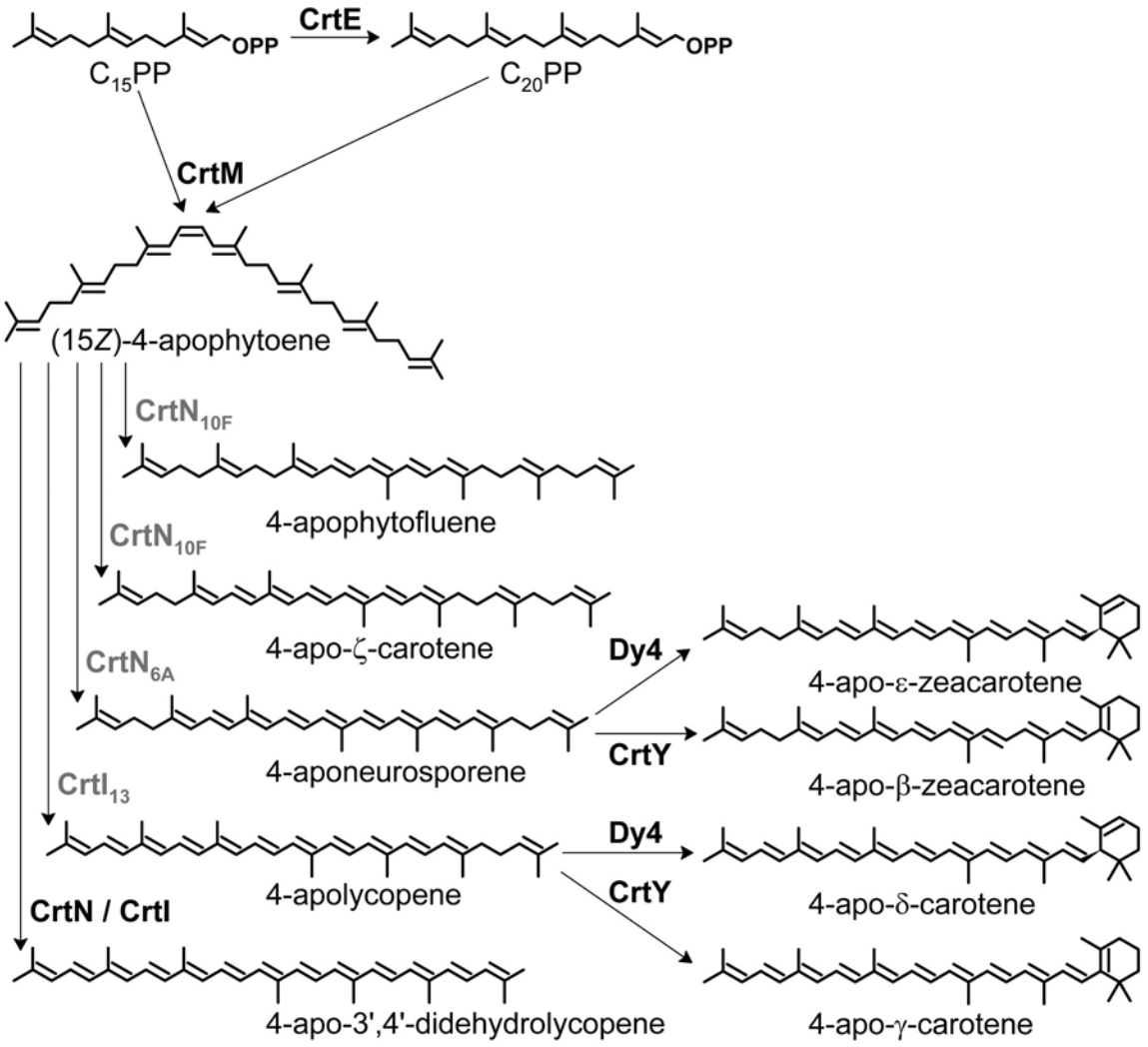


Figure 1.11. Generation of a C_{35} carotenoid pathway by gene assembly and directed enzyme evolution. (211) When supplied with C_{15}PP and C_{20}PP, the C_{30} carotenoid synthase CrtM produces the C_{35} carotenoid backbone, 4-apophytoene. By coexpressing C_{30} or C_{40} desaturases and mutants thereof (gray), different clones could be found that produce each possible acyclic C_{35} carotenoid as the main product. *E. uredovora* (*P. ananatis)* β-cyclase (CrtY) and *L. sativa* ε-cyclase (Dy4) were shown to cyclize two of the C_{35} substrates, leading to four different cyclic C_{35} carotenoids. All ten C_{35} carotenoids in this figure had never been previously reported.

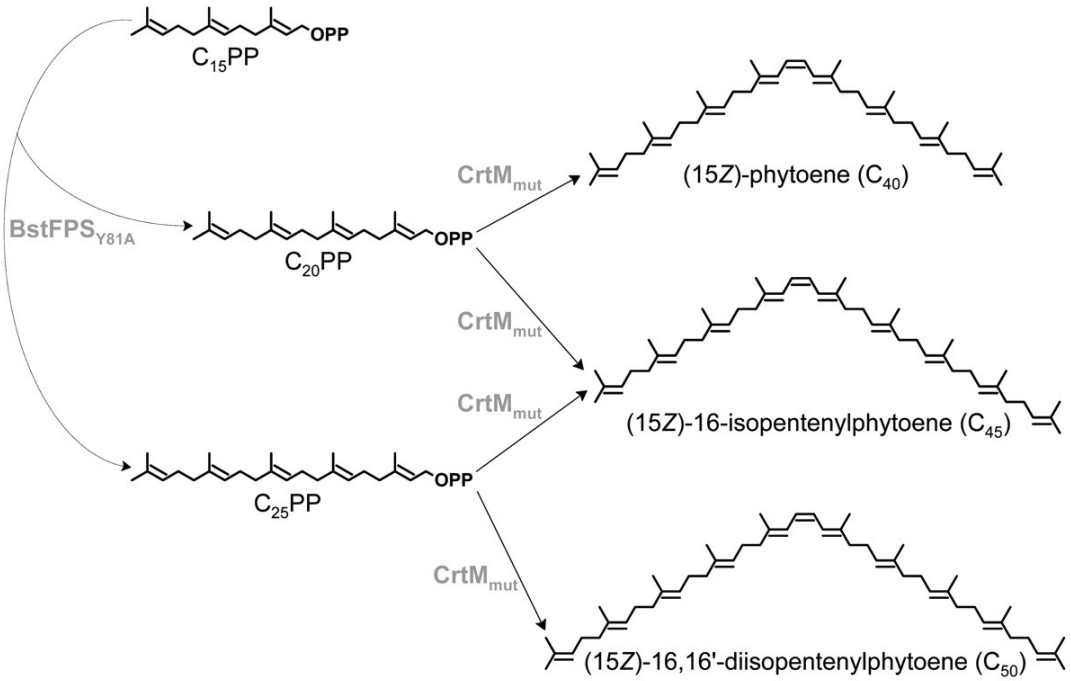


Figure 1.12. Biosynthesis of C_{45} and C_{50} carotenoid backbones. (212) Expression of the Y81A variant of the C_{15}PP synthase from *B. stearothermophilus* ($BstFPS_{Y81A}$) led to the production of C_{20}PP and C_{25}PP in *E. coli*. Additional coexpression of variants of the C_{30} carotenoid synthase CrtM mutated at F26 and/or W38 ($CrtM_{mut}$) resulted in biosynthesis of the C_{45} carotenoid backbone 16-isopentenylphytoene (C_{20}+C_{25}) and the C_{50} backbone 16,16'-diisopentenylphytoene (C_{25}+C_{25}). Both of these larger carotenoid backbones had never been previously reported. These backbones can be further metabolized by the desaturase CrtI, leading to new families of unnatural carotenoid pigments (see Chapter 3).

REFERENCES

1. **Abecassis, V., D. Pompon, and G. Truan.** 2000. High efficiency family shuffling based on multi-step PCR and *in vivo* DNA recombination in yeast: statistical and functional analysis of a combinatorial library between human cytochrome P450 1A1 and 1A2. Nucleic Acids Res. **28:**E88.
2. **Abecassis, V., D. Pompon, and G. Truan.** 2003. Producing chimeric genes by CLERY: *in vitro* and *in vivo* recombination, p. 165-173. *In* F. H. Arnold and G. Georgiou (ed.), Methods Mol. Biol., vol. 231. Directed Evolution Library Creation Methods and Protocols. Humana Press, Totowa, N.J.
3. **Aharoni, A., L. C. Keizer, H. J. Bouwmeester, Z. Sun, M. Alvarez-Huerta, H. A. Verhoeven, J. Blaas, A. M. van Houwelingen, R. C. De Vos, H. van der Voet, R. C. Jansen, M. Guis, J. Mol, R. W. Davis, M. Schena, A. J. van Tunen, and A. P. O'Connell.** 2000. Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. Plant Cell **12:**647-62.
4. **Albrecht, M., S. Takaichi, N. Misawa, G. Schnurr, P. Böger, and G. Sandmann.** 1997. Synthesis of atypical cyclic and acyclic hydroxy carotenoids in *Escherichia coli* transformants. J. Biotechnol. **58:**177-185.
5. **Albrecht, M., S. Takaichi, S. Steiger, Z. Y. Wang, and G. Sandmann.** 2000. Novel hydroxycarotenoids with improved antioxidative properties produced by gene combination in *Escherichia coli*. Nat. Biotechnol. **18:**843-846.
6. **Aragón, C. M., F. J. Murillo, M. D. de la Guardia, and E. Cerdá-Olmedo.** 1976. An enzyme complex for the dehydrogenation of phytoene in *Phycomyces*. Eur. J. Biochem. **63:**71-5.
7. **Armstrong, G. A.** 1999. Carotenoid genetics and biochemistry, p. 321-352. *In* D. E. Cane (ed.). Isoprenoids including carotenoids and steroids, vol. 2., Elsevier, Amsterdam.
8. **Armstrong, G. A., and J. E. Hearst.** 1996. Carotenoids 2: Genetics and molecular biology of carotenoid pigment biosynthesis. Faseb J. **10:**228-237.
9. **Arnold, F. H., and G. Georgiou (ed.).** 2003. Methods Mol. Biol., vol. 231. Directed Evolution Library Creation Methods and Protocols. Humana Press, Totowa, N.J.
10. **Arnold, F. H., P. C. Wintrode, K. Miyazaki, and A. Gershenson.** 2001. How enzymes adapt: Lessons from directed evolution. Trends Biochem. Sci. **26:**100-106.
11. **Arrach, N., R. Fernández-Martin, E. Cerdá-Olmedo, and J. Ávalos.** 2001. A single gene for lycopene cyclase, phytoene synthase, and regulation of carotene biosynthesis in *Phycomyces*. Proc. Natl. Acad. Sci. U.S.A. **98:**1687-1692.
12. **Arrach, N., T. J. Schmidhauser, and J. Ávalos.** 2002. Mutants of the carotene cyclase domain of *al-2* from *Neurospora crassa*. Mol. Genet. Genomics **266:**914-21.
13. **Ausich, R. L.** 1994. Production of carotenoids by recombinant-DNA technology. Pure Appl. Chem. **66:**1057-1062.
14. **Ávalos, J., J. Casadesús, and E. Cerdá-Olmedo.** 1985. *Gibberella fujikuroi* mutants obtained with UV radiation and N-methyl-N'-nitro-N-nitrosoguanidine. Appl. Environ. Microbiol. **49:**187-91.

15. **Ávalos, J., and E. Cerdá-Olmedo.** 1987. Carotenoid mutants of *Gibberella fujikuroi*. Curr. Genet. **11:**505-511.
16. **Badenhop, F., S. Steiger, M. Sandmann, and G. Sandmann.** 2003. Expression and biochemical characterization of the 1-HO-carotenoid methylase CrtF from *Rhodobacter capsulatus*. FEMS Microbiol. Lett. **222:**237-42.
17. **Barkovich, R., and J. C. Liao.** 2001. Metabolic engineering of isoprenoids. Metab. Eng. **3:**27-39.
18. **Bejarano, E. R., N. S. Govind, and E. Cerdá-Olmedo.** 1987. Zeta-carotene and other carotenes in a *Phycomyces* mutant. Phytochemistry **26:**2251-2254.
19. **Bendich, A.** 1994. Recent advances in clinical research involving carotenoids. Pure Appl. Chem. **66:**1017-1024.
20. **Bergquist, P. L., M. D. Gibbs, D. D. Morris, V. S. Te'o, D. J. Saul, and H. W. Moran.** 1999. Molecular diversity of thermophilic cellulolytic and hemicellulolytic bacteria. FEMS Microbiol. Ecol. **28:**99-110.
21. **Bertram, J. S.** 1999. Carotenoids and gene regulation. Nutr. Rev. **57:**182-191.
22. **Blagg, B. S. J., M. B. Jarstfer, D. H. Rogers, and C. D. Poulter.** 2002. Recombinant squalene synthase. A mechanism for the rearrangement of presqualene diphosphate to squalene. J. Am. Chem. Soc. **124:**8846-8853.
23. **Blanch, H. W., and D. S. Clark.** 1997. Product Recovery, p. 453-577. Biochemical Engineering. Marcel Dekker, Inc., New York.
24. **Bleicher, K. H., H. J. Bohm, K. Muller, and A. I. Alanine.** 2003. Hit and lead generation: beyond high-throughput screening. Nat. Rev. Drug Discov. **2:**369-78.
25. **Bohlmann, J., M. Phillips, V. Ramachandiran, S. Katoh, and R. Croteau.** 1999. cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the Tpsd gene family from grand fir (*Abies grandis*). Arch. Biochem. Biophys. **368:**232-43.
26. **Boswell, H. D., B. Drager, W. R. McLauchlan, A. Portsteffen, D. J. Robins, R. J. Robins, and N. J. Walton.** 1999. Specificities of the enzymes of N-alkyltropane biosynthesis in *Brugmansia* and *Datura*. Phytochemistry **52:**871-8.
27. **Boucher, Y., and W. F. Doolittle.** 2000. The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. Mol. Microbiol. **37:**703-716.
28. **Bouvier, F., O. Dogbo, and B. Camara.** 2003. Biosynthesis of the food and cosmetic plant pigment bixin (annatto). Science **300:**2089-2091.
29. **Bramley, P. M.** 1985. The *in vitro* biosynthesis of carotenoids. Adv. Lipid Res. **21:**243-279.
30. **Breinbauer, R., I. R. Vetter, and H. Waldmann.** 2002. From protein domains to drug candidates—natural products as guiding principles in the design and synthesis of compound libraries. Angew. Chem. Int. Ed. Engl. **41:**2879-90.
31. **Breitenbach, J., and G. Sandmann.** 2005. Zeta-carotene *cis* isomers as products and substrates in the plant poly-*cis* carotenoid biosynthetic pathway to lycopene. Planta **220:**785-793.
32. **Britton, G.** 1998. Overview of carotenoid biosynthesis, p. 13-147. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids, vol. 3: Biosynthesis and Metabolism. Birkhäuser Verlag, Basel.
33. **Britton, G.** 1995. UV/Visible Spectroscopy, p. 13-62. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids, vol. 1B: Spectroscopy. Birkhäuser

Verlag, Basel.

34. **Brodelius, M., A. Lundgren, P. Mercke, and P. E. Brodelius.** 2002. Fusion of farnesyldiphosphate synthase and epi-aristolochene synthase, a sesquiterpene cyclase involved in capsidiol biosynthesis in *Nicotiana tabacum*. Eur. J. Biochem. **269:**3570-3577.
35. **Buckingham, J.** 2004. Dictionary of Natural Products. Chapman & Hall/CRC Press. http://www.chemnetbase.com/scripts/dnpweb.exe.
36. **Bu'Lock, J. D.** 1965. The Biosynthesis of Natural Products. McGraw Hill, London.
37. **Bulter, T., and M. Alcalde.** 2003. Preparing libraries in *Saccharomyces cerevisiae*, p. 17-22. *In* F. H. Arnold and G. Georgiou (ed.), Methods Mol. Biol., vol. 231. Directed Evolution Library Creation Methods and Protocols. Humana Press, Totowa, N.J.
38. **Bulter, T., M. Alcalde, V. Sieber, P. Meinhold, C. Schlachtbauer, and F. H. Arnold.** 2003. Functional expression of a fungal laccase in *Saccharomyces cerevisiae* by directed evolution. Appl. Environ. Microbiol. **69:**987-95.
39. **Cadwell, R. C., and G. F. Joyce.** 1994. Mutagenic PCR. PCR Methods Appl. **3:**S136-S140.
40. **Candau, R., E. R. Bejarano, and E. Cerdá-Olmedo.** 1991. *In vivo* channeling of substrates in an enzyme aggregate for beta-carotene biosynthesis. Proc. Natl. Acad. Sci. U.S.A. **88:**4936-40.
41. **Cerdá-Olmedo, E.** 1985. Carotene mutants of *Phycomyces*. Methods Enzymol. **110:**220-243.
42. **Chappell, J.** 1995. Biochemistry and molecular biology of the isoprenoid biosynthetic pathway in plants. Annu. Rev. Plant Physiol. Plant Molec. Biol. **46:**521-547.
43. **Chasan-Taber, L., W. C. Willett, J. M. Seddon, M. J. Stampfer, B. Rosner, G. A. Colditz, F. E. Speizer, and S. E. Hankinson.** 1999. A prospective study of carotenoid and vitamin A intakes and risk of cataract extraction in US women. Am. J. Clin. Nutr. **70:**509-516.
44. **Chen, K., and F. H. Arnold.** 1993. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. Proc. Natl. Acad. Sci. U.S.A. **90:**5618-22.
45. **Chen, Z., B. S. Katzenellenbogen, J. A. Katzenellenbogen, and H. Zhao.** 2004. Directed evolution of human estrogen receptor variants with significantly enhanced androgen specificity and affinity. J. Biol. Chem. **(advance electronic publication)**.
46. **Cherry, J. R., M. H. Lamsa, P. Schneider, J. Vind, A. Svendsen, A. Jones, and A. H. Pedersen.** 1999. Directed evolution of a fungal peroxidase. Nat. Biotechnol. **17:**379-84.
47. **Colby, S. M., J. Crock, B. Dowdle-Rizzo, P. G. Lemaux, and R. Croteau.** 1998. Germacrene C synthase from *Lycopersicon esculentum* cv. VFNT cherry tomato: cDNA isolation, characterization, and bacterial expression of the multiple product sesquiterpene cyclase. Proc. Natl. Acad. Sci. U.S.A. **95:**2216-21.
48. **Collins, A. R.** 2001. Carotenoids and genomic stability. Mutat. Res.-Fundam. Mol. Mech. Mutagen. **475:**21-28.

49. **Cornish, K., and D. J. Siler.** 1996. Alternative natural rubber. Chemtech **26:**38-44.
50. **Crameri, A., G. Dawes, E. Rodriguez, Jr., S. Silver, and W. P. Stemmer.** 1997. Molecular evolution of an arsenate detoxification pathway by DNA shuffling. Nat. Biotechnol. **15:**436-8.
51. **Crock, J., M. Wildung, and R. Croteau.** 1997. Isolation and bacterial expression of a sesquiterpene synthase cDNA clone from peppermint (*Mentha x piperita*, L.) that produces the aphid alarm pheromone *(E)*-beta-farnesene. Proc. Natl. Acad. Sci. U.S.A. **94:**12833-12838.
52. **Croteau, R., F. Karp, K. C. Wagschal, D. M. Satterwhite, D. C. Hyatt, and C. B. Skotland.** 1991. Biochemical characterization of a spearmint mutant that resembles peppermint in monoterpene content. Plant Physiol. **96:**744-752.
53. **Cunningham, F. X., and E. Gantt.** 2001. One ring or two? Determination of ring number in carotenoids by lycopene epsilon-cyclases. Proc. Natl. Acad. Sci. U.S.A. **98:**2905-2910.
54. **Cunningham, F. X., G. C. T. Jiang, and E. Gantt.** 1997. Analysis of structure and function of carotenoid cyclase enzymes. Plant Physiol. **114:**926-926.
55. **Cunningham, F. X., B. Pogson, Z. R. Sun, K. A. McDonald, D. DellaPenna, and E. Gantt.** 1996. Functional analysis of the beta and epsilon lycopene cyclase enzymes of *Arabidopsis* reveals a mechanism for control of cyclic carotenoid formation. Plant Cell **8:**1613-1626.
56. **Davies, B. H.** 1977. C_{30} Carotenoids, p. 51-100. *In* T. W. Goodwin (ed.), vol. 14. Biochemistry of Lipids II. University Park Press, Baltimore.
57. **Davisson, V. J., and C. D. Poulter.** 1993. Farnesyl-diphosphate synthase—Interplay between substrate topology, stereochemistry, and regiochemistry in electrophilic alkylations. J. Am. Chem. Soc. **115:**1245-1260.
58. **De la Guardia, M. D., C. M. Aragón, F. J. Murillo, and E. Cerdá-Olmedo.** 1971. A carotenogenic enzyme aggregate in *Phycomyces*: evidence from quantitive complementation. Proc. Natl. Acad. Sci. U.S.A. **68:**2012-5.
59. **Dewick, P. M.** 2002. The biosynthesis of C_5-C_{25} terpenoid compounds. Nat. Prod. Rep. **19:**181-222.
60. **Facchini, P. J., and J. Chappell.** 1992. Gene family for an elicitor-induced sesquiterpene cyclase in tobacco. Proc. Natl. Acad. Sci. U.S.A. **89:**11088-92.
61. **Fernandez, S. M. S., B. A. Kellogg, and C. D. Poulter.** 2000. Farnesyl diphosphate synthase. Altering the catalytic site to select for geranyl diphosphate activity. Biochemistry **39:**15316-15321.
62. **Firn, R. D., and C. G. Jones.** 2000. The evolution of secondary metabolism—a unifying model. Mol. Microbiol. **37:**989-994.
63. **Firn, R. D., and C. G. Jones.** 1996. An explanation of secondary product "redundancy," p. 295-312. *In* J. T. Romeo, J. A. Saunders, and P. Barbosa (ed.), Recent Advances in Phytochemistry, vol. 30. Phytochemical Diversity and Redundancy in Ecological Interactions. Plenum Press, New York.
64. **Firn, R. D., and C. G. Jones.** 2003. Natural products—a simple model to explain chemical diversity. Nat. Prod. Rep. **20:**382-391.
65. **Fraser, P. D., N. Misawa, H. Linden, S. Yamano, K. Kobayashi, and G. Sandmann.** 1992. Expression in *Escherichia coli*, purification, and reactivation

of the recombinant *Erwinia uredovora* phytoene desaturase. J. Biol. Chem. **267:**19891-19895.

66. **Fraser, P. D., M. J. Ruiz-Hidalgo, M. A. Lopez-Matas, M. I. Alvarez, A. P. Eslava, and P. M. Bramley.** 1996. Carotenoid biosynthesis in wild type and mutant strains of *Mucor circinelloides*. Biochim. Biophys. Acta **1289:**203-8.
67. **Fujisaki, S., H. Hara, Y. Nishimura, K. Horiuchi, and T. Nishino.** 1990. Cloning and nucleotide-sequence of the *ispA* gene responsible for farnesyl diphosphate synthase activity in *Escherichia coli*. J. Biochem. (Tokyo) **108:**995-1000.
68. **Fujisaki, S., T. Nishino, H. Katsuki, H. Hara, Y. Nishimura, and Y. Hirota.** 1989. Isolation and characterization of an *Escherichia coli* mutant having temperature-sensitive farnesyl diphosphate synthase. J. Bacteriol. **171:**5654-5658.
69. **Garcia-Asua, G., R. J. Cogdell, and C. N. Hunter.** 2002. Functional assembly of the foreign carotenoid lycopene into the photosynthetic apparatus of *Rhodobacter sphaeroides*, achieved by replacement of the native 3-step phytoene desaturase with its 4-step counterpart from *Erwinia herbicola*. Mol. Microbiol. **44:**233-244.
70. **Garcia-Asua, G., H. P. Lang, R. J. Cogdell, and C. N. Hunter.** 1998. Carotenoid diversity: a modular role for the phytoene desaturase step. Trends Plant Sci. **3:**445-449.
71. **Giuliano, G., S. Al-Babili, and J. von Lintig.** 2003. Carotenoid oxygenases: cleave it or leave it. Trends Plant Sci. **8:**145-9.
72. **Giuliano, G., L. Giliberto, and C. Rosati.** 2002. Carotenoid isomerase: a tale of light and isomers. Trends Plant Sci. **7:**427-9.
73. **Goodwin, T. W., and L. A. Griffiths.** 1952. Studies in carotenogenesis. V. Carotene production by various mutants of *Phycomyces blakesleeanus* and by *Phycomyces nitens*. Biochem. J. **52:**499-501.
74. **Goto, S., K. Kogure, K. Abe, Y. Kimata, K. Kitahama, E. Yamashita, and H. Terada.** 2001. Efficient radical trapping at the surface and inside the phospholipid membrane is responsible for highly potent antiperoxidative activity of the carotenoid astaxanthin. Biochim. Biophys. Acta-Biomembr. **1512:**251-258.
75. **Guo, R. T., C. J. Kuo, C. C. Chou, T. P. Ko, H. L. Shr, P. H. Liang, and A. H. Wang.** 2004. Crystal structure of octaprenyl pyrophosphate synthase from hyperthermophilic *Thermotoga maritima* and mechanism of product chain length determination. J. Biol. Chem. **279:**4903-12.
76. **Harker, M., and J. Hirschberg.** 1997. Biosynthesis of ketocarotenoids in transgenic cyanobacteria expressing the algal gene for beta-C-4-oxygenase, *crtO*. FEBS Lett. **404:**129-134.
77. **Hemmi, H., M. Noike, T. Nakayama, and T. Nishino.** 2003. An alternative mechanism of product chain-length determination in type III geranylgeranyl diphosphate synthase. Eur. J. Biochem. **270:**2186-94.
78. **Hemmi, H., M. Noike, T. Nakayama, and T. Nishino.** 2002. Change of product specificity of hexaprenyl diphosphate synthase from *Sulfolobus solfataricus* by introducing mimetic mutations. Biochem. Biophys. Res. Commun. **297:**1096-101.
79. **Hirayama, O., K. Nakamura, S. Hamada, and Y. Kobayasi.** 1994. Singlet oxygen quenching ability of naturally-occurring carotenoids. Lipids **29:**149-150.

80. **Hirooka, K., T. Kato, J. Matsu-ura, H. Hemmi, and T. Nishino.** 2000. The role of histidine-114 of *Sulfolobus acidocaldarius* geranylgeranyl diphosphate synthase in chain-length determination. FEBS Lett. **481:**68-72.
81. **Hirooka, K., S. Ohnuma, A. Koike-Takeshita, T. Koyama, and T. Nishino.** 2000. Mechanism of product chain length determination for heptaprenyl diphosphate synthase from *Bacillus stearothermophilus*. Eur. J. Biochem. **267:**4520-4528.
82. **Hornero-Méndez, D., and G. Britton.** 2002. Involvement of NADPH in the cyclization reaction of carotenoid biosynthesis. FEBS Lett. **515:**133-136.
83. **Hughes, D. A.** 1999. Effects of carotenoids on human immune function. Proc. Nutr. Soc. **58:**713-718.
84. **Hugueney, P., A. Badillo, H. C. Chen, A. Klein, J. Hirschberg, B. Camara, and M. Kuntz.** 1995. Metabolism of cyclic carotenoids—a model for the alteration of this biosynthetic pathway in *Capsicum annuum* chromoplasts. Plant J. **8:**417-424.
85. **Hunter, C. N., B. S. Hundle, J. E. Hearst, H. P. Lang, A. T. Gardiner, S. Takaichi, and R. J. Cogdell.** 1994. Introduction of new carotenoids into the bacterial photosynthetic apparatus by combining the carotenoid biosynthetic pathways of *Erwinia herbicola* and *Rhodobacter sphaeroides*. J. Bacteriol. **176:**3692-3697.
86. **Isaacson, T., I. Ohad, P. Beyer, and J. Hirschberg.** 2004. Analysis *in vitro* of the enzyme CRTISO establishes a poly-*cis*-carotenoid biosynthesis pathway in plants. Plant Physiol. **136:**4246-4255.
87. **Iwata-Reuyl, D., S. K. Math, S. B. Desai, and C. D. Poulter.** 2003. Bacterial phytoene synthase: molecular cloning, expression, and characterization of *Erwinia herbicola* phytoene synthase. Biochemistry **42:**3359-65.
88. **Jarstfer, M. B., B. S. J. Blagg, D. H. Rogers, and C. D. Poulter.** 1996. Biosynthesis of squalene. Evidence for a tertiary cyclopropylcarbinyl cationic intermediate in the rearrangement of presqualene diphosphate to squalene. J. Am. Chem. Soc. **118:**13089-13090.
89. **Jarstfer, M. B., D. L. Zhang, and C. D. Poulter.** 2002. Recombinant squalene synthase. Synthesis of non-head-to-tail isoprenoids in the absence of NADPH. J. Am. Chem. Soc. **124:**8834-8845.
90. **Jensen, R. A.** 1976. Enzyme recruitment in evolution of new function. Annu. Rev. Microbiol. **30:**409-25.
91. **Jones, C. G., and R. D. Firn.** 1991. On the evolution of plant secondary chemical diversity. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. **333:**273-280.
92. **Kainou, T., K. Okada, K. Suzuki, T. Nakagawa, H. Matsuda, and M. Kawamukai.** 2001. Dimer formation of octaprenyl-diphosphate synthase (IspB) is essential for chain length determination of ubiquinone. J. Biol. Chem. **276:**7876-83.
93. **Karrer, P., and C. H. Eugster.** 1951. Carotinoidsynthesen 8. Synthese des dodecapreno-beta-carotins. Helv. Chim. Acta **34:**1805-1814.
94. **Kato, J. I., S. Fujisaki, K. I. Nakajima, Y. Nishimura, M. Sato, and A. Nakano.** 1999. The *Escherichia coli* homologue of yeast Rer2, a key enzyme of dolichol synthesis, is essential for carrier lipid formation in bacterial cell wall

synthesis. J. Bacteriol. **181:**2733-2738.

95. **Kaur, C., and H. C. Kapoor.** 2001. Antioxidants in fruits and vegetables—the millennium's health. Int. J. Food Sci. Technol. **36:**703-725.
96. **Kawasaki, T., Y. Hamano, T. Kuzuyama, N. Itoh, H. Seto, and T. Dairi.** 2003. Interconversion of the product specificity of type I eubacterial farnesyl diphosphate synthase and geranylgeranyl diphosphate synthase through one amino acid substitution. J. Biochem. (Tokyo) **133:**83-91.
97. **Kellogg, B. A., and C. D. Poulter.** 1997. Chain elongation in the isoprenoid biosynthetic pathway. Curr. Opin. Chem. Biol. **1:**570-578.
98. **Kiefer, C., S. Hessel, J. M. Lampert, K. Vogt, M. O. Lederer, D. E. Breithaupt, and J. von Lintig.** 2001. Identification and characterization of a mammalian enzyme catalyzing the asymmetric oxidative cleavage of provitamin A. J. Biol. Chem. **276:**14110-14116.
99. **Kim, B. S., D. H. Sherman, and K. A. Reynolds.** 2004. An efficient method for creation and functional analysis of libraries of hybrid type I polyketide synthases. Prot. Eng. Des. Sel. **17:**277-284.
100. **Kleinig, H., and R. Schmitt.** 1982. On the biosynthesis of C_{30} carotenoic acid glucosyl esters in *Pseudomonas rhodos*—Analysis of *car*-mutants. Z. Naturforsch. (C) **37:**758-760.
101. **Kleinig, H., R. Schmitt, W. Meister, G. Englert, and H. Thommen.** 1979. New C_{30} Carotenoic Acid Glucosyl Esters from *Pseudomonas rhodos*. Z. Naturforsch. (C) **34:**181-185.
102. **Kobayashi, M., T. Koyama, K. Ogura, S. Seto, F. J. Ritter, and I. E. M. Bruggemannrotgans.** 1980. Bioorganic synthesis and absolute-configuration of faranal. J. Am. Chem. Soc. **102:**6602-6604.
103. **Kobayashi, M., and Y. Sakamoto.** 1999. Singlet oxygen quenching ability of astaxanthin esters from the green alga *Haematococcus pluvialis*. Biotechnol. Lett. **21:**265-269.
104. **Kojima, N., W. Sitthithaworn, E. Viroonchatapan, D. Y. Suh, N. Iwanami, T. Hayashi, and U. Sankaw.** 2000. Geranylgeranyl diphosphate synthases from *Scoparia dulcis* and *Croton sublyratus*. cDNA cloning, functional expression, and conversion to a farnesyl diphosphate synthase. Chem. Pharm. Bull. (Tokyo). **48:**1101-3.
105. **Komori, M., R. Ghosh, S. Takaichi, Y. Hu, T. Mizoguchi, Y. Koyama, and M. Kuki.** 1998. A null lesion in the rhodopin 3,4-desaturase of *Rhodospirillum rubrum* unmasks a cryptic branch of the carotenoid biosynthetic pathway. Biochemistry **37:**8987-8994.
106. **Koyama, T., K. Ogura, and S. Seto.** 1977. Construction of a chiral center by use of stereospecificity of prenyltransferase. J. Am. Chem. Soc. **99:**1999-2000.
107. **Koyama, T., A. Saito, K. Ogura, and S. Seto.** 1980. Substrate-specificity of farnesylpyrophosphate synthetase. Application to asymmetric synthesis. J. Am. Chem. Soc. **102:**3614-3618.
108. **Krinsky, N. I.** 1994. The Biological properties of carotenoids. Pure Appl. Chem. **66:**1003-1010.
109. **Krubasik, P., M. Kobayashi, and G. Sandmann.** 2001. Expression and functional analysis of a gene cluster involved in the synthesis of decaprenoxanthin

reveals the mechanisms for C_{50} carotenoid formation. Eur. J. Biochem. **268:**3702-3708.

110. **Krubasik, P., and G. Sandmann.** 2000. A carotenogenic gene cluster from *Brevibacterium linens* with novel lycopene cyclase genes involved in the synthesis of aromatic carotenoids. Mol. Gen. Genet. **263:**423-432.
111. **Krubasik, P., and G. Sandmann.** 2000. Molecular evolution of lycopene cyclases involved in the formation of carotenoids with ionone end groups. Biochem. Soc. Trans. **28:**806-810.
112. **Krubasik, P., S. Takaichi, T. Maoka, M. Kobayashi, K. Masamoto, and G. Sandmann.** 2001. Detailed biosynthetic pathway to decaprenoxanthin diglucoside in *Corynebacterium glutamicum* and identification of novel intermediates. Arch. Microbiol. **176:**217-23.
113. **Kushwaha, S. C., M. Kates, and H. J. Weber.** 1980. Exclusive formation of all-trans-phytoene by a colorless mutant of *Halobacterium halobium*. Can. J. Microbiol. **26:**1011-4.
114. **Kwon, S. J., A. L. de Boer, R. Petri, and C. Schmidt-Dannert.** 2003. High-level production of porphyrins in metabolically engineered *Escherichia coli*: Systematic extension of a pathway assembled from overexpressed genes involved in heme biosynthesis. Appl. Environ. Microbiol. **69:**4875-83.
115. **Lakshman, M. R.** 2004. Alpha and omega of carotenoid cleavage. J. Nutr. **134:**241S-245S.
116. **Lange, B. M., T. Rujan, W. Martin, and R. Croteau.** 2000. Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. Proc. Natl. Acad. Sci. U.S.A. **97:**13172-13177.
117. **Lauro, G. J.** 1991. A primer on natural colors. Cereal Foods World **36:**949-953.
118. **Lee, P. C., A. Z. R. Momen, B. N. Mijts, and C. Schmidt-Dannert.** 2003. Biosynthesis of structurally novel carotenoids in *Escherichia coli*. Chem. Biol. **10:**453-462.
119. **Lee, P. C., and C. Schmidt-Dannert.** 2002. Metabolic engineering towards biotechnological production of carotenoids in microorganisms. Appl. Microbiol. Biotechnol. **60:**1-11.
120. **Lee, T. C., D. B. Rodriguez, I. Karasawa, T. H. Lee, K. L. Simpson, and C. O. Chichester.** 1975. Chemical alteration of carotene biosynthesis in *Phycomyces blakesleeanus* and mutants. Appl. Microbiol. **30:**988-93.
121. **Leuenberger, M. G., C. Engeloch-Jarret, and W. D. Woggon.** 2001. The reaction mechanism of the enzyme-catalyzed central cleavage of beta-carotene to retinal. Angew. Chem. Int. Ed. Engl. **40:**2613-2617.
122. **Liang, P. H., T. P. Ko, and A. H. J. Wang.** 2002. Structure, mechanism and function of prenyltransferases. Eur. J. Biochem. **269:**3339-3354.
123. **Lindqvist, A., and S. Andersson.** 2002. Biochemical properties of purified recombinant human beta-carotene 15,15'-monooxygenase. J. Biol. Chem. **277:**23942-23948.
124. **LinGoerke, J. L., D. J. Robbins, and J. D. Burczak.** 1997. PCR-based random mutagenesis using manganese and reduced dNTP concentration. Biotechniques **23:**409-&.
125. **Maki, Y., M. Komabayashi, Y. Gotoh, N. Ohya, H. Hemmi, K. Hirooka, T.**

Nishino, and T. Koyama. 2002. Dramatic changes in the substrate specificities of prenyltransferase by a single amino acid substitution. J. Mol. Catal. B-Enzym. **19:**431-436.

126. **Maki, Y., A. Masukawa, H. Ono, T. Endo, T. Koyama, and K. Ogura.** 1995. Substrate-specificity of farnesyl diphosphate synthase from *Bacillus stearothermophilus*. Bioorg. Med. Chem. Lett. **5:**1605-1608.
127. **Mann, J.** 1987. Secondary Metabolism. Clarendon Press, Oxford.
128. **Mann, V., M. Harker, I. Pecker, and J. Hirschberg.** 2000. Metabolic engineering of astaxanthin production in tobacco flowers. Nat. Biotechnol. **18:**888-892.
129. **Marshall, J. H., and G. J. Wilmoth.** 1981. Pigments of *Staphylococcus aureus*, a series of triterpenoid carotenoids. J. Bacteriol. **147:**900-13.
130. **Marshall, J. H., and G. J. Wilmoth.** 1981. Proposed pathway of triterpenoid carotenoid biosynthesis in *Staphylococcus aureus*: evidence from a study of mutants. J. Bacteriol. **147:**914-9.
131. **Masamoto, K., H. Wada, T. Kaneko, and S. Takaichi.** 2001. Identification of a gene required for *cis*-to-*trans* carotene isomerization in carotenogenesis of the cyanobacterium *Synechocystis* sp. PCC 6803. Plant Cell Physiol. **42:**1398-402.
132. **Mehta, G., and V. Singh.** 2002. Hybrid systems through natural product leads: an approach towards new molecular entities. Chem. Soc. Rev. **31:**324-334.
133. **Misawa, N., S. Kajiwara, K. Kondo, A. Yokoyama, Y. Satomi, T. Saito, W. Miki, and T. Ohtani.** 1995. Canthaxanthin biosynthesis by the conversion of methylene to keto groups in a hydrocarbon beta-carotene by a single gene. Biochem. Biophys. Res. Commun. **209:**867-876.
134. **Misawa, N., M. Nakagawa, K. Kobayashi, S. Yamano, Y. Izawa, K. Nakamura, and K. Harashima.** 1990. Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. J. Bacteriol. **172:**6704-6712.
135. **Misawa, N., Y. Satomi, K. Kondo, A. Yokoyama, S. Kajiwara, T. Saito, T. Ohtani, and W. Miki.** 1995. Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway proposed at the gene level. J. Bacteriol. **177:**6575-6584.
136. **Misawa, N., S. Yamano, H. Linden, M. R. Defelipe, M. Lucas, H. Ikenaga, and G. Sandmann.** 1993. Functional expression of the *Erwinia uredovora* carotenoid biosynthesis gene *crtl* in transgenic plants showing an increase of beta-carotene biosynthesis activity and resistance to the bleaching herbicide norflurazon. Plant J. **4:**833-840.
137. **Montenegro, M. A., M. A. Nazareno, E. N. Durantini, and C. D. Borsarelli.** 2002. Singlet molecular oxygen quenching ability of carotenoids in a reverse-micelle membrane mimetic system. Photochem. Photobiol. **75:**353-361.
138. **Murillo, F. J., S. Torres-Martinez, C. M. Aragón, and E. Cerdá-Olmedo.** 1981. Substrate transfer in carotene biosynthesis in *Phycomyces*. Eur. J. Biochem. **119:**511-6.
139. **Nagaki, M., S. Sato, Y. Maki, T. Nishino, and T. Koyama.** 2000. Artificial substrates for undecaprenyl diphosphate synthase from *Micrococcus luteus* B-P 26. J. Mol. Catal. B-Enzym. **9:**33-38.

140. **Nagaki, M., A. Takaya, Y. Maki, J. Ishibashi, Y. Kato, T. Nishino, and T. Koyama.** 2000. One-pot syntheses of the sex pheromone homologs of a codling moth, *Laspeyresia promonella* L. J. Mol. Catal. B-Enzym. **10:**517-522.
141. **Narita, K., S. Ohnuma, and T. Nishino.** 1999. Protein design of geranyl diphosphate synthase. Structural features that define the product specificities of prenyltransferases. J. Biochem. (Tokyo) **126:**566-571.
142. **Newman, D. J., G. M. Cragg, and K. M. Snader.** 2003. Natural products as sources of new drugs over the period 1981-2002. J. Nat. Prod. **66:**1022-1037.
143. **Nishino, H.** 1997. Cancer prevention by natural carotenoids. J. Cell. Biochem.**:**86-91.
144. **Nishino, H., H. Tokuda, Y. Satomi, M. Masuda, P. Bu, M. Onozuka, S. Yamaguchi, Y. Okuda, J. Takayasu, J. Tsuruta, M. Okuda, E. Ichiishi, M. Murakoshi, T. Kato, N. Misawa, T. Narisawa, N. Takasuka, and M. Yano.** 1999. Cancer prevention by carotenoids. Pure Appl. Chem. **71:**2273-2278.
145. **Ogura, K., and T. Koyama.** 1998. Enzymatic aspects of isoprenoid chain elongation. Chem. Rev. **98:**1263-1276.
146. **Ogura, K., T. Koyama, and H. Sagami.** 1997. Polyprenyl diphosphate synthases, p. 57-87. *In* R. Bittman (ed.). Subcellular Biochemistry 28: Cholesterol: Its function and metabolism in biology and medicine., Plenum Press, New York.
147. **Ohnuma, S., H. Hemmi, T. Koyama, K. Ogura, and T. Nishino.** 1998. Recognition of allylic substrates in *Sulfolobus acidocaldarius* geranylgeranyl diphosphate synthase: Analysis using mutated enzymes and artificial allylic substrates. J. Biochem. (Tokyo) **123:**1036-1040.
148. **Ohnuma, S., K. Hirooka, H. Hemmi, C. Ishida, C. Ohto, and T. Nishino.** 1996. Conversion of product specificity of archaebacterial geranylgeranyl-diphosphate synthase—Identification of essential amino acid residues for chain length determination of prenyltransferase reaction. J. Biol. Chem. **271:**18831-18837.
149. **Ohnuma, S., K. Hirooka, C. Ohto, and T. Nishino.** 1997. Conversion from archaeal geranylgeranyl diphosphate synthase to farnesyl diphosphate synthase—Two amino acids before the first aspartate-rich motif solely determine eukaryotic farnesyl diphosphate synthase activity. J. Biol. Chem. **272:**5192-5198.
150. **Ohnuma, S., K. Hirooka, N. Tsuruoka, M. Yano, C. Ohto, H. Nakane, and T. Nishino.** 1998. A pathway where polyprenyl diphosphate elongates in prenyltransferase—Insight into a common mechanism of chain length determination of prenyltransferases. J. Biol. Chem. **273:**26705-26713.
151. **Ohnuma, S., K. Narita, T. Nakazawa, C. Ishida, Y. Takeuchi, C. Ohto, and T. Nishino.** 1996. A role of the amino acid residue located on the fifth position before the first aspartate-rich motif of farnesyl diphosphate synthase on determination of the final product. J. Biol. Chem. **271:**30748-30754.
152. **Ohnuma, S. I., T. Nakazawa, H. Hemmi, A. M. Hallberg, T. Koyama, K. Ogura, and T. Nishino.** 1996. Conversion from farnesyl diphosphate synthase to geranylgeranyl diphosphate synthase by random chemical mutagenesis. J. Biol. Chem. **271:**10087-10095.
153. **Okada, K., T. Kainou, K. Tanaka, T. Nakagawa, H. Matsuda, and M. Kawamukai.** 1998. Molecular cloning and mutational analysis of the ddsA gene encoding decaprenyl diphosphate synthase from *Gluconobacter suboxydans*. Eur.

J. Biochem. **255:**52-9.

154. **Okada, K., M. Minehira, X. F. Zhu, K. Suzuki, T. Nakagawa, H. Matsuda, and M. Kawamukai.** 1997. The *ispB* gene encoding octaprenyl diphosphate synthase is essential for growth of *Escherichia coli*. J. Bacteriol. **179:**3058-3060.
155. **Ootaki, T., A. C. Lighty, M. Delbrück, and W. J. Hsu.** 1973. Complementation between mutants of *Phycomyces* deficient with respect to carotenogenesis. Mol. Gen. Genet. **121:**57-70.
156. **Ortiz de Montellano, P. R., J. S. Wei, W. A. Vinson, R. Castillo, and A. S. Boparai.** 1977. Substrate selectivity of squalene synthetase. Biochemistry **16:**2680-2685.
157. **Ouchane, S., M. Picaud, C. Vernotte, F. ReissHusson, and C. Astier.** 1997. Pleiotropic effects of puf interposon mutagenesis on carotenoid biosynthesis in *Rubrivivax gelatinosus*—A new gene organization in purple bacteria. J. Biol. Chem. **272:**1670-1676.
158. **Paik, J., A. During, E. H. Harrison, C. L. Mendelsohn, K. Lai, and W. S. Blaner.** 2001. Expression and characterization of a murine enzyme able to cleave beta-carotene—The formation of retinoids. J. Biol. Chem. **276:**32160-32168.
159. **Pandit, J., D. E. Danley, G. K. Schulte, S. Mazzalupo, T. A. Pauly, C. M. Hayward, E. S. Hamanaka, J. F. Thompson, and H. J. Harwood.** 2000. Crystal structure of human squalene synthase—A key enzyme in cholesterol biosynthesis. J. Biol. Chem. **275:**30610-30617.
160. **Perry, K. L., T. A. Simonitch, K. J. Harrisonlavoie, and S. T. Liu.** 1986. Cloning and regulation of *Erwinia herbicola* pigment genes. J. Bacteriol. **168:**607-612.
161. **Pfeifer, B. A., and C. Khosla.** 2001. Biosynthesis of polyketides in heterologous hosts. Microbiol. Mol. Biol. Rev. **65:**106-18.
162. **Powls, R., and G. Britton.** 1977. A series of mutant strains of *Scenedesmus obliquus* with abnormal carotenoid compositions. Arch. Microbiol. **113:**275-80.
163. **Raisig, A., and G. Sandmann.** 2001. Functional properties of diapophytoene and related desaturases of C_{30} and C_{40} carotenoid biosynthetic pathways. Biochim. Biophys. Acta Mol. Cell Biol. Lipids **1533:**164-170.
164. **Rau, W., and U. Mitzka-Schnabel.** 1985. Carotenoid synthesis in *Neurospora crassa*. Methods Enzymol. **110:**253-67.
165. **Redmond, T. M., S. Gentleman, T. Duncan, S. Yu, B. Wiggert, E. Gantt, and F. X. Cunningham.** 2001. Identification, expression, and substrate specificity of a mammalian beta-carotene 15,15'-dioxygenase. J. Biol. Chem. **276:**6560-6565.
166. **Reetz, M. T., M. H. Becker, H.-W. Klein, and D. Stöckigt.** 1999. A method for high-throughput screening of enantioselective catalysts. Angew. Chem. Int. Ed. Engl. **38:**1758-1761.
167. **Reetz, M. T., A. Eipper, P. Tielmann, and R. Mynott.** 2002. A practical NMR-based high-throughput assay for screening enantioselective catalysts and biocatalysts. Adv. Synth. Catal. **344:**1008-1016.
168. **Reetz, M. T., C. Torre, A. Eipper, R. Lohmer, M. Hermes, B. Brunner, A. Maichele, M. Bocola, M. Arand, A. Cronin, Y. Genzel, A. Archelas, and R. Furstoss.** 2004. Enhancing the enantioselectivity of an epoxide hydrolase by directed evolution. Org. Lett. **6:**177-80.

169. **Reeves, C. D.** 2003. The enzymology of combinatorial biosynthesis. Crit. Rev. Biotechnol. **23:**95-147.
170. **Rice-Evans, C. A., J. Sampson, P. M. Bramley, and D. E. Holloway.** 1997. Why do we expect carotenoids to be antioxidants *in vivo*? Free Radic. Res. **26:**381-398.
171. **Rison, S. C., and J. M. Thornton.** 2002. Pathway evolution, structurally speaking. Curr. Opin. Struct. Biol. **12:**374-82.
172. **Rouhi, A. M.** 2003. Rediscovering Natural Products. Chem. Eng. News **81(41):**77-91.
173. **Sager, R., and M. Zalokar.** 1958. Pigments and photosynthesis in a carotenoid-deficient mutant of *Chlamydomonas*. Nature **182:**98-100.
174. **Sandmann, G.** 1994. Carotenoid biosynthesis in microorganisms and plants. Eur. J. Biochem. **223:**7-24.
175. **Sandmann, G.** 2002. Combinatorial biosynthesis of carotenoids in a heterologous host: A powerful approach for the biosynthesis of novel structures. ChemBioChem **3:**629-635.
176. **Sandmann, G.** 2002. Molecular evolution of carotenoid biosynthesis from bacteria to plants. Physiol. Plantarum **116:**431-440.
177. **Sandmann, G., M. Albrecht, G. Schnurr, O. Knorzer, and P. Boger.** 1999. The biotechnological potential and design of novel carotenoids by gene combination in *Escherichia coli*. Trends Biotechnol. **17:**233-237.
178. **Saperstein, S., and M. P. Starr.** 1954. The ketonic carotenoid canthaxanthin isolated from a colour mutant of *Corynebacterium michiganense*. Biochem. J. **57:**273-5.
179. **Satomi, Y., T. Yoshida, K. Aoki, N. Misawa, M. Masuda, M. Murakoshi, N. Takasuka, T. Sugimura, and H. Nishino.** 1995. Production of phytoene, an oxidative stress protective carotenoid, in mammalian cells by introduction of phytoene synthase gene *crtB* isolated from a bacterium *Erwinia uredovora*. Proc. Jpn. Acad. Ser. B-Phys. Biol. Sci. **71:**236-240.
180. **Schmidt-Dannert, C.** 2000. Engineering novel carotenoids in microorganisms. Curr. Opin. Biotechnol. **11:**255-261.
181. **Schmidt-Dannert, C., D. Umeno, and F. H. Arnold.** 2000. Molecular breeding of carotenoid biosynthetic pathways. Nat. Biotechnol. **18:**750-753.
182. **Schrader, W., A. Eipper, D. J. Pugh, and M. T. Reetz.** 2002. Second-generation MS-based high-throughput screening system for enantioselective catalysts and biocatalysts. Can. J. Chem. **80:**626-632.
183. **Schroeder, W. A., and E. A. Johnson.** 1995. Carotenoids protect *Phaffia rhodozyma* against singlet oxygen damage. J. Ind. Microbiol. **14:**502-507.
184. **Schwab, W.** 2003. Metabolome diversity: too few genes, too many metabolites? Phytochemistry **62:**837-49.
185. **Schwartz, S. H., X. Qin, and J. A. Zeevaart.** 2001. Characterization of a novel carotenoid cleavage dioxygenase from plants. Journal of Biological Chemisry **276:**25208-11.
186. **Schwartz, S. H., B. C. Tan, D. A. Gage, J. A. D. Zeevaart, and D. R. McCarty.** 1997. Specific oxidative cleavage of carotenoids by VP14 of maize. Science **276:**1872-1874.

187. **Shafikhani, S., R. A. Siegel, E. Ferrari, and V. Schellenberger.** 1997. Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization. Biotechniques **23:**304-10.

188. **Shalit, M., I. Guterman, H. Volpin, E. Bar, T. Tamari, N. Menda, Z. Adam, D. Zamir, A. Vainstein, D. Weiss, E. Pichersky, and E. Lewinsohn.** 2003. Volatile ester formation in roses. Identification of an acetyl-coenzyme A. Geraniol/Citronellol acetyltransferase in developing rose petals. Plant Physiol. **131:**1868-76.

189. **Smolke, C. D., V. J. J. Martin, and J. D. Keasling.** 2001. Controlling the metabolic flux through the carotenoid pathway using directed mRNA processing and stabilization. Metab. Eng. **3:**313-321.

190. **Steele, C. L., J. Crock, J. Bohlmann, and R. Croteau.** 1998. Sesquiterpene synthases from grand fir (*Abies grandis*)—Comparison of constitutive and wound-induced activities, and cDNA isolation, characterization and bacterial expression of delta-selinene synthase and gamma-humulene synthase. J. Biol. Chem. **273:**2078-2089.

191. **Steiger, S., A. Mazet, and G. Sandmann.** 2003. Heterologous expression, purification, and enzymatic characterization of the acyclic carotenoid 1,2-hydratase from *Rubrivivax gelatinosus*. Arch. Biochem. Biophys. **414:**51-8.

192. **Steiger, S., S. Takaichi, and G. Sandmann.** 2002. Heterologous production of two unusual acyclic carotenoids, 1,1'-dihydroxy-3,4-didehydrolycopene and 1-hydroxy-3,4,3',4'-tetradehydrolycopene by combination of the *crtC* and *crtD* genes from *Rhodobacter* and *Rubrivivax*. J. Biotechnol. **97:**51-8.

193. **Stemmer, W. P.** 1994. Rapid evolution of a protein *in-vitro* by DNA shuffling. Nature **370:**389-391.

194. **Stickforth, P., S. Steiger, W. R. Hess, and G. Sandmann.** 2003. A novel type of lycopene epsilon-cyclase in the marine cyanobacterium *Prochlorococcus marinus* MED4. Arch. Microbiol. **179:**409-415.

195. **Su, V., and B. D. Hsu.** 2003. Cloning and expression of a putative cytochrome P450 gene that influences the colour of *Phalaenopsis* flowers. Biotechnol. Lett. **25:**1933-9.

196. **Subden, R. E., and S. F. Threlkeld.** 1968. Genetic and complementation studies of a new carotenoid mutant of *Neurospora crassa*. Can. J. Genet. Cytol. **10:**351-6.

197. **Sun, Z. R., E. Gantt, and F. X. Cunningham.** 1996. Cloning and functional analysis of the beta-carotene hydroxylase of *Arabidopsis thaliana*. J. Biol. Chem. **271:**24349-24352.

198. **Takaichi, S., K. Inoue, M. Akaike, M. Kobayashi, H. Ohoka, and M. T. Madigan.** 1997. The major carotenoid in all known species of heliobacteria is the C_{30} carotenoid 4,4'-diaponeurosporene, not neurosporene. Arch. Microbiol. **168:**277-281.

199. **Takaichi, S., G. Sandmann, G. Schnurr, Y. Satomi, A. Suzuki, and N. Misawa.** 1996. The carotenoid 7,8-dihydro-psi end group can be cyclized by the lycopene cyclases from the bacterium *Erwinia uredovora* and the higher plant *Capsicum annuum*. Eur. J. Biochem. **241:**291-296.

200. **Takaichi, S., and K. Shimada.** 1999. Pigment composition of two pigment-protein complexes derived from anaerobically and semi-aerobically grown

Rubrivivax gelatinosus, and identification of a new keto-carotenoid, 2-ketospirilloxanthin. Plant Cell Physiol. **40:**613-617.

201. **Takaichi, S., Y. Tamura, T. Miyamoto, K. Azegami, Y. Yamamoto, and J. Ishidsu.** 1999. Carotenogenesis pathway of novel carotenoid glucoside mycolic acid esters in *Rhodococcus rhodochrous* using carotenogenesis mutants and inhibitors. Biosci. Biotechnol. Biochem. **63:**2014-2016.
202. **Takatsuji, H., T. Nishino, K. Izui, and H. Katsuki.** 1982. Formation of dehydrosqualene catalyzed by squalene synthetase in *Saccharomyces cerevisiae*. J. Biochem. (Tokyo) **91:**911-921.
203. **Tarshis, L. C., P. J. Proteau, B. A. Kellogg, J. C. Sacchettini, and C. D. Poulter.** 1996. Regulation of product chain length by isoprenyl diphosphate synthases. Proc. Natl. Acad. Sci. U.S.A. **93:**15018-15023.
204. **Tarshis, L. C., M. J. Yan, C. D. Poulter, and J. C. Sacchettini.** 1994. Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-angstrom resolution. Biochemistry **33:**10871-10877.
205. **Taylor, R. F.** 1984. Bacterial triterpenoids. Microbiol. Rev. **48:**181-198.
206. **Taylor, R. F., and B. H. Davies.** 1976. Triterpenoid carotenoids and related lipids—Triterpenoid carotenoid aldehydes from *Streptococcus faecium* Unh 564p. Biochem. J. **153:**233-239.
207. **Teramoto, M., S. Takaichi, Y. Inomata, H. Ikenaga, and N. Misawa.** 2003. Structural and functional analysis of a lycopene beta-monocyclase gene isolated from a unique marine bacterium that produces myxol. FEBS Lett. **545:**120-126.
208. **Tietze, L. F., H. P. Bell, and S. Chandrasekhar.** 2003. Natural product hybrids as new leads for drug discovery. Angew. Chem. Int. Ed. Engl. **42:**3996-4028.
209. **Timberlake, C. F., and B. S. Henry.** 1986. Plant pigments as natural food colors. Endeavour **10:**31-36.
210. **Tuveson, R. W., R. A. Larson, and J. Kagan.** 1988. Role of cloned carotenoid genes expressed in *Escherichia coli* in protecting against inactivation by near-UV light and specific phototoxic molecules. J. Bacteriol. **170:**4675-4680.
211. **Umeno, D., and F. H. Arnold.** 2003. A C_{35} carotenoid biosynthetic pathway. Appl. Environ. Microbiol. **69:**3573-3579.
212. **Umeno, D., and F. H. Arnold.** 2004. Evolution of a pathway to novel long-chain carotenoids. J. Bacteriol. **186:**1531-1536.
213. **Umeno, D., K. Hiraga, and F. H. Arnold.** 2003. Method to protect a targeted amino acid residue during random mutagenesis. Nucleic Acids Res. **31:**e91.
214. **Umeno, D., A. V. Tobias, and F. H. Arnold.** 2002. Evolution of the C_{30} carotenoid synthase CrtM for function in a C_{40} pathway. J. Bacteriol. **184:**6690-6699.
215. **Valadon, L. R. G., and R. S. Mummery.** 1977. Natural beta-apo-4'-carotenoic acid methyl ester in the fungus *Verticillium agaricinum*. Phytochemistry **16:**613-614.
216. **Van Dien, S. J., C. J. Marx, B. N. O'Brien, and M. E. Lidstrom.** 2003. Genetic characterization of the carotenoid biosynthetic pathway in *Methylobacterium extorquens* AM1 and isolation of a colorless mutant. Appl. Environ. Microbiol. **69:**7563-6.
217. **Vesterlund, S., J. Paltta, A. Laukova, M. Karp, and A. C. Ouwehand.** 2004.

Rapid screening method for the detection of antimicrobial substances. J. Microbiol. Methods **57:**23-31.

218. **Viljanen, K., S. Sundberg, T. Ohshima, and M. Heinonen.** 2002. Carotenoids as antioxidants to prevent photooxidation. Eur. J. Lipid Sci. Technol. **104:**353-359.
219. **von Lintig, J., and K. Vogt.** 2000. Filling the gap in vitamin A research—Molecular identification of an enzyme cleaving beta-carotene to retinal. J. Biol. Chem. **275:**11915-11920.
220. **von Lintig, J., and A. Wyss.** 2001. Molecular analysis of vitamin a formation: Cloning and characterization of beta-carotene 15,15'-dioxygenases. Arch. Biochem. Biophys. **385:**47-52.
221. **Walters, W. P., and M. Namchuk.** 2003. Designing screens: how to make your hits a hit. Nat. Rev. Drug Discov. **2:**259-266.
222. **Wang, C. W., and J. C. Liao.** 2001. Alteration of product specificity of *Rhodobacter sphaeroides* phytoene desaturase by directed evolution. J. Biol. Chem. **276:**41161-41164.
223. **Wang, C. W., M. K. Oh, and J. C. Liao.** 2000. Directed evolution of metabolically engineered *Escherichia coli* for carotenoid production. Biotechnol. Prog. **16:**922-926.
224. **Wang, C. W., M. K. Oh, and J. C. Liao.** 1999. Engineered isoprenoid pathway enhances astaxanthin production in *Escherichia coli*. Biotechnol. Bioeng. **62:**235-241.
225. **Wang, K., and S. Ohnuma.** 1999. Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. Trends Biochem. Sci. **24:**445-451.
226. **Wieland, B., C. Feil, E. Gloriamaercker, G. Thumm, M. Lechner, J. M. Bravo, K. Poralla, and F. Gotz.** 1994. Genetic and biochemical analyses of the biosynthesis of the yellow carotenoid 4,4'-diaponeurosporene of *Staphylococcus aureus*. J. Bacteriol. **176:**7719-7726.
227. **Windhovel, U., B. Geiges, G. Sandmann, and P. Boger.** 1994. Expression of *Erwinia uredovora* phytoene desaturase in *Synechococcus* Pcc7942 leading to resistance against a bleaching herbicide. Plant Physiol. **104:**119-125.
228. **Wirtz, G. M., C. Bornemann, A. Giger, R. K. Muller, H. Schneider, G. Schlotterbeck, G. Schiefer, and W. D. Woggon.** 2001. The substrate specificity of beta,beta-carotene 15,15'-monooxygenase. Helv. Chim. Acta **84:**2301-2315.
229. **Wise, M. L., T. J. Savage, E. Katahira, and R. Croteau.** 1998. Monoterpene synthases from common sage (*Salvia officinalis*). cDNA isolation, characterization, and functional expression of (+)-sabinene synthase, 1,8-cineole synthase, and (+)-bornyl diphosphate synthase. J. Biol. Chem. **273:**14891-9.
230. **Woodall, A. A., S. W. M. Lee, R. J. Weesie, M. J. Jackson, and G. Britton.** 1997. Oxidation of carotenoids by free radicals: relationship between structure and reactivity. Biochim. Biophys. Acta-Gen. Subj. **1336:**33-42.
231. **Wyss, A., G. Wirtz, W. D. Woggon, R. Brugger, M. Wyss, A. Friedlein, H. Bachmann, and W. Hunziker.** 2000. Cloning and expression of beta,beta-carotene 15,15'-dioxygenase. Biochem. Biophys. Res. Commun. **271:**334-336.
232. **Yan, W. M., G. F. Jang, F. Haeseleer, N. Esumi, J. H. Chang, M. Kerrigan, M. Campochiaro, P. Campochiaro, K. Palczewski, and D. J. Zack.** 2001. Cloning

and characterization of a human beta,beta-carotene-15,15'-dioxygenase that is highly expressed in the retinal pigment epithelium. Genomics **72:**193-202.

233. **Ycas, M.** 1974. On earlier states of the biochemical system. J. Theor. Biol. **44:**145-60.

234. **Ye, X. D., S. Al-Babili, A. Kloti, J. Zhang, P. Lucca, P. Beyer, and I. Potrykus.** 2000. Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science **287:**303-305.

235. **Yokoyama, A., Y. Shizuri, and N. Misawa.** 1998. Production of new carotenoids, astaxanthin glucosides, by *Escherichia coli* transformants carrying carotenoid biosynthetic genes. Tetrahedron Lett. **39:**3709-3712.

236. **Zeevaart, J. A. D., and R. A. Creelman.** 1988. Metabolism and physiology of abscisic acid. Annu. Rev. Plant Physiol. Plant Molec. Biol. **39:**439-473.

237. **Zhang, D. L., and C. D. Poulter.** 1995. Biosynthesis of non-head-to-tail isoprenoids—Synthesis of 1'-1-structures and 1'-3-structures by recombinant yeast squalene synthase. J. Am. Chem. Soc. **117:**1641-1642.

238. **Zhang, Y. W., X. Y. Li, and T. Koyama.** 2000. Chain length determination of prenyltransferases: Both heteromeric subunits of medium-chain (E)-prenyl diphosphate synthase are involved in the product chain length determination. Biochemistry **39:**12717-12722.

239. **Zhang, Y. W., X. Y. Li, H. Sugawara, and T. Koyama.** 1999. Site-directed mutagenesis of the conserved residues in component I of *Bacillus subtilis* heptaprenyl diphosphate synthase. Biochemistry **38:**14638-14643.

240. **Zhao, H., and F. H. Arnold.** 1999. Directed evolution converts subtilisin E into a functional equivalent of thermitase. Protein Eng. **12:**47-53.

CHAPTER 2

Evolution of the C_{30} Carotenoid Synthase CrtM for Function in a C_{40} Pathway

Material from this chapter appears in *Evolution of the C_{30} Carotenoid Synthase CrtM for Function in a C_{40} Pathway*, Daisuke Umeno, Alexander V. Tobias, and Frances H. Arnold, Journal of Bacteriology 184(23): 6690-6699 (2002) and is reprinted with permission from the American Society for Microbiology.

SUMMARY

The C_{30} carotene synthase CrtM from *Staphylococcus aureus* and the C_{40} carotene synthase CrtB from *Erwinia uredovora* were swapped into their respective foreign C_{40} and C_{30} biosynthetic pathways (heterologously expressed in *Escherichia coli*) and evaluated for function. Each displayed negligible ability to synthesize the natural carotenoid product of the other. After one round of mutagenesis and screening, we isolated 116 variants of CrtM able to synthesize C_{40} carotenoids. In contrast, we failed to find a single variant of CrtB with detectable C_{30} activity. Subsequent analysis revealed that the best CrtM mutants performed comparably to CrtB in an *in vivo* C_{40} pathway and that Phe 26 of CrtM is a key specificity-determining residue. The CrtM mutants showed significant variation in performance in their original C_{30} pathway, indicating the emergence of enzymes with broadened substrate specificity as well as those with shifted specificity. We also report for the first time the isolation of carotenoid pigments based on a C_{35} carbon backbone formed by condensation of farnesyl diphosphate (FPP) and geranylgeranyl dipohosphate (GGPP) by wild-type and variants of CrtM. The expanded substrate and product range of the CrtM mutants reported herein highlights the potential for creating further new carotenoid backbone structures.

INTRODUCTION

A common feature of small molecule natural product biosynthetic pathways is the extensive use of enzymes with broad substrate and product specificities. Many isoprenoid biosynthetic enzymes, for example, accept a variety of both natural and unnatural substrates (27, 28, 31). Some have been shown to synthesize an impressively large number of compounds (sometimes >50) from a single substrate (9, 12, 45). Natural product biosynthetic pathways also use enzymes with remarkably stringent specificity. Such enzymes are frequently seen in key determinant positions, usually in the very early steps of a pathway, while promiscuous enzymes tend to be located further downstream. Thus, many natural product pathways have a "reverse tree" topology (3, 42) where the backbone structures of metabolites are dictated by a small number of stringent, upstream enzymes. When the substrate or product preferences of these key upstream enzymes are altered, pathway branches leading to sets of novel compounds may be opened (10). Our interest is to achieve the same by applying methods of directed enzyme evolution to recombinant pathways in *Escherichia coli* (43, 50).

Among the most widespread of all small molecule natural products, carotenoids are natural pigments that play important biological roles. Some are accessory light-harvesting components of photosynthetic systems, while others are photo-protecting antioxidants or regulators of membrane fluidity. Recent studies advocate their effectiveness in preventing cancer and heart disease (26) as well as their potential hormonal activity (4, 18). Such diverse molecular functions justify exploring rare or novel carotenoid structures. At present, ~700 carotenoids from the naturally occurring C_{30} and C_{40} carotenoid biosynthetic pathways have been characterized (17). Most natural

carotenoid diversity arises from differences in types and levels of desaturation and other modifications of the C_{40} backbone. C_{40} carotenoids are also much more widespread in nature than their C_{30} counterparts. The former are synthesized by thousands of plant and microbial species, whereas the latter are known only in a select few bacteria (46, 48). Homocarotenoids (carotenoids with >40 carbon atoms) and apocarotenoids (carotenoids with <40 carbon atoms), which result from the action of downstream enzymes on a C_{40} substrate, are also known. Although these structures do not have 40 carbon atoms, they are nonetheless derived from C_{40} carotenoid precursors (5).

The first committed step in carotenoid biosynthesis is the head-to-head condensation of two prenyl diphosphates catalyzed by a synthase enzyme (**Figure 2.1**). The C_{40} carotenoid phytoene is synthesized by the condensation of two molecules of geranylgeranyl diphosphate (GGPP, $C_{20}PP$) catalyzed by the synthase CrtB. (Most phytoene synthases produce the 15*Z* isomer of phytoene (5)). C_{30} carotenoids are synthesized via an independent route whereby two molecules of farnesyl diphosphate (FPP, $C_{15}PP$) are condensed to (15*Z*)-4,4'-diapophytoene by CrtM (48). The various downstream modification enzymes possess broad substrate specificity and therefore represent potential targets for generating biosynthetic routes to novel carotenoids. For example, when the three-step phytoene desaturase CrtI from *Rhodobacter sphaeroides* was replaced with a four-step enzyme from *Erwinia herbicola*, the cells accumulated a series of carotenoids produced neither by *Erwinia* nor *Rhodobacter* (16). Carotene desaturases (39), carotene cyclases (44), and β,β-carotene cleavage enzyme (40) have also been shown to accept a broad range of substrates. Combinatorial expression of such enzymes can create unusual and sometimes previously unidentified carotenoids (1, 2, 21).

Nevertheless, the greatest potential to further extend carotenoid biosynthetic diversity lies in creating whole new backbone structures, and therefore with engineering the carotene synthases.

The C_{30} and C_{40} pathways are very similar except in the sizes of their precursor molecules and their distributions in nature, and it is clear that they diverged from a common ancestral pathway. We would like to determine the minimal genetic change required in key carotenoid biosynthetic enzymes to create such new pathway branches. Can the enzymes that synthesize one carotenoid be modified in a laboratory evolution experiment to synthesize others? How much of carotenoid diversity can be accessed in this way? And, can *novel* pathways to different, even unnatural structures (e.g., C_{35}, C_{45}, C_{50}, or larger carotenoids) be accessed by using C_{30} or C_{40} enzymes as a starting point? To begin to answer these questions, we studied the performance of the C_{30} carotene synthase CrtM from *Staphylococcus aureus* in a C_{40} pathway and the C_{40} carotene synthase CrtB from *Erwinia uredovora* in a C_{30} pathway. We then examined the ability of these enzymes to adapt to their respective "foreign" pathways in order to assess the ease and uncover the mechanisms by which this might be accomplished.

RESULTS AND DISCUSSION

Constructing pathways for C_{30} and C_{40} carotenoids

To establish a recombinant C_{40} pathway in *E. coli*, we subcloned *crtE* encoding GGPP (C_{20}PP) synthase, *crtB* encoding phytoene synthase, and *crtI* encoding phytoene desaturase, all from *E. uredovora,* into a vector derived from pUC18, resulting in plasmid pUC-*crtE-crtB-crtI* (**Figure 2.2a**). For C_{30} carotenoid production, plasmid pUC-*fps-crtM-*

crtN was constructed by integrating the *E. coli* farnesyl diphosphate (C_{15}PP) synthase (FPS) gene, *fps,* with *crtM* (4,4'-diapophytoene synthase gene) and *crtN* (4,4'-diapophytoene desaturase gene) from *S. aureus* into the same vector. Plasmid pUC-*crtM-crtN* was constructed in an identical fashion, but it lacks *fps*. Each plasmid shares the cloning sites for the corresponding enzyme genes: *Eco*RI and *Xba*I sites flank isoprenyl diphosphate synthase (IDS) genes (*crtE* and *fps*), *Xba*I and *Xho*I sites flank carotene synthase genes (*crtB* and *crtM*), and *Xho*I and *Apa*I sites flank carotene desaturase genes (*crtI* and *crtN*) (**Figure 2.2a**). With this arrangement, corresponding genes could be easily swapped in order to evaluate their function in the other pathway.

E. coli cells harboring pUC-*crtE-crtB-crtI* plasmids (C_{40} pathway) developed a characteristic red-pink color, whereas cells possessing pUC-*fps-crtM-crtN* and pUC-*crtM-crtN* plasmids (C_{30} pathway) were yellow (**Figure 2.2b**). HPLC analysis of extracted pigments showed that XL1-Blue(pUC-*crtE-crtB-crtI*) cells produce mostly lycopene (four-step desaturation) along with a small amount (5 to 10% of total pigment) of 3,4,3',4'-tetradehydrolycopene (six-step desaturation) under the conditions described. Although CrtI is classified as a four-step desaturase, production of 3,4-didehydrolycopene (five-step desaturation) by CrtI both *in vivo* and *in vitro* has been reported (15, 23).

Several groups have expressed CrtM and CrtN from *S. aureus* in *E. coli* and observed almost exclusive production of 4,4'-diaponeurosporene (39, 51). However, in our XL1-Blue(pUC-*crtM-crtN*) expression system, the cells accumulated a significant amount of 4,4'-diapolycopene (~30% of total carotenoids). When a constitutive *lac* promoter (lacking the operator) was used for operon expression, the amount of 4,4'-

diapolycopene increased further and reached 50% of carotenoids produced (D. Umeno, unpublished data). This phenomenon was observed in all *E. coli* strains we tested, BL21, BL21(DE3), JM109, JM101, DH5α, HB101, SCS110, and XL10-Gold, and was insensitive to growth temperature and plasmid copy number. Thus, it is clear that the apparent desaturation step number of CrtN depends on its expression level and the effective concentration of substrates.

In *E. coli*, FPP (C_{15}PP) is a precursor to a variety of important housekeeping molecules such as respiratory quinones, prenylated tRNA, and dolichol. Concerned about retarding or preventing the growth of our recombinant C_{30} cultures due to depletion of FPP, we first expressed the *fps* gene along with *crtM* and *crtN*. However, we observed no difference in growth rate, pigmentation level, or carotenoid composition between XL1-Blue cells harboring the pUC-*crtM-crtN* plasmid and those harboring the pUC-*fps-crtM-crtN* plasmid. This demonstrates that endogenous FPP levels in *E. coli* suffice to support both growth and synthesis of C_{30} carotenoids.

Functional analysis of CrtM and CrtB swapped into their respective foreign pathways

To assess the function of wild-type CrtM in a C_{40} pathway, pUC-*crtE-crtM-crtI* was constructed and transformed into *E. coli* XL1-Blue. Under these circumstances, CrtM is supplied with GGPP (C_{20}PP) produced by CrtE. If CrtM were able to synthesize the C_{40} carotenoid phytoene from GGPP, subsequent desaturation by CrtI to lycopene would cause the cells to develop a red-pink color. However, while cells expressing the *crtE-crtB-crtI* operon exhibited the characteristic red-pink of lycopene (**Figure 2.2b**) and synthesized this carotenoid in liquid culture (**Figure 2.4**), cells expressing the *crtE-crtM-*

crtI operon had only very subtle pink-orange color on agar plates and synthesized much less lycopene in liquid culture (**Figure 2.4**). As did Raisig and Sandmann (39), we thus conclude that CrtM fails to complement CrtB in a C_{40} pathway and has very poor ability compared to CrtB to synthesize the C_{40} carotenoid backbone. This observation cannot be explained by simple competition between FPP and GGPP for access to CrtM (coupled with poor ability of CrtI to desaturate the C_{30} product) because XL1-Blue cells expressing the *crtE-crtM-crtN* operon showed only very minor yellow color development. This indicates that the availability of FPP for carotenoid biosynthesis is significantly reduced upon expression of CrtE.

The function of CrtB in a C_{30} pathway was examined by analyzing the pigmentation of XL1-Blue cells transformed with pUC-*crtB-crtN*. In this case, endogenous FPP is the only available prenyl diphosphate substrate for CrtB, since the level of GGPP is very low in *E. coli* compared with that of FPP. If CrtB could synthesize C_{30} carotenoids from FPP, the 4,4'-diapophytoene produced would be desaturated by CrtN and the cells would develop a yellow color. In contrast to the intense yellow of XL1-Blue transformed with pUC-*crtM-crtN*, XL1-Blue(pUC-*crtB-crtN*) showed no color development (**Figure 2.2b**). When expressed alone or with CrtN, CrtB synthesized C_{30} carotenoids very poorly in liquid culture (see **Figures 2.5 and 2.6**). We thus conclude that CrtB fails to complement CrtM in a C_{30} pathway.

We also analyzed the pigment produced by XL1-Blue carrying pUC-*fps-crtB-crtN*. In this case, the cells displayed a yellow color similar to that of XL1-Blue transformed with pUC-*fps-crtM-crtN* (**Figure 2.2b**). HPLC analysis of carotenoid extracts from XL1-Blue(pUC-*fps-crtB-crtN*) revealed the C_{40} carotenoid neurosporene,

with trace amounts of the C_{30} carotenoids 4,4'-diapolycopene and 4,4'-diaponeurosporene. Thus, CrtB seems to have at least some C_{30} activity in the presence of high levels of FPP. Production of the C_{40} carotenoid neurosporene as the major pigment is explained by the promiscuous nature of both FPS and CrtN. This was verified by our observation that cells expressing an *fps-crtB-crtI* operon had a weak pink hue and accumulated lycopene and 3,4,3',4'-tetradehydrolycopene. Thus, FPS can produce significant amounts of GGPP in *E. coli* when overexpressed. It is known that avian FPS also possesses weak ability to synthesize GGPP (41). Indeed, other studies have shown a varied product distribution and strong dependence on conditions for FPS enzymes (25, 34). Additionally, CrtN can accept phytoene as a substrate and introduce two or three double bonds (39). When we transformed XL1-Blue with pUC-*crtE-crtB-crtN*, the resulting cells exhibited significant yellow color (**Figure 2.2b**) and accumulated neurosporene, thus demonstrating the promiscuity of CrtN.

Screening for synthase function in a foreign pathway

In the previous section, we confirmed that the carotene synthases CrtB and CrtM show negligible activity in their respective foreign pathways. We next proceeded to evolve the two enzymes, with the goal of improving the function of each in the other's native pathway. To uncover CrtB variants with significant CrtM-like ability to synthesize C_{30} carotenoids, we constructed pUC-*[crtB]-crtN* libraries and transformed them into XL1-Blue cells. Here GGPP is not available, so cells with the wild-type *crtB-crtN* operon fail to develop color. Any variants of CrtB able to convert FPP (C_{15}PP) into 4,4'-diapophytoene would produce yellow colonies. Similarly, we searched for CrtM mutants with improved C_{40} activity by transforming four pUC-*crtE-[crtM]-crtI* libraries into XL1-

Blue. As described in the previous section, the amount of FPP available for carotenoid biosynthesis becomes significantly depleted when CrtE is overexpressed, resulting in negligible production of C_{30} carotenoids, even by native C_{30} enzymes. Cells expressing wild-type CrtM from the *crtE-crtM-crtI* operon showed a weak pink-orange color due to trace production of lycopene as well as C_{30} and C_{35} carotenoids, but CrtM mutants able to complement CrtB via acquisition of enhanced C_{40} activity would be expected to yield intensely red-pink colonies and thus be distinguishable on the plates.

Evolution of 4,4'-diapophytoene synthase (CrtM) for function in a C_{40} pathway

Four different mutagenic libraries of *crtM* corresponding to four different mutation rates were generated by performing error-prone PCR (54) on the entire 843-nucleotide *crtM* gene. PCR products from each reaction were ligated into the *Xba*I-*Xho*I site of pUC-*crtE-crtM-crtI* (**Figure 2.2a**), resulting in pUC-*crtE-[crtM]-crtI* libraries. **Figure 2.3** shows a typical agar plate covered with a nitrocellulose membrane upon which lie colonies of *E. coli* XL1-Blue expressing a *crtE-[crtM]-crtI* library. In each library, colonies with pale orange, pale yellow, or virtually no color dominated the population. The pale orange colonies express variants of CrtM with phenotypes similar to that of the wild-type enzyme in this context, while the pale and colorless colonies express severely or completely inactivated synthase mutants. More rare were colonies that developed an intense red-pink color, indicating significant production of C_{40} carotenoids and hence improved CrtB-like activity of CrtM. On average, about 0.5% of the colonies screened showed intense red-pink color. The highest frequency of positives was obtained from the library prepared by PCR with the lowest $MnCl_2$ concentration (0.02 mM). In this library, approximately 1 out of every 120 colonies was red-pink (0.8%). We screened

over 23,000 CrtM mutants from the four libraries and picked 116 positive clones. These clones were re-screened by stamping them onto an agar plate covered with a white nitrocellulose membrane. All 116 stamped clones exhibited red-pink coloration. We sequenced the 10 most intensely red stamped clones, as determined by visual assessment. Mutations found in these variants (**Table 2.1**) were heavily biased toward transitions: 36 versus only 3 transversions. Additionally, 87% of the base substitutions found by sequencing were A/T→G/C. This is a typical observation for PCR mutagenesis with $MnCl_2$ (7, 24). Out of 39 total nucleotide substitutions, 24 resulted in amino acid substitutions. Most notably, 9 of the 10 sequenced CrtM mutants had a mutation at phenylalanine 26.

Carotenoid production of evolved CrtM variants

We analyzed in detail the *in vivo* carotenoid production of variant M_8, which has the F26L mutation only; variant M_9, which has the F26S mutation; and variant M_{10}, which has no mutation at F26. To confirm the newly acquired CrtB-like function of these three sequenced variants of CrtM, XL1-Blue cultures carrying each of the three plasmids pUC-*crtE-M*$_8$-*crtI*, pUC-*crtE-M*$_9$-*crtI*, and pUC-*crtE-M*$_{10}$-*crtI* (collectively referred to as pUC-*crtE-M*$_{8\text{-}10}$-*crtI*) were cultivated in TB medium. Extracted pigments were analyzed by HPLC with a photodiode array detector (**Figure 2.4**). These analyses revealed that all three clones produced the C_{40} carotenoids lycopene (peak 5) and 3,4,3',4'-tetradehydrolycopene (peak 4) as major products, whereas cells harboring the parent pUC-*crtE-crtM-crtI* plasmid produced mainly C_{30} and trace amounts of C_{40} and C_{35} carotenoids. Two carotenoids with C_{35} backbone structures were detected in extracts from cells harboring pUC-*crtE-crtM-crtI* and pUC-*crtE-M*$_{8\text{-}10}$-*crtI*. Elution profiles, UV-visible

spectra, and mass spectra confirm that one of the structures is the fully conjugated C_{35} carotenoid 4-apo-3',4'-didehydrolycopene. We also detected C_{35} carotenoids with 11 conjugated double bonds. Because C_{35} carotenoids are asymmetric, there are two possible C_{35} structures that possess 11 conjugated double bonds: 4-apolycopene and 4-apo-3',4'-didehydro-7,8-dihydrolycopene. At present, we have not determined whether the cells synthesize one (and if so, which one) or both of these C_{35} carotenoids.

Direct product distribution of CrtM variants in the presence of GGPP (C_{20}PP)

To directly evaluate the product specificities of the three CrtM mutants, we constructed pUC-*crtE-M*$_{8\text{-}10}$ plasmids and transformed the plasmids into XL1-Blue cells. Here the CrtM variants are supplied with GGPP, but the carotenoid products cannot be desaturated. Because all three possible products in this scenario, 4,4'-diapophytoene (C_{30}), 4-apophytoene (C_{35}), and phytoene (C_{40}), have an identical chromophore structure consisting of three conjugated double bonds, the molecular extinction coefficients for all three can be assumed to be equivalent, irrespective of the total number of carbon atoms in the carotenoid molecule (49). Thus, the product distribution of the CrtM variants can be discerned from the HPLC chromatogram peak heights (at 286 nm) for each product. As can be seen in **Figure 2.5**, cultures expressing CrtB along with CrtE produced approximately 230 nmol of phytoene/g of dry cell mass but did not detectably synthesize C_{30} or C_{35} carotenoids. CrtE-CrtM cultures produced about 20 to 35 nmol of each of the C_{30}, C_{35}, and C_{40} carotenoids/g. In stark contrast, CrtE-M_8 and CrtE-M_9 cultures, which express synthases mutated at F26, generated over 300 nmol of phytoene/g. These cultures also produced 40 to 75% more C_{35} carotenoids but 70 to 90% fewer C_{30} carotenoids than CrtE-CrtM cultures. Cultures expressing M_{10}, which has no mutation at F26, along with

CrtE synthesized roughly 100 nmol of phytoene/g as well as about 20 and 14 nmol of C_{35} and C_{30} carotenoids/g, respectively.

Comparison of the C_{40} and C_{30} performance of CrtM variants

To compare the acquired CrtB-like C_{40} function of the CrtM mutants with their CrtM-like ability to synthesize C_{30} carotenoids, the three mutants were placed back into the original C_{30} pathway, resulting in pUC-*$M_{8\text{-}10}$-crtN* plasmids. Pigmentation analysis of cells carrying these as well as pUC-*crtE-$M_{8\text{-}10}$-crtI* plasmids (**Figure 2.6**) revealed that the CrtM variants retained C_{30} activity, although the C_{30} pathway performance of cells expressing these mutants varied. Acetone extracts of cultures expressing variant M_8, which has the F26L mutation alone, along with CrtE and CrtI (CrtE-M_8-CrtI cultures) had more than threefold-higher C_{40} carotenoid absorbance (475 nm) than extracts of CrtE-CrtM-CrtI cultures. However, the C_{30} carotenoid absorbance (470 nm) of extracts of cultures expressing variant M_8 and CrtN was only about 40% that of CrtM-CrtN cultures. Cultures expressing variant M_9, which has the F26S mutation, along with CrtE and CrtI also yielded extracts with over three times the C_{40} signal of CrtE-CrtM-CrtI culture extracts. Yet the C_{30} signal of M_9-CrtN culture extracts was only about 10% that of CrtM-CrtN culture extracts. Cultures expressing mutant M_{10}, which has no mutation at F26, along with CrtE and CrtI generated extracts with approximately 70% higher C_{40} absorbance than extracts of CrtE-CrtM-CrtI cultures. Interestingly, cultures expressing M_{10} and CrtN showed no reduction in C_{30} pathway performance compared to CrtM-CrtN cultures. M_{10} was the only one of the 10 sequenced variants that gave this result (data not shown).

Analysis of mutations

The most significant and only recurring mutations found in the 10 sequenced CrtM variants were those at F26. In seven variants, phenylalanine is replaced by leucine, while two have serine at this position. The F26L substitution alone is sufficient for acquisition of C_{40} activity by CrtM (M_8; **Table 2.1**). Thus, we conclude that mutation at residue 26 of CrtM directly alters the enzyme's specificity. Changing enzyme expression level or stability would not lead to increased C_{40} performance and decreased C_{30} performance of cultures expressing CrtM variants compared to those expressing wild-type CrtM (**Figure 2.6**). Variant M_{10} possesses no mutation at amino acid position 26. Thus, it is apparent that mutation at this residue is not the only means by which CrtM can acquire CrtB-like activity. Of all 10 sequenced variants, M_{10} is the only one with no substitution for phenylalanine at position 26 and is also the only one whose cultures did not have decreased C_{30} pathway performance compared to CrtM cultures.

Structural considerations: mapping mutations onto human squalene synthase

Most structurally characterized isoprenoid biosynthetic enzymes, including FPS, squalene synthase (SqS), and terpene cyclases, have the same "isoprenoid synthase fold," consisting predominantly of α-helices (22). In addition, secondary structure prediction (11) and sequence alignment (8) of CrtM and CrtB with their related enzymes also suggest that the enzymes have a common fold. Given this, and the lack of a crystal structure for a carotenoid synthase, we mapped the amino acid substitutions in our CrtM variants onto the crystal structure of human SqS (37).

SqSs catalyze the first committed step in cholesterol biosynthesis. As with carotene synthases, this is the head-to-head condensation of two identical prenyl

diphosphates (FPP for SqS). The condensation reaction catalyzed by SqS proceeds in two distinct steps (38). The first half-reaction generates the stable intermediate presqualene diphosphate, which forms upon abstraction of a diphosphate group from a prenyl donor, followed by 1-1' condensation of the donor and acceptor molecules. In the second half-reaction, the intermediate undergoes a complex rearrangement followed by a second removal of diphosphate and a final carbocation-quenching process (**Figure 2.7**). SqSs catalyze the additional reduction of the central double bond of 4,4'-diapophytoene (also called dehydrosqualene) by NADPH to form squalene, a reaction not performed by carotene synthases. Because the SqS and carotene synthase enzymes share clusters of conserved amino acids and catalyze essentially identical reactions, it is probable that they also have the same reaction mechanism. Indeed, when NADPH is in short supply, SqS produces 4,4'-diapophytoene, the natural product of CrtM (19, 52).

Sequence alignment of CrtM with related enzymes implies that F26 in CrtM corresponds to I58 in human SqS, which is located in helix B and points into the pocket that accommodates the second half-reaction. This residue is located four amino acids downstream of a flexible “flap” region in SqS that is believed to form a “lid” that shields intermediates in the reaction pocket from water (37). The amino acids constituting the flap are almost completely conserved among all known head-to-head isoprenoid synthase enzymes that catalyze 1-1' condensation.

It is noteworthy that a single mutation at F26 of CrtM is sufficient to permit this enzyme to synthesize C_{40} carotenoids. Because this position is thought to lie in the site of the second half-reaction (rearrangement and quenching of a cyclopropylcarbinyl intermediate) (37), and because wild-type CrtM produces trace amounts of phytoene

(indicating that initially accepting two molecules of GGPP is not impossible), it is likely that wild-type CrtM is able to perform the first half-reaction of phytoene synthesis (condensation of two molecules of GGPP to form a presqualene diphosphate-like structure). We hypothesize that the F26 residue prevents the second half-reaction from going to completion by acting as a steric or electrostatic inhibitor of intermediate rearrangement. When this bulky phenylalanine residue is replaced with a smaller or more flexible amino acid such as serine or leucine, the second half-reaction is permitted to proceed and phytoene is produced.

Similar results for a variety of short-chain isoprenyl diphosphate synthases (IDSs) have been reported. In this class of enzymes, the size of the fifth amino acid upstream of the first aspartate-rich motif determines product length (29, 32, 33, 35, 36). Based on a very strong correlation between average product length and surface area of amino acids in this position (36), as well as the available crystal structure for avian FPS (47), it was hypothesized that this residue forms a steric barrier or wall that controls the size of the products (30). This model has been successfully applied to a variety of other enzymes in this family, including medium-chain IDSs (53). It is unknown why so many IDSs differing so greatly in sequence share a single key residue that determines product specificity.

Evolution of phytoene synthase (CrtB) in a C_{30} pathway

We constructed mutant libraries of *crtB* to search for variants with C_{30} activity. Four mutagenic PCR libraries differing in $MnCl_2$ concentration were ligated into the *Xba*I-*Xho*I site of pUC-*crtB*-*crtN*, resulting in four pUC-*[crtB]*-*crtN* plasmid libraries. Upon transformation into XL1-Blue, pUC-*crtB*-*crtN*, containing wild-type *crtB*, gave no

discernible pigmentation, while the pUC-*crtM-crtN* cells produced had intense yellow pigmentation. Among the ~43,000 colonies expressing pUC-*[crtB]-crtN* variants screened, not a single one showed distinguishable yellow (or other) pigmentation. Thus, we found no CrtB mutants with improved C_{30} activity. We also constructed five additional pUC-*[crtB]-crtN* libraries by using the Genemorph PCR mutagenesis kit (Stratagene), which enables the construction of randomly mutagenized gene libraries with a mutational spectrum different from that of those generated by mutagenic PCR with $MnCl_2$ (7). We screened an additional ~10,000 variants, but again found no variants of CrtB able to synthesize C_{30} carotenoids at appreciable levels.

An explanation for the relative difficulty in acquiring C_{30} function for CrtB may be that accepting a smaller-than-natural substrate is a more difficult task for this type of enzyme than accepting a larger-than-natural substrate. However, an evolutionary explanation may also be in order. The contexts in which the two enzymes evolved differ markedly. FPP is a precursor for many life-supporting compounds and is present in all organisms. Thus, C_{40} enzymes such as CrtB have evolved in the presence of FPP throughout their history. It is not unreasonable to infer that CrtB has evolved under a nontrivial selection pressure to minimize consumption of FPP, for accepting FPP as a substrate and bypassing GGPP could be detrimental to the host organism's fitness. In stark contrast, C_{30} synthases have evolved in an environment essentially devoid of GGPP. Consequently, no pressure to reject this substrate has been placed on these enzymes.

CONCLUSIONS

We have demonstrated for the first time that the specificity of a carotenoid synthase is easily and profoundly altered by directed evolution. In our experiments, the carotene synthases CrtM and CrtB failed to complement one another in their respective foreign pathways. However, upon random mutagenesis of *crtM* followed by C_{40}-specific color complementation screening, we isolated 116 mutants of the enzyme (~0.5% of the total screened) with significant C_{40} activity. The *in vivo* C_{40} pathway performance of the best CrtM variants is comparable to that of CrtB, the native C_{40} synthase. That CrtM, a key biosynthetic enzyme that determines the size of the carotenoids in the downstream pathway, is capable of synthesizing C_{35} carotenoid backbones and is one amino acid substitution from becoming a C_{40} phytoene synthase is a finding of evolutionary and technological significance.

Firn and Jones have suggested that natural product biosynthetic pathways have evolved traits that enhance their ability to generate and retain chemical diversity in order to maximize their likelihood of discovering new biologically active molecules (13, 14, 20). The two major traits that Firn and Jones have identified in these pathways are the use of enzymes with broad substrate or product specificity and a high degree of pathway branching or bifurcation. Our work has revealed a special type of combinatorial pathway branching that exploits the precursor-fusion chemistry of carotenoid synthases. The native reaction catalyzed by CrtM is the condensation of two molecules of FPP to form a C_{30} carotenoid backbone. Here we have shown that wild-type CrtM also possesses the ability to condense one molecule of FPP with one molecule of GGPP to form a C_{35} carotenoid backbone. However, the enzyme is very poor at condensing two molecules of

GGPP to form the C_{40} carotenoid backbone, phytoene. By broadening the specificity of CrtM through point mutagenesis, we have created variants that can also synthesize C_{40} carotenoids. That these mutants can synthesize two non-native products as a result of the addition of only one new precursor (GGPP) to their substrate repertoires is a consequence of both their broadened specificity as well as the nature of the coupling reaction catalyzed by carotenoid synthases. If a (mutant) carotenoid synthase does not discriminate between *n* different prenyl diphosphate substrates, it will synthesize $n(n+1)/2$ different carotenoid backbones by virtue of the combinatorics of substrate coupling. Therefore, carotenoid synthases, by virtue of their special chemistry, can act as key "diversity-multiplying" enzymes in carotenoid biosynthetic pathways if suitable mutants able to accept non-native precursors can be found. In the following chapter, we describe the continuation of our efforts to further expand carotenoid biosynthesis by creating additional backbones that can serve as new branch points for carotenoid pathway diversification.

MATERIALS AND METHODS

Materials

Erwinia uredovora (current approved name: *Pantoea ananatis*) C_{40} carotenoid pathway genes *crtE* encoding GGPP (C_{20}PP) synthase, *crtB* encoding phytoene synthase, and *crtI* encoding phytoene desaturase were obtained by genomic PCR as previously described (43). The *Escherichia coli* farnesyl diphosphate (FPP, C_{15}PP) synthase gene *fps* was cloned from *E. coli* strain JM109. The C_{30} pathway genes *crtM* (4,4'-diapophytoene synthase) and *crtN* (4,4'-diapophytoene desaturase) were cloned by PCR from *Staphylococcus aureus* genomic DNA (ATCC no. 35556D). We used *E. coli* XL1-Blue

supercompetent cells (Stratagene, La Jolla, CA) for cloning, screening, and carotenoid biosynthesis. AmpliTaq polymerase (Perkin-Elmer, Boston, MA) was employed for mutagenic PCR, while Vent polymerase (New England Biolabs, Beverly, MA) was used for cloning PCR. All chemicals and reagents used were of the highest available grade.

Plasmid construction

Plasmid maps and sequences can be found in Appendix A. Plasmid pUC18m was constructed by removing the entire *lacZ* fragment and multi-cloning site from pUC18 and inserting the multi-restriction site sequence 5'-CATATG-GAATTC-TCTAGA-CTCGAG-GGGCCC-GGCGCC-3' (*Nde*I-*Eco*RI-*Xba*I-*Xho*I-*Apa*I-*Ehe*I). Each open reading frame following a Shine-Dalgarno ribosomal binding sequence (boldface) and a spacer (**AGGAGG**ATTACAAA) was cloned into pUC18m (see Appendix A) to form artificial operons for acyclic C_{40} carotenoids (pUC-*crtE-crtB-crtI*) or acyclic C_{30} carotenoids (pUC-*fps-crtM-crtN* or pUC-*crtM-crtN*) (the genes in plasmids and operons are always listed in transcriptional order). To facilitate exchange between the two pathways, corresponding genes were flanked by the same restriction sites: isoprenyl diphosphate synthase genes (*fps* and *crtE*) were flanked by *Eco*RI and *Xba*I sites, carotene synthase genes (*crtM* and *crtB*) were flanked by *Xba*I and *Xho*I sites, and carotene desaturase genes (*crtN* and *crtI*) were flanked by *Xho*I and *Apa*I sites (**Figure 2.2a**).

Error-prone PCR mutagenesis and screening

A pair of primers (5'-GCTGCCGTCAGTTAATCTAGAAGGAGG-3') and (5'-AGACGAATTGCCAGTGCCAGGCCACCG-3') flanking *crtM* were designed to amplify the 0.85-kb gene by PCR under mutagenic conditions: 5 U AmpliTaq (100 µl

total volume); 20 ng of template DNA (entire plasmid); 50 pmol of each primer; 0.2 mM dATP; 1.0 mM (each) dTTP, dGTP, and dCTP; and 5.5 mM $MgCl_2$. Four different mutagenic libraries were made by using four different $MnCl_2$ concentrations: 0.2, 0.1, 0.05, and 0.02 mM. The temperature cycling scheme was 95 °C for 4 min followed by 30 cycles of 95 °C for 30 s, 40 °C for 45 s, and 72 °C for 2 min and by a final stage of 72 °C for 10 min. PCR yields for the 0.85-kb amplified fragment were 5 μg, corresponding to an amplification factor of ~1,000 or ~10 effective cycles. The PCR product from each library was purified with a Zymoclean gel purification kit (Zymo Research, Orange, CA), followed by digestion with *Xho*I and *Xba*I and *Dpn*I treatment to digest the template. The PCR products were ligated into the carotene synthase gene site of vector pUC-*crtE-crtM-crtI*, resulting in pUC-*crtE-[crtM]-crtI* libraries (square brackets indicate the randomly mutagenized gene). PCR mutagenesis of *crtB* on plasmid pUC-*crtB-crtN* was performed with primers 5'-CTTTACACTTTATGCTTCCGG-3' and 5'-TCCTGTGACACCTG CACCAATTACTGC-3' under the same conditions used for mutagenesis of *crtM*. The PCR products were purified, digested, and ligated as described above into the carotene synthase gene site of pUC-*crtB-crtN*, resulting in four pUC-*[crtB]-crtN* libraries.

The ligation mixtures were transformed into *E. coli* XL1-Blue supercompetent cells. Colonies were grown on Luria-Bertani (LB) plates containing carbenicillin (50 mg/l) as a selective marker at 37 °C for 12 h. Colonies were lifted onto white nitrocellulose membranes (Pall, Port Washington, NY), transferred onto LB+carbenicillin plates, and visually screened for color variants after an additional 12 to 24 h at room temperature. Selected colonies were picked and cultured overnight in 96-well plates, each well containing 0.5 ml of liquid LB medium supplemented with carbenicillin (50 mg/l).

Pigment analysis

Among the strains we tested as expression hosts, XL1-Blue showed the best results in terms of stability and intensity of the color developed by colonies on agar plates. Although all of the genes assembled in each plasmid are grouped under a single *lac* operator/promoter, our expression system showed no response in terms of pigmentation levels to different IPTG (isopropyl-β-D-thiogalactopyranoside) concentrations. Thus, leaky transcription from the *lac* promoter was sufficient for carotenoid production in *E. coli*. Based on this observation, all experiments described in this report were performed without IPTG induction.

To measure the relative amounts of carotenoids synthesized (see **Figure 2.6**), single colonies were inoculated into 3-ml precultures (LB medium containing 50 mg/l of carbenicillin) and shaken at 250 rpm and 37 °C overnight. Twenty microliters of each preculture were inoculated into 3 ml of Terrific broth (TB) medium (also containing 50 mg/l of carbenicillin) and shaken for 24 (C_{30} carotenoid cultures) or 30 h (C_{40} carotenoid cultures) at 250 rpm and 30 °C. The optical density at 600 nm (OD_{600}) of each culture was measured immediately before harvesting. Then, 2 ml of each TB culture was centrifuged, the liquid was decanted, and the resulting cell pellet was extracted with 1 ml of acetone. The absorbance spectrum of each extract was measured with a SpectraMax Plus 384 microplate spectrophotometer (Molecular Devices, Sunnyvale, CA). Pigmentation levels in the culture extracts were determined from the height of absorption maxima (λ_{max}): 470 nm for C_{30} carotenoids and 475 nm for C_{40} carotenoids.

For accurate determination of the carotenoids produced by the cultures, 500 µl of LB preculture was inoculated into 50 ml of TB+carbenicillin (50 mg/l) and shaken in a

250-ml tissue culture flask (Becton Dickinson-Falcon, Bedford, MA) at 170 rpm and 30 °C for 24 to 30 h. Cultures expressing only CrtE plus a carotenoid synthase (CrtB, CrtM, or a mutant CrtM) were cultivated in 50 ml of TB+carbenicillin (50 mg/l) for 40 h at 160 rpm and 28 °C in 250-ml tissue culture flasks. The OD_{600} of each culture was measured immediately before harvesting, and the dry cell mass was determined from this measurement by using a calibration curve generated for similar cultures. After centrifugation, the cell pellets were extracted with 10 ml of an acetone-methanol mixture (2:1 [vol/vol]). Pigments were concentrated, and the solvent was replaced with 20 ml of hexane. Then, an equal volume of aqueous NaCl (100 g/l) was added, and the mixture was shaken vigorously to remove oily lipids. The upper phase containing the carotenoids was dewatered with anhydrous $MgSO_4$ and concentrated in a rotary evaporator. The final volume of extract from each 50-ml culture was 1 ml. A 30- to 50-µl aliquot of extract was passed through a Spherisorb ODS2 column (4.6 × 250 mm, 5-µm particle size; Waters, Milford, MA) and eluted with an acetonitrile-isopropanol mixture (93:7 or 80:20 [vol/vol]) at a flow rate of 1 ml/min using an Alliance high-pressure liquid chromatography (HPLC) system (Waters) equipped with a photodiode array detector. Mass spectra were obtained with a series 1100 HPLC-mass spectrometer (Hewlett-Packard/Agilent, Palo Alto, CA) coupled with an atmospheric pressure chemical ionization interface.

The molar quantities of the carotenoid backbones shown in **Figure 2.5** were determined by comparing their HPLC chromatogram peak heights (at 286 nm) to a calibration curve generated using known amounts of a β,β-carotene standard (Sigma), and then multiplying by $\varepsilon_{\beta,\beta\text{-carotene (450 nm)}}/\varepsilon_{\text{phytoene (286 nm)}}$. The values of the molar

extinction coefficients (ε) used in the calculation were 138,900 and 49,800, respectively (6). The molar quantities of carotenoids were then normalized to the dry cell mass of each culture.

Table 2.1. Mutations found in sequenced CrtM variants

Mutant designation	Nonsynonymous mutations (amino acid change)	Synonymous mutations
M_1	T76C (F26L) A364T (T122S)	A561G
M_2	A58G (K20E) T77C (F26S)	A471T
M_3	T76C (F26L) G127A (V43M)	A408G
M_4	T76C(F26L) A446G (E149G)	T150C G489A A726G
M_5	T76C (F26L) T800C (F267S)	
M_6	A35G (H12R) T76C (F26L) A80G (D27G) A290G (K97R) A620G (H207R)	A688G
M_7	T76C (F26L)	A186G A447G
M_8	T78A (F26L)	A345G
M_9	T77C (F26S) T119C (I40T)	T135C T141C
M_{10}	A10G (M4V) A35G (H12R) T176C (F59S) A242G (Q81R) A539G (E180G)	A39G

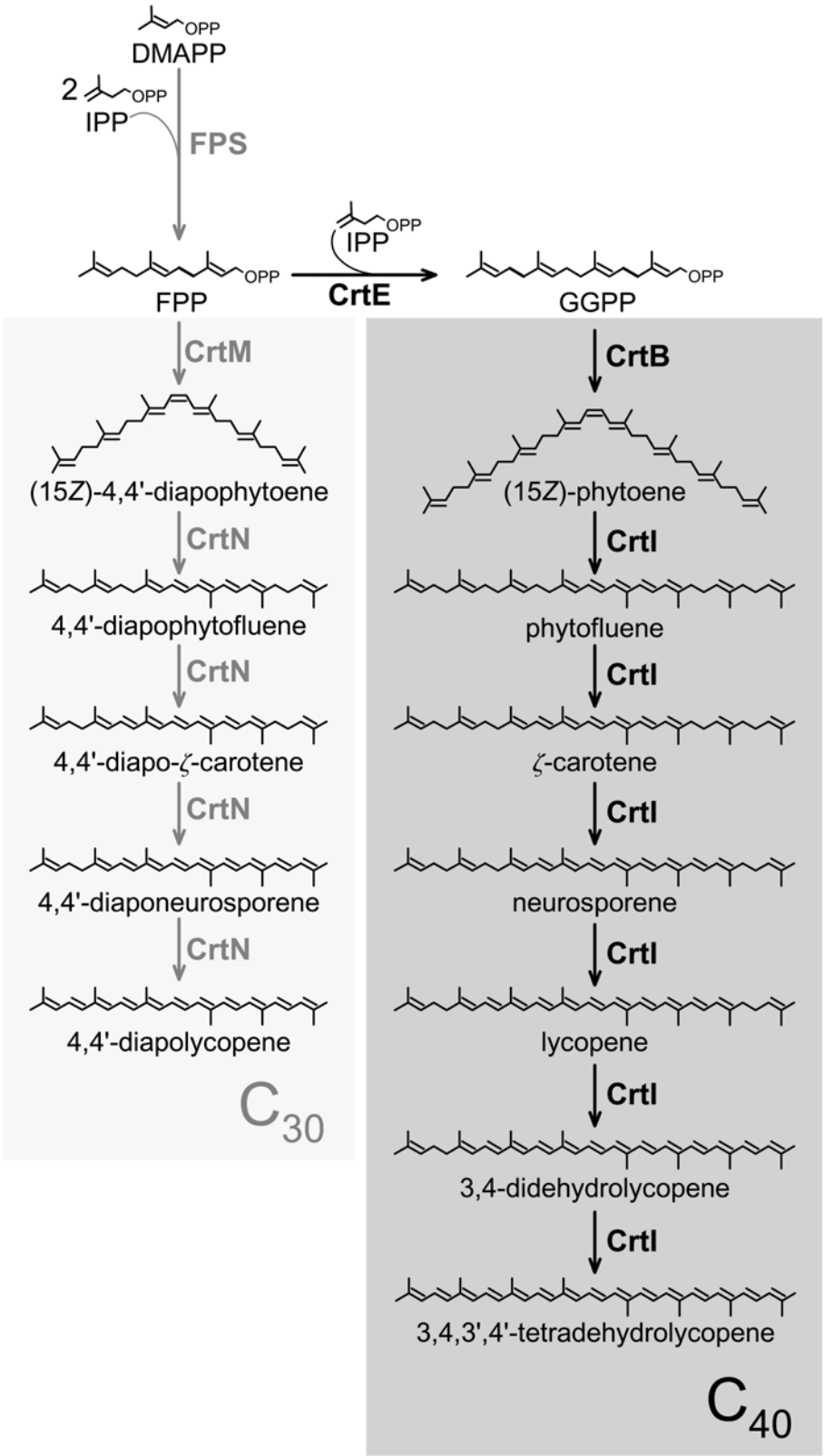

OPP
DMAPP
2
OPP
IPP
FPS
OPP
IPP
FPP
CrtE
GGPP
CrtM
(15Z)-4,4'-diapophytoene
CrtN
4,4'-diapophytofluene
CrtN
4,4'-diapo-ζ-carotene
CrtN
4,4'-diaponeurosporene
CrtN
4,4'-diapolycopene
C30
CrtB
(15Z)-phytoene
CrtI
phytofluene
CrtI
ζ-carotene
CrtI
neurosporene
CrtI
lycopene
CrtI
3,4-didehydrolycopene
CrtI
3,4,3',4'-tetradehydrolycopene
C40

Figure 2.1

Figure 2.1. Natural carotenoid biosynthetic pathways. Carotenoid pathways are branches of the general isoprenoid pathway. In nature, two distinctive routes to carotenoid structures are known. C_{40} pathways, which start from the head-to-head condensation of two molecules of GGPP (C_{20}PP), are found in a variety of plant and microbial species. C_{30} pathways, which begin with the condensation of two molecules of FPP (C_{15}PP), have been identified only in a small number of bacterial species. Shown are the 15*Z* isomers of 4,4'-diapophytoene and phytoene, which are produced by CrtM and CrtB, respectively. Bacterial desaturases CrtN and CrtI catalyze *Z*-*E* isomerization of the central double bond in addition to desaturation of their carotenoid backbone substrates (15). The enzyme FPS is an FPP synthase; CrtE is a bacterial GGPP synthase; CrtM and CrtB are bacterial carotenoid synthases. DMAPP, dimethylallyl diphosphate; IPP, isopentenyl diphosphate.

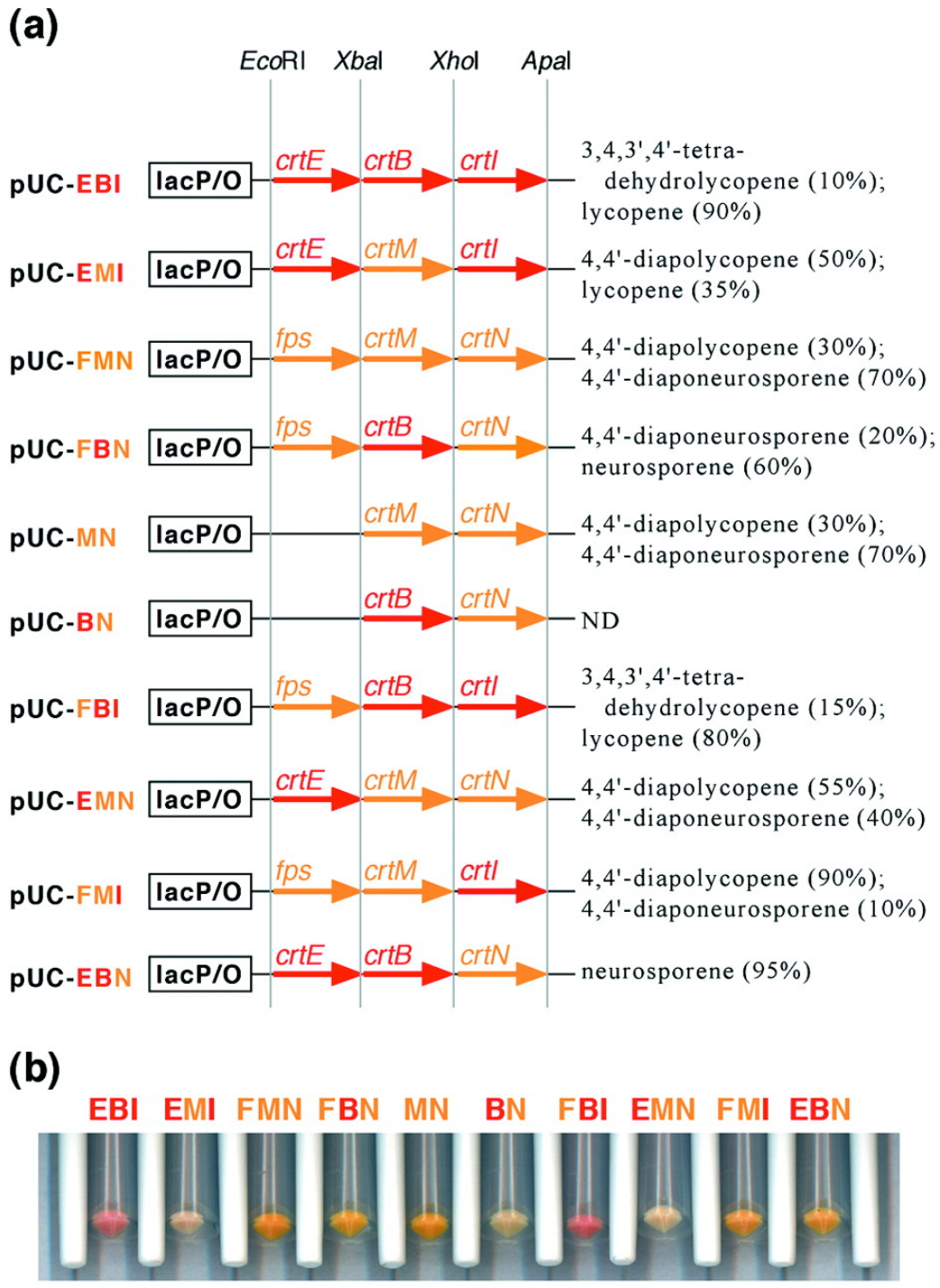

(a)
EcoRI
XbaI
XhoI
ApaI
pUC-EBI
lacP/O
crtE
crtB
crtI
3,4,3',4'-tetra-dehydrolycopene (10%); lycopene (90%)
pUC-EMI
lacP/O
crtE
crtM
crtI
4,4'-diapolycopene (50%); lycopene (35%)
pUC-FMN
lacP/O
fps
crtM
crtN
4,4'-diapolycopene (30%); 4,4'-diaponeurosporene (70%)
pUC-FBN
lacP/O
fps
crtB
crtN
4,4'-diaponeurosporene (20%); neurosporene (60%)
pUC-MN
lacP/O
crtM
crtN
4,4'-diapolycopene (30%); 4,4'-diaponeurosporene (70%)
pUC-BN
lacP/O
crtB
crtN
ND
pUC-FBI
lacP/O
fps
crtB
crtI
3,4,3',4'-tetra-dehydrolycopene (15%); lycopene (80%)
pUC-EMN
lacP/O
crtE
crtM
crtN
4,4'-diapolycopene (55%); 4,4'-diaponeurosporene (40%)
pUC-FMI
lacP/O
fps
crtM
crtI
4,4'-diapolycopene (90%); 4,4'-diaponeurosporene (10%)
pUC-EBN
lacP/O
crtE
crtB
crtN
neurosporene (95%)
(b)
EBI
EMI
FMN
FBN
MN
BN
FBI
EMN
FMI
EBN

Figure 2.2

Figure 2.2. C_{30} and C_{40} production systems used in this work. (a) Plasmids used in this work. Under the control of the *lac* promoter and operator of a pUC18-derived plasmid, carotene synthase genes (*crtB* and *crtM*) were cloned into an *Xba*I-*Xho*I site, along with the isoprenyl diphosphate synthase genes (*crtE* and *fps*; *Eco*RI-*Xba*I site) and the carotene desaturase genes (*crtN* and *crtI*; *Xho*I-*Apa*I site). Listed on the right are the approximate mole percentages (as a fraction of total carotenoids) of the main carotenoids produced by *E. coli* XL1-Blue cells transformed with each plasmid. Only major species consisting of at least 10% of total carotenoids are shown. ND, none detected. (b) Cell pellets of XL1-Blue harboring various carotenogenic plasmids.

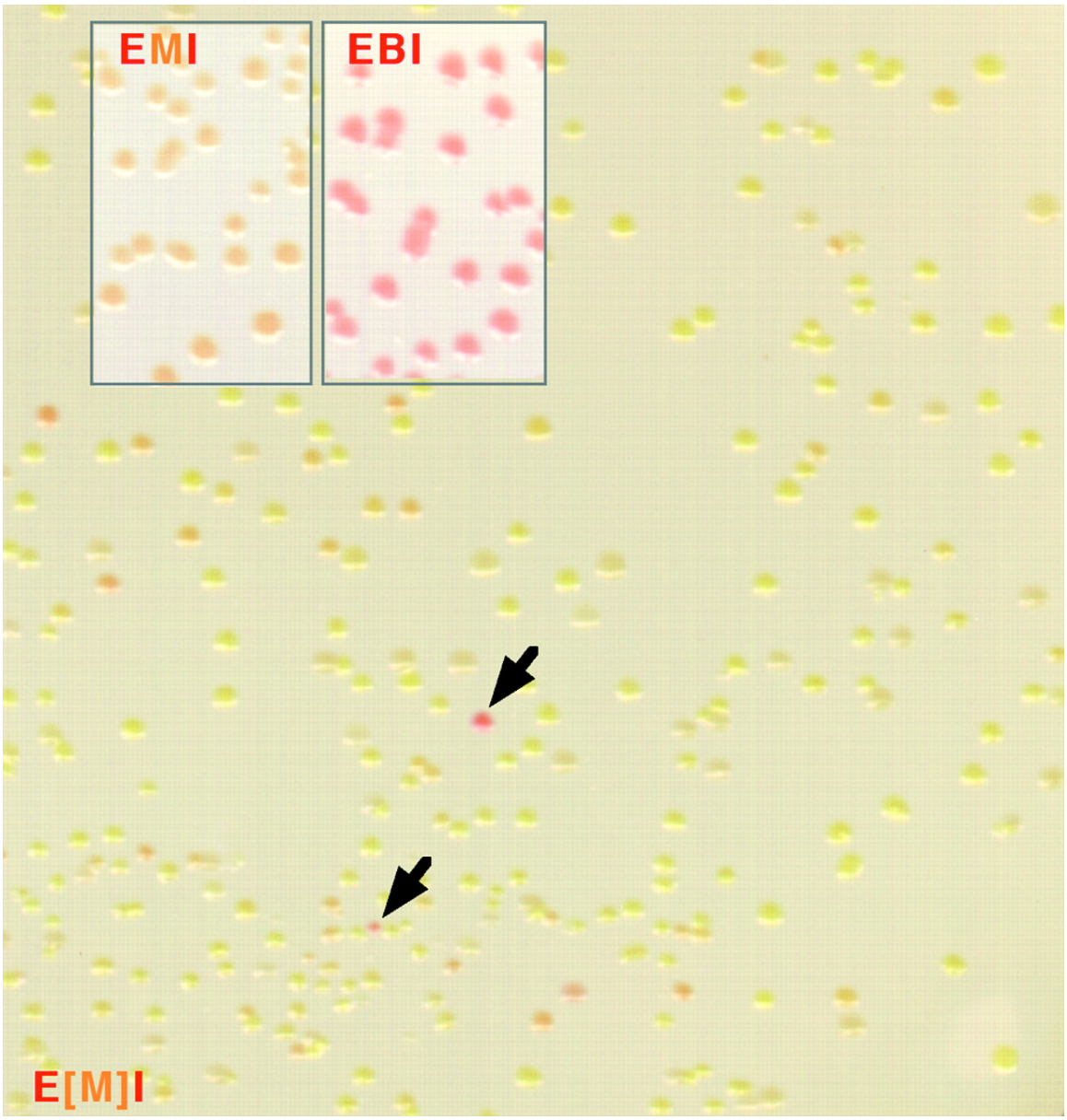


Figure 2.3

Figure 2.3. Typical plate with *E. coli* XL1-Blue colonies expressing a mutagenic library of CrtM together with CrtI and CrtE. XL1-Blue cells were transformed with pUC-*crtE-[crtM]-crtI* (E[M]I), where *[crtM]* represents a mutagenic library of *crtM*. Among a majority of pale colonies can be seen deep red-pink colonies (arrows) expressing CrtM variants that have acquired C_{40} pathway functionality. EMI and EBI, XL1-Blue cells transformed with pUC-*crtE-crtM-crtI* and pUC-*crtE-crtB-crtI*, respectively.

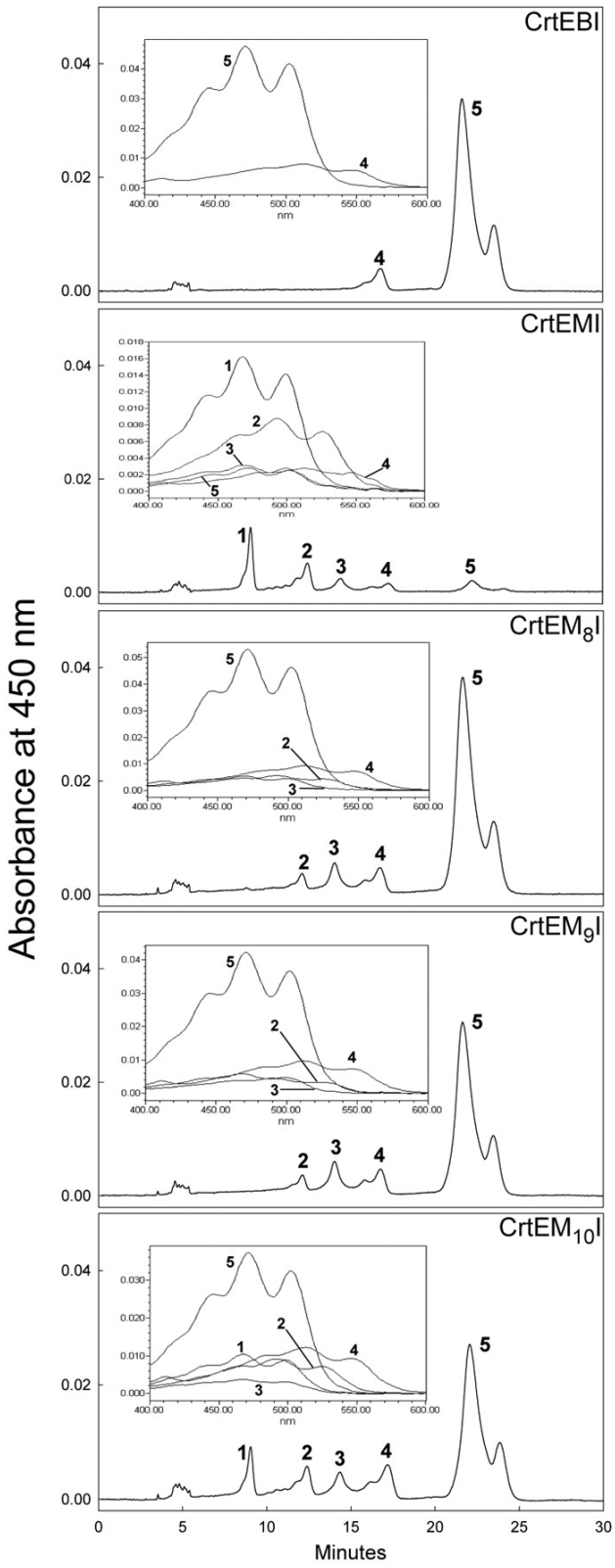

CrtEBI
CrtEMI
CrtEM8I
CrtEM9I
CrtEM10I
Absorbance at 450 nm
Minutes
nm

Figure 2.4

Figure 2.4. HPLC-photodiode array analysis of carotenoid extracts of *E. coli* transformants carrying plasmids pUC-*crtE-crtB-crtI*, pUC-*crtE-crtM-crtI*, and pUC-*crtE-M*$_{8\text{-}10}$*-crtI*. The following carotenoids were identified: peak 1, 4,4'-diapolycopene (λ_{max} [nm]: 501, 470, 444, M^+ at *m/z* = 400.4); peak 2, 4-apo-3',4'-didehydrolycopene (λ_{max} [nm]: 527, 490, 465, M^+ at *m/z* = 466.4); peak 3, 4-apolycopene or 4-apo-3',4'-didehydro-7,8-dihydrolycopene (λ_{max} [nm]: 500, 470, 441, M^+ at *m/z* = 468.4); peak 4, 3,4,3',4'-tetradehydrolycopene (λ_{max} [nm]: 540, 510, 480, M^+ at *m/z* = 532.4); peak 5, lycopene (λ_{max} [nm]: 502, 470, 445, M^+ at *m/z* = 536.5). Double peaks indicate different geometrical isomers of the same compound. Insets, recorded absorption spectra of individual HPLC peaks.

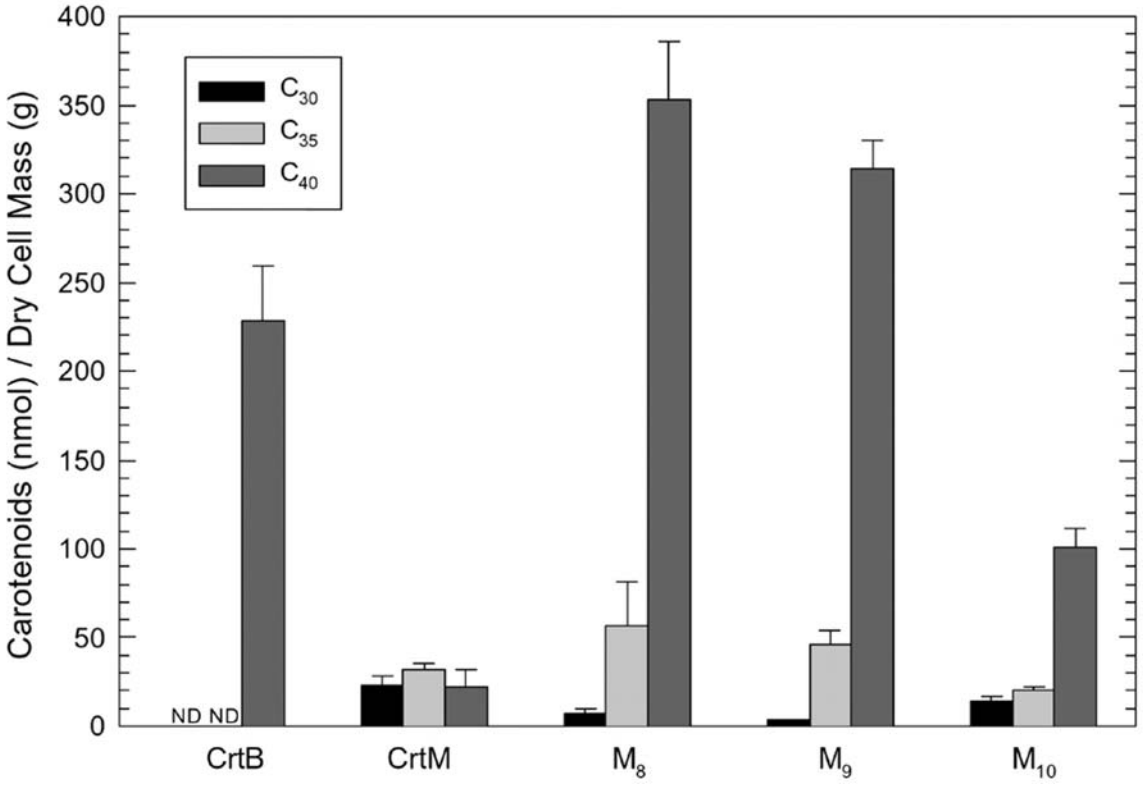


Figure 2.5. Direct product distribution of CrtM and its mutants in the presence of CrtE (GGPP supply). Carotenoid extracts of XL1-Blue cells carrying plasmids pUC-*crtE-crtB*, pUC-*crtE-crtM*, and pUC-*crtE-M*$_{8\text{-}10}$ were analyzed by HPLC with a photodiode array detector. Peaks for 4,4'-diapophytoene (C_{30}), 4-apophytoene (C_{35}), and phytoene (C_{40}) were monitored at 286 nm. Molar quantities of the various carotenoids were determined as described in Materials and Methods. Bar heights are normalized to dry cell mass, each represents the average of three independent cultures; error bars, standard deviations. ND, not detected.

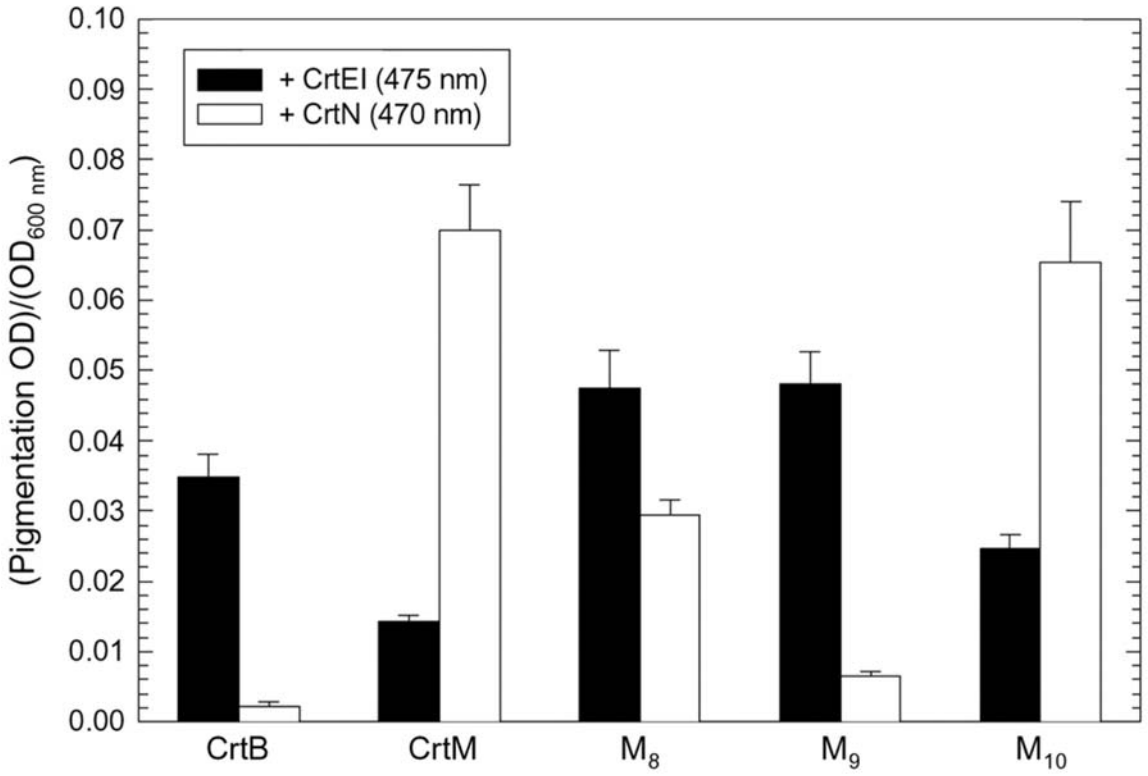


Figure 2.6. Pigmentation produced by CrtM variants in C_{40} and C_{30} pathways. XL1-Blue cells were transformed with either pUC-*crtE*-$M_{8\text{-}10}$-*crtI* (C_{40} pathway) or pUC-$M_{8\text{-}10}$-*crtN* (C_{30} pathway) and cultured in a test tube (3 ml of TB) as described in Materials and Methods. Pigmentation levels in the culture extracts were determined from the absorption peak height of λ_{max} (470 nm for C_{30} carotenoids, 475 nm for C_{40} carotenoids) of each sample. Bar heights are normalized to OD_{600} and represent the averages of at least three replicates; error bars, standard deviations.

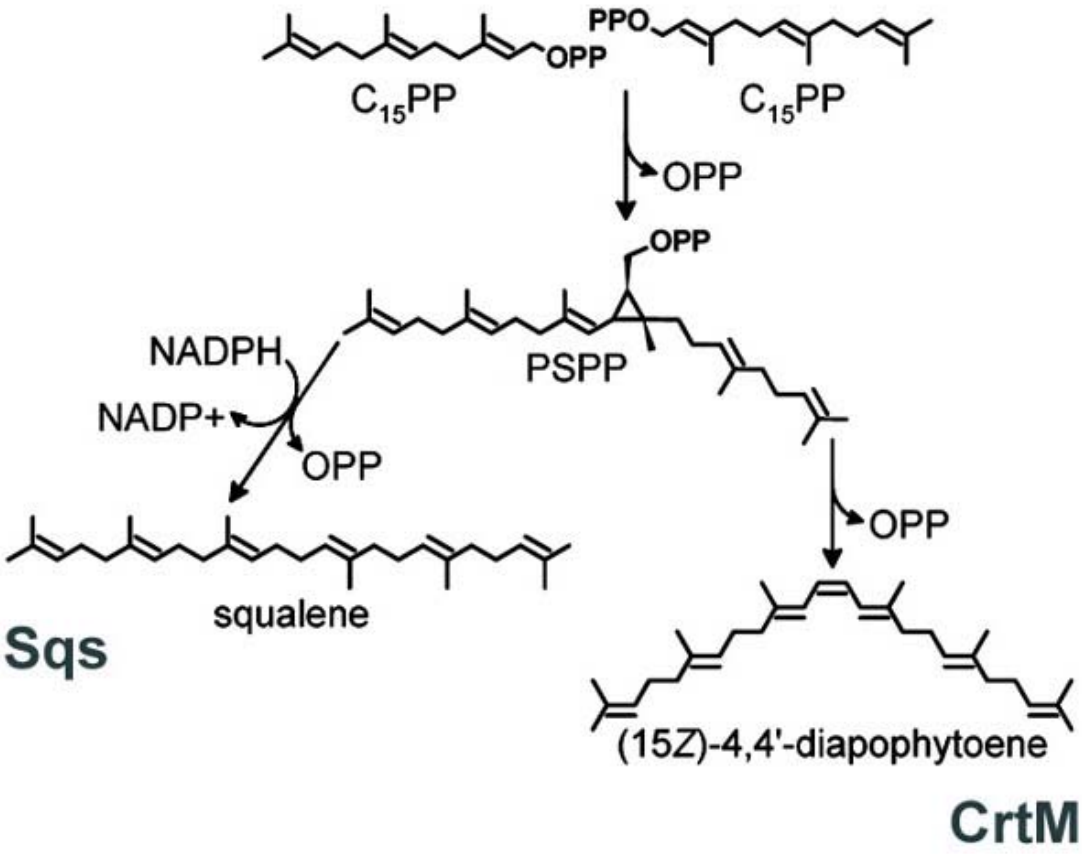


Figure 2.7. Reaction schemes for SqS and CrtM. PSPP, presqualene diphosphate.

REFERENCES

1. **Albrecht, M., S. Takaichi, N. Misawa, G. Schnurr, P. Boger, and G. Sandmann.** 1997. Synthesis of atypical cyclic and acyclic hydroxy carotenoids in *Escherichia coli* transformants. J. Biotechnol. **58:**177-185.
2. **Albrecht, M., S. Takaichi, S. Steiger, Z. Y. Wang, and G. Sandmann.** 2000. Novel hydroxycarotenoids with improved antioxidative properties produced by gene combination in *Escherichia coli*. Nat. Biotechnol. **18:**843-846.
3. **Armstrong, G. A., and J. E. Hearst.** 1996. Carotenoids 2: Genetics and molecular biology of carotenoid pigment biosynthesis. Faseb J **10:**228-37.
4. **Ben-Dor, A., A. Nahum, M. Danilenko, Y. Giat, W. Stahl, H. D. Martin, T. Emmerich, N. Noy, J. Levy, and Y. Sharoni.** 2001. Effects of acyclo-retinoic acid and lycopene on activation of the retinoic acid receptor and proliferation of mammary cancer cells. Arch. Biochem. Biophys. **391:**295-302.
5. **Britton, G.** 1998. Overview of carotenoid biosynthesis, p. 13-140. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids vol. 3: Biosynthesis and Metabolism. Birkhäuser Verlag, Basel.
6. **Britton, G.** 1995. UV/Visible Spectroscopy, p. 13-62. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids, vol. 1B: Spectroscopy. Birkhäuser Verlag, Basel.
7. **Cline, J., and H. Hogrefeo.** 1999. Randomize gene sequences with new PCR mutagenesis kit. Stratagies **13:**157-162.
8. **Corpet, F.** 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. **16:**10881-90.
9. **Crock, J., M. Wildung, and R. Croteau.** 1997. Isolation and bacterial expression of a sesquiterpene synthase cDNA clone from peppermint (*Mentha x piperita*, L.) that produces the aphid alarm pheromone *(E)*-beta-farnesene. Proc. Natl. Acad. Sci. USA **94:**12833-12838.
10. **Croteau, R., F. Karp, K. C. Wagschal, D. M. Satterwhite, D. C. Hyatt, and C. B. Skotland.** 1991. Biochemical characterization of a spearmint mutant that resembles peppermint in monoterpene content. Plant Physiol. **96:**744-752.
11. **Cuff, J. A., M. E. Clamp, A. S. Siddiqui, M. Finlay, and G. J. Barton.** 1998. JPred: a consensus secondary structure prediction server. Bioinformatics **14:**892-3.
12. **Facchini, P. J., and J. Chappell.** 1992. Gene family for an elicitor-induced sesquiterpene cyclase in tobacco. Proc. Natl. Acad. Sci. USA **89:**11088-11092.
13. **Firn, R. D., and C. G. Jones.** 2000. The evolution of secondary metabolism—a unifying model. Mol. Microbiol. **37:**989-994.
14. **Firn, R. D., and C. G. Jones.** 2003. Natural products—a simple model to explain chemical diversity. Nat. Prod. Rep. **20:**382-391.
15. **Fraser, P. D., N. Misawa, H. Linden, S. Yamano, K. Kobayashi, and G. Sandmann.** 1992. Expression in *Escherichia coli*, purification, and reactivation of the recombinant *Erwinia uredovora* phytoene desaturase. J. Biol. Chem. **267:**19891-19895.
16. **Garcia-Asua, G., H. P. Lang, C. N. Hunter, and R. J. Cogdell.** 1998. Carotenoid diversity: a modular role for the phytoene desaturase step. Trends Plant Sci. **3:**445-449.

17. **Hornero-Méndez, D., and G. Britton.** 2002. Involvement of NADPH in the cyclization reaction of carotenoid biosynthesis. FEBS Lett. **515:**133-136.
18. **Ishimi, Y., M. Ohmura, X. X. Wang, M. Yamaguchi, and S. Ikegami.** 1999. Inhibition by carotenoids and retinoic acid of osteoclast-like cell formation induced by bone-resorbing agents *in vitro*. J. Clin. Biochem. Nutr. **27:**113-122.
19. **Jarstfer, M. B., B. S. J. Blagg, D. H. Rogers, and C. D. Poulter.** 1996. Biosynthesis of squalene. Evidence for a tertiary cyclopropylcarbinyl cationic intermediate in the rearrangement of presqualene diphosphate to squalene. J. Am. Chem. Soc. **118:**13089-13090.
20. **Jones, C. G., and R. D. Firn.** 1991. On the evolution of plant secondary chemical diversity. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. **333:**273-280.
21. **Komori, M., R. Ghosh, S. Takaichi, Y. Hu, T. Mizoguchi, Y. Koyama, and M. Kuki.** 1998. A null lesion in the rhodopin 3,4-desaturase of *Rhodospirillum rubrum* unmasks a cryptic branch of the carotenoid biosynthetic pathway. Biochemistry **37:**8987-8994.
22. **Lesburg, C. A., J. M. Caruthers, C. M. Paschall, and D. W. Christianson.** 1998. Managing and manipulating carbocations in biology: terpenoid cyclase structure and mechanism. Curr. Opin. Struct. Biol. **8:**695-703.
23. **Linden, H., N. Misawa, D. Chamovitz, I. Pecker, J. Hirschberg, and G. Sandmann.** 1991. Functional complementation in *Escherichia coli* of different phytoene desaturase genes and analysis of accumulated carotenes. Z. Naturforsch. (C) **46:**1045-1051.
24. **LinGoerke, J. L., D. J. Robbins, and J. D. Burczak.** 1997. PCR-based random mutagenesis using manganese and reduced dNTP concentration. Biotechniques **23:**409-412.
25. **Matsuoka, S., H. Sagami, A. Kurisaki, and K. Ogura** 1991. Variable product specificity of microsomal dehydrodolichyl diphosphate synthase from rat liver. J. Biol. Chem. **266:**3464-3468.
26. **Mayne, S. T.** 1996. Beta-carotene, carotenoids, and disease prevention in humans. Faseb J. **10:**690-701.
27. **Nagaki, M., S. Sato, Y. Maki, T. Nishino, and T. Koyama.** 2000. Artificial substrates for undecaprenyl diphosphate synthase from *Micrococcus luteus* B-P 26. J. Mol. Catal. B-Enzym. **9:**33-38.
28. **Nagaki, M., A. Takaya, Y. Maki, J. Ishibashi, Y. Kato, T. Nishino, and T. Koyama.** 2000. One-pot syntheses of the sex pheromone homologs of a codling moth, *Laspeyresia promonella* L. J. Mol. Catal. B-Enzym. **10:**517-522.
29. **Narita, K., S. Ohnuma, and T. Nishino.** 1999. Protein design of geranyl diphosphate synthase. Structural features that define the product specificities of prenyltransferases. J Biochem (Tokyo) **126:**566-71.
30. **Ogura, K., and T. Koyama.** 1998. Enzymatic aspects of isoprenoid chain elongation. Chem. Rev. **98:**1263-1276.
31. **Ohnuma, S., H. Hemmi, T. Koyama, K. Ogura, and T. Nishino.** 1998. Recognition of allylic substrates in *Sulfolobus acidocaldarius* geranylgeranyl diphosphate synthase: Analysis using mutated enzymes and artificial allylic substrates. J. Biochem. (Tokyo) **123:**1036-1040.

32. **Ohnuma, S., K. Hirooka, H. Hemmi, C. Ishida, C. Ohto, and T. Nishino.** 1996. Conversion of product specificity of archaebacterial geranylgeranyl-diphosphate synthase—identification of essential amino acid residues for chain length determination of prenyltransferase reaction. J. Biol. Chem. **271:**18831-18837.
33. **Ohnuma, S., K. Hirooka, C. Ohto, and T. Nishino.** 1997. Conversion from archaeal geranylgeranyl diphosphate synthase to farnesyl diphosphate synthase—Two amino acids before the first aspartate-rich motif solely determine eukaryotic farnesyl diphosphate synthase activity. J. Biol. Chem. **272:**5192-5198.
34. **Ohnuma, S., T. Koyama, and K. Ogura** 1993. Alternation of the product specificities of prenyltransferases by metal ions. Biochem. Biophys. Res. Commun. **192:**407-412.
35. **Ohnuma, S., T. Nakazawa, H. Hemmi, A. M. Hallberg, T. Koyama, K. Ogura, and T. Nishino.** 1996. Conversion from farnesyl diphosphate synthase to geranylgeranyl diphosphate synthase by random chemical mutagenesis. J. Biol. Chem. **271:**10087-10095.
36. **Ohnuma, S., K. Narita, T. Nakazawa, C. Ishida, Y. Takeuchi, C. Ohto, and T. Nishino.** 1996. A role of the amino acid residue located on the fifth position before the first aspartate-rich motif of farnesyl diphosphate synthase on determination of the final product. J. Biol. Chem. **271:**30748-30754.
37. **Pandit, J., D. E. Danley, G. K. Schulte, S. Mazzalupo, T. A. Pauly, C. M. Hayward, E. S. Hamanaka, J. F. Thompson, and H. J. Harwood.** 2000. Crystal structure of human squalene synthase—A key enzyme in cholesterol biosynthesis. J. Biol. Chem. **275:**30610-30617.
38. **Poulter, C. D., T. L. Capson, M. D. Thompson, and R. S. Bard.** 1989. Squalene synthetase—inhibition by ammonium analogs of carbocationic intermediates in the conversion of presqualene diphosphate to squalene. J. Am. Chem. Soc. **111:**3734-3739.
39. **Raisig, A., and G. Sandmann.** 2001. Functional properties of diapophytoene and related desaturases of C_{30} and C_{40} carotenoid biosynthetic pathways. Biochim. Biophys. Acta Mol. Cell Biol. Lipids **1533:**164-170.
40. **Redmond, T. M., S. Gentleman, T. Duncan, S. Yu, B. Wiggert, E. Gantt, and F. X. Cunningham.** 2001. Identification, expression, and substrate specificity of a mammalian beta-carotene 15,15'-dioxygenase. J. Biol. Chem. **276:**6560-6565.
41. **Reed, B. C., and H. C. Rilling.** 1976. Substrate binding of avian liver prenyltransferase. Biochemistry **15:**3739-45.
42. **Sacchettini, J. C., and C. D. Poulter.** 1997. Creating isoprenoid diversity. Science **277:**1788-1789.
43. **Schmidt-Dannert, C., D. Umeno, and F. H. Arnold.** 2000. Molecular breeding of carotenoid biosynthetic pathways. Nat Biotechnol **18:**750-3.
44. **Schnurr, G., N. Misawa, and G. Sandmann.** 1996. Expression, purification and properties of lycopene cyclase from *Erwinia uredovora*. Biochem. J. **315:**869-74.
45. **Steele, C. L., J. Crock, J. Bohlmann, and R. Croteau.** 1998. Sesquiterpene synthases from grand fir (*Abies grandis*)—comparison of constitutive and wound-induced activities, and cDNA isolation, characterization and bacterial expression of delta-selinene synthase and gamma-humulene synthase. J. Biol. Chem. **273:**2078-2089.

46. **Takaichi, S., K. Inoue, M. Akaike, M. Kobayashi, H. Ohoka, and M. T. Madigan.** 1997. The major carotenoid in all known species of heliobacteria is the C_{30} carotenoid 4,4'-diaponeurosporene, not neurosporene. Arch. Microbiol. **168:**277-281.
47. **Tarshis, L. C., M. J. Yan, C. D. Poulter, and J. C. Sacchettini.** 1994. Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-angstrom resolution. Biochemistry **33:**10871-10877.
48. **Taylor, R. F.** 1984. Bacterial triterpenoids. Microbiol. Rev. **48:**181-198.
49. **Taylor, R. F., and B. H. Davies.** 1974. Triterpenoid carotenoids of *Streptococcus faecium* UNH 564P. Biochem. J. **139:**751-760.
50. **Wang, C. W., and J. C. Liao.** 2001. Alteration of product specificity of *Rhodobacter sphaeroides* phytoene desaturase by directed evolution. J. Biol. Chem. **276:**41161-41164.
51. **Wieland, B., C. Feil, E. Gloriamaercker, G. Thumm, M. Lechner, J. M. Bravo, K. Poralla, and F. Gotz.** 1994. Genetic and biochemical analyses of the biosynthesis of the yellow carotenoid 4,4'-diaponeurosporene of *Staphylococcus aureus*. J. Bacteriol. **176:**7719-7726.
52. **Zhang, D. L., and C. D. Poulter.** 1995. Biosynthesis of non-head-to-tail isoprenoids—Synthesis of 1'-1-structures and 1'-3-structures by recombinant yeast squalene synthase. J. Am. Chem. Soc. **117:**1641-1642.
53. **Zhang, Y. W., X. Y. Li, and T. Koyama.** 2000. Chain length determination of prenyltransferases: both heteromeric subunits of medium-chain *(E)*-prenyl diphosphate synthase are involved in the product chain length determination. Biochemistry **39:**12717-12722.
54. **Zhao, H., J. C. Moore, A. Volkov, and F. H. Arnold.** 1999. Methods for optimizing industrial enzymes by directed evolution, p. 597-604. *In* A. L. Demain and J. E. Davies (ed.). Methods of Industrial Microbiology and Biotechnology. 2nd ed, ASM Press, Washington DC.

CHAPTER 3

Taking Natural Products to New Lengths:

Biosynthesis of Novel Carotenoid Families

Based on Unnatural Carbon Scaffolds

SUMMARY

In an extension of previous work in which biosynthetic routes to novel C_{45} and C_{50} carotenoid backbones were established in recombinant *E. coli*, we demonstrate the capacity of the C_{40} carotenoid desaturase CrtI from *Erwinia uredovora* to accept and desaturate these unnatural substrates. Desaturation step number in the C_{45} and C_{50} pathways was not very specific, resulting in the production of at least 10 more C_{45} and C_{50} carotenoids. These compounds have never before been identified in nature or chemically synthesized. We also present evidence of the biosynthesis of a novel asymmetric C_{40} backbone formed by condensation of farnesyl diphosphate (C_{15}PP) with farnesylgeranyl diphosphate (C_{25}PP), and the subsequent desaturation of this backbone by CrtI in an unusual manner. Under some conditions, we found that C_{40}, C_{45}, and C_{50} carotenoid backbones synthesized in *E. coli* were oxidized by unknown chemical or enzymatic means, giving monohydroxylated backbone derivatives. Some of these hydroxylated species served as substrates for CrtI *in vitro*, leading to the production of still more novel carotenoids. The ability to supply CrtI with unnaturally large substrates *in vivo* has allowed us to show that this enzyme regulates its desaturation step number by sensing the end groups of its substrate, unlike certain fungal carotenoid desaturases whose step number is apparently determined by their particular multimeric state. Analysis of the different molecular mechanisms by which chemical diversity is generated and then propagated through our nascent pathways provides insight into how this occurs in nature and the selective pressures that have shaped the evolution of natural product biosynthetic enzymes and pathways.

INTRODUCTION

Nature's wealth of small molecules has proven extremely useful to humankind as a source of medicines, fragrances, pigments, toxins, and other functional compounds. Many of these natural products have captivated scientists for decades with their highly complex structures and the elaborate biosynthetic routes by which they are generated. As products of evolution, the ~170,000 natural products characterized thus far (10) are a collective demonstration of the power of this simple algorithm to generate immense chemical diversity.

Inspired by the importance of natural products to health, culture, chemistry, and biology, we have been conducting our own evolution experiments with natural product biosynthetic pathways in the laboratory. By mimicking and accelerating some of the very same processes that drive natural evolution—mutation, gene transfer, and selection—we and others have shown that it is possible to evolve carotenoid biosynthetic pathways, a model system of choice, in new, unnatural directions in standard laboratory bacteria (3, 4, 31, 35, 42, 43, 45, 46, 48, 53, 54, 56). (See Chapter 1 for a detailed description of how biosynthetic pathways can be evolved in the laboratory.)

Carotenoids are ancient natural pigments that play vital roles in key biological processes such as photosynthesis, quenching of free radicals, and vision (57). The ~700 carotenoids identified in nature branch from only two major pathways. Over 95% of natural carotenoids are biosynthesized from the symmetric C_{40} backbone phytoene, which is formed by condensation of two molecules of geranylgeranyl diphosphate (GGPP, C_{20}PP) (**Figure 3.1**). The C_{40} pathway, in addition to being the most diverse carotenoid pathway, is also the most widespread in nature, appearing in thousands of species of

bacteria, archaea, algae, fungi, and plants. A separate C_{30} pathway that begins with the fusion of two molecules of farnesyl diphosphate (FPP, C_{15}PP) (**Figure 3.1**) accounts for the remainder of natural carotenoid diversity. C_{30} carotenoids are known in only a small group of bacteria such as *Staphylococcus*, *Streptococcus*, *Methylobacterium*, and *Heliobacterium* species (52). (See Chapter 1 for a more thorough description of the key enzymatic steps in carotenoid biosynthesis.)

Previously, our group reported the expansion of carotenoid biosynthesis with the generation of a novel C_{35} carotenoid pathway in recombinant *E. coli* (53, 56). This unusual pathway, which begins with the heterocondensation of C_{15}PP and C_{20}PP to form an unnatural, asymmetric C_{35} carotenoid backbone, was further diversified by coexpression of downstream carotenoid desaturases and cyclases, some of which were evolved in our laboratory for a particular function in the new pathway (see Chapter 1). This work led to the biosynthesis of at least ten new carotenoids never before identified in nature or chemically synthesized (53).

Our laboratory also reported the biosynthesis of novel C_{45} and C_{50} carotenoid backbones in recombinant *E. coli* expressing the Y81A mutant of the farnesyl diphosphate (C_{15}PP) synthase from *Bacillus stearothermophilus*, BstFPS$_{Y81A}$ (37), and the F26A+W38A double mutant of the C_{30} carotenoid synthase CrtM from *Staphylococcus aureus*, CrtM$_{F26A,W38A}$ (54). The first enzyme, BstFPS$_{Y81A}$, synthesizes a mixture of C_{15}PP, C_{20}PP, and the very rare (see Discussion) farnesylgeranyl diphosphate (FGPP, C_{25}PP). The next enzyme, CrtM$_{F26A,W38A}$, has an expanded substrate and product range compared to wild-type CrtM, which can only synthesize C_{30} and C_{35} carotenoid backbones (53, 56). When BstFPS$_{Y81A}$ and CrtM$_{F26A,W38A}$ were coexpressed in *E. coli*, a

mixture of C_{35}, C_{40}, C_{45}, and C_{50} carotenoid backbones was produced (**Figure 3.1**, ref. (54)).

The above work did not investigate whether these unnaturally large C_{45} and C_{50} carotenoid backbones could be metabolized by carotenoid desaturases, leading to the introduction of an extended conjugated system (chromophore) and the resulting emergence of pigmentation in the C_{45} and C_{50} pathways. In addition, the lack of an asymmetric C_{40} carotenoid backbone detected in extracts of *E. coli* cultures coexpressing $BstFPS_{Y81A}$ and $CrtM_{F26A,W38A}$ warranted further examination. From homo- or heterocondensation of the three precursors $C_{15}PP$, $C_{20}PP$, and $C_{25}PP$, six different carotenoid backbones are theoretically possible: C_{30}, C_{35}, symmetric C_{40} (phytoene, C_{20}+C_{20}), asymmetric C_{40} (C_{15}+C_{25}), C_{45}, and C_{50} (see **Figure 3.1**). However, only one population of C_{40} backbones was isolated, and it was assumed to be all phytoene (54). Since $C_{15}PP$, $C_{20}PP$, and $C_{25}PP$ are all present in the cells, as indicated by the biosynthesis of C_{35}, C_{45}, and C_{50} carotenoid backbones ($CrtM_{F26A,W38A}$ has lost most of its C_{30} synthesis capability (54)), it was unclear if asymmetric C_{40} backbones were not discovered because they were not synthesized or because they could not be separated and distinguished from phytoene.

In this work, we demonstrate the remarkable substrate promiscuity of the carotenoid desaturase CrtI from *Erwinia uredovora* by showing that it can desaturate C_{45} and C_{50} backbones, leading to an array of pigmented C_{45} and C_{50} carotenoids. We reveal that the production of specific desaturation products and the absence of others provides insight into the mechanism by which CrtI regulates its desaturation step number. We also present evidence that the asymmetric C_{40} carotenoid backbone is indeed synthesized by *E.*

coli expressing BstFPS$_{Y81A}$ and variants of CrtM, and that this unnatural backbone is processed by CrtI in an unusual manner. Finally, we show that large carotenoid backbones are further diversified in an unexpected manner by *in vivo* hydroxylation under certain conditions, and that some of these hydroxylated backbones can serve as alternative substrates for CrtI.

In addition to the potential technological uses of the new carotenoids we report herein, the biosynthetic routes by which they are synthesized provide a convenient laboratory model to study emergent metabolic pathways in nature. We show how analysis of the modes by which chemical diversity is propagated in the new carotenoid pathways can enhance our understanding of the mechanisms used by evolution to continually discover new small molecules.

RESULTS

CrtI desaturates unnatural C_{45}, C_{50}, and asymmetric C_{40} carotenoid backbones, resulting in the biosynthesis of numerous novel carotenoids

E. coli XL1-Blue cells coexpressing BstFPS$_{Y81A}$, a CrtM variant (either the F26L single mutant CrtM$_{F26L}$ or the F26A+W38A double mutant CrtM$_{F26A,W38A}$), and wild-type CrtI from *E. uredovora* synthesize at least ten novel desaturated carotenoids with C_{45} or C_{50} backbones. Structures **3**, **4**, **5**, **6**, **10**, **11a**, **11b**, **12a**, **12b/c**, and **13**, reported here for the first time, were identified by their high-performance liquid chromatography (HPLC) retention times, UV-visible spectra, and mass spectra (**Table 3.1**; **Figures 3.2**, **3.3**, and **3.6**). We also isolated an unusual 2-step desaturated C_{40} carotenoid that is likely based on

an asymmetric C_{40} backbone (structure **1**, see below). **Table 3.2** lists trivial and IUPAC-IUB semi-systematic names for the structures depicted in **Figures 3.2-3.4**.

E. coli cultures harboring the plasmid pUCmodII-*crtM*$_{F26L}$-*crtI*-*bstFPS*$_{Y81A}$, pUCmodII-*crtM*$_{F26A,W38A}$-*crtI*-*bstFPS*$_{Y81A}$, or the plasmids pUC18m-*bstFPS*$_{Y81A}$ and pAC-*crtM*$_{F26A,W38A}$-*crtI* together synthesized mixtures of all of these novel desaturated carotenoids in different proportions and titers, depending on the expression plasmid(s). With all of these expression systems, we observed almost 100% conversion of the C_{45} carotenoid backbone 16-isopentenylphytoene to desaturated C_{45} carotenoids, while only about 25% of the C_{50} backbone 16,16'-diisopentenylphytoene was converted to desaturated products (**Table 3.1**).

Whereas *E. uredovora* CrtI primarily catalyzes four desaturation steps on phytoene (24, 33, 56), acts predominantly as a 4-step desaturase in a C_{30} pathway (56), and performs 4-5 desaturation steps on C_{35} carotenoids (53), the step number of CrtI is less well defined on C_{45} and C_{50} substrates. In the C_{45} pathway, 2-, 3-, 5-, and 6-step products were isolated from *E. coli* cultures harboring the plasmids listed above; in the same cultures, 2-, 3-, 4-, and 6-step C_{50} products were found. In both pathways, there was no clear majority product (**Table 3.1**). We refer to this imprecise desaturase behavior as "stuttering." **Figures 3.2** and **3.3** depict the desaturation isomers of these C_{45} and C_{50} products whose biosynthesis is supported by the HPLC and MS data (see below).

In general, the UV-visible absorption spectra of the desaturated C_{45} and C_{50} carotenoids are hypsochromically shifted compared to those of C_{40} carotenoids with the same number of conjugated double bonds (**Figure 3.6**: **3** and **10** vs. ζ-carotene, **4** and **11a** vs. neurosporene, **12a** vs. lycopene; see **Figure 2.1** for C_{40} structures). The spectra shown

in **Figure 3.6** are the most bathochromic of those measured before and after iodine-catalyzed photoisomerization, and most of these spectra exhibit *cis*-peaks (9) that are less or only slightly more intense than the corresponding C_{40} standard (e.g., **11a** vs. neurosporene, **10** vs. ζ-carotene). We therefore hypothesize that the hypsochromic shifts in the spectra of **3**, **4**, **10**, **11a**, and **12a** are not caused by a high proportion of *Z*- (*cis*) isomers, but rather by unfavorable interactions between these highly non-polar carotenoids and the much more polar, mostly-acetonitrile solvent.

Of special interest are C_{50} carotenoids **11b** and **12b/c**. Compared with **11a**, the former has a spectrum that is hypsochromically shifted by 11 nm at the wavelength of maximum absorption, λ_{max} (**Figure 3.6**), elutes slightly earlier in reverse-phase HPLC, is only one-tenth as abundant, and has an identical molecular ion at *m/z* = 674.2 (**Table 3.1**). We believe that these properties are best explained by an unusual desaturation pattern in **11b** in which all three desaturation steps are on one side of the molecule (**Figure 3.3**) (we denote this as "3+0" desaturation). We are not aware of other carotenoids with a 3+0 desaturation pattern whose spectra we can compare with that of **11b**. However, 7,8,11,12-tetrahydrolycopene, an isomer of ζ-carotene with a 2+0 desaturation pattern instead of the 1+1 pattern of ζ-carotene (see **Figure 3.4**), has an absorption spectrum that is hypsochromically shifted by ~5 nm compared with that of ζ-carotene (9, 14, 15, 51). Product **12b/c** has a spectrum that is hypsochromically shifted by 4 nm at λ_{max} compared with that for **12a** (**Figure 3.6**), and it elutes slightly before **12a** in HPLC (**Table 3.1**). This species is only slightly less abundant than **12a** and has the same molecular ion at *m/z* = 672.3 (**Table 3.1**). These properties suggest that **12b/c** is a 4-step desaturated C_{50} carotenoid with a 3+1 desaturation pattern (**12b** in **Figure 3.3**). However, we cannot rule

out the possibility that **12b/c** has a 4+0 desaturation pattern (**12c**) or is a mixture of **12b** and **12c**, which is why we have designated its spectrum in **Figure 3.6** as "**12b/c**."

Carotenoid **5** has the mass of a 5-step desaturated C_{45} carotenoid, and its absorption spectrum corresponds closely with that of 3,4-didehydrolycopene, which has a 3+2 desaturation pattern (9, 25). There are two possible C_{45} carotenoids with this desaturation pattern (**Figure 3.2**), and we could not distinguish by HPLC and MS analysis whether carotenoid **5** has specific structure **5a** or **5b**, or represents a mixture of the two. Similarly, it is not possible to confirm whether carotenoid **4** has precise structure **4a** or **4b** without synthetic standards or the use of advanced nuclear magnetic resonance techniques requiring up to tens of milligrams of sample (18, 32).

Spectra **6** and **13** are similar to the spectrum of 3,4,3',4'-tetradehydrolycopene with 6 desaturation steps in a 3+3 pattern (9). Spectrum **6** is slightly hypsochromically shifted compared to spectrum **13**; this effect is likely due to a significant proportion of *Z*-isomers in our sample of carotenoid **6**, evidenced by the greater relative absorbance of its *cis* peaks at 396 and 415 nm. Having no reason to believe that either **6** or **13** has an unusual desaturation pattern, we have depicted these structures in **Figures 3.2** and **3.3**, respectively, with symmetrical 3+3 desaturation.

Notably absent in all the extracts we analyzed of cultures expressing $BstFPS_{Y81A}$, $CrtM_{F26L}$ or $CrtM_{F26A,W38A}$, and CrtI were 4-step desaturated C_{45} and 5-step desaturated C_{50} carotenoids. We believe these observations, which may initially seem merely curious, can help to clarify the means by which this desaturase regulates its step number (see Discussion).

Carotenoid **1** was initially surprising to find in the above extracts. This molecule and ζ-carotene have the same molecular ion at *m/z* = 540.2 (**Table 3.1**), but compared with ζ-carotene, carotenoid **1** has a slightly longer HPLC retention time (~19 vs. ~18 min) and a UV-visible spectrum that is hypsochromically shifted by 5 nm at λ_{max} (**Figure 3.6**). Because both molecules are C_{40} carotenoids with equal masses and virtually identical polarities, this wavelength shift cannot be due to a change in solvent–analyte interaction. Rather, we believe the shift is due to a 2+0 desaturation pattern, and that carotenoid **1** is based on an asymmetric C_{40} (C_{15}+C_{25}) backbone as shown in **Figure 3.4** (see Discussion).

E. coli expressing only BstFPS$_{Y81A}$ and a variant of CrtM accumulate hydroxylated carotenoid backbones

When *E. coli* XL1-Blue cells were transformed with the desaturase-free plasmids pUCmodII-*crtM*$_{F26L}$-*bstFPS*$_{Y81A}$ or pUCmodII-*crtM*$_{F26A,W38A}$-*bstFPS*$_{Y81A}$ and grown in liquid culture, they accumulated novel monohydroxylated C_{45} and C_{50} carotenoid backbones **7** and **14** as well as monohydroxylated C_{40} backbones, which may be novel depending on the location of the OH-group and whether the C_{40} backbone is symmetric or asymmetric (**Figures 3.5**, **3.7** and **3.8**). The biosynthesis of these hydroxylated carotenoids was confirmed by MS and chemical derivatization (**Figure 3.9**), and their proportions relative to each other and the unmodified backbones were quite reproducible (**Figure 3.8**). On the other hand, similar XL1-Blue cultures harboring the plasmids pUC18m-*bstFPS*$_{Y81A}$ and pAC-*crtM*$_{F26A,W38A}$ did not produce hydroxylated backbones. Some possible reasons for this disparity are presented in the Discussion.

Although we could not identify the specific locations of the hydroxy groups by HPLC and MS analysis, we can exclude several possible regioisomers. The *in vitro* acetylation reactions with acetic anhydride described in Materials and Methods were positive for all the hydroxylated C_{40}, C_{45}, and C_{50} backbones, with conversions above 90% in all cases. This indicates that the hydroxy groups are primary or secondary (19). Also, the propensity of the hydroxylated and acetylated carotenoids to lose water or acetate, respectively, in atmospheric pressure chemical ionization mass spectrometry (APCI-MS) (**Figure 3.9**) is evidence that the substituents are located in an allylic position (2). Finally, that some of these hydroxylated backbones are desaturated by CrtI *in vitro* (see below) suggests the OH group is located far from the center of the molecule. A hydroxy group located close to the center of a carotenoid backbone, which is where desaturases initiate their catalytic action, might interfere with the ability of a desaturase to process such a substrate.

Only one C_{40} backbone fraction was seen in HPLC; likewise, hydroxylated C_{40} backbones eluted as a single peak (**Figure 3.7**). Attempts to further separate these fractions using a linear gradient of acetonitrile:isopropanol (98:2 to 93:7 over 30 min.) also failed. However, because of the strong evidence that carotenoid **1** is based on the asymmetric C_{40} carotenoid backbone 16-isopentenyl-4'-apophytoene, we nevertheless believe that the C_{40} backbone fraction is a mixture of phytoene and 16-isopentenyl-4'-apophytoene, and that the C_{40}-OH fraction may also be a mixture of hydroxylated versions of these backbones. Indeed, the subtle structural differences between symmetric and asymmetric backbones of the same length should have minimal effects on

chromatographic retention. Therefore, co-elution under our separation conditions would not be unexpected.

Cultures of *E. coli* strain HB101 carrying either plasmid pUCmodII-*crtM*$_{F26L}$-*bstFPS*$_{Y81A}$ or pUCmodII-*crtM*$_{F26A,W38A}$-*bstFPS*$_{Y81A}$ also synthesized hydroxylated C_{40}, C_{45}, and C_{50} carotenoids, but in different relative proportions compared with the XL1-Blue cultures (data not shown). Surprisingly, significant proportions of acetylated C_{40} and C_{45} carotenoid backbones (~10 and ~19 mol% of total carotenoids, respectively) were detected in cultures of HB101(pUC18m-*bstFPS*$_{Y81A}$ + pAC-*crtM*$_{F26A,W38A}$) in addition to smaller amounts of hydroxylated C_{40} and C_{45} carotenoid backbones (~2 and ~4 mol%, respectively). The *in vivo*-acetylated carotenoids behaved identically in HPLC and APCI-MS to their counterparts produced by *in vitro* derivatization of hydroxylated carotenoids with acetic anhydride. These acetylated backbones are probably formed by the reaction of hydroxylated carotenoids with acetyl-CoA, a process that is somehow promoted in HB101 cells carrying these plasmids.

In vitro desaturation experiments with CrtI

In an effort to learn more about the substrate specificity of *E. uredovora* CrtI and to potentially access new desaturated carotenoids, we carried out *in vitro* desaturation reactions in which an invariant amount of *E. coli* lysate containing CrtI was incubated with an individual carotenoid backbone (see Materials and Methods). These experiments, in which substrates dissolved in a small amount of acetone were added to cell lysate, yielded markedly different results compared with carotenoid desaturation in living cells, where carotenoid backbones already present in intact cell membranes are converted by membrane-bound desaturases. *E. uredovora* CrtI converts all the available phytoene to

lycopene (and even tetradehydrolycopene) *in vivo*, leaving no intermediates (33, 56). *In vitro*, on the other hand, the enzyme converted an average of 19% of the phytoene to a mixture of mainly lycopene and intermediate products (**Table 3.3**). A similarly reduced *in vitro* efficiency on both native and non-native substrates was also seen for purified CrtI from *Rhodobacter capsulatus* and CrtN from *S. aureus* (39, 40), and appears to be a general feature of *in vitro* compared with *in vivo* carotenoid desaturation. This reduced efficiency is probably a consequence of suboptimal reaction conditions and substrate delivery *in vitro*. *In vivo*, carotenoids are sequestered together with desaturases in cell membranes, increasing the effective concentration of both enzyme and substrate. *In vitro*, proper association of carotenoid backbones and desaturases may be hampered. Additionally, inactivated desaturases are replenished with newly synthesized copies *in vivo* but not *in vitro*.

Of the intermediates detected and quantified by HPLC, 1-step phytofluene and 2-step ζ-carotene were each present at approximately twice the level of 3-step neurosporene (**Table 3.3**, structures of these carotenoids are shown in **Figure 3.4**). This result qualitatively agrees with that of Fraser et al., who detected phytofluene and ζ-carotene but not neurosporene in the product mixture when they supplied purified *E. uredovora* CrtI with phytoene *in vitro* (24).

If we take the *in vitro* conversion statistics as a measure of the degree of compatibility between a substrate and CrtI, then comparing the fractional conversion of other backbones to the benchmark of ~19% conversion of phytoene can allow us to assess the substrate range of the enzyme, at least with respect to this relevant parameter. In this context, the C_{30} diapophytoene is a much less favored substrate for CrtI, while the

C_{45} backbone ranks between phytoene and diapophytoene in acceptability (**Table 3.3**). The C_{50} backbone diisopentenylphytoene was not converted at all *in vitro*, reaffirming our *in vivo* results showing that a substantial proportion of this backbone synthesized in *E. coli* cells remained unmetabolized by CrtI. These results also further highlight the reduced efficiency that carotenoid desaturases display *in vitro*.

In addition to its much lower conversion *in vitro* (~5% vs. ~100% *in vivo*), the C_{45} backbone also underwent fewer desaturation steps *in vitro*, with the terminal *in vitro* product being the 2-step carotenoid **3**. The predominant C_{45} desaturation product *in vitro* was the 1-step carotenoid **2**, which was not detected at all *in vivo*. Therefore, the reduced step number displayed by CrtI *in vitro* on the C_{45} backbone allowed access to a novel carotenoid that did not accumulate in cultured cells.

CrtI also desaturated hydroxylated C_{40} and C_{45} backbones, resulting in the production of (at least) two more new carotenoids, **8** and **9**. (The desaturated monohydroxy C_{40} carotenoids OH-phytofluene, OH-ζ-carotene, OH-neurosporene, and OH-lycopene (see **Figure 3.5**) may also be novel, depending on the location of the hydroxy group.) Although the fungal carotenoid desaturase Al-1 from *Neurospora crassa* was reported to desaturate 1-OH-neurosporene and 1-OH-lycopene (25) and the 3-step desaturase from *R. capsulatus* was shown to efficiently accept 1,2-epoxyphytoene as a substrate (39), this is to our knowledge the first report of the desaturation of oxygenated carotenoids by a bacterial desaturase. Because the precise position of the hydroxy group in the C_{45}-OH substrate **7** is unknown, we do not know the exact structures of products **8** and **9**. Accordingly, we have represented these carotenoids only by name in **Figure 3.5**. Interestingly, the average fractional conversion of the hydroxylated C_{45} backbone was

almost three times that of the unmodified C_{45} backbone, implying that the hydroxy group promotes desaturation *in vitro*. We are uncertain of the reason for this, but hypothesize that the increased polarity brought about by hydroxylation may enhance the ability of the substrate to be solubilized and therefore mix with the enzyme in the reaction medium. The hydroxy group on the C_{50}-OH substrate **14**, however, did not appear to assist desaturation by CrtI *in vitro*. Neither it nor its unmodified counterpart was desaturated in these experiments. As with the C_{45} backbone, the C_{45}-OH backbone **7** was desaturated two steps *in vitro*, with a greater proportion of 2-step products accumulating in the latter reactions. The hydroxy group clearly does not interfere with the desaturation sequence of CrtI, at least up to the first two steps. We interpret this as evidence that the OH group is located distal to the center of the molecule.

Hydroxylated C_{40} carotenoid backbones were also converted at a relatively high level (**Table 3.3**). It is not certain whether this substrate pool was a mixture of hydroxylated phytoene and hydroxylated 16-isopentenyl-4'-apophytoene, and if so, the proportion of each. However, judging by the relatively similar percentages of 1- to 4-step products in the phytoene and C_{40}-OH desaturation experiments, we surmise that most of the C_{40}-OH pool was hydroxylated phytoene. Therefore, we did not depict any hydroxylated asymmetric C_{40} pathway products in **Figure 3.5**. We are not sure of the reason(s) for the increased proportion of 3-step products in the reactions with C_{40}-OH backbones compared to phytoene. This may be somehow due to the hydroxy group on the substrate, but many other explanations are plausible. HPLC analysis of the 2-step desaturation products from the C_{40}-OH reactions did not yield convincing evidence of a hydroxylated derivative of carotenoid **1**. This can be taken as additional evidence that

hydroxylated phytoene dominated the C_{40}-OH backbone population. As with the C_{45}-OH backbone, the hydroxy group on the C_{40}-OH substrate does not inhibit the desaturation sequence of CrtI, as evidenced by the high proportion of 4-step desaturation products. We believe this result points to a terminal or near-terminal position of the OH group in the hydroxylated C_{40} backbones.

Carotenoid biosynthesis in cultures expressing the C_{25}PP synthase from *Aeropyrum pernix*

In an attempt to increase the proportion of C_{50} carotenoids synthesized by our recombinant *E. coli*, we substituted BstFPS$_{Y81A}$ with the C_{25}PP synthase from the thermophilic archaeon *Aeropyrum pernix*, ApFGS. Farnesylgeranyl diphosphate (FGPP, C_{25}PP) is a very rare molecule in nature, and *apFGS* is the only natural C_{25}PP synthase gene sequenced to date (50). We cloned *apFGS* from *A. pernix* genomic DNA and coexpressed it with *crtM*$_{F26A,W38A}$ on one plasmid as well as on separate plasmids.

ApFGS turned out to be disappointing as a C_{25}PP synthase and mainly behaved as a C_{20}PP synthase in our cells, as shown in **Figure 3.10**. When CrtM$_{F26A,W38A}$ is expressed alone in *E. coli* XL1-Blue(pUC18m-*crtM*$_{F26A,W38A}$), only C_{30} and C_{35} carotenoids are synthesized (**Figure 3.10**, first data set). Expression of wild-type CrtM alone in *E. coli* results in almost exclusive production of C_{30} diapophytoene (54), but CrtM$_{F26A,W38A}$ is a much poorer C_{30} synthase and apparently a very effective scavenger of C_{20}PP. Therefore, even though the endogenous level of C_{20}PP is very low in *E. coli* and cannot support C_{40} phytoene synthesis by the native C_{40} synthase CrtB (see Chapter 2) or by CrtM$_{F26A,W38A}$ (**Figure 3.10**, first data set), the latter enzyme condenses abundant C_{15}PP with much less abundant C_{20}PP to such an extent that C_{30} backbones outnumber C_{35} backbones by only a

factor of 2 in XL1-Blue(pUC18m-*crtM*$_{F26A,W38A}$). When ApFGS was coexpressed with CrtM$_{F26A,W38A}$ together on plasmid pUC18m-*apFGS-crtM*$_{F26A,W38A}$, approximately equal proportions of C_{35} and C_{40} carotenoid backbones were synthesized (**Figure 3.10**, second data set). This indicates that ApFGS primarily synthesizes C_{20}PP and also that a substantial amount of C_{15}PP remains available for carotenoid biosynthesis in XL1-Blue(pUC18m-*apFGS-crtM*$_{F26A,W38A}$) cells. When the same two genes were expressed on separate plasmids, XL1-Blue cultures harboring pUC18m-*apFGS* and pAC-*crtM*$_{F26A,W38A}$ synthesized the C_{40} backbone phytoene almost exclusively, with traces of C_{45} and C_{50} backbones also being detected (**Figure 3.10**, fourth data set). This difference is most likely due to the differential copy number of the two plasmids. Being on a high-copy pUC-based plasmid, *apFGS* should be more highly expressed than *crtM*$_{F26A,W38A}$, which is on a medium-copy pAC-based plasmid (both genes are under the control of identical *lac* promoter and operator sequences). The higher relative ApFGS level should therefore allow this enzyme to more completely consume the available C_{15}PP before it can be incorporated into carotenoids by CrtM$_{F26A,W38A}$. However, even when given this opportunity, ApFGS still releases primarily C_{20}PP and only small amounts of C_{25}PP.

ApFGS was shown to prefer C_{20}PP as a substrate for synthesizing C_{25}PP (50). Therefore, to see if increasing the supply of C_{20}PP in the cells would result in a higher proportion of C_{45} or C_{50} carotenoids, we co-transformed XL1-Blue cells with pUC18m-*apFGS-crtM*$_{F26A,W38A}$ and pAC-*crtE* containing the C_{20}PP synthase gene from *E. uredovora*. This resulted in a ~3-fold increase in total carotenoids, and some C_{45} backbones were detected, but phytoene represented the vast majority of the carotenoid

backbone pool in these cells (**Figure 3.10**, third data set). Evidently, ApFGS could not effectively compete with $CrtM_{F26A,W38A}$ for the additional $C_{20}PP$ supplied by CrtE.

A. pernix is a thermophilic organism, and ApFGS probably has an optimum temperature much higher than that at which *E. coli* cultures can be grown. However, cultivating ApFGS-expressing cultures at 37 °C instead of the usual 28 °C at which we grow our carotenogenic *E. coli* cultures did not increase the proportion of C_{45} or C_{50} carotenoids (data not shown). When $CrtM_{F26A,W38A}$ was coexpressed with ApFGS at 37 °C, almost no carotenoid backbones were synthesized. This may be due to compromised stability of this CrtM double mutant.

The distribution of carotenoids synthesized by cells expressing $BstFPS_{Y81A}$ is strongly influenced by Idi overexpression and physiological state

It is known that the biosynthesis of isoprenoids in *E. coli* is limited by the supply of the C_5PP "starter unit" dimethylallyl diphosphate (DMAPP, structure shown in **Figure 1.4**), and that overexpression of isopentenyl diphosphate isomerase (Idi) can increase carotenoid titers by approximately an order of magnitude (59). In an effort to increase our carotenoid titers, we transformed XL1-Blue cells with plasmids pUC18m-$bstFPS_{Y81A}$-*idi* (containing the *idi* gene from *E. coli*) and pAC-$crtM_{F26A,W38A}$. After culturing these cells in TB medium, we found that Idi overexpression resulted in a ~10-fold increase in the quantity of carotenoids synthesized by the cells, but also dramatically shifted the distribution of carotenoid backbones toward production of phytoene (cf. first and second data sets in **Figure 3.11**). This result is likely due to alteration of the relative levels of isoprenyl diphosphate precursors of different lengths brought about by Idi overexpression (see Discussion).

In the course of our experiments involving coexpression of *bstFPS*$_{Y81A}$, *crtM*$_{F26L}$ or *crtM*$_{F26A,W38A}$, and *crtI*, we noticed that colonies expressing these genes from one or two plasmids reproducibly displayed a dramatically deeper red color than cell pellets from liquid cultures harboring the same plasmid(s). To investigate this further, we collected a sufficient quantity of the colonies (~0.5 g wet cells) to permit carotenoid extraction and HPLC analysis, and compared the carotenoid content of the colonies with that of the cell pellets. We discovered that the colonies primarily synthesized C_{40} carotenoids and barely made any C_{45} and C_{50} products, while the liquid cultures inoculated by the very same colonies made substantial amounts of the larger carotenoids. This effect was also observed with no desaturase present (cf. first and third data sets in **Figure 3.11**) and occurred whether colonies were grown on LB- or TB-agar plates. Therefore, the dissimilar carotenoid distributions do not result from "defective" colonies, but rather, from differences in the physiology of *E. coli* colonies and liquid cultures. The strikingly similar shift toward an increased proportion of C_{40} carotenoids at the expense of larger products caused by Idi overexpression and by cell growth in colonies is examined in more detail in the Discussion.

Attempted creation of a C_{60} carotenoid biosynthetic pathway

To ascertain whether our best C_{50} carotenoid-producing CrtM variants or the C_{40} synthase CrtB could also synthesize C_{60} carotenoid backbones, we coexpressed the hexaprenyl diphosphate (C_{30}PP) synthase HexPS from *Micrococcus luteus* (47) with CrtM$_{F26L}$, CrtM$_{F26A,W38A}$, or CrtB on plasmids pUC18m-*hexPS-crtM*$_{F26L}$, pUC18m-*hexPS-crtM*$_{F26A,W38A}$, and pUC18m-*hexPS-crtB*, respectively. However, we did not detect any carotenoids at all in cultures of XL1-Blue transformed with these plasmids. This

result indicates that HexPS is an active enzyme that consumes C_{15}PP, since expression of *crtM*$_{F26A,W38A}$ alone on a pUC18m-based plasmid yields C_{30} and C_{35} carotenoids (**Figure 3.10**), and also that these synthases cannot synthesize C_{60} carotenoids. Additional evidence also demonstrates that HexPS efficiently consumes C_{15}PP and specifically synthesizes C_{30}PP. XL1-Blue colonies harboring pUC18m-*crtM-crtN* are bright yellow due to the synthesis of large amounts of desaturated C_{30} carotenoids (**Figure 3.12a**, panel 1). However, colonies transformed with pUC18m-*hexPS-crtM-crtN* are colorless, indicating considerable consumption of C_{15}PP by HexPS (**Figure 3.12a**, panel 2). Coexpression of the C_{40} carotenoid genes CrtB and CrtI further demonstrates that HexPS does not release C_{20}PP. Whereas colonies harboring pUC18m-*crtE-crtB-crtI* are red-pink due to production of desaturated C_{40} carotenoids beginning with the synthesis of C_{20}PP by CrtE (**Figure 3.12a**, panel 3), colonies transformed with plasmid pUC18m-*hexPS-crtB-crtI* are colorless because *hexPS* does not supply CrtB with C_{20}PP (**Figure 3.12a**, panel 4). Finally, *in vitro* experiments performed by Adam Hartwick show that when supplied with C_{15}PP and isopentenyl diphosphate (IPP, structures shown in **Figures 3.1-3.5**), CrtE synthesizes the expected C_{20}PP product (**Figure 3.12b**, lane 1), BstFPS$_{Y81A}$ synthesizes the expected C_{25}PP product (**Figure 3.12b**, lane 2), and HexPS specifically synthesizes an even larger product (**Figure 3.12b**, lane 3). Similar experiments performed by Shimizu et al. confirmed with authentic standards that *M. luteus* HexPS indeed synthesizes primarily C_{30}PP from C_{15}PP and IPP (47).

No known carotenoid synthase, including our CrtM mutants, can convert C_{30}PP to a C_{60} carotenoid backbone. It may be possible to evolve one of these mutants or another carotenoid synthase in the laboratory for C_{60} carotenoid synthesis. However, our results

showing poor conversion of the C_{50} backbone to desaturated carotenoids by CrtI (**Tables 3.1** and **3.3**) and no conversion by the C_{30} desaturase CrtN (data not shown) indicate that, should a biosynthetic route to C_{60} backbones be established, desaturation of this substrate would be even less efficient by these enzymes. Therefore, no simple colorimetric assay exists for the screening of synthase variants in a directed evolution experiment aimed at evolving a C_{60} carotenoid synthase.

DISCUSSION

Determination of desaturation step number by *Erwinia* CrtI

Elegant genetic complementation experiments with heterokaryons of the fungus *Phycomyces blakesleeanus* have provided convincing evidence for the existence of multienzyme complexes that function as assembly lines for carotenoid biosynthesis in that organism (5, 12, 16, 36). In these complexes, carotenoid substrates undergo stepwise chemical transformations as they are processed by one enzyme and then passed on to the next one in the complex, hence the analogy to an industrial assembly line. Therefore, phytoene undergoes four desaturation steps in *Phycomyces* because four desaturase subunits are present in that organism's carotenoid biosynthetic enzyme complexes (5). Carotenogenic enzyme complexes are believed to be widespread in other organisms as well, although there are many uncertainties regarding the factors that determine the extent of carotenoid desaturation (7).

Our results on the *in vivo* desaturation of the C_{45} backbone isopentenylphytoene by *E. uredovora* CrtI provide evidence against the idea that the number of subunits in a complex determines the desaturation step number. As mentioned, *E. uredovora* CrtI is

primarily a 4-step desaturase in the symmetric C_{40} pathway, converting phytoene to lycopene as the majority product (24, 33, 56). If the primary determinant of this 4-step product specificity were the association of CrtI subunits as tetramers, we should also expect primarily 4-step products with other carotenoid backbones as well, assuming these substrates are accepted and processed by the complex. However, not only is this not the case in the C_{45} pathway, but 4-step C_{45} products do not accumulate at all, even though 5- and 6-step C_{45} carotenoids are detected (**Table 3.1**, **Figure 3.2**). Therefore, 4-step C_{45} carotenoids are "skipped over" by CrtI *in vivo*, indicating a preference for synthesizing the higher step-number C_{45} products **5** (5 steps, majority product) and **6** (6 steps, second-most abundant).

We believe this observation is related to two others. First, in the C_{50} pathway, we observed no accumulation of 5-step products *in vivo* (**Table 3.1**, **Figure 3.3**), even though 6-step C_{50} products were formed. Second, *in vitro* desaturation experiments with *E. uredovora* CrtI on phytoene have shown that the 3-step neurosporene is the least abundant intermediate, accumulating at low levels (**Table 3.3**) or not at all (24). (*In vivo*, the enzyme leaves behind no intermediates and even catalyzes six desaturation steps on phytoene (33, 56).)

These seemingly disparate phenomena are all connected by a common trait shared by neurosporene and the likeliest possible 4-step C_{45} and 5-step C_{50} products: all have one ψ-end and one dihydro-ψ-end. This structural feature (see **Figure 1.8**) shared by all three disfavored products implies that *E. uredovora* CrtI has a strong propensity to avoid terminating its desaturation sequence at products with this combination of ends, regardless of the size of the carotenoid backbone substrate or the number of steps

required. Therefore, CrtI appears to regulate desaturation step number by sensing its substrate's end groups, with particular preference for carotenoids with one ψ- and one dihydro-ψ-end, which are desaturated with high efficiency. The enzyme may accomplish this by having a higher affinity for substrates with these ends, but other strategies are possible. Whatever the specific mechanism, our results on C_{45} and C_{50} backbones clarify that the step number of this enzyme is significantly influenced by the size and end groups of its substrate. Thus, nature has apparently solved the problem of regulating carotenoid desaturation step number in at least two very different ways, as demonstrated by the distinct biochemical strategies employed by the phytoene desaturases of *E. uredovora* and *P. blakesleeanus*.

These results also shed light on the functional plasticity displayed by bacterial carotenoid desaturases in directed evolution experiments aimed at altering desaturation step number (Chapter 1, refs. (46, 53, 58)). If a change in desaturation step number required a change in multimeric state, then converting a 4-step desaturase into a 6-step enzyme would require conversion of a tetrameric enzyme into a hexameric one. This would have to result from only a small number of mutations. While mutations can abolish the ability of a protein to form multimers, it seems much less plausible that the number of subunits in a complex could be so finely-tuned by minimal mutation. It is easier to envision how a desaturase's catalytic rate or tendency to synthesize products with particular end groups could be modified by mutation. Therefore, the relative ease and high frequency with which desaturation mutants have been discovered by directed evolution of bacterial desaturases are more easily rationalized in the light of altered specificity rather than multimeric state.

Evidence of a novel asymmetric C_{40} carotenoid biosynthetic pathway

We stated earlier that carotenoid **1** is likely the 2+0 desaturation product of an asymmetric C_{40} (C_{15}+C_{25}) carotenoid backbone, as shown in **Figure 3.4**. Although **1** has the same mass and chromophore size as ζ-carotene, it is unlikely that **1** is ζ-carotene with the hypsochromic shift in its spectrum resulting from *Z*-isomerization. The spectrum of **1** (**Figure 3.6**) shows the hallmarks of a majority all-*E* sample population: the *cis* peaks at 286 and 295 are low (~10% of λ_{max}) and the ratio of the height of the longest-wavelength absorption band (419 nm) to the absorption at λ_{max} (the so-called "III/II" ratio (9)) is 1.0. In fact, the spectrum shown for ζ-carotene in **Figure 3.6** is that of a sample with more *Z*-isomer content than **1**. In that spectrum, III/II is only 0.83 (for all-*trans* ζ-carotene, III/II is between 1.0 and 1.028 (14, 15)).

A 2+0 desaturation pattern also results in a hypsochromic shift of 5 nm, as shown by comparison of the spectrum of the 2+0 desaturation product of phytoene, 7,8,11,12-tetrahydrolycopene with that of ζ-carotene, a 1+1 desaturated carotenoid (9, 14, 15, 51). However, although the absorption spectrum of 7,8,11,12-tetrahydrolycopene is strikingly similar to that of **1** (9, 14, 15, 51), it is also unlikely that carotenoid **1** is 7,8,11,12-tetrahydrolycopene, because *E. uredovora* CrtI desaturates phytoene to lycopene via ζ-carotene (**Table 3.3**, ref. (24)) and has not been shown to synthesize 7,8,11,12-tetrahydrolycopene. Furthermore, a 2-step desaturation product of phytoene like ζ-carotene or 7,8,11,12-tetrahydrolycopene is not expected to accumulate in a culture expressing *E. uredovora* CrtI. As mentioned previously, this enzyme desaturates all the available phytoene to lycopene and even 3,4,3',4'-tetradehydrolycopene *in vivo*, with no intermediates being detected. Indeed, in the same cultures in which **1** was found,

lycopene was also present and was usually at least twice as abundant; tetradehydrolycopene was also detected in smaller amounts (data not shown). Finally, the presence of a 2-step C_{40} carotenoid cannot be explained by a "phytoene overload" that overwhelms CrtI, for even in engineered *E. coli* that accumulate orders of magnitude more carotenoids than XL1-Blue, this desaturase is capable of efficiently converting the vastly increased phytoene supply to lycopene despite being expressed from a low-copy plasmid (20).

The most reasonable conclusion from the above evidence is that carotenoid **1** has an asymmetric C_{40} (C_{15}+C_{25}) carbon backbone that has undergone two desaturation steps on its C_{15}-side. Although we were unable to separate the asymmetric C_{40} backbone 16-isopentenyl-4'-apophytoene from phytoene by HPLC of culture extracts of *E. coli* cultures expressing only BstFPS$_{Y81A}$ and a mutant of CrtM (**Figures 3.7** and **3.8**), it is reasonable to expect that the former backbone is made by these cultures, whose additional synthesis of C_{35}, C_{45}, and C_{50} carotenoid backbones proves that both C_{15}PP and C_{25}PP are present in the cells. Furthermore, it is not surprising that CrtI would desaturate 16-isopentenyl-4'-apophytoene differently than phytoene. Although equal in size, the former is the most asymmetric of the six possible carotenoid backbones shown in **Figure 3.1**, and this likely affects the catalytic action of CrtI. When presented with the asymmetric C_{40} backbone, CrtI apparently catalyzes two desaturation steps on the C_{15}-side of the substrate (which is all that this side can accommodate) rather easily, but has trouble desaturating the C_{25}-side of the molecule; therefore, product **1** accumulates. (This scenario seems more probable than 2 steps on the C_{25}-side and none on the C_{15}-side since CrtI can desaturate the C_{30} (C_{15}+C_{15}) backbone more efficiently than the C_{50} (C_{25}+C_{25})

backbone (**Table 3.3**).) We are unsure why we did not detect any asymmetric C_{40} pathway products with longer chromophores than **1**. Given that CrtI can desaturate C_{45} and C_{50} backbones, it is possible that some 4-6 step asymmetric C_{40} pathway products were made but were not distinguished from symmetric C_{40} carotenoids in our analysis, leaving only product **1** to stand out because of the lack of ζ-carotene produced. Because of this possibility, we could not quantify the relative proportions of symmetric and asymmetric C_{40} carotenoids made in the cultures.

Potential utility of novel desaturated C_{45} and C_{50} carotenoids

In this chapter, we report the biosynthesis of at least 14 new carotenoids based on unnatural carbon backbones. However, it remains to be seen whether any of these new products will prove technologically useful. Carotenoids are primarily sold for use as pigments, nutriceuticals, and as food and animal feed supplements (34). In some circumstances, a specific natural carotenoid is required, e.g., astaxanthin for salmon feed. However, in other current and envisioned applications such as antioxidant research and even microelectronics (11), "designer" molecules with specific properties not seen in natural carotenoids may prove extremely useful.

The chromophore of a carotenoid is the primary determinant of its color and chemical properties (8, 61, 63). C_{45} and C_{50} carotenoid backbones can in principle accommodate chromophores of 17 and 19 conjugated carbon-carbon double bonds, respectively, which is longer than the 15 possible on a C_{40} backbone. Carotenoids with such large chromophores can absorb light at wavelengths above 600 nm, and their colors range from dark red to bluish purple (29, 35). Such carotenoids should be extremely reactive toward free radicals and could prove useful as potent antioxidants or colorants.

Although we did not generate carotenoids with more than 15 conjugated double bonds in this work, we have made progress toward this goal, and we believe that it will be possible to discover a desaturase in nature or to evolve one in the laboratory that can fully desaturate a C_{45} or C_{50} backbone. The insights we have gained about CrtI show that to evolve a 7- or 8-step desaturase will not require the discovery of extraordinarily rare mutants able to form heptamers or octamers, but merely variants with an altered preference for synthesizing carotenoids with fully desaturated end groups. Also, it may be possible to chemically desaturate 6-step C_{45} and C_{50} carotenoids even further (64), leading to carotenoids with chromophores of 17 or 19 conjugated double bonds.

Desaturated C_{45} carotenoids **2**, **3**, **4**, **5**, and **6** (**Figure 3.2**) and desaturated C_{50} carotenoids **10**, **11**, **12**, and **13** (**Figure 3.3**) reported in this work may be of interest for research purposes. Studies comparing the reactivity of these molecules toward free radicals in artificial lipid membranes could help researchers dissect the individual contributions of a carotenoid's chromophore, polarity, and end groups in determining this reactivity. For example, it is expected that the novel C_{45} and C_{50} carotenoids produced in this work should orient themselves deep within the hydrophobic core of a lipid bilayer, and such orientation effects are believed to strongly influence the *in vivo* antioxidant properties of a carotenoid (63).

Finally, it has been demonstrated that the presence of a single terminal hydroxy group in a carotenoid can dramatically improve its antioxidative properties (4). Therefore, the hydroxylation of carotenoid backbones we have observed in *E. coli* may be advantageous for the generation of monohydroxylated C_{45} and C_{50} carotenoids with desirable antioxidative abilities. Although we only observed hydroxylation of carotenoid

backbones *in vivo*, we showed that the hydroxylated C_{45} backbone could be desaturated *in vitro*, generating 1- and 2-step monohydroxy C_{45} structures **8** and **9**. Our results showing increased desaturation step number *in vivo* compared with *in vitro* (cf. **Tables 3.1** and **3.3**) point to a promising strategy for generating hydroxylated C_{45} and C_{50} carotenoids with longer chromophores: co-transform cells harboring pUCmodII-*crtM*$_{F26L}$-*bstFPS*$_{Y81A}$ or pUCmodII-*crtM*$_{F26A,W38A}$-*bstFPS*$_{Y81A}$ with another plasmid containing *crtI* under the control of a tightly-regulated promoter, and induce desaturase expression late in the culture. This approach would allow the cells to first accumulate monohydroxylated C_{45} and C_{50} backbones, which would then be desaturated *in vivo* upon desaturase expression. In nature, carotenoids are desaturated before they are hydroxylated. Therefore, it is interesting to note that this scheme could potentially produce highly desaturated monohydroxy C_{45} and C_{50} carotenoids with beneficial characteristics by an inverse biocatalytic route.

Origins of *in vivo*-hydroxylated carotenoids

We remain uncertain of the mechanism by which the carotenoid backbones are hydroxylated *in vivo*. Krubasik et al. observed some hydroxylation of the C_{50} flavuxanthin in recombinant *E. coli* cells (30) and C_{30} carotenoids synthesized in *E. coli* have also been reported to be hydroxylated by unknown processes (31, 40). Lee et al. postulated that their C_{30} diapocarotenoids were hydroxylated by free peroxy radicals present in *E. coli* membranes (31). Like us (56), they did not observe hydroxylation of C_{40} carotenoids in *E. coli*. Albrecht et al. reported a significant amount of epoxidation and hydroxylation of phytoene in cells of the green algae *Scenedesmus acutus* whose phytoene desaturase was inhibited by the herbicide norflurazon (2). Their report

suggested that the origins of the oxygenated phytoene derivatives might be enzymatic, but the oxidation was later attributed to reaction with free radicals (G. Sandmann, personal communication).

As mentioned, XL1-Blue cultures harboring the plasmids pUC18m-*bstFPS*$_{Y81A}$ and pAC-*crtM*$_{F26A,W38A}$ produced no hydroxylated backbones (**Figure 3.8**). These cultures also produced fewer total carotenoids than the single-plasmid cultures (see **Figure 3.8** caption). This expression vector-dependency of carotenoid backbone hydroxylation suggests that *E. coli* cells selectively induce carotenoid oxidation only under certain conditions. Perhaps the lack of a *lac* operator in the pUCmodII-based plasmids and the resulting high-level constitutive expression of the carotenoid biosynthetic genes therein elicits this activation in *E. coli*. Plasmids pUC18m-*bstFPS*$_{Y81A}$ and pAC-*crtM*$_{F26A,W38A}$ have their carotenogenic genes under the control of a *lac* promoter and operator, which should limit gene expression somewhat and possibly avoid the same response in the cells. Alternatively, the induction of carotenoid backbone hydroxylation might be triggered when the quantity of large carotenoid backbones in the cell membrane reaches a critical threshold. It is possible that, in response to excessive loss of membrane fluidity brought on by the accumulation of highly nonpolar C_{45} and C_{50} carotenoid backbones, the cells up-regulate an oxidase or radical-producing enzyme that helps to rectify this problem by effecting hydroxylation of the offending molecules, thereby increasing their polarity. Similarly, accumulation of these carotenoids may disrupt a portion of the respiratory electron transport chain only when their concentration reaches a certain value, causing electron leakage and the production of reactive oxygen species. These hypotheses, although attractive, are only weakly supported by the fact that the carotenoid titer of

XL1-Blue(pUCmodII-*crtM*$_{F26L}$-*bstFPS*$_{Y81A}$) is a mere ~50% higher than that of XL1-Blue(pUC18m-*bstFPS*$_{Y81A}$ + pAC-*crtM*$_{F26A,W38A}$) (see **Figure 3.8** caption).

Coexpressing CrtI along with BstFPS$_{Y81A}$ and either CrtM$_{F26L}$ or CrtM$_{F26A,W38A}$ abolished almost all carotenoid backbone hydroxylation in the cells even though C_{50} backbones were quite abundant and accounted for ~50 mol% of total carotenoids. In these cultures, hydroxylated C_{50} backbones represented only ~1 mol% of the carotenoid pool. A possible reason for this is that free radicals preferentially react with and are quenched by desaturated carotenoids, whose extended chromophores are much more reactive toward radical species compared with undesaturated carotenoid backbones (17, 57, 61, 63).

The above evidence points to radical oxygen species as the culprit in carotenoid backbone hydroxylation. However, the sharp, single HPLC peaks in **Figure 3.7** suggest a high degree of regiospecificity of the hydroxylation reactions and therefore direct enzymatic oxidation. If multiple positions of these backbones were hydroxylated, as might be expected if the mechanism were chemical oxidation by reactive oxygen species, we should observe multiple clustered HPLC peaks, each corresponding to a different positional isomer of a hydroxylated backbone. (Different hydroxylated regioisomers should have much more distinct chromatographic properties than symmetric or asymmetric versions of a C_{40} backbone.) Indeed, an HPLC chromatogram showing multiple clustered peaks corresponding to different regioisomers of hydroxylated phytoene was generated by Albrecht et al. under chromatographic conditions similar to ours (2). On the other hand, the apparent regiospecific hydroxylation of our carotenoid backbones may result from their orientation in the cell membranes. For example, it is

possible that the backbones are positioned in such a manner that one end of the molecule protrudes from the membrane and consequently becomes prone to free-radical oxidation.

Clearly, additional research will be required to uncover the origins of carotenoid backbone hydroxylation in *E. coli*. It would probably be simplest to first conduct *in vitro* investigations of the reactions of the various backbones with free radicals generated *in situ* in a manner similar to that described by Woodall et al. (61). If hydroxylated products from these reactions closely match those we have reported, this result would support the hypothesis that similar radical reactions are responsible for carotenoid hydroxylation *in vivo*. On the other hand, if the products are not a close match, and especially if multiple clustered peaks are observed under the same HPLC conditions, then the possibility of enzymatic hydroxylation might warrant further investigation. We anticipate, however, that the search for an enzyme responsible for these transformations would be a very challenging experiment. For one, screening of genomic libraries would be very difficult since this could not be done in *E. coli*, and selection of another host organism does not seem straightforward. However, the fact that carotenoid hydroxylation only occurs with certain strain and plasmid combinations suggests that microarray identification of differentially expressed genes might be fruitful.

Properties of nascent biosynthetic pathways—lessons for biology and metabolic pathway engineering

In addition to producing several novel carotenoids based on unnatural carbon scaffolds, this work has allowed us to investigate the precise biomolecular phenomena that resulted in the evolution of three novel biosynthetic pathways to carotenoid pigments. In this section, we analyze the specific mechanisms by which novel chemical diversity

was generated and propagated along these new pathways by their constituent enzymes. We then discuss key properties our pathways likely share with nascent biosynthetic pathways in general.

The Y81A amino acid substitution in $BstFPS_{Y81A}$ dramatically broadens the specificity of this enzyme, converting it from a strict $C_{15}PP$ synthase into an enzyme that synthesizes a mixture of $C_{15}PP$, $C_{20}PP$, and $C_{25}PP$. This mutation has been studied in detail, and its effect on the enzyme's product range has been shown to result from an enlarged product elongation pocket (37). The Y81A mutation transforms BstFPS into a "statistical" biocatalyst. Although capable of synthesizing products up to $C_{25}PP$ by catalyzing two additional condensations of IPP with the growing prenyl chain, $BstFPS_{Y81A}$ also releases intermediates $C_{15}PP$ and $C_{20}PP$; hence a distribution of prenyl diphosphate products with 3, 4, and 5 isoprene units is generated. The diversity of this product mixture is then amplified by the next enzyme in the pathway, a mutant of the *S. aureus* C_{30} carotenoid synthase CrtM.

In Chapter 1, we discussed the molecular effects of mutations at F26 and W38 in CrtM. A homology model with human squalene synthase allowed us to map these residues to the site of the second half-reaction and led us to propose that their substitution with smaller amino acids increases the size of the cavity where cyclopropyl intermediates are rearranged into carotenoid backbones. From this analysis, we conjectured that wild-type CrtM could accept $C_{20}PP$ and even $C_{25}PP$ as substrates and form cyclopropyl intermediates from these precursors, but cannot rearrange intermediates larger than C_{35} into carotenoid backbones. When additional space is created in the rearrangement pocket by replacement of F26 and/or W38 with smaller amino acids, conversion of the larger

intermediates becomes possible, resulting in the synthesis of carotenoid backbones up to C_{50}. Like $BstFPS_{Y81A}$, these CrtM mutants have lost much of their capacity to regulate the size of their products. When coexpressed with $BstFPS_{Y81A}$ *in vivo*, the CrtM mutants condense every possible pairwise combination among $C_{15}PP$, $C_{20}PP$, and $C_{25}PP$ except, intriguingly, the wild-type's combination of two molecules of $C_{15}PP$ (a consequence of the relatively low $C_{15}PP$ concentration combined with the intrinsically poor ability of the CrtM variants to make C_{30} carotenoids). The result is a mixture of C_{35}, symmetric (C_{20}+C_{20}), and asymmetric (C_{15}+C_{25}) C_{40}, C_{45}, and C_{50} carotenoid backbone products.

The desaturase CrtI then further diversifies the five carotenoid backbones generated by the two previous enzymes. CrtI possesses no mutations, but its inherent promiscuity allows it to accept all of the above backbones to some degree. In the C_{45} and C_{50} pathways, the step number of CrtI is not very well defined (except for the absence of certain products discussed above), and there was no clear majority desaturation product in these pathways *in vivo* (**Table 3.1**). Therefore, the stuttering action of CrtI on the C_{45} and C_{50} backbones is the primary catalytic basis for the branching of these new pathways. As previously discussed, CrtI very efficiently catalyzes four desaturation steps on its native substrate phytoene *in vivo*. Under less optimal *in vitro* conditions, however, substantial stuttering occurs with the same substrate, and all possible intermediates in the desaturation sequence are detected (**Table 3.3**). In the more favorable milieu of a cell membrane, however, CrtI can still be made to stutter when presented with unnaturally large carotenoid backbones. We do not know the exact reason for this, but expect it is due to a decreased catalytic efficiency on these substrates. CrtI also produces carotenoids with unusual desaturation patterns like **1**, **11b**, and **12b/c** (**Figures 3.3**, **3.4**, and **3.6**;

Table 3.1), indicating that unnaturally large or asymmetric substrates can cause the enzyme to become slightly "confused" and carry out non-standard desaturation sequences. It is interesting that the unusual catalytic behavior that can result when an iterative enzyme is challenged with a new substrate can serve as a diversity-generating mechanism in the evolution of novel biosynthetic pathways.

This work demonstrates the rapid product diversification that is possible when biosynthetic pathways are constructed from multiple broad-specificity enzymes arranged consecutively. Such arrangements were likely very common in the early evolution of metabolism. The widely accepted "patchwork" model states that the first metabolic pathways were assembled from promiscuous enzymes whose specificities were then narrowed by natural selection, yielding the relatively specific pathways we see in nature today, especially in so-called primary metabolism (27, 41, 62). Our laboratory-generated carotenoid biosynthetic pathways also probably share key features with recently evolved metabolic pathways in nature. Most notably, a lack of overall pathway specificity characterized by the generation of multiple products is a trait likely common to nascent pathways created both in nature and in the laboratory. Directed evolution experiments on a wide range of enzymes demonstrate that it is rare for enzymes to completely shift their substrate or product specificity after accumulating only a small number (1-2) of mutations. Rather, mutants with broadened specificity are the norm, even if the selection pressure of the experiment requires only one particular reaction to be catalyzed (1). In nature, enzymes that have recently evolved the ability to accept non-native substrates such as man-made antibiotics also tend to have broadened rather than shifted specificities, hence the emergence of so-called "extended-spectrum" beta-lactamases (13). Therefore,

new biosynthetic pathways (or pathway branches) that have emerged in nature as a result of enzyme mutation probably possess low product specificity, similar to our nascent carotenoid pathways, especially when multiple enzymes have been mutated. In many if not most cases, it remains to be seen how many more mutations would be required and how strong the selection pressures would have to be in order to convert newly emerged pathways into mature ones capable of specifically synthesizing a particular complex end product.

Analysis of the means by which our novel carotenoid pathways have emerged also helps to clarify some aspects of a popular theory about the evolution of natural product biosynthetic pathways. The "screening hypothesis" put forth by Jones and Firn argues that natural product pathways have evolved under selection for particular traits, such as pathway branching and enzymatic promiscuity, because such traits promote the production and retention of product diversity at minimal cost (Chapter 1, refs. (21-23, 28)). However, our highly branched carotenoid pathways comprising laboratory-evolved promiscuous enzymes emerged despite a lack of selection for these properties. In no instance was selection for a diverse array of products applied in the laboratory evolution and subsequent site-directed mutagenesis experiments that resulted in the discovery of mutants $BstFPS_{Y81A}$, $CrtM_{F26L}$, and $CrtM_{F26A,W38A}$ (37, 38, 54-56). Yet, in all these cases, as is typical in directed evolution (see previous paragraph), enzyme variants with broadened specificity were obtained. This implies that traits such as enzyme promiscuity and the types of pathway branching to which it can lead do not require selection to arise or be maintained, but merely represent the default state of enzymes and metabolic pathways in the absence of strong and sustained selective pressure to be highly specific.

Therefore, we believe that the low specificity of many natural product biosynthetic enzymes and pathways is not the outcome of positive selection for promiscuity, but is more aptly attributed to genetic drift caused by weak selection pressure for narrow specificity, or to selection pressures that change rapidly and hence do not allow sufficient evolutionary time for any particular narrow specificity to fully develop.

Challenges for achieving new pathway specificity

In the Results section, we described several unsuccessful experiments aimed at establishing more specific pathways in *E. coli* to particular novel-backboned carotenoids. For example, it should be possible to increase the proportion of C_{50} carotenoids by increasing the proportion of C_{25}PP relative to C_{15}PP and C_{20}PP. If it were possible to drive the fraction of C_{25}PP to nearly 100%, it would not be necessary to further evolve a carotenoid synthase for increased C_{50} product specificity because only one substrate would be available and therefore only the C_{50} carotenoid backbone would be possible. As mentioned earlier, C_{25}PP is a very rare molecule in nature. Its only biological role found to date is as a precursor for the C_{25}-C_{25} and C_{20}-C_{25} ether-linked membrane lipids of some archaea such as *Aeropyrum pernix* and *Natronomonas pharaonis* (6, 49, 50). C_{25}PP synthase enzymes are correspondingly very rare in nature, and only one such enzyme has been cloned to date: the farnesylgeranyl diphosphate synthase from *Aeropyrum pernix*, ApFGS (50). We cloned the *apFGS* gene from *A. pernix* genomic DNA, but found it inferior to BstFPS$_{Y81A}$ for producing C_{25}PP in *E. coli* cells. ApFGS actually behaved more like a C_{20}PP synthase with peripheral C_{25}PP synthase activity in our carotenogenic *E. coli* (**Figure 3.10**).

In their paper describing the first cloning of ApFGS, Tachibana et al. showed that the enzyme produces a mixture of C_{10}PP to C_{25}PP products, with C_{20}PP being the most abundant, when supplied with the C_5PP starter unit DMAPP and the C_5PP extender unit IPP *in vitro* (50). (See **Figure 1.4** for structures.) Only when supplied with C_{20}PP did ApFGS exhibit specific C_{25}PP synthase ability, but neither the rate nor the extent of the reaction was reported. The extensive production of C_{40} carotenoids in *E. coli* expressing ApFGS and $CrtM_{F26A,W38A}$ (**Figure 3.10**) appears to stem from the ability of the latter enzyme to incorporate C_{20}PP into carotenoids faster than it can be readmitted to ApFGS for further prenyl chain elongation to C_{25}PP. A similar effect likely occurs with $BstFPS_{Y81A}$. When supplied with C_{15}PP and IPP *in vitro*, this enzyme synthesizes C_{25}PP as the major product, with much smaller amounts of C_{20}PP being detected (**Figure 3.12** and ref. (37)). However, when $BstFPS_{Y81A}$ is coexpressed with a CrtM variant in *E. coli*, significant amounts of C_{40} and C_{45} carotenoids are synthesized (**Figure 3.8**), indicating that C_{20}PP is frequently released by $BstFPS_{Y81A}$ *in vivo*. Production of C_{20}PP by $BstFPS_{Y81A}$ is more common in the elongation of the 5-carbon DMAPP (see below), a substrate also present in *E. coli*. Whereas *in vitro*, released C_{20}PP is readily readmitted to $BstFPS_{Y81A}$ and extended to C_{25}PP (37), in the above cells, $BstFPS_{Y81A}$ must compete with a CrtM variant for C_{20}PP. Therefore, in *E. coli* cells also expressing $CrtM_{F26L}$ or $CrtM_{F26A,W38A}$, much of the C_{20}PP released by $BstFPS_{Y81A}$ is instead used to synthesize C_{40} and C_{45} carotenoids. An important lesson from these observations is that the performance of an enzyme *in vitro* may not be representative of its performance as part of an *in vivo* biosynthetic pathway.

In the course of other work aimed at increasing the total quantity of carotenoids synthesized by our cultures, we overexpressed *E. coli* isopentenyl diphosphate isomerase (Idi) along with $BstFPS_{Y81A}$ and a CrtM variant on plasmids (**Figure 3.11**). The isomerization of IPP to DMAPP is a rate-limiting step for the production of isoprenoids in *E. coli* (59). IPP is the basic isoprene unit synthesized by both the mevalonate and non-mevalonate pathways of isoprenoid biosynthesis in nature (7). However, IPP serves as an extender unit in prenyl chain elongation, and its isomer, DMAPP, which serves as the starter unit, is only synthesized by Idi-catalyzed isomerization of IPP (7). Overexpression of Idi increases carbon flux to DMAPP, which in turn increases flux through downstream isoprenoid biosynthetic pathways (59).

When we transformed *E. coli* XL1-Blue cells with the plasmids pUC18m-$bstFPS_{Y81A}$-*idi* and pAC-$crtM_{F26A,W38A}$, the resulting liquid cultures did indeed synthesize about an order of magnitude more total carotenoids compared with XL1-Blue(pUC18m-$bstFPS_{Y81A}$ + pAC-$crtM_{F26A,W38A}$) cultures (**Figure 3.11** caption). However, the distribution of carotenoid backbones was also dramatically shifted toward the production of C_{40} carotenoids (**Figure 3.11**). This result can be rationalized by the "statistical" nature of $BstFPS_{Y81A}$ and the increased level of DMAPP in cells overexpressing Idi. Five-carbon DMAPP, the smallest isoprenyl diphosphate starter unit, is in addition to $C_{15}PP$ and $C_{20}PP$, an acceptable substrate for $BstFPS_{Y81A}$ (37). With the isoprenyl diphosphate pool of cells overexpressing Idi enriched in DMAPP, $BstFPS_{Y81A}$ will admit this substrate more often. As a "statistical" iterative biocatalyst, $BstFPS_{Y81A}$ tends to discharge shorter products on average when supplied with shorter substrates, and indeed, the enzyme frequently releases $C_{20}PP$ as it catalyzes prenyl chain elongation starting

from DMAPP (37). It follows that an increased level of DMAPP in the cells will result in increased production of C_{20}PP by BstFPS$_{Y81A}$. Since the CrtM variants can condense a variety of prenyl diphosphates into different carotenoids, the larger proportion of C_{20}PP in their substrate pool results in the observed shift toward the production of C_{40} carotenoids. This example illuminates the importance of the composition of the first enzyme's substrate pool in determining the mixture of end-products of biosynthetic pathways composed of a succession of mutant enzymes with low substrate specificity.

As stated previously, C_{50} carotenoids were only desaturated to a relatively small extent *in vivo* by CrtI, and unmetabolized C_{50} carotenoid backbones comprised approximately 50 mol% of the carotenoid pool in cultures of XL1-Blue(pUCmodII-*crtM*$_{F26A,W38A}$-*crtI-bstFPS*$_{Y81A}$). If we could evolve a desaturase to more efficiently convert C_{50} backbones, it might be possible to use colorimetric screening of bacterial colonies in order to further evolve either BstFPS$_{Y81A}$ for more specific production of C_{25}PP, or to evolve a CrtM variant for more specific synthesis of either C_{45} or C_{50} carotenoid backbones. Such a screen would require that a given desaturase variant could proficiently desaturate both C_{45} and C_{50} backbones, and also that the major desaturation product of each pathway be a different color. However, before we could attempt to evolve CrtI or another desaturase for improved performance on a C_{50} substrate, we encountered an unexpected roadblock. The first indication of this issue occurred when we noticed that *E. coli* colonies expressing BstFPS$_{Y81A}$, a CrtM variant, and CrtI from one or two plasmids reproducibly displayed a dramatically deeper red color than cell pellets from liquid cultures harboring the same plasmid(s). HPLC analysis revealed that the distribution of carotenoids in our recombinant *E. coli* colonies was shifted toward C_{40}

products, with C_{50} and C_{45} carotenoids each representing less than 10% of total carotenoids (**Figure 3.11**). The similarity of this shift to that brought on by Idi overexpression in liquid cultures leads us to hypothesize that our *E. coli* cells grown in colonies on agar plates contain a higher proportion of DMAPP than cells harboring the same plasmids grown in liquid culture. As we have seen with Idi, such a difference in the relative proportions of C_5-C_{25} prenyl diphosphates between the two physiological states would be sufficient to explain the different carotenoid distributions. However, we can only speculate on the origins of this putative difference in the isoprenyl diphosphate pools of cells cultivated in liquid medium versus those grown on agar plates.

One of our early goals was to generate biosynthetic pathways to C_{60} carotenoids. When we formulated that objective, the C_{45} and C_{50} pathways had not yet been generated, and it was not known for certain which aspects of the C_{60} project would be the most problematic. Now, in retrospect, we have substantially more understanding of the experimental bottlenecks to expect. The first step of the pathway, generation of a C_{30}PP precursor, has already been discovered by nature. Unlike C_{25}PP, C_{30}PP is found in bacteria, archaea, and eukaryotes, serving as a precursor to the side chain of quinones (26, 60). Isoprenyl diphosphate synthases that synthesize C_{30}PP are therefore rather widespread in nature. The results of our work as well as that of Shimizu et al. with the C_{30}PP synthase HexPS from *M. luteus* demonstrate that it is quite efficient and specific at converting C_{15}PP to C_{30}PP (**Figure 3.12** and ref. (47)).

Despite the ability of HexPS to supply C_{30}PP in *E. coli*, enzymes capable of catalyzing C_{60} carotenoid backbone synthesis and desaturation remain undiscovered. We found that neither the C_{40} synthase CrtB nor CrtM mutants capable of C_{50} carotenoid

backbone synthesis could convert C_{30}PP into a C_{60} backbone. Furthermore, screening for a carotenoid synthase variant capable of producing C_{60} backbones will likely not be a simple procedure. Our preferred screening method for evolving carotenoid biosynthetic enzymes, colorimetric screening of colonies, is probably not possible at present in a C_{60} pathway, since it cannot be assumed that any known carotenoid desaturase will impart pigmentation on a C_{60} backbone. We believe it is unlikely that any wild-type carotenoid desaturase would detectably desaturate a C_{60} substrate, since even processing of C_{50} backbones poses a great challenge to the broad-specificity desaturase CrtI. On the other hand, if a C_{60} carotenoid synthase were to be discovered first, screening for a desaturase variant capable of accepting C_{60} backbones should be relatively straightforward. Since C_{30}PP is the only substrate available for carotenoid biosynthesis in *E. coli* expressing HexPS, C_{60} carotenoid backbones should be the only carotenoids possible in cells expressing HexPS and a C_{60} synthase. Therefore, screening *E. coli* colonies additionally transformed with a library of variant desaturases for the emergence of color should be a reliable method to search for desaturase variants capable of converting C_{60} backbones into pigmented C_{60} products.

The production of hydroxylated carotenoid backbones we have reported in this work is an example of another type of hurdle to pathway specificity that can be encountered by metabolic pathway engineers. Although assembling and coexpressing a series of foreign biosynthetic genes in a heterologous host may look simple on paper, it is not possible to predict all the unwanted side reactions that might occur in the complex environment of a living cell. As we have shown, even *E. coli*, the most studied of all organisms, will carry out unexpected chemical modification of the products of foreign

biosynthetic genes it is forced to express. This is probably especially common when such products are harmful to the cell in some way and elicit poorly understood "emergency" responses. Therefore, given the current state of the art in the heterologous production of small molecules in microorganisms, the old adage, "expect the unexpected" is definitely sage advice.

CONCLUSIONS

The complexity and variety of small molecules in nature is a testament to the power of evolution to innovate and diversify. Our experiments on carotenoid biosynthetic enzymes and pathways represent an attempt to simultaneously harness and study these remarkable features of evolution. By capturing entire new pathways in the very earliest stages of their emergence, we have learned that new enzyme specificities are discovered rather easily upon mutation, and that new chemical diversity generated early in a pathway can be multiplied by downstream enzymes via many different chemical mechanisms. It is likely that our nascent pathways resemble their counterparts in nature in their rapid gain of access to a variety of new products upon very limited genetic change. However, the evolutionary "pruning" process by which the unwanted majority of a nascent pathway's products are discarded in favor of a desired few remains unreproduced in the laboratory and, consequently, less well understood. Therefore, the generation of new pathways in the laboratory by directed evolution of biosynthetic enzymes, especially those located early in a pathway, remains primarily a tool for molecular discovery. It is hoped that future directed evolution experiments will shed light on how biosynthetic pathways shift their focus toward the synthesis of advantageous new products. If such research proves

fruitful, the exciting prospect of a commercial-scale whole-cell process for the specific production of a useful new molecule originally discovered in the laboratory by directed evolution may become a reality.

MATERIALS AND METHODS

Genes and plasmids

Plasmid maps and sequences can be found in Appendix A. Genes are listed in transcriptional order in the names of all the plasmids in this report. Plasmids based on the high-copy pUCmodII vector (46, 54) are designated by names beginning with "pUCmodII." The carotenoid biosynthetic genes on these plasmids are expressed as an operon under the control of a single *lac* promoter with no *lac* operator. In plasmids pUCmodII-$crtM_{F26L}$-$bstFPS_{Y81A}$ and pUCmodII-$crtM_{F26A,W38A}$-$bstFPS_{Y81A}$, a variant of the C_{30} carotenoid synthase gene *crtM* from *Staphylococcus aureus* ($crtM_{F26L}$ or $crtM_{F26A,W38A}$, respectively (54-56)) is flanked by *Xba*I and *Xho*I restriction sites. $bstFPS_{Y81A}$, which encodes the Y81A mutant of the farnesyl diphosphate synthase from *Bacillus stearothermophilus* (37, 54), is flanked by *Eco*RI and *Nco*I restriction sites and directly follows the *crtM* variant. Plasmids pUCmodII-$crtM_{F26L}$-*crtI*-$bstFPS_{Y81A}$ and pUCmodII-$crtM_{F26A,W38A}$-*crtI*-$bstFPS_{Y81A}$ additionally contain the gene encoding the carotenoid desaturase CrtI from *Erwinia uredovora* (current approved name: *Pantoea ananatis*), which is inserted between the *crtM* and *bstFPS* variants and is flanked by *Xho*I and *Eco*RI restriction sites. Plasmid pUCmodII-*crtI*, used to express *E. uredovora* CrtI alone for *in vitro* desaturation experiments, contains *crtI* flanked by *Eco*RI and *Nco*I restriction sites.

The genes on plasmids based on the high-copy vector pUC18m (56) are expressed as an operon under the control of a *lac* promoter and operator. Plasmid pUC18m-*bstFPS*$_{Y81A}$ contains the *bstFPS*$_{Y81A}$ gene flanked by *Xba*I and *Xho*I restriction sites. Plasmid pUC18m-*bstFPS*$_{Y81A}$-*idi* additionally contains the *idi* gene encoding *E. coli* isopentenyl diphosphate isomerase, which is inserted between *Xho*I and *Apa*I restriction sites. Plasmid pUC18m-*crtM*$_{F26A,W38A}$ contains *crtM*$_{F26A,W38A}$ flanked by *Xba*I and *Xho*I sites. We cloned *apFGS*, the FGPP (C_{25}PP) synthase gene from *Aeropyrum pernix* (50) from *A. pernix* genomic DNA (ATCC no. 700893D) by PCR. pUC18m-*apFGS* contains the *apFGS* gene flanked by *Nde*I and *Eco*RI sites. Plasmid pUC18m-*apFGS*-*crtM*$_{F26A,W38A}$ additionally contains *crtM*$_{F26A,W38A}$ flanked by *Xba*I and *Xho*I sites. Plasmid pUC18m-*hexPS* contains the operon from *Micrococcus luteus* strain B-P 26 encoding the two polypeptide components of hexaprenyl diphosphate (C_{30}PP) synthase (47). This operon, referred to as *"hexPS,"* is flanked by *Eco*RI and *Xba*I restriction sites. pUC18m-*hexPS-crtM-crtN* additionally includes wild-type *crtM* flanked by *Xba*I and *Xho*I sites and *crtN*, the *S. aureus* C_{30} carotenoid desaturase gene, between *Xho*I and *Apa*I sites. In plasmid pUC18m-*crtM-crtN*, *crtM* and *crtN* are flanked by the same sites as in pUC18m-*hexPS-crtM-crtN*. Plasmid pUC18m-*hexPS-crtB-crtI* has the same construction as pUC18m-*hexPS-crtM-crtN* but contains *crtB*, the C_{40} carotenoid synthase gene from *E. uredovora*, in place of *crtM* and *crtI* from the same bacterium in place of *crtN*. The plasmids pUC18m-*hexPS-crtM*$_{F26L}$ and pUC18m-*hexPS-crtM*$_{F26A,W38A}$ have the same construction as pUC18m-*hexPS-crtM-crtN* but lack *crtN*. pUC18m-*hexPS-crtB* is the same as pUC18m-*hexPS-crtB-crtI* but lacks *crtI*, while plasmid pUC18m-*crtE-crtB-crtI*

has the same construction as pUC18m-*hexPS-crtB-crtI* but has *crtE*, the gene encoding the GGPP (C_{20}PP) synthase from *E. uredovora*, in place of *hexPS*.

Plasmids based on the medium-copy vector pACmod (pACYC184 with the *Xba*I site removed) (46) were constructed by insertion of a fragment containing the entire operon from a pUC18m-based plasmid (including the *lac* promoter and operator) into pACmod. In plasmid pAC-$crtM_{F26A,W38A}$, $crtM_{F26A,W38A}$ is inserted between *Xba*I and *Xho*I restriction sites. In pAC-$crtM_{F26A,W38A}$-*crtI*, *crtI* from *E. uredovora* is also present and is flanked by *Xho*I and *Apa*I sites. Plasmid pAC-*crtE* contains *crtE* flanked by *Eco*RI and *Xba*I sites.

In all plasmids listed above, an optimized Shine-Dalgarno ribosomal binding sequence (AGGAGG) followed by eight spacer nucleotides is situated directly upstream of each gene's start codon. Even for the plasmids with *lac* operator sequences, leaky transcription was sufficient to effect the expression of carotenogenic genes. Therefore, all experiments described in this report were performed without IPTG (isopropyl-β-D-thiogalactopyranoside) induction.

Bacterial cultures and carotenoid extraction

E. coli XL1-Blue cells (Stratagene, La Jolla, CA) were transformed with plasmid DNA and plated on Luria-Bertani (LB)-agar plates supplemented with 50 mg/l of each appropriate antibiotic (carbenicillin for pUC-based plasmids, chloramphenicol for pAC-based plasmids). Colonies were usually visible after 12 h at 37 °C, at which point 2-ml precultures of LB+antibiotics (50 mg/l each) were inoculated with single colonies and shaken at 37 °C and 250 rpm for 12-18 h. One milliliter of preculture was inoculated into 1 liter of Terrific broth (TB)+antibiotics (50 mg/l each) in a 2.8-liter Erlenmeyer flask.

One-liter cultures were shaken at 28 °C and 250 rpm for 48 h. Optical densities at 600 nm of 1-liter cultures (OD_{600nm}) were measured after 10-fold dilution into fresh TB medium. Dry cell masses were determined from OD_{600nm} values using a calibration curve generated for XL1-Blue cultures.

Cells were pelleted by centrifugation at 2500 × *g* for 15 min. For accurate quantification of total carotenoid titers, a known amount of β,β-carotene (Fluka/Sigma, St. Louis, MO) was added to pellets as an internal standard before extraction. Cell lysis and initial carotenoid extraction was achieved by addition of 100 ml of acetone supplemented with 30 mg/l of 3,5-di-*tert*-butyl-4-hydroxytoluene (BHT; Sigma) followed by vigorous shaking. After this point, to avoid difficulties with mass spectrometry (MS), contact with plasticware was avoided and only glassware was used. After filtration and concentration to ~5 ml under a stream of N_2, lipids including carotenoids were partitioned twice into 8 ml of hexanes and then concentrated to ~4 ml under a stream of N_2. This extract was then washed 5 times with 2 ml of deionized water, dewatered with anhydrous $MgSO_4$, filtered, and then evaporated to dryness under a stream of N_2. After resuspension in 2 ml of acetone + BHT (30 mg/l), a significant amount of fatty material usually precipitated on the sides of the vial. The mixture was sonicated at 40 °C for 10-20 min. to dissolve carotenoids trapped by the precipitate and was then filtered and evaporated to dryness under a stream of N_2. Dried extracts were stored under argon gas at -20 °C. Extracts were dissolved into 0.1-1 ml of acetonitrile prior to separation and analysis.

For analysis of carotenoids synthesized by *E. coli* colonies, LB-agar plates were removed from 37 °C after colonies had formed (usually 12 h). The plates were incubated

at 25 °C for 48 h to promote carotenoid biosynthesis. Colonies were lifted onto white nitrocellulose membranes (Pall, Port Washington, NY) for imaging on a flatbed scanner. For extraction of carotenoids from colonies, TB medium was used to suspend the colonies, after which the mixture was centrifuged at 2500 × *g* for 15 min. After decanting the liquid, dry cell masses were determined from measured wet cell masses using a calibration curve generated for XL1-Blue cells. The same steps listed above for carotenoid extraction from liquid culture pellets were followed but all volumes were scaled by a factor of 0.2 due to the smaller cell pellets obtained from scraping colonies compared with pellets from liquid cultures. We generally found that ~0.5 g of wet cells was sufficient to obtain enough extract for HPLC analysis.

Separation and analysis of carotenoids

The above extracts dissolved in acetonitrile were injected into an Xterra MS C_{18} column (3.0 × 150 mm, 3.5-μm particle size; Waters, Milford, MA) outfitted with a guard cartridge (3.0 × 20 mm, identical particles) using an Alliance 2690 HPLC system (Waters) equipped with a photodiode array detector (PDA) set to 1.2-nm resolution. For isocratic elution, the mobile phase was acetonitrile:isopropanol 93:7 (by volume) and the flowrate was 0.4 ml/min. For more rapid elution of highly nonpolar C_{45} and C_{50} carotenoids, a gradient method with the same flowrate was employed: 0-35 min, acetonitrile:isopropanol 93:7; 35-37 min, linear gradient to acetonitrile:isopropanol 50:50; 37-50 min, acetonitrile:isopropanol 50:50. Carotenoids injected to HPLC were quantified by comparing their chromatogram peak areas (determined at the wavelength of maximum absorption for each carotenoid in the sample) to a calibration curve generated using known amounts of β,β-carotene and then multiplying by the ratio of molar

extinction coefficients ($\varepsilon_{\beta,\beta\text{-carotene}}/\varepsilon_{\text{sample carotenoid}}$) (9). The calibration curve and internal standard peak areas enabled the calculation of scaling factors, which permitted the computation of total carotenoid titers from the HPLC data.

Fractions of the HPLC eluent containing separated carotenoids were collected and evaporated to dryness under N_2. Mass spectra of carotenoids were obtained using either an 1100 Series HPLC-PDA-MS system (Hewlett Packard/Agilent, Palo Alto, CA) equipped with an atmospheric pressure chemical ionization (APCI) interface or a Finnigan LCQ mass spectrometer equipped with an electrospray ionization (ESI) source (Thermo Electron, San Jose, CA). Previously separated carotenoids were resuspended in either hexanes (for APCI-MS) or methanol:dichloromethane 85:15 (by volume) (for ESI-MS) and flow-injected into the mobile phase, which was the same as the solvent used to dissolve the sample.

Carotenoids were identified and the putative structures of novel carotenoids were assigned based on their HPLC retention times, UV-visible spectra, and mass spectra.

Acetylation of hydroxylated carotenoid backbones

To confirm the presence of a hydroxy group on carotenoid backbones, the acetylation protocol suggested by Eugster (19) was followed with slight modification. Briefly, 5-10 nmol of a hydroxylated carotenoid backbone was dissolved in 0.5 ml of pyridine (dried over BaO powder). Fifty microliters of acetic anhydride was then added, and the reactions were carried out at room temperature for 1 h. The reactions were terminated by addition of 4 ml of deionized water, and carotenoids were extracted by partitioning twice with 1 ml of hexanes. After evaporation of the solvent under a stream of N_2, the product mixture was analyzed and separated by HPLC as described above. The

separated product and unreacted fractions were dried under N_2, resuspended in hexanes, and analyzed by APCI-MS as described above.

In vitro desaturation of carotenoid backbones

Cell lysate of *E. coli* cultures expressing CrtI was prepared for *in vitro* desaturation reactions as follows. 150-ml cultures of XL1-Blue(pUCmodII-*crtI*) were shaken in TB medium supplemented with carbenicillin (50 mg/l) at 28 °C and 113 rpm in upright 175 cm^2 tissue culture flasks (BD Falcon, Bedford, MA) until they reached an OD_{600nm} of 5-6 (~48 h). Cells were centrifuged at 2500 × *g* for 15 min., and each pellet from a 150-ml culture was resuspended in 20 ml of Tris-HCl buffer (50 mM, pH 8) containing 1 mM phenylmethylsulfonylfluoride (PMSF). Cells were lysed by French press, and the lysate was stored at -20 °C in 1-ml aliquots until use.

Carotenoid backbone substrates for *in vitro* desaturation reactions were separated and purified as described above from cultures of XL1-Blue(pUCmodII-$crtM_{F26L}$-$bstFPS_{Y81A}$), XL1-Blue(pUCmodII-$crtM_{F26A,W38A}$-$bstFPS_{Y81A}$), or XL1-Blue(pUC18m-*crtE-crtB*) (synthesizing authentic phytoene only (56)). After resuspension in a known amount of hexanes, carotenoid backbones were quantified by their absorbance at 286 nm using a Cary 100 Bio UV-visible spectrophotometer (Varian, Palo Alto, CA). For each *in vitro* desaturation reaction, 10 nmol of a carotenoid backbone was resuspended in 10 µl of acetone and added to a thawed 1-ml aliquot of XL1-Blue(pUCmodII-*crtI*) lysate supplemented with additional flavin adenine dinucleotide (FAD, 1mM) and $MgCl_2$ (4 mM). Reaction mixtures were then incubated at 30 °C with gentle end-over-end rotation for 12-18 h.

To extract carotenoids from *in vitro* reactions for HPLC analysis, the reaction mixtures were centrifuged at 14,000 × *g* for 5 min. and the pellets were separated from the supernatants. Pellets were extracted with 1 ml of acetone + BHT (30 mg/l), thoroughly vortexed, and then filtered. Carotenoids were extracted from the aqueous supernatants by partitioning twice with 1 ml of hexanes. The hexane and acetone extracts of each reaction mixture were pooled and evaporated to dryness under a stream of N_2. Carotenoids were then resuspended in 150 µl of acetonitrile and analyzed by HPLC as described above.

Iodine-catalyzed photoisomerization

To enrich mixtures of *E*- and *Z*-carotenoids in all-*E* (*trans*) isomers or verify the all-*E* configuration of specific carotenoids, individual carotenoids or culture extracts were subjected to iodine-catalyzed photoisomerization following the protocol suggested by Schiedt and Liaaen-Jensen (44). After photoisomerization, samples were analyzed by HPLC and changes in *E/Z* isomeric configuration were identified by the bathochromic (or hypsochromic) shifts in their absorption spectra and the intensity change of *cis*-peaks (9) compared with control samples not subjected to photoisomerization.

Thin-layer chromatography and autoradiography analysis of isoprenyl diphosphate synthase reactions

In vitro investigations of isoprenyl diphosphate synthase reactions were performed by Adam Hartwick as follows. *E.coli* XL1-Blue cells transformed with pUC18m-*crtE*, pUC18m-*bstFPS*$_{Y81A}$, or pUC18m-*hexPS* were cultured for 24 h in 100 ml of TB medium supplemented with carbenicillin (50 mg/l) at 37 °C in 500-ml Erlenmeyer

flasks shaken at 250 rpm for 24 h. Each culture was divided into 4 equal aliquots, which were centrifuged at 4 °C for 10 min. at 3000 × *g*. The cell pellets were stored at -80 °C. Pellets were thawed on ice and then resuspended in 5 ml of Tris-HCl buffer (25 mM, pH 8.5) supplemented with 1 mM EDTA and 10 mM 2-mercaptoethanol. The resuspended cells were lysed by microtip sonication and then centrifuged at 4500 × *g* for 20 min. The supernatants were stored in 250-µl aliquots at -80 °C until use.

Enzyme reactions were carried out by mixing the following: 100 µl of reaction buffer (50 mM HEPES buffer, pH 8.2 supplemented with 40 mM $MgCl_2$, 40 mM $MnCl_2$, and 200 mM 2-mercaptoethanol), 100 µl of FPP ($C_{15}PP$) buffer (100 µM FPP in 25 mM HEPES buffer, pH 8.2), 100 µl of ^{14}C-IPP buffer (12 µM ^{14}C-labeled IPP in 25 mM HEPES buffer, pH 8.2), and 100 µl of one of the above lysates. The reaction mixtures were shaken at 37 °C and 1000 rpm for 4 h.

Radiolabeled isoprenyl diphosphate products were extracted from the reactions by addition of 200 µl of 1-butanol followed by vortexing and centrifugation at 14000 × *g* for 30 sec. The butanol epiphase was withdrawn and used directly in the subsequent dephosphorylation reaction by potato acid phosphatase (Sigma). The butanol extracts were combined with 400 µl of methanol, 100 µl of acetate buffer (1 M, pH 5.6), 100 µl of 1% (w/v) Triton X-100, and 200 µl of acid phosphatase buffer (44 units enzyme/ml in acetate buffer). The dephosphorylation reactions were then shaken at 37 °C and 1000 rpm for 24 h.

Following dephosphorylation, ^{14}C-labeled isoprenyl alcohols were extracted with 0.1 ml of n-pentane. After vortexing and centrifugation, the organic layer was removed by pipette. A portion of each pentane extract was spotted onto a C_{18} reverse-phase TLC

plate, which was developed with a mobile phase of acetone:water 9:1 (by volume). TLC plates were then exposed to a phosphor plate for ~24 h. Autoradiogram images were obtained by scanning the phosphor plates with a Typhoon Imager (Amersham).

Table 3.1. Properties of novel desaturated carotenoids synthesized by XL1-Blue(pUCmodII-*crtM*$_{F26A,W38A}$-*crtI-bstFPS*$_{Y81A}$)

Backbone	Conjugated double bonds	Retention time[a] (min)	Molecular ion (*m/z*)[b]	Mol%[c]	UV-visible spectrum (Fig. 3.6) and putative structure (Figs. 3.2-3.4)
C_{40}	7	19.2	540.2	5	**1**
C_{45}	7	30.0	608.4	0.5	**3**
C_{45}	9	28.5	606.2	1	**4**
C_{45}	13	15.6, 16.4	602.0	4	**5**
C_{45}	15	13.9	600.2	2	**6**
C_{50}	7	42.0	676.2	2	**10**
C_{50}	9	39.6, 40.4	674.2	5	**11a**
C_{50}	9	38.0	674.2	0.5	**11b**
C_{50}	11	35.9	672.3	3	**12a**
C_{50}	11	34.8	672.3	2	**12b/c**
C_{50}	15	20.3, 21.4	668.1	5	**13**

a. Using the column and gradient method described in Materials and Methods. Multiple times refer to *E/Z* isomers of the same carotenoid.

b. Determined by electrospray mass spectrometry. Carotenoids underwent direct electrochemical oxidation to give the radical molecular cation $[M\cdot]^+$.

c. Approximate, of all carotenoids detected in the culture extract by HPLC. Unmetabolized 16,16'-diisopentenylphytoene (C_{50} backbone) represented ~50 mol% of this population.

Table 3.2. Names of numbered carotenoid structures depicted in Figures 3.2-3.4

Structure number	Trivial name[a]	IUPAC-IUB semi-systematic name
1	16-isopentenyl-7,8,11,12-tetrahydro-4'-apolycopene	16-(3-methylbut-2-enyl)-7,8,11,12-tetrahydro-4'-apo-ψ,ψ-carotene
2a	C_{45}-phytofluene (9-13')	16-(3-methylbut-2-enyl)-7,8,7',8',11',12'-hexahydro-ψ,ψ-carotene
2b	C_{45}-phytofluene (13-9')	16-(3-methylbut-2-enyl)-7,8,11,12,7',8'-hexahydro-ψ,ψ-carotene
3	C_{45}-ζ-carotene	16-(3-methylbut-2-enyl)-7,8,7',8'-tetrahydro-ψ,ψ-carotene
4a	C_{45}-neurosporene (5-9')	16-(3-methylbut-2-enyl)-7',8'-dihydro-ψ,ψ-carotene
4b	C_{45}-neurosporene (9-5')	16-(3-methylbut-2-enyl)-7,8-dihydro-ψ,ψ-carotene
5a	C_{45}-didehydrolycopene (1-5')	16-(3-methylbut-2-enyl)-3,4-didehydro-ψ,ψ-carotene
5b	C_{45}-didehydrolycopene (5-1')	16-(3-methylbut-2-enyl)-3',4'-didehydro-ψ,ψ-carotene
6	C_{45}-tetradehydrolycopene (1-1')	16-(3-methylbut-2-enyl)-3,4,3',4'-tetradehydro-ψ,ψ-carotene
10	C_{50}-ζ-carotene	16,16'-di-(3-methylbut-2-enyl)-7,8,7',8'-tetrahydro-ψ,ψ-carotene
11a	C_{50}-neurosporene (5-9')	16,16'-di-(3-methylbut-2-enyl)-7,8-dihydro-ψ,ψ-carotene
11b	C_{50}-neurosporene (1-13')	16,16'-di-(3-methylbut-2-enyl)-3,4-didehydro-7',8',11',12'-tetrahydro-ψ,ψ-carotene
12a	C_{50}-lycopene (5-5')	16,16'-di-(3-methylbut-2-enyl)-ψ,ψ-carotene
12b	C_{50}-lycopene (9-1')	16,16'-di-(3-methylbut-2-enyl)-3,4-didehydro-7',8'-dihydro-ψ,ψ-carotene
12c	C_{50}-lycopene (13-3")	16-(3-methylbut-2-enylidene),16'-(3-methylbut-2-enyl)-3,4-didehydro-7',8',11',12'-dihydro-ψ,ψ-carotene
13	C_{50}-tetradehydrolycopene	16,16'-di-(3-methylbut-2-enyl)-3,4,3',4'-tetradehydro-ψ,ψ-carotene

a. Numbers in brackets refer to the carbon atoms that mark the ends of the conjugated double-bond system. The double prime in the trivial name for carotenoid **12c** refers to the numbering of the 3-methylbut-2-enylidene group on the right side of the molecule as shown in Figure 3.3.

Table 3.3. Results of *in vitro* desaturation experiments with XL1-Blue(pUCmodII-*crtI*) lysate on various carotenoid backbones[a]

Backbone	Conversion (%)	Desaturation products (mol% of product mixture)			
		1-step	2-step	3-step	4-step
C_{30}	1	28	41	9	22
phytoene[b]	19	14	13[c]	7	66
C_{40}-OH[d]	30	11	12	19	58
C_{45}	5	71 [**2**][e]	29 [**3**]	–	–
C_{45}-OH [**7**]	14	48 [**8**]	52 [**9**]	–	–
C_{50}	0	–	–	–	–
C_{50}-OH [**14**]	0	–	–	–	–

Numbers in square brackets refer to products shown in Figures 3.2 and 3.5.

a. Values represent the average of at least two independent experiments rounded to the closest percentage point. Coefficients of variation were less than 0.25.

b. The native substrate of CrtI and positive control. Phytoene was purified from cultures of XL1-Blue(pUC18m-*crtE-crtB*) (56), which produces authentic phytoene only.

c. Identified by its absorbance spectrum as ζ-carotene.

d. Most likely predominantly OH-phytoene but may have contained some hydroxylated 16-isopentenyl-4'-apophytoene (asymmetric C_{40} backbone).

e. This product was detected only from *in vitro* desaturation of 16-isopentenylphytoene (C_{45} backbone) and was not found *in vivo*.

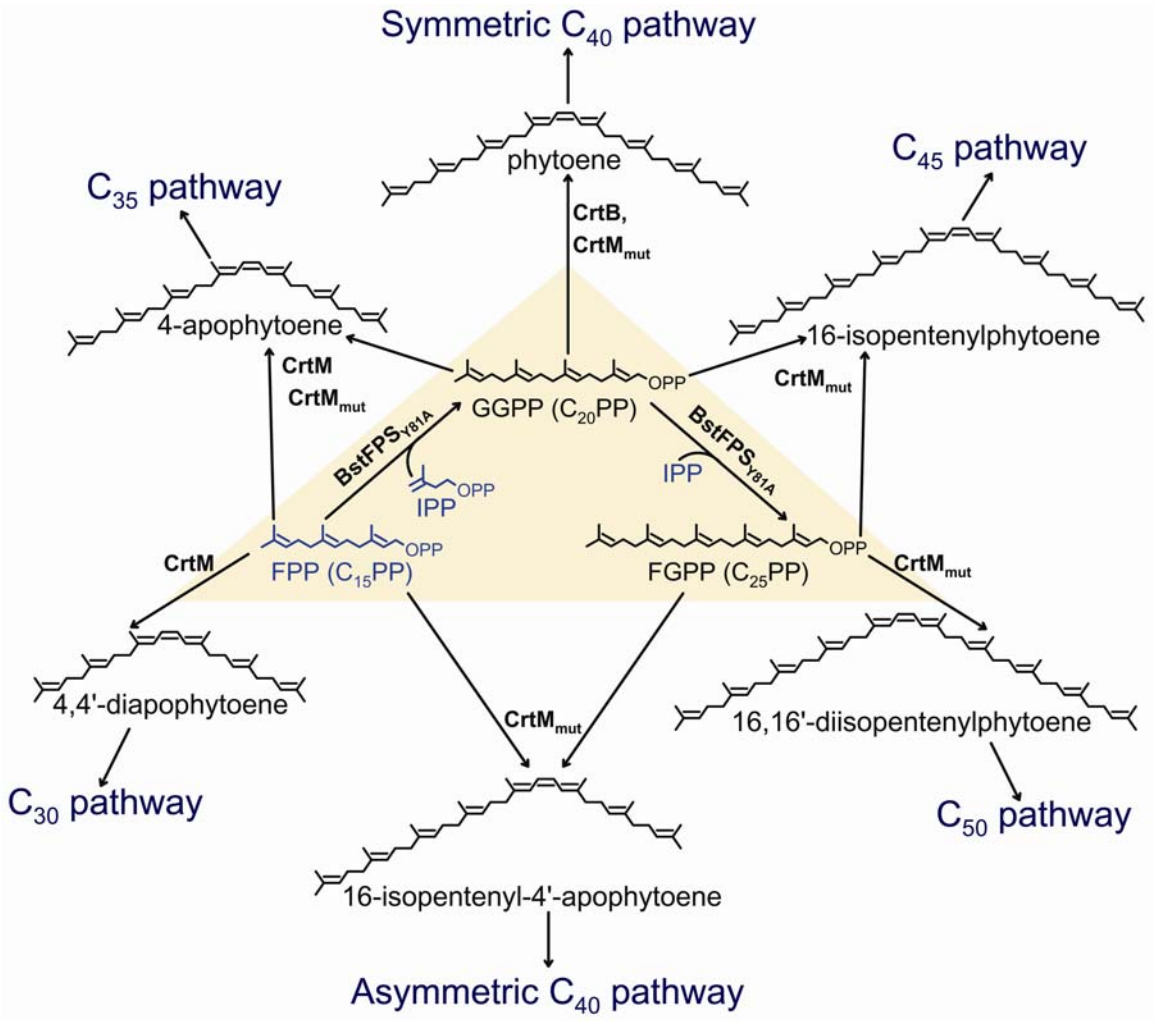


Figure 3.1. Six carotenoid biosynthetic pathways from three prenyl diphosphate precursors. Only the C_{30} and symmetric C_{40} pathways synthesized from C_{15}PP and C_{20}PP, respectively, have been identified in nature. *E. coli* cells expressing $BstFPS_{Y81A}$ synthesize the rare C_{25}PP precursor FGPP, which can be incorporated into three unnatural carotenoid backbones—C_{45}, C_{50} (54), and asymmetric C_{40}—upon coexpression of one of a number of CrtM mutants with altered specificity, $CrtM_{mut}$. Wild-type and variants of CrtM can also synthesize C_{35} carotenoids by fusion of C_{15}PP and C_{20}PP as reported previously (53, 56). Prenyl diphosphate precursors endogenous to *E. coli* are depicted in blue. Carotenoid backbones are depicted as the 15*Z* isomer synthesized by these bacterial carotenoid synthases. CrtB, phytoene synthase; IPP, isopentenyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; FGPP, farnesylgeranyl diphosphate.

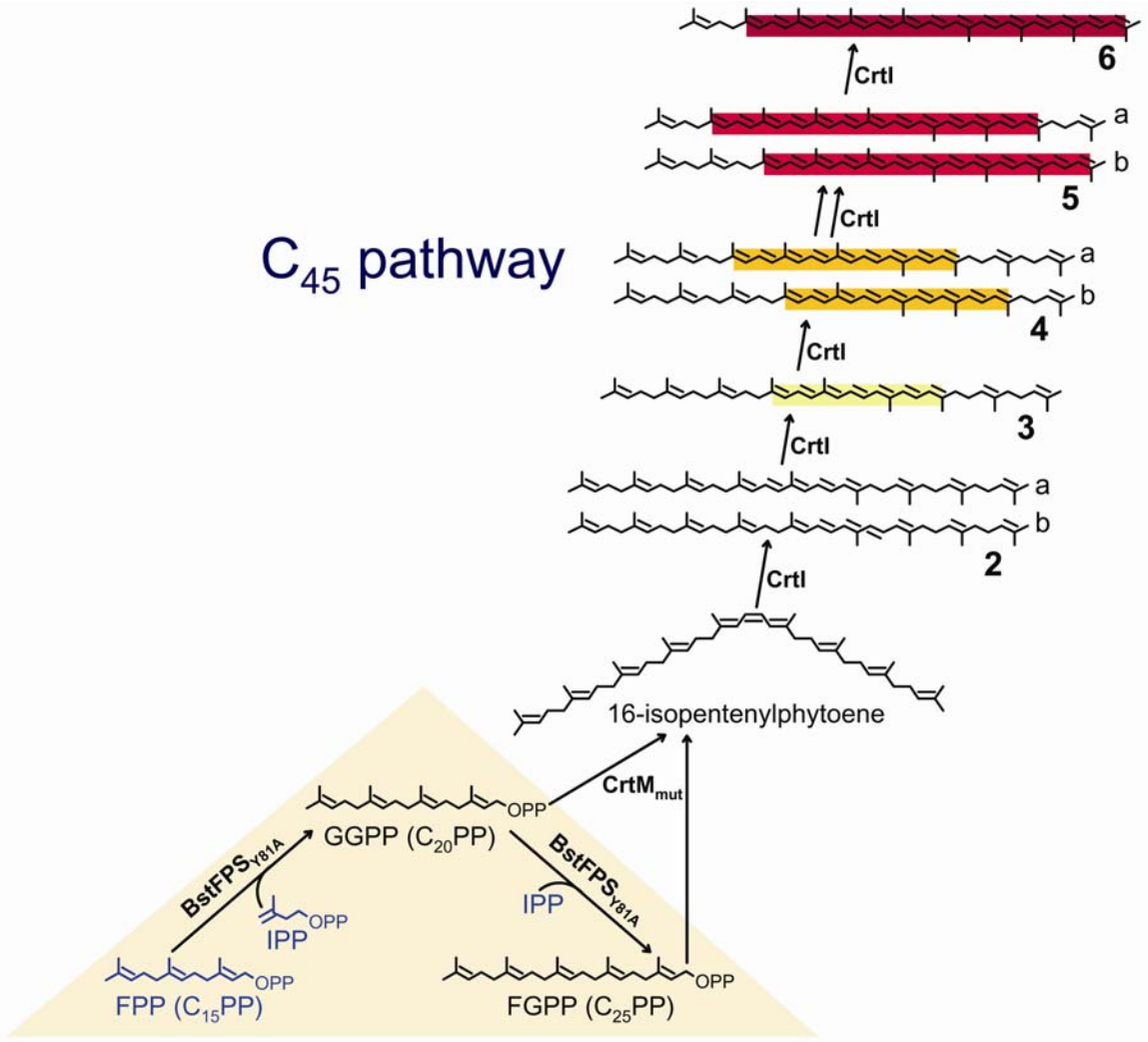


Figure 3.2. Desaturation in the C_{45} pathway. Simultaneous coexpression of $BstFPS_{Y81A}$, $CrtM_{F26L}$ or $CrtM_{F26A,W38A}$ ($CrtM_{mut}$), and the carotenoid desaturase CrtI from *E. uredovora* resulted in the biosynthesis of desaturated C_{45} carotenoids with putative structures **3**, **4**, **5**, and **6**, reported for the first time in this work. Carotenoid **2** was only detected from *in vitro* desaturation of 16-isopentenylphytoene. Colored boxes highlighting carotenoid chromophores depict the approximate color of the molecule in white light. Double arrows signify two desaturation steps. Structure numbers (and letters for different desaturation isomers) are used throughout the text and other figures. Names for the numbered structures are listed in Table 3.2. Prenyl diphosphate precursors endogenous to *E. coli* are depicted in blue. IPP, isopentenyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; FGPP, farnesylgeranyl diphosphate.

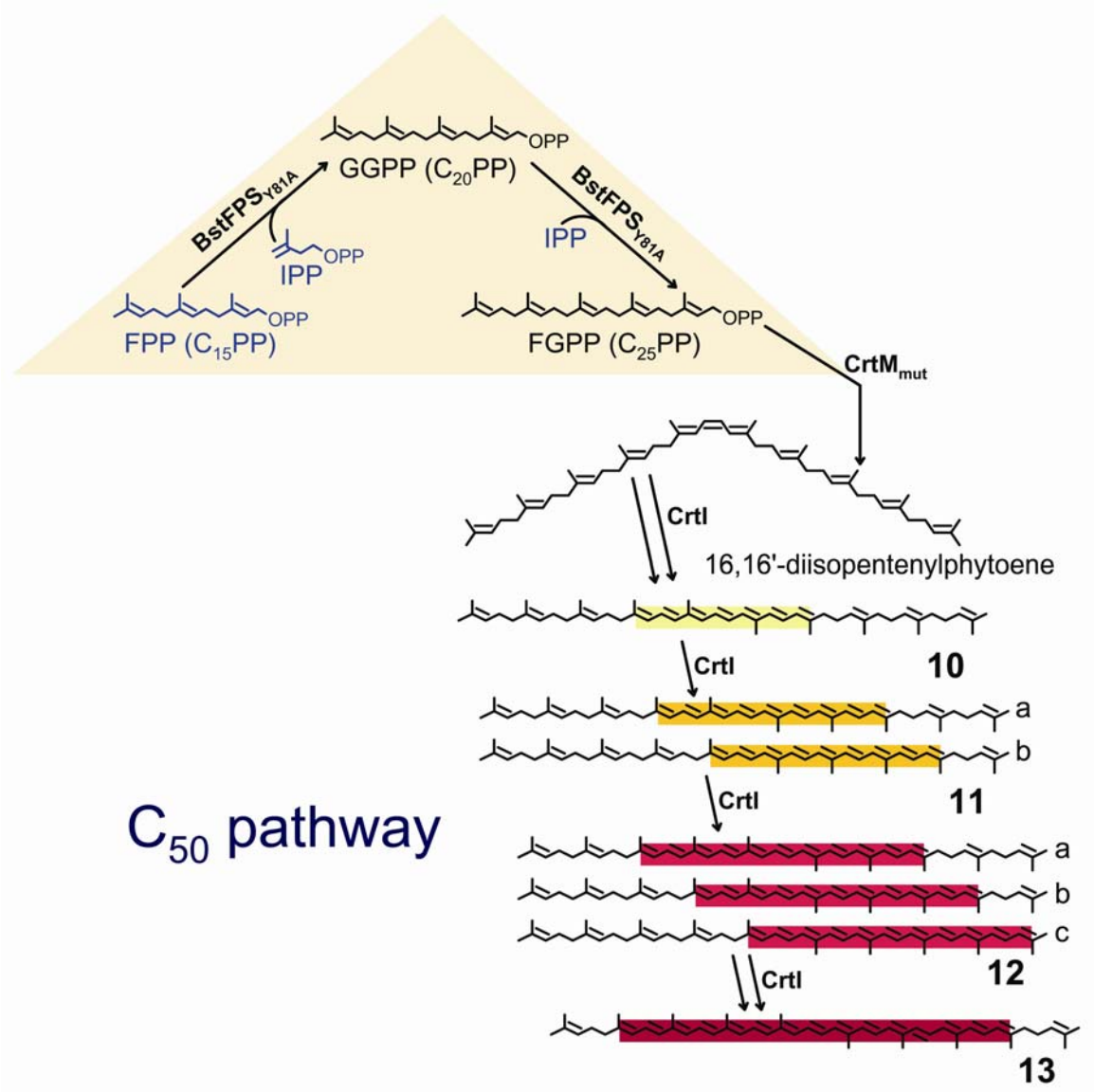

OPP
GGPP (C_{20}PP)
BstFPS$_{Y81A}$
IPP
OPP
IPP
BstFPS$_{Y81A}$
OPP
FPP (C_{15}PP)
OPP
FGPP (C_{25}PP)
CrtM$_{mut}$
CrtI
16,16'-diisopentenylphytoene
CrtI
10
a
b
CrtI
11
a
b
c
CrtI
12
13
C_{50} pathway

Figure 3.3

Figure 3.3. Desaturation in the C_{50} pathway. Simultaneous coexpression of $BstFPS_{Y81A}$, $CrtM_{F26L}$ or $CrtM_{F26A,W38A}$ ($CrtM_{mut}$), and the carotenoid desaturase CrtI from *E. uredovora* resulted in the biosynthesis of desaturated C_{50} carotenoids with putative structures **10**, **11**, **12**, and **13**, reported for the first time in this work. Colored boxes highlighting carotenoid chromophores depict the approximate color of the molecule in white light. Double arrows signify two desaturation steps. Structure numbers (and letters for different desaturation isomers) are used throughout the text and other figures. Names for the numbered structures are listed in Table 3.2. Prenyl diphosphate precursors endogenous to *E. coli* are depicted in blue. IPP, isopentenyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; FGPP, farnesylgeranyl diphosphate.

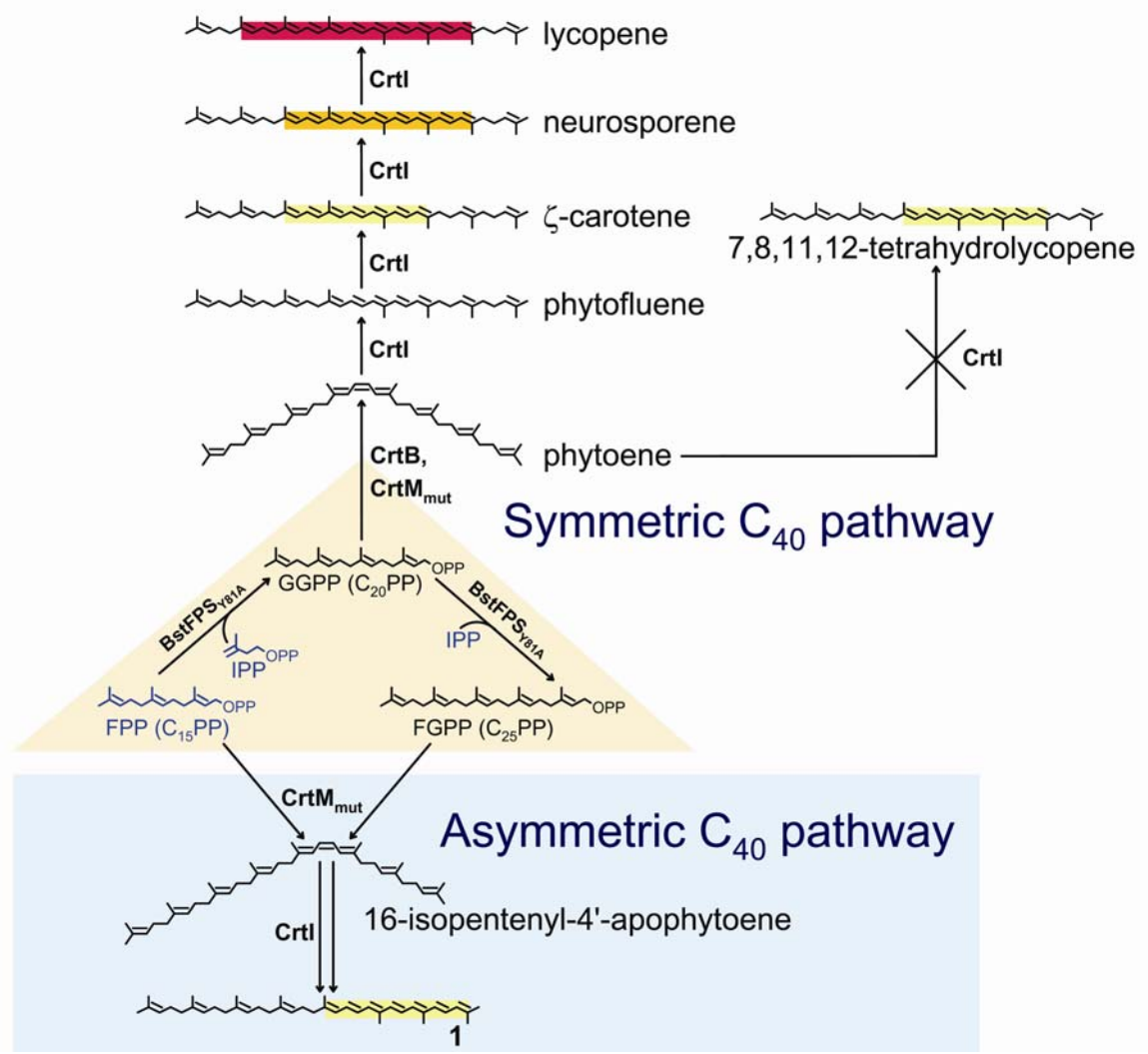

lycopene
CrtI
neurosporene
CrtI
ζ-carotene
CrtI
phytofluene
CrtI
phytoene
7,8,11,12-tetrahydrolycopene
CrtI
CrtB,
CrtM$_{mut}$
Symmetric C$_{40}$ pathway
GGPP (C$_{20}$PP)
OPP
BstFPS$_{Y81A}$
IPP
OPP
FPP (C$_{15}$PP)
OPP
IPP
BstFPS$_{Y81A}$
FGPP (C$_{25}$PP)
OPP
CrtM$_{mut}$
Asymmetric C$_{40}$ pathway
CrtI
16-isopentenyl-4'-apophytoene
1

Figure 3.4

Figure 3.4. An asymmetric C_{40} pathway. Simultaneous coexpression of $BstFPS_{Y81A}$, $CrtM_{F26L}$ or $CrtM_{F26A,W38A}$ ($CrtM_{mut}$), and the carotenoid desaturase CrtI from *E. uredovora* resulted in the biosynthesis of 2-step carotenoid **1**, putatively the first carotenoid from the novel asymmetric C_{40} carotenoid biosynthetic pathway (blue shading) to be isolated. Several lines of evidence support the precise structure of **1** (named in Table 3.2) shown in the figure, including its biosynthesis via 2+0 desaturation of an asymmetric C_{40} carotenoid backbone formed by condensation of $C_{15}PP$ and $C_{25}PP$ (see Discussion). Symmetric C_{40} pathway carotenoids phytoene, phytofluene, ζ-carotene, neurosporene, and lycopene are shown for comparison; lycopene was also detected in the same cultures as **1**. The structure of 7,8,11,12-tetrahydrolycopene, with the same 2+0 desaturation pattern as **1**, is also shown for comparison. This carotenoid is not known to be made by *E. uredovora* CrtI (crossed-out arrow). Colored boxes highlighting carotenoid chromophores depict the approximate color of the molecule in white light. Prenyl diphosphate precursors endogenous to *E. coli* are depicted in blue. CrtB, phytoene synthase; IPP, isopentenyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; FGPP, farnesylgeranyl diphosphate.

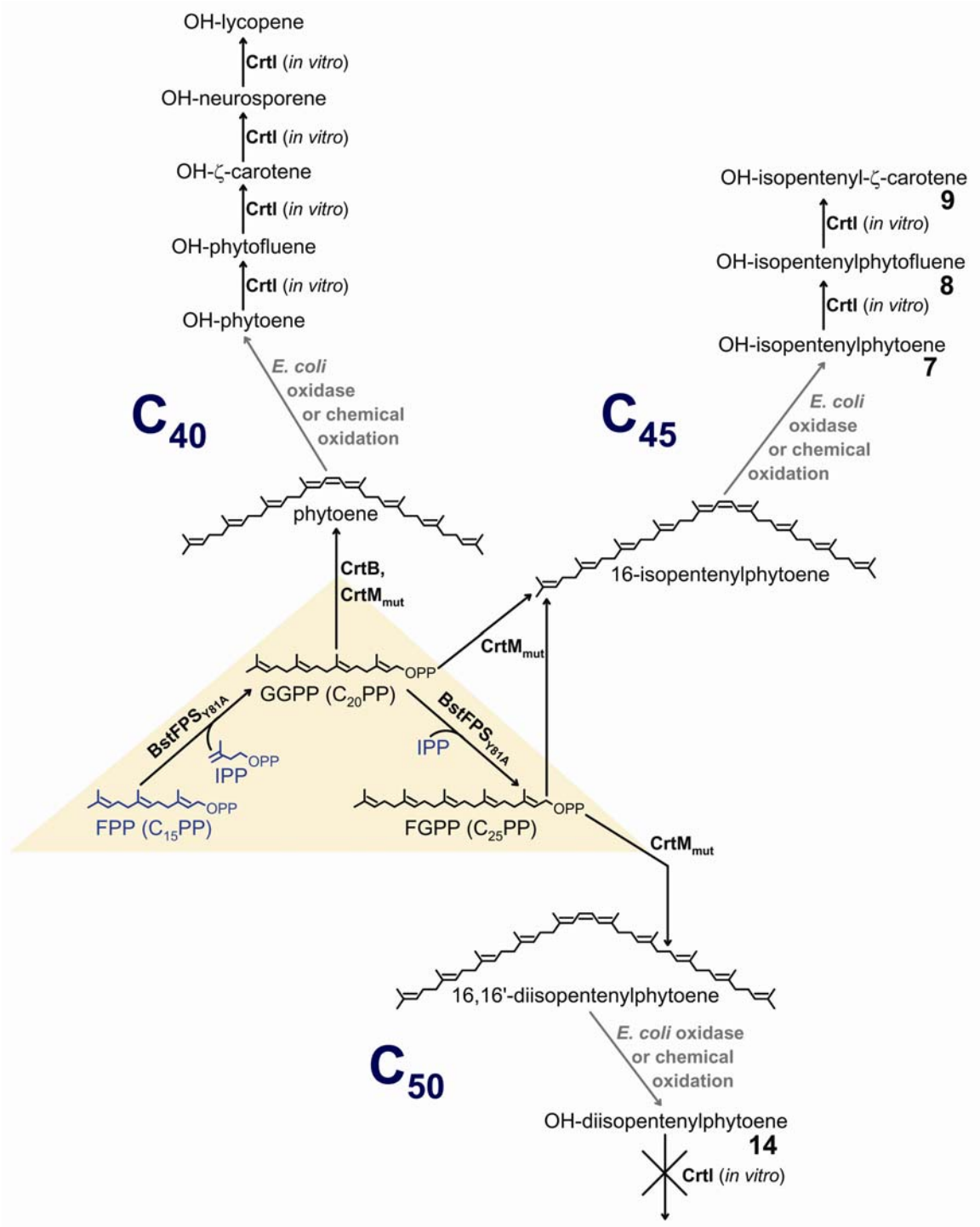

OH-lycopene
CrtI (in vitro)
OH-neurosporene
CrtI (in vitro)
OH-ζ-carotene
CrtI (in vitro)
OH-phytofluene
CrtI (in vitro)
OH-phytoene
E. coli oxidase or chemical oxidation
C40
phytoene
CrtB, CrtMmut
OPP
GGPP (C20PP)
BstFPSY81A
IPP
OPP
FPP (C15PP)
CrtMmut
16-isopentenylphytoene
E. coli oxidase or chemical oxidation
C45
OH-isopentenyl-ζ-carotene 9
CrtI (in vitro)
OH-isopentenylphytofluene 8
CrtI (in vitro)
OH-isopentenylphytoene 7
BstFPSY81A
IPP
OPP
FGPP (C25PP)
CrtMmut
16,16'-diisopentenylphytoene
E. coli oxidase or chemical oxidation
C50
OH-diisopentenylphytoene 14
CrtI (in vitro)

Figure 3.5

Figure 3.5. Hydroxylation and subsequent *in vitro* desaturation of carotenoid backbones. *E. coli* cells harboring plasmid pUCmodII-*crtM*$_{F26L}$-*bstFPS*$_{Y81A}$ or pUCmodII-*crtM*$_{F26A,W38A}$-*bstFPS*$_{Y81A}$ accumulate monohydroxylated phytoene as well as monohydroxylated C_{45} and C_{50} carotenoid backbones **7** and **14**, respectively (Figures 3.7-3.9). Neither the mechanism nor the specific sites of hydroxylation are known. *In vitro*, *E. uredovora* CrtI catalyzes 4 desaturation steps on the hydroxylated C_{40} backbone, with each intermediate in the desaturation sequence being detected (Table 3.3). *In vitro* desaturation of the hydroxylated C_{45} backbone **7** yielded 1-step product **8** and 2-step product **9**. The hydroxylated C_{50} backbone **14** was not detectably desaturated by CrtI *in vitro* (crossed-out arrow). CrtB, phytoene synthase; $CrtM_{mut}$, mutant synthase $CrtM_{F26L}$ or $CrtM_{F26A,W38A}$; IPP, isopentenyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; FGPP, farnesylgeranyl diphosphate.

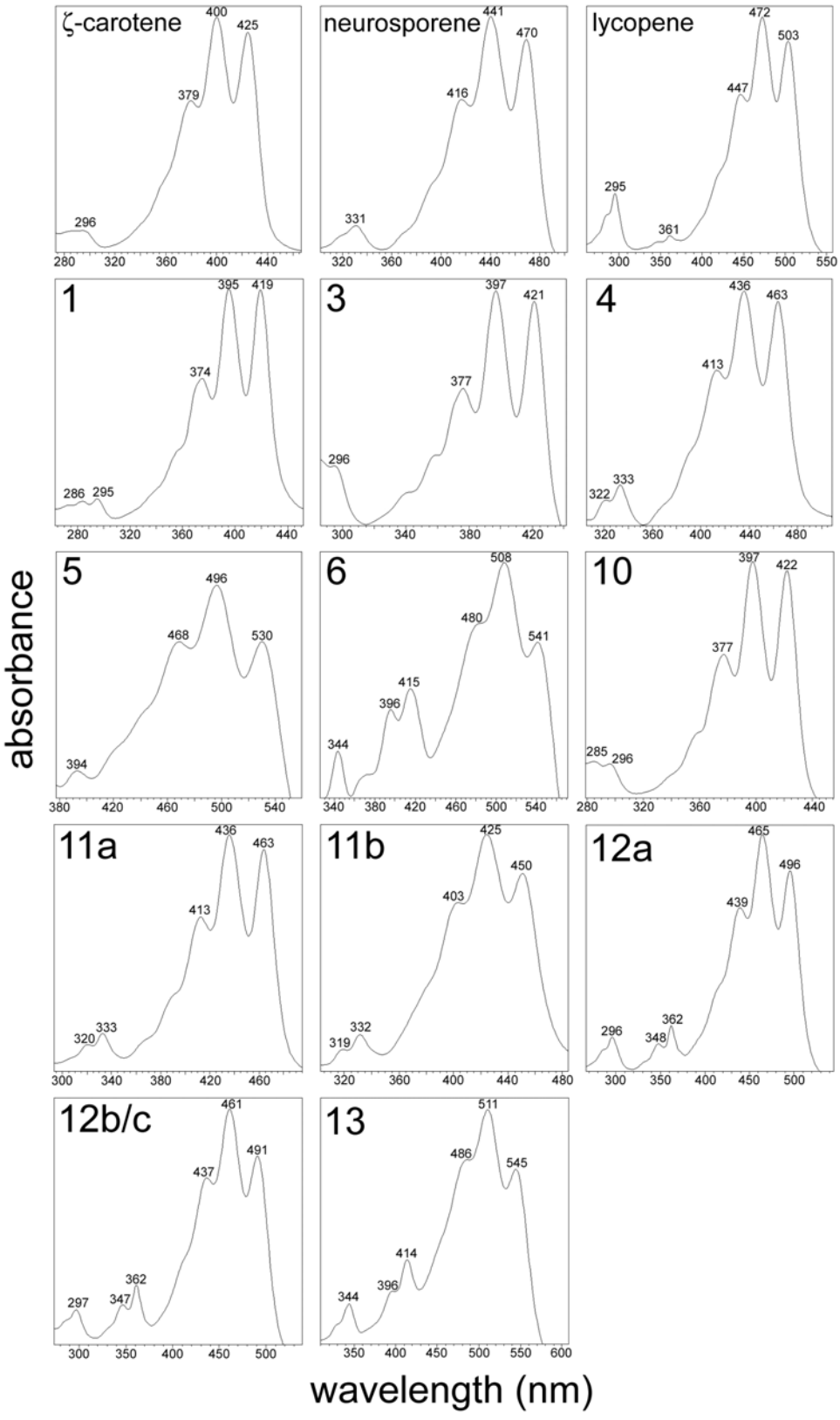

ζ-carotene
400
425
379
296
neurosporene
441
470
416
331
lycopene
472
503
447
295
361
1
395
419
374
286
295
3
397
421
377
296
4
436
463
413
322
333
5
496
468
530
394
6
508
480
541
415
396
344
10
397
422
377
285
296
11a
436
463
413
320
333
11b
425
450
403
319
332
12a
465
496
439
296
362
348
12b/c
461
491
437
297
347
362
13
511
486
545
414
396
344
absorbance
wavelength (nm)

Figure 3.6

Figure 3.6. UV-visible absorption spectra of novel carotenoids biosynthesized by recombinant *E. coli*. Numeric labels refer to structures depicted in Figures 3.2-3.4; multiple letters indicate that the spectrum may be that of either or both structures. Spectra were measured by photodiode array directly after HPLC separation, with a mobile phase of acetonitrile:isopropanol 93:7 (by volume, see Materials and Methods). Spectra were taken before and after iodine-catalyzed photoisomerization; each spectrum shown is the most bathochromically shifted of the two. Absorption spectra of ζ-carotene, neurosporene, and lycopene generated by *in vitro* desaturation of authentic phytoene are shown for comparison (See Figure 3.4 for structures of these C_{40} carotenoids). Absorption maxima are labeled with the corresponding wavelength. Table 3.1 lists additional properties of these molecules; Table 3.2 lists their names.

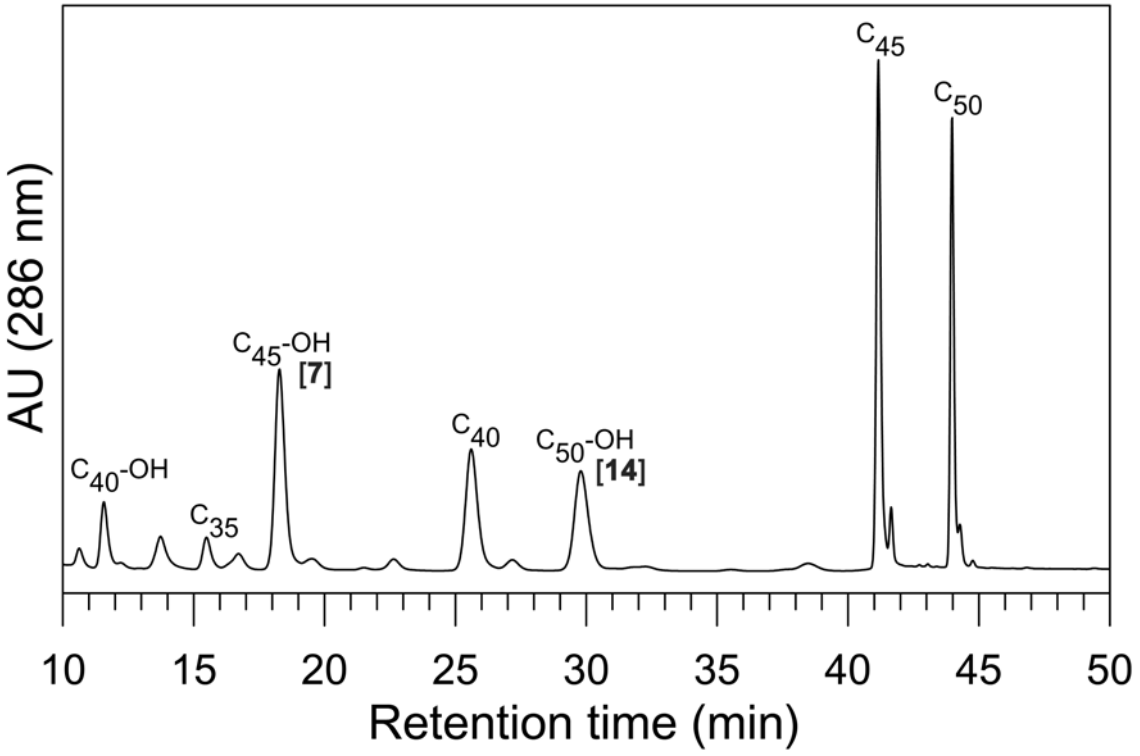


Figure 3.7. HPLC trace of carotenoid backbones extracted from a culture of XL1-Blue(pUCmodII-*crtM*$_{F26L}$-*bstFPS*$_{Y81A}$). Carotenoids were eluted using the gradient method described in Materials and Methods. The identity of each peak was confirmed by MS; the presence of monohydroxy substituents was further confirmed by acetylation and subsequent MS (see Figure 3.9). The C_{40} and C_{40}-OH peaks likely represent mixtures of symmetric and asymmetric C_{40} backbones that did not separate under these elution conditions. Labeled major peaks represent 15*Z* isomers; minor peaks eluting just after major peaks represent all-*E* isomers. Numbers in square brackets refer to products shown in Figure 3.5.

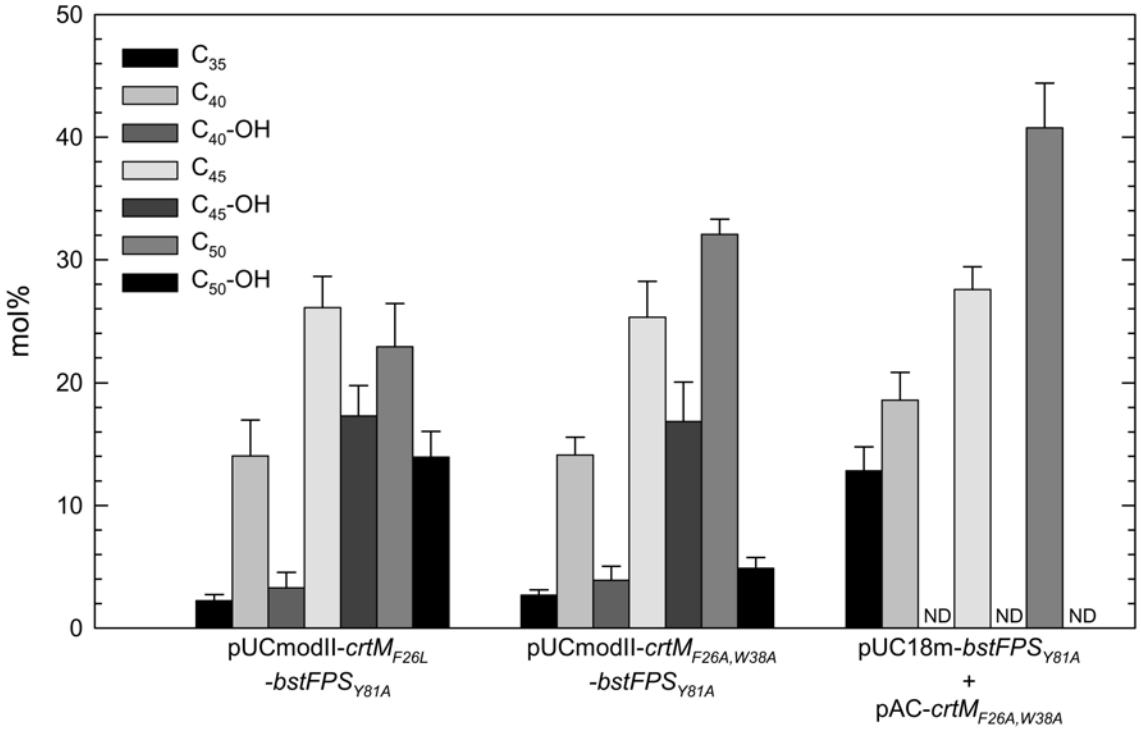


Figure 3.8. Relative molar quantities of the carotenoid backbones produced by recombinant XL1-Blue cultures expressing BstFPS$_{Y81A}$ and a CrtM variant. Transformed plasmid(s) are shown on the horizontal axis. Bar heights represent the average of measurements of at least three independent cultures; error bars, standard deviations. Total carotenoid titers varied to a greater extent than relative molar quantities, approximate values were: pUCmodII-*crtM*$_{F26L}$-*bstFPS*$_{Y81A}$, ~100 nmol/g dry cells; pUCmodII-*crtM*$_{F26A,W38A}$-*bstFPS*$_{Y81A}$, ~300 nmol/g dry cells; pUC18m-*bstFPS*$_{Y81A}$ + pAC-*crtM*$_{F26A,W38A}$, ~70 nmol/g dry cells. ND, not detected. The C_{40} and C_{40}-OH data series likely represent mixtures of symmetric and asymmetric C_{40} backbones that did not separate by the HPLC methods described in Materials and Methods.

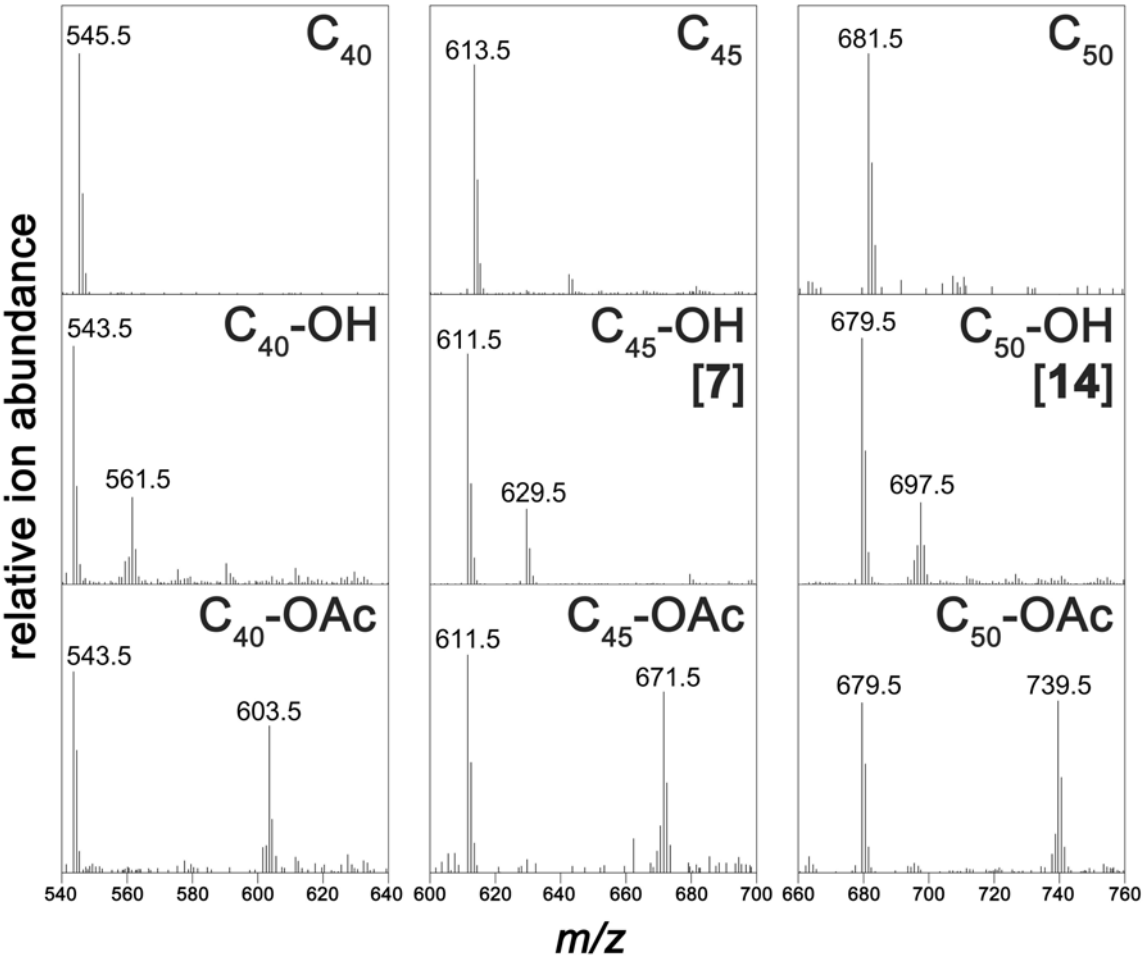

relative ion abundance
545.5
C40
543.5
C40-OH
561.5
543.5
C40-OAc
603.5
540 560 580 600 620 640
613.5
C45
611.5
C45-OH
[7]
629.5
611.5
C45-OAc
671.5
600 620 640 660 680 700
681.5
C50
679.5
C50-OH
[14]
697.5
679.5
C50-OAc
739.5
660 680 700 720 740 760
m/z

Figure 3.9

Figure 3.9. APCI mass spectra of unmodified, hydroxylated, and acetylated carotenoid backbones. Unmodified carotenoid backbones ionize by protonation in APCI-MS, giving $[M+H]^+$ quasimolecular ions (calculated monoisotopic masses for carotenoid backbones: C_{40}, 544.5 Da; C_{45}, 612.6 Da; C_{50}, 680.6 Da). Hydroxylated carotenoid backbones ionize by protonation followed by loss of water, yielding an $[(M+H)-H_2O]^+$ ion (*m/z* = 543.9 for C_{40}-OH, 611.5 for C_{45}-OH, and 679.5 for C_{50}-OH) in addition to the $[M+H]^+$ quasimolecular ion (*m/z* = 561.5 for C_{40}-OH, 629.5 for C_{45}-OH, and 697.5 for C_{50}-OH). Hydroxylated backbones were acetylated by reaction with acetic anhydride. Mass spectra of the acetylated products reveal a similar mode of ionization to their hydroxylated counterparts, yielding the $[(M+H)-CH_3COOH]^+$ ion (identical to the $[(M+H)-H_2O]^+$ ions listed above) in addition to the $[M+H]^+$ quasimolecular ion (*m/z* = 603.9 for C_{40}-OAc, 671.5 for C_{45}-OAc, and 739.5 for C_{50}-OAc). Numbers in square brackets refer to products shown in Figure 3.5.

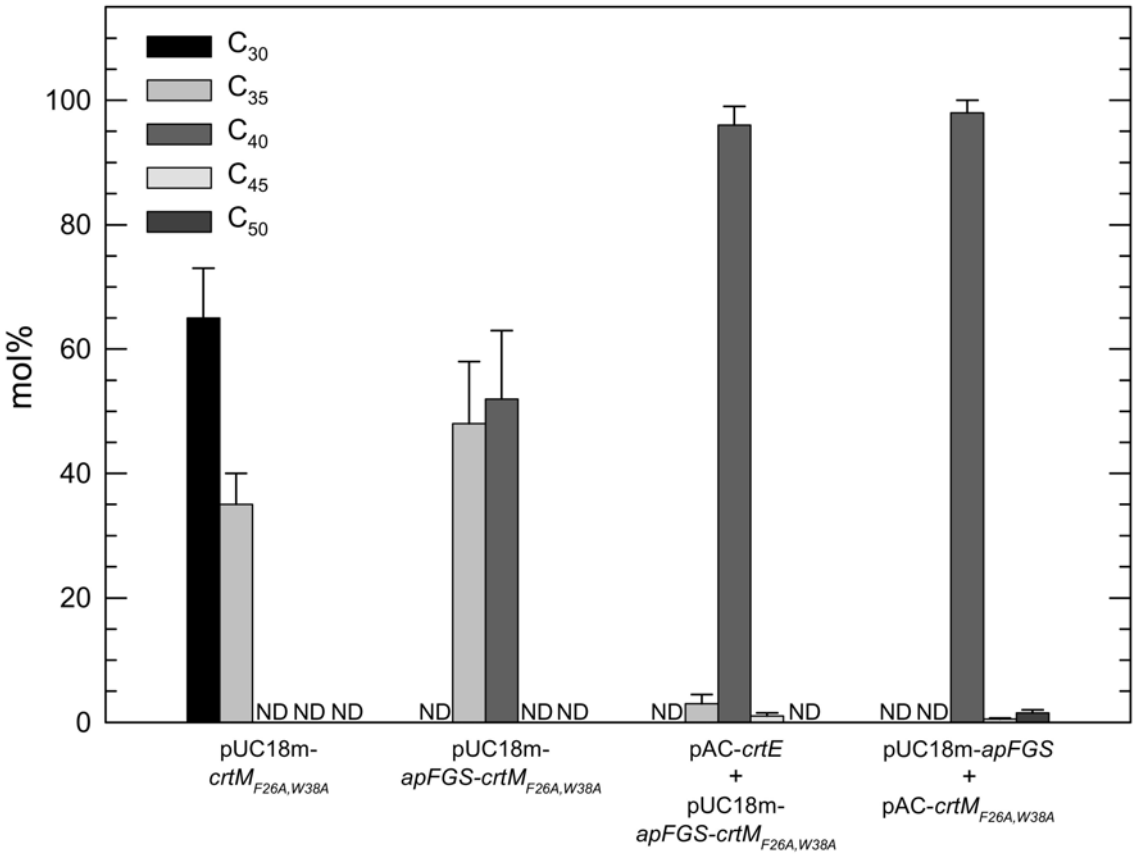


Figure 3.10. Relative molar quantities of the carotenoid backbones produced by recombinant XL1-Blue cultures expressing ApFGS and CrtM$_{F26A,W38A}$. Transformed plasmid(s) are shown on the horizontal axis. Bar heights represent the average of measurements on two independent cultures; error bars, standard deviations. Total carotenoid titers varied to a greater extent than relative molar quantities, approximate values were: pUC18m-*crtM*$_{F26A,W38A}$, ~30 nmol/g dry cells; pUC18m-*apFGS-crtM*$_{F26A,W38A}$, ~100 nmol/g dry cells; pAC-*crtE* + pUC18m-*apFGS-crtM*$_{F26A,W38A}$, ~300 nmol/g dry cells; pUC18m-*apFGS* + pAC-*crtM*$_{F26A,W38A}$, ~100 nmol/g dry cells. ND, not detected. The relatively low proportion of C_{50} backbones in the cultures expressing ApFGS indicates a low C_{25}PP supply in the cells. Therefore, ApFGS acts predominantly as a C_{20}PP synthase under these conditions, and the C_{40} backbone population should thus consist almost exclusively of phytoene.

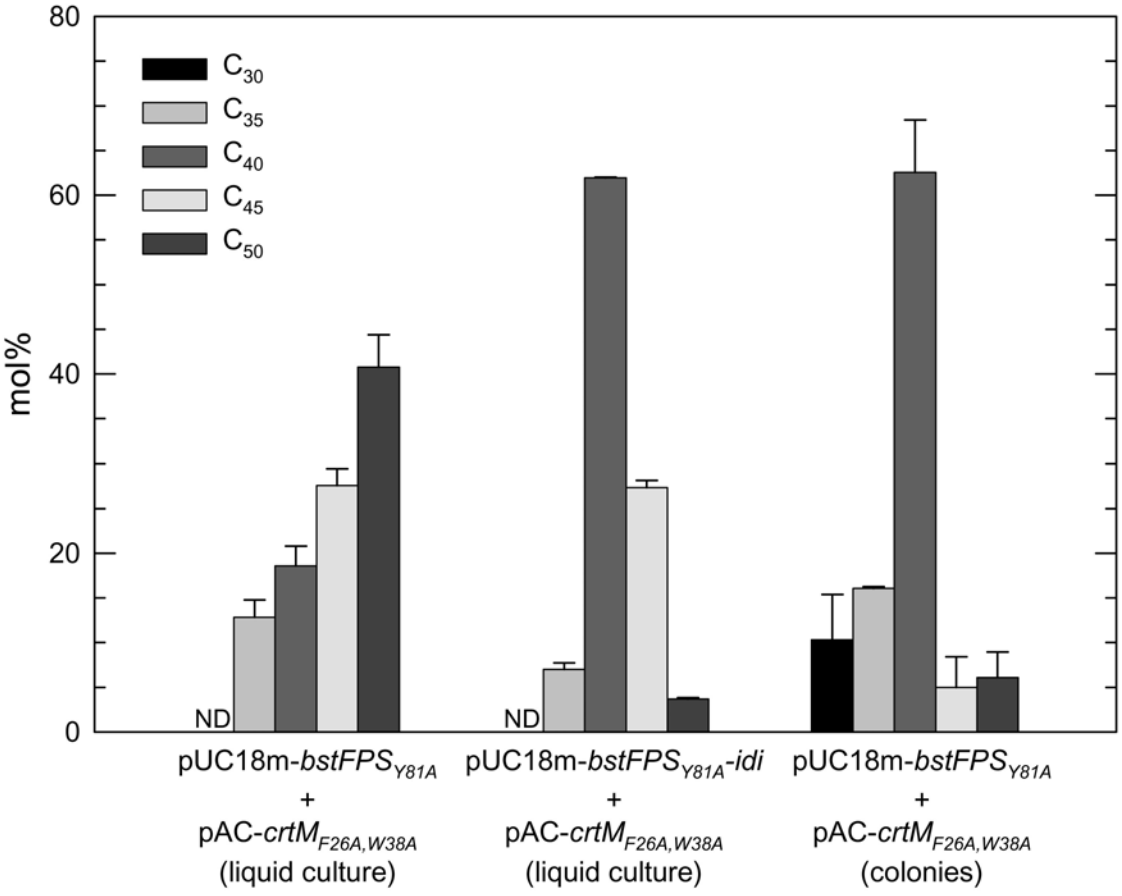

80
60
40
20
0
mol%
C30
C35
C40
C45
C50
ND
ND
pUC18m-bstFPSY81A + pAC-crtMF26A,W38A (liquid culture)
pUC18m-bstFPSY81A-idi + pAC-crtMF26A,W38A (liquid culture)
pUC18m-bstFPSY81A + pAC-crtMF26A,W38A (colonies)

Figure 3.11

Figure 3.11. Idi overexpression and physiological state strongly influence the distribution of carotenoid backbones in recombinant XL1-Blue cells expressing BstFPS$_{Y81A}$ and CrtM$_{F26A,W38A}$. Transformed plasmid(s) and physiological state (liquid culture or colonies) are shown on the horizontal axis. For the first data set, which is repeated from Figure 3.8, bar heights represent the average of measurements on four replicate cultures. For the other sets, bar heights represent the average of measurements on two independent TB liquid cultures (second set) or on collected colonies from two independent LB-agar plates (third set). Total carotenoid titers varied to a greater extent than relative molar quantities, approximate values were: pUC18m-*bstFPS*$_{Y81A}$ + pAC-*crtM*$_{F26A,W38A}$ (liquid culture), ~70 nmol/g dry cells; pUC18m-*bstFPS*$_{Y81A}$-*idi* + pAC-*crtM*$_{F26A,W38A}$ (liquid culture), ~700 nmol/g dry cells; pUC18m-*bstFPS*$_{Y81A}$ + pAC-*crtM*$_{F26A,W38A}$ (colonies), ~70 nmol/g dry cells. ND, not detected. The relatively low proportion of C_{50} backbones in the second and third data sets indicates a low C_{25}PP supply in the cells; therefore, the C_{40} backbone population should consist predominantly of phytoene.

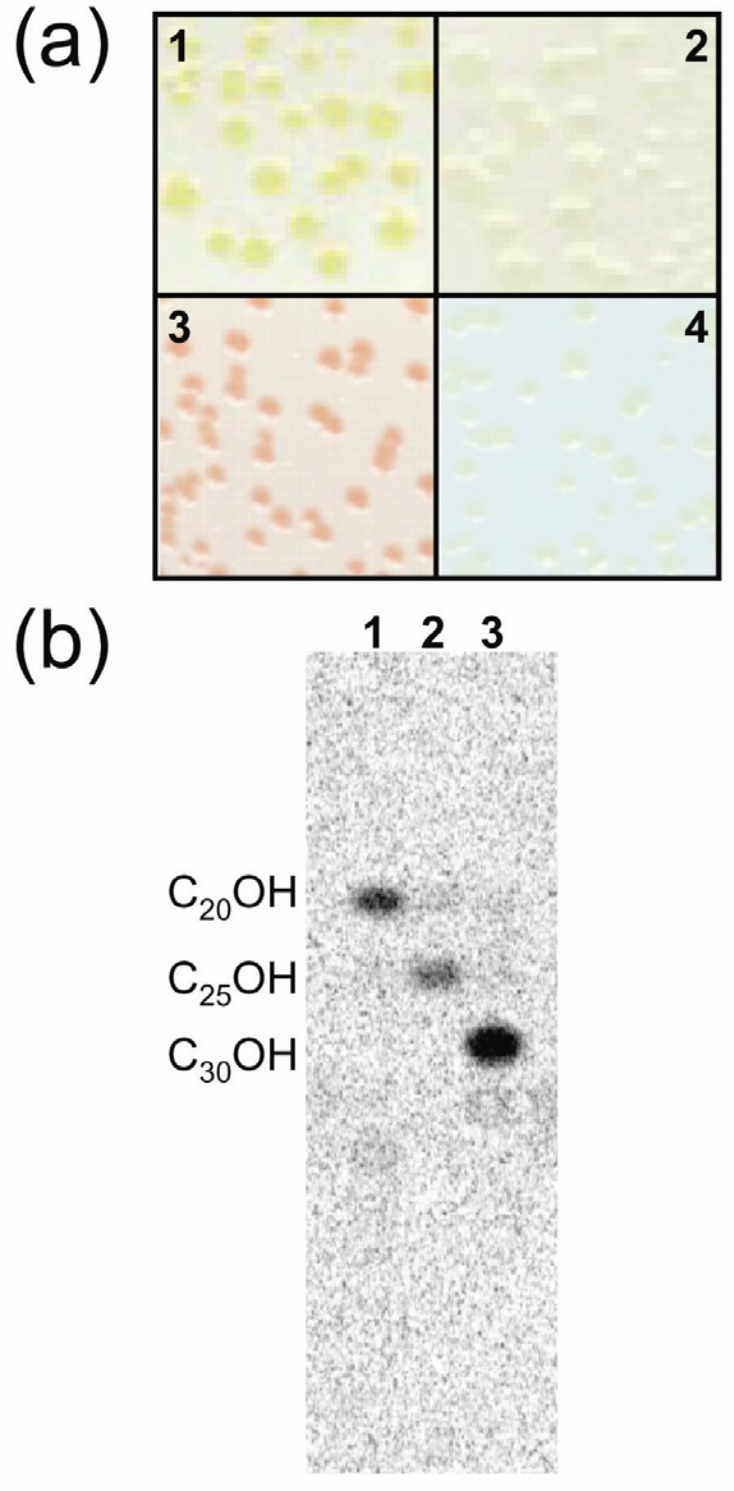


Figure 3.12

Figure 3.12. Product specificity of *M. luteus* hexaprenyl diphosphate synthase (HexPS). ***(a)*** Pigmentation of *E. coli* XL1-Blue colonies expressing different plasmids and imaged on white nitrocellulose: panel 1, pUC18m-*crtM-crtN*; panel 2, pUC18m-*hexPS-crtM-crtN*; panel 3, pUC18m-*crtE-crtB-crtI*; panel 4, pUC18m-*hexPS-crtB-crtI*. ***(b)*** TLC-autoradiogram of the hydrolyzed products from *in vitro* reactions of FPP (C_{15}PP) with ^{14}C-labeled IPP catalyzed by various isoprenyl diphosphate synthases: lane 1, *E. uredovora* CrtE; lane 2, BstFPS$_{Y81A}$; lane 3, HexPS. Prenyl diphosphate products were hydrolyzed with acid phosphatase to the corresponding alcohol and then separated by reverse-phase TLC (see Materials and Methods). The direction of mobile phase flow was bottom to top. This autoradiogram was obtained by Adam Hartwick.

REFERENCES

1. **Aharoni, A., L. Gaidukov, O. Khersonsky, Q. G. S. Mc, C. Roodveldt, and D. S. Tawfik.** 2005. The 'evolvability' of promiscuous protein functions. Nat. Genet. **37:**73-6.
2. **Albrecht, M., G. Sandmann, D. Musker, and G. Britton.** 1991. Identification of epoxy- and hydroxyphytoene from norflurazon-treated *Scenedesmus*. J. Agric. Food Chem. **39:**566-569.
3. **Albrecht, M., S. Takaichi, N. Misawa, G. Schnurr, P. Böger, and G. Sandmann.** 1997. Synthesis of atypical cyclic and acyclic hydroxy carotenoids in *Escherichia coli* transformants. J. Biotechnol. **58:**177-185.
4. **Albrecht, M., S. Takaichi, S. Steiger, Z. Y. Wang, and G. Sandmann.** 2000. Novel hydroxycarotenoids with improved antioxidative properties produced by gene combination in *Escherichia coli*. Nat. Biotechnol. **18:**843-846.
5. **Aragón, C. M., F. J. Murillo, M. D. de la Guardia, and E. Cerdá-Olmedo.** 1976. An enzyme complex for the dehydrogenation of phytoene in *Phycomyces*. Eur. J. Biochem. **63:**71-5.
6. **Boucher, Y., M. Kamekura, and W. F. Doolittle.** 2004. Origins and evolution of isoprenoid lipid biosynthesis in archaea. Mol. Microbiol. **52:**515-27.
7. **Britton, G.** 1998. Overview of carotenoid biosynthesis, p. 13-147. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids, vol. 3: Biosynthesis and Metabolism. Birkhäuser Verlag, Basel.
8. **Britton, G.** 1995. Structure and properties of carotenoids in relation to function. Faseb J. **9:**1551-8.
9. **Britton, G.** 1995. UV/Visible Spectroscopy, p. 13-62. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids, vol. 1B: Spectroscopy. Birkhäuser Verlag, Basel.
10. **Buckingham, J.** 2004. Dictionary of Natural Products. Chapman & Hall/CRC Press. http://www.chemnetbase.com/scripts/dnpweb.exe.
11. **Burch, R. R., Y. Dong, C. Fincher, M. Goldfinger, and P. Rouviere.** 2004. Electrical properties of polyunsaturated natural products: field effect mobility of carotenoid polyenes. Synth. Met. **146:**43-46.
12. **Candau, R., E. R. Bejarano, and E. Cerdá-Olmedo.** 1991. *In vivo* channeling of substrates in an enzyme aggregate for beta-carotene biosynthesis. Proc. Natl. Acad. Sci. U.S.A. **88:**4936-40.
13. **Colodner, R.** 2005. Extended-spectrum beta-lactamases: a challenge for clinical microbiologists and infection control specialists. Am. J. Infect. Control **33:**104-7.
14. **Davies, B. H., C. J. Hallett, R. A. London, and A. F. Rees.** 1974. The nature of "theta-carotene." Phytochemistry **13:**1209-1217.
15. **Davis, J. B., L. M. Jackman, P. T. Siddons, and B. C. L. Weedon.** 1966. Carotenoids and related compounds Part XV: The structure and synthesis of phytoene, phytofluene, zeta-carotene, and neurosporene. J. Chem. Soc. (C)**:**2154-2165.

16. **De la Guardia, M. D., C. M. Aragón, F. J. Murillo, and E. Cerdá-Olmedo.** 1971. A carotenogenic enzyme aggregate in *Phycomyces*: evidence from quantitive complementation. Proc. Natl. Acad. Sci. U.S.A. **68:**2012-5.
17. **Edge, R., D. J. McGarvey, and T. G. Truscott.** 1997. The carotenoids as anti-oxidants—a review. J. Photochem. Photobiol. B. **41:**189-200.
18. **Englert, G.** 1995. NMR Spectroscopy, p. 147-260. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids, vol. 1B: Spectroscopy. Birkhäuser Verlag, Basel.
19. **Eugster, C. H.** 1995. Chemical Derivatization: Microscale Tests for the Presence of Common Functional Groups in Carotenoids, p. 71-80. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids, vol. 1A: Isolation and Analysis. Birkhäuser Verlag, Basel.
20. **Farmer, W. R., and J. C. Liao.** 2000. Improving lycopene production in *Escherichia coli* by engineering metabolic control. Nat. Biotechnol. **18:**533-537.
21. **Firn, R. D., and C. G. Jones.** 2000. The evolution of secondary metabolism—a unifying model. Mol. Microbiol. **37:**989-994.
22. **Firn, R. D., and C. G. Jones.** 1996. An explanation of secondary product "redundancy," p. 295-312. *In* J. T. Romeo, J. A. Saunders, and P. Barbosa (ed.), Recent Advances in Phytochemistry, vol. 30. Phytochemical Diversity and Redundancy in Ecological Interactions. Plenum Press, New York.
23. **Firn, R. D., and C. G. Jones.** 2003. Natural products—a simple model to explain chemical diversity. Nat. Prod. Rep. **20:**382-391.
24. **Fraser, P. D., N. Misawa, H. Linden, S. Yamano, K. Kobayashi, and G. Sandmann.** 1992. Expression in *Escherichia coli*, purification, and reactivation of the recombinant *Erwinia uredovora* phytoene desaturase. J. Biol. Chem. **267:**19891-19895.
25. **Hausmann, A., and G. Sandmann.** 2000. A single five-step desaturase is involved in the carotenoid biosynthesis pathway to beta-carotene and torulene in *Neurospora crassa*. Fungal Genetics and Biology **30:**147-153.
26. **Hemmi, H., S. Ikejiri, S. Yamashita, and T. Nishino.** 2002. Novel medium-chain prenyl diphosphate synthase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. J. Bacteriol. **184:**615-20.
27. **Jensen, R. A.** 1976. Enzyme recruitment in evolution of new function. Annu. Rev. Microbiol. **30:**409-25.
28. **Jones, C. G., and R. D. Firn.** 1991. On the evolution of plant secondary chemical diversity. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. **333:**273-280.
29. **Karrer, P., and C. H. Eugster.** 1951. Carotinoidsynthesen 8. Synthese des dodecapreno-beta-carotins. Helv. Chim. Acta **34:**1805-1814.
30. **Krubasik, P., M. Kobayashi, and G. Sandmann.** 2001. Expression and functional analysis of a gene cluster involved in the synthesis of decaprenoxanthin reveals the mechanisms for C_{50} carotenoid formation. Eur. J. Biochem. **268:**3702-3708.
31. **Lee, P. C., A. Z. R. Momen, B. N. Mijts, and C. Schmidt-Dannert.** 2003. Biosynthesis of structurally novel carotenoids in *Escherichia coli*. Chem. Biol. **10:**453-462.

32. **Liaaen-Jensen, S.** 1995. Combined Approach: Identification and Structure Elucidation of Carotenoids, p. 343-354. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids, vol. 1B: Spectroscopy. Birkhäuser Verlag, Basel.
33. **Linden, H., N. Misawa, D. Chamovitz, I. Pecker, J. Hirschberg, and G. Sandmann.** 1991. Functional complementation in *Escherichia coli* of different phytoene desaturase genes and analysis of accumulated carotenes. Z. Naturforsch. (C) **46:**1045-1051.
34. **Lorenz, R. T., and G. R. Cysewski.** 2000. Commercial potential for *Haematococcus* microalgae as a natural source of astaxanthin. Trends Biotechnol. **18:**160-7.
35. **Mijts, B. N., P. C. Lee, and C. Schmidt-Dannert.** 2005. Identification of a carotenoid oxygenase synthesizing acyclic xanthophylls: combinatorial biosynthesis and directed evolution. Chem. Biol. **12:**453-60.
36. **Murillo, F. J., S. Torres-Martinez, C. M. Aragón, and E. Cerdá-Olmedo.** 1981. Substrate transfer in carotene biosynthesis in *Phycomyces*. Eur. J. Biochem. **119:**511-6.
37. **Ohnuma, S., K. Narita, T. Nakazawa, C. Ishida, Y. Takeuchi, C. Ohto, and T. Nishino.** 1996. A role of the amino acid residue located on the fifth position before the first aspartate-rich motif of farnesyl diphosphate synthase on determination of the final product. J. Biol. Chem. **271:**30748-30754.
38. **Ohnuma, S. I., T. Nakazawa, H. Hemmi, A. M. Hallberg, T. Koyama, K. Ogura, and T. Nishino.** 1996. Conversion from farnesyl diphosphate synthase to geranylgeranyl diphosphate synthase by random chemical mutagenesis. J. Biol. Chem. **271:**10087-10095.
39. **Raisig, A., G. Bartley, P. Scolnik, and G. Sandmann.** 1996. Purification in an active state and properties of the 3-step phytoene desaturase from *Rhodobacter capsulatus* overexpressed in *Escherichia coli*. J. Biochem. (Tokyo) **119:**559-64.
40. **Raisig, A., and G. Sandmann.** 2001. Functional properties of diapophytoene and related desaturases of C_{30} and C_{40} carotenoid biosynthetic pathways. Biochim. Biophys. Acta Mol. Cell Biol. Lipids **1533:**164-170.
41. **Rison, S. C., and J. M. Thornton.** 2002. Pathway evolution, structurally speaking. Curr. Opin. Struct. Biol. **12:**374-82.
42. **Sandmann, G.** 2002. Combinatorial biosynthesis of carotenoids in a heterologous host: A powerful approach for the biosynthesis of novel structures. ChemBioChem **3:**629-635.
43. **Sandmann, G., M. Albrecht, G. Schnurr, O. Knorzer, and P. Boger.** 1999. The biotechnological potential and design of novel carotenoids by gene combination in *Escherichia coli*. Trends Biotechnol. **17:**233-237.
44. **Schiedt, K., and S. Liaaen-Jensen.** 1995. Isolation and Analysis, p. 81-108. *In* G. Britton, S. Liaaen-Jensen, and H. Pfander (ed.). Carotenoids, vol. 1A: Isolation and Analysis. Birkhäuser Verlag, Basel.
45. **Schmidt-Dannert, C.** 2000. Engineering novel carotenoids in microorganisms. Curr. Opin. Biotechnol. **11:**255-261.
46. **Schmidt-Dannert, C., D. Umeno, and F. H. Arnold.** 2000. Molecular breeding of carotenoid biosynthetic pathways. Nat. Biotechnol. **18:**750-753.

47. **Shimizu, N., T. Koyama, and K. Ogura.** 1998. Molecular cloning, expression, and characterization of the genes encoding the two essential protein components of *Micrococcus luteus* B-P 26 hexaprenyl diphosphate synthase. J. Bacteriol. **180:**1578-81.

48. **Steiger, S., S. Takaichi, and G. Sandmann.** 2002. Heterologous production of two unusual acyclic carotenoids, 1,1'-dihydroxy-3,4-didehydrolycopene and 1-hydroxy-3,4,3',4'-tetradehydrolycopene by combination of the *crtC* and *crtD* genes from *Rhodobacter* and *Rubrivivax*. J. Biotechnol. **97:**51-8.

49. **Tachibana, A.** 1994. A novel prenyltransferase, farnesylgeranyl diphosphate synthase, from the haloalkaliphilic archaeon, *Natronobacterium pharaonis*. FEBS Lett. **341:**291-4.

50. **Tachibana, A., Y. Yano, S. Otani, N. Nomura, Y. Sako, and M. Taniguchi.** 2000. Novel prenyltransferase gene encoding farnesylgeranyl diphosphate synthase from a hyperthermophilic archaeon, *Aeropyrum pernix*: Molecular evolution with alteration in product specificity. Eur. J. Biochem. **267:**321-8.

51. **Takaichi, S.** 2000. Characterization of carotenes in a combination of a C_{18} HPLC column with isocratic elution and absorption spectra with a photodiode-array detector. Photosynth. Res. **65:**93-99.

52. **Tao, L., A. Schenzle, J. M. Odom, and Q. Cheng.** 2005. Novel carotenoid oxidase involved in biosynthesis of 4,4'-diapolycopene dialdehyde. Appl. Environ. Microbiol. **71:**3294-301.

53. **Umeno, D., and F. H. Arnold.** 2003. A C_{35} carotenoid biosynthetic pathway. Appl. Environ. Microbiol. **69:**3573-3579.

54. **Umeno, D., and F. H. Arnold.** 2004. Evolution of a pathway to novel long-chain carotenoids. J. Bacteriol. **186:**1531-1536.

55. **Umeno, D., K. Hiraga, and F. H. Arnold.** 2003. Method to protect a targeted amino acid residue during random mutagenesis. Nucleic Acids Res. **31:**e91.

56. **Umeno, D., A. V. Tobias, and F. H. Arnold.** 2002. Evolution of the C_{30} carotenoid synthase CrtM for function in a C_{40} pathway. J. Bacteriol. **184:**6690-6699.

57. **Vershinin, A.** 1999. Biological functions of carotenoids—diversity and evolution. Biofactors **10:**99-104.

58. **Wang, C. W., and J. C. Liao.** 2001. Alteration of product specificity of *Rhodobacter sphaeroides* phytoene desaturase by directed evolution. J. Biol. Chem. **276:**41161-41164.

59. **Wang, C. W., M. K. Oh, and J. C. Liao.** 1999. Engineered isoprenoid pathway enhances astaxanthin production in *Escherichia coli*. Biotechnol. Bioeng. **62:**235-241.

60. **Wang, K., and S. Ohnuma.** 1999. Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. Trends Biochem. Sci. **24:**445-451.

61. **Woodall, A. A., S. W. M. Lee, R. J. Weesie, M. J. Jackson, and G. Britton.** 1997. Oxidation of carotenoids by free radicals: relationship between structure and reactivity. Biochim. Biophys. Acta-Gen. Subj. **1336:**33-42.

62. **Ycas, M.** 1974. On earlier states of the biochemical system. J. Theor. Biol. **44:**145-60.

63. **Young, A. J., and G. M. Lowe.** 2001. Antioxidant and prooxidant properties of carotenoids. Arch. Biochem. Biophys. **385:**20-7.

64. **Zechmeister, L., and B. K. Koe.** 1954. Stepwise dehydrogenation of the colorless polyenes phytoene and phytofluene with N-bromosuccinimide to carotenoid pigments. J. Am. Chem. Soc. **76:**2923-2926.

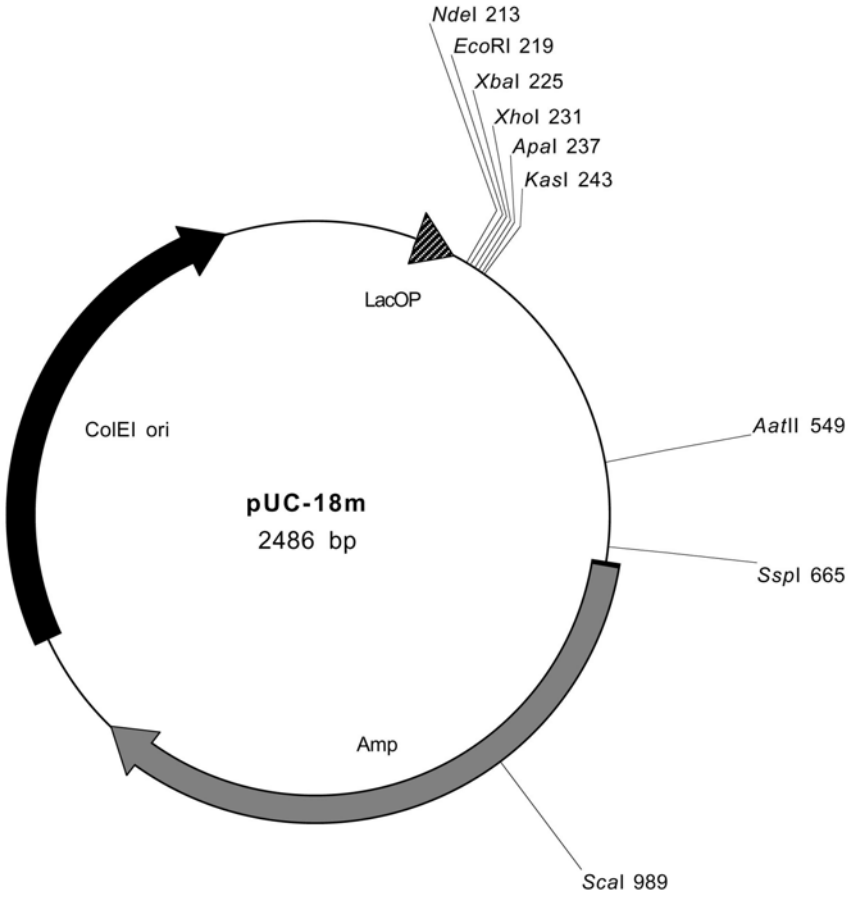
APPENDIX A

Plasmid Sequences and Maps

Plasmid name pUC18m (empty vector); 2486 bp

Construction
The *lacZ* gene from pUC18 was cut out and replaced by a *lac* promoter/operator and the multi-cloning site: *Nde*I-*Eco*RI-*Xba*I-*Xho*I-*Apa*I-*Ehe*I/*Kas*I.

Plasmid map



pUC18m sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGGAATTCTCTAGACTCGAGGG
GCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatc
tgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggc
atccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacg
aaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatg
tgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataata
ttgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcaccca

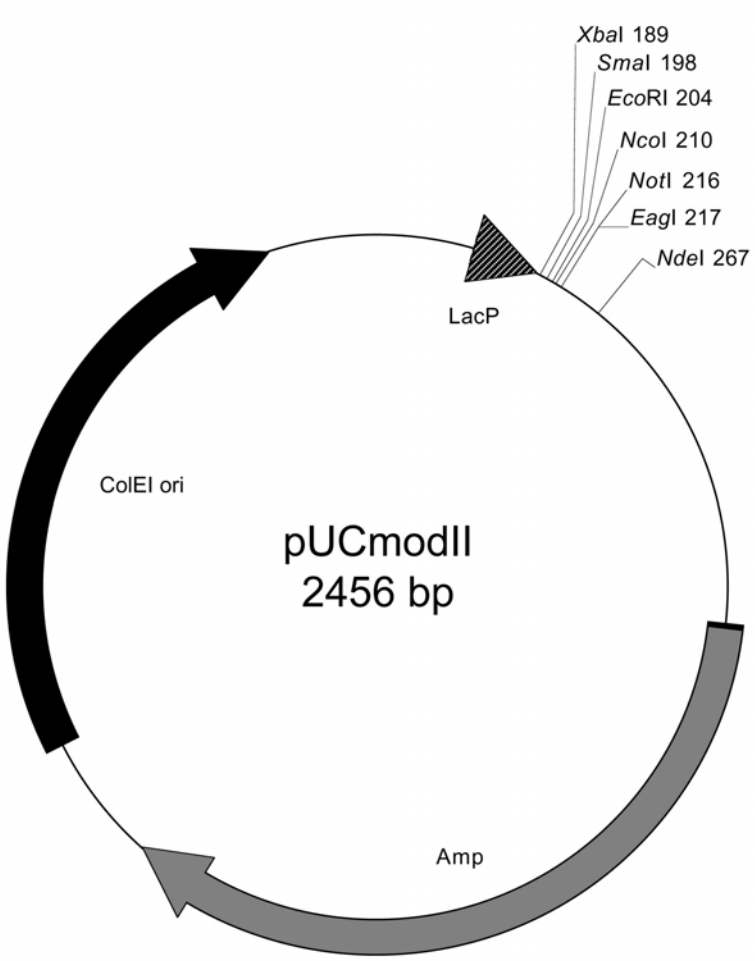
gaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaa
gatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattg
acgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatc
ttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaac
gatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagct
gaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcg
aactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttt
ccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggta
agccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagatagg
tgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaagga
tctaggtgaagatccttttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaa
gatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtt
tgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtag
ccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagt
ggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggggttc
gtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgctt
cccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcgg
agcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatc
ccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtc
agtgagcgaggaagcggaaga

Plasmid name pUCmodII (empty vector); 2456 bp

Construction
Based on pUC19. The *lac* operator and multi cloning site from pUC19 were removed and replaced with multi cloning site: *Xba*I-*Sma*I-*Eco*RI-*Nco*I-*Not*I/*Eag*I (see plasmid map below).

Remarks
High copy number (several hundred). Since the *lac* operator has been removed, the genes cloned into the MCS are expressed constitutively in *E. coli*.

Plasmid map



pUCmodII sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcgTCTAGAgcgCCCGGGGAATTCCCATGGGCGGCCGCtgcggtattttct
ccttacgcatctgtgcggtatttcacaccgCATATGgtgcactctcagtacaatctgctctgatgccgcatagttaagccagc
cccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtc
tccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttat

aggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttct
aaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattca
acatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgct
gaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaa
cgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgc
cgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaatt
atgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccg
ctttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcg
tgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaat
taatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatct
ggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacg
acggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtca
gaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatg
accaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttc
cgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactct
gtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgga
ctcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaac
gacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtat
ccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcg
ggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcgg
cctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcct
ttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pACmod (empty vector); 4245 bp; medium copy

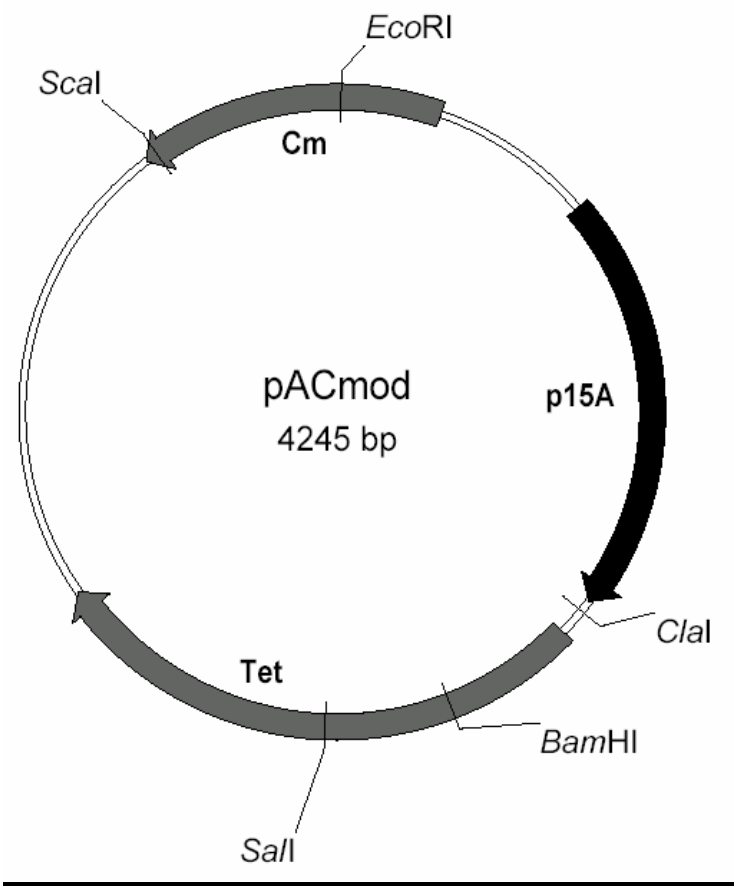
Construction

Identical to commercially available pACYC184 except the *Xba*I site in that plasmid (TCTAGA) has been mutated to TCAAGA. Although this mutation is in the p15A origin of replication, the plasmid seems fine.

Sequence landmarks

Chloramphenicol resistance gene 3805-219 (reverse orientation); tetracycline resistance gene 1581-2771; p15A origin 581-1493

Plasmid map



pACmod sequence

gaattccggatgagcattcatcaggcgggcaagaatgtgaataaaggccggataaaacttgtgcttatttttctttacggtctttaaa
aaggccgtaatatccagctgaacggtctggttataggtacattgagcaactgactgaaatgcctcaaaatgttctttacgatgccat
tgggatatatcaacggtggtatatccagtgatttttttctccattttagcttccttagctcctgaaaatctcgataactcaaaaaatacgc
ccggtagtgatcttatttcattatggtgaaagttggaacctcttacgtgccgatcaacgtctcattttcgccaaaagttggcccaggg
cttcccggtatcaacagggacaccaggatttatttattctgcgaagtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcg
ggtgatgctgccaacttactgatttagtgtatgatggtgtttttgaggtgctccagtggcttctgtttctatcagctgtccctcctgttca
gctactgacggggtggtgcgtaacggcaaaagcaccgccggacatcagcgctagcggagtgtatactggcttactatgttggc

actgatgagggtgtcagtgaagtgcttcatgtggcaggagaaaaaaggctgcaccggtgcgtcagcagaatatgtgatacagg
atatattccgcttcctcgctcactgactcgctacgctcggtcgttcgactgcggcgagcggaaatggcttacgaacggggcgga
gatttcctggaagatgccaggaagatacttaacagggaagtgagagggccgcggcaaagccgtttttccataggctccgcccc
cctgacaagcatcacgaaatctgacgctcaaatcagtggtggcgaaacccgacaggactataaagataccaggcgtttccccct
ggcggctccctcgtgcgctctcctgttcctgcctttcggtttaccggtgtcattccgctgttatggccgcgtttgtctcattccacgcc
tgacactcagttccgggtaggcagttcgctccaagctggactgtatgcacgaaccccccgttcagtccgaccgctgcgccttatc
cggtaactatcgtcttgagtccaacccggaaagacatgcaaaagcaccactggcagcagccactggtaattgatttagaggagtt
agtcttgaagtcatgcgccggttaaggctaaactgaaaggacaagttttggtgactgcgctcctccaagccagttacctcggttca
aagagttggtagctcagagaaccttcgaaaaaccgccctgcaaggcggtttttcgttttcagagcaagagattacgcgcagacc
aaaacgatctcaagaagatcatcttattaatcagataaaatatttcaagatttcagtgcaatttatctcttcaaatgtagcacctgaagt
cagccccatacgatataagttgtaattctcatgtttgacagcttatcATCGATaagctttaatgcggtagtttatcacagttaaatt
gctaacgcagtcaggcaccgtgtatgaaatctaacaatgcgctcatcgtcatcctcggcaccgtcaccctggatgctgtaggcat
aggcttggttatgccggtactgccgggcctcttgcgggatatcgtccattccgacagcatcgccagtcactatggcgtgctgcta
gcgctatatgcgttgatgcaatttctatgcgcacccgttctcggagcactgtccgaccgctttggccgccgcccagtcctgctcgc
ttcgctacttggagccactatcgactacgcgatcatggcgaccacacccgtcctgtGGATCCtctacgccggacgcatcgt
ggccggcatcaccggcgccacaggtgcggttgctggcgcctatatcgccgacatcaccgatggggaagatcgggctcgcca
cttcgggctcatgagcgcttgtttcggcgtgggtatggtggcaggccccgtggccgggggactgttgggcgccatctccttgcat
gcaccattccttgcggcggcggtgctcaacggcctcaacctactactgggctgcttcctaatgcaggagtcgcataagggagag
cGTCGACcgatgcccttgagagccttcaacccagtcagctccttccggtgggcgcggggcatgactatcgtcgccgcactt
atgactgtcttctttatcatgcaactcgtaggacaggtgccggcagcgctctgggtcattttcggcgaggaccgctttcgctggag
cgcgacgatgatcggcctgtcgcttgcggtattcggaatcttgcacgccctcgctcaagccttcgtcactggtcccgccaccaaa
cgtttcggcgagaagcaggccattatcgccggcatggcggccgacgcgctgggctacgtcttgctggcgttcgcgacgcgag
gctggatggccttccccattatgattcttctcgcttccggcggcatcgggatgcccgcgttgcaggccatgctgtccaggcaggt
agatgacgaccatcagggacagcttcaaggatcgctcgcggctcttaccagcctaacttcgatcactggaccgctgatcgtcac
ggcgatttatgccgcctcggcgagcacatggaacgggttggcatggattgtaggcgccgccctataccttgtctgcctccccgc
gttgcgtcgcggtgcatggagccgggccacctcgacctgaatggaagccggcggcacctcgctaacggattcaccactccaa
gaattggagccaatcaattcttgcggagaactgtgaatgcgcaaaccaaccttggcagaacatatccatcgcgtccgccatctc
cagcagccgcacgcggcgcatctcgggcagcgttgggtcctggccacgggtgcgcatgatcgtgctcctgtcgttgaggacc
cggctaggctggcggggttgccttactggttagcagaatgaatcaccgatacgcgagcgaacgtgaagcgactgctgctgcaa
aacgtctgcgacctgagcaacaacatgaatggtcttcggtttccgtgtttcgtaaagtctggaaacgcggaagtcccctacgtgct
gctgaagttgcccgcaacagagagtggaaccaaccggtgataccacgatactatgactgagagtcaacgccatgagcggcct
catttcttattctgagttacaacagtccgcaccgctgtccggtagctccttccggtgggcgcggggcatgactatcgtcgccgcac
ttatgactgtcttctttatcatgcaactcgtaggacaggtgccggcagcgcccaacagtcccccggccacggggcctgccaccat
acccacgccgaaacaagcgccctgcaccattatgttccggatctgcatcgcaggatgctgctggctaccctgtggaacacctac
atctgtattaacgaagcgctaaccgtttttatcaggctctgggaggcagaataaatgatcatatcgtcaattattacctccacgggg
agagcctgagcaaactggcctcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccagca
atagacataagcggctatttaacgaccctgccctgaaccgacgaccgggtcgaatttgctttcgaatttctgccattcatccgcttat
tatcacttattcaggcgtagcaccaggcgtttaagggcaccaataactgccttaaaaaaattacgccccgccctgccactcatcgc
AGTACTgttgtaattcattaagcattctgccgacatggaagccatcacagacggcatgatgaacctgaatcgccagcggca
tcagcaccttgtcgccttgcgtataatatttgcccatggtgaaaacgggggcgaagaagttgtccatattggccacgtttaaatcaa
aactggtgaaactcacccagggattggctgagacgaaaaacatattctcaataaaccctttagggaaataggccaggttttcacc
gtaacacgccacatcttgcgaatatatgtgtagaaactgccggaaatcgtcgtggtattcactccagagcgatgaaaacgtttcag
tttgctcatggaaaacggtgtaacaagggtgaacactatcccatatcaccagctcaccgtctttcattgccatacg

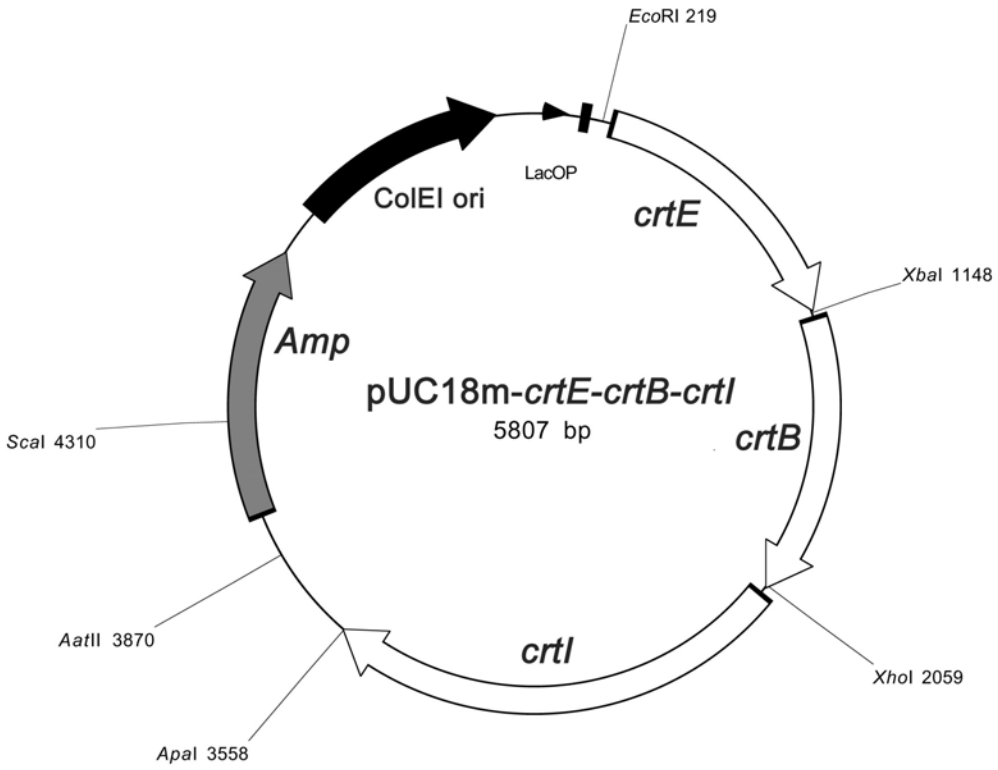
Plasmid name pUC18m-*crtE-crtB-crtI*; 5807 bp

Based on vector pUC18m.

Construction

All 3 genes are on a single operon, regulated by a single *lac* promoter and operator. Just upstream of each gene is an optimized Shine-Dalgarno site (AGGAGG) followed by 8 spacer nucleotides. The gene *crtE* from *Erwinia uredovora*, (GGPP synthase) is inserted between the *Eco*RI and *Xba*I sites. The gene *crtB* from *Erwinia uredovora* (C_{40} carotenoid synthase) is inserted between the *Xba*I and *Xho*I sites. The gene *crtI* from *Erwinia uredovora* (C_{40} carotenoid desaturase) is inserted between the *Xho*I and *Apa*I sites.

Plasmid map



---

pUC18m-*crtE-crtB-crtI* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaa
cgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcacaca
ggaaacaCATATGGAATTCaggaggattacaaaatgacggtctgcgcaaaaaaacacgttcatctcactcgcgatgctgcggagcagttactg
gctgatattgatcgacgccttgatcagttattgcccgtggagggagaacgggatgttgtgggtgccgcgatgcgtgaaggtgcgctggcaccgggaaaac
gtattcgccccatgttgctgttgctgaccgcccgcgatctgggttgcgctgtcagccatgacggattactggatttggcctgtgcggtggaaatggtccacgc
ggcttcgctgatccttgacgatatgccctgcatggacgatgcgaagctgcggcgcggacgccctaccattcattctcattacggagagcatgtggcaatact
ggcggcggttgccttgctgagtaaagcctttggcgtaattgccgatgcagatggcctcacgccgctggcaaaaaatcgggcggtttctgaactgtcaaacg
ccatcggcatgcaaggattggttcagggtcagttcaaggatctgtctgaaggggataagccgcgcagcgctgaagctattttgatgacgaatcactttaaaa
ccagcacgctgttttgtgcctccatgcagatggcctcgattgttgcgaatgcctccagcgaagcgcgtgattgcctgcatcgtttttcacttgatcttggtcagg

catttcaactgctggacgatttgaccgatggcatgaccgacaccggtaaggatagcaatcaggacgccggtaaatcgacgctggtcaatctgttaggcccg
agggcggttgaagaacgtctgagacaacatcttcagcttgccagtgagcatctctctgcggcctgccaacacgggcacgccactcaacattttattcaggcc
tggtttgacaaaaaactcgctgccgtcagttaaTCTAGAaggaggattacaaaatggcagttggctcgaaaagttttgcgacagcctcaaagttatttga
tgcaaaaacccggcgcagcgtactgatgctctacgcctggtgccgccattgtgacgatgttattgacgatcagacgctgggctttcaggcccggcagcctg
ccttacaaacgcccgaacaacgtctgatgcaacttgagatgaaaacgcgccaggcctatgcaggatcgcagatgcacgaaccggcgtttgcggcttttcag
gaagtggctatggctcatgatatcgccccggcttacgcgtttgatcatctggaaggcttcgccatggatgtacgcgaagcgcaatacagccaactggatgat
acgctgcgctattgctatcacgttgcaggcgttgtcggcttgatgatggcgcaaatcatgggcgtgcgggataacgccacgctggaccgcgcctgtgacctt
gggctggcatttcagttgaccaatattgctcgcgatattgtggacgatgcgcatgcgggccgctgttatctgccggcaagctggctggagcatgaaggtctg
aacaaagagaattatgcggcacctgaaaaccgtcaggcgctgagccgtatcgcccgtcgtttggtgcaggaagcagaaccttactatttgtctgccacagc
cggcctggcagggttgcccctgcgttccgcctgggcaatcgctacggcgaagcaggtttaccggaaaataggtgtcaaagttgaacaggccggtcagca
agcctgggatcagcggcagtcaacgaccacgcccgaaaaattaacgctgctgctggccgcctctggtcaggcccttacttcccggatgcgggctcatcct
ccccgccctgcgcatctctggcagcgcccgctctagCTCGAGaggaggattacaaaatgaaaccaactacggtaattggtgcaggcttcggtggcct
ggcactggcaattcgtctacaagctgcggggatccccgtcttactgcttgaacaacgtgataaacccggcggtcgggcttatgtctacgaggatcaggggtt
tacctttgatgcaggcccgacggttatcaccgatcccagtgccattgaagaactgtttgcactggcaggaaaacagttaaaagagtatgtcgaactgctgcc
ggttacgccgttttaccgcctgtgttgggagtcagggaaggtctttaattacgataacgatcaaacccggctcgaagcgcagattcagcagtttaatccccgc
gatgtcgaaggttatcgtcagtttctggactattcacgcgcggtgtttaaagaaggctatctaaagctcggtactgtccctttttatcgttcagagacatgcttcg
cgccgcacctcaactggcgaaactgcaggcatggagaagcgtttacagtaaggttgccagttacatcgaagatgaacatctgcgccaggcgttttctttcca
ctcgctgttggtgggcggcaatcccttcgccacctcatccatttatacgttgatacacgcgctggagcgtgagtggggcgtctggtttccgcgtggcggcac
cggcgcattagttcaggggatgataaagctgtttcaggatctgggtggcgaagtcgtgttaaacgccagagtcagccatatggaaacgacaggaaacaag
attgaagccgtgcatttagaggacggtcgcaggttcctgacgcaagccgtcgcgtcaaatgcagatgtggttcatacctatcgcgacctgttaagccagcac
cctgccgcggttaagcagtccaacaaactgcagactaagcgcatgagtaactctctgtttgtgctctattttggtttgaatcaccatcatgatcagctcgcgcat
cacacggtttgtttcggcccgcgttaccgcgagctgattgacgaaatttttaatcatgatggcctcgcagaggacttctcactttatctgcacgcgccctgtgtc
acggattcgtcactggcgcctgaaggttgcggcagttactatgtgttggcgccggtgccAcatttaggcaccgcgaacctcgactggacggttgaggggc
caaaactacgcgaccgtatttttgcgtaccttgagcagcattacatgcctggcttacggagtcagctggtcacgcaccggatgtttacgccgtttgattttcgcg
accagcttaatgcctatcatggctcagccttttctgtggagcccgttcttacccagagcgcctggtttcggccgcataaccgcgataaaaccattactaatctct
acctggtcggcgcaggcacgcatcccggcgcaggcattcctggcgtcatcggctcggcaaaagcgacagcaggtttgatgctggaggatctgatttgaG
GGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcat
agttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgg
gagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatgg
tttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccct
gataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcaccca
gaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgc
cccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatac
actattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagt
gataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgtt
gggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaac
tacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgct
gataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggag
tcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagat
tgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagac
cccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccg
gatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaa
gaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacag
cgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgaggga
gcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctat
ggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtatta
ccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

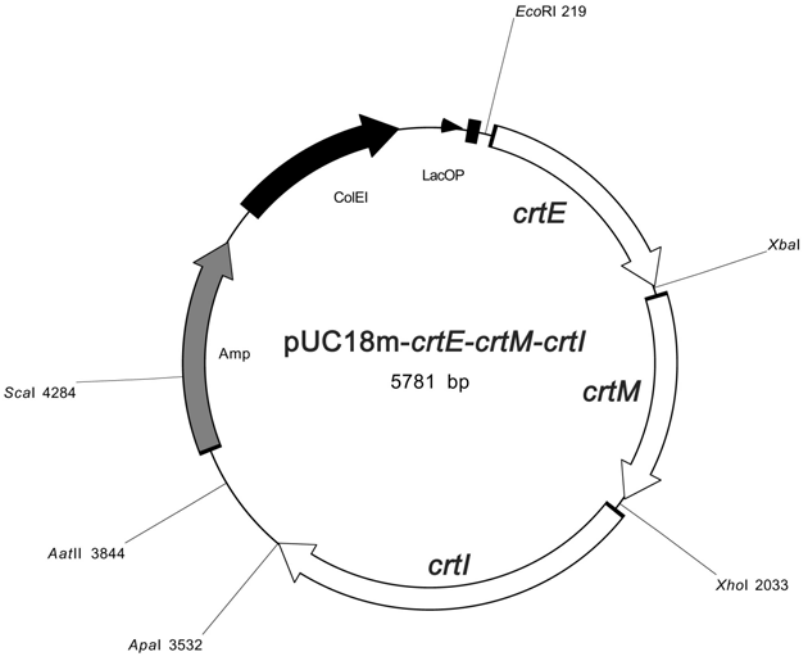
Plasmid name pUC18m-*crtE-crtM-crtI*; 5781 bp

Based on vector pUC18m

Construction

All 3 genes are on a single operon, regulated by a single *lac* promoter and operator. Just upstream of each gene is an optimized Shine-Dalgarno site (AGGAGG) followed by 8 spacer nucleotides. The gene *crtE* from *Erwinia uredovora* encoding a GGPP synthase is inserted between the *Eco*RI and *Xba*I sites. The gene *crtM* from *Staphylococcus aureus* encoding a C_{30} carotenoid synthase is inserted between the *Xba*I and *Xho*I sites. The gene *crtI* from *Erwinia uredovora* encoding a C_{40} carotenoid desaturase (4-step) is inserted between the *Xho*I and *Apa*I sites.

Plasmid map



pUC18m-*crtE-crtM-crtI* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagt
gagcgcaacgcaattaatgtgagttagctcactcattaggcacccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagc
ggataacaatttcacacaggaaacaCATATGGAATTCaggaggattacaaaatgacggtctgcgcaaaaaaacacgttcatctca
ctcgcgatgctgcggagcagttactggctgatattgatcgacgccttgatcagttattgcccgtggagggagaacgggatgttgtgggtgccg
cgatgcgtgaaggtgcgctggcaccgggaaaacgtattcgccccatgttgctgttgctgaccgcccgcgatctgggttgcgctgtcagccat
gacggattactggatttggcctgtgcggtggaaatggtccacgcggcttcgctgatccttgacgatatgccctgcatggacgatgcgaagctg
cggcgcggacgccctaccattcattctcattacggagagcatgtggcaatactggcggcggttgccttgctgagtaaagcctttggcgtaatt
gccgatgcagatggcctcacgccgctggcaaaaaatcgggcggtttctgaactgtcaaacgccatcggcatgcaaggattggttcagggtc
agttcaaggatctgtctgaaggggataagccgcgcagcgctgaagctattttgatgacgaatcactttaaaaccagcacgctgttttgtgcctc
catgcagatggcctcgattgttgcgaatgcctccagcgaagcgcgtgattgcctgcatcgtttttcacttgatcttggtcaggcatttcaactgct
ggacgatttgaccgatggcatgaccgacaccggtaaggatagcaatcaggacgccggtaaatcgacgctggtcaatctgttaggcccgag
ggcggttgaagaacgtctgagacaacatcttcagcttgccagtgagcatctctctgcggcctgccaacacgggcacgccactcaacattttat

tcaggcctggtttgacaaaaaactcgctgccgtcagttaaTCTAGAaggaggattacaaaatgacaatgatgaatatgaatttttaaatatt
gtcataaaatcatgaagaaacattcaaaaagcttttcttacgcttttgacttgttaccagaagatcaaagaaaagcggtttgggcaatttatgctgt
gtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaatttttaaatcaaataaaagaagatatacaatctattgaaaaatacccatatg
aacatcatcactttcaaagtgatcgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatatcgcctttcaatctttttataatctcattg
atactgtatataaagatcaacattttacaatgtttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattg
acgccgattttaagtgatcatgaaacacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgt
cggtgaagattttgacaatgaacggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaa
taatcattatattgacttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagcac
aaccaatcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgtggata
agaggaaaaaggcaaagttgtttcatgaaaataaatagtaaatatcatagaatatagCTCGAGaggaggattacaaaatgaaaccaact
acggtaattggtgcaggcttcggtggcctggcactggcaattcgtctacaagctgcggggatccccgtcttactgcttgaacaacgtgataaa
cccggcggtcgggcttatgtctacgaggatcaggggtttacctttgatgcaggcccgacggttatcaccgatcccagtgccattgaagaactg
tttgcactggcaggaaaacagttaaaagagtatgtcgaactgctgccggttacgccgttttaccgcctgtgttgggagtcagggaaggtcttta
attacgataacgatcaaacccggctcgaagcgcagattcagcagtttaatccccgcgatgtcgaaggttatcgtcagtttctggactattcacg
cgcggtgtttaaagaaggctatctaaagctcggtactgtccctttttatcgttcagagacatgcttcgcgccgcacctcaactggcgaaactgc
aggcatggagaagcgtttacagtaaggttgccagttacatcgaagatgaacatctgcgccaggcgttttctttccactcgctgttggtgggcgg
caatcccttcgccacctcatccatttatacgttgatacacgcgctggagcgtgagtggggcgtctggtttccgcgtggcggcaccggcgcatt
agttcaggggatgataaagctgtttcaggatctgggtggcgaagtcgtgttaaacgccagagtcagccatatggaaacgacaggaaacaag
attgaagccgtgcatttagaggacggtcgcaggttcctgacgcaagccgtcgcgtcaaatgcagatgtggttcatacctatcgcgacctgtta
agccagcaccctgccgcggttaagcagtccaacaaactgcagactaagcgcatgagtaactctctgtttgtgctctattttggtttgaatcacca
tcatgatcagctcgcgcatcacacggtttgtttcggcccgcgttaccgcgagctgattgacgaaatttttaatcatgatggcctcgcagaggact
tctcactttatctgcacgcgccctgtgtcacggattcgtcactggcgcctgaaggttgcggcagttactatgtgttggcgccggtgccAcattt
aggcaccgcgaacctcgactggacggttgaggggccaaaactacgcgaccgtatttttgcgtaccttgagcagcattacatgcctggcttac
ggagtcagctggtcacgcaccggatgtttacgccgtttgattttcgcgaccagcttaatgcctatcatggctcagccttttctgtggagcccgttc
ttacccagagcgcctggtttcggccgcataaccgcgataaaaccattactaatctctacctggtcggcgcaggcacgcatcccggcgcagg
cattcctggcgtcatcggctcggcaaaagcgacagcaggtttgatgctggaggatctgatttgaGGGCCCGGCGCCtgatgcggt
attttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccg
acacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctg
catgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataa
tggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatga
gacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcatttt
gccttcctgttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatct
caacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatccc
gtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatctta
cggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggagga
ccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagc
cggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcag
gcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatact
ttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttcc
actgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccg
ctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtcc
ttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgcc
agtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgca
cacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggaga
aaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttata
gtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcg
gccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtg
agctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*crtE-M*$_8$-*crtI*; 5781 bp

Based on vector pUC18m

Construction
This plasmid is identical to pUC18m-*crtE-crtM-crtI* except in the place of wild-type *crtM* is M_8, a mutant of the C_{30} carotenoid synthase *crtM* from *Staphylococcus aureus* that is capable of synthesizing C_{40} carotenoids.

Plasmid map
See pUC18m-*crtE-crtM-crtI* plasmid map.

Mutations
Point mutations in M_8 (compared with wild-type *crtM*) are shown below in the plasmid sequence in bold, underlined capital letters and are summarized below (numbered according to their position in the *crtM* ORF; amino acid substitutions are listed in brackets).
T78A (F26L)
A345G (silent)

pUC-*crtE-M*$_8$-*crtI* sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggattacaaaatgacggtct
gcgcaaaaaaacacgttcatctcactcgcgatgctgcggagcagttactggctgatattgatcgacgccttgatcagttattgccc
gtggagggagaacgggatgttgtgggtgccgcgatgcgtgaaggtgcgctggcaccgggaaaacgtattcgccccatgttgct
gttgctgaccgcccgcgatctgggttgcgctgtcagccatgacggattactggatttggcctgtgcggtggaaatggtccacgcg
gcttcgctgatccttgacgatatgccctgcatggacgatgcgaagctgcggcgcggacgccctaccattcattctcattacggag
agcatgtggcaatactggcggcggttgccttgctgagtaaagcctttggcgtaattgccgatgcagatggcctcacgccgctgg
caaaaaatcgggcggtttctgaactgtcaaacgccatcggcatgcaaggattggttcagggtcagttcaaggatctgtctgaagg
ggataagccgcgcagcgctgaagctattttgatgacgaatcactttaaaaccagcacgctgttttgtgcctccatgcagatggcct
cgattgttgcgaatgcctccagcgaagcgcgtgattgcctgcatcgtttttcacttgatcttggtcaggcatttcaactgctggacga
tttgaccgatggcatgaccgacaccggtaaggatagcaatcaggacgccggtaaatcgacgctggtcaatctgttaggcccga
gggcggttgaagaacgtctgagacaacatcttcagcttgccagtgagcatctctctgcggcctgccaacacgggcacgccactc
aacattttattcaggcctggtttgacaaaaaactcgctgccgtcagttaaTCTAGAaggaggattacaaaatgacaatgatga
atatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagctttcttacgctt**A**gacttgttaccagaagatcaaaga
aaagcggtttgggcaatttatgctgtgtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaatttttaaatcaaataaaa
gaagatatacaatctattgaaaaatacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgcttcagcatgttg
cacaacataaaaatatcgcctttcaatctttttataatctcattgatactgtatataaagatca**G**cattttacaatgtttgaaacggacg
ctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacgccgattttaagtgatcatgaaacacatcagaca
tacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaagattttgacaatgaacggatat
attttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcattatattgacttatggg
aatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagcacaaccaatcataga
attagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgtggataagag
gaaaaaggcaaagttgtttcatgaaaataaatagtaaatatcatagaatatagCTCGAGaggaggattacaaaatgaaacca
actacggtaattggtgcaggcttcggtggcctggcactggcaattcgtctacaagctgcggggatccccgtcttactgcttgaaca

acgtgataaacccggcggtcgggcttatgtctacgaggatcaggggtttacctttgatgcaggcccgacggttatcaccgatccc
agtgccattgaagaactgtttgcactggcaggaaaacagttaaaagagtatgtcgaactgctgccggttacgccgttttaccgcct
gtgttgggagtcagggaaggtctttaattacgataacgatcaaacccggctcgaagcgcagattcagcagtttaatccccgcgat
gtcgaaggttatcgtcagtttctggactattcacgcgcggtgtttaaagaaggctatctaaagctcggtactgtccctttttatcgttc
agagacatgcttcgcgccgcacctcaactggcgaaactgcaggcatggagaagcgtttacagtaaggttgccagttacatcgaa
gatgaacatctgcgccaggcgttttctttccactcgctgttggtgggcggcaatcccttcgccacctcatccatttatacgttgatac
acgcgctggagcgtgagtggggcgtctggtttccgcgtggcggcaccggcgcattagttcaggggatgataaagctgtttcag
gatctgggtggcgaagtcgtgttaaacgccagagtcagccatatggaaacgacaggaaacaagattgaagccgtgcatttaga
ggacggtcgcaggttcctgacgcaagccgtcgcgtcaaatgcagatgtggttcatacctatcgcgacctgttaagccagcaccc
tgccgcggttaagcagtccaacaaactgcagactaagcgcatgagtaactctctgtttgtgctctattttggtttgaatcaccatcat
gatcagctcgcgcatcacacggtttgtttcggcccgcgttaccgcgagctgattgacgaaatttttaatcatgatggcctcgcaga
ggacttctcactttatctgcacgcgccctgtgtcacggattcgtcactggcgcctgaaggttgcggcagttactatgtgttggcgcc
ggtgccAcatttaggcaccgcgaacctcgactggacggttgaggggccaaaactacgcgaccgtatttttgcgtaccttgagca
gcattacatgcctggcttacggagtcagctggtcacgcaccggatgtttacgccgtttgattttcgcgaccagcttaatgcctatcat
ggctcagccttttctgtggagcccgttcttacccagagcgcctggtttcggccgcataaccgcgataaaaccattactaatctctac
ctggtcggcgcaggcacgcatcccggcgcaggcattcctggcgtcatcggctcggcaaaagcgacagcaggtttgatgctgg
aggatctgatttgaGGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatgg
tgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacg
ggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcac
cgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtg
gcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccct
gataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgcct
tcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactgg
atctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgc
ggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccag
tcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgt
tgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgca
aactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccactt
ctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcact
ggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacaga
tcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcattt
ttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga
ccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctacc
agcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactg
tccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagt
ggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggct
gaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgaga
aagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagg
gagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgt
caggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttc
tttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccga
gcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*crtE-M*$_9$-*crtI*; 5781 bp

Based on vector pUC18m

Construction
This plasmid is identical to pUC18m-*crtE-crtM-crtI* except in the place of wild-type *crtM* is *M*$_9$, a mutant of the C_{30} carotenoid synthase *crtM* from *Staphylococcus aureus* that is capable of synthesizing C_{40} carotenoids.

Plasmid map
See pUC18m-*crtE-crtM-crtI* plasmid map.

Mutations
Point mutations in *M*$_9$ (compared with wild-type *crtM*) are shown below in the plasmid sequence in bold, underlined capital letters and are summarized below (numbered according to their position in the *crtM* ORF; amino acid substitutions are listed in brackets).
T77C (F26S)
T119C (I40T)
T135C (silent)
T141C (silent)
A850G (after stop codon)

pUC-*crtE-M*$_9$-*crtI* sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggattacaaaatgacggtct
gcgcaaaaaaacacgttcatctcactcgcgatgctgcggagcagttactggctgatattgatcgacgccttgatcagttattgccc
gtggagggagaacgggatgttgtgggtgccgcgatgcgtgaaggtgcgctggcaccgggaaaacgtattcgccccatgttgct
gttgctgaccgcccgcgatctgggttgcgctgtcagccatgacggattactggatttggcctgtgcggtggaaatggtccacgcg
gcttcgctgatccttgacgatatgccctgcatggacgatgcgaagctgcggcgcggacgccctaccattcattctcattacggag
agcatgtggcaatactggcggcggttgccttgctgagtaaagcctttggcgtaattgccgatgcagatggcctcacgccgctgg
caaaaaatcgggcggtttctgaactgtcaaacgccatcggcatgcaaggattggttcagggtcagttcaaggatctgtctgaagg
ggataagccgcgcagcgctgaagctattttgatgacgaatcactttaaaaccagcacgctgttttgtgcctccatgcagatggcct
cgattgttgcgaatgcctccagcgaagcgcgtgattgcctgcatcgtttttcacttgatcttggtcaggcatttcaactgctggacga
tttgaccgatggcatgaccgacaccggtaaggatagcaatcaggacgccggtaaatcgacgctggtcaatctgttaggcccga
gggcggttgaagaacgtctgagacaacatcttcagcttgccagtgagcatctctctgcggcctgccaacacgggcacgccactc
aacattttattcaggcctggtttgacaaaaaactcgctgccgtcagttaaTCTAGAaggaggattacaaaatgacaatgatga
atatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagcttttcttacgctt**C**tgacttgttaccagaagatcaaaga
aaagcggtttgggcaa**C**ttatgctgtgtgtcg**C**aaaat**C**gatgacagtatagatgtttatggcgatattcaatttttaaatcaaata
aaagaagatatacaatctattgaaaaatacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgcttcagcat
gttgcacaacataaaaatatcgcctttcaatctttttataatctcattgatactgtatataaagatcaacattttacaatgtttgaaacgga
cgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacgccgattttaagtgatcatgaaacacatcag
acatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaagattttgacaatgaacgg
atatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcattatattgacttat
gggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagcacaaccaatcat

agaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgtggataa
gaggaaaaaggcaaagttgtttcatgaaaataaatagtaaGtatcatagaatatagCTCGAGaggaggattacaaaatgaa
accaactacggtaattggtgcaggcttcggtggcctggcactggcaattcgtctacaagctgcggggatccccgtcttactgcttg
aacaacgtgataaacccggcggtcgggcttatgtctacgaggatcaggggtttacctttgatgcaggcccgacggttatcaccg
atcccagtgccattgaagaactgtttgcactggcaggaaaacagttaaaagagtatgtcgaactgctgccggttacgccgttttac
cgcctgtgttgggagtcagggaaggtctttaattacgataacgatcaaacccggctcgaagcgcagattcagcagtttaatcccc
gcgatgtcgaaggttatcgtcagtttctggactattcacgcgcggtgtttaaagaaggctatctaaagctcggtactgtccctttttta
tcgttcagagacatgcttcgcgccgcacctcaactggcgaaactgcaggcatggagaagcgtttacagtaaggttgccagttac
atcgaagatgaacatctgcgccaggcgttttctttccactcgctgttggtgggcggcaatcccttcgccacctcatccatttatacgt
tgatacacgcgctggagcgtgagtggggcgtctggtttccgcgtggcggcaccggcgcattagttcaggggatgataaagctg
tttcaggatctgggtggcgaagtcgtgttaaacgccagagtcagccatatggaaacgacaggaaacaagattgaagccgtgcat
ttagaggacggtcgcaggttcctgacgcaagccgtcgcgtcaaatgcagatgtggttcatacctatcgcgacctgttaagccagc
accctgccgcggttaagcagtccaacaaactgcagactaagcgcatgagtaactctctgtttgtgctctattttggtttgaatcacca
tcatgatcagctcgcgcatcacacggtttgtttcggcccgcgttaccgcgagctgattgacgaaatttttaatcatgatggcctcgc
agaggacttctcactttatctgcacgcgccctgtgtcacggattcgtcactggcgcctgaaggttgcggcagttactatgtgttggc
gccggtgccAcatttaggcaccgcgaacctcgactggacggttgaggggccaaaactacgcgaccgtatttttgcgtaccttga
gcagcattacatgcctggcttacggagtcagctggtcacgcaccggatgtttacgccgtttgattttcgcgaccagcttaatgccta
tcatggctcagccttttctgtggagcccgttcttacccagagcgcctggtttcggccgcataaccgcgataaaaccattactaatct
ctacctggtcggcgcaggcacgcatcccggcgcaggcattcctggcgtcatcggctcggcaaaagcgacagcaggtttgatg
ctggaggatctgatttgaGGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatat
atggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctg
acgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcat
caccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcag
gtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataac
cctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttg
ccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaac
tggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtgg
cgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcac
cagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcgg
ccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgat
cgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgc
gcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggacc
acttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcag
cactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaataga
cagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaactt
cattttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgt
cagaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgc
taccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaat
actgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttac
cagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctat
gagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcac
gagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatg
ctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcaca
tgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgac
cgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*crtE-M*$_{10}$*-crtI*; 5781 bp

Based on vector pUC18m

Construction
This plasmid is identical to pUC18m-*crtE-crtM-crtI* except in the place of wild-type *crtM* is M_{10}, a mutant of the C_{30} carotenoid synthase *crtM* from *Staphylococcus aureus* that is capable of synthesizing C_{40} carotenoids.

Plasmid map
See pUC18m-*crtE-crtM-crtI* plasmid map.

Mutations
Point mutations in M_{10} (compared with wild-type *crtM*) are shown below in the plasmid sequence in bold, underlined capital letters and are summarized below (numbered according to their position in the *crtM* ORF; amino acid substitutions are listed in brackets).
A10G (M4V)
A35G (H12R)
A39G (silent)
T176C (F59S)
A242G (Q81R)
A539G (E180G)

pUC-*crtE-M*$_{10}$*-crtI* sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcggg
cagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga
attgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggattacaaaatgacggtctgcgcaaaaaa
acacgttcatctcactcgcgatgctgcggagcagttactggctgatattgatcgacgccttgatcagttattgcccgtggagggagaac
gggatgttgtgggtgccgcgatgcgtgaaggtgcgctggcaccgggaaaacgtattcgccccatgttgctgttgctgaccgcccgcg
atctgggttgcgctgtcagccatgacggattactggatttggcctgtgcggtggaaatggtccacgcggcttcgctgatccttgacgata
tgccctgcatggacgatgcgaagctgcggcgcggacgccctaccattcattctcattacggagagcatgtggcaatactggcggcgg
ttgccttgctgagtaaagcctttggcgtaattgccgatgcagatggcctcacgccgctggcaaaaaatcgggcggtttctgaactgtca
aacgccatcggcatgcaaggattggttcagggtcagttcaaggatctgtctgaaggggataagccgcgcagcgctgaagctattttga
tgacgaatcactttaaaaccagcacgctgttttgtgcctccatgcagatggcctcgattgttgcgaatgcctccagcgaagcgcgtgatt
gcctgcatcgtttttcacttgatcttggtcaggcatttcaactgctggacgatttgaccgatggcatgaccgacaccggtaaggatagca
atcaggacgccggtaaatcgacgctggtcaatctgttaggcccgagggcggttgaagaacgtctgagacaacatcttcagcttgcca
gtgagcatctctctgcggcctgccaacacgggcacgccactcaacattttattcaggcctggtttgacaaaaaactcgctgccgtcagtt
aaTCTAGAaggaggattacaaaatgacaatg**G**tgaatatgaattttaaatattgtc**G**taa**G**atcatgaagaaacattcaaaaagc
ttttcttacgcttttgacttgttaccagaagatcaaagaaaagcggtttgggcaatttatgctgtgtgtcgtaaaattgatgacagtatagatg
tttatggcgatattcaat**C**tttaaatcaaataaaagaagatatacaatctattgaaaaatacccatatgaacatcatcactttc**G**aagtgat
cgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatatcgccttttcaatctttttataatctcattgatactgtatataaagatc
aacattttacaatgtttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacgccgattttaa
gtgatcatgaaacacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaag
attttgacaatg**G**acggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataat
cattatattgacttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagca
caaccaatcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgt

ggataagaggaaaaaggcaaagttgtttcatgaaaataaatagtaaatatcatagaatatagCTCGAGaggaggattacaaaatg
aaaccaactacggtaattggtgcaggcttcggtggcctggcactggcaattcgtctacaagctgcggggatccccgtcttactgcttga
acaacgtgataaacccggcggtcgggcttatgtctacgaggatcaggggtttacctttgatgcaggcccgacggttatcaccgatccc
agtgccattgaagaactgtttgcactggcaggaaaacagttaaaagagtatgtcgaactgctgccggttacgccgttttaccgcctgtgt
tgggagtcagggaaggtctttaattacgataacgatcaaacccggctcgaagcgcagattcagcagtttaatccccgcgatgtcgaag
gttatcgtcagtttctggactattcacgcgcggtgtttaaagaaggctatctaaagctcggtactgtccctttttatcgttcagagacatgct
tcgcgccgcacctcaactggcgaaactgcaggcatggagaagcgtttacagtaaggttgccagttacatcgaagatgaacatctgcg
ccaggcgttttctttccactcgctgttggtgggcggcaatcccttcgccacctcatccatttatacgttgatacacgcgctggagcgtgag
tggggcgtctggtttccgcgtggcggcaccggcgcattagttcaggggatgataaagctgtttcaggatctgggtggcgaagtcgtgtt
aaacgccagagtcagccatatggaaacgacaggaaacaagattgaagccgtgcatttagaggacggtcgcaggttcctgacgcaag
ccgtcgcgtcaaatgcagatgtggttcatacctatcgcgacctgttaagccagcaccctgccgcggttaagcagtccaacaaactgca
gactaagcgcatgagtaactctctgtttgtgctctattttggtttgaatcaccatcatgatcagctcgcgcatcacacggtttgtttcggccc
gcgttaccgcgagctgattgacgaaatttttaatcatgatggcctcgcagaggacttctcactttatctgcacgcgccctgtgtcacggat
tcgtcactggcgcctgaaggttgcggcagttactatgtgttggcgccggtgccAcatttaggcaccgcgaacctcgactggacggttg
aggggccaaaactacgcgaccgtatttttgcgtaccttgagcagcattacatgcctggcttacggagtcagctggtcacgcaccggat
gtttacgccgtttgattttcgcgaccagcttaatgcctatcatggctcagccttttctgtggagcccgttcttacccagagcgcctggtttcg
gccgcataaccgcgataaaaccattactaatctctacctggtcggcgcaggcacgcatcccggcgcaggcattcctggcgtcatcgg
ctcggcaaaagcgacagcaggtttgatgctggaggatctgatttgaGGGCCCGGCGCCtgatgcggtattttctccttacgca
tctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgcca
acacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtc
agaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataatgtggt
ttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatg
agacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcg
gcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcg
aactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtgg
cgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccag
tcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaactta
cttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaacc
ggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactgg
cgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttc
cggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagcc
ctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactg
attaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatc
ctttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttg
agatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctacca
actcttttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaaga
actctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgg
actcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgac
ctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggta
agcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgcca
cctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcct
ggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgatacc
gctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

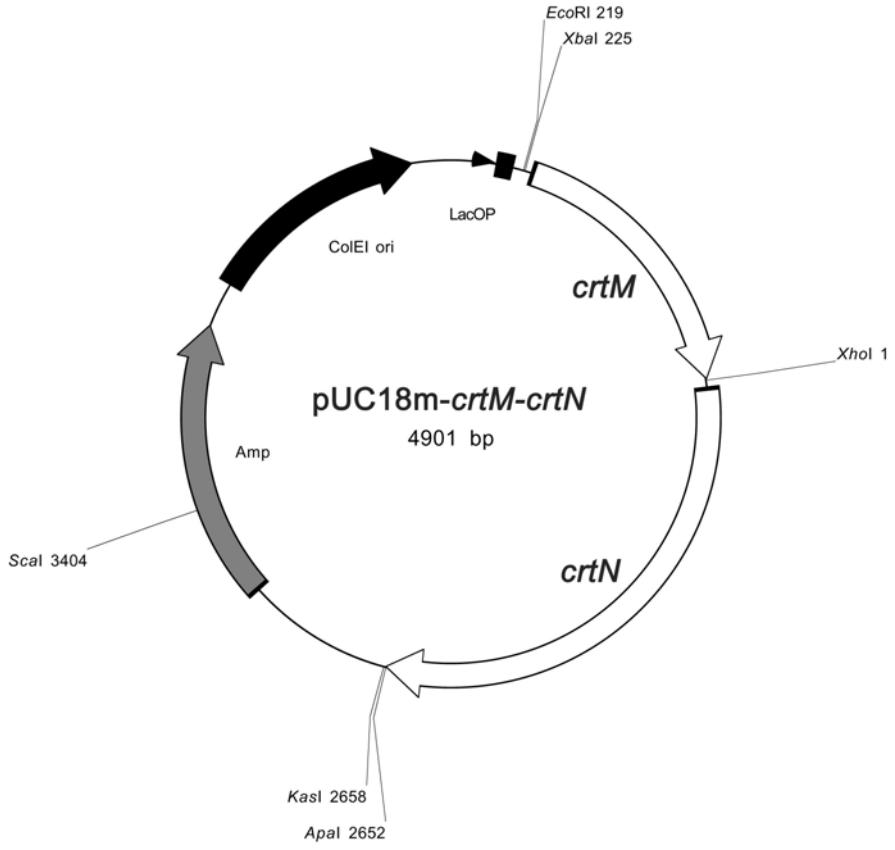
Plasmid name pUC18m-*crtM-crtN*; 4901 bp

Based on vector pUC18m

Construction

crtM, encoding the C_{30} carotenoid synthase from *Staphylococcus aureus* is inserted between the *Xba*I and *Xho*I sites. *crtN*, also from *S. aureus*, encoding a C_{30} carotenoid desaturase, is inserted between the *Xho*I and *Apa*I sites. Just upstream of each ORF is an optimized Shine-Dalgarno site (AGGAGG) followed by 8 spacer nucleotides. A single *lac* operator/promoter site regulates the entire operon, as shown in the plasmid map.

Plasmid map



pUC18m-*crtM-crtN* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcggg
cagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga
attgtgagcggataacaatttcacacaggaaacaCATATGGAATTCTCTAGAaggaggattacaaaatgacaatgat
gAatatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagcttttcttacgcttttgacttgttaccagaagatcaaagaaa
agcggtttgggcaatttatgctgtgtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaatttttaaAtcaaataaaagaaga
tatacaatctattgaaaaatacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgcttcagcatgttgcacaacata
aaaatatcgcctttcaatctttttataatctcattgatactgtatataaagatcaacattttacaatgtttgaaacggacgctgaattattcggat

attgttatggtgttgctggtacagtaggtgaagtattgacgccgattttaagtgatcatgaaacacatcagacatacgatgtcgcaagaag
acttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaagattttgacaatgaacggatatattttagtaagcaacgattaaa
gcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcattatattgacttatgggaatattatgcagctatcgcagaaaa
agattttcaagatgttatggatcaaatcaaagtatttagtattgaagcacaaccaatcatagaattagcagcacgtatatatattgaaatact
ggacgaagtgagacaggctaactatacattacatgaacgtgtttttgtggataagaggaaaaaggcaaagttgtttcatgaaaAtaaat
agtaaatatcatagaatatagCTCGAGaggaggattacaaaatgaagattgcagtaattggtgcaggtgtcacaggattagcagc
ggcagcccgtattgcttctcaaggtcatgaagtgacgatatttgaaaaaaataataatgtaggcgggcgtatgaatcaattaaagaaag
acggctttacatttgatatgggtcccacaattgtcatgatgccagatgtttataaagatgtttttacagcgtgtggtaaaaattatgaagatta
tattgaattgagacaattacgttatatttacgatgtgtattttgaccacgatgatcgtataacggtgcctacagatttagctgaattacagcaa
atgctagaaagtatagaacctggttcaacgcatggttttatgtccttttaacggatgtttataaaaaatatgaaattgcacgtcgctatttctt
agaaagaacgtatcgcaaaccgagtgacttttataatatgacgtcacttgtgcaaggtgctaagttaaaaacgttaaatcatgcagatca
gctaattgaacattatattgataacgaaaagatacaaaagcttttagcgtttcaaacgttatacataggaattgatccaaaacgaggcccg
tcactatattcaattattcctatgattgaaatgatgtttggtgtgcattttattaaaggcggtatgtatggcatggctcaagggctagcgcaat
taaataaagacttaggcgttaatattgaactaaatgctgaaattgagcaaattattattgatcctaaattcaaacgggccgatgcgataaaa
gtgaatggtgacataagaaaatttgataaaatttttatgtacggctgatttccctagtgttgcggaatcattaatgccagattttgcacctatta
aaaagtatccaccacataaaattgcagacttagattactcttgttcagcatttttaatgtatatcggtatagatattgatgtgacagatcaagt
gagacttcataatgttatttttttcagatgactttagaggcaatattgaagaaatatttgagggacgtttatcatatgatccttctatttatgtgtat
gtaccagcggtcgctgataaatcacttgcgccagaaggcaaaactggtatttatgtgctaatgccgacgccggaacttaaaacaggta
gcggaatcgattggtcagatgaagctttgacgcaacaaataaaggaaattatttatcgtaaattagcaacgattgaagtatttgaagatat
aaaatcgcatattgtttcagaaacaatctttacgccaaatgattttgagcaaacgtatcatgcgaaatttggttcggcattcggtttaatgcc
aactttagcgcaaagtaattattatcgtccacaaaatgtatcgcgagattataaagatttatattttgcaggtgcaagtacgcatccaggtg
caggcgttcctattgtcttaacgagtgcgaaaataactgtagatgaaatgattaaagatattgagcggggcgtataagggagtagtctaa
GGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatc
tgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatcc
gcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggc
ctcgtgatacgcctatttttataggttaatgtcatgataatgtggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacc
cctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaatatattgaaaaaggaaga
gtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagta
aaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccga
agaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcg
ccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatg
cagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgc
acaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacg
atgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatg
gaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtg
ggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactat
ggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttag
attgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttttgataatctcatgaccaaaatcccttaacgtgagttttcgttcc
actgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaa
atactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttacca
gtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctga
acgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcg
ccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttcca
gggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcgg
agcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccct
gattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagc
gaggaagcggaaga

Plasmid name pUC18m-*M_8-crtN*; 4901 bp

Based on vector pUC18m

Construction
This plasmid is identical to pUC18m-*crtM-crtN* except in the place of wild-type *crtM* is *M_8*, a mutant of the C_{30} carotenoid synthase *crtM* from *Staphylococcus aureus* that is capable of synthesizing C_{40} carotenoids.

Plasmid Map
See pUC18m-*crtM-crtN* plasmid map.

Mutations
Point mutations in *M_8* (compared with wild-type *crtM*) are shown below in the plasmid sequence in bold, underlined capital letters and are summarized below (numbered according to their position in the *crtM* ORF; amino acid substitutions are listed in brackets).
T78A (F26L)
A345G (silent)

pUC18m-*M_8-crtN* sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacacatatgGAATTCTCTAGAaggaggattacaaaatga
caatgatgaatatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagcttttcttacgctt**A**gacttgttaccagaag
atcaaagaaaagcggtttgggcaatttatgctgtgtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaatttttaaatc
aaataaaagaagatatacaatctattgaaaaatacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgcttca
gcatgttgcacaacataaaaatatcgcctttcaatctttttataatctcattgatactgtatataaagatca**G**cattttacaatgtttgaa
acggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacgccgattttaagtgatcatgaaacac
atcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaagattttgacaatga
acggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcattatattga
cttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagcacaacca
atcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgtgg
ataagaggaaaaaggcaaagttgtttcatgaaaataaatagtaaatatcatagaatatagCTCGAGaggaggattacaaaat
gaagattgcagtaattggtgcaggtgtcacaggattagcagcggcagcccgtattgcttctcaaggtcatgaagtgacgatatttg
aaaaaaataataatgtaggcgggcgtatgaatcaattaaagaaagacggctttacatttgatatgggtcccacaattgtcatgatgc
cagatgtttataaagatgtttttacagcgtgtggtaaaaattatgaagattatattgaattgagacaattacgttatatttacgatgtgtat
tttgaccacgatgatcgtataacggtgcctacagatttagctgaattacagcaaatgctagaaagtatagaacctggttcaacgcat
ggttttatgtccttttaacggatgtttataaaaaatatgaaattgcacgtcgctatttcttagaaagaacgtatcgcaaaccgagtgac
ttttataatatgacgtcacttgtgcaaggtgctaagttaaaaacgttaaatcatgcagatcagctaattgaacattatattgataacga
aaagatacaaaagcttttagcgtttcaaacgttatacataggaattgatccaaaacgaggcccgtcactatattcaattattcctatga
ttgaaatgatgtttggtgtgcattttattaaaggcggtatgtatggcatggctcaagggctagcgcaattaaataaagacttaggcgt
taatattgaactaaatgctgaaattgagcaaattattattgatcctaaattcaaacgggccgatgcgataaaagtgaatggtgacata
agaaaatttgataaaatttttatgtacggctgatttccctagtgttgcggaatcattaatgccagattttgcacctattaaaaagtatcca
ccacataaaattgcagacttagattactcttgttcagcatttttaatgtatatcggtatagatattgatgtgacagatcaagtgagactt
cataatgttatttttttcagatgactttagaggcaatattgaagaaatatttgagggacgtttatcatatgatccttctatttatgtgtatgta

ccagcggtcgctgataaatcacttgcgccagaaggcaaaactggtatttatgtgctaatgccgacgccggaacttaaaacaggta
gcggaatcgattggtcagatgaagctttgacgcaacaaataaaggaaattatttatcgtaaattagcaacgattgaagtatttgaag
atataaaatcgcatattgtttcagaaacaatctttacgccaaatgattttgagcaaacgtatcatgcgaaatttggttcggcattcggtt
taatgccaactttagcgcaaagtaattattatcgtccacaaaatgtatcgcgagattataaagatttatattttgcaggtgcaagtacg
catccaggtgcaggcgttcctattgtcttaacgagtgcgaaaataactgtagatgaaatgattaaagatattgagcggggcgtata
agggagtagtctaaGGGCCCggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtg
cactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacggg
cttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccg
aaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggc
acttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctga
taaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttc
ctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggat
ctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcg
gtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagt
cacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaa
cttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgtt
gggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgca
aactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccactt
ctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcact
ggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacaga
tcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcattt
ttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga
ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctacc
agcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactg
tccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagt
ggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggct
gaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgaga
aagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagg
gagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgt
caggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttc
tttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccga
gcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*M_9-crtN*; 4901 bp

Based on vector pUC18m

Construction
This plasmid is identical to pUC18m-*crtM-crtN* except in the place of wild-type *crtM* is *M_9*, a mutant of the C_{30} carotenoid synthase *crtM* from *Staphylococcus aureus* that is capable of synthesizing C_{40} carotenoids.

Plasmid Map
See pUC18m-*crtM-crtN* plasmid map.

Mutations
Point mutations in *M_9* (compared with wild-type *crtM*) are shown below in the plasmid sequence in bold, underlined capital letters and are summarized below (numbered according to their position in the *crtM* ORF; amino acid substitutions are listed in brackets).
T77C (F26S)
T119C (I40T)
T135C (silent)
T141C (silent)
A850G (after stop codon)

pUC18m-*M_9-crtN* sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacacatatgGAATTCTCTAGAaggaggattacaaaatga
caatgatgaatatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagcttttcttacgctt**C**tgacttgttaccagaag
atcaaagaaaagcggtttgggcaa**C**ttatgctgtgtgtcg**C**aaaat**C**gatgacagtatagatgtttatggcgatattcaatttttaa
atcaaataaaagaagatatacaatctattgaaaaatacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgct
tcagcatgttgcacaacataaaaatatcgcctttcaatctttttataatctcattgatactgtatataaagatcaacattttacaatgtttg
aaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacgccgattttaagtgatcatgaaac
acatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaagattttgacaat
gaacggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcattatatt
gacttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagcacaac
caatcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgt
ggataagaggaaaaaggcaaagttgtttcatgaaaataaatagtaa**G**tatcatagaatatagCTCGAGaggaggattacaa
aatgaagattgcagtaattggtgcaggtgtcacaggattagcagcggcagcccgtattgcttctcaaggtcatgaagtgacgatat
ttgaaaaaaataataatgtaggcgggcgtatgaatcaattaaagaaagacggctttacatttgatatgggtcccacaattgtcatga
tgccagatgtttataaagatgtttttacagcgtgtggtaaaaattatgaagattatattgaattgagacaattacgttatatttacgatgt
gtattttgaccacgatgatcgtataacggtgcctacagatttagctgaattacagcaaatgctagaaagtatagaacctggttcaac
gcatggttttatgtccttttaacggatgtttataaaaaatatgaaattgcacgtcgctatttcttagaaagaacgtatcgcaaaccgag
tgacttttataatatgacgtcacttgtgcaaggtgctaagttaaaaacgttaaatcatgcagatcagctaattgaacattatattgataa
cgaaaagatacaaaagcttttagcgtttcaaacgttatacataggaattgatccaaaacgaggcccgtcactatattcaattattcct
atgattgaaatgatgtttggtgtgcattttattaaaggcggtatgtatggcatggctcaagggctagcgcaattaaataaagacttag
gcgttaatattgaactaaatgctgaaattgagcaaattattattgatcctaaattcaaacgggccgatgcgataaaagtgaatggtg

acataagaaaatttgataaaattttatgtacggctgatttccctagtgttgcggaatcattaatgccagattttgcacctattaaaaagt
atccaccacataaaattgcagacttagattactcttgttcagcatttttaatgtatatcggtatagatattgatgtgacagatcaagtga
gacttcataatgttatttttttcagatgactttagaggcaatattgaagaaatatttgagggacgtttatcatatgatccttctatttatgtgt
atgtaccagcggtcgctgataaatcacttgcgccagaaggcaaaactggtatttatgtgctaatgccgacgccggaacttaaaac
aggtagcggaatcgattggtcagatgaagctttgacgcaacaaataaaggaaattatttatcgtaaattagcaacgattgaagtatt
tgaagatataaaatcgcatattgtttcagaaacaatctttacgccaaatgattttgagcaaacgtatcatgcgaaatttggttcggcat
tcggtttaatgccaactttagcgcaaagtaattattatcgtccacaaaatgtatcgcgagattataaagatttatattttgcaggtgcaa
gtacgcatccaggtgcaggcgttcctattgtcttaacgagtgcgaaaataactgtagatgaaatgattaaagatattgagcggggc
gtataagggagtagtctaaGGGCCCggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatat
ggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctga
cgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatc
accgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcagg
tggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataacc
ctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgc
cttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaact
ggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggc
gcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcacc
agtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggc
caacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatc
gttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcg
caaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggacca
cttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagc
actggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagac
agatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttc
atttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtc
agaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgct
accagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaata
ctgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttacc
agtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgg
gctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatg
agaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacg
agggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgct
cgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacat
gttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgac
cgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*M_{10}-crtN*; 4901 bp

Based on vector pUC18m

Construction
This plasmid is identical to pUC18m-*crtM-crtN* except in the place of wild-type *crtM* is *M_{10}*, a mutant of the C_{30} carotenoid synthase *crtM* from *Staphylococcus aureus* that is capable of synthesizing C_{40} carotenoids.

Plasmid Map
See pUC18m-*crtM-crtN* plasmid map.

Mutations
Point mutations in *M_{10}* (compared with wild-type *crtM*) are shown below in the plasmid sequence in bold, underlined capital letters and are summarized below (numbered according to their position in the *crtM* ORF; amino acid substitutions are listed in brackets).
A10G (M4V)
A35G (H12R)
A39G (silent)
T176C (F59S)
A242G (Q81R)
A539G (E180G)

pUC18m-*M_{10}-crtN* sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacacatatgGAATTCTCTAGAaggaggattacaaaatga
caatg**G**tgaatatgaattttaaatattgtc**G**taa**G**atcatgaagaaacattcaaaaagcttttcttacgcttttgacttgttaccaga
agatcaaagaaaagcggtttgggcaatttatgctgtgtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaat**C**ttta
aatcaaataaaagaagatatacaatctattgaaaaatacccatatgaacatcatcactttc**G**aagtgatcgtagaatcatgatggc
gcttcagcatgttgcacaacataaaaatatcgcctttcaatctttttataatctcattgatactgtatataaagatcaacattttacaatgt
ttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacgccgattttaagtgatcatgaa
acacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaagattttgac
aatg**G**acggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcatt
atattgacttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagca
caaccaatcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtt
tttgtggataagaggaaaaaggcaaagttgtttcatgaaaataaatagtaaatatcatagaatatagCTCGAGaggaggatta
caaaatgaagattgcagtaattggtgcaggtgtcacaggattagcagcggcagcccgtattgcttctcaaggtcatgaagtgacg
atatttgaaaaaaataataatgtaggcgggcgtatgaatcaattaaagaaagacggctttacatttgatatgggtcccacaattgtca
tgatgccagatgtttataaagatgtttttacagcgtgtggtaaaaattatgaagattatattgaattgagacaattacgttatatttacga
tgtgtattttgaccacgatgatcgtataacggtgcctacagatttagctgaattacagcaaatgctagaaagtatagaacctggttca
acgcatggttttatgtccttttaacggatgtttataaaaaatatgaaattgcacgtcgctatttcttagaaagaacgtatcgcaaaccg
agtgacttttataatatgacgtcacttgtgcaaggtgctaagttaaaaacgttaaatcatgcagatcagctaattgaacattatattgat
aacgaaaagatacaaaagcttttagcgtttcaaacgttatacataggaattgatccaaaacgaggcccgtcactatattcaattattc
ctatgattgaaatgatgtttggtgtgcattttattaaaggcggtatgtatggcatggctcaagggctagcgcaattaaataaagactt

aggcgttaatattgaactaaatgctgaaattgagcaaattattattgatcctaaattcaaacgggccgatgcgataaaagtgaatgg
tgacataagaaaatttgataaaattttatgtacggctgatttccctagtgttgcggaatcattaatgccagattttgcacctattaaaaa
gtatccaccacataaaattgcagacttagattactcttgttcagcatttttaatgtatatcggtatagatattgatgtgacagatcaagt
gagacttcataatgttatttttttcagatgactttagaggcaatattgaagaaatatttgagggacgtttatcatatgatccttctatttatgt
gtatgtaccagcggtcgctgataaatcacttgcgccagaaggcaaaactggtatttatgtgctaatgccgacgccggaacttaaa
acaggtagcggaatcgattggtcagatgaagctttgacgcaacaaataaaggaaattatttatcgtaaattagcaacgattgaagt
atttgaagatataaaatcgcatattgtttcagaaacaatctttacgccaaatgattttgagcaaacgtatcatgcgaaatttggttcgg
cattcggtttaatgccaactttagcgcaaagtaattattatcgtccacaaaatgtatcgcgagattataaagatttatattttgcaggtg
caagtacgcatccaggtgcaggcgttcctattgtcttaacgagtgcgaaaataactgtagatgaaatgattaaagatattgagcgg
ggcgtataagggagtagtctaaGGGCCCggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcat
atatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccc
tgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtc
atcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataatatggtttcttagacgtc
aggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaat
aaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcat
tttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcg
aactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatg
tggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtact
caccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactg
cggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgcct
tgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacg
ttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcagg
accacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattg
cagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaat
agacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaa
acttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactga
gcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacca
ccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc
aaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctg
ttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcgg
tcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagc
tatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgc
acgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtga
tgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctca
catgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacg
accgagcgcagcgagtcagtgagcgaggaagcggaaga

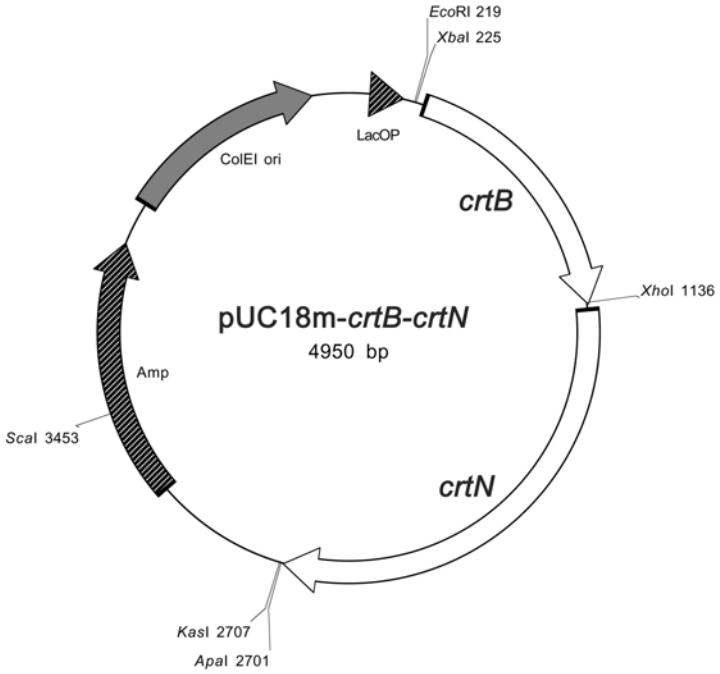
Plasmid name pUC18m-*crtB*-*crtN*; 4950 bp

Based on vector pUC18m

Construction

The gene *crtB* from *Erwinia uredovora* (C_{40} carotenoid synthase) is inserted between the *Xba*I and *Xho*I sites. The gene *crtN* from *Staphylococcus aureus* (C_{30} carotenoid desaturase) is inserted between the *Xho*I and *Apa*I sites. Just upstream of each ORF is an optimized Shine-Dalgarno site (AGGAGG) followed by 8 spacer nucleotides. A single *lac* operator/promoter site regulates the entire operon, as shown in the plasmid map. *E. coli* carrying this plasmid synthesize NO carotenoids, as CrtB accepts only GGPP as a substrate, and this is present only at low levels in *E. coli* that do not express an additional GGPP synthase.

Plasmid map



pUC18m-*crtB*-*crtN* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcggg
cagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga
attgtgagcggataacaatttcacacaggaaacaCATATGGAATTCTCTAGAaggaggattacaaaatggcagttgg
ctcgaaaagttttgcgacagcctcaaagttatttgatgcaaaaacccggcgcagcgtactgatgctctacgcctggtgccgccattgtg
acgatgttattgacgatcagacgctgggctttcaggcccggcagcctgccttacaaacgcccgaacaacgtctgatgcaacttgagat
gaaaacgcgccaggcctatgcaggatcgcagatgcacgaaccggcgtttgcggcttttcaggaagtggctatggctcatgatatcgc
cccggcttacgcgtttgatcatctggaaggcttcgccatggatgtacgcgaagcgcaatacagccaactggatgatacgctgcgctatt
gctatcacgttgcaggcgttgtcggcttgatgatggcgcaaatcatgggcgtgcgggataacgccacgctggaccgcgcctgtgacc
ttgggctggcatttcagttgaccaatattgctcgcgatattgtggacgatgcgcatgcgggccgctgttatctgccggcaagctggctgg
agcatgaaggtctgaacaaagagaattatgcggcacctgaaaaccgtcaggcgctgagccgtatcgcccgtcgtttggtgcaggaag

cagaaccttactatttgtctgccacagccggcctggcagggttgcccctgcgttccgcctgggcaatcgctacggcgaagcaggttta
ccggaaaataggtgtcaaagttgaacaggccggtcagcaagcctgggatcagcggcagtcaacgaccacgcccgaaaaattaacg
ctgctgctggccgcctctggtcaggcccttacttcccggatgcgggctcatcctccccgccctgcgcatctctggcagcgcccgctcta
gCTCGAGaggaggattacaaaatgaagattgcagtaattggtgcaggtgtcacaggattagcagcggcagcccgtattgcttctc
aaggtcatgaagtgacgatatttgaaaaaaataataatgtaggcgggcgtatgaatcaattaaagaaagacggctttacatttgatatgg
gtcccacaattgtcatgatgccagatgtttataaagatgtttttacagcgtgtggtaaaaattatgaagattatattgaattgagacaattacg
ttatatttacgatgtgtattttgaccacgatgatcgtataacggtgcctacagatttagctgaattacagcaaatgctagaaagtatagaacc
tggttcaacgcatggttttatgtccttttaacggatgtttataaaaaatatgaaattgcacgtcgctatttcttagaaagaacgtatcgcaaa
ccgagtgacttttataatatgacgtcacttgtgcaaggtgctaagttaaaaacgttaaatcatgcagatcagctaattgaacattatattgat
aacgaaaagatacaaaagcttttagcgtttcaaacgttatacataggaattgatccaaaacgaggcccgtcactatattcaattattcctat
gattgaaatgatgtttggtgtgcattttattaaaggcggtatgtatggcatggctcaagggctagcgcaattaaataaagacttaggcgtta
atattgaactaaatgctgaaattgagcaaattattattgatcctaaattcaaacgggccgatgcgataaaagtgaatggtgacataagaaa
atttgataaaattttatgtacggctgatttccctagtgttgcggaatcattaatgccagattttgcacctattaaaaagtatccaccacataaa
attgcagacttagattactcttgttcagcatttttaatgtatatcggtatagatattgatgtgacagatcaagtgagacttcataatgttatttttt
cagatgactttagaggcaatattgaagaaatatttgagggacgtttatcatatgatccttctatttatgtgtatgtaccagcggtcgctgata
aatcacttgcgccagaaggcaaaactggtatttatgtgctaatgccgacgccggaacttaaaacaggtagcggaatcgattggtcagat
gaagctttgacgcaacaaataaaggaaattatttatcgtaaattagcaacgattgaagtatttgaagatataaaatcgcatattgtttcaga
aacaatctttacgccaaatgattttgagcaaacgtatcatgcgaaatttggttcggcattcggtttaatgccaactttagcgcaaagtaatta
ttatcgtccacaaaatgtatcgcgagattataaagatttatattttgcaggtgcaagtacgcatccaggtgcagcgcttcctattgtcttaac
gagtgcgaaaataactgtagatgaaatgattaaagatattgagcgggcgtataagggagtagtctaagagaaagatgtgagaaaagta
taaGGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtaca
atctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggca
tccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaag
ggcctcgtgatacgcctatttttataggttaatgtcatgataatagtggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcgg
aacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaagga
agagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaa
gtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccc
cgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcgg
tcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaatt
atgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttt
tgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacca
cgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactgga
tggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgt
gggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaacta
tggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatacttta
gattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttttgataatctcatgaccaaaatcccttaacgtgagttttcgtt
ccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaa
ccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc
aaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttac
cagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggct
gaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaag
cgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttc
caggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggc
ggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatccc
ctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtga
gcgaggaagcggaaga

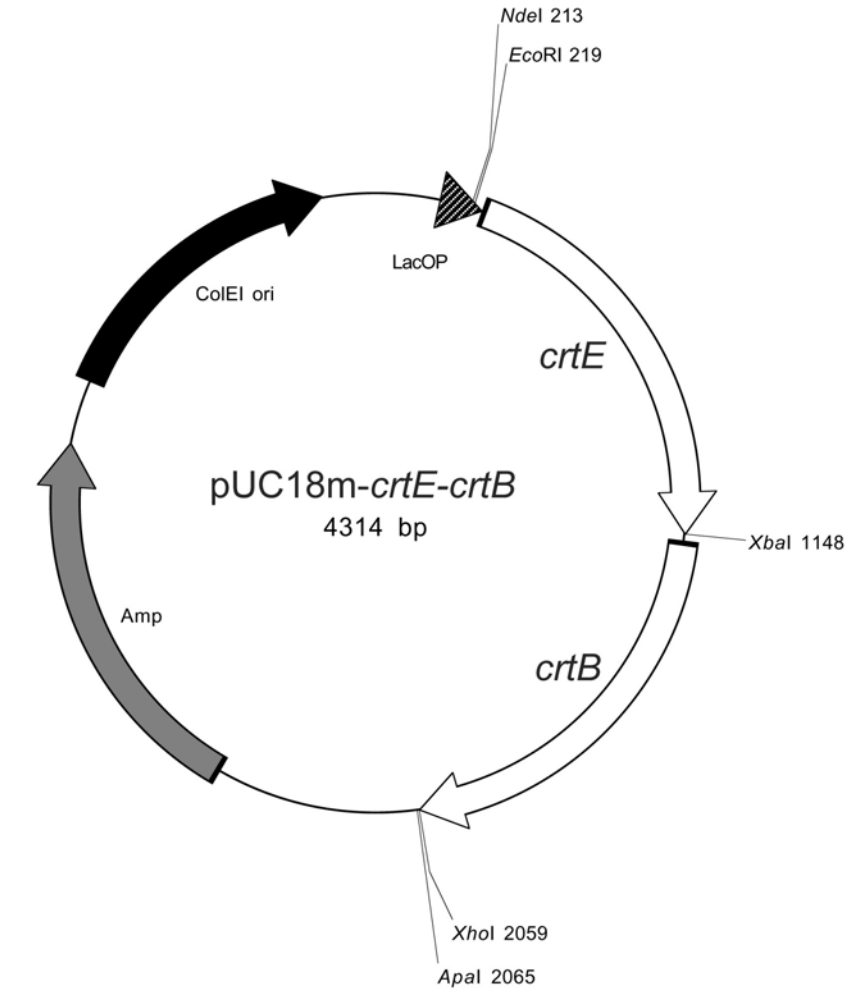
Plasmid name pUC18m-*crtE-crtB*; 4314 bp

Based on vector pUC18m

Construction

Both genes are on a single operon, regulated by a single *lac* promoter and operator. Just upstream of each gene is an optimized Shine-Dalgarno site (AGGAGG) followed by 8 spacer nucleotides. The gene *crtE* from *Erwinia uredovora* (GGPP synthase) is inserted between the *Eco*RI and *Xba*I sites. The gene *crtB* from *Erwinia uredovora* (C_{40} carotenoid synthase) is inserted between the *Xba*I and *Xho*I sites.

Plasmid map



pUC18m-*crtE-crtB* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcggg
cagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga
attgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggattacaaaatgacggtctgcgcaaaaaa
acacgttcatctcactcgcgatgctgcggagcagttactggctgatattgatcgacgccttgatcagttattgcccgtggagggagaac

gggatgttgtgggtgccgcgatgcgtgaaggtgcgctggcaccgggaaaacgtattcgccccatgttgctgttgctgaccgcccgcg
atctgggttgcgctgtcagccatgacggattactggatttggcctgtgcggtggaaatggtccacgcggcttcgctgatccttgacgata
tgccctgcatggacgatgcgaagctgcggcgcggacgccctaccattcattctcattacggagagcatgtggcaatactggcggcgg
ttgccttgctgagtaaagcctttggcgtaattgccgatgcagatggcctcacgccgctggcaaaaaatcgggcggtttctgaactgtca
aacgccatcggcatgcaaggattggttcagggtcagttcaaggatctgtctgaaggggataagccgcgcagcgctgaagctattttga
tgacgaatcactttaaaaccagcacgctgttttgtgcctccatgcagatggcctcgattgttgcgaatgcctccagcgaagcgcgtgatt
gcctgcatcgtttttcacttgatcttggtcaggcatttcaactgctggacgatttgaccgatggcatgaccgacaccggtaaggatagca
atcaggacgccggtaaatcgacgctggtcaatctgttaggcccgagggcggttgaagaacgtctgagacaacatcttcagcttgcca
gtgagcatctctctgcggcctgccaacacgggcacgccactcaacattttattcaggcctggtttgacaaaaaactcgctgccgtcagtt
aaTCTAGAaggaggattacaaaatggcagttggctcgaaaagttttgcgacagcctcaaagttatttgatgcaaaaacccggcgc
agcgtactgatgctctacgcctggtgccgccattgtgacgatgttattgacgatcagacgctgggctttcaggcccggcagcctgcctt
acaaacgcccgaacaacgtctgatgcaacttgagatgaaaacgcgccaggcctatgcaggatcgcagatgcacgaaccggcgtttg
cggcttttcaggaagtggctatggctcatgatatcgccccggcttacgcgtttgatcatctggaaggcttcgccatggatgtacgcgaag
cgcaatacagccaactggatgatacgctgcgctattgctatcacgttgcaggcgttgtcggcttgatgatggcgcaaatcatgggcgtg
cgggataacgccacgctggaccgcgcctgtgaccttgggctggcatttcagttgaccaatattgctcgcgatattgtggacgatgcgc
atgcgggccgctgttatctgccggcaagctggctggagcatgaaggtctgaacaaagagaattatgcggcacctgaaaaccgtcag
gcgctgagccgtatcgcccgtcgtttggtgcaggaagcagaaccttactatttgtctgccacagccggcctggcagggttgcccctgc
gttccgcctgggcaatcgctacggcgaagcaggtttaccggaaaataggtgtcaaagttgaacaggccggtcagcaagcctgggat
cagcggcagtcaacgaccacgcccgaaaaattaacgctgctgctggccgcctctggtcaggcccttacttcccggatgcgggctcat
cctccccgccctgcgcatctctggcagcgcccgctctagCTCGAGGGGCCCGGCGCCtgatgcggtattttctccttac
gcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccg
ccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatg
tgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataa
tggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctc
atgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttg
cggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacat
cgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgt
ggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcacc
agtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaac
ttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaact
ggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggccct
tccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagc
cctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcact
gattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagat
cctttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttctt
gagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctacc
aactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaag
aactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttg
gactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacga
cctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggt
aagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgcc
acctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttc
ctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgatac
cgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*crtE*; 3409 bp

Based on vector pUC18m

Construction
Identical to pUC18m-*crtE-crtB* but without the *crtB* gene.

pUC18m-*crtE* sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcggg
cagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga
attgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggattacaaaatgacggtctgcgcaaaaaa
acacgttcatctcactcgcgatgctgcggagcagttactggctgatattgatcgacgccttgatcagttattgcccgtggagggagaac
gggatgttgtgggtgccgcgatgcgtgaaggtgcgctggcaccgggaaaacgtattcgccccatgttgctgttgctgaccgcccgcg
atctgggttgcgctgtcagccatgacggattactggatttggcctgtgcggtggaaatggtccacgcggcttcgctgatccttgacgata
tgccctgcatggacgatgcgaagctgcggcgcggacgccctaccattcattctcattacggagagcatgtggcaatactggcggcgg
ttgccttgctgagtaaagcctttggcgtaattgccgatgcagatggcctcacgccgctggcaaaaaatcgggcggtttctgaactgtca
aacgccatcggcatgcaaggattggttcagggtcagttcaaggatctgtctgaaggggataagccgcgcagcgctgaagctattttga
tgacgaatcactttaaaaccagcacgctgttttgtgcctccatgcagatggcctcgattgttgcgaatgcctccagcgaagcgcgtgatt
gcctgcatcgtttttcacttgatcttggtcaggcatttcaactgctggacgatttgaccgatggcatgaccgacaccggtaaggatagca
atcaggacgccggtaaatcgacgctggtcaatctgttaggcccgagggcggttgaagaacgtctgagacaacatcttcagcttgcca
gtgagcatctctctgcggcctgccaacacgggcacgccactcaacattttattcaggcctggtttgacaaaaaactcgctgccgtcagtt
aaTCTAGACTCGAGGGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatat
atggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacg
ggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccga
aacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttt
cggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgctt
caataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcac
ccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaag
atccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgcc
gggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatg
gcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggacc
gaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccatacca
aacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttccc
ggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgat
aaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctaca
cgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcag
accaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaa
aatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaat
ctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactgg
cttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacata
cctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccgga
taaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagataccta
cagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacagg
agagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttt
gtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctc
acatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgac
cgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*crtM*$_{F26A,W38A}$; 3365 bp

Based on vector pUC18m

Construction

The gene encoding the F26A, W38A double mutant of CrtM is inserted between the *Xba*I and *Xho*I sites of pUC18m. The codon mutations encoding the F26A and W38A amino acid substitutions are shown in bold, underlined capital letters in the sequence below.

pUC18m-*crtM*$_{F26A,W38A}$ sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagt
gagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagc
ggataacaatttcacacaggaaacaCATATGGAATTCTCTAGAaggaggattacaaaatgacaatgatgaatatgaattttaa
atattgtcataaaatcatgaagaaacattcaaaaagcttttcttacgct**GCA**gacttgttaccagaagatcaaagaaaagcggtt**GCA**gca
atttatgctgtgtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaatttttaaAtcaaataaaagaagatatacaatctattgaaaa
atacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatatcgcctttcaatcttttt
ataatctcattgatactgtatataaagatcaacattttacaatgtttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtag
gtgaagtattgacgccgattttaagtgatcatgaaacacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatat
taagagatgtcggtgaagattttgacaatgaacggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaa
atggtgttaataatcattatattgacttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagta
ttgaagcacaaccaatcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtt
tttgtggataagaggaaaaaggcaaagttgtttcatgaaaAtaaatagtaaatatcatagaatatagCTCGAGGGGCCCGGCG
CCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagtta
agccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtc
tccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatg
tcatgataatgtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgt
atccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctt
ttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacat
cgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcg
cggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacag
aaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaa
cgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaa
gccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagc
ttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgat
aaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgac
ggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagttt
actcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgt
gagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaa
aaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatac
caaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccag
tggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggg
gggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcc
cgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcct
ggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgcc
agcaacgcggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattacc
gcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

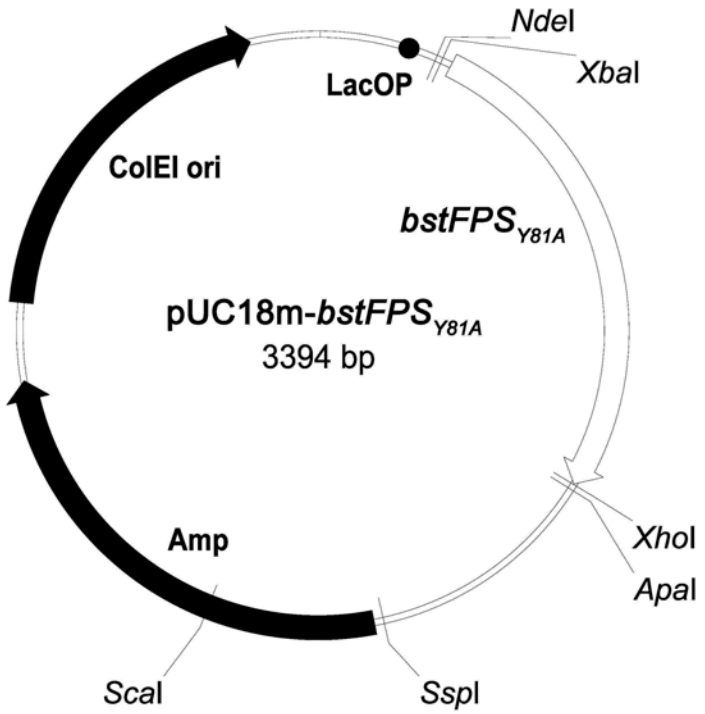
Plasmid name pUC18m-*bstFPS*$_{Y81A}$; 3394 bp

Based on vector pUC18m

Construction

The farnesyl diphosphate synthase gene from *Bacillus stearothermophilus* (with a codon mutation resulting in the amino acid substitution Y81A) is inserted in the *Xba*I and *Xho*I sites of the vector. This ORF is named *bstFPS*$_{Y81A}$ for short. Just upstream of the ORF is an optimized Shine-Dalgarno site (AGGAGG) followed by an 8-nucleotide spacer.

Plasmid map



pUC18m-*bstFPS*$_{Y81A}$ sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGGAATTCTCTAGAaggaggagtaagc
gatggcgcagctttcagttgaacagtttctcaacgagcaaaaacaggcggtggaaacagcgctctcccgttatatagagcgctta
gaagggccggcgaagctgaaaaaggcgatggcgtactcattggaggccggcggcaaacgaatccgtccgttgctgcttctgt
ccaccgttcgggcgctcggcaaagacccggcggtcggattgcccgtcgcctgcgcgattgaaatgatccatacggcatctttga
tccatgatgatttgccgagcatggacaacgatgatttgcggcgcggcaagccgacgaaccataaagtgttcggcgaggcgatg
gccatcttggcgggggacgggttgttgacgtacgcgtttcaattgatcaccgaaatcgacgatgagcgcatccctccttccgtcc
ggcttcggctcatcgaacggctggcgaaagcggccggtccggaagggatggtcgccggtcaggcagccgatatggaagga
gaggggaaaacgctgacgctttcggagctcgaatacattcatcggcataaaaccgggaaaatgctgcaatacagcgtgcacgc
cggcgccttgatcggcggcgctgatgcccggcaaacgcgggagcttgacgaattcgccgcccatctaggccttgcctttcaaat

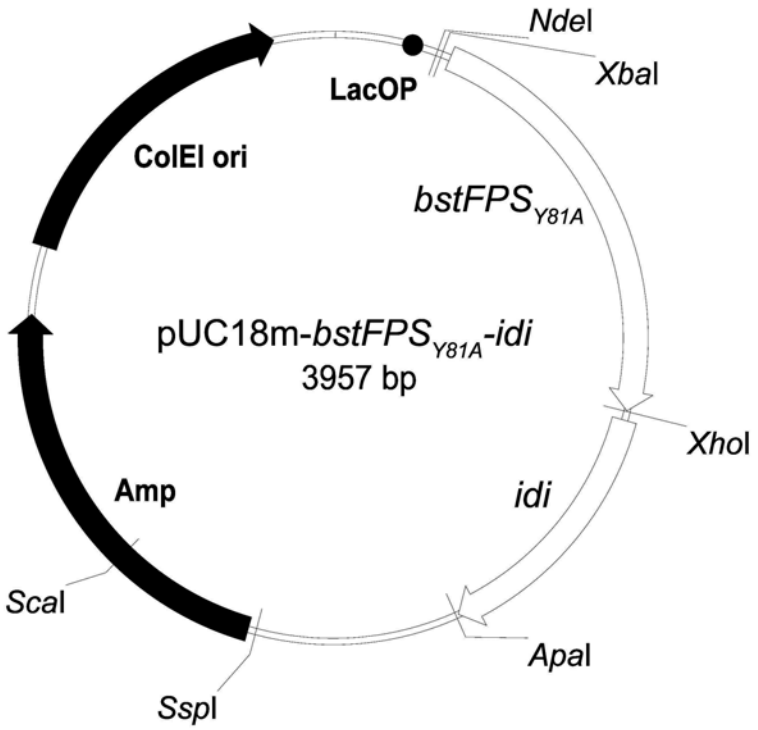
tcgcgatgatattctcgatattgaaggggcagaagaaaaaatcggcaagccggtcggcagcgaccaaagcaacaacaaagcg
acgtatccagcgttgctgtcgcttgccggcgcgaaggaaaagttggcgttccatatcgaggcggcgcagcgccatttacggaa
cgctgacgttgacggcgccgcgctcgcctatatttgcgaactggtcgccgcccgcgaccattaaCTCGAGGGGCCC
GGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctg
atgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgct
tacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaaggg
cctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcg
gaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaa
aaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcacccagaaacg
ctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttg
agagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgg
gcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggat
ggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggag
gaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaag
ccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactactta
ctctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctgg
ctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcc
cgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactg
attaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaa
gatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaagg
atcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatc
aagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttag
gccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataag
tcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacaca
gcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaaggg
agaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcct
ggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatgga
aaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgt
ggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgag
gaagcggaaga

Plasmid name pUC18m-*bstFPS*$_{Y81A}$-*idi*; 3957 bp

Based on vector pUC18m

Construction
Identical to pUC18m-*bstFPS*$_{Y81A}$ but with the addition of the *idi* gene encoding *E. coli* isopentenyl diphosphate isomerase between *Xho*I and *Apa*I sites.

Plasmid map



pUC18m-*bstFPS*$_{Y81A}$-*idi* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGGAATTCTCTAGAaggaggagtaagc
gatggcgcagctttcagttgaacagtttctcaacgagcaaaaacaggcggtggaaacagcgctctcccgttatatagagcgctta
gaagggccggcgaagctgaaaaaggcgatggcgtactcattggaggccggcggcaaacgaatccgtccgttgctgcttctgt
ccaccgttcgggcgctcggcaaagacccggcggtcggattgcccgtcgcctgcgcgattgaaatgatccatacggcatctttga
tccatgatgatttgccgagcatggacaacgatgatttgcggcgcggcaagccgacgaaccataaagtgttcggcgaggcgatg
gccatcttggcgggggacgggttgttgacgtacgcgtttcaattgatcaccgaaatcgacgatgagcgcatccctccttccgtcc
ggcttcggctcatcgaacggctggcgaaagcggccggtccggaagggatggtcgccggtcaggcagccgatatggaagga
gaggggaaaacgctgacgctttcggagctcgaatacattcatcggcataaaaccgggaaaatgctgcaatacagcgtgcacgc
cggcgccttgatcggcggcgctgatgcccggcaaacgcgggagcttgacgaattcgccgcccatctaggccttgcctttcaaat
tcgcgatgatattctcgatattgaaggggcagaagaaaaaatcggcaagccggtcggcagcgaccaaagcaacaacaaagcg
acgtatccagcgttgctgtcgcttgccggcgcgaaggaaaagttggcgttccatatcgaggcggcgcagcgccatttacggaa

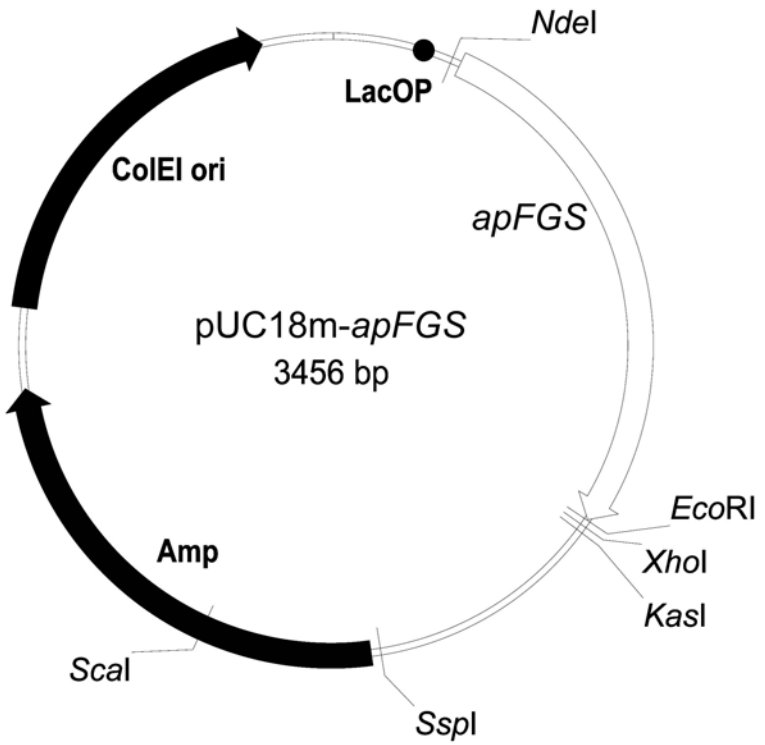
cgctgacgttgacggcgccgcgctcgcctatatttgcgaactggtcgccgcccgcgaccattaaCTCGAGaggaggatta
caaaatgcaaacggaacacgtcattttattgaatgcacagggagttcccacgggtacgctggaaaagtatgccgcacacacgg
cagacacccgcttacatctcgcgttctccagttggctgtttaatgccaaaggacaattattagttacccgccgcgcactgagcaaa
aaagcatggcctggcgtgtggactaactcggtttgtgggcacccacaactgggagaaagcaacgaagacgcagtgatccgcc
gttgccgttatgagcttggcgtggaaattacgcctcctgaatctatctatcctgactttcgctaccgcgccaccgatccgagtggca
ttgtggaaaatgaagtgtgtccggtatttgccgcacgcaccactagtgcgttacagatcaatgatgatgaagtgatggattatcaat
ggtgtgatttagcagatgtattacacggtattgatgccacgccgtgggcgttcagtccgtggatggtgatgcaggcgacaaatcg
cgaagccagaaaacgattatctgcatttacccagcttaaataaGGGCCCGGCGCCtgatgcggtattttctccttacgca
tctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacaccc
gccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagct
gcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtc
atgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattca
aatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtg
tcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagt
tgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatg
atgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacacta
ttctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctg
ccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaa
catgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacga
tgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactgg
atggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtg
agcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtc
aggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagttta
ctcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatccc
ttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatct
gctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaact
ggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgc
ctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccg
aactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcg
gcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccac
ctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggtt
cctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagct
gataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*apFGS*; 3456 bp

Based on vector pUC18m

Construction
The gene encoding farnesylgeranyl diphosphate (C_{25}PP) synthase from *Aeropyrum pernix* is inserted between the *Nde*I and *Eco*RI sites of pUC18m.

Plasmid map



pUC18m-*apFGS* sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGaggaggatactcgatgctcatagaccactatata
atggattttatgagcataactcctgacaggcttagcggggcatcgctacacctcattaaggccggcgggaagaggctcagaccc
ctgataacactcttgacggcgaggatgctgggggtctagaggctgaggcccgtgcaatccctcttgcagcatctatagagacg
gcccacacgttttcgcttatacacgatgacatcatggatagggatgaggtgaggaggggcgtccccacaacccatgtagtctac
ggcgatgactgggctatactggctggcgacacgctgcacgcggcggccttcaagatgatagccgattctagggagtgggggga
tgagccatgaacaggcctaccgggccttcaaagttctttctgaggctgctatacagatatccaggggtcaggcgtacgacatgct
cttcgaggagacctgggatgttgatgtagccgactacctcaacatggtcaggctcaagactggcgctctcatagaggcggcgg
cgcgcataggagccgtcgctgcgggggctggcagcgagattgagaagatgatgggtgaggtgggtatgaacgctggcatcg
ccttccagataagggatgacatactgggcgttataggtgatccaaaggttactggcaagcctgtttacaacgacttgaggagggg
gaagaagacgctgctggtgatatacgctgtcaagaaggcgggtcggagggagatcgtcgaccttattgggcccaaggcgagt

gaggatgatttgaagagggcggcatcgataatagtggactcgggtgccctcgactacgcggagtccagggctaggttctatgtt
gaaagggctagggacatactatccagggtgccggcagttgacgcggagtcgaaggagcttctcaaccttcttctagactatattg
tggagagggtgaagtagGAATTCTCTAGACTCGAGGGGCCCGGCGCCtgatgcggtattttctcctt
acgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccga
cacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgg
gagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggtta
atgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatac
attcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttc
cgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagat
cagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttcc
aatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatac
actattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagt
gctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgc
acaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacc
acgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaataga
ctggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagcc
ggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgggg
agtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaa
gtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaa
atcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgt
aatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggg
taactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagca
ccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaag
acgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgaccta
caccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggt
aagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttc
gccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttt
acggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagt
gagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*apFGS-crtM*$_{F26A,W38A}$; 4335 bp

Based on vector pUC18m

Construction

Identical to pUC18m-*apFGS* but with the gene encoding the F26A, W38A double mutant of CrtM is inserted between the *Xba*I and *Xho*I sites. The codon mutations encoding the F26A and W38A amino acid substitutions are shown in bold, underlined capital letters in the sequence below.

pUC18m-*apFGS-crtM*$_{F26A,W38A}$ sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGaggaggatactcgatgctcatagaccactatata
atggattttatgagcataactcctgacaggcttagcggggcatcgctacacctcattaaggccggcgggaagaggctcagaccc
ctgataacactcttgacggcgaggatgctgggggtctagaggctgaggcccgtgcaatccctcttgcagcatctatagagacg
gcccacacgttttcgcttatacacgatgacatcatggatagggatgaggtgaggaggggcgtccccacaacccatgtagtctac
ggcgatgactgggctatactggctggcgacacgctgcacgcggcggccttcaagatgatagccgattctagggagtgggggga
tgagccatgaacaggcctaccgggccttcaaagttctttctgaggctgctatacagatatccaggggtcaggcgtacgacatgct
cttcgaggagacctgggatgttgatgtagccgactacctcaacatggtcaggctcaagactggcgctctcatagaggcggcgg
cgcgcataggagccgtcgctgcgggggctggcagcgagattgagaagatgatgggtgaggtgggtatgaacgctggcatcg
ccttccagataagggatgacatactgggcgttataggtgatccaaaggttactggcaagcctgtttacaacgacttgaggagggg
gaagaagacgctgctggtgatatacgctgtcaagaaggcgggtcggagggagatcgtcgaccttattgggcccaaggcgagt
gaggatgatttgaagagggcggcatcgataatagtggactcgggtgccctcgactacgcggagtccagggctaggttctatgtt
gaaagggctagggacatactatccagggtgccggcagttgacgcggagtcgaaggagcttctcaaccttcttctagactatattg
tggagagggtgaagtagGAATTCTCTAGAaggaggattacaaaatgacaatgatgaatatgaattttaaatattgtcat
aaaatcatgaagaaacattcaaaaagcttttcttacgct**GCA**gacttgttaccagaagatcaaagaaaagcggtt**GCA**gcaa
tttatgctgtgtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaatttttaaAtcaaataaaagaagatatacaatctat
tgaaaaatacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatat
cgcctttcaatctttttataatctcattgatactgtatataaagatcaacattttacaatgtttgaaacggacgctgaattattcggatatt
gttatggtgttgctggtacagtaggtgaagtattgacgccgattttaagtgatcatgaaacacatcagacatacgatgtcgcaagaa
gacttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaagattttgacaatgaacggatatattttagtaagcaacga
ttaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcattatattgacttatgggaatattatgcagctatc
gcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagcacaaccaatcatagaattagcagcacgtatat
atattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgtggataagaggaaaaaggcaaagttg
tttcatgaaaAtaaatagtaaatatcatagaatatagCTCGAGGGGCCCGGCGCCtgatgcggtattttctccttac
gcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgaca
cccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccggga
gctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaat
gtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacatt
caaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccg
tgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatca
gttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaa
tgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacac
tattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgct
gccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgctttttgcaca

acatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacg
atgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactg
gatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggt
gagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagt
caggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagttt
actcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcc
cttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatc
tgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaac
tggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgc
ctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccg
aactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcg
gcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccac
ctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggtt
cctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagct
gataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*hexPS*; 4629 bp

Based on vector pUC18m

Construction

The operon from *Micrococcus luteus* strain B-P 26 encoding the two polypeptide components of hexaprenyl diphosphate (C_{30}PP) synthase (*hexPS*) is inserted between the *Eco*RI and *Xba*I sites of pUC18m.

pUC18m-*hexPS* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggTAGATCAatgcgt
tatttacataaaattgaactagaattaaaccgacttacaagtcgatatccatttttcaaaaaaattgcatttgatgctgaaatcataaag
ctcgttgatgacctaaatgtcgatgaaaatgtaaaatgtgcgattgttgccattgacacgagtatgcgtatgcaggattttatcaatg
aagataataaagacagttttgtactatcaacggatgttttgagtgctttattttataagtatttatcacagccattttatcagcatgatttttt
agtactgacggattgtgtaagtcgtatcaatgaattaaaatcaataagagcaacgattacagacgaaattgctttgcataatattaat
aaacaaattcattatatgttcatacaaccttatatgaacaatgagaaagtggtgtcttatgagtaaacagttaaatggacaggaaaaa
agtgagcttgtacataatgtattccagaatgtatcgacaaagtatgaccgcctcaacgatatcataagttttaatcagcataaatcct
ggcgtaaatatacgatgaaacagatgaatgttaaaaaagggtcgaaagcacttgatgtatgctgcggtacaggcgactggacaa
ttcagatggcacaggctgtcggtaaaaatggtcatgttattggtcttgatttcagtgagaatatgttaagtgttgcacaaggaaaaac
gaatcatatacaaaatattgaattaattcatggtaatgcgatggaattaccatttgaagataatatatttgattatacaacgattggtttt
ggtttacgtaacttaccggattataaaaaaggattagaagaaatgtatcgtgtattaaaacctggcggcatgattgttgttttagaaa
cgagccatccaacaatgccagtatttaaacaaggttacaaattatatttcaaatacgttatgcccctgtttgggaaagtatttgctaag
tctatgaaggaatatagctggttacagcaaagtgcttttgaatttcctgataagtacacgttagcacttttaatggctgaaactggattt
acacacattaaatttaaaggttttactggtggcgtgagtgcgatgcatcttgcatacaagccgaaagaaaaatagaatggatgatt
gctttgagttataaagcgtttttaaacccatatatcattgaagttgaaaaaaggttatatgagtgtattcagagtgattctgaaacgata
aacaaggcggcacaccatattttaagttcaggaggaaagcgcgtacgtccgatgtttgtattattaagtggttttctgaatgatacac
aaaaggatgacttgattcgtacagcagtatctctggagctcgttcatatggcaagtctcgttcatgatgattacatcgataatagtga
tatgcgtcgtggtaatacttcggttcatatagcttttgataaagacacagcaattcgcacaggacatttttattagcacgtgcgttac
aaaatattgcaactatcaataattcgaaattccatcaaatttttagtaaaacgatacttgaagtttgttttggtgaatttgaccagatgg
cagatcgatttaattatcctgtatcctttactgcatatttaagacgtattaatcgtaaaacagcgatactgatagaagcaagctgtcatt
taggggctctcagctcacagcttgatgaacaatctacatatcatataaaacaatttgggcattgtattggaatgagttatcaaattatt
gatgatattctcgattacacgagtgacgaagcaacactcggtaaacctgtcggtagcgatataagaaacggtcatattacgtatcc
gcttatggccgctatcgctaatttgaaagagcaagatgacgataaacttgaagcagttgttaaacatttaacatcaacatcagatga
tgaagtgtatcaatatattgtttcgcaagttaaacaatatggaattgaacctgcagaattgctgagcagaaaatatggtgataaagc
gaaatatcacttgagtcaattacaggatagtaatattaaagattatttagaagaaatccacgaaaaaatgttaaaacgtgtttattaaT
CTAGACTCGAGGGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatat
atggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctg
acgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcat
caccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcag
gtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataac
cctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttg
ccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaac
tggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtgg
cgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcac

cagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcgg
ccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgat
cgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgc
gcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggacc
acttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcag
cactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaataga
cagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaactt
cattttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgt
cagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgc
taccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaat
actgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttac
cagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcg
ggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctat
gagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcac
gagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatg
ctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcaca
tgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgac
cgagcgcagcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*hexPS-crtM-crtN*; 7044 bp

Based on vector pUC18m

Construction

Identical to plasmid pUC18m-*crtM-crtN* but with the additional operon from *Micrococcus luteus* strain B-P 26 encoding the two polypeptide components of hexaprenyl diphosphate (C_{30}PP) synthase (*hexPS*) inserted between the *Eco*RI and *Xba*I sites.

pUC18m-*hexPS-crtM-crtN* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcggg
cagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga
attgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggtagatcaatgcgttatttacataaaattgaa
ctagaattaaaccgacttacaagtcgatatccatttttcaaaaaaattgcatttgatgctgaaatcataaagctcgttgatgacctaaatgtc
gatgaaaatgtaaaatgtgcgattgttgccattgacacgagtatgcgtatgcaggattttatcaatgaagataataaagacagttttgtact
atcaacggatgttttgagtgctttatttataagtatttatcacagccattttatcagcatgatttttttagtactgacggattgtgtaagtcgtatc
aatgaattaaaatcaataagagcaacgattacagacgaaattgctttgcataatattaataaacaaattcattatatgttcatacaaccttata
tgaacaatgagaaagtggtgtcttatgagtaaacagttaaatggacaggaaaaaagtgagcttgtacataatgtattccagaatgtatcg
acaaagtatgaccgcctcaacgatatcataagttttaatcagcataaatcctggcgtaaatatacgatgaaacagatgaatgttaaaaaa
gggtcgaaagcacttgatgtatgctgcggtacaggcgactggacaattcagatggcacaggctgtcggtaaaaatggtcatgttattgg
tcttgatttcagtgagaatatgttaagtgttgcacaaggaaaaacgaatcatatacaaaatattgaattaattcatggtaatgcgatggaatt
accatttgaagataatatatttgattatacaacgattggttttggtttacgtaacttaccggattataaaaaaggattagaagaaatgtatcgt
gtattaaaacctggcggcatgattgttgttttagaaacgagccatccaacaatgccagtatttaaacaaggttacaaattatatttcaaatac
gttatgcccctgtttgggaaagtatttgctaagtctatgaaggaatatagctggttacagcaaagtgcttttgaatttcctgataagtacacg
ttagcacttttaatggctgaaactggatttacacacattaaatttaaaggttttactggtggcgtgagtgcgatgcatcttgcatacaagccg
aaagaaaaatagaatggatgattgctttgagttataaagcgtttttaaacccatatatcattgaagttgaaaaaaggttatatgagtgtattc
agagtgattctgaaacgataaacaaggcggcacaccatattttaagttcaggaggaaagcgcgtacgtccgatgtttgtattattaagtg
gttttctgaatgatacacaaaaggatgacttgattcgtacagcagtatctctggagctcgttcatatggcaagtctcgttcatgatgattaca
tcgataatagtgatatgcgtcgtggtaatacttcggttcatatagcttttgataaagacacagcaattcgcacaggacatttttattagcac
gtgcgttacaaaatattgcaactatcaataattcgaaattccatcaaatttttagtaaaacgatacttgaagtttgttttggtgaatttgaccag
atggcagatcgatttaattatcctgtatcctttactgcatatttaagacgtattaatcgtaaaacagcgatactgatagaagcaagctgtcatt
taggggctctcagctcacagcttgatgaacaatctacatatcatataaaacaatttgggcattgtattggaatgagttatcaaattattgatg
atattctcgattacacgagtgacgaagcaacactcggtaaacctgtcggtagcgatataagaaacggtcatattacgtatccgcttatgg
ccgctatcgctaatttgaaagagcaagatgacgataaacttgaagcagttgttaaacatttaacatcaacatcagatgatgaagtgtatca
atatattgtttcgcaagttaaacaatatggaattgaacctgcagaattgctgagcagaaaatatggtgataaagcgaaatatcacttgagt
caattacaggatagtaatattaaagattatttagaagaaatccacgaaaaaatgttaaaacgtgtttattaaTCTAGAaggaggatta
caaaatgacaatgatgaatatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagcttttcttacgcttttgacttgttacca
gaagatcaaagaaaagcggtttgggcaatttatgctgtgtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaatttttaaAt
caaataaaagaagatatacaatctattgaaaaatacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgcttcagc
atgttgcacaacataaaaatatcgcctttcaatctttttataatctcattgatactgtatataaagatcaacattttacaatgtttgaaacggac
gctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacgccgattttaagtgatcatgaaacacatcagacata
cgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaagattttgacaatgaacggatatatttta
gtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcattatattgacttatgggaatattatg
cagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagcacaaccaatcatagaattagcagcacg
tatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgtggataagaggaaaaaggcaaagttg
tttcatgaaaAtaaatagtaaatatcatagaatatagCTCGAGaggaggattacaaaatgaagattgcagtaattggtgcaggtgt

cacaggattagcagcggcagcccgtattgcttctcaaggtcatgaagtgacgatatttgaaaaaaataataatgtaggcgggcgtatga
atcaattaaagaaagacggctttacatttgatatgggtcccacaattgtcatgatgccagatgtttataaagatgtttttacagcgtgtggta
aaaattatgaagattatattgaattgagacaattacgttatatttacgatgtgtattttgaccacgatgatcgtataacggtgcctacagattta
gctgaattacagcaaatgctagaaagtatagaacctggttcaacgcatggttttatgtccttttaacggatgtttataaaaaaatatgaaatt
gcacgtcgctatttcttagaaagaacgtatcgcaaaccgagtgacttttataatatgacgtcacttgtgcaaggtgctaagttaaaaacgtt
aaatcatgcagatcagctaattgaacattatattgataacgaaaagatacaaaagcttttagcgtttcaaacgttatacataggaattgatc
caaaacgaggcccgtcactatattcaattattcctatgattgaaatgatgtttggtgtgcattttattaaaggcggtatgtatggcatggctc
aagggctagcgcaattaaataaagacttaggcgttaatattgaactaaatgctgaaattgagcaaattattattgatcctaaattcaaacgg
gccgatgcgataaaagtgaatggtgacataagaaaatttgataaaattttatgtacggctgatttccctagtgttgcggaatcattaatgcc
agattttgcacctattaaaaagtatccaccacataaaattgcagacttagattactcttgttcagcatttttaatgtatatcggtatagatattg
atgtgacagatcaagtgagacttcataatgttatttttttcagatgactttagaggcaatattgaagaaatatttgagggacgtttatcatatga
tccttctatttatgtgtatgtaccagcggtcgctgataaatcacttgcgccagaaggcaaaactggtatttatgtgctaatgccgacgccg
gaacttaaaacaggtagcggaatcgattggtcagatgaagctttgacgcaacaaataaaggaaattatttatcgtaaattagcaacgatt
gaagtatttgaagatataaaatcgcatattgtttcagaaacaatctttacgccaaatgattttgagcaaacgtatcatgcgaaatttggttcg
gcattcggtttaatgccaactttagcgcaaagtaattattatcgtccacaaaatgtatcgcgagattataaagatttatattttgcaggtgca
agtacgcatccaggtgcaggcgttcctattgtcttaacgagtgcgaaaataactgtagatgaaatgattaaagatattgagcggggcgt
ataagggagtagtctaaGGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatgg
tgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggctt
gtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgc
gcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcgggg
aaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataa
tattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccaga
aacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcctt
gagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggc
aagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatg
acagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagg
agctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgac
gagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctg
gagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagt
ttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatccctt
aacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgct
tgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagca
gagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctc
tgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg
cagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtg
agctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgc
acgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgct
cgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttct
ttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgc
agcgagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*hexPS-crtB-crtI*; 7027 bp

Based on vector pUC18m

Construction

Identical to plasmid pUC18m-*crtE-crtB-crtI* but with the operon from *Micrococcus luteus* strain B-P 26 encoding the two polypeptide components of hexaprenyl diphosphate (C_{30}PP) synthase (*hexPS*) in place of *crtE* between the *Eco*RI and *Xba*I sites.

pUC18m-*hexPS-crtB-crtI* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcggg
cagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga
attgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggtagatcaatgcgttatttacataaaattgaa
ctagaattaaaccgacttacaagtcgatatccatttttcaaaaaaattgcatttgatgctgaaatcataaagctcgttgatgacctaaatgtc
gatgaaaatgtaaaatgtgcgattgttgccattgacacgagtatgcgtatgcaggattttatcaatgaagataataaagacagttttgtact
atcaacggatgttttgagtgctttatttataagtatttatcacagccattttatcagcatgatttttttagtactgacggattgtgtaagtcgtatc
aatgaattaaaatcaataagagcaacgattacagacgaaattgctttgcataatattaataaacaaattcattatatgttcatacaaccttata
tgaacaatgagaaagtggtgtcttatgagtaaacagttaaatggacaggaaaaaagtgagcttgtacataatgtattccagaatgtatcg
acaaagtatgaccgcctcaacgatatcataagttttaatcagcataaatcctggcgtaaatatacgatgaaacagatgaatgttaaaaaa
gggtcgaaagcacttgatgtatgctgcggtacaggcgactggacaattcagatggcacaggctgtcggtaaaaatggtcatgttattgg
tcttgatttcagtgagaatatgttaagtgttgcacaaggaaaaacgaatcatatacaaaatattgaattaattcatggtaatgcgatggaatt
accatttgaagataatatatttgattatacaacgattggttttggtttacgtaacttaccggattataaaaaaggattagaagaaatgtatcgt
gtattaaaacctggcggcatgattgttgttttagaaacgagccatccaacaatgccagtatttaaacaaggttacaaattatatttcaaatac
gttatgcccctgtttgggaaagtatttgctaagtctatgaaggaatatagctggttacagcaaagtgcttttgaatttcctgataagtacacg
ttagcacttttaatggctgaaactggatttacacacattaaatttaaaggttttactggtggcgtgagtgcgatgcatcttgcatacaagccg
aaagaaaaatagaatggatgattgctttgagttataaagcgtttttaaacccatatatcattgaagttgaaaaaaggttatatgagtgtattc
agagtgattctgaaacgataaacaaggcggcacaccatattttaagttcaggaggaaagcgcgtacgtccgatgtttgtattattaagtg
gttttctgaatgatacacaaaaggatgacttgattcgtacagcagtatctctggagctcgttcatatggcaagtctcgttcatgatgattaca
tcgataatagtgatatgcgtcgtggtaatacttcggttcatatagcttttgataaagacacagcaattcgcacaggacatttttattagcac
gtgcgttacaaaatattgcaactatcaataattcgaaattccatcaaatttttagtaaaacgatacttgaagtttgttttggtgaatttgaccag
atggcagatcgatttaattatcctgtatcctttactgcatatttaagacgtattaatcgtaaaacagcgatactgatagaagcaagctgtcatt
taggggctctcagctcacagcttgatgaacaatctacatatcatataaaacaatttgggcattgtattggaatgagttatcaaattattgatg
atattctcgattacacgagtgacgaagcaacactcggtaaacctgtcggtagcgatataagaaacggtcatattacgtatccgcttatgg
ccgctatcgctaatttgaaagagcaagatgacgataaacttgaagcagttgttaaacatttaacatcaacatcagatgatgaagtgtatca
atatattgtttcgcaagttaaacaatatggaattgaacctgcagaattgctgagcagaaaatatggtgataaagcgaaatatcacttgagt
caattacaggatagtaatattaaagattatttagaagaaatccacgaaaaaatgttaaaacgtgtttattaaTCTAGAaggaggatta
caaaatggcagttggctcgaaaagttttgcgacagcctcaaagttatttgatgcaaaaacccggcgcagcgtactgatgctctacgcct
ggtgccgccattgtgacgatgttattgacgatcagacgctgggctttcaggcccggcagcctgccttacaaacgcccgaacaacgtct
gatgcaacttgagatgaaaacgcgccaggcctatgcaggatcgcagatgcacgaaccggcgtttgcggcttttcaggaagtggctat
ggctcatgatatcgccccggcttacgcgtttgatcatctggaaggcttcgccatggatgtacgcgaagcgcaatacagccaactggat
gatacgctgcgctattgctatcacgttgcaggcgttgtcggcttgatgatggcgcaaatcatgggcgtgcgggataacgccacgctgg
accgcgcctgtgaccttgggctggcatttcagttgaccaatattgctcgcgatattgtggacgatgcgcatgcgggccgctgttatctgc
cggcaagctggctggagcatgaaggtctgaacaaagagaattatgcggcacctgaaaaccgtcaggcgctgagccgtatcgcccgt
cgtttggtgcaggaagcagaaccttactatttgtctgccacagccggcctggcagggttgcccctgcgttccgcctgggcaatcgctac
ggcgaagcaggtttaccggaaaataggtgtcaaagttgaacaggccggtcagcaagcctgggatcagcggcagtcaacgaccacg
cccgaaaaattaacgctgctgctggccgcctctggtcaggcccttacttcccggatgcgggctcatcctccccgccctgcgcatctctg
gcagcgcccgctctagCTCGAGaggaggattacaaaatgaaaccaactacggtaattggtgcaggcttcggtggcctggcact

ggcaattcgtctacaagctgcggggatccccgtcttactgcttgaacaacgtgataaacccggcggtcgggcttatgtctacgaggatc
aggggtttacctttgatgcaggcccgacggttatcaccgatcccagtgccattgaagaactgtttgcactggcaggaaaacagttaaaa
gagtatgtcgaactgctgccggttacgccgttttaccgcctgtgttgggagtcagggaaggtctttaattacgataacgatcaaacccgg
ctcgaagcgcagattcagcagtttaatccccgcgatgtcgaaggttatcgtcagtttctggactattcacgcgcggtgtttaaagaaggc
tatctaaagctcggtactgtccctttttatcgttcagagacatgcttcgcgccgcacctcaactggcgaaactgcaggcatggagaagc
gtttacagtaaggttgccagttacatcgaagatgaacatctgcgccaggcgttttctttccactcgctgttggtgggcggcaatcccttcg
ccacctcatccatttatacgttgatacacgcgctggagcgtgagtggggcgtctggtttccgcgtggcggcaccggcgcattagttcag
gggatgataaagctgtttcaggatctgggtggcgaagtcgtgttaaacgccagagtcagccatatggaaacgacaggaaacaagatt
gaagccgtgcatttagaggacggtcgcaggttcctgacgcaagccgtcgcgtcaaatgcagatgtggttcatacctatcgcgacctgtt
aagccagcaccctgccgcggttaagcagtccaacaaactgcagactaagcgcatgagtaactctctgtttgtgctctattttggtttgaat
caccatcatgatcagctcgcgcatcacacggtttgtttcggcccgcgttaccgcgagctgattgacgaaatttttaatcatgatggcctcg
cagaggacttctcactttatctgcacgcgccctgtgtcacggattcgtcactggcgcctgaaggttgcggcagttactatgtgttggcgc
cggtgccgcatttaggcaccgcgaacctcgactggacggttgaggggccaaaactacgcgaccgtatttttgcgtaccttgagcagc
attacatgcctggcttacggagtcagctggtcacgcaccggatgtttacgccgtttgattttcgcgaccagcttaatgcctatcatggctc
agccttttctgtggagcccgttcttacccagagcgcctggtttcggccgcataaccgcgataaaaccattactaatctctacctggtcgg
cgcaggcacgcatcccggcgcaggcattcctggcgtcatcggctcggcaaaagcgacagcaggtttgatgctggaggatctgatttg
aGGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaat
ctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatc
cgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaaggg
cctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaac
ccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaaga
gtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagta
aaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccga
agaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcg
ccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatg
cagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgc
acaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacg
atgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatg
gaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtg
ggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactat
ggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttag
attgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttcc
actgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaa
atactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttacca
gtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctga
acgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcg
ccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttcca
gggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcgg
agcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccct
gattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagc
gaggaagcggaaga

Plasmid name pUC18m-*hexPS-crtB*; 5534 bp

Based on vector pUC18m

Construction
Identical to plasmid pUC18m-*hexPS-crtB-crtI* but lacking *crtI*.

pUC18m-*hexPS-crtB* sequence
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggtagatcaatgcgttattta
cataaaattgaactagaattaaaccgacttacaagtcgatatccatttttcaaaaaaattgcatttgatgctgaaatcataaagctcgtt
gatgacctaaatgtcgatgaaaatgtaaaatgtgcgattgttgccattgacacgagtatgcgtatgcaggattttatcaatgaagata
ataaagacagttttgtactatcaacggatgttttgagtgctttatttttataagtatttatcacagccattttatcagcatgatttttttagtact
gacggattgtgtaagtcgtatcaatgaattaaaatcaataagagcaacgattacagacgaaattgctttgcataatattaataaacaa
attcattatatgttcatacaaccttatatgaacaatgagaaagtggtgtcttatgagtaaacagttaaatggacaggaaaaaagtgag
cttgtacataatgtattccagaatgtatcgacaaagtatgaccgcctcaacgatatcataagttttaatcagcataaatcctggcgta
aatatacgatgaaacagatgaatgttaaaaaagggtcgaaagcacttgatgtatgctgcggtacaggcgactggacaattcagat
ggcacaggctgtcggtaaaaatggtcatgttattggtcttgatttcagtgagaatatgttaagtgttgcacaaggaaaaacgaatca
tatacaaaatattgaattaattcatggtaatgcgatggaattaccatttgaagataatatatttgattatacaacgattggttttggtttac
gtaacttaccggattataaaaaaggattagaagaaatgtatcgtgtattaaaacctggcggcatgattgttgttttagaaacgagcc
atccaacaatgccagtatttaaacaaggttacaaattatatttcaaatacgttatgcccctgtttgggaaagtatttgctaagtctatga
aggaatatagctggttacagcaaagtgcttttgaatttcctgataagtacacgttagcacttttaatggctgaaactggatttacacac
attaaatttaaaggttttactggtggcgtgagtgcgatgcatcttgcatacaagccgaaagaaaaatagaatggatgattgctttga
gttataaagcgtttttaaacccatatatcattgaagttgaaaaaaggttatatgagtgtattcagagtgattctgaaacgataaacaag
gcggcacaccatattttaagttcaggaggaaagcgcgtacgtccgatgtttgtattattaagtggttttctgaatgatacacaaaag
gatgacttgattcgtacagcagtatctctggagctcgttcatatggcaagtctcgttcatgatgattacatcgataatagtgatatgcg
tcgtggtaatacttcggttcatatagcttttgataaagacacagcaattcgcacaggacatttttttattagcacgtgcgttacaaaatat
tgcaactatcaataattcgaaattccatcaaatttttagtaaaacgatacttgaagtttgttttggtgaatttgaccagatggcagatcg
atttaattatcctgtatcctttactgcatatttaagacgtattaatcgtaaaacagcgatactgatagaagcaagctgtcatttaggggc
tctcagctcacagcttgatgaacaatctacatatcatataaaacaatttgggcattgtattggaatgagttatcaaattattgatgatatt
ctcgattacacgagtgacgaagcaacactcggtaaacctgtcggtagcgatataagaaacggtcatattacgtatccgcttatgg
ccgctatcgctaatttgaaagagcaagatgacgataaacttgaagcagttgttaaacatttaacatcaacatcagatgatgaagtgt
atcaatatattgtttcgcaagttaaacaatatggaattgaacctgcagaattgctgagcagaaaatatggtgataaagcgaaatatc
acttgagtcaattacaggatagtaatattaaagattatttagaagaaatccacgaaaaaatgttaaaacgtgtttattaaTCTAG
Aaggaggattacaaaatggcagttggctcgaaaagttttgcgacagcctcaaagttatttgatgcaaaaacccggcgcagcgta
ctgatgctctacgcctggtgccgccattgtgacgatgttattgacgatcagacgctgggctttcaggcccggcagcctgccttaca
aacgcccgaacaacgtctgatgcaacttgagatgaaaacgcgccaggcctatgcaggatcgcagatgcacgaaccggcgttt
gcggcttttcaggaagtggctatggctcatgatatcgccccggcttacgcgtttgatcatctggaaggcttcgccatggatgtacg
cgaagcgcaatacagccaactggatgatacgctgcgctattgctatcacgttgcaggcgttgtcggcttgatgatggcgcaaatc
atgggcgtgcgggataacgccacgctggaccgcgcctgtgaccttgggctggcatttcagttgaccaatattgctcgcgatattg
tggacgatgcgcatgcgggccgctgttatctgccggcaagctggctggagcatgaaggtctgaacaaagagaattatgcggca
cctgaaaaccgtcaggcgctgagccgtatcgcccgtcgtttggtgcaggaagcagaaccttactatttgtctgccacagccggc
ctggcagggttgcccctgcgttccgcctgggcaatcgctacggcgaagcaggtttaccggaaaataggtgtcaaagttgaaca
ggccggtcagcaagcctgggatcagcggcagtcaacgaccacgcccgaaaaattaacgctgctgctggccgcctctggtcag
gcccttacttcccggatgcgggctcatcctccccgccctgcgcatctctggcagcgcccgctctagCTCGAGGGGCC

CGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctc
tgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccg
cttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaag
ggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcg
cggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattga
aaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcacccagaaa
cgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcc
ttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgcc
gggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacgg
atggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcgg
aggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatga
agccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactac
ttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggct
ggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccct
cccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcac
tgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtg
aagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaag
gatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggat
caagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagtta
ggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataa
gtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcacac
agcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagg
gagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgc
ctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatgg
aaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattct
gtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcg
aggaagcggaaga

Plasmid name pUC18m-*hexPS-crtM*$_{F26L}$; 5508 bp

Based on vector pUC18m

Construction

Identical to plasmid pUC18m-*hexPS-crtB* but with the gene *crtM*$_{F26L}$ encoding the F26L mutant of the C_{30} carotenoid synthase CrtM from *Staphylococcus aureus* in place of *crtB* between the *Xba*I and *Xho*I sites. The codon mutation encoding the F26L amino acid substitution in *crtM*$_{F26L}$ is shown in bold, underlined capital letters in the sequence below.

pUC18m-*hexPS-crtM*$_{F26L}$ sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggtagatcaatgcgttattta
cataaaattgaactagaattaaaccgacttacaagtcgatatccatttttcaaaaaaattgcatttgatgctgaaatcataaagctcgtt
gatgacctaaatgtcgatgaaaatgtaaaatgtgcgattgttgccattgacacgagtatgcgtatgcaggattttatcaatgaagata
ataaagacagttttgtactatcaacggatgttttgagtgctttattttataagtatttatcacagccattttatcagcatgatttttagtact
gacggattgtgtaagtcgtatcaatgaattaaaatcaataagagcaacgattacagacgaaattgctttgcataatattaataaacaa
attcattatatgttcatacaaccttatatgaacaatgagaaagtggtgtcttatgagtaaacagttaaatggacaggaaaaaagtgag
cttgtacataatgtattccagaatgtatcgacaaagtatgaccgcctcaacgatatcataagttttaatcagcataaatcctggcgta
aatatacgatgaaacagatgaatgttaaaaaagggtcgaaagcacttgatgtatgctgcggtacaggcgactggacaattcagat
ggcacaggctgtcggtaaaaatggtcatgttattggtcttgatttcagtgagaatatgttaagtgttgcacaaggaaaaacgaatca
tatacaaaatattgaattaattcatggtaatgcgatggaattaccatttgaagataatatatttgattatacaacgattggttttggtttac
gtaacttaccggattataaaaaaggattagaagaaatgtatcgtgtattaaaacctggcggcatgattgttgttttagaaacgagcc
atccaacaatgccagtatttaaacaaggttacaaattatatttcaaatacgttatgcccctgtttgggaaagtatttgctaagtctatga
aggaatatagctggttacagcaaagtgcttttgaatttcctgataagtacacgttagcacttttaatggctgaaactggatttacacac
attaaatttaaaggttttactggtggcgtgagtgcgatgcatcttgcatacaagccgaaagaaaaatagaatggatgattgctttga
gttataaagcgtttttaaacccatatatcattgaagttgaaaaaaggttatatgagtgtattcagagtgattctgaaacgataaacaag
gcggcacaccatattttaagttcaggaggaaagcgcgtacgtccgatgtttgtattattaagtggttttctgaatgatacacaaaag
gatgacttgattcgtacagcagtatctctggagctcgttcatatggcaagtctcgttcatgatgattacatcgataatagtgatatgcg
tcgtggtaatacttcggttcatatagcttttgataaagacacagcaattcgcacaggacatttttttattagcacgtgcgttacaaaatat
tgcaactatcaataattcgaaattccatcaaatttttagtaaaacgatacttgaagtttgttttggtgaatttgaccagatggcagatcg
atttaattatcctgtatcctttactgcatatttaagacgtattaatcgtaaaacagcgatactgatagaagcaagctgtcatttaggggc
tctcagctcacagcttgatgaacaatctacatatcatataaaacaatttgggcattgtattggaatgagttatcaaattattgatgatatt
ctcgattacacgagtgacgaagcaacactcggtaaacctgtcggtagcgatataagaaacggtcatattacgtatccgcttatgg
ccgctatcgctaatttgaaagagcaagatgacgataaacttgaagcagttgttaaacatttaacatcaacatcagatgatgaagtgt
atcaatatattgtttcgcaagttaaacaatatggaattgaacctgcagaattgctgagcagaaaatatggtgataaagcgaaatatc
acttgagtcaattacaggatagtaatattaaagattatttagaagaaatccacgaaaaaatgttaaaacgtgtttattaaTCTAG
Aaggaggattacaaaatgacaatgatgaatatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagcttttcttacg
ct**CTC**gacttgttaccagaagatcaaagaaaagcggtttgggcaatttatgctgtgtgtcgtaaaattgatgacagtatagatgtt
tatggcgatattcaatttttaaatcaaataaaagaagatatacaatctattgaaaaatacccatatgaacatcatcactttcaaagtgat
cgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatatcgcctttcaatctttttataatctcattgatactgtatataaa
gatcaacattttacaatgtttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacgcc
gattttaagtgatcatgaaacacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagat
gtcggtgaagattttgacaatgaacggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaa
atggtgttaataatcattatattgacttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaag

tatttagtattgaagcacaaccaatcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactata
cattacatgaacgtgtttttgtggataagaggaaaaaggcaaagttgtttcatgaaaataaatagtaaatatcatagaatatagCT
CGAGGGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactc
tcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtc
tgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacg
cgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttc
ggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatg
cttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgttttt
gctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaac
agcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattat
cccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacaga
aaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttc
tgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaacc
ggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaa
ctggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctc
ggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggcca
gatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgag
ataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaa
aaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgta
gaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtg
gtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttcta
gtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgct
gccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggg
gggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcc
acgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttc
cagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcagggg
ggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgc
gttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagc
gagtcagtgagcgaggaagcggaaga

Plasmid name pUC18m-*hexPS-crtM*$_{F26A,W38A}$; 5508 bp

Based on vector pUC18m

Construction

Identical to plasmid pUC18m-*hexPS-crtB* but with the gene *crtM*$_{F26A,W38A}$ encoding the F26A, W38A mutant of the C_{30} carotenoid synthase CrtM from *Staphylococcus aureus* in place of *crtB* between the *Xba*I and *Xho*I sites. The codon mutations encoding the F26A and W38A amino acid substitutions in *crtM*$_{F26A,W38A}$ are shown in bold, underlined capital letters in the sequence below.

pUC18m-*hexPS-crtM*$_{F26A,W38A}$ sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcggataacaatttcacacaggaaacaCATATGGAATTCaggaggtagatcaatgcgttattta
cataaaattgaactagaattaaaccgacttacaagtcgatatccatttttcaaaaaaattgcatttgatgctgaaatcataaagctcgtt
gatgacctaaatgtcgatgaaaatgtaaaatgtgcgattgttgccattgacacgagtatgcgtatgcaggattttatcaatgaagata
ataaagacagttttgtactatcaacggatgttttgagtgctttattttataagtatttatcacagccattttatcagcatgatttttagtact
gacggattgtgtaagtcgtatcaatgaattaaaatcaataagagcaacgattacagacgaaattgctttgcataatattaataaacaa
attcattatatgttcatacaaccttatatgaacaatgagaaagtggtgtcttatgagtaaacagttaaatggacaggaaaaaagtgag
cttgtacataatgtattccagaatgtatcgacaaagtatgaccgcctcaacgatatcataagttttaatcagcataaatcctggcgta
aatatacgatgaaacagatgaatgttaaaaaagggtcgaaagcacttgatgtatgctgcggtacaggcgactggacaattcagat
ggcacaggctgtcggtaaaaatggtcatgttattggtcttgatttcagtgagaatatgttaagtgttgcacaaggaaaaacgaatca
tatacaaaatattgaattaattcatggtaatgcgatggaattaccatttgaagataatatatttgattatacaacgattggttttggtttac
gtaacttaccggattataaaaaaggattagaagaaatgtatcgtgtattaaaacctggcggcatgattgttgttttagaaacgagcc
atccaacaatgccagtatttaaacaaggttacaaattatatttcaaatacgttatgcccctgtttgggaaagtatttgctaagtctatga
aggaatatagctggttacagcaaagtgcttttgaatttcctgataagtacacgttagcacttttaatggctgaaactggatttacacac
attaaatttaaaggttttactggtggcgtgagtgcgatgcatcttgcatacaagccgaaagaaaaatagaatggatgattgctttga
gttataaagcgtttttaaacccatatatcattgaagttgaaaaaaggttatatgagtgtattcagagtgattctgaaacgataaacaag
gcggcacaccatattttaagttcaggaggaaagcgcgtacgtccgatgtttgtattattaagtggttttctgaatgatacacaaaag
gatgacttgattcgtacagcagtatctctggagctcgttcatatggcaagtctcgttcatgatgattacatcgataatagtgatatgcg
tcgtggtaatacttcggttcatatagcttttgataaagacacagcaattcgcacaggacatttttttattagcacgtgcgttacaaaatat
tgcaactatcaataattcgaaattccatcaaatttttagtaaaacgatacttgaagtttgttttggtgaatttgaccagatggcagatcg
atttaattatcctgtatcctttactgcatatttaagacgtattaatcgtaaaacagcgatactgatagaagcaagctgtcatttaggggc
tctcagctcacagcttgatgaacaatctacatatcatataaaacaatttgggcattgtattggaatgagttatcaaattattgatgatatt
ctcgattacacgagtgacgaagcaacactcggtaaacctgtcggtagcgatataagaaacggtcatattacgtatccgcttatgg
ccgctatcgctaatttgaaagagcaagatgacgataaacttgaagcagttgttaaacatttaacatcaacatcagatgatgaagtgt
atcaatatattgtttcgcaagttaaacaatatggaattgaacctgcagaattgctgagcagaaaatatggtgataaagcgaaatatc
acttgagtcaattacaggatagtaatattaaagattatttagaagaaatccacgaaaaaatgttaaaacgtgtttattaaTCTAG
Aaggaggattacaaaatgacaatgatgaatatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagcttttcttacg
ct**GCA**gacttgttaccagaagatcaaagaaaagcggtt**GCA**gcaatttatgctgtgtgtcgtaaaattgatgacagtatagat
gtttatggcgatattcaatttttaaAtcaaataaaagaagatatacaatctattgaaaaatacccatatgaacatcatcactttcaaagt
gatcgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatatcgcctttcaatctttttataatctcattgatactgtatata
aagatcaacattttacaatgtttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacg
ccgattttaagtgatcatgaaacacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagag
atgtcggtgaagattttgacaatgaacggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtacca

aaatggtgttaataatcattatattgacttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaa
agtatttagtattgaagcacaaccaatcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaacta
tacattacatgaacgtgtttttgtggataagaggaaaaaggcaaagttgtttcatgaaaAtaaatagtaaatatcatagaatatagC
TCGAGGGGCCCGGCGCCtgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcact
ctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgt
ctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaac
gcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttt
tcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaa
tgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtt
tttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctca
acagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtatt
atcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcaca
gaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttac
ttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttggga
accggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaacta
ttaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcg
ctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactgggg
ccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgct
gagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatt
taaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagacccc
gtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcg
gtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtcctt
ctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggct
gctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaac
gggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagc
gccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagc
ttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcagg
ggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcct
gcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgca
gcgagtcagtgagcgaggaagcggaaga

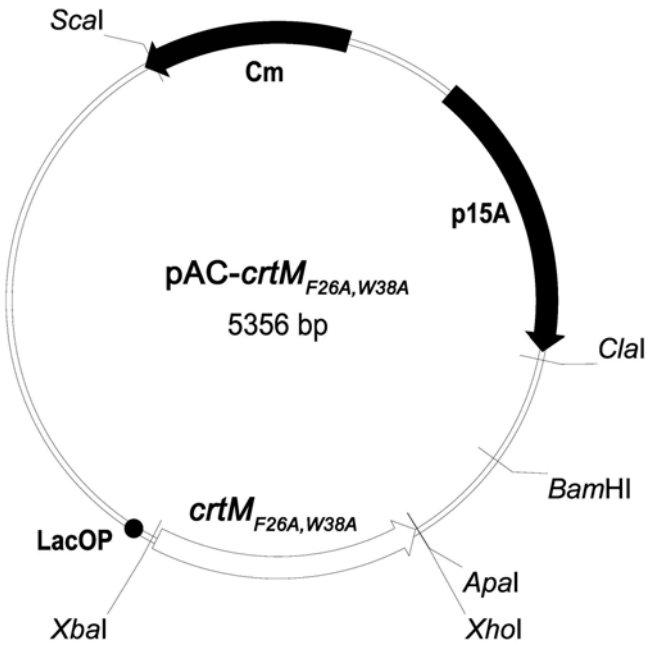
Plasmid name pAC-*crtM*$_{F26A,W38A}$; 5356 bp

Based on vector pACmod

Construction

A fragment containing a *lac* promoter/operator followed by an optimized Shine-Dalgarno sequence, 8 spacer nucleotides, and then *crtM*$_{F26A,W38A}$ (encoding the F26A, W38A variant of the C_{30} carotenoid synthase CrtM from *S. aureus*) was inserted in the *Sal*I site of pACmod (disrupting the tetracycline resistance gene on the plasmid backbone). The resulting plasmid was sequenced to confirm the direction of the insertion, which is as shown on the plasmid map. *crtM*$_{F26A,W38A}$ is directly flanked by unique *Xba*I and *Xho*I restriction sites. The codon mutations encoding the F26A and W38A amino acid substitutions are shown in bold, underlined capital letters in the sequence below. Note that the sequence shows the reverse strand of *crtM*$_{F26A,W38A}$.

Plasmid map



pAC-*crtM*$_{F26A,W38A}$ sequence

gaattccggatgagcattcatcaggcgggcaagaatgtgaataaaggccggataaaacttgtgcttatttttctttacggtctttaaaaaggccg
taatatccagctgaacggtctggttataggtacattgagcaactgactgaaatgcctcaaaatgttctttacgatgccattgggatatatcaacgg
tggtatatccagtgatttttttctccattttagcttccttagctcctgaaaatctcgataactcaaaaaatacgcccggtagtgatcttatttcattatgg
tgaaagttggaacctcttacgtgccgatcaacgtctcattttcgccaaaagttggcccagggcttcccggtatcaacagggacaccaggattta
tttattctgcgaagtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcgggtgatgctgccaacttactgatttagtgtatgatggtgttt
ttgaggtgctccagtggcttctgtttctatcagctgtccctcctgttcagctactgacggggtggtgcgtaacggcaaaagcaccgccggacat
cagcgctagcggagtgtatactggcttactatgttggcactgatgagggtgtcagtgaagtgcttcatgtggcaggagaaaaaaggctgcac
cggtgcgtcagcagaatatgtgatacaggatatattccgcttcctcgctcactgactcgctacgctcggtcgttcgactgcggcgagcggaaa
tggcttacgaacggggcggagatttcctggaagatgccaggaagatacttaacagggaagtgagagggccgcggcaaagccgtttttccat

aggctccgccccccctgacaagcatcacgaaatctgacgctcaaatcagtggtggcgaaacccgacaggactataaagataccaggcgtttc
cccctggcggctccctcgtgcgctctcctgttcctgcctttcggtttaccggtgtcattccgctgttatggccgcgtttgtctcattccacgcct
gacactcagttccgggtaggcagttcgctccaagctggactgtatgcacgaaccccccgttcagtccgaccgctgcgccttatccggtaact
atcgtcttgagtccaacccggaaagacatgcaaaagcaccactggcagcagccactggtaattgatttagaggagttagtcttgaagtcatgc
gccggttaaggctaaactgaaaggacaagttttggtgactgcgctcctccaagccagttacctcggttcaaagagttggtagctcagagaacc
ttcgaaaaaccgccctgcaaggcggttttttcgttttcagagcaagagattacgcgcagaccaaaacgatctcaagaagatcatcttattaatca
gataaaatatttcaagatttcagtgcaatttatctcttcaaatgtagcacctgaagtcagccccatacgatataagttgtaattctcatgtttgacag
cttatcATCGATaagctttaatgcggtagtttatcacagttaaattgctaacgcagtcaggcaccgtgtatgaaatctaacaatgcgctcat
cgtcatcctcggcaccgtcaccctggatgctgtaggcataggcttggttatgccggtactgccgggcctcttgcgggatatcgtccattccga
cagcatcgccagtcactatggcgtgctgctagcgctatatgcgttgatgcaatttctatgcgcacccgttctcggagcactgtccgaccgctttg
gccgccgcccagtcctgctcgcttcgctacttggagccactatcgactacgcgatcatggcgaccacacccgtcctgtGGATCCtctac
gccggacgcatcgtggccggcatcaccggcgccacaggtgcggttgctggcgcctatatcgccgacatcaccgatggggaagatcgggc
tcgccacttcgggctcatgagcgcttgtttcggcgtgggtatggtggcaggccccgtggccgggggactgttgggcgccatctccttgcatg
caccattccttgcggcggcggtgctcaacggcctcaacctactactgggctgcttcctaatgcaggagtcgcataagggagagcGTCG
ACgtgaaataccgcacagatgcgtaaggagaaaataccgcatcaggcgccGGGCCCCTCGAGctatattctatgatatttactat
ttattttcatgaaacaactttgccttttttcctcttatccacaaaaacacgttcatgtaatgtatagttagcctgtctcacttcgtccagtatttcaatatat
atacgtgctgctaattctatgattggttgtgcttcaatactaaatactttgatttgatccataacatcttgaaaatctttttctgcgatagctgcataata
ttcccataagtcaatataatgattattaacaccattttggtacacttcagcaatatcaacttcatattgctttaatcgttgcttactaaaatatatccgtt
cattgtcaaaatcttcaccgacatctcttaatatattaatcaattgcaacgattcaccaagtcttcttgcgacatcgtatgtctgatgtgtttcatgatc
acttaaaatcggcgtcaatacttcacctactgtaccagcaacaccataacaatatccgaataattcagcgtccgtttcaaacattgtaaaatgttg
atctttatatacagtatcaatgagattataaaaagattgaaaggcgatatttttatgttgtgcaacatgctgaagcgccatcatgattctacgatcac
tttgaaagtgatgatgttcatatgggtatttttcaatagattgtatatcttctttttatttgatttaaaaattgaatatcgccataaacatctatactgtcatc
aattttacgacacacagcataaattgc**TGC**aaccgcttttctttgatcttctggtaacaagtc**TGC**agcgtaagaaaagctttttgaatgtttc
ttcatgattttatgacaatatttaaaattcatattcatcattgtcattttgtaatcctcctTCTAGAGAATTCCATATGtgtttcctgtgt
gaaattgttatccgctcacaattccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacat
taattgcgttgcgctcactgcccgctttccagtcgggaaacctGTCGACcgatgcccttgagagccttcaacccagtcagctccttccgg
tgggcgcggggcatgactatcgtcgccgcacttatgactgtcttctttatcatgcaactcgtaggacaggtgccggcagcgctctgggtcattt
tcggcgaggaccgctttcgctggagcgcgacgatgatcggcctgtcgcttgcggtattcggaatcttgcacgccctcgctcaagccttcgtca
ctggtcccgccaccaaacgtttcggcgagaagcaggccattatcgccggcatggcggccgacgcgctgggctacgtcttgctggcgttcg
cgacgcgaggctggatggccttccccattatgattcttctcgcttccggcggcatcgggatgcccgcgttgcaggccatgctgtccaggcag
gtagatgacgaccatcagggacagcttcaaggatcgctcgcggctcttaccagcctaacttcgatcactggaccgctgatcgtcacggcgat
ttatgccgcctcggcgagcacatggaacgggttggcatggattgtaggcgccgccctataccttgtctgcctccccgcgttgcgtcgcggtg
catggagccgggccacctcgacctgaatggaagccggcggcacctcgctaacggattcaccactccaagaattggagccaatcaattcttg
cggagaactgtgaatgcgcaaaccaaccctggcagaacatatccatcgcgtccgccatctccagcagccgcacgcggcgcatctcgggc
agcgttgggtcctggccacgggtgcgcatgatcgtgctcctgtcgttgaggacccggctaggctggcggggttgccttactggttagcagaa
tgaatcaccgatacgcgagcgaacgtgaagcgactgctgctgcaaaacgtctgcgacctgagcaacaacatgaatggtcttcggtttccgtg
tttcgtaaagtctggaaacgcggaagtcccctacgtgctgctgaagttgcccgcaacagagagtggaaccaaccggtgataccacgatacta
tgactgagagtcaacgccatgagcggcctcatttcttattctgagttacaacagtccgcaccgctgtccggtagctccttccggtgggcgcgg
ggcatgactatcgtcgccgcacttatgactgtcttctttatcatgcaactcgtaggacaggtgccggcagcgcccaacagtcccccggccac
ggggcctgccaccatacccacgccgaaacaagcgccctgcaccattatgttccggatctgcatcgcaggatgctgctggctaccctgtgga
acacctacatctgtattaacgaagcgctaaccgtttttatcaggctctgggaggcagaataaatgatcatatcgtcaattattacctccacgggg
agagcctgagcaaactggcctcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccagcaatagacat
aagcggctatttaacgaccctgccctgaaccgacgaccgggtcgaatttgctttcgaatttctgccattcatccgcttattatcacttattcaggc
gtagcaccaggcgtttaagggcaccaataactgccttaaaaaaattacgcccgccctgccactcatcgcAGTACTgttgtaattcatta
agcattctgccgacatggaagccatcacagacggcatgatgaacctgaatcgccagcggcatcagcaccttgtcgccttgcgtataatatttg
cccatggtgaaaacgggggcgaagaagttgtccatattggccacgtttaaatcaaaactggtgaaactcacccagggattggctgagacga
aaaacatattctcaataaaccctttagggaaataggccaggttttcaccgtaacacgccacatcttgcgaatatatgtgtagaaactgccggaa
atcgtcgtggtattcactccagagcgatgaaaacgtttcagtttgctcatggaaaacggtgtaacaagggtgaacactatcccatatcaccagc
tcaccgtctttcattgccatacg

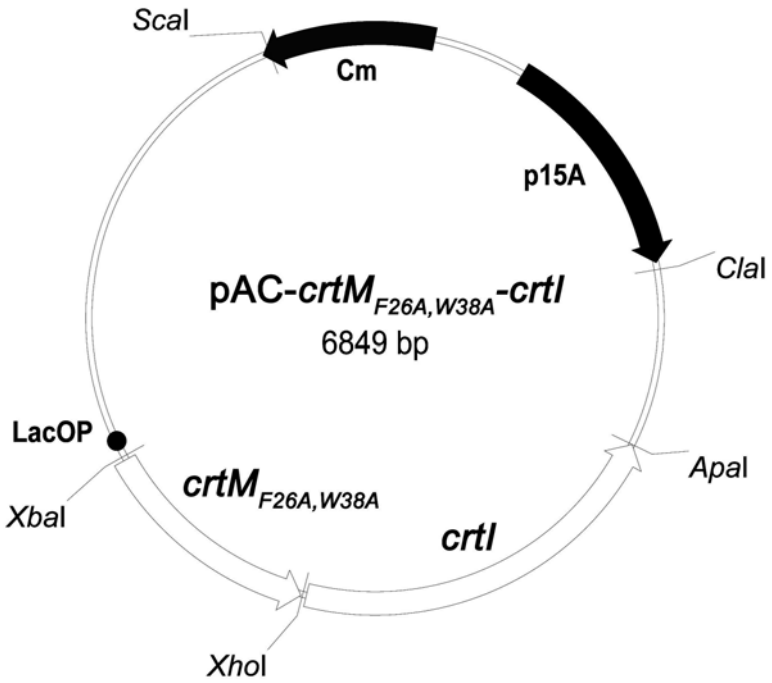
Plasmid name pAC-*crtM*$_{F26A,W38A}$-*crtI*; 6849 bp

Based on vector pACmod

Construction

Derived from pAC-*crtM*$_{F26A,W38A}$ but with the *crtI* gene (C_{40} carotenoid desaturase) from *Erwinia uredovora* inserted after the mutant allele of *crtM*. *crtI* is flanked by unique *Xho*I and *Apa*I restriction sites, which were used to insert it into pAC-*crtM*$_{F26A,W38A}$ with known orientation (see plasmid map). Each gene is preceded by a Shine-Dalgarno ribosomal binding sequence (AGGAGG) followed by 8 spacer nucleotides. The codon mutations encoding the F26A and W38A amino acid substitutions in *crtM*$_{F26A,W38A}$ are shown in bold, underlined capital letters in the sequence below. Note that the sequence shows the reverse strand of *crtM*$_{F26A,W38A}$ and *crtI*.

Plasmid map



pAC-*crtM*$_{F26A,W38A}$-*crtI* sequence

gaattccggatgagcattcatcaggcgggcaagaatgtgaataaaggccggataaaacttgtgcttatttttctttacggtctttaaa
aaggccgtaatatccagctgaacggtctggttataggtacattgagcaactgactgaaatgcctcaaaatgttctttacgatgccat
tgggatatatcaacggtggtatatccagtgatttttttctccattttagcttccttagctcctgaaaatctcgataactcaaaaaatacgc
ccggtagtgatcttatttcattatggtgaaagttggaacctcttacgtgccgatcaacgtctcattttcgccaaaagttggcccaggg
cttcccggtatcaacagggacaccaggatttatttattctgcgaagtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcg
ggtgatgctgccaacttactgatttagtgtatgatggtgtttttgaggtgctccagtggcttctgtttctatcagctgtccctcctgttca
gctactgacggggtggtgcgtaacggcaaaagcaccgccggacatcagcgctagcggagtgtatactggcttactatgttggc
actgatgagggtgtcagtgaagtgcttcatgtggcaggagaaaaaaggctgcaccggtgcgtcagcagaatatgtgatacagg

atatattccgcttcctcgctcactgactcgctacgctcggtcgttcgactgcggcgagcggaaatggcttacgaacggggcgga
gatttcctggaagatgccaggaagatacttaacagggaagtgagagggccgcggcaaagccgtttttccataggctccgcccc
cctgacaagcatcacgaaatctgacgctcaaatcagtggtggcgaaacccgacaggactataaagataccaggcgtttccccct
ggcggctccctcgtgcgctctcctgttcctgcctttcggtttaccggtgtcattccgctgttatggccgcgtttgtctcattccacgcc
tgacactcagttccgggtaggcagttcgctccaagctggactgtatgcacgaacccccccgttcagtccgaccgctgcgccttatc
cggtaactatcgtcttgagtccaacccggaaagacatgcaaaagcaccactggcagcagccactggtaattgatttagaggagtt
agtcttgaagtcatgcgccggttaaggctaaactgaaaggacaagttttggtgactgcgctcctccaagccagttacctcggttca
aagagttggtagctcagagaaccttcgaaaaaccgccctgcaaggcggtttttcgttttcagagcaagagattacgcgcagacc
aaaacgatctcaagaagatcatcttattaatcagataaaatatttcaagatttcagtgcaatttatctcttcaaatgtagcacctgaagt
cagccccatacgatataagttgtaattctcatgtttgacagcttatcATCGATaagctttaatgcggtagtttatcacagttaaatt
gctaacgcagtcaggcaccgtgtatgaaatctaacaatgcgctcatcgtcatcctcggcaccgtcaccctggatgctgtaggcat
aggcttggttatgccggtactgccgggcctcttgcgggatatcgtccattccgacagcatcgccagtcactatggcgtgctgcta
gcgctatatgcgttgatgcaatttctatgcgcacccgttctcggagcactgtccgaccgctttggccgccgcccagtcctgctcgc
ttcgctacttggagccactatcgactacgcgatcatggcgaccacacccgtcctgtggatcctctacgccggacgcatcgtggcc
ggcatcaccggcgccacaggtgcggttgctggcgcctatatcgccgacatcaccgatggggaagatcgggctcgccacttcg
ggctcatgagcgcttgtttcggcgtgggtatggtggcaggccccgtggccgggggactgttgggcgccatctccttgcatgcac
cattccttgcggcggcggtgctcaacggcctcaacctactactgggctgcttcctaatgcaggagtcgcataagggagagcGT
CGACgtgaaataccgcacagatgcgtaaggagaaaataccgcatcaggcgccGGGCCCtcaaatcagatcctccagc
atcaaacctgctgtcgcttttgccgagccgatgacgccaggaatgcctgcgccgggatgcgtgcctgcgccgaccaggtagag
attagtaatggttttatcgcggttatgcggccgaaaccaggcgctctgggtaagaacgggctccacagaaaaggctgagccatg
ataggcattaagctggtcgcgaaaatcaaacggcgtaaacatccggtgcgtgaccagctgactccgtaagccaggcatgtaatg
ctgctcaaggtacgcaaaaatacggtcgcgtagttttggcccctcaaccgtccagtcgaggttcgcggtgcctaaatgtggcacc
ggcgccaacacatagtaactgccgcaaccttcaggcgccagtgacgaatccgtgacacagggcgcgtgcagataaagtgaga
agtcctctgcgaggccatcatgattaaaaatttcgtcaatcagctcgcggtaacgcgggccgaaacaaaccgtgtgatgcgcga
gctgatcatgatggtgattcaaaccaaaatagagcacaaacagagagttactcatgcgcttagtctgcagtttgttggactgcttaa
ccgcggcagggtgctggcttaacaggtcgcgataggtatgaaccacatctgcatttgacgcgacggcttgcgtcaggaacctg
cgaccgtcctctaaatgcacggcttcaatcttgtttcctgtcgtttccatatggctgactctggcgtttaacacgacttcgccaccca
gatcctgaaacagctttatcatcccctgaactaatgcgccggtgccgccacgcggaaaccagacgccccactcacgctccagc
gcgtgtatcaacgtataaatggatgaggtggcgaagggattgccgcccaccaacagcgagtggaaagaaaacgcctggcgca
gatgttcatcttcgatgtaactggcaaccttactgtaaacgcttctccatgcctgcagtttcgccagttgaggtgcggcgcgaagca
tgtctctgaacgataaaaaagggacagtaccgagctttagatagccttctttaaacaccgcgcgtgaatagtccagaaactgacg
ataaccttcgacatcgcggggattaaactgctgaatctgcgcttcgagccgggtttgatcgttatcgtaattaaagaccttccctga
ctcccaacacaggcggtaaaacggcgtaaccggcagcagttcgacatactcttttaactgttttcctgccagtgcaaacagttcttc
aatggcactgggatcggtgataaccgtcgggcctgcatcaaaggtaaaccccctgatcctcgtagacataagcccgaccgccgg
gtttatcacgttgttcaagcagtaagacggggatccccgcagcttgtagacgaattgccagtgccaggccaccgaagcctgcac
caattaccgtagttggtttcattttgtaatcctcctCTCGAGctatattctatgatatttactatttatttttcatgaaacaactttgccttt
ttcctcttatccacaaaaacacgttcatgtaatgtatagttagcctgtctcacttcgtccagtatttcaatatatatacgtgctgctaattc
tatgattggttgtgcttcaatactaaatactttgatttgatccataacatcttgaaaatctttttctgcgatagctgcataatattcccataa
gtcaatataatgattattaacaccattttggtacacttcagcaatatcaacttcatattgctttaatcgttgcttactaaaatatatccgttc
attgtcaaaatcttcaccgacatctcttaatatattaatcaattgcaacgattcaccaagtcttcttgcgacatcgtatgtctgatgtgttt
catgatcacttaaaatcggcgtcaatacttcacctactgtaccagcaacaccataacaatatccgaataattcagcgtccgtttcaa
acattgtaaaatgttgatctttatatacagtatcaatgagattataaaaagattgaaaggcgatatttttatgttgtgcaacatgctgaa
gcgccatcatgattctacgatcactttgaaagtgatgatgttcatatgggtatttttcaatagattgtatatcttctttttatttgatttaaaaa
ttgaatatcgccataaacatctatactgtcatcaattttacgacacacagcataaattgc**TGC**aaccgcttttctttgatcttctggta
acaagtc**TGC**agcgtaagaaaagctttttgaatgtttcttcatgattttatgacaatatttaaaattcatattcatcattgtcattttgta
atcctcctTCTAGAgaattccatatgtgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgagccggaag

cataaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaa
acctGTCGACcgatgcccttgagagccttcaacccagtcagctccttccggtgggcgcggggcatgactatcgtcgccgc
acttatgactgtcttctttatcatgcaactcgtaggacaggtgccggcagcgctctgggtcattttcggcgaggaccgctttcgctg
gagcgcgacgatgatcggcctgtcgcttgcggtattcggaatcttgcacgccctcgctcaagccttcgtcactggtcccgccacc
aaacgtttcggcgagaagcaggccattatcgccggcatggcggccgacgcgctgggctacgtcttgctggcgttcgcgacgc
gaggctggatggccttccccattatgattcttctcgcttccggcggcatcgggatgcccgcgttgcaggccatgctgtccaggca
ggtagatgacgaccatcagggacagcttcaaggatcgctcgcggctcttaccagcctaacttcgatcactggaccgctgatcgt
cacggcgatttatgccgcctcggcgagcacatggaacgggttggcatggattgtaggcgccgccctataccttgtctgcctccc
cgcgttgcgtcgcggtgcatggagccgggccacctcgacctgaatggaagccggcggcacctcgctaacggattcaccactc
caagaattggagccaatcaattcttgcggagaactgtgaatgcgcaaaccaacccttggcagaacatatccatcgcgtccgccat
ctccagcagccgcacgcggcgcatctcgggcagcgttgggtcctggccacgggtgcgcatgatcgtgctcctgtcgttgagga
cccggctaggctggcggggttgccttactggttagcagaatgaatcaccgatacgcgagcgaacgtgaagcgactgctgctgc
aaaacgtctgcgacctgagcaacaacatgaatggtcttcggtttccgtgtttcgtaaagtctggaaacgcggaagtcccctacgtg
ctgctgaagttgcccgcaacagagagtggaaccaaccggtgataccacgatactatgactgagagtcaacgccatgagcggc
ctcatttcttattctgagttacaacagtccgcaccgctgtccggtagctccttccggtgggcgcggggcatgactatcgtcgccgc
acttatgactgtcttctttatcatgcaactcgtaggacaggtgccggcagcgcccaacagtcccccggccacggggcctgccac
catacccacgccgaaacaagcgccctgcaccattatgttccggatctgcatcgcaggatgctgctggctaccctgtggaacacc
tacatctgtattaacgaagcgctaaccgtttttatcaggctctgggaggcagaataaatgatcatatcgtcaattattacctccacgg
ggagagcctgagcaaactggcctcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccag
caatagacataagcggctatttaacgaccctgccctgaaccgacgaccgggtcgaatttgctttcgaatttctgccattcatccgct
tattatcacttattcaggcgtagcaccaggcgtttaagggcaccaataactgccttaaaaaaattacgccccgccctgccactcatc
gcAGTACTgttgtaattcattaagcattctgccgacatggaagccatcacagacggcatgatgaacctgaatcgccagcgg
catcagcaccttgtcgccttgcgtataatatttgcccatggtgaaaacgggggcgaagaagttgtccatattggccacgtttaaatc
aaaactggtgaaactcacccagggattggctgagacgaaaaacatattctcaataaaccctttagggaaataggccaggttttca
ccgtaacacgccacatcttgcgaatatatgtgtagaaactgccggaaatcgtcgtggtattcactccagagcgatgaaaacgtttc
agtttgctcatggaaaacggtgtaacaagggtgaacactatcccatatcaccagctcaccgtctttcattgccatacg

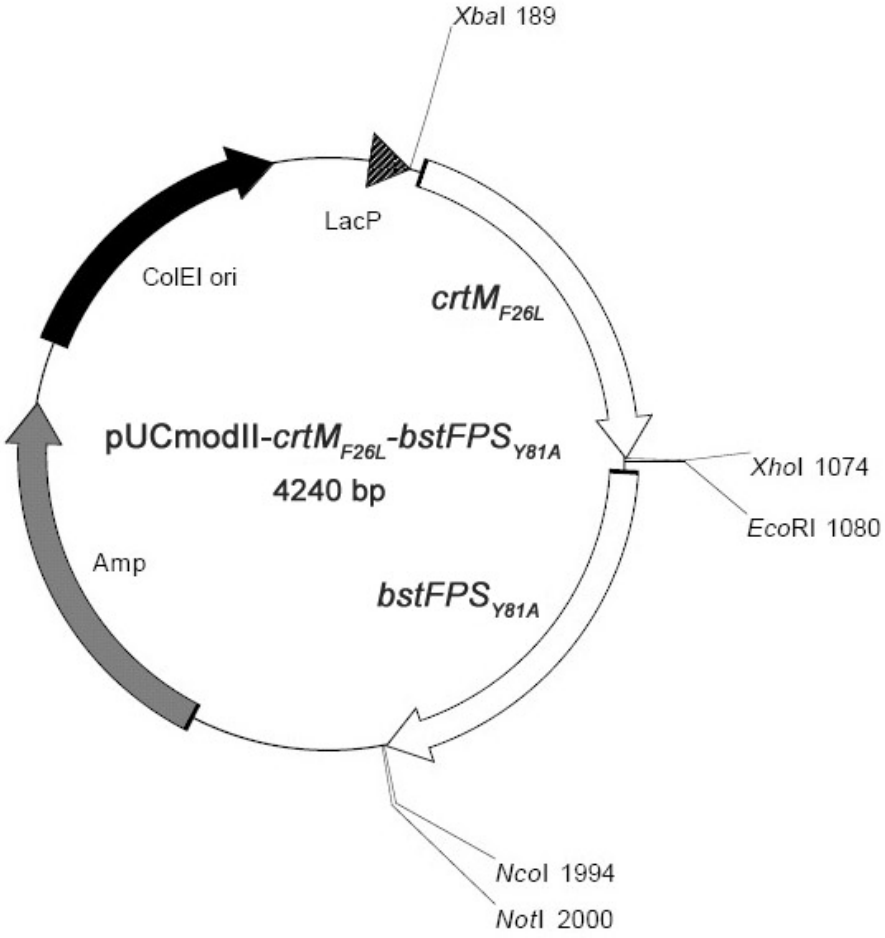
Plasmid name pUCmodII-*crtM*$_{F26L}$-*bstFPS*$_{Y81A}$; 4240 bp

Based on vector pUCmodII

Construction

The gene encoding the F26L mutant of the C_{30} carotenoid synthase CrtM from *Staphylococcus aureus* (*crtM*$_{F26L}$) is cloned between *Xba*I and *Xho*I sites. The gene encoding the Y81A mutant of the farnesyl diphosphate synthase from *Bacillus stearothermophilus* (*bstFPS*$_{Y81A}$) is cloned between *Eco*RI and *Nco*I sites. The genes are part of a single operon under the control of a single *lac* promoter (no operator). Each gene is preceded by a Shine-Dalgarno ribosomal binding sequence (AGGAGG) followed by 8 spacer nucleotides. The codon mutation encoding the F26L amino acid substitution in *crtM*$_{F26L}$ is shown in bold, underlined capital letters in the sequence below.

Plasmid map



pUCmodII-*crtM*$_{F26L}$-*bstFPS*$_{Y81A}$ sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcggg
cagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga
attgtgagcgTCTAGAaggaggattacaaaatgacaatgatgaatatgaattttaaatattgtcataaaatcatgaagaaacattcaa
aaagcttttcttacgct**CTC**gacttgttaccagaagatcaaagaaaagcggtttgggcaatttatgctgtgtgtcgtaaaattgatgaca

gtatagatgtttatggcgatattcaatttttaaAtcaaataaaagaagatatacaatctattgaaaaatacccatatgaacatcatcactttca
aagtgatcgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatatcgcctttcaatctttttataatctcattgatactgtatata
aagatcaacattttacaatgtttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacgccg
attttaagtgatcatgaaacacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcgg
tgaagattttgacaatgaacggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgtta
ataatcattatattgacttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattga
agcacaaccaatcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgt
ttttgtggataagaggaaaaaggcaaagttgtttcatgaaaAtaaatagtaaatatcatagaatatagCTCGAGGAATTCag
gaggagtaagcgatggcgcagctttcagttgaacagtttctcaacgagcaaaaacaggcggtggaaacagcgctctcccgttatatag
agcgcttagaagggccggcgaagctgaaaaaggcgatggcgtactcattggaggccggcggcaaacgaatccgtccgttgctgctt
ctgtccaccgttcgggcgctcggcaaagacccggcggtcggattgcccgtcgcctgcgcgattgaaatgatccatacggcatctttga
tccatgatgatttgccgagcatggacaacgatgatttgcggcgcggcaagccgacgaaccataaagtgttcggcgaggcgatggcc
atcttggcgggggacgggttgttgacgtacgcgtttcaattgatcaccgaaatcgacgatgagcgcatccctccttccgtccggcttcg
gctcatcgaacggctggcgaaagcggccggtccggaagggatggtcgccggtcaggcagccgatatggaaggagaggggaaaa
cgctgacgctttcggagctcgaatacattcatcggcataaaaccgggaaaatgctgcaatacagcgtgcacgccggcgccttgatcg
gcggcgctgatgcccggcaaacgcgggagcttgacgagtttgccgcccatctaggccttgcctttcaaattcgcgatgatattctcgat
attgaaggggcagaagaaaaaatcggcaagccggtcggcagcgaccaaagcaacaacaaagcgacgtatccagcgttgctgtcgc
ttgccggcgcgaaggaaaagttggcgttccatatcgaggcggcgcagcgccatttacggaacgctgacgttgacggcgccgcgctc
gcctatatttgcgaactggtcgccgcccgcgaccattaaCCATGGGCGGCCGCtgcggtattttctccttacgcatctgtgc
ggtatttcacaccgcatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgc
tgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtttt
caccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataatgtttcttagac
gtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaata
accctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgc
cttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggat
ctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtatt
atcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaa
agcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgaca
acgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctga
atgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactac
ttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggc
tggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtat
cgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcat
tggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgata
atctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttt
tttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttc
cgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgta
gcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaaga
cgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccg
aactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcag
ggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgactt
gagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcctttttacggttcctggccttttg
ctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccg
cagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

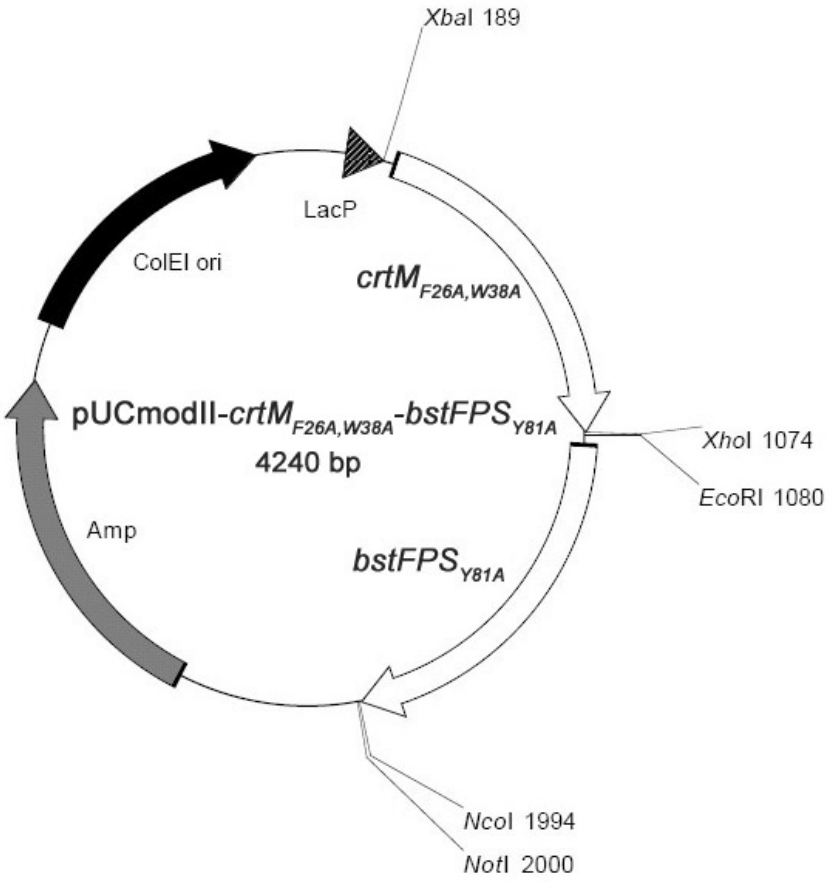
Plasmid name pUCmodII-*crtM*$_{F26A,W38A}$-*bstFPS*$_{Y81A}$; 4240 bp

Based on vector pUCmodII

Construction

The gene encoding the F26A, W38A double mutant of the C_{30} carotenoid synthase CrtM from *Staphylococcus aureus* (*crtM*$_{F26A,W38A}$) is cloned between *Xba*I and *Xho*I sites. The gene encoding the Y81A mutant of the farnesyl diphosphate synthase from *Bacillus stearothermophilus* (*bstFPS*$_{Y81A}$) is cloned between *Eco*RI and *Nco*I sites. The genes are part of a single operon under the control of a single *lac* promoter (no operator). Each gene is preceded by a Shine-Dalgarno ribosomal binding sequence (AGGAGG) followed by 8 spacer nucleotides. The codon mutations encoding the F26A and W38A amino acid substitutions in *crtM*$_{F26A,W38A}$ are shown in bold, underlined capital letters in the sequence below.

Plasmid map



pUCmodII-*crtM*$_{F26A,W38A}$-*bstFPS*$_{Y81A}$ sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcggg
cagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga

attgtgagcgTCTAGAaggaggattacaaaatgacaatgatgaatatgaattttaaatattgtcataaaatcatgaagaaacattcaa
aaagcttttcttacgct**<u>GCA</u>**gacttgttaccagaagatcaaagaaaagcggtt**<u>GCA</u>**gcaatttatgctgtgtgtcgtaaaattgatg
acagtatagatgtttatggcgatattcaatttttaaAtcaaataaaagaagatatacaatctattgaaaaatacccatatgaacatcatcact
ttcaaagtgatcgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatatcgcctttcaatctttttataatctcattgatactgta
tataaagatcaacattttacaatgtttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtattgacg
ccgattttaagtgatcatgaaacacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgt
cggtgaagattttgacaatgaacggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggt
gttaataatcattatattgacttatgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtat
tgaagcacaaccaatcatagaattagcagcacgtatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaac
gtgtttttgtggataagaggaaaaaggcaaagttgtttcatgaaaAtaaatagtaaatatcatagaatatagCTCGAGGAATT
Caggaggagtaagcgatggcgcagctttcagttgaacagtttctcaacgagcaaaaacaggcggtggaaacagcgctctcccgttat
atagagcgcttagaagggccggcgaagctgaaaaaggcgatggcgtactcattggaggccggcggcaaacgaatccgtccgttgc
tgcttctgtccaccgttcgggcgctcggcaaagacccggcggtcggattgcccgtcgcctgcgcgattgaaatgatccatacggcatc
tttgatccatgatgatttgccgagcatggacaacgatgatttgcggcgcggcaagccgacgaaccataaagtgttcggcgaggcgatg
gccatcttggcgggggacgggttgttgacgtacgcgtttcaattgatcaccgaaatcgacgatgagcgcatccctccttccgtccggct
tcggctcatcgaacggctggcgaaagcggccggtccggaagggatggtcgccggtcaggcagccgatatggaaggagagggga
aaacgctgacgctttcggagctcgaatacattcatcggcataaaaccgggaaaatgctgcaatacagcgtgcacgccggcgccttgat
cggcggcgctgatgcccggcaaacgcgggagcttgacgagtttgccgcccatctaggccttgcctttcaaattcgcgatgatattctcg
atattgaaggggcagaagaaaaaatcggcaagccggtcggcagcgaccaaagcaacaacaaagcgacgtatccagcgttgctgtc
gcttgccggcgcgaaggaaaagttggcgttccatatcgaggcggcgcagcgccatttacggaacgctgacgttgacggcgccgcg
ctcgcctatatttgcgaactggtcgccgcccgcgaccattaaCCATGGGCGGCCGCtgcggtattttctccttacgcatctgt
gcggtatttcacaccgcatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacaccc
gctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggt
tttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataatgtttcttag
acgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaa
taaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttg
ccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactgg
atctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggt
attatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacag
aaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctg
acaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggag
ctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaa
ctacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggc
tggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctccc
gtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaa
gcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt
gataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatc
ctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctt
tttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactct
gtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactca
agacgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctaca
ccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcgg
cagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctg
acttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggcct
tttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgc
cgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

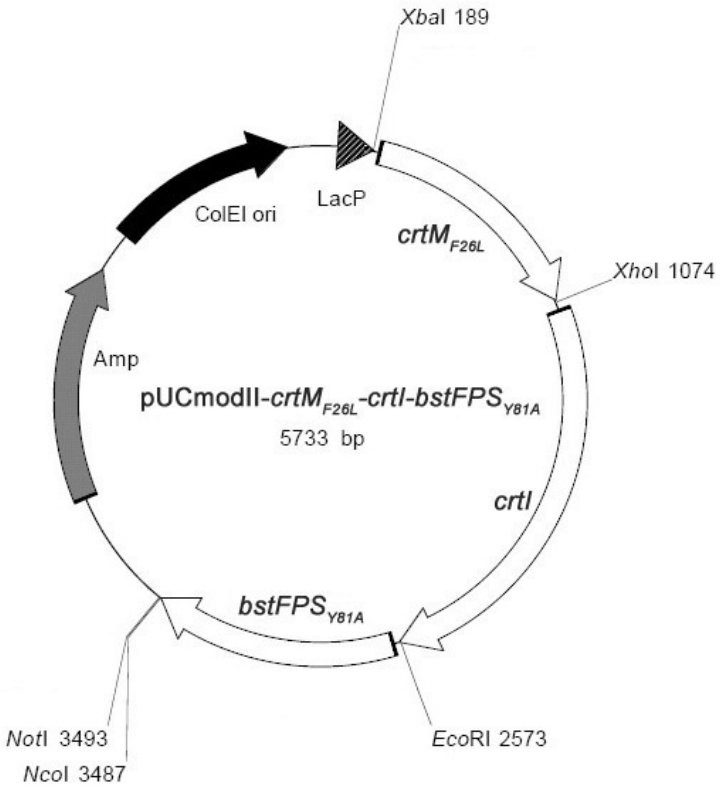
Plasmid name pUCmodII-*crtM*$_{F26L}$-*crtI*-*bstFPS*$_{Y81A}$; 5733 bp

Based on vector pUCmodII

Construction

The gene encoding the F26L mutant of the C_{30} carotenoid synthase CrtM from *Staphylococcus aureus* (*crtM*$_{F26L}$) is cloned between *Xba*I and *Xho*I sites. The gene encoding the C_{40} carotenoid desaturase CrtI from *Erwinia uredovora* (*crtI*) is cloned between *Xho*I and *Eco*RI sites. The gene encoding the Y81A mutant of the farnesyl diphosphate synthase from *Bacillus stearothermophilus* (*bstFPS*$_{Y81A}$) is cloned between *Eco*RI and *Nco*I sites. The genes are part of a single operon under the control of a single *lac* promoter (no operator). Each gene is preceded by a Shine-Dalgarno ribosomal binding sequence (AGGAGG) followed by 8 spacer nucleotides. The codon mutation encoding the F26L amino acid substitution in *crtM*$_{F26L}$ is shown in bold, underlined capital letters in the sequence below.

Plasmid map



pUCmodII-*crtM*$_{F26L}$-*crtI*-*bstFPS*$_{Y81A}$ sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaa
cgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcgTCTAGAaggagg
attacaaaatgacaatgatgaatatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagcttttcttacgct**CTC**gacttgttaccagaagatc
aaagaaaagcggtttgggcaatttatgctgtgtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaatttttaaatcaaataaaagaagatatacaa
tctattgaaaaatacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatatcgcctttcaatctttt
tataatctcattgatactgtatataaagatcaacattttacaatgtttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagtat

tgacgccgattttaagtgatcatgaaacacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcggtgaa
gattttgacaatgaacggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcattatattgactta
tgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagcacaaccaatcatagaattagcagcacgt
atatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgtggataagaggaaaaaggcaaagttgtttcatgaaaataaa
tagtaaatatcatagaatatagCTCGAGaggaggattacaaaatgaaaccaactacggtaattggtgcaggcttcggtggcctggcactggcaattcgt
ctacaagctgcggggatccccgtcttactgcttgaacaacgtgataaacccggcggtcgggcttatgtctacgaggatcaggggtttacctttgatgcaggc
ccgacggttatcaccgatcccagtgccattgaagaactgtttgcactggcaggaaaacagttaaaagagtatgtcgaactgctgccggttacgccgttttacc
gcctgtgttgggagtcagggaaggtctttaattacgataacgatcaaacccggctcgaagcgcagattcagcagtttaatccccgcgatgtcgaaggttatc
gtcagtttctggactattcacgcgcggtgtttaaagaaggctatctaaagctcggtactgtccctttttatcgttcagagacatgcttcgcgccgcacctcaact
ggcgaaactgcaggcatggagaagcgtttacagtaaggttgccagttacatcgaagatgaacatctgcgccaggcgttttctttccactcgctgttggtgggc
ggcaatcccttcgccacctcatccatttatacgttgatacacgcgctggagcgtgagtggggcgtctggtttccgcgtggcggcaccggcgcattagttcag
gggatgataaagctgtttcaggatctgggtggcgaagtcgtgttaaacgccagagtcagccatatggaaacgacaggaaacaagattgaagccgtgcattt
agaggacggtcgcaggttcctgacgcaagccgtcgcgtcaaatgcagatgtggttcatacctatcgcgacctgttaagccagcaccctgccgcggttaagc
agtccaacaaactgcagactaagcgcatgagtaactctctgtttgtgctctattttggtttgaatcaccatcatgatcagctcgcgcatcacacggtttgtttcgg
cccgcgttaccgcgagctgattgacgaaatttttaatcatgatggcctcgcagaggacttctcactttatctgcacgcgccctgtgtcacggattcgtcactgg
cgcctgaaggttgcggcagttactatgtgttggcgccggtgccgcatttaggcaccgcgaacctcgactggacggttgaggggccaaaactacgcgaccg
tatttttgcgtaccttgagcagcattacatgcctggcttacggagtcagctggtcacgcaccggatgtttacgccgtttgattttcgcgaccagcttaatgcctat
catggctcagccttttctgtggagcccgttcttacccagagcgcctggtttcggccgcataaccgcgataaaaccattactaatctctacctggtcggcgcag
gcacgcatcccggcgcaggcattcctggcgtcatcggctcggcaaaagcgacagcaggtttgatgctggaggatctgatttgaGAATTCaggagga
gtaagcgatggcgcagctttcagttgaacagtttctcaacgagcaaaaacaggcggtggaaacagcgctctcccgttatatagagcgcttagaagggccg
gcgaagctgaaaaaggcgatggcgtactcattggaggccggcggcaaacgaatccgtccgttgctgcttctgtccaccgttcgggcgctcggcaaagac
ccggcggtcggattgcccgtcgcctgcgcgattgaaatgatccatacggcatctttgatccatgatgatttgccgagcatggacaacgatgatttgcggcgc
ggcaagccgacgaaccataaagtgttcggcgaggcgatggccatcttggcgggggacgggttgttgacgtacgcgtttcaattgatcaccgaaatcgacg
atgagcgcatccctccttccgtccggcttcggctcatcgaacggctggcgaaagcggccggtccggaagggatggtcgccggtcaggcagccgatatgg
aaggagaggggaaaacgctgacgctttcggagctcgaatacattcatcggcataaaaccgggaaaatgctgcaatacagcgtgcacgccggcgccttga
tcggcggcgctgatgcccggcaaacgcgggagcttgacgagtttgccgcccatctaggccttgcctttcaaattcgcgatgatattctcgatattgaagggg
cagaagaaaaaatcggcaagccggtcggcagcgaccaaagcaacaacaaagcgacgtatccagcgttgctgtcgcttgccggcgcgaaggaaaagttg
gcgttccatatcgaggcggcgcagcgccatttacggaacgctgacgttgacggcgccgcgctcgcctatatttgcgaactggtcgccgcccgcgaccatt
aaCCATGGGCGGCCGCtgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatggtgcactctcagtacaatctgctctgatgccg
catagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctc
cgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataata
atggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctca
cccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttt
tcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgc
atacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccat
gagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttga
tcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggc
gaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggttta
ttgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactt
tagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtc
agaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgttt
gccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccac
ttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaag
acgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacc
tacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacga
gggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggag
cctatggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataacc
gtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

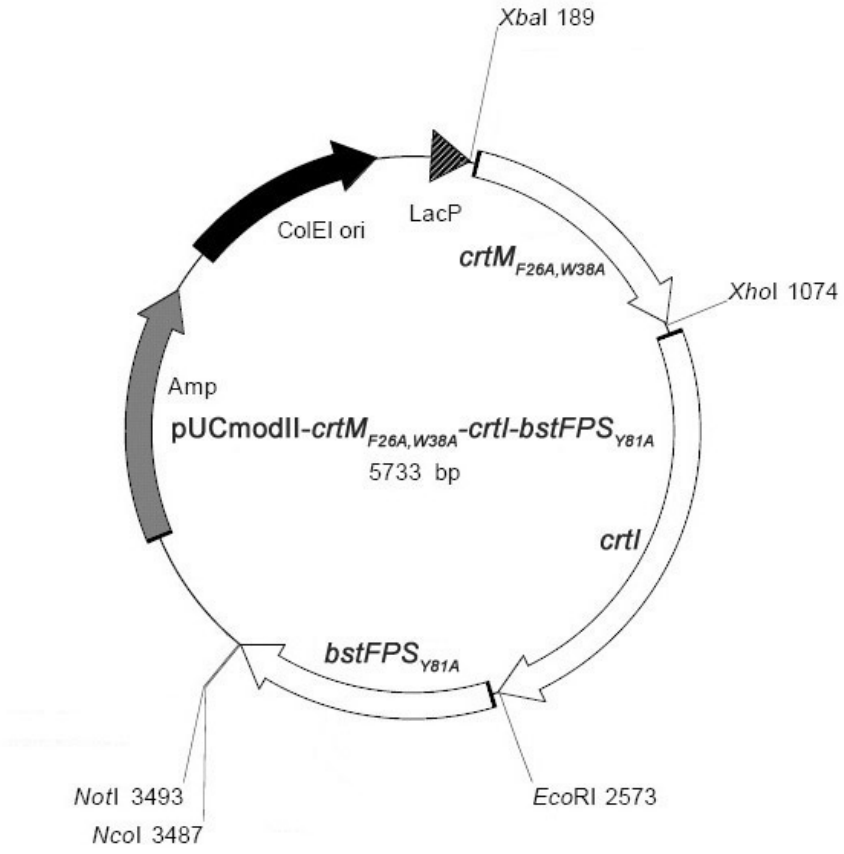
Plasmid name pUCmodII-*crtM*$_{F26A,W38A}$-*crtI-bstFPS*$_{Y81A}$; 5733 bp

Based on vector pUCmodII

Construction

The gene encoding the F26A, W38A double mutant of the C_{30} carotenoid synthase CrtM from *Staphylococcus aureus* (*crtM*$_{F26A,W38A}$) is cloned between *Xba*I and *Xho*I sites. The gene encoding the C_{40} carotenoid desaturase CrtI from *Erwinia uredovora* (*crtI*) is cloned between *Xho*I and *Eco*RI sites. The gene encoding the Y81A mutant of the farnesyl diphosphate synthase from *Bacillus stearothermophilus* (*bstFPS*$_{Y81A}$) is cloned between *Eco*RI and *Nco*I sites. The genes are part of a single operon under the control of a single *lac* promoter (no operator). Each gene is preceded by a Shine-Dalgarno ribosomal binding sequence (AGGAGG) followed by 8 spacer nucleotides. The codon mutations encoding the F26A and W38A amino acid substitutions in *crtM*$_{F26A,W38A}$ are shown in bold, underlined capital letters in the sequence below.

Plasmid map



pUCmodII-*crtM*$_{F26A,W38A}$-*crtI-bstFPS*$_{Y81A}$ sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaa
cgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcgTCTAGAaggagg
attacaaaatgacaatgatgaatatgaattttaaatattgtcataaaatcatgaagaaacattcaaaaagcttttcttacgct**GCA**gacttgttaccagaagatc
aaagaaaagcggtt**GCA**gcaatttatgctgtgtgtcgtaaaattgatgacagtatagatgtttatggcgatattcaatttttaaatcaaataaaagaagatatac
aatctattgaaaaatacccatatgaacatcatcactttcaaagtgatcgtagaatcatgatggcgcttcagcatgttgcacaacataaaaatatcgcctttcaatct

ttttataatctcattgatactgtatataaagatcaacattttacaatgtttgaaacggacgctgaattattcggatattgttatggtgttgctggtacagtaggtgaagt
attgacgccgattttaagtgatcatgaaacacatcagacatacgatgtcgcaagaagacttggtgaatcgttgcaattgattaatatattaagagatgtcggtga
agattttgacaatgaacggatatattttagtaagcaacgattaaagcaatatgaagttgatattgctgaagtgtaccaaaatggtgttaataatcattatattgactt
atgggaatattatgcagctatcgcagaaaaagattttcaagatgttatggatcaaatcaaagtatttagtattgaagcacaaccaatcatagaattagcagcacg
tatatatattgaaatactggacgaagtgagacaggctaactatacattacatgaacgtgtttttgtggataagaggaaaaaggcaaagttgtttcatgaaaataa
atagtaaatatcatagaatatagCTCGAGaggaggattacaaaatgaaaccaactacggtaattggtgcaggcttcggtggcctggcactggcaattcg
tctacaagctgcggggatcccgtcttactgcttgaacaacgtgataaacccggcggtcgggcttatgtctacgaggatcaggggtttaccttgatgcaggc
ccgacggttatcaccgatcccagtgccattgaagaactgtttgcactggcaggaaaacagttaaaagagtatgtcgaactgctgccggttacgccgttttacc
gcctgtgttgggagtcagggaaggtctttaattacgataacgatcaaacccggctcgaagcgcagattcagcagtttaatccccgcgatgtcgaaggttatc
gtcagtttctggactattcacgcgcggtgtttaaagaaggctatctaaagctcggtactgtccctttttatcgttcagagacatgcttcgcgccgcacctcaact
ggcgaaactgcaggcatggagaagcgtttacagtaaggttgccagttacatcgaagatgaacatctgcgccaggcgttttctttccactcgctgttggtgggc
ggcaatcccttcgccacctcatccatttatacgttgatacacgcgctggagcgtgagtggggcgtctggtttccgcgtggcggcaccggcgcattagttcag
gggatgataaagctgtttcaggatctgggtggcgaagtcgtgttaaacgccagagtcagccatatggaaacgacaggaaacaagattgaagccgtgcattt
agaggacggtcgcaggttcctgacgcaagccgtcgcgtcaaatgcagatgtggttcatacctatcgcgacctgttaagccagcaccctgccgcggttaagc
agtccaacaaactgcagactaagcgcatgagtaactctctgtttgtgctctattttggtttgaatcaccatcatgatcagctcgcgcatcacacggtttgtttcgg
cccgcgttaccgcgagctgattgacgaaatttttaatcatgatggcctcgcagaggacttctcactttatctgcacgcgccctgtgtcacggattcgtcactgg
cgcctgaaggttgcggcagttactatgtgttggcgccggtgccgcatttaggcaccgcgaacctcgactggacggttgaggggccaaaactacgcgaccg
tatttttgcgtaccttgagcagcattacatgcctggcttacggagtcagctggtcacgcaccggatgtttacgccgtttgattttcgcgaccagcttaatgcctat
catggctcagccttttctgtggagcccgttcttacccagagcgcctggtttcggccgcataaccgcgataaaaccattactaatctctacctggtcggcgcag
gcacgcatcccggcgcaggcattcctggcgtcatcggctcggcaaaagcgacagcaggtttgatgctggaggatctgatttgaGAATTCaggagga
gtaagcgatggcgcagctttcagttgaacagtttctcaacgagcaaaaacaggcggtggaaacagcgctctcccgttatatagagcgcttagaagggccg
gcgaagctgaaaaaggcgatggcgtactcattggaggccggcggcaaacgaatccgtccgttgctgcttctgtccaccgttcgggcgctcggcaaagac
ccggcggtcggattgcccgtcgcctgcgcgattgaaatgatccatacggcatctttgatccatgatgatttgccgagcatggacaacgatgatttgcggcgc
ggcaagccgacgaaccataaagtgttcggcgaggcgatggccatcttggcgggggacgggttgttgacgtacgcgtttcaattgatcaccgaaatcgacg
atgagcgcatccctccttccgtccggcttcggctcatcgaacggctggcgaaagcggccggtccggaagggatggtcgccggtcaggcagccgatatgg
aaggagaggggaaaacgctgacgctttcggagctcgaatacattcatcggcataaaaccgggaaaatgctgcaatacagcgtgcacgccggcgccttga
tcggcggcgctgatgcccggcaaacgcgggagcttgacgagtttgccgcccatctaggccttgcctttcaaattcgcgatgatattctcgatattgaagggg
cagaagaaaaaatcggcaagccggtcggcagcgaccaaagcaacaacaaagcgacgtatccagcgttgctgtcgcttgccggcgcgaaggaaaagttg
gcgttccatatcgaggcggcgcagcgccatttacggaacgctgacgttgacggcgccgcgctcgcctatatttgcgaactggtcgccgcccgcgaccatt
aaCCATGGGCGGCCGCtgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatggtgcactctcagtacaatctgctctgatgccg
catagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctc
cgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataata
atggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaaccccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctca
cccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttt
tcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgc
atacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccat
gagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttga
tcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggc
gaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggttta
ttgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgg
ggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactt
tagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtc
agaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgttt
gccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccac
ttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaag
acgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacc
tacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacga
gggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggag
cctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataacc
gtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaga

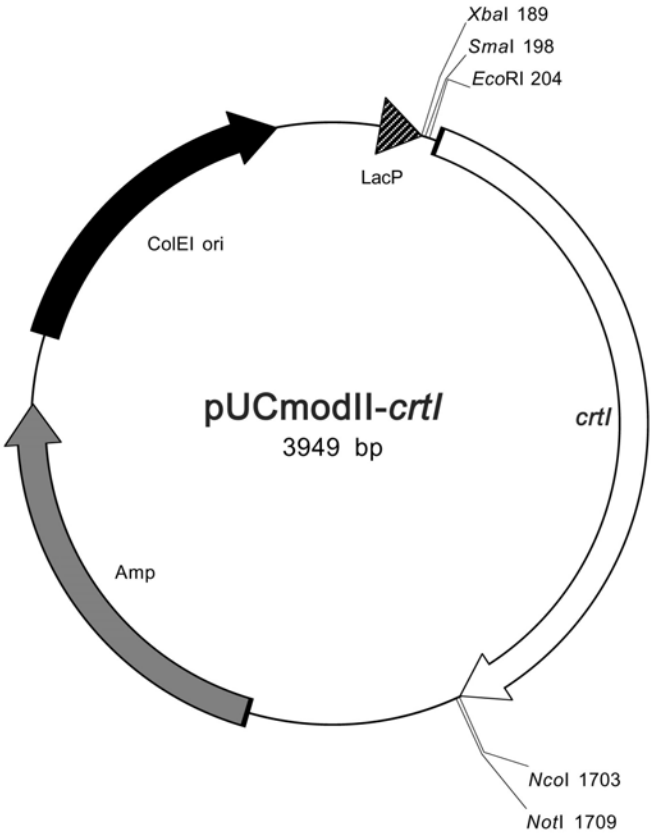
Plasmid name pUCmodII-*crtI*; 3949 bp

Based on vector pUCmodII

Construction

The gene encoding the C_{40} carotenoid desaturase CrtI from *Erwinia uredovora* (*crtI*) is cloned between the *Eco*RI and *Nco*I restriction sites of pUCmodII. The gene is preceded by a Shine-Dalgarno ribosomal binding sequence (AGGAGG) followed by 8 spacer nucleotides.

Plasmid map



pUCmodII-*crtI* sequence

gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaag
cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgt
tgtgtggaattgtgagcgTCTAGAgcgCCCGGGGAATTCaggaggattacaaaatgaaaccaactacggtaatt
ggtgcaggcttcggtggcctggcactggcaattcgtctacaagctgcggggatccccgtcttactgcttgaacaacgtgataaac
ccggcggtcgggcttatgtctacgaggatcaggggtttacctttgatgcaggcccgacggttatcaccgatcccagtgccattga
agaactgtttgcactggcaggaaaacagttaaaagagtatgtcgaactgctgccggttacgccgttttaccgcctgtgttgggagt
cagggaaggtctttaattacgataacgatcaaacccggctcgaagcgcagattcagcagtttaatccccgcgatgtcgaaggtta
tcgtcagtttctggactattcacgcgcggtgtttaaagaaggctatctaaagctcggtactgtccctttttatcgttcagagacatgct
tcgcgccgcacctcaactggcgaaactgcaggcatggagaagcgtttacagtaaggttgccagttacatcgaagatgaacatct
gcgccaggcgttttctttccactcgctgttggtgggcggcaatcccttcgccacctcatccatttatacgttgatacacgcgctgga

gcgtgagtggggcgtctggtttccgcgtggcggcaccggcgcattagttcaggggatgataaagctgtttcaggatctgggtgg
cgaagtcgtgttaaacgccagagtcagccatatggaaacgacaggaaacaagattgaagccgtgcatttagaggacggtcgca
ggttcctgacgcaagccgtcgcgtcaaatgcagatgtggttcatacctatcgcgacctgttaagccagcaccctgccgcggttaa
gcagtccaacaaactgcagactaagcgcatgagtaactctctgtttgtgctctattttggtttgaatcaccatcatgatcagctcgcg
catcacacggtttgtttcggcccgcgttaccgcgagctgattgacgaaatttttaatcatgatggcctcgcagaggacttctcacttt
atctgcacgcgccctgtgtcacggattcgtcactggcgcctgaaggttgcggcagttactatgtgttggcgccggtgccgcattta
ggcaccgcgaacctcgactggacggttgaggggccaaaactacgcgaccgtatttttgcgtaccttgagcagcattacatgcct
ggcttacggagtcagctggtcacgcaccggatgtttacgccgtttgattttcgcgaccagcttaatgcctatcatggctcagccttt
ctgtggagcccgttcttacccagagcgcctggtttcggccgcataaccgcgataaaaccattactaatctctacctggtcggcgc
aggcacgcatcccggcgcaggcattcctggcgtcatcggctcggcaaaagcgacagcaggtttgatgctggaggatctgattt
gaCCATGGGCGGCCGCtgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatggtgcactctcagta
caatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcc
cggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcga
gacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcgggg
aaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttca
ataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctc
acccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagc
ggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatccc
gtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaa
gcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctga
caacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccgg
agctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaact
ggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcgg
cccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccaga
tggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagat
aggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaa
ggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtaga
aaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtt
tgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtg
tagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgcc
agtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggggg
gttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccac
gcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttcca
gggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggg
cggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgtt
atcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcga
gtcagtgagcgaggaagcggaaga